



ELSEVIER

Contents lists available at ScienceDirect

The Lancet Regional Health - Western Pacific

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

Research paper

Settings of virus exposure and their implications in the propagation of transmission networks in a COVID-19 outbreak

Ngai Sze Wong^{a,1}, Shui Shan Lee^{a,1}, Tsz Ho Kwan^a, Eng-Kiong Yeoh^{b,c,*}^a Stanley Ho Centre for Emerging Infectious Diseases, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong, China^b Centre for Health Systems and Policy Research, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong, China^c Jockey Club School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong, China

ARTICLE INFO

Article history:

Received 16 September 2020

Revised 15 October 2020

Accepted 27 October 2020

Available online 9 November 2020

Keywords:

COVID-19

Outbreaks

Network

Exposure setting

ABSTRACT

Background: Transmission dynamics of SARS-CoV-2 varied by the settings of virus exposure. Understanding the inter-relationship between exposure setting and transmission networks would provide a basis for informing public health control strategies.

Methods: Surveillance and clinical data from the first wave of COVID-19 outbreaks in Hong Kong were accessed. Twelve exposure setting types were differentiated – household, neighbourhood, eateries, entertainment, parties, shopping, personalised service, workplace, education, worship, healthcare, transport. Clustering was investigated followed by reconstructing the transmission cascades of clustered cases using social networking approach. Linked and unlinked cases were compared in statistical analyses.

Findings: Between 23 January and 19 June 2020, 1128 cases were reported. Among 324 cases related to local transmission, 123 clusters comprising two or more epidemiologically linked cases were identified. Linked cases had lower Ct value ($p < 0.001$) than unlinked cases. Households accounted for 63% of all clusters with half as primary setting, while entertainment accounted for the highest number of primary setting transmission cases. There were altogether 19 cascades involving >1 exposure setting, with a median reproduction number of 3 (IQR: 2–4), versus 1 (IQR: 1–2) for cascades involving a single setting ($n = 36$ cascades). The longest cascade featured a bar (entertainment) as primary setting, with propagation through 30 non-primary exposure settings from seven setting types, reflecting propensity for widespread dispersion and difficulty in containment.

Interpretation: There was marked heterogeneity in the characteristics of SARS-CoV-2 transmission cascades which differed by exposure setting. Network epidemiological analyses of transmission cascades can be applied as a risk assessment tool in decision-making for calibrating social distancing measures.

Funding: Health and Medical Research Fund

© 2020 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

1. Introduction

Since its emergence in late December 2019, the COVID-19 pandemic has been spreading swiftly and extensively around the world. As of the end of September 2020, there were over 34 million confirmed cases resulting in over 1 million deaths (<https://covid19.who.int/>). With a daily increase of over 0.2 million new cases, the ultimate toll of the pandemic may surpass many of the previous influenza pandemics. Amidst the widespread dissemination of COVID-19, marked heterogeneity of SARS-CoV-2's transmis-

sion pattern is noted, [1] as has been described for the related virus causing SARS over a decade ago [2]. Modelling studies suggested the phenomenon of overdispersion, with 80% of the secondary transmission caused by a very small fraction of proportion of SARS-CoV-2 infected patients, [3] which explains the characteristics of outbreaks attributed to a single patient exposure, [4,5] or inferred from some big clusters as reported in South Korea [6] and Germany [7].

Heterogeneity of SARS-CoV-2 transmission poses challenges to the implementation of an effective containment strategy, and its adjustment in response to the evolving pandemic. Heterogeneous virus spread was postulated to be associated with high viral shedding of superspreaders, [8] as shown in reported cases [5]. The pattern of superspreading events may however vary between

* Corresponding author.

E-mail address: yeoh_ek@cuhk.edu.hk (E.-K. Yeoh).¹ Contributed equally as first authors

outbreaks, and superspreaders could be unidentified or identified only after the event. An analysis of epidemiological data of 101 cases collected early in the epidemic as of 3 March 2020 in Hong Kong, Japan and Singapore suggested transmission was not overdispersed and did not find strong evidence for the presence of superspreading events [9]. However subsequent studies including one done in Hong Kong analysing 1,037 cases as of 28 April showed substantial individual heterogeneity in transmissibility ($k = 0.45, 95\%; 0.3-0.72$) [10]. On the other hand, heterogeneity may also be related to environmental and behavioural and social factors of individuals that influence transmission dynamics of the virus, which could be dependent on the settings of the respective outbreaks [11]. Many clusters were linked to indoor settings, and it is speculated that closed environment could promote secondary virus transmission. [11,12] Droplet transmission was the predominant mode of virus spread, though airborne and fomite transmission could account for infections under special circumstances, as reviewed by CDC (<https://www.cdc.gov/coronavirus/2019-ncov/more/scientific-brief-sars-cov-2.html>). Host factors, notably asymptomatic transmission, have been reported to account for 10 to 23% of SARS-CoV-2 infection. [13–15] Detection of asymptomatic infection requires considerable efforts and financial resources. For each city/country, assessment of the clustering patterns of SARS-CoV-2 transmission and their association with asymptomatic infection would enhance the development of an effective intervention strategy.

In Hong Kong, a 7.5-million population Asia-pacific metropolitan city, the first case of COVID-19 was reported in late January 2020, a month after the virus first detected in Wuhan, China [16]. Corresponding with the growth of the epidemic in Europe and USA, imported infections were increasingly detected as from February, which precipitated the occurrence of local outbreaks. This first wave of local transmissions lasted through June 2020. The aim of this study was to delineate the virus exposure settings underlying the heterogeneity of SARS-CoV-2 transmission, so as to reconstruct the cascades of transmission clusters in evaluating their dynamics through the adoption of a network epidemiological approach.

2. Methods

2.1. Data source and variables

Surveillance data of COVID-19 cases reported during the first wave of the outbreak between 23 January and 19 June 2020 were collected from the Centre for Health Protection, Department of Health, Hong Kong Special Administrative Region Government. Case reporting criteria for COVID-19 were derived from World Health Organization, which included presentation with fever or acute respiratory illness/pneumonia, in conjunction with either relevant travel history or contact with a confirmed case in the preceding 14 days (https://cdis.chp.gov.hk/CDIS_CENO_ONLINE/ceno.html). Confirmed diagnosis was made in all reported cases, with the detection of SARS-CoV-2 nucleic acid. The datasets comprised socio-demographic and epidemiological data. Matched data from clinical and laboratory records of all reported cases, with follow-up data through 2 August were retrieved from the Hospital Authority which manages all public hospitals where all COVID-19 patients were isolated and receive care. The data included socio-demographics (age, gender, ethnicity, residency, underlying medical conditions), diagnosis and reporting (dates and mode of detection, reporting, minimum PCR Ct [polymerase chain reaction cycle threshold] value, with a lower value suggesting higher viral load), clinical characteristics and outcome (symptoms before or at diagnosis, clinical status, days from onset to admission, deaths), epidemiological characteristics (travel history, imported/local trans-

mission, number of close contacts, epidemiological linkages, history of quarantine for close contacts, and isolation for infected persons).

Ethical approval was obtained from the Survey and Behavioural Research Ethics Committee of the Chinese University of Hong Kong (Ref no. SBRE-19-595). Data access approval was granted by the Hong Kong Special Administrative Region Government.

2.2. Exposure settings

Twelve exposure setting types were differentiated for all local transmission cases reported during the first wave of SARS-CoV-2 outbreak in Hong Kong: household, neighbourhood, entertainment, eateries, parties, shopping, personalised services, workplace, education, worship, healthcare, and transport. Characteristics of these settings varied by the number of individuals involved, indoor or outdoor location, gathering of persons known and/or unknown to one another, repeated versus one-off event, and the duration (Appendix Table S1). Briefly, household referred to the residence for co-living individuals or close contacts who always meet each other but not living together e.g. couples in romantic relationship. Neighbourhood was the living environment of neighbours not from the same household. There were three different settings for social activities: eateries referred to places where meals were offered, e.g. restaurants and cafeteria; entertainment were social activities at premises like bar and karaoke patronised by members of the public; and parties were private gatherings with known persons. Shopping referred to activities in shops, markets and department stores, which were open to public. Personalised services were attended by individuals outside residence as those in gymnasium and beauty parlour. Workplace was activity space for staff. Education referred to school and classroom-based activities. Worship comprised designated venues for religious activities. Healthcare referred to institutions for caring for sick people. Transport covered vehicles and stations for passengers, drivers, or staff.

2.3. Transmission networks, linkages of cases and cascades

SARS-CoV-2 is transmitted in networks, with cases forming clusters composing of epidemiologically linked persons, either directly through inter-personal contacts or indirectly as a result of sharing the same exposure setting during specific time intervals. All cases could present either as isolates without epidemiological linkages (unlinked) to other cases, or in clusters when two or more persons were epidemiologically linked. Clustering of cases in the transmission networks was inferred from the epidemiological history derived from clinical and surveillance datasets. Network diagrams were constructed using UCINET 6 for Windows.

A transmission cascade represents the structure of the chained propagation of the infections through nodes in the network [17]. By applying a two-mode network approach, SARS-CoV-2 infected persons in clusters were represented as nodes in one mode, while the second mode was the respective setting of virus exposure. These cascades were illustrated in two-mode network diagrams. The reproduction number R of a cascade was the average number of secondary cases generated by an index case in the cascade. To evaluate the network property of the clusters, two-mode networks were transformed to one-mode networks.

2.4. Data analysis

Descriptive statistics were used to describe the epidemiologic characteristics of all SARS-CoV-2 cases reported during the first wave of the outbreak. Comparison between imported (marked as 0) and local (marked as 1) cases were performed in bivariable logistic regression models in SPSS 25 (Appendix Figure S1). Contin-

Table 1

Comparison between imported cases (n = 690) and local transmission cases (n = 438) during the first wave of SARS-CoV-2 outbreak in Hong Kong, between January and June 2020

	imported n	(n = 690) %	local n	(n = 438) %	Total n	%	OR (U-test)	95%C.I. (p-value)
Gender								
female	313	45%	209	48%	522	46%	ref	
male	377	55%	229	52%	606	54%	0.91	0.72–1.16
Median age at onset, IQR	28	20–48	40	31–56	35	22–52	(98427)	(<0.001*)
Ethnicity								
non-Chinese	221	32%	134	31%	355	31%	ref	
Chinese	469	68%	304	69%	773	69%	1.07	0.83–1.38
Hong Kong residency								
no	20	3%	1	0.2%	21	2%	ref	
yes	670	97%	437	99.8%	1107	98%	13.04	1.74–97.55*
Underlying illness								
no	602	87%	358	82%	960	85%	ref	
yes	88	13%	80	18%	168	15%	1.53	1.10–2.13*
Symptomatic								
No	205	30%	48	11%	253	22%	ref	
Yes	485	70%	390	89%	875	78%	3.43	2.44–4.83*

*p < 0.05

IQR: interquartile range; OR: odds ratio; U-test: Mann-Whitney U test

ous independent variables such as age and Ct value were compared by Mann-Whitney U test. Local cases as well as imported cases causing local transmission (referred as cases related to local transmission) were selected, and comparison between unlinked (marked as 0) and linked (marked as 1) cases was performed in bivariable logistic regression. Complete-case analyses were performed.

The pattern of transmission in linked cases was described by the size of cascades, intervals between the earliest onset date and last reporting date in the cascade, and type of exposure setting. From the transformed one-mode networks, degree centrality was measured for all linked cases by setting to evaluate the number of ties each case had. Normalised degree centrality was calculated by dividing the number of edges by the maximum number of possible edges in UCINET.

2.5. Role of the funding source

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

3. Results

3.1. General characteristics

Between 23 January and 19 June 2020, a total of 1128 cases of SARS-CoV-2 infection were reported in Hong Kong. The distribution of cases by reporting channels followed the description in a published report [16]. Travel restrictions were introduced by the government since January/February, and enhanced in addition to legal regulations on social distancing enacted (see Fig. 1a) in late March when the transmission in the current wave reached its peak. Three-quarters of these cases (843/1128) were reported during a 26-day period between 17 March and 11 April 2020 (Fig. 1a). Around 60% were imported cases, and 69% of all cases reported travel history 14-day prior the reporting date. Compared with imported cases (n = 690), local cases (n = 438) were more likely to be older (40 year-old (yo) vs 28 yo, p < 0.001), have underlying illness (OR = 1.53, 95%CI. = 1.10–2.13), and be symptomatic (OR = 3.43, 95%CI. = 2.44–4.83) at presentation (Table 1). Asymptomatic cases (n = 253) were identified in the course of quarantine (n = 227), contact tracing (n = 20) and the introduction of

enhanced surveillance (n = six) by the screening of citizens with risk of exposure, the latter undertaken in private clinics, government outpatient clinics (GOPC) and accident and emergency departments (AED). Six (0.5%) of the reported cases have died.

3.2. Comparison of linked and unlinked cases related to local transmission

There were 450 cases related to local transmission, including 20 imported infections causing subsequent secondary infections. Among them, 324/450 (72%) were linked cases while 126/450 (28%) were unlinked cases. There was no significant difference in socio-demographic characteristics between the two groups. A lower proportion of linked cases were symptomatic at presentation (88% vs 94%) (Table 2). During hospitalisation, linked cases recorded lower Ct value (23.6 vs 29.1, p < 0.001) than unlinked cases. Epidemiologically, linked cases gave a shorter interval from onset to hospital admission (3.9 days vs 5.7 days, p < 0.001).

3.3. Settings of virus exposure and clustering

Table 3 shows the distribution of all linked cases related to local transmission and the number of clusters by exposure setting. A total of 123 clusters were identified among 324 linked cases related to local transmission. Transmission in household settings accounted for 63% of clusters and 56% of all linked cases (Table 3). Each of the remaining settings accounted for less than 10% of clusters and cases, except entertainment settings (28% of cases). Demographically, a high proportion of cases were aged 25–49 in all settings, except neighbourhood (more in age group of 50–64) and worship (more in age group of ≥65). Transport, household and neighbourhood had the lowest proportion of clusters in primary settings. The proportion of asymptomatic infections varied between 0% and 20% of cases by setting. The mean Ct value by setting ranged from 18 to 35 for index cases (n = 45), and between 18 and 31 for clusters. Statistical comparison could not be made because of the small numbers per setting.

The exposure settings could be further differentiated into primary (initial transmission) and non-primary (subsequent transmission) by their sequence of occurrence in a cascade. Table 4 shows the propagation of linked transmission from primary to subsequent non-primary settings. The primary setting with the highest number of linked transmissions was household, forming 37 clusters

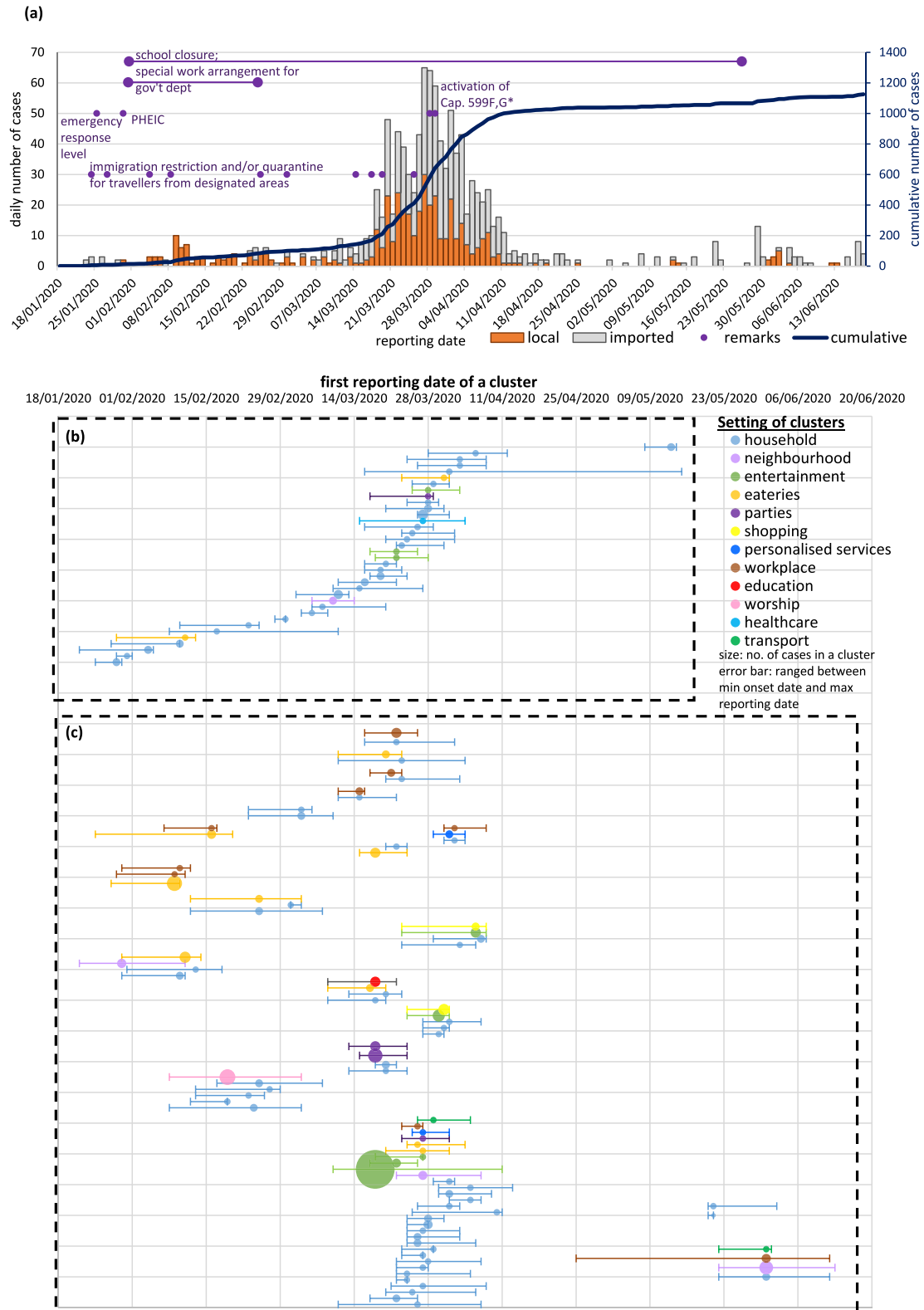


Fig. 1. (a) epidemic curve of SARS-CoV-2 transmission in Hong Kong (n=1128), with interventions shown on the same timeline; temporal distribution of transmission 324 cases: (b) by cluster of two or more cases; and (c) by cascade composing of linked clusters
 Remarks: *Cap. 599F Prevention and Control of Disease (Requirements and Directions) (Business and Premises) Regulation came into operation on 28 March 2020. Under the regulation, customers at eateries should wear mask at all times and only to remove when they consume food or drink; *Cap. 599G – Prevention and Control of Disease (Prohibition on Group Gathering) Regulation came into operation on 29 March 2020. On Fig. 1(b) and (c), the bubble size represents the number of cases in each cluster of linked transmissions, coloured by settings, the location of bubble is the first reporting date, and error bar from the onset date of the first case to the reporting date of the last case.

Table 2
Comparison between linked (n = 332) and unlinked cases (n = 126) related to local transmission of SARS-CoV-2 in Hong Kong, January to June 2020

	unlinked n	(n = 126) %	linked n	(n = 324) %	OR (U-test)	95%CI. (p-value)
Socio-demographics						
Gender						
female	55	44%	159	49%	ref	
male	71	56%	165	51%	0.80	0.53–1.22
Median age at onset, IQR	39	32–51	40	31–57	(19919)	(0.69)
Ethnicity						
non-Chinese	41	33%	98	30%	ref	
Chinese	85	67%	226	70%	1.11	0.72–1.73
Underlying illness						
no	103	82%	264	81%	ref	
yes	23	18%	60	19%	1.02	0.60–1.73
Symptomatic at diagnosis						
No	7	6%	40	12%	ref	
Yes	119	94%	284	88%	0.42	0.18–0.96*
fever	65	52%	175	54%	1.10	0.73–1.66
cough	62	49%	141	44%	0.80	0.53–1.20
sore throat	28	22%	64	20%	0.86	0.52–1.42
diarrhoea	8	6%	18	6%	0.87	0.37–2.05
loss of taste and/or smell	7	6%	10	3%	0.54	0.20–1.46
Detection and outcome						
onset and PCR Ct value date, median days, IQR	12	6–17	8	4–14	(7348)	(<0.001*)
median min. PCR Ct value, IQR	29.1	22.7–32.7	23.6	18.7–30.9	(8774)	(<0.001*)
onset to admission, median days, IQR	5.7	3.7–8.3	3.9	1.8–7.0	(12992.5)	(<0.001*)
onset to isolation, median days, IQR	5.0	3.0–9.0	4.0	2.0–9.0	(3403.5)	(0.54)
onset to report, median days, IQR	6.0	3.0–9.0	4.0	2.0–8.0	(13641)	(0.002*)
Epidemiologic characteristics						
Travel history						
no	50	40%	288	89%	ref	
yes	76	60%	36	11%	0.80	0.05–0.14*
Classification						
Close contact of imported case	0	0%	23	7%	/	
Close contact of local case	0	0%	178	55%		
Close contact of possibly local case	0	0%	61	19%		
Imported source case	0	0%	20	6%		
Local source case	41	33%	27	8%		
Possibly local source case	85	67%	15	5%		

*p < 0.05

IQR: interquartile range; OR: odds ratio; PCR Ct value: polymerase chain reaction cycle threshold value; U-test: Mann–Whitney U test

among a total of 87 persons, followed by eateries with 8 clusters and 33 cases. Entertainment, while accounting for the third highest number of clusters (n = seven), involved the highest number of 94 infected persons. Primary entertainment setting infection was epidemiologically linked with a diverse range of subsequent non-primary setting transmissions, including household, neighbourhood, eateries, parties, personalised services, workplace, and transport. Neighbourhood as a primary setting was linked only with two non-primary setting transmissions – household and eateries. Linkage of eateries was limited to the non-primary settings of household and workplace. Similarly, workplace was linked only with three non-primary settings, namely, household, neighbourhood, and transport.

Temporally, clusters in entertainment setting propagated through a short period of time (about 1.5 months), compared with those in household, neighbourhood, and workplace settings (almost spanning over the whole study period) (Fig. 1b and 1c). With social distancing regulations gazetted by the government on 28 March, reporting of clusters in settings of eateries, entertainment, personalised services, shopping, parties, education, and worship ceased by 11 April 2020. The last case of a cluster outbreak was reported on 2 June 2020.

3.4. Transmission cascades and network properties

To profile the transmission cascades, two-mode network diagrams were drawn to demonstrate the propagation of the viruses

through settings and persons (Fig. 2a). There were altogether 19 cascades involving more than one infection setting, with the remainder (36) associated with only one setting. The size of these cascades ranged between two and 103 by the number of persons, and one and 31 by the number of different settings. Almost three-quarters (239/324) of the linked cases were involved in 35% (19/55) of transmission cascades with more than one setting (Table 5). Among all transmission cascades involving only one setting, 78% (28/36) were in households, and 8% in entertainments. Among transmission cascades with more than one setting, a higher proportion were in household (7/19) and eateries (6/19). The median reproduction number R of cascades involving one or more setting was 3 (IQR: 2–4), versus 1 (IQR: 1–2) for single setting. A high proportion of cascades involved households as non-primary setting (10/19), while none of the cascades involved entertainment, shopping, education, worship, and healthcare as non-primary settings. Transport only existed as secondary setting for transmission cascades.

On a temporal scale, the median time interval of onset dates between the first and the last case in the same cascade was 15 days (IQR 9–24) in cascades with more than one setting, and seven days (IQR 4–10) in cascade with only one setting. The longest cascade by the number of exposure settings was associated with a bar (entertainment setting type) as primary setting (72 cases reported), which propagated through 30 non-primary settings (31 cases reported), spanning over 26 days by reporting dates (week 13–14). The cascade's reproduction number was 4.25. Comparing

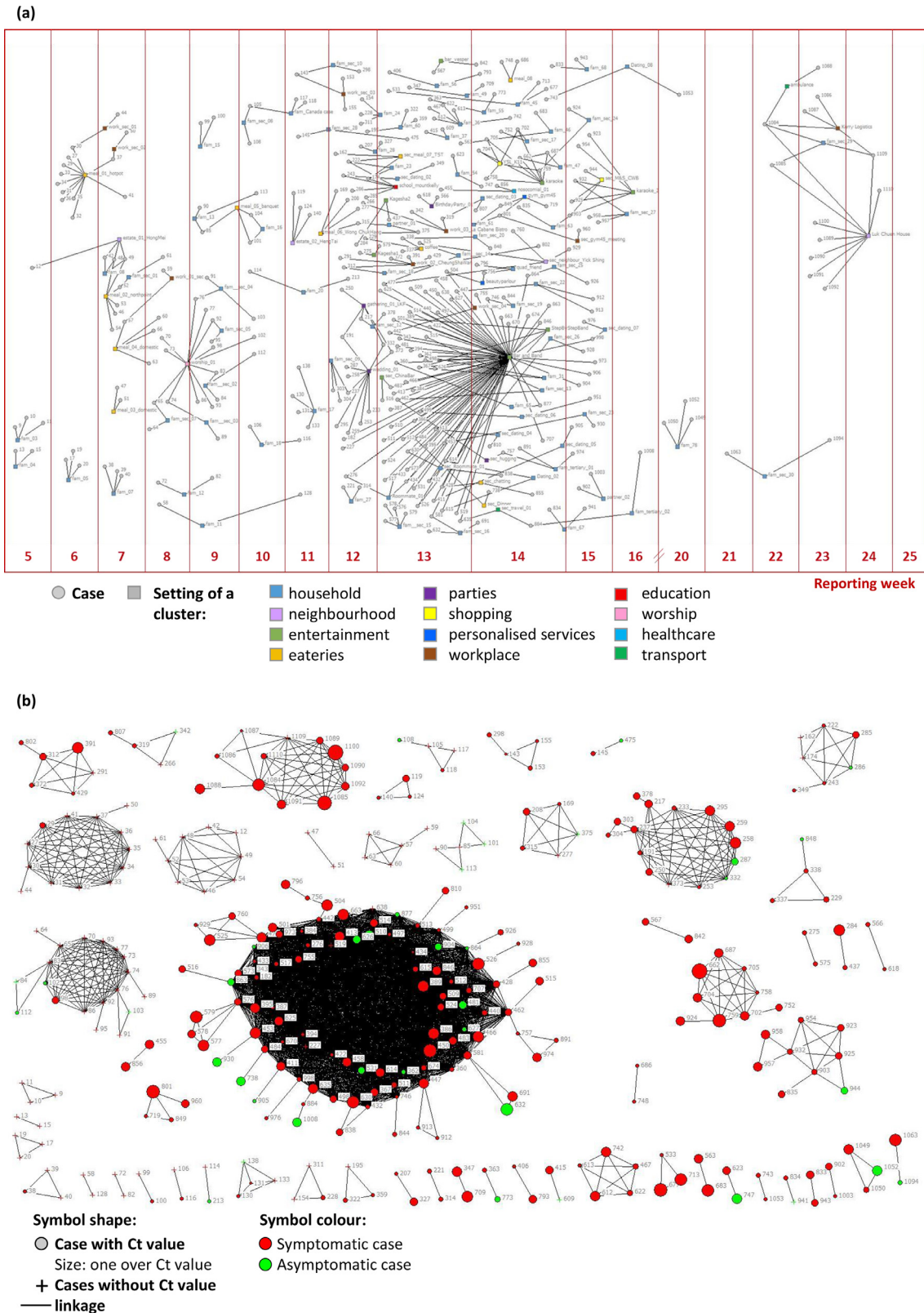


Fig. 2. Transmission cascades of 324 linked cases related to local transmission illustrated in networks (a) Two-mode networks with cases linked with cluster name - cases shown as circles and linked to other cases sharing the same cluster name, classified by the type of exposure settings and aligned by reporting week (from week 5 (29 January 2020) to week 24 (13 June 2020)); (b) One-mode network diagram of epidemiologically linked cases - cases with Ct value shown as circles with 1/Ct value represented by graduated symbol size and symbol colour (symptomatic in red, asymptomatic in green) while cases without Ct value records are shown as 'plus' signs.

Table 3
SARS-CoV-2 transmission characteristics by setting of virus exposure, n = 324

	household	neighbourhood	entertainment	eateries	parties	shopping	personalised services	workplace	education	worship	healthcare	transport
# of clusters (n = 123)	78	4	7	10	3	2	2	10	1	1	1	2
# of cases^ (n = 324)	180	20	91	43	16	9	5	25	5	12	2	4
age group												
< 15	6%	0%	0%	0%	6%	0%	0%	0%	0%	0%	0%	0%
15-24	4%	5%	9%	9%	0%	11%	0%	0%	0%	0%	0%	25%
25-49	51%	20%	76%	49%	88%	89%	60%	48%	100%	17%	50%	75%
50-64	26%	55%	13%	23%	6%	0%	40%	48%	0%	17%	0%	0%
>=65	13%	20%	2%	19%	0%	0%	0%	4%	0%	67%	50%	0%
number of index case with Ct value	22	2	6	3	2	2	1	4	1	1	1	0
Ct value of index (mean)	26.40	27.61	24.39	31.71	33.57	21.66	32.89	32.69	34.73	35.04	18.12	NA
Ct value of cluster (mean)	25.70	22.98	23.35	27.00	26.10	22.56	23.72	26.02	31.00	28.88	18.39	20.71
% of asymptomatic cases	14%	0%	12%	9%	13%	11%	0%	4%	20%	8%	0%	0%
% of clusters in primary settings	47%	35%	99%	77%	88%	100%	60%	60%	100%	100%	100%	0%
% of clusters in non-primary settings	53%	65%	1%	23%	12%	0%	40%	40%	0%	0%	0%	100%
Normalised degree ^{&} (median; IQR)	0.01; 0.003-0.02	0.03; 0.01-0.03	0.22; 0.22-0.22	0.02; 0.01-0.03	0.03; 0.01-0.04	0.03; 0.02-0.04	0.01; 0.01-0.01	0.01; 0.01-0.01	0.02; 0.01-0.02	0.01; 0.01-0.01	0.003; NA	0.02; 0.004-0.18

[^]case counts may be duplicated if the same cases belong to more than one setting

[&] case centrality in terms of normalised degree is calculated in a complete network, which could be formed by a few connected clusters (Fig. 2b). The normalised degree – degree (i.e. number of edges) divided by the maximum possible edges;

Table 4
Propagation of SARS-CoV-2 linked transmission from primary to subsequent (non-primary) settings of virus exposure, n = 324

Cascades' primary settings	No. of clusters (cases) in cascades' primary setting	No. of clusters (cases) in cascades' non-primary setting	Breakdown of non-primary settings									
			eateries	entertainment	personalised services	household	neighbourhood	parties	transport	workplace		
entertainment, shopping	1 (7)	3 (6)				3 (6)						
entertainment, household, entertainment, shopping	5 (82)	29 (67)	2 (4)	1 (2)	1 (2)	21 (49)	1 (4)	1 (2)	1 (2)	1 (2)	1 (2)	
household, eateries	1 (7)	1 (3)				1 (2)						
household, eateries, education	1 (5)	1 (2)				1 (2)						
household, workplace	1 (6)	1 (2)				1 (2)						
household, eateries, workplace	2 (10)											
household, eateries, workplace, neighbourhood	32 (76)											
eateries	6 (27)	5 (10)				2 (4)						3 (6)
workplace	2 (7)	4 (16)				2 (5)						
neighbourhood	2 (7)	3 (11)	1 (6)			2 (5)	1 (9)				1 (2)	
personalised services	1 (3)	2 (4)				1 (2)						1 (2)
parties	3 (17)	2 (5)				2 (5)						
healthcare	1 (2)											
worship	1 (12)	5 (12)										

Note: case counts may be duplicated if the same cases belong to more than one cluster

Table 5
Properties of 55 transmission cascades with only one and more than one virus exposure setting

	One setting	More than one setting
No. of cascades (no. of cases)	36 (85 cases)	19 (239 cases)
No. of cascades in the primary (non-primary) settings		
household	28	7 (10)
neighbourhood	1	1 (2)
entertainment	3	3 (0)
eateries	2	6 (2)
parties	1	1 (1)
shopping	0	2 (0)
personalised services	0	1 (1)
workplace	0	4 (4)
education	0	1 (0)
worship	0	1 (0)
healthcare	1	0 (0)
transport	0	0 (2)
% of cases in non-primary settings (median; IQR in cascades)	/	28% (22%; 14%–30%)
Median days from onset date of 1 st case to the last case*; IQR	7; 4–10	15; 9–24
Median days from onset date of 1 st case in primary setting to onset date of 1 st new case in non-primary setting*; IQR	/	9; 5–18
Median reproduction number R of cascade [§] ; IQR	1; 1–2	3; 2–4

*for asymptomatic cases without onset date, reporting date is used

[§] average number of secondary cases generated by an index case in the cascade

secondary cases between cascades with entertainment and household as primary setting, secondary cases in entertainment cascades were less likely to be asymptomatic (13% vs 28%, $p = 0.018$). Among symptomatic secondary cases, the interval between onset date and isolation date was marginally significantly different between the two settings (median 3 days for entertainment vs median 2 days for household, $p = 0.056$). All cases in cascades of entertainment as primary setting (median 37 yo) was younger than those in household cascade (median 48 yo, $p = 0.04$). To assess the importance of the component nodes (infected persons) in the one-mode network analysis, the normalised degree centrality of clusters in entertainment setting was highest (Table 3, Fig. 2b), reflecting the highest number of linkages.

4. Discussion

Heterogeneity is the hallmark of the transmission pattern of SARS-CoV-2, as inferred from global molecular epidemiologic studies. [18] Its propagation varies by the rate and extent of secondary transmission in the population, which is in turn dependent on the settings of virus exposure. Our results reported, for the first time, the association between exposure settings and the propagation of transmission clusters using real world surveillance and contact tracing data. Our results showed that among all settings implicated in the transmission of SARS-CoV-2 in the first wave outbreak in Hong Kong, entertainment constituted the main setting characterised by rapidity of propagation, linkage with multiple secondary transmission settings and long cascades involving multiple settings. The entertainment setting often involved the participation of customers with unknown risks who might not be known to one another, resulting in clusters that featured higher centrality, reflecting closer and more linkages between persons. Such transmission appeared to be effectively contained through regulatory measures. Households accounted for over half of all reported transmissions, but a majority of the linked infections in households involved two linked cases only, without leading to long cascades which can be explained by easier identification and resultant early intervention. Predominance of household setting for SARS-CoV-2 transmission

has also been reported in other Chinese cities, [19,20] and highlighted in a recent systematic review [11]. On the other hand, transmission clusters from worship and personalised service (like fitness centres) setting were documented, as have been reported in South Korea at a larger scale. [21,22] Eateries, personalised services and workplace led to further transmission outside households, but the cascades were short. Healthcare and transport setting transmissions were self-limiting without further transmission in secondary settings. Likewise, workplace and neighbourhood settings accounted for a small proportion of the transmission clusters. Overall, long cascades and multiple exposure settings reflected the propensity for widespread transmission and dispersion in the community and the difficulty in containment.

To understand the factors potentially contributing to the propagation of transmission clusters, we examined the virus concentration by evaluating the Ct values from PCR testing. In our study, the linked cases gave a lower Ct, therefore implying higher concentration, during hospitalisation. The Ct results and asymptomatic proportion by exposure setting did not allow meaningful comparison because of the small number per setting available for analysis. The heterogeneous transmission pattern by exposure setting type could however be assessed by the difference in the characteristics of the interpersonal relationship and environment implicated. The prolonged and repeated exposure to infected family members explained the risk of transmission in households [11]. Gathering in closed environments was a feature of entertainment settings, posing higher risk of virus transmission [12]. Parties, which involved friends and relatives facilitated early identification in contact tracing, gave shorter cascades compared to entertainment which was often participated by people unknown to one another and with unknown risk and challenges in contact tracing. Airborne and fomite transmission might occur in the spread of the virus in settings with no evidence of direct inter-personal contacts, for example, specific workplaces, neighbourhood, and transport.(Appendix Table S1). Their role in the propagation of long cascades like that in entertainment setting would require further investigations. The Japanese Government advised the population to avoid the “Three Cs”: closed spaces with poor ventilation, crowded places, and close-contact settings [23]. Such characteristics were common for entertainments in Hong Kong, which led to larger scale (multiple non-primary settings) of spread and longer cascades. Elsewhere, hostels were important location for virus spread as reported in Singapore, where 19.4% of the residents in one major dormitory housing – migrant workers were infected [24]. Similar scale of such accommodation for workers has not been introduced in Hong Kong, and no reported transmission has occurred in similar settings. Limitation of spread from healthcare settings in this study could be related to the effective infection control practice adopted, while extensive spread has been widely reported elsewhere [25]. Transport setting characteristics involved limited short-term contact despite the large population implicated. The observation underlines the importance of profiling settings of local importance, as interpersonal networks could differ from place to place.

The network epidemiologic delineation of transmission cascades represents one practical means of comprehending the dynamic pattern of SARS-CoV-2 outbreak. When COVID-19 first emerged in the beginning of the year, its monitoring had relied on the use of population level metrics including doubling time and basic reproduction numbers. [26,27] The latter has subsequently been modified as estimated time-varying basic reproduction numbers or effective reproduction number (Rt) for tracking temporal changes, while control measures continued to be enforced [28]. Timely development of an exit strategy that is based on a risk assessment framework is crucial to enable societies to return to normal when outbreaks are deemed to be under control [29]. Taking reference to the heterogeneity of the transmission cascades, con-

trol measures could be adjusted by setting which reflects their relative risk in likelihood of contributing to a resurgence of transmission. As illustrated in this study, activities embodied in settings associated with rapid virus dissemination, multiple secondary transmission and longer cascades which would be more difficult to contain could be prioritised for maintaining enhanced social distancing control. On the contrary, clusters in settings with limited spread and short cascades are less likely to lead to extensive spread. For example, social distancing measures in households and neighbourhood settings could be lifted when community spread is subsiding.

Our study carried some limitations. The data used in the study were derived from the COVID-19 surveillance system, contact tracing and case detection, and clinical records of reported infections. It is possible that some asymptomatic infections or mildly symptomatic patients not seeking medical attention might have been missed. The linkage between cases were founded on reports from epidemiological investigation without validation by other objective methodologies like molecular studies for all cases. The sources of infection of the first detected case in most clusters were unknown and assumed to be from the local community. Likewise, in the assignment of settings for virus exposure, recall bias and subjective perception might have affected the ultimate epidemic history. In the delineation of the transmission networks, omitted infections were assumed to have played little role in the reconstructed cascades. In this study, the epidemiologic analyses have covered only exposure settings involved in the spread of reported SARS-CoV-2 infections during the first wave of Hong Kong's outbreak. The relatively short cascades of most transmission networks might have resulted from the government's system of contact tracing and quarantine, and early introduction of social distancing measures. [16,30] The analyses shown in this study could therefore be the result of both setting-based transmission and also the effects of intervention. Extrapolation to other cities/countries and at different epidemic stage would need to be cautioned. Nevertheless network epidemiological analysis in conjunction with contact tracing can be effective in targeting interventions which would be relevant in different contexts, in the control of transmission of SARS-CoV-2 and avert the need for population wide quarantine.

In conclusion, through a network-based approach, we have reconstructed the transmission cascades of the first wave of Hong Kong's SARS-CoV-2 outbreak. There was marked heterogeneity in the characteristics of the transmission cascades which varied with the setting of virus exposure. Entertainment stood out as the primary setting associated with the widest spread and longest cascades of linked transmission. While the transmission dynamics of SARS-CoV-2 infection may vary geographically and be impacted by different intervention strategies, the focus on exposure setting is of epidemiological importance. By evaluating the transmission cascades, activities of settings implicated in rapid virus dissemination, multiple secondary transmission and longer cascades should be targeted for evaluating and adjusting intervention strategies in reducing the risk of rebound and achieving public health control of the epidemics.

Contributors

SSL, NSW, and EKY designed the study. EKY collected the data. NSW and THK conducted data analysis under supervision of SSL and EKY. SSL and NSW wrote the draft of the manuscript. All authors contributed to and approved the final version of the manuscript.

Research In Context

Evidence before this study

Transmission dynamics of SARS-CoV-2 infection was reviewed from literature accessed on PubMed as of the end of July 2020. Marked heterogeneity of COVID-19 epidemiology was noted, highlighting the potential influences of superspreading events. Modelling studies suggested the phenomenon of overdispersion, with 80% of the secondary transmission caused by a very small fraction of infected patients, which explained the characteristics of outbreaks attributed to single patient exposure, or inferred from clusters reported in China, South Korea, and Germany. The settings of virus exposure were postulated to be associated with the risk of virus transmission, influenced by environmental and behavioural factors. The relationship between the settings and the epidemiology of the transmission networks has however not been specifically investigated.

Added value of this study

For the first time we used a network approach to analyse real world surveillance data to correlate with virus exposure setting. During the first wave of the SARS-CoV-2 outbreak in Hong Kong, twelve exposure settings were defined. A total of 123 clusters of two or more cases among 324 infected persons related to local transmission were identified. The transmission cascades constructed from two-mode networks showed that their dynamics was related to the primary exposure setting. Specifically, entertainment setting exposure was epidemiologically linked with the widest extent of secondary transmissions, while transport, neighbourhood and households accounted for the lowest proportion of clustered transmission in primary setting. Linkage of eateries was limited to the non-primary setting transmissions of household and workplace. The results highlighted the phenomenon of setting-specific heterogeneity of SARS-CoV-2 transmission, and supported the application of network analyses for assessing the population risk of COVID-19.

Implications of all the available evidence

By defining exposure settings in the population in conjunction of SARS-CoV-2 cluster analyses, its transmission dynamics could be assessed for staging epidemiologic risks. Activities of settings implicated in rapid virus dissemination, multiple secondary transmission and longer cascades could be targeted for adjusting intervention strategies like relaxation of social distancing measures, enabling better public health control of the epidemics to be achieved.

Disclaimer

The opinions and assertions contained herein are private views of the authors and do not necessarily reflect those of the Centre for Health Protection, Hong Kong Special Administrative Region Government Department of Health.

Data sharing statement

The dataset cannot be included in a public repository because the data are owned by third parties. Access to these data and permission could be inquired through the Hospital Authority and Department of Health, Hong Kong SAR Government.

Editor note: The Lancet Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

Declaration of Competing Interest

We have no conflict of interest to declare.

Acknowledgements

This work was supported by Health and Medical Research Fund of the Food and Health Bureau, The Hong Kong Special Administrative Region [grant numbers COVID190105]. The authors thank Ms

Carrie Yam, Ms Mandy Li and Ms Sharon Chung for their assistance in data collection. Li Ka Shing Institute of Health Sciences, and The Centre for Health Systems and Policy Research funded by Tung Foundation, The Chinese University of Hong Kong, is acknowledged for rendering technical support throughout the conduction of this study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.lanwpc.2020.100052](https://doi.org/10.1016/j.lanwpc.2020.100052).

References

- [1] Skums P, Kirpich A, Icer Baykal P, Zelikovsky A, Chowell G. Global transmission network of SARS-CoV-2: from outbreak to pandemic. medRxiv 2020;03:22 Mar 27:202020041145. doi:[10.1101/2020.03.22.20041145](https://doi.org/10.1101/2020.03.22.20041145).
- [2] Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature* 2005;438:355–9.
- [3] Endo A. Centre for the mathematical modelling of infectious diseases COVID-19 working group. In: Abbott S, Kucharski AJ, Funk S, editors. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res*; 2020. p. 67.
- [4] Bao C, Pan E, Ai J, et al. COVID-19 Outbreak following a single patient exposure at an entertainment site: an epidemiological study. *Transbound Emerg Dis* 2020 Jul 28. doi:[10.1111/tbed.13742](https://doi.org/10.1111/tbed.13742).
- [5] Lin J, Yan K, Zhang J, Cai T, Zheng J. A super-spreader of COVID-19 in Ningbo city in China. *J Infect Public Health* 2020;13:935–7.
- [6] Shim E, Tariq A, Choi W, Lee Y, Chowell G. Transmission potential and severity of COVID-19 in South Korea. *Int J Infect Dis* 2020;93:339–44.
- [7] Walker A, Houwaart T, Wienemann T, et al. Genetic structure of SARS-CoV-2 reflects clonal superspreading and multiple independent introduction events, North-Rhine Westphalia, Germany, February and March 2020. *Euro Surveill* 2020;25:2000746.
- [8] Beldomenico PM. Do superspreaders generate new superspreaders? A hypothesis to explain the propagation pattern of COVID-19. *Int J Infect Dis* 2020;96:461–3.
- [9] Kwok KO, Chan HHH, Huang Y, et al. Inferring super-spreading from transmission clusters of COVID-19 in Hong Kong, Japan and Singapore. *J Hosp Infect* 2020;105:682–5.
- [10] Adam D, Wu P, Wong J, et al. Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in Hong Kong. *Res Sq* 2020 May 21. doi:[10.21203/rs.3.rs-29548/v1](https://doi.org/10.21203/rs.3.rs-29548/v1).
- [11] Leclerc QJ, Fuller NM, Knight LE. CMMID COVID-19 Working Group. In: Funk S, Knight GM, editors. What settings have been linked to SARS-CoV-2 transmission clusters?, 5. *Wellcome Open Res*; 2020. p. 83.
- [12] Nishiura H, Oshitani H, Kobayashi T, et al. Closed environments facilitate secondary transmission of coronavirus disease 2019 (COVID-19). medRxiv 2020;02:28 Mar 3: 202020029272. doi:[10.1101/2020.02.28.20029272](https://doi.org/10.1101/2020.02.28.20029272).
- [13] Liu Y. Centre for mathematical modelling of infectious diseases nCoV working group. In: Funk S, Flasche S, editors. The contribution of pre-symptomatic infection to the transmission dynamics of COVID-2019, 5. *Wellcome Open Res*; 2020. p. 58.
- [14] Lee JY, Hong SW, Hyun M, et al. Epidemiological and clinical characteristics of Coronavirus Disease 2019 in Daegu, South Korea. *Int J Infect Dis* 2020;98:462–6.
- [15] Tsou T.P., Chen W.C., Huang A.S., Chang S.C.; Taiwan COVID-19 Outbreak Investigation Team. Epidemiology of the first 100 cases of COVID-19 in Taiwan and its implications on outbreak control. *J Formos Med Assoc* 2020;S0929-6646(20)30326-0.
- [16] Lam HY, Lam TS, Wong CH, et al. The epidemiology of COVID-19 cases and the successful containment strategy in Hong Kong-January to May 2020. *Int J Infect Dis* 2020;98:51–8.
- [17] Braunstein A, Ingresso A, Muntoni AP. Network reconstruction from infection cascades. *J R Soc Interface* 2019;16:20180844.
- [18] Skums P, Kirpich A, Icer Baykal P, Zelikovsky A, Chowell G. Global transmission network of SARS-CoV-2: from outbreak to pandemic. medRxiv. 2020;03 Mar 27:202022.20041145. doi:[10.1101/2020.03.22.20041145](https://doi.org/10.1101/2020.03.22.20041145).
- [19] Luo L, Liu D, Liao X, et al. Contact settings and risk for transmission in 3410 close contacts of patients with COVID-19 in Guangzhou, China: a prospective cohort study. *Ann Intern Med* 2020 10.7326/M20-2671.
- [20] Bi Q, Wu Y, Mei S, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect Dis*. 2020;20:911–19.
- [21] Shim E, Tariq A, Choi W, Lee Y, Chowell G. Transmission potential and severity of COVID-19 in South Korea. *Int J Infect Dis* 2020;93:339–44.
- [22] Bae S, Kim H, Jung TY, et al. Epidemiological Characteristics of COVID-19 Outbreak at Fitness Centers in Cheonan, Korea. *J Korean Med Sci* 2020;35:e288.
- [23] Furuse Y, Sando E, Tsuchiya N, et al. Clusters of coronavirus disease in communities, Japan, January–April 2020. *Emerg Infect Dis* 2020;26:2176–9.
- [24] Koh D. Migrant workers and COVID-19. *Occup Environ Med* 2020;77:634–6.
- [25] Zhou Q, Gao Y, Wang X, et al. Nosocomial infections among patients with COVID-19, SARS and MERS: a rapid review and meta-analysis. *Ann Transl Med* 2020;8:629.
- [26] 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.
- [27] Zhang L, Tao Y, Wang J, et al. Early characteristics of the COVID-19 outbreak predict the subsequent epidemic scope. *Int J Infect Dis* 2020;97:219–24.
- [28] Anderson RM, Hollingsworth TD, Baggaley RF, Madder R, Vegvari C. COVID-19 spread in the UK: the end of the beginning? *Lancet* 2020;396:587–90.
- [29] Petersen E, Wasserman S, Lee SS, et al. COVID-19-we urgently need to start developing an exit strategy. *Int J Infect Dis* 2020;96:233–9.
- [30] Cowling BJ, Ali ST, Ng TWY, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *Lancet Public Health* 2020;5:e279–88.