

Original Paper

Characterization of Unlinked Cases of COVID-19 and Implications for Contact Tracing Measures: Retrospective Analysis of Surveillance Data

Ka Chun Chong^{1*}, PhD; Katherine Jia^{2*}, MSc; Shui Shan Lee³, MBBS; Chi Tim Hung¹, MBBS; Ngai Sze Wong³, PhD; Francisco Tsz Tsun Lai⁴, PhD; Nancy Chau¹, MSc; Carrie Ho Kwan Yam¹, PhD; Tsz Yu Chow¹, BSc; Yuchen Wei¹, PhD; Zihao Guo¹, MSc; Eng Kiong Yeoh¹, MBBS

¹Centre for Health Systems and Policy Research, Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, Hong Kong

²MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, United Kingdom

³Stanley Ho Centre for Emerging Infectious Diseases, The Chinese University of Hong Kong, Hong Kong, Hong Kong

⁴Department of Pharmacology and Pharmacy, The University of Hong Kong, Hong Kong, Hong Kong

*these authors contributed equally

Corresponding Author:

Eng Kiong Yeoh, MBBS

Centre for Health Systems and Policy Research

Jockey Club School of Public Health and Primary Care

The Chinese University of Hong Kong

Central Avenue

Hong Kong

Hong Kong

Phone: 852 22528716

Email: yeoh_ek@cuhk.edu.hk

Abstract

Background: Contact tracing and intensive testing programs are essential for controlling the spread of COVID-19. However, conventional contact tracing is resource intensive and may not result in the tracing of all cases due to recall bias and cases not knowing the identity of some close contacts. Few studies have reported the epidemiological features of cases not identified by contact tracing (“unlinked cases”) or described their potential roles in seeding community outbreaks.

Objective: For this study, we characterized the role of unlinked cases in the epidemic by comparing their epidemiological profile with the linked cases; we also estimated their transmission potential across different settings.

Methods: We obtained rapid surveillance data from the government, which contained the line listing of COVID-19 confirmed cases during the first three waves in Hong Kong. We compared the demographics, history of chronic illnesses, epidemiological characteristics, clinical characteristics, and outcomes of linked and unlinked cases. Transmission potentials in different settings were assessed by fitting a negative binomial distribution to the observed offspring distribution.

Results: Time interval from illness onset to hospital admission was longer among unlinked cases than linked cases (median 5.00 days versus 3.78 days; $P < .001$), with a higher proportion of cases whose condition was critical or serious (13.0% versus 8.2%; $P < .001$). The proportion of unlinked cases was associated with an increase in the weekly number of local cases ($P = .049$). Cluster transmissions from the unlinked cases were most frequently identified in household settings, followed by eateries and workplaces, with the estimated probability of cluster transmissions being around 0.4 for households and 0.1-0.3 for the latter two settings.

Conclusions: The unlinked cases were positively associated with time to hospital admission, severity of infection, and epidemic size—implying a need to design and implement digital tracing methods to complement current conventional testing and tracing. To minimize the risk of cluster transmissions from unlinked cases, digital tracing approaches should be effectively applied in high-risk socioeconomic settings, and risk assessments should be conducted to review and adjust the policies.

(*JMIR Public Health Surveill* 2021;7(11):e30968) doi: [10.2196/30968](https://doi.org/10.2196/30968)

KEYWORDS

COVID-19; contact tracing; unlinked; superspreading; dispersion; surveillance; monitoring; digital health; testing; transmission; epidemiology; outbreak; spread

Introduction

As COVID-19 cases are still rising around the world and new variants are emerging, nonpharmaceutical interventions (NPIs) are essential for controlling the spread of COVID-19 [1,2] in many countries, especially when vaccination programs are impeded by factors such as vaccine hesitancy, reduced efficacy against some new variants, and vaccine shortage. NPIs such as personal protective equipment, social distancing, contact tracing followed by quarantine, border controls, travel restrictions, and enforced or recommended “stay-at-home” policies and “lockdowns” reduce transmission arising from individual contacts [3].

Hong Kong is a densely populated cosmopolitan city with extensive connections with mainland China and the rest of the world. The city experienced and successfully controlled three waves of COVID-19 in mid-January, March, and July 2020, respectively. Early in the first wave, the government closed its borders with the mainland and enforced a 14-day mandatory quarantine for all arrivals from mainland China and the close contacts of confirmed cases. Quarantine was then extended to arrivals from high-risk countries, and eventually to all arrivals, and entry was denied for all nonresidents during the second wave [1,2]. Unlike the first two waves, during which cases were predominantly imported cases and their close contacts, the third wave was characterized by community transmission in local clusters [3]. In response, the government stepped up public health measures, by implementing mandates for physical distancing, expanded community testing, enhanced case detection, contact tracing, and quarantine [4].

Active identification and isolation of infected persons and quarantine of close contacts reduced subsequent transmissions and brought the third wave under control within two months [3]. The importance and effectiveness of contact tracing have been documented elsewhere [5-8]. According to Aleta et al [5], aggressive social distancing measures, robust testing, contact tracing, and quarantine kept the disease within the capacity of the health care system in the absence of herd immunity.

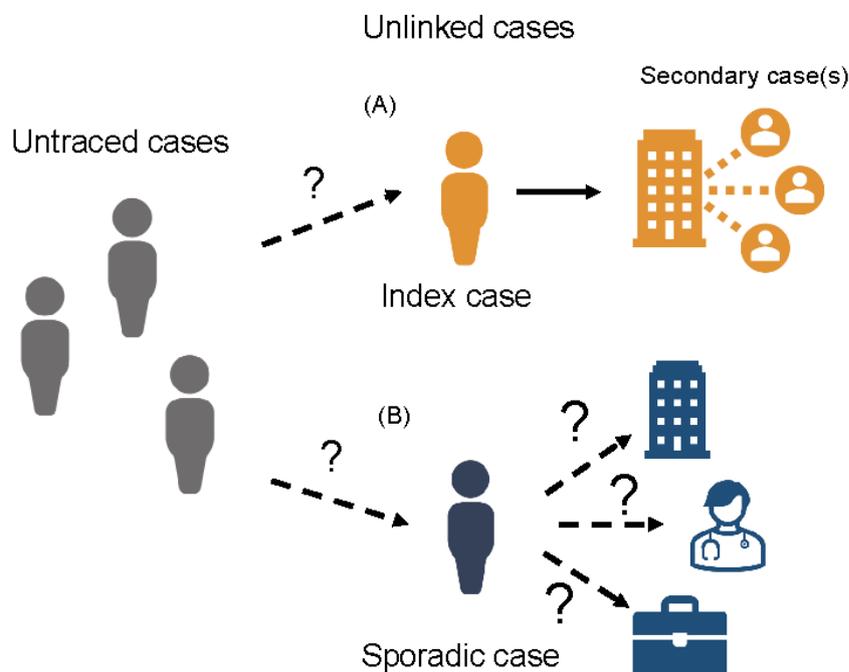
Ideally, all local cases can be epidemiologically linked, forming a meta-cluster of a closed transmission network in the absence of imported cases. However, in reality, resources for contact tracing are limited, and cases are not able to recollect (ie, recall

bias) or choose not to disclose all their close contacts. Close contacts may also not be known to the index case and therefore cannot be traced by conventional methods. As a consequence, some of the cases in a cluster outbreak or in unrecognized chains of transmissions could eventually emerge as “unlinked cases,” either as the index case in a new cluster outbreak or as a sporadic case without known involvement in any identified cluster (Figure 1). As shown by a modelling study [7], unlinked secondary cases can markedly impede outbreak control due to delayed isolation. Another simulation study predicted that at least 71% of close contacts of infectors needed to be traced to control the epidemic (ie, reproduction number reduced to <1) [8].

Hong Kong has implemented stringent containment measures since January 2020 [9,10]. Individuals who contacted a confirmed case up to two days before the initial case’s symptom onset would be put under mandatory quarantine and be tested [11]. Contact tracing became even more important during the third wave, which was characterized by more local transmissions, the impact of which should be evaluated.

It is critical to obtain empirical evidence on the effect of contact tracing on the containment of the epidemic in jurisdictions with control strategies (eg, Singapore, South Korea, and Hong Kong), where intensive contact tracing systems are applied to suppress transmission, to try to avoid the need for extensive restrictions to socioeconomic life (eg, lockdowns). Such data will also be invaluable for informing policies regarding the development of digital tracing technologies that have the potential to minimize the proportion of epidemiologically unlinked cases as well as the perpetuation of outbreaks. In this study, we investigated the clinical and epidemiological characteristics of the unlinked COVID-19 cases in Hong Kong, and estimated the transmission potential of the unlinked cases in different socioeconomic settings. Our study period covers the first three waves of the COVID-19 pandemic in Hong Kong. The findings provide empirical evidence on the value of contact tracing by comparing the epidemiological characteristics of the linked and unlinked cases, and estimating the probability of unlinked cases seeding community outbreaks. We also identified “hot spots” of outbreaks that were triggered by unlinked cases, which can inform risk assessments for designing future public health interventions to control transmission and prevent the health care system from being overwhelmed.

Figure 1. A schematic of the types of unlinked cases in a community. All local cases are supposed to be epidemiologically linked. Untraced cases (in grey) have silent transmissions (dashed arrows), which are the unlinked cases with no information about prior transmission chains. An unlinked case can either be (A) the index case of another cluster or (B) a sporadic case with no identified secondary transmissions. The transmission potential of the unlinked cases can be inferred by the size of the secondary transmission cluster triggered by one case (in orange). For sporadic unlinked cases (in blue), there is no information on whether they caused secondary transmissions or their location history. For each setting u , we estimated the transmission potential of the unlinked cases in that setting by considering a range of probabilities p_u (5%-95%) that the sporadic unlinked cases may have visited that setting before.



Methods

Data Sources

Rapid surveillance data containing the line listing of SARS-CoV-2 infections was obtained from the Centre of Health Protection of the Government of the Hong Kong Special Administrative Region. Individual-level data from January 23 to September 18, 2020, were extracted, including demographics (age, gender, residency), presence of chronic diseases (eg, hypertension, diabetes, and stroke), epidemiological characteristics (imported/local transmission, epidemiological linkages, hospital admission/detection date, mode of detection, types of exposure settings for clusters, any secondary generations linked, and cross-border travel history), clinical characteristics, and outcomes (symptomatic or asymptomatic status, illness onset date, clinical condition, death event). The polymerase chain reaction (PCR) cycle threshold (Ct) value collected on the date closest to the admission date after a 7-day window period was obtained. A lower Ct value suggests a higher viral load.

Asymptomatic cases were those diagnosed without any symptoms of COVID-19 at the time of detection. Modes of detection included the following: (1) medical surveillance, (2) general outpatient clinics (GOPC) or Accident & Emergency (A&E) departments, (3) enhanced surveillance in the private sector, and (4) others (eg, testing for inbound travelers or high-risk groups, and community screening programs). Close contacts of an infected case were quarantined in designated

locations and medical surveillance was arranged for other contacts. Time from illness onset to hospital admission was defined as the time interval (days) from illness onset date to hospital admission date.

Ethics approval was obtained from the Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee.

Clusters and Epidemiological Linkages

In accordance with definitions from our previous study, transmission clusters identified by contact tracing may belong to one of the following settings: household, dormitory, workplace, eatery, party, shopping, health care, entertainment, and education. Detailed descriptions of settings can be found in [12]. Household refers to the residential setting where individuals live together most of the time. A dormitory is a room or apartment where a group of unrelated residents reside. A workplace is any working space for and shared by staff. An eatery refers to places where customers stay for meals (eg, cafeterias and restaurants). A party refers to a private social gathering, while a place of entertainment refers to social activities at premises including bars and karaoke bars. A shopping venue can be a market or a department store. A health care setting is where health care services are provided (long-term care facilities are included in this category). In an education setting, teaching and learning activities are carried out—this category includes primary and secondary schools.

Case Definitions

A linked case is defined as a local secondary case (infectee) epidemiologically linked to a confirmed case (local/imported), either via personal contact or exposure to the same setting at the same time. An unlinked case refers to a local case without any source of infection identified by epidemiological linkage. An unlinked case could be detected either as an index case that has led to further secondary transmission(s) in different socioeconomic settings, or as a sporadic unlinked case without any secondary cases identified (Figure 1).

Statistical Analysis

As the third wave was predominated by local cases, we divided the data into two epochs: epoch one (from January 23 to June 18, 2020), covering the first and second waves, and epoch two (from June 19 to September 18, 2020), covering the third wave. For each epoch, we compared the categorical variables (ie, age group, sex, symptomatic/asymptomatic, presence of chronic conditions, mode of case detection, ever consulted private clinics prior to admission, critical/serious condition, and death event) of linked and unlinked cases by using chi-square tests, whereas we compared the PCR Ct values of linked and unlinked cases as a continuous variable by using independent *t* tests.

We examined the correlation between the weekly proportion of unlinked cases among all local infections and the weekly number of new local cases using the Spearman correlation coefficient (*r*). For the time interval from illness onset to hospital admission, we compared the survival functions between the linked and unlinked cases on their time to admission using the Kaplan-Meier method. The time to admission was further compared between the two groups when adjusted for age and gender in a Cox proportional hazard model.

The probability that an unlinked case would generate one or more secondary case(s) in a particular setting is estimated by assuming that the offspring distribution (*Z*) follows a negative binomial distribution [13,14]. Compared with other distributions, such as Poisson distribution, the dispersed distributional assumption can account for transmission heterogeneity via specification of the dispersion parameter (*k*) and effective reproductive number (*R*), which have been found to be more rigorous for modeling the offspring distribution when epidemics are characterized by superspreading events [15,16]. Following Lloyd-Smith et al [15], we assumed Z_u , the number of secondary cases generated by a case in a setting *u* (eg, household), follows a negative binomial distribution with mean=*R* and dispersion=*k*. $Z_u=0$ if a local case did not have any identified secondary cases and thus contributed to a cluster of size *j*=1. For $Z_u \geq 1$, *j* would be ≥ 2 (Figure 1).

By employing a branching process, we formulated the likelihood function (L_u) as follows:

$$L_u = [P(Z_u = 0)]^{N_{Z_u=0}} \prod_{z_u=2}^{N_{Z_u=z_u}} [P(Z_u = z_u)]^{N_{Z_u=z_u}}$$

where $N_{z_u=z_u}$ is the number of clusters of size j_u . $N_{z_u=z_u}$ can only be observed when $j_u \geq 2$ (ie, when the case was an index case with traced secondary cases that form a cluster at a specific setting *u*; Figure 1A). Conversely, $N_{z_u=0}$ is unobserved, as there is no information indicating whether the sporadic unlinked case had been in the particular setting, but only that no further secondary cases were known to be traced or linked to the sporadic case (Figure 1B). Therefore, for each setting *u*, we computed the number of sporadic unlinked cases that have been in or associated with that particular setting by multiplying the total number of sporadic unlinked cases with a percentage p_u (which varied from 5%-95% to account for uncertainty). Since places visited were not mutually exclusive, we assumed p_u to be independent of each other and considered each setting separately. We determined the probability of secondary transmission per unlinked case for each setting *u*, $P(Z_u \geq 1 | p_u, \hat{R}, \hat{k})$, by solving for the likelihood using the Markov chain Monte Carlo method.

Results

Characteristics of Unlinked Cases

There were 438 and 3345 confirmed local cases during epoch one (first and second wave) and epoch two (third wave), respectively. In each epoch, around one-third of the local cases were unlinked cases (ie, not epidemiologically linked with any clusters). Table 1 shows the comparison of the epidemiological characteristics of the linked and unlinked cases. A higher proportion of the unlinked cases were ≥ 18 years old ($P < .001$) when data from the two epochs were combined. More unlinked cases were male in epoch one ($P = .02$), but this association was not observed in epoch two.

In both epochs, the proportion of asymptomatic cases was lower among the unlinked cases than among the linked cases (9.2% versus 21.2%; $P < .001$). The unlinked cases also had higher Ct values (median 24.8 versus 23.3; $P < .001$), indicating a lower viral load at presentation. Linked cases were more likely to be detected through medical surveillance, whereas unlinked cases were more likely to be detected in GOPCs, A&E departments, and private clinics ($P < .001$). The unlinked cases were also more likely to have consulted private doctors prior to an admission (10.4% versus 4.9%; $P < .001$). For disease outcomes, the mortality rate did not differ significantly between linked and unlinked cases, but unlinked cases were more likely to progress to a critical or serious clinical condition (13.0% versus 8.2%; $P < .001$).

Table 1. Comparison of epidemiological characteristics between the linked and unlinked cases in epoch 1 (wave 1 and 2) and epoch 2 (wave 3).

Variables	Epoch 1 ^a (N=438)		P value ^c	Epoch 2 ^b (N=3345)		P value	Overall (N=3783)		
	Linked cases (N=267)	Unlinked cases (N=171)		Linked cases (N=2058)	Unlinked cases (N=1287)		Linked cases (N=2325)	Unlinked cases (N=1458)	P value
Age group (in years), n (%)			.07			<.001			<.001
<18	12 (4.5)	1 (0.6)		203 (9.9)	24 (1.9)		215 (9.2)	25 (1.7)	
18-49	161 (60.3)	114 (66.7)		855 (41.5)	552 (42.9)		1016 (43.7)	666 (45.7)	
50-64	62 (23.2)	41 (24.0)		565 (27.5)	419 (32.6)		627 (27.0)	460 (31.6)	
≥65	32 (12.0)	15 (8.8)		435 (21.1)	292 (22.7)		467 (20.1)	307 (21.1)	
Male, n (%)	128 (47.9)	101 (59.1)	.02	988 (48.0)	616 (47.9)	.94	1116 (48.0)	717 (49.2)	.48
Asymptomatic, n (%)	39 (14.6)	9 (5.3)	.002	450 (22.1)	124 (9.7)	<.001	489 (21.2)	133 (9.2)	<.001
Presence of chronic conditions, n (%)	88 (33.0)	54 (31.6)	.76	241 (11.7)	172 (13.4)	.16	329 (14.2)	226 (15.5)	.25
Mode of case detection, n (%)			<.001			<.001			<.001
Medical surveillance	145 (54.3)	9 (5.3)		828 (40.2)	65 (5.1)		973 (41.8)	74 (5.1)	
GOPC ^d and A&E	66 (24.7)	71 (41.5)		1004 (48.8)	764 (59.4)		1070 (46)	835 (57.3)	
Private clinics	7 (2.6)	35 (20.5)		196 (9.5)	375 (29.1)		203 (8.7)	410 (28.1)	
Others ^e	49 (18.4)	56 (32.7)		30 (1.5)	83 (6.4)		79 (3.4)	139 (9.5)	
Ever consulted private clinics prior to admission, n (%)	47 (17.6)	41 (24)	.10	67 (3.3)	111 (8.6)	<.001	114 (4.9)	152 (10.4)	<.001
PCR Ct ^f value, median (25th percentile to 75th percentile)	24.4 (19.1-31.1)	29.3 (22.3-33.8)	<.001	23.2 (18.8-28.8)	24.5 (19.8-29.7)	<.001	23.3 (18.8-29.0)	24.8 (19.9-30.1)	<.001
Critical/serious condition, n (%)	17 (6.4)	17 (9.9)	.17	173 (8.5)	171 (13.4)	<.001	190 (8.2)	188 (13.0)	<.001
Death, n (%)	3 (1.1)	2 (1.2)	.96	61 (3.0)	34 (2.6)	.59	64 (2.8)	36 (2.5)	.60

^aJanuary 23 to June 18, 2020, covering the first and second wave.

^bJune 19 to September 23, 2020, covering the third wave.

^cP values were from chi-square tests and independent two-sample *t* tests for categorical and continuous variables, respectively.

^dGOPC: general outpatient clinic; A&E: Accident & Emergency department.

^eOthers includes testing for inbound travelers, testing for high-risk groups, and community screening programs.

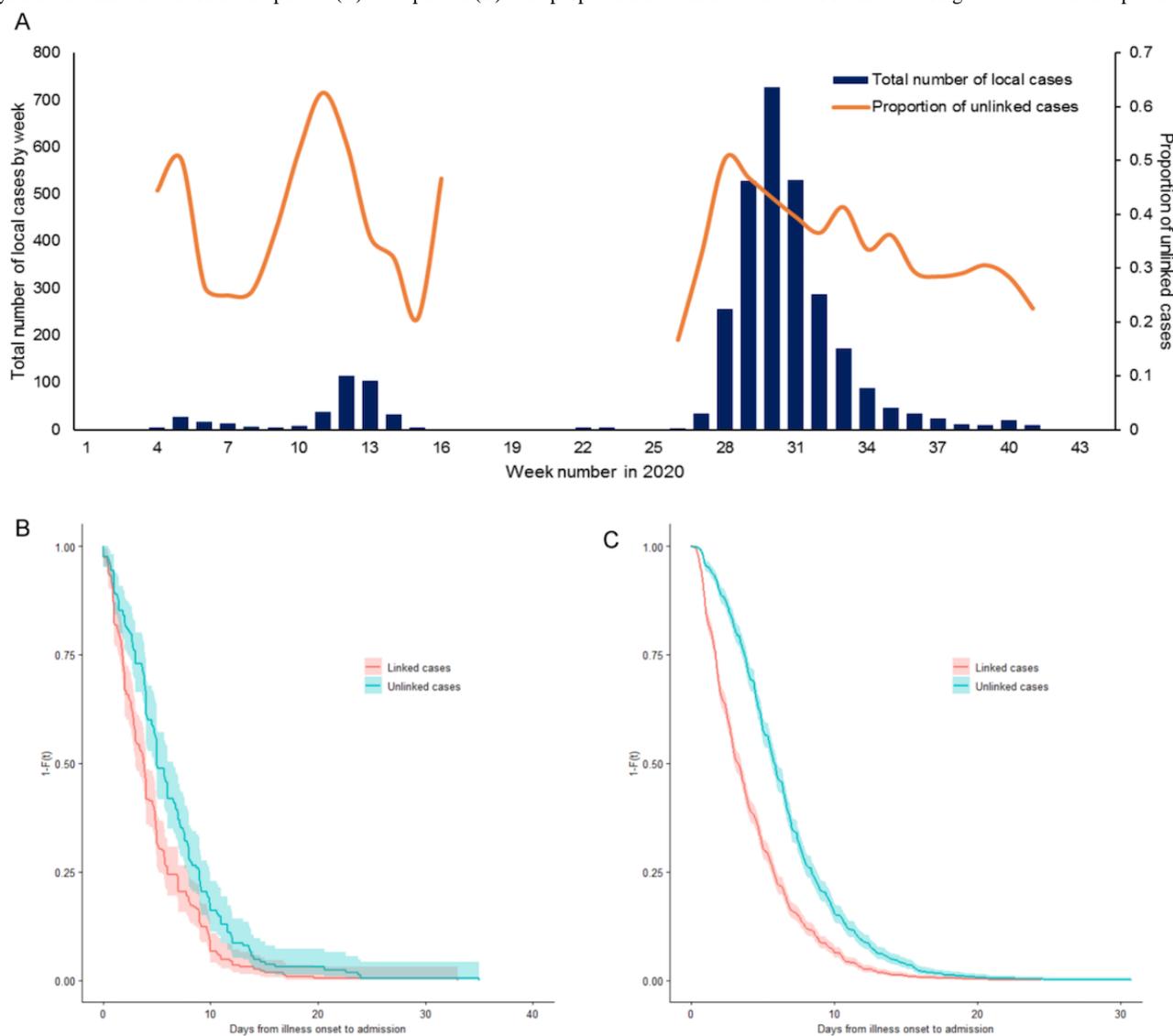
^fPCR Ct: polymerase chain reaction cycle threshold.

Relationship of Linked Cases With Epidemic Size and Time to Hospital Admission

The proportion of unlinked cases was significantly associated with the number of new local cases by week ($r=0.39$; $P=.049$; [Figure 2A](#)). In both epochs, the unlinked cases had significantly

longer time to hospital admission than the linked cases ($P<.001$, sex and age adjusted; [Figure 2B](#) and [Figure 2C](#)). The median time to admission was 3.78 days for the linked cases (95% CI 3.30-4.27) and 5.00 days (95% CI 4.37-5.63) for the unlinked cases in epoch one, and 3.26 days (95% CI 3.06-3.47) and 4.46 days (95% CI 4.28-4.64) in epoch two, respectively.

Figure 2. (A) Relationship between the proportion of unlinked cases and weekly number of new local cases (in the week of illness onset). A survival analysis of the time to admission in epoch 1 (B) and epoch 2 (C). The proportion of unlinked cases was smoothed using a 1-week window period.



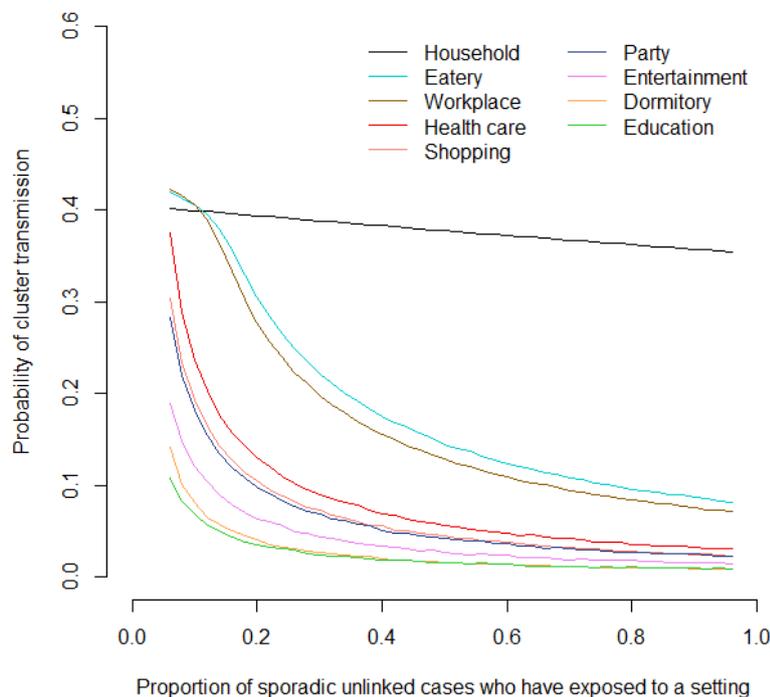
Transmission Potential of Unlinked Cases in Different Settings

Of the unlinked cases, 678 (46.5%) were identified as index cases with further secondary transmission. Settings where these cluster transmissions were most likely to be observed were households (493/678, 72.7%), eateries (54/678, 8.0%), and workplaces (48/678, 7.1%).

Households had the highest probability of outbreaks—each unlinked case had around 0.4 probability of causing a cluster of size $j \geq 2$ in household settings (Figure 3). In general, outbreak potential was sensitive to the values of p_u in all settings except households, where the probability of secondary transmission

was insensitive to p_u , consistent with a previous study showing a comparatively low potential for dissemination from a household to other settings [12]. Eateries and workplaces also had high outbreak potentials compared to other settings, with probabilities ranging from around 0.1-0.4 under different values of p_u . Health care, shopping, and party settings were particularly sensitive to p_u and could have a secondary transmission probability of >0.3 if a small proportion of sporadic unlinked cases had been involved in these settings (ie, low p_u). An outbreak potential probability as high as 0.3 (or any threshold defining a “high” probability) informs decisions regarding the public health measures needed to mitigate the risk of transmission in those settings.

Figure 3. Probabilities of cluster transmissions in different settings across the proportions of unlinked sporadic cases assumed to have visited each setting. The probability of cluster transmissions is defined as the probability of having caused a cluster size >1 for each unlinked case, based on the offspring distribution that follows a negative binomial distribution given the estimated dispersion parameter and effective reproductive number. Since there is no location history for the sporadic unlinked cases, the number of these cases at a particular setting u was simulated by multiplying the total number of sporadic unlinked cases with a factor pu that varied from 5%-95%.



Discussion

In this study, we characterized the epidemiological profile of local unlinked cases in the first three epidemic waves of COVID-19 in Hong Kong, a Special Administrative Region with a well-established case detection and contact tracing system used as a key containment strategy. Nevertheless, conventional contact tracing is unable to identify all close contacts at risk of infection, resulting in a significant number of unlinked cases, detected either as sporadic cases or as index cases of clusters of infections. We showed that unlinked cases were positively correlated with time to hospital admission, serious/critical clinical condition, and epidemic size. The higher proportion of patients progressing to a serious/critical condition among the unlinked cases was particularly pronounced in the third wave when cases were mostly due to local transmission.

Compared with the linked cases that were quickly identified through contact tracing and could be diagnosed before symptom onset, unlinked cases—for which no information regarding epidemiological linkage had been obtained through conventional contact tracing methods—could only be ascertained when they sought medical consultation upon becoming symptomatic. This could result in a longer delay in receiving appropriate care and treatment, which could result in an increased likelihood of progressing to a serious/critical clinical condition, which is consistent with reports regarding other respiratory infectious diseases [17,18]. However, the overall mortality rate was not found to be significantly different between linked and unlinked cases. One possible reason is that the health care system provides effective clinical management of patients with COVID-19. In Hong Kong, the mortality rate has been relatively low [19] and

the health care system still had the capacity to provide optimal care for patients with COVID-19 as the epidemic was brought under control rapidly with physical distancing measures during both epochs of the epidemic.

Another important finding was the association between the unlinked infections and the local epidemic situation. As indicated by Kretzschmar et al [6], the delay in detection and isolation would reduce the effectiveness of contact tracing measures for COVID-19 control. Across the different settings, the probability of an outbreak was highest in households, a phenomenon reported by studies in mainland China [20] and Korea [21]. However, it should be noted that household members are usually less subject to recall bias than others, leading to an overrepresentation of household secondary cases in our data set and an upward bias in the transmission potential for household settings. As shown by Adam et al [13], frequencies of secondary transmission were highest in households, while clusters were largest in social settings like restaurants, where superspreading events occurred. In line with our finding, a previous study showed that eateries and workplaces were the settings more likely to generate transmission cascades across more than a single setting in the second wave of the epidemic [12]. Such enclosed settings where large numbers of people were in close contact were a risk factor for transmission, while the congregation of people not known to each other was a risk factor for not being able to identify close contacts, allowing the dispersion of infections in a wide range of the settings.

Globally, it was estimated that 10% of cases outside China had caused 80% of secondary cases at the start of the pandemic [22]. Superspreading events have driven the pandemic significantly

and could quickly overwhelm the contact tracing system; for example, for a single large cluster, the epidemic investigation team traced up to 1000 people in Hong Kong [23]. Although venues like gymnasiums and sport venues were closed in Hong Kong and the government encouraged people to work from home during the pandemic, most businesses were operational. Dining in restaurants was permitted with restrictions in numbers, operational hours, and capacity. Unsurprisingly, we found a higher probability of cluster transmissions in eateries and workplaces compared to other settings. However, considering the scenario in which sporadic unlinked cases are not likely to have been involved in clusters in health care, shopping, and party settings, these settings would also have high potentials for cluster transmission. Our analysis thus provides a retrospective risk assessment for optimizing the relaxation of social/physical distancing restrictions in different settings; in these settings, risk can be managed with a robust and strictly enforced digital tracing system. Resurgences of infections could be prevented, obviating the need for a complete lockdown and the associated social and economic costs.

Many studies have focused on identifying hot spots for superspreading events. Although indoor environments with poor ventilation are generally riskier, human activities and behaviors (whether people have taken precautions) are even more crucial [12,24]. Therefore, health care, shopping, and party settings should have strictly implemented precautions for preventing cluster transmission, because of the risk and large number of contacts associated with such settings. Our findings offer empirical evidence for the differential influence of unlinked cases on epidemic control, which could help prioritize mitigation strategies.

Many jurisdictions in the Western Pacific Region heavily relied on stringent case detection, contact tracing, isolation, quarantine, and intensive testing programs together with border restrictions for COVID-19 control [19]. However, the conventional contact tracing approach is labor-intensive and feasible for low-incidence settings only [7], capturing only known contacts and being subject to cases' recall bias. Mobile phone data are a promising complementary tool. First, mobility data can inform incidence forecasts—for example, modeling combined with local mobility data showed that restaurants, bars, and gyms were hot spots for transmission [25]. Second, apps can be used to alert users instantaneously if their contacts are confirmed positive and recorded in the system [26,27]. This would prompt the alerted users to seek testing, especially for presymptomatic or asymptomatic transmissions [28-31]. In this study, we found that the percentage of asymptomatic infections was significantly lower in the unlinked cases than in the linked cases, primarily due to the fact that the unlinked cases would only be identified when they became symptomatic and sought medical consultation. However, even though contact tracing can identify some asymptomatic cases, others are missed by the system. App-based tools could prove useful for tracing these infections by providing improved coverage and timeliness and have the

potential to address the challenge of anonymous close contacts that the index patients do not know. Digital contact tracing has been deployed in many countries [32], including India [33], Switzerland [26], and the United Kingdom [27]. Nevertheless, we acknowledge that the effectiveness of app-based monitoring tools on outbreak control remains controversial when variations in exposure relative to the infectiveness period, testing accuracy, isolation adherence, and coverage of the app-based technology are taken into account [34,35]. Most importantly, the use of smartphone data for digital contact tracing in such contexts raises a number of ethical and privacy concerns [36]. In Hong Kong, for example, residents were skeptical of the LeaveHomeSafe app implemented after the third wave of the epidemic, despite the government's reassurance that user registration was not required and the check-in data would not be uploaded to the government's system or any other systems [37,38]. The critical role of digital technology in complementing conventional contact tracing needs to be evaluated, and ethical and privacy concerns must be addressed.

One major limitation of this study is that the location history for the sporadic unlinked cases was unavailable. We estimated the transmission probability of the unlinked cases by considering a range of scenarios that varied by the proportion of sporadic unlinked cases. In addition, we did not quantify the frequency of outbreaks (eg, the number of clusters per 1000 unlinked cases), which would be of interest for decision-making due to the unknown number of sporadic unlinked cases in each setting. Instead, we provided the probability that an unlinked case could result in secondary transmission to convey the outbreak potential in that setting. This indicates a need for an exposure tracking system that automatically computes the risk level for different settings. Second, cases were linked epidemiologically without validation through other approaches such as the routine use of phylogenetic analysis due to a lack of viral sequencing data [39]. Fingerprinting of SARS-CoV-2 via phylogenetic information can assist in the identification of chains of transmission; therefore, such analyses should be performed in future investigations. Third, the retrospective data cannot infer transmission potential for some settings that were already closed during the study period (eg, kindergartens that were closed since the early phase of the first wave). Fourth, some of the unlinked cases may have been infected by the imported cases exempted from quarantine and testing [40], limiting the generalizability of our results to other places with different border control measures.

In conclusion, our findings suggest a need to promote the use of digital tracing methods on top of current conventional testing and tracing. Contact tracing measures can potentially shorten the time to admission, reduce serious cases, and prevent further spread. With the considerable probability of secondary transmission from the unlinked cases, digital tracing measures should be strictly enforced in high-risk social settings and the local epidemic should be closely monitored such that the measures can be adjusted in a timely manner.

Acknowledgments

This work was funded by the Health and Medical Research Fund by the Food and Health Bureau, The Hong Kong Special Administrative Region (grant numbers COVID190105, INF-CUHK-1, 19181132), and the Collaborative Research Fund by the Research Grants Council (grant number C4139-20G). The Centre for Health Systems and Policy Research funded by The Tung Foundation is acknowledged for their support throughout the conduct of this study. The funder of the study had no role in study design, data collection, data analysis, data interpretation, writing of the manuscript, or the decision to submit for publication.

We would like to thank Kirran N Mohammad for editing the manuscript.

Authors' Contributions

KCC and EKY were responsible for the study concept and design. FTTL, NNSC, CHKY, and TYC were responsible for the literature search. FTTL, NNSC, CHKY, and TYC were responsible for acquisition of data. KCC and KMJ were responsible for statistical analysis and produced the output. KCC was responsible for figures. KCC, KMJ, SSL, CTH, NSW, and EKY were responsible for analysis and interpretation of data. KCC, KMJ, and EKY were responsible for drafting of the manuscript. SSL, CTH, NSW, FTTL, and YW were responsible for critical revision of the manuscript for important intellectual content. All authors contributed to revision of the manuscript and approved the final version for submission. All authors had full access to all the data in the study and were responsible for the decision to submit the manuscript for publication.

Conflicts of Interest

None declared.

References

1. Cowling B, Ali S, Ng T, Tsang T, Li J, Fong M, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *Lancet Public Health* 2020 May;5(5):e279-e288 [FREE Full text] [doi: [10.1016/S2468-2667\(20\)30090-6](https://doi.org/10.1016/S2468-2667(20)30090-6)] [Medline: [32311320](https://pubmed.ncbi.nlm.nih.gov/32311320/)]
2. Lam HY, Lam TS, Wong CH, Lam WH, Leung CME, Au KWA, 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 Sep;98:51-58 [FREE Full text] [doi: [10.1016/j.ijid.2020.06.057](https://doi.org/10.1016/j.ijid.2020.06.057)] [Medline: [32579906](https://pubmed.ncbi.nlm.nih.gov/32579906/)]
3. Centre for Health Protection. Latest situation of cases of COVID-19. 2021 May 2. URL: https://www.chp.gov.hk/files/pdf/local_situation_covid19_en.pdf [accessed 2021-05-26]
4. Government stays vigilant to cope with next wave of COVID-19 epidemic 2021. The Government of the Hong Kong Special Administrative Region. 2021 May 03. URL: <https://www.info.gov.hk/gia/general/202009/18/P2020091800996.htm> [accessed 2021-05-26]
5. Aleta A, Martín-Corral D, Pastore Y Piontti A, Ajelli M, Litvinova M, Chinazzi M, et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nat Hum Behav* 2020 Sep 05;4(9):964-971 [FREE Full text] [doi: [10.1038/s41562-020-0931-9](https://doi.org/10.1038/s41562-020-0931-9)] [Medline: [32759985](https://pubmed.ncbi.nlm.nih.gov/32759985/)]
6. Kretzschmar ME, Rozhnova G, Bootsma MCJ, van Boven M, van de Wijgert JHHM, Bonten MJM. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *The Lancet Public Health* 2020 Aug 06;5(8):e452-e459 [FREE Full text] [doi: [10.1016/s2468-2667\(20\)30157-2](https://doi.org/10.1016/s2468-2667(20)30157-2)]
7. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health* 2020 Apr;8(4):e488-e496. [doi: [10.1016/s2214-109x\(20\)30074-7](https://doi.org/10.1016/s2214-109x(20)30074-7)]
8. Keeling MJ, Hollingsworth TD, Read JM. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). *J Epidemiol Community Health* 2020 Oct 23;74(10):861-866 [FREE Full text] [doi: [10.1136/jech-2020-214051](https://doi.org/10.1136/jech-2020-214051)] [Medline: [32576605](https://pubmed.ncbi.nlm.nih.gov/32576605/)]
9. Lai C, Ng R, Wong M, Chong K, Yeoh Y, Chen Z, et al. Epidemiological characteristics of the first 100 cases of coronavirus disease 2019 (COVID-19) in Hong Kong Special Administrative Region, China, a city with a stringent containment policy. *Int J Epidemiol* 2020 Aug 01;49(4):1096-1105 [FREE Full text] [doi: [10.1093/ije/dyaa106](https://doi.org/10.1093/ije/dyaa106)] [Medline: [32601677](https://pubmed.ncbi.nlm.nih.gov/32601677/)]
10. Wong MCS, Ng RWY, Chong KC, Lai CKC, Huang J, Chen Z, et al. Stringent containment measures without complete city lockdown to achieve low incidence and mortality across two waves of COVID-19 in Hong Kong. *BMJ Glob Health* 2020 Oct 07;5(10):e003573 [FREE Full text] [doi: [10.1136/bmjgh-2020-003573](https://doi.org/10.1136/bmjgh-2020-003573)] [Medline: [33028700](https://pubmed.ncbi.nlm.nih.gov/33028700/)]
11. Centre for Health Protection. Frequently Asked Questions for Close Contacts under Mandatory Quarantine at Quarantine Centres. 2020. URL: https://www.chp.gov.hk/files/pdf/faq_for_close_contacts_eng.pdf [accessed 2021-05-26]
12. Wong NS, Lee SS, Kwan TH, Yeoh E. Settings of virus exposure and their implications in the propagation of transmission networks in a COVID-19 outbreak. *The Lancet Regional Health - Western Pacific* 2020 Nov;4:100052. [doi: [10.1016/j.lanwpc.2020.100052](https://doi.org/10.1016/j.lanwpc.2020.100052)]
13. Adam DC, Wu P, Wong JY, Lau EHY, Tsang TK, Cauchemez S, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nat Med* 2020 Nov 17;26(11):1714-1719. [doi: [10.1038/s41591-020-1092-0](https://doi.org/10.1038/s41591-020-1092-0)] [Medline: [32943787](https://pubmed.ncbi.nlm.nih.gov/32943787/)]

14. He D, Zhao S, Xu X, Lin Q, Zhuang Z, Cao P, et al. Low dispersion in the infectiousness of COVID-19 cases implies difficulty in control. *BMC Public Health* 2020 Oct 16;20(1):1558 [FREE Full text] [doi: [10.1186/s12889-020-09624-2](https://doi.org/10.1186/s12889-020-09624-2)] [Medline: [33066755](https://pubmed.ncbi.nlm.nih.gov/33066755/)]
15. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature* 2005 Nov 17;438(7066):355-359 [FREE Full text] [doi: [10.1038/nature04153](https://doi.org/10.1038/nature04153)] [Medline: [16292310](https://pubmed.ncbi.nlm.nih.gov/16292310/)]
16. Zhao S, Shen M, Musa SS, Guo Z, Ran J, Peng Z, et al. Inferencing superspreading potential using zero-truncated negative binomial model: exemplification with COVID-19. *BMC Med Res Methodol* 2021 Feb 10;21(1):30 [FREE Full text] [doi: [10.1186/s12874-021-01225-w](https://doi.org/10.1186/s12874-021-01225-w)] [Medline: [33568100](https://pubmed.ncbi.nlm.nih.gov/33568100/)]
17. Sopena N, Force L, Pedro-Botet ML, Barrufet P, Sauca G, García-Núñez M, et al. Sporadic and epidemic community legionellosis: two faces of the same illness. *Eur Respir J* 2007 Jan 27;29(1):138-142 [FREE Full text] [doi: [10.1183/09031936.00077206](https://doi.org/10.1183/09031936.00077206)] [Medline: [17005576](https://pubmed.ncbi.nlm.nih.gov/17005576/)]
18. Wu Z, Sha J, Yu Z, Zhao N, Cheng W, Chan T, et al. Epidemiological and virological differences in human clustered and sporadic infections with avian influenza A H7N9. *Int J Infect Dis* 2016 Aug;49:9-17 [FREE Full text] [doi: [10.1016/j.ijid.2016.05.022](https://doi.org/10.1016/j.ijid.2016.05.022)] [Medline: [27235087](https://pubmed.ncbi.nlm.nih.gov/27235087/)]
19. Yeoh EK, Chong KC, Chiew CJ, Lee VJ, Ng CW, Hashimoto H, et al. Assessing the impact of non-pharmaceutical interventions on the transmissibility and severity of COVID-19 during the first five months in the Western Pacific Region. *One Health* 2021 Jun;12:100213 [FREE Full text] [doi: [10.1016/j.onehlt.2021.100213](https://doi.org/10.1016/j.onehlt.2021.100213)] [Medline: [33506086](https://pubmed.ncbi.nlm.nih.gov/33506086/)]
20. World Health Organization. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). 2020 Feb 16. URL: <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf> [accessed 2021-05-26]
21. COVID-19 National Emergency Response Center, Epidemiology Case Management Team, Korea Centers for Disease Control Prevention. Coronavirus Disease-19: Summary of 2,370 Contact Investigations of the First 30 Cases in the Republic of Korea. *Osong Public Health Res Perspect* 2020 Apr;11(2):81-84 [FREE Full text] [doi: [10.24171/j.phrp.2020.11.2.04](https://doi.org/10.24171/j.phrp.2020.11.2.04)] [Medline: [32257773](https://pubmed.ncbi.nlm.nih.gov/32257773/)]
22. Endo A, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res* 2020 Jul 10;5:67 [FREE Full text] [doi: [10.12688/wellcomeopenres.15842.3](https://doi.org/10.12688/wellcomeopenres.15842.3)] [Medline: [32685698](https://pubmed.ncbi.nlm.nih.gov/32685698/)]
23. Ng KC, Sun F. Hong Kong races to contain "super-spreader" Covid-19 cluster from dance clubs, with police helping to trace up to 1,000 people. *South China Morning Post*. URL: <https://www.scmp.com/news/hong-kong/health-environment/article/3110962/hong-kong-facing-73-new-cases-covid-19-health> [accessed 2021-05-26]
24. Chu D, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet* 2020 Jun 27;395(10242):1973-1987 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)31142-9](https://doi.org/10.1016/S0140-6736(20)31142-9)] [Medline: [32497510](https://pubmed.ncbi.nlm.nih.gov/32497510/)]
25. Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* 2021 Jan 10;589(7840):82-87. [doi: [10.1038/s41586-020-2923-3](https://doi.org/10.1038/s41586-020-2923-3)] [Medline: [33171481](https://pubmed.ncbi.nlm.nih.gov/33171481/)]
26. Salathé M, Althaus C, Anderegg N, Antonioni D, Ballouz T, Bugnon E, et al. Early evidence of effectiveness of digital contact tracing for SARS-CoV-2 in Switzerland. *Swiss Med Wkly* 2020 Dec 14;150:w20457 [FREE Full text] [doi: [10.4414/smw.2020.20457](https://doi.org/10.4414/smw.2020.20457)] [Medline: [33327003](https://pubmed.ncbi.nlm.nih.gov/33327003/)]
27. Wymant C, Ferretti L, Tsallis D, Charalambides M, Abeler-Dörner L, Bonsall D. The epidemiological impact of the NHS COVID-19 App 2021. GitHub. 2021 May 08. URL: https://github.com/BDI-pathogens/covid-19_instant_tracing/blob/master/Epidemiological_Impact_of_the_NHS_COVID_19_App_Public_Release_V1.pdf [accessed 2021-05-26]
28. Liu Y, Funk S, Flasche S. The contribution of pre-symptomatic infection to the transmission dynamics of COVID-2019. *Wellcome Open Res* 2020 Apr 1;5:58. [doi: [10.12688/wellcomeopenres.15788.1](https://doi.org/10.12688/wellcomeopenres.15788.1)]
29. Yu X, Yang R. COVID-19 transmission through asymptomatic carriers is a challenge to containment. *Influenza Other Respir Viruses* 2020 Jul 15;14(4):474-475 [FREE Full text] [doi: [10.1111/irv.12743](https://doi.org/10.1111/irv.12743)] [Medline: [32246886](https://pubmed.ncbi.nlm.nih.gov/32246886/)]
30. He D, Zhao S, Lin Q, Zhuang Z, Cao P, Wang MH, et al. The relative transmissibility of asymptomatic COVID-19 infections among close contacts. *Int J Infect Dis* 2020 May;94:145-147 [FREE Full text] [doi: [10.1016/j.ijid.2020.04.034](https://doi.org/10.1016/j.ijid.2020.04.034)] [Medline: [32315808](https://pubmed.ncbi.nlm.nih.gov/32315808/)]
31. Moghadas SM, Fitzpatrick MC, Sah P, Pandey A, Shoukat A, Singer BH, et al. The implications of silent transmission for the control of COVID-19 outbreaks. *Proc Natl Acad Sci U S A* 2020 Jul 28;117(30):17513-17515 [FREE Full text] [doi: [10.1073/pnas.2008373117](https://doi.org/10.1073/pnas.2008373117)] [Medline: [32632012](https://pubmed.ncbi.nlm.nih.gov/32632012/)]
32. Nature Editorial. Show evidence that apps for COVID-19 contact-tracing are secure and effective. *Nature* 2020 Apr;580(7805):563. [doi: [10.1038/d41586-020-01264-1](https://doi.org/10.1038/d41586-020-01264-1)] [Medline: [32350479](https://pubmed.ncbi.nlm.nih.gov/32350479/)]
33. Garg S, Bhatnagar N, Gangadharan N. A Case for Participatory Disease Surveillance of the COVID-19 Pandemic in India. *JMIR Public Health Surveill* 2020 Apr 16;6(2):e18795 [FREE Full text] [doi: [10.2196/18795](https://doi.org/10.2196/18795)] [Medline: [32287038](https://pubmed.ncbi.nlm.nih.gov/32287038/)]

34. Braithwaite I, Callender T, Bullock M, Aldridge RW. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. *The Lancet Digital Health* 2020 Nov;2(11):e607-e621. [doi: [10.1016/s2589-7500\(20\)30184-9](https://doi.org/10.1016/s2589-7500(20)30184-9)]
35. Burgess M. Why the NHS Covid-19 contact tracing app failed. *Wired*. 2021 May 05. URL: <https://www.wired.co.uk/article/nhs-tracing-app-scrapped-apple-google-uk> [accessed 2021-05-26]
36. Parker MJ, Fraser C, Abeler-Dörner L, Bonsall D. Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic. *J Med Ethics* 2020 Jul 04;46(7):427-431 [FREE Full text] [doi: [10.1136/medethics-2020-106314](https://doi.org/10.1136/medethics-2020-106314)] [Medline: [32366705](https://pubmed.ncbi.nlm.nih.gov/32366705/)]
37. Ma M. Public health concerns more than privacy. *The Standard*. 2020 Nov 26. URL: <https://www.thestandard.com.hk/section-news/section/17/225095/Public-health-concerns-more-than-privacy> [accessed 2021-05-26]
38. The Government of the Hong Kong Special Administrative Region. Launch of "LeaveHomeSafe" COVID-19 exposure notification mobile app. 2021 May 5. URL: <https://www.info.gov.hk/gia/general/202011/11/P2020111100367.htm> [accessed 2021-05-26]
39. Yong SEF, Anderson DE, Wei WE, Pang J, Chia WN, Tan CW, et al. Connecting clusters of COVID-19: an epidemiological and serological investigation. *The Lancet Infectious Diseases* 2020 Jul;20(7):809-815. [doi: [10.1016/s1473-3099\(20\)30273-5](https://doi.org/10.1016/s1473-3099(20)30273-5)]
40. Expert says third wave likely originated from imported cases or quarantine exemptions. *The Standard*. URL: <https://tinyurl.com/2h4kx4ha> [accessed 2021-05-26]

Abbreviations

A&E: Accident & Emergency

Ct: cycle threshold

GOPC: general outpatient clinic

NPI: nonpharmaceutical intervention

PCR: polymerase chain reaction

Edited by G Eysenbach; submitted 04.06.21; peer-reviewed by VP Bongolan; comments to author 28.06.21; revised version received 12.07.21; accepted 19.09.21; published 16.11.21

Please cite as:

Chong KC, Jia K, Lee SS, Hung CT, Wong NS, Lai FTT, Chau N, Yam CHK, Chow TY, Wei Y, Guo Z, Yeoh EK

Characterization of Unlinked Cases of COVID-19 and Implications for Contact Tracing Measures: Retrospective Analysis of Surveillance Data

JMIR Public Health Surveill 2021;7(11):e30968

URL: <https://publichealth.jmir.org/2021/11/e30968>

doi: [10.2196/30968](https://doi.org/10.2196/30968)

PMID: [34591778](https://pubmed.ncbi.nlm.nih.gov/34591778/)

©Ka Chun Chong, Katherine Jia, Shui Shan Lee, Chi Tim Hung, Ngai Sze Wong, Francisco Tsz Tsun Lai, Nancy Chau, Carrie Ho Kwan Yam, Tsz Yu Chow, Yuchen Wei, Zihao Guo, Eng Kiong Yeoh. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 16.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.