Evolving epidemiological characteristics of COVID-19 in Hong Kong, January to August 2020

Kin On KWOK, Wan In WEI, Ying HUANG, Kai Man KAM, Ying Yang Emily CHAN, Steven RILEY, Ho Hin Henry CHAN, David Shu Cheong HUI, Samuel Yeung Shan WONG, Eng Kiong YEOH

Submitted to: Journal of Medical Internet Research
on: December 19, 2020

Table of Contents

Evolving epidemiological characteristics of COVID-19 in Hong Kong, January to August 2020

Kin On KWOK[1, 2, 3] PhD; Wan In WEI[1] MPhil; Ying HUANG[1] MPH; Kai Man KAM[1, 2] FRCPath; Ying Yang Emily CHAN[1] MD; Steven RILEY[4] PhD; Ho Hin Henry CHAN[1] PhD; David Shu Cheong HUI[1, 2] MD; Samuel Yeung Shan WONG[1] MD; Eng Kiong YEOH[1] MD

[1]JC School of Public Health and Primary Care The Chinese University of Hong Kong Hong Kong HK
[2]Stanley Ho Centre for Emerging Infectious Diseases The Chinese University of Hong Kong Hong Kong HK
[3]Shenzhen Research Institute The Chinese University of Hong Kong Hong Kong HK
[4]School of Public Health Imperial College London United Kingdom London GB

Corresponding Author:
Kin On KWOK PhD
JC School of Public Health and Primary Care
The Chinese University of Hong Kong
Room 419, 4/F, JC School of Public Health and Primary Care Building
Prince of Wales Hospital, Shatin, N.T.
Hong Kong
HK

Abstract

Background: COVID-19 plagued the globe, with non-pharmaceutical interventions and multiple SARS-CoV-2 clusters hinting on its evolving epidemiology.

Objective: To guide interventions, we studied such evolvement in Hong Kong in January-August 2020. We focused on containment delays (CDs) and serial intervals (SIs).

Methods: We retrieved the official case series and the Apple mobility data. The empirical CDs and SIs were fitted to theoretical distributions, and multivariate linear regression models were used to examine their associated factors. Effective reproductive number (Rt) was estimated with the best fitted distribution for SIs.

Results: We identified two epidemic waves, featured by imported cases and clusters of local cases respectively. Rt rose to peak at 2.39 (wave 1) and 3.04 (wave 2) respectively. Log-normal distribution best fitted the 1574 CDs (mean:5.18 days; SD:3.04) and the 558 SIs (17 negative) (mean:4.74 days; SD:4.24). CDs increased with older age (?60 years) (aOR:1.10; 95%CI:1.03,1.17), but decreased with general symptoms (aOR:0.89; 95%CI:0.85,0.94), involvement in a cluster (aOR:0.82-0.91) or detection in the public healthcare sector (aOR:0.76; 95%CI:0.72,0.80). SIs were shorter in wave 2 (aOR:0.79; 95%CI:0.72,0.86) and in tertiary transmission or beyond (aOR:0.68-0.89), but was lengthened by mobility (aOR:1.01; 95%CI:1.00,1.01).

Conclusions: Pre-symptomatic transmission and asymptomatic cases reminded the importance of remaining vigilant about COVID-19. The time-varying epidemiological parameters suggest the need to incorporate their temporal variations when depicting the epidemic trajectory. The slowing-down of the epidemic in late August 2020 suggested prompt government actions were crucial in suppressing resurgence. Clinical Trial: Not applicable

(JMIR Preprints 19/12/2020:26645)
DOI: https://doi.org/10.2196/preprints.26645

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
Please make my preprint PDF available to anyone at any time (recommended).
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Evolving epidemiological characteristics of COVID-19 in Hong Kong, January to August 2020

Kin On KWOK[1,2,3]*, Wan In WEI[1]*, Ying HUANG[1], Kai Man KAM[1,2], Emily Ying Yang CHAN[1], Steven RILEY[4,5], Henry Ho Hin CHAN[1], David Shu Cheong HUI[1,2], Samuel Yeung Shan WONG[1#], Eng Kiong YEOH[1#]

[1] JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong Special Administrative Region

[2] Stanley Ho Centre for Emerging Infectious Diseases, The Chinese University of Hong Kong, Hong Kong Special Administrative Region

[3] Shenzhen Research Institute, The Chinese University of Hong Kong, Hong Kong Special Administrative Region

[4] School of Public Health, Imperial College London, United Kingdom

[5] MRC Centre for Global infectious Disease Analysis and Abdul Latif Jameel Institute for Disease and Emergency Analytics, Imperial College London, United Kingdom

*Joint correspondence

Joint senior authors

KEYWORDS

SARS-CoV-2; COVID-19; evolving epidemiology; containment delay; serial interval; Hong Kong

INFORMATION FOR CORRESPONDENCE (I)

Name: Kin On KWOK

Email: kkokwok@cuhk.edu.hk

Telephone: +852 22528405

Address: Room 419, 4/F, JC School of Public Health and Primary Care Building, Prince of Wales Hospital, Shatin, N.T., Hong Kong

Fax: Nil

ORCID: 0000-0002-2804-5433

INFORMATION FOR CORRESPONDENCE (II)

Name: Wan In WEI

Email: vivian1628@cuhk.edu.hk

Telephone: +852 22528767

Address: Room 427, 4/F, JC School of Public Health and Primary Care Building, Prince of Wales Hospital, Shatin, N.T., Hong Kong

Fax: Nil

ORCID: Nil

HIGHEST DEGREE OBTAINED BY AUTHORS

Kin On KWOK - PhD

Wan In WEI - MPhil

Ying HUANG - MPH

Kai Man KAM - FRCPath

Emily Ying Yang CHAN - MD

Steven RILEY - PhD

Henry Ho Hin CHAN - PhD

David Shu Cheong HUI - MD

Samuel Yeung Shan WONG - MD

Eng Kiong YEOH - FRCP

ABSTRACT

Background: COVID-19 plagued the globe, with non-pharmaceutical interventions and multiple SARS-CoV-2 clusters hinting on its evolving epidemiology. Since the disease course is governed by important epidemiological parameters, including containment delays (CDs) (i.e. time between symptom onset and mandatory isolation) and serial intervals (SIs) (i.e. time between symptom onsets of infector-infectee pairs), understanding their temporal changes helps to guide interventions.

Objective: To characterize the epidemiology of the first two epidemic waves of COVID-19 in Hong Kong by (i) estimating the CDs, SIs, effective reproductive number (R_t), and the proportion of asymptomatic cases; (ii) identifying factors associated with the temporal changes of the CDs and SIs; and (iii) depicting the COVID-19 transmission by age assortativity and types of social settings.

Methods: We retrieved the official case series and the Apple mobility data of Hong Kong in January-August 2020. The empirical CDs and SIs were fitted to theoretical distributions, and factors associated with their temporal changes were quantified in terms of percentage contribution (PC), which was the percentage change in the predicted outcome (from multivariable regression models) relative to a predefined comparator. R_t was estimated with the best fitted distribution for SIs.

Results: The two epidemic waves were featured by imported cases and clusters of local cases respectively. R_t rose to peak at 2.39 (wave 1) and 3.04 (wave 2). The proportion of asymptomatic cases decreased from 36.4% (0-9 years) to 12.9% (≥80 years). Log-normal distribution best fitted the 1574 CDs (mean:5.18 days; standard deviation [SD]:3.04) and the 558 SIs (17 negative) (mean:4.74 days; SD:4.24). CDs decreased with involvement in a cluster (PC:10.08-20.73%) and case detection in the public healthcare sector (PC:27.56%; 95% confidence interval [CI]:22.52, 32.33). SIs decreased over time (6.70 days [wave 1] versus 4.35 days [wave 2]) and with the tertiary transmission or beyond (PC: -17.31% to -50.75%), but was lengthened by mobility (PC: 0.83%). Transmission within the same age band was high (18.1%); and households (69.9%), followed by social settings (20.3%), were two most common settings at which the transmission happened.

Conclusion: First, the factors associated with reduced CDs suggested government-enacted interventions (such as public mode of case detection) were useful to achieve outbreak control and should be further encouraged. Second, shorter SIs associated with the composite mobility index called for empirical surveys that disentangle the role of different contact dimensions in disease transmission. Third, the pre-symptomatic transmission and asymptomatic cases reminded the importance of remaining vigilant about COVID-19. Fourth, the time-varying epidemiological parameters suggest the need to incorporate their temporal variations when depicting the epidemic trajectory. Fifth, the high proportion of transmission events within the same age groups supported the ban on gatherings outside of households, and that in household pushed forward the need for residence-center preventive measures.

Word count = 450

INTRODUCTION

The novel coronavirus (SARS-CoV-2), causing Coronavirus Disease 2019 (COVID-19), first appeared in Wuhan, China in late December 2019 and quickly plagued the globe. The World Health Organisation declared COVID-19 a pandemic on 12 March 2020. As of 27 September 2020, there have been 32.7 million cases and almost one million deaths worldwide **[1]**. Countries are experiencing the resurgence of COVID-19. For example, in the week of 21-27 September 2020 alone, there were about 4.2 hundred million new cases in Europe **[1]**, triggering another round of lockdown measures **[2]**. Researchers have promptly summarized the case epidemiology during the early phase of the pandemic **[3,4]**. However, the enactment of non-pharmaceutical interventions and the presence of multiple genetic SARS-CoV-2 clusters **[5]** hint on important changes in the COVID-19 epidemiology.

Hong Kong is no exception to COVID-19. The **first epoch of the first wave** of COVID-19 in Hong Kong took off after the first imported case reported on 23 January 2020 **[6,7]**. Initially, the epidemic was under control after prompt bundled public health interventions **[8]**. With the number of infections surging worldwide in mid-March 2020, Hong Kong faced the **second epoch of the first wave** of infection. The compulsory laboratory tests to all passenger arrivals followed by 14-day compulsory quarantine triggered flocks of overseas residents to return, resulting in a small influx of imported cases. After peaking in March 2020, the number of cases remained low until a surge of local cases in July 2020, signifying **the second wave** of the epidemic in Hong Kong. This second wave represented the largest local outbreak in Hong Kong, which was likely attributable to initiation by imported cases coupled with the easing of social-distancing measures in July 2020.

The disease course of COVID-19 is governed by important epidemiological parameters, including containment delay and serial interval. The former has been shown to be associated with the infection source and number of doctor consultations **[6]**, which in turn vary as the epidemic progresses; whereas the latter varies with the specific virus types and subtypes **[9,10]**, the contact patterns between susceptible and infectious individuals **[10]**, as well as the enactment of non-pharmaceutical interventions during the epidemics **[11]**.

Parameterization of mathematical models that account for the temporal variation of epidemiological characteristics would improve the decision of mitigation strategies. Moreover, the containment delay increases opportunities for transmission and affects the effectiveness of control measures, whereas investigation of ways to reduce it could enhance control measures **[12]**. As such, we analyzed laboratory-confirmed COVID-19 case series in Hong Kong between January 2020 and August 2020 to quantify and identify the associated factors for the containment delay and serial intervals.

METHODS

Data retrieval

We analyzed the case series provided by the Hong Kong Centre for Health Protection (HKCHP) from 23 January 2020 to 2 August 2020, from which we extracted: demographics, case classification, travel history, epidemiological links among cases, date of symptom onset, date of isolation and the report date. Based on the order of settings embedded in a cluster and the case classification (cases, or close contact of cases), we compiled a line-list database of infector-infectee pairs (thereafter denoted as “paired data”).

Definitions

A laboratory-confirmed case (thereafter denoted as "a case") and a cluster were defined previously **[6]**. In short, **a case** refers to an individual with SARS-CoV-2 detected in a clinical specimen; and **a cluster** refers to at least two cases who are epidemiologically linked. Further, **a local cluster** is defined as a cluster which consists of at least one local case. A cluster encompasses one or more order of settings, referred to as **primary, secondary, tertiary, and quaternary settings**.

Containment delay and serial interval were defined previously **[6]**. In short, **containment delay** is the time elapsed between the first onset of symptoms and mandatory isolation of a case **[6]**; and **serial interval** is the time interval between the symptom onset of an infector and an infectee **[13]**. Further, the **secondary transmission** refers to the first generation of infections induced by a case; and infections caused by infectees of a secondary transmission are referred to as **tertiary transmission**. Accordingly, subsequent orders of transmission are named based on the aforementioned rationale. **Effective reproductive number (R_t)** is the average number of secondary cases generated by a primary case at any given time. It measured the real-time transmissibility in response to control measures. The epidemic will die down if R_t is consistently smaller than 1, and vice versa.

Classifications of cases

The HKCHP classified cases into six types according to their likely source of infection: imported cases, local cases, possibly local cases, and cases with epidemiological linkage with imported/ local/ possibly local cases. Based on their travel history during the 14 days preceding the first symptom onset and their involvement in local clusters, the latter four types of cases were re-classified **(Table 1)** such that there were only three types of cases: (i) imported; (ii) local and (iii) unclassified.

Symptom profile

Symptoms manifested by cases, if any, were grouped based on the International Statistical Classification of Diseases and Related Health Problems (ICD-10) and the national ambulatory medical care survey **[14]** into eight categories **(Table S1)**, including general (such as fever and headache) and respiratory (such as cough and sore throat) symptoms.

Mobility index

It is generated by Apple Maps, and represents the relative volume of routing requests (walking) for directions in Hong Kong to the baseline volume which on 13 January 2020 **[15].** The higher the index from the baseline indicates higher level of mobility. For ease of presentation, this index was further normalized with the value on 18 January 2020 as 100.

Statistical analysis

Characteristics of cases and epidemiological parameters were summarized with mean, standard deviation (SD), percentage, frequency and bootstrapped confidence interval. Missing isolation dates were replaced with rounding up the mean of available isolation days in a strata, which was made up of cases having the same likely source of infection, asymptomatic indicator, quarantine status, mode of detection and report dates. If there was no available isolation date in a strata, the mean day-differences between report dates and isolation dates were calculated, and the missing isolation dates were imputed as: the corresponding report dates minus the mean day-difference. For containment delay, only local cases who (i) were neither quarantined nor under medical surveillance and (ii) had non-negative containment delay were analyzed. For serial intervals, only settings which (i) linked two to four cases, (ii) had identifiable infectors and (iii) were related to local transmission were used to generate infector-infectee pairs for analysis. For R_t, only symptomatic local cases were considered.

A Markov Chain Monte Carlo with doubly interval censored likelihood **[16]** was adopted to fit the empirical containment delay and serial interval with four candidate distributions (gamma, lognormal, weibull and normal) with credible intervals (CrI) computed. The smallest value of the leave-one-out cross-validation information criterion (LOOIC) **[17]** indicates the best fitted distribution. Assuming the best fitted distribution for serial intervals as the weighted infectivity function, incidence data of local cases by dates of symptom onset with a 8-day window period (for reflecting the time between exposure to SARS-CoV-2 and symptom onset) was fitted by a novel and statistically robust tool in estimating R_t **[18]**.

Assuming the respective best fitted distributions as the likelihood function, multivariate linear regression models were used to examine the associated factors for containment delay and serial intervals. The effect of each factor was measured in terms of percentage contribution as previously described **[19]**, which is defined as the percentage change in the predicted value of the outcome (from the regression models) relative to a comparator set of covariate values (referred to as "the comparator"). The comparator referred to a transmission event fulfilling these conditions: the case (for containment delay) or the infector (for serial interval) was female, aged 0-30 years, free of chronic diseases, free of general and respiratory symptoms; the transmission happened in households, wave 1 and is of the secondary generation; and (for serial interval only) the mobility index being the baseline value on the onset date of the infector. Specifically, $\ln(\alpha_i + j + 1)$ and $\ln(\beta_i + k + 1)$ were defined as the response variables for the multivariable regression analyses, where α_i and β_i are the containment delay and serial interval of the i^{th} case respectively, j and k are the maximum absolute values for the non-positive containment delays and the non-positive serial intervals respectively.

A statistical significance of 0.05 was specified. All analyses were performed in R (version 3.6.3) **[20]** with RStan package (version 2.19.3).

Ethical statement

This study was approved by the Survey and Behavioral Research Ethics Committee of The Chinese University of Hong Kong (reference: SBRE-19-595).

Data statement

The availability of the dataset is subject to the approval from HKCHP and relevant government departments.

RESULTS

Characteristics of the two epidemic waves

The report date, 14 June 2020, defined two epidemic waves in Hong Kong: (i) 23 January 2020 - 14 June 2020 **(wave 1)**; and (ii) 15 June 2020- 2 August 2020 **(wave 2)**. This cut-off date separated the clusters in both waves without overlapping. Two epochs were further defined within wave 1: (i) 23 January 2020 - 29 February 2020 **(epoch 1)**; and (ii) 1 March 2020 - 14 June 2020 **(epoch 2)**. Wave 1 was initially dominated by imported cases from mainland China and local cases (epoch1) and subsequently by imported cases from European and American countries (epoch 2); whereas wave 2 was composed mainly of local cases **(Figure 1)**.

Hong Kong government has enacted multi-pronged interventions, and the major ones include school closures, work-from-home arrangements, and limiting customer flow and time for dine-in services **(Figure 1)**. On the other hand, the community had voluntarily reduced their mobility substantially, that the mobility index dropped from 100.2 initially to as far as 28.0 as the epidemic progressed **(Figure 1)**.

R_t was generally below one throughout the epidemic, but peaked at 2.39 in wave 1 and 3.04 in wave 2. This elevated R_t, coupled with short doubling time of the epidemic size in wave 2 **(Figure S1)**, suggested Hong Kong was on the verge of an uncontrolled outbreak during wave 2.

Characteristics of cases

As of 2 August 2020, 3512 cases were reported. The mean age was 43.91 years (SD: 20.27), with half of them being male (1748/3512), 6.2% (216/3512) had had doctor consultation before case confirmation, and 9.7% (339/3512) with chronic conditions **(Table S2)**. Most cases (75.4%; 2649/3512) were symptomatic, with fever (35.1%; 929/2649) and cough (30.4%; 805/2649) being the most common symptoms **(Table S3);** and the proportion of asymptomatic cases decreased from the younger age groups to the older ones (0-9 years: 36.4%; 10-19 years: 34.9%; 20-29 years: 21.9%; 30-39 years: 23.0%; 40-49 years: 19.5%; 50-59 years: 16.4%; 60-69 years: 12.9%; 70-79 years: 21.7%; and 80 years or above: 12.9%). In terms of the probable source of infection, 67.4% (2366/3512) were locally-acquired cases **(Table 1)**, amongst whom 1575 cases were symptomatic and who were neither quarantined nor under medical surveillance **(Table 2)**.

Containment delay

After excluding one case with negative containment delay, 1574 cases were included in the estimation. Log-normal distribution (mean: 5.18 days; SD: 3.04) fitted the empirical containment delay best (LOOIC:7525.6) **(Table S4)**, which ranged from 4.38 days (95% CI: 3.80, 4.95) for cases belonging to the secondary settings or beyond in a cluster to 6.48 days (95% CI: 6.15, 6.82) for cases identified through private mode of detection **(Table 3).**

Containment delay varied by several factors **(Table 3; Figure S2A)**. Notably, the percentage contribution (relative to the comparator) significantly **increases** with older age (≥61 years) by 11.29% (95% CI: 3.05, 20.09). On the contrary, the percentage contribution **decreases** with (i) the presence of general symptoms by 12.45% (95% CI: -1739, -7.27); (ii) the transmission event originating from a primary setting of a cluster by 10.08% (95% CI: 4.43, 15.44); (iii) the transmission event originating from a secondary setting or beyond of a cluster by 20.73% (95% CI: 10.93, 29.61); and (iv) the private mode of case detection by 27.56% (95% CI: 22.52, 32.33). There was a marginal difference (*P*=0.11) in wave 2 than in wave 1: the percentage contribution by wave 2 was -7.23% (95% CI: -15.58, 1.82).

Serial interval

Initially there were 847 settings linking at least two cases. After removing 143 settings consisting of solely imported cases, 72 settings with five or more cases (as simultaneous presence of large number of cases in a setting obscures the transmission link between cases) and 104 settings from which the infectors could not be identified (for example, there was no, or more than one, index case), 528 settings remained. From these 528 settings, 757 infector-infectee pairs were generated. After further removing four pairs with duplicated infectees and 195 pairs with missing onset dates of infectors/infectees, 558 paired data were included in the analysis. The mean number of infectees per infector was 1.46.

There were 17 negative serial intervals (range: -5 to -1 days). Log-normal distribution (mean: 4.74 days; SD: 4.24) fitted the overall empirical serial intervals best (LOOIC: 3095.5) **(Table S4)**. The serial intervals were 6.70 days (95% CI: 5.45, 7.95) and 4.35 days (95% CI: 4.00, 4.70) respectively in wave 1 and wave 2. Further, the subgroup estimates of serial intervals ranged from 2.18 days (95% CI: -0.52, 4.88) among quaternary transmission or beyond to 5.85 days (95% CI: 4.57, 7.13) among the infector-infectee pairs with chronic conditions among infectors **(Table 4)**.

Serial intervals varied by several factors **(Table 4; Figure S2B)**. They were significantly shorter in wave 2 than in wave 1: the percentage contribution by wave 2 was -33.9% (95% CI: -45.08, -21.63). Serial intervals also decreased with a tertiary transmission (percentage contribution: -17.31% [95% CI: -28.84, -4.75]), and a quaternary transmission or beyond (percentage contribution: -50.75% [95% CI: -23.45, -72.53]). On the other hand, the 8-day lagged relative mobility index, assuming an incubation period of 7.76 days **[21]** which reflected the possible exposure dates for each infector-infectee pair, lengthened the serial interval (percentage contribution: 0.83% [95% CI: 0.32, 1.33]).

Transmission events

Transmission events which occurred within the same age band was high (18.1%, number not shown) even on a five-year age band **(Figure 2A)**. Similar age transmission patterns were also observed in households **(Figure 2B)** and in social settings **(Figure 2C)**, but less obvious in work **(Figure 2D)** and institutional **(Figure 2E)** settings. In households setting, the transmission from infectors in older age groups (50-70 years) to infectees in younger age groups (15-40 years) was commonly observed **(Figure 2B)**. The age transmission matrix asymmetry in all settings well reflected this phenomenon but not in the reverse direction (Figure 2A). **(Figure 2A)**.

Importantly, households (69.9%; 390/558) and social settings (20.3%) were the two most common settings for transmission to take place in Hong Kong **(Table 4)** in both epidemic waves **(Table S5)**.

DISCUSSION

Summary of study findings

Understanding the evolving epidemiology is the groundwork to guide infection control policies. From the case-series between 23 January 2020 and 2 August 2020, we identified two waves of epidemics. R_t was in general below the outbreak threshold, suggesting that interventions, whether adopted voluntarily by the community **[22]** or institutionalized by the government **[8]**, successfully interrupted the transmission of COVID-19. As the epidemic sprawled in Hong Kong, containment delay and serial intervals were shortened over time. The shortening of the former was associated with the manifestation of general symptoms, involvement in clusters, and public mode of case detection; whereas that of the latter was associated with the later transmission generation and lower mobility. The occurrence of intra-age-group transmission outweigh that of inter-age-group transmission; and households and social settings altogether account for 90.14% of all identified transmission events.

Result implications

Our results have five implications. **First**, the factors associated with reduction in containment delay suggested that government-enacted interventions were useful to achieve COVID-19 outbreak control in Hong Kong and should be further encouraged. Containment delay had a major role in deciding whether an outbreak was controllable, that (assuming 80% of contacts can be traced) the chance of having it controlled fell from 89% to 31% if the containment delay increased from 3.43 days to 8.09 days **[23]**. The difference in containment delay experienced by cases detected from different modes (private: 6.48 days; public :4.80 days) pinpointed the worthiness of continual investment in public modes of case detection, including setting up community testing centers and mobile specimen collection stations. **Further**, as reflected from the percentage contribution by cases involved in a cluster (-10.08% to -20.73%), contact tracing and the follow-up quarantine on individuals with epidemiological links with cases were useful in reducing containment delay, that they should be enacted with high tracing ratio (preferably ≥80%) and when there were only few initial cases **[23]**.

Second, the association between decreasing serial intervals over time (from 6.52 days in wave 1 to 4.35 days in wave 2) and lower mobility aligned with the contention that serial intervals are shortened by non-pharmaceutical interventions **[11]**. Reduced serial intervals indicated faster case generation replacement, which may be attributable to the institutionalized social-distancing policies **(Figure 1)** that diminished the geographical stretch of citizens. This is a tradeoff between time spent in a place versus number of places visited in limited time, such that citizens may stay in confined locations (such as home) longer. Although the underlying mechanisms remain undetermined, we hypothesize that confined geographical outstretch would, in reality, intensify the proximity of contacts between successive case generations. This hypothesis is in line with earlier findings that more time spent with the index case closely shortened serial intervals of influenza **[10]**. While the adopted mobility index remained composite, empirical contact surveys, such as the ones by Mossong and colleagues **[24]** and Kwok and colleagues **[19]**, that disentangle the interplay of contact dimensions will advance research in the area of evolving epidemiology. On one hand, it is of interest to dissect the role of different contact dimensions in disease transmission; on the other, the potential of other composite social mobility measures, such as The Twitter Social Mobility Index **[25]**, in estimating epidemiological parameters should be explored.

that disentangle the interplay of contact dimensions will advance research in the area of evolving epidemiology.

Third, the presence of negative serial intervals, suggestive of pre-symptomatic transmission, and asymptomatic cases reminded the community to stay on guard against the resurgence of COVID-19. Infections happening before symptom onset would impede the effectiveness of control measures. This pre-symptomatic fraction appeared to be low so far (3.05%; 17/558), but it can be up to 12.6% **[26]**. Meanwhile, the proportion of asymptomatic local cases was 12.0% (285/2366) and the accumulated infection so far (that is, 3512 cases among more than seven million population in Hong Kong) is not high enough to confer herd immunity when compared with the conservative 5.66% previously suggested **[27]**). Together, coupled with the fact that COVID-19 vaccine is not yet available, these results suggested that the community should remain vigilant against any resurgence.

Fourth, temporal variations of key epidemiological parameters should be considered. It is common to assume the empirical data (of the epidemiological parameters) to resemble theoretical distributions. However, the correlation of (i) lower mobility (as proxy to voluntary or compulsory social-distancing en masse), (ii) involvement in local clusters (as proxy to early government actions on case-tracing), and (iii) government-level case identification with either containment delay or serial intervals observed in this study suggested that epidemiological parameters are dynamic throughout the epidemic. Furthermore, with SARS-CoV-2 transcending international borders, it mutated and diverged into different clusters or subtypes **[28–30]**. These subtypes differ by their intrinsic properties, exhibiting variations in COVID-19 epidemiology **[30]**. Therefore, caution must be taken to interpret findings from infectious disease models which assumed static epidemiological parameters.

Fifth, the high proportion of transmission events within the same age groups supported the ban on gatherings outside of households. The tendency of intra-age-group transmission of COVID-19 echoed with the largely contact assortativity by age of the Hong Kong population **[19]**, and is in line with an earlier conclusion **[31]**: social contact should be considered together with age when it comes to the driving force for the incidence of respiratory infections. Further, the asymmetric age transmission matrix revealed that children rarely infected others in the first two epidemic waves. This phenomenon may be attributable to the continual school closure, resulting in less-than-expected social interactions among children in Hong Kong. With the resumption of schools, hence that of more social-mixing among children, transmission chain branching from children is possible such that elderly, who are prone to COVID-19 mortality **[32]**, contacting children frequently should be prioritized to receive the COVID-19 vaccine. In addition, the abundance of transmission events in households pushed forward the need for residence-centered preventive measures, as per the lesson from the aerosol transmission (through defective plumbing) of the 2003 severe acute respiratory syndrome virus in the Amoy Gardens housing complex in Hong Kong **[33]**.

Study limitations

There are two study limitations which bear mentioning. **First**, the data in this study, including self-reported symptoms and contact history, were subject to recall bias. However, the medical surveillance in place to monitor contacts of cases may lessen the data uncertainties. **Second**, some cases might not have been captured by the present surveillance system in Hong Kong due to the under-diagnosis of mild cases and asymptomatic individuals, who might play a role in the transmission chain involving unlinked local cases.

Word count = 3546

LEGEND

Table 1. Re-classification regime of cases

Table 2. Classification of 3512 cases in Hong Kong, as of 2 August 2020

Table 3. Estimates of and factors associated with containment delay based on 1574 cases

Table 4. Estimates of and factors associated with serial interval based on 558 infector-infectee pairs

Figure 1. Epidemic curve of COVID-19 and timeline for major interventions in Hong Kong

Figure 2. Age-specific transmission events (A) in all settings; (B) in households; (C) in social settings; (D) in work settings; and (E) in institutions.

Table S1. Grouping of symptoms

Table S2. Frequency of chronic diseases among the 339 cases with chronic conditions

Table S3. Frequency of symptoms among the 2649 symptomatic cases

Table S4. Non-truncated and bootstrapped estimates for containment delay and serial interval

Table S5. Number of infector-infectee pairs included in the serial interval analysis, stratified by order of transmissions and settings

Figure S1. Cumulative number of COVID-19 cases by number of days elapsed since the first reported case(s) in two epidemic waves

Figure S2. Percentage contribution of covariates in changing the length of (A) containment delay; and (B) serial interval.

DECLARATION OF INTEREST

None declared.

FUNDING

This study is funded by Direct Grant for Research (number: 2019.020) of The Chinese University of Hong Kong.

ACKNOWLEDGEMENT

The authors would like to thank Hong Kong Centre for Health Protection for providing the data, and Margaret TSOI for technical support. Kin On KWOK acknowledges support by CUHK Direct grant for research (2019.020), Health and Medical Research Fund (reference number: INF-CUHK-1, 17160302, 18170312), General Research Fund (reference number: 14112818), and Wellcome Trust (UK, 200861/Z/16/Z).

AUTHOR CONTRIBUTION

Kin On KWOK (KOK), Samuel Yeung Shan WONG (SYSW), and Eng Kiong YEOH (EKY) conceptualized the study and collected the data; KOK, Wan In WEI (WIW), Y Huang (YH), and Henry Ho Hin CHAN (HHHC) analyzed the data; KOK and WIW wrote the first draft the manuscript; and all authors interpreted the data, edited the manuscript and provided critical comments.

ABBREVIATIONS

COVID-19: Coronavirus Disease 2019

HKCHP: Hong Kong Centre for Health Protection

CrI: credible interval

LOOIC: leave-one-out cross-validation information criterion

SD: standard deviation

CI: confidence interval

REFERENCES

1. World Health Organization. Weekly Epidemiological Update - Coronavirus disease 2019 (COVID-19) [Internet]. 2020 Sep [cited 2020 Sep 30]. Available from: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200928-weekly-epi-update.pdf?sfvrsn=9e354665_6
2. BBC News. Europe lockdown: New coronavirus rules country-by-country [Internet]. BBC News. 2020 [cited 2020 Sep 30]. Available from: https://www.bbc.com/news/explainers-53640249
3. Park M, Cook AR, Lim JT, Sun Y, Dickens BL. A Systematic Review of COVID-19 Epidemiology Based on Current Evidence. J Clin Med Res [Internet]. 2020 Mar 31;9(4). Available from: http://dx.doi.org/10.3390/jcm9040967
4. Khalili M, Karamouzian M, Nasiri N, Javadi S, Mirzazadeh A, Sharifi H. Epidemiological characteristics of COVID-19: a systematic review and meta-analysis. Epidemiol Infect. 2020 Jun 29;148:e130.
5. Yang X, Dong N, Chan EW-C, Chen S. Genetic cluster analysis of SARS-CoV-2 and the identification of those responsible for the major outbreaks in various countries. Emerg Microbes Infect. 2020 Dec;9(1):1287–99.
6. Kwok KO, Wong VWY, Wei WI, Wong SYS, Tang JW-T. Epidemiological characteristics of the first 53 laboratory-confirmed cases of COVID-19 epidemic in Hong Kong, 13 February 2020. Euro Surveill [Internet]. 2020 Apr;25(16). Available from: http://dx.doi.org/10.2807/1560-7917.ES.2020.25.16.2000155
7. Kwok KO, Wong V, Wei VWI, Wong SYS, Tang JW-T. Novel coronavirus (2019-nCoV) cases in Hong Kong and implications for further spread [Internet]. Vol. 80, Journal of Infection. 2020. p. 671–93. Available from: http://dx.doi.org/10.1016/j.jinf.2020.02.002
8. Wong SYS, Kwok KO, Chan FKL. What can countries learn from Hong Kong's response to the

COVID-19 pandemic? [Internet]. Vol. 192, Canadian Medical Association Journal. 2020. p. E511–5. Available from: http://dx.doi.org/10.1503/cmaj.200563

9. Vink MA, Bootsma MCJ, Wallinga J. Serial intervals of respiratory infectious diseases: a systematic review and analysis. Am J Epidemiol. 2014 Nov 1;180(9):865–75.
10. Levy JW, Cowling BJ, Simmerman JM, Olsen SJ, Fang VJ, Suntarattiwong P, et al. The serial intervals of seasonal and pandemic influenza viruses in households in Bangkok, Thailand. Am J Epidemiol. 2013 Jun 15;177(12):1443–51.
11. Ali ST, Wang L, Lau EHY, Xu X-K, Du Z, Wu Y, et al. Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. Science. 2020 Aug 28;369(6507):1106–9.
12. Bonačić Marinović A, Swaan C, van Steenbergen J, Kretzschmar M. Quantifying reporting timeliness to improve outbreak control. Emerg Infect Dis. 2015 Feb;21(2):209–16.
13. Fine PEM. The interval between successive cases of an infectious disease. Am J Epidemiol. 2003 Dec 1;158(11):1039–47.
14. Meads S, McLemore T. The national ambulatory medical care survey: symptom classification. Vital Health Stat 2. 1974 May;(63):1–41.
15. Apple [Internet]. Mobility trends reports. [cited 2020 Aug 20]. Available from: https://www.apple.com/covid19/mobility
16. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. Int J Infect Dis [Internet]. 2020 Mar 4; Available from: http://dx.doi.org/10.1016/j.ijid.2020.02.060
17. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat Comput. 2016 Aug 30;27(5):1413–32.
18. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am J Epidemiol. 2013 Nov

1;178(9):1505–12.

19. Kwok KO, Cowling B, Wei V, Riley S, Read JM. Temporal variation of human encounters and the number of locations in which they occur: a longitudinal study of Hong Kong residents. J R Soc Interface [Internet]. 2018 Jan;15(138). Available from: http://dx.doi.org/10.1098/rsif.2017.0838
20. Core Team R, Others. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available. 2013;
21. Qin J, You C, Lin Q, Hu T, Yu S, Zhou X-H. Estimation of incubation period distribution of COVID-19 using disease onset forward time: a novel cross-sectional and forward follow-up study. medRxiv [Internet]. 2020 Mar 10; Available from: http://dx.doi.org/10.1101/2020.03.06.20032417
22. Kwok KO, Li KK, Chan HHH, Yi YY, Tang A, Wei WI, et al. Community Responses during Early Phase of COVID-19 Epidemic, Hong Kong. Emerg Infect Dis. 2020 Jul;26(7):1575–9.
23. 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. Lancet Glob Health. 2020 Apr;8(4):e488–96.
24. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 2008 Mar 25;5(3):e74.
25. Xu P, Dredze M, Broniatowski DA. The Twitter Social Mobility Index: Measuring Social Distancing Practices With Geolocated Tweets. J Med Internet Res. 2020 Dec 3;22(12):e21499
26. Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. Emerg Infect Dis. 2020 Jun;26(6):1341–3.
27. Kwok KO, Lai F, Wei WI, Wong SYS, Tang JWT. Herd immunity - estimating the level required to halt the COVID-19 epidemics in affected countries. J Infect. 2020 Jun;80(6):e32–3.
28. Forster P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-2

genomes. Proc Natl Acad Sci U S A. 2020 Apr 28;117(17):9241–3.

29. To KK-W, Chan W-M, Ip JD, Chu AW-H, Tam AR, Liu R, et al. Unique SARS-CoV-2 clusters causing a large COVID-19 outbreak in Hong Kong. Clin Infect Dis [Internet]. 2020 Aug 5; Available from: http://dx.doi.org/10.1093/cid/ciaa1119
30. Tang X, Wu C, Li X, Song Y, Yao X, Wu X, et al. On the origin and continuing evolution of SARS-CoV-2. Natl Sci Rev [Internet]. 2020 Mar 3 [cited 2020 May 9]; Available from: https://academic.oup.com/nsr/article-abstract/doi/10.1093/nsr/nwaa036/5775463
31. Kwok KO, Cowling BJ, Wei VWI, Wu KM, Read JM, Lessler J, et al. Social contacts and the locations in which they occur as risk factors for influenza infection. Proc Biol Sci. 2014 Aug 22;281(1789):20140709.
32. Ho FK, Petermann-Rocha F, Gray SR, Jani BD, Katikireddi SV, Niedzwiedz CL, et al. Is older age associated with COVID-19 mortality in the absence of other risk factors? General population cohort study of 470,034 participants. PLoS One. 2020 Nov 5;15(11):e0241824.
33. Yu ITS, Li Y, Wong TW, Tam W, Chan AT, Lee JHW, et al. Evidence of airborne transmission of the severe acute respiratory syndrome virus. N Engl J Med. 2004 Apr 22;350(17):1731–9.

Table 1. Re-classification regime of cases

Original classification	Reclassified classification	Number of cases (N = 3512)
Imported cases	Imported cases	1045
Local cases	Local cases	901
Possibly local cases	Local cases [a]	2
	Imported cases [b]	71
	Unclassified cases [c]	30
Close contact of imported cases	Local	31
Close contact of local cases	Local	1370
Close contact of possibly local cases	Local	62

[a] Cases without travel history during the 14 days before their first symptom onset.

[b] Cases had travel history during the 14 days before their first symptom onset and were not involved in any local cluster.

[c] Cases who either (i) had travel history during the 14 days before their first symptom onset and were linked to a local cluster; or (ii) did not have onset/arrival date.

Table 2. Classification of 3512 cases in Hong Kong, as of 2 August 2020

	Wave 1 (n=1110) (%)	Wave 2 (n=2402) (%)	Total (n=3512) (%)
Imported case			
Symptomatic	554 (49.9)	94 (3.9)	648 (18.5)
Asymptomatic	189 (17.0)	231 (9.6)	420 (12.0)
Missing	0 (0.0)	48 (2.0)	48 (1.4)
Local case			
Symptomatic			
Quarantine[a]	31 (2.8)	158 (6.6)	189 (5.4)
Non-quarantine			
Medical surveillance	91 (8.2)	125 (5.2)	216 (6.2)
Others[b]	176 (15.9)	1399 (58.2)	1575 (44.8)
Asymptomatic	39 (3.5)	246 (10.2)	285 (8.1)
Missing	0 (0.0)	101 (4.2)	101 (2.9)
Unclassified	30 (2.7)	0 (0.0)	30 (0.9)

[a]Included home/hotel confinee cases, camp/center quarantine cases and cases released from home/hotel quarantine before getting infected.

[b]Cases with these modes of case detection: enhanced lab surveillance, enhanced surveillance in private, enhanced surveillance at GOPC and AED, meeting reporting criteria, diagnosed in private clinic, test at private clinic and being under Tier 7 classification.

Table 3. Estimates of and factors associated with containment delay based on 1574 cases.

	Number of cases (%)	Subgroup-specific estimates (95% empirical CI)	Percentage contribution (95% CI)
Age group (years)			
0-30	301 (19.1)	4.87 (4.55, 5.18)	referent [c]
31-45	358 (22.7)	4.86 (4.58, 5.14)	0.04 (-7.82, 8.48)
46-60	448 (28.5)	5.27 (4.99, 5.55)	6.86 (-1.14, 15.41)
61+	467 (29.7)	5.52 (5.21, 5.82)	11.29 (3.05, 20.09)
Sex			
Female	834 (53.0)	5.15 (4.95, 5.35)	referent [c]
Male	740 (47.0)	5.20 (4.98, 5.42)	2.11 (-3.17, 7.63)
Number of general symptoms			
0	985 (62.6)	5.38 (5.20, 5.56)	referent [c]
≥1	589 (37.4)	4.83 (4.58, 5.08)	-12.45 (-17.39, -7.27)
Number of respiratory symptoms			
0	1002 (63.7)	5.17 (4.99, 5.35)	referent [c]
1	459 (29.2)	5.20 (4.91, 5.49)	-1.25 (-7.09, 4.91)
≥2	113 (7.2)	5.08 (4.54, 5.62)	1.86 (-8.77, 13.57)
Order of settings			
None	423 (26.9)	5.71 (5.43, 6.00)	referent [c]
Primary	1045 (66.4)	5.03 (4.85, 5.22)	-10.08 (-15.44, -4.43)
Secondary or beyond	106 (6.7)	4.38 (3.80, 4.95)	-20.73 (-29.61, -10.93)
Mode of case detection			
Private [a]	347 (22.0)	6.48 (6.15, 6.82)	referent [c]
Public [b]	1227 (78.0)	4.80 (4.64, 4.96)	-27.56 (-32.33, -22.52)
Wave			
One	176 (11.2)	5.29 (4.78, 5.80)	referent [c]
Two	1398 (88.8)	5.16 (5.00, 5.31)	-7.23 (-15.58, 1.82)

[a] Private mode includes diagnosis in private, enhanced surveillance in private and private test.

[b] Public mode includes enhanced lab surveillance, enhanced surveillance at GOPC and AED, meeting reporting criteria, Tier 6 and Tier 7.

[c] The referent is a common comparator across all variables.

Table 4. Estimates of and factors associated with serial interval based on 558 infector-infectee pairs

	Number of pairs (%)	Subgroup-specific estimates (95% empirical CI)	Percentage contribution (95% CI)
Age group of infector (years)			
0-30	77 (13.8)	4.12 (3.41, 4.83)	referent [c]
31-45	121 (21.7)	4.43 (3.53, 5.33)	-0.24 (-16.50, 17.88)
46-60	179 (32.1)	4.84 (4.26, 5.43)	10.47 (-6.17, 28.92)
61+	181 (32.4)	5.12 (4.42, 5.82)	7.57 (-9.28, 26.33)
Sex of infector			
Female	295 (52.9)	4.83 (4.34, 5.33)	referent [c]
Male	263 (47.1)	4.64 (4.10, 5.18)	-5.52 (-14.81, 4.37)
Presence of chronic conditions among infectors			
No	485 (86.9)	4.58 (4.20, 4.95)	referent [c]
Yes	73 (13.1)	5.85 (4.57, 7.13)	0.41 (-15.17, 17.67)
Number of general symptoms presented by infectors			
0	338 (60.6)	4.78 (4.32, 5.25)	referent [c]
1+	220 (39.4)	4.69 (4.09, 5.28)	-6.34 (-16.48, 4.53)
Number of respiratory symptoms presented by infectors			
0	336 (60.2)	4.22 (3.83, 4.61)	referent [c]
1+	222 (39.8)	5.54 (4.84, 6.23)	8.29 (-3.07, 20.48)
Types of setting			
Household	390 (69.9)	4.83 (4.41, 5.25)	referent [c]
Institution	10 (1.8)	5.00 (2.59, 7.41)	4.05 (-30.27, 47.56)
Social activity	113 (20.3)	4.27 (3.54, 4.99)	-9.17 (-20.78, 3.41)
Work	45 (8.1)	5.16 (3.12, 7.19)	-4.00 (-21.02, 15.12)
Wave			
One	94 (16.8)	6.70 (5.45, 7.95)	referent [c]
Two	464 (83.2)	4.35 (4.00, 4.70)	-33.90 (-45.08, -21.63)
Order of transmission			
Secondary	445 (79.7)	4.89 (4.49, 5.30)	referent [c]
Tertiary	102 (18.3)	4.36 (3.46, 5.27)	-17.31 (-28.84, -4.75)
Quaternary or beyond	11 (2.0)	2.18 (-0.52, 4.88)	-50.75 (-72.53, -23.45)
Relative mobility index [a,b,c]	Nil	Nil	0.83 (0.32, 1.33)

[a] A 8-day lag was assumed to account for the time between exposure to SARS-CoV-2 and the first symptom onset, based on the estimated 7.76-day incubation period **[19]**.
[b] The original mobility index was further adjusted relative to 18 January 2020, which has the value of 100.
[c] The referent is a common comparator across all variables.