Check for updates

RESEARCH ARTICLE

REVISED

Transmission dynamics, serial interval and epidemiology of COVID-19 diseases in Hong Kong under different control measures [version 2; peer review: 3 approved with reservations]

Yung-Wai Desmond Chan[1,2], Stefan Flasche[2], Tin-Long Terence Lam[1], Mei-Hung Joanna Leung[1], Miu-Ling Wong[1], Ho-Yeung Lam[1], Shuk-Kwan Chuang[1]

[1]Communicable Disease Branch, Centre for Health Protection, Department of Health, 147C Argyle Street, Kowloon, Hong Kong
[2]Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

v2 **First published:** 11 May 2020, **5**:91
https://doi.org/10.12688/wellcomeopenres.15896.1
Latest published: 09 Nov 2020, **5**:91
https://doi.org/10.12688/wellcomeopenres.15896.2

Abstract
Background: The outbreak of coronavirus disease 2019 (COVID-19) started in Wuhan, China in late December 2019, and subsequently became a pandemic. Hong Kong had implemented a series of control measures since January 2020, including enhanced surveillance, isolation and quarantine, border control and social distancing. Hong Kong recorded its first case on 23 January 2020, who was a visitor from Wuhan. We analysed the surveillance data of COVID-19 to understand the transmission dynamics and epidemiology in Hong Kong.
Methods: We constructed the epidemic curve of daily COVID-19 incidence from 23 January to 6 April 2020 and estimated the time-varying reproduction number (R_t) with the R package EpiEstim, with serial interval computed from local data. We described the demographic and epidemiological characteristics of reported cases. We computed weekly incidence by age and residential district to understand the spatial and temporal transmission of the disease.
Results: COVID-19 disease in Hong Kong was characterised with local cases and clusters detected after two waves of importations, first in late January (week 4 to 6) and the second one in early March (week 9 to 10). The R_t increased to approximately 2 95% credible interval (CI): 0.3-3.3) and approximately 1 (95%CI: 0.2-1.7), respectively, following these importations; it decreased to below 1 afterwards from weeks 11 to 13, which coincided with the implementation, modification and intensification of different control measures. Compared to local cases, imported cases were younger (mean age: 52 years among local cases

Open Peer Review

Reviewer Status ? ? ?

	Invited Reviewers		
	1	**2**	**3**
version 2 (revision) 09 Nov 2020			? report
version 1 11 May 2020	? report	? report	

1. **Jing Yuan**, Rutgers University, Piscataway, USA
2. **Lindsay Keegan**, University of Utah, Salt Lake City, USA
3. **Philip Arevalo**, University of Chicago, Chicago, USA

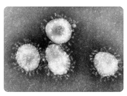
Any reports and responses or comments on the article can be found at the end of the article.

vs 35 years among imported cases), had a lower proportion of underlying disease (9% vs 5%) and severe outcome (13% vs 5%). Cases were recorded in all districts but the incidence was highest in those in the Hong Kong Island region.

Conclusions: Stringent and sustained public health measures at population level could contain the COVID-19 disease at a relatively low level.

Keywords

COVID-19, SARS-CoV-2, transmission potential, reproduction number, serial interval



This article is included in the Coronavirus (COVID-19) collection.

Corresponding author: Yung-Wai Desmond Chan (yung-wai.chan@lshtm.ac.uk)

Author roles: Chan YWD: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Flasche S**: Conceptualization, Formal Analysis, Methodology, Software, Supervision, Writing – Review & Editing; **Lam TLT**: Conceptualization, Formal Analysis, Software, Visualization, Writing – Review & Editing; **Leung MHJ**: Data Curation, Validation, Writing – Review & Editing; **Wong ML**: Conceptualization, Data Curation, Methodology, Validation, Writing – Review & Editing; **Lam HY**: Data Curation, Investigation, Validation, Writing – Review & Editing; **Chuang SK**: Investigation, Resources, Supervision, Validation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: Stefan Flasche is funded through a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant Number 208812).

How to cite this article: Chan YWD, Flasche S, Lam TLT *et al.* **Transmission dynamics, serial interval and epidemiology of COVID-19 diseases in Hong Kong under different control measures [version 2; peer review: 3 approved with reservations]** Wellcome Open Research 2020, **5**:91 https://doi.org/10.12688/wellcomeopenres.15896.2

First published: 11 May 2020, **5**:91 https://doi.org/10.12688/wellcomeopenres.15896.1

REVISED **Amendments from Version 1**

Thanks for the comments from both reviewers and the commenter, we made several changes in our article. The key changes are listed below. In the Methods we clarified how importation status was taken into account in the estimation of time-varying reproduction number (R_t), described in more detail the method of R_t estimation we adopted from Cori et al. and Thompson *et al*., and also expanded how the data available for analysis was collected. In the Results (and abstracts) we added confidence intervals to R_t estimates and serial interval. We expanded our Discussion on the interpretation of the R_t (especially on the inclusion/ exclusion of imported cases) and how using date of report as date of onset for asymptomatic would impact our analyses. We also corrected some typing errors and revised the hyperlinks to the data repository.

Any further responses from the reviewers can be found at the end of the article

Introduction

In December 2019, a cluster of pneumonia cases with unidentified etiology was reported in Wuhan, China[1]. The cause was later identified in early January as a novel type of coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), for which the World Health Organization (WHO) later named the disease as coronavirus disease 2019 (COVID-19)[2]. It sparked an outbreak in Wuhan and spread to other parts of China. The WHO declared this outbreak a Public Health Emergency of International Concern on 30 January 2020 and further characterised it as a pandemic on 11 March 2020[3]. As at 14 April 2020, there were over 1,800,000 confirmed cases with over 117,000 deaths affecting many countries and areas around the world[4].

Hong Kong recorded the first imported case from Wuhan, China on 23 January 2020 and over 700 cases were recorded by the end of March 2020. Local cases and clusters were identified after different waves of importation, first from Mainland China in late January and other parts of the world in March. A series of control measures had been implemented and intensified over the course of the epidemic including enhanced surveillance, isolation and quarantine, border control and social distancing (Table 1). This report summarises the transmission dynamics and the epidemiology of COVID-19 in Hong Kong from late January to early April 2020. We estimated the daily reproduction number to understand the transmission dynamics of COVID-19 under different waves of importation and various non-pharmaceutical public health measures implemented.

Methods

In Hong Kong, the Centre for Health Protection (CHP) of the Department of Health conducted surveillance for COVID-19 for early detection of cases through systems implemented at different levels, including statutory notification by frontline doctors[5], enhanced health surveillance at border control points (e.g. temperature screening) and health declaration by visitors, enhanced laboratory surveillance on hospital inpatients, emergency departments, outpatient clinics, and asymptomatic inbound travellers[6–8]. All cases should be tested for SARS-CoV-2 by quantitative PCR (qPCR) at the Public Health Laboratory Service Branch (PHLSB) of the CHP. Seroconversion or four-fold or greater increase in antibody titre to SARS-CoV-2 in paired serum is another laboratory criterium of confirmation. Irrespective of clinical presentation, cases tested positive were classified ‘confirmed’ and those with inconclusive laboratory result but had epidemiological linkage to confirmed COVID-19 cases were considered as ‘probable’. Contact tracing would be carried out for every confirmed and probable case to identify any exposure persons who had contact with the case. They would be put under quarantine and medical surveillance. As part of the case-based investigation, the CHP also collected information on cases’ demographics and clinical symptoms (including date of onset), based on patient recalls and information provided by reporting doctors. The CHP used these data to generate reports of daily situation updates on these confirmed and probable cases[9]. Using the data from these daily updates and the Government’s press releases[10], we collated a line-list of COVID-19 cases with basic demographic, epidemiological and clinical information including age, gender, residential status (i.e. whether the case was a Hong Kong resident or not) and residential district (grossly divided into 18 districts in Hong Kong), date of symptom onset and date of reporting (notified to the CHP), importation status and discharge status (whether the case was still under admission and isolation, discharged home or death). The CHP assigned an official classification of importation status to each case based on their travel history and contact with confirmed/ probable COVID-19 cases during the incubation period. These importation status include: “imported”, “close contact of imported case”, “local”, “close contact of local case”, “possibly local” and “close contact of possibly local case”. To facilitate our data analyses including estimation of the reproduction number, we regrouped the importation status based on the official case classification[9] and travel history of the cases before onset of symptom into “imported” and “local” (i.e. non-imported) cases (Table 2). In brief, imported cases were those with travel history who had spent a significant amount of their incubation period (for instance half of their incubation period or more) at a place(s) where there was evidence of local transmission, according to the WHO situation reports[4] and ECDC global case update[11]. A total of 20 of 915 cases (2.2%) remained unclassified as they had travel history but also spent time in Hong Kong when there was local transmission.

Estimation of time-varying reproduction number

We included confirmed and probable cases reported between 23 January and 6 April 2020 into an analysis estimating the effective reproduction number and its change during the study period. We included all confirmed and probable cases. We constructed an epidemic curve of these cases. The daily number of new cases are expected to increase over time if R exceeds 1 whilst an epidemic is expected to decline if R is consistently reduced to below 1. The time-varying reproduction number (R_t) refers to the effective reproduction number at different time points in an epidemic and it was estimated based on the daily COVID-19 incidence using the R package EpiEstim version 2.2-1[12]. Grossly, there are two variations of the effective

Table 1. **Key control measures implemented to combat COVID-19 in Hong Kong from Jan to Mar 2020.**

2020	Jan	Feb	Mar
No. of cases in HK (imported)	**13 (11)**	**82 (12)**	**620 (436)**
Key events	**- Outbreaks in Wuhan and other parts of China started**	**- Outbreaks in South Korea, Iran and part of Italy**	**- Increasing number of cases worldwide - WHO declared pandemic**
1. Enhanced surveillance and management of cases			
a. Case based detection			
Reporting criteria	On 31 Dec 2019, a reporting criteria was developed and a surveillance system has been launched immediately to trace suspected cases. The reporting criteria has been revised from time to time according to the latest situation.		
Statutory reporting	Legal Ordinance was amended to include "Severe Respiratory Disease associated with a Novel Infectious Agent" as a statutorily notifiable infectious disease since 8 Jan 2020. New regulations were made for compulsory quarantine and information disclosure to facilitate investigation and control measures.		
b. Enhanced laboratory surveillance at population level			
Free COVID-19 testing at different settings	• Since mid-Jan, certain groups of in-patients with pneumonia (e.g. ICU) were tested for COVID-19. • Since late Jan, testing was further extended to all inpatients with pneumonia (irrespective of travel history).	• Since mid-Feb, symptomatic patients attending General Outpatient Clinics (GOPC) and Accident and Emergency Departments (AED) were offered tests for COVID-19.	• Since early Mar, testings were provided to symptomatic patients attending private medical practitioners. • Since late Mar, testings were provided to asymptomatic inbound travellers from UK, Europe and US; later extended to cover all asymptomatic inbound travellers.
c. Isolation and quarantine			
Confirmed or symptomatic suspected cases	Isolation and treatment were provided in hospital until physically fit for discharge and having negative PCR results in 2 consecutive days.		
Close contacts	14-day quarantine at quarantine centres		
Persons from places with high risk of COVID-19 transmission	• 14-day quarantine at quarantine centres or home quarantines; • The list of areas/places expanded from Hubei province in late Jan to Mainland China on 8 Feb; further included South Korea, Iran and parts of Italy in late Feb; expanded to parts of France, Germany, Spain and Japan in early Mar; and later included all overseas places in mid-Mar.		
2. Border control			
Reducing flow of people from Mainland China and other places	• Since 27 Jan, Hubei residents have been restricted from entering HK • Since 28 Jan, transport services and border control point services to-and-from Mainland China were reduced or suspended gradually (flights, railway, ferry, cruise terminal, ocean terminal, land-based cross-boundary transport).	• Since 8 Feb, all people entering HK from the Mainland were required to undergo 14-day mandatory quarantine. • Since 24 Feb, Red Outbound Travel Alert (OTA) was issued on South Korea and later expanded to parts of Italy; travellers from there would be subject to compulsory quarantine/medical surveillance.	• Since 10 Mar, Red OTA was expanded to other countries/ territories gradually. • Since 19 Mar, people from all overseas places in the past 14 days would be subject to compulsory quarantine. • Since 25 Mar, all non-HK residents coming from overseas would be denied entry. All transit services at the airport were suspended. All travellers coming from Macao and Taiwan would be subject to compulsory quarantine.

2020	Jan	Feb	Mar
No. of cases in HK (imported)	**13 (11)**	**82 (12)**	**620 (436)**
Key events	**- Outbreaks in Wuhan and other parts of China started**	**- Outbreaks in South Korea, Iran and part of Italy**	**- Increasing number of cases worldwide - WHO declared pandemic**
Enhanced health surveillance measures at border control points	• Since 1 Jan, inbound travellers with relevant symptoms and travel or exposure history identified would be referred to public hospitals for isolation. • Additional temperature screening of inbound travellers from Wuhan; later expanded to inbound travellers from Mainland China, including mandatory health declaration for all passengers arrived by high speed rail since 26 Jan.	• Since 1 Feb, travellers departing and on transit from the airport were required to undergo temperature checks; exit screening was later expanded to Cruise and Ocean Terminal. • Since late Feb, additional temperature checks had been conducted for travellers from Korea.	• Enhanced temperature screening to all inbound travellers.
Health declaration	• Since 21 Jan, inbound passengers from Wuhan by air are required to complete a health declaration form, later expanded to high speed rail and then all passengers from Mainland China.	• Since 1 Feb, health declaration arrangements were expanded to other border control points (land, sea) gradually. • Since 25 Feb, health declaration form is required for all inbound flights from Korea.	• Since 8 Mar, the health declaration arrangements have been expanded to all flights arriving HK.
3. Non-pharmaceutical public health measures			
a. Infection Control			
Public	• Since 1 Jan, health promotion and education activities to enhance personal, respiratory, environmental hygiene and travel health advice through website, social media, TV and radio APIs, advertisements, press releases, health talks, etc. • Health education materials (e.g. infographics, pamphlets, posters) are translated and made available to nine ethnic minorities, which have been disseminated to different stakeholders and sectors. • Telephone hotline was set up to provide information and answers to questions.		
Community	• Cleaning was stepped up in premises where cases have stayed. Cleaning of public area and facilities, such as lifts and escalators, was enhanced. • Infection control measures were implemented at government buildings and offices, such as temperature checking and providing alcohol hand sanitizers, etc. • Infection control guidelines were developed for community settings (e.g. schools, workplace, institutions, hotel industry, public transport, property management, etc.) • Letters/emails to schools, institutions, private hospitals and clinics, trade and industry, non-government organisations and other community partners on infection control measures.		
Healthcare	• Infection control measures were strengthened in hospitals, clinics and health service centres. • Infection control trainings were provided to healthcare professionals in public and private sectors. • Infection control guidelines were developed for different healthcare settings and healthcare professionals. • Since 26 Jan, visiting, volunteer activities and clinical placement were suspended in public hospitals; all persons in public hospital areas required to wear mask.		
b. Social distancing			
Work	• From late Jan to early Mar, civil servants were allowed to work from home (except emergency and essential public services). The Government also appealed to other employers to make flexible work arrangements for employees to reduce social contacts. • From early Mar to late Mar, public services were gradually resumed, but flexible working hours and roster adopted. • Since late Mar, special work arrangement was made for civil servants. All government departments would provide essential, emergency and limited public services only.		
Education	• Since late Jan, schools including tutorial centres were suspended until further notice. Online teaching was adopted. • The Hong Kong Diploma of Secondary Education (HKDSE) examinations were postponed from Mar to late Apr.		
Community	• Since late Jan, mass gathering events and recreational facilities have been cancelled or closed, e.g. the Hong Kong marathon, large-scale concerts, public libraries, Disneyland and Ocean Park, etc. • Since late Mar, more outdoor leisure facilities were closed, e.g. sports facilities, playgrounds, etc. • In late Mar, new regulations were made to impose temporary restriction on seating capacity and distance between tables at catering premises, temporary closure of scheduled premises (e.g. cinema), and prohibit any group gathering of more than 4 persons in any public place.		

Table 2. Classification of importation status for cases included in the analyses (n = 915).

Official case classification	Regrouped importation status	N
Imported	Imported	526
Local case	Local[1]	63
Close contact of imported case		21
Close contact of local case		155
Close contact of possibly local case		48
Close contact of local case	Unclassified[2]	3
Possibly local	Imported[3]	80
	Local[4]	2
	Unclassified[2]	17

[1]Cases had contracted infections in Hong Kong with various source of infection and had no travel history 14 days before onset of symptoms/confirmation of infection. All were regrouped as local cases.

[2]Cases remained unclassified as they had travel history but also spent time in Hong Kong when there was local transmission.

[3]Cases had travel history during their incubation period to places with high risk of COVID-19 transmission.

[4]One case did not have any travel history during his incubation period and the other case with travel history was linked to a local cluster.

reproduction number, namely the case reproduction number and the instantaneous reproduction number. The case reproduction number, defined as the expected number of secondary infections that an infected individual at time t will eventually cause, was first described by Wallinga *et al.*[13]. In contrast, the instantaneous reproduction number is the expected number of secondary infections resulting from the number of infected individuals and their relative infectiousness at time t. It was first described by Cori *et al.*[14] and an improved framework taking into account importation status of incidence data was later proposed by Thompson *et al.*[15]. Both iterations of effective reproduction number can be estimated using the EpiEstim package with time-series incidence data. However, these incidence data in an on-going epidemic are right truncated (i.e. infections occurring near the end of observation period were not yet detected or reported). As estimation of the case reproduction number relies on the cases observed after time t, it is more prone to under-estimation towards the end of the observed time series, as detailed in Gostic *et al.*[16]. Therefore, we used the EpiEstim package to estimate the instantaneous reproduction number as R_t in this study since the COVID-19 epidemic was still ongoing at the time of analysis.

In brief, the average incidence of cases at time t is following a Poisson distribution mean

$$E[I_t] = R_t \sum_{s=1}^{t} I_{t-s} w_s$$

where R_t is the effective reproduction number at time t, I_{t-s} is the infection incidence at time $t - s$ and w_s is the weighted infectivity function. Since the day of infection is unknown in almost all cases, we instead use the incidence of symptomatic cases based on the reported onset date. Note that as a result our estimates of R_t for infection incidence are therefore shifted by about 5 days[16].

We fit the incidence data by date of onset with a time window of 7 days to compute the R_t. For the incidence data included in our study, date of onset was available for all cases except those with no clinical symptom at the time of report and no case was fatal at the time of report. For asymptomatic cases, their date of reporting was used instead (17% of the cases included in this analyses were asymptomatic at the time of report to CHP). Case ID 503, who had been residing in Spain and returned to Hong Kong on 26 March, was tested positive upon arrival. As he had chronic cough since January but no other symptoms at the time of confirmation his date of arrival was used as date of onset. The EpiEstim package caters the probability of a secondary case arising from incident case that could be either local or imported[15]. Thus we fit the daily incidence stratified by the importation status of cases (imported and likely locally acquired). For unclassified cases after regrouping, we conducted separate estimates by including them either as imported or local. For the 915 cases recorded during the study period, 67 cases picked up at borders were not included in the R_t estimation, as their risk of transmission to the community was considered minimal. We reviewed reported cases among different clusters and identified 47 pairs of symptomatic infectors and infectees with clear epidemiological evidence of a direct transmission link (Table 3). All infectees identified had direct exposure to the respective infectors and no exposure with other known confirmed cases. The corresponding serial interval was computed as the difference in days between disease onset of these successive cases, which approximated the interval between generations of infections[18]. A gamma distribution was fit to the serial interval data and was used as the weighted infectivity function (w_s) in the R_t estimation. Four sensitivity analyses were carried out to assess the impact of data source (using all incidence cases with and without stratification into imported and local, using local and unclassified cases only as well as using local cases only), source of serial interval (using local serial interval data and serial intervals reported in two other studies[19,20]), exclusion of cases identified by contact tracing and size of time window (4, 7 and 14 days).

Descriptive epidemiology

We described the demographic and epidemiological characteristics of these cases. We compared the age- and sex-specific

Table 3. Infector-infectee pairs used to compute local serial interval data.

Setting	Case No		Onset date[1]		Serial interval (days)
	From	To	From	To	
Household	9	11	25-01-20	28-01-20	3
Household	13	15	29-01-20	01-02-20	3
Household	17	19	22-01-20	30-01-20	8
Workplace	30	44	29-01-20	01-02-20	3
Workplace	37	50	30-01-20	03-02-20	4
Household	52	61	31-01-20	02-02-20	2
Workplace	57	59	07-02-20	12-02-20	5
Household	58	128	08-02-20	08-03-20	29
Household	66	63	25-01-20	12-02-20	18
Social	66	57	25-01-20	07-02-20	13
Social	66	60	25-01-20	08-02-20	14
Household	72	82	10-02-20	20-02-20	10
Household	74	89	13-02-20	25-02-20	12
Household	76	91	08-02-20	20-02-20	12
Household	83	84	17-02-20	25-02-20	8
Household	85	90	12-02-20	16-02-20	4
Household	92	95	13-02-20	29-02-20	16
Client and personal driver	106	116	04-03-20	07-03-20	3
Client and personal driver	143	155	11-03-20	15-03-20	4
Household	143	298	11-03-20	20-03-20	9
Household	154	228	11-03-20	18-03-20	7
Household	154	311	11-03-20	19-03-20	8
Social	169	208	16-03-20	18-03-20	2
Social	174	162	09-03-20	12-03-20	3
Household	174	222	09-03-20	18-03-20	9
Workplace	174	243	09-03-20	13-03-20	4
Workplace	174	285	09-03-20	20-03-20	11
Workplace	174	243	09-03-20	13-03-20	4
Workplace	174	285	09-03-20	20-03-20	11
Household	195	322	17-03-20	21-03-20	4
Dating	243	349	13-03-20	19-03-20	6
Social	338	229	11-03-20	18-03-20	7
Social	338	337	11-03-20	19-03-20	8
Household	347	709	22-03-20	29-03-20	7
Workplace	442	504	23-03-20	25-03-20	2
Household	447	912	24-03-20	27-03-20	3
Household	447	913	24-03-20	27-03-20	3
Household	462	515	23-03-20	25-03-20	2
Social	462	757	23-03-20	30-03-20	7
Household	477	525	22-03-20	24-03-20	2
Social	483	516	18-03-20	25-03-20	7
Beauty salon	504	756	25-03-20	29-03-20	4
Household	513	810	24-03-20	02-04-20	9
Neighbour	525	760	24-03-20	31-03-20	7
Household	581	691	24-03-20	29-03-20	5
Household	746	844	26-03-20	02-04-20	7
Social	756	796	29-03-20	30-03-20	1

[1]Onset date in the format “dd-mm-yy”.

proportion of cases with the Hong Kong demographics in 2019[21]. To understand the spatial and temporal distribution of the disease in the community, we computed incidence by age and residential district in different weeks of disease onset. We excluded cases picked up at borders as well as those with unknown residential district from the spatial analysis. We described the temporal trend of the onset, report, death or discharge of cases and utility of hospital beds with data provided by the Hong Kong SAR Government[9]. We computed the duration between onset to report with time, as well as the duration between report to discharge over time and age.

Results

Estimation of R_t

From 23 January to 6 April 2020, 915 COVID-19 cases were recorded with 606 (66%) being imported (Table 4). The epidemic started with imported cases from the Mainland China (Figure 1), followed by cases with no travel history and no known contacts with imported cases. The first case was a febrile 39-year-old visitor from Wuhan intercepted at the border by the Port Health Division (PHD) of the DH and was later confirmed for COVID-19 on 23 January (Week 3). One week later the first non-imported case with symptom onset on

Table 4. Characteristics of COVID-19 cases in Hong Kong reported from 23 January to 6 April 2020 (n = 915).

	Local	Imported	Unclassified	Overall
n (%)	289	606	20	915
Male gender	151 (52.2)	327 (54.0)	15 (75.0)	493 (53.9)
Age (mean (SD))	43.4 (17.2)	35.3 (17.5)	45.8 (16.7)	38.1 (17.8)
Age group				
0 to 9	7 (2.4)	3 (0.5)	0 (0.0)	10 (1.1)
10 to 19	3 (1.0)	123 (20.3)	0 (0.0)	126 (13.8)
20 to 29	53 (18.3)	167 (27.6)	3 (15.0)	223 (24.4)
30 to 39	80 (27.7)	102 (16.8)	6 (30.0)	188 (20.5)
40 to 49	47 (16.3)	59 (9.7)	3 (15.0)	109 (11.9)
50 to 59	44 (15.2)	72 (11.9)	4 (20.0)	120 (13.1)
60 to 69	34 (11.8)	61 (10.1)	1 (5.0)	96 (10.5)
70 to 79	13 (4.5)	16 (2.6)	2 (10.0)	31 (3.4)
80+	8 (2.8)	3 (0.5)	1 (5.0)	12 (1.3)
Hong Kong residents				
HK resident	288 (99.7)	586 (96.7)	20 (100.0)	894 (97.7)
Non-HK resident	1 (0.3)	20 (3.3)	0 (0.0)	21 (2.3)
Asymptomatic when reported	32 (11.1)	116 (19.1)	3 (15.0)	151 (16.5)
Any underlying disease	27 (9.3)	33 (5.4)	5 (25.0)	65 (7.1)
Hospitalisation status				
Deceased	2 (0.7)	1 (0.2)	1 (5.0)	4 (0.4)
Discharged	91 (31.5)	120 (19.8)	5 (25.0)	216 (23.6)
Still hospitalised	196 (67.8)	476 (78.5)	14 (70.0)	686 (75.0)
Pending admission	0 (0.0)	9 (1.5)	0 (0.0)	9 (1.0)
Ever serious/under ICU care for discharged or deceased cases[1]	12 (12.9)	6 (5.0)	1 (16.7)	19 (7.7)
Duration between onset and report (median [IQR])	4.00 [2.00, 8.00]	5.00 [3.00, 8.00]	6.00 [4.00, 8.00]	5.00 [3.00, 8.00]

[1]According to information reported by the Hospital Authority (HA) during the daily press conference. Serious cases were those who required high-flow oxygen.

22 January was reported, followed by emergence of first wave of cases around the period of Chinese New Year holiday (week 4 to 6). These included two large clusters involving family gatherings, each led to onward transmission in households or workplace with 13 and 6 cases being eventually identified[22]. In February (week 6 to 9), a majority of cases had no travel history (51 of 67, or 76%), with a large cluster of 19 persons associated with Buddhist temple in Eastern district (12 persons worked/ visited this temple who then led to 7 additional secondary cases)[22]. For this wave, the R_t increased from around 1.5 (95% credible interval (CI): 0.2-4.0) in week 3 to around 2 (95% CI: 0.6-3.3) in week 4, it then gradually decreased to under 1 (95%CI: 0.1-1.5) along with decrease in reported number of cases between week 6 to 8. The second wave started in early March (week 9 to 10), with the R_t increased to around 1 (95%CI: 0.2 to 1.7) with an increasing number of reported cases. Among the 436

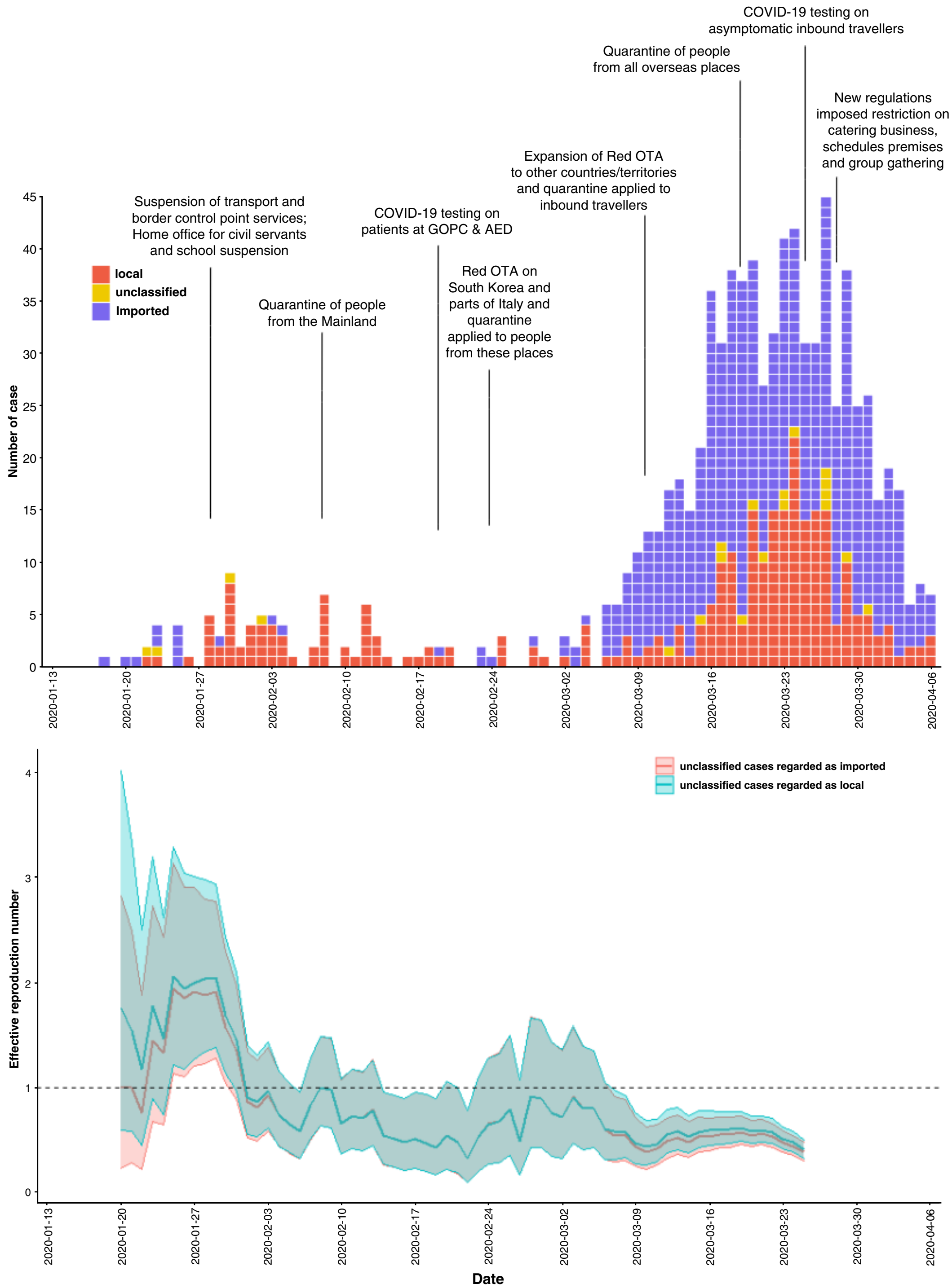


Figure 1. Epidemic curve and time-varying reproduction number estimates for COVID-19 cases in Hong Kong by date of onset (n = 848). Note: 915 cases reported from 23 Jan to 6 Apr 2020. 67 cases picked up at borders were not included in the R_t estimation and was not shown in the epicurve in the first plot; R_t estimates of last 7 days was not shown as it was likely under-estimated due to the delay between disease onset and report.

imported cases in March, most had travel history to UK and other parts of Europe (321 or 74%), followed by the US (38 or 9%) and Canada (20 or 5%). More local cases and clusters associated with social activities and gatherings were reported following influx of such imported cases. For example a large cluster with over 90 persons involving a group of band players, staff and customers of restaurants and pubs, as well as onward transmission to their friends and families were identified in late

March[23,24]. Two other clusters involved social gatherings in Karaoke were also identified[25,26]. The R_t decreased to around 0.5 from weeks 11 to 13. The mean serial interval among 47 infector-infectee pairs identified was 6.5 days [standard deviation (SD) 4.7; median: 5, 95% quantiles: 0, 18)] (Figure 2).

Age, spatial and temporal distribution

Compared to local cases, imported cases were younger (mean age 35 years for imported cases vs 43 years for local cases) and had a higher proportion of being asymptomatic at the time of reporting (19% for imported cases vs 11% for local cases) (Table 4). Up to 6 April, 216 cases were discharged while 4 cases passed away. Among those already discharged/ deceased, 8% were under intensive care unit (ICU) care or considered serious. Serious/ critical condition of patients were classified according to clinical judgement of frontline doctors, but generally patients with deteriorating condition requiring high-flow oxygen and/ other supportive care would be considered serious. The four fatal cases involved three males and one female aged 39 to 80 years, all were reported to have underlying illnesses.

Compared to the whole population, there was a considerably higher proportion of cases among those aged 10 to 39 years for both genders (Figure 3). In contrast, those aged under 10 and those aged 80+ only contributed to 1% each among the reported cases, which was substantially lower than their respective distribution in the population. The age-specific incidence varied greatly by weeks of disease onset (Figure 4b). The incidence for those aged 60 to 79 was higher among those in weeks 4 to 7 and remained stable until a surge in weeks 11 and 13 which coincided with a rapid increase in incidence for those aged 10 to 39, for which the incidence had been low. The incidence for those aged under 10 was all along lowest as no cases were recorded until week 11 and 12.

Geographically, the incidence per 100,000 was highest in the districts in the Hong Kong Island region of the city (Wan Chai: 480 and Central and Western: 357) and was considerably lower in the New Territories region (under 40 for Yuen Long and Tsuen Wan) and lowest in the Wong Tai Sin district[11] of the Kowloon Peninsula region (Figure 4a). Despite of the overall geographical difference, the increase in incidence after week 10 was observed in most districts (Figure 4c). The incidence for the extremes of age was still low even in Hong Kong Island region, despite the overall high incidence in this area (Figure 4d).

Trend of case onset, confirmation and discharge/death

All cases reported from week 4 to 7 were symptomatic. The proportion of asymptomatic cases increased to 20% in week 13. The rates of disease onset and report remained stable until the upsurge after week 10. The stable rate of discharge could not catch up with the rapid increase in the report of confirmed cases in weeks 11 and 12 (Figure 5a). Correspondingly, the daily bed usage for investigation of suspected cases and management for confirmed COVID-19 cases increased from under 50 to nearly 200 by week 4 and became stable at around 200 beds until week 10. The daily hospital bed usage doubled to over 400 per day in week 11, and further increased to around 600 in week 12 and 800 in week 13.

Both the median duration from onset to reporting and from reporting to discharge decreased with time as the epidemic evolved (Figure 6a, b). The duration from report to discharge

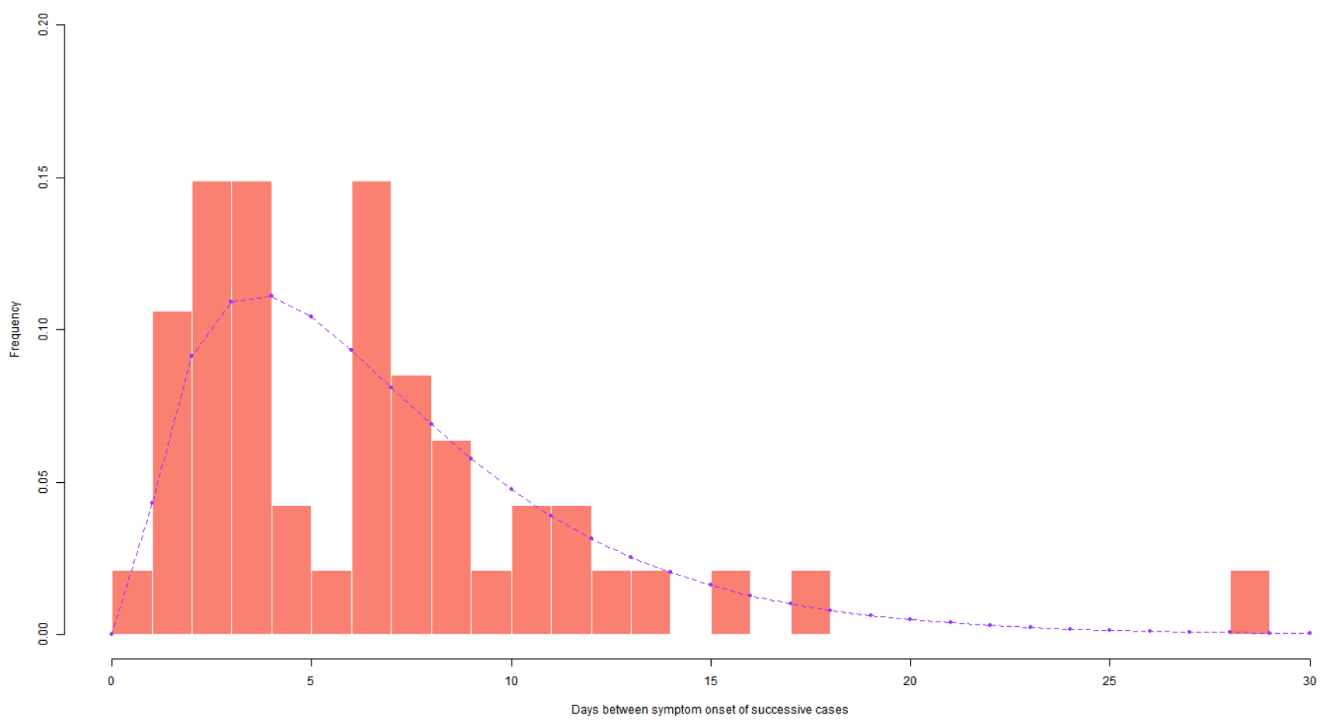


Figure 2. Gamma distribution fitted to serial interval data of COVID-19 cases in Hong Kong.

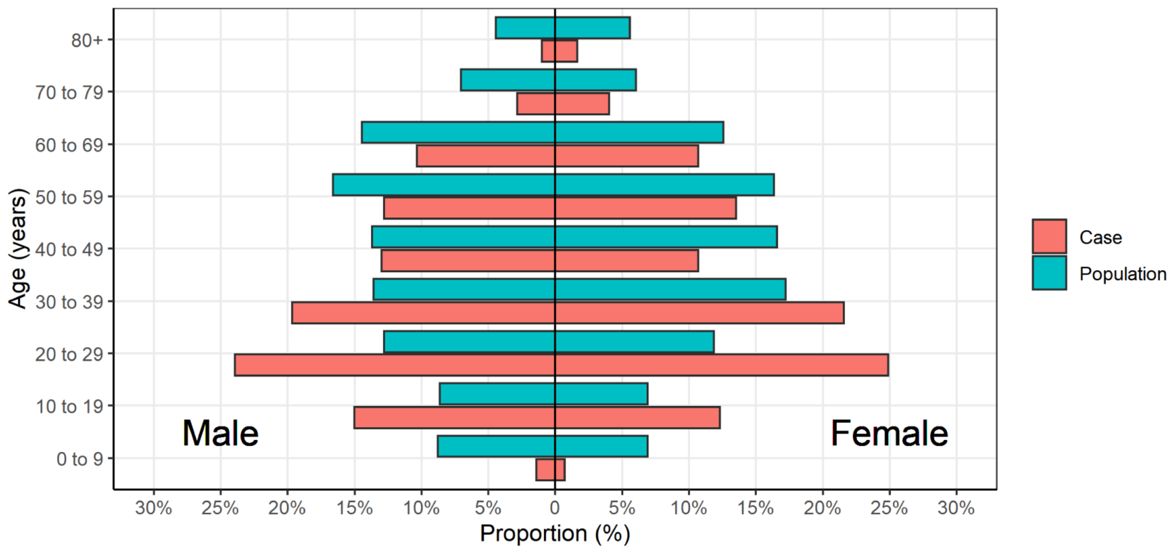


Figure 3. Proportion of COVID-19 cases in Hong Kong by age and sex in comparison to the population, reported from 23 January to 6 April 2020 (n = 915). The proportion of cases by age and sex was compared to those in the whole population. Those aged 10 to 39 years attributed to a higher proportion for both sexes among the cases reported (59% of cases were aged 10 to 39 years for both sexes, compared to 35 to 36% of the general population aged 10 to 39 years).

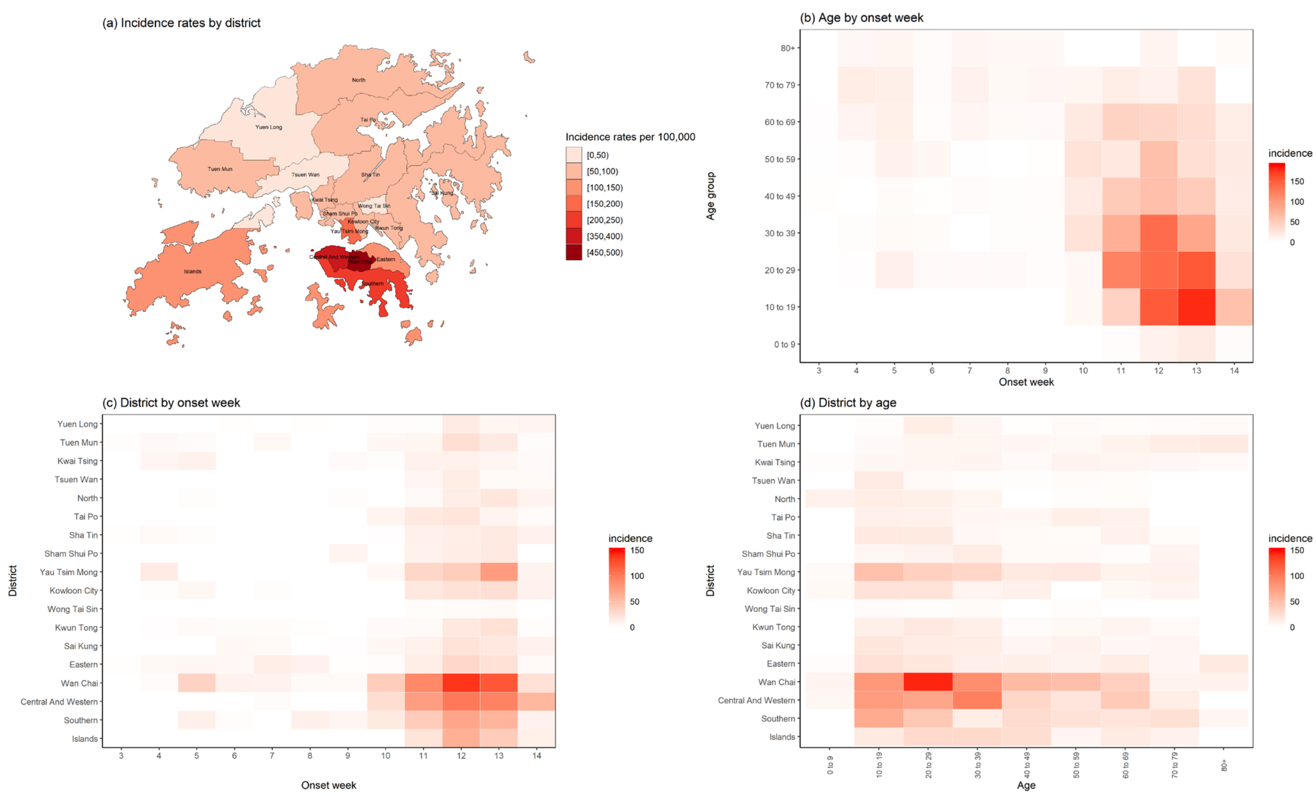


Figure 4. Incidence of COVID-19 cases reported from 23 January to 6 April 2020 by age, week of onset and residential district (n = 837). In total, 837 of 915 cases (91%) were included in the above analyses. Cases detected at border or those with no known information on residential district were excluded. Hong Kong is divided into 18 districts and these districts are grossly grouped into different regions: New Territories West (Yuen Long, Tuen Mun, Kwai Tsing, Tsuen Wan), New Territories East (North, Tai Po, Sha Tin, Sai Kung), Kowloon West (Sham Shui Po, Yau Tsim Mong), Kowloon East (Kowloon City, Wong Tai Sin, Kwun Tong), Hong Kong Island (Eastern, Wan Chai, Central and Western, Southern) and Islands.

decreased from 34 days in week 4 to less than 15 days from week 8 onwards (Figure 6b). This duration varied by age with a median of 8 days for those aged 20 to 29 years as compared to over 15 days for those aged 50 years and above (Figure 6c).

Discussion

We analysed the transmission dynamic and epidemiological characteristics of COVID-19 cases in Hong Kong in the first three months of the epidemic. The median estimated R_t varied from 0.3 to 2.0 in this period and it showed that there were two waves of transmission in Hong Kong for which the R_t was close to or exceeded 1, the first one following cases imported from Mainland China and the second one after imported cases from overseas places like Europe and the US. The first wave of importations occurred right before the Chinese New Year holidays and coincided with the outbreak in Wuhan and other parts of Mainland China. Transmission in the first wave was most intense at week 5 around the Chinese New Year holiday when there were increased social contacts (e.g. family gatherings). The Government implemented different measures to control the epidemic starting from January, including work from home for civil servants[27], school suspension, closure of recreational facilities, reducing importation pressure by closure of some border control points, enhanced laboratory surveillance for early detection of cases, etc. The transmission reduced subsequently as R_t was largely below 1 in February. However, the Government departments[27] and other companies gradually scaled back to normal services in March and social distancing might have relaxed in the community[28]. When compared to the baseline, Google mobility data in Hong Kong showed reductions in visits to retail and recreation, parks, transit stations and workplaces in February and March[29]. However, these indicators appeared to be lower in mid February and late March but higher in early March. These could reflect a relaxed social distancing in the community and coincided with the increase in R_t in early March. Coupled with the rapid increase in number of confirmed imported cases from early March onwards, the occurrence of large clusters originated from social venues including bars and restaurants affecting over 90 persons hinted a grave risk of community transmission originated from importations. While the source of infection for many of these clusters was yet to be determined (e.g. before availability of genetic sequence data), the incidence was the highest among those aged 10 to 39 years for both genders, coinciding with the high number of imported cases related to returning travellers, workers and students for these highly sociable age groups in the same period. Although young adults were less likely to develop severe complications of COVID-19[30], they posed a risk of onward transmission

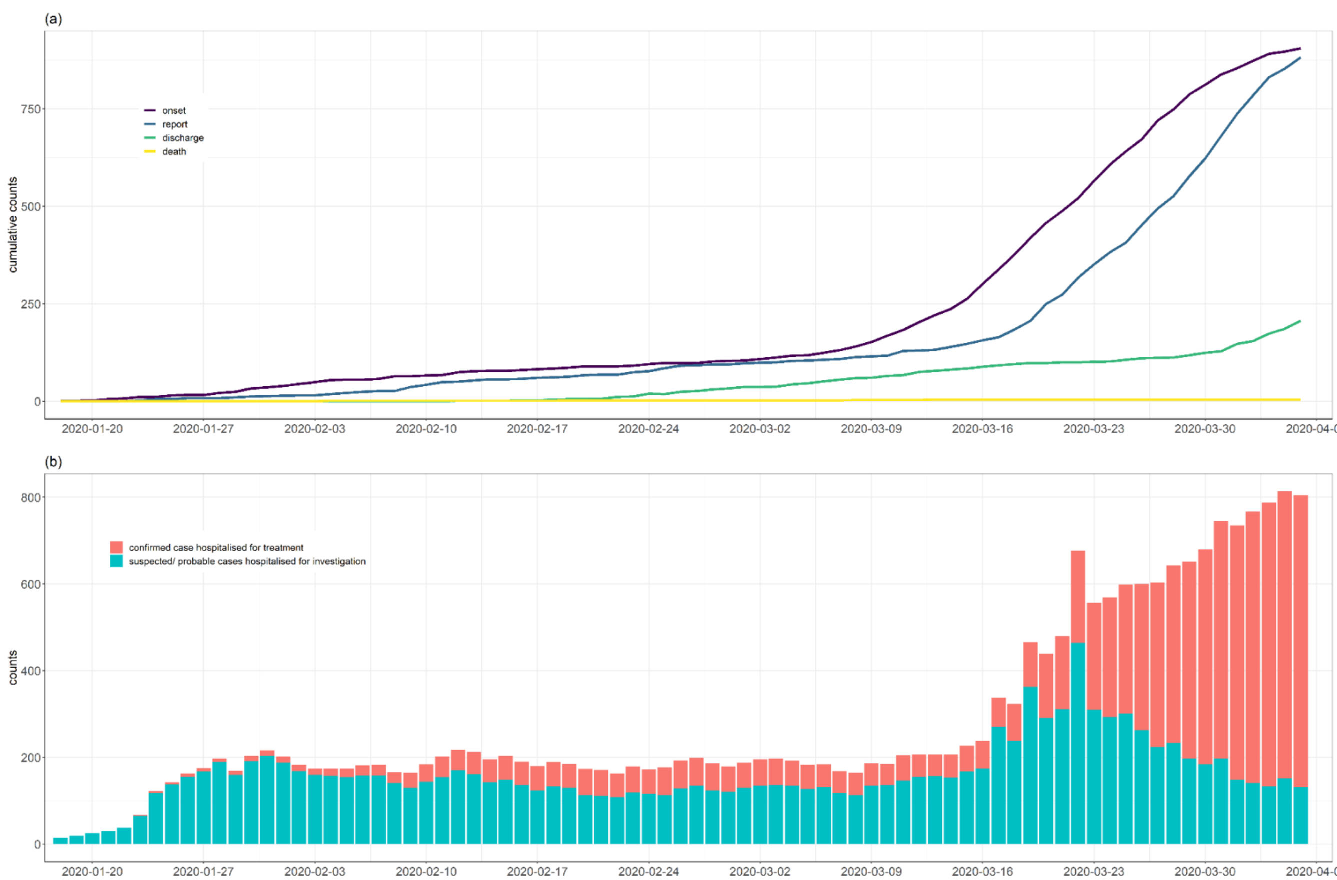


Figure 5. COVID-19 cases in Hong Kong reported from 23 January to 6 April 2020. (**a**) Cumulative number of confirmed case by date and status and (**b**) use of hospital bed due to COVID-19 cases in public hospitals by date and case status.

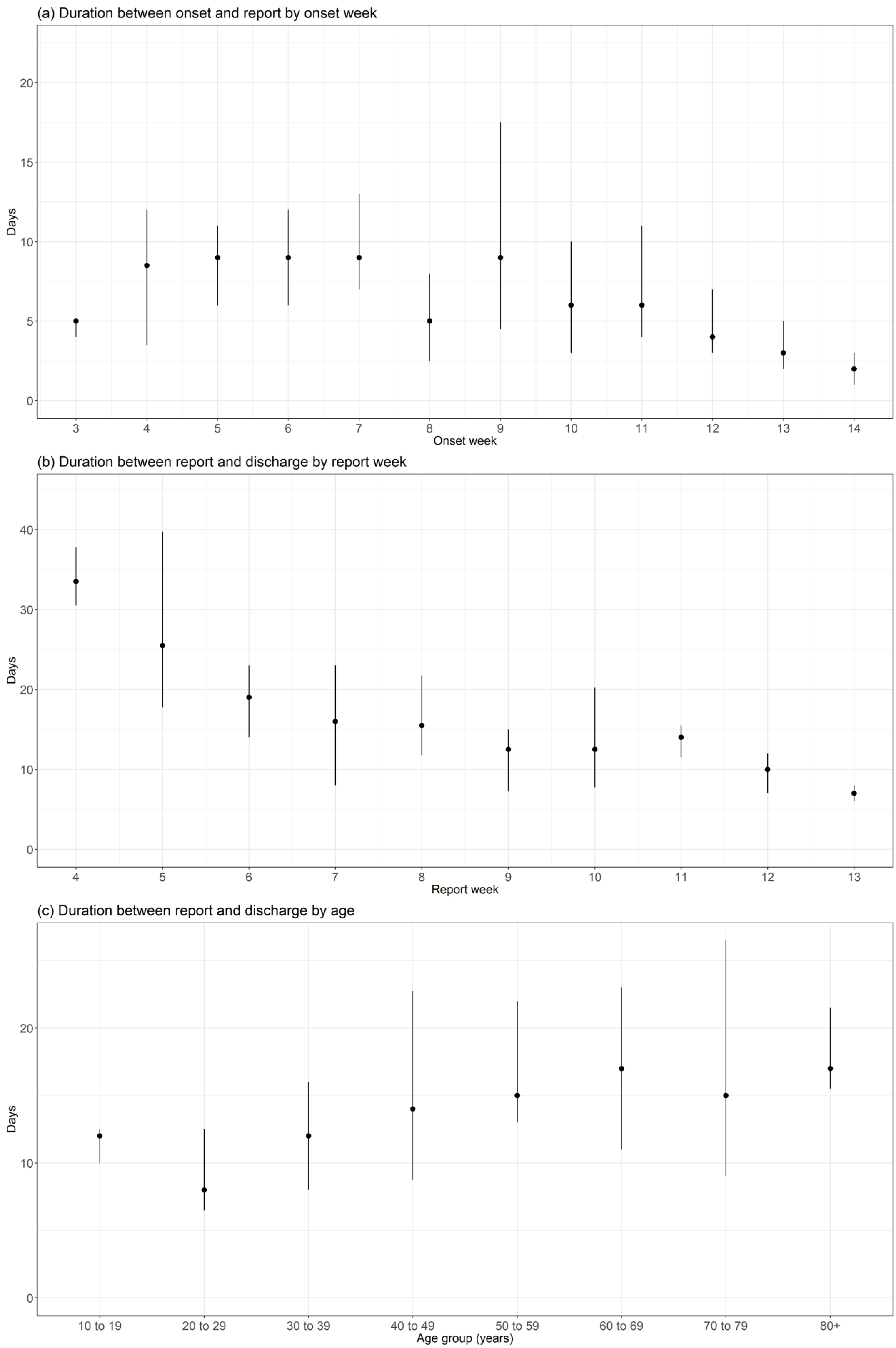


Figure 6. Change in (**a**) duration between onset to report by onset week, (**b**) duration between report and discharge/ death by report week and (**c**) by age groups. Asymptomatic cases were not included in (**a**). Report week 14 and those aged 0 to 9 were not shown in (**b**) in (**c**) respectively due to small numbers.

to the community, including older individuals who are at higher risk of severe COVID-19 diseases. A local case of infected nurse working in a residential care home for the elderly as part of large social venue cluster further highlighted such potential risk. The R_t decreased from around 1 in early March to around 0.5 and stabilized in mid March, despite a second peak contributed mainly by imported cases. As these imported cases would not be regarded as secondary cases infected locally, their occurrence should not infer to local transmission and hence the R_t estimated was relatively low in this period despite of an apparently larger second peak. However, this second wave of epidemic led to a rapid increase in hospital bed usage for both confirmed and suspected cases. Despite a shortened duration from report to discharge and increased proportion of mild (or asymptomatic) cases, demand for isolation facility continued to stretch beyond limit along with the rapid case upsurge as two consecutive negative PCR results is one of the discharge-from-isolation of confirmed COVID-19 cases. Keeping a low level of community transmission is thus crucial in protecting the healthcare system from being overwhelmed.

The *R* estimates in Hong Kong were considerably lower than that estimated in Wuhan, which ranged from 2 to 4 when there was widespread transmission[19,31–34]. The lower *R* estimates could be attributed to different factors, such as difference in contact patterns, different methodologies and serial intervals/ generation time adopted in estimation of *R*, as well as lower transmissibility of SARS-CoV-2 in Hong Kong in our study period. The lower transmissibility might be attributed by different control measures in place since January, though further studies is needed to understand how these measures had impacted on the viral transmission. The R_t in Wuhan also decreased substantially after implementation of a series of public health interventions[34]. Of note, the majority of the Hong Kong community have been wearing masks in public areas since late January, although the effectiveness of this and other measures in reducing viral transmission at population level was still unclear[35]. Phylogenetic analyses of 50 of the first 93 cases from late January to February by Leung *et al*. showed that COVID-19 cases and clusters in Hong Kong was mostly propagated from two ancestors, indicating limited importation and propagation of the virus in the community during this period[22]. Spatial clustering of cases around affluent and commercial areas in the Hong Kong Island region since March might be evidence that transmission might initiate from persons with more frequent overseas travel and an increased risk of importation. Secondary transmission was still relatively confined and not yet widespread in the entire territory.

Our R_t estimates in Hong Kong were comparable to those from Abbott *et al*.[17] but lower than that from Cowling *et al*. in March[36]. Both studies also employed the R package EpiEstim. The multi-countries estimates from Abbott *et al*. used publicly available data with only reporting dates available and the onset dates was estimated from a reporting delay distribution[17]. Our R_t estimates was largely comparable to that from Cowling *et al.*[36] with the exception from mid March onwards for which their R_t maintained at around 1 while ours was around 0.5. As they assumed completeness of reporting imported and local cases to be 99% and 80% and adjusted both data before model fitting, the proportion of local cases increased. This should lead to higher R_t estimates as local cases would account for a more significant proportion of new incident cases, as shown in our sensitivity analysis of including only local cases (Figure 7). In addition, the longer serial interval they adopted from Li *et al*.[19] would also result in higher R_t estimate (Figure 8). Our trend of effective reproduction number dropping from around 2 in January to less than 1 in late February was consistent with the estimate from a meta-population SEIR model by Yuan *et al.*[37]. As a sensitivity analysis we examined the impact of excluding cases detected by contact tracing (Figure 9). We found that this would lead to more fluctuation for the estimate in March as bigger clusters were detected. As containment effort had been maintained since the start of the epidemic, we expected limited change in surveillance sensitivity due to improvement in contact tracing and opted for only excluding cases detected at border in our main analysis. Our sensitivity analyses also showed that a smaller time window led to more abrupt changes in *R* estimates but was also subject to more statistical noise and a wider confidence bound (Figure 10). To balance the sensitivity and confidence of *R* estimates, we adopted a 7-day time window.

The mean serial interval estimated from our local data was shorter than the estimate by Li *et al*. in Wuhan (mean 7.5 days [SD 3.4])[19] but longer than the estimate by Nishiura *et al*. (mean 4.8 days [SD 2.3])[20] which collated data from a readily available database. Considering the serial interval of COVID-19 is shorter than those of SARS (mean 8.4 days [SD 3.8])[38] and MERS (mean 12.4 days [SD 2.8] in South Korea[39] and mean 6.8 days [SD 4.1] in Saudi Arabia[40]), early identification of cases is required to interrupt transmission. Among countries that implemented large scale testing, South Korea had successfully contained a large outbreak in February and the current daily number of new infections remains stably at less than 100 since early March[41]; Iceland[42] and Germany recorded much smaller number of cases as compared to other European countries such as Italy and Spain[11]. Wider availability of laboratory testing will shorten the delay from onset to confirmation, thus reducing the risk of onward transmission from an infectious person. Of note, similar to our observations, South Korea also detected a high incidence of young adult cases with less severe clinical manifestation[41]. The role of early detection among these age groups and its implication in the overall prevention and control of COVID-19 disease remains an unanswered research question[43].

There were several limitations to our study. First, the R_t estimates are subject to changes in surveillance sensitivity. With increasing testing capacity and enhanced surveillance gradually being implemented, it was likely that the surveillance system had become more sensitive and the R_t in earlier period might be underestimated. Second, the R_t estimate in the most recent days was likely under-estimated as cases which were infected recently would not be captured by the surveillance system. Third, we used date of reporting for asymptomatic cases in computing the incidence for R_t estimation and this might have affected the R_t estimates in different directions, depending on the clinical progression of these asymptomatic cases. The R_t is a measure

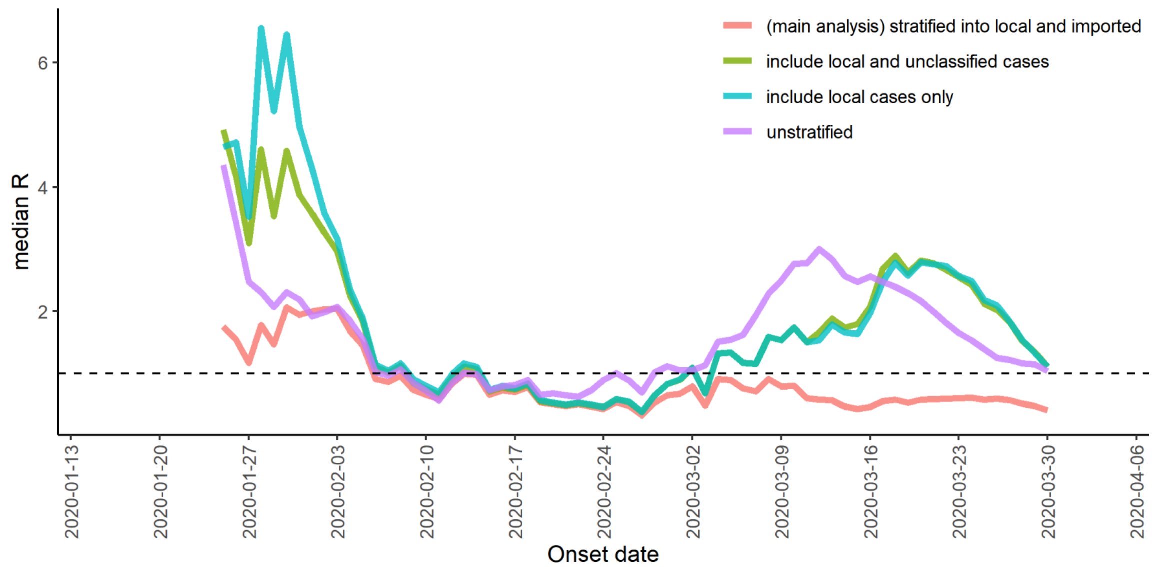


Figure 7. Sensitivity analysis on source of incidence data. We tested the effect of including different types of incidence data in this sensitivity analysis. Including only local (and unclassified) cases resulted in higher R_t estimate (purple and blue lines), as the risk of local transmission resulting from imported cases was ignored. Similarly, if all cases were included but not stratified by importation status, they would result in higher R_t estimates.

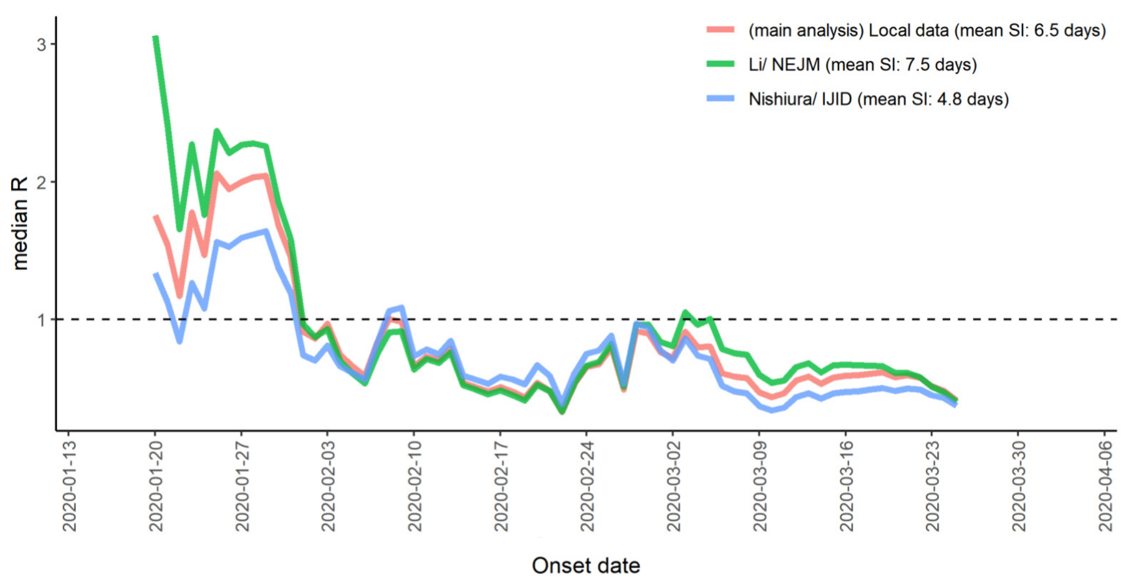


Figure 8. Sensitivity analysis on using different serial interval data. R_t estimates increased with serial interval.

of the transmission dynamics at the time of infection, inferred from observed time-series incidence data. A number of events occurred between the time of infection and report/ observation, namely symptom onset, doctor visit, specimen collection, testing, confirmation and reporting, isolation/ hospitalisation, recovery/ death[16]. Using incidence data based on date of disease onset aims at shortening this delay when compared to incidence data based on date of report/ observation. However, cases being asymptomatic at the time of observation (when they were reported and being interviewed/ investigated) might or might not develop clinical symptoms afterwards. For cases that remained asymptomatic all along, this simple imputation would have no effect on the distribution of the delay from infection to observation. For asymptomatic cases that developed symptoms afterwards, their actual date of onset should be later than the date of report (observation). Imputing their date

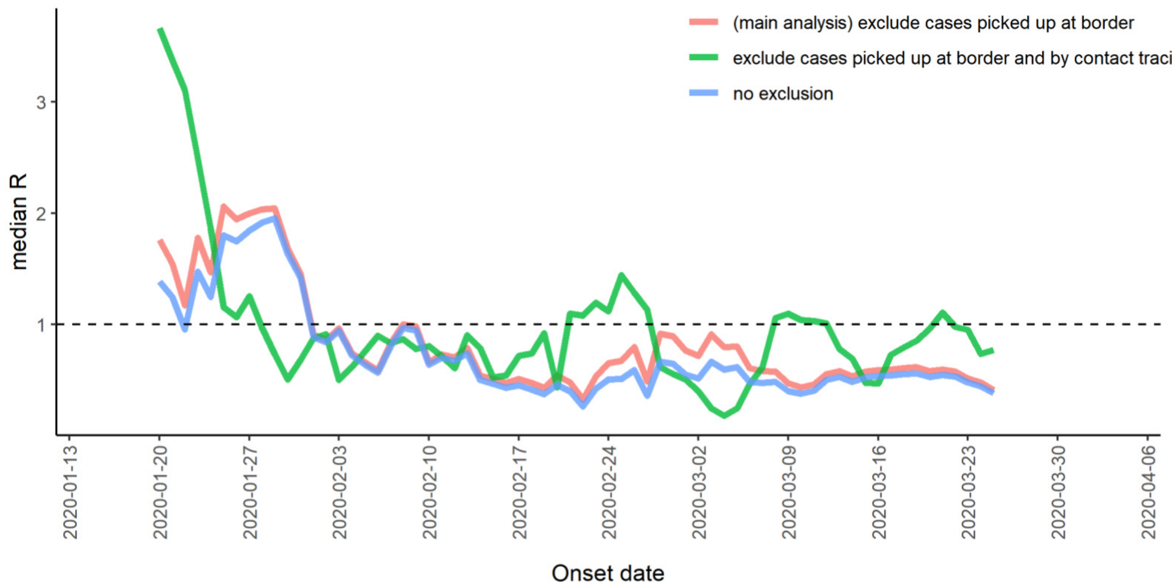


Figure 9. Sensitivity analysis on exclusion of cases detected by different systems. We tested the effect of excluding cases detected by different systems. Excluding cases detected by contact tracing would lead to more fluctuation of the R_t estimates in March as more cases from different clusters were detected through contact tracing during this period.

Figure 10. Sensitivity analysis on size of time window. R_t estimates become more stable with longer time windows as overlapping of incidence cases included in subsequent time points increased.

of onset with date of report would result in prematurely assigning these cases at earlier time of the incidence time-series than the actual date of onset and skewed the delay distribution slightly. This would result in under-estimation of R_t near the end of the observation period and over-estimation of R_t at earlier time. However, the dates of onset for these cases were likely not available until the cases were discharged, which would only be beneficial to a retrospective estimation of R_t but not for near real-time nowcasting. Jointly inferring the delay distribution from time of infection to onset (i.e. the incubation period) and the delay from onset to report/ observation, for both symptomatic and asymptomatic cases, might overcome this issue. However more complex computation is needed and this area is under active development as an ongoing effort for a more accurate R_t estimation[16] [personal communication with Dr Sam Abbott]. Fourth, we only used residential district to study the spatial distribution of the diseases but one could have been infected or infected others in places not close to their residential districts. Incorporating local movement data to study spatial distribution would require careful consideration of the time spent and the risk of being infected/passed along the infection. Regardless, residential district should be a good proxy of the geographical spread of the transmission.

In conclusion, we studied the transmission dynamics of COVID-19 in Hong Kong under two waves of importation pressure. The R_t served as an indicator to the transmission potential of COVID-19 in the community and allowed us to understand the impact of control measures during an epidemic. We showed that timely implementation of control measures such as social distancing was associated with the reduction of transmissibility of the infections. Further extension of the current analyses may include short-term forecasting once the number of imported cases stabilized. The same principle for R_t estimation has been extended to include contact and genetic sequence data to infer transmission network in outbreaks, which would be useful to enhance our understanding of the role of persons with different characteristics in the transmission chain.

Data availability

Source data

Data on COVID-19 cases in Hong Kong are available in the website of Government of Hong Kong SAR as csv files. Please visit "Data in Coronavirus Disease (COVID-19)" under data.gov.hk or use the following hyperlink:

https://data.gov.hk/en-data/dataset/hk-dh-chpsebcddr-novel-infectious-agent

Data used in the manuscript was downloaded on 14 Apr 2020. The following data files were used.

Details of confirmed and probable COVID-19 cases:
http://www.chp.gov.hk/files/misc/enhanced_sur_covid_19_eng.csv

Number of suspected cases under investigation, number of confirmed cases still under isolation and number of cases discharged:
http://www.chp.gov.hk/files/misc/latest_situation_of_reported_cases_covid_19_eng.csv

Residential buildings resided by confirmed and probable COVID-19 cases:
http://www.chp.gov.hk/files/misc/building_list_eng.csv

Population figures by age and district: https://www.bycensus2016.gov.hk/en/bc-dp.html.

Acknowledgements

We thanked all colleagues in the Department of Health Hong Kong involved in the investigation, control, data collection and sharing knowledge regarding COVID-19. We thanked Dr Sebastian Funk and Dr Sam Abbott of the London School of Hygiene and Tropical Medicine on the helpful discussion on estimation of reproduction number.

References

1. World Health Organization: **Coronavirus (COVID-19) events as they happen**. [cited 2020 Apr 12]. Reference Source
2. World Health Organization: **Naming the coronavirus disease (COVID-19) and the virus that causes it**. [cited 2020 Apr 12]. Reference Source
3. World Health Organization: **WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020**. [cited 2020 Apr 12]. Reference Source
4. World Health Organization: **Coronavirus disease (COVID-2019) situation reports**. [cited 2020 Apr 15]. Reference Source
5. Centre for Health Protection of the Department of Health, Hong Kong SAR: **Letters-to-doctors on 7 January 2020: Statutory Reporting of "Severe Respiratory Disease associated with a Novel Infectious Agent"**. [cited 2020 Apr 12]. Reference Source
6. Centre for Health Protection of the Department of Health, Hong Kong SAR: **Letters-to-doctors on 23 February 2020: Updated Situation of COVID-19 and Enhanced Surveillance in Out-patient Settings in the Hospital Authority**. 2020; [cited 2020 Apr 12]. Reference Source
7. Centre for Health Protection of the Department of Health, Hong Kong SAR: **Letters-to-doctors on 6 March 2020: Enhanced Laboratory Surveillance with Testing for COVID-19 at Clinics of Private Medical Practitioners**. [cited 2020 Apr 12]. Reference Source
8. Centre for Health Protection of the Department of Health, Hong Kong SAR: **Letters-to-doctors on 11 March 2020: Update on the list of areas with active community transmission of COVID-19 and further enhancement on the surveillance at clinics of Private Medical Practitioners**. [cited 2020 Apr 12]. Reference Source
9. Centre for Health Protection of the Department of Health, Hong Kong SAR: **Latest situation of cases of COVID-19**. Reference Source
10. The Government of the Hong Kong SAR: **Press Releases on Coronavirus Disease (COVID-19)**. [cited 2020 Apr 12]. Reference Source
11. European Centre for Disease Prevention and Control: **Situation update worldwide, as of 15 April 2020**. [cited 2020 Apr 16]. Reference Source
12. Cori MA: **Package "EpiEstim" Title Estimate Time Varying Reproduction Numbers from Epidemic Curves**. 2019; [cited 2020 Apr 12]. Reference Source
13. Wallinga J, Teunis P: **Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal.** *Am J Epidemiol.* 2004; **160**(6): 509–16. PubMed Abstract | Publisher Full Text | Free Full Text
14. Cori A, Ferguson NM, Fraser C, *et al.*: **A new framework and software to estimate time-varying reproduction numbers during epidemics.** *Am J Epidemiol.* 2013; **178**(9): 1505–12. PubMed Abstract | Publisher Full Text | Free Full Text
15. Thompson RN, Stockwin JE, van Gaalen RD, *et al.*: **Improved inference of time-varying reproduction numbers during infectious disease outbreaks.** *Epidemics.* 2019; **29**: 100356. PubMed Abstract | Publisher Full Text | Free Full Text

16. Gostic KM, McCough L, Baskerville E, *et al.*: **Practical considerations for measuring the effective reproductive number, R_t.** *medRxiv.* 2020.
PubMed Abstract | Publisher Full Text | Free Full Text

17. Sam A, Joel H, James MD, *et al.*: **Temporal variation in transmission during the COVID-19 outbreak**. [cited 2020 Apr 15].
Reference Source

18. Fine PE: **The interval between successive cases of an infectious disease.** *Am J Epidemiol.* 2003; **158**(11): 1039–47.
PubMed Abstract | Publisher Full Text

19. Li Q, Guan X, Wu P, *et al.*: **Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia.** *N Engl J Med.* 2020; **382**(13): 1199–207.
PubMed Abstract | Publisher Full Text | Free Full Text

20. Nishiura H, Linton NM, Akhmetzhanov AR: **Serial interval of novel coronavirus (COVID-19) infections.** *Int J Infect Dis.* 2020; **93**: 284–6.
PubMed Abstract | Publisher Full Text | Free Full Text

21. Census and Statistics Department, Hong Kong SAR: **Population**. [cited 2020 Apr 17].
Reference Source

22. Leung KSS, Ng TTL, Wu AKL, *et al.*: **A Territory-wide study of COVID-19 cases and clusters with unknown source in Hong Kong community: A clinical, epidemiological and phylogenomic investigation.** *medRxiv.* 2020; 2020.03.30.20045740.
Publisher Full Text

23. The Government of the Hong Kong SAR press release: **Quarantine advisory on bars' cluster**. [cited 2020 Apr 20].
Reference Source

24. The Government of the Hong Kong SAR press release: **Restrictions on bars gazetted**. [cited 2020 Apr 20].
Reference Source

25. Centre for Health Protection: **Department of Health Hong Kong SAR press release on confirmed COVID-19 cases (1 Apr 2020)**. [cited 2020 Apr 12].
Reference Source

26. Centre for Health Protection: **Department of Health Hong Kong SAR press release on COVID-19**. 2020; [cited 2020 Apr 15].
Reference Source

27. The Government of the Hong Kong SAR press release: **Special work arrangement extended**. [cited 2020 Apr 15].
Reference Source

28. HKU: **Medicine facebook post-Attitudes towards preventive measures against COVID-19**. [cited 2020 Apr 15].
Reference Source

29. Google: **COVID-19 Community Mobility Reports**. [cited 2020 Apr 16].
Reference Source

30. CDC COVID-19 Response Team: **Preliminary Estimates of the Prevalence of Selected Underlying Health Conditions Among Patients with Coronavirus Disease 2019 — United States. February 12-March 28, 2020.** *MMWR Morb Mortal Wkly Rep.* 2020; **69**(13): 382–6.
PubMed Abstract | Publisher Full Text | Free Full Text

31. Zhao S, Lin Q, Ran J, *et al.*: **Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak.** *Int J Infect Dis.* 2020; **92**: 214–7.
PubMed Abstract | Publisher Full Text | Free Full Text

32. Kucharski AJ, Russell TW, Diamond C, *et al.*: **Early dynamics of transmission and control of COVID-19: a mathematical modelling study.** *Lancet Infect Dis.* 2020; **20**(5): 553–558.
PubMed Abstract | Publisher Full Text | Free Full Text

33. Riou J, Althaus CL: **Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020.** *Euro Surveill.* 2020; **25**(4): 2000058.
PubMed Abstract | Publisher Full Text | Free Full Text

34. Pan A, Liu L, Wang C, *et al.*: **Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China.** *JAMA.* 2020; **323**(19): 1915–1923.
PubMed Abstract | Publisher Full Text | Free Full Text

35. Chan KH, Yuen KY: **COVID-19 epidemic: disentangling the re-emerging controversy about medical facemasks from an epidemiological perspective.** *Int J Epidemiol.* 2020; dyaa044.
PubMed Abstract | Publisher Full Text | Free Full Text

36. Cowling BJ, Ali ST, Ng TW, *et al.*: **Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study.** *Lancet Public Health.* 2020; **5**(5): e279–e288.
PubMed Abstract | Publisher Full Text | Free Full Text

37. Yuan HY, Mao A, Han G, *et al.*: **Effectiveness of quarantine measure on transmission dynamics of COVID-19 in Hong Kong**. [cited 2020 Apr 18];
Publisher Full Text

38. Lipsitch M, Cohen T, Cooper B, *et al.*: **Transmission dynamics and control of severe acute respiratory syndrome.** *Science.* 2003; **300**(5627): 1966–70.
PubMed Abstract | Publisher Full Text | Free Full Text

39. Cowling BJ, Park M, Fang VJ, *et al.*: **Preliminary epidemiological assessment of MERS-CoV outbreak in South Korea, May to June 2015.** *Euro Surveill.* 2015; **20**(25): 7–13.
PubMed Abstract | Publisher Full Text | Free Full Text

40. Cauchemez S, Nouvellet P, Cori A, *et al.*: **Unraveling the drivers of MERS-CoV transmission.** *Proc Natl Acad Sci U S A.* 2016; **113**(32): 9081–6.
PubMed Abstract | Publisher Full Text | Free Full Text

41. COVID-19 National Emergency Response Center, Epidemiology and Case Management Team, Korea Centers for Disease Control and Prevention: **Coronavirus Disease-19: The First 7,755 Cases in the Republic of Korea.** *Osong Public Health Res Perspect.* 2020; **11**(2): 85–90.
PubMed Abstract | Publisher Full Text | Free Full Text

42. Gudbjartsson DF, Helgason A, Jonsson H, *et al.*: **Spread of SARS-CoV-2 in the Icelandic Population.** *N Engl J Med.* 2020; **382**(24): 2302–2315.
PubMed Abstract | Publisher Full Text | Free Full Text

43. Davies NG, Klepac P, Liu Y, *et al.*: **Age-dependent effects in the transmission and control of COVID-19 epidemics | CMMID Repository**. 2020; [cited 2020 Apr 15].
Reference Source

Open Peer Review

Current Peer Review Status: ? ? ?

Version 2

Reviewer Report 14 June 2021

https://doi.org/10.21956/wellcomeopenres.17678.r44034

?

Philip Arevalo
Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA

The authors present an estimation of the time-varying reproductive number (R(t)) of SARS-CoV-2 in Hong Kong from late January to early April 2020. Importantly, they explicitly adjust for imported cases in order to more accurately estimate the true local transmission rate. The demographic characteristics of the infected and the way in which delays from onset to testing, discharge, and death changed over the course of the time period under study were also reported. The authors found that the second wave of the epidemic was driven largely by importation and that local transmission, while close to 1, was largely contained. However, hospital usage increased rapidly during the second wave despite low inferred local transmission.

I only have one main comment relating to the possibility that the true fraction of asymptomatic cases may also have varied over time. There are hints of this from the asymptomatic fraction of *reported* cases increasing over time. If the asymptomatic fraction increased over time (for example, from the epidemic shifting more toward younger individuals), then the relationship between observed cases and true infections might also have changed and confounded estimates of R(t). I acknowledge that adjusting for the asymptomatic fraction here may be impractical, but this limitation should be addressed in the discussion.

One minor comment I have is that it would be useful to provide a citation for the incubation period used to shift cases for R(t) estimation (5 days) although it does match with my knowledge of the literature.

Otherwise, I find that the manuscript is scientifically sound.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemic modeling, ecology and evolution of infectious disease, influenza evolution.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Version 1

Reviewer Report 07 August 2020

https://doi.org/10.21956/wellcomeopenres.17436.r38916

? **Lindsay Keegan**
University of Utah, Salt Lake City, UT, USA

Overall this presents a nice epidemiological study of COVID-19 in Hong Kong. The authors set out to describe the transmission dynamics and estimate the impact of control measures on transmission.

Specific comments:

- The authors describe using the package EpiEstim to estimate Rt and cite 3 different method EpiEstim supports. However, these are three distinct methods and yet the authors only present a single Rt estimate. The authors should either state which method for estimating Rt they use or how they are combining the three outputs for a single Rt curve.

- Related to the above comment, there has been a lot of recent discussion about the viability of different methods to estimate Rt, the authors should detail why they think the method they chose (or how they combine the three methods) is appropriate.

- The authors use date of symptom onset to estimate Rt, the authors should discuss how this data was collected, was it based on patients remembering when they first felt ill? How are missing data handled or do all COVID-19 patients remember when they fell ill? How was this collected for patients who died?

- The authors state that “for asymptomatic cases, their date of reporting was used instead.” The authors should discuss how this might this impact the estimates of Rt?

- The authors should specify confidence intervals for each mean estimate reported in the text. E.g., Rt, serial interval, etc.

- The first paragraph in methods describes the protocol for COVID testing and surveillance. It would be beneficial if the authors had a break-out paragraph describing the data they use in the analyses. As written it’s difficult to extract what you actually used.

- The authors say “Compared to the whole population, there was a considerably higher proportion of cases among those aged 10 to 39 years for both genders” what is “considerably higher”?

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Infectious Disease Dynamics, Epidemiology, Mathematical Modelling, Theoretical Ecology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have

significant reservations, as outlined above.

Author Response 09 Nov 2020

YUNG WAI CHAN, Department of Health, Hong Kong

We thank Dr Keegan for her valuable comments on the manuscript. We have updated various parts of the manuscript in the second version. Please find our responses to the specific comments below.

Specific comments:

Comment 1: The authors describe using the package EpiEstim to estimate Rt and cite 3 different method EpiEstim supports. However, these are three distinct methods and yet the authors only present a single Rt estimate. The authors should either state which method for estimating Rt they use or how they are combining the three outputs for a single Rt curve.

Related to the above comment, there has been a lot of recent discussion about the viability of different methods to estimate Rt, the authors should detail why they think the method they chose (or how they combine the three methods) is appropriate.

Response: We estimated Rt by using the method for estimating the instantaneous reproduction number proposed by Cori A *et al* and Thompson RN *et al* implemented in the EpiEstim package, for its more real-time estimation compared to the method proposed by Wallinga *et al*. Gostic KM *et al* has a recent pre-print comparing the different methods of measuring the effective reproduction number, for which we have added in the text. We expanded the Methods under sub-heading “Estimation of time-varying reproduction number” which now reads:

“The daily number of new cases are expected to increase over time if R exceeds 1 whilst an epidemic is expected to decline if R is consistently reduced to below 1. The time-varying reproduction number (R_t) refers to the effective reproduction number at different time points in an epidemic and it was estimated based on the daily COVID-19 incidence using the R package EpiEstim version 2.2-1 [Cori MA et al EpiEstim package]. Grossly, there are two variations of the effective reproduction number, namely the case reproduction number and the instantaneous reproduction number. The case reproduction number, defined as the expected number of secondary infections that an infected individual at time t will eventually cause, was first described by Wallinga et al. In contrast, the instantaneous reproduction number is the expected number of secondary infections resulting from the number of infected individuals and their relative infectiousness at time t. It was first described by Cori et al. and an improved framework taking into account importation status of incidence data was later proposed by Thompson et al. Both iterations of effective reproduction number can be estimated using the EpiEstim package with time-series incidence data. However, these incidence data in an on-going epidemic are right truncated (i.e. infections occurring near the end of observation period were not yet detected or reported). As estimation of the case reproduction number relies on the cases observed after time t, it is more prone to under-estimation towards the end of the observed time series, as detailed in Gostic et al. Therefore, we used the EpiEstim package to estimate the instantaneous reproduction number as R_t in this study as the COVID-19 epidemic was still ongoing at the time of analysis.”

Wellcome Open Research 2020, 5:91 Last updated: 14 JUN 2021

Comment 2: The authors use date of symptom onset to estimate Rt, the authors should discuss how this data was collected, was it based on patients remembering when they first felt ill? How are missing data handled or do all COVID-19 patients remember when they fell ill? How was this collected for patients who died?

Response: The date of symptom onset was collected as part of the case-based investigation, based on patient recall and supplemented by clinical records provided by reporting physicians. For the cases included in this analysis, there was no missing data on date of onset for symptomatic cases, and no COVID-19 cases passed away before date of report. We supplemented Methods in the article accordingly:

"As part of the case-based investigation, the CHP also collected information on cases' demographics and clinical symptoms (including date of onset), based on patient recalls and information provided by reporting doctors. The CHP used these data to generate reports of daily situation updates on these confirmed and probable cases."

Comment 3: The authors state that "for asymptomatic cases, their date of reporting was used instead." The authors should discuss how this might this impact the estimates of Rt?

Response: The effect of this simple imputation might affect the incidence time-series data for Rt estimation, depending on the clinical progress of these asymptomatic cases. Briefly, for asymptomatic cases that would eventually develop clinical symptoms, this imputation might have mis-assigned cases to earlier dates of the incidence time series, resulting in under-estimation of Rt near the end of the observation period (and over-estimation in earlier time). On the other hand, some cases would remain asymptomatic all along, and there is indeed no date of onset for these cases. The distribution of delay from onset to observation affects the incidence time series and Rt estimation. We are keeping track of the development in this area and hope to improve this simple imputation later with joint estimation of both the incubation period and the delay from onset to report, ideally for both symptomatic and asymptomatic cases. We supplemented the para regarding this limitation under the Discussion as follow. It now reads as follow:

"Third, we used date of reporting for asymptomatic cases in computing the incidence for R t estimation and this might have affected the R t estimates in different directions, depending on the clinical progression of these asymptomatic cases. The Rt is a measure of the transmission dynamics at the time of infection, inferred from observed time-series incidence data. A number of events occurred between the time of infection and report/ observation, namely symptom onset, doctor visit, specimen collection, testing, confirmation and reporting, isolation/ hospitalisation, recovery/ death [Gostic KM et al.]. Using incidence data based on date of disease onset aims at shortening this delay when compared to incidence data based on date of report/ observation. However, cases being asymptomatic at the time of observation (when they were reported and being interviewed/ investigated) might or might not develop clinical symptoms afterwards. For cases that remained asymptomatic all along, this simple imputation would have no effect on the distribution of the delay from infection to observation. For asymptomatic cases that developed symptoms afterwards, their actual date of onset should be later than the date of report (observation). Imputing their date of onset with date of report would result in prematurely

assigning these cases at earlier time of the incidence time-series than the actual date of onset and skewed the delay distribution slightly. This would result in under-estimation of Rt near the end of the observation period and over-estimation of Rt at earlier time. The dates of onset for these cases were likely not available until the cases were discharged, which would only be beneficial to a retrospective estimation of R t but not for near real-time nowcasting. Jointly inferring the delay distribution from time of infection to onset (i.e. the incubation period) and the delay from onset to report/ observation, for both symptomatic and asymptomatic cases, might overcome this issue. However more complex computation is needed and this area is under active development as an ongoing effort for a more accurate Rt estimation [Gostic KM et al. and personal communication with Dr Sam Abbott].”

Comment 4: The authors should specify confidence intervals for each mean estimate reported in the text. E.g., Rt, serial interval, etc.

Response: We have added confidence intervals for Rt estimates. In addition to mean and SD, we also added the median and 95% quantile of serial interval estimated from fitting to discretized gamma distribution.

Comment 5: The first paragraph in methods describes the protocol for COVID testing and surveillance. It would be beneficial if the authors had a break-out paragraph describing the data they use in the analyses. As written it’s difficult to extract what you actually used.

Response: We expanded the description of the data used in our analyses under the first two para in Methods. It now reads as follow:

“Contact tracing would be carried out for every confirmed and probable case to identify any exposure persons who had contact with the case. They would be put under quarantine and medical surveillance. As part of the case-based investigation, the CHP also collected information on cases’ demographics and clinical symptoms (including date of onset), based on patient recalls and information provided by reporting doctors. The CHP used these data to generate reports of daily situation updates on these confirmed and probable cases [Reference local situation report of CHP]. Using the data from these daily updates and the Government’s press releases, we collated a line-list of COVID-19 cases with basic demographic, epidemiological and clinical information including age, gender, residential status (i.e. whether the case was a Hong Kong resident or not) and residential district (grossly divided into 18 districts in Hong Kong), date of symptom onset and date of reporting (notified to the CHP), importation status and discharge status (whether the case was still under admission and isolation, discharged home or death). The CHP assigned an official classification of importation status was to each case based on their travel history and contact with confirmed/ probable COVID-19 cases during the incubation period. These importation status include: “imported”, “close contact of imported case”, “local”, “close contact of local case”, “possibly local” and “close contact of possibly local case”. To facilitate our data analyses including estimation of the reproduction number, we regrouped the importation status based on the official case classification 9 and travel history of the cases before onset of symptom into “imported” and “local” (i.e. non-imported) cases (Table 2).”

Comment 6: The authors say “Compared to the whole population, there was a considerably higher proportion of cases among those aged 10 to 39 years for both genders” what is

"considerably higher"?
Is the work clearly and accurately presented and does it cite the current literature?

Response: We added figures in the relevant para. It now reads:

"The proportion of cases by age and sex was compared to those in the whole population. Those aged 10 to 39 years attributed to a higher proportion for both sexes among the cases reported (59% of cases were aged 10 to 39 years for both sexes, compared to 35 to 36% of the general population aged 10 to 39 years)."

Competing Interests: No competing interests were disclosed.

Reviewer Report 17 June 2020

https://doi.org/10.21956/wellcomeopenres.17436.r38687

?

Jing Yuan
Department of Pharmacy Practice and Administration, Rutgers University, Piscataway, NJ, USA

Overall, this study evaluated the transmission dynamic of COVID-19 in Hong Kong, also tried to examine the impact of different control measures on containing the spread of COVID-19.

1. Estimation of time-varying reproduction number. The traditional methods, e.g., exponential growth or maximum likelihood can only be used in the exponential stage of outbreak. There were a large number imported cases since late March, therefore, the statistical method used to estimate Rt should take into account of the imported cases. Please clarify if time-varying Rt could measure imported cases.

2. Discussion. "The transmission reduced subsequently as R_t was largely below 1 in February. " It looks like the Rt was continuously decreasing since late February, despite of the second peak of COVID-19.

3. "The lower transmissibility might be attributed by different control measures in place since January." Many factors influence the estimation of Rt, e.g., generation time, statistical methods, etc. It is hard to say the lower transmissibility observed in Hong Kong was attributed by the control measures. Please revise.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiology, statistical modelling, infectious disease.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 25 Jun 2020

YUNG WAI CHAN, Department of Health, Hong Kong

We thank the review and thoughtful comments by Dr Jing Yuan. After going through the comments, we have made changes in version 2 of the article to clarify on the estimation of time-varying reproduction number in Methods and its interpretation in the Discussion. Below please find our responses to the specific comments.

Comment 1: Estimation of time-varying reproduction number. The traditional methods, e.g., exponential growth or maximum likelihood can only be used in the exponential stage of outbreak. There were a large number imported cases since late March, therefore, the statistical method used to estimate Rt should take into account of the imported cases. Please clarify if time-varying Rt could measure imported cases.

Response to comment 1: Our estimation of time-varying reproduction number (R_t) took into account imported and local cases. The EpiEstim package used to estimate R_t in this study caters the probability of a secondary case arising from incident case that could either be local or imported (for details, please refer to Thompson RN *et al.* Epidemics 2019 which was quoted as reference 15 of this article). In our R_t estimation, we fit the daily incidence stratified by the importation status of cases (imported and likely locally acquired). Cases without clear epidemiological evidence to be classified as“imported”or“local”were considered as “unclassified”(Table 2) and separate estimates was conducted by including them either as imported or local (red and green lines in the lower plot of Figure 1). The classification of importation status and fitting of imported and local incident cases to obtain the R_t were detailed in Methods. We amended the relevant para in Methods to clarify this and it now reads as follow:

*"**The EpiEstim package caters the probability of a secondary case arising from incident case that could be either local or imported [15]**. Thus we fit the daily incidence stratified by the importation status of cases (imported and likely locally acquired)........."*

Comment 2: Discussion. "The transmission reduced subsequently as Rt was largely below 1 in February. " It looks like the Rt was continuously decreasing since late February, despite of the second peak of COVID-19.

Response to comment 2: We agreed that the R_t estimate was indeed decreasing from late February/ early March to mid March, despite the second peak of COVID-19. As detailed in Methods and our response to comment 1, the R_t estimation by the EpiEstim package readily took into account the probability that secondary cases could arise from both imported and local incident cases. Since majority of the cases reported in March were imported and they would not be regarded as secondary cases infected locally, their occurrence should not infer to local transmission and hence the R_t estimated was low in this period despite an apparently larger second peak. We beefed up relevant para in the Discussion which now reads as follow:

*".........A local case of infected nurse working in a residential care home for the elderly as part of large social venue cluster further highlighted such potential risk. **The R_t decreased from around 1 in early March to around 0.5 and stabilized in mid March, despite a second peak contributed mainly by imported cases. As these imported cases would not be regarded as secondary cases infected locally, their occurrence should not infer to local transmission and hence the R_t estimated was relatively low in this period with an apparently larger second peak.** However, this second wave of epidemic led to a rapid increase in hospital bed usage for both confirmed and suspected cases........."*

Comment 3: "The lower transmissibility might be attributed by different control measures in place since January." Many factors influence the estimation of Rt, e.g., generation time, statistical methods, etc. It is hard to say the lower transmissibility observed in Hong Kong was attributed by the control measures. Please revise.

Response to comment 3: We agreed that other factors might also influence the R_t estimates and we have amended this para in the discussion and included other plausible explanations of lower R_t estimates. This para now reads as follow:

*"The R estimates in Hong Kong were considerably lower than that estimated in Wuhan, which ranged from 2 to 4 when there was widespread transmission [18, 30–33] . T**he lower R estimates could be attributed to different factors, such as difference in contact patterns, different methodologies and serial intervals/ generation time adopted in estimation of R, as well as lower transmissibility of SARS-CoV-2 in Hong Kong in our study period.** The lower transmissibility might be attributed by different control measures in place since January**, though further studies is needed to understand how these measures had impacted on the viral transmission.** The R_t in Wuhan also decreased substantially after implementation of a series of public health interventions [33] ...."*

Competing Interests: We have no competing interests to disclose.

Comments on this article

Version 1

Reader Comment 29 Jun 2020

Andrei R. Akhmetzhanov, Hokkaido University, Sapporo, Japan

In my opinion, the manuscript of Chan *et al.* provides a good overview of the first three months of the COVID-19 pandemic in Hong Kong. The results are consistent with other studies, and the present Table and graphs are quite informative. However, I think some aspects could be improved and I have some remarks/critique listed below. My main concern on how the authors address the reporting delay in their estimates of Rt and serial interval.

Remarks:

- Abstract, Results: The sentence "The Rt increased to approximately 2 and approximately 1…" does not sound very rigorously. Is it possible to rephrase and give more rigor to it? As the authors obtained some estimates, they could provide the median and 95% CI. They could also indicate the dates when the increased values of Rt were observed.
- Abstract, Conclusions: this section could be better linked to the Results. At the moment it is quite general, and is not really connected to the main subject of the paper.
- Introduction: I think the authors could make it clearer that the second wave was caused by importation events, not by undocumented or documented chain of local transmissions. This could be done by one or two additional sentences.
- Introduction: I think also that the authors could specify better their research agenda for the report. It is good to have some descriptive analysis as stated in the last sentence, but it could be probably stronger if they specify one or two specific questions (they do study the dynamics of Rt, then they write about that also).
- Methods: "In brief, imported cases were those with travel history who had spent a significant amount of their incubation period at a place(s) where there was evidence of local transmission…" - this sentence is a bit tricky because it is yes and no. There should be also no active chain of transmission at the place of residence, and it would be nice to have some specificity about how much is "significant".
- Methods: "For asymptomatic cases, their date of reporting was used instead." - what is the date of reporting here? The date of lab confirmation or date of notification by CHP? Could the authors also briefly specify what the percentage of asymptomatic cases was detected? (for example in the parentheses)
- Table 1: I remember there was some critique that the passengers using a bullet train from Mainland China were not screened for some long time. The authors slightly specify this by writing "Since 1 Feb, health declaration arrangements were expanded … gradually", but this

does not give much clarity. I may recall that there was a particular date for introducing the screening of passengers from bullet trains. It would be great to have some more information on this aspect because the rail is the main connection of Hong Kong with the Mainland. Apart of this, I found the Table 1 quite informative and organized.
- Page 10: “For this wave, the Rt increased from around 1.5 in week 3 to around 2 in week 4, it then gradually decreased to under 1 along with decrease in reported number of cases between week 6 to 8.” The authors can be more specific here. The median Rt was around 2, but the CIs were quite wide. So it would be better to give the exact value of the estimate of Rt on week 3 and report it as XX (95% CI: XX, XX) (or in any other form preferable to the authors). The same applies to other estimated values in the Results section.
- Page 10: “The mean serial interval among 47 infector-infectee pairs identified was 6.5 days [standard deviation (SD) 4.7 days]” - What are the CIs for these estimates?
- Page 10: “Among those already discharged/ deceased, 8% were under intensive care unit (ICU) care or considered serious.” - I wonder, what are those cases considered serious and not under ICU care? Could the authors specify their definition?
- Page 11: “Both the median duration from onset to reporting and from reporting to discharge decreased with time as the epidemic evolved” Did the authors take into account the right truncation in the data? People who had a longer period from onset to reporting may not have been reported yet in the last week.
- Rt: the same applies to the estimation of the effective reproduction number: how do the authors account for the reporting delay in their estimation of Rt? Some cases could have been not reported yet at the time of data collection.
- Page 13, The observed difference in age between imported and local cases: I think it is a bit missing from the study, but the imported cases could be younger than local, because younger individuals are more likely to get screened/tested while travelling. From another side, the locally acquired infections among young individuals are likely to be missed by the surveillance system if they are mild and not a part of the investigated clusters. In my opinion, the authors could mention this a bit more in their manuscript.

Typos, etc.:

- Abstract: “from Wu**ha**n”
- Methods: SARS-CoV-2 not SARS-Co-V-2
- Table 3: It would be good to specify that the Onset date is in the format “dd-mm-yy”
- Table 3: what is the difference between “Dating” and “Social”?
- Page 8, 5th line from above, left panel: to be consistent, “Among the 436 important cases in March” not “in Mar”

Competing Interests: No competing interests to be declared.