Brief Report

Accounting for imported cases in estimating the time-varying reproductive number of COVID-19

Tim K. Tsang¹, Peng Wu¹,², Eric H. Y. Lau¹,², Benjamin J. Cowling¹,²

Affiliations:
1. WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China.
2. Laboratory of Data Discovery for Health Limited, Hong Kong Science Park, New Territories, Hong Kong.

Corresponding author:
Tim Tsang, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 21 Sassoon Road, Pokfulam, Hong Kong.
Tel: +852 3917 6761; Fax: +852 3520 1945; email: timtsang@connect.hku.hk

Running head: Imported cases and COVID-19 transmissibility

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
SOURCES OF FINANCIAL SUPPORT

This project was supported by the Health and Medical Research Fund, Food and Health Bureau, Government of the Hong Kong Special Administrative Region (grant no. COVID190118), and the Collaborative Research Fund from the University Grants Committee of Hong Kong (project no. C7123-20G).

COMPETING INTERESTS STATEMENT

BJC consults for Roche and Sanofi Pasteur. The authors report no other potential conflicts of interest.

ACKNOWLEDGMENTS

The authors thank Stephanie Gao and Faith Ho for technical assistance.
Accounting for imported cases in estimating the time-varying reproductive number of COVID-19

ABSTRACT

**Background:** Estimating the time-varying reproductive number, $R_t$, is critical for monitoring transmissibility of an emerging infectious disease during outbreaks. When local transmission is effectively suppressed, imported cases could substantially impact transmission dynamics.

**Methods:** We developed methodology to estimate separately the $R_t$ for local cases and imported cases, since certain public health measures aim only to reduce onwards transmission from imported cases. We applied the framework to data on COVID-19 outbreaks in Hong Kong.

**Results:** We estimated that the $R_t$ for local cases decreased from above one in the early phase of outbreak to below one after tightening of public health measures. Assuming the same infectiousness of local and imported cases underestimated $R_t$ for local cases due to control measures targeting travelers.

**Conclusions:** When a considerable proportion of all cases are imported, the impact of imported cases in estimating $R_t$ is critical. The methodology described here can allow for differential infectiousness of local imported cases.
During an emerging infectious disease epidemics, monitoring transmissibility is important for guiding the implementation of public health measures and providing situational awareness on the effectiveness of interventions [1-3]. This can be achieved by estimating the time-varying reproductive number, $R_t$, a measure of the transmissibility of the virus over time. A wide range of methods have been proposed [3-8] but the impact of imported cases on estimation of $R_t$ is rarely explored [5].

Imported cases are cases that have been infected outside an area but travelled into and identified in the local area during their incubation period or symptomatic period. They have not been infected locally, but could potentially generate local transmissions. It is expected that the impact of imported cases on estimation of $R_t$ would be limited when a local epidemic is underway as most infections occur locally. However, imported cases can play a more important role in the early stages of a local epidemic, or when local transmission is being suppressed by public health measures. While one study [5] proposed a modification by assuming the same infectiousness among imported and local cases, such modification may not be accurate if there are specific interventions targeting imported cases, such as testing on arrival at the border, or quarantine for inbound travelers [9].

To address such issues related to imported cases, we extend the framework in Cori et al [4] to estimate the $R_t$ for local cases and imported cases separately. Using the first two waves of COVID-19 outbreaks in Hong Kong as an example,
we illustrate the use of inappropriate methods to estimate $R_t$ would lead to a biased assessment of transmissibility.

**METHODS**

**Sources of Data**

COVID-19 has been a notifiable disease in Hong Kong since 8 January 2020. Data on laboratory-confirmed COVID-19 cases were obtained from the webpage of Centre for Health Protection. Cases are classified as “imported cases”, “local cases epidemiologically linked with imported cases”, “unlinked local cases” and local cases epidemiologically linked with local cases” according to their epidemiological characteristics and location of infection [10].

**Statistical analysis**

We examined different methods to handle imported cases in the estimation of $R_t$. First, when we ignored the fact that imported cases were infected overseas and assumed them as local cases, $R_t$ could be estimated in the framework described by Cori et al. [4]. In this approach, $R_t$ was the ratio between the number of new cases at time $t$ and the total infectiousness of cases at time $t$, given by $\sum_{s=1}^{t-1} l_{t-s} w_s$, where $l_t$ was the number of infections at time $t$ and $w_t$ was probability distribution of infectiousness profile. Second, assuming equal infectiousness for imported cases and local cases [5], the total infectiousness of cases could be represented by $\sum_{s=1}^{t-1} (l_{t-s}^{local} + l_{t-s}^{imported}) w_s$, where $l_{t-s}^{local}$ and $l_{t-s}^{imported}$ were the numbers of local and imported cases at time $t-s$. 


To account for the different infectiousness of local cases and imported cases due to travel-related measures [11], we extended the framework to estimate the $R_t$ of local cases and imported cases. We modified the estimation so that $R_t$ for imported cases was the ratio of the number of new local cases infected by imported cases at time $t$ and the total infectiousness of imported case at time $t$, given by $\sum_{s=1}^{t-1} I_{t-s}^{Imported} w_s$. The $R_t$ for local cases could be calculated by using the same approach.

For unlinked local cases, their source of infection was unclear. In addition, because of the potential for pre-symptomatic infectiousness of COVID-19 [12], it is preferable to conduct inference on the epidemic curve by infection date (unobserved) instead of by case confirmation date (observed). Therefore, we used a bootstrap approach that involved sampling of the source of unlinked local cases (linked with local cases or imported cases), and a deconvolution approach [13] to reconstruct the epidemic curve by infection date from the epidemic curve by confirmation date, with the distribution of delay from infection to confirmation, which was the convolution distribution of the incubation period distribution, with mean 5.2 days (SD 3.9) [14], and the empirical distribution of delay from onset to reporting. Our inference was based on a Bayesian framework and we developed Markov chain Monte Carlo algorithm to estimate the model parameters (Appendix). We conducted simulation to validate that our approach could obtain unbiased estimates of $R_t$ for imported and local cases.

We used these three approaches to fit the data for the first two waves of COVID-19 outbreak in Hong Kong. We compared the estimates of $R_t$ from the alternative
All analyses were conducted in R version 3.5.2 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

From 18 January to 8 May 2020, there were 719 imported cases and 326 local cases detected in Hong Kong. Among the 326 local cases, there were 62 (19%), 73 (22%) and 191 (59%) cases epidemiologically linked with imported cases, unlinked local cases, and epidemiologically linked with local cases respectively (Figure 1A).

Based on our approach, we estimated that the surge in local case numbers corresponded to an estimated $R_t$ for local cases greater than one during March 14 to 26, which was the period before public health measures were tightened (Figure 1B). During the period with implemented interventions in the community (civil servant special work arrangement and additional social distancing measure), the estimated $R_t$ for local cases was decreasing. The $R_t$ for imported cases were well below one even when there were more than 10 daily imported cases during early March (Figure 1C), consistent with the implementation of a 14-day quarantine for inbound travelers.

We explored two alternative approaches to account for imported cases when estimating $R_t$. Assuming same infectiousness for local cases and imported cases would have $R_t$ less than one all along since late February, which would miss the local outbreaks in mid-March (Figure 2A). Assuming all cases were local cases in
the estimation led to a peak in $R_t$ in early March, which was inconsistent with the low number of local cases in that period (Figure 2B).

Simulation studies suggested that our model could obtain unbiased estimate of $R_t$ for local cases and imported cases with constant $R_t$ (Figure S1), and with an intervention controlling for local cases only so that $R_t$ decreased from 2.0 to 0.8 (Figure S2).

**DISCUSSIONS**

In this study, we extended the current framework for estimation of $R_t$ to account for differential transmission from imported cases. In this framework, we separately estimated the $R_t$ for local cases and imported cases, so that when control measures were targeting the imported cases, our framework would be feasible to allow $R_t$ for imported cases to be lower than for local cases.

We compared the extended framework with the current two approaches, namely assuming all cases were local cases, or assuming equal infectiousness for imported cases and local cases [5]. These two approaches would have little impact when local transmission was substantial, but would bias the estimated $R_t$ in measuring transmissibility in the community, when the number of local cases and imported cases were comparable. In particular, assuming all cases were local cases would overestimate the number of local transmissions as imported cases could not be infected locally. This was improved in the framework assuming equal infectiousness of local and imported cases [5]. However, this would be inappropriate if there were targeted control measures for inbound travelers.
Given that pre-symptomatic transmission was substantial for COVID-19, using
serial intervals as a proxy of infectiousness profile would be inappropriate [4,
15]. Hence, we used a deconvolution approach to infer the epidemic curve by
infection date from the observed epidemic curve that was by confirmation date,
and using the infectiousness profile since infection [16]. Misspecification of the
infectiousness profile since infection, or using a wrong approach to obtain the
epidemic curve by infection dates such as back-shifting based on the delay
distribution would bias the estimated $R_t$ [8, 17]. We also properly accounted for
the uncertainty in each step by using a bootstrap approach [18].

Our study has some limitations. First, we did not account for incomplete
observations. Mild or asymptomatic cases would likely be undetected. Also,
testing availability and criteria would also be likely changing over time. If the
proportion of undetected cases were constant over the outbreak, the $R_t$ would
still be unbiased [17]. Further investigation would be necessary to develop
method to account for the changing proportion of undetected cases. Finally,
rapid changes in the parameters could affect the estimated $R_t$, such as the sudden
change for the delay distribution from infection to confirmation caused by
overwhelmed healthcare systems, changing case definition during the epidemic
[19], or superspreading events during the outbreaks [20].

In conclusion, we developed methodology to estimate separately the $R_t$ for local
and imported cases, to account for the potential differential infectiousness
between them, due to control measures targeting inbound travelers. We
compared the developed framework with the two current approaches to account
for imported cases and illustrate the importance of properly accounting for the
differential infectiousness in the estimation of $R_t$. Accurate estimation of $R_t$
allows situational awareness of transmission and control in the community.
REFERENCES


FIGURE LEGENDS

Figure 1. Epidemic curve of the COVID-19 transmission in Hong Kong from January through May 2020 (Panel A) and the estimated time-varying reproductive number for local cases (Panel B) and imported cases (Panel C).

Figure 2. Comparison of estimation of time-varying reproductive number using our proposed framework, with alternative approaches that assuming equal infectiousness for local cases and imported cases (Panel A), or assuming all cases were local cases in the analysis (Panel B).
A

Epidemic curve

14–day mandatory quarantine for arriving persons

Civil servants special work arrangements

Additional social distancing measures

B

Time varying reproductive number for local cases

Civil servants special work arrangements

Additional social distancing measures

C

Time varying reproductive number for imported cases

14–day mandatory quarantine for arriving persons
Assuming equal infectiousness for imported and local cases

A

Assuming all cases were local cases

B

Civil servants special work arrangements

Additional social distancing measures

Estimates based on our proposed method

Estimates assumed equal infectiousness

Estimates mixed local or imported cases

Effective reproductive number

Date

January

February

March

April

May

January

February

March

April

May