

1 **Relative role of border restrictions, case finding and contact tracing in controlling SARS-**
2 **CoV-2 in the presence of undetected transmission**

3 Rachael Pung, Hannah E. Clapham, Vernon J. Lee, Adam J Kucharski on behalf of the CMMID
4 COVID-19 working group

5
6 Corresponding Author:

7 Rachael Pung

8 Communicable Disease Division

9 Ministry of Health, Singapore

10 12 College Road, Singapore 169852

11 Email: Rachael_PUNG@moh.gov.sg

12 Abstract

13 Background

14 Several countries have controlled the spread of COVID-19 through varying combinations of
15 border restrictions, case finding, contact tracing and careful calibration on the resumption of
16 domestic activities. However, evaluating the effectiveness of these measures based on observed
17 cases alone is challenging as it does not reflect the transmission dynamics of missed infections.

18
19 Methods

20 Combining data on notified local COVID-19 cases with known and unknown sources of infections
21 (i.e. linked and unlinked cases) in Singapore in 2020 with a transmission model, we reconstructed
22 the incidence of missed infections and estimated the relative effectiveness of different types of
23 outbreak control. We also examined implications for estimation of key real-time metrics — the
24 reproduction number and ratio of unlinked to linked cases, using observed data only as compared
25 to accounting for missed infections.

26
27 Findings

28 Prior to the partial lockdown in Singapore, initiated in April 2020, we estimated 89% (95%CI 75–
29 99%) of the infections caused by notified cases were contact traced, but only 12.5% (95%CI 2–
30 69%) of the infections caused by missed infectors were identified. We estimated that the
31 reproduction number was 1.23 (95%CI 0.98–1.54) after accounting for missed infections but was
32 0.90 (95%CI 0.79–1.1) based on notified cases alone. At the height of the outbreak, the ratio of
33 missed to notified infections was 34.1 (95%CI 26.0–46.6) but the ratio of unlinked to linked
34 infections was 0.81 (95%CI 0.59–1.36). Our results suggest that when case finding and contact
35 tracing identifies at least 50% and 20% of the infections caused by missed and notified cases
36 respectively, the reproduction number could be reduced by more than 14%, rising to 20% when
37 contact tracing is 80% effective.

38
39 Interpretation

40 Depending on the relative effectiveness of border restrictions, case finding and contact tracing,
41 unobserved outbreak dynamics can vary greatly. Commonly used metrics to evaluate outbreak

1 control — typically based on notified data — could therefore misrepresent the true underlying
2 outbreak.

3 4 Funding

5
6 Ministry of Health, Singapore.

7 8 **Research in context**

9 10 **Evidence before this study**

11
12 We searched PubMed, BioRxiv and MedRxiv for articles published in English up to Mar 20, 2021
13 using the terms: (2019-nCoV OR “novel coronavirus” OR COVID-19 OR SARS-CoV-2) AND
14 (border OR travel OR restrict* OR import*) AND (“case finding” OR surveillance OR test*) AND
15 (contact trac*) AND (model*). The majority of modelling studies evaluated the effectiveness of
16 various combinations of interventions in the absence of outbreak data. For studies that
17 reconstructed the initial spread of COVID-19 with outbreak data, they further simulated
18 counterfactual scenarios in the presence or absence of these interventions to quantify the impact
19 to the outbreak trajectory. None of the studies disentangled the effects of case finding, contact
20 tracing, introduction of imported cases and the reproduction number, in order to reproduce an
21 observed SARS-CoV-2 outbreak trajectory.

22 23 **Added value of this study**

24
25 Notified COVID-19 cases with unknown and known sources of infection are identified through
26 case finding and contact tracing respectively. Their respective daily incidence and the growth rate
27 over time may differ. By capitalising on these differences in the outbreak data and the use of a
28 mathematical model, we could identify the key drivers behind the growth and decline of both
29 notified and missed COVID-19 infections in different time periods — e.g. domestic transmission
30 vs external introductions, relative role of case finding and contact tracing in domestic
31 transmission. Estimating the incidence of missed cases also allows us to evaluate the usefulness
32 of common surveillance metrics that rely on observed cases.

33 34 **Implications of all the available evidence**

35
36 Comprehensive outbreak investigation data integrated with mathematical modelling helps to
37 quantify the strengths and weaknesses of each outbreak control intervention during different
38 stages of the pandemic. This would allow countries to better allocate limited resources to
39 strengthen outbreak control. Furthermore, the data and modelling approach allows us to estimate
40 the extent of missed infections in the absence of population wide seroprevalence surveys. This
41 allows us to compare the growth dynamics of notified and missed infections as reliance on the
42 observed data alone may create the illusion of a controlled outbreak.

43
44

1 **Background**

2 The COVID-19 pandemic has resulted in substantial disruption to international travel and trade
3 due to widespread border restrictions that have been enacted by countries as a key strategy to
4 reduce the importation and spread of COVID-19.¹ In addition to border controls, which do not
5 totally prevent the importation of cases,² case finding and contact tracing have formed a central
6 part of the response in many countries. Case finding has helped in early identification and isolation
7 of new infections that are not associated with other known cases through testing of suspected
8 cases and surveillance in target groups.^{3,4} Meanwhile, contact tracing identifies potential
9 transmission routes and new infections among contacts of a known case.^{3,4} Local COVID-19
10 cases with known and unknown sources of infection can therefore be categorised as 'linked' or
11 'unlinked' respectively.^{3,5}

12
13 The occurrence of unlinked COVID-19 cases implies that the pandemic is partly unobserved. This
14 could be attributed to the importation and transmission from asymptomatic or mildly symptomatic
15 infections who do not require medical attention⁶ and underreporting of symptomatic cases⁷.
16 Furthermore, failure to trace secondary cases arising from a notified case would also create gaps
17 in the observed transmission chains. The respective case counts and ratio of unlinked to linked
18 cases are often used as a metric for the effectiveness of outbreak control, and are closely
19 monitored in many countries to assess the potential for resuming social and economic activities
20 and the lifting of border restrictions.⁸⁻¹⁰ However, methods to establish the relative role of border
21 restrictions, case finding, contact tracing to the trajectory of linked and unlinked cases, remain
22 elusive.

23
24 In addition, the pandemic trajectory is often measured by the effective reproduction number which
25 relies on notified cases or cases extrapolated from observed deaths.¹¹⁻¹³ With effective case
26 isolation and contact tracing, the time spent in the community while infectious for a notified case
27 is truncated, shortening the observed serial interval (a proxy for generation interval).¹⁴
28 Consequently, the observed chains of transmission are short-lived and, on average, less than
29 one secondary case is generated.¹⁵ As such, it is currently unclear whether metrics based on
30 observed features such as notified cases and their linkage give an accurate picture on the
31 underlying outbreak dynamics.

32
33 As countries progressively reopen, it is important for policy makers to understand the parameters
34 that contributes to the spread of cases and the growth dynamics in undetected cases. Combining
35 the daily incidence of imported, and local linked and unlinked COVID-19 cases in Singapore with
36 a mathematical model, we aim to characterise the effectiveness of detection and control
37 measures over the course of the pandemic and estimate the incidence of missed cases, even
38 when longitudinal serology surveys are absent. Furthermore, we compared the growth patterns
39 in missed and notified cases and investigated the implications of making inferences on the
40 pandemic trajectory based on observed data alone.

1 **Methods**

2 Data

3 Cases of COVID-19 (confirmed with a respiratory sample positive for SARS-CoV-2 on PCR)
4 notified to the Ministry of Health, Singapore from Jan 23 to Dec 31, 2020 in Singapore were used.
5 Extensive epidemiological investigations are conducted for each case to establish their exposure
6 history prior to notification.^{16–19} A local linked case was a case with at least one known source of
7 infection while a local unlinked case was a case with an unknown source of infection. Case finding
8 measures (i.e. active surveillance) identify local unlinked cases through testing of persons with
9 acute respiratory infection or pneumonia, routine testing of targeted groups at high-risk of
10 acquiring or transmitting disease, and ad-hoc testing of sub-populations of interest.³

11 Imported cases are confirmed COVID-19 cases with travel history to a country with ongoing
12 COVID-19 outbreak in the preceding 14 days. They were stratified into two main categories, those
13 quarantined in dedicated facilities upon arrival, and those undergoing home-based quarantine or
14 who were not quarantined at all, prior to detection.²⁰ The former were tested at the start and end
15 of their quarantine period and were assumed to be incapable of introducing infections into the
16 community while the latter, to whom the community were potentially exposed to, could generate
17 local infections.

18 All confirmed cases were conveyed to secured isolation facilities and discharged after 21 days
19 from date of confirmation if assessed to be clinically well, or with sequential negative tests. Cases
20 occurring in persons residing in a foreign-worker dormitory and notified from Apr 7 to Oct 31, 2020
21 were omitted from analysis as these dormitories were placed under lockdown for an extended
22 period of time. As workers were subjected to restricted movements, the opportunity to interact
23 with the community during this period was minimal and hence they were assumed to be incapable
24 of driving community-level transmission. Furthermore, about 0.2% of the confirmed cases
25 occurred in persons providing care to confirmed cases and as these secondary infections were
26 not community-acquired infections, they were omitted from analysis.

27 Transmission Model

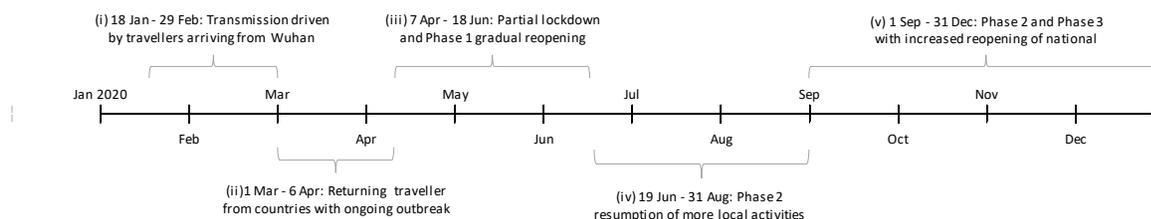
28 We simulated disease transmission through a branching process to compute the expected
29 incidence over time. The model was fitted using a Poisson likelihood to the daily incidence of
30 linked and unlinked local cases to reconstruct the incidence of missed infections (see
31 Supplementary Information for details).

32 Infections were introduced into the population by either notified imported cases who potentially
33 had contact with community individuals (referred to as notified imported cases with community
34 contact in the remaining text) or through missed imported infections. We modelled the number of
35 missed imported infections relative to the number of notified imported cases with community
36 contact using a factor ρ . Both types of imported cases could generate community infections from
37 the time of arrival to isolation or end of their infectiousness respectively. Community infections
38 were identified through varying effectiveness of contract tracing of notified cases (i.e. probability
39 of detecting linked cases, ϵ_{link}) or case finding (i.e. probability of detecting unlinked cases, ϵ_{unlink}).

1 The potential reproduction number of an infected individual, R , was defined as the average
2 number of secondary cases generated by a single infectious individual over the entire infectious
3 period in the absence of quarantine/isolation. This is analogous to the reproduction number of a
4 missed infected individual, R_{missed} . For a notified case, the amount of time spent in the community
5 while infectious is generally shorter as compared to a missed infection, either as they sought
6 medical attention and were isolated, or when a secondary case was identified through contact
7 tracing and quarantined before being tested positive. The reproduction number of a notified case,
8 R_{notified} , was defined as the average number of secondary cases generated by a single infectious
9 individual till the time of quarantine/isolation. Notified cases were assumed to be incapable of
10 generating offspring infections once isolated. Overall, the effective reproduction number, R_{eff} , is
11 an aggregate measure of the reproduction number of both missed and notified cases.

12 We assumed the generation time was gamma distributed with mean 7.5 days (SD 3.4).²¹ The
13 distribution was left-truncated for all imported cases to exclude the infectious period while
14 overseas and right-truncated for all notified cases due to early case isolation.

15 Due to varying effectiveness of detection and control measures, and changing social behaviour
16 over the course of the pandemic, we defined five different time periods with a different value
17 estimated for the four unknown parameters (ρ , ϵ_{link} , ϵ_{unlink} , R) for each period – when transmission
18 was mainly driven by (i) travellers from China, and (ii) travellers from other countries with ongoing
19 outbreak; the implementation of (iii) partial lockdown within Singapore; (iv) resumption of local
20 activities, and (v) increased reopening of national borders. Results for ρ were converted to the
21 average daily number of missed imported cases for each time period to provide an estimate of
22 the magnitude of missed imported infections and we report the posterior median with 95% credible
23 intervals (CI) for all model outputs.



24
25 Figure 1 Time periods considered in transmission model

26 Independent model validation

27 We validated the model outcomes against two independent population level surveys involving (i)
28 41,852 preschool staff who underwent a PCR test between May 15 and 29, 2020 followed by
29 serological testing for confirmed cases and (ii) 1,578 participants of a cross-sectional
30 seroprevalence survey conducted from Oct 7 to 31, 2020 with participants randomly selected from
31 the general population.²² Incorporating PCR detection and seroconversion probabilities,^{23,24} we
32 estimated PCR and serology positive cases from the model to compare against the incidence rate
33 in preschool staff, and seroconverted cases from the model to compare against the
34 seroprevalence rate in the general population (Supplementary Information). Observed data were
35 presented as the mean and the 95% confidence intervals for binomial proportions were computed

1 using Wilson's method.²⁵ We performed a z-test to evaluate the difference between the observed
2 and modelled rates and p values <0.05 were considered statistically significant.

3

4 Modelling varying effectiveness of detecting linked and unlinked cases and impact to epidemic 5 growth dynamics

6 We further explored the epidemic growth dynamics in the presence of missed and notified SARS-
7 CoV-2 infections via the use of a next-generation matrix mathematical framework (Supplementary
8 Information). In general, a notified case spends a reduced amount of time spent in the community
9 while infectious as compared to a missed infection. This creates a heterogeneity in the
10 reproduction number of a missed and notified case. As such, we studied the impact of the R_{eff}
11 across a range of values for ϵ_{link} , ϵ_{unlink} and R . We assumed a Weibull distributed time of infection
12 to isolation with mean 9.2 days (SD 4.4) for notified cases (derived from observed data in
13 symptomatic cases notified from Mar 1 to Apr 6, 2020, Supplementary Figure 1 and Table 1).

14 The ratio of missed to notified cases established during exponential growth was also modelled to
15 characterise the extent of case ascertainment and compared against the ratio of unlinked to linked
16 infections. Lastly, we analysed how different ratios of missed to notified imported cases with
17 community contact influences the time taken to achieve exponential growth and implications of
18 inferring outbreak trajectory from transient growth patterns.

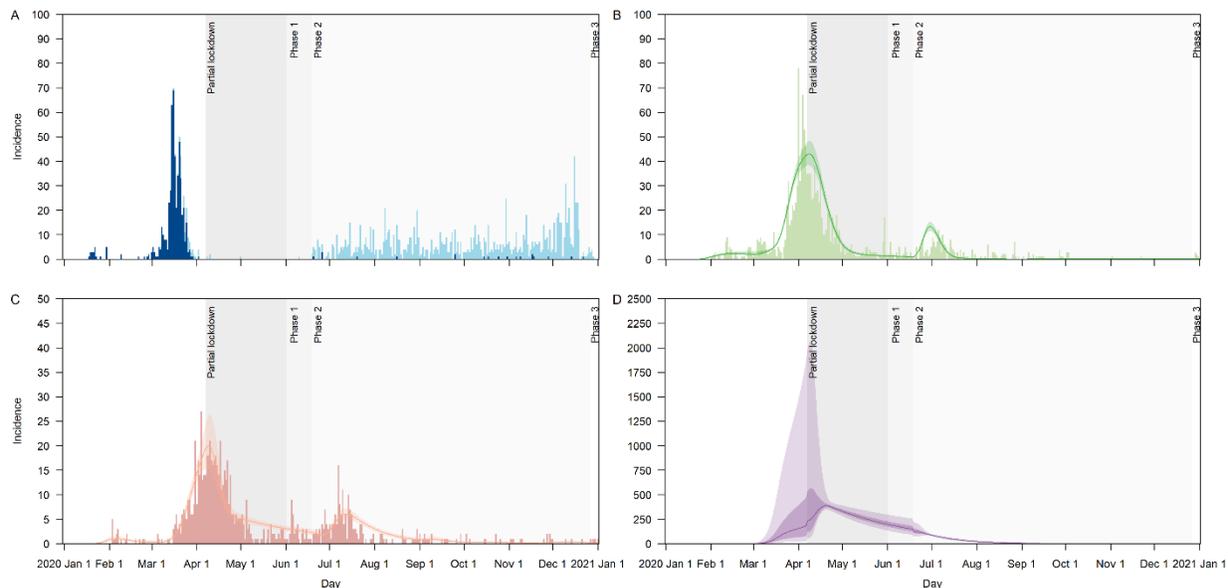
19 All data and code required to reproduce the analysis is available online.²⁶

20

21

1 Results

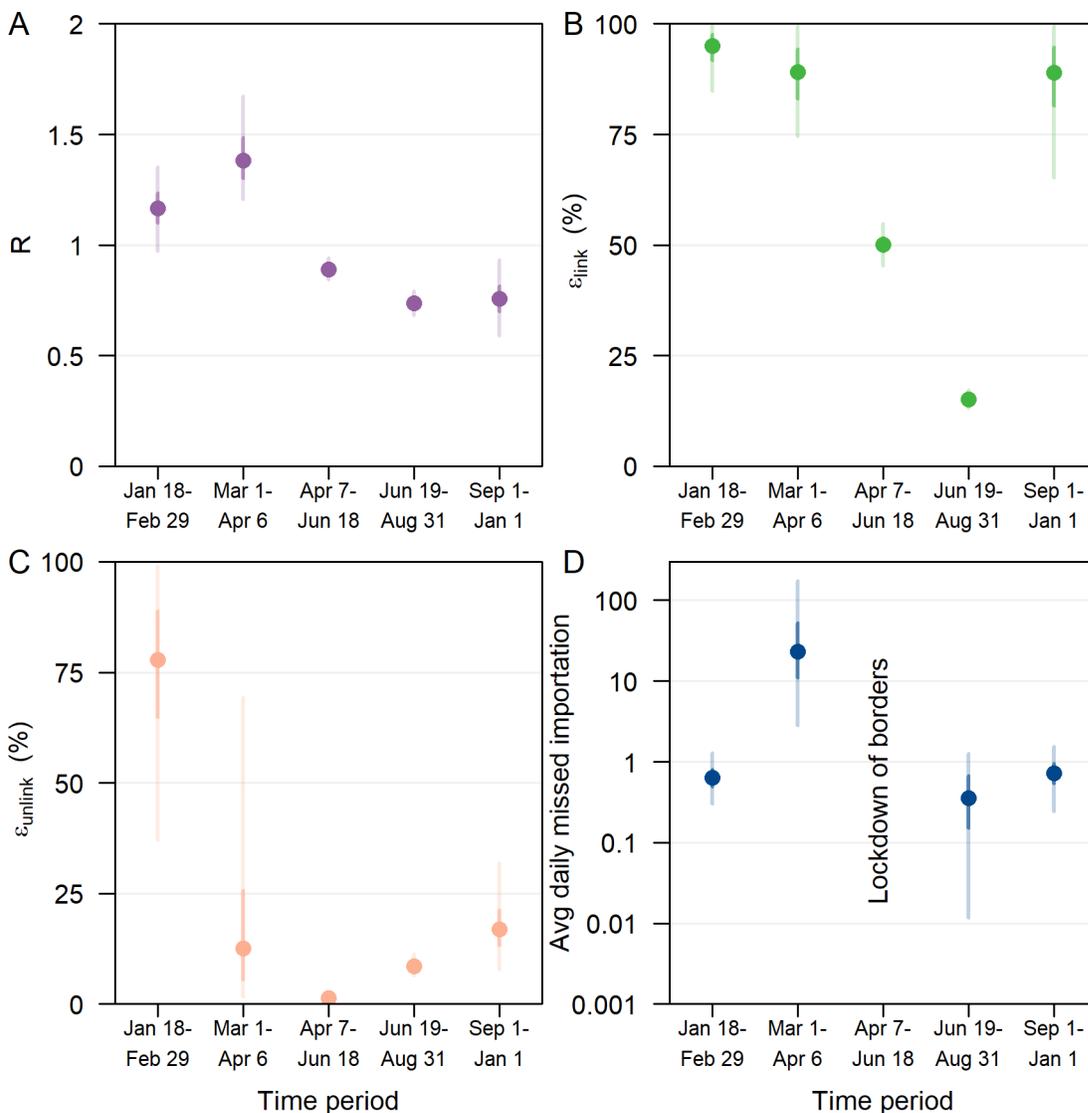
2 In Singapore, initial border and outbreak control measures from Jan 18 to Feb 29, 2020 aimed to
3 reduce the spread of SARS-CoV-2 by infected persons arriving from China. These measures
4 were progressively expanded from Mar 1, 2020 in response to a surge of imported cases returning
5 from countries with ongoing outbreaks (figure 2a). Despite the decline in notified imported cases
6 from Mar 16 to Apr 1, 2020, community transmissions were rising (figure 2b and c). By the time a
7 nationwide partial lockdown was implemented on Apr 7, 2020, 810 community infections had been
8 notified. However, our model estimated the majority of the infections had gone undetected, with
9 2,940 (95% CI 414–26,500) missed (figure 2d). During the partial lockdown period and Phase 1
10 of reopening (Apr 7–Jun 18, 2020), the estimated daily number of missed cases decreased but
11 remained above 100 and cumulatively, 20,000 infections (95% CI 15,200–34,400) were missed
12 (figure 2d). As border restrictions were gradually lifted and economic and social activities
13 resumed, the number of imported cases with community contact were kept low and there was no
14 evidence of sustained community transmission (figure 2a–c). Overall, we estimated that 26,700
15 cases (95% CI 19,400–64,000) were missed in 2020 (figure 2d).



16

17 Figure 2 Daily incidence of COVID-19 cases in Singapore, (A) notified imported cases with
18 community contact (dark blue bars) and notified imported cases who were quarantined upon
19 arrival (light blue bars), (B) notified local linked cases (green bars) and the posterior median
20 (green line) and 95% CI (green shaded area), (C) notified local unlinked cases (pink bars) and
21 the posterior median (pink line) and 95% CI (pink shaded area), (D) modelled posterior median
22 local missed infections (line), 50% CI (dark purple shaded area) and 95% CI (light purple shaded
23 area).

1 We estimated that R was 1.17 (95% CI 0.97–1.35) at the start of the pandemic. From Mar 1 to
 2 Apr 6, 2020, this increased to 1.38 (95% CI 1.21–1.67) prior to the partial lockdown (figure 3a).
 3 However, R_{notified} , was lower at 0.90 (95% CI 0.79–1.1) due to the reduced amount of time spent
 4 in the community while infectious compared to a missed infection (supplementary figure 5). After
 5 accounting for missed infections, the R_{eff} was 1.23 (95% CI 0.98–1.54) (supplementary figure 5).
 6 While the country’s contact tracing system was able to detect 89% of the secondary cases (ϵ_{link} ,
 7 95% CI 75–99%) arising from a notified case, only 12.5% of the infections (ϵ_{unlink} , 95%CI 2-69%)
 8 caused by a missed infected individual were identified through case finding — a sharp decline
 9 from 78% (ϵ_{unlink} , 95% CI 37–99%) as estimated from Jan 18 to Feb 29 (figure 3b and c). We
 10 observed a peak of 70 imported cases per day being isolated and we estimated that 23 imported
 11 cases (95% CI 3–175) were missed daily (figure 3d).



12
 13 Figure 3 Posterior median (dot), 50% CI (dark vertical lines) and 95% CI (light vertical lines) of
 14 model parameters. (A) potential reproduction number, R , of a COVID-19 case, (B) effectiveness

1 of detecting a linked case, $\varepsilon_{\text{link}}$, (C) effectiveness of detecting an unlinked case, $\varepsilon_{\text{unlink}}$, (D) average
2 daily number of missed imported cases in log scale.

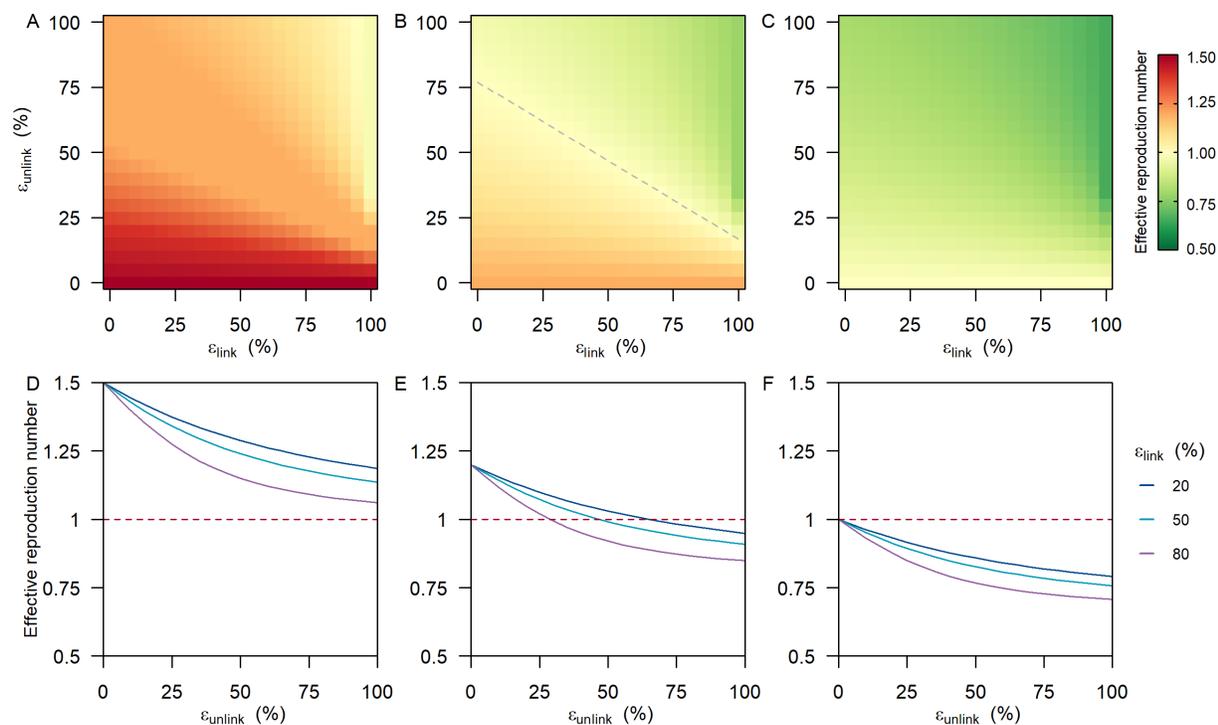
3 During the partial lockdown, although we estimated R to be below 1, which signalled a controlled
4 outbreak, the effectiveness of detecting linked and unlinked cases was low at 50% (95%CI 45–
5 55%) and 1% (95%CI 1–2%) respectively (figure 3b and c). As such, it took about two months to
6 reach a daily observed incidence of less than 10 (figure 2b and c). As social and economic
7 activities were progressively resumed from Jun 19, 2020 onwards, the effectiveness of detecting
8 linked and unlinked cases remained low at 15% (95% CI 13–17%) and 8% (95% CI 6–11%)
9 respectively. However, with strict quarantine of incoming travellers and continued enforcement of
10 outbreak control measures, the average daily number of missed imported infections was low at
11 0.35 (95% CI 0.01–1.27) (figure 3c and d) and R was approximately 0.74 (95% CI 0.68–0.79)
12 (figure 3a and d). By Sep 1, 2020 the contact tracing system's effectiveness in identifying
13 secondary cases from a notified case had recovered to 89% (95%CI 65–99%) (figure 3b).

14 Independent model validation

15
16 While the transmission model was able to reproduce the observed temporal trends, we sought to
17 further validate the model outputs against independent population level surveys. Firstly, of the
18 41,852 preschool staff tested between May 15 and 29, 2020, 13 staff were PCR and serology
19 positive for SARS-CoV-2 infection, which translates to an observed incidence rate of 311
20 infections per million population (95%CI 181–531) compared to our model estimate of 355 per
21 million population (95%CI 302–421). Secondly, in a cross-sectional seroprevalence survey from
22 Oct 7 to 31, 2020, SARS-CoV-2 antibodies were detected in 7 out of 1,578 participants.²⁰ The
23 observed seroprevalence was 0.44% (95%CI 0.22–0.91%) compared to our model estimate of
24 0.43% (95%CI 0.31–1.03%). The difference between the observed and modelled rates in both
25 surveys were statistically insignificant ($p=0.63$ and $p=0.93$).
26

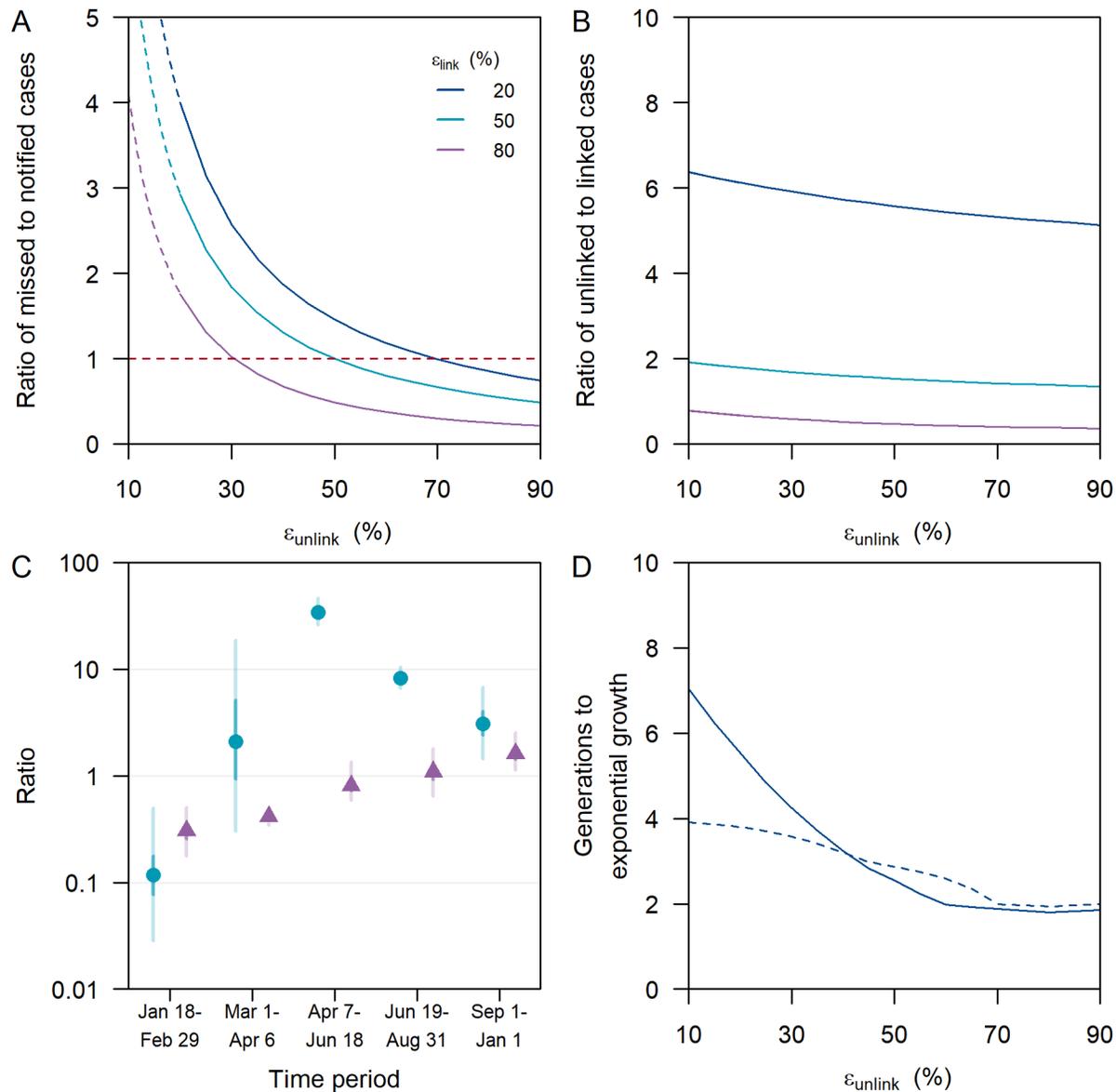
27 Varying effectiveness of detecting linked and unlinked cases and implications for epidemic growth

28 When the effectiveness of detecting linked cases is low ($\varepsilon_{\text{link}}=20\%$), an effectiveness of detecting
29 unlinked cases of at least 50% results in R_{eff} approximately 14.1% lower than R (figure 4). As the
30 testing capacity and effectiveness of case finding increases, with a further increase in the ability
31 to ring fence secondary infections arising from notified cases, $\varepsilon_{\text{link}}$, of up to 80%, this could further
32 reduce the R_{eff} to at least 23.3% lower than R (figure 4). As a country's case finding and contact
33 tracing system strengthens and the time from infection to isolation is reduced, the reduction in the
34 R_{eff} relative to R is expected to increase (supplementary figure 6 and 7).



1
 2 Figure 4 Effect of detecting linked and unlinked cases on the effective reproduction number, R_{eff} .
 3 For a potential reproduction number, R of 1.5 (A, D), 1.2 (B, E) and 1.0 (C, F) and a fixed
 4 distribution of time from infection to isolation (Weibull distributed with mean 9.2 days (SD 4.4)),
 5 increasing levels of ϵ_{link} and ϵ_{unlink} creates a heterogenous pool of missed and notified infections
 6 with the latter being subjected to early case isolation and hence lowers the R_{eff} . (B) Grey dashed
 7 line represents R_{eff} of 1.

8
 9
 10 To characterise the extent of case ascertainment, we estimate the ratio of missed to notified
 11 cases. When ϵ_{link} is more than 80%, this ratio remains below 1 if ϵ_{unlink} is greater than 30%. In
 12 other words, aggressive ring fencing of contacts whenever a confirmed case is detected, helps to
 13 ensure that the underlying outbreak is mostly observed (figure 5a). For the same values of ϵ_{link}
 14 and ϵ_{unlink} , the ratio of missed to notified cases can be vastly different from the ratio of unlinked to
 15 linked cases — typically used to characterise the extent of outbreak control. When ϵ_{unlink} is low,
 16 the epidemic becomes increasingly obscure and the ratio of missed to notified cases grows very
 17 large but the ratio of the unlinked to linked cases tends to a fixed value dependent on ϵ_{link} (figure
 18 5a and b).



1
 2 Figure 5 Comparison of unobserved outbreak dynamics with common surveillance metrics. (A)
 3 Ratio of missed to notified cases for varying ϵ_{link} and ϵ_{unlink} . Dashed lines implies that the ratio
 4 extends to infinity as ϵ_{unlink} tends to zero; (B) Ratio of unlinked to linked cases for varying ϵ_{link} and
 5 ϵ_{unlink} ; (C) Ratio of missed to notified cases (turquoise dots) and ratio of unlinked to linked cases
 6 (purple triangle) in different time periods of the pandemic. Posterior median (dot/triangle), 50% CI
 7 (dark vertical lines) and 95% CI (light vertical lines). (D) Generations to exponential growth for an
 8 outbreak with 5 missed imported cases and 95 notified imported cases with community contact,
 9 R of 1.5, ϵ_{link} of 80%, for varying ϵ_{unlink} .

10
 11 To put these results into the Singapore context, as the effectiveness of detecting unlinked cases
 12 declined in March during the surge of imported cases and further declined during the partial
 13 lockdown, the ratio of missed to notified cases increased to 34.1 (95%CI 26.0–46.6) and was

1 many times higher than the ratio of unlinked to linked cases of 0.81 (95%CI 0.59–1.36) (figure
2 5c). As such, metrics derived from observed data alone do not always accurately reflect the
3 underlying outbreak.

4
5 Under certain conditions, the time to achieve exponential growth in notified and missed infections
6 could differ. When ϵ_{unlink} is low and when the initial ratio of missed to notified imported cases with
7 community contact is low (5 missed:95 notified), the number of generations required to achieve
8 exponential growth in notified cases is greater than that in missed cases (figure 5d). This time to
9 exponential growth is reduced when either the ϵ_{unlink} is high or when the initial ratio of missed to
10 notified imported cases with community contact increases (figure 5d and supplementary figure 8).

11

12 Discussion

13 Using the growth patterns in the daily incidence of local linked and unlinked cases, and imported
14 cases with community contact, our model was capable of disentangling the effects of case finding
15 and contact tracing (ϵ_{unlink} and ϵ_{link}) from other outbreak interventions that affect the potential
16 reproduction number of a case (R), at a time before vaccination roll out. In spite of a strong
17 capability to contact trace, without a tight control on the number of imported cases coupled with
18 low ability to detect new cases and R_{eff} exceeding 1, community transmission was sustained, as
19 witnessed in Singapore's daily incidence of COVID-19 cases from Mar 1 to Apr 6, 2020. This
20 surge in community cases affected the ability of the contact tracing system to ring fence notified
21 cases in the following months (figure 3b) but the partial lockdown with strong enforcement of non-
22 pharmaceutical interventions such as mask wearing, physical distancing, and movement
23 restrictions helped to reverse the pandemic trajectory (figure 2 and 3).

24

25 As countries progressively resume economic and social activities in partially vaccinated
26 populations, the potential reproduction number of an infectious individual engaged in these
27 activities may exceed unity. Case finding and contact tracing help in the early identification and
28 isolation of (secondary) cases thereby minimising the duration of infectious period spent in the
29 community. Even if contact tracing capacity is low, if more than half of the infections arising from
30 a previously undetected infection present to the healthcare system for early testing and isolation,
31 the R of a case could be reduced by more than 14% (figure 4). Increasing the effectiveness of
32 case finding (ϵ_{unlink}) implies casting a wider surveillance net but ultimately this measure depends
33 on compliance with testing regimes. We estimated a sharp drop in ϵ_{unlink} during the partial
34 lockdown period (figure 3c) and this behaviour was corroborated by behavioural surveys
35 documenting diminished health-seeking behaviour.²⁷ Factors driving avoidance of testing
36 warrants further studies as, no matter how many infections can be detected from contact tracing,
37 the healthcare system relies on the testing and identification of cases in the first instance.

38

39 As contact tracing devices are progressively rolled out to speed up contact tracing and coupled
40 with effective quarantine and testing of close contacts, for a given ϵ_{unlink} of 50% and ϵ_{link} of 80%,
41 R is lowered by more than 20% (figure 4). Even if a healthcare system is only able to detect 30–
42 50% of the infections arising from previously undetected infections (ϵ_{unlink}), a high ϵ_{link} ensures

1 aggressive ring fencing of secondary infections arising from an unlinked case and keeps the
2 outbreak in check (figure 5a).

3
4 Current methods to estimate the effective reproduction number, R_{eff} , of SARS-CoV-2 rely on the
5 notified case incidence or deaths, some with appropriate adjustments to capture the right-
6 censoring of data.^{11–13,15} However, in the presence of asymptomatic infections⁶ and
7 underreporting of symptomatic cases⁷, these methods neglect the growth dynamics of a non-
8 negligible number of missed infections which can result in misleading inferences. From the model,
9 we estimated that R was about 1.38 from Mar 1 to Apr 6, 2020 (figure 3a). With an effective
10 contact tracing system that quarantines close contacts and detects close to 90% of the secondary
11 cases from notified cases, this results in reduced time from infection to isolation and the
12 reproduction number of a notified case was 0.90, similar to previous modelling estimates.¹⁵
13 Collectively, the R_{eff} was approximately 1.2, which exceeds 1 and signalled sustained
14 transmissions (supplementary figure 5).

15
16 By reconstructing the daily incidence of missed infections, we could infer the time-varying level of
17 case under-ascertainment. We estimated that nearly 90% of the missed infections occurred
18 between Mar 1 and Jun 18, 2020 and the ratio of missed to notified infections was more than 30
19 (figure 5c). This overall level of case under-ascertainment was comparable to the estimated
20 under-ascertainment in symptomatic cases alone in other countries during similar phases of the
21 pandemic and the under-ascertainment levels in many countries are expected to be higher when
22 accounting for asymptomatic infections.⁷ Furthermore, we showed that using the ratio of unlinked
23 to linked cases — calculated from observed data alone, does not adequately characterise the
24 extent of outbreak control. The ratio of unlinked to linked cases showed little variation over the
25 course of the pandemic as compared to the ratio of missed to notified infections and does not
26 provide warning of a runaway outbreak.

27
28 With the reopening of borders, missed and notified imported cases with community contact could
29 initiate the first generation of local infections and if there is a low ratio of missed to notified
30 imported case of 1:19 coupled with a low ability to detect unlinked cases ($\epsilon_{\text{unlink}} = 10\%$), it takes 3
31 more generations for the notified cases to achieve exponential growth as compared to the missed
32 cases (figure 5d). The deviation from exponential growth in notified cases at the early stages of
33 the outbreak creates an illusion that outbreak can be controlled based on observed data, while
34 the number of undetected infections continues to escalate. One effective but resource intensive
35 method of minimising the number of missed imported infections is to implement strict quarantine
36 and testing of incoming travellers, especially those arriving from countries with high COVID-19
37 incidence. This reduces the overall number of new infections introduced into the population and
38 is especially important for countries that have managed to stabilise their epidemic numbers after
39 the initial wave(s) of infections.

40
41 There are some limitations to our study. Firstly, the model assumes that each of the four
42 parameters remains constant in a specified time period. As such, we are unable to provide a time-
43 varying measure to characterise the impact of different outbreak detection and control measures

1 that were progressively rolled out in the population at a granular level. Instead, time periods were
2 chosen based on prior knowledge of major policies that would affect at least one of the four model
3 parameters. Secondly, those imported cases subjected to home quarantine and no quarantine
4 were assumed to have the same potential of making contact with members of the community as
5 household transmission might occur, while in reality the amount of community contact would be
6 different as those under home quarantine could potentially only contact household members. In
7 the absence of data, we made a conservative assumption that the imported cases under home
8 quarantine were capable of generating local infections. Further model calibration would require
9 data on the outcome of the various quarantine measures. Finally, missed infections could arise
10 from asymptomatic or mildly symptomatic infections, or underreporting of symptomatic cases. We
11 assumed that R is the same among these cases. More information is needed to determine the
12 temporal variation in the types of cases to account for lowered transmission potential in
13 asymptomatic or mildly symptomatic cases. At the same time, a key strength to our analysis is
14 that our model was able to reproduce independent observations in two separate population level
15 surveys and this lends support to our assumption of a homogeneous R among all missed
16 infections.

17
18 The SARS-CoV-2 pandemic has generated new forms of data collection and many new ways to
19 reconstruct outbreak dynamics and evaluate the extent of missed infections arising from high
20 asymptomatic rates and underreporting of cases. The daily incidence of linked and unlinked cases
21 could help countries evaluate their performance in case finding, contact tracing and the
22 effectiveness of their border restrictions. Missed and notified infections bring about a
23 heterogeneity in the reproduction number and the mixture of these factors can create an illusion
24 of a controlled outbreak. As countries progressively reopen borders or plan for pandemics in the
25 future, it is important to have an integrated surveillance and modelling analysis system to
26 overcome the challenges of undetected transmissions.

27 28 **References**

- 29
30 1 World Tourism Organization. COVID-19 related travel restrictions. a global review for tourism.
31 Second report. 2020; published online April 28. [https://www.unwto.org/news/covid-19-travel-](https://www.unwto.org/news/covid-19-travel-restrictions)
32 [restrictions](https://www.unwto.org/news/covid-19-travel-restrictions) (accessed March 25, 2021).
- 33 2 Dickens BL, Koo JR, Lim JT, *et al.* Strategies at points of entry to reduce importation risk of
34 COVID-19 cases and reopen travel. *Journal of Travel Medicine* 2020; **27**.
35 DOI:10.1093/jtm/taaa141.
- 36 3 Ng Y, Li Z, Chua YX, *et al.* Evaluation of the Effectiveness of Surveillance and Containment
37 Measures for the First 100 Patients with COVID-19 in Singapore — January 2–February 29,
38 2020. *MMWR Morb Mortal Wkly Rep* 2020; **69**. DOI:10.15585/mmwr.mm6911e1.
- 39 4 Kucharski AJ, Klepac P, Conlan AJK, *et al.* Effectiveness of isolation, testing, contact tracing,
40 and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a
41 mathematical modelling study. *The Lancet Infectious Diseases* 2020; **20**: 1151–60.

- 1 5 Lam HY, Lam TS, Wong CH, *et al.* The epidemiology of COVID-19 cases and the successful
2 containment strategy in Hong Kong—January to May 2020. *Int J Infect Dis* 2020; **98**: 51–8.
- 3 6 Yanes-Lane M, Winters N, Fregonese F, *et al.* Proportion of asymptomatic infection among
4 COVID-19 positive persons and their transmission potential: A systematic review and meta-
5 analysis. *PLOS ONE* 2020; **15**: e0241536.
- 6 7 Russell TW, Golding N, Hellewell J, *et al.* Reconstructing the early global dynamics of under-
7 ascertained COVID-19 cases and infections. *BMC Medicine* 2020; **18**: 332.
- 8 8 Channel News Asia. Singapore, Hong Kong to defer air travel bubble launch. CNA. 2020;
9 published online Dec 1. [https://www.channelnewsasia.com/news/singapore/covid-19-](https://www.channelnewsasia.com/news/singapore/covid-19-singapore-hong-kong-defer-travel-bubble-december-13675222)
10 [singapore-hong-kong-defer-travel-bubble-december-13675222](https://www.channelnewsasia.com/news/singapore/covid-19-singapore-hong-kong-defer-travel-bubble-december-13675222) (accessed March 14, 2021).
- 11 9 New Zealand Government. COVID-19 media conference — 7pm, 14 February 2021 | Unite
12 against COVID-19. 2021; published online Feb 14. [https://covid19.govt.nz/updates-and-](https://covid19.govt.nz/updates-and-resources/latest-updates/covid-19-media-conference-7pm-14-february-2021/)
13 [resources/latest-updates/covid-19-media-conference-7pm-14-february-2021/](https://covid19.govt.nz/updates-and-resources/latest-updates/covid-19-media-conference-7pm-14-february-2021/) (accessed
14 March 14, 2021).
- 15 10 Government of the Hong Kong Special Administrative Region. COVID-19 testing to intensify.
16 Hong Kong's Information Services Department. 2021; published online Feb 1.
17 http://www.news.gov.hk/eng/2021/02/20210201/20210201_172308_389.html (accessed
18 March 14, 2021).
- 19 11 Flaxman S, Mishra S, Gandy A, *et al.* Estimating the effects of non-pharmaceutical
20 interventions on COVID-19 in Europe. *Nature* 2020; **584**: 257–61.
- 21 12 Cori A, Ferguson NM, Fraser C, Cauchemez S. New Framework and Software to Estimate
22 Time-Varying Reproduction Numbers During Epidemics | American Journal of Epidemiology |
23 Oxford Academic. *Am J Epidemiol* 2013; **178**: 1505–12.
- 24 13 Abbott S, Hellewell J, Thompson RN, *et al.* Estimating the time-varying reproduction number
25 of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Res* 2020; **5**:
26 112.
- 27 14 Ali ST, Wang L, Lau EHY, *et al.* Serial interval of SARS-CoV-2 was shortened over time by
28 nonpharmaceutical interventions. *Science* 2020; **369**: 1106–9.
- 29 15 Pung R, Cook AR, Chiew CJ, *et al.* Effectiveness of Containment Measures Against COVID-
30 19 in Singapore: Implications for Other National Containment Efforts. *Epidemiology* 2021; **32**:
31 79–86.
- 32 16 Yong SEF, Anderson DE, Wei WE, *et al.* Connecting clusters of COVID-19: an
33 epidemiological and serological investigation. *The Lancet Infectious Diseases* 2020; **20**: 809–
34 15.
- 35 17 Channel News Asia. 'Like an invisible criminal': How police helped find missing link between
36 COVID-19 church clusters in a day - CNA. CNA. 2020; published online March 8.
37 [https://www.channelnewsasia.com/news/singapore/police-missing-link-church-clusters-](https://www.channelnewsasia.com/news/singapore/police-missing-link-church-clusters-covid19-coronavirus-12509492)
38 [covid19-coronavirus-12509492](https://www.channelnewsasia.com/news/singapore/police-missing-link-church-clusters-covid19-coronavirus-12509492) (accessed March 14, 2021).

- 1 18Channel News Asia. COVID-19: MOH conducting 'swab operations' as part of stepped up
2 efforts. CNA. 2020; published online July 24.
3 [https://www.channelnewsasia.com/news/singapore/covid-19-moh-swab-operations-testing-](https://www.channelnewsasia.com/news/singapore/covid-19-moh-swab-operations-testing-ringfence-12961220)
4 [ringfence-12961220](https://www.channelnewsasia.com/news/singapore/covid-19-moh-swab-operations-testing-ringfence-12961220) (accessed March 14, 2021).
- 5 19Channel News Asia. Behind the scenes of a COVID-19 contact tracer's work, tough cases
6 and how TraceTogether has helped. CNA. 2021; published online Jan 27.
7 [https://www.channelnewsasia.com/news/singapore/covid-19-contact-tracing-work-](https://www.channelnewsasia.com/news/singapore/covid-19-contact-tracing-work-tracetogogether-moh-14052632)
8 [tracetogogether-moh-14052632](https://www.channelnewsasia.com/news/singapore/covid-19-contact-tracing-work-tracetogogether-moh-14052632) (accessed March 14, 2021).
- 9 20Immigration & Checkpoints Authority, Singapore. Stay Home Notice (SHN).
10 <https://safetravel.ica.gov.sg/health/shn> (accessed April 15, 2021).
- 11 21Li Q, Guan X, Wu P, *et al.* Early Transmission Dynamics in Wuhan, China, of Novel
12 Coronavirus–Infected Pneumonia. *N Engl J Med* 2020; **382**: 1199–207.
- 13 22Chen MI-C, Lim VWX. Updates on the sero-epidemiology of SARS-CoV-2 in Singapore, and
14 reflections on the role of post-vaccine sero-surveillance [Unpublished data]. 2020.
- 15 23Borremans B, Gamble A, Prager K, *et al.* Quantifying antibody kinetics and RNA detection
16 during early-phase SARS-CoV-2 infection by time since symptom onset. *eLife* 2020; **9**:
17 e60122.
- 18 24Young BE, Ong SWX, Ng LFP, *et al.* Viral Dynamics and Immune Correlates of Coronavirus
19 Disease 2019 (COVID-19) Severity. *Clinical Infectious Diseases* 2020; published online Aug
20 28. DOI:10.1093/cid/ciaa1280.
- 21 25Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. *Journal of*
22 *the American Statistical Association* 1927; **22**: 209–12.
- 23 26Pung R. [rachaelpung/covid_missed_infections](https://github.com/rachaelpung/covid_missed_infections). 2021
24 https://github.com/rachaelpung/covid_missed_infections (accessed May 5, 2021).
- 25 27National Centre for Infectious Diseases (NCID), National University of Singapore (NUS),
26 Nanyang Technological University, Singapore (NTU Singapore). NCID, NUS and NTU
27 Studies Highlight the Role of Socio-Behavioural Factors in Managing COVID-19. 2020;
28 published online May 21.
29 [http://news.ntu.edu.sg/pages/newsdetail.aspx?URL=http://news.ntu.edu.sg/news/Pages/NR2](http://news.ntu.edu.sg/pages/newsdetail.aspx?URL=http://news.ntu.edu.sg/news/Pages/NR2020_May21.aspx&Guid=221cc9a8-f0cd-40ac-a837-c1caf19897b7&Category=All)
30 [020_May21.aspx&Guid=221cc9a8-f0cd-40ac-a837-c1caf19897b7&Category=All](http://news.ntu.edu.sg/pages/newsdetail.aspx?URL=http://news.ntu.edu.sg/news/Pages/NR2020_May21.aspx&Guid=221cc9a8-f0cd-40ac-a837-c1caf19897b7&Category=All) (accessed
31 March 14, 2021).

32

