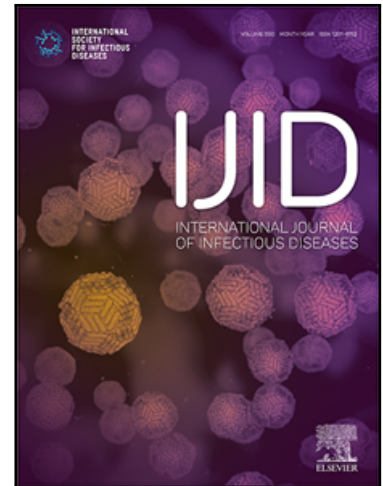


Journal Pre-proof

A hospital-related outbreak of SARS-CoV-2 associated with a novel variant Cal.20C (B.1.429) in Taiwan: transmission potential and outbreak containment under intensified contact tracing, January–February 2021

Andrei R. Akhmetzhanov , Sung-mok Jung , Hao-Yuan Cheng , Robin N. Thompson

PII: S1201-9712(21)00517-8
DOI: <https://doi.org/10.1016/j.ijid.2021.06.028>
Reference: IJID 5498



To appear in: *International Journal of Infectious Diseases*

Received date: 23 April 2021
Revised date: 11 June 2021
Accepted date: 12 June 2021

Please cite this article as: Andrei R. Akhmetzhanov , Sung-mok Jung , Hao-Yuan Cheng , Robin N. Thompson , A hospital-related outbreak of SARS-CoV-2 associated with a novel variant Cal.20C (B.1.429) in Taiwan: transmission potential and outbreak containment under intensified contact tracing, January–February 2021, *International Journal of Infectious Diseases* (2021), doi: <https://doi.org/10.1016/j.ijid.2021.06.028>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2021 The Author(s). Published by Elsevier Ltd on behalf of International Society for Infectious Diseases.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Highlights

- This outbreak was contained by extensive contact tracing and proactive isolation
- These rigorous measures allowed the outbreak end to be declared quickly
- Less stringent control would have meant lower confidence that the outbreak was over
- Public vigilance was still required for several weeks after the final case
- Later identification of the outbreak could have led to more infections

Journal Pre-proof

A hospital-related outbreak of SARS-CoV-2 associated with a novel variant Cal.20C (B.1.429) in Taiwan: transmission potential and outbreak containment under intensified contact tracing, January–February 2021

Andrei R. Akhmetzhanov^{a,*}, Sung-mok Jung^{b,c}, Hao-Yuan Cheng^d, Robin N. Thompson^{e,f}

^a College of Public Health, National Taiwan University, 17 Xu-Zhou Road, Zhongzheng District, Taipei, 10055, Taiwan

^b School of Public Health, Kyoto University, Yoshidakonoecho, Sakyo-ku, Kyoto, 606-8501, Japan

^c Graduate School of Medicine, Hokkaido University, Kita 15 Jo Nishi 7 Chome, Kita-ku, Sapporo-shi, Hokkaido, 060-8638, Japan

^d Epidemic Intelligence Center, Taiwan Centers for Disease Control, 6 Linsen South Road, Zhongzheng District, Taipei, 10050, Taiwan

^e Mathematics Institute, University of Warwick, Coventry, CV4 7AL, U.K.

^f Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research, University of Warwick, Coventry, CV4 7AL, U.K.

* Corresponding author at: Global Health Program & Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, 17 Xu-Zhou Rd, Taipei 10055, Taiwan.

Emails: akhmetzhanov@ntu.edu.tw (A.R. Akhmetzhanov), seductmd@gmail.com (S-m. Jung),
drhao@cdc.gov.tw (H-Y. Cheng), robin.n.thompson@warwick.ac.uk (R.N. Thompson)

Journal Pre-proof

Abstract

Objectives: A hospital-related cluster of 22 COVID-19 cases occurred in Taiwan in January–February 2021. Rigorous control measures were introduced and could only be relaxed once the outbreak was declared over. Each day after the apparent outbreak end, we estimated the risk of future cases occurring, to inform decision-making.

Methods: Probabilistic transmission networks were reconstructed and transmission parameters (the reproduction number R and overdispersion parameter k) were estimated. We estimated the reporting delay during the outbreak (Scenario 1). We also considered a counterfactual scenario with less effective interventions characterized by a longer reporting delay (Scenario 2). Each day, we estimated the risk of future cases under both scenarios.

Results: The values of R and k were estimated to be 1.30 (95% credible interval: 0.57,3.80) and 0.38 (0.12,1.20), respectively. The mean reporting delays considered were 2.5 days (Scenario 1) and 7.8 days (Scenario 2). The inferred probability of future cases occurring declined more quickly in Scenario 1 than Scenario 2.

Conclusions: Following outbreak containment, rigorous control measures allowed the outbreak to be declared over quickly. This highlights the need for effective interventions, not only to reduce cases during outbreaks but also to allow outbreaks to be declared over with confidence.

Introduction

As of 9 March 2021, Taiwan had confirmed fewer than 1,000 SARS-CoV-2 infections, of which 77 were locally acquired ([Taiwan Centers for Disease Control, 2021](#)). Following stringent border control measures, proactive contact tracing and case isolation, Taiwan's largest individual outbreak to date was a hospital-related outbreak that involved 22 cases and occurred in January-

February 2021. Despite successful containment of that outbreak, some aspects were concerning. First, the custom of wearing face masks, especially in hospital, was unable to fully prevent transmission in several instances ([Central Epidemic Command Center, 2021b](#)). Second, the source of infection for one of the infected inpatients was undetermined: that individual attended a hospital ward that was not included in a so-called “red zone”, and he did not interact with other detected infected individuals ([Central Epidemic Command Center, 2021a](#)). Third, having implemented control measures at the time of the first suspected cases, additional cases continued to be seen 1–2 weeks afterwards. This led to further investigations into possible causes of the outbreak and required an end-of-outbreak determination (i.e., assessment of the probability that the outbreak was over—or, conversely, the probability that additional reported cases would occur in future) after the last case was reported ([Djaafara et al., 2021](#), [Nishiura et al., 2016](#), [Parag et al., 2020](#)).

Here, we provide a descriptive analysis of the outbreak and quantify viral transmissibility during that outbreak. We also present estimates of the probability that additional reported cases would occur in future, as obtained in real-time after the final case had been observed. As the time since the last observed case increases, the certainty that the outbreak is over increases. We consider two distinct scenarios when estimating the probability of future cases. Scenario 1 describes containment of the outbreak under intensified contact tracing, as was the situation during this outbreak. Under Scenario 1, proactive testing and quarantine of all close contacts of detected cases (and suspected cases) after epidemiological investigations was considered, so that cases are found quickly and transmission beyond individuals that attended hospital and their contacts was unlikely. Scenario 2 describes a situation with reduced contact tracing and testing, increasing the

risk of transmission into the wider community with some chains of transmission potentially remaining untraced. In our analysis, these scenarios are assessed by implementing two different reporting delays, which represent the time periods from symptom onset to case confirmation. The reporting delay under Scenario 1 is shorter than under Scenario 2 due to efficient case identification and isolation (Tian et al., 2021).

Materials and Methods

Outbreak investigation

A cluster of locally acquired SARS-CoV-2 infections occurred in Taiwan in January–February 2021. This cluster originated in a hospital and involved 22 reported cases (Figure 1). The first two cases to be detected, a doctor (B1.1) and his household and work contact (B1.2), were suspected positive and tested on 11 January. They were then confirmed positive the following day. The authorities acted proactively by testing their close contacts on 11 January, ordering a two-week home isolation of all close contacts, restricting hospital admissions, and arranging for a second round of health inspections three days later. Regular press conferences raised public awareness and ensured that the local community remained vigilant throughout the outbreak.

The index case (A0) was a Taiwanese female in her 60s who travelled to the United States in October 2020 and returned to Taiwan on 27 December 2020. Having tested negative for SARS-CoV-2 infection within three days before her flight, she developed initial symptoms on 29 December while in quarantine. She was later hospitalized and was placed on a ventilator. During

her treatment, a doctor (B1.1) was exposed to the virus on 4 January 2021 and experienced initial mild symptoms on 8 January. The virus further spread to his household contact B1.2 and other medical personnel, most likely due to work-related interactions between B1.1, B2, and B3 on 10 January. The chains of transmission that followed then included three other work-related infections (B4–6), three infections of attending inpatients (C1.1, D1.1, E1) and transmission in their households. Household transmission accounted for 12 cases (57%), with the family cluster of B4.1 involving all seven family members including one death. In total, two deaths (B4.4, D1.1) occurred.

All cases were epidemiologically linked through contact tracing, except for an inpatient (D1.1) who had no record of contact with any known infected individual in the hospital. This suggests that his infection was likely due to either indirect transmission (e.g. via a contaminated surface from a known source) or from an undetected case. The same route of transmission could have occurred for infection of B2 by B1.1, since both individuals were wearing masks during their interaction (one of which was a highly effective surgical N95 mask).

One individual (C1.2) was pre-symptomatic when testing positive, with onset of symptoms two days later. Two infected individuals remained asymptomatic throughout infection. At least one pre-symptomatic transmission occurred: a foreign nurse (B6) was exposed to the virus on 7 January while interacting with B1.1, one day before B1.1 developed symptoms. Unlike the family cluster of B4.1, where the secondary attack rate was 100%, the employer of B6 and all his family members tested negative despite their close contact with infected case B6. Genetic

sequencing of a subset of cases from the outbreak identified a novel variant Cal.20C (lineage B.1.429), originated in Southern California in 2020 ([GISAID, 2021](#), [Zhang et al., 2021](#)).

Reconstruction of the transmission network and estimation of the transmission potential

We characterized the transmission potential of this novel variant by analyzing the offspring distribution, which describes the number of secondary infections per primary case. We fitted a negative-binomial distribution with mean R and overdispersion parameter k ([Riou and Althaus, 2020](#)). The reproduction number R describes the average number of secondary infections per primary case, while k measures variability in the number of secondary infections and quantifies the potential for superspreading (which is more likely to occur for lower values of k).

First, we applied the Wallinga-Teunis (WT) method ([Wallinga and Teunis, 2004](#)) to resolve the uncertainty in transmission patterns in family clusters by incorporating the serial interval distribution from ([Nishiura et al., 2020](#)), which is similar to other reported estimates ([Biggerstaff et al., 2020](#), [Hart et al., 2021](#)). The pairings of infectees to their infectors were known for 12 secondary cases as a result of epidemiological investigations: $A0 \rightarrow B1.1$, $B1.1 \rightarrow \{B1.2, B2, B3, B6, D1.1\}$, $B1.2 \rightarrow B1.3$, $B2 \rightarrow C1.1$, $B4.1 \rightarrow \{B4.2, B4.3\}$, and $D1.1 \rightarrow \{D1.2, E1\}$. We assigned the infection of case D1.1, an inpatient who attended the hospital in the first week of the outbreak, to case B1.1, given the timing and that only case B1.1 was symptomatic at that time. All other potential infectors (B2, B3, B4.1, B5, and B6) developed initial symptoms more than two days after the visit of D1.1 to the hospital (Figure 1). The

infectors of the other nine cases (excluding the index case A0) were uncertain, with the following possibilities: $\{B1.2, B1.3\} \rightarrow B1.4$, $\{B1.1, B3, B2\} \rightarrow B4.1$, $\{B4.1, B4.5, B4.6\} \rightarrow B4.4$, $\{B4.1, B4.4, B4.6\} \rightarrow B4.5$, $\{B4.1, B4.4, B4.5\} \rightarrow B4.6$, $\{B4.4, B4.5, B4.6\} \rightarrow B4.7$, $\{B1.1, B2, B3, B4.1\} \rightarrow B5$, $\{C1.1, C1.3\} \rightarrow C1.2$, and $\{C1.1, C1.2\} \rightarrow C1.3$. Almost all of these transmissions (except for infection of case B5) may have been due to household transmission, and so exact determination of who-infected-whom is impossible. The infector of case B5 could not be identified precisely as that transmission likely occurred in the workplace, where case B5 contacted multiple possible infectors. Under the WT method, for each of those nine infectees i , we selected an infector j from their lists of potential infectors J_i based on probabilistic sampling. The likelihood that case j (with symptoms onset at time t_j) infected case i , relative to the likelihood that any other potential infector infected case i , was given by:

$$p_{ij} = \frac{g(t_i - t_j | \{\mu_{SI}, k_{SI}\})}{\sum_{v \in J_i} g(t_i - t_v | \{\mu_{SI}, k_{SI}\})}, \quad (1)$$

where $g(\circ | \{\mu_{SI}, k_{SI}\})$ represents the serial interval distribution modeled by a Weibull distribution with the mean $\mu_{SI} = 4.8 \pm 0.6$ days (i.e. $\mu_{SI} \sim \mathcal{N}(4.8, 0.6)$) and shape parameter $k_{SI} = 2.3 \pm 0.4$ (Nishiura et al., 2020).

Second, we determined the number of transmissions from each primary case in any probabilistic realization of the transmission network. We fitted a negative binomial probability mass function to each resulting distribution, with mean R and overdispersion parameter k .

Generation-based reproduction number

The reconstructed transmission networks allowed us to make a probabilistic assignment of generation membership to cases and derive the generation-based reproduction number, R_m (Akhmetzhanov et al., 2018, Worden et al., 2020). Given a particular network, each node (i.e. each case) was assigned to a generation m , where the value of m represents the number of links from that node to the index case A0. The node A0 was placed at the root of the network and assigned to generation 0. To derive the generation-based reproduction number, R_m , we divided the number of transmissions generated by cases in generation m by the number of cases in that generation. Hence, the reproduction number for the final generation M ($m \leq M$) was exactly zero. The reproduction number for generation zero was equal to one. Because the transmission networks were generated probabilistically, each R_m was also characterized by a posterior distribution.

Estimation of the reporting delay

Fitting the reporting delay distribution with a mixture of three shifted distributions (gamma, Weibull, and lognormal), we estimated the mean reporting delay for this outbreak (under the intensive measures that were in place for this outbreak – Scenario 1). We also considered a counterfactual scenario (Scenario 2) in which public health measures are less rigorous. Rather than attempting to model the wide range of possible effects of less rigorous contact tracing and case isolation, in Scenario 2 we simply set the reporting delay to be longer than in Scenario 1. In Scenario 2, we set the mean reporting delay by estimating its value using data for all local cases reported in Taiwan since the beginning of 2020.

Specifically, for each scenario, we extracted data describing dates of symptoms onset and confirmation for all symptomatic cases. The number of extracted cases was $N = 20$ for Scenario 1 and $N = 68$ for Scenario 2. The likelihood was given by a mixture of three component likelihoods with respective weights w_l ($l = 1,2,3$), and right truncation at the time of the latest update T :

$$L(\theta | \{\Delta_i\}) = \sum_{l=\{1,2,3\}} w_l L^{(l)}(\theta | \{\Delta_i\}), \quad (2)$$

$$L^{(l)}(\theta | \{\Delta_i\}) = \prod_{k=1 \dots N} \frac{f_l(\Delta_i | \theta)}{F_l(T - o_i | \theta)}, \quad (3)$$

where $\Delta_i = c_i - o_i$ is the time difference between confirmation c_i and symptom onset o_i for case i . Because the extracted data contained only the dates of symptom onset O_i and confirmation C_i , we assumed that the priors for the times of symptom onset o_i and confirmation c_i were uniformly distributed within those days: $o_i \sim \mathcal{U}(O_i, O_i + 1)$ and $c_i \sim \mathcal{U}(C_i, C_i + 1)$. Some observed Δ_k were negative, so that the reporting delay distributions were modeled by either shifted gamma, Weibull, or lognormal distributions ($l = 1,2,3$). The function $f_l(\Delta_i | \theta)$ denoted the probability density function (PDF):

$$f_l(\Delta_i | \theta = \{\tau, \mu, \sigma\}) = \text{distribution}_l^{\text{PDF}}(\Delta_i + \tau | \mu, \sigma), \quad (4)$$

where τ is the shift of distribution l ($\tau > 0$), μ and σ are the mean and standard deviation of the distribution l . To improve the convergence of the mixture model, we assumed that parameters $\{\tau, \mu, \sigma\}$ were common to the three distributions, as has been proposed elsewhere for Bayesian

model averaging (Akhmetzhanov, 2021, Keller and Kamary, 2018). $F_l(T - o_i | \theta)$ denoted the cumulative distribution function (CDF). The truncation time T was set to 12 April 2021.

The relative weightings of the different component distributions l were defined using the formula:

$$q_l = \frac{w_l L^{(l)}(\theta | \{\Delta_i\})}{L(\theta | \{\Delta_i\})}. \quad (5)$$

End-of-outbreak probability

On a given day, to estimate whether or not the outbreak was already over, we used a previously described method devised by Linton *et al.* (Linton *et al.*, 2021). First, we considered the epidemic curve up to the time of report t with dates of symptom onsets $o_i < t$ for all symptomatic cases $i = 1 \dots N$. The probability that one or more new cases will be reported after day t is defined by the following expression:

$$\Pr(X(t) > 0) = 1 - \prod_{i=1}^N \sum_{y=1}^{\infty} p_y [H_i(t|\theta)]^y. \quad (6)$$

In this expression, $X(t)$ is the number of cases reported on day t and p_y is the probability of y transmissions occurring from a primary case i , which follows a negative binomial distribution with mean R and overdispersion parameter k as described above. The function $H_i(t|\theta)$ represents the CDF for the probability that an individual infected by case i reports infection by time t . This function is therefore the CDF of a convolution of the serial interval and the reporting

delay. For each potential infectee, the reporting delay was selected at random from the three distributions described above according to the probabilities q_i (5).

Technical details

R 4.1.0 (R Development Core Team, 2021) and CmdStan 2.27.0 (Stan Development Team, 2021) were used to conduct the main analysis; Python 3.6 was used for statistical inference of the generation-based reproduction number. Reproducible code for this study is available on GitHub at <https://github.com/aakhmetz/Taiwan-COVID19-end-of-outbreak-JanFeb2021>. All derived estimates of model parameters and the results of sensitivity analyses can also be found in Supplementary Materials (Supplementary Tables 1–2, Supplementary Figures 1–3).

Results

Our statistical inference of the offspring distribution identified the median R estimate to be 1.30 (95% credible interval (CI): 0.57, 3.80) and median k estimate to be 0.38 (95% CI: 0.12, 1.20). The generation-based reproduction number (i.e., the expected number of transmissions arising from an infector in a specific generation of the transmission chain, where patient A0 represents generation 0) declined throughout the outbreak from generation 1 onwards. In generation 1, the generation-based reproduction number was estimated to be 6, falling below 1 by generation 3 (Supplementary Figures 4–5; Supplementary Table 2). Inspection of probabilistic transmission networks (Supplementary Figure 6) confirmed a high value of the case reproduction number, R ,

for B1.1, but also supported sequential transmission of the virus within households resulting in a greater estimated value of k compared to earlier studies ([Bi et al., 2020](#), [Endo et al., 2020](#), [Nakajo and Nishiura, 2021](#)). A previous study by Ng *et al.* ([Ng et al., 2020](#)) involved an analysis of data from Taiwan from 2020, and found an estimated value of k that was substantially larger, in part due to the small sample size in their analysis (the posterior mean of k was 19.20).

We estimated the mean reporting delay for the outbreak (under the intensive measures that were in place for this outbreak – Scenario 1) to be 2.5 days (95% CI: 1.8, 3.5) with standard deviation (SD) 1.6 days (95% CI: 1.1, 2.9). Under counterfactual Scenario 2, where we instead estimated the mean reporting delay for all local cases reported in Taiwan since the beginning of 2020, the mean reporting delay was 7.8 days (95% CI: 6.2, 10.1) with SD 7.8 days (95% CI: 5.7, 13.2). Because of a small number of negative delays (i.e. some individuals were detected prior to developing symptoms), the distributions were shifted approximately 1 day earlier as a result of the model fitting (1.0 day (95% CI: 0.2, 2.9) for Scenario 1 and 0.8 days (95% CI: 0.1, 2.0) for Scenario 2). The observed difference in mean reporting delays between Scenario 1 and Scenario 2 can be attributed to different ways in which cases were detected. Under Scenario 1, cases were detected quickly by contact tracing, whereas under Scenario 2 cases were detected by both contact tracing and symptom-based surveillance ([Bi et al., 2020](#)).

Incorporating the posterior distributions for R , k , the serial interval ([Nishiura et al., 2020](#)) and the reporting delay in the formula for the end-of-outbreak probability (equation (6)), following the final case reported in this outbreak we observed a sharper decline in the estimated probability that new cases will be reported in future under Scenario 1 than Scenario 2 (green and black lines

and regions in Figure 2). Ten days after the last reported case, on 19 February this probability (reported here as a percentage) dropped to 24.7% under Scenario 1 compared to 79.7% under Scenario 2. Depending on the policy-maker's "acceptable risk", different thresholds in this probability could be chosen before declaring an outbreak over (Thompson et al., 2019). For instance, if a threshold of 10% is chosen, the outbreak could have been declared over on 24 February under Scenario 1 compared to a later date of 18 March under Scenario 2. Sensitivity analyses are presented for different values of R and k , as well as different reporting delay distributions for Scenario 2, in the Supplementary Material, indicating qualitatively similar results. In each case, more rigorous control measures (characterized by a shorter reporting delay) allow policy-makers to be confident that the outbreak is over sooner after the final reported case.

Conclusions

In summary, our results suggest that the rigorous public health measures that were in place allowed the outbreak end to be declared around three weeks earlier than if these intensive measures were not introduced. Stringent control measures allowed policy-makers to be confident that the epidemic was over earlier compared to a scenario with less intense measures. However, even with strict control measures, public vigilance was required for 2–3 weeks after the final reported case until total confidence that the outbreak was over was achieved (Figure 2).

We conclude that proactive countermeasures and high public compliance contributed to efficient containment and a high confidence that the hospital-related outbreak in Taiwan was over by mid-February 2021. We note that later identification of the outbreak could have led to larger number

of infections ([Akhmetzhanov, 2020](#), [Liu et al., 2020](#)), and therefore potentially a later end-of-outbreak declaration.

Ethical Approval

The present study used publicly available data, and thus, did not require ethical approval.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Taiwan public health authorities and institutions for surveillance, laboratory testing, epidemiological investigations, and data collection. We are grateful to two anonymous reviewers for their helpful comments. A.R.A. also thanks Yin-Chin Fan and Yun-Chun Wu (National Taiwan University) for discussions about this topic.

Conflict of Interest

We declare that we have no conflict of interest.

Funding Source

S-m.J. received funding from the Japan Society for the Promotion of Science (JSPS) via a KAKENHI grant (20J2135800).

References

Akhmetzhanov AR. Large SARS-CoV-2 Outbreak Caused by Asymptomatic Traveler, China.

Emerg Infect Dis 2020;26(12):3106.

Akhmetzhanov AR. Estimation of delay-adjusted all-cause excess mortality in the USA: March–

December 2020 (Submitted). *Epidemiol Infect* 2021.

Akhmetzhanov AR, Lee H, Jung S-m, Kinoshita R, Shimizu K, Yoshi K, et al. Real time

forecasting of measles using generation-dependent mathematical model in Japan, 2018. *PLoS*

Curr Outbreaks 2018;10.

Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. Epidemiology and transmission of COVID-19

in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study.

Lancet Infect Dis 2020;20(8):911-9.

Biggerstaff M, Cowling BJ, Cucunubá ZM, Dinh L, Ferguson NM, Gao H, et al. Early insights

from statistical and mathematical modeling of key epidemiologic parameters of COVID-19.

Emerg Infect Dis 2020;26(11).

Central Epidemic Command Center. COVID-19 press conference of 5 February 2021 (in

Chinese); 2021a. Available from: <https://youtu.be/SLMJIV6ncBI>. [Accessed 12 April 2021].

Central Epidemic Command Center. COVID-19 press conference of 17 January 2021 (in Chinese); 2021b. Available from: <https://youtu.be/LvGoNhe2lXw>. [Accessed 12 April 2021].

Djaafara BA, Imai N, Hamblion E, Impouma B, Donnelly CA, Cori A. A quantitative framework for defining the end of an infectious disease outbreak: application to Ebola virus disease. *Am J Epidemiol* 2021;190(4):642-51.

Endo A, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China [version 3; peer review: 2 approved]. *Wellcome Open Res* 2020;5.

GISAID. Accession numbers of first 3 samples: EPI_ISL_956329, EPI_ISL_956330, EPI_ISL_1020315. 2021.

Hart WS, Maini PK, Thompson RN. High infectiousness immediately before COVID-19 symptom onset highlights the importance of continued contact tracing. *eLife* 2021;10.

Keller M, Kamary K. Bayesian model averaging via mixture model estimation. *arXiv* 2018.

Linton NM, Akhmetzhanov AR, Nishiura H. Localized end-of-outbreak determination for coronavirus disease 2019 (COVID-19): examples from clusters in Japan. *Int J Infect Dis* 2021;105:286-92.

Liu J, Huang J, Xiang D. Large SARS-CoV-2 outbreak caused by asymptomatic traveler, China. *Emerg Infect Dis* 2020;26:2260-3.

Nakajo K, Nishiura H. Transmissibility of asymptomatic COVID-19: Data from Japanese clusters. *Int J Infect Dis* 2021;105:236-8.

Ng T, Cheng H, Chang H, Liu C, Yang C, Jian S, et al. Effects of case- and population-based COVID-19 interventions in Taiwan. *medrxiv* 2020.

Nishiura H, Linton NM, Akhmetzhanov AR, 2020;113332. Serial interval of novel coronavirus (COVID-19) infections. *Int J Infect Dis* 2020;93:284-6.

Nishiura H, Miyamatsu Y, Mizumoto K. Objective determination of end of MERS outbreak, South Korea, 2015. *Emerg Infect Dis* 2016;22(1):146-8.

Parag KV, Donnelly CA, Jha R, Thompson RN. An exact method for quantifying the reliability of end-of-epidemic declarations in real time. *PLoS Comput Biol* 2020;16(11):e1008478.

R Development Core Team. R: a language and environment for statistical computing; 2021. Available from: <https://www.r-project.org>.

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).

Stan Development Team. Stan modeling language users guide and reference manual, 2.27.0; 2021. Available from: <https://mc-stan.org>. [Accessed 3 June 2021].

Taiwan Centers for Disease Control. 2021. Available from: <https://www.cdc.gov.tw>. [Accessed 12 April 2021].

Thompson RN, Morgan OW, Jalava K. Rigorous surveillance is necessary for high confidence in end-of-outbreak declarations for Ebola and other infectious diseases. *Phil Trans Roy Soc B* 2019;374(1776).

Tian L, Li X, Qi F, Tang Q-Y, Tang V, Liu J, et al. Harnessing peak transmission around symptom onset for non-pharmaceutical intervention and containment of the COVID-19 pandemic. *Nat Commun* 2021;12(1).

Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am J Epidemiol* 2004;160:509-16.

Worden L, Ackley SF, Zipprich J, Harriman K, Enanoria WTA, Wannier R, et al. Measles transmission during a large outbreak in California. *Epidemics* 2020;30.

Zhang W, Davis BD, Chen SS, Sincuir Martinez JM, Plummer JT, Vail E. Emergence of a novel SARS-CoV-2 variant in Southern California. *JAMA* 2021;325(13):1324-6.

Figures

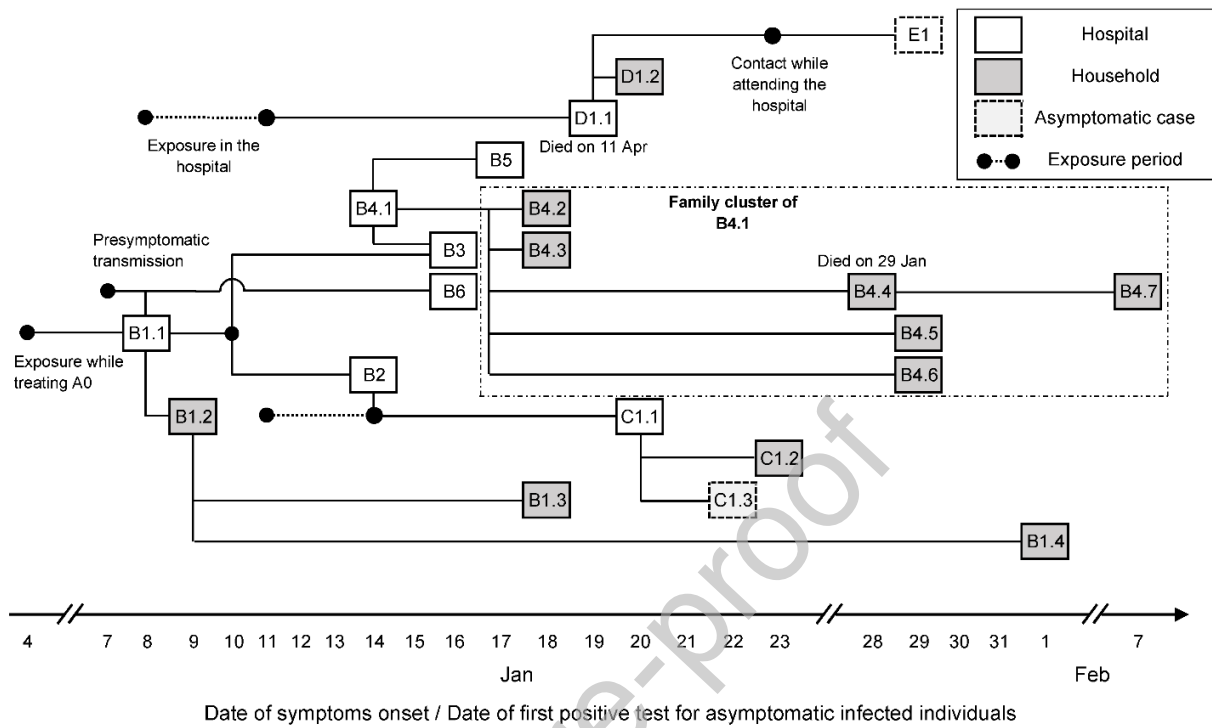


Figure 1. Timeline of exposure and possible connections between reported cases. Connections shown here were determined either by identifying the most probable infector via epidemiological investigation or by the earliest time of symptom onset among all close contacts if the most likely pair could not be determined (such as in family clusters of B1.2, C1.1, and B4.1).

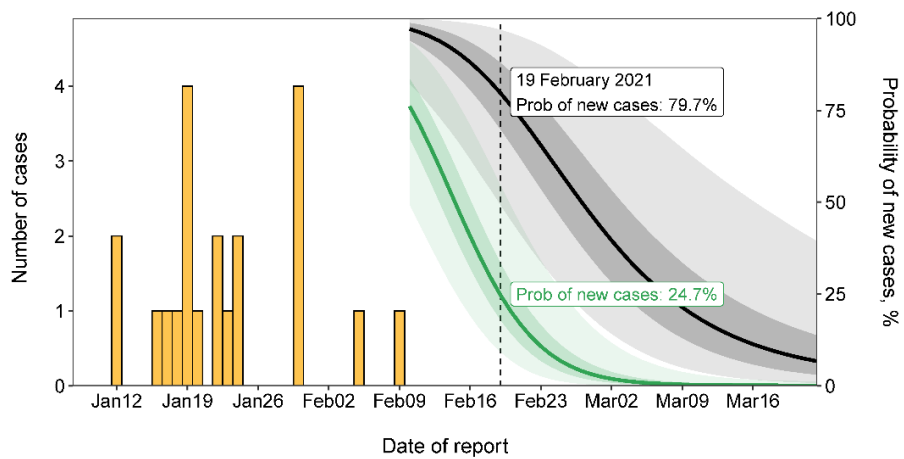


Figure 2. The estimated risk of cases being reported in future for Scenario 1 (under intensified contact tracing; green) and Scenario 2 (less rigorous public health measures; black). Bars in orange indicate the incidence of COVID-19 by confirmation date.