Surveillance testing of SARS-CoV-2

Daniel B. Larremore[†1,2], Bryan Wilder[3], Evan Lester[6,5], Soraya Shehata[4,5], James M. Burke[6], James A. Hay[7,8], Milind Tambe[3], Michael J. Mina[‡7,8,9,*], and Roy Parker[§4,6,10,2,*]

[1]*Department of Computer Science, University of Colorado Boulder*
[2]*BioFrontiers Institute, University of Colorado at Boulder*
[3]*Center for Research on Computation & Society, Harvard John A Paulson School of Engineering and Applied Sciences, Harvard University*
[4]*Department of Molecular, Cellular and Developmental Biology, University of Colorado*
[5]*Medical Scientist Training Program, University of Colorado Anschutz Medical Campus*
[6]*Department of Biochemistry, University of Colorado Boulder*
[7]*Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health*
[8]*Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health*
[9]*Department of Pathology, Brigham and Women's Hospital, Harvard Medical School*
[10]*Howard Hughes Medical Institute*
[*]*These authors contributed equally.*

The COVID-19 pandemic has created a public health crisis. Because SARS-CoV-2 can spread from individuals with pre-symptomatic, symptomatic, and asymptomatic infections [1, 2, 3], the re-opening of societies and the control of virus spread will be facilitated by robust surveillance, for which virus testing will often be central. After infection, individuals undergo a period of incubation during which viral titers are usually too low to detect, followed by an exponential growth of virus, leading to a peak viral load and infectiousness, and ending with declining viral levels and clearance [4]. Given the pattern of viral load kinetics [4], we model surveillance effectiveness considering test sensitivities, frequency, and sample-to-answer reporting time. These results demonstrate that effective surveillance, including time to first detection and outbreak control, depends largely on frequency of testing and the speed of reporting, and is only marginally improved by high test sensitivity. We therefore conclude that surveillance should prioritize accessibility, frequency, and sample-to-answer time; analytical limits of detection should be secondary.

[†]daniel.larremore@colorado.edu
[‡]mmina@hsph.harvard.edu
[§]roy.parker@colorado.edu

The reliance on testing as a means to safely reopen societies has placed a microscope on the analytical sensitivity of virus assays, with a gold-standard of quantitative real-time polymerase chain reaction (qPCR). These assays have analytical limits of detection that are usually within around 10^3 viral RNA copies per ml (cp/ml) [5]. However, qPCR remains expensive and as a laboratory based assay often have sample-to-result times of 24-48 hours. New developments in SARS-CoV-2 diagnostics have the potential to reduce cost significantly, allowing for expanded testing or greater frequency of testing and can reduce turnaround time to minutes. These assays however largely do not meet the gold standard for analytical sensitivity, which has encumbered translation of these assays for widescale use [6].

Three features of the viral increase, infectivity, and decline during SARS-CoV-2 infection led us to hypothesize that there might be minimal differences in effective surveillance using viral detection tests of different sensitivities, such as RT-qPCR with a limit of detection (LOD) at 10^3 cp/ml [5] compared to often cheaper or faster assays with higher limits of detection (i.e., around 10^5 cp/ml) such as point-of-care nucleic acid LAMP and rapid antigen tests (Figure 1A). First, since filtered samples collected from patients displaying less than 10^6 N or E RNA cp/ml contain minimal or no measurable infectious virus [7, 8, 9], either class of test should detect individuals who are currently infectious. The absence of infectious particles at viral RNA concentrations $< 10^6$ cp/ml is likely due to (i) the fact that the N and E RNAs are also present in abundant subgenomic mRNAs, leading to overestimation of the number of actual viral genomes by $\sim$100-1000X [10], (ii) technical artifacts of RT-PCR at Ct values > 35 due to limited template [11, 12], and (iii) the production of non-infectious viral particles as is commonly seen with a variety of RNA viruses [13]. Second, during the exponential growth of the virus, the time difference between 10^3 and 10^5 cp/ml is short, allowing only a limited window in which only the more sensitive test could diagnose individuals. For qPCR, this corresponds to the time required during viral growth to go from Ct values of 40 to $\sim$34. While this time window for SARS-CoV-2 is not yet rigorously defined, for other respiratory viruses such as influenza, and in ferret models of SARS-CoV-2 transmission, it is on the order of a day [14, 15]. Finally, high-sensitivity screening tests, when applied during the viral decline accompanying recovery, are unlikely to substantially impact transmission because such individuals detected have low, if any, infectiousness [10].

To examine how surveillance testing would reduce the average infectiousness of individuals, we first modeled the viral loads and infectiousness curves of 10,000 simulated individuals using the predicted viral trajectories of SARS-CoV-2 infections based on key features of latency, growth, peak, and decline identified in the literature (Figure 1A; see Methods). Accounting for these within-host viral kinetics, we calculated what percentage of their total infectiousness would be removed by surveillance and isolation (Figure 1B) with tests at LOD of 10^3 and 10^5, and at different frequencies. Here, infectiousness was taken to be proportional to the logarithm of viral load (with alternative assumptions addressed in Supplemental Materials), consistent with the observation that pre-symptomatic patients are most infectious just prior to the onset of symptoms [4], and evidence that the efficiency of viral transmission coincides with peak viral loads, which was also identified during the related 2003 SARS outbreak [16, 17]. We considered that 20% of patients would undergo symptomatic isolation near their peak viral load if they had not been tested and

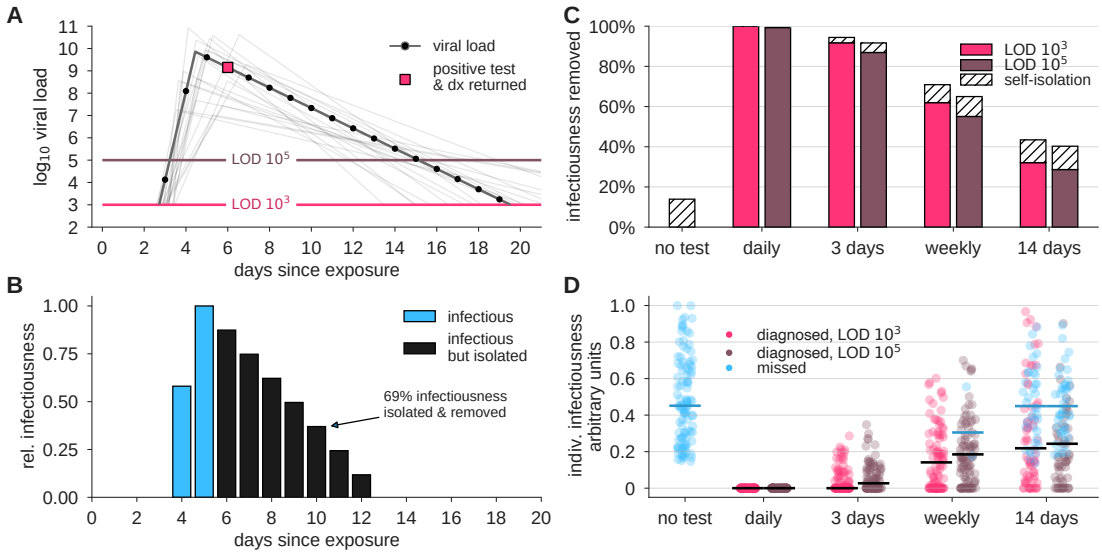

Figure 1: **Surveillance testing effectiveness depends on frequency.** (A) An example viral load trajectory is shown with LOD thresholds of two tests, and a hypothetical positive test on day 6, two days after peak viral load. 20 other stochastically generated viral loads are shown to highlight trajectory diversity (light grey; see Methods). (B) Relative infectiousness for the viral load shown in panel A pre-test, totaling 31% (blue) and post-isolation, totaling 69% (black). (C) Surveillance programs using tests at LODs of 10^3 and 10^5 at frequencies indicated were applied to $10,000$ individuals trajectories of whom 20% would undergo symptomatic isolation near their peak viral load if they had not been tested and isolated first. Total infectiousness removed during surveillance (colors) and self isolation (hatch) are shown for surveillance as indicated, relative to total infectiousness with no surveillance or self-isolation. (D) The impact of surveillance on the infectiousness of 100 individuals is shown for each surveillance program and no testing, as indicated, with each individual colored by test if their infection was detected during infectiousness (medians, black lines) or colored blue if their infection was missed by surveillance or detected positive *after* their infectious period (medians, blue lines). Units are arbitrary and scaled to the maximum infectiousness of sampled individuals.

isolated first, and 80% would have sufficiently mild or no symptoms such that they would not isolate unless they were detected by surveillance testing. This analysis demonstrated that there was little difference in averting infectiousness between the two classes of test. Dramatic reductions in total infectiousness of the individuals were observed by testing daily or every third day, $\sim 60\%$ reduction when testing weekly, and $< 40\%$ under biweekly testing (Figure 1C). Because viral loads and infectiousness vary across individuals, we also analyzed the impact of different surveillance regimes on the distribution of individuals' infectiousness (Figure 1D).

Above, we assumed that each infection was independent. To investigate the effects of surveillance testing strategies at the population level, we used simulations to monitor whether epidemics were contained or became uncontrolled, while varying the frequencies at which the test was administered, ranging from daily testing to testing every 14 days, and considering tests with LOD of 10^3 and 10^5, analogous to RT-qPCR and RT-LAMP / rapid antigen tests, respectively. We used two different epidemiological models to ensure that important observations were independent of the specific modeling approach. The first model is a previously described agent-based model with

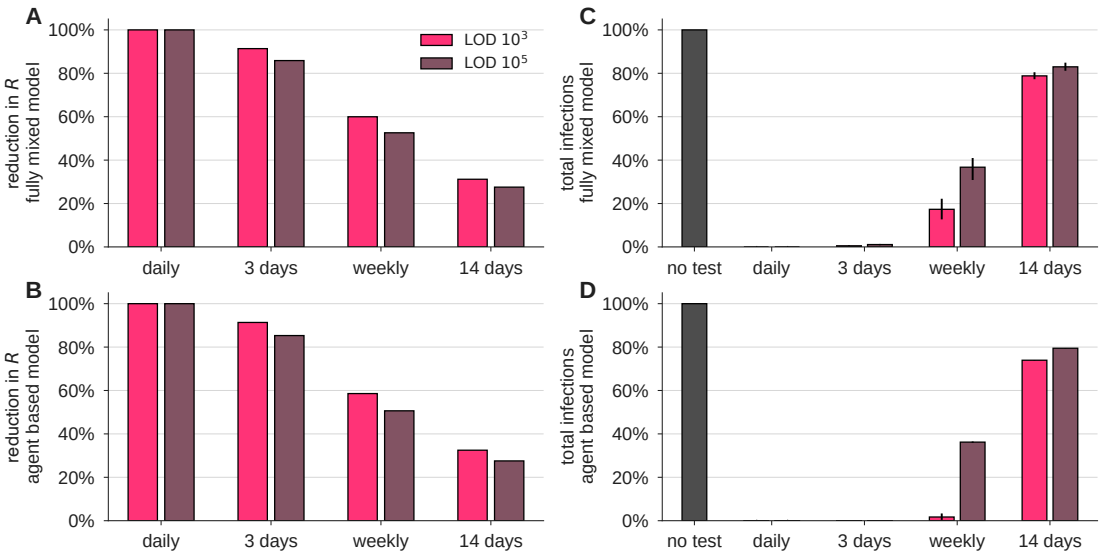

Figure 2: **Surveillance testing affects disease dynamics.** Both the fully-mixed compartmental model (top row) and agent based model (bottom row) are affected by surveillance programs. (A, B) More frequent testing reduces the effective reproductive number R, shown as the percentage by which R_0 is reduced, $100 \times (R_0 - R)/R_0$. Values of R were estimated from 50 independent simulations of dynamics (see Methods). (C, D) Relative to no testing (grey bars), surveillance suppresses the total number of infections in both models when testing every day or every three days, but only partially mitigates total cases for weekly or bi-weekly testing. Error bars indicate inner 95% quantiles of 50 independent simulations each.

both within-household and age-stratified contact structure based on census microdata in a city representative of New York City [18], and initialized with 100 cases without additional external infections. The second model is a simple fully mixed model representing a population of 20,000, similar to a large university setting, with a constant rate of external infection approximately equal to one new import per day. Individual viral loads were simulated for each infection, and individuals who received a positive test result were isolated, but contact tracing and monitoring was not included to more conservatively estimate the impacts of surveillance alone [19, 20]. Model details and parameters are fully described in Methods.

We observed that a surveillance program administering either test with high frequency limited viral spread, measured by both a reduction in the reproductive number R (Figures 2A and B; see Methods for calculation procedure) and by the total infections that persisted in spite of different surveillance programs, expressed relative to no surveillance (Figures 2C and D). Testing frequency was found to be the primary driver of population-level epidemic control, with only a small margin of improvement provided by using a more sensitive test. Direct examination of simulations showed that with no surveillance or biweekly testing, infections were uncontrolled, whereas surveillance testing weekly with either LOD $= 10^3$ or 10^5 effectively attenuated surges of infections (examples shown in Figure S1).

The relationship between test sensitivity and the frequency of testing required to control outbreaks in both the fully mixed model and the agent-based model generalize beyond the examples shown in

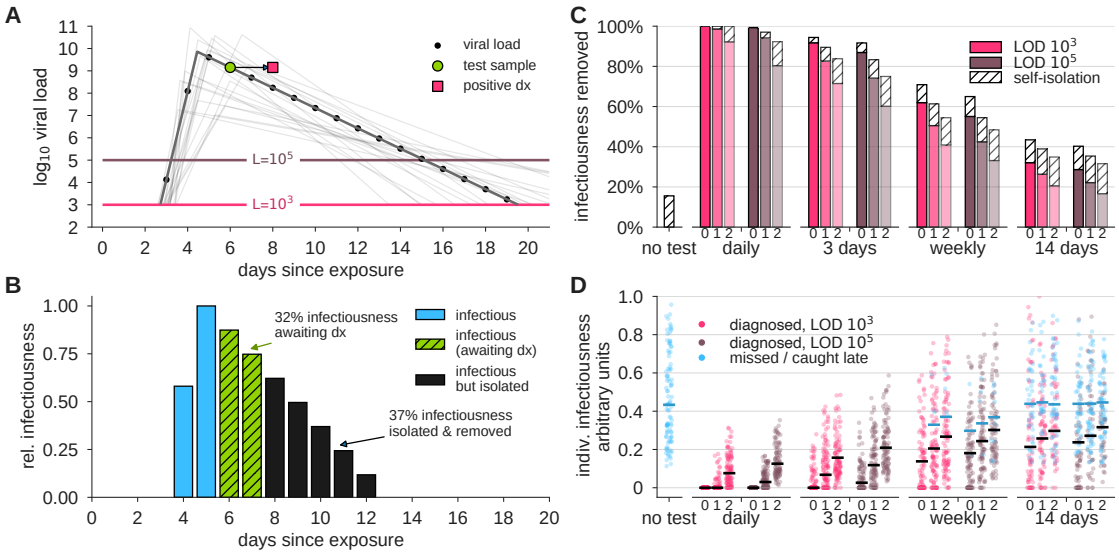

Figure 3: **Effectiveness of surveillance testing is compromised by delays in reporting.** (A) An example viral load trajectory is shown with LOD thresholds of two tests, and a hypothetical positive test on day 6, but with results reported on day 8. 20 other stochastically generated viral loads are shown to highlight trajectory diversity (light grey; see Methods). (B) Relative infectiousness for the viral load shown in panel A pre-test (totaling 31%; blue) and post-test but pre-diagnosis (totaling 32%; green), and post-isolation (totaling 37%; black). (C) Surveillance programs using tests at LODs of 10^3 and 10^5 at frequencies indicated, and with results returned after 0, 1, or 2 days (indicated by small text beneath bars) were applied to $10,000$ individuals trajectories of whom 20% were symptomatic and self-isolated after peak viral load if they had not been tested and isolated first. Total infectiousness removed during surveillance (colors) and self isolation (hatch) are shown, relative to total infectiousness with no surveillance or self-isolation. Delays substantially impact the fraction of infectiousness removed. (D) The impact of surveillance with delays in returning diagnosis of 0, 1, or 2 days (small text beneath axis) on the infectiousness of 100 individuals is shown for each surveillance program and no testing, as indicated, with each individual colored by test if their infection was detected during infectiousness (medians, black lines) or colored blue if their infection was missed by surveillance or diagnosed positive *after* their infectious period (medians, blue lines). Units are arbitrary and scaled to the maximum infectiousness of sampled individuals.

Figure 2 and are also seen at other testing frequencies and sensitivities. We simulated both models at LODs of 10^3, 10^5, and 10^6, and for testing ranging from daily to every 14 days. For those, we measured each surveillance policy's impact on total infections (Figure S2A and B) and on R (Figure S2C and D). In Figure 2, we modeled infectiousness as proportional to $\log_{10}$ of viral load. To address whether these finding are sensitive to this modeled relationship, we performed similar simulations with infectiousness proportional to viral load (Figure S3), or uniform above 10^6/ml (Figure S4). We found that results were robust to these large variations in the modeled relationship between infectiousness and viral load.

An important variable in surveillance testing is the time between a test's sample collection and the reporting of a diagnosis. To examine how time to reporting affected epidemic control, we re-analyzed both the reduction in individuals' infectiousness, as well as the epidemiological sim-

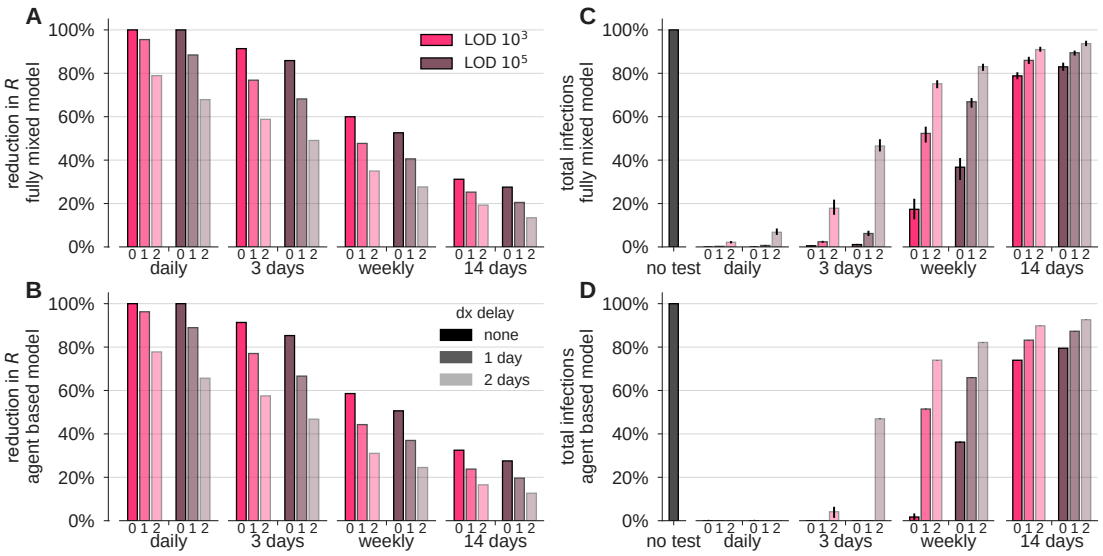

Figure 4: **Delays in reporting decrease the epidemiological impact of surveillance-driven isolation.** The effectiveness of surveillance programs are dramatically diminished by delays in reporting in both the fully-mixed compartmental model (top row) and agent based model (bottom row). (A, B) The impact of surveillance every day, 3 days, weekly, or biweekly, on the reproductive number R, calculated as $100 \times (R_0 - R)/R_0$, is shown for LODs 10^3 and 10^5 and delays of 0, 1, or 2 days (small text below axis). Values of R were estimated from 50 independent simulations of dynamics (see Methods). (C, D) Relative to no testing (grey bars), surveillance suppresses the total number of infections in both models when testing every day or every three days, but delayed results lead to only partial mitigation of total cases, even for testing every day or 3 days. Error bars indicate inner 95% quantiles of 50 independent simulations each.

ulations, comparing the results of instantaneous reporting (reflecting a rapid point-of-care assay), one day delay, and two day delay (Figure 3A and B). Delays in reporting dramatically decreased the reduction in infectiousness in individuals as seen by the total infectiousness removed (Figure 3C), the distribution of infectiousness in individuals (Figure 3D), or the dynamics of the epidemiological models (Figure 4). This result was robust to the modeled relationship between infectiousness and viral load in both simulation models and for various test sensitivities and frequencies (Figure S5). These results highlight that delays in reporting lead to dramatically less effective control of viral spread and emphasize that fast reporting of results is critical in any surveillance testing. These results also reinforce the relatively smaller benefits of improved limits of detection.

Communities vary in their transmission dynamics, due to difference in rates of imported infections and in the basic reproductive number R_0, both of which will influence the frequency and sensitivity with which surveillance testing must occur. We performed two analyses to illustrate this point. First, we varied the rate of external infection in our fully mixed model, and confirmed that when the external rate of infection is higher, more frequent surveillance is required to prevent outbreaks (Figure S6A). Second, we varied the reproductive number R_0 between infected individuals in both models, and confirmed that at higher R_0, more frequent surveillance is also required (Figure S6B and C). This may be relevant to institutions like college campuses or military bases wherein fre-

quent classroom setting or dormitory living are likely to increase contact rates. Thus, the specific strategy for successful surveillance will depend on the current community infection prevalence and transmission rate.

Our results lead us to conclude that surveillance testing of asymptomatic individuals can be used to limit the spread of SARS-CoV-2. However, our findings are subject to a number of limitations. First, the sensitivity of a test may depend on factors beyond LOD, including manufacturer variation and improper clinical sampling [21], though the latter may be ameliorated by different approaches to sample collection, such as saliva-based testing [22]. Second, the exact performance differences between testing schemes will depend on whether our model truly captures viral kinetics and infectiousness profiles [4], particularly during the acceleration phase between exposure and peak viral load. Continued clarification of these within-host dynamics would increase the impact and value of this, and other [19, 20] modeling studies.

A critical point is that the requirements for surveillance testing are distinct from clinical testing. Clinical diagnoses target symptomatic individuals, need high accuracy and sensitivity, and are not limited by cost. Because they focus on symptomatic individuals, those individuals can isolate such that a diagnosis delay does not lead to additional infections. In contrast, results from the surveillance testing of asymptomatic individuals need to be returned quickly, since even a single day diagnosis delay compromises the surveillance program's effectiveness. Indeed, at least for viruses with infection kinetics similar to SARS-CoV-2, we find that speed of reporting is much more important than sensitivity, although more sensitive tests are nevertheless somewhat more effective.

The difference between clinical and surveillance testing highlights the need for additional tests to be approved and utilized for surveillance. Such tests should not be held to the same degree of sensitivity as clinical tests, in particular if doing so encumbers rapid deployment of faster cheaper SARS-CoV-2 assays. We suggest that the FDA, other agencies, or state governments, encourage the development and use of alternative faster and lower cost tests for surveillance purposes, even if they have poorer limits of detection. If the availability of point-of-care or self-administered surveillance tests leads to faster turnaround time or more frequent testing, our results suggest that they would have high epidemiological value.

Our modeling suggests that some types of surveillance will subject some individuals to unnecessary quarantine days. For instance, the infrequent use of a sensitive test will not only identify (i) those with a low viral load in the beginning of the infection, who must be isolated to limit viral spread, but (ii) those in the recovery period, who still have detectable virus or RNA but are below the infectious threshold [9, 10]. Isolating this second group of patients will have no impact on viral spread but will incur costs of isolation. The use of serology, repeat testing 24 or 48 hours apart, or some other test, to distinguish low viral load patients on the upslope of infection from those in the recovery phase could allow for more effective quarantine decisions.

Acknowledgements

The authors wish to thank the BioFrontiers Institute IT HPC group. This work was supported by grants NIH F32 AI145112 (James Burke), NIH F30 AG063468 (Evan Lester), MURI W911NF-17-1-0370 (Milind Tambe), an NIH directors DP5 award 1DP5OD028145-01 (Michael Mina), and the Howard Hughes Medial Institute (Roy Parker).

References

[1] Melissa M Arons, Kelly M Hatfield, Sujan C Reddy, Anne Kimball, Allison James, Jesica R Jacobs, Joanne Taylor, Kevin Spicer, Ana C Bardossy, Lisa P Oakley, et al. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *New England Journal of Medicine*, 2020.

[2] Desmond Sutton, Karin Fuchs, Mary D'alton, and Dena Goffman. Universal screening for SARS-CoV-2 in women admitted for delivery. *New England Journal of Medicine*, 2020.

[3] Daniel P Oran and Eric J Topol. Prevalence of asymptomatic SARS-CoV-2 infection: A narrative review. *Annals of Internal Medicine*, 2020.

[4] Xi He, Eric HY Lau, Peng Wu, Xilong Deng, Jian Wang, Xinxin Hao, Yiu Chung Lau, Jessica Y Wong, Yujuan Guan, Xinghua Tan, et al. Temporal dynamics in viral shedding and transmissibility of covid-19. *Nature Medicine*, 26(5):672–675, 2020.

[5] Chantal BF Vogels, Anderson F Brito, Anne Louise Wyllie, Joseph R Fauver, Isabel M Ott, Chaney C Kalinich, Mary E Petrone, Marie-Louise Landry, Ellen F Foxman, and Nathan D Grubaugh. Analytical sensitivity and efficiency comparisons of sars-cov-2 qrt-pcr assays. *medRxiv*, 2020.

[6] Coronavirus (COVID-19) update: FDA informs public about possible accuracy concerns with Abbott ID NOW Point-of-Care Test. https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-informs-public-about-possible-accuracy-concerns-abbott-id-now-point, May 14, 2020.

[7] Kendra Quicke, Emily Gallichote, Nicole Sexton, Michael Young, Ashley Janich, Gregory Gahm, Elizabeth J Carlton, Nicole Ehrhart, and Gregory D Ebel. Longitudinal surveillance for SARS-CoV-2 rna among asymptomatic staff in five colorado skilled nursing facilities: Epidemiologic, virologic and sequence analysis. *medRxiv*, 2020.

[8] Roman Wölfel, Victor M Corman, Wolfgang Guggemos, Michael Seilmaier, Sabine Zange, Marcel A Müller, Daniela Niemeyer, Terry C Jones, Patrick Vollmar, Camilla Rothe, et al. Virological assessment of hospitalized patients with covid-2019. *Nature*, 581(7809):465–469, 2020.

[9] Bernard La Scola, Marion Le Bideau, Julien Andreani, Van Thuan Hoang, Clio Grimaldier, Philippe Colson, Philippe Gautret, and Didier Raoult. Viral rna load as determined by cell

culture as a management tool for discharge of SARS-CoV-2 patients from infectious disease wards. *European Journal of Clinical Microbiology & Infectious Diseases*, 39(6):1059, 2020.

[10] Soren Alexandersen, Anthony Chamings, and Tarka Raj Bhatta. Sars-cov-2 genomic and subgenomic rnas in diagnostic samples are not an indicator of active replication. *medRxiv*, 2020.

[11] Charles GB Caraguel, Henrik Stryhn, Nellie Gagné, Ian R Dohoo, and K Larry Hammell. Selection of a cutoff value for real-time polymerase chain reaction results to fit a diagnostic purpose: analytical and epidemiologic approaches. *Journal of Veterinary Diagnostic Investigation*, 23(1):2–15, 2011.

[12] Adrián Ruiz-Villalba, Elizabeth van Pelt-Verkuil, Quinn D Gunst, Jan M Ruijter, and Maurice JB van den Hoff. Amplification of nonspecific products in quantitative polymerase chain reactions (qpcr). *Biomolecular detection and quantification*, 14:7–18, 2017.

[13] PJ Klasse. Molecular determinants of the ratio of inert to infectious virus particles. In *Progress in molecular biology and translational science*, volume 129, pages 285–326. Elsevier, 2015.

[14] Amber M Smith and Alan S Perelson. Influenza A virus infection kinetics: quantitative data and models. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 3(4):429–445, 2011.

[15] Mathilde Richard, Adinda Kok, Dennis de Meulder, Theo M Bestebroer, Mart M Lamers, Nisreen MA Okba, Martje Fentener van Vlissingen, Barry Rockx, Bart L Haagmans, Marion PG Koopmans, et al. Sars-cov-2 is transmitted via contact and via the air between ferrets. *bioRxiv*, 2020.

[16] Zhuang Shen, Fang Ning, Weigong Zhou, Xiong He, Changying Lin, Daniel P Chin, Zonghan Zhu, and Anne Schuchat. Superspreading sars events, beijing, 2003. *Emerging Infectious Diseases*, 10(2):256, 2004.

[17] Joseph Sriyal Malik Peiris, Chung-Ming Chu, Vincent Chi-Chung Cheng, KS Chan, IFN Hung, Leo LM Poon, Kin-Ip Law, BSF Tang, TYW Hon, CS Chan, et al. Clinical progression and viral load in a community outbreak of coronavirus-associated sars pneumonia: a prospective study. *The Lancet*, 361(9371):1767–1772, 2003.

[18] Bryan Wilder, Marie Charpignon, Jackson A Killian, Han-Ching Ou, Aditya Mate, Shahin Jabbari, Andrew Perrault, Angel Desai, Milind Tambe, and Maimuna S Majumder. Modeling between-population variation in COVID-19 dynamics in hubei, lombardy, and new york city. *Available at SSRN 3564800*, 2020.

[19] Corey M Peak, Rebecca Kahn, Yonatan H Grad, Lauren M Childs, Ruoran Li, Marc Lipsitch, and Caroline O Buckee. Individual quarantine versus active monitoring of contacts for the mitigation of covid-19: a modelling study. *The Lancet Infectious Diseases*, 2020.

[20] Adam J Kucharski, Petra Klepac, Andrew Conlan, Stephen M Kissler, Maria Tang, Hannah Fry, Julia Gog, John Edmunds, CMMID COVID-19 Working Group, et al. Effectiveness of isolation, testing, contact tracing and physical distancing on reducing transmission of sars-cov-2 in different settings. *medRxiv*, 2020.

[21] Yicheng Fang, Huangqi Zhang, Jicheng Xie, Minjie Lin, Lingjun Ying, Peipei Pang, and Wenbin Ji. Sensitivity of chest ct for covid-19: comparison to rt-pcr. *Radiology*, page 200432, 2020.

[22] Anne Louise Wyllie, John Fournier, Arnau Casanovas-Massana, Melissa Campbell, Maria Tokuyama, Pavithra Vijayakumar, Bertie Geng, M Catherine Muenker, Adam J Moore, Chantal BF Vogels, et al. Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs. *Medrxiv*, 2020.

[23] Stephen A Lauer, Kyra H Grantz, Qifang Bi, Forrest K Jones, Qulu Zheng, Hannah R Meredith, Andrew S Azman, Nicholas G Reich, and Justin Lessler. The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, 172(9):577–582, 2020.

[24] Abishek Chandrashekar, Jinyan Liu, Amanda J Martinot, Katherine McMahan, Noe B Mercado, Lauren Peter, Lisa H Tostanoski, Jingyou Yu, Zoltan Maliga, Michael Nekorchuk, et al. Sars-cov-2 infection protects against rechallenge in rhesus macaques. *Science*, 2020.

[25] Jingyou Yu, Lisa H Tostanoski, Lauren Peter, Noe B Mercado, Katherine McMahan, Shant H Mahrokhian, Joseph P Nkolola, Jinyan Liu, Zhenfeng Li, Abishek Chandrashekar, et al. Dna vaccine protection against sars-cov-2 in rhesus macaques. *Science*, 2020.

[26] Benny Borremans, Amandine Gamble, KC Prager, Sarah K Helman, Abby M McClain, Caitlin Cox, Van Savage, and James O Lloyd-Smith. Quantifying antibody kinetics and rna shedding during early-phase sars-cov-2 infection. *medRxiv*, 2020.

[27] Jeroen J.A. van Kampen, David A.M.C. van de Vijver, Pieter L.A. Fraaij, Bart L. Haagmans, Mart M. Lamers, Nisreen Okba, Johannes P.C. van den Akker, Henrik Endeman, Diederik A.M.P.J. Gommers, Jan J. Cornelissen, Rogier A.S. Hoek, Menno M. van der Eerden, Dennis A. Hesselink, Herold J. Metselaar, Annelies Verbon, Jurriaan E.M. de Steenwinkel, Georgina I. Aron, Eric C.M. van Gorp, Sander van Boheemen, Jolanda C. Voermans, Charles A.B. Boucher, Richard Molenkamp, Marion P.G. Koopmans, Corine Geurtsvankessel, and Annemiek A. van der Eijk. Shedding of infectious virus in hospitalized patients with coronavirus disease-2019 (covid-19): duration and key determinants. *medRxiv*, 2020.

[28] Hitoshi Kawasuji, Yusuke Takegoshi, Makito Kaneda, Akitoshi Ueno, Yuki Miyajima, Koyomi Kawago, Yasutaka Fukui, Yoshihiko Yoshida, Miyuki Kimura, Hiroshi Yamada, et al. Viral load dynamics in transmissible symptomatic patients with covid-19. *medRxiv*, 2020.

[29] Kelvin Kai-Wang To, Owen Tak-Yin Tsang, Wai-Shing Leung, Anthony Raymond Tam, Tak-Chiu Wu, David Christopher Lung, Cyril Chik-Yan Yip, Jian-Piao Cai, Jacky Man-Chun

Chan, Thomas Shiu-Hong Chik, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by sars-cov-2: an observational cohort study. *The Lancet Infectious Diseases*, 2020.

[30] Jin Yong Kim, Jae-Hoon Ko, Yeonjae Kim, Yae-Jean Kim, Jeong-Min Kim, Yoon-Seok Chung, Heui Man Kim, Myung-Guk Han, So Yeon Kim, and Bum Sik Chin. Viral load kinetics of sars-cov-2 infection in first two patients in korea. *Journal of Korean medical science*, 35(7), 2019.

[31] Ai Tang Xiao, Yi Xin Tong, and Sheng Zhang. Profile of rt-pcr for sars-cov-2: a preliminary study from 56 covid-19 patients. *Clinical Infectious Diseases*, 2020.

[32] Yang Liu, Li-Meng Yan, Lagen Wan, Tian-Xin Xiang, Aiping Le, Jia-Ming Liu, Malik Peiris, Leo LM Poon, and Wei Zhang. Viral dynamics in mild and severe cases of covid-19. *The Lancet Infectious Diseases*, 2020.

[33] Minnesota Population Center. Integrated public use microdata series, international: Version 7.2 [dataset], 2019. https://doi.org/10.18128/D020.V7.2.

[34] Yang Liu, Rosalind Eggo, and Adam Kucharski. Secondary attack rate and superspreading events for SARS-CoV-2. *The Lancet*, 2020.

[35] Kiesha Prem, Alex Cook, and Mark Jit. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Computational Biology*, 13(9):e1005697, 2017.

Supplemental Figures and Tables

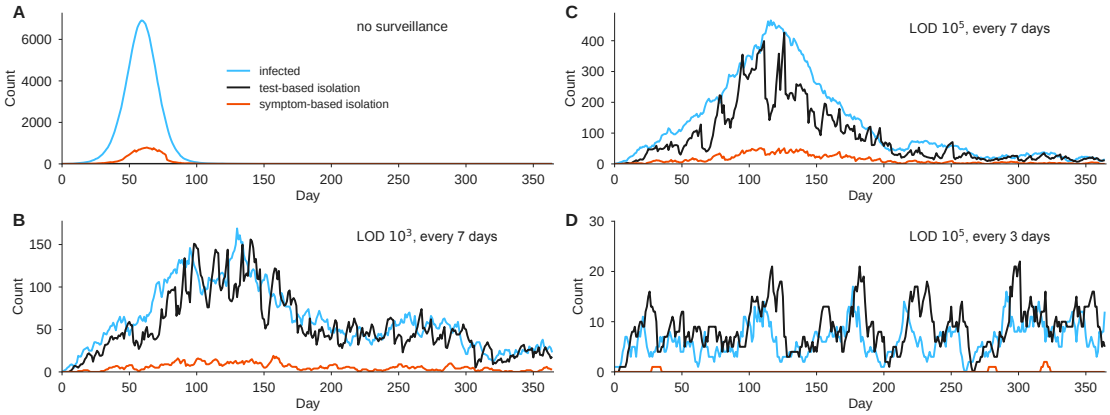

Figure S1: **Sample simulation trajectories from fully mixed model.** Simulation trajectories show the number of infected individuals in a population of $N = 20,000$ with a constant rate of external infection set to $1/N$ per person per day, i.e. around 1 imported case per day. Infections (blue), test-based isolation (black), and symptom-based isolation (red) are shown for four scenarios, with $R_0 = 2.5$. (A) No surveillance. (B) Weekly testing at LOD 10^3. (C) Weekly testing at LOD 10^5. (D) Testing every 3 days with LOD 10^5. Note the variation in the vertical axis scales. The model is fully described in Methods.

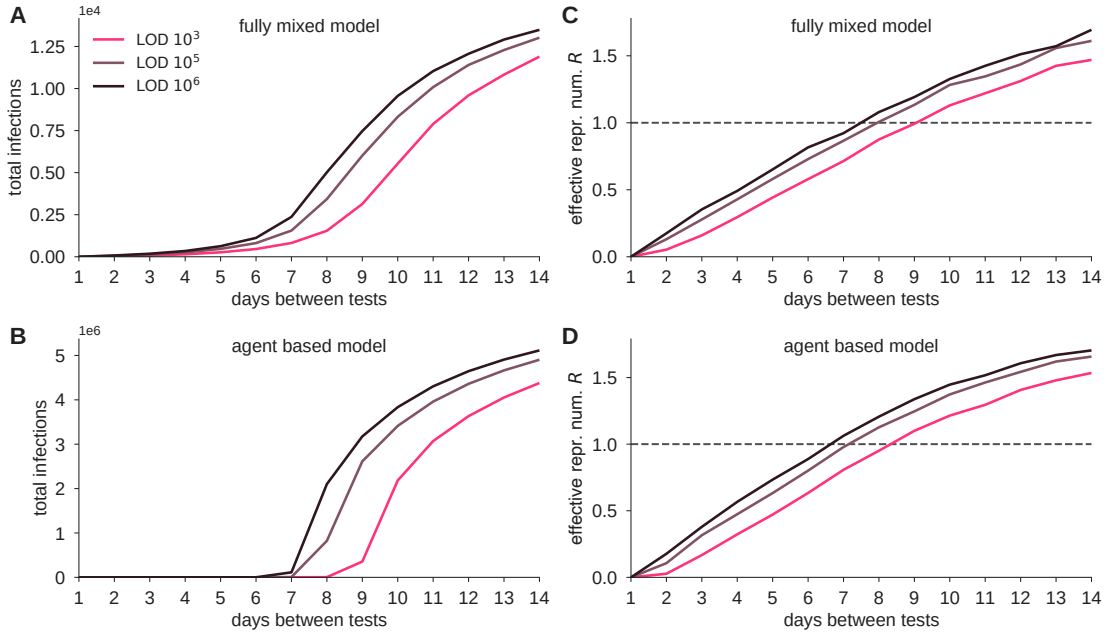

Figure S2: **Epidemiological model outcomes for various test LODs and frequencies.** The fully mixed model (top row) and agent based model (bottom row) were simulated (Methods) with various test frequencies, ranging from daily to once every 14 days, and with LODs of 10^3, 10^5, and 10^6. Modeling results show mean outcomes from 50 independent simulations at each point, expressed as (A, B) total infections and (C, D) effective reproductive number R, from a baseline of $R_0 = 2.5$. For the fully mixed model, only secondary infections are shown, excluding imported infections. Total population sizes were $N = 2 \times 10^4$ for the fully mixed model and 8.4×10^6 for the agent based model. Dashed lines indicate $R = 1$ for reference.

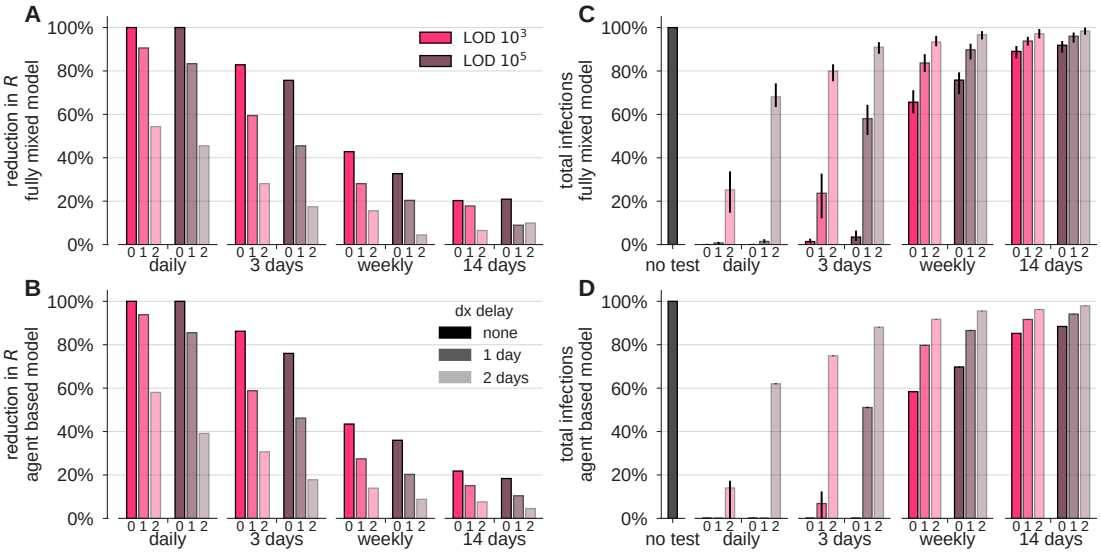

Figure S3: **Delays in reporting decrease the epidemiological impact of surveillance-driven isolation (proportional infectiousness).** This figure presents results from simulations which were identical to those shown in the main text Figure 4, but in which infectiousness was assumed to be *directly proportional* to viral load. Compare with threshold (binary) infectiousness in Fig. S4 and log-proportional infectiousness in Fig 4. See Methods.

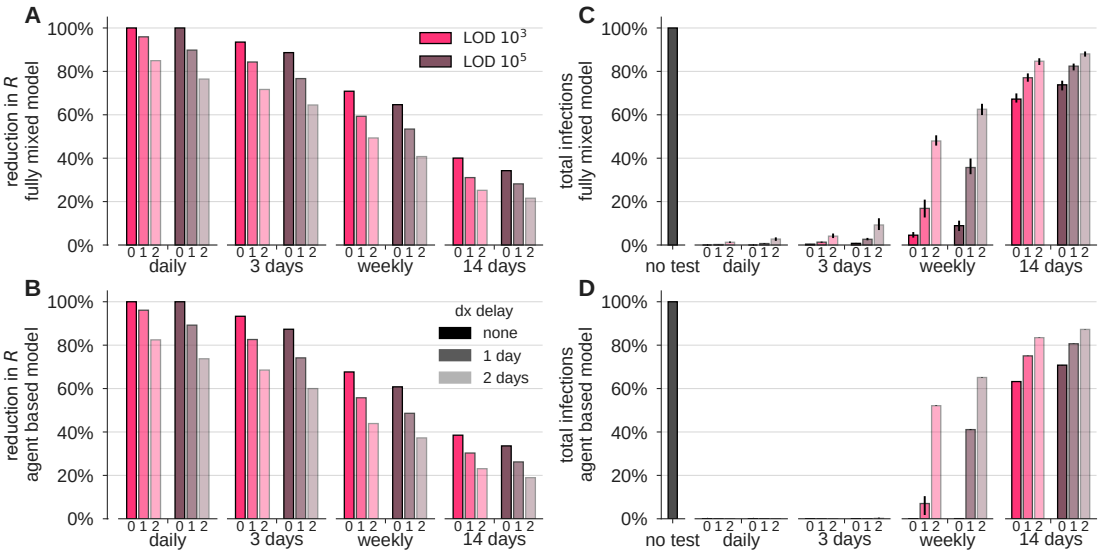

Figure S4: **Delays in reporting decrease the epidemiological impact of surveillance-driven isolation (threshold infectiousness).** This figure presents results from simulations which were identical to those shown in the main text Figure 4, but in which infectiousness was assumed to be *binary*, i.e. no infectiousness below 10^6 and equal infectiousness for any viral load above 10^6. Compare with proportional infectiousness in Fig. S3 and log-proportional infectiousness in Fig 4. See Methods.

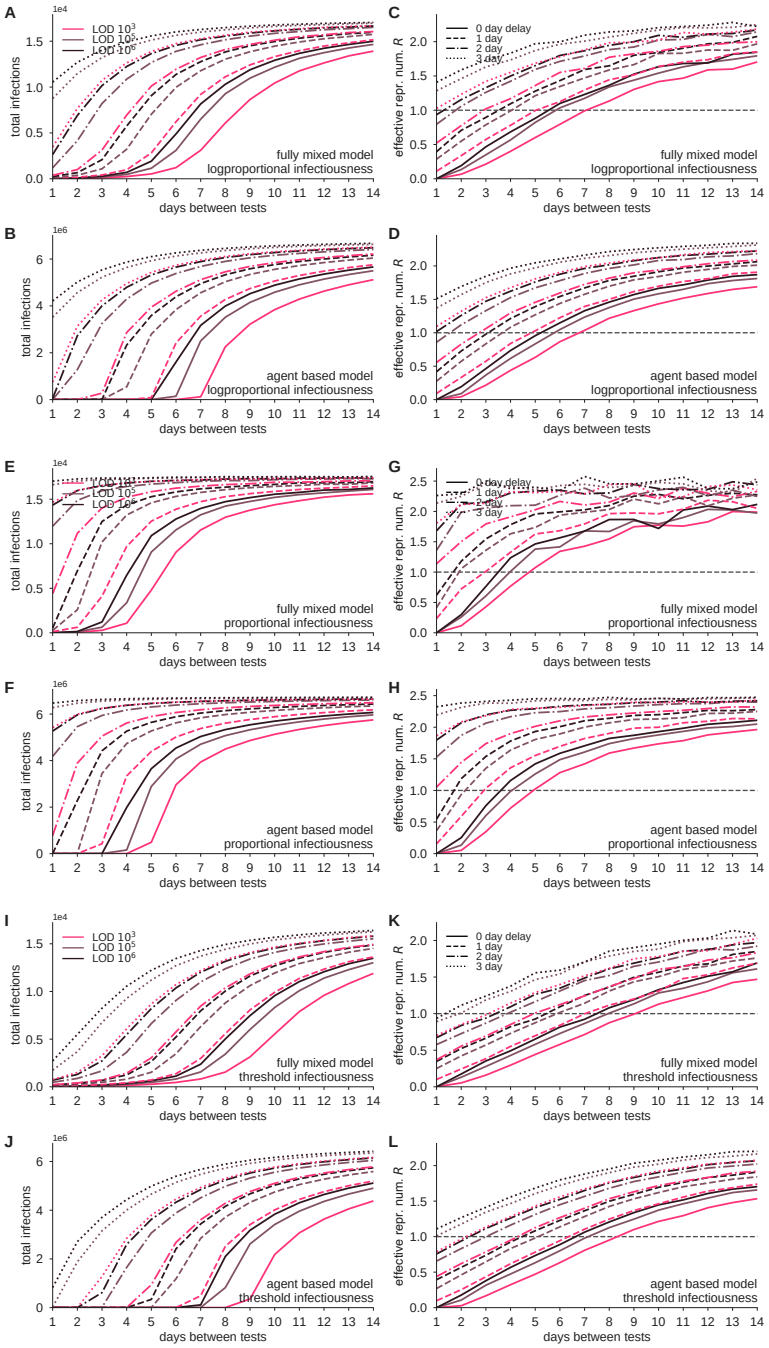

Figure S5: **Epidemiological model outcomes for various test LODs, frequencies, infectiousness functions, and with reporting delays.** The fully mixed model and agent based model were simulated (Methods) with various test frequencies, ranging from daily to once every 14 days, with LODs of 10^3, 10^5, and 10^6, and with delays of 0, 1, 2, or 3 days, for log-proportional, proportional, and threshold infectiousness functions (see Methods). Legends in panels A and B indicate LODs and delays, and in-plot annotations describe various conditions. Modeling results show mean outcomes from 50 independent simulations at each point, expressed as total infections and effective reproductive number R, from a baseline of $R_0 = 2.5$. For the fully mixed model, only secondary infections are shown, excluding imported infections. Total population sizes were $N = 2 \times 10^4$ for the fully mixed model and 8.4×10^6 for the agent based model. Dashed lines indicate $R = 1$ for reference.

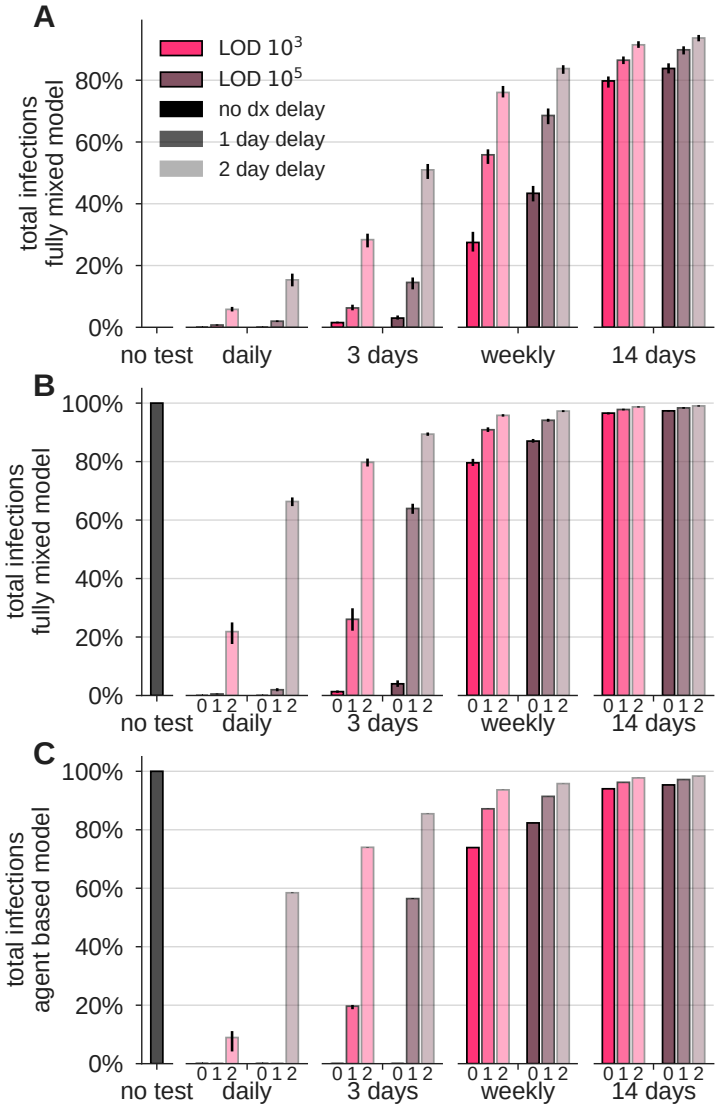

Figure S6: **Robustness of surveillance effectiveness to epidemiological model parameters.** (A) Results from the fully-mixed simulation with a tripled rate of external infection, i.e. $3/N$ per person per day. (B) Results from the fully mixed simulation with R_0 doubled, i.e. $R_0 = 5$. (C) Results from the agent-based simulation with R_0 doubled, i.e. $R_0 = 5$.

Methods

Viral Loads

Viral loads were drawn from a simple viral kinetics model intended to capture (1) a variable latent period, (2) a rapid growth phase from the lower limit of PCR detectability to a peak viral load, and (3) a slower decay phase. These dynamics were based on the following observations.

Latent periods prior to symptoms have been estimated to be around 5 day [23]. Viral load appears to peak prior to symptom onset [4], and peaks within 2 days of challenge in a macaque model [24, 25], though it should be noted that macaque challenge doses were high. Viral load decreases monotonically from the time of symptom onset [4, 26, 27, 28, 29], but may be high and detectable 3 or more days before symptom onset [1, 30]. Peak viral loads are difficult to measure due to lack of prospective sampling studies of individuals prior to exposure and infection, but viral loads have been reported in the range of $\mathcal{O}(10^4)$ to $\mathcal{O}(10^9)$ copies per ml [8, 28, 29]. Viral loads appear to become undetectable by PCR within 3 weeks of symptom onset [26, 29, 31], but detectability and timing may differ depending on the degree or presence of symptoms [31, 32]. Finally, we note that the general understanding of viral kinetics may vary depending on the mode of sampling, as demonstrated via a comparison between sputum and swab samples [8].

To mimic growth and decay, $\log_{10}$ viral loads were specified by a continuous piecewise linear "hinge" function, specified uniquely with three control points: $(t_0, 3)$, $(t_{\text{peak}}, V_{\text{peak}})$,$(t_f, 6)$ (Figure S7; green squares). The first point represents the time at which an individual's viral load first crosses 10^3, with $t_0 \sim \text{unif}[2.5, 3.5]$, measured in days since exposure. The second point represents the peak viral load. Peak height was drawn $V_{\text{peak}} \sim \text{unif}[7, 11]$, and peak timing was drawn with respect to the start of the exponential growth phase, $t_{\text{peak}} - t_0 \sim 0.2 + \text{gamma}(1.8)$. The third point represents the time at which an individual's viral load crosses beneath the 10^6 threshold, at which point viral loads no longer cause active cultures in laboratory experiments [], and was drawn with respect to peak timing, $t_f - t_{\text{peak}} \sim \text{unif}[5, 10]$. In simulations, each viral load's parameters were drawn independently of others, and the continuous function described here was evaluated at 21 integer time points (Figure S7; black dots) representing a three week span of viral load values.

Infectiousness

Infectiousness F was assumed to be directly related to viral load V in one of three ways. In the main text, each individual's relative infectiousness was proportional $\log_{10}$ of viral load's excess beyond 10^6, i.e. $F \propto \log_{10}(V) - 6$. In the supplementary sensitivity analyses, we investigated two opposing extremes. To capture a more extreme relationship between infectiousness and viral load, we considered F to be directly proportional to viral load's excess above 10^6, i.e. $F \propto 10^{\log_{10}(V)-6} = V \times 10^{-6}$, and to capture a more extreme relationship, but in the opposing direction, we considered F to simply be a constant when viral load exceeded 10^6, i.e. $F \propto \mathbf{1}_{V>10^6}$. We call these three functions *log-proportional*, *proportional*, and *threshold* throughout the text and

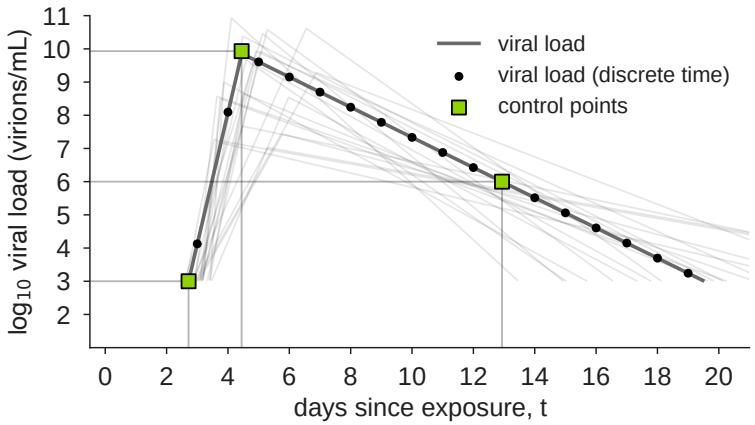

Figure S7: Example viral load (line) with stochastic control points highlighted (squares). Because simulations took place in discrete time, dots show points at which this example viral load would have been sampled. Light grey lines show 20 alternative trajectories to illustrate the diversity of viral loads drawn from the simple model.

supplemental materials.

Recently, He et al [4] published an analysis of infectiousness relative to symptom onset. Among our infectiousness functions, this inferred relationship bears the greatest similarity, over time, to the log-proportional infectiousness function, as visualized in Figs. 1 and 3. The proportional and threshold models therefore represent one of many types of sensitivity analysis. Results for those models can be found in Figures S3, S4, and S5.

In all simulations, the value of the proportionality constant implied by the infectiousness functions above was chosen to achieve the targeted value of R_0 for that simulation, and confirmed via simulation as described below.

Disease Transmission Models

Overview

Two models were used to simulate SARS-CoV-2 dynamics, both based on a typical compartmental framework. The first model was a fully-mixed model of $N = 20,000$ individuals with all-to-all contact structure, zero initial infections, and a constant $1/N$ per-person probability of becoming infected from an external source. This model could represent, for instance, a large college campus with high mixing, situated within a larger community with low-level disease prevalence. The second model was an agent-based model of $N = 8.4$ million agents representing the population and contact structure of New York City, as previously described [18]. Contact patterns were based on a combination of individual-level household contacts drawn from census microdata and age-stratified contact matrices which describe outside of household contacts. This model was initialized with 100 initial infections and no external sources of infection.

Both the fully-mixed and agent-based models tracked discrete individuals who were Susceptible (S), Infected (I), Recovered (R), Isolated (Q), and Self-Isolated (SQ) at each discrete one day timestep. Upon becoming infected ($S \rightarrow I$), a viral load trajectory $V(t)$ was drawn which included a latent period, growth, and decay. Each day, an individual's viral load trajectory was used to determine whether their diagnostic test would be positive if administered, as well as their infectiousness to susceptible individuals. Based on a schedule of testing each person every D days, if an individual happened to be tested on a day when their viral load exceeded the limit of detection L of the test, their positive result would cause them to isolate ($I \rightarrow Q$), but with the possibility of a delay in turnaround time. A fraction $1 - f$ of individuals self-isolate on the first day after peak viral load, to mimic symptom-driven isolation ($I \rightarrow SQ$), with $f = 0.8$ for the fully mixed model and $f = 1$ for the agent based model. When an individual's viral load dropped below 10^3, that individual recovered ($I, Q, SQ \rightarrow R$). Details follow.

Testing, Isolation, and Sample-to-Answer Turnaround Times

All individuals were tested every D days, so that they could be moved into isolation if their viral load exceeded the test's limit of detection $V(t) > L$. Each person was deterministically tested exactly every D days , but testing days were drawn uniformly at random such that not all individuals were tested on the same day. To account for delays in returning test results, we included a sample-to-answer turnaround time T, meaning that an individual with a positive test on day t would isolate on day $t + T$.

Transmission, Population Structure, and Mixing Patterns: Fully-mixed model

Simulations were initialized with all individuals susceptible, $S = N$. Each individual was chosen to be symptomatic independently with probability f, and each individual's first test day (e.g. the day of the week that their weekly test would occur) was chosen uniformly at random between 1 and D. Relative infectiousness was scaled up or down to achieve the specified R_0 in the absence of any testing policy, but inclusive of any assumed self-isolation of symptomatics.

In each timestep, those individuals who were marked for testing that day were tested, and a counter was initialized to T, specifying the number of days until that individual received their results. Next, individuals whose test results counters were zero were isolated, $I \rightarrow Q$. Then, symptomatic individuals whose viral load had declined relative to the previous day were self-isolated, $I \rightarrow SQ$. Next, each susceptible individual was spontaneously (externally) infected independently with probability $1/N$, $S \rightarrow I$. Then, all infected individuals contacted all susceptible individuals, with the probability of transmission based on that day's viral load $V(t)$ for each person and the particular infectiousness function, described above, $S \rightarrow I$.

To conclude each time step, individuals' viral loads and test results counters were advanced, with those whose infectious period had completely passed moved to recovery, $I, Q, SQ \rightarrow R$.

Transmission, Population Structure, and Mixing Patterns: Agent-based model

The agent-based model added viral kinetics and testing policies (as described above) to an existing model for SARS-CoV-2 transmission in New York City. A full description of the agent-based model is available [18]; here we provide an overview of the relevant transmission dynamics.

Simulations were initialized with all individuals susceptible, except for 100 initially infected individuals, $S = N - 100$. As in the fully-mixed model, each individual's test day was chosen uniformly at random and relative infectiousness was scaled to achieve the specified R_0.

In each timestep, those individuals who were marked for testing that day were tested, and a counter was initialized to T, specifying the number of days until that individual received their results. Next, individuals whose test results counters were zero were isolated, $I \to Q$. There was no self-isolation in this model (and accordingly, the model did not label individuals as symptomatic or asymptomatic).

Then, transmission from infected individuals to susceptible individuals was simulated both within and outside households. To model within-household transmission, each individual had a set of other individuals comprising their household. Household structures, along with the age of each individual, were sampled from census microdata for New York City [33]. The probability for an infectious individual to infect each of their household members each day was determined by scaling the relative infectiousness values to match the estimated secondary attack rate for close household contacts previously reported in case cluster studies [34].

Outside of household transmission was simulated using age-stratified contact matrices, which describe the expected number of daily contacts between an individual in a given age group and those in each other age group. Each infectious individual of age i drew Poisson(M_{ij}) contacts with individuals in age group j, where M is the contact matrix. The contacted individuals were sampled uniformly at random from age group j. We use a contact matrix for the United States estimated by [35]. Each contact resulted in infection, $S \to I$, with probability proportional to the relative infectiousness of the infected individual on that day, scaled to obtain the specified value of R_0.

To conclude each time step, individuals' viral loads and test results counters were advanced, with those whose infectious period had completely passed moved to recovery, $I, Q \to R$.

Calibration to achieve targeted R_0 and estimation of R

As a consistency check, each simulation's R_0 was estimated as follows, to ensure that simulations were properly calibrated to their intended values. Note that to vary R_0, the proportionality constant in the function that maps viral load to infectiousness need only be adjusted up or down. In a typical SEIR model, this would correspond to changing the infectiousness parameter which governs the rate at which I-to-S contacts cause new infections β.

For the fully-mixed, the value of R_0 was numerically estimated by running single-generation simulations in which a 50 infected individual were placed in a population of $N - 50$ others. The number of secondary infections from those initially infected was recorded and used to directly estimate R_0.

For the agent-based model, the value of R_0 depends on the distribution of infected agents due to stratification by age and household. We numerically estimate R_0 by averaging over the number of secondary infections caused by each agent who was infected in the first 15 days of the simulation (at which point the population is still more than 99.99% susceptible).

Estimations of R proceeded exactly as estimations of R_0 for both models, except with interventions applied to the the viral loads and therefore the dynamics.