Comparing the efficiency of forward and backward contact tracing

Juul, Jonas L; Strogatz, Steven H

Published in:
Physical Review E

Link to article, DOI:
10.1103/PhysRevE.108.034308

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Comparing the efficiency of forward and backward contact tracing

Jonas L. Juul [1,2,*] and Steven H. Strogatz [1]

[1]*Center for Applied Mathematics, Cornell University, Ithaca, New York 14853, USA*

[2]*Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby 2800, Denmark*

(Received 30 June 2022; revised 2 May 2023; accepted 28 June 2023; published 18 September 2023)

Tracing potentially infected contacts of confirmed cases is important when fighting outbreaks of many infectious diseases. The COVID-19 pandemic has motivated researchers to examine how different contact tracing strategies compare in terms of effectiveness (ability to mitigate infections) and cost efficiency (number of prevented infections per isolation). Two important strategies are so-called forward contact tracing (tracing to whom disease spreads) and backward contact tracing (tracing from whom disease spreads). Recently, Kojaku and colleagues reported that backward contact tracing was "profoundly more effective" than forward contact tracing, that contact tracing effectiveness "hinges on reaching the 'source' of infection," and that contact tracing outperformed case isolation in terms of cost efficiency. Here we show that these conclusions are not true in general. They were based in part on simulations that vastly overestimated the effectiveness and efficiency of contact tracing. Our results show that the efficiency of contact tracing strategies is highly contextual; faced with a disease outbreak, the disease dynamics determine whether tracing infection sources or new cases is more impactful. Our results also demonstrate the importance of simulating disease spread and mitigation measures in parallel rather than sequentially.

DOI: 10.1103/PhysRevE.108.034308

I. INTRODUCTION

To combat the spread of an infectious disease, contact tracing can be a useful mitigation tool. Contact tracing aims to identify infectious individuals through their social contacts; when an infected person has been identified, their close social contacts are traced and asked to test and possibly get treatment or quarantine. By proactively tracing and testing potentially infected people, public health officials can reduce the transmission of the infectious disease. For diseases such as SARS-CoV-2, where many transmissions take place before the infectious person develops symptoms [1–3], contact tracing is critical in reducing further spread [4–6].

During the last two decades, mathematical epidemiologists have sought to understand in what circumstances contact tracing is effective [7]. Among other things, researchers have investigated how the ability of contact tracing to curb the spread of disease is influenced by the network structure of social interactions [8–16], disease characteristics such as infectiousness profiles or case report statistics [17–21], different choices in how contact contact tracing is carried out [13,14,22], and specific outbreak scenarios (real or imagined) [23–25]. The findings of these studies have led to both concrete recommendations in case of bioterrorism [23], and more qualitative insights such as that contact tracing effectiveness can be affected by clustering and heterogeneity in population contact network structures [13,15].

The COVID-19 pandemic spurred a range of theoretical epidemiological investigations. Topics included vaccination prioritization strategies [26], the importance of accuracy and turnaround time in diagnostic tests [21,27,28], how standard confidence intervals can obscure extremes of epidemic projections (and what the alternatives are when presenting such projections to decision makers) [29], how schools and other institutions can be reopened safely following lockdowns [30–32], and how superspreading influences disease transmission and control measures [33–35]. Attention was also paid to contact tracing and its effectiveness. Arguably, the most influential theoretical contact tracing study during the pandemic was that authored by Kojaku *et al.* [36].

Kojaku *et al.* [36] investigated what makes contact tracing efficient in networked populations. They showed that backward contact tracing preferentially leads to high-degree infected individuals—superspreaders—with a sampling bias stronger than the celebrated friendship paradox [36]. The friendship paradox states that "your friends have on average more friends than you do" [37]; the reason being that (not considering possible degree correlations) following a random edge from a node leads to a degree-k node with probability proportional to kp_k, where p_k is the probability that a uniformly random node has degree k. The statistical arguments presented by Kojaku *et al.* [36] show that tracing backward in the transmission tree (a rooted, directed tree illustrating who infected whom) leads to degree k-nodes with probability proportional to $k^2 p_k$—a much stronger bias than that underlying the friendship paradox. This important result formalizes how effective backward tracing is at uncovering superspreaders.

In addition to the derivation of the backward-tracing sampling bias, Kojaku *et al.* [36] also used numerical simulations to make general conclusions about the effectiveness of different disease mitigation strategies. They devised a contact

*jlju@dtu.dk

tracing protocol that preferentially traced sources of infection. The contact tracing protocol is useful if only a limited number of nodes can be traced each timestep. Thus, Kojaku *et al.* demonstrate that contact tracing backward or forward can be a strategic choice and of high practical relevance if contact tracing resources are limited compared to the number of positive cases identified. One of their main conclusions was that "compared to 'forward' contact tracing (tracing to whom disease spreads), 'backward' contact tracing (tracing from whom disease spreads) is profoundly more effective." Indeed, this quote was the center of most coverage of the study in news outlets and on social media [38]. They also reported that the efficiency of contact tracing "hinges on reaching the source of infection" and that contact tracing beats case isolation in terms of cost efficiency. In this context, "case isolation" and "contact tracing" refer to two different mitigation strategies. "Case isolation" is the strategy where positive cases are isolated upon identification. "Contact tracing" is the strategy where positive cases are isolated upon identification and the contacts of the identified positive case are then proactively traced and tested in an effort to identify nearby active cases.

In this paper, we demonstrate that Kojaku *et al.*'s conclusions about the superiority of backward tracing over forward tracing and case isolation are not correct in general. Their conclusions rely on simulations of compartmental models on networks that systematically overestimate the efficiency and effectiveness of contact tracing. The overestimate stems mainly from the *sequential* nature of how Kojaku *et al.* [36] implement contact tracing and case isolation in their simulations: they first run their model epidemic to completion; only after that do they perform the model versions of case isolation and contact tracing. But in reality, these two processes occur contemporaneously rather than sequentially [39].

More precisely, Kojaku *et al.* [36] simulate epidemics involving four classes of people relative to the disease, namely those who are susceptible (S), exposed (E), infected (I), and recovered or removed (R). They begin by simulating unconstrained SEIR epidemics spreading on various network topologies. From these simulations, they then obtain transmission trees. Finally, contact tracing and case isolation are simulated using these transmission trees: If an infectious node is successfully identified and isolated, then Kojaku *et al.* [36] assume that all the nodes downstream in the transmission tree (the "descendants" that would have otherwise been infected) are thereby precluded from getting infected. But this assumption ignores the fact that a downstream node may well have other infected neighbors, each of which could potentially pass the infection on to it, even if the traced ancestor has been isolated.

The overestimation of contact tracing efficiency and effectiveness is likely to be large because descendant distributions are heavy-tailed [40]. In reality, the epidemic unfolds side-by-side with mitigation interventions like contact tracing. For this reason, simulating disease mitigation and epidemic spreading in parallel would provide a more accurate estimate of the efficacy of contact tracing.

In this paper, we examine contact tracing efficiency by simulating disease spread and mitigation measures in parallel rather than sequentially. This choice reduces the estimated efficiency of contact tracing by an order of magnitude compared to Kojaku *et al.*'s estimates [36], and demonstrates that backward contact tracing can be less efficient than forward contact tracing. Correcting another shortcoming of Kojaku *et al.*'s simulations—that they never release susceptible nodes that were traced and quarantined—further decreases contact tracing efficiency and demonstrates that case isolation can be more efficient than contact tracing.

II. SIMULATING EPIDEMICS AND MITIGATION STRATEGIES

We first simulate epidemics unfolding on a Barabási–Albert network [41] like the one used by Kojaku *et al.* [36]. The network has 250 000 nodes with average degree 4. Our simulation progresses in discrete time steps. We assume that each infected node first spends t_E days as preinfectious and then is infectious for t_I days. Each of these times is drawn from probability distributions of our choice. Unless otherwise stated, we draw t_E from an exponential distribution with mean 4 days and t_I from a lognormal distribution recently reported to resemble the incubation period of COVID-19 [42]. We note that our results are robust to changes in these distributions: excluding the exposed compartment entirely or drawing t_E and t_I from the same exponential distribution with mean 4 days yields similar results (see Appendix A).

In our model, we assume that an infectious node infects each susceptible neighbor with probability $p_I = q(t - t_0)\bar{p}_I$ at each time step. Here $\bar{p}_I$ is an average infectivity and $q(t - t_0)$ is a normalized function that expresses the relative likelihood of the node infecting a neighbor each day after time step t_0 when the node itself became infectious. We choose $\bar{p}_I = \frac{1.0}{(\bar{k}-1)}$, where $\bar{k}$ is the mean degree of the network. We choose this $\bar{p}_I$ value because this is the value that would result in a single secondary case on average if the disease were spreading unmitigated in a tree with the same mean degree (in this case, a newly infected node would have mean degree $\bar{k}$ on average and a single infected neighbor, so $\bar{k} - 1$ susceptible neighbors that would each get infected with independent probability $\bar{p}_I$).

In addition to the disease dynamics described above, we also simulate disease mitigation measures. Contrary to Kojaku *et al.*'s simulations of compartmental models and mitigation measures on networks, we simulate the mitigation measures in parallel with the disease spreading; This choice effectively makes contact tracing a dynamical process that competes for nodes with the spreading disease [11,22,43] Like Kojaku *et al.* [36], we assume that an infected node is identified with probability p_s at symptom onset and that each of its neighbors (infectious or not) is successfully traced with probability p_t. When a node is traced, we add it to a contact list with some weight w. To facilitate simulation of various contact tracing strategies, we treat w as adjustable. If $w = 1$ for all nodes, then the contact tracing scheme reduces to that of Kojaku *et al.* [36]; other choices of w allow us to implement backward contact tracing and forward contact tracing, as we discuss below.

We compare the impact of backward contact tracing and forward contact tracing by carrying out two kinds of simulations. In one, we simulate backward contact tracing by allowing only the direct source of infection to be traced ($w = w_{\text{parent}} = 1$ if a traced node is the direct source of infection, and $w = 1 - w_{\text{parent}} = 0$ otherwise). In the other, we simulate

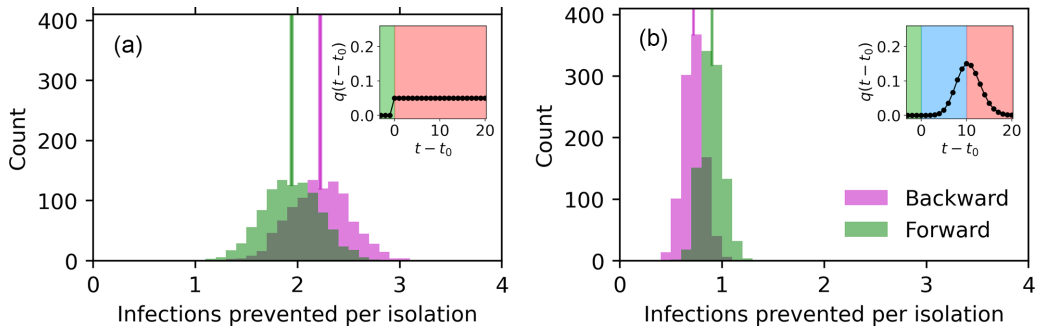

FIG. 1. Efficiency of contact tracing for two epidemic models with different patterns of infectiousness. Figure insets illustrate the assumed infectiousness pattern. Vertical values illustrate infectiousness on day $t - t_0$ after the node became infectious, and inset background color illustrates the disease state of a node on that day (Green: Exposed; Blue: Asymptomatic and infectious; Red: Symptomatic and infectious). In these illustrations, the infectious period lasts 20 days. We emphasize that we draw the infectious period from a probability distribution and that 20 days is a rare draw; we use a 20-day infectious period in this illustration because it makes the change of infectiousness as a function of time easy to see. After its infectious period, the node stops being infectious. In the main panels of the figure, we plot the number of infections that were prevented by two types of contact tracing for each person that was isolated in a simulation. Magenta histograms show results for an idealized backward contact tracing scheme in which only the direct source of infection (the "parent" node) of an identified infected node can be traced; in the green histograms, any neighbor except the direct source of the infection can be traced. We take the number of prevented infections to be the difference in nodes that got exposed to the disease when simulating the epidemic with parameters $p_s = 0$, $p_t = 0$ and $p_s = 0.05$, $p_t = 0.50$ (and otherwise identical initial conditions). The histograms show values obtained for 1 000 different simulations. The horizontal span of each of the vertical lines covers the interval between histogram mean value minus the error of this mean and histogram mean value plus the error of this mean value. (a) Constant infectiousness model. An infectious node is always symptomatic and infects each susceptible neighbor with equal probability on each of its infectious days. Backward contact tracing is the more efficient mitigation strategy, as shown by the relative positions of the magenta and green histograms. (b) Skewed infectiousness model. An infected node is asymptomatic during its first half of its infectious period and its infectiousness peaks around symptom onset. Backward contact tracing is the less efficient strategy here.

forward contact tracing by allowing any neighbor other than the direct source of infection to be traced ($w = w_{\text{parent}} = 0$ if a traced node is the direct source of infection, and $w = 1 - w_{\text{parent}} = 1$ otherwise). At each time step, after the above-described contact tracing has been carried out, we sum up the weights that each node is listed with in the contact list. Like Kojaku *et al.* [36], we then isolate the $n = 30$ nodes with the highest sum of weights in that list. We clear the contact list at the beginning of each time step.

III. RESULTS

We now present results indicating that (i) backward contact tracing is not necessarily more effective or efficient than forward contact tracing, (ii) contact tracing is not necessarily more efficient than case isolation and (iii) contact tracing efficiency can depend on network structure and in particular whether the network has many hubs or not. Our simulations also indicate that Kojaku *et al.*'s simulation choices cause their estimate of contact tracing efficiency to be an order of magnitude too high.

A. Comparing the efficiency of backward and forward tracing

To compare backward contact tracing to forward contact tracing, we simulate the mitigation of disease outbreaks with each of these contact tracing strategies. As detailed in the previous section, we implement backward contact tracing by setting $w_{\text{parent}} = 1$ and $w_{\text{nonparent}} = 0$ and forward contact tracing by setting $w_{\text{parent}} = 0$ and $w_{\text{nonparent}} = 1$.

For the simulation of each contact tracing strategy, we use the parameters $p_s = 0.05$ and $p_t = 0.50$, and simulate two different epidemic models: a "constant infectiousness" model and a "skewed infectiousness" model. The constant infectiousness model assumes that each infected node has a constant probability of passing the infection along to each of its susceptible neighbors during the days when the node is infectious and symptomatic [Fig. 1(a)]. In the skewed infectiousness model, inspired by known properties of COVID-19, each infected node is asymptomatic during the first half of its infectious period, and its infectiousness peaks around symptom onset [Fig. 1(b)]. Furthermore, for the skewed infectiousness model, we use an empirically estimated time-dependent transmissibility of COVID-19 for the function $q(t - t_0)$ [1,44].

Figures 1(a) and 1(b) compare the efficiency of both kinds of contact tracing (tracing only parents or no parents) for the two infectiousness models. For the constant infectiousness model [Fig. 1(a)], each isolation results in around 2 prevented infections (histogram mean is indicated by the center of the vertical line with the same color as the histogram bars; the horizontal span of the line marks the histogram mean plus and minus the standard error of the mean). For the skewed infectiousness model [Fig. 1(b)], each isolation results in around 1 prevented infection. These numbers are an order of magnitude smaller than the approximately 20 prevented infections per isolation estimated by Kojaku *et al.* [36] for a constant infectiousness model.

Another point to notice is the relative positions of the magenta and green histograms in Figs. 1(a) and 1(b). Backward contact tracing is the more efficient disease-mitigation

strategy in the constant infectiousness model [Fig. 1(a)], but the less efficient strategy in the skewed infectiousness model [Fig. 1(b)]. This reversal confirms the intuitive arguments made above.

The observations presented above for the efficiency of the two contact tracing schemes can also be made for the effectiveness of the strategies. For the constant infectiousness model, the mean number of infected in a simulation and error of this mean is $94\,304 \pm 22$ for forward contact tracing and $93\,454 \pm 24$ for backward contact tracing. This makes backward contact tracing the more effective strategy, just like it was the more efficient strategy. For the skewed infectiousness model, however, the number of infected in a simulation changes to $92\,851 \pm 23$ for forward contact tracing and $94\,357 \pm 22$ for backward contact tracing, making forward contact tracing the more effective strategy in this case.

B. Comparing case isolation and contact tracing

Kojaku *et al.* overestimate the efficiency of contact tracing by simulating disease spread and mitigation measures sequentially rather than in parallel. As we shall later demonstrate, this sequential implementation is what causes an overestimation of contact tracing efficiency by an order of magnitude. On top of that, another assumption incorrectly inflates Kojaku *et al.*'s estimate of contact tracing efficiency. Correcting that assumption reveals that case isolation can be more cost efficient than contact tracing, contrary to Kojaku *et al.*'s conclusion.

A key unrealistic assumption in Kojaku *et al.*'s [36] simulation of contact tracing and case isolation is that quarantined nodes remain in quarantine for all time—even if they are susceptible. This assumption, again, increases the estimated efficiency of contact tracing. If we relax this unrealistic assumption, and instead assume that quarantined nodes are released after 4 days if they are neither exposed nor infectious, then case isolation becomes more efficient than contact tracing in terms of the number of prevented infections per isolation. This result is shown in Fig. 2. Thus, Kojaku *et al.* [36] not only overestimated the numerical value of the estimates of contact tracing efficiency. They also overestimated the efficiency of contact tracing relative to other mitigation measures.

One might wonder whether releasing quarantined susceptible nodes changes the results presented in Fig. 1(a). Figure 3 demonstrates that this is not the case; releasing quarantined susceptible nodes does not change that backward tracing is more efficient than forward tracing, at least when simulating the constant infectiousness disease model on a Barabási–Albert network with mean degree 4. In Fig. 3, we plot the estimated number of infections prevented per isolation for different choices of contact tracing strategy. On the horizontal axis, we gradually decrease the weight that a direct source of infections is given in the contact list; from $w_{\text{parent}} = 1$ (and $w_{\text{nonparent}} = 1 - w_{\text{parent}} = 0$) at the leftmost point, to $w_{\text{parent}} = 0$ at the rightmost point. The background color changes from magenta to green as the contact tracing scheme changes from pure backward contact tracing to pure forward contact tracing tracing as w_{parent} decreases from 1 at the leftmost point of the plot to 0 at the rightmost point (magenta and green being the face colors of the histograms

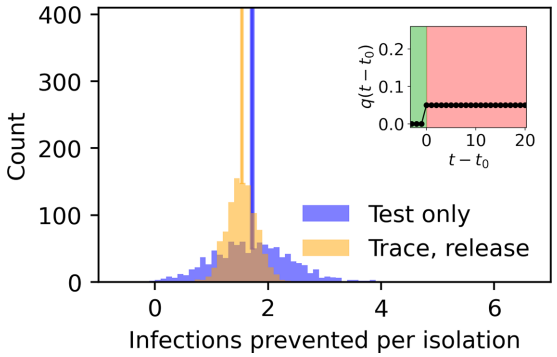

FIG. 2. Efficiency of case isolation vs contact tracing when susceptible quarantined nodes are eventually released. Figure inset illustrates the assumed infectiousness pattern. Vertical values illustrate infectiousness on day $t - t_0$ after the node became infectious, and inset background color illustrates the disease state of a node on that day (Green: Exposed; Red: Symptomatic and infectious). In the main panel of the figure, the yellow histogram plots the number of infections that were prevented by contact tracing for each person that was isolated in a simulation. We allow for both forward and backward contact tracing in this figure and release a traced and isolated node after 4 days of isolation if it is not infectious. The blue histogram plots the number of infections that were prevented by case isolation for each person that was isolated in a simulation. We take the number of prevented infections to be the difference in nodes that got infected by the disease when simulating the epidemic with parameters $p_s = 0$, $p_t = 0$ and $p_s = 0.05$, $p_t = x$ (where x is 0.50 for the yellow histogram and 0 for the blue histogram) (and otherwise identical initial conditions). The histograms show values obtained for 1 000 different simulations. In this case, the mean infections prevented per isolation is 1.72 ± 0.02 for case isolation (interval marked by the blue vertical line for the blue histogram) and 1.54 ± 0.01 (interval marked by the yellow vertical line for the yellow histogram), indicating that for these combinations of disease and mitigation measures, case isolation is more efficient than contact tracing.

in Fig. 1). For the Barabási-Albert network, the efficiency of contact tracing decreases monotonically as parent nodes are given less weight. Changing the network to a fully connected Erdős–Rényi, network with 245 046 nodes and a mean degree of 4.08 makes contact tracing efficiency monotonically increase as parent nodes are given less weight in the contact list. It is interesting that Fig. 3 shows that contact tracing efficiency decreases with the weight children are given in the tracing procedure for Barabási–Albert networks, whereas the contact tracing efficiency does not decrease as children are traced more frequently for Erdős–Rényi, networks. (A Mann-Kendall statistical test yields a probability of 0.06 that the observed increases in contact tracing efficiency with increased weight of children in the tracing procedure on Erdős–Rényi, networks could be due to random effects.) One possible explanation for this is the sampling bias derived by Kojaku *et al.*: Barabási–Albert networks have more hub structure and therefore includes many potential super spreaders. Tracing backwards can help identifying these super spreaders and prevent further spread from these hubs. For Erdős–Rényi, networks, however, super spreaders are much less likely to exist. So the backward tracing more rarely identifies very

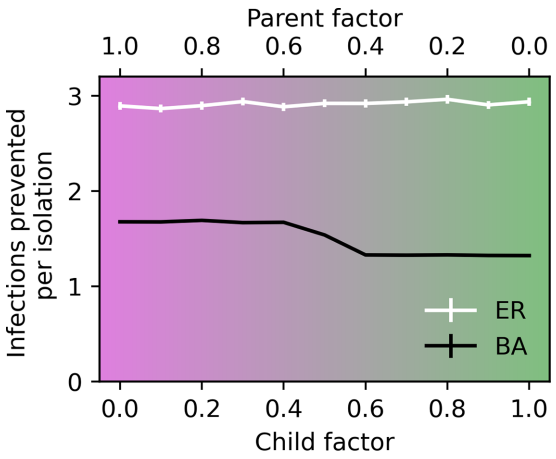

FIG. 3. Efficiency of contact tracing for combinations of backward and forward contact tracing. We simulate a constant infectiousness model on two different networks with similar mean degree: A Barabási–Albert network and an Erdős–Rényi, network. With parameters $p_s = 0.05$ and $p_t = 0.50$, we simulate different combinations of forward and backward tracing. We combine backward and forward tracing by varying a parameter w_{parent} ("parent factor" on the top horizontal axis). When a node is traced, we add it to the contact list with weight w_{parent} if it was the direct source of infection for the node it was traced through; otherwise, we add it to the contact list with weight $1 - w_{\text{parent}}$ ("Child factor" on the bottom horizontal axis). After each time step, we add the weights of each node in the contact list. We then isolate the 30 nodes with the highest sum of weights and clear the contact list before the next time step. For each choice of parent factor, we simulate 1 000 different outbreaks and show the mean infections prevented per isolation (and standard error of the mean) in these simulations as compared to simulations on the same networks, with no contact tracing and the same set of initially infected nodes. For the Barabási–Albert network, the contact tracing efficiency drops monotonically as forward contact tracing becomes a more important part of the contact tracing strategy. For the Erdős–Rényi, network, the curve is not decreasing.

high-degree nodes and therefore forward tracing may be more beneficial for these networks.

Figure 4 tests this hypothesis. We first use two networks consisting of approximately 250 000 nodes: The Erdős–Rényi, network and the Barabási–Albert network introduced above. From the Erdős–Rényi, network, we then create new networks by rewiring each edge in the network with identical probability p_r. If the edge connects nodes i and j in the Erdős–Rényi network and is selected for rewiring, then we rewire the edge to instead connect two nodes l and k. The nodes l and k are chosen with probability proportional to the degree of these nodes in the Barabási–Albert network. If either i or j has degree 2 or less, then the low-degree node keeps the edge, and we only rewire the other end. With this procedure, we can create networks with adjustable amount of hub structure: A higher p_r means more rewiring and more hub structure. We then simulate spread of a disease with constant infectiousness on networks created with p_r values 0, 0.1, 0.3, 0.5, and 1. Figure 4 shows that forward contact tracing is more efficient for the Erdős–Rényi, network, but is overtaken in efficiency by backward contact tracing as the networks get more hub structure. This supports the hypothesis

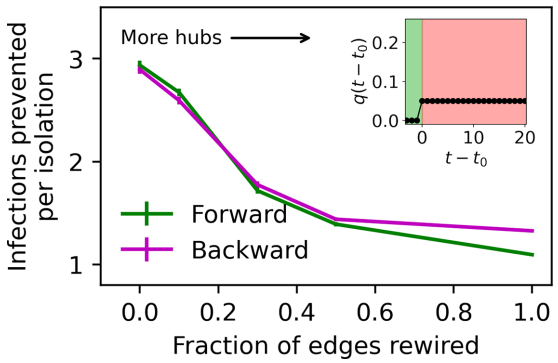

FIG. 4. Efficiency of backward and forward contact tracing on networks with adjustable amount of hub structure. We simulate a constant infectiousness model on five different networks created by rewiring each edge in an Erdős–Rényi, network with probability p_r (horizontal axis) to connect two other nodes each chosen proportional to their degree in a separate Barabási–Albert network. A higher p_r results in more hub structure. We create five networks in this way; on each, we simulate the constant infectiousness model with parameters $p_s = 0.05$ and $p_t = 0.50$ and release isolated susceptible nodes after 4 days. For each network, we simulate 1 000 different outbreaks and show the mean infections prevented per isolation (and standard error of the mean) in these simulations as compared to simulations on the same networks, with no contact tracing and the same set of initially infected nodes. As the networks get more hub structure (moving right on the horizontal axis), forward contact tracing goes from being the preferred contact tracing strategy to being the least favored contact tracing strategy.

that hub structure combined with the sampling bias in contact tracing explain the observations made in Fig. 3. In Appendices B and C, we provide results for simulations on bipartite people-gathering networks and branching processes with poisson and power-law degree distributions. In all cases, the infectiousness model heavily affects the efficiency of forward and backward contact tracing including their relative efficiency.

C. Replicating high estimates of contact-tracing efficiency

In the previous sections, we showed that Kojaku *et al.*'s [36] estimates of contact tracing efficiency are an order of magnitude higher than ours. Our epidemic model is different from Kojaku *et al.*'s [36] in a number of ways. For example, time progresses in discrete time steps in our model, and is continuous in Kojaku *et al.*'s [36]. Such differences can impact results significantly (see Ref. [45] for a demonstration of how offspring distributions are different for the susceptible-infected-recovered model and the Independent Cascade model). It is therefore natural to wonder whether the large difference in estimates of contact tracing efficiency could be due to such model differences, and not the reasons outlined above. To test this possibility, we simulated contact tracing like Kojaku *et al.* [36] did, by first running the epidemic to completion (using our constant infectiousness model and the Barabási–Albert network used above) and then implementing mitigation measures on the resulting transmission trees. We obtained 1 000 different simulated transmission trees and implemented simulated contact tracing 10 times

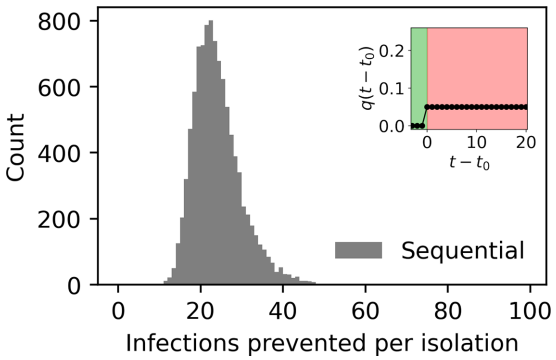

FIG. 5. Estimates of infections prevented per isolation when disease spread and mitigation measures are simulated sequentially rather than in parallel. We simulated an infectious disease spreading to completion 1 000 times, obtaining 1 000 transmission trees in the process. On each tree, we simulated case isolation and contact tracing 10 different times. Each time, we estimated the number of infections prevented per isolation. The figure shows the distribution of the 10 000 estimates. The mean value is 24.14 ± 0.06, a value similar to that obtained by Kojaku *et al.*, but an order magnitude larger than the more realistic estimates obtained when simulating disease spread and mitigation measures in parallel.

on each. This gave us 10 000 estimates of contact tracing efficiency using this sequential method.

Figure 5 plots the results in a histogram. The distribution has a thick right tail and the resulting mean estimate of contact tracing efficiency was 24.14 ± 0.06 infections prevented per isolation. This is of the same order of magnitude as Kojaku *et al.*'s estimate, and thus strongly supports the claim that the sequential nature of how Kojaku *et al.* [36] simulate contact tracing causes their estimate of contact tracing efficiency to be much too large.

IV. DISCUSSION

Contact tracing is a central component in mitigation strategies for many infectious diseases. Contact tracing can be used for different purposes, e.g., proactively finding cases to start their treatment early and prevent severe disease, or as a tool to reduce the number of secondary infections of newly infected people. Here, we have focused on contact tracing as a tool for reducing the number of secondary infections. A better understanding of how different contact tracing strategies compare in terms of efficiency could translate into saved lives and decreased economic losses when faced with an epidemic. In this paper, we investigated recent influential claims that backward contact tracing is "profoundly more effective" than forward contact tracing, that contact tracing effectiveness "hinges on reaching the 'source' of infection" and that contact tracing beats case isolation in terms of cost efficiency [36].

By correcting shortcomings in how disease spread and mitigation measures were simulated in the study [36], we showed that the above-mentioned findings do not hold up. We conclude that contact tracing is not necessarily more cost efficient than case isolation, and that whether backward tracing or forward tracing is superior depends on the disease in question and the structure of the underlying contact network. For COVID-19-like disease dynamics simulated on several kinds of contact networks (Barabási–Albert, Erdős–Rényi, and people-gathering networks), we found that backward tracing could actually be the less efficient strategy. Even so, backward tracing could still be valuable as a means to uncover new branches of the transmission tree that could then be forward traced. How efficient this strategy would be remains an open question.

We have simulated an idealized model for disease spread and mitigation on static Barabási–Albert networks. The idealized nature of both the choices of models and networks means that simulations could be made more realistic by choosing empirically observed network structures, rather than the synthetic Barabási–Albert networks, or by choosing disease models and parameters more carefully to match those of the COVID-19 pandemic. We stress, however, that the point of this work is not to present a realistic simulation of disease spread and disease mitigation. Instead, our main contribution is to show that recent conclusions about the superiority of backward tracing to forward tracing and case isolation are flawed and to demonstrate that the flaw stems not from unrealistic models of disease spread or mitigation measures, but from the sequential nature of Kojaku *et al.*'s [36] simulations. These sequential simulations of disease spread and mitigation efforts cause both the quantitative estimates of backward tracing efficiency and the qualitative results comparing the efficiency of different mitigation strategies to be unreliable.

One of the things we have shown is that disease characteristics such as the infectiousness profile can greatly impact the effectiveness and relative effectiveness of forward contact tracing and backward tracing. A related effect that can have similar impact is waiting times for testing or contact tracing. A delay in the testing or contact tracing will cause cases to be identified later in their infectious period. Depending on the length of the delay and incubation times, this could again make backward contact tracing less favorable as compared to forward contact tracing. However, very fast testing and contact tracing could make it possible to ring-fence active cases with contact tracing [21,46,47]. For discussions about test turnaround times and contact tracing strategies see Refs. [21,46].

There are many interesting directions for future research on the efficiency of contact tracing strategies. As already mentioned, one direction is to investigate mixed strategies of strategically combined backward and forward contact tracing. Using a branching process framework Endo *et al.* [48] recently showed that a combination of backward and forward contact tracing can curb the spread of diseases with overdispersion better than forward tracing alone. Among other things, their analysis did not investigate the effect of time-varying infectiousness on contact tracing effectiveness, and many interesting questions remain. How should resources for backward and forward contact tracing be allocated when faced with a specific disease? How many forward tracings should be carried out for each person that was successfully backward traced? Under what circumstances is backward contact tracing so inefficient that it should not be prioritized explicitly? Questions like these abound.

Another interesting direction would be to estimate contact tracing efficiency at different stages of an epidemic. Contact

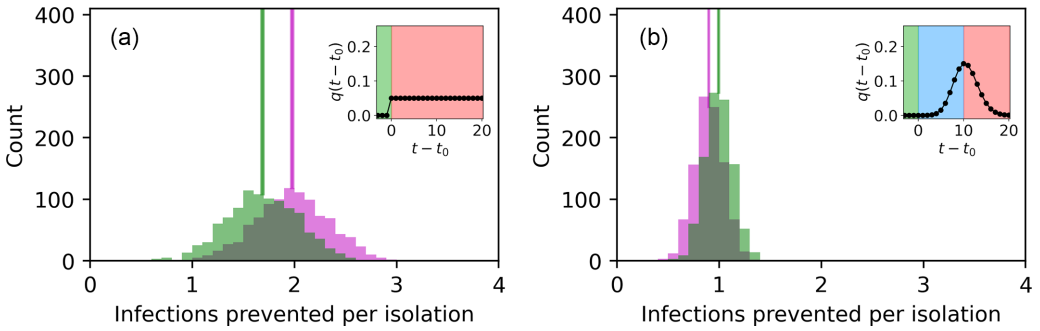

FIG. 6. Efficiency of contact tracing for two epidemic models with different patterns of infectiousness and identical distributions for t_E and t_I simulated on Barabási–Albert networks. Figure insets illustrate the assumed infectiousness pattern. Vertical values illustrate infectiousness on day $t - t_0$ after the node became infectious, and inset background color illustrates the disease state of a node on that day (Green: Exposed; Blue: Asymptomatic and infectious; Red: Symptomatic and infectious). In the main panels of the figure, we plot the number of infections that were prevented by two types of contact tracing for each person that was isolated in a simulation. Magenta histograms show results for an idealized backward contact tracing scheme in which only the direct source of infection (the "parent" node) of an identified infected node can be traced; in the green histograms, any neighbor except the direct source of the infection can be traced. We take the number of prevented infections to be the difference in nodes that got exposed to the disease when simulating the epidemic with parameters $p_s = 0$, $p_t = 0$ and $p_s = 0.05$, $p_t = 0.50$ (and otherwise identical initial conditions). We draw t_E and t_I from exponential distributions with mean 4 days. The histograms show values obtained for 1 000 different simulations. The horizontal span of each of the vertical lines covers the interval between histogram mean value minus the error of this mean and histogram mean value plus the error of this mean value. (a) Constant infectiousness model. An infectious node is always symptomatic and infects each susceptible neighbor with equal probability on each of its infectious days. Backward contact tracing is the more efficient mitigation strategy, as shown by the relative positions of the magenta and green histograms. (b) Skewed infectiousness model. An infected node is asymptomatic during its first half of its infectious period and its infectiousness peaks around symptom onset. Backward contact tracing is the less efficient strategy here.

tracing is of course most efficient in the very beginning of an epidemic, when there is still hope that the spreading can be stopped altogether. But how fast does the expected number of prevented cases per isolation drop as the epidemic infects more and more people? Is there a time when authorities should prioritize carrying out more contact tracing and another where resources are better spent increasing population screening capabilities? Answers to such questions would be useful for decision makers next time the world is faced with a pandemic.

The code and data necessary to reproduce our results are available at [49].

ACKNOWLEDGMENTS

J.L.J.'s work presented here is supported by the Carlsberg Foundation, Grant No. CF21-0342.

APPENDIX A: EXPONENTIAL DISTRIBUTIONS FOR t_E AND t_I

In simulating the spread of diseases, we had to make choices for the distributions of t_E—the time that an exposed node spends in the the exposed compartment—and t_I—the time an infectious node stays infectious. In the main text, we assumed that t_E was exponentially distributed and t_I lognormally distributed. Our results are robust to changing these distributions. Figure 6 shows the obtained efficiency of backward and contact tracing when drawing t_E and t_I from the same exponential distribution with mean 4 days. In Fig. 6(a), we simulate a constant infectiousness model. In this case, backward contact tracing is the more efficient strategy, as was the case in Fig. 1(a) in the main text. In Fig. 6(b), we simulate the skewed infectiousness model. In this case, forward contact tracing is the more efficient strategy. Again, this is in line with our results in Fig. 1(b), where t_I was drawn from a lognormal distribution.

APPENDIX B: PEOPLE-GATHERING NETWORKS

Figure 7 shows simulations of epidemics and contact tracing strategies on bipartite people-gathering networks. The degrees of people and gatherings are both drawn from power-law distributions with exponent -3 (gatherings having minimum degree 2, whereas people have minimum degree 1). Because it is unlikely that two random draws from probability distributions result in the same degree sums, we make the degree sums equal by increasing the degree of one node until degree sums are equal for people and gatherings. The stubs of people and gatherings are then paired uniformly randomly.

Figure 7 illustrates the impact of changing the infectiousness model when simulating an epidemic in the people-gathering networks. For the constant-infectiousness model backward contact tracing is the most efficient [Fig. 7(a)], whereas forward contact tracing is the most efficient for the empirically informed skewed-infectiousness model [Fig. 7(b)].

APPENDIX C: BRANCHING PROCESSES

A popular way of studying infectious diseases in populations is branching processes. One reason for this is that branching processes often allows for analytical or numerical results in infinite populations. Here, we numerically analyze contact tracing efficiency in a branching process context. We first create a branching process network starting from N_{seeds}

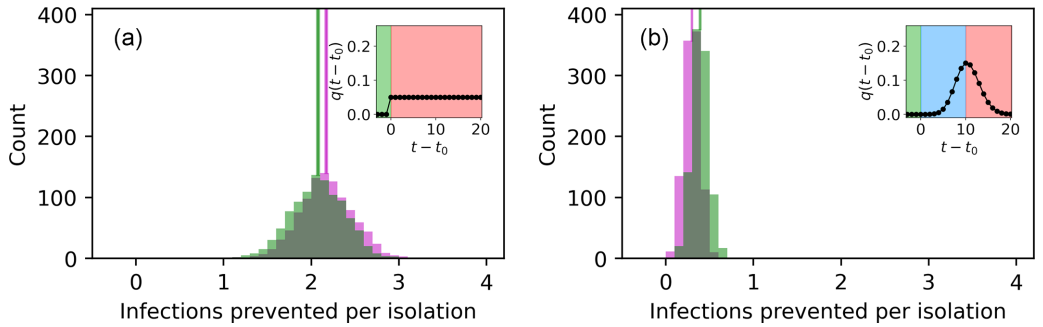

FIG. 7. Efficiency of contact tracing for two epidemic models with different patterns of infectiousness simulated on bipartite people-gathering networks. Figure insets illustrate the assumed infectiousness pattern. Vertical values illustrate infectiousness on day $t - t_0$ after the node became infectious, and inset background color illustrates the disease state of a node on that day (Green: Exposed; Blue: Asymptomatic and infectious; Red: Symptomatic and infectious). In the main panels of the figure, we plot the number of infections that were prevented by two types of contact tracing for each person that was isolated in a simulation. Magenta histograms show results for an idealized backward contact tracing scheme in which only the direct source of infection (the 'parent' node) of an identified infected node can be traced; in the green histograms, any neighbor except the direct source of the infection can be traced. We take the number of prevented infections to be the difference in nodes that got exposed to the disease when simulating the epidemic with parameters $p_s = 0$, $p_t = 0$ and $p_s = 0.05$, $p_t = 0.50$ (and otherwise identical initial conditions). The histograms show values obtained for 1 000 different simulations. The horizontal span of each of the vertical lines covers the interval between histogram mean value minus the error of this mean and histogram mean value plus the error of this mean value. (a) Constant infectiousness model. An infectious node is always symptomatic and infects each susceptible neighbor with equal probability on each of its infectious days. Backward contact tracing is the more efficient mitigation strategy, as shown by the relative positions of the magenta and green histograms. (b) Skewed infectiousness model. An infected node is asymptomatic during its first half of its infectious period and its infectiousness peaks around symptom onset. Backward contact tracing is the less efficient strategy here.

isolated generation-zero nodes. For each of these seeds, we draw their out-degree, k, from an out-degree distribution $p(k)$, and connect this many generation-one nodes to the seed in question. For each generation-one node, we then draw an out-degree from $p(k)$, connect it to that number of generation-two nodes, and so on. We continue the process until node number $N_{\max}$ has been drawn. The result is a network consisting of N_{seeds} different trees with at most $g_{\max}$ generations of nodes where all but some nodes in the final two generations have out-degree distribution $p(k)$.

Having created a branching-process network, we can now simulate epidemics unfolding on top of these using the proce-

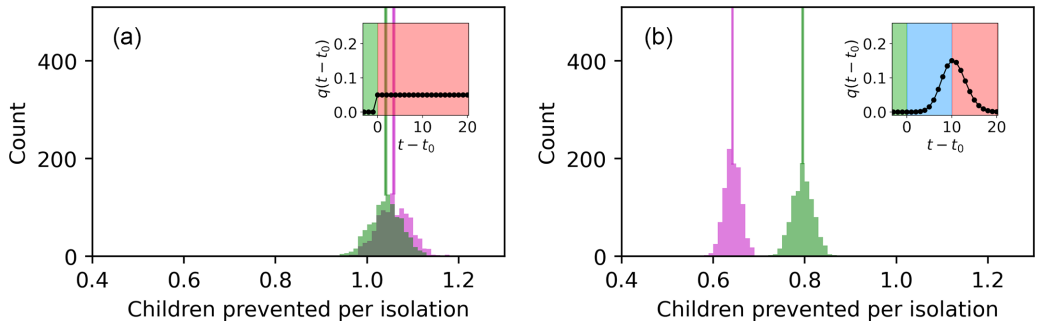

FIG. 8. Efficiency of contact tracing for two epidemic models with different patterns of infectiousness simulated on branching process networks with poisson degree distributions. Figure insets illustrate the assumed infectiousness pattern. Vertical values illustrate infectiousness on day $t - t_0$ after the node became infectious, and inset background color illustrates the disease state of a node on that day (Green: Exposed; Blue: Asymptomatic and infectious; Red: Symptomatic and infectious). In the main panels of the figure, we plot the number of infections that were prevented by two types of contact tracing for each person that was isolated in a simulation. Magenta histograms show results for an idealized backward contact tracing scheme in which only the direct source of infection (the 'parent' node) of an identified infected node can be traced; in the green histograms, any neighbor except the direct source of the infection can be traced. We take the number of prevented children to be the decrease in direct neighbors of infected nodes that got infected in simulations where identified infected were isolated as compared to counter factual simulations where the identified infected were not isolated. We simulate the epidemic with parameters $p_s = 0$, $p_t = 0$ and $p_s = 0.05$, $p_t = 0.50$ (and otherwise identical initial conditions). The histograms show values obtained for 1 000 different simulations. The horizontal span of each of the vertical lines covers the interval between histogram mean value minus the error of this mean and histogram mean value plus the error of this mean value. (a) Constant infectiousness model. An infectious node is always symptomatic and infects each susceptible neighbor with equal probability on each of its infectious days. Backward contact tracing is the more efficient mitigation strategy, as shown by the relative positions of the magenta and green histograms. (b) Skewed infectiousness model. An infected node is asymptomatic during its first half of its infectious period and its infectiousness peaks around symptom onset. Backward contact tracing is the less efficient strategy here.

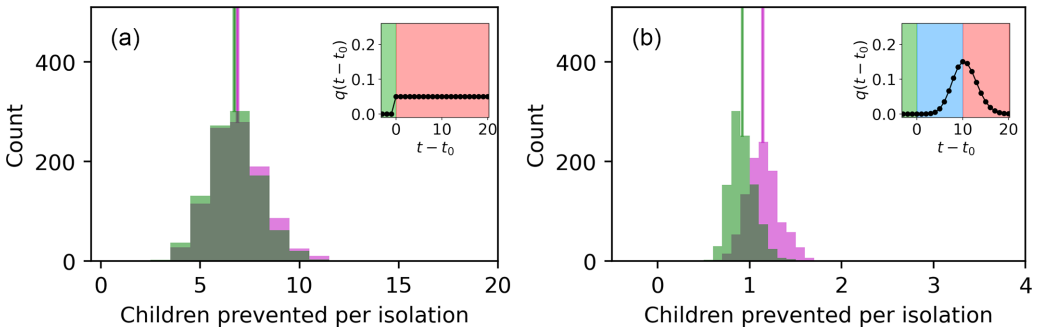

FIG. 9. Efficiency of contact tracing for two epidemic models with different patterns of infectiousness simulated on branching process networks with power law degree distributions. Figure insets illustrate the assumed infectiousness pattern. Vertical values illustrate infectiousness on day $t - t_0$ after the node became infectious, and inset background color illustrates the disease state of a node on that day (Green: Exposed; Blue: Asymptomatic and infectious; Red: Symptomatic and infectious). In the main panels of the figure, we plot the number of infections that were prevented by two types of contact tracing for each person that was isolated in a simulation. Magenta histograms show results for an idealized backward contact tracing scheme in which only the direct source of infection (the "parent" node) of an identified infected node can be traced; in the green histograms, any neighbor except the direct source of the infection can be traced. We take the number of prevented children to be the decrease in direct neighbors of infected nodes that got infected in simulations where identified infected were isolated as compared to counter factual simulations where the identified infected were not isolated. We simulate the epidemic with parameters $p_s = 0$, $p_t = 0$ and $p_s = 0.05$, $p_t = 0.50$ (and otherwise identical initial conditions). The histograms show values obtained for 1 000 different simulations. The horizontal span of each of the vertical lines covers the interval between histogram mean value minus the error of this mean and histogram mean value plus the error of this mean value. (a) Constant infectiousness model. An infectious node is always symptomatic and infects each susceptible neighbor with equal probability on each of its infectious days. Backward contact tracing is the more efficient mitigation strategy, as shown by the relative positions of the magenta and green histograms. (b) Skewed infectiousness model. An infected node is asymptomatic during its first half of its infectious period and its infectiousness peaks around symptom onset. Backward contact tracing is the less efficient strategy here.

dure described in the main text. We use all the generation-zero nodes as seeds, define an infectiousness profile, an average infectiousness $\bar{p}_I$, the two contact tracing parameters p_s and p_t and a contact tracing strategy, and simulate 1 000 different epidemics with these settings. To evaluate the efficiency of forward and backward tracing, we need another metric than number of cases avoided per isolate. This is because the branching process framework ideally mimics an infinite-size system, and descendant distributions are often heavy-tailed and can even have a diverging mean [40]. For this reason, we use another metric: The number of prevented *children* per isolation. To quantify the number of prevented children (secondary cases), for each isolated node in a simulation we run a counterfactual scenario where the node was not isolated. In this counterfactual scenario, the node's course of disease continues until recovery, and it gets the same chances to infect its neighbors as it would have if the infectious node was not discovered or isolated. We count the number of neighbors the node successfully infects in this counterfactual scenario; this number, we take to be the number of children prevented by the isolation of the node. We note that this is an approximation – the node could have been traced and isolated by another neighbor at a later point. We also note that we only run counterfactual scenarios for nodes in generations 0 to $g_{\max} - 2$, since these have out-degrees described by the distribution $p(k)$.

For all simulations, we choose $N_{\text{seeds}} = 250$, $N_{\max} = 1\,000\,000$, $p_s = 0.05$, and $p_t = 0.50$. We note that the branching process networks are trees, and so do not at all resemble actual human contact networks which contain, e.g., nontrivial clustering [50]. Figure 8 shows results obtained when simulating disease spread on a branching-process network with a Poisson out-degree distribution $p(k)$ with mean 4. For the constant-infectiousness model, backward tracing is the most efficient [Fig. 8(a)], whereas forward contact tracing is much to be preferred in the skewed-infectiousness model [Fig. 8(b)]. Figure 9 shows corresponding results for simulations on branching-process networks with a power-law out-degree distribution (with exponent -2.1015 and maximum-allowed degree of 1000, resulting in mean degree 4). These trees are very different from the Barabási–Albert networks, which contain considerable clustering and more complicated network structure. For simulations on these trees, backward contact tracing is the more efficient contact tracing strategy for both the constant-infectiousness [Fig. 9(a)] and skewed-infectiousness model [Fig. 9(b)].

[1] X. He, E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang *et al.*, Temporal dynamics in viral shedding and transmissibility of COVID-19, Nat. Med. **26**, 672 (2020).

[2] D. P. Oran and E. J. Topol, Prevalence of asymptomatic SARS-CoV-2 Infection: A narrative review, Ann. Intern. Med. **173**, 362 (2020).

[3] M. M. Arons, K. M. Hatfield, S. C. Reddy, A. Kimball, A. James, J. R. Jacobs, J. Taylor, K. Spicer, A. C. Bardossy, L. P.

Oakley, S. Tanwar, J. W. Dyal, J. Harney, Z. Chisty, J. M. Bell, M. Methner, P. Paul, C. M. Carlson, H. P. McLaughlin, N. Thornburg *et al.*, Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility, N. Engl. J. Med. **382**, 2081 (2020).

[4] S. M. Moghadas, M. C. Fitzpatrick, P. Sah, A. Pandey, A. Shoukat, B. H. Singer, and A. P. Galvani, The implications of silent transmission for the control of COVID-19 outbreaks, Proc. Natl. Acad. Sci. USA **117**, 17513 (2020).

[5] N. C. Grassly, M. Pons-Salort, E. P. K. Parker, P. J. White, N. M. Ferguson, K. Ainslie, M. Baguelin, S. Bhatt, A. Boonyasiri, N. Brazeau, L. Cattarino, H. Coupland, Z. Cucunuba, G. Cuomo-Dannenburg, A. Dighe, C. Donnelly, S. L. V. Elsland, R. FitzJohn, S. Flaxman, K. Fraser *et al.*, Comparison of molecular testing strategies for COVID-19 control: A mathematical modelling study, Lancet Infect. Dis. **20**, 1381 (2020).

[6] Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19), Technical report, World Health Organization (2020), https://www.who.int/publications/i/item/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19).

[7] J. Müller and M. Kretzschmar, Contact tracing—Old models and new challenges, Infect. Disease Model. **6**, 222 (2021).

[8] R. Huerta and L. S. Tsimring, Contact tracing and epidemics control in social networks, Phys. Rev. E **66**, 056115 (2002).

[9] K. T. D. Eames and M. J. Keeling, Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases, Proc. Natl. Acad. Sci. USA **99**, 13330 (2002).

[10] K. T. D. Eames and M. J. Keeling, Contact tracing and disease control, Proc. R. Soc. London B **270**, 2565 (2003).

[11] I. Z. Kiss, D. M. Green, and R. R. Kao, Disease contact tracing in random and clustered networks, Proc. R. Soc. Ser. B Biol. Sci. **272**, 1407 (2005).

[12] I. Z. Kiss, D. M. Green, and R. R. Kao, Infectious disease control using contact tracing in random and scale-free networks, J. R. Soc. Interface. **3**, 55 (2006).

[13] K. T. D. Eames, Contact tracing strategies in heterogeneous populations, Epidemiol. Infect. **135**, 443 (2007).

[14] I. Z. Kiss, D. M. Green, and R. R. Kao, The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing, J. R. Soc. Interface. **5**, 791 (2008).

[15] T. House and M. J. Keeling, The impact of contact tracing in clustered populations, PLoS Comput. Biol. **6**, e1000721 (2010).

[16] A. Okolie and J. Müller, Exact and approximate formulas for contact tracing on random trees, Math. Biosci. **321**, 108320 (2020).

[17] D. Klinkenberg, C. Fraser, and H. Heesterbeek, The effectiveness of contact tracing in emerging epidemics, PLoS ONE **1**, e12 (2006).

[18] C. Browne, H. Gulbudak, and G. Webb, Modeling contact tracing in outbreaks with application to Ebola, J. Theor. Biol. **384**, 33 (2015).

[19] C. M. Peak, L. M. Childs, Y. H. Grad, and C. O. Buckee, Comparing nonpharmaceutical interventions for containing emerging epidemics, Proc. Natl. Acad. Sci. USA **114**, 4023 (2017).

[20] L. Baumgarten and S. Bornholdt, Epidemics with asymptomatic transmission: Subcritical phase from recursive contact tracing, Phys. Rev. E **104**, 054310 (2021).

[21] J. L. Juul and K. Græsbøll, Are fast test results preferable to high test sensitivity in contact-tracing strategies? medRxiv 2021.02.17.21251921, doi:10.1101/2021.02.17.21251921.

[22] M. Meister and J. Kleinberg, Optimizing the order of actions in contact tracing, PNAS Nexus, **2**, pgad003 (2023).

[23] E. H. Kaplan, D. L. Craft, and L. M. Wein, Emergency response to a smallpox attack: The case for mass vaccination, Proc. Natl. Acad. Sci. USA **99**, 10935 (2002).

[24] K. T. D. Eames, C. Webb, K. Thomas, J. Smith, R. Salmon, and J. M. F. Temple, Assessing the role of contact tracing in a suspected H7N2 influenza A outbreak in humans in Wales, BMC Infect. Diseases **10**, 141 (2010).

[25] K. C. Swanson, C. Altare, C. S. Wesseh, T. Nyenswah, T. Ahmed, N. Eyal, E. L. Hamblion, J. Lessler, D. H. Peters, and M. Altmann, Contact tracing performance during the Ebola epidemic in Liberia, 2014-2015, PLoS Negl. Trop. Dis. **12**, e0006762 (2018).

[26] K. M. Bubar, K. Reinholt, S. M. Kissler, M. Lipsitch, S. Cobey, Y. H. Grad, and D. B. Larremore, Model-informed COVID-19 vaccine prioritization strategies by age and serostatus, Science **371**, 916 (2021).

[27] M. J. Mina, R. Parker, and D. B. Larremore, Rethinking Covid-19 Test Sensitivity — A Strategy for Containment, N. Engl. J. Med. **383**, e120 (2020).

[28] D. B. Larremore, B. Wilder, E. Lester, S. Shehata, J. M. Burke, J. A. Hay, M. Tambe, M. J. Mina, and R. Parker, Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening, Sci. Adv. **7**, eabd5393 (2021).

[29] J. L. Juul, K. Græsbøll, L. E. Christiansen, and S. Lehmann, Fixed-time descriptive statistics underestimate extremes of epidemic curve ensembles, Nat. Phys. **17**, 5 (2021).

[30] O. Karin, Y. M. Bar-On, T. Milo, I. Katzir, A. Mayo, Y. Korem, B. Dudovich, E. Yashiv, A. J. Zehavi, N. Davidovitch, R. Milo, and U. Alon, Cyclic exit strategies to suppress COVID-19 and allow economic activity, medRxiv 2020.04.04.20053579, doi:10.1101/2020.04.04.20053579.

[31] R. S. McGee, J. R. Homburger, H. E. Williams, C. T. Bergstrom, and A. Y. Zhou, Model-driven mitigation measures for reopening schools during the COVID-19 pandemic, Proc. Natl. Acad. Sci. USA **118**, e2108909118 (2021).

[32] P. I. Frazier, J. M. Cashore, N. Duan, S. G. Henderson, A. Janmohamed, B. Liu, D. B. Shmoys, J. Wan, and Y. Zhang, Modeling for COVID-19 college reopening decisions: Cornell, a case study, Proc. Natl. Acad. Sci. USA **119**, e2112532119 (2022).

[33] B. F. Nielsen, L. Simonsen, and K. Sneppen, COVID-19 Superspreading Suggests Mitigation by Social Network Modulation, Phys. Rev. Lett. **126**, 118301 (2021).

[34] K. Sneppen, B. F. Nielsen, R. J. Taylor, and L. Simonsen, Overdispersion in COVID-19 increases the effectiveness of limiting nonrepetitive contacts for transmission control, Proc. Natl. Acad. Sci. USA **118**, e2016623118 (2021).

[35] B. M. Althouse, E. A. Wenger, J. C. Miller, S. V. Scarpino, A. Allard, L. Hébert-Dufresne, and H. Hu, Superspreading events in the transmission dynamics of SARS-CoV-2: Opportunities for interventions and control, PLOS Biol. **18**, e3000897 (2020).

[36] S. Kojaku, L. Hébert-Dufresne, E. Mones, S. Lehmann, and Y.-Y. Ahn, The effectiveness of backward contact tracing in networks, Nat. Phys. **17**, 652 (2021).

[37] S. L. Feld, Why your friends have more friends than you do, Amer. J. Sociol. **96**, 1464 (1991).

[38] Altmetric—The effectiveness of backward contact tracing in networks.

[39] We note that Kojaku *et al.* also present a branching process analysis in their paper. In this branching process analysis, contact tracing and case isolation are not implemented sequentially. The branching process analysis, however, is not used to argue that backward contact tracing is more efficient than forward contact tracing, and so we focus entirely on Kojaku *et al.*'s simulations of compartmental models on networks in the following.

[40] J. S. Juul and S. H. Strogatz, Descendant distributions for the impact of mutant contagion on networks, Phys. Rev. Res. **2**, 033005 (2020).

[41] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, Science **286**, 509 (1999).

[42] C. McAloon, Á. Collins, K. Hunt, A. Barber, A. W. Byrne, F. Butler, M. Casey, J. Griffin, E. Lane, D. McEvoy, P. Wall, M. Green, L. O'Grady, and S. J. More, Incubation period of COVID-19: A rapid systematic review and meta-analysis of observational research, BMJ Open **10**, e039652 (2020).

[43] A. Reyna-Lara, D. Soriano-Paños, S. Gómez, C. Granell, J. T. Matamalas, B. Steinegger, A. Arenas, and J. Gómez-Gardeñes, Virus spread versus contact tracing: Two competing contagion processes, Phys. Rev. Res. **3**, 013163 (2021).

[44] P. Ashcroft, J. S. Huisman, S. Lehtinen, J. A. Bouman, C. L. Althaus, R. R. Regoes, and S. Bonhoeffer, COVID-19 infectivity profile correction, Swiss Med. Wkly. **150**, w20336 (2020).

[45] J. L. Juul and J. Ugander, Comparing information diffusion mechanisms by matching on cascade size, Proc. Natl. Acad. Sci. USA **118**, e2100786118 (2021).

[46] M. E. Kretzschmar, G. Rozhnova, M. C. J. Bootsma, M. V. Boven, J. H. H. M. Van de Wijgert, and M. J. M. Bonten, Impact of delays on effectiveness of contact tracing strategies for COVID-19: A modelling study, Lancet Public Health **5**, e452 (2020).

[47] S. H. S. Lai, C. Q. Y. Tang, A. Kurup, and G. Thevendran, The experience of contact tracing in Singapore in the control of COVID-19: Highlighting the use of digital technology, Int. Orthopaed. **45**, 65 (2021).

[48] A. Endo, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Q. J. Leclerc, G. M. Knight, G. F. Medley, K. E. Atkins, S. Funk, and A. J. Kucharski, Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks, Wellcome Open Res. **5**, 239 (2021).

[49] https://github.com/jonassjuul/backward-tracing.

[50] M. E. J. Newman and J. Park, Why social networks are different from other types of networks, Phys. Rev. E **68**, 036122 (2003).