Population-scale testing can suppress the spread of COVID-19

Jussi Taipale*†, Paul Romer#, and Sten Linnarsson†✉

*† Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden. *Applied Tumor Genomics Research Program, Faculty of Medicine, University of Helsinki, Finland and Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom. # New York University*

✉e-mail: sten.linnarsson@ki.se

We propose an additional intervention that would contribute to the control of the COVID-19 pandemic, offer more protection for people working in essential jobs, and help guide an eventual reopening of society. The intervention is based on: (1) testing every individual (2) repeatedly, and (3) self-quarantine of infected individuals. Using a standard epidemiological model (SIR), we show here that by identification and isolation of the majority of infectious individuals, including those who may be asymptomatic, the reproduction number R_0 of SARS-CoV-2 would be reduced well below 1.0, and the epidemic would collapse. We replicate these observations in a more complex stochastic dynamic model on a social network graph. We also find that the testing regime would be additive to other interventions, and be effective at any level of prevalence. If adopted as a policy, any industrial society could sustain the regime for as long as it takes to find a safe and effective cure or vaccine. Our model also indicates that unlike sampling-based tests, population-scale testing does not need to be very accurate: false negative rates up to 15% could be tolerated if 80% comply with testing every ten days, and false positives can be almost arbitrarily high when a high fraction of the population is already effectively quarantined. Testing at the required scale would be feasible if existing qPCR-based methods are scaled up and multiplexed. A mass produced, low throughput field test kit could also be carried out at home. Economic analysis also supports the feasibility of the approach: current reagent costs for tests are in the range of a dollar or less, and the estimated benefits for population-scale testing are so large that the policy would be cost-effective even if the costs were larger by more than two orders of magnitude. To identify both active and previous infections, both viral RNA and antibodies could be tested. All technologies to build such test kits, and to produce them in the scale required to test the entire worlds' population exist already. Integrating them, scaling up production, and implementing the testing regime will require resources and planning, but at a scale that is very small compared to the effort that every nation would devote to defending itself against a more traditional foe.

The ongoing pandemic spread of SARS-CoV-2 is the cause of widespread and accelerating outbreaks of the respiratory syndrome COVID-19. As of April 18, 2020, a total of 2 160 207 persons have been confirmed to be infected, and 146 088 have died[1]. Currently, the virus is present on all continents, growing rapidly in Europe and the USA, and is a major threat to world order. It now seems likely that unemployment in most countries will exceed the levels reached during the depth of the Great Depression of the 1930s. Detailed models of the epidemic dynamics of SARS-CoV-2 are available that take into account population structure, and can provide estimates of the magnitude and duration of the peak of infection (www.naherlab.org/covid19 and Ref. [2]). However, very broadly speaking, an outbreak of a novel virus in a naïve population can have only two outcomes. As long as the reproduction

number R_0 remains greater than 1, the virus spreads rapidly until most people have been infected (**Fig. 1A**), creating a temporary surge of infected individuals. If, using pharmacological or social interventions, R_0 can be reduced below 1, then the epidemic collapses (**Fig. 1B**), and most people remain uninfected (but still susceptible). Because of the exponential nature of epidemics, the outcomes are nearly binary. Even when R_0 exceeds one by only a small amount the disease spreads at an accelerating pace, whereas as soon as R_0 falls just below one it rapidly collapses. These two outcomes correspond to two distinct strategies for epidemic control, suppression and mitigation, close variants of which are currently attempted by several Asian countries (with different political systems) and Western democracies, respectively.

In the mitigation model, the goal is to reduce R as much as possible but not below 1.0, hoping to end up with a population that is largely immune, without overwhelming the healthcare system in the process (as in **Fig. 1A**, but attempting to flatten the temporary surge of infected individuals). This could (but is not guaranteed to) lead to "herd immunity" (see, for example Refs. [3,4]), which would limit spread in future epidemics caused by variants of the same virus. However, exponential processes are notoriously difficult to control, particularly in the absence of accurate real-time data and when the effect of policy changes is uncertain. The choice is stark: allowing the disease to spread to a large fraction of a population, however slowly, greatly increases the total number of infected people and would cause a loss of life that most societies will not accept. Furthermore, given the difficulties in controlling exponential processes using limited information, even a strongly enforced mitigation strategy runs the risk of overwhelming the health care system and significantly increasing the mortality rate due to the failure to treat every patient optimally (primarily due to the lack of intensive care capacity and sufficient numbers of ventilators). If the healthcare system is overwhelmed, patients must be triaged as in wartime, potentially for extended periods of time.

Notably, both suppression and mitigation are unstable: the mitigation model might first wreck the health-care system and then (as the public demands harsher controls when mortality rises) also wreck the economy. The suppression model might first wreck the economy and then as public pressure forces a relaxation of control, the virus re-emerges. For many months, both approaches are likely to force a large fraction of the population into quarantine. This is because of the large number of asymptomatic carriers of covid-19; in the absence of population-scale testing, the measures need to be implemented in an indiscriminate manner, affecting the whole population. Over time, this will result in severe and unequal economic deprivation. Our estimate is that in the United States, GDP per capita is already lower by about 1000 USD per month. Redistribution can offer some protection for the most vulnerable families, but if a loss of income of this magnitude persists for six or twelve months, it could generate a backlash against the social distance measures that are currently our only weapon for fighting the disease. As a result, epidemiologists are giving serious consideration to scenarios that alternate between lockdown and relaxation that will lead to more loss of life and add to even more economic uncertainty.

We can use what we know about the dynamics of disease to suppress this pandemic in a way that is far less disruptive than indiscriminate lockdowns and social distancing. Because it is less disruptive, a nation can sustain this approach for as long as it takes to find a safe and effective vaccine or a cure. Reducing the disruption will thereby save lives.

We know that at low levels of prevalence, testing, contact tracing and quarantine (TTQ) is a very effective means of suppression[3,5], because it reduces the effective rate of reproduction close to zero. It is not a feasible strategy for suppressing the virus in the current,

higher prevalence conditions faced by most countries because it would demand resources that would overwhelm any health department. In addition, the TTQ approach suffers from its own instability. Unless it identifies every single person who becomes infected, asymptomatic individuals that are not identified will generate clusters that will not be detected until someone develops a severe infection that requires medical care. If only 10% of cases are severe enough to be tested after two weeks, a single missed case will lead to an average cluster of 100 new cases before it is found. As a result, once the rate of new cases exceeds the capacity of tracing, even briefly, the epidemic runs out of control and the exponential dynamics make it almost impossible to catch up without imposing a lock-down.

Some have suggested that an updated version of TTQ that relies on modern surveillance technology could be viable because it will not make the same resource demands on the public health system. To be sure, it would be useful for innovators to work toward a working prototype that members of the public could voluntarily adopt. But because no such system has been deployed, even as a prototype, it would be dangerous for policy makers to count on the availability in the coming months of a system that is both effective in slowing the spread of disease and acceptable to most members of society.

Here, we propose a radically simpler strategy: just test everyone, repeatedly. When someone tests positive, ask them to self-quarantine and provide them public assistance that reduces the burden this imposes on them. This approach relies on a key observation that has not been widely appreciated, namely that what matters is the fraction of all individuals that are identified and quarantined. It follows that testing a small number of individuals with a highly accurate test can be much less effective than testing everyone with a less accurate test. In fact, there is a quantifiable relationship between the reproduction number of a virus, and the efficiency of a population-scale testing strategy that brings the effective reproduction number below 1. Below we use analytical models to derive both an upper and a lower bound on the effectiveness of testing, and demonstrate their real-world relevance using more realistic stochastic models.

The approach has several important advantages. First, it will work no matter how high the prevalence of infection might be. Second, it does not suffer from the inherent instability of contact tracing. The offsetting disadvantage is that it is a challenge to test at the required scale, but this is not as difficult as it might at first seem. It could be implemented using mass distribution (e.g. regular mail) without returning samples to a central testing site. In fact, the tests required do not even have to be properly "diagnostic." They will not be the basis for any decision about medical care. They only influence the decision to self-quarantine. In the worst case, they may cause people who are not infectious to be quarantined, but this is already true for most people (including the authors) in the baseline lockdown scenario. **This is an important feature, as it relaxes the demands on the quality of the test.** The test can tolerate many false positives, because the result of a provisionally positive test is that someone self-quarantines for two weeks when they did not have to. False negatives are also acceptable as long as people are retested frequently.

Although this strategy should be introduced alongside of existing measures, it is a useful exercise to ask what level of testing would be required for this strategy by itself to contain any level of infection. Clearly, if a perfectly accurate test were applied to the entire population at once, and those who tested positive were fully quarantined, the epidemic would immediately collapse with no new infections (**Fig. 1C**).

To examine the effects of false negatives and noncompliance, we first make the best case assumption about the timing of the tests: Every person who is infected is tested before encountering someone who is susceptible. This limit can be approached, for example, by a very effective form of contact-tracing. For coronavirus, it has been estimated[7] that $R_0 = 2.4$ and uncertain data from quarantine in Wuhan[8] suggests that $R_q = 0.3$. Using the standard (continuous, deterministic) SIR model, the equations in **Fig. 2A** and **Methods** show that the optimal population-scale testing strategy will succeed if at least two thirds of all new COVID-19 cases are immediately identified and quarantined. If p is the true positive rate of the test and c is the fraction of the public that complies in the sense that they agree to be tested and follow any instruction to go into quarantine, this bound means that the product cp must be greater than 2/3.

Next we do the opposite — assume that the test and compliance are perfect, so $p = c = 1$, and consider the worst-case assumption on the timing of the tests: each day, a randomly selected fraction of the population is tested. Under that strategy, we find that testing at a rate equal to $(R_0 - 1)$ percent of the population per infectious period will ensure that R < 1 (**Fig. 2B**, **Methods**). Using $R_0 = 2.4$ and a two-week infectious period for COVID-19, this implies that at least 10% of the population would have to be tested each day. Real-world testing strategies could do much better than test at random. for example by implementing procedures that test individuals concurrently within a region; that run the screen as a sweep across a country; that slice the population into groups that are tested in a cycle; or use other variables to predict who is more likely to be infected and to test them more frequently (see **Methods**).

Because herd immunity and other interventions — including the use of masks or reliance on social distancing — are additive with respect to the testing, any of these effects can lower the required frequency of the tests. For example, **Fig. 1D** shows the required compliance rate as a function of the strength of other interventions, assuming a fixed false negative testing rate of 15%.

The standard but simple and deterministic Susceptible-Infectious-Recovered (SIR) models used to calculate these bounds is based on strong assumptions and approximations, such as random mixing of all individuals. To relax those assumptions, we implemented a more realistic numerical simulation using a stochastic model on a social interaction graph (i.e. a stochastic network SEIR model) to model two realistic scenarios. We focused on the initial exponential growth phase of an epidemic. **Fig. 3A** shows a simulation that starts with 100 infected individuals and assumes that the product of the compliance and true positive rate $cp = 0.8$. Population-scale testing using random weekly tests was started on day 20, and immediately suppressed the epidemic, which was fully stopped by day 100. In contrast, without testing, viral spread caused a surge of infections. Death rates were 0.19% with testing, and 0.66% without, i.e. a more than three-fold improvement corresponding to 1.5 million lives saved in a US-sized population. This demonstrates the power of population-scale testing and quarantine for the suppression of novel viruses.

The second scenario modeled a country that exits from a lockdown that has suppressed the growth of a pandemic and a with a small fraction of individuals who are immune (**Fig. 3B**). In such cases, new outbreaks will inevitably happen. TTQ and periodic lockdowns can suppress these outbreaks, but so can an ongoing process of testing and isolating. In the model, a lock-down was applied from day 20 which nearly extinguished the virus by day 100. At that point, the lockdown was lifted and social interactions returned to normal, but population-scale testing and quarantine was applied as above. Once again, the epidemic was suppressed indefinitely, with total deaths limited to 0.16% of the population. If

the lockdown was lifted without population-scale testing, a powerful second wave was generated leading to the death of 0.57% of the population overall. The decreased mortality due to testing after lifting lockdown corresponds to more than one million lives saved in a US-sized population. This demonstrates that population-scale testing can be an effective replacement for periodic lockdown as a sustainable way to prevent the resurgence of the virus.

Finally, systematic simulation of the parameters of the stochastic model (**Supplemental Figure 1**) showed that with $cp = 2/3$ (for example, a compliance of 80% and test sensitivity of 85%) testing at least every 11 days on average was sufficient to suppress the epidemic, whereas testing less frequently was not.

The scale of screening required for the approach is not unimaginable, as it is within an order of magnitude of the level of screening that is critical for protection of the segment of the workforce that is employed in essential sectors such as healthcare, elderly care, public order and delivery of foods and medicines. Furthermore, building a test that could be applied at population scale is clearly feasible using current technology. It could be based on detection of antibodies to the virus[9]. However, as it takes time for antibodies to build, the antibody-test cannot detect cases early. A single test also does not discriminate between current and past infection. Making sure that someone is not currently infected requires that an antibody test is performed twice over a period of three weeks, during which the individual must be held in strict quarantine. Alternatively, the test can be combined with an RNA or antigen test. Despite these drawbacks, an antibody test will clearly be part of the solution, as it can detect immune individuals that can continue to work safely in health care or with risk groups. However, in countries where the epidemic is successfully suppressed, the fraction of immune individuals is far too small (e.g. 1-10%) for restoring normal levels of economic activity. Furthermore, deploying antibody test at a population scale will be more difficult than using an RNA test, as the current approach requires blood samples, which decreases compliance and makes self-testing more difficult.

A population-scale test can also be based on viral proteins (technically more difficult but possible[10]), or viral RNA, like the current state-of-the-art diagnostic tests (for example Ref. [11]). Technically, few measurements are easier and/or cheaper in biochemistry than determining whether a particular RNA species is present in a particular sample. The main technical concern relates to false positives caused by contamination of input samples by the amplified DNA from previous tests; this can be simply prevented by well-established procedures. Despite the technical simplicity, detecting viral RNA in the field at population-scale is difficult to achieve using the same design and strict regulatory framework that is used for tests designed for medical diagnostic purposes. Current diagnostic tests for SARS-CoV-2 are qRT-PCR assays that require (1) nasopharyngeal swab collected by a trained nurse, (2) sample collection in viral transport media, (3) RNA purification, (4) reverse transcription and quantitative PCR. The test is highly accurate, and the total cost is in the order of $100. Such highly accurate testing is critical for accurate diagnosis of cases in a hospital setting. However, due to the very detailed and specific regulation, specialized staff and equipment, and centralized testing facilities, such tests have proven difficult to rapidly scale above thousands of assays in each location. A distributed system of sample collection and testing could, however, conceivably be used to scale qRT-PCR to population levels, particularly when using a regional sweeping approach to limit the number of simultaneous tests needed. The capacity could also be increased 10 to 100-fold by group testing[24], a method with a long history of use in public health that was originally designed for Syphilis tests, and now commonly also used for optimally efficient detection of defective components in industrial

production. Although some components of tests are currently limiting due to sudden high demand, scaling up their production is not difficult, as the methodology is based on raw materials that are not scarce (e.g. plastic, sand) and biological molecules (enzymes, nucleotide triphosphates and short nucleic acid primers) that are easily produced either industrially or locally using simple biotechnological processes. We also note that qPCR instruments are currently in short supply, but isothermal tests are available that require only a waterbath (see below).

A parallel relatively centralized testing method based on existing DNA sequencing technology could also be fielded rapidly. In this approach, viral RNA in the samples is used to generate DNA sequences containing the virus sequences, a sample DNA barcode (to identify each case) and two unique molecular identifiers[23] at both ends of the resulting DNA fragment (to count the number of virus RNAs per sample and to ensure that patient samples do not get mixed in the reaction), and then sequenced using a massively parallel sequencer. This approach is very scalable, as in principle, a single sequencing instrument that is routinely used in scientific research can report more than a billion results per day. Furthermore, in the future, a test based on sequencing[19-21] that covers many acute infections could also be used to suppress or even eradicate a large number of infectious diseases simultaneously. This would have significant benefits to humanity, and would be very difficult to achieve using vaccines or drugs that target each infectious agent separately.

Alternatively, we envisage supplementing the current testing regime with a mass-produced home test kit that could be used by anyone, result in a simple easily-understood readout, and be performed without specialized equipment. The test should be as easy to use as a pregnancy test, to ensure maximal compliance. Boxes of e.g. 50 tests would be mass-mailed to all citizens, and a national information campaign would encourage everyone to test themselves weekly. In an infected individual, viral RNA is present at reasonably high levels in nasopharyngeal swabs, throat swabs, sputum, and stool for up to two weeks[12], with the greatest amounts in sputum and stool. Sputum might be the ideal source for a home test kit, given the ease of sampling. Compliance of the home testing could be increased by both rewards and penalties, and potentially enforced by adding a serial number to each test that needs to be reported together with the test result to collect the rewards. The test result can be open in such a way that the result is clear to everyone. It can also be designed so that it maintains privacy. Here, the result (e.g. resulting color, number of bars that are visible) needs to be reported together with the serial number to a central facility to get the answer and/or the cash reward. The open and private approaches can also be combined, to design a test that is open and contains an encoded part that needs to be reported to collect a reward. Such designs may complicate the approach, but would allow the healthcare system to obtain data that would facilitate monitoring of the outbreak and large-scale contact tracing. The cash rewards could also be contingent on being regularly tested. Anyone found positive would be compelled to self-quarantine, possibly under monetary or criminal sanctions, or using additional rewards for compliance. Provided that the test is sufficiently quick, testing could be performed in workplaces, or even in checkpoints exiting areas with high infection rate that are currently under lockdown.

Our approach is not something that can only be fielded in the far future. In fact, tests suitable for home use have already been developed. In contrast to the commonly used polymerase chain reaction (PCR), which optimally requires an instrument that repeatedly changes the temperature of the reaction, many other "isothermal" detection methods have been developed that operate using a single set temperature, and do not require special equipment beyond what is available in every kitchen (e.g. hot water). For example, an

isothermal and colorimetric test has been described[13,14], based on reverse transcription-loop mediated amplification (RT-LAMP) technology. This test has several desirable properties: unlike PCR, it does not require temperature cycling; the readout is binary and can be achieved by simple observation; and it can start from crude samples[15]. Many other technologies also have the potential to detect viral RNA rapidly and isothermally[16-18]; these include recombinase polymerase amplification (RPA), transcription mediated amplification, nicking enzyme amplification reaction (NEAR), rolling circle replication, and *in vitro* viral replication assays. Although a population-scale test does not need to be as accurate as a clinical-grade qRT-PCR test (see above), apart from a potential increase in errors due to sample collection, there is no theoretical reason why a self-test based on isothermal amplification would not achieve the false negative and positive rates that are equivalent to the current state-of-the-art methodology. Making the necessary reagents at scale is also not difficult. Per 10 million people, a 100 µL test requires 1,000 liters of reagents, consisting of primers, nucleotides, pH-sensitive dye and enzyme, all of which are easy to make at the required scale.

A population-scale strategy has the potential to save many lives, and to buy precious time for a vaccine or an effective drug to be developed. It's important to recognize that in contrast to vaccine and drug development, scaling up testing does not depend on any new scientific discoveries; it is a matter of engineering and logistics only. A field test would synergize with drug treatment, as many antivirals act more effectively when they are given at an early stage of the infection. Furthermore, development of field-applicable tests needed for rapid population-level screening will have great benefits in combating epidemics in countries with less developed healthcare systems, and would also help in responding to future epidemics, or variants of the current one. The costs of mobilization of scientific and industrial resources for rapid development of such a test are considerable; however, in our opinion, they are still orders of magnitude lower than the costs of the current suppression and mitigation strategies. Although balancing collective goods such as economic activity and public health commonly involve very difficult trade-offs, we believe that such a trade-off is not relevant in this case, and that a population-scale screening policy can be implemented in such a way that will both save more lives *and* cause less economic and social disruption than the current approach. As there is little overlap with other industrial mobilization efforts, such as scaling up current testing regime, building of ventilators or developing drugs or a vaccine, the increased effort for the development of tests would also have very limited opportunity cost. The development of capacity for population-scale testing would also be an important and relatively inexpensive insurance against other pandemics, or re-emergence of Covid-19 once herd immunity is lost or the virus mutates to evade immunity, vaccines and/or antiviral drugs.

Acknowledgments

We thank many colleagues for comments on the early version of the work. We are especially grateful to Drs. Minna Taipale, Mikko Taipale and Paul Pharoah for review of the draft of manuscript. A draft of this paper was initially released as a public preprint (Ref. 25), and supporting, independently developed model reported by P.R. on www.paulromer.net. We also note that during the writing of this work, we became aware of two independent analyses, one by Julian Peto (Ref. 26), and the other by a team consisting of David Berger, Kyle Hirkenhoff and Simon Mongey that report similar conclusions (Ref. 27).

Competing interests
The authors declare no competing interests.

METHODS

The epidemic was first modelled with a standard (continuous, deterministic) susceptible, infected, removed (SIR) model. In addition to the very general assumption that there are a relatively large number of cases, which allows modeling of a partially discrete system using a continuous model, the SIR model is based on the following standard assumptions: (1) the population is fixed, (2) it mixes homogenously, (3) the only way a person can leave the susceptible group is to become infected, (4) the only way a person can leave the infected group is to recover from the disease, (5) recovered persons become immune, (6) age, sex, social status, genetics etc. do not affect the probability of being infected, (7) there is no inherited immunity, and (8) the other mitigation strategies and testing are independent of each other (for **Fig. 1D**). The assumption (2) leads the SIR model to overestimate viral spread, as in reality population has substructure (e.g. families, workplaces) and is geographically separated and contacts are more likely between subsets of the population; this is not expected to materially affect our analysis as our conclusions are not based on the absolute rate of the spread, only on its exponential nature.

In addition, we modeled the effect of testing in two ways. The first, maximally effective testing strategy assumed that every individual was tested before they infected another person, leading to the upper bound on testing performance in **Fig. 2A-B**. Under this model, the requirement for collapsing the epidemic is that the weighted average of the basic reproduction number R_0 and the reproduction number in quarantine R_q must be less than one:

$$pcR_q + (1 - pc)R_0 < 1$$

$$pc > \frac{R_0 - 1}{R_0 - R_q}$$

Here, p is the true positive rate of the test and c is the compliance (fraction of all tested individuals who actually self-quarantine).

Using $R_0 = 2.4$ and $R_q = 0.3$ for COVID-19, the product of the true positive rate and compliance must be greater than two thirds:

$$pc > {}^{2}/_{3}$$

The second, lower bound testing strategy (**Fig. 2C-D**) was modelled by adding an additional 'detected' state to the model, and adding transitions from infected to detected (with rate τI) and from detected to recovered (with rate γD). This corresponds to continuous random testing of the population at a fixed rate τ per person per day. Here, the requirement for successful collapse of the epidemic is given by the basic reproduction number (assuming perfect quarantine; **Fig. 2D**), as follows. First, the rate equations for the SIR model with testing are:

$$\frac{dS}{dt} = \frac{-\beta SI}{N}$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - (\gamma + \tau)I$$

$$\frac{dD}{dt} = \tau I - \gamma D$$

$$\frac{dR}{dt} = \gamma(I + D)$$

Rewriting the second equation above as follows:

$$\frac{dI}{dt} = \left(\frac{\beta}{\gamma + \tau}\frac{S}{N} - 1\right)(\gamma + \tau)I$$

makes it clear that dI/dt will be negative (i.e. the epidemic will collapse) only if:

$$\frac{\beta}{\gamma + \tau} < \frac{N}{S}$$

Note that the ratio $\beta/\gamma$is the basic reproduction number R_0, so that the previous inequality can be rewritten as follows:

$$\tau > \gamma\left(\frac{R_0 S}{N} - 1\right)$$

In other words, the testing rate must exceed a threshold given by the recovery rate γ and the fraction of susceptible individuals S/N. In the beginning of an epidemic in a naïve population, when all individuals are susceptible, this reduces further to

$$\tau > \gamma(R_0 - 1)$$

For SARS-CoV-2, assuming $\gamma = 1/14$ (i.e. an infectious interval of two weeks) and $R_0 = 2.4$, the required minimal testing rate would be 10% of the population per day. As the

epidemic progresses, the required testing rate drops as fewer and fewer individuals remain susceptible and herd immunity kicks in.

To understand the difference between (1) testing everyone at the same time, (2) testing everyone in a time separated manner, or (3) testing the population by random sampling, it is helpful to consider an extreme case of certainly and completely collapsing an epidemic by testing and quarantine, using a perfect test that detects all infected individuals, and complete quarantine. For optimally achieving this, it is necessary to identify everyone who is infected before they have infected anyone else (denoted efficiency, e = 1) and quarantining every infected individual (c = 1). This requires obtaining a minimum of n bits of information for a population of size n.

In case (1), this is achieved by testing everyone at the same time with a perfectly accurate test that returns one bit (positive or negative). In this case, e = 1 and c = 1. However, when tests are separated in time (2), the order of testing becomes important. The optimal strategy discussed in the preceding section, testing everyone at different times, but before they have had the chance of infecting anyone else also works optimally and collapses the epidemic over a single infectious period. However, most strategies for testing n individuals during time $t_{test_interval}$ before t_0 have e < 1, and are not sufficient to completely collapse the epidemic using one testing round, as e depends on the relationship between the order of testing and the order of infections. For example, using a random order of testing allows some individuals that have already been tested negative to become infected during the $t_{test_interval}$ (the mutual information between test results and person being infected at t_0 is less than one bit). However, some other regimens using a perfectly sensitive test can collapse the epidemic (but not always prevent all future infections): for example, a geographical sweep where infections (individuals) are prevented from crossing a moving test front can be used to identify every infected individual in the population by performing a single round of n tests.

In case (3), random sampling of n individuals, e is always less than 1. The testing becomes less efficient than testing each of the n individuals at the same time, because some individuals are tested twice, and some not at all; some information is thus not obtained, and some tests do not return information that is completely independent of information returned by other tests (sum of mutual information between all pairs of tests is not 0 bits). In other words, if individuals are selected randomly, during a given time interval, the tests will miss some individuals, and some individuals are tested more than once (this increases true positive rate for those individuals, but this does not make up for failing to catch some individuals entirely).

Considering the extreme case of immediate collapse, it may appear that testing in a time separated manner or by using random sampling will not work because non-concurrent testing can permit infections to cross the testing boundary, and random sampling clearly leaves some cases undetected. However, this very intuitive idea is incorrect, as collapsing an epidemic only requires that the rate of generation of new cases per current case is less than one. The limit for random testing can be obtained using the SIR model extended with testing (SIR+T), which abstracts away individuals and thus can (only) be used to investigate the effect of random, time-separated testing. Analytically from this model, as shown above, the $R < 1$ condition is true when tests are performed at a rate that is higher than $R_0 - 1$ tests per mean infectious period. The same limit results from the following consideration: reducing R_0

to less than one using the method representing the lower bound – a completely random testing regime – requires that an infected individual has less than an equal probability of (a) infecting another individual over (b) being tested and quarantined or recovering from the infection (analogously, in SIR+T, the combined testing and recovery rate needs to be higher than the rate of new infections). Events (a) can recur, but either event (b) terminates the chain. Therefore, at R = 1 there will be on average one (a) event, which requires that the order of the infectious and protective events are randomly ordered with respect to each other, with equal density. This yields $R_0 - 1$ tests and one recovery per R_0 infections per infectious period, and an upper limit of R_0 tests per infectious period at infinity (because as $R_0 \to \infty$ the expectation value for R becomes the geometric series $\sum_{n=1}^{\infty} 2^{-n} = 1$).

Outside of the theoretical consideration of $e = 1$, multiple population-scale tests are always required to collapse the epidemic in the absence of other interventions that achieve the same aim. Performing multiple tests over time imposes an additional constrain on optimality – the allocation of tests to each transmission interval. As described above, best performance of continuous testing and quarantine is thus achieved when testing is performed immediately after infection for each individual, or as requirement for exiting quarantine. Testing blood before transfusion to prevent transmission of HIV or hepatitis C, testing at border crossings, conditional opening of lockdown, or some regimes that apply contact-tracing may come close to approximating this limit, which for Covid-19 is $pc = (R_0 - 1)/R_0 = 0.57$ per mean infectious period; **Fig. 2**). However, in most scenarios, such testing efficacy is difficult to maintain over time (because contact is lost, and the unknown infectious intervals rapidly become randomly distributed over time). This level can thus be considered an upper limit of performance of any scenario applied at population scale.

Using a test whose true positive rate of 1 and testing everyone at the same time performs as well as the optimal strategy. As test sensitivity decreases, the performance of the concurrent regime becomes lower than optimal. However, concurrent testing still performs well above the lower limit obtained from the random testing model. The required pc rate to bring R0 < 1 using concurrent tests has a simple relationship with the exponential growth of infectious cases. Over interval t-t_0, pc > 1-(infectious cases at t_0)/(infectious cases at t). However, it is not as simple to relate this to original R0, because the relationship between R0 and growth rate is a function of the distribution of the generation intervals (Ref. 28). Estimating at R0 = 2.4 using even probability distribution of infections over time, the infected population becomes approx. eight times larger in a single 14 day infectious period. This means that a testing regime that is regularly spaced at 14 day intervals should have pc value of > 7/8 = 0.875 to bring R0 < 1. This is confirmed using empirical simulations to assess the rate of exponential growth in the complete absence of immunity and all other types of interventions; the limit $R = 1$ at $R_0 = 2.35$ with testing every infectious period (14 days) is reached when $pc \approx 0.85$ (compared to 0.58 for testing each individual directly after infection). The required testing interval at $R_0 = 2.35$ and $pc = 0.8$ in the absence of other interventions and immunity is 11, 8 and 5.5 days for concurrent testing, testing each individual randomly once during each testing period, and continuous random testing, respectively.

These considerations can be summarized as follows: the order of testing efficacies is: everyone before they have had a chance to infect anyone > everyone at the same time > everyone once during a period > testing by random sampling – with population-scale testing

remaining feasible and cost-effective by one or more orders of magnitude across all these regimens.

Stochastic SEIR model on a social graph

The SIR model fails to account for several key properties of real epidemics, such as social and geographical population structure, the discrete and stochastic nature of infection and disease progression, and the fact that testing cannot be instantaneous. To account for such more complex real-world phenomena, we implemented a stochastic network model using the Gillespie algorithm for accurate numerical simulation of the stochastic dynamics. We used the seirsplus Python package (https://github.com/ryansmcgee/seirsplus), which models an epidemic on a social graph, where each individual transitions between six states: susceptible, exposed, detected-exposed, infectious, detected-infectious, and recovered. The two detected states are used to model the effectiveness of testing and quarantine, and social distancing is modelled by removing edges from the initial social graph. We used a random social graph of mean degree 13 (median 10) and two-sided exponential tails, which was reduced to mean degree 2 for social distancing (lockdown) and quarantine. The population consisted of 10,000 individuals. In both shown scenarios we assumed that the test had 80% sensitivity, and epidemic parameters were modelled loosely after COVID-19. Detailed source code with comments and parameter settings for each model are available in the accompanying Jupyter notebook at https://github.com/paulromer149/ubiquitous-testing.

REFERENCES

1. Organization, W.H. Coronavirus disease 2019 (COVID-19) Situation report - 59. Vol. 2020 (2020).
2. Ferguson, N.M. & al., e. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. *imperial.ac.uk* **1**, 1-20 (2020).
3. Madhav, N.*, et al.* Pandemics: Risks, Impacts, and Mitigation. in *Disease Control Priorities: Improving Health and Reducing Poverty* (eds. rd*, et al.*) (Washington (DC), 2017).
4. Guerra, F.M.*, et al.* The basic reproduction number (R0) of measles: a systematic review. *The Lancet. Infectious diseases* **17**, e420-e428 (2017).
5. McCloskey, B., Dar, O., Zumla, A. & Heymann, D.L. Emerging infectious diseases and pandemic potential: status quo and reducing risk of global spread. *The Lancet. Infectious diseases* **14**, 1001-1010 (2014).
6. Tondo, L. Scientists say mass test in Italian town have halted Covid-19 there. in *The Guardian* (The Guardian, UK, 2020).
7. Li, R.*, et al.* Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science* (2020).
8. Wang, C.*, et al.* Evolving epidemiology and impact of non-pharmaceutical interventions on the outbreak of coronavirus disease 2019 in Wuhan, China. *medRxiv* **2020.03.03**, 20030593v20030591 (2020).
9. Meyer, B., Drosten, C. & Muller, M.A. Serological assays for emerging coronaviruses: challenges and pitfalls. *Virus research* **194**, 175-183 (2014).
10. Loynachan, C.N.*, et al.* Platinum Nanocatalyst Amplification: Redefining the Gold Standard for Lateral Flow Immunoassays with Ultrabroad Dynamic Range. *ACS nano* **12**, 279-288 (2018).

11. Corman, V.M., *et al.* Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **25**(2020).
12. Woelfel, R., *et al.* Clinical presentation and virological assessment of hospitalized cases of coronavirus disease 2019 in travel-associated transmission cluster. *medRxiv* **2020.03.05**(2020).
13. Yu, L., *et al.* Rapid colorimetric detection of COVID-19 coronavirus using a reverse-transcriptional loop-mediated isothermal amplification (RT-LAMP) diagnostic platform: iLACO. *medRxiv* **2020.02.20**(2020).
14. Zhang, Y., *et al.* Rapid molecular detection of SARS-CoV-2 (COVID-19) virus RNA using colorimetric LAMP. *medRxiv* **2020.02.26**(2020).
15. Silva, S., *et al.* Development and Validation of Reverse Transcription Loop-Mediated Isothermal Amplification (RT-LAMP) for Rapid Detection of ZIKV in Mosquito Samples from Brazil. *Scientific reports* **9**, 4494 (2019).
16. Bodulev, O.L. & Sakharov, I.Y. Isothermal Nucleic Acid Amplification Techniques and Their Use in Bioanalysis. *Biochemistry. Biokhimiia* **85**, 147-166 (2020).
17. Qi, H., Yue, S., Bi, S., Ding, C. & Song, W. Isothermal exponential amplification techniques: From basic principles to applications in electrochemical biosensors. *Biosensors & bioelectronics* **110**, 207-217 (2018).
18. Zhao, Y., Chen, F., Li, Q., Wang, L. & Fan, C. Isothermal Amplification of Nucleic Acids. *Chemical reviews* **115**, 12491-12545 (2015).
19. Deng, X., *et al.* Metagenomic sequencing with spiked primer enrichment for viral diagnostics and genomic surveillance. *Nature microbiology* **5**, 443-454 (2020).
20. Erlich, Y. A vision for ubiquitous sequencing. *Genome research* **25**, 1411-1416 (2015).
21. Hoper, D., Wylezich, C. & Beer, M. Loeffler 4.0: Diagnostic Metagenomics. *Advances in virus research* **99**, 17-37 (2017).
22. Kermack, W.O. & McKendrick, A.G. Contributions to the mathematical theory of epidemics--I. 1927. *Bulletin of mathematical biology* **53**, 33-55 (1991).
23. Kivioja, T., et al. Counting absolute numbers of molecules using unique molecular identifiers. Nature methods 9, 72-74 (2011).
24. Dorfman, R. The detection of defective members of large populations. Ann Math Statist 14, 436-440 (1943).
25. Taipale, J. & Linnarsson, S. Test everyone, repeatedly to defeat Covid-19. *Medium* medium.com/@sten.linnarsson/to-stop-covid-19-test-everyone-373fd80eb03b, (2020).
26. Peto, J. Covid-19 mass testing facilities could end the epidemic rapidly. Bmj 368, m1163 (2020).
27. Berger, D.W., Herkenhoff, K.F. & Mongey, S. An SEIR Infectious Disease Model with Testing and Conditional Quarantine. National Bureau of Economic Research Working Paper March, 26901 (2020).
28. Wallinga, J. & Lipstch, M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc R Soc B* **274**, 599-604 (2006).

FIGURES

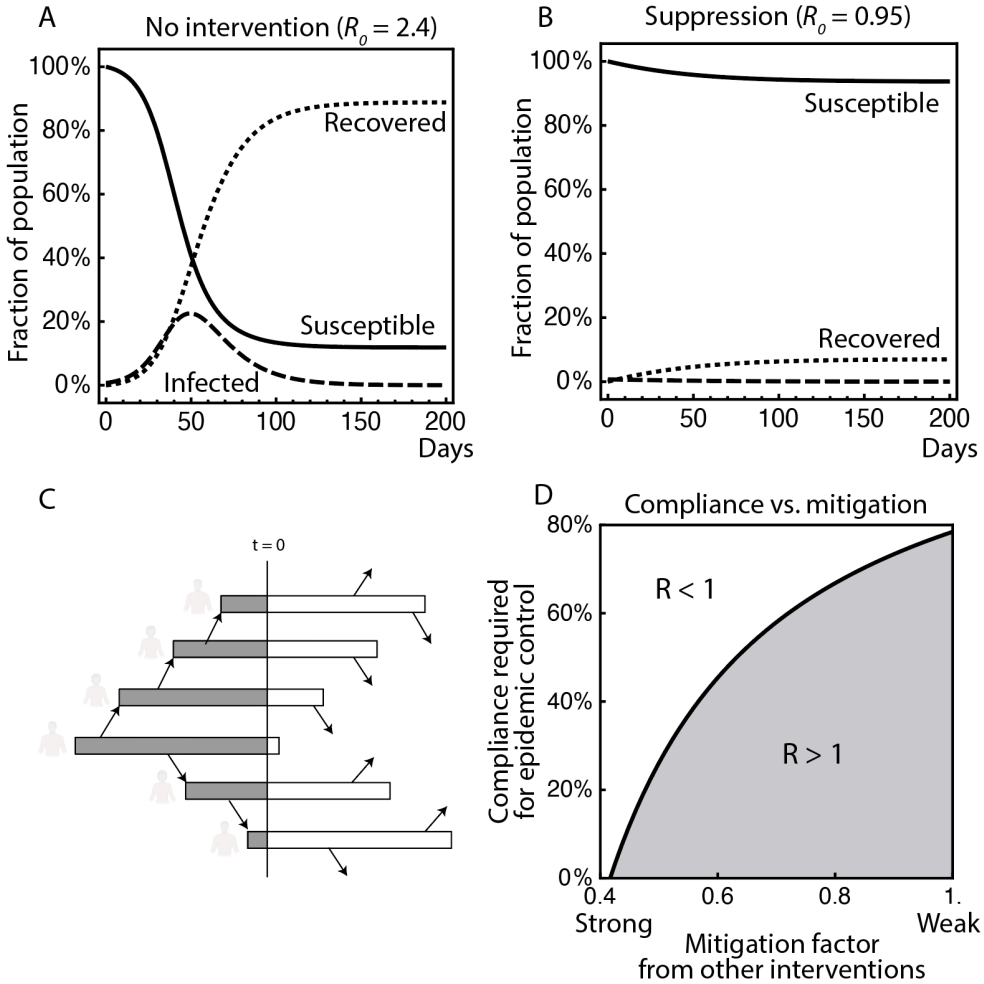


Fig. 1 | Requirements for epidemic control of COVID-19 using population-scale testing and self-quarantine. (A) An SIR model[22] with $R_0 = 2.4$ leads to infection of the majority of the population, with a massive peak of active infection that overwhelms the health care system. (B) With strong interventions that reduce R_0 to 0.95, many deaths are avoided. (C) Testing everyone simultaneously cuts all chains of transmission. (D) The required level of compliance (fraction of all individuals) for effective control of the epidemic, as a function of the strength of other interventions and assuming a test with 85% efficacy (fraction of future infections detected). With moderate social distancing, epidemic control can be achieved even with low levels of test compliance.

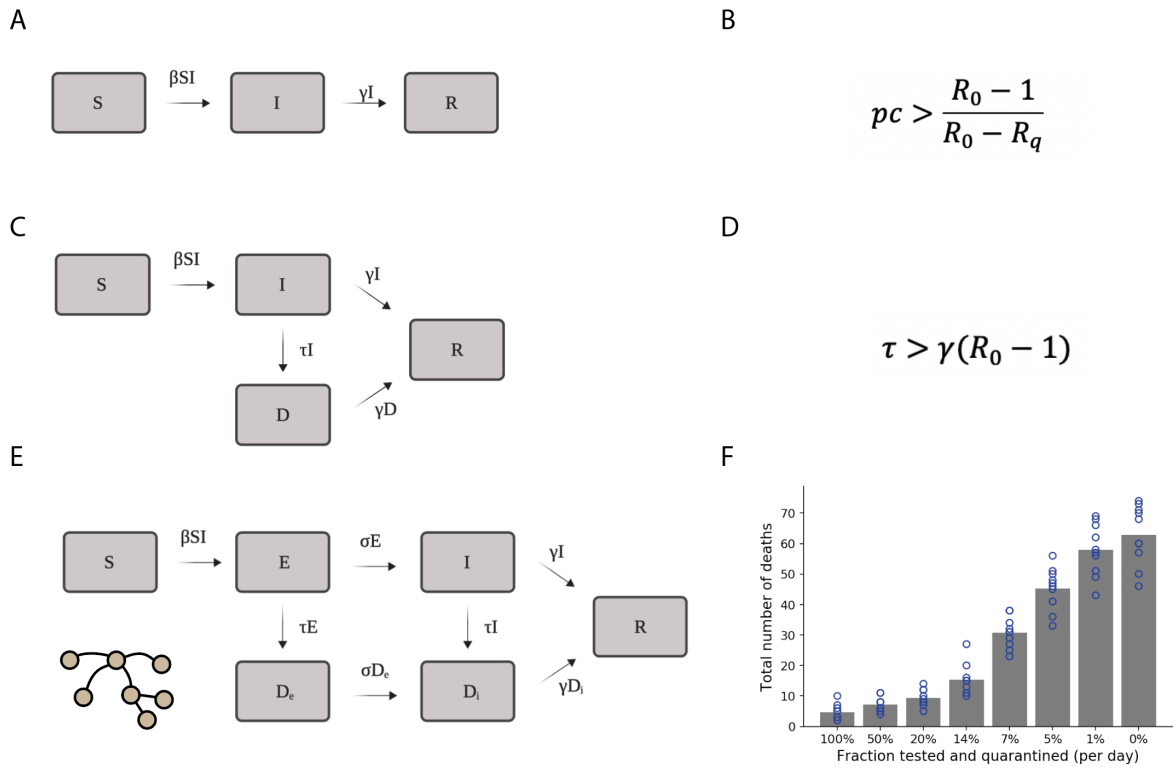


Fig. 2 | Compartment models and requirements for suppression by testing. (A) Parameters for a standard SIR model. (B) Inequality that must be true to suppress transmission. For the epidemic to collapse, the weighted average of the natural reproduction number R_0 and the reproduction number in self-quarantine R_q must be less than one. Here, p represents the test true positive rate (fraction of all infectious individuals detected), and c the rate of compliance. (C) Parameters for a SIR model with testing and a detected state. (D) Requirements for testing to collapse an epidemic in the SIR model with testing, expressed in terms of the testing rate t required in a population where all individuals are susceptible, with inverse infectious interval λ. (E) Parameters for the discrete, stochastic SEIR model on a social graph. Each compartment was modelled for every individual on the social graph. (F) Outcomes of ten simulation runs of the stochastic SEIR model on a social graph, showing total number of deaths as a function of the fraction tested every day, assuming compliance and true positive rate $pc = 2/3$.

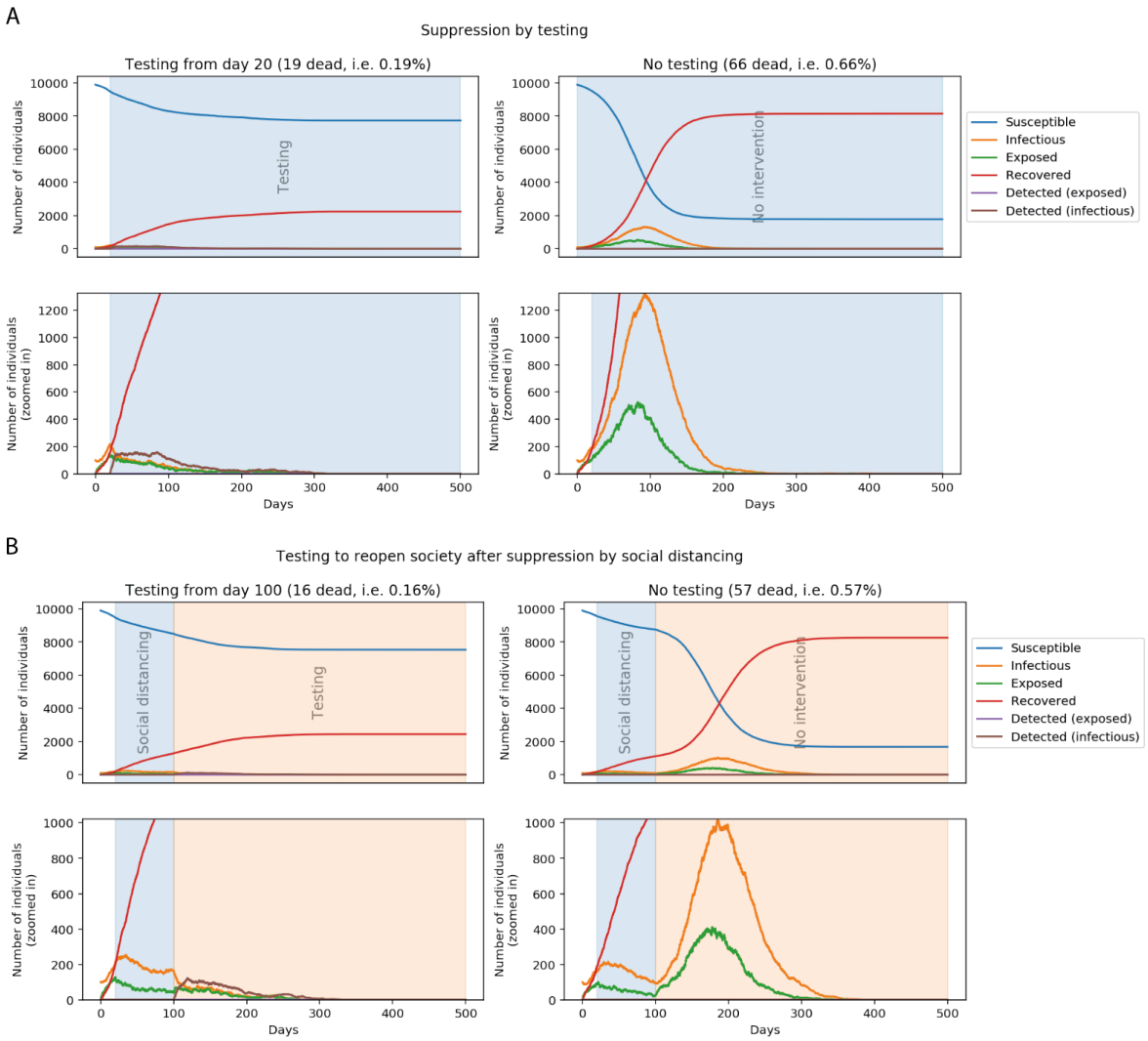
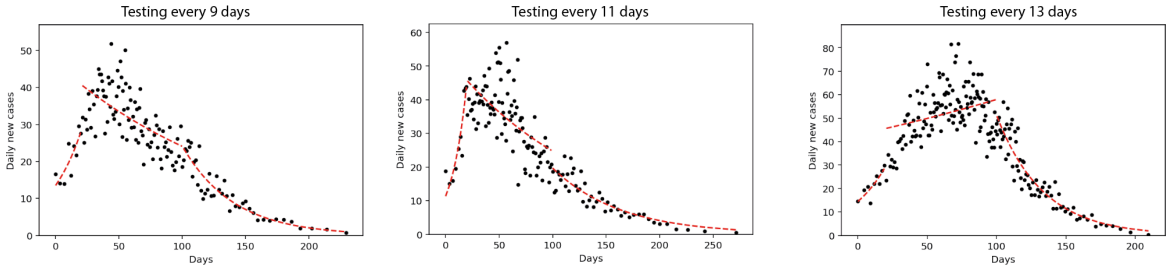


Fig 3 | Scenario simulation using a stochastic network SEIR model. (A) Simulation of a population of 10,000 individuals, 100 of which are infected at the outset. Population-scale testing is implemented from day 20 (left), or not implemented (right). Lower panels show the same plots but zoomed in on the vertical (population) scale. (B) Simulation of a population of 10,000 individuals, 100 of which are infected at the outset. Severe social distancing (lockdown) is applied after 20 days, and then lifted after 100 days. Population-scale testing is implemented after day 100 (left) or not implemented (right). In all panels, shaded colored regions indicate policy regimes, and the total number of dead individuals is indicated in the sub-panel titles. For both panels, the product of compliance and test efficacy was set to $ce = 0.8$ and the testing rate was set to $\tau = 1/7$.



Supplemental Figure 1 | Successful testing strategies under the stochastic SEIR model on a social graph. Growth curves (daily new cases) showed that given $cp = 2/3$, testing at least every 11 days successfully flipped the sign of the exponential growth curve from day 20 (when testing started), whereas testing every 13 days was insufficient. Red dashed curves, piecewise exponential fits for days 0 – 20, 20 – 100, and 100 – 250.