Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening

Daniel B. Larremore1,2*, Bryan Wilder3, Evan Lester4,5, Soraya Shehata5,6, James M. Burke4, James A. Hay7,8, Milind Tambe3, Michael J. Mina7,8,9*†, Roy Parker2,4,6,10*†

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

INTRODUCTION

Successful population screening testing of SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) depends on understanding both the dynamics of spread between individuals and the dynamics of the virus within the human body. Critically, the ability of SARS-CoV-2 to spread from individuals who are symptomatic, asymptomatic, or essentially asymptomatic (1–3) means that diagnosis and isolation based on symptoms alone will be unable to prevent ongoing spread (4, 5). As a consequence, the use of population screening testing to identify infectious individuals presents one possible means to break enough transmission chains to suppress the ongoing pandemic and reopen societies, with or without a vaccine.

The reliance on testing as a means to safely reopen societies has placed a microscope on the analytical sensitivity of virus assays, with a gold standard of quantitative real-time polymerase chain reaction (RT-qPCR). These assays have analytical limits of detection that are usually within around 10^3 viral RNA copies per milliliter (cp/ml) (6). However, RT-qPCR remains expensive and, as a laboratory-based assay, often has sample-to-result times of 24 to 48 hours. New developments in SARS-CoV-2 diagnostics have the potential to reduce cost substantially, allowing for expanded testing or greater frequency of testing and can reduce turnaround time to minutes (7–9). These assays, however, largely do not meet the gold standard for analytical sensitivity, which has encompassed the widespread use of these assays (10).

Three features of the viral increase, infectivity, and decline during SARS-CoV-2 infection led us to hypothesize that there might be minimal differences in effective screening regimens using viral detection tests of different sensitivities, such as RT-qPCR with a limit of detection (LOD) at 10^3 cp/ml (6) compared to often cheaper or faster assays with higher LODs [i.e., around 10^5 cp/ml (7–9)] such as point-of-care nucleic acid loop-mediated isothermal amplification (LAMP) and rapid antigen tests (Fig. 1A).

First, since filtered samples collected from patients displaying less than 10^5 N or E RNA cp/ml contain minimal or no measurable infectious virus (11–13), either class of test should detect individuals who are currently infectious. The absence of infectious particles at viral RNA concentrations <10^5 cp/ml is likely due to (i) the fact that the nucleocapsid and envelope RNAs are also present in abundant subgenomic mRNAs, leading to overestimation of the number of actual viral genomes by ~100 to 1000× (14); (ii) technical artifacts of RT-PCR at C_t values >35 due to limited template (15, 16); and (iii) the production of noninfectious viral particles as is commonly seen with a variety of RNA viruses (17). Second, during the exponential growth of the virus, the time difference between 10^3 and 10^5 cp/ml is short, allowing only a limited window in which only the more sensitive test could diagnose individuals. For qPCR, this corresponds to the time required during viral growth to go from C_t values of 40 to ~34. While this time window for SARS-CoV-2 is not yet rigorously defined (18), for other respiratory viruses such as influenza, and in ferret models of SARS-CoV-2 transmission, it is on the order of a day (19, 20). Last, high-sensitivity screening tests, when applied during the viral decline accompanying recovery, are unlikely to substantially affect transmission because such individuals detected have low, if any, infectiousness (14). A recent review by Cevik et al. (18) notes that no study to date has successfully cultured live virus more than 9 days after the onset of symptoms.

RESULTS

Impact of repeated population screening on individuals

To examine how repeated population screening would reduce the average infectiousness of individuals, we first modeled the viral loads and infectiousness curves of 10,000 simulated individuals using the
predicted viral trajectories of SARS-CoV-2 infections based on key features of latency, growth, peak, and decline identified in the literature (Fig. 1A; see Materials and Methods). Accounting for these within-host viral kinetics, we calculated what percentage of their total infectiousness would be removed by screening and isolation (Fig. 1B) with tests at LOD of $10^3$ and $10^5$ and at different testing frequencies. Here, infectiousness was taken to be proportional to the logarithm of viral load in excess of $10^6$ cp/ml (with alternative assumptions addressed in sensitivity analyses; see the Supplementary Materials), consistent with the observation that presymptomatic patients are most infectious just before the onset of symptoms (21) and evidence that the efficiency of viral transmission coincides with peak viral loads, which was also identified during the related 2003 SARS outbreak (22, 23). We considered that 35% of patients would undergo asymptomatic isolation within 3 days of their peak viral load if they had not been tested and isolated first, and 65% would have sufficiently mild or no symptoms such that they would not isolate unless they were detected by testing. On the basis of recent results, we modeled asymptomatic and symptomatic infections as having the same initial viral loads (1, 24–26) but with faster clearance among asymptomatics (see Materials and Methods) (24, 26–29).

This analysis demonstrated that there was little difference in averting infectiousness between the two classes of test. Marked reductions in total infectiousness of the individuals were observed by testing daily or every third day, 62 to 66% reduction when testing weekly, and 45 to 47% under biweekly testing (Fig. 1C). Because viral loads and infectiousness vary across individuals, we also analyzed the impact of different screening regimens on the distribution of individuals’ infectiousness, revealing that more sporadic testing leads to an increased likelihood that individuals will test positive after they are no longer infectious or be missed by testing entirely (Fig. 1D).

**Impact of repeated screening on a population**

Above, we assumed that each infection was independent. To investigate the effects of population screening strategies at the population level, we used simulations to monitor whether epidemics were contained or became uncontrolled while varying the frequencies at which the test was administered, ranging from daily testing to testing every 14 days, and considering tests with LOD of $10^3$ and $10^5$, analogous to RT-qPCR and RT-LAMP/rapid antigen tests, respectively. We integrated individual viral load trajectories into two different epidemiological models to ensure that important observations were independent of the specific modeling approach. The first model is a simple fully mixed model representing a population of 20,000, similar to a large university setting, with a constant rate of external infection approximately equal to one new import per day. The second model is a previously described agent-based model with both within-household and age-stratified contact structure based on census microdata in a city representative of New York City (30), which we initialized with 100 cases without additional external infections. Individual viral loads were simulated for each infection, and individuals who received a positive test result were isolated, but contact tracing and monitoring were not included to more conservatively estimate the impacts of screening alone (31, 32). Model details and parameters are fully described in Materials and Methods.

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

The relationship between test sensitivity and the frequency of testing required to control outbreaks in both the fully mixed model and the agent-based model generalize beyond the examples shown in Fig. 2 and are also seen at other testing frequencies, sensitivities, and asymptomatic fractions. We simulated both models at LODs of $10^3$, $10^5$, and $10^6$ and for testing ranging from daily to every 14 days. For those, we measured each population screening policy’s impact on total infections (Fig. S1, A and B) and on $R$ (Fig. S1, C and D). In Fig. 2, we modeled infectiousness as proportional to $\log_{10}$ of viral load. To address whether these findings are sensitive to this modeled relationship, we performed similar simulations with infectiousness proportional to viral load (Fig. S2) or uniform above $10^6$/ml (Fig. S3). We found that results were robust to these large variations in the modeled relationship between infectiousness and viral load. To further address whether our results depended on the exact 35% fraction of individuals assumed to be behaviorally symptomatic, we performed sensitivity analyses with fewer (20%) or more (50%) symptomatic individuals and found no meaningful difference in results (Fig. S4).

**Impact of delayed test results**

An important variable in testing is the time between a test’s sample collection and the reporting of a diagnosis. To examine how time to reporting affected epidemic control, we reanalyzed both the reduction in individuals’ infectiousness and the epidemiological simulations, comparing the results of instantaneous reporting (reflecting a rapid point-of-care assay), 1-day delay, and 2-day delay (Fig. 4, A and B). Delays in reporting markedly decreased the reduction in infectiousness in individuals as seen by the total infectiousness removed (Fig. 4C), the distribution of infectiousness in individuals (Fig. 4D), or the dynamics of the epidemiological models (Fig. 5). This result was robust to the modeled relationship between infectiousness and viral load in both simulation models and for various test sensitivities and frequencies (Fig. S5). These results highlight that delays in reporting lead to markedly less effective control of viral spread and emphasize that fast reporting of results is critical in any screening regimen. These results also reinforce the relatively smaller benefits of improved LODs.

**Generality of findings to changes in modeling assumptions**

Communities vary in their transmission dynamics because of difference in rates of imported infections and in the basic reproductive number $R_0$, both of which will influence the frequency and sensitivity with which screening tests must occur. We performed two analyses to illustrate this point. First, we varied the rate of external infection in our fully mixed model and confirmed that when the external rate of infection is higher, more frequent screening is required to prevent outbreaks (Fig. S6A). Second, we varied the reproductive number $R_0$ between infected individuals in both models and confirmed that at higher $R_0$, more frequent screening is also required (Fig. S6, B and C). This may be relevant to institutions like college campuses or military bases wherein frequent classroom setting or dormitory living is likely to increase contact rates. Thus, the specific strategy for successful population screening will depend on the current community infection prevalence and transmission rate.

The generality of our findings to different epidemiological parameters (Fig. S6), relationships between viral load and infectiousness

**Fig. 2. Repeated population screening affects disease dynamics.** Both the fully mixed compartmental model (top row) and agent-based model (bottom row) are affected by repeated population screening. (A and B) More frequent testing reduces the effective reproductive number $R$, shown as the percentage by which $R_0$ is reduced, $100 \times (R_0 - R)/R_0$. Values of $R$ were estimated from 50 independent simulations of dynamics with 100% of the population participating (see Materials and Methods). (C and D) Relative to no testing (gray bars), screening suppresses the total number of infections in both models when testing every day or every 3 days but only partially mitigates total cases for weekly or biweekly testing. Error bars indicate inner 95% quantiles of 50 independent simulations each.

Fig. 3. Example simulation trajectories from fully mixed model with repeated population screening. Simulation trajectories show the number of infected individuals in a population of $N = 20,000$ with a constant rate of external infection set to $1/N$ per person per day, i.e., around one imported case per day, and full participation in the testing regimen. Infections are classified as freely mixing in the population (blue), isolated because of a positive test (black), or isolated because of symptoms (red) in four simulated example scenarios with $R_0 = 2.5$. (A) No screening. (B) Weekly testing at LOD $10^3$. (C) Weekly testing at LOD $10^5$. (D) Testing every 3 days with LOD $10^5$. Note the variation in the vertical axis scales. The models are fully described in Materials and Methods.

Fig. 4. Effectiveness of screening is compromised by delays in reporting. (A) An example viral load trajectory is shown with LOD thresholds of two tests and a hypothetical positive test on day 6 but with results reported on day 8. Twenty other stochastically generated viral loads are shown to highlight trajectory diversity (light gray; see Materials and Methods). (B) Relative infectiousness for the viral load shown in (A) pretest (totaling 35%; blue) and posttest but before diagnosis (totaling 34%; green) and after isolation (totaling 31%; black). (C) Population screening programs using tests at LODs of $10^3$ and $10^5$ at frequencies indicated and with results returned after 0, 1, or 2 days (indicated by small text beneath bars) were applied to 10,000 individuals’ trajectories of whom 35% were symptomatic and self-isolated after peak viral load if they had not been tested and isolated first. Total infectiousness removed during screening (colors) and self-isolation (hatch) is shown, relative to total infectiousness with no screening or self-isolation. Delays substantially affect the fraction of infectiousness removed. (D) The impact of screening with delays in returning diagnosis of 0, 1, or 2 days (small text beneath the axis) on the infectiousness of 100 individuals is shown for each population screening regimen and no testing, as indicated, with each individual colored by test if their infection was detected during infectiousness (medians, black lines) or colored blue if their infection was missed by screening or diagnosed positive after their infectious period (medians, blue lines). Units are arbitrary and scaled to the maximum infectiousness of sampled individuals. a.u., arbitrary units.
Repetitive population screening to mitigate an ongoing epidemic

The impact of repetitive population screening on transmission dynamics led us to hypothesize that testing could be used as an active tool to mitigate an ongoing epidemic. To test this idea, we simulated an outbreak situation using both the fully mixed and agent-based models but with three additional conditions. First, we assumed that in an ongoing pandemic, other mitigating interventions would cause the reproductive number to be lower although nevertheless larger than one. Second, we considered that not all individuals would want to or be able to participate in a SARS-CoV-2 screening program. Third, we assumed that the collection of samples for testing, if performed on a large scale, could result in imperfect sample collection, causing an increase in the false-negative rate, independent of an assay’s analytical sensitivity. These modifications are fully described in Materials and Methods.

We simulated epidemics in which screening began only at the point when uncontrolled infections reached 4% prevalence. On the basis of results from our previous analyses, we considered a less sensitive but rapid test with LOD $10^5$ cp/ml and a 0-day delay in results. We then examined scenarios of testing every 3 days and every 7 days, with either 50 or 75% of individuals participating, starting from a partially mitigated $R_0 = 1.5$. We found that testing 75% of individuals every 3 days was sufficient to drive the epidemic toward extinction within 6 weeks and reduce cumulative incidence by 88% and that other combinations also had successful but less rapid mitigating impacts, particularly when compared with no intervention (Fig. 6). Notably, even weekly testing with 50% participation was able to reduce the peak length of the outbreak, illustrating how even partial screening using a test with 100× lower molecular sensitivity than PCR can have public health benefits when used frequently (Fig. 6). Repeating these simulations using a test with LOD $10^6$ led to similar results (Fig. S8). To further generalize these results, we modified our mathematical formula to predict the impacts of per-individual test refusal and per-test sampling-related sensitivity on the reproductive number $R$ (see text S1).

**DISCUSSION**

Our results lead us to conclude that repetitive population screening of asymptomatic individuals can be used to limit the spread of SARS-CoV-2. However, our findings are subject to a number of limitations. First, the sensitivity of a test may depend on factors beyond LOD, including manufacturer variation and improper clinical sampling (33), although the latter may be ameliorated by different approaches to sample collection, such as saliva-based testing (34). Second, the exact performance differences between testing schemes will depend on whether our model truly captures viral kinetics and infectiousness profiles (31), particularly during the acceleration phase between exposure and peak viral load. Continued clarification of these within-host dynamics would increase the impact and value of this, and other (31, 32, 35, 36) modeling studies. Last, we modeled participation in screening regimens (or refusal thereof) as statistically independent between individuals, but health-related behaviors...
Fig. 6. Repeated population screening suppresses an ongoing epidemic. Widespread testing and isolation of infected individuals drive prevalence downward for both (A) the fully mixed compartmental model and (B) the agent-based model. Time series of prevalence, measured as the total number of infectious individuals, are shown for no intervention (solid) and population screening scenarios (various dashed lines; see legend) for individual stochastic simulations. Screening began only when prevalence reached 4% (box), and time series are shifted such that testing begins at \( t = 0 \). Scenarios show the impact of a test with LOD \( 10^3 \), no delay in results, and with 10% of samples assumed to be incorrectly collected (and therefore negative) to reflect decreased sensitivity incurred at sample collection in a mass testing scenario. Annotations show total number of post-intervention infections, as a percentage of the no-intervention scenario, labeled as 100% (see fig. S8 for identical simulations using a test with LOD \( 10^3 \)).

Fig. 7. Example asymptomatic and symptomatic viral loads with model control points. Examples of model viral loads (lines) and corresponding stochastically drawn control points (squares and circles) are shown for (A) an asymptomatic viral load trajectory and (B) a symptomatic viral load trajectory. Because simulations took place in discrete time, black dots show points at which this example viral load would have been sampled. Light gray lines show 20 alternative trajectories in each panel to illustrate the diversity of viral loads drawn from the simple model. Red circles indicate the control points that are modified in symptomatic trajectories to account for symptom onset and prolonged time until clearance.

have been shown to be socially (37) and geographically (38, 39) correlated. Clustered refusal of testing or refusal to isolate upon testing positive could present challenging barriers to implementation.

Our findings show that the impact of repeated population screening can be expressed as a reduction of the reproductive number \( R \). By mapping a given testing regimen to a reduction in \( R \), the impact of testing regimen can be approximated and generalized without complicated simulations. For instance, one could estimate the maximum allowable turnaround time delays or the minimum testing frequency required to bring \( R \) below one on the basis of user-specified and scenario-specific assumptions. To facilitate such generalizations and scenario planning, open-source calculation tools accompany this manuscript.

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

The difference between clinical and screening tests highlights the need for additional tests to be approved and used for screening. These tests should not be held to the same degree of sensitivity as clinical tests, particularly if doing so encumbers rapid deployment of faster cheaper SARS-CoV-2 assays. We suggest that the U.S. Food and Drug Administration, other agencies, or state governments encourage the development and use of alternative faster and lower-cost tests for public health and repeated population screening purposes, even if they have poorer LODs. If the availability of point-of-care or self-administered screening tests leads to faster turnaround time or more frequent testing, then our results suggest that they would have high epidemiological value.

Our modeling suggests that some types of repeated population screening will subject some individuals to unnecessary quarantine days. For instance, the infrequent use of a sensitive test will identify not only (i) those with a low viral load in the beginning of the infection, who must be isolated to limit viral spread, but also (ii) those in the recovery period, who still have detectable virus or RNA but are below the infectious threshold (13, 14). Isolating this second group of patients will have no impact on viral spread but will incur costs of isolation, as would the isolation of individuals who received a false-positive test result due to imperfect test specificity. The use of serology, repeat testing 24 or 48 hours apart, or some other test to distinguish low-viral load patients on the upslope of infection from those in the recovery phase could allow for more effective quarantine decisions.

MATERIALS AND METHODS

Viral loads

Viral loads were drawn from a simple viral kinetics model intended to capture (i) a variable latent period, (ii) a rapid growth phase from
the lower limit of PCR detectability to a peak viral load, (iii) a slower decay phase, and (iv) prolonged clearance for symptomatic infections versus asymptomatic infections. These dynamics were based on the following observations.

Latent periods before symptoms have been estimated to be around 5 days (40). Latent periods before detection via virological tests at secondary sites of replication or shedding have been estimated to be up to 4 days (41), corresponding to a latent or eclipse phase observed with other viruses (42). Viral load appears to peak before symptom onset (21) and peaks within 2 days of challenge in a macaque model (43, 44), although it should be noted that macaque challenge doses were high. Viral load decreases monotonically from the time of symptom onset (21, 45–48) but may be high and detectable 3 or more days before symptom onset (1, 49). Peak viral loads are difficult to measure because of lack of prospective sampling studies of individuals before exposure and infection, but viral loads have been reported in the range of $O(10^5)$ to $O(10^3)$ cp/ml (12, 47, 48). Viral loads appear to become undetectable by PCR within 3 weeks of symptom onset (45, 48, 50), but detectability and timing may differ, depending on the degree of presence of symptoms (50, 51). The majority of studies reviewed by Cevik et al. (18) found initial viral loads to be similar between symptomatic and asymptomatic infections (1, 24–26), but viral clearance was significantly and substantially faster among asymptomatic infections (24, 26–29).

Last, we note that the general understanding of viral kinetics may vary depending on the mode of sampling, as demonstrated via a comparison between sputum and swab samples (12). For a comprehensive review of viral load dynamics, duration of shedding, and infectiousness, see (18).

To mimic growth and decay, $\log_{10}$ viral loads were specified by a continuous piecewise linear “hinge” function, specified uniquely with three control points: $(t_0, 3), (t_{peak}, V_{peak}), (t_6, 6)$ (Fig. 7A, green squares). The first point represents the time at which an individual’s viral load first crosses $10^3$, with $t_0 \sim \text{unif}[2.5, 3.5]$, measured in days since exposure. The second point represents the peak viral load. Peak height was drawn $V_{peak} \sim \text{unif}[7, 11]$, and peak timing was drawn with respect to the start of the exponential growth phase, $t_{peak} - t_0 \sim \text{gamma}(1.5)$ with a maximum of 3. The third point represents the time at which an individual’s viral load crosses beneath the $10^6$ threshold, at which point viral loads no longer cause active cultures in laboratory experiments (11–13, 18). For asymptomatic infections, this point was drawn with respect to peak timing, $t_f - t_{peak} \sim \text{unif}[4, 9]$. For symptomatic infections, a symptom onset time was first drawn with respect to peak timing, $t_{symptoms} - t_{peak} \sim \text{unif}[0.3]$, and then, the third control point was drawn with respect to symptom onset, $t_f - t_{symptoms} \sim \text{unif}[4.9]$. Thus, symptomatic trajectories are systematically longer, in both duration of infectiousness (see below) and duration of viral shedding, reflecting the documented prolonged clearance and relationship with viral culture experiments (Fig. 7B, red circles). In simulations, each viral load’s parameters were drawn independently of others, and the continuous function described here was evaluated at 28 integer time points (Fig. 7, black dots) representing a 4-week span of viral load values.

Infectiousness

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

We note that a comprehensive review of viral loads, shedding, and infectiousness (18) found that, across the surveyed literature, no virus could be cultured beyond 9 days after symptoms. Thus, the choice of the final control point in our symptomatic viral load model (Fig. 7B), which corresponds to the latest time at which an individual is infectious, is, at most, 9 days after symptom onset.

Recently, He et al. (21) published an analysis of infectiousness relative to symptom onset that was corrected by Bonhoeffer et al. [see (21) for details]. Among our infectiousness functions, this inferred relationship bears the greatest similarity, over time, to the log-proportional infectiousness function, as visualized in Figs. 1 and 4. The proportional and threshold models therefore represent one of many types of sensitivity analysis. Results for those models can be found in figs. S2, S3, and S5. In all simulations, the value of the proportionality constant implied by the infectiousness functions above was chosen to achieve the targeted value of $R_0$ for that simulation and confirmed via simulation as described below.

Disease transmission models

Overview

Two models were used to simulate SARS-CoV-2 dynamics, both based on a typical compartmental framework. The first model was a fully mixed model of $N = 20,000$ individuals with all-to-all contact structure, zero initial infections, and a constant $1/N$ per-person probability of becoming infected from an external source. This model could represent, for instance, a large college campus with high mixing, situated within a larger community with low-level disease prevalence. The second model was an agent-based model of $N = 8.4$ million agents representing the population and contact structure of New York City, as previously described (30). Contact patterns were based on a combination of individual-level household contacts drawn from census microdata and age-stratified contact matrices, which describe outside of household contacts. This model was initialized with 100 initial infections and no external sources of infection.

Both the fully mixed and agent-based models tracked discrete individuals who were Susceptible (S), Infected (I), Recovered (R), Isolated (I), and Self-Isolated (SQ) at each discrete 1-day time step. Upon becoming infected ($S \rightarrow I$), a viral load trajectory $V(t)$ was drawn, which included a latent period, growth, and decay. Each day, an individual’s viral load trajectory was used to determine whether their diagnostic test would be positive if administered, as well as their infectiousness to susceptible individuals. On the basis of a schedule of testing each person every $D$ days, if an individual happened to be tested on a day when their viral load exceeded the LOD of the test, then their positive result would cause them to isolate ($I \rightarrow Q$), but with the possibility of a delay in turnaround time. A fraction $1-f$ of individuals self-isolate on the day of symptom onset, which occurs 0 to 3 days after peak viral load, to mimic symptom-driven isolation ($I \rightarrow SQ$), with $f = 0.65$ for both models and with $f = 0.8$ and $f = 0.5$ explored in sensitivity analyses (fig. S4).
Presymptomatic individuals were isolated before symptom onset only if they received positive test results. When an individual’s viral load dropped below $10^3$, that individual recovered ($I, Q, SQ \rightarrow R$). Details follow.

**Testing, isolation, and sample-to-answer turnaround times**

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

**Transmission, population structure, and mixing patterns: Fully mixed model**

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

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

**Transmission, population structure, and mixing patterns: Agent-based model**

The agent-based model added viral kinetics and testing policies (as described above) to an existing model for SARS-CoV-2 transmission in New York City. A full description of the agent-based model is available (30); here, we provide an overview of the relevant transmission dynamics.

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

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

Then, transmission from infected individuals to susceptible individuals was simulated both within and outside households. To model within-household transmission, each individual had a set of other individuals comprising their household. Household structures, along with the age of each individual, were sampled from census microdata for New York City (52). The probability for an infectious individual to infect each of their household members each day was determined by scaling the relative infectiousness values to match the estimated secondary attack rate for close household contacts previously reported in case cluster studies (53).

Outside of household transmission was simulated using age-stratified contact matrices, which describe the expected number of daily contacts between an individual in a given age group and those in each other age group. Each infectious individual of age $i$ drew Poisson($M_{ij}$) contacts with individuals in age group $j$, where $M$ is the contact matrix. The contacted individuals were sampled uniformly at random from age group $j$. We use a contact matrix for the United States estimated in (54). Each contact resulted in infection, $S \rightarrow I$, with probability proportional to the relative infectiousness of the infected individual on that day, scaled to obtain the specified value of $R_0$. To conclude each time step, individuals’ viral loads and test results counters were advanced, with those whose infectious period had completely passed moved to recovery, $I, Q \rightarrow R$.

**Calibration to achieve targeted $R_0$ and estimation of $R_0$**

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

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

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

Estimations of $R$ proceeded exactly as estimations of $R_0$ for both models, except with interventions applied to the viral loads and, therefore, the dynamics. Prediction of $R$ without direct simulation is described in text S1.

**SUPPLEMENTARY MATERIALS**

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/sciadv.abd5393/DC1

**View/request a protocol for this paper from Bio-protocol.**

**REFERENCES AND NOTES**


Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening

Daniel B. Larremore, Bryan Wilder, Evan Lester, Soraya Shehata, James M. Burke, James A. Hay, Milind Tambe, Michael J. Mina and Roy Parker

Sci Adv 7 (1), eabd5393.
DOI: 10.1126/sciadv.abd5393 originally published online November 20, 2020

Use of this article is subject to the Terms of Service.