medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Controlling COVID-19 via test-trace-quarantine

Cliff C. Kerr[1*], Dina Mistry[1†], Robyn M. Stuart[2,3†], Katherine Rosenfeld[1], Gregory R. Hart[1], Rafael C. Núñez[1], Jamie A. Cohen[1], Prashanth Selvaraj[1], Romesh G. Abeysuriya[3], Michał Jastrzębski[4], Lauren George[1], Brittany Hagedorn[1], Jasmina Panovska-Griffiths[5,6], Meaghan Fagalde[7], Jeffrey Duchin[7], Michael Famulare[1], and Daniel J. Klein[1]

[1] *Institute for Disease Modeling, Global Health Division, Bill and Melinda Gates Foundation, Seattle, WA, USA*
[2] *Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark*
[3] *Burnet Institute, Melbourne, VIC, Australia*
[4] *GitHub, Inc., San Francisco, CA, USA*
[5] *Department of Applied Health Research, University College London, London, UK*
[6] *Wolfson Centre for Mathematical Biology and The Queen's College, Oxford University, Oxford, UK*
[7] *Seattle-King County Health Authority, Seattle, WA, USA*

** Corresponding author. E-mail: ckerr@idmod.org*
† Contributed equally

Abstract

Initial COVID-19 containment in the United States focused on limiting mobility, including school and workplace closures. However, these interventions have had enormous societal and economic costs. Here we demonstrate the feasibility of an alternative control strategy, test-trace-quarantine: routine testing of primarily symptomatic individuals, tracing and testing their known contacts, and placing their contacts in quarantine. We performed this analysis using Covasim, an open-source agent-based model, which was calibrated to detailed demographic, mobility, and epidemiological data for the Seattle region from January through June 2020. With current levels of mask use and schools remaining closed, we found that high but achievable levels of testing and tracing are sufficient to maintain epidemic control even under a return to full workplace and community mobility and with low vaccine coverage. The easing of mobility restrictions in June 2020 and subsequent scale-up of testing and tracing programs through September provided real-world validation of our predictions. Although we show that test-trace-quarantine can control the epidemic in both theory and practice, its success is contingent on high testing and tracing rates, high quarantine compliance, relatively short testing and tracing delays, and moderate to high mask use. Thus, in order for test-trace-quarantine to control transmission with a return to high mobility, strong performance in all aspects of the program is required.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Introduction

Within 12 months of the world first becoming aware of COVID-19, the total number of diagnosed cases exceeded 80 million (*1*), with the true number of infections likely much higher. As the pandemic has evolved, so too have global public health responses. Many of the initial efforts to contain the spread focused on border controls, but when these proved insufficient to prevent community transmission, the focus turned to broad non-pharmaceutical interventions, such as lockdowns and other physical distancing measures. Whilst effective in controlling infection rates (*2*), such measures have come at enormous cost (*3*). Consequently, governments are increasingly relaxing lockdowns in favor of more targeted "test-and-trace" strategies, whereby only those most likely to have COVID-19 – for example, those who have symptoms, or who have been in contact with a confirmed case – are required to quarantine. Such strategies have the potential to offer the epidemiological benefits of a large-scale lockdown with smaller economic and societal costs.

Several studies have examined test-and-trace-based containment strategies of COVID-19 in different contexts. Modeling studies have provided evidence that the success of such strategies depends on the proportion of symptomatic cases, the speed and completeness of contact tracing, and adherence to isolation and quarantine (*4–8*). Given an estimated basic reproduction number (R_0) of 2.4–5.6 (*9*), the number of effective contacts must be reduced by at least 60–80% to achieve epidemic control. Despite this stringent requirement, China successfully demonstrated the feasibility of epidemic control through mandatory home-based quarantine and isolation of those with confirmed infections. This strategy was successfully followed by South Korea, Singapore, and other countries (*10*). However, success has not been universal, and some countries have had to re-impose restrictions due to epidemic resurgence after relaxing social distancing restrictions (*11*, *12*).

These studies show that there are global precedents for the success of containment strategies based on (a) high rates of routine testing (including both symptomatic and asymptomatic), (b) rapid return of test results, (c) high rates of contact tracing, and (d) social support for people who have been diagnosed or quarantined – a strategy we refer to as test-trace-quarantine (TTQ). However, the success of this strategy depends on how effectively each component is implemented. To date, the COVID-19 response in the United States (and other Western countries, such as the United Kingdom) has been marked by insufficient quantities of test kits and associated supplies, along with challenges in implementing contact tracing at scale and imperfect adherence to isolation guidelines (*13*). Increasing coverage of infection-blocking vaccines will make control easier (*14*), but this effect is likely to be at least partially countered by the higher transmissibility, and potential immune escape, of emerging SARS-CoV-2 variants (*15*).

This study investigates what the requirements are for a high-income, urbanized setting to successfully transition from a policy of mobility restrictions towards TTQ-based containment prior to high vaccine coverage. To answer this question, used Covasim, a detailed, data-driven, agent-based model of COVID-19, and applied it to the Seattle context (specifically King County, which includes Seattle and the surrounding metropolitan area).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Results

Mobility restrictions achieved initial epidemic control

The first case of SARS-CoV-2 in the USA was diagnosed on 20 January 2020 in the Seattle metropolitan area (*16*); the first documented COVID-19 death in the USA was on 27 February, a resident of a long-term care facility (LTCF) also in the Seattle area (*17*). Local and state governments began issuing a series of measures to control the expanding epidemic, including school closures on 12 March and a shelter-in-place order ("Stay Home, Stay Healthy") from 23 March until 31 May 2020 (*18*).

We fit the Covasim model to age-stratified data on COVID-19 diagnosed cases and deaths in Seattle from January through June 2020 using Optuna, a parameter optimization library (*19*). Detailed demographic information (including population age structure and contact patterns, school enrolment, employment, and LTCF residency), mobility data (provided by SafeGraph; see http://safegraph.com), and COVID-19 testing data were used as inputs. As shown in Fig. 1, Covasim was able to accurately reproduce the detailed time trends of both diagnoses and deaths (Fig. 1A and B), including the age distribution of each (insets). We estimate that approximately 100,000 SARS-CoV-2 infections (95% confidence interval: 80,000–115,000 infections) occurred in Seattle between 27 January and 9 June 2020 (Fig. 1C), out of a total population of 2.25 million, for an attack rate (cumulative infections divided by population size) of 3.5–5.1%. A total of 8,548 cases had been diagnosed by June 9, for an overall diagnosis rate of 9% (95% CI: 7–11%). The effective reproduction number, R_e, is estimated to have been 2.3 (95% CI: 2.0–2.6) prior to policy interventions, consistent with previous estimates (*20*), and to have dropped below 1 as the shelter-in-place order took effect (Fig. 1D). This period also coincided with the peak number of active infections, 16,000, with model projections validated by prevalence data from the Seattle Coronavirus Assessment Network (Fig. 1C).

Calibrated model parameters, which provide estimates of transmission dynamics and intervention effectiveness, are shown in Fig. 1E. The parameters used for calibration were: overall transmissibility β, defined as the probability of transmission between an infectious and susceptible adult on a single day in a typical household setting; transmission relative to baseline, which may change due to mask usage, hygiene, physical distancing, and other measures; and the odds ratio for people with COVID-19 symptoms being tested versus people without symptoms (i.e., uninfected, asymptomatic, or presymptomatic people). To determine the impact of mobility-related changes in transmission, we calibrated the model both using reductions in the number of work and community contacts based on SafeGraph weekly mobility data (M, blue), and using no mobility data (N, red). Including the mobility trends, we found that relative transmissibility was reduced by 12±5% compared to its initial value, reflecting the impact of other NPI, including interpersonal distancing, hygiene, and mask use; this drop is consistent with increasing trends of protective health behaviors (*21*). To verify the calibration, we excluded mobility data and recalibrated the model, finding that relative transmissibility dropped by 71±3% compared to its initial value. All other parameters had consistent values between the two calibrations, including the change in transmission at LTCFs (estimated to have dropped by 80–92%) and overall transmissibility (estimated at 4.3–4.5% per household contact per day). The symptomatic testing odds ratio, reflecting the much higher rate at which people with COVID-like symptoms test, was estimated to be 17–24. While the testing odds ratio remained constant, the routine testing yield (the number of diagnoses divided by the number of tests) showed a decline from 10–15% in March to 1.5–2.5% in early June, due to the much lower number of active infections.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

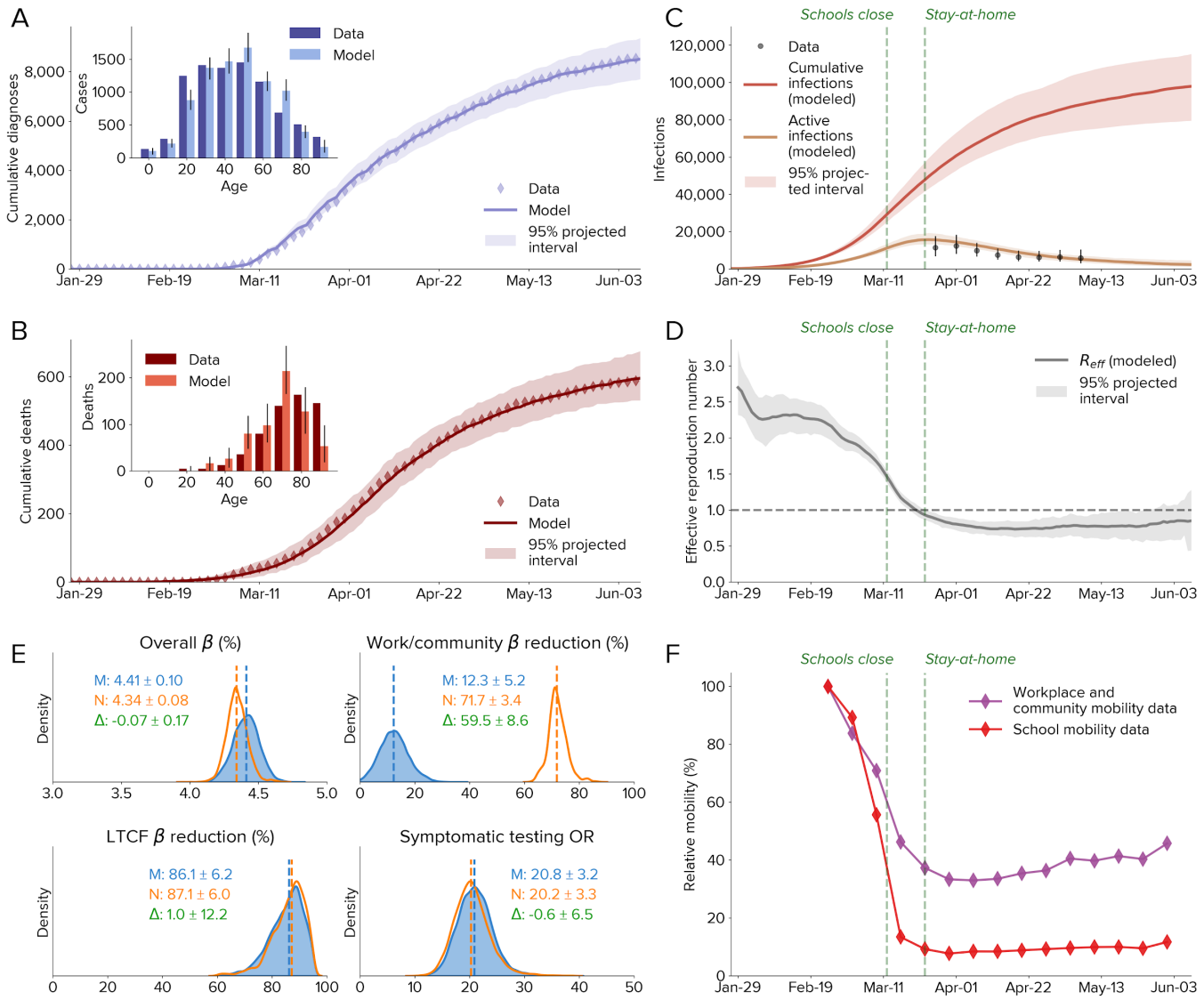

Fig. 1: Calibration of the model to data from Seattle-King County, Washington, from 27 January to 9 June 2020. A–B: The cumulative number of diagnosed cases and deaths, over time and by age. **C:** Estimated numbers of cumulative and active infections. Dashed lines show policy interventions; data are from the Seattle Coronavirus Assessment Network. **D:** Effective reproduction number, showing a drop consistent with policy interventions. **E:** Calibration of model parameters with SafeGraph mobility data (M, blue) and with no mobility data (N, red); differences (Δ, green) are only significant for work/community transmission direction. **F:** SafeGraph mobility data for workplaces and the community and for schools. LTCF, long-term care facility; OR, odds ratio.

Since Covasim includes intra-host viral dynamics and a detailed demographic model, it can be used to investigate mechanisms of transmission, as shown in Fig. 2. We find that infections were primarily driven by transmission in workplace and community contact layers (accounting for approximately 58% of the total) prior to interventions. Surprisingly, even though distancing interventions led to a roughly two-thirds drop in workplace and community mobility (Fig. 1F), the total proportion of infections due to workplace and community transmission reduced only slightly, to 52% (Fig. 2A). This is in part due to the high overdispersion of SARS-CoV-2 infections (Fig. 2B): a majority of people infected do not transmit, while 50% of infections are caused by just 10% of people infected; these 10% infect, on average, 6.3 other people (Fig. 2C). Thus, a relatively small proportion of highly infectious individuals are likely responsible

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

for a majority of ongoing COVID-19 spread. Preliminary data from the contact tracing program in Seattle provides further evidence for this: of the 44% of household contacts who received a COVID-19 test, 43% of them tested positive (i.e., 19% of traced household contacts were positive); of the 31% of non-household contacts who were tested, 28% tested positive (i.e., 9% of traced workplace contacts were positive). High-risk index cases and contacts were preferentially both traced and tested, so these estimates represent an upper bound on the attack rate, and international estimates on household secondary attack rate have been even lower, ranging from 5% (*22*) to 19% (*23*). To be consistent with our estimated value of R_e, these relatively low household attack rates require high dispersion and significant non-household transmission. We also find that 54% of transmissions are from symptomatic individuals, similar to previous estimates (*24*, *25*).

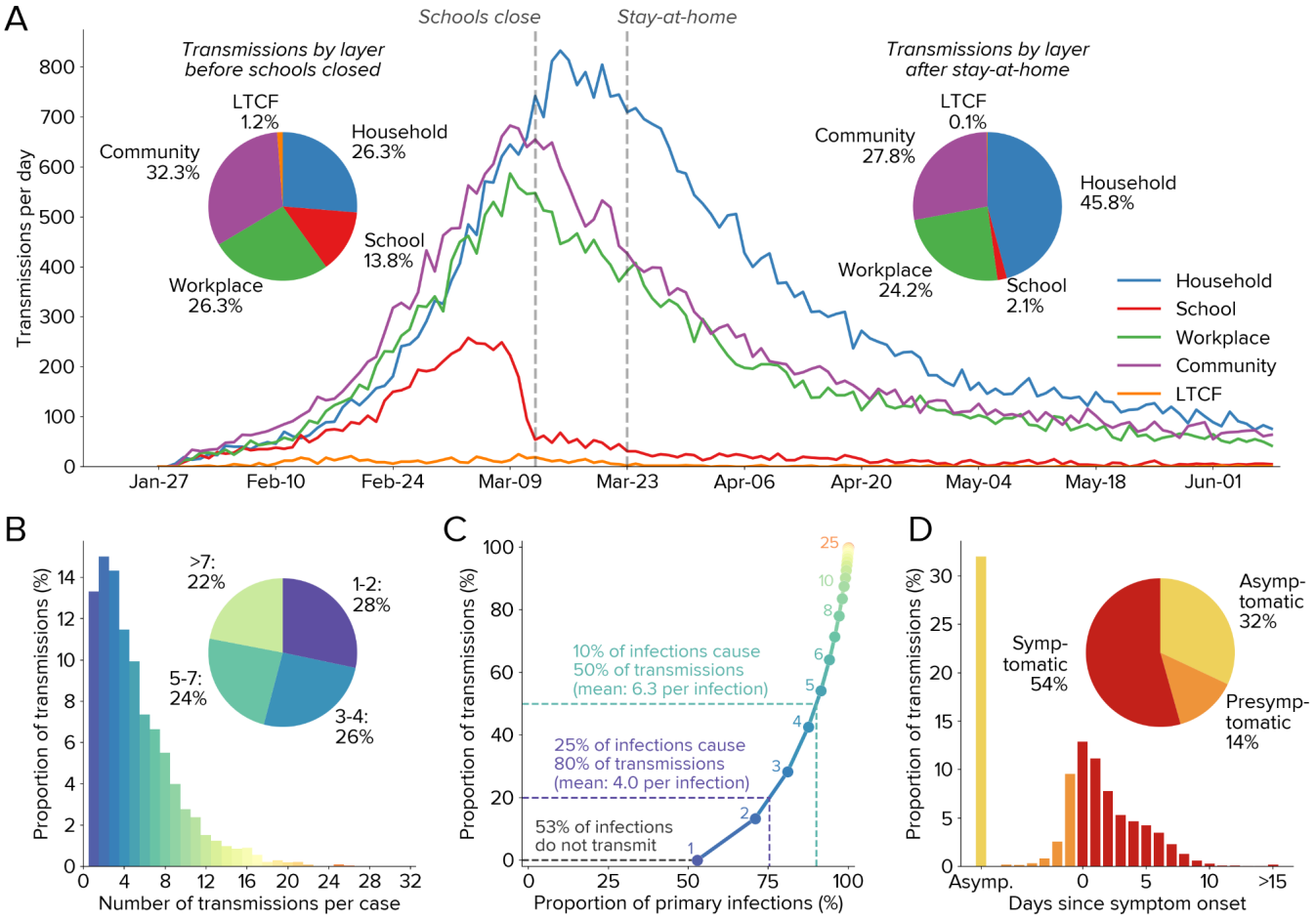

Fig. 2: Modeled transmission dynamics. A: Infections over time by contact layer. **B:** Overdispersion of infections (up until school closures on 12 March), with roughly equal numbers of infections attributable to individuals who transmit to 1–2 others, 3–4 others, 5–7 others, or more than 7 others. **C:** Due to overdispersion, 10% of primary infections are responsible for 50% of secondary infections, while 53% of all primary infections do not cause any secondary infections. Annotations show the number of transmissions per primary infection, corresponding to each bar of panel B. **D:** Infections as a function of symptom onset, showing that slightly over half of infections are transmitted by symptomatic individuals.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Idealized test-trace-quarantine results in self-limiting epidemic dynamics

Before investigating TTQ in the Seattle context, we first consider how TTQ impacts SARS-CoV-2 transmission in a hypothetical population. Consider an idealized TTQ scenario, where all contacts are traced, all traced contacts are tested and enter into 14-day quarantine (regardless of test result), and combined testing and tracing delays are less than the duration of infectiousness (which is also assumed to be the time period when a person would test positive). In this idealized scenario, epidemic control can be achieved even for high values of R_0, regardless of the stage of the epidemic at which the intervention begins. This is because as a branch from a cluster of infections grows, the probability increases that someone from that branch will be diagnosed. When this occurs, idealized contact tracing would identify that branch via a series of steps, including both backwards ("upstream") and forwards ("downstream") infections (*26*), hence removing that branch from the infectious pool.

Since each traced contact who tests positive results in additional traced contacts, contact tracing can be thought of as an "infectious" process on the network. Specifically, if (a) the sum of the testing and tracing delays is less than the average serial interval of SARS-CoV-2; and if (b) the majority of secondary transmissions are successfully traced, diagnosed, and isolated, then the number of traced and diagnosed contacts will spread locally on the network faster than SARS-CoV-2 infections, extinguishing that branch. The number of backwards steps that can be taken is approximately the duration for which someone returns a positive test following infection divided by the sum of testing and tracing delays. Assuming the former is approximately 10–14 days and the latter is approximately 2–4 days, roughly 2–5 backward steps should be achievable, though in practice false negative tests would likely break the chain sooner. However, even with just forward tracing, epidemic control is still theoretically achievable. Several recent studies have produced similar findings (*26–28*).

Figure 3 shows an illustrative example of idealized TTQ resulting in epidemic control. In a hypothetical population of 100 people without interventions, infections continue until herd immunity prevents further spread (Fig. 3A). If a high level of testing and isolation is introduced (15% probability of testing per day for people with symptoms, coupled with 80% effective isolation), the number of infections is only modestly decreased despite nearly half of cases being diagnosed, since a large proportion of transmission occurs before cases are diagnosed. Adding a moderate level of tracing (70% of household contacts, 10% of workplace and school contacts) significantly reduces the number of infections (Fig. 3C), due to the rapid diagnosis of traced contacts and the preventative effect of quarantine.

Crucially, we find that the effectiveness of contact tracing is proportional to incidence, thereby resulting in self-limiting epidemic dynamics. Fig. 3D shows a hypothetical population of 30,000 people in a medium transmission scenario (R_0 = 2.5), where borderline epidemic control ($R_e \approx 1$) can be achieved through either moderate physical distancing alone (i.e., 60% reduction in β), high levels of routine testing and isolation alone (75% daily probability of people with symptoms testing and isolating), or TTQ (8% daily probability of people with symptoms testing, 90% of contacts of diagnosed individuals being traced and quarantined, and 75% probability of testing on entering quarantine). In a low transmission setting (Fig. 3E, R_0 = 2.0), both physical distancing and testing lead to rapid epidemic extinction, while TTQ maintains $R_e \approx 1$. Conversely, in a high transmission setting (Fig. 3F, R_0 = 3.0), TTQ again maintains $R_e \approx 1$, while physical distancing and testing do not achieve epidemic control. This is because distancing and routine testing act like constant multipliers on transmission; they will achieve epidemic control if and only if they bring $R_e < 1$. In contrast, in a TTQ scenario with no capacity constraints, more infections will

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

result in more diagnoses, more contacts being traced, more people placed in quarantine, and more people tested in quarantine. This mechanism limits the size of individual clusters of infections in a TTQ setting, as well as placing an upper bound on epidemic growth even with extremely high transmission scenarios (e.g., R_0 = 5). However, this phenomenon only occurs with sufficient levels of contact tracing: analogous to R_e, self-limiting dynamics only occur if at least one new case is detected on average for each new index case whose contacts are traced. Whether or not this occurs is determined by the probability of contact tracing, the probability of testing in quarantine, the population network structure, and the rate of transmission.

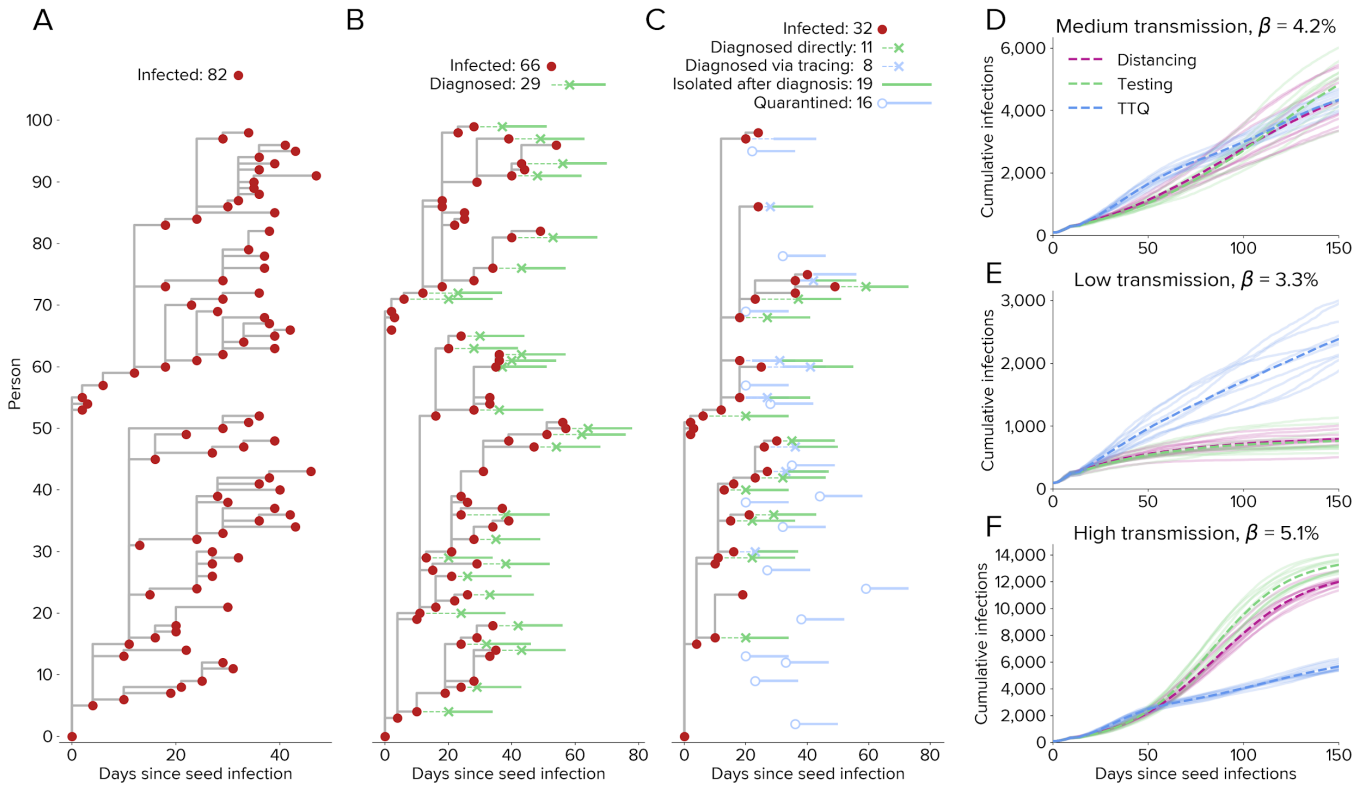

Fig. 3: Epidemic dynamics differ depending on the intervention. A–C: Transmission trees for a cluster of 100 people under three scenarios: (**A**) no interventions, (**B**) testing and isolation only (starting on day 20), and (**C**) test-trace-quarantine. **D–F:** Comparison of interventions for different levels of transmissibility. For medium baseline transmission (**D**), moderate distancing, high testing, or high tracing each result in $R_e \approx 1$. For low transmission (**E**), the same distancing and testing interventions both result in $R_e < 1$, while the same tracing intervention maintains $R_e \approx 1$. For high transmission (**F**), the same distancing and testing interventions both result in $R_e > 1$, while the same tracing intervention continues to maintain $R_e \approx 1$.

Realistic test-trace-quarantine scenarios allow high mobility

Implementing a successful TTQ strategy requires solving a challenging prioritization problem: whom to test, whom to trace, how to ensure people safely isolate and quarantine, how to quickly return test results, and how to quickly trace contacts. Barriers to high performance in each of these areas are varied, including limited budgets, shortages of equipment and staff, behavioral compliance, and racial and economic inequalities.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

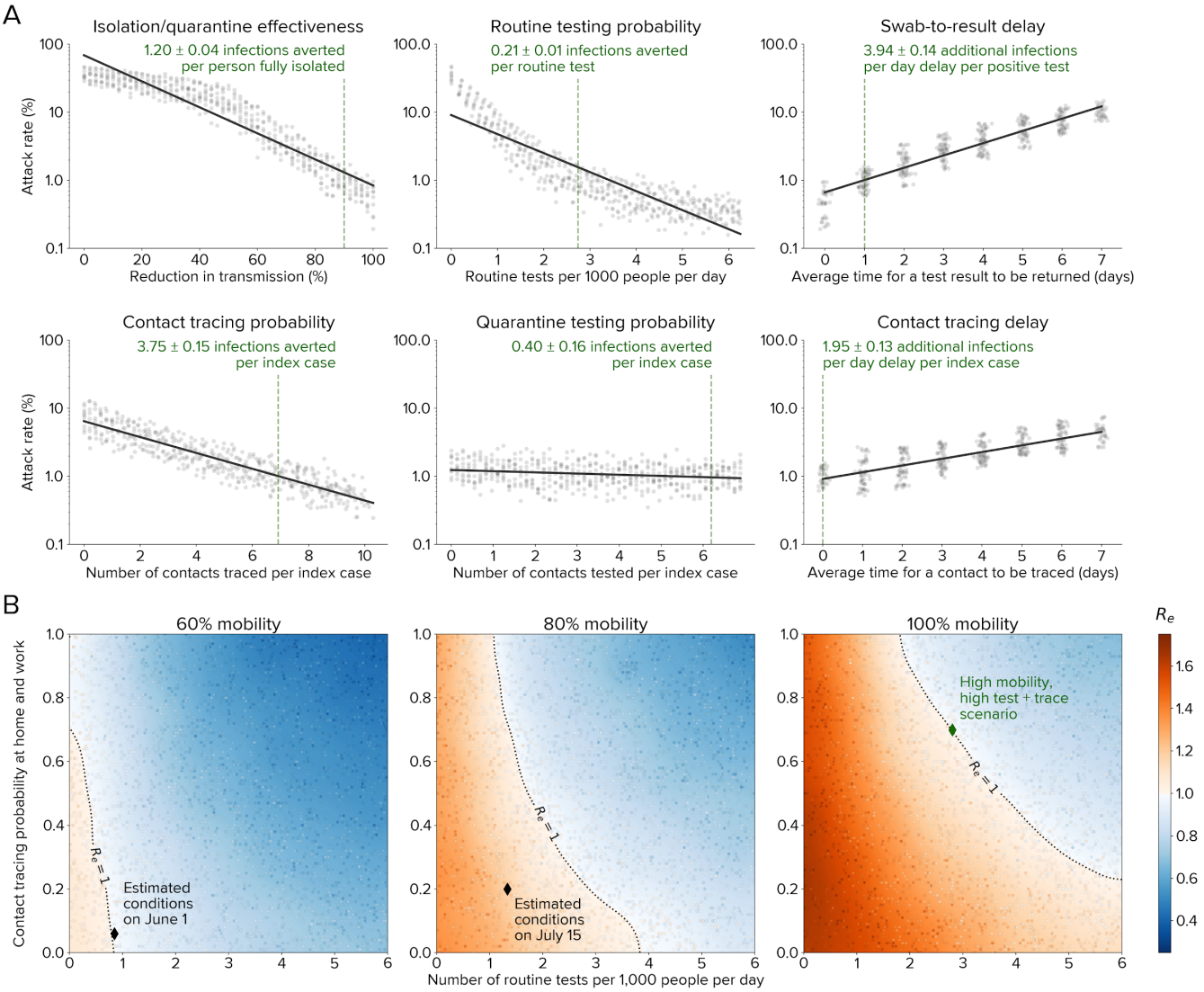

Fig. 4: Impact of testing, tracing, and quarantine. A: Relative importance of different aspects of the TTQ strategy for a scenario of high mobility (full return to baseline workplace and community movement patterns), high testing, and high tracing in Seattle (dashed green lines). Each dot shows a simulation, with other parameters held constant. Isolation/quarantine effectiveness has the greatest impact, with 2.2 infections averted for each person fully isolated, although all parameters have a significant impact on epidemic outcomes. **B:** Countering the effects of increased mobility via testing, tracing, and quarantine. Current interventions (black diamonds) were estimated to keep $R_e < 1$ for 60% of baseline mobility level (left). Subsequently, increased transmission rates exceeded intervention scale-up, leading to $R_e > 1$ temporarily (center). For a return to full mobility (right), high levels of both testing and tracing are required to maintain epidemic control (green diamond, corresponding to the dashed lines in panel A). Dots show individual simulations.

Figure 4A shows how six aspects of the TTQ strategy affect the estimated numbers of infections in Seattle: (1) effectiveness of isolation and quarantine (i.e., relative reduction in transmission during the 14-day isolation/quarantine period), (2) the number of routine tests per 1,000 people per day, (3) the probability of a person's household and workplace contacts being traced following diagnosis, (4) the proportion of contacts who are tested after they are traced, (5) the delay between when a person is tested and when they receive their test result, and (6) the delay between when a person receives a

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

positive test result and when their contacts are traced. Using the model calibrated to the Seattle region up until June 1 and projecting forward for a 90-day period, we consider a hypothetical baseline scenario of high mobility (100%), high testing (~6,000 routine tests per day, or 2.7 per 1,000 people per day, compared to ~1,800 routine tests per day as of June 1 and ~3,500 routine tests per day as of July 15), and high tracing (70% of all household and workplace contacts traced within 2 days, compared to roughly 30% of household contacts and close to 0% of workplace contacts as of June 1). This scenario was chosen as the most realistic means of achieving an R_e value of below 1 given a return to 100% mobility (Fig. 4B, green diamond). We then vary individual aspects of the response relative to this baseline.

All six aspects of the TTQ strategy had a significant impact on epidemic outcomes. The most important aspects, in terms of their impact on overall attack rate, were isolation/quarantine effectiveness and routine testing probability. Each diagnosed or quarantined person who fully isolates is estimated to avert 1.2 subsequent infections over the 90-day period, while each routine test conducted is estimated to avert 0.2 infections. However, speed also matters: an additional one-day delay in returning test results is estimated to result in nearly 4 additional infections per person who tests positive, while a one-day delay in tracing contacts is estimated to result in nearly 2 additional infections for every index case whose contacts are traced. While quarantine testing had the smallest overall impact on the attack rate, an additional 0.4 infections are still estimated to be averted for every index case whose contacts are tested. This is because quarantine testing is highly efficient at identifying infections: while the test positivity rate for routine testing in early June was roughly 1.5%, traced contacts had a test positivity rate of 34%.

To explore the practical implications of these results for Seattle, we simulated three reopening scenarios (Fig. 4B): 60%, 80%, and 100% mobility in the workplace and community contact layers, relative to baseline, across a range of testing and contact tracing rates. We assume these mobility changes occur in the context of continued use of masks and other non-mobility-related NPI, which we estimate have together reduced transmission by roughly 10–15% (see Fig. 1E), which is consistent with roughly 40% efficacy (*29*) coupled with 25–40% compliance. Unsurprisingly, a return to 100% workplace and community mobility in the absence of other interventions would increase R_e to well above 1, leading to a large second wave of infections. High testing and high tracing are both required to maintain epidemic control with full reopening. For example, even a four-fold increase in testing rates would fail to control the epidemic without an increasing in contact tracing.

After the stay-at-home order was lifted on June 1, there was a large increase in transmission rates, especially among younger age groups, which was followed by a large scale-up of both testing and tracing programs. Figure 5 shows that the model calibrated until June 1 was able to replicate observed epidemic trends over this period, providing validation that it is capturing the underlying transmission dynamics in Seattle, as well as the impact of the testing and tracing interventions on the epidemic. Figure 5 also illustrates what would have been achieved had a "high testing and tracing" scenario, as described above, been implemented instead: while the peak numbers of people tested (Fig. 5A) and contacts traced (Fig. 5B) would have been much higher (5,000 vs. 3,500 tests per day and 900 vs. 200 contacts traced per day), active infections would have declined much more rapidly Fig. 5C). By Aug. 31, due to the low number of new infections, we estimate that the number of contacts traced (Fig. 5B) and people diagnosed (Fig. 5D) in the "high testing and tracing" scenario would be comparable to or lower than the true number.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

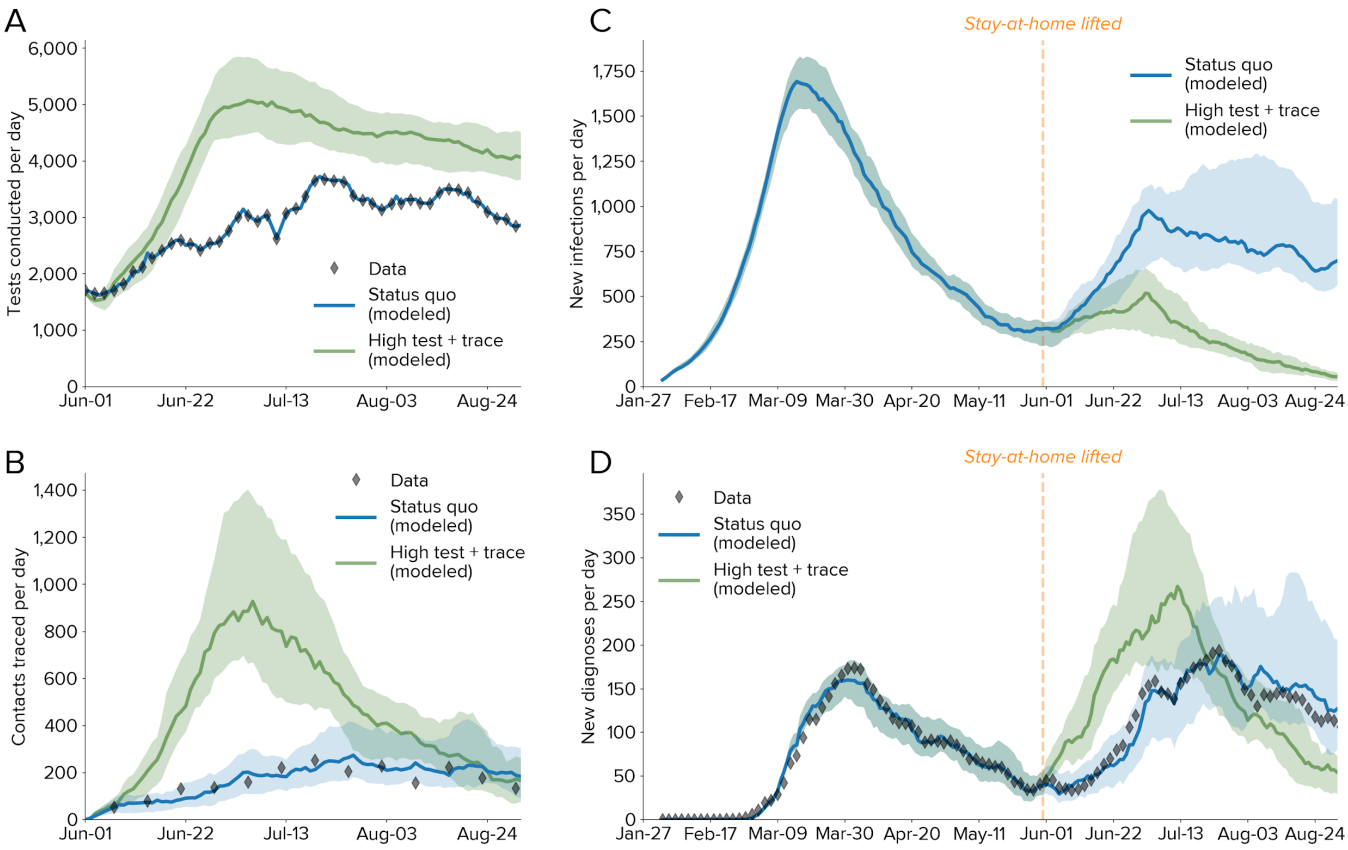

Fig. 5: Comparison between observed epidemic trends and projected scenarios from June 1 to August 31. **A:** Numbers of tests conducted per day, with modeled values for the status quo (using the data as an input) and a counterfactual scenario with high testing and high tracing. Lines show medians; shaded regions show 80% confidence intervals. **B:** Number of contacts traced per day. **C:** Estimated numbers of new infections, with a significant rise in infections observed shortly after the stay-at-home order was lifted. **D:** Number of diagnoses per day, showing consistency between the model and the data both for the calibrated period (Jan. 27 – May 31) and the projected period (Jun. 1 – Aug. 31).

Discussion

Seattle achieved epidemic control between mid March and early June as a result of (a) greatly reduced mobility; (b) adoption of additional NPI, including interpersonal distancing, hand washing, and face masks; and (c) moderate rates of routine testing, plus low but increasing rates of contact tracing. We separated the effects of the first two components using detailed mobility data, and were therefore able to assess the impact of reopening whilst assuming that distancing, hand washing, and face mask use will continue. We found that with mobility at 60% of its pre-COVID levels, epidemic control could be attained despite relatively low testing and tracing. Returning to full mobility would require identifying and isolating many more cases, by significantly increasing both the number of routine tests conducted and the proportion of contacts traced. From June through September, both mobility and interventions increased, and data from this period were used to validate these findings.

The handful of countries that achieved control of their COVID-19 epidemics without relying on extensive lockdowns have done so using a diversity of approaches, but those that have used a broader set of interventions have achieved more stable epidemic control. Taiwan and South Korea used high rates of testing, contact tracing, mask compliance, and other interventions to quickly bring their epidemics under

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

control (*30*). Japan has had high mask compliance and a relatively high rate of contact tracing, but relatively low testing rates; after early control, reported new cases increased from late May through early July (*31*). Australia and New Zealand achieved early epidemic control via strong travel restrictions and high rates of testing and contact tracing. However, mask use was low, and a rapid increase in cases in the state of Victoria, Australia, began in late June, and was only brought under control after reimposed mobility restrictions, mandated mask use, and high testing rates. In contrast, the neighboring state of New South Wales was able to maintain epidemic control without reimposing lockdowns by combining high testing rates (2.9 tests per 1,000 people per day, similar to the high testing and tracing scenario presented here) with near-perfect tracing of close contacts (*32*). This observational evidence for a diverse range of interventions being required for epidemic control is consistent with our finding that each aspect of the response is roughly equally important: shortfalls in one area (e.g., low rates of testing or mask usage) may be partly offset by high performance in another (e.g., high rates of contact tracing). However, these examples suggest that epidemic control may be only fleeting unless performance is high in all three areas (testing, tracing, and either mask usage or mobility restrictions).

As shown in Fig. 3, control via TTQ produces strong positive feedback dynamics. This means that it is an especially effective control strategy when case numbers are low; however, even relatively small spikes in cases can cause the system to be overwhelmed, a result that has been seen both in theory (*33*) and in practice (*11*, *12*). If this occurs, then a return to mobility restrictions remains the only consistently effective strategy for regaining control.

There are several limitations of this study. First, we do not consider geographical clustering or day-of-the-week changes in mobility, so cannot model hotspots or outbreaks in specific areas or on specific days. Although this may not be crucial given that interventions were set at a county-wide level, subsequent phases of the response may require more localized policy actions. Second, there is continued debate around how susceptibility and transmissibility vary by age and with comorbidities; the model parameters reflect the best available evidence to date (*34*), but new evidence is continually coming to light. Third, there is considerable uncertainty around other crucial characteristics of both SARS-CoV-2 transmission (including the extent to which it is seasonal, and the proportion of asymptomatic and presymptomatic transmission) and the impact of interventions (such as mask efficacy). We have handled these uncertainties by calibrating extensively to data (Fig. 1A-C), and by propagating remaining uncertainties in parameters (Fig. 1E) through all scenarios. However, additional data on transmission characteristics would nonetheless help refine our understanding of the most important transmission pathways, as well as how to disrupt them.

In summary, we have shown that (a) agent-based models can be fit to detailed epidemic time trends and age distributions (Fig. 1), as well as make accurate forecasts (Fig. 5); (b) an idealized test-trace-quarantine program with no capacity constraints can control an epidemic even at high rates of transmission (Fig. 3); and (c) high rates of testing and tracing, short delays, and high quarantine compliance are all important for maintaining epidemic control (Fig. 4A), but the levels required for each are likely to be achievable even under a return to full mobility (Fig. 4B). Thus, we believe the example of Seattle provides strong evidence for test-trace-quarantine as a feasible control strategy – a strategy that other jurisdictions may wish to invest in more heavily.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Methods

Model

We performed modeling using Covasim, an agent-based model of COVID-19 transmission and interventions. Covasim's methodology is described in detail in (*35*). Briefly, the Covasim model is designed to capture the nuances of realistic COVID-19 transmission, including: age and population structure, including relative susceptibility and mortality rates; transmission networks in different social layers, including households, schools, workplaces, communities, and long-term care facilities; and viral dynamics reflecting variable infectivity within and between hosts. Covasim also supports an extensive set of interventions: non-pharmaceutical interventions (NPI), such as distancing and masks; testing interventions, such as symptomatic and asymptomatic testing, contact tracing, isolation, and quarantine; and pharmaceutical interventions, such as therapeutics and vaccines. While Covasim was originally developed to inform policy decisions in Washington State, it has since been adapted for use in more than a dozen countries, including Australia (*36*) and the United Kingdom (*37*). The code is fully open source (available at http://github.com/institutefordiseasemodeling/covasim), and includes extensive documentation (available at http://docs.covasim.org), tutorials, and software tests. Analyses were performed with Covasim version 2.0.0. The following sections describe the customizations to Covasim that were used in this study.

Data and sources

As an agent-based model, Covasim can make use of rich data sources. Mechanistic representations of individuals, contacts, and infections enables physical parameter values to be input directly or used as priors during model fitting. Default values for most inputs come from publicly available data and literature, and are described in Tables 1 and 2 of (*35*).

To model the Seattle-King County region for this analysis, we used epidemiological data provided by the Washington State Department of Health (WA-DoH) under a use agreement. (Note that "Seattle" and "King County" are used interchangeably, i.e. the analysis is not restricted to the city of Seattle, nor does it include counties in the greater Seattle metropolitan area such as Snohomish and Pierce counties.) WA-DoH maintains all COVID-19 data as a line-list in the Washington Disease Reporting System (WDRS), and has provided weekly exports to the study team for the purpose of conducting this and other analyses in support of model-based decision making. We aggregated line-list entries to daily totals by 10-year age bins to produce target data for model calibration. The resulting dataset includes the number of positive and negative tests (by date of sample collection) as well as the number of deaths (by date of death) in King County. The WDRS records also enabled us to characterize the distribution of delays between symptom onset on diagnostic swab; we used these data to validate the implementation of the testing intervention (Fig. S11).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

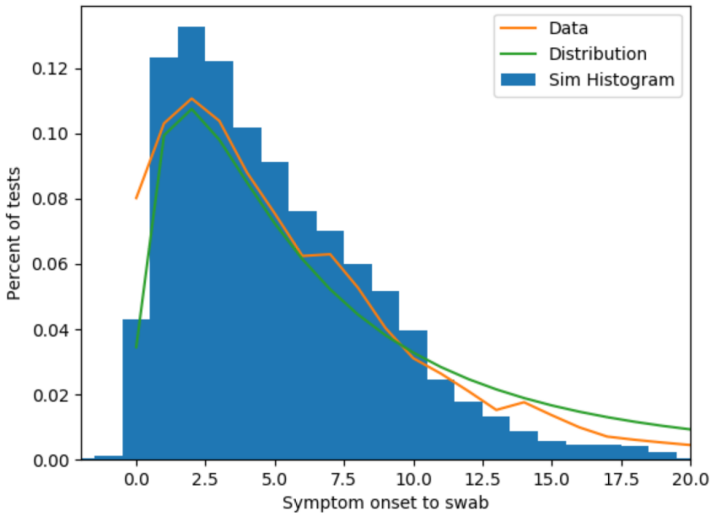

Fig. 6: Comparison of empirical (orange), statistically modeled (green), and simulated (blue) symptom-to-swab delay distributions.

Using data on the number of tests and number of diagnoses, we were able to calibrate the model to testing yield. By combining this with estimated numbers of infections, which we know from both data on deaths as well as independent seroprevalence surveys (as shown in Fig. 1C), we were able to estimate the testing rates of people both with and without COVID-19. According to the most recent data available at the time the analyses were performed (June 9), roughly 1,800 tests per day were being conducted in King County; using this method, we found that a roughly 8% probability per day of testing for people experiencing symptoms, and 0.1% probability per day of testing for people who without symptoms (uninfected, asymptomatic, and presymptomatic), allowed us to match observed values for (a) the overall number of tests, (b) the test positivity rate, and (c) the symptom-to-swab delay.

The partnership with Public Health Seattle & King County (PHSKC) has provided additional context to ensure the model captures transmission, testing, care, and contact tracing in this setting. Public health co-authors and other county officials have provided data, under use agreement, and insights on testing campaigns, focal outbreaks, schools, hospitalization, and congregate settings such as long-term care facilities (LTCF). Most of these data are available publicly on daily summary and LTCF dashboards. Insights on how the testing program has evolved over the course of the epidemic has been used to set change points for testing parameters, which are identified during the calibration process.

Washington is a home-rule state, meaning that laws can be set at the local level, and as such the contact tracing programs are led by each local health jurisdiction. Seattle-King County was one of the first jurisdictions in the state to pilot contact tracing using local health resources. Today, the county has sufficient human resources to trace approximately half of the cases, while the other half are handled by the Washington State contact tracing program. Data for this analysis were provided by PHSKC.

We used data on weekly foot traffic patterns obtained through SafeGraph to model the degree of mobility in the workplace and community layers of the model. This publicly available dataset is based on anonymized cell phone data, which connects foot traffic counts with points of interest visits for over 5 million unique locations in the United States. This rich dataset enables a detailed view of hourly visits to specific locations. Using an aggregation of visits in King County across all industries starting at the end of January, we classified visits with a dwell-time less than four hours as community-associated mobility,

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

and those visits with a dwell-time of more than four hours as workplace-associated mobility. Using the last week of January as a baseline value for "pre-COVID mobility", we assigned a weekly mobility level for community and work relative to the baseline value.

Population and network model

To model King County with detailed information on the demographics and network structures of the King County population, we used SynthPops, an open-source data-driven model for generating realistic synthetic contact networks for populations. Further details are provided below. For the population of King County, we used a combination of data sources at the county, state, and country resolution with the SynthPops model (version 0.7.2). Specifically, data from the 2018 American Community Survey (ACS) at the county resolution (*38*) was used to estimate age and household size distributions. The US Census Bureau (*39*) provided data for the age of reference individuals by household size at the country resolution, and age mixing contact matrices for the US are drawn from (*40*). For the network layer of schools, we used 2018 ACS 1-year estimates for county enrollment rates by age (*38*), municipal records on school enrollment numbers (*41*), student-teacher ratios, and the average class size for schools (*42*). For the network layer of workplaces, we used 2018 ACS 1-year estimates for county employment rates by age (*43*) and 2015 county estimates for workplace sizes (*44*).

In King County, a significant percent of COVID-19 cases and deaths have occurred to date within the LTCF population. However, most COVID models to date have not explicitly included LTCFs, which has been identified as a major limitation (*45*). To capture the dynamics of transmission in this setting, the SynthPops model was extended to include the contact layer of those living and working in long term care facilities to reflect the initial outbreaks that occurred in these facilities (*46*). Data on the demographics of LTCF residents for Washington state (*47*) was used to estimate the number of residents and their ages for Seattle area facilities. From these data, we estimate approximately 15,000 individuals aged 60 and older are residents of long term care facilities within the Seattle area. Additional King County data on the number of residents per facility and resident-to-staff ratios were used to sample facility sizes and populate the facilities with both residents and staff members. There were an average of 123 residents per facility and 132 staff per facility, but there is wide variation in both total numbers of residents and staff and in the resident-to-staff ratios during any shift. For the purposes of this model, residents considered to be living in these facilities were not assigned any additional outside contacts (household, school, workplace, or community). Staff members were drawn from the labor force of the population under 60 years of age. With large facility sizes, we modeled close contacts in facilities by sampling for each resident and staff member a subset of 20 contacts, ensuring that each resident is in contact with at least one staff member.

Five synthetic populations of 225,000 modeled agents (representing 10% of King County's population) were generated using SynthPops, with dynamic scaling used to rescale this population to represent the full 2.25 million population of King County (see Section 2.6.2 of (*35*)). These synthetic populations had roughly 500,000 household contacts, 1 million school contacts, 1.9 million workplace contacts, 4.5 million community contacts, and 31,000 LTCF contacts. To reflect the relative amount of time spent with each contact across different layers, averaged across a typical week, relative transmission weights per layer were set to be 100% for households (as a reference value), 50% for LTCFs, 20% for schools and workplaces, and 10% for community contacts. These values were chosen for consistency with both time-use surveys (*48*) and studies of infections with known contact types (*32*). A subset of simulations were also run with 2.25 million agents to verify that results were consistent with and without rescaling.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Calibration methodology

We calibrated model parameters using Optuna, a Python-based optimization library (*19*), using the tree-structured Parzen estimator (TPE) sampler (*49*). This sampler trains models of $p(\theta|y)$ and $p(y)$, where θ is a set of parameters and y is a (scalar) output of a loss (objective) function, to find the region of the parameter space that minimizes y. We defined the loss function to be the sum of the absolute differences between observed data and the corresponding model predictions for seven different target outputs, namely: (a) cumulative diagnoses per day, (b) cumulative deaths per day, (c) 7-day rolling average diagnoses, (d) 7-day rolling average deaths, (e) total diagnoses by age (using 10-year age bins), (f) average test positivity rate by age, and (g) total deaths by age. To equalize the weight given to each point in each of these five different data types, each data type was normalized by the maximum value in the data. In addition, cumulative vs. rolling average data were weighted in the ratio 4:1, which was found to most efficiently optimize the tradeoff between accurate fitting of long-term trends (driven by fits to cumulative data) and short-term trends (driven by fits to rolling average data). Deaths and diagnoses by age were given a relative weighting of 2, while test positivity by age was given a weight of 0.4 (since it was highly correlated with diagnoses by age, given that the total number of tests was fixed). These weights were chosen through an iterative process to determine algorithm convergence; final results are not sensitive to them, since the final simulations used for the analysis were good fits to all seven target outputs, and since they do not represent independent degrees of freedom (e.g. a good fit to rolling average diagnoses is necessarily at least a reasonable fit to cumulative diagnoses). Similarly, excluding a given target (e.g., diagnoses by age) did not always result in a significantly worse fit to that target, as long as at least one comparable target was included in the calibration (e.g., cumulative diagnoses).

We used 104,000 simulation runs during the calibration process to ensure broad exploration of parameter space. To determine parametric uncertainty, we used a cutoff value for the loss function of 30, which corresponded to no more than roughly 2% average relative error per point in the diagnoses and deaths time series, and roughly 10% average relative error for diagnoses, yield, and deaths age distributions; this cutoff was also roughly a factor of 2 larger than the single best-fitting simulations (which had total losses of 15.6 and 15.9 for calibrations with and without mobility data, respectively). Using this cutoff, the posterior distribution consisted of the 15,092 best-fitting parameter sets for the calibrations that used SafeGraph mobility data, and the 8,821 best-fitting parameter sets for calibrations that did not. Median values and 95% confidence intervals for epidemic projections and parameter distributions (Fig. 1) were produced using these parameter sets (based on a uniform sample of 200 simulations). Detailed transmission characteristics (Fig. 2) were based on the single best fit with mobility data. For scenario analyses (Figs. 4 and 5), the top 10 best-fitting parameter sets for the calibrations that included mobility data were used.

Calibrating using four parameters was found to be sufficient to allow sufficient flexibility to capture observed epidemic trends, both with and without using mobility data as input. These parameters are shown in Table 1; uniform priors were used. Simulations were initialized with 300 seed infections, distributed at random throughout the population, on January 27. This initialization was chosen by calibrating the number of seed infections and overall transmission rate (β) to the subset of data prior to major policy or mobility changes (i.e., February 27), and for consistency with other estimates of the initial reproduction number in King County. In the calibration, larger numbers of seed infections were compensated for by smaller transmission rates; 300 seed infections was the fewest that could be used (reflecting the highest baseline transmission rate) that provided a reasonable match to the data. We

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

used SafeGraph data to determine the proportion of network edges in workplace and community layers that should be removed or restored over time based on observed changes in the mobility. Other model parameters were set to use Covasim defaults, as described below. The model was calibrated to data from January 27 until June 9. Scenarios began on June 1; we used the 9-day period of overlap to ensure consistency between calibrated and projected estimates of new infections, tests, diagnoses, and deaths.

Table 1. Model parameters and calibrated values determined via fitting model outputs to King County data.

Parameter	Primarily constrained by	Calibrated value with SafeGraph data: median (95% CI)	Calibrated value without SafeGraph data: median (95% CI)	Search interval
Probability of transmission per contact per day (β, %)	Initial rate of epidemic growth in observed diagnoses and deaths	4.4 (4.2, 4.6)	4.3 (4.2, 4.5)	[3.3, 4.8]
Relative reduction in transmission rate in work and community layers from March 23 onwards (%)	Numbers of diagnoses and deaths	12.3 (2.9, 23.5)	71.7 (65.0, 79.9)	[0, 90]
Relative reduction in transmission rate in LTCFs from March 23 onwards (%)	Age distribution of deaths and diagnoses; ratio of deaths to diagnoses; time trend of deaths	86.1 (70.6, 94.2)	87.1 (71.0, 94.3)	[60, 95]
Odds ratio of people with symptoms testing	Test positivity rate; number of diagnoses	20.9 (15.2, 27.7)	20.2 (14.8, 27.7)	[10, 60]

Idealized test-trace-quarantine scenarios

For the illustrative transmission trees shown in Fig. 3A–C, we used a hypothetical population of 100 people with a single seed infection simulated for 100 days. Population demographics were based on Seattle, Washington, USA, but contact networks were generated using the "hybrid" algorithm rather than SynthPops; this algorithm is described in Section 2.4.3 of (*35*). Testing and tracing interventions began on day 20 of the simulation. The testing intervention used 15% daily probabilities of testing for people with symptoms; people without symptoms were not tested, and all people were tested upon entering quarantine. After consulting with Public Health Seattle & King County on estimated behavioral norms, contact tracing probabilities for the household, school, work, and community layers were 70%, 10%, 10%, and 0%, respectively (note that long-term care facilities are not included in these scenarios). People who were diagnosed and isolated were assumed to reduce their transmission rates by 70% for household contacts, and 90% for school, workplace, and community contacts. People who were contact traced and quarantined were assumed to reduce their transmission rates by 40% for household contacts, and 80% for school, workplace, and community contacts.

To explore the theoretical properties of test-trace-quarantine (Fig. 3D–F), we used a hypothetical population of 30,000 people with 100 seed infections simulated for 150 days. As above, a hybrid network was used. Simulations were run with 10 different random seeds, for three different transmission levels: medium transmission (β = 4.2% per household contact per day, consistent with estimated β for Seattle), low transmission (β = 3.3%), and high transmission (β = 5.1%).

Parameters for each of the three intervention scenarios (physical distancing, testing, and testing plus tracing) were chosen to bring $R_e \approx 1$ for the medium transmission scenario. These intervention

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

parameters were held constant for the low and high transmission scenarios. The interventions that began on day 15 of the simulation for each of the three scenarios were:

- Physical distancing scenario: 60% reduction in β, no testing or contact tracing;
- Testing scenario: no reduction in β; daily probability of testing of 75% and 7.5% for people with and without symptoms, respectively, with no testing delay (test results returned same day); no contact tracing;
- Test-and-trace scenario: no reduction in β; daily probability of testing of 8%, 0.8%, and 75% for people with symptoms, without symptoms, and in quarantine, respectively, with no testing delay; tracing probability of 90% across all layers with no tracing delay.

While zero delays were used here, we also ran a sensitivity analysis with nonzero delays (1 day for testing and 1-2 days for contact tracing). Note that even with zero delays, there is a minimum one-day delay per step in the contact tracing process (since people who are placed into quarantine cannot test until the next timestep, i.e. the following day). In the model we assume that the duration of infectiousness is equivalent to the period during which a person would test positive. For idealized TTQ to succeed for high rates of transmission, the average delay for a single step of contact tracing must be less than the average serial interval, i.e. the average delay between a primary infection and a secondary infection. However, it is not necessary for the delays to be less than the shortest serial interval; if a secondary infection occurs prior to contact tracing, the cluster can still be contained as long as the average tracing delay is less than the average serial interval. An example of this is shown in Fig. 3C: person 71 is not diagnosed until after they have infected person 74 (day 37), and person 74 is not traced until after they have infected person 75 (day 40). However, person 75 is quarantined before they transmit further (day 42), and the cluster is contained.

Realistic test-trace-quarantine scenarios

We used the 10 parameter configurations with the best fit to the data over the period February 1 to June 10 as the basis for the test-trace-quarantine (TTQ) scenarios shown in Fig. 4A. Scenarios were also run with other sets of calibrations (including the top 100, and using the same goodness-of-fit threshold used for the distributions shown in Fig. 1E). These results did not differ qualitatively and only modest quantitative differences were observed; the top 10 calibrations were chosen to ensure the best fit to data while still capturing both parametric and stochastic uncertainty. To explore the relative importance of different intervention parameters, we ran a sweep of 50 points for each of the six parameters (described below), for each of the 10 parameter configurations, for a total of 3,000 simulations. Each scenario began on June 10 and ended on August 30, which was chosen as the period prior to the potential reopening of schools. Scenarios began with an immediate return to 100% mobility in work and community (from a baseline value of 43% from the last reported SafeGraph data on June 1) as well as immediate implementation of the testing and contact tracing interventions with the parameter values described below. Relative β (compared to baseline), reflecting mask use and other non-pharmaceutical interventions, remained constant throughout the scenarios at its last calibrated value, which varied from 73% to 82%.

For each scenario, only one parameter at a time was varied. Note that the parameters interact nonlinearly; for example, the impact of the contact tracing delay depends on the amount of contact tracing. Additionally, the impact of interventions is dependent on the epidemic dynamics: with very low mobility and hence baseline transmission, the impact of interventions will be reduced. Thus, the baseline scenario was chosen to reflect (a) a situation where $R_e \approx 1$, which is the point most sensitive to small

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

differences in intervention effectiveness; and (b) a balance between testing and contact tracing that is intended to reflect a realistic scale-up of both current programs. While other baseline points would be possible, this scenario aims to reflect a potentially achievable point by which Seattle-King County could maintain $R_e \approx 1$ with full reopening.

The six intervention parameters are defined as follows:

1. **Isolation/quarantine effectiveness:** The relative change in transmission following either diagnosis (isolation) or after being notified as a known contact (quarantine). While in practice (and for the assumptions used during the calibration period) the effectiveness would differ between isolation and quarantine (with isolation expected to have higher effectiveness), as well as between contact layers (with a greater reduction in workplace and community transmission compared to household), here we used a single weighted average value to ensure that the slope has meaningful units (i.e., infections averted per person fully isolated or quarantined). The default value chosen was 80% efficacy, which is a weighted average between workplace and community contacts (where isolation efficacy is likely to be higher, e.g. 90–95% effectiveness) and household contacts (where isolation efficacy is likely to be lower, e.g. 40–70% effectiveness). This parameter was varied from 0% (no impact of isolation/quarantine) to 100% (zero transmission during isolation/quarantine).
2. **Contact tracing probability:** The proportion of household, workplace, and LTCF contacts of a person who has been diagnosed who are reached by contact tracers. (The proportion of community contacts reached is assumed to be zero for this analysis; schools are closed for the scenario period so there are no school contacts to trace.) The default value chosen was 50%, which again reflects a weighted average between household and LTCF contact tracing (where probabilities well above 80% are achievable) and workplace contact tracing (where probabilities in Seattle-King County are currently low). This parameter was varied from 0% (no contacts traced) to 100% (all household and workplace contacts traced).
3. **Quarantine testing probability:** The probability that a known contact, once traced, will be tested for COVID-19. The default value used was 90%, regardless of symptoms. This parameter was varied from 0% (no testing of people in quarantine) to 100% (including uninfected, asymptomatic, and presymptomatic contacts). Upon testing negative, contacts were *not* released from quarantine, due to the possibility that they would become infected due to continued contact with the index case (as is often the case for household contacts), or in case they were exposed but had not yet started shedding at detectable levels.
4. **Routine testing probability:** The probability per person per day of a person receiving a test for COVID-19. The default values chosen were 16% per day for a person with active symptoms, and 0.16% for people who are uninfected or who do not have symptoms. These values correspond to an approximate doubling of the number of daily tests relative to June 10; the ratio of probabilities for people with and without symptoms was set to be 100, which was chosen to be consistent with the observed testing yield in the data (approximately 1.5–2.5%). This parameter was varied from 0% (no routine testing) to 50% daily symptomatic testing and 0.5% daily non-symptomatic testing, corresponding to a roughly 4-fold increase in testing rates compared to June 10.
5. **Swab-to-result delay:** The average number of days between when a person receives a COVID-19 swab to when they are notified of their result. The default value chosen was 1 day, reflecting a slight improvement on practice in Seattle-King County as of June 10 (approximately 1.5 days). This parameter was varied from 0 days (immediate return of test results) to 7 days.
6. **Contact tracing delay:** The average number of days between when a person receives a positive result from a COVID-19 test and when their contacts are traced and notified. The default value

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

chosen was 2 days, which is somewhat shorter than estimates as of June 10 (3-5 days). This parameter was varied from 0 days (immediate notification of all contacts, although the swab-to-result delay is still present) to 7 days.

Because epidemic growth is an exponential process, the attack rate varied widely between scenarios, from less than 0.1% to nearly 50%. The attack rate had nonlinear dependence on all intervention parameters. Thus, the attack rate was log-transformed prior to fitting. The ordinary least squares method from the Python package *statsmodels* was used for the fit. The uncertainty interval shown is the 95% confidence interval from the ordinary least squares fit. Because of the log transform, the slope of the line depends on the point of evaluation; in all cases, it was evaluated at the default value for each parameter. Since the dependent variable in the regression is attack rate, the slope is also dependent on the period of integration (here, 91 days); a longer integration period, for example, would lead to a higher cumulative attack rate and thus larger slopes.

For the reopening sweeps (Fig. 4B), all parameters except for routine testing probability and contact tracing probability were fixed at the default values described above. We simulated eight different reopening levels (60% to 100%, in 5% increments), and show 60%, 80%, and 100% to represent low, medium, and high transmission scenarios. Each sweep consists of 12,000 simulations, with each simulation drawn from a uniform random distribution for (a) routine testing probability and (b) contact tracing probability, with each simulation also drawing from one of the 10 best calibrations as described above.

For reopening scenarios (Fig. 5), the status quo model was implemented using the baseline calibration to data until June 1, using input data on observed numbers of tests performed and contacts traced. In addition, mobility rates were increased to 80% on June 1, representing the lifting of the "Stay Home, Stay Healthy" measure; a subsequent β reduction of 25% was applied on July 1, reflecting the statewide mask requirements that were mandated on June 23 and July 7, as described in the August 13 Situation Report.

Acknowledgements and contributions

Acknowledgements

Review of data and results was provided by Matthew Golden, Cathy Wasserman, and Ian Painter. Literature and code reviews were performed by Anna Palmer, Dominic Delport, Caroline S. Bennette, Bradley Wagner, Stewart Chang, Jasmina Panovska-Griffiths, and Edward Wenger. Schematics were produced by Amanda Izzo. Additional contributors to the Covasim model and this study include: from GitHub, William Fitzgerald, Hamel Husain, Cory Gwin, Julian Nadeau, Rasmus Wriedt Larsen, Aditya Sharad, and Oege de Moor; from Microsoft, William Chen, Scott Ayers, and Rolf Harms; from the Institute for Disease Modeling, Mary Fisher, Jennifer Schripsema, Dennis Chao, Christian Wiswell, Samuel Buxton, Christopher Lorton, Clinton Collins, Christopher Jones, Charles Eliot, Svetlana Titova, Dejan Lukacevic, Jeffrey Steinkraus, John Sheppard, Niket Thakkar, Roy Burstein, Robert Hart, Guillaume Chabot-Couture, Caitlin Bever, Helen Olsen, Greer Fowler, and Natalia Corona; from the Allen Institute, Natalia Orlova; from the Jet Propulsion Laboratory, Casey Handmer; from the QIMR Berghofer Medical Research Institute, Paula Sanz-Leon and James Roberts; from the Kirby Institute, Richard Gray; and from the Burnet Institute, Nick Scott and Sherrie Kelly. We also wish to thank the participants of the Covasim Users Group, including Julie Maher, Dean Sidelinger, and Erik Everson from the Oregon Health Authority; André Lin Ouédraogo from the Institute for Disease Modeling; and David P. Wilson from the Bill and Melinda Gates Foundation.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Funding

Bill and Melinda Gates Foundation. Institutional support, including high-performance computing resources and library access, was provided by the Burnet Institute and the University of Sydney School of Physics

Author contributions

Covasim model development was led by Cliff Kerr, Robyn Stuart, Romesh Abeysuriya, and Daniel Klein, with additional support by Gregory Hart, Katherine Rosenfeld, Prashanth Selvaraj, Rafael Núñez, Jamie Cohen, Lauren George, and Michał Jastrzębski. The SynthPops model was developed by Dina Mistry, with additional support by Cliff Kerr, Romesh Abeysuriya, Daniel Klein, and Lauren George. The health systems component was based on a model developed by Brittany Hagedorn. Data were provided and curated by Meaghan Fagalde and Jeffrey Duchin. Analyses were performed by Cliff Kerr, Dina Mistry, Robyn Stuart, Katherine Rosenfeld, Gregory Hart, Rafael Núñez, Prashanth Selvaraj, and Jamie Cohen. Supervision was provided by Jeffrey Duchin, Jasmina Panovska-Griffiths, Michael Famulare, and Daniel Klein. The manuscript was written by Cliff Kerr, Dina Mistry, Robyn Stuart, Rafael Núñez, and Daniel Klein. Manuscript review and editing was performed by Katherine Rosenfeld, Jamie Cohen, Jasmina Panovska-Griffiths, and Michael Famulare.

Competing interests

Authors declare no competing interests.

Data availability

All data used in this study are available via GitHub. Most data used in this study are also available from the King County Data Dashboard.

Code availability

The Covasim model code is fully open source and available via GitHub. Analysis and plotting scripts to reproduce the results of this study are also available via GitHub. A webapp that renders interactive versions of the figures is available at https://ttq-app.covasim.org.

References

1. WHO Coronavirus Disease (COVID-19) Dashboard, (available at https://covid19.who.int).
2. B. F. Maier, D. Brockmann, Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*. **368**, 742–746 (2020).
3. O. Coibion, Y. Gorodnichenko, M. Weber, "The Cost of the Covid-19 Crisis: Lockdowns, Macroeconomic Expectations, and Consumer Spending," *Working Paper Series* (Working Paper 27141, National Bureau of Economic Research, 2020), , doi:10.3386/w27141.
4. M. E. Kretzschmar, G. Rozhnova, M. C. J. Bootsma, M. van Boven, J. H. H. M. van de Wijgert, M. J. M. Bonten, Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *The Lancet Public Health*. **5**, e452–e459 (2020).
5. A. J. Kucharski, P. Klepac, A. J. K. Conlan, S. M. Kissler, M. L. Tang, H. Fry, J. R. Gog, W. J. Edmunds, J. C. Emery, G. Medley, J. D. Munday, T. W. Russell, Q. J. Leclerc, C. Diamond, S. R. Procter, A. Gimma, F. Y. Sun, H. P. Gibbs, A. Rosello, K. van Zandvoort, S. Hué, S. R. Meakin, A. K. Deol, G. Knight, T. Jombart, A. M. Foss, N. I. Bosse, K. E. Atkins, B. J. Quilty, R. Lowe, K. Prem, S. Flasche, C. A. B. Pearson, R. M. G. J. Houben, E. S. Nightingale, A. Endo, D. C. Tully, Y. Liu, J. Villabona-Arenas, K. O'Reilly, S. Funk, R. M. Eggo, M. Jit, E. M. Rees, J. Hellewell, S. Clifford, C. I. Jarvis, S. Abbott, M.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Auzenbergs, N. G. Davies, D. Simons, Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *The Lancet Infectious Diseases*. **20**, 1151–1160 (2020).

6. J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski, W. J. Edmunds, F. Sun, S. Flasche, B. J. Quilty, N. Davies, Y. Liu, S. Clifford, P. Klepac, M. Jit, C. Diamond, H. Gibbs, K. van Zandvoort, S. Funk, R. M. Eggo, Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health*. **8** (2020), pp. e488–e496.
7. A. Bilinski, F. Mostashari, J. A. Salomon, Modeling Contact Tracing Strategies for COVID-19 in the Context of Relaxed Physical Distancing Measures. *JAMA Netw Open*. **3**, e2019217 (2020).
8. A. Aleta, D. Martin-Corral, A. P. y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. Elizabeth Halloran, I. M. Longini, S. Merler, A. Pentland, A. Vespignani, E. Moro, Y. Moreno, Modeling the impact of social distancing, testing, contact tracing and household quarantine on second-wave scen- arios of the COVID-19 epidemic (2020).
9. S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, M. Monod, A. C. Ghani, C. A. Donnelly, S. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, S. Bhatt, Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*. **584**, 257–261 (2020).
10. E. Gibney, Whose coronavirus strategy worked best? Scientists hunt most effective policies. *Nature*. **581**, 15–16 (2020).
11. T. Blakely, J. Thompson, N. Carvalho, L. Bablani, N. Wilson, M. Stevenson, The probability of the 6-week lockdown in Victoria (commencing 9 July 2020) achieving elimination of community transmission of SARS-CoV-2. *The Medical Journal of Australia*. **213**, 349-351.e1 (2020).
12. S. Knapton, At least 11 countries have reimposed restrictions amid fears of coronavirus second wave. *The Telegraph* (2020), (available at https://www.telegraph.co.uk/news/2020/06/04/least-11-countries-have-reimposed-restrictions-amid-fears-coronavirus/).
13. S. Griffin, Covid-19: Lack of test and trace data is frustrating government scrutiny. *BMJ*. **369** (2020), doi:10.1136/bmj.m2239.
14. E. Mahase, Covid-19: Oxford vaccine could be 59% effective against asymptomatic infections, analysis shows. *BMJ*, m4777 (2020).
15. N. G. Davies, S. Abbott, R. C. Barnard, C. I. Jarvis, A. J. Kucharski, J. Munday, C. A. B. Pearson, T. W. Russell, D. C. Tully, A. D. Washburne, T. Wenseleers, A. Gimma, W. Waites, K. L. Wong, K. van Zandvoort, J. D. Silverman, C. C.-19 W. Group, K. Diaz-Ordaz, R. Keogh, R. M. Eggo, S. Funk, M. Jit, K. E. Atkins, W. J. Edmunds, *medRxiv*, in press, doi:10.1101/2020.12.24.20248822.
16. M. L. Holshue, C. DeBolt, S. Lindquist, K. H. Lofy, J. Wiesman, H. Bruce, C. Spitters, K. Ericson, S. Wilkerson, A. Tural, G. Diaz, A. Cohn, L. Fox, A. Patel, S. I. Gerber, L. Kim, S. Tong, X. Lu, S. Lindstrom, M. A. Pallansch, W. C. Weldon, H. M. Biggs, T. M. Uyeki, S. K. Pillai, First Case of 2019 Novel Coronavirus in the United States. *New England Journal of Medicine*. **382**, 929–936 (2020).
17. CDC, Washington State Report First COVID-19 Death | CDC Online Newsroom | CDC (2020), (available at https://www.cdc.gov/media/releases/2020/s0229-COVID-19-first-death.html).
18. A. Lasry, Timing of Community Mitigation and Changes in Reported COVID-19 and Community Mobility — Four U.S. Metropolitan Areas, February 26–April 1, 2020. *MMWR Morb Mortal Wkly Rep*. **69** (2020), doi:10.15585/mmwr.mm6915e2.
19. T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, in *"Optuna: A Next-generation Hyperparameter Optimization Framework", Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Association for Computing Machinery, Anchorage, AK, USA, 2019; https://doi.org/10.1145/3292500.3330701), *KDD '19*, pp. 2623–2631.
20. N. Thakkar, G. Huynh, R. Etzioni, I. Painter, J. Lavista Ferres, M. Famulare, Situation Report 7: COVID-19 transmission across Washington State. *Institute for Disease Modeling. https://covid.idmod.org/data/WA_Situation_Report_7_COVID-19_transmission_across_Washington_State.pdf* (2020).
21. A. Kapteyn, M. Angrisani, D. Bennett, W. B. de Bruin, J. Darling, T. Gutsche, Y. Liu, E. Meijer, F.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Perez-Arce, S. Schaner, K. Thomas, B. Weerman, Tracking the Effect of the COVID-19 Pandemic on the Lives of American Households. *Survey Research Methods*. **14**, 179–186 (2020).

22. H.-Y. Cheng, S.-W. Jian, D.-P. Liu, T.-C. Ng, W.-T. Huang, H.-H. Lin, Contact Tracing Assessment of COVID-19 Transmission Dynamics in Taiwan and Risk at Different Exposure Periods Before and After Symptom Onset. *JAMA Intern Med* (2020), doi:10.1001/jamainternmed.2020.2020.
23. Q.-L. Jing, M.-J. Liu, Z.-B. Zhang, L.-Q. Fang, J. Yuan, A.-R. Zhang, N. E. Dean, L. Luo, M.-M. Ma, I. Longini, E. Kenah, Y. Lu, Y. Ma, N. Jalali, Z.-C. Yang, Y. Yang, Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study. *The Lancet Infectious Diseases*. **0** (2020), doi:10.1016/S1473-3099(20)30471-0.
24. 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, B. J. Cowling, F. Li, G. M. Leung, Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*, 1–4 (2020).
25. L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, C. Fraser, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*. **368**, eabb6936 (2020).
26. W. J. Bradshaw, E. C. Alley, J. H. Huggins, A. L. Lloyd, K. M. Esvelt, Bidirectional contact tracing could dramatically improve COVID-19 control. *Nature Communications*. **12**, 232 (2021).
27. S. Kojaku, L. Hébert-Dufresne, E. Mones, S. Lehmann, Y.-Y. Ahn, The effectiveness of backward contact tracing in networks. *Nature Physics*, 1–7 (2021).
28. 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, A. J. Kucharski, Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks. *Wellcome Open Res*. **5**, 239 (2021).
29. R. C. Reiner, R. M. Barber, J. K. Collins, P. Zheng, C. Adolph, J. Albright, C. M. Antony, A. Y. Aravkin, S. D. Bachmeier, B. Bang-Jensen, M. S. Bannick, S. Bloom, A. Carter, E. Castro, K. Causey, S. Chakrabarti, F. J. Charlson, R. M. Cogen, E. Combs, X. Dai, W. J. Dangel, L. Earl, S. B. Ewald, M. Ezalarab, A. J. Ferrari, A. Flaxman, J. J. Frostad, N. Fullman, E. Gakidou, J. Gallagher, S. D. Glenn, E. A. Goosmann, J. He, N. J. Henry, E. N. Hulland, B. Hurst, C. Johanns, P. J. Kendrick, A. Khemani, S. L. Larson, A. Lazzar-Atwood, K. E. LeGrand, H. Lescinsky, A. Lindstrom, E. Linebarger, R. Lozano, R. Ma, J. Månsson, B. Magistro, A. M. M. Herrera, L. B. Marczak, M. K. Miller-Petrie, A. H. Mokdad, J. D. Morgan, P. Naik, C. M. Odell, J. K. O'Halloran, A. E. Osgood-Zimmerman, S. M. Ostroff, M. Pasovic, L. Penberthy, G. Phipps, D. M. Pigott, I. Pollock, R. E. Ramshaw, S. B. Redford, G. Reinke, S. Rolfe, D. F. Santomauro, J. R. Shackleton, D. H. Shaw, B. S. Sheena, A. Sholokhov, R. J. D. Sorensen, G. Sparks, E. E. Spurlock, M. L. Subart, R. Syailendrawati, A. E. Torre, C. E. Troeger, T. Vos, A. Watson, S. Watson, K. E. Wiens, L. Woyczynski, L. Xu, J. Zhang, S. I. Hay, S. S. Lim, C. J. L. Murray, IHME COVID-19 Forecasting Team, Modeling COVID-19 scenarios for the United States. *Nature Medicine*. **27**, 94–105 (2021).
30. C. J. Wang, C. Y. Ng, R. H. Brook, Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing. *JAMA*. **323**, 1341–1342 (2020).
31. T. Hale, A. Petherick, T. Phillips, S. Webster, Variation in government responses to COVID-19. *Blavatnik school of government working paper*. **31** (2020).
32. New South Wales Government Department of Health, "COVID-19 Weekly Surveillance in NSW: Epidemiological Week 36" (2020), (available at https://www.health.nsw.gov.au/Infectious/covid-19/Documents/covid-19-surveillance-report-20200905.pdf).
33. S. Contreras, J. Dehning, M. Loidolt, J. Zierenberg, F. P. Spitzner, J. H. Urrea-Quintero, S. B. Mohr, M. Wilczek, M. Wibral, V. Priesemann, The challenges of containing SARS-CoV-2 via test-trace-and-isolate. *Nature Communications*. **12**, 378 (2021).
34. J. Zhang, M. Litvinova, Y. Liang, Y. Wang, W. Wang, S. Zhao, Q. Wu, S. Merler, C. Viboud, A. Vespignani, M. Ajelli, H. Yu, Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*. **368**, 1481–1486 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.15.20154765; this version posted March 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

35. C. C. Kerr, R. M. Stuart, D. Mistry, R. G. Abeysuriya, K. Rosenfeld, G. R. Hart, R. C. Nunez, J. A. Cohen, P. Selvaraj, B. Hagedorn, L. George, M. Jastrzebski, A. Izzo, G. Fowler, A. Palmer, D. Delport, N. Scott, S. Kelly, C. Bennette, B. Wagner, S. Chang, A. P. Oron, E. Wenger, J. Panovska-Griffiths, M. Famulare, D. J. Klein, *medRxiv*, in press, doi:10.1101/2020.05.10.20097469.
36. N. Scott, A. Palmer, D. Delport, R. G. Abeysuriya, R. M. Stuart, C. C. Kerr, D. Mistry, D. J. Klein, R. Sacks-Davis, K. Heath, S. Hainsworth, A. Pedrana, M. Stoove, D. P. Wilson, M. Hellard, Modelling the impact of reducing control measures on the COVID-19 pandemic in a low transmission setting. *Med J Aust* (2020), doi:10.1101/2020.06.11.20127027v1.
37. J. Panovska-Griffiths, C. C. Kerr, R. M. Stuart, D. Mistry, D. J. Klein, R. M. Viner, C. Bonell, Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study. *The Lancet Child & Adolescent Health*, S2352464220302509 (2020).
38. Census profile: Seattle-Tacoma-Bellevue, WA Metro Area. *Census Reporter*, (available at http://censusreporter.org/profiles/31000US42660-seattle-tacoma-bellevue-wa-metro-area/).
39. U. C. Bureau, America's Families and Living Arrangements: 2019. *The United States Census Bureau*, (available at https://www.census.gov/data/tables/2019/demo/families/cps-2019.html).
40. K. Prem, A. R. Cook, M. Jit, Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Comput. Biol.* **13**, e1005697 (2017).
41. Data Portal | OSPI, (available at https://www.k12.wa.us/data-reporting/data-portal).
42. Home - Washington State Report Card, (available at https://washingtonstatereportcard.ospi.k12.wa.us/).
43. Census - Table Results, (available at https://data.census.gov/cedsci/table?q=Employment%20and%20Labor%20Force%20Status&tid=ACSST1Y2018.S2301&g=0500000US53033&vintage=2018&hidePreview=true&t=Employment%20and%20Labor%20Force%20Status).
44. U. C. Bureau, Custom Tabulations. *The United States Census Bureau*, (available at https://www.census.gov/programs-surveys/susb/data/custom-tabulations.html).
45. K. Pillemer, L. Subramanian, N. Hupert, The Importance of Long-term Care Populations in Models of COVID-19. *JAMA*. **324**, 25 (2020).
46. T. M. McMichael, D. W. Currie, S. Clark, S. Pogosjans, M. Kay, N. G. Schwartz, J. Lewis, A. Baer, V. Kawakami, M. D. Lukoff, J. Ferro, C. Brostrom-Smith, T. D. Rea, M. R. Sayre, F. X. Riedo, D. Russell, B. Hiatt, P. Montgomery, A. K. Rao, E. J. Chow, F. Tobolowsky, M. J. Hughes, A. C. Bardossy, L. P. Oakley, J. R. Jacobs, N. D. Stone, S. C. Reddy, J. A. Jernigan, M. A. Honein, T. A. Clark, J. S. Duchin, Epidemiology of Covid-19 in a Long-Term Care Facility in King County, Washington. *New England Journal of Medicine* (2020), doi:10.1056/NEJMoa2005412.
47. Long-Term Care Providers and Services Users in the United States--State Estimates Supplement: National Study of Long-Term Care Providers, 2015-2016, 155.
48. D. Lader, S. Short, J. Gershuny, "The time use survey, 2005" (Office for National Statistics, London, 2006).
49. J. S. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, in *"Algorithms for Hyper-Parameter Optimization", Advances in neural information processing systems* (2011), pp. 2546–2554.