Supplementary Material

Supplementary methods

Model description

Supplementary Figure 1 shows the structure of the discrete-time stochastic individual-level susceptible-exposed-infectious-recovered (SEIR) model used to simulate transmission in the shelter. On any given day t each individual is in one of the seven states shown in the flow diagram and defined in Supplementary Table 1.

The probabilities that individual i is infected on day t or avoids infection on day t given that they are susceptible to infection are:

$$p_i(t) = 1 - e^{-\lambda_i(t)}, \quad i \in \mathcal{S}(t)$$

$$1 - p_i(t) = e^{-\lambda_i(t)}, \quad i \in \mathcal{S}(t)$$

where $\lambda_i(t)$ is the force of infection on each individual i on day t. The force of infection on each individual on each day is equal to the force of infection they are exposed to when inside the shelter, which is proportional to the prevalence of infectious individuals inside the shelter on that day and the infectiousness of these individuals, plus a background force of infection they are exposed to when outside the shelter in the community:

$$\lambda_i(t) = \frac{\beta \mathbb{I}_i(t) \left(h \left(\alpha I_{s,1}(t) + I_{s,2}(t) \right) + \alpha I_{c,1}(t) + I_{c,2}(t) \right)}{\sum_j \mathbb{I}_j(t)} + \epsilon. \quad (1)$$

Here β is the transmission rate coefficient within the shelter (assumed constant); $\mathbb{I}_i(t)$ is an indicator function for whether the individual is present in the shelter on day t; the $I(t)$'s represent the number of infectious individuals inside the shelter on day t in different states of infection – subclinical and clinical, denoted by subscripts s and c respectively, and early and late stage, denoted by subscripts 1 and 2 respectively; h is the infectiousness of subclinically infected individuals relative to those with clinical symptoms; α is the infectiousness of the early infectious stage relative to the late stage; and ϵ is the transmission rate outside of the shelter. Mixing of infectious and susceptible individuals in the shelter is assumed to be homogeneous due to a lack of contact data from the shelter outbreaks with which to parameterize inhomogeneous mixing. Transmission from the external community is modeled assuming homogeneous mixing of the shelter residents and staff with the community outside of the shelter during each day and negligible impact of infected individuals entering the community from the shelter on the overall background transmission rate. The background transmission rate is treated as constant given the relatively short durations of the outbreaks, and is estimated from the incidence of confirmed COVID-19 cases for the city of each shelter with adjustments for reporting delay, infection-to-onset time and relative risk of infection for homeless individuals as described below (see *Estimation of background infection rate* below).

We note that the formulation of the force of infection in equation (1) corresponds to frequency-dependent transmission, i.e. assumes that the number of "contacts" per infectious individual in the shelter per day is approximately constant regardless of the number of individuals present in the shelter. We make this assumption in common with other authors [1] because it is believed that the main mode of SARS-CoV-2 transmission is from person to person via respiratory droplets containing virus particles [2], i.e. occurs over short distances predominantly among close contacts of infectious individuals.

The duration D_E of the latent infection stage $\mathcal{E}$, is assumed to be negative-binomial, with mean $\mu_E = 3$ days and shape parameter $r_E = 4$, i.e. probability mass function:

$$\mathbb{P}(D_E = d) = \frac{\Gamma(d + r_E)}{\Gamma(r_E)d!}\left(\frac{r_E}{r_E + \mu_E - 1}\right)^{r_E}\left(\frac{\mu_E - 1}{r_E + \mu_E - 1}\right)^{d-1}, \qquad d = 1,2, \ldots,$$

where $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$ is the gamma function. After passing throught the latent stage, individuals enter an early (presymptomatic) infectious stage ($\mathcal{I}_{c,1}$) leading to clinical symptoms with age-dependent probability $\rho(a_i)$, where a_i is the age group (<60 years/≥60 years) of individual i, or an early infectious stage ($\mathcal{I}_{s,1}$) leading to subclinical infection (no symptoms/very mild symptoms) with probability $1 - \rho(a_i)$. After a negative-binomial number of days ($NB(r_1 = 4, \mu_1 = 2.3)$) of early stage infectiousness individuals progress to late stage subclinical ($\mathcal{I}_{s,2}$) or clinical ($\mathcal{I}_{c,2}$) infectiousness.

Following a further negative-binomially-distributed duration ($NB(r_2 = 4, \mu_2 = 8)$ with mean 8 days) subclinical cases recover and are no longer infectious and clinical cases either recover or are hospitalized, and therefore no longer contribute to transmission in the shelter ($\mathcal{R}$). See Supplementary Table 1 for the key attributes of each infection state and Supplementary Table 5 for a full list of model parameters and their values. The probability of hospitalization for clinical cases is age- and co-morbidity dependent (Supplementary Table 2), and hospitalized cases are assumed to have a 26.1% risk of requiring intensive care based on data from Wuhan, China [3]. Cases admitted to the intensive care unit (ICU) have an age- and co-morbidity-dependent risk of death estimated from ICU data from Wuhan, China [4,5].

Basic reproduction number, R_0

The basic reproduction number for the model, defined as the average number of secondary infections caused by the average infectious individual in an entirely susceptible shelter population (in the absence of interventions) can be calculated from first principles as:

$$R_0 = \text{probability of infection given "contact"} \times \text{"contact" rate} \times \text{duration of infectiousness}$$
$$= \frac{\beta \sum_i \mathbb{I}_i(0)\left(\left(1 - \rho(a_i)\right)h + \rho(a_i)\right)(\alpha\mu_1 + \mu_2)}{\sum_i \mathbb{I}_i(0)}$$

where "contact" is defined as susceptible individuals coming into contact with infectious material from infected individuals, and the sums are over all individuals in the shelter (residents and staff).

Relative infectiousness of early infectious stage and subclinical infection

Evidence suggests that infectiousness of symptomatic COVID-19 cases is not constant over time, but peaks at or shortly before symptom onset, such that pre-symptomatic individuals are more infectious than symptomatic individuals [6,7]. We approximated this variation in infectiousness over time by treating individuals' infectiousness as constant during each of the early and late infectious stages, but higher during the early infectious stage. The relative infectiousness of the early infectious stage to the late infectious stage, $\alpha = 2$, was chosen to approximately match the estimates of the proportion of pre-symptomatic transmission of He et al [6] and the lower range of those of Casey et al [7] of 44% and 34% respectively, assuming a mean duration of the early infectious stage of $\mu_1 = 2.3$ days[6]:

$$\text{proportion of transmission from early infectious stage} = \frac{\alpha\mu_1}{\alpha\mu_1 + \mu_2} = \frac{2 \times 2.3}{2 \times 2.3 + 8} = 37\%,$$

We used the lower range of the estimates of Casey et al for the base case analysis, as we believe that estimates of pre-symptomatic infectiousness are likely to be biased upward due to behavior changes (e.g. self-isolation) upon symptom onset that reduce transmission from symptomatic individuals in general settings but are less likely to apply in congregate shelter settings. However, we also considered higher relative infectiousness (up to 3) in the sensitivity analysis. We assumed the same relative infectiousness of early stage infection for clinical infection and subclinical infection.

Subclinically infected individuals were assumed to be as infectious as clinical cases in the base case analysis, due to limited data on the relative infectiousness of subclinical infection and the detection of similar viral load in asymptomatic individuals as symptomatic individuals in several studies [6,8]. Lower relative infectiousness of subclinical infection (down to 50%) was considered in the sensitivity analysis.

Duration of detectable viral load

Studies that have measured the viral load of individuals infected with SARS-CoV-2 over time since symptom onset suggest that the virus remains detectable from throat and nasal swabs and sputum and stool samples for longer (~20 days after symptom onset) than individuals remain infectious (~7 or 8 days after symptom onset) [6,9]. We therefore modelled the duration of detectable viral load for each infected individual by assigning a random draw from a truncated discretized normal distribution that characterizes the variation in this duration (Supplementary Figure 3) to each individual when they enter the late infectious stages $\mathcal{I}_{s,2}$ and $\mathcal{I}_{c,2}$. We chose the parameters of the distribution based on data on the variation in times after symptom onset at which individuals' viral loads reach the PCR detection limit from several studies [6,9–13], with minimum and maximum durations of 5 days and 37 days.

We assumed that individuals in the early infectious stages $\mathcal{I}_{s,1}$ and $\mathcal{I}_{c,1}$ always have detectable viral loads and that in the latent infection stage the viral load is undetectable. We treated the duration of detectable viral load as being the same for subclinical and clinical infection and independent of the individual's duration of infectiousness. This means that individuals can still have a detectable viral load when they are no longer infectious and are in the recovered compartment ($\mathcal{R}$). It is likely that there is some correlation between viral load and symptom

severity and infectious duration (and other factors such as age), but as of yet there is insufficient data with which to parameterize these relationships and studies measuring viral loads over time by age and severity have shown mixed results [6,14–16].

Sensitivity and specificity of PCR tests

Although there is some evidence to suggest that the sensitivity of PCR tests varies with time since infection (i.e. that it is lower during early and late infection) [17], the data currently available to accurately characterize this variation is very limited. We therefore made the simplifying assumption that the sensitivity of PCR tests is constant with time since the start of infectiousness (time of entering the early infectious stages $\mathcal{I}_{s,1}$ and $\mathcal{I}_{c,1}$) and use a fixed sensitivity of 75% based on available data [17–20] for the base case analysis. We assumed perfect specificity of PCR tests for the base case analysis, i.e. no false positive test results, as current evidence suggests that they have high specificity (~99%) [17,20], but also considered lower specificity in the sensitivity analysis.

Estimation of background infection rate during shelter outbreaks

We used publicly available data on daily numbers of new confirmed COVID-19 cases in the city of each shelter from county public health departments to estimate the background infection rate [21–23]. We assumed a fixed delay from infection to reporting of 7 days, corresponding to 2 days of pre-symptomatic infection and 5 days of symptoms before reporting [24,25], and so used case counts 7 days ahead as an estimate of the number of new reported infections on each day. We estimated the infection incidence in the community outside the shelter during the period of each shelter outbreak, i_a, from the delay-adjusted case counts for the city of the shelter over the 3 weeks prior to the end date of data collection, T_{end}, as:

$$i_a = \frac{\gamma \sum_{t=T_{end}-14}^{T_{end}+7} n_t}{P}, \quad (2)$$

where n_t is the number of new confirmed cases reported on day t, P is the city population [26], and γ is an under-reporting factor to account for only a proportion of infections being reported. For the period of the shelter outbreaks we used $\gamma = 10$, based on estimated infection-to-reported-case ratios from seroprevalence data for the period of late March–early May from 10 different localities across the US [27]. For the intervention simulations, we used $\gamma = 4$, based on lower estimates from seroprevalence data for late May-June from a continuation of the same study [28]. We then estimated the background infection rate in homeless individuals outside of the shelter as:

$$\epsilon = r i_a,$$

where r is a relative risk of infection for homeless individuals. Based on data from Seattle & King County, WA, where 1.5% of the population is homeless (11199/753675)[26,29] and homeless individuals account for 2.5% (445/18130) of confirmed COVID-19 cases [21,30], we use an approximate relative risk for homeless individuals of $r = 2$.

Model calibration

We calibrated the model by fitting to data on numbers of PCR-positive and negative individuals in testing conducted in outbreaks in 5 homeless shelters in 3 different cities: San Francisco (n=1), Boston (n=1) and Seattle (n=3). Prevalence of PCR positivity among residents and staff during mass testing events in these shelters was markedly different, ranging from 2.6%–51.6%. This is likely partly due to testing being conducted at different times after the outbreaks started, but likely also reflects differences in transmissibility due to other factors, such as variation in infectiousness between individuals, shelter living density, bed spacing, air ventilation quality, and differences in shared washing and eating facilities. We therefore fitted the model separately to the data from each of these outbreaks to estimate the basic reproduction number, R_0, for each setting. Given uncertainty in when the first infected individuals entered each shelter and how many of them there were, due to asymptomatic infection and incomplete detection of early cases, we also estimated the time since introduction of infection into the shelter at the end of the data collection period, T, and the initial number of latently infected individuals who entered the shelter, E_0. We assumed that the individual(s) who initially introduced infection into the shelter were all latently infected when they entered, since a large range of scenarios in which they were in a later infection stage or mixture of infection stages are covered by the flexibility in the introduction time and the initial number infected. We also assumed that following the initial introduction into the shelter any further introductions were solely as a result of residents and staff being infected when mixing with the community outside the shelter and then returning to the shelter.

Since more detailed individual-level data on PCR test results and symptom onset times for clinical cases was available for the outbreak in the San Francisco shelter, we also used the numbers of early symptomatic cases who tested PCR-positive and daily numbers of new symptom onsets when calibrating the model to this outbreak (see Supplementary Figure 4 and S5A Fig). As exact dates of testing during the cross-sectional surveys in the Seattle and Boston shelters were not available, we assumed that all testing occurred on the last day of each survey. For all the outbreaks, we assumed that PCR tests took a day to be processed, such that results were returned and positive individuals removed the day following testing (except on April 10th at the San Francisco shelter, when negative individuals were removed instead).

Aggregate data (by age and co-morbidity risk group) on movement of individuals in and out of the shelter over time was only available for the San Francisco shelter, so we ensured that the population of this shelter matched that in the shelter registry (see *Demographics and movement of individuals in and out of the shelter* below) and assumed that the populations of the other shelters remained approximately constant over the period of data collection. Residents' age categories in the Seattle and Boston shelters were set according to data on the age distributions from [31] and [32]. For all shelters, staff were assumed to be all <60-years-old and at low risk of hospitalization and death.

Approximate Bayesian Computation algorithm

We fitted the model to the data using an approximate Bayesian computation sequential Monte Carlo (ABC-SMC) algorithm [33–35] to estimate R_0, E_0, and T. Since E_0 and T are discrete parameters, we adapted the model selection algorithm of Toni et al [34,35] by replacing the model index as the discrete parameter with the discrete parameter pair $m = (E_0, T)$ (i.e. the model indices by the different possible combinations of E_0 and T). The algorithm starts by

sampling pairs $m^* = (E_0^*, T^*)$ from the prior distribution $\pi(m)$, and corresponding R_0 values, R_0^{**}, from the prior distribution $\pi(R_0)$; simulating outbreaks with these parameter values (particles); and accepting those for which the simulated number of PCR-positives D^* falls within a certain pre-specified tolerance ϵ_1 of the observed number of PCR-positives D according to a distance measure $d(\cdot)$, i.e. $d(D, D^*) \leq \epsilon_1$. A sequence of distributions (generations) is then constructed by repeating this process with a set of decreasing tolerances $\boldsymbol{\epsilon} = \{\epsilon_g\}_{g=1,2,\ldots}$, proposing R_0 values for each value of m in each generation by perturbing the particles (R_0 values) specific to m from the previous generation using a perturbation kernel $K(R_0|R_0^*)$. In this way, the particles in successive generations converge towards the joint posterior distribution of the parameters given the data. Pseudocode for the algorithm is as follows:

1. Set the number of generations G and number of particles N.
2. Set the tolerance schedule $\epsilon_1 > \epsilon_2 > \cdots > \epsilon_G$. Set the generation index $g = 1$.
3. Set the particle index i to 1.
4. Sample m^* from the prior distribution $\pi(m)$. If $g = 1$, sample R_0^{**} from the prior distribution $\pi(R_0)$. If $g > 1$, sample R_0^* from the previous generation $\{R_0(m^*)_{g-1}\}$ with weights $w(m^*)_{g-1}$, and perturb the particle R_0^* to obtain $R_0^{**} \sim K(R_0|R_0^*)$.
5. If $\pi(R_0^{**}) = 0$, return to step 4.
6. Run a simulation of the outbreak and PCR testing for the sampled values (R_0^{**}, m^*) to generate a candidate dataset D^{**}.
7. If $\widehat{\mathbb{P}}(D|D^*) = \mathbb{I}\big(d(D, D^*) < \epsilon_g\big) = 0$, return to step 4.
8. Set $m_g^{(i)} = m^*$, and add R_0^{**} to the population of particles $\{R_0(m^*)_g\}$ and calculate its weight as:

$$w_g^{(i)} = \begin{cases} 1, & \text{if } g = 1 \\ \dfrac{\pi(R_0^{**})}{\sum_{i;m_{g-1}=m^*} \dfrac{w_{g-1}^{(i)} K(R_0^{**}|R_{0,g-1}^{(i)})}{\mathbb{P}(m_{g-1} = m^*)}}, & \text{if } g > 1 \end{cases}$$

9. If $i < N$, set $i = i + 1$ and go to step 4.
10. Normalize the weights w_g such that $\sum_{i=1}^{N} w_g^{(i)} = 1$.
11. Calculate the marginal probabilities for the combinations $m = (E_0, T)$, by summing the weights for each combination:

$$\mathbb{P}\big(m_g = m\big) = \sum_{i;m_g^{(i)}=m} w_g^{(i)}\Big(R_{0,g}^{(i)}, m_g^{(i)}\Big).$$

12. If $g < G$, set $g = g + 1$ and go to step 3.

We used $G = 10$ generations, with $N = 1000$ particles in each generation, and a normal perturbation kernel for R_0, $K(R_0|R_0^*) \sim N(R_0^*, \sigma^2)$, with standard deviation $\sigma = 1$. We used broad uniform prior distributions due to a lack of information to support more informative prior distributions: $R_0 \sim U(1,8)$, $m = (E_0, T) \sim U(1,5) \times U(14,30)$ (for all shelters except Seattle shelter B, for which $m \sim U(1,10) \times U(13,20)$), where the prior for m is discrete with integer support. The wide bounds for the prior for R_0 were chosen based on the large range of basic reproduction numbers reported in the literature[24] and demonstrated potential of COVID-19 for superspreading events [36–40]. The 14-day lower bound of the prior for T was chosen based on the symptom onsets of the first cases identified in each of the shelters being at least 9 days before

the end of data collection and the mean incubation period being approximately 5 days, such that the first cases were unlikely to have been infected later than 14 days before the end of data collection. Thirty days was taken as the upper bound for T based on it giving the earliest plausible time for introduction of infection into the shelters without earlier occurrence and detection of symptomatic cases. Lower bounds were used for the prior for T for Seattle shelter B to reflect the fact that it did not report any symptomatic cases before the first mass testing event on March 30–April 1 and had a very low prevalence of infection at that survey.

We used the sum of squared differences between the numbers of PCR-positives on the testing days, $\boldsymbol{T}_T$, in the simulations, D^*, and those in the observed data, D, for the distance metric $d(D, D^*)$:

$$d(D, D^*) = \sqrt{\sum_{t \in \boldsymbol{T}_T} (D_t - D_t^*)^2}$$

For the tolerance schedule $\boldsymbol{\epsilon}^T = (\epsilon_1, \ldots, \epsilon_G)$ we used regular steps decreasing from a discrepancy of twice the daily number of PCR-positives, $\epsilon_1 = \sqrt{\sum_t (2D_t)^2}$, to half the width of the exact binomial confidence interval on the number of PCR-positives on each day. For the San Francisco shelter, we used additional distance metrics, $d^S(D^S, D^{S*})$ and $d^C(D^C, D^{C*})$, and tolerance schedules, $\boldsymbol{\epsilon}^S$ and $\boldsymbol{\epsilon}^C$, for the differences in the simulated and observed numbers of PCR-positives among early symptomatic cases tested, D^{S*} and D^S, on days $\boldsymbol{T}_S$ (3/30/20–4/7/20) and the differences in the simulated and observed daily numbers of symptom onsets, D^{C*} and D^C:

$$d^S(D^S, D^{S*}) = \sqrt{\sum_{t \in \boldsymbol{T}_S} (D_t^S - D_t^{S*})^2}$$

$$d^C(D^C, D^{C*}) = \sqrt{\sum_{t=1}^{T} (D_t^C - D_t^{C*})^2}$$

Since the symptom onset data is less reliable than the PCR test data and potentially incomplete, we used less strict tolerances for $\boldsymbol{\epsilon}^S$ and $\boldsymbol{\epsilon}^C$, decreasing from discrepancies of twice the daily observed number of PCR-positives among early symptomatic cases and twice the daily number of new symptom onsets to 2/3 and 1.3 times the observed daily numbers respectively:

$$\boldsymbol{\epsilon}^S = (11,10,9,9,8,7,6,5,5,4)$$
$$\epsilon^C = (49,47,45,43,41,39,37,36,34,32).$$

Proposed values of (R_0, E_0, T) were accepted at each generation g only if all tolerances were satisfied, i.e. only if $d(D, D^*) \leq \epsilon_g$ and $d^S(D^S, D^{S*}) \leq \epsilon_g^S$ and $d^C(D^C, D^{C*}) \leq \epsilon_g^C$.

We assessed the performance of the algorithm by calculating the effective sample size (ESS) of the final generation of particles, $ESS = 1/\sum_{i=1}^{N} (w_G^{(i)})^2$, and the acceptance rate of proposed particles in each generation.

Details of San Francisco shelter outbreak

Full details of the outbreak in the San Franciso shelter are provided elsewhere [41]. Briefly, the data consisted of individual-level information on age, PCR test date and result, and partial data on symptom status, co-morbidity, and health outcome. The first two clinical cases identified in the shelter were confirmed on April 5, 2020 from PCR tests on April 4 and 5, 2020. The first case had symptom onset on March 31, the second on April 2. However, several individuals who later tested PCR-positive reported that they had symptom onset around these dates (Supplementary Table 3 and S5A Fig). After the first cases were identified on April 4, contact tracing, symptom screening and PCR testing of symptomatic individuals was performed up to April 7. On April 8 and 9 mass testing of residents and staff was performed. As of April 10, 2020, 89 individuals out of 175 tested in the shelter were PCR-positive, of whom 65 were pre-symptomatic/symptomatic (2 unknown status), 4 required hospitalization (2 unknown outcome), and 1 died. Supplementary Table 3 shows the numbers of positive test results returned from the different testing that was conducted by day of testing and number of new symptom onsets each day.

Demographics and movement of individuals in and out of the shelter

Even before the first cases in the shelter in San Francisco were identified, the shelter was not running at capacity (340 residents) and efforts were made to move high-risk individuals (those aged 60 and over or with co-morbidities) out of the shelter. Following identification of the first cases on April 5, progressively more and more individuals were removed from the shelter into isolation and quarantine at various sites; first those identified as close contacts and bedmates of the first cases and those suspected of being infected, then later on April 10 PCR-negative individuals, particularly high-risk individuals. Efforts were made to cohort the remaining residents into those who were PCR-positive and those with unknown COVID-status, but this was hampered by individuals returning to the shelter on the evening of April 10. The shelter was disbanded and all residents and staff were moved to isolation and quarantine sites on April 11.

According to the shelter register, there were a total of 255 residents who were present at some point from March 29 to April 10, 2020. Supplementary Figure 2 shows the breakdown of the number of residents in each of the different risk groups (no co-morbidities and under-60, co-morbidities and under-60, no co-morbidities and 60 or over, co-morbidities and 60 or over) present on each day from March 29 to April 10. We initialized the resident population of the shelter in the simulations such that the numbers in the different risk groups matched those present on March 29, and assumed the numbers in the different groups remained constant prior to March 29. We assumed that the proportions of the remaining individuals not present on March 29 in the different risk groups were the same as among those present. The movement of residents in and out of the shelter each day was simulated by randomly drawing individuals from each risk group to remove/add such that the number present in each risk group matched that in the register, accounting for the removal of symptomatic PCR-positive individuals from the testing from April 4–8 and their risk group.

A total of 64 staff, covering general running of the shelter, support services, maintenance, laundry and food services over 3 shifts per day with approximately 20-25 staff on each shift, were present at some point from March 29 to April 10. Due to a lack of detailed information on staff demographics and movement we assumed that all staff had low risk of clinical symptoms,

hospitalization and death (were all under-60 without co-morbidities), and were present in the shelter each day for approximately the same amount of time as the average resident, such that they had the same risk of infection as residents.

Estimation of impact of different interventions

We estimated the impact of six different intervention strategies, listed in Supplementary Table 6 with their component interventions, on the probability of averting an outbreak and the total numbers of infections, clinical cases, hospitalizations and deaths over 30 days in a shelter of 250 residents and 50 staff into which one latently infected individual is introduced by comparing output of simulations in which there were interventions with counterfactual simulations without any interventions. An outbreak was defined as ≥3 cases that originated within the shelter within any 14-day period, which we determined by probabilistically assigning an infection source (background transmission vs infectious individuals within the shelter) to each infected individual upon infection in the simulations and tracking the number of infections whose source was within the shelter over time. We ran 1000 simulations for each intervention strategy and the counterfactual scenario, and calculated the probability of averting an outbreak from pairs of counterfactual and intervention simulations as the proportion of simulation pairs with an outbreak in the no-intervention scenario in which there was no outbreak in the intervention scenario:

$$\mathbb{P}(\text{outbreak averted}) = \frac{\mathbb{P}(\text{no outbreak with intervention}|\text{outbreak without intervention})}{\mathbb{P}(\text{outbreak without intervention})}$$

$$= \frac{\#(\text{pairs with no outbreak in intervention simulation \& an outbreak in counterfactual simulation})}{\#(\text{counterfactual simulations with an outbreak})}$$

Reductions in cumulative incidence of infections and clinical cases under each intervention strategy were calculated as the median percentage reduction in total number of infections/clinical cases between the counterfactual simulation and the intervention simulation across all simulation pairs, where the percentage reduction was treated as 0 if there were no cases in the counterfactual simulation.

Scenario and sensitivity analyses

To assess the effect of the transmission potential within the shelter (R_0) and the background infection rate (i_a) on intervention impact, we predicted the impact of the different intervention strategies for the different R_0 estimates from the calibration ($R_0 = 2.9$ (Seattle A), 3.9 (Boston), 6.2 (San Francisco)) and $R_0 = 1.5$ (representing a lower-risk setting) for different background infection rates estimated from recent incidence of confirmed cases in Seattle, Boston and San Francisco. The background infection rates were estimated as in Equation (2) but with the limits in the sum replaced by July 4 and July 17, 2020 (to represent reported incidence for June 27–July 10, 2020, with a 7-day infection-to-reporting delay) and an infection-to-reported-case ratio, γ, of 4 [28]. This gave background infection rates ranging from 122/1,000,000/day for Boston to 439/1,000,000/day for San Francisco. We used the limits of this range and the mean across the three cities, along with a zero background infection rate, for the scenario analyses. The results are provided in Table 2 in the main text and Supplementary Tables 9 and 10. We also assessed the variation in the probability of averting an outbreak under each intervention strategy for a larger number of background infection rates over the same range (Figure 1 in the main text).

We assessed the sensitivity of the intervention impact estimates to uncertainty in key natural history and intervention parameters (relative infectiousness of subclinical infection and the early infectious stage, sensitivities and specificities of symptom screening and PCR tests, testing and masking compliances, and mask effectiveness) by simulating each intervention strategy with all combinations of minimum and maximum values of these parameters over their uncertainty ranges (Supplementary Table 5) for the base case background infection rate of 122/1,000,000/day. We then calculated the minimum and maximum values of the outcome measures (probability of averting an outbreak and reduction in total numbers of infections and clinical cases) over all parameter combinations to generate uncertainty intervals around the base case estimates of the outcome measures (Table 2 in the main text). The sensitivity of the probability of averting an outbreak to variation in the different parameter values is shown in Supplementary Figure 9 and discussed in the main text.

Supplementary results

Model calibration

Supplementary Figure 10 shows the posterior distributions and pairwise correlation plots for the calibrated parameters R_0, E_0 and T for each of the shelters. The considerable uncertainty in the parameter estimates due to the predominantly cross-sectional aggregate nature of the data is reflected in the broad posterior distributions, covering most of the range of the prior distributions for the parameters for all the shelters except the San Francisco shelter, and the strong correlation between R_0 and T for the Boston and San Francisco shelters.
The effective sample sizes of the output for the different shelters ranged between 640 for Seattle shelter C and 951 for the San Francisco shelter, indicating that a sufficient number of particles was used to estimate the posterior distributions. The acceptance rates varied across shelters and decreased over successive generations, remaining above 50% for all of the Seattle shelters but decreasing to 4% for the San Francisco shelter, but overall suggest that the algorithm sampled efficiently from the posterior distributions.

Impact of infection control strategies

The relative impact of the different infection control strategies on reducing cumulative infection incidence followed the same pattern as the probability of averting an outbreak (cf. Supplementary Table 10 with Table 2 in the main text and Supplementary Table 9). However, the percentage reduction in cumulative incidence varied non-linearly with R_0 due to the bimodal nature of the outbreak size distribution (Supplementary Figures 6–8), such that the highest percentage reductions were achieved for the $R_0 = 2.9$ scenario. Daily symptom screening alone, and daily symptom screening with relocation of high-risk individuals led to reductions in cumulative incidence of 9% for $R_0 = 6.2$ to 45% and 43% for $R_0 = 2.9$. Twice-weekly PCR testing of staff provided modest additional benefit, increasing the percentage reduction to 13-54%. Reductions under universal masking and twice-weekly PCR testing of all residents and staff were much greater (51–75% and 37–81%), though the impact of PCR testing attenuated more than that of masking with increasing R_0. The highest percentage reductions of 71–90% were achieved under the combination strategy, with the biggest gain from combining interventions occurring in the highest transmissibility setting ($R_0 = 6.2$).

The pattern of impact of the intervention strategies in terms of reduction in total numbers of clinical cases was the same (Supplementary Table 10), except for relocation of high-risk individuals, which led to greater percentage reductions in clinical cases than symptom screening and routine PCR testing of staff. Total numbers of hospitalizations and deaths over 30 days were small with or without interventions (medians ≤ 4) and therefore not considered relevant at the scale of a single shelter.

Supplementary Table 1. Definition of states in the transmission model

State	Symbol	Infectious	Symptomatic	Detectable viral load	Immune
Susceptible	$\mathcal{S}(t)$	×	×	×	×
Exposed to infection	$\mathcal{E}(t)$	×	×	×	×
Early subclinical infection	$\mathcal{I}_{s,1}(t)$	✓	×	✓	×
Late subclinical infection	$\mathcal{I}_{s,2}(t)$	✓	×/✓ (no/mild symptoms)	✓	×
Early clinical infection	$\mathcal{I}_{c,1}(t)$	✓	×	✓	×
Late clinical infection	$\mathcal{I}_{c,2}(t)$	✓	✓	✓	×
Recovered	$\mathcal{R}(t)$	×	×	✓/×	✓

Supplementary Table 2. Risk of clinical symptoms and hospitalization by age group and co-morbidity status

Risk group	Probability of developing clinical symptoms, $\rho(a_i)$	Probability of hospitalization for clinical cases	Probability of death for hospitalized cases admitted to ICU
Low risk: age <60 yrs + no co-morbidities	0.473	0.040	0.22
Moderate risk: age <60 yrs + co-morbidities	0.473	0.085	0.58
High risk: age ≥60 yrs + no co-morbidities	0.747	0.289	0.52
Very high risk: age ≥60 yrs + co-morbidities	0.747	0.618	1

Supplementary Table 3. Numbers of PCR-positive individuals by day of test result and daily new symptom onsets in San Francisco shelter March 28–April 10, 2020

Date	Number tested in random testing	Number PCR-positive in random testing	Number of early symptomatic cases tested	Number of early symptomatic cases PCR-positive	Number of new symptom onsets
Mar 28	1	0	0	-	2
Mar 29	1	0	0	-	0
Mar 30	2	0	1	0	3
Mar 31	0	-	0	-	2
Apr 1	0	-	0	-	3
Apr 2	0	-	0	-	1
Apr 3	0	-	0	-	1
Apr 4	1	0	1	1	2
Apr 5	0	-	1	1	3
Apr 6	0	-	3	2	7
Apr 7	1	0	5	5	3
Apr 8	89	35	-	-	4
Apr 9	64	44	-	-	16
Apr 10	5	1	-	-	15

Supplementary Table 4. Numbers of residents and staff PCR tested and PCR positive at three shelters in Seattle during two testing events March 30–April 1, 2020 and April 7–8, 2020

Shelter	Testing event 1 (Mar 30–Apr 1, 2020)		Testing event 2 (Apr 7–8, 2020)	
	No. tested	No. (%) positive	No. tested	No. (%) positive
Seattle A*				
Residents	43	7 (16.3)	-	-
Staff	15	4 (26.7)	-	-
Seattle B				
Residents	74	2 (2.7)	52	4 (7.7)
Staff	2	0 (0)	8	1 (12.5)
Seattle C				
Residents	37	6 (16.2)	44	10 (22.7)
Staff	10	0 (0)	7	1 (14.3)

* Shelter A closed April 5, 2020, so data from testing event 2 was not used.

Supplementary Table 5. Input parameters for microsimulation of COVID-19 transmission in homeless shelters

Parameter	Symbol	Base case value	Range in sensitivity analysis	References
Demography				
Number of residents				
Seattle A		43	-	[31]
Seattle B		109	-	[31]
Seattle C		93	-	[31]
Boston		408	-	[42]
San Francisco		Time-varying (see Supplementary Figure 2)	-	
Number of staff				
Seattle A		15	-	[31]
Seattle B		8	-	[31]
Seattle C		10	-	[31]
Boston		50	-	[42]
San Francisco		64	-	
Age group of individual i (<60 years, ≥60 years)	a_i	Shelter-specific (see text)		[31,32]
Natural history				
Mean duration of latent infection period, days	μ_E	3	-	[6]
Shape parameter of negative-binomially-distributed latent infection period	r_E	4		[1]
Mean duration of early infectious stage (subclinical/clinical), days	μ_1	2.3	-	[6]
Shape parameter of negative-binomially-distributed early infectious stage (subclinical/clinical)	r_1	4		[1]
Mean duration of late infectious stage (subclinical/clinical), days	μ_2	8	-	[6,9,43,44]
Shape parameter of negative-binomially-distributed late infectious stage (subclinical/clinical)	r_2	4		[1]

Relative infectiousness of subclinical infection to clinical infection	h	1	0.5–1	[8,45,46]
Relative infectiousness of early infectious stage to late infectious stage	α	2	1–3	[6,7]
Probability of developing clinical symptoms	$\rho(a_i)$	Age-dependent (see Supplementary Table 2)	-	[1]
Mean duration of detectable viral load from start of late infectious stage, days		20	-	[6,9–13]
Minimum duration of detectable viral load from start of late infectious stage, days		5	-	[6,9–13]
Maximum duration of detectable viral load from start of late infectious stage, days		37	-	[11,13]
Mean time from symptom onset to hospitalization, days		8		Assumed same as duration of late infectious stage [24,47]
Probability of hospitalization for clinical cases		Age- and co-morbidity dependent (see Supplementary Table 2)	-	[4]
Probability of ICU admission among hospitalized cases		0.261	-	[3]
Probability of death for hospitalized cases admitted to ICU		Age- and co-morbidity dependent (see Supplementary Table 2)	-	[4]
Infection-to-reported-case ratio	γ	10		[27]
Background infection rate in community outside shelter, infections/1,000,000 person-days	i_a			
Seattle A		561		[21]
Seattle B		543		[21]
Seattle C		543		[21]
Boston		2018		[22]
San Francisco		445		[23]
Relative risk of infection for homeless individuals	r	2		
Background infection rate in homeless community outside shelter, infections/1,000,000 person-days	$\epsilon = ri_a$			

Intervention scenarios				
Simulation duration, days		30		Assumed
Number of residents		250		Assumed
Number of staff		50		Assumed
Initial number of infected individuals		1 (assumed latent)		Assumed
Age- and co-morbidity stratification		Same as for San Francisco shelter		Assumed
Basic reproduction number	R_0			
“low-risk”		1.5		Assumed
“Seattle”		2.9		Calibrated
“Boston”		3.9		Calibrated
“San Francisco”		6.2		Calibrated
Infection-to-reported-case ratio	γ	4		[28]
Background infection rate in community outside shelter, infections/1,000,000 person-days	i_a		0–439	
No background infection		0		
Low		122		[22]
Moderate		253		[21–23]
High		439		[23]
Symptom screening				
Sensitivity		0.4	0.3–0.5	Assumed based on [48]
Specificity		0.9	0.8–0.9	Assumed
Compliance of symptomatic individuals with PCR testing, %		80	50–100	Assumed
PCR testing				
Sensitivity		0.75	0.6–0.9	[17–20]
Specificity		1	0.95–1	[17,20]
Frequency		Twice weekly	Daily–Monthly	[49–51]

Compliance, %	80	50–100	Assumed
Masks			
Effectiveness (reduction in transmission)	30	10-50	[52–57]
Compliance, %	80	50–100	Assumed

Supplementary Table 6. Different intervention strategies tested

Strategy	Interventions				
	Daily symptom screening	Twice-weekly PCR testing of residents	Twice-weekly PCR testing of staff	Universal masking	Relocation of high-risk individuals
(1) Symptom screening	✓	×	×	×	×
(2) Routine PCR testing	✓	✓	✓	×	×
(3) Universal mask wearing	✓	×	×	✓	×
(4) Relocation of high-risk individuals	✓	×	×	×	✓
(5) Routine PCR testing of staff only	✓	×	✓	×	×
(6) Combination strategy	✓	✓	✓	✓	✓

Supplementary Table 7. Estimated epidemiologic parameters based on observed outbreak data from homeless shelters in Seattle, Boston and San Francisco

Shelter	Basic reproduction number R_0, median (95% CI)*	Number of latently infected individuals who initially entered shelter E_0, median (95%CI)*	Number of days prior to first reported infection that infected individuals entered shelter† D, median (95% CI)*
Seattle A	2.9 (1.1–7.3)	3 (1–5)	16 (9–25)
Seattle B	2.9 (1.1–6.7)	3 (1–5)	10 (7–14)
Seattle C	3.0 (1.2–7.2)	3 (1–5)	15 (8–24)
Boston	3.9 (2.2–7.6)	3 (1–5)	15 (9–23)
San Francisco	6.2 (4.0–7.9)	3 (1–5)	21 (17–26)

CI = credible interval.

Data was available for three shelters in Seattle (labeled A-C).

* 95% CIs calculated as 2.5%–97.5% quantile interval of posterior distribution.

† D is calculated from the estimated time since introduction of infection into the shelter at the end of data collection, T, as:
$D = \text{date first case identified} - (\text{end date of data collection} - T) + 1$

Supplementary Table 8. Estimated cumulative infection incidence at the end of the PCR testing period in homeless shelters in Seattle, Boston and San Francisco

Shelter	Cumulative infection incidence, median (95% CI)*, %
Seattle A	40 (10–81)
Seattle B	14 (1–41)
Seattle C	37 (12–71)
Boston	64 (52–78)
San Francisco	83 (72–92)

CI = credible interval.

* 95% CIs calculated as 2.5%–97.5% quantile interval of posterior distribution.

Supplementary Table 9. Probability of averting an outbreak over a 30-day period in a generalized homeless shelter* with simulated infection control strategies for different background infection rates in the community outside the shelter

	Probability of averting an outbreak			
Infection control strategy	**$R_0 = 1.5$ (low-risk)**	**$R_0 = 2.9$ (Seattle)**	**$R_0 = 3.9$ (Boston)**	**$R_0 = 6.2$ (San Francisco)**
Background infection rate = 0				
(1) Symptom screening	0.61	0.31	0.21	0.12
(2) Routine twice-weekly PCR testing	0.78	0.40	0.28	0.12
(3) Universal mask wearing	0.73	0.44	0.31	0.16
(4) Relocation of high-risk individuals	0.61	0.31	0.21	0.11
(5) Routine twice-weekly PCR testing of staff only	0.67	0.34	0.24	0.13
(6) Combination strategy	0.86	0.56	0.41	0.20
Background infection rate = 253/1,000,000/day				
(1) Symptom screening	0.20	0.04	0.02	0.01
(2) Routine twice-weekly PCR testing	0.36	0.14	0.03	0.02
(3) Universal mask wearing	0.31	0.08	0.04	0.01
(4) Relocation of high-risk individuals	0.19	0.04	0.02	0.01
(5) Routine twice-weekly PCR testing of staff only	0.21	0.05	0.02	0.00
(6) Combination strategy	0.51	0.23	0.09	0.02
Background infection rate = 439/1,000,000/day				
(1) Symptom screening	0.07	0.01	0.01	0.00
(2) Routine twice-weekly PCR testing	0.22	0.03	0.01	0.00
(3) Universal mask wearing	0.14	0.02	0.01	0.01
(4) Relocation of high-risk individuals	0.07	0.01	0.01	0.00
(5) Routine twice-weekly PCR testing of staff only	0.08	0.01	0.00	0.00
(6) Combination strategy	0.34	0.08	0.03	0.01

R_0 = basic reproduction number.
* Generalized homeless shelter defined as 250 residents and 50 staff.

Supplementary Table 10. Reductions in the total number of infections and symptomatic cases over a 30-day period in a generalized homeless shelter* with simulated infection control strategies for different background infection rates in the community outside the shelter

Infection control strategy	Median reduction in total infections, %				Median reduction in total symptomatic cases, %			
	$R_0 = 1.5$ (low-risk)	$R_0 = 2.9$ (Seattle)	$R_0 = 3.9$ (Boston)	$R_0 = 6.2$ (San Francisco)	$R_0 = 1.5$ (low-risk)	$R_0 = 2.9$ (Seattle)	$R_0 = 3.9$ (Boston)	$R_0 = 6.2$ (San Francisco)
Background infection rate = 0								
(1) Symptom screening	50	60	41	10	40	55	43	24
(2) Routine PCR testing	86	85	78	38	60	75	75	53
(3) Universal mask wearing	75	87	83	56	59	77	79	67
(4) Relocation of high-risk individuals	50	60	40	10	44	58	48	31
(5) Routine PCR testing of staff only	71	65	56	17	50	54	52	35
(6) Combination strategy	93	95	94	81	75	88	90	86
Background infection rate = 122/1,000,000/day								
(1) Symptom screening	38	45	40	9	33	41	42	23
(2) Routine PCR testing	62	81	79	37	50	72	76	50
(3) Universal mask wearing	55	75	74	51	49	67	71	62
(4) Relocation of high-risk individuals	38	43	38	9	40	47	48	33
(5) Routine PCR testing of staff only	50	54	45	13	33	47	47	28
(6) Combination strategy	71	89	90	74	63	86	89	82
Background infection rate = 253/1,000,000/day								
(1) Symptom screening	33	43	31	5	29	39	32	19
(2) Routine PCR testing	55	73	64	24	44	67	60	42
(3) Universal mask wearing	53	71	65	39	44	63	63	56
(4) Relocation of high-risk individuals	32	40	30	5	35	45	40	27
(5) Routine PCR testing of staff only	39	49	38	7	33	44	39	21
(6) Combination strategy	61	84	82	62	58	81	80	75
Background infection rate = 439/1,000,000/day								
(1) Symptom screening	28	37	25	4	20	36	30	15
(2) Routine PCR testing	52	66	60	20	45	61	62	42
(3) Universal mask wearing	45	65	61	28	38	60	62	48

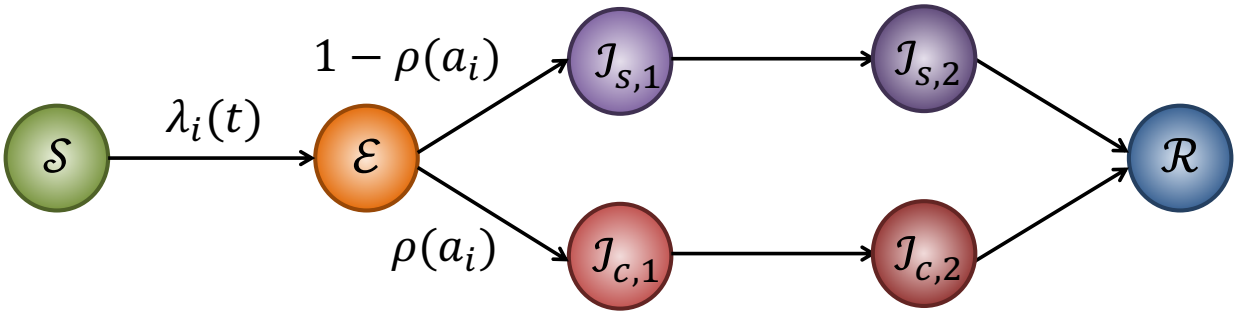
(4) Relocation of high-risk individuals	27	35	24	3	29	42	38	26
(5) Routine PCR testing of staff only	34	42	33	4	27	39	38	19
(6) Combination strategy	61	79	81	56	58	76	81	74

R_0 = basic reproduction number.
* Generalized homeless shelter defined as 250 residents and 50 staff.

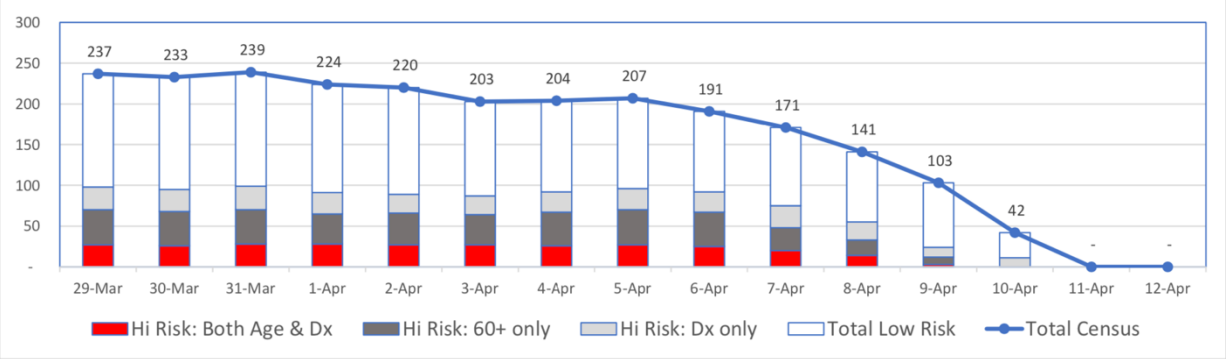
Supplementary Table 11. Numbers of PCR tests used under each infection control strategy

Infection control strategy	Mean number of tests used per person per month*
(1) Symptom screening	2.0
(2) Routine PCR testing	6.6
(3) Universal mask wearing	2.0
(4) Relocation of high-risk individuals	2.0
(5) Routine PCR testing of staff only	2.8
(6) Combination strategy	6.6

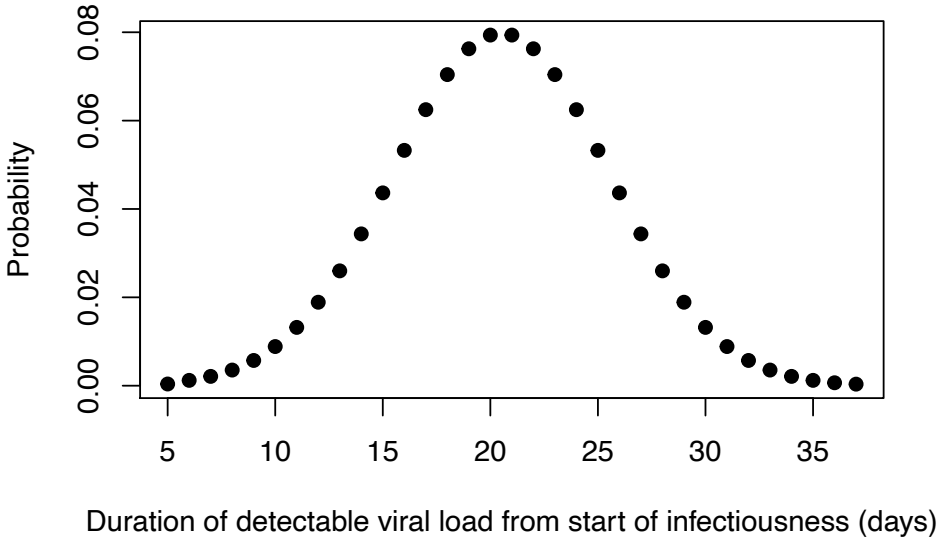
* All strategies use tests as they all include daily symptom screening.
Numbers only shown for $R_0 = 2.9$ (Seattle), chosen as representative example, and a background infection rate of 122/1,000,000/day as numbers vary little with R_0 and background rate.



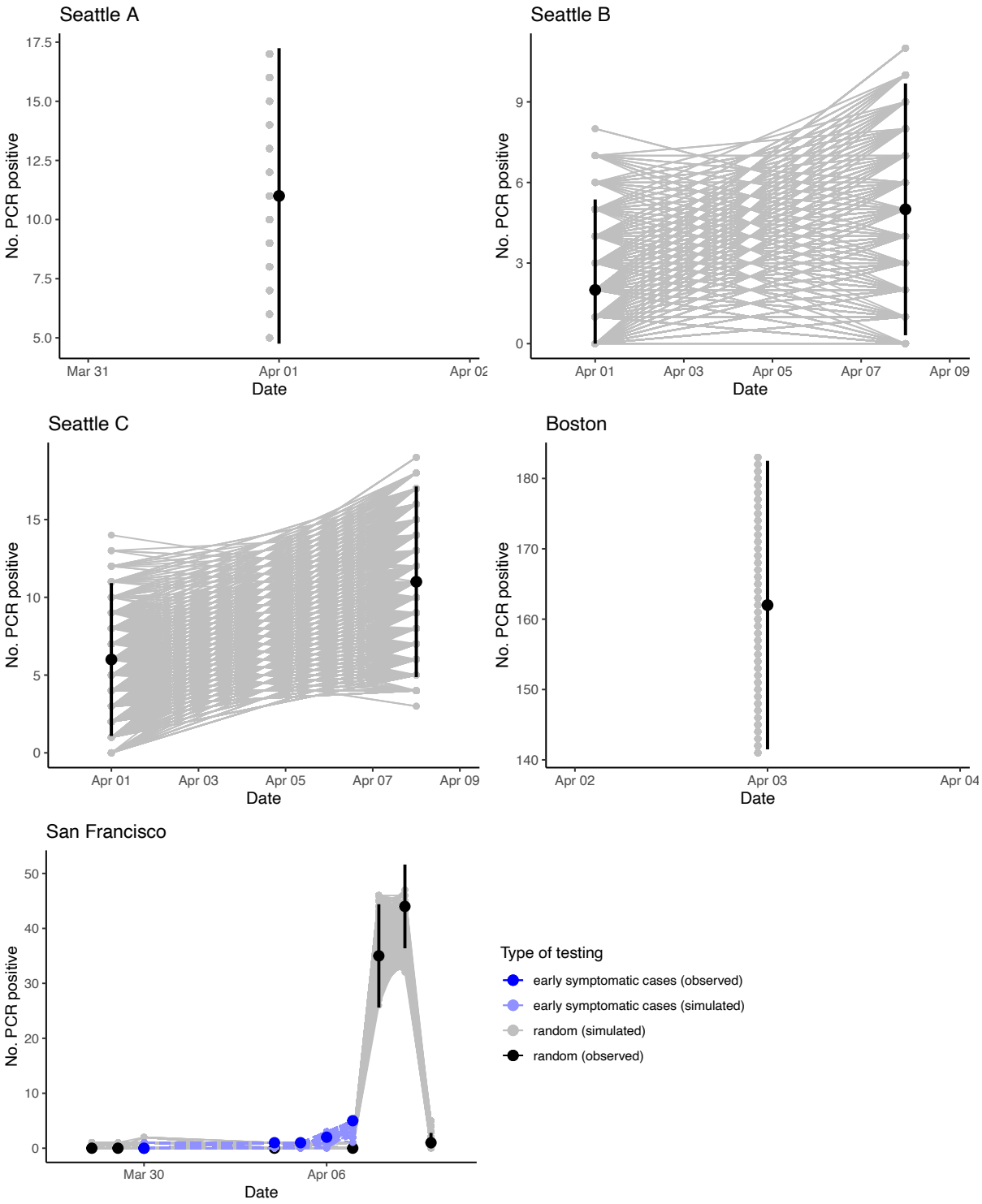
Supplementary Figure 1. Structure of stochastic individual-level susceptible-exposed-infectious-recovered ($\mathcal{S}$-$\mathcal{E}$-$\mathcal{I}$-$\mathcal{R}$) model of COVID-19 transmission in homeless shelter. Notation as defined in Supplementary Tables 1 and 5 and Equation (1).



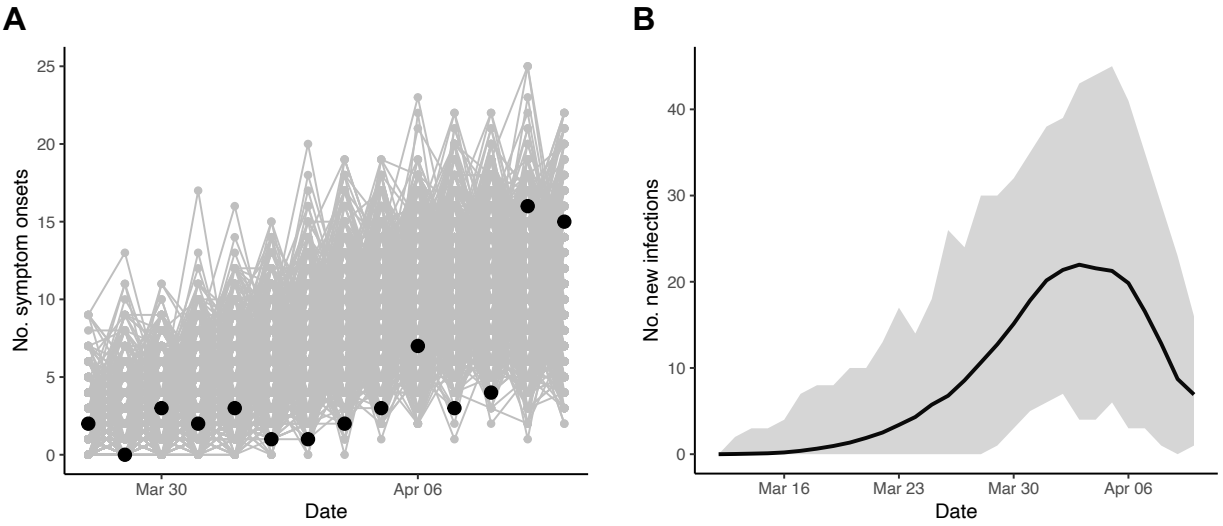
Supplementary Figure 2. Daily numbers of residents by risk group present in the San Francisco shelter March 29–April 10, 2020. Shelter was disbanded April 11.



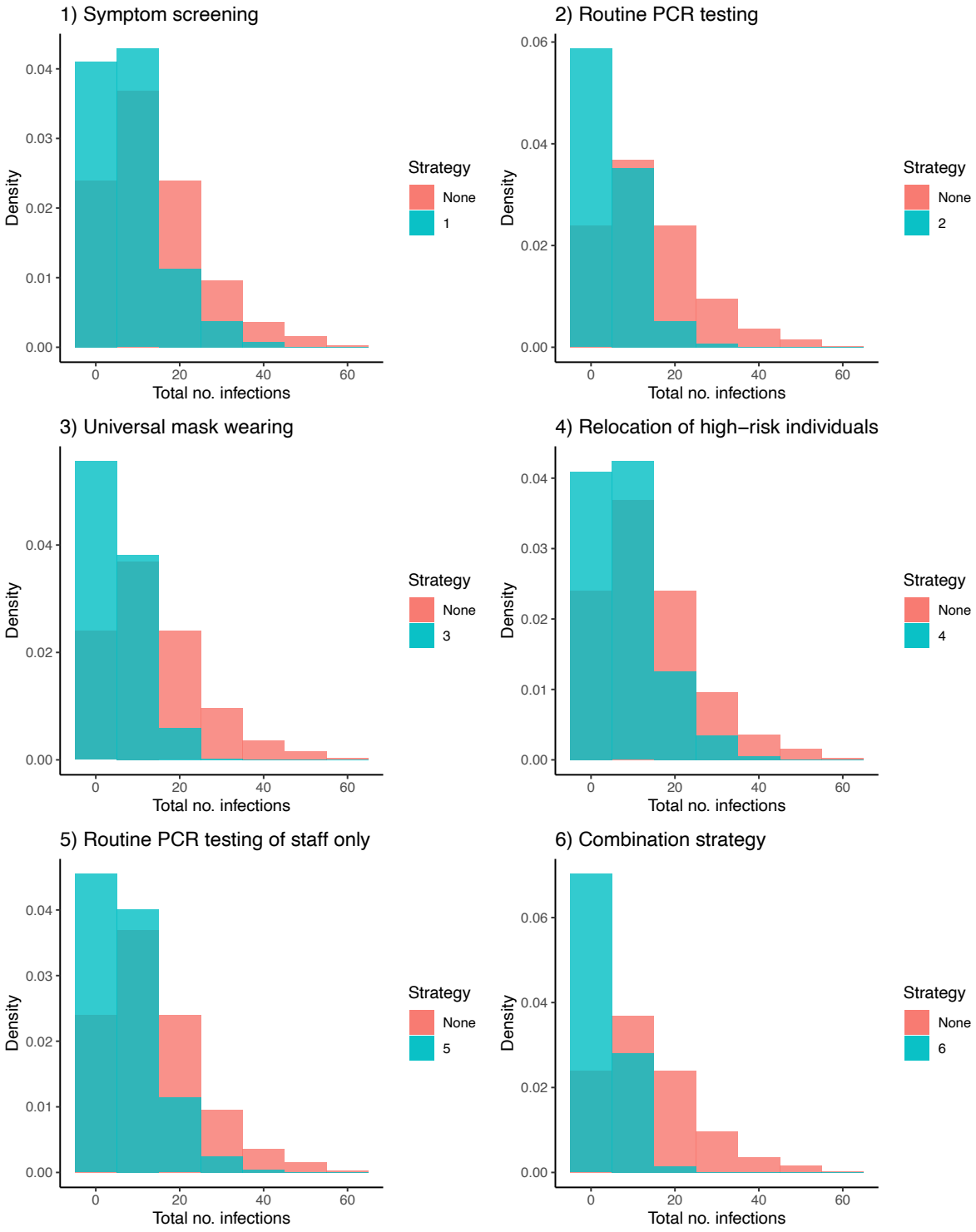
Supplementary Figure 3. Distribution of duration of detectable viral load from start of late infectious stage



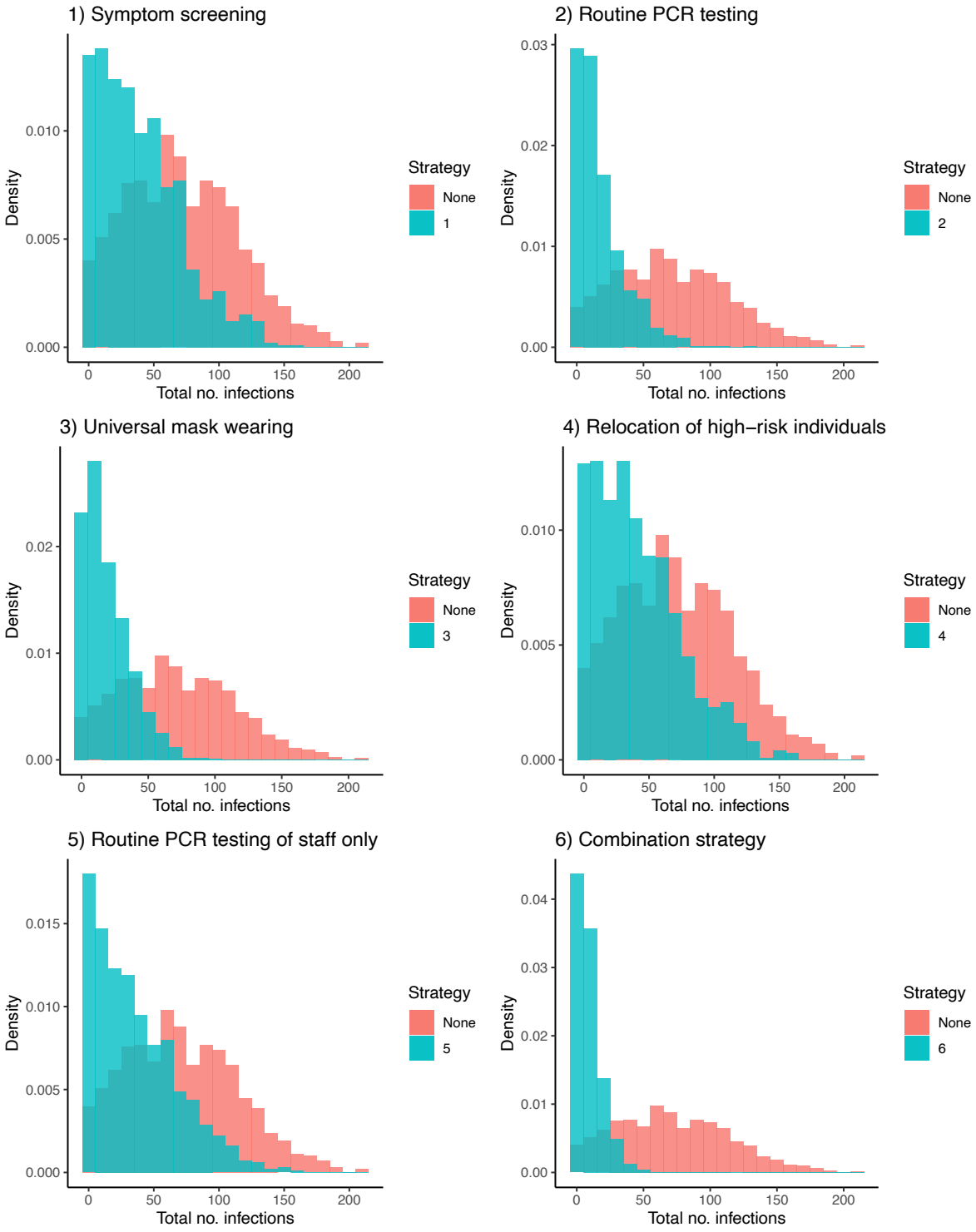
Supplementary Figure 4. Calibration of microsimulation to observed PCR testing data from outbreaks in homeless shelters in Seattle, Boston and San Francisco. Data was available for three shelters in Seattle (labeled A-C). Vertical black lines show exact binomial 95% confidence intervals for observed numbers of PCR-positive individuals in random testing.



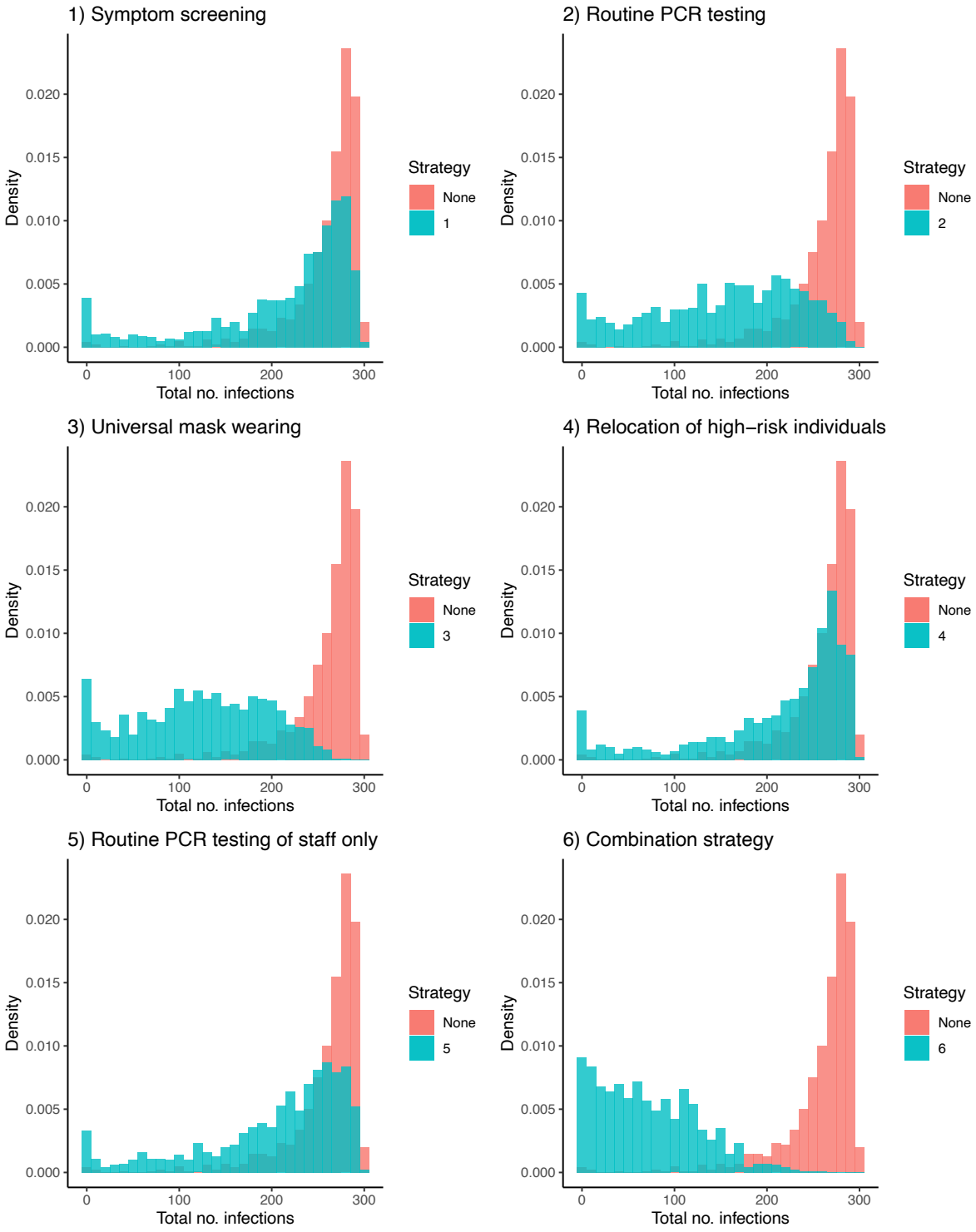
Supplementary Figure 5. Calibration of microsimulation to additional data from San Francisco shelter outbreak. (A) Calibration to daily numbers of symptom onsets. Black dots show reported number of symptom onsets, grey lines show simulation output. (B) Estimated daily numbers of new infections over time in 1000 calibrated simulations. Black line and grey shaded region show mean and range respectively across all simulations.



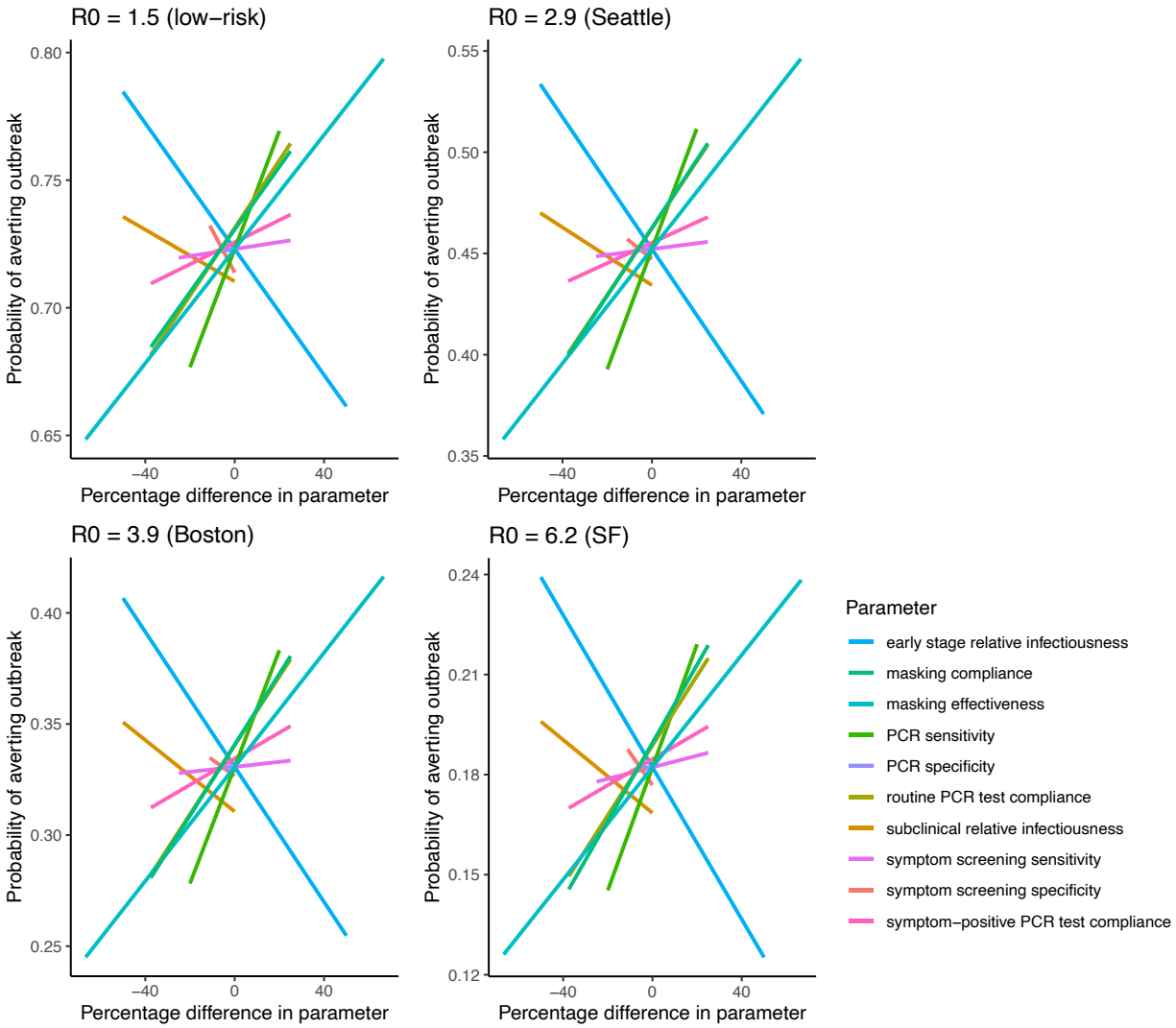
Supplementary Figure 6. Outbreak size distributions 30 days after introduction of infection in a generalized homeless shelter under different infection control strategies for $R_0 = 1.5$ (low-risk setting). Red and green histograms show outbreak size distributions with no interventions and under the intervention strategy respectively. Background infection rate of 122/1,000,000 person-days.



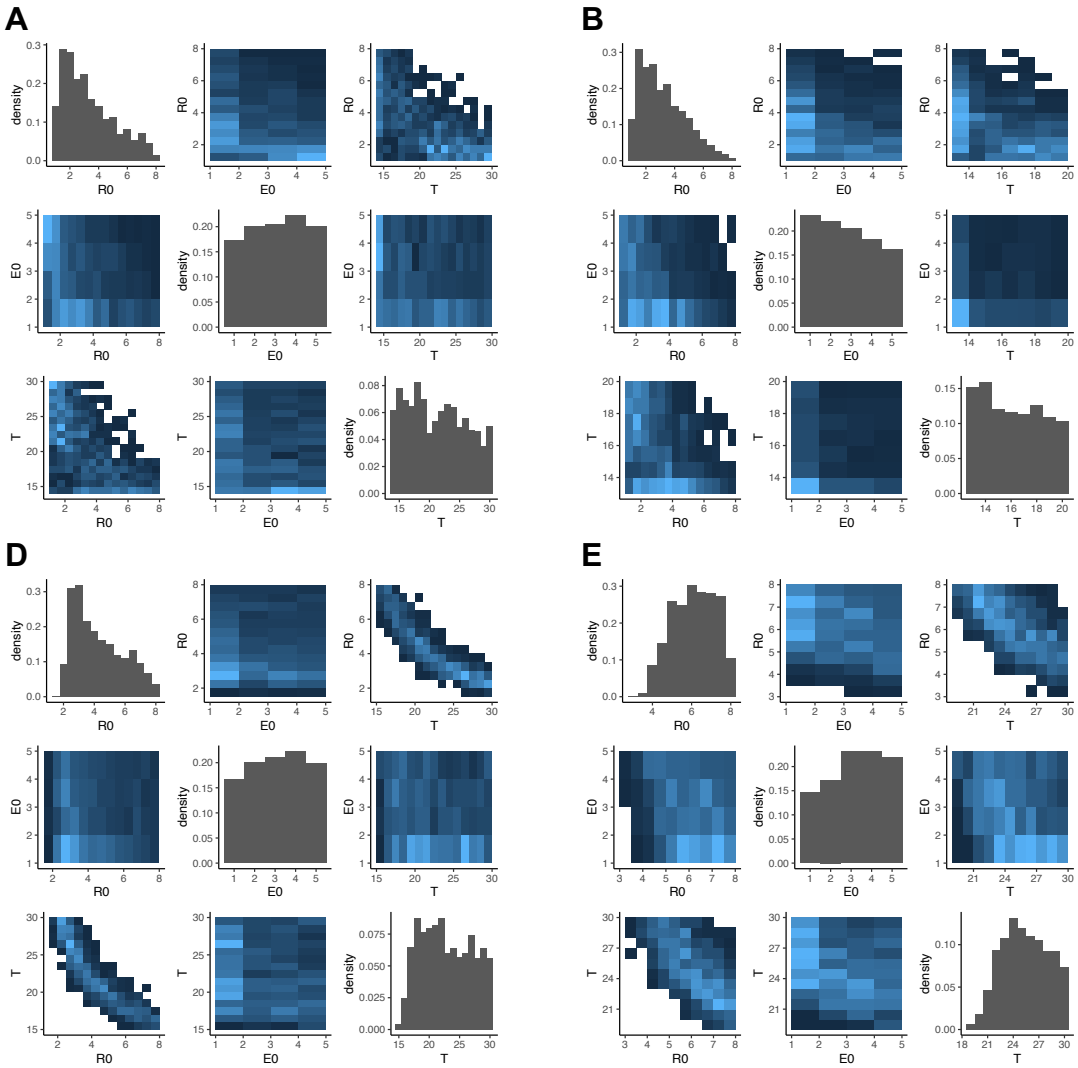
Supplementary Figure 7. Outbreak size distributions 30 days after introduction of infection in a generalized homeless shelter under different infection control strategies for $R_0 = 2.9$ (Seattle). Red and green histograms show outbreak size distributions with no interventions and under the intervention strategy respectively. Background infection rate of 122/1,000,000 person-days.



Supplementary Figure 8. Outbreak size distributions 30 days after introduction of infection in a generalized homeless shelter under different infection control strategies for $R_0 = 6.2$ (San Francisco). Red and green histograms show outbreak size distributions with no interventions and under the intervention strategy respectively. Background infection rate of 122/1,000,000 person-days.



Supplementary Figure 9. Spider diagrams showing the sensitivity of the estimated probability of averting an outbreak to variation in key natural history and intervention parameters for different R_0 values. The horizontal axis shows the percentage variation in each parameter relative to its base case value (see Supplementary Table 5). The vertical axis shows the mean probability of averting an outbreak for each parameter across all combinations of the minimums and maximums of the other parameter values.



Supplementary Figure 10. Posterior distributions and pairwise correlation plots for calibrated model parameters – R_0, E_0 and T – for (A)-(C) Seattle shelters A–C, (D) Boston shelter and (E) San Francisco shelter. R_0 = basic reproduction number; E_0 = number of latently infected individuals who initially entered the shelter; T = number of days before the end of data collection that these individuals entered the shelter.

References

1. Davies NG, Klepac P, Liu Y, Prem K, Jit M, Eggo RM. Age-dependent effects in the transmission and control of COVID-19 epidemics. medRxiv **2020**;
2. World Health Organization. Transmission of SARS-CoV-2: implications for infection prevention precautions. 2020. Available at: https://www.who.int/publications/i/item/modes-of-transmission-of-virus-causing-covid-19-implications-for-ipc-precaution-recommendations. Accessed 11 August 2020.
3. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. JAMA - J Am Med Assoc **2020**; 323:1061–1069.
4. Tuite AR, Fisman DN, Greer AL. Mathematical modelling of COVID-19 transmission and mitigation strategies in the population of Ontario, Canada. Cmaj **2020**; 192:E497–E505.
5. Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. Lancet Respir Med **2020**; 8:475–481. Available at: http://dx.doi.org/10.1016/S2213-2600(20)30079-5.
6. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med **2020**; 26:672–675.
7. Casey M, Griffin J, McAloon CG, et al. Estimating pre-symptomatic transmission of COVID-19: a secondary analysis using published data. medRxiv **2020**; :2020.05.08.20094870. Available at: http://medrxiv.org/lookup/doi/10.1101/2020.05.08.20094870.
8. Long QX, Tang XJ, Shi QL, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. Nat Med **2020**;
9. Wölfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized patients with COVID-2019. Nature **2020**; 581:465–469.
10. To KKW, Tsang OTY, Leung WS, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis **2020**; 20:565–574. Available at: http://dx.doi.org/10.1016/S1473-3099(20)30196-1.
11. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet **2020**; 395:1054–1062. Available at: http://dx.doi.org/10.1016/S0140-6736(20)30566-3.
12. Zou L, Ruan F, Huang M, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. N Engl J Med **2020**; 382:1177–1179.
13. Xiao AT, Tong YX, Gao C, Zhu L, Zhang YJ, Zhang S. Dynamic profile of RT-PCR findings from 301 COVID-19 patients in Wuhan, China: A descriptive study. J Clin Virol **2020**; 127:0–6.
14. Zhou R, Li F, Chen F, et al. Viral dynamics in asymptomatic patients with COVID-19. Int J Infect Dis **2020**; 96:288–290. Available at: https://doi.org/10.1016/j.ijid.2020.05.030.
15. Liu Y, Yan LM, Wan L, et al. Viral dynamics in mild and severe cases of COVID-19. Lancet Infect Dis **2020**; 20:656–657. Available at: http://dx.doi.org/10.1016/S1473-3099(20)30232-2.
16. Heald-Sargent T, Muller WJ, Zheng X, Rippe J, Patel AB, Kociolek LK. Age-Related Differences in Nasopharyngeal Severe Acute Respiratory Syndrome Coronavirus 2

(SARS-CoV-2) Levels in PatientsWith Mild to Moderate Coronavirus Disease 2019 (COVID-19) Children. JAMA - J Am Med Assoc **2020**; :E1–E2.
17. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction–Based SARS-CoV-2 Tests by Time Since Exposure. Ann Intern Med **2020**;
18. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, et al. False-Negative Results of Initial RT-PCR Assays for Covid-19: a Systematic Review. **2020**; :1–26.
19. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 test result. BMJ **2020**; 369:1–7. Available at: http://dx.doi.org/doi:10.1136/bmj.m1808.
20. Padhye NS. Reconstructed diagnostic sensitivity and specificity of the RT-PCR test for COVID-19. medRxiv **2020**; 19:2020.04.24.20078949. Available at: http://medrxiv.org/lookup/doi/10.1101/2020.04.24.20078949.
21. King County Department of Public Health. Daily COVID-19 outbreak summary - King County. 2020. Available at: https://www.kingcounty.gov/depts/health/covid-19/data/daily-summary.aspx. Accessed 13 August 2020.
22. City of Boston. COVID-19 Case Tracker. 2020. Available at: https://dashboard.cityofboston.gov/t/Guest_Access_Enabled/views/COVID-19/Dashboard1?:showAppBanner=false&:display_count=n&:showVizHome=n&:origin=viz_share_link&:isGuestRedirectFromVizportal=y&:embed=y. Accessed 13 August 2020.
23. San Francisco Department of Public Health. COVID-19 Cases Summarized by Date, Transmission and Case Disposition - City and County of San Francisco. 2020. Available at: https://data.sfgov.org/COVID-19/COVID-19-Cases-Summarized-by-Date-Transmission-and/tvq9-ec9w. Accessed 13 August 2020.
24. MIDAS Network. COVID-19 parameter estimates. Available at: https://github.com/midas-network/COVID-19/tree/master/parameter_estimates/2019_novel_coronavirus. Accessed 11 August 2020.
25. Abbott S, Hellewell J, Munday J, Funk S. The transmissibility of novel Coronavirus in the early stages of the 2019-20 outbreak in Wuhan: Exploring initial point-source exposure sizes and durations using scenario analysis. Wellcome Open Res **2020**; 5:1–12.
26. U.S. Census Bureau. U.S. Census Bureau QuickFacts. 2020. Available at: https://www.census.gov/quickfacts/fact/table/seattlecitywashington,bostoncitymassachusetts,sanfranciscocitycalifornia/PST045219. Accessed 1 September 2020.
27. Havers FP, Reed C, Lim T, et al. Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020. JAMA Intern Med **2020**; 30329:1–11.
28. Centers for Disease Control. Interactive Serology Dashboard for Commercial Laboratory Surveys. 2020. Available at: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/commercial-labs-interactive-serology-dashboard.html. Accessed 21 August 2020.
29. Andrews D, Salcedo F, Donahue L, et al. Seattle/King County Point-in-time Count of Persons Experiencing Homelessness. 2019. Available at: www.allhomekc.orgwww.appliedsurveyresearch.org. Accessed 20 August 2020.
30. King County Department of Public Health. Homelessness and COVID-19. 2020. Available at: https://www.kingcounty.gov/depts/health/covid-19/data/homeless.aspx. Accessed 21 August 2020.
31. Tobolowsky FA, Gonzales E. COVID-19 Outbreak Among Three Affiliated Homeless Service Sites — King County, Washington, 2020. Morb Mortal Wkly Rep **2020**; 69:523–526.

32. Baggett TP, Keyes H, Sporn N, Gaeta JM. Prevalence of SARS-CoV-2 Infection in Residents of a Large Homeless Shelter in Boston. JAMA - J Am Med Assoc **2020**; 323:2191–2192.
33. Minter A, Retkute R. Approximate Bayesian Computation for infectious disease modelling. Epidemics **2019**; 29:100368. Available at: https://doi.org/10.1016/j.epidem.2019.100368.
34. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. J R Soc Interface **2009**; 6:187–202.
35. Toni T, Stumpf MPH. Simulation-based model selection for dynamical systems in systems and population biology. Bioinformatics **2009**; 26:104–110.
36. Lau MS, Grenfell B, Nelson K, Lopman B. Characterizing super-spreading events and age-specific infectivity of COVID-19 transmission in Georgia, USA. medRxiv **2020**; :2020.06.20.20130476. Available at: https://www.medrxiv.org/content/10.1101/2020.06.20.20130476v2%0Ahttps://www.medrxiv.org/content/10.1101/2020.06.20.20130476v2.abstract.
37. Endo A, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. **2020**; :1–13.
38. James A, Eagle L, Phillips C, et al. High COVID-19 Attack Rate Among Attendees at Events at a Church — Arkansas, March 2020. Morb Mortal Wkly Rep **2020**; 69:632–635.
39. County S, Hamner L, Dubbel P, et al. High SARS-CoV-2 Attack Rate Following Exposure at a Choir Practice. Morb Mortal Wkly Rep **2020**; 69:606–610. Available at: https://www.cdc.gov/mmwr/volumes/69/wr/mm6919e6.htm.
40. Szablewski CM, Chang KT, Brown MM, et al. SARS-CoV-2 Transmission and Infection Among Attendees of an Overnight Camp - Georgia, June 2020. Morb Mortal Wkly Rep **2020**; 69:1023–1025. Available at: http://www.ncbi.nlm.nih.gov/pubmed/32759921.
41. Imbert E, Kinley PM, Scarborough A, et al. Coronavirus Disease 2019 (COVID-19) Outbreak in a San Francisco Homeless Shelter. Clin Infect Dis **2020**; Available at: https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa1071/5879965. Accessed 21 August 2020.
42. Mosites E, Parker EM, N Clarke KE, et al. Assessment of SARS-CoV-2 Infection Prevalence in Homeless Shelters — Four U.S. Cities, March 27–April 15, 2020. **2019**; 69. Available at: https://www.medrxiv.org/content/10.1101/2020.04.12.20059618v1.
43. Bullard J, Dust K, Funk D, et al. Predicting infectious SARS-CoV-2 from diagnostic samples. Clin Infect Dis **2020**; Available at: https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa638/5842165. Accessed 11 August 2020.
44. Chang D, Mo G, Yuan X, et al. Time Kinetics of Viral Clearance and Resolution of Symptoms in Novel Coronavirus Infection. Am J Respir Crit Care Med **2020**; 201:1150–1152.
45. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science (80-) **2020**; 368:489–493.
46. Chen Y, Wang A, Yi B, et al. Epidemiological characteristics of infection in COVID-19 close contacts in Ningbo city. Chinese J Epidemiol **2020**; 41:667–671.
47. Centers for Disease Control. COVID-19 Pandemic Planning Scenarios. 2020. Available at: https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html. Accessed 21 August 2020.

48. Gostic KM, Gomez ACR, Mummah RO, Kucharski AJ, Lloyd-Smith JO. Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19. Elife **2020**; 9:1–18.
49. Larremore DB, Wilder B, Lester E, et al. Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance. medRxiv **2020**; :2020.06.22.20136309. Available at: https://www.medrxiv.org/content/10.1101/2020.06.22.20136309v2%0Ahttps://www.medrxiv.org/content/10.1101/2020.06.22.20136309v2.abstract.
50. Paltiel AD, Zheng A, Walensky RP. Assessment of SARS-CoV-2 Screening Strategies to Permit the Safe Reopening of College Campuses in the United States. JAMA Netw Open **2020**; 3:e2016818. Available at: http://www.ncbi.nlm.nih.gov/pubmed/32735339.
51. Chin ET, Huynh BQ, Murrill M, Basu S, Lo NC. Frequency of routine testing for COVID-19 in high-risk environments to reduce workplace outbreaks. medRxiv **2020**; :2020.04.30.20087015. Available at: http://medrxiv.org/content/early/2020/06/22/2020.04.30.20087015.abstract.
52. Leung NHL, Chu DKW, Shiu EYC, et al. Respiratory virus shedding in exhaled breath and efficacy of face masks. Nat Med **2020**; 26:676–680. Available at: http://dx.doi.org/10.1038/s41591-020-0843-2.
53. MacIntyre CR, Chughtai AA. Facemasks for the prevention of infection in healthcare and community settings. BMJ **2015**; 350:1–12.
54. MacIntyre CR, Chughtai AA. A rapid systematic review of the efficacy of face masks and respirators against coronaviruses and other respiratory transmissible viruses for the community, healthcare workers and sick patients. Int J Nurs Stud **2020**; 108:0–5.
55. Brainard JS, Jones N, Lake I, Hooper L, Hunter P. Facemasks and similar barriers to prevent respiratory illness such as COVID-19: A rapid systematic review. **2020**;
56. Chan JF-WW, Yuan S, Zhang AJ, et al. Surgical mask partition reduces the risk of non-contact transmission in a golden Syrian hamster model for Coronavirus Disease 2019 (COVID-19). Clin Infect Dis **2020**; Available at: https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa644/5848814. Accessed 11 August 2020.
57. Mueller A, Eden M, Oakes J, Bellini C, Fernandez L. Quantitative Method for Comparative Assessment of Particle Removal Efficiency of Fabric Masks as Alternatives to Standard Surgical Masks for PPE. Matter **2020**; Available at: https://doi.org/10.1016/j.matt.2020.07.006.