Measuring Basic Reproduction Number to Assess Effects of Nonpharmaceutical Interventions on Nosocomial SARS-CoV-2 Transmission

George Shirreff, Jean-Ralph Zahar, Simon Cauchemez, Laura Temime,[1] Lulla Opatowski,[1]
EMEA-MESuRS Working Group on the Nosocomial Modelling of SARS-CoV-2[2]

Outbreaks of SARS-CoV-2 infection frequently occur in hospitals. Preventing nosocomial infection requires insight into hospital transmission. However, estimates of the basic reproduction number (R_0) in care facilities are lacking. Analyzing a closely monitored SARS-CoV-2 outbreak in a hospital in early 2020, we estimated the patient-to-patient transmission rate and R_0. We developed a model for SARS-CoV-2 nosocomial transmission that accounts for stochastic effects and undetected infections and fit it to patient test results. The model formalizes changes in testing capacity over time, and accounts for evolving PCR sensitivity at different stages of infection. R_0 estimates varied considerably across wards, ranging from 3 to 15 in different wards. During the outbreak, the hospital introduced a contact precautions policy. Our results strongly support a reduction in the hospital-level R_0 after this policy was implemented, from 8.7 to 1.3, corresponding to a policy efficacy of 85% and demonstrating the effectiveness of nonpharmaceutical interventions.

Despite sweeping control measures, SARS-CoV-2 continues to pose a major threat to older persons and persons with comorbidities, both of whom can have poorer clinical outcomes (*1*,*2*). Thus, hospitals and long-term care facilities (LTCFs) must be particularly vigilant to prevent the spread of SARS-CoV-2 infection among their patients. Nosocomial spread has been an issue since the pandemic began in 2020, and many outbreaks have occurred in hospitals and healthcare facilities, often with high attack and mortality rates (*3*).

To control nosocomial spread, healthcare facilities have progressively implemented preventive measures, such as generalized masking, testing campaigns among patients and staff, isolation, visitor restrictions (*3*), and more recently vaccination (*4*). However, the risk for viral transmission among hospital patients and staff and the effectiveness of control measures remain unclear, and outbreaks still occur (*3*,*5*,*6*).

The basic reproduction number (R_0) refers to the number of secondary infections caused by a single index infection in an otherwise susceptible population. R_0 has been widely used as an indicator of SARS-CoV-2 epidemic risk and has also proved valuable for evaluating testing strategies and other preventive measures within healthcare settings (*7*,*8*). R_0 likely varies between types of healthcare facilities and differs considerably from estimates in the general community (*9*). However, estimating R_0 in healthcare settings is more challenging than estimating R_0 in the community. The populations in institutions are small and epidemics are highly stochastic. More data usually are available from hospitals or wards that have more cases. Healthcare facilities rarely test patients randomly or at multiple times during their hospitalizations. Most available data from hospital outbreaks consist of distributions of positive tests over time in a context of evolving testing policy and capacity.

Author affiliations: Institut Pasteur, Université Paris Cité, Paris, France (G. Shirreff, S. Cauchemez, L. Opatowski); Conservatoire National des Arts et Métiers, Paris (G. Shirreff, L. Temime); University of Versailles Saint-Quentin-en-Yvelines, Montigny-le-Bretonneux, France (G. Shirreff, L. Opatowski); Assistance Publique–Hôpitaux de Paris, Bobigny, France (J.-R. Zahar); PACRI Unit, Paris (L. Temime)

DOI: https://doi.org/10.3201/eid2807.212339

[1]These authors were co-principal investigators.

[2]Members of the workgroup are listed at the end of this article.

At the beginning of the pandemic, most countries had no standard strategy or recommendation on how surveillance should be carried out and tests distributed. Testing was mostly conducted on symptomatic patients, and surveillance consisted of possible contact tracing around detected cases. However, unreported asymptomatic cases could represent a substantial fraction of transmissions, and little data on the testing policy are available to estimate how many cases fell through the gaps.

Here, we propose a new framework to analyze detailed hospital test data by using a stochastic transmission model explicitly accounting for testing policy. We estimated R_0 in the context of a large SARS-CoV-2 outbreak in a LTCF. The outbreak had a high initial R_0, and we reconstructed the unobserved epidemic to assess effectiveness of nonpharmaceutical interventions.

Methods

Hospital and Patient Information

Available data came from a LTCF in Paris, France. The hospital has 3 buildings (A, B, and C), each of which has 4 floors (0–3) that we considered as separate wards. The results of all valid PCR tests were available for each patient identification number during March 1–April 30, 2020 (61 days). Patient information also included the ward to which they were admitted or transferred, admission and discharge dates, and any symptoms they had at first positive test. All dates we provide are relative to the date of the first positive sample in the facility. We censored the data from day 51 onward because the hospital began to change the containment policy after that point. We excluded 23 patients from any ward-level analysis because the ward in which they were tested was unknown (Appendix, https://wwwnc.cdc.gov/EID/article/28/7/21-2339-App1.pdf). We only used anonymized, aggregated patient data and did not collect additional patient data beyond those for clinical use. The Comité Local d'Ethique pour la Recherche Clinique des HUPSSD Avicenne-Jean Verdier-René Muret approved the study as protocol no. CLEA-2021-190.

Laboratory Testing

The LTCF collected all nasopharyngeal swab samples from patients. Reasons for testing included having symptoms characteristic of SARS-CoV-2, having had contact with a positive case, or patient transfer between wards or into or out of the hospital (Appendix).

Model Description

We modeled the spread of infection within the LTCF population by using a modified stochastic susceptible-exposed-infected-recovered model (Figure 1; Appendix, Appendix Table 1). We defined the force of infection at a given time, $\lambda(t)$, as the per-capita rate at which susceptible persons become infected, which we determined by the transmission rate, β, and the proportion of infectious patients at that time (Appendix). On the date the epidemic began (t_{init}), we considered a specific number (E_{init}) of persons infected. We assumed persons in infectious incubation had reduced infectiousness by a factor of ε, compared with symptomatic infected persons. Similarly, we assumed asymptomatic infectious persons had lower infectiousness by a factor of κ_1.

To fully determine transmission over the outbreak period, we compared 2 distinct models. In the primary model, we assumed a single transmission rate, β, throughout the study period. However, based on knowledge of changing practices within the hospital, we defined a more complex, 2-phase model in which each phase had its own transmission rate, β_1 and β_2, and was delimited by an inflection date, $t_{inflect}$. Potential values for $t_{inflect}$ ranged from day 1, which was the date of the first positive sample, through day 16, which was >1 week after the facility introduced contact precautions and France implemented a generalized lockdown.

We directly computed R_0 for each stage of infection from the transmission rate, duration of each infectious stage, and the probability infected persons would become symptomatic (Appendix). For the 2-phase model, we computed the average R_0 by weighting each phase by its duration (Appendix).

Observation Model

Because of asymptomatic infections, imperfect test sensitivity, and irregular availability of tests, the facility could not identify all infected patients. To account for the imperfect reporting, we added an observation model to the transmission model (Appendix, Appendix Figure 1). The observation model assumes all persons are initially untested, but upon testing, the model moves them to an equivalent tested state. Any patient can be retested in the model, but retesting occurs at a reduced relative rate, ϕ, estimated directly from the number of tests and retests in the available data (Appendix). When a person in the model develops symptoms, they lose their tested status and rejoin the untested compartment, I_s (Figure 1), enabling the model to account for increased testing when symptoms appear in a patient.

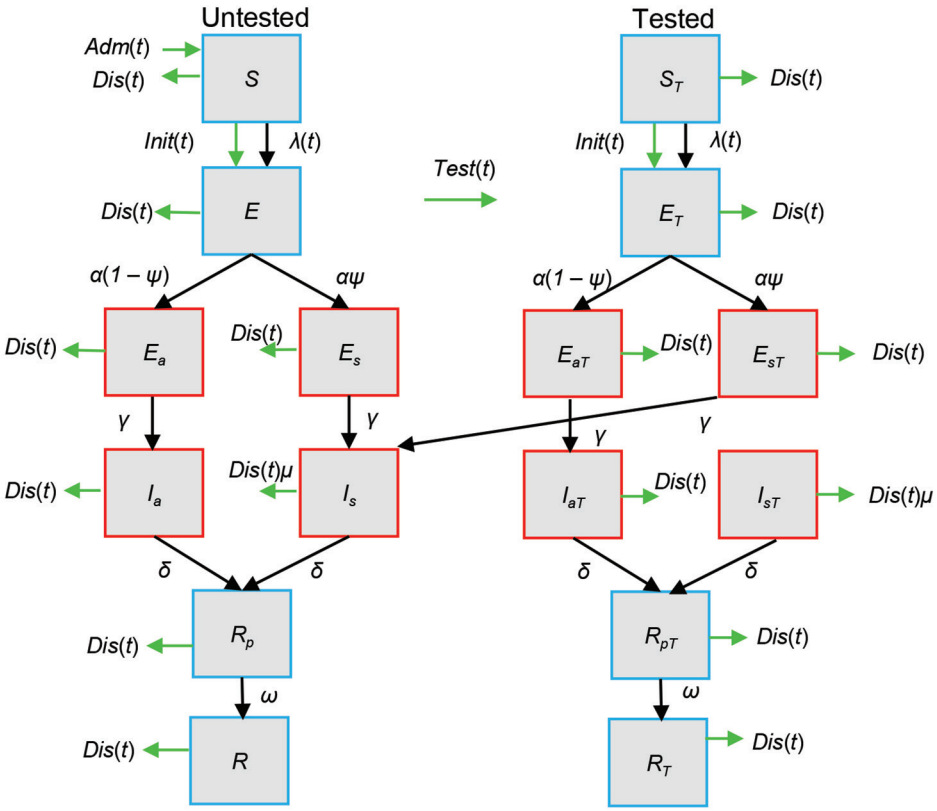


Figure 1. Compartmental susceptible-exposed-infectious-recovered model used to estimate nosocomial SARS-CoV-2 transmission rates on the basis of data for a long-term care facility in France. Red boxes indicate SARS-CoV-2 infectious compartments and blue boxes indicate noninfectious compartments. The left side shows the trajectory of untested persons, the right side shows tested persons. If untested persons are tested at any point in state *X*, they will enter the equivalent tested compartment (X_T, right panel), which is epidemiologically identical except for the testing rate. Patients in the susceptible state (*S*) can become infected by contact with infectious patients. When infected, patients move to the noninfectious incubation (*E*) compartment, after which they can either enter an asymptomatic or a symptomatic pathway of infectiousness. Each pathway has an infectious incubation period (E_a, E_s) before asymptomatic (I_a) or symptomatic (I_s) infection begins. After full infection, patients recover into a noninfectious state (R_p) where they are still likely to test positive before full recovery (*R*) when the probability of testing positive diminishes to (1 – test specificity). Green arrows refer to processes, initiation (*Init*), admission (*Adm*), discharge (*Dis*), and testing (*Test*), that occur a specified number of times on a given day according to model inputs. Black arrows indicate processes that are natural for infection and are entirely stochastic (Appendix Methods, Figure 1). *E*, exposed; E_a, asymptomatic exposed; E_{aT}, asymptomatic exposed and tested; E_s, symptomatic exposed; E_{sT}, symptomatic exposed and tested; E_T, exposed and tested; *I*, infectious; I_a, asymptomatic infectious; I_{aT}, asymptomatic infectious and tested; I_s, symptomatic infectious; I_{sT}, symptomatic infectious and tested; I_T, infectious and tested; *R*, recovered; R_p, recovered to noninfectious state; R_{pT}, recovered to noninfectious state and tested; R_T, recovered and tested; *S*, susceptible; *t*, time; α, rate of progression from noninfectious incubation; ψ, proportion of patients entering symptomatic pathway; λ(t), force of infection at time t; α, rate of progression from infectious incubation; δ, rate of progression from symptomatic infection; μ, relative rate of discharge for symptomatic patients relative to any nonsymptomatic patient; ω, rate at which viral shedding ceases during recovery.

However, testing does not change the rates of infectiousness or disease progression.

We used hospital data on the number of admissions, discharges, and tests per day as inputs (Appendix, Appendix Figure 2). The model considers admitted patients are in a susceptible untested state and are discharged at random from any state with a relative rate, μ, for symptomatic patients. For any day that tests are performed, the model prioritizes patients who have not been tested since becoming symptomatic and conducts any remaining tests at random on the rest of the population (Appendix, Appendix Figure 1). We used the sensitivity and specificity of the PCR test at the stage of infection to determine whether patients test positive or negative for SARS-CoV-2.

Statistical Inference

We calculated the likelihood by comparing the observed numbers of positive and negative cases on each day with the expected numbers generated by the internal model state via the observation process, assuming a binomial distribution (Appendix). We used iterative filtering in the pomp package (*10*) in R (R Foundation for Statistical Computing, https://www.r-project.org) to estimate parameters. In addition to estimating transmission rates, β, or β_1 and β_2, we also estimated the virus introduction time, t_{init}, and fixed the initial number of infections, E_{init}, to 1. For each analysis comprising the same model, dataset, and fixed parameter values, we used profile likelihood to calculate 95% CI for the estimated parameters (Appendix). We

compared models by calculating the Akaike information criterion (AIC).

Model Inference Validation

As a preliminary step, we tested the model and inference methodology on synthetic data. We used this test to ensure that known simulated transmission rates (β, or β_1 and β_2) and t_{init} could be recovered by statistical inference (Appendix).

Hospital- and Ward-Level Analyses

We first analyzed data at the hospital level, assuming homogeneous mixing across all buildings and wards. We then analyzed the data and estimated parameters for each ward separately. After parameter estimation, we conducted simulations of the visible and undetected parts of the epidemic at both the hospital and ward levels (Appendix).

Sensitivity Analysis and Time-Varying Reproduction Number

We conducted a sensitivity analysis to identify parameters with variations that most affected our estimated parameters. We perturbed the input parameters, using the lower and upper bound of the CI reported in the literature, and replicated the analysis. For comparison, we used incident cases to calculate the time-varying reproduction number (R_t) across the entire hospital by using the EpiEstim package (https://CRAN.R-project.org/package=EpiEstim) (Appendix).

Results

A total of 459 patients were in the hospital during the study period. PCR testing began on day −6; we consider day 1 as the first positive sample was collected. By the end of day 50, 152/312 patients sampled tested positive (Figure 2, panels A, B). The secondary attack

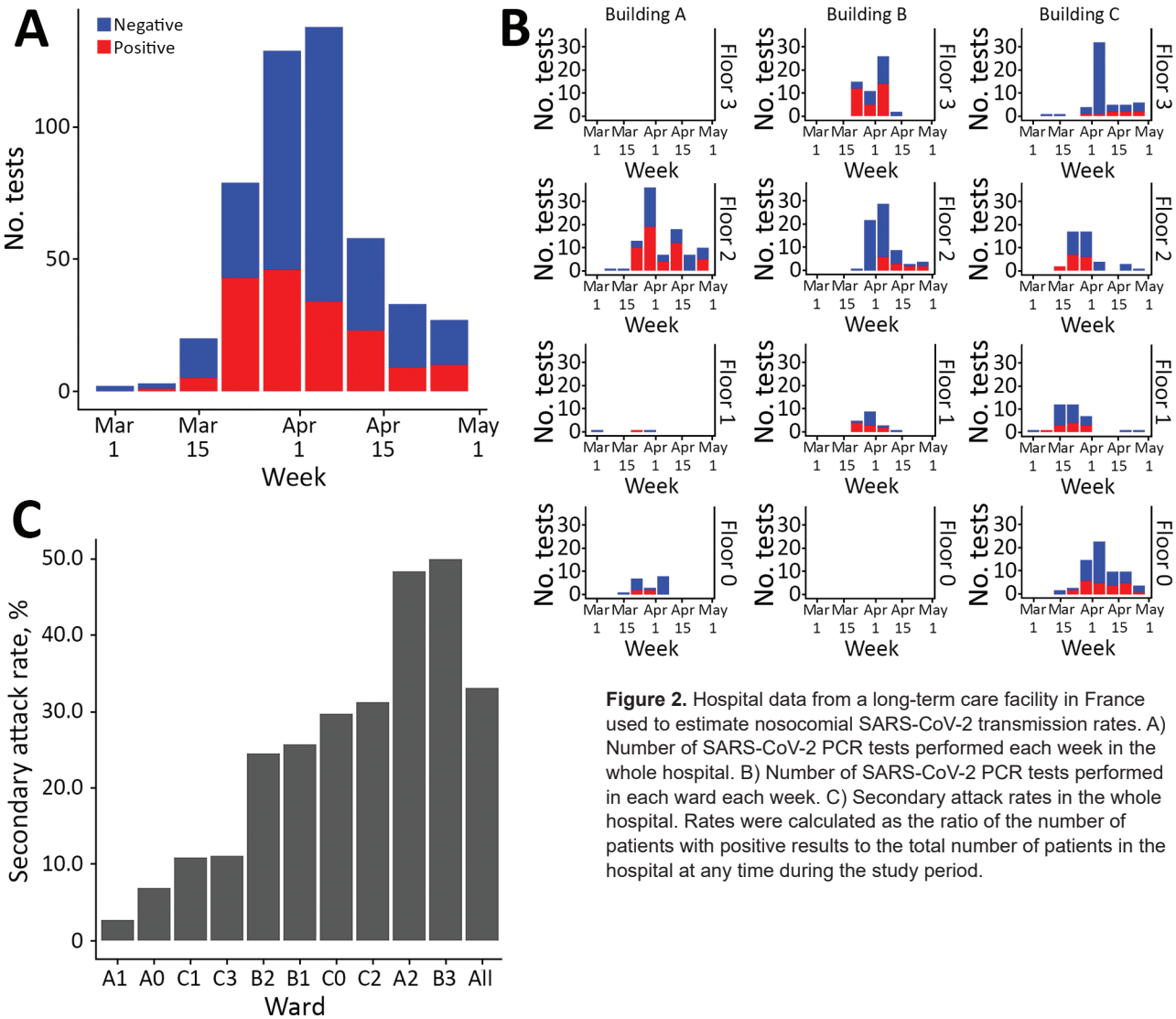


Figure 2. Hospital data from a long-term care facility in France used to estimate nosocomial SARS-CoV-2 transmission rates. A) Number of SARS-CoV-2 PCR tests performed each week in the whole hospital. B) Number of SARS-CoV-2 PCR tests performed in each ward each week. C) Secondary attack rates in the whole hospital. Rates were calculated as the ratio of the number of patients with positive results to the total number of patients in the hospital at any time during the study period.

rate differed substantially between wards (Figure 2, panel C), ranging from 3% to 50%, and the overall secondary attack rate was 33%.

Model Inference Validation Results

The results of the validation of parameter inference on synthetic data suggest that sufficient power was available at the hospital level to recover parameters with relatively good accuracy (Appendix, Appendix Figures 4, 5). However, power was not always sufficient at the ward level, and we restricted our subsequent analysis of wards to only those where the recovered estimates did not deviate excessively in the estimates of β (Appendix Figures 6,7).

Whole-Hospital Analysis

We calculated estimations of transmission rates at the whole hospital level (Table 1; Appendix). In the 2-phase model, using day 12 as $t_{inflect}$ gave the best model fit (Appendix Table 4), which is 6 days after the facility officially introduced an obligatory mask-wearing policy and cancellation of all group activities between patients. This model proved a better fit to the data than the 1-phase model, as measured by the AIC (Table 1). Simulated curves from the observed epidemic produced by the models show that the 2-phase model captured the early peak in cases better than the 1-phase model (Figure 3, panels A, B).

In the 2-phase model where $t_{inflect} = 12$, we observed a notable difference between the transmission rates estimated before and after $t_{inflect}$, which we assume to be attributable to the new contact precautions. The transmission rate fell from 1.3 (95% CI 0.8–2.4) to 0.19 (95% CI 0.10–0.30) infections/patient/day in symptomatic infection, corresponding to a drop in R_0 from 8.7 (95% CI 5.1–16.3) to 1.3 (95% CI 0.7–2.0). This result translates to an 85% (95% CI 66%–94%) decrease of the transmission risk after generalized implementation of contact precautions. Although the value of $t_{inflect}$ had a substantial effect on the absolute values of the transmission rates, the size of the decrease in transmission rate was relatively stable, ranging from 81%–89% (Appendix Table 4). At peak prevalence of infectious patients, we estimated the proportion of undetected infections at 60%, and overall, ≈25% of cases were undetected over the entire study period (Figure 4, panel A).

Ward-Level Analysis

We calculated estimates and corresponding fits for each individual ward for which the 1-phase model could be validated (Table 2; Figure 3, panel B). We reconstructed the undetected parts of the epidemic (Figure 4, panel B). We also conducted ward-level analysis using the 2-phase model but this did not improve the fit (Appendix, Appendix Table 5).

Point estimates for β ranged from 0.42 to 2.13 across the studied wards. We were only able to calculate an upper bound for the transmission rate in 1 ward, C3; the resulting range estimate of 0.42 (0.11–1.30) infections/patient/day corresponds to an R_0 of 2.87 (0.75–8.84). However, we could estimate a lower bound for each ward; the highest value, 0.51 infections/patient/day in ward A2, corresponds to a minimum R_0 of 3.47.

Sensitivity Analysis Results

For most parameters, perturbing had relatively minor effects on the estimated transmission rates for the 2 phases, or on t_{init} (Appendix, Appendix Figure 8). The transmission rate in the second phase, β_2, was the most sensitive, and most markedly sensitive to the duration of symptomatic infection ($1/\delta$).

R_t Results

We calculated R_t estimates by using EpiEstim (Appendix, Appendix Figure 9). The value was initially 10, then fell to <3, before a second peak.

Discussion

We developed a specific framework to analyze SARS-CoV-2 data from a hospital outbreak using a transmission model of patient-to-patient infection. We estimated transmission rates from a LTCF during March–April 2020, across the entire hospital and in individual wards. We assessed 1 or 2

Table 1. Best estimates and ranges for parameters from 2 models applied to hospital data from a long-term care facility in France to estimate nosocomial transmission rates of SARS-CoV-2*

	Model	
Parameter	1-phase	2-phase†
β	0.38 (0.30–0.60)	NA
β_1	NA	1.28 (0.76–2.40)
β_2	NA	0.19 (0.10–0.30)
R_0	2.6 (2.0–4.1)	NA
R_0 before	NA	8.72 (5.14–16.32)
R_0 after	NA	1.33 (0.68–2.04)
R_0 combined	NA	5.72 (3.62–8.70)
Intervention efficacy‡	NA	0.85 (0.66–0.94)
t_{init}	−22 (−39 to −4)	−4 (−25 to −1)
AIC	657.33	628.85

*The value of E_{init} was fixed at day 1 and the value of $t_{inflect}$ at day 12. The R_0 values were calculated by using equations 4 and 5 (Appendix). AIC, Akaike information criterion; NA, not applicable; R_0, basic reproduction number; β, current transmission rate per day; β_1, transmission rate per day before inflection date; β_2, transmission rate per day after inflection date; E_{init}, number of initial infections at date t_{init}; R_0, basic reproduction number; t_{init}, date on which the initial infection occurs.
†R_0 was calculated before and after inflection date in the 2-phase model.
‡The intervention efficacy was calculated as $1 - \beta_2/\beta_1$. Days for t_{init} are relative to the first positive sample on day 1.

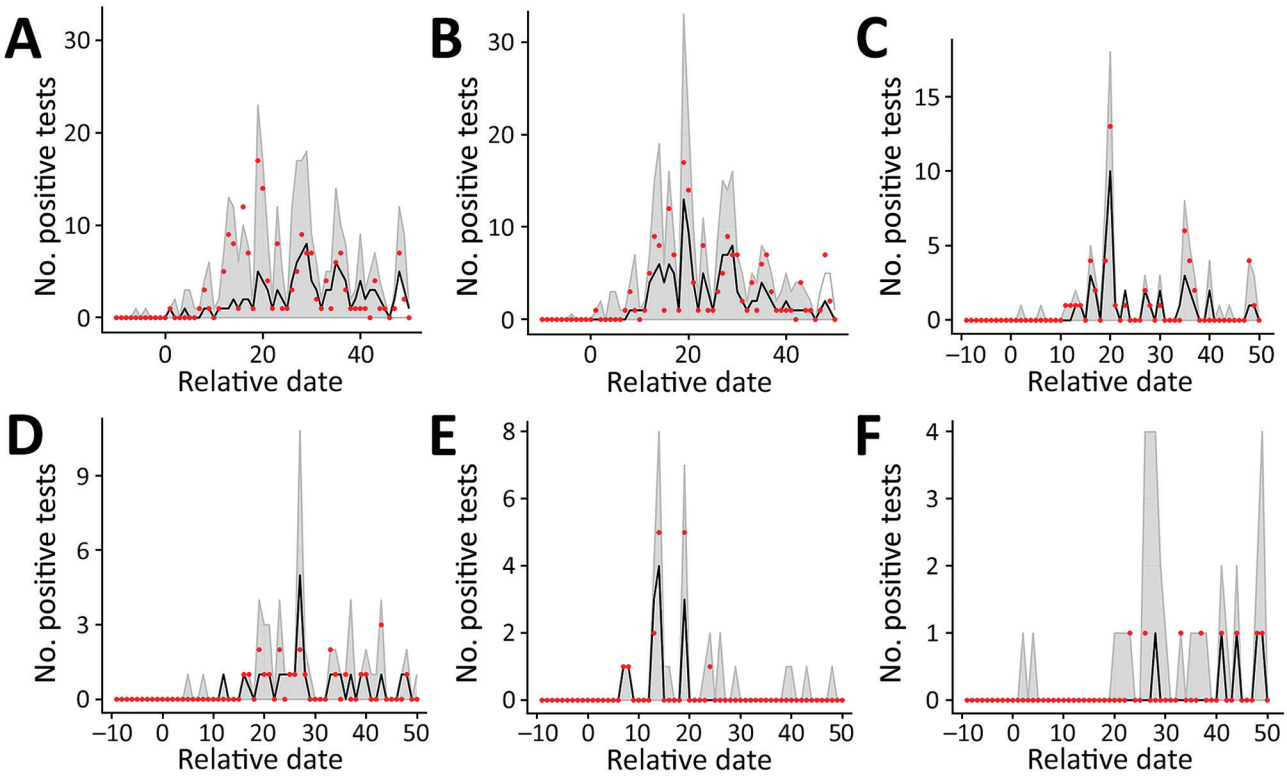


Figure 3. Results of simulated epidemics in a model of nosocomial SARS-CoV-2 transmission using estimated parameters determined on the basis of data from a long-term care facility in France. A) 1-phase model for the whole hospital. B) 2-phase model for the whole hospital. C–F) 1-phase model for individual wards: A2 (C), C0 (D), C2 (E), and C3 (F). Red dots show the observed number of positive tests in the data, black dashed lines indicate the median across that date for all simulations, and gray shading indicates the 95% CI range of the simulated values. Input parameter sets were included if their likelihood fell within the 95% CI relative to the maximum likelihood for 1- and 2-phase models for the whole hospital and individual wards. Estimated parameters are from Tables 1, 2. Extinct epidemics (i.e., those having <3 cumulative cases) were excluded from the distribution.

phases of transmission delimited by a specific change date ($t_{inflect}$) corresponding to implementation of contact precautions, including obligatory mask-wearing for patients and staff, and the cessation of group activities.

We found that the 2-phase model was better supported by the data aggregated across the entire hospital than a model with a single transmission rate, and the 2-phase model better captured the early peak in cases. Model validation suggested sufficient power to estimate transmission rates in 2 phases. The early phase rate (1.3 transmissions/patient/day) corresponded to an early R_0 of 8.7 and the late phase rate (0.19 transmissions/patient/day) corresponded to a late R_0 of 1.3. This change in transmission rate can largely be explained by the initial absence of preventive measures after the policy recommendation on day 6 and its gradual implementation over the next week. Under this assumption, the measures introduced were 85% (95% CI 66%–94%) effective at reducing transmission. The high estimates in the first phase suggest an explosive outbreak or superspreading event, which is consistent with the high secondary attack rate (33%). The estimates in the second phase, after the updated policy, might be more representative of current transmission rates in hospitals, which can provide and encourage the use of personal protective equipment.

Little research is available for the effect of contact precautions against SARS-CoV-2 transmission in healthcare settings. A meta-analysis of the effect of mask use against nosocomial transmission of coronaviruses found 67% protective efficacy of facemasks and 96% efficacy of N95 respirators (*11*), but the 1 study involving SARS-CoV-2 only examined a protective effect for healthcare workers (HCWs), which was unquantifiable because no infections were reported in the masked group (*12*). Several modeling studies have quantified the level of mask wearing that would prevent epidemic spread of SARS-CoV-2 in the community (*13*–*15*; D. Kai et al., unpub. data, http://arxiv.org/abs/2004.13553), but studies of interventions

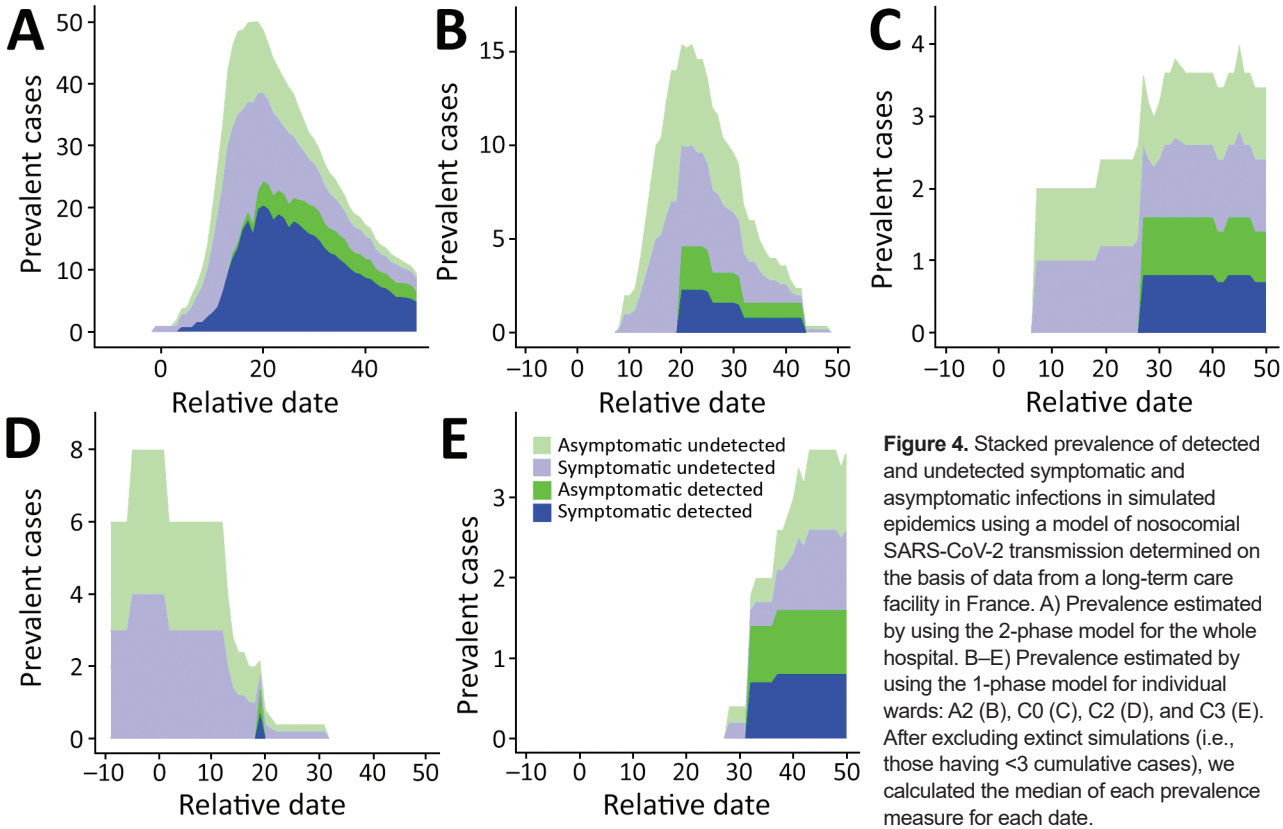


Figure 4. Stacked prevalence of detected and undetected symptomatic and asymptomatic infections in simulated epidemics using a model of nosocomial SARS-CoV-2 transmission determined on the basis of data from a long-term care facility in France. A) Prevalence estimated by using the 2-phase model for the whole hospital. B–E) Prevalence estimated by using the 1-phase model for individual wards: A2 (B), C0 (C), C2 (D), and C3 (E). After excluding extinct simulations (i.e., those having <3 cumulative cases), we calculated the median of each prevalence measure for each date.

for prevention of patient-to-patient transmission in healthcare environments are lacking.

Few other studies have published estimates of R_0 in healthcare settings. By analyzing the initial exponential growth phase of a hospital epidemic, one study computed an expedient estimate of R_0 for patients (1.13) and hospital staff (1.21) (*16*), but that study did not account for asymptomatic infections and did not provide a range for the R_0 estimates (*17*). In another study, the authors estimated an R_0 of 1.021 (95% CI 1.018–1.024) across 12 nursing homes based on a single introduction per floor of each institution and a secondary attack rate of 4.1% among 930 residents (B. Reyné et al., unpub. data, https://doi.org/10.1101/2020.11.27.20239913). The heterogeneity of transmission between different wards was also demonstrated in a previous review and meta-analysis in which the authors calculated an average observed reproduction number of 1.18 across 4 different healthcare settings (*18*), but showed much heterogeneity between settings; 1 was 4.5, and 3 were <0.25. A fourth study analyzed several hospitals in Canada by using incident cases and estimated an R_0 of 2.51, which ranged from 0.56 to 9.17 in individual facilities (*19*). However, the authors of that study did not model asymptomatic infection or account for negative test results or the outcomes of testing at different infectious stages (*19*).

To assess how estimates vary when looking at smaller subpopulations, we separately fit a 1-phase model to data from each ward. Using this method, we

Table 2. Characteristics and parameter estimates in hospital wards in a long-term care facility in France used to estimate nosocomial transmission rates of SARS-CoV-2*

Ward	No. beds	Total no. patients	Day of first positive case	No. cases	β	R_0†	t_{init}
A2	48	62	11	30	1.29 (0.51–NE)	8.76 (3.47–NE)	2 (−14 to 29)
C0	37	74	16	22	0.56 (0.22–NE)	3.79 (1.50–NE)	4 (−39 to 9)
C2	37	48	7	15	2.13 (0.29–NE)	14.46 (1.97–NE)	−8 (−39 to −14)
C3	37	63	24	7	0.42 (0.11–1.30)	2.87 (0.75–8.84)	19 (−9 to 21)

*Estimates and 95% CI for β, R_0, and t_{init} are from the fitting the 1-phase model to data from each ward (E_{init} = 1). In many instances, the upper bound of the 95% CI for β, and in the most likely value of β for some wards, could not be estimated due to a flat likelihood surface, in which case the value is given as NE. NE, not estimated; β, current transmission rate per day; E_{init}, number of initial infections at date t_{init}; R_0, basic reproduction number; t_{init}, date on which the initial infection occurs.

†The R_0 values were calculated using equation 4 (Appendix).

could not always estimate upper bounds of the transmission rates, probably because of strong stochasticity and scarcity of observed cases, an inherent feature of SARS-CoV-2 in which a large proportion of infected persons remain asymptomatic. However, our validation analyses suggested that point estimates for transmission rates across the wards could be consistently estimated. Applied to our dataset, estimated transmission rates ranged from 0.4 to 2.1, corresponding to an R_0 of 2.9–14.5. This heterogeneity might have been driven by differences in the timing of and compliance with preventive measures or by differences in contact patterns between staff and patients.

Calibrating models to real hospital outbreaks and estimating transmission rates provides more realistic transmission models to evaluate scenarios with alternative surveillance or control measures. We estimated the response to introducing barrier interventions at the beginning of the COVID-19 pandemic, when population immunity was minimal. Investigating alternative scenarios involving contemporary levels of population immunity or other viral variants could be easily achieved by updating the model parameters, such as the initial level of immunity or transmission rates. Updating parameters would enable prediction of the probability and size of hospital outbreaks and evaluation of testing strategies to prevent spread. As mentioned, a major challenge in analyzing outbreaks in hospitals or other small, closed environments lies in the consideration of imperfect testing practice, which we addressed through the observation model. First, a substantial proportion of infectious persons were not symptomatic; therefore, they were less likely to be tested, and we accounted for this difference in the model testing policy. Second, PCR test sensitivity is imperfect and depends on the time from infection, which is we also reflected in our evolving test sensitivity for different stages of infection. Finally, testing procedures were not regular and might have been affected by many factors not directly related to the epidemiologic situation, such as the day of the week, the available testing capacity, or changing strategies at the local scale. We addressed irregular testing procedures by using the number of tests per day directly described in the data rather than determining the number of tests performed from the number of infected persons. The model also tracked testing status to include realistic probabilities for testing and retesting of patients.

We compared our results with R_t from the commonly used EpiEstim package, which demonstrated the additional value of our approach. Ignoring negative tests and the complexity of testing policies, this simpler approach captured the high initial R_0 and subsequent fall but also showed a second peak that likely resulted from increased testing rather than an actual increase in transmission rate.

Our analysis has several limitations resulting from simplifying assumptions. First, we did not account for the possibility of imported infections other than the index case or cases; instead, we assumed that the force of infection from other patients would substantially outweigh that from the community. Second, because we had no data on infectious status for HCWs during the study period, we focused on patients and did not explicitly model acquisition by nor transmission from HCWs, although HCWs were implicitly considered potential vectors of patient-to-patient transmission. Rates of transmission from infectious patients to HCWs are relatively low (*20*,*21*), as are transmission rates from HCWs to patients (*22*), although these rates might have been higher in the early stages of the pandemic, considering low levels of hand hygiene (*23*). Ignoring the contribution of HCWs to new infections in the analysis suggests that we might have overestimated the transmission risk from infectious patients, but our estimates can still be interpreted as valid measures of the nosocomial risk to patients. Third, the model relies on parameters taken from the literature, which may be inaccurate. However, we conducted a sensitivity analysis to measure the sensitivity of transmission rates to appropriate variation in these parameters, and our main results remained unaffected. Finally, we note that the decision to analyze data from this hospital is partly due to the size of the outbreak, implying a selection bias toward a higher transmission rate than would be typical across all hospitals. However, >44,000 nosocomial infections were reported in France by February 14, 2021 (*24*), most of which consisted of clusters of cases; thus, our results can be interpreted as plausible for a hospital at risk for an outbreak. In addition, the model framework we propose is suitable for estimating transmission rates in any healthcare environment, and we provide some guidance for adaptation (Appendix).

In conclusion, the novel dynamic modeling framework we propose realistically simulates evolving testing policies and could easily be used on similar nosocomial COVID-19 datasets. The model also could be adapted for specific epidemiologic features, such as patient isolation. Overall, our results underline both the substantial potential effect of protective interventions introduced in healthcare settings and the considerable heterogeneity in transmission rates between hospital wards.

Additional members of EMEA-MESuRS Working Group on the Nosocomial Modelling of SARS-CoV-2: Sophie Chervet, Audrey Duval, Kévin Jean, Sofía Jijón, Ajmal Oodally, David R.M. Smith, and Cynthia Tamandjou.

Acknowledgments

We thank Sandrine Jacques for data collation, Niels Hendrickx for discussion of the results, and Matthieu Domenech de Cellès for advice on the statistical inference.

This work was supported directly by internal resources from the French National Institute for Health and Medical Research, the Institut Pasteur, the Conservatoire National des Arts et Métiers, and the University of Versailles-Saint-Quentin-en-Yvelines/University of Paris-Saclay. This study received funding through the MODCOV project from the Fondation de France (grant no. 106059) as part of the alliance framework "Tous unis contre le virus," the Université Paris-Saclay (no. AAP Covid-19 2020) and the French government through its National Research Agency (project no. SPHINX-17-CE36-0008-01).

G.S. constructed the model, conducted the analysis, and produced the draft and graphics. J.R.Z. produced the data and provided the medical perspective to inform model assumptions. S.C. made substantial contributions to the interpretation of results. L.T. and L.O. conceived the study, had regular input on analysis and interpretation, and contributed to the writing. All authors read, provided comments on, and approved the final manuscript.

About the Author

Dr. Shirreff is a postdoctoral researcher in nosocomial infectious disease at Conservatoire Nationale des Arts et Métiers, University of Versailles, and Institut Pasteur, Paris, France. His research focuses on understanding infectious disease transmission using mathematical models and modern tools for data capture.

References

1. Guan WJ, Liang WH, Zhao Y, Liang HR, Chen ZS, Li YM, et al.; China Medical Treatment Expert Group for COVID-19. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. Eur Respir J. 2020;55:2000547. https://doi.org/10.1183/13993003.00547-2020
2. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med. 2020;180:934–43. https://doi.org/10.1001/jamainternmed.2020.0994
3. Abbas M, Robalo Nunes T, Martischang R, Zingg W, Iten A, Pittet D, et al. Nosocomial transmission and outbreaks of coronavirus disease 2019: the need to protect both patients and healthcare workers. Antimicrob Resist Infect Control. 2021;10:7. https://doi.org/10.1186/s13756-020-00875-7
4. Hall VJ, Foulkes S, Saei A, Andrews N, Oguti B, Charlett A, et al.; SIREN Study Group. COVID-19 vaccine coverage in health-care workers in England and effectiveness of BNT162b2 mRNA vaccine against infection (SIREN): a prospective, multicentre, cohort study. Lancet. 2021;397:1725–35. https://doi.org/10.1016/S0140-6736(21)00790-X
5. Cheng VC–C, Fung KS-C, Siu GK-H, Wong S-C, Cheng LS-K, Wong M-S, et al. Nosocomial outbreak of COVID-19 by possible airborne transmission leading to a superspreading event. Clin Infect Dis. 2021;73:e1356–64. https://doi.org/10.1093/cid/ciab313
6. Du Q, Zhang D, Hu W, Li X, Xia Q, Wen T, et al. Nosocomial infection of COVID-19: A new challenge for healthcare professionals. Int J Mol Med. 2021;47:31. https://doi.org/10.3892/ijmm.2021.4864
7. Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? Lancet. 2020;395:931–4. https://doi.org/10.1016/S0140-6736(20)30567-5
8. Smith DRM, Duval A, Pouwels KB, Guillemot D, Fernandes J, Huynh B-T, et al.; AP-HP/Universities/Inserm COVID-19 research collaboration. Optimizing COVID-19 surveillance in long-term care facilities: a modelling study. BMC Med. 2020;18:386. https://doi.org/10.1186/s12916-020-01866-6
9. Temime L, Gustin M-P, Duval A, Buetti N, Crépey P, Guillemot D, et al. A conceptual discussion about the basic reproductive number of severe acute respiratory syndrome coronavirus 2 in healthcare settings. Clin Infect Dis. 2020;72:141–3. https://doi.org/10.1093/cid/ciaa682
10. King AA, Nguyen D, Ionides EL. Statistical inference for partially observed Markov processes via the R package pomp. J Stat Softw. 2016;69:1–43. https://doi.org/10.18637/jss.v069.i12
11. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, et al.; COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet. 2020;395:1973–87. https://doi.org/10.1016/S0140-6736(20)31142-9
12. Wang X, Pan Z, Cheng Z. Association between 2019-nCoV transmission and N95 respirator use. J Hosp Infect. 2020;105:104–5. https://doi.org/10.1016/j.jhin.2020.02.021
13. Tian L, Li X, Qi F, Tang Q-Y, Tang V, Liu J, et al. Harnessing peak transmission around symptom onset for non-pharmaceutical intervention and containment of the COVID-19 pandemic. Nat Commun. 2021;12:1147. https://doi.org/10.1038/s41467-021-21385-z
14. Stutt ROJH, Retkute R, Bradley M, Gilligan CA, Colvin J. A modelling framework to assess the likely effectiveness of facemasks in combination with 'lock-down' in managing the COVID-19 pandemic. Proc R Soc Math Phys Eng Sci. 2020;476:20200376. https://doi.org/10.1098/rspa.2020.0376
15. Fisman DN, Greer AL, Tuite AR. Bidirectional impact of imperfect mask use on reproduction number of COVID-19: a next generation matrix approach. Infect Dis Model. 2020;5:405–8. https://doi.org/10.1016/j.idm.2020.06.004
16. Tang JW, Young S, May S, Bird P, Bron J, Mohamedanif T, et al. Comparing hospitalised, community and staff COVID-19 infection rates during the early phase of the

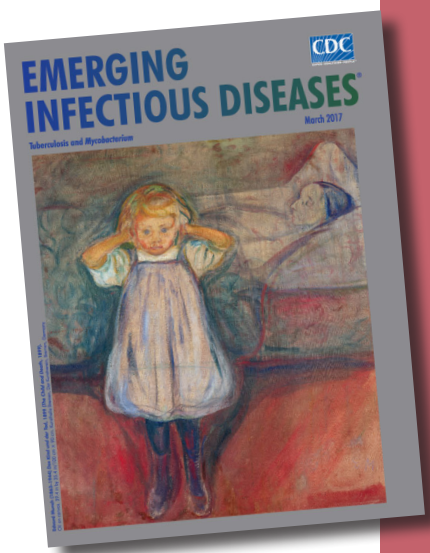
evolving COVID-19 epidemic. J Infect. 2020;81:647–79. https://doi.org/10.1016/j.jinf.2020.05.029
17. Accorsi E, Qiu X, Rumpler E, Kennedy-Shaffer L, Kahn R, Joshi K, et al. How to detect and reduce potential sources of biases in studies of SARS-CoV-2 and COVID-19. Eur J Epidemiol. 2021;36:179–96. https://doi.org/10.1007/s10654-021-00727-7
18. Thompson HA, Mousa A, Dighe A, Fu H, Arnedo-Pena A, Barrett P, et al. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) setting-specific transmission rates: a systematic review and meta-analysis. Clin Infect Dis. 2021;73:e754–64. https://doi.org/10.1093/cid/ciab100
19. Stockdale JE, Anderson SC, Edwards AM, Iyaniwura SA, Mulberry N, Otterstatter MC, et al. Quantifying transmissibility of SARS-CoV-2 and impact of intervention within long-term healthcare facilities. R Soc Open Sci. 2022;9:211710. https://doi.org/10.1098/rsos.211710
20. Basso T, Nordbø SA, Sundqvist E, Martinsen TC, Witsø E, Wik TS. Transmission of infection from non-isolated patients with COVID-19 to healthcare workers. J Hosp Infect. 2020;106:639–42. https://doi.org/10.1016/j.jhin.2020.08.015
21. Tong X, Ning M, Huang R, Jia B, Yan X, Xiong Y, et al. Surveillance of SARS-CoV-2 infection among frontline health care workers in Wuhan during COVID-19 outbreak. Immun Inflamm Dis. 2020;8:840–3. https://doi.org/10.1002/iid3.340
22. Baker MA, Fiumara K, Rhee C, Williams SA, Tucker R, Wickner P, et al. Low risk of coronavirus disease (COVID-19) among patients exposed to infected healthcare workers. Clin Infect Dis. 2021;73:e1878–80. https://doi.org/10.1093/cid/ciaa1269
23. Huang F, Armando M, Dufau S, Florea O, Brouqui P, Boudjema S. COVID-19 outbreak and healthcare worker behavioural change toward hand hygiene practices. J Hosp Infect. 2021;111:27–34. https://doi.org/10.1016/j.jhin.2021.03.004
24. Public Health France. COVID-19 epidemiologic update for February 18, 2021 [in French] [cited 2021 Jul 12]. https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-18-fevrier-2021

Address for correspondence: George Shirreff, Institut Pasteur, National des Arts et Métiers, 25–28 rue Dr. Roux, Paris 75015, France; email: george.shirreff@pasteur.fr

etymologia revisited

Mycobacterium chimaera

[mi'ko-bak-tēr'e-əm ki-mēr'ə]



Originally published in March 2017

Formerly an unnamed *Mycobacterium* sequevar within the *M. avium–M. intracellulare–M. scrofulaceum* group (MAIS), *M. chimaera* is an emerging opportunistic pathogen that can cause infections of heart valve prostheses, vascular grafts, and disseminated infections after open-heart surgery. Heater–cooler units used to regulate blood temperature during cardiopulmonary bypass have been implicated, although most isolates are respiratory. In 2004, Tortoli et al. proposed the name M. chimaera for strains that a reverse hybridization–based line probe assay suggested belonged to MAIS but were different from *M. avium*, *M. intracellulare*, or *M. scrofulaceum*. The new species name comes from the chimera, a mythological being made up of parts of 3 different animals.

Sources:

1. Schreiber PW, Kuster SP, Hasse B, Bayard C, Rüegg C, Kohler P, et al. Reemergence of *Mycobacterium chimaera* in heater–cooler units despite
intensified cleaning and disinfection protocol. Emerg Infect Dis. 2016;22:1830–3.
2. Struelens MJ, Plachouras D. *Mycobacterium chimaera* infections associated with heater-cooler units (HCU): closing another loophole in patient safety. Euro Surveill. 2016;21:1–3.
3. Tortoli E, Rindi L, Garcia MJ, Chiaradonna P, Dei R, Garzelli C, et al. Proposal to elevate the genetic variant MAC-A, included in the *Mycobacterium avium* complex, to species rank as *Mycobacterium chimaera* sp. nov. Int J Syst Evol Microbiol. 2004;54:1277–85.

https://wwwnc.cdc.gov/eid/article/23/3/et-2303_article

Article DOI: https://doi.org/10.3201/eid2807.212339

Measuring Basic Reproduction Number to Assess Effects of Nonpharmaceutical Interventions on Nosocomial SARS-CoV-2 Transmission

Appendix

Additional Methods

Data Input and Parameter Estimation

SARS-CoV-2 RNA extraction was performed on a NucliSENS easyMAG (bioMérieux, https://www.biomerieux.com) device, strictly following the manufacturer's recommendations. Reverse transcription PCR (RT-PCR) was performed on an ABI QuantStudio 7 (Thermo Fisher Scientific, https://www.thermofisher.com) device, using the commercial RealStar SARS-CoV-2 RT-PCR Kit 1.0 (Altona Diagnostics, https://www.altona-diagnostics.com) test. Briefly, 10 µL of RNA is added to the 20 µL RT-PCR mix. Two targets are detected, one specific for betacoronavirus, and one specific for the SARS-COV-2 strains. Internal control was added in the lysis buffer to validate both extraction and amplification steps.

All patients were included in the study if they were in the hospital during the study period from day −10 to day 50. Daily data on tests, admissions, and discharges were input directly into the model from the hospital data (Appendix Figure 2). In most cases, patients were recorded as being in a particular ward, and these were used to create ward-specific datasets. In case of a gap in the patient record between recorded stays in different wards the transfer was assumed to occur at the midpoint of the gap. All data analyzed for this study at the whole hospital and ward levels, along with the R scripts used to conduct the analysis, are available at github.com/georgeshirreff/Hospital_R0_C19.

Parameters Estimated Directly from Longitudinal Hospital Data

The parameter $1/\omega$, the duration of the R_p stage during which persons who have recovered from infectious disease but continue to frequently test positive through PCR, was selected by calculating the likelihood of each duration according to the results of repeat tests. The data used were repeat tests taken after a patient had an earlier positive test (Appendix Figure 3, panel A). We assumed that the R_p stage began 7 days after the first positive test ($1/\delta$), the probability of testing positive during this stage was 30% (Z_{Rp}) and the probability of testing positive in the recovery (R) stage afterwards was 1% ($1 - v$) (Table 1). The likelihood reached a plateau where $1/\omega = 23$, so a value of 25 days was subsequently used as consistent with this result (Appendix Figure 3, panel B).

The parameter ϕ, the relative rate of retesting, has been crudely estimated from the data by counting the number of repeat tests, i.e., on persons who have been tested again without having symptoms develop (n = 211) divided by the sum of the number of first tests (n = 314) plus the number who were retested upon developing symptoms (n = 34), giving a retesting rate of $\phi = 60\%$. A bootstrap analysis was conducted on the dataset to estimate of 50% CI and 70% CI.

Hospital Prevention and Contact Policy

At the beginning of the study period, hospital policy did not specify the use of any masks during contact between healthcare workers (HCWs) and patients. Patients normally participated in group activities including leisure activities, meals, and joint physiotherapy sessions. Gloves and gowns were required by staff during any contact with bodily fluids. After March 17, 2020, the policy changed to require the wearing of surgical masks during proximity contact between HCWs and patients, and between staff members, as well as cancellation of all visits and group activities.

PCR testing, where capacity was available, was conducted on any patients with typical COVID-19 symptoms, namely persistent cough, fever, anosmia, or diarrhea. PCR was also conducted on any patients after suspected infectious contact with other patients, as well as on any patients being admitted to the hospital or moved between wards. Where capacity was lacking, patients displaying new symptoms were tested with priority. This testing procedure continued throughout the study period.

Mathematical Model

Observation Model and Differential Equations

We show the observational model that defines how testing and retesting is conducted for each stage of infection (Appendix Figure 1). We also show the structure of the transmission model (Figure 1). The observation model can be represented using differential equations that describe the change in state of each compartment in each time step (equation 1), in which $\lambda(t)$ is the force of infection (equation 2). The bold terms in equation 1 above refer to flows that are determined in part by the available data from a given day.

Equation 1

$$\frac{dS}{dt} = \mathbf{Admission}(\boldsymbol{t}) - \mathbf{Initiation}(\boldsymbol{S},\boldsymbol{t}) - \lambda(t)S - \mathbf{Discharge}(\boldsymbol{S},\boldsymbol{t}) - \mathbf{Test}(\boldsymbol{t},\boldsymbol{S})$$

$$\frac{dS_T}{dt} = -\mathbf{Initiation}(\boldsymbol{t},\boldsymbol{S_T}) - \lambda(t)S_T - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{S_T}) + \mathbf{Test}(\boldsymbol{t},\boldsymbol{S})$$

$$\frac{dE}{dt} = \mathbf{Initiation}(\boldsymbol{t},\boldsymbol{S}) + \lambda(t)S - \alpha E - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{E}) - \mathbf{Test}(\boldsymbol{t},\boldsymbol{E})$$

$$\frac{dE_T}{dt} = \mathbf{Initiation}(\boldsymbol{t},\boldsymbol{S_T}) + \lambda(t)S_T - \alpha E_T - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{E_T}) + \mathbf{Test}(\boldsymbol{t},\boldsymbol{E})$$

$$\frac{dE_a}{dt} = \alpha E(1-\psi) - E_a\gamma\kappa_2 - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{E_a}) - \mathbf{Test}(\boldsymbol{t},\boldsymbol{E_a})$$

$$\frac{dE_{aT}}{dt} = \alpha E_T(1-\psi) - E_{aT}\gamma\kappa_2 - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{E_{aT}}) + \mathbf{Test}(\boldsymbol{t},\boldsymbol{E_a})$$

$$\frac{dE_s}{dt} = \alpha E\psi - E_s\gamma - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{E_s}) - \mathbf{Test}(\boldsymbol{t},\boldsymbol{E_s})$$

$$\frac{dE_{sT}}{dt} = \alpha E_T\psi - E_{sT}\gamma - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{E_{sT}}) + \mathbf{Test}(\boldsymbol{t},\boldsymbol{E_s})$$

$$\frac{dI_a}{dt} = E_a\gamma\kappa_2 - I_a\delta\kappa_3 - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{I_a}) - \mathbf{Test}(\boldsymbol{t},\boldsymbol{I_a})$$

$$\frac{dI_{aT}}{dt} = E_{aT}\gamma\kappa_2 - I_{aT}\delta\kappa_3 - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{I_{aT}}) + \mathbf{Test}(\boldsymbol{t},\boldsymbol{I_a})$$

$$\frac{dI_s}{dt} = E_{sT}\gamma + E_s\gamma - I_s\delta - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{I_s}) - \mathbf{Test}(\boldsymbol{t},\boldsymbol{I_s})$$

$$\frac{dI_{sT}}{dt} = -I_{sT}\delta - \mathbf{Discharge}(\boldsymbol{t},\boldsymbol{I_{sT}}) + \mathbf{Test}(\boldsymbol{t},\boldsymbol{I_s})$$

$$\frac{dR_p}{dt} = I_s\delta + I_a\delta\kappa_3 - \omega R_p - \mathbf{Discharge}(\boldsymbol{t}, \boldsymbol{R_p}) - \mathbf{Test}(\boldsymbol{t}, \boldsymbol{R_p})$$

$$\frac{dR_{pT}}{dt} = I_{sT}\delta + I_{aT}\delta\kappa_3 - \omega R_{pT} - \mathbf{Discharge}(\boldsymbol{t}, \boldsymbol{R_{pT}}) + \mathbf{Test}(\boldsymbol{t}, \boldsymbol{R_p})$$

$$\frac{dR}{dt} = \omega R_p - \mathbf{Discharge}(\boldsymbol{t}, \boldsymbol{R}) - \mathbf{Test}(\boldsymbol{t}, \boldsymbol{R})$$

$$\frac{dR_T}{dt} = \omega R_{pT} - \mathbf{Discharge}(\boldsymbol{t}, \boldsymbol{R_T}) + \mathbf{Test}(\boldsymbol{t}, \boldsymbol{R})$$

Force of Infection and Basic Reproduction Number Calculation

The force of infection acting on susceptible patients (equation 2) is defined by the infectious populations, the transmission rate, β, and the total population size N (equation 3).

Equation 2

$$\lambda(t) = \frac{\beta\left(I_s + I_{sT} + \varepsilon(E_s + E_{sT}) + \kappa_1(I_a + I_{aT}) + \varepsilon\kappa_1(E_a + E_{aT})\right)}{N}$$

Equation 3

$$N = S + S_T + E + E_T + E_a + E_{aT} + E_s + E_{sT} + I_a + I_{aT} + I_s + I_{sT} + R_p + R_{pT} + R + R_T$$

The basic reproduction number (R_0) value of can be calculated directly from the parameters according to equation 4, which takes into account the full-blown symptomatic transmission rate (β), the probability of entering the symptomatic (ψ) or asymptomatic pathway ($1 - \psi$), the relative transmission rate of each stage of infection (ε, κ_1), and the rate of leaving each stage (γ and δ). In the 2-phase model, 2 R_0 values, $R_{0\ \text{before}}$ and $R_{0\ \text{after}}$, were calculated independently for each phase based on the different transmission rates, β_1 and β_2, using equation 4; a combined R_0 value was calculated as an average weighted by the duration of each phase using equation 5, with the final date being the end of the study period, day 50.

Equation 4

$$R_0 = \beta\left(\psi\left(\frac{\varepsilon}{\gamma} + \frac{1}{\delta}\right) + (1 - \psi)\kappa_1\left(\frac{\varepsilon}{\gamma} + \frac{1}{\delta}\right)\right)$$

Equation 5

$$R_{0\text{ combined}} = \frac{R_{0\text{ before}}\max\left(t_{inflect} - t_{init}, 0\right) + R_{0\text{ after}}\left(\text{final_date} - \max\left(t_{inflect}, t_{init}\right)\right)}{\text{final_date} - \min\left(t_{inflect}, t_{init}\right)}$$

Deterministic Processes

The bold terms in equation 1 above refer to flows that are determined in part by the available data from a given day (d), including number of admissions $A(d)$, discharges $D(d)$, and of tests $T(d)$, as shown in weekly aggregate (Appendix Figure 2) or by specific parameter values, such as SARS-CoV-2 introduction date (t_{init}) and size (E_{init}), which together determine the number of daily new infectees $C(d)$, in the case of epidemic initiation. This calculation ensures that admissions, discharges, and number of tests each occur with the same frequency in the model as they do in the data, and that the timing and size of the start of the epidemic is determined by parameter values.

Admission

Admissions only occur into the susceptible untested (S) group, and so on a given day (d) the number of admissions into this group is exactly determined by the number of admissions on that day, $A(d)$:

Equation 6

$$\int_{t=d}^{t=d+1} \text{Admission}(t)dt = A(d)$$

Initiation

The expectation of the number of initial infectees from each susceptible group (S, S_T) is determined by equations 7 and 8, where $C(d)$ is equal to E_{init} on day t_{init}, and otherwise 0, with the total in both groups being equal to $C(d)$ (equation 10).

Equation 7

$$E\left(\text{Initiation}(d, S)\right) = C(d)\frac{S}{S + S_T}$$

Equation 8

$$E\big(\text{Initiation}(d, S_T)\big) = C(d)\frac{S_T}{S + S_T}$$

For simplification, we denote the number of infections initiated on day *d* in compartment *X* by equation 9, which is equal to the integral of all initiations in all (both) compartments across the whole of the day. The total initiations across all compartments is equal to *C*(*d*) (equation 10).

Equation 9

$$\text{Initiation}(d, X) = \int_{t=d}^{t=d+1} \text{Initiation}(t, X)dt$$

Equation 10

$$\sum\nolimits_{X \in S, S_T} \text{Initiation}(d, X) = C(d)$$

Discharge

The expectation of the number of discharges for a compartment *X* on each day (equation 11) has different values for symptomatically infected patients (I_s and I_{sT}), who have a discharge rate modified by the parameter μ. The denominator of the expectation, *W*, is the total dischargeable population, adjusting for these differences in rate (equation 12).

Equation 11

$$E\big(\text{Discharge}(d, X)\big) = \begin{cases} \dfrac{\mu X}{W} D(d) & \text{for } X \in I_S, I_{ST} \\ \dfrac{X}{W} D(d) & \text{for } X \notin I_S, I_{ST} \end{cases}$$

Equation 12

$$W = S + S_T + E + E_T + E_a + E_{aT} + E_s + E_{sT} + I_a + I_{aT} + \mu(I_s + I_{sT}) + R_p + R_{pT} + R + R_T$$

As with initiations above, we denote the number of patients discharged on day d in compartment X by equation 13, which is equal to the integral of all discharges in that compartment across the whole of the day. The total discharges on day d across all compartments U is equal to $D(d)$ (equation 14).

Equation 13

$$\text{Discharge}(d, X) = \int_{t=d}^{t=d+1} \text{Discharge}(t, X)dt$$

Equation 14

$$\sum_{X \in U} \text{Discharge}(d, X) = D(d)$$

Testing Model

The expectation of the number of tests to occur in a compartment X on day d is given as follows, with $T(d)$ referring to the number of tests occurring on day d in the data (equation 15).

The untested symptomatic patients are tested as a priority and so the number of tests they receive is determined as the minimum of their size (I_s) and the number of tests available ($T(d)$), so the expected number of these (in equation 15 for I_s) is the same as their total (equation 18).

The remaining tests are distributed randomly throughout the remaining compartments. Expectations are shown in equation 15, derived from the number of tests left over after first symptomatic tests, the size of the compartment, the parameter φ for the already tested compartments, and a denominator M (equation 16) which represents the total testable population (including the adjustment for retesting).

Equation 15

$$E\big(\text{Test}(d, X)\big) = \begin{cases} \min(T(d), I_S) & \text{for } X \in I_S \\ \min(T(d) - I_S, 0)\dfrac{X}{M} & \text{for } X \in S, E, E_a, E_s, I_a, R_p, R \\ \min(T(d) - I_S, 0)\dfrac{\varphi X}{M} & \text{for } X \in S_T, E_T, E_{aT}, E_{sT}, I_{aT}, I_{sT}, R_{pT}, R_T \end{cases}$$

Equation 16

$$M = \big(S + E + E_a + E_s + I_a + R_p + R\big) + \varphi(S_T + E_T + E_{aT} + E_{sT} + I_{aT} + I_{sT} + R_{pT} + R_T)$$

As with initiations above, we denote the number of patients tested on day d in compartment X by equation 17, which is equal to the integral of all tests in that compartment across the whole of the day. The total tests across all compartments other than I_s are equal to the number of remaining tests (equation 19).

Equation 17

$$\mathrm{Test}(d, X) = \int_{t=d}^{t=d+1} \mathrm{Test}(t, X)dt$$

Equation 18

$$\mathrm{Test}(d, I_S) = \min(T(d), I_S)$$

Equation 19

$$\sum\nolimits_{X \notin I_S} \mathrm{Test}(d, X) = \min(T(d) - I_S, 0)$$

When simulating the observed tests on a given day using the function *rmeasure* in *pomp* (https://CRAN.R-project.org/package=pomp) for R (R Foundation for Statistical Computing, https://www.r-project.org), the number of positive and negative tests for a given day is drawn from the number of tests occurring in each compartment (equation 15) according to the probability that a sample taken from a patient in that compartment would test positive, which is governed by specificity, v, for virus-free compartments, and sensitivity z_X for each compartment X (equations 20 and 21). These expected distributions are also used for evaluating the likelihood of observed data using *dmeasure* in *pomp*.

Equation 20

$$\begin{aligned} E\big(\mathrm{Positives}(d)\big) \\ &= \big(\mathrm{Test}(d, S) + \mathrm{Test}(d, S_T)\big)(1 - v) + \big(\mathrm{Test}(d, E) + \mathrm{Test}(d, E_T)\big)z_E \\ &+ \big(\mathrm{Test}(d, E_a) + \mathrm{Test}(d, E_{aT})\big)z_{Ea} + \big(\mathrm{Test}(d, E_s) + \mathrm{Test}(d, E_{sT})\big)z_{Es} \\ &+ \big(\mathrm{Test}(d, I_a) + \mathrm{Test}(d, I_{aT})\big)z_{Ia} + \big(\mathrm{Test}(d, I_s) + \mathrm{Test}(d, I_{sT})\big)z_{Is} \\ &+ \Big(\mathrm{Test}\big(d, R_p\big) + \mathrm{Test}\big(d, R_{pT}\big)\Big)\, z_{Rp} + \big(\mathrm{Test}(d, R) + \mathrm{Test}(d, R_T)\big)(1 - v) \end{aligned}$$

Equation 21

$$
\begin{aligned}
E\big(\text{Negatives}(d)\big) \\
&= \big(\text{Test}(d,S) + \text{Test}(d,S_T)\big)v + \big(\text{Test}(d,E) + \text{Test}(d,E_T)\big)(1 - z_E) \\
&+ \big(\text{Test}(d,E_a) + \text{Test}(d,E_{aT})\big)(1 - z_{Ea}) \\
&+ \big(\text{Test}(d,E_s) + \text{Test}(d,E_{sT})\big)(1 - z_{Es}) + \big(\text{Test}(d,I_a) + \text{Test}(d,I_{aT})\big)(1 - z_{Ia}) \\
&+ \big(\text{Test}(d,I_s) + \text{Test}(d,I_{sT})\big)(1 - z_{Is}) \\
&+ \Big(\text{Test}\big(d,R_p\big) + \text{Test}\big(d,R_{pT}\big)\Big)\big(1 - z_{Rp}\big) + \big(\text{Test}(d,R) + \text{Test}(d,R_T)\big)v
\end{aligned}
$$

Implementation of the Stochastic Model

The hybrid stochastic model was implemented by using the *rprocess* function in *pomp* using the Gillespie algorithm because the number of events on a given day is relatively small (around 15–30) given the population of <400 patients. The algorithm calculates a rate for each possible type of event to occur, determines the time and type of the next event accordingly, and then recalculates the rates after each event. To ensure that deterministic transitions (i.e., events determined by model input) occur with certainty within the framework of the Gillespie algorithm, the rate of such events was set to an arbitrarily large number, $L = 10^6$, if further instances of that event are still to occur on day d. Once all required instances have occurred, the rate is set to zero.

Statistical Inference

Likelihood Calculation

The infection model was linked to observed data (number of positive and negative tests per day) using the framework of a partially observed Markov process (POMP) in which the modified susceptible-exposed-infected-recovered (SEIR) model governs the underlying infection dynamics, and each day the observation process provides a likelihood of observing the data given the internal state.

The likelihood for the observation of several negative and positive tests on a particular day d is given in equation 22. The set of parameters is represented by θ, and normally only β and t_{init} would vary within the 1-phase model, but in the 2-phase model β_1, β_2, and t_{init} are estimated. The expected numbers of positives and negatives according to the model and θ are given by

equations 20 and 21. The total likelihood is the product of the likelihood values across all time points d.

Equation 22

$$\text{Likelihood}_d = p(\text{testing negative}(d)|\text{model}, \theta)^{\text{Negatives}(d)} p(\text{testing positive}(d)|\text{model}, \theta)^{\text{Positives}(d)}$$

$$= \left(\frac{E\big(\text{negatives}(d)\big)}{\Big(E\big(\text{negatives}(d)\big) + E\big(\text{positives}(d)\big)\Big)}\right)^{\text{Negatives}(d)} \left(\frac{E\big(\text{positives}(d)\big)}{\Big(E\big(\text{negatives}(d)\big) + E\big(\text{positives}(d)\big)\Big)}\right)^{\text{Positives}(d)}$$

Parameter Estimation through Stochastic Model Fitting

The inference of parameters (transmission rates β, or β_1 and β_2, and for t_{init} or E_{init}) was conducted according to the methodology proposed by King et al. (*6*). The first step was an initial search for the values of all parameters to be estimated, using 500 iterations of 500 particles, and a cooling fraction of 50% every 50 steps. This was repeated 10 times for each of 1,000 different starting points of the parameters to be estimated.

Subsequently a likelihood profile was estimated for β (in the 1-phase model) or β_1 (in the 2-phase model) by repeating the analysis above but using starting points for the parameter to be profiled across its relevant range (e.g., 0.1–10 in steps of 0.1), and holding this parameter constant while estimating the other(s) using the same inference methodology.

In all iterative filtering analyses, transmission rates were allowed to vary during an iteration, while t_{init} (or E_{init}) was only varied at the beginning of an iteration as an initial value parameter. During inference of the 2-phase model, each β-value was only allowed to vary during the phase in which it took direct effect, meaning β_1 would only vary before $t_{inflect}$ and β_2 would only vary afterwards.

Both the initial search and likelihood profiling were conducted using 500 iterations of 500 particles in each analysis, each of which was repeated 10 times for each of 1,000 different starting points of the parameters to be estimated.

The likelihood of the final parameter combination in each analysis (whether initial search or profiling) was then estimated by performing 10 repetitions of particle filtering with 100,000 particles, from which a linear average of the 10 likelihoods was taken.

Confidence Intervals for Estimated Parameters

Confidence intervals for estimated parameters were established by identifying sets of parameters values with a likelihood above a threshold relative to the highest likelihood for each analysis. The threshold was the maximum value of the likelihood minus half of the 95% quantile of the χ-square distribution with degrees of freedom corresponding to the number of parameters to be estimated, typically 2 for the 1-phase model (β and t_{init}) and 3 for the 2-phase model (β_1, β_2, and t_{init}).

Model Inference Validation

We conducted validation by using synthetic data to test the effectiveness of the model and statistical inference to recover known values of the parameters. Several datasets, representing numbers of "observed" positive cases, were generated from model simulations to represent the transmission within the whole hospital and the individual wards using known parameter values. The parameter values estimated from these datasets were then compared with the known values used to create them. Values of the parameters of interest were varied simultaneously, while all other parameters were fixed as depicted (Appendix Table 1).

Datasets were generated using *rmeasure* in *pomp* for each set of known parameter values over a 3-month observation period (day −39 to day 50). Multiple (n = 10) datasets were generated for each ward, or the whole hospital, and each set of known parameter values. In each case, the real data on numbers of daily tests, admissions, discharges, and number of patients were used. Iterative filtering to estimate the relevant parameters was then conducted on each dataset. The 1-phase model was validated by using data at the scale of both the whole-hospital and individual ward levels, while the 2-phase model was validated only at the whole-hospital level.

Unlike in the analysis on the true data, the value of β or β_1 was only estimated by iterative filtering directly, without the systematic likelihood profile for a range of values. For each parameter search, 500 iterations of 500 particles were used. For each known set of parameter

values, the median value of the estimated parameters was identified and compared with their known values.

To systematically identify wards with sufficient power to be analyzed using our inference methodology, the resulting estimated values were compared with true values. For each dataset, a deviation was estimated as the ratio of the estimated to the true value. If the median value of this deviation across all analyses for the ward was ≤ 1.15, it was considered that the ward had sufficient power to be analyzed.

Simulated Epidemic Curves

After identification of sets of parameter values with likelihoods within the 95% CI relative to maximum likelihood, these sets of parameter values were sampled with replacement 1,000 times, and each time an epidemic was simulated. Those parameters that went to extinction (having <3 cumulative infections) were excluded, and the remaining epidemics were used to calculate the median and 95% CI for relevant epidemic variables (number of positive tests, detected and undetected symptomatic and asymptomatic prevalence) for each date.

Due to repeat testing, it is not possible to calculate exactly the number of infections that were detected and undetected in simulated data, but equation 23 provides an approximation, where undetected includes both untested patients as well as false negatives. Symbols are described in Appendix Table 1.

Equation 23

Prevalence of undetected asymptomatics $= \left(E_a + I_a + E_{aT}(1 - Z_{Ea}) + I_{aT}(1 - Z_{Ia})\right)$

Prevalence of undetected symptomatics $= \left(E_s + I_s + E_{sT}(1 - Z_{Es}) + I_{sT}(1 - Z_{Is})\right)$

Prevalence of detected asymptomatics $= (E_{aT}Z_{Ea} + I_{aT}Z_{Ia})$

Prevalence of detected symptomatics $= (E_{sT}Z_{Es} + I_{sT}Z_{Is})$

Calculating Time-Varying Reproduction Number

We calculated the time-varying reproduction number, R_t, across the entire hospital and from the date of the first positive test until the 50th day, using the EpiEstim package

(https://CRAN.R-project.org/package=EpiEstim). The number of tests per day that were the first positive test for each patient was used as the daily incidence. We assumed a serial interval of 5.8 (range 4.8–6.8) days (*3*).

Considerations in Adapting to Alternative Scenarios

The model code is available on Github (github.com/georgeshirreff/Hospital_R0_C19) with the intention that this model can be applied to other healthcare environments. With minimal adjustment, the could be directly adapted to a dataset aggregating positive and negative tests for active infection, admissions, and discharges each day. It could easily be adapted to different types of SARS-CoV-2 tests by adjusting the sensitivity and specificity for different stages of infection. Testing for both active infection and serology could also be included with consideration of the outcomes from testing at each stage of infection and adjustment to the likelihood function to account for both testing streams.

Our model rests on an assumption of free mixing between persons, so ward-level analysis might be more appropriate where this is unrealistic. Simultaneous modeling of different subpopulations, such as HCWs and patients in different wards, could be straightforwardly achieved, requiring consideration of the relative contact rates between these groups. The number of iterations and particles used for the statistical inference, as well as the number of repetitions of the analysis, might need to be adjusted to the size of the dataset and the desired level of precision.

Supplementary Results

Validation of Statistical Inference of Model Parameters

Simultaneous estimation of known parameter values was conducted for both the 1- and 2-phase models on data at the scale of the whole hospital. We describe the relationship between known and estimated values of β and t_{init} for the 1-phase model (Appendix Figure 4). The value of β was well estimated throughout the range, but the estimate of t_{init} was slightly over-estimated from days −25 to −16, but not after this point.

We also describe the relationship for the 2-phase model between known values of β_1, β_2, and t_{init} and their estimates (Appendix Figure 5). The first phase transmission rate, β_1, was reliably estimated up to 1.0 with a slight overestimation after that point. The second phase

transmission rate, β_2, was correctly estimated, except for values wrongly estimated to be close to zero. However, in the absence of estimates close to zero in the analysis of true data, the estimates could be considered reliable. As with the 1-phase model, the estimate of t_{init} deviated slightly when it fell in the first weeks but was more reliable after day −15.

The ability of our framework to correctly estimate parameter values at the ward level was limited due to much smaller population sizes and specific distributions of tests. We illustrate the ability to estimate values of β and t_{init} simultaneously using the 1-phase model on individual wards (Appendix Figures 6,7). Our results suggest that only data corresponding to wards A2, C0, C1, C2, and C3 provided sufficient power to be analyzed through our framework, and C1 was also excluded due to the lack of visible relationship between true and estimated values.

Results of Whole-Hospital Analysis

We show results for all analyses of the 1-phase model (Appendix Table 2) and for E_{init} = 1 only (Table 1). We show results for all analyses of the 2-phase model with $t_{inflect}$ = day 12 (Appendix Table 3) and those for E_{init} = 1 only (Table 1). We also show analyses exploring the effect of changing $t_{inflect}$ (Appendix Table 4).

Results of Ward-Level Analysis

We compared results from each ward (Appendix Table 5). These results demonstrate that the Akaike information criterion (AIC) for the 1-phase model is lower or equal for 3 of 4 wards.

Sensitivity Analysis

We calculated best estimates and 95% CI for β_1, β_2, and t_{init} (Appendix Figure 8). Many parameters affect the upper ranges of β_1, most markedly ε which also affects the best estimate. However, the relative effects on β_2 are much greater; δ and the sensitivity parameters of form Z_x have a large effect on both the mean estimate and range. An early inflection point, $t_{inflect}$, serves to suggest the possibility of an earlier epidemic initiation point, t_{init}, but the biggest effect on t_{init} comes from perturbing the number of index cases, E_{init}, with a larger number of indices pointing to a later introduction.

Time-Varying Reproduction Number

The estimated time-varying reproduction number (R_t) had an initially high value, 10 (1.8–23.7), which is consistent with our own estimate of R_0 in the first phase (Appendix Figure 9). R_t drops to a low of 2.8 (0.7–5.8), which reflects the fall in our own estimate in the second

phase of the analysis, representing a decrease of 72%. A similar magnitude is observed by calculating the mean R_t before, 8.0 (1.5–18.9), and after, 2.1 (1.3–3.1), our estimated change point, $t_{inflect}$. The analysis also displays a second peak after the first fall in R_t. Because this method only accounts for incident cases and ignores the evolving testing rate, it is likely that this peak reflects that substantial increase in testing rate rather than a true increase in transmission rate.

References

1. Buitrago-Garcia D, Egli-Gany D, Counotte MJ, Hossmann S, Imeri H, Ipekci AM, et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. PLoS Med. 2020;17:e1003346. PubMed https://doi.org/10.1371/journal.pmed.1003346

2. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science. 2020;368:489–93. PubMed https://doi.org/10.1126/science.abb3221

3. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med. 2020;26:672–5. PubMed https://doi.org/10.1038/s41591-020-0869-5

4. 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;173:262–7. PubMed https://doi.org/10.7326/M20-1495

5. Ra SH, Lim JS, Kim G, Kim MJ, Jung J, Kim S-H. Upper respiratory viral load in asymptomatic individuals and mildly symptomatic patients with SARS-CoV-2 infection. Thorax. 2021;76:61–3. PubMed https://doi.org/10.1136/thoraxjnl-2020-215042

6. King AA, Nguyen D, Ionides EL. Statistical inference for partially observed Markov processes via the R package pomp. J Stat Softw. 2016;69:1–43. https://doi.org/10.18637/jss.v069.i12

Appendix Table 1. Symbols, parameters, values, and state variables used in a model of basic reproduction number of nosocomial SARS-CoV-2 transmission

Symbol	Parameter	Value (95% CI)	Source
t_{init}	Date on which the initial infection occurs	Estimated from the modeling analysis	NA
$t_{inflect}$	Date on which the value of β changes in the 2-phase model	Estimated within range 1–16 (relative to date of first positive sample)	NA
E_{init}	The number of initial infections at date t_{init}	1 (default), 3 or 10 cases	NA
β_1	Transmission rate per day before the inflection date	Estimated from the modeling analysis	
β_2	Transmission rate per day after the inflection date	Estimated from the modeling analysis	NA
β	The current transmission rate per day, or the single transmission rate in the 1-phase model	Estimated from the modeling analysis	
ε	Relative transmission rate from during pre-symptomatic infection compared with symptomatic infection	0.63 (0.18–2.26)	(*1*)
$1/\alpha$	Mean duration of non-infectious incubation in days	3.4 (3.3–4.0)	(*2*) Latent period
$1/\gamma$	Mean duration of infectious incubation stage in days	2.3 (0.8–3.0)	(*3*) Duration of pre-symptomatic infection
$1/\delta$	Mean duration of full-blown infection in days	7 (2.4–9.1)	(*3*) with uncertainty proportional to that of duration of infectious incubation
$1/\omega$	Mean duration of viral shedding following recovery from infectious stage in days	20 (0–60)	Hospital data; see Appendix Methods; Appendix Figure 3
Ψ	Proportion entering symptomatic pathway	0.69 (0.62–0.76)	(*1*)
κ_1	Relative infectivity of asymptomatics in full infection relative to full symptomatic infection	0.35 (0.1–1.27)	(*1*)
κ_2, κ_3	Relative rates of progression to full infection (E_a to I_a) and recovery in asymptomatic pathway, relative to symptomatic pathway	1	Assumption
μ	Relative rate of discharge for symptomatic patients relative to any non-symptomatic patient	1	Assumption
Z_E, Z_{Ea}, Z_{Es}, Z_{Ia}, Z_{Is}, Z_{Rp}	PCR test sensitivity for E, E_a, E_s, I_a, I_s, or R_p states	0.1 (0–0.5), 0.7, 0.7 (0.25–0.85), 0.8, 0.8 (0.65–0.9), 0.3 (0.2–0.5)	(*4*) with (*5*) confirming that viral loads in symptomatic and asymptomatic infection are similar
v	PCR test specificity	0.99 (0.96–0.992)	Assumption (lower bound of range comes from A.N. Cohen et al., unpub. data, http://medrxiv.org/lookup/doi/10.1101/2020.04.26.20080911)
φ	Relative rate of retesting compared to testing for the first time	0.6 (0.5–0.7)	Hospital data; see Appendix Methods
$A(d)$	The number of admissions in the data on day d		Hospital data
$D(d)$	The number of discharges in the data on day d		Hospital data
$T(d)$	The number of tests in the data on day d		Hospital data
$C(d)$	The number of infections initiated (moved from S to E or S_T to E_T on day d)		t_{init} and E_{init}
$\lambda(t)$	Force of infection at time t		State variable
			State variable
W	The total rate adjusted number of dischargeable individuals across all compartments at a given time		State variable
M	The total rate adjusted number of testable individuals across all compartments including those eligible for retesting, but excluding untested I_s, at a given time		State variable
N	The total population size at a given time		State variable
U	The universal set of compartments		State variable
Admission(t)	Admissions occurring into compartment S at time t		State variable

Initiation(X,t)	Infection initiations occurring from compartment X at time t	State variable
Discharge(X,t)	Discharges occurring from compartment X at time t	State variable
Test(X,t)	Tests occurring from compartment X at time t	State variable
S	Susceptible untested at a given time	State variable
E	Infected uninfectious untested at a given time	State variable
E_a	Early infectious infection on asymptomatic pathway, untested at a given time	State variable
E_s	Pre-symptomatic infectious infection on symptomatic pathway, untested at a given time	State variable
I_a	Full-blown infection on asymptomatic pathway, untested at a given time	State variable
I_s	Full-blown symptomatic infection, untested at a given time	State variable
R_p	Recovered but still shedding virus, untested at a given time	State variable
R	Recovered and no longer shedding virus, untested at a given time	State variable
S_T	Susceptible tested at a given time	State variable
E_T	Infected uninfectious tested at a given time	State variable
E_{aT}	Early infectious infection on asymptomatic pathway, tested at a given time	State variable
E_{sT}	Pre-symptomatic infectious infection on symptomatic pathway, tested at a given time	State variable
I_{aT}	Full-blown infection on asymptomatic pathway, tested at a given time	State variable
I_{sT}	Full-blown symptomatic infection, tested at a given time	State variable
R_{pT}	Recovered but still shedding virus, tested at a given time	State variable
R_T	Recovered and no longer shedding virus, tested at a given time	State variable

Appendix Table 2. Best estimates and the ranges for β in the 1-phase model and corresponding R_0 values to assess nosocomial SARS-CoV-2 transmission*

Estimate	β	R_0	E_{init}	t_{init}†	AIC
β, t_{init}	0.38 (0.30–0.60)	2.6 (2.0- 4.1)	**1**	−22 (−39 to −4)	657.3257
	0.40 (0.29–0.62)	2.7 (2.0- 4.2)	**3**	−8 (−38 to −2)	656.5639
	0.38 (0.26–0.60)	2.6 (1.8- 4.1)	**10**	−4 (−11 to 0)	653.7993
β, E_{init}	0.37 (0.27–0.61)	2.5 (1.8- 4.1)	2.7 (1.5–19.9)	**–6**	654.4575
	0.37 (0.26–0.61)	2.5 (1.8- 4.1)	2.5 (0.5–11.3)	**–13**	656.2111
	0.40 (0.29–0.57)	2.7 (2.0- 3.9)	1.8 (0.5- 7.8)	**–20**	655.3993

*R_0 values are calculated using equation 4. Bold text indicates fixed values. AIC, Akaike information criterion; NE, not estimated; β_1, transmission rate per day before the inflection date; β_2, transmission rate per day after the inflection date; E_{init}, number of initial infections at date initial infection occurs; R_0, basic reproduction number; t_{init}, date initial infection occurs.
†Values for t_{init} are relative to the day of the first positive sample.

Appendix Table 3. Best estimates and their ranges for β_1, β_2 from the 2-phase model and corresponding R_0 values to assess nosocomial SARS-CoV-2 transmission*

Estimate	E_{init}	β_1	β_2	$R_{0\ before}$	$R_{0\ after}$	t_{init}	AIC
β_1, β_2, t_{init}	**1**	1.28 (0.76–2.40)	0.19 (0.10–0.30)	8.72 (5.14–16.32)	1.33 (0.68–2.04)	−4 (−24 to 0)	628.85
	3	1.23 (0.68–2.20)	0.19 (0.10–0.28)	8.37 (4.65–14.96)	1.31 (0.66–1.89)	−2 (−13 to 4)	628.4
	10	1.03 (0.59–2.80)	0.20 (0.10–0.27)	7.03 (4.00–19.04)	1.39 (0.67–1.85)	0 (−6 to 6)	631.21
β_1, β_2,	0.70 (0.50–2.49)	1.10 (0.63–2.00)	0.20 (0.11–0.30)	7.48 (4.26–13.60)	1.39 (0.76–2.02)	**–6**	634.51
E_{init}	0.63 (0.50–2.43)	1.14 (0.69–2.10)	0.20 (0.10–0.29)	7.78 (4.68–14.28)	1.36 (0.69–1.96)	**–13**	631.74
	1.12 (0.50–7.47)	1.24 (0.70–2.20)	0.19 (0.10–0.29)	8.42 (4.73–14.96)	1.32 (0.67–1.95)	**–20**	629.24

*R_0 values were calculated using equation 4, substituting the corresponding β value. Bold text indicates fixed values. AIC, Akaike information criterion; NE, not estimated; β_1, transmission rate per day before the inflection date; β_2, transmission rate per day after the inflection date; E_{init}, number of initial infections at date initial infection occurs; R_0, basic reproduction number; t_{init}, date initial infection occurs.

Appendix Table 4. Effect of changing $t_{inflect}$ on estimated parameter values in the 2-phase model to assess nosocomial SARS-CoV-2 transmission*

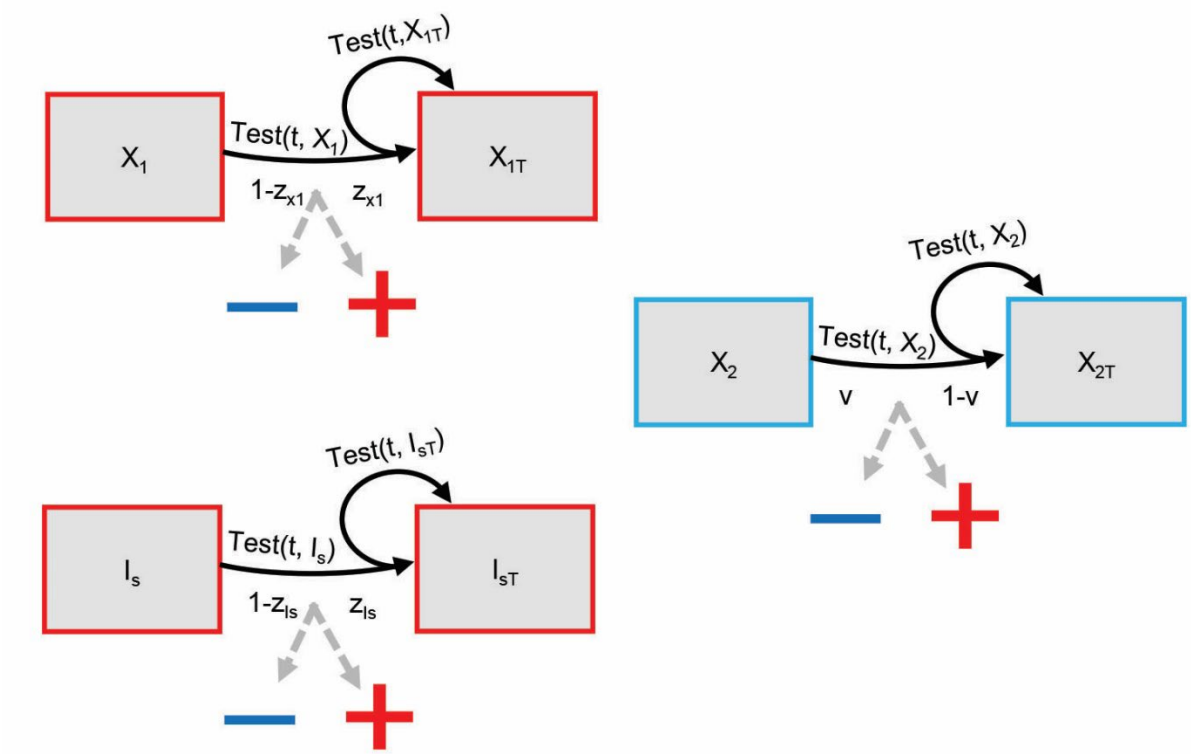
$t_{inflect}$	β_1	β_2	$R_{0\ before}$	$R_{0\ after}$	$R_{0\ combined}$†	Intervention efficacy	t_{init}	AIC
1	3.44 (0.29–49.50)	0.36 (0.23–0.57)	23.38 (2.00–336.67)	2.47 (1.60–3.88)	13.59 (2.41–180.07)	0.89 (−0.47 to 0.99)	−4 (−37 to 1)	654.27
6	2.68 (0.99–49.50)	0.29 (0.17–0.40)	18.21 (6.71–336.67)	1.98 (1.16–2.70)	10.61 (4.50–179.87)	0.89 (0.64–1.00)	−2 (−19 to 6)	639.97
8	2.24 (0.99–12.10)	0.25 (0.16–0.34)	15.22 (6.73–82.30)	1.68 (1.07–2.35)	8.87 (4.39–44.43)	0.89 (0.68–0.98)	−2 (−20 to 3)	633.56
10	1.62 (0.91–4.00)	0.22 (0.14–0.32)	11.01 (6.18–27.21)	1.49 (0.96–2.14)	6.55 (4.14–15.12)	0.86 (0.67–0.95)	−3 (−22 to 1)	629.89
12	1.28 (0.76–2.00)	0.19 (0.10–0.30)	8.72 (5.14–13.60)	1.33 (0.68–2.04)	5.26 (3.38–7.94)	0.85 (0.66–0.94)	−4 (−24 to 1)	628.85
14	1.02 (0.70–1.50)	0.17 (0.08–0.26)	6.91 (4.73–10.20)	1.15 (0.56–1.77)	4.22 (3.06–6.09)	0.83 (0.63–0.94)	−5 (−27 to 2)	630.19
16	0.84 (0.54–1.08)	0.16 (0.07–0.26)	5.69 (3.65–7.31)	1.07 (0.48–1.78)	3.53 (2.61–4.29)	0.81 (0.61–0.93)	−5 (−28 to 2)	634.83

*Best estimates and ranges for β_1, β_2, corresponding R_0 values, and t_{init}. R_0 values are calculated using equation 4, substituting the corresponding β value. The risk ratio is calculated for each point estimate as β_1/β_2. AIC, Akaike information criterion; β_1, transmission rate per day before the inflection date; β_2, transmission rate per day after the inflection date; R_0, basic reproduction number; $t_{inflect}$, date on which the value of β changes in the 2-phase model; t_{init}, date initial infection occurs.
†The combined R_0 is an average R_0 in each phase weighted by phase duration as in equation 5.

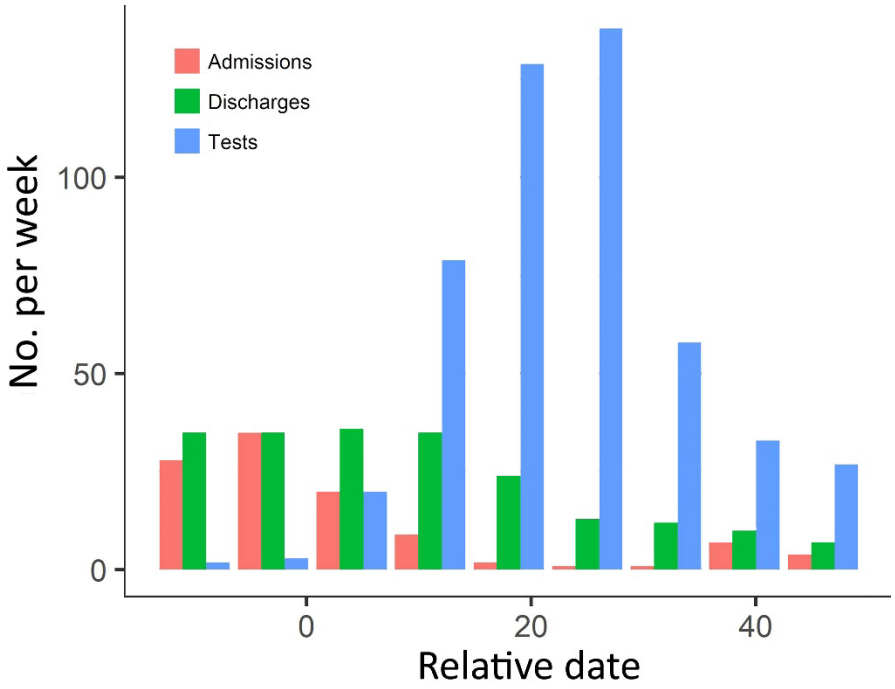
Appendix Table 5. Best estimates and ranges for β_1, β_2, and R_0 in each phase, combined, and t_{init} for each hospital ward in a 2-phase model to assess nosocomial SARS-CoV-2 transmission*

	2-phase, value (95% CI)†						1-phase
Ward	β_1	β_2	Risk ratio‡	$R_{0\ combined}$§	t_{init}	AIC	AIC
A2	2.16 (0.30–NE)	0.70 (0.31–4.42)	0.33 (0.04–11.39)	10.41 (4.77–49.01)	4 (−20 to 7)	139.4	138.25
C0	NE	0.35 (0.26–4.89)	0.04 (0.03–7.20)	39.44 (1.75–50.23)	10 (−38 to 11)	89.91	91.59
C2	NE	0.00 (0.00–0.08)	0.00 (0.00–0.06)	37.25 (1.16–38.52)	−14 (−33 to −12)	57.62	57.92
C3	6.50 (0.00–NE)	0.41 (0.23–0.64)	0.06 (0.03–NE)	26.43 (1.03–39.82)	20 (−26 to 21)	47.25	45.25

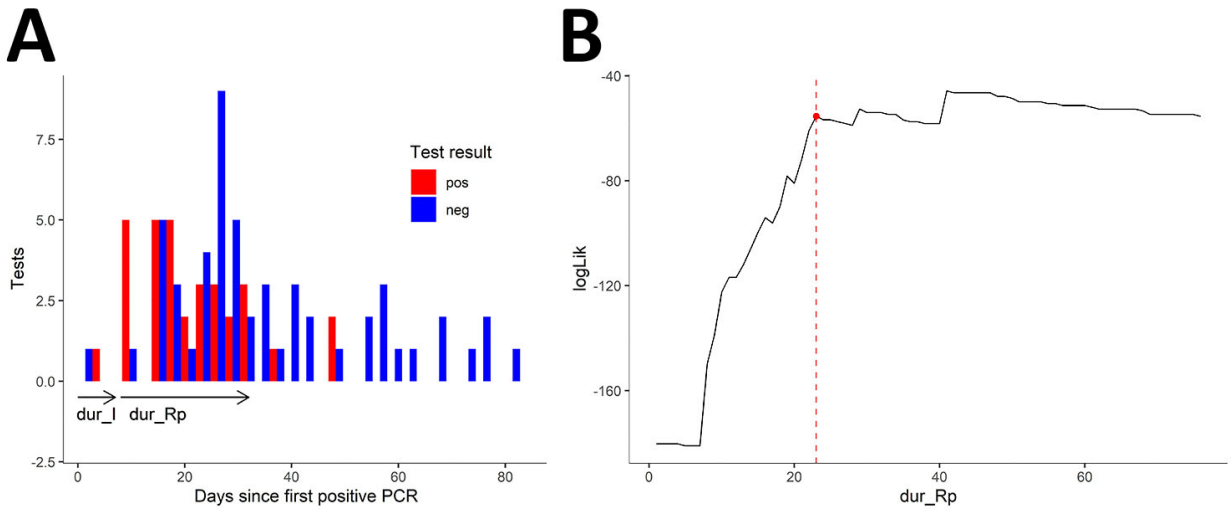
*The values of E_{init} was fixed at day 1 and $t_{inflect}$ was fixed at day 11. AIC, Akaike information criterion; NE, not estimated; β_1, transmission rate per day before the inflection date; β_2, transmission rate per day after the inflection date; E_{init}, number of initial infections at date initial infection occurs; R_0, basic reproduction number; $t_{inflect}$, date on which the value of β changes in the 2-phase model; t_{init}, date initial infection occurs.
†In many instances, the upper bound of the 95% CI for β_1, and in some also the most likely value of β_1, could not be estimated due to a flat likelihood surface, in which case the value is given as NE.
‡The risk ratio is calculated for each point estimate as β_1/β_2.
§R_0 values were calculated using equations 4 and 5.



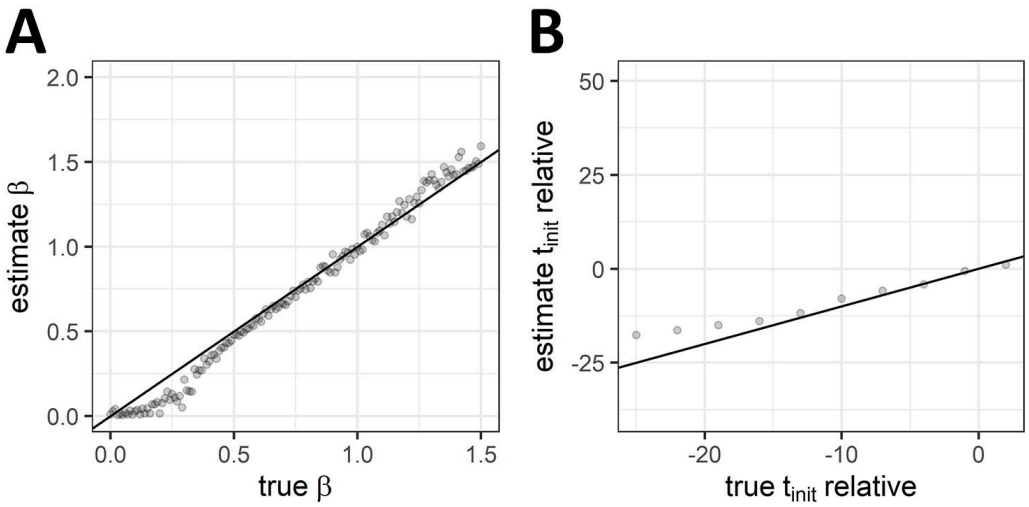
Appendix Figure 1. Illustration of the observation process for a model of nosocomial SARS-CoV-2 transmission. X_1 represents any compartment of untested persons who are shedding virus (E, E_a, E_s, I_a, R_p) who test positive at their compartment-specific sensitivity rate (z_{X1}), and X_2 represents any compartment of persons not recently tested who are not shedding virus (S, R) who test negative at rate v. X_{1T}, I_{sT}, and X_{2T} represent tested counterparts. The symptomatic persons (I_s and I_{sT}) are shown separately because testing is conducted first on the non-recently tested symptomatic group, but retesting is equally likely for symptomatic persons as for asymptomatic persons. Upon testing or retesting, the dotted arrows indicate the probabilities of the possible observed outcomes, positive and negative. E, exposed; E_a, asymptomatic exposed; E_s, symptomatic exposed; I_a, asymptomatic infected; I_s, symptomatic infected; I_{sT}, symptomatic infected and tested; R, recovered; R_p, recovered but shedding virus; S, susceptible; +, positive; – negative.



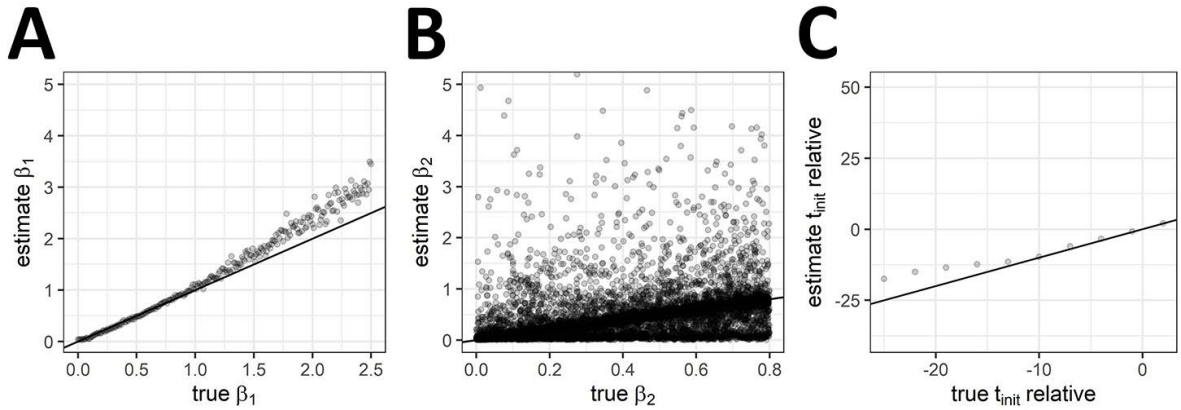
Appendix Figure 2. Weekly aggregated number of admissions, discharges, and PCR tests reported over the study periodin a hospital used for developing a model of basic reproduction number of nosocomial SARS-CoV-2 transmission. The daily disaggregated data are used in the model as $A(d)$, admissions/day; $D(d)$, discharges/day; and $T(d)$, tests/day.



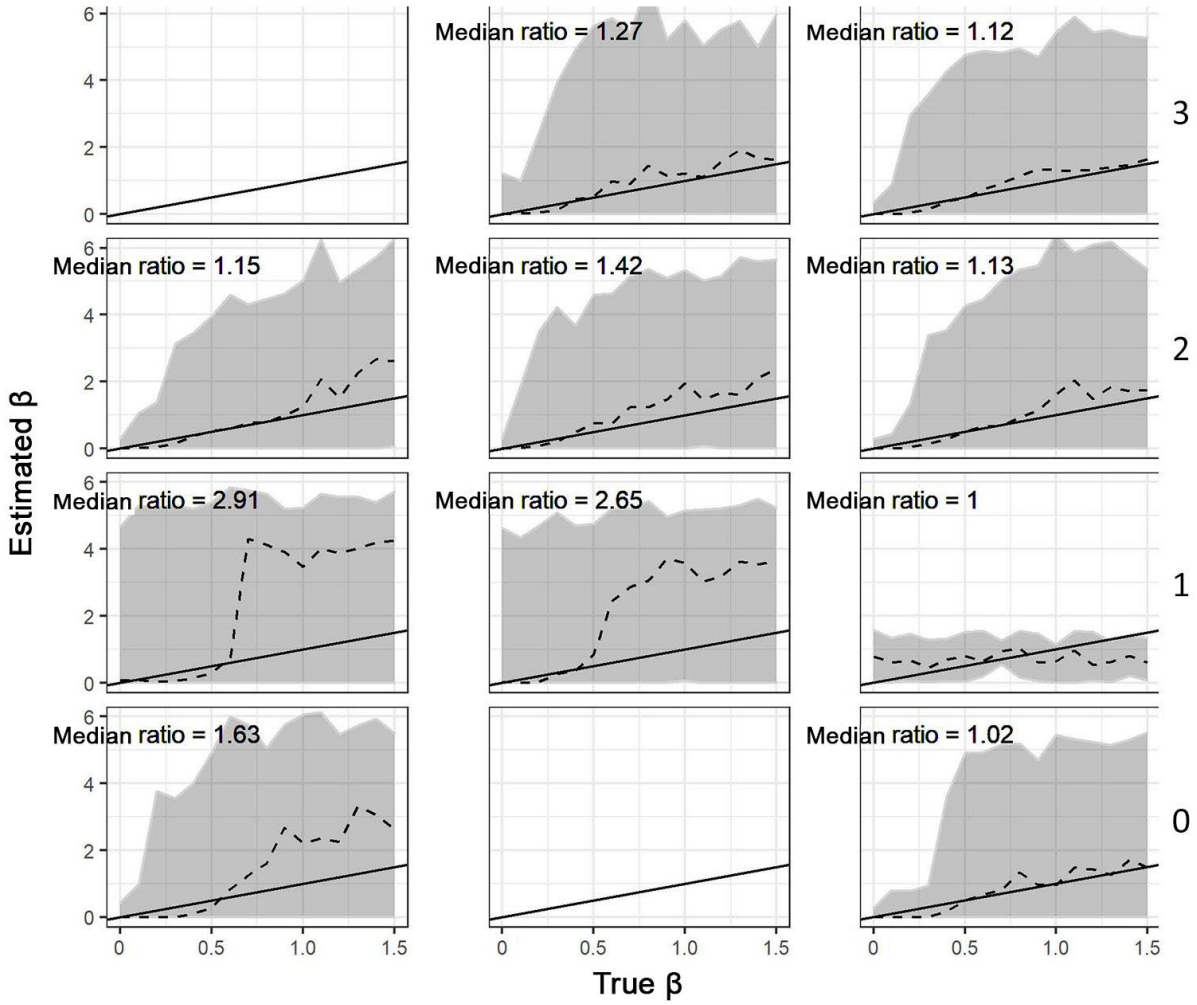
Appendix Figure 3. Results from repeat tests taken after a first PCR–positive test among patients in a hospital used to develop a model of basic reproduction number of nosocomial SARS-CoV-2 transmission. A) Number of positive (pos) and negative (neg) PCR tests reported. The duration of the full-blown infection stage was 7 days (Table 1). B) Likelihood for each potential duration (in days) of the R_p stage according to the sensitivity of the subsequent stages and the number of tests from each. dur_I, duration of infectiousness; dur_R_p, duration of viral shedding after recovered; R_p, recovered but shedding virus.



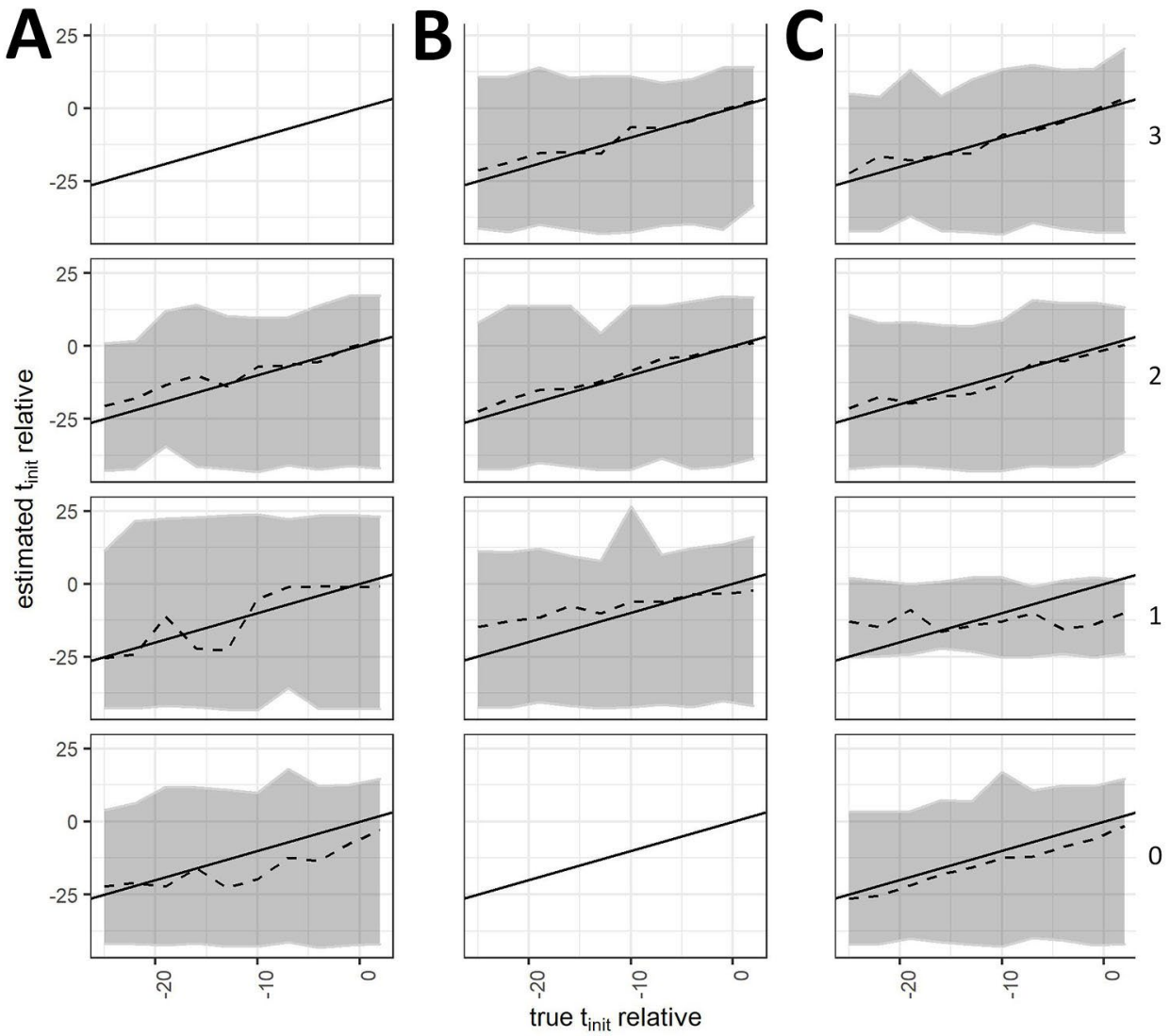
Appendix Figure 4. Validation of simultaneous estimation of transmission rate and initial infection date using the 1-phase model of nosocomial SARS-CoV-2 transmission on datasets at the whole-hospital level. A) Estimation of β. B) Estimation of t_{init}. Each point represents a true value of the parameter on its x-axis, with the value on the y-axis being the median across 10 attempts to estimate the true value using particle filtering. The solid black line indicates where the true and estimated values are equal. The value of E_{init} was fixed at 1. E_{init}, number of initial infections at date initial infection occurs; t_{init}, date initial infection occurs; β, current transmission rate per day.



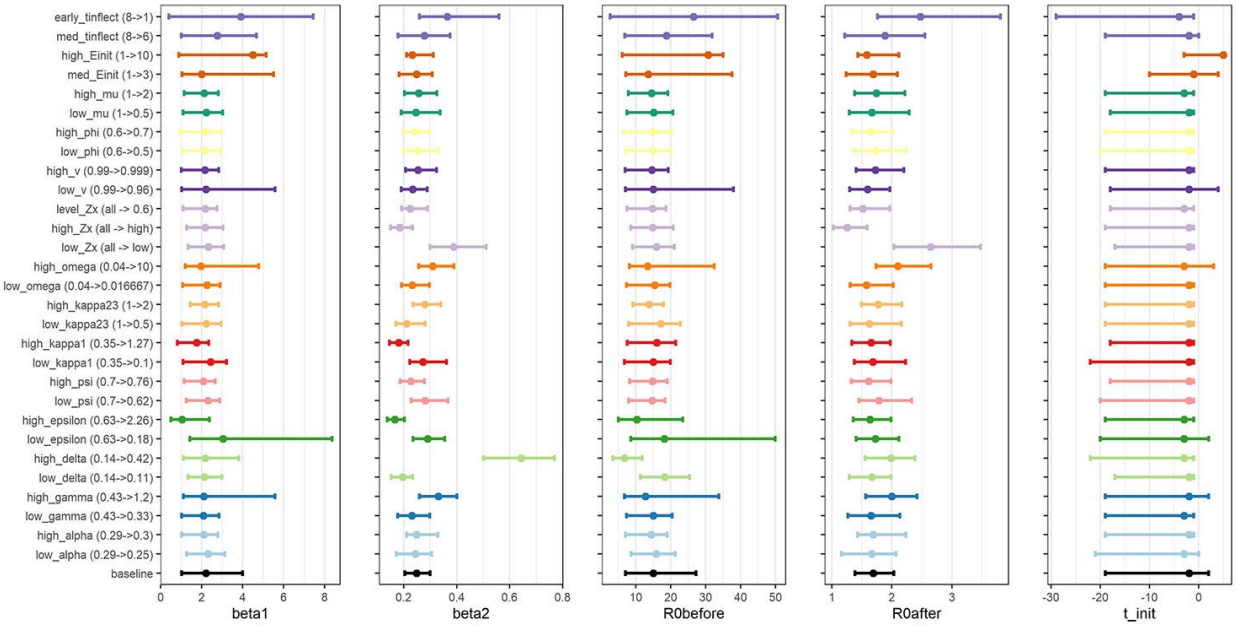
Appendix Figure 5. Validation of simultaneous estimations of transmission rates and initial infection date using a 2-phase model of nosocomial SARS-CoV-2 transmission on the datasets at the scale of the whole hospital. A) Validation for β_1. B) Validation for β_2. C) Validation for t_{init}. Each point represents a true value of the parameter on its x-axis; the value on the y-axis is the median across 10 attempts to estimate the true value using particle filtering. Values for E_{init} were fixed at day 1 and values for $t_{inflect}$ were fixed at day 12. Solid black line indicates where the true and estimated values are equal. E_{init}, number of initial infections on the date initial infection occurs; $t_{inflect}$, date on which the value of β changes in the 2-phase model; t_{init}, date initial infection occurs; β_1, transmission rate per day before the inflection date; β_2, transmission rate per day after the inflection date.



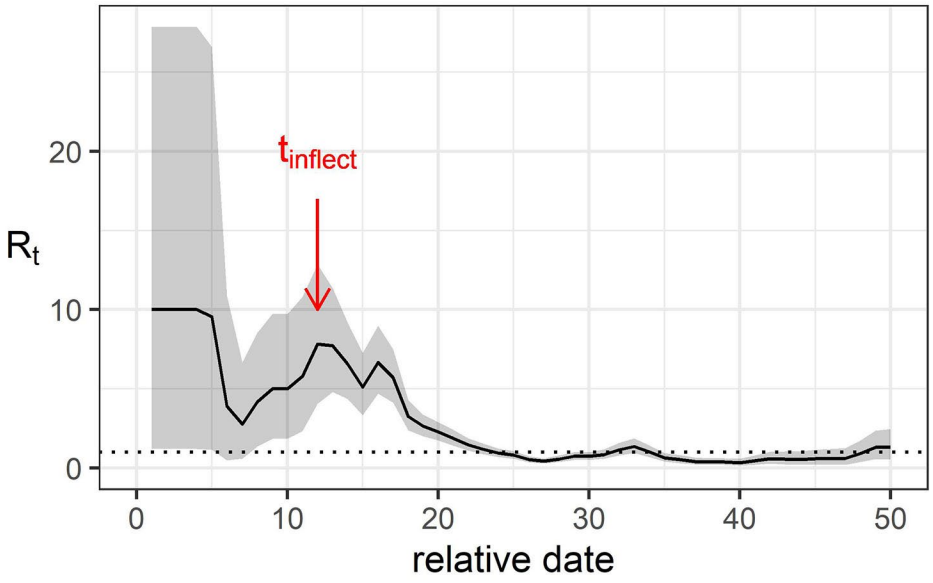
Appendix Figure 6. Validation of the estimation of *β* using the 1-phase model of nosocomial SARS-CoV-2 transmission on datasets at the ward level in a hospital used for developing a model to measure basic reproduction number. Columns represent the hospital buildings A–C (left–right); rows 0–3 represent the floors in each building. The black dashed line represents the median estimate on the y-axis for each true value of the parameter on the x-axis. The gray area represents the 95% range of estimates for each value of the true parameter. The solid black line indicates where the true and estimated values are equal. The value of E_{init} was fixed at 1. The numerical value given in the corner is the median ratio between the estimated and true values. E_{init}, number of initial infections at date initial infection occurs; *β*, current transmission rate per day.



Appendix Figure 7. Validation of the estimation of initial infection date (t_{init}) using the 1-phase model of nosocomial SARS-CoV-2 transmission on datasets at the ward level. Columns A–C represent the hospital buildings; rows 0–3 represent the floors in each building. The black dashed line represents the median estimate on the y-axis for each true value of the parameter on the x-axis. The gray area represents the 95% range of estimates for each value of the true parameter. The solid black line indicates where the true and estimated values are equal. The value of E_{init} was fixed at 1. E_{init}, number of initial infections at date initial infection occurs.



Appendix Figure 8. Results of sensitivity analysis of a 2-phase model of nosocomial SARS-CoV-2 transmission. Analysis shows beta1 (β_1), beta2 (β_2), and their corresponding R_0 values $R_{0before}$ (corresponds to β_1), and R_{0after} (corresponds to β_2); and t_{init} under parameter values perturbed according to their uncertainty ranges (Table 1; Appendix Methods), with all other parameters held at baseline values. Error bars indicate 95% CI, the corresponding dot shows the values of the parameter values that had the highest likelihood. In the scenario modifying Z_x, all Z_x parameters were modified simultaneously, to their lower or upper bounds, or to 0.6. In the scenario kappa23, both κ_2, the relative rates of progression from stage E_a, and κ_3, the relative rate of progression from stage I_a, compared to the equivalent symptomatic stage, were modified by the same factor. R_0, basic reproduction number; E_a, asymptomatic exposed; E_{init}, number of initial infections at date initial infection occurs; I_a, asymptomatic infected; $t_{inflect}$, date on which the value of β changes in the 2-phase model; t_{init}, date initial infection occurs; β_1, transmission rate per day before the inflection date; β_2, transmission rate per day after the inflection date, Z_x, PCR test sensitivity.



Appendix Figure 9. Time-varying reproduction number of nosocomial SARS-CoV-2 transmission. Estimation performed by using the EpiEstim package (https://CRAN.R-project.org/package=EpiEstim), and on the basis of incident cases, using a serial interval mean of 5.8 days and standard deviation of 0.51. The solid black line indicates the median estimate. The gray area indicates the 95% credibility interval. The red arrow indicates our best estimate for the transmission rate change point in our 2-phase model analysis. $t_{inflect}$, date on which the value of β changes in the 2-phase model; β, current transmission rate per day.