The intensity of COVID-19 outbreaks is modulated by variation in SARS-CoV-2 free-living survival and environmental transmission

C. Brandon Ogbunugafor[1,2]*, Miles Miller-Dickson[1], Victor A. Meszaros[1], Lourdes M. Gomez[1], Anarina L. Murillo[3,4], and Samuel V. Scarpino[5]

[1]Department of Ecology and Evolutionary Biology, Brown University
[2]Center for Computational Molecular Biology, Brown University
[3]Department of Pediatrics, Warren Alpert Medical School at Brown University
[4]Center for Statistical Sciences, Brown University School of Public Health
[5]Network Science Institute, Northeastern University
*Corresponding author (brandon_ogbunu@brown.edu)

ABSTRACT

COVID-19 has circled the globe, rapidly expanding into a pandemic within a matter of weeks. While early studies revealed important features of SARS-CoV-2 transmission, the role of variation in free-living virus survival in modulating the dynamics of outbreaks remains unclear. Using an empirically determined understanding of SARS-CoV-2 natural history and detailed, country-level case data, we elucidate how variation in free-living virus survival influences key features of COVID-19 epidemics. Our findings suggest that COVID-19's basic reproductive number ($\mathcal{R}_0$) and other key signatures of outbreak intensity are defined by transmission between infected individuals and the environment. Summarizing, we propose that variation in environmental transmission may explain observed differences in disease dynamics from setting to setting, and can inform public health interventions.

INTRODUCTION

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), the etiological agent of coronavirus disease 2019 (COVID-19), has caused one of the most devasting pandemics of the last century. The complex set of epidemiological characteristics defining COVID-19 outbreaks presents a number of challenges for controlling this disease. As a consequence, countries have achieved varying levels of success in reducing transmission and protecting vulnerable populations, often with dramatic variation from setting to setting in the epidemic growth rate, intensity, and/or severity. Developing a mechanistic understanding of how SARS-CoV-2 is transmitted in different settings is thus essential for guiding ongoing and future non-pharmaceutical interventions.

The basic reproductive number ($\mathcal{R}_0$) (1-3), fatality rate (4-5), incubation period (4, 6-8), transmission interval (9), prevalence of super-spreading events (10-11) and other relevant aspects of COVID-19 epidemiology that provide a mechanistic window into how SARS-CoV-2 is transmitted in different settings. However, one feature of SARS-CoV-2 transmission that was validated in laboratory settings (12), but whose epidemiological role remains uncertain, is SARS-CoV-2 free-living survival. Specifically, while several laboratory and epidemiological findings have suggested that environmental transmission may play a role in some settings (11, 13-14), none have fully investigated how this route of transmission may influence outbreaks.

In this study, we identify key laboratory and epidemiologically validated parameters associated with SARS-CoV-2 environmental transmission. We then integrate these findings into a mechanistic transmission model of SARS-CoV-2 to evaluate the potential for variability in environmentally-mediated transmission to explain variability in COVID-19 outbreak intensity. This framework includes parameters corresponding to the transmission of the virus from both presymptomatic/asymptomatic and clinical (symptomatic) carriers of SARS-CoV-2, and the possibility that susceptible hosts can acquire infection through environmental reservoirs. We examine how outbreak dynamics are influenced by differences in SARS-CoV-2 free-living survival that include

values corresponding to empirical values for survival on various abiotic surfaces (e.g. plastic, copper, steel, cardboard) (12).

Using confirmed case data from seventeen countries from around the world with large outbreaks, we find evidence that the role of environmental SARS-CoV-2 transmission on COVID-19 epidemics can vary from setting to setting. Our findings highlight the need to incorporate the potential for environmental transmission into COVID-19 non-pharmaceutical intervention plans and forecasting models. Specifically, we propose that environmental transmission—including the particulars of how free-living virus survives on different abiotic surfaces—should be a greater focus in emerging infectious disease outbreaks, as undertreatment of this route can obfuscate essential properties of how the disease is spreading and potential avenues for intervention.

A *Waterborne, Abiotic,* and other *Indirectly Transmitted* (WAIT) model of SARS-CoV-2 transmission. Several models have been engineered to explore aspects of COVID-19 dynamics. For example, models have been used to investigate the role of social distancing (2, 15), social mixing (16), the importance of undocumented infections (17), the role of mobility in the early spread of disease in China (18), and the potential for contact tracing as a solution (19). Only a few notable models of SARS-CoV-2 transmission incorporate features of indirect or environmental transmission (13-14, 19), and none consider the dynamical properties of viral free-living survival in the environment. Such a model structure would provide an avenue towards exploring how variation in free-living survival influences disease outbreaks. Environmental transmission models are aplenty in the literature and serve as a theoretical foundation for exploring similar concepts in SARS-CoV-2 transmission (20-29).

Here, we parameterize and validate an SEIR-W model: Susceptible (S), Exposed (E), Infectious (I), Recovered (R), and WAIT (W) model. Here W represents the environmental component of the transmission cycle during the early stage of the SARS CoV-2 pandemic. This model is derived from a framework previously developed called the “WAIT” modeling framework—which stands for *Waterborne, Abiotic,* and other

Indirectly Transmitted—that incorporates an environmental reservoir where a pathogen can sit and *wait* for hosts to interact with it (30-31).

METHODS

Building the SEIR-W model framework for SARS-CoV-2. Here *W* represents the environmental component of the early stage of the SARS CoV-2 pandemic (Figure 1). In the context of SARS CoV-2, this environmental compartment refers to surfaces that people may have contact with on a daily basis, such as doorknobs, tables, chairs, or mail packages. The *W* compartment of our model represents the fraction of these environmental reservoirs that house some sufficiently transmissible amount of infectious virus. We emphasize that the *W* compartment is meant to only represent surfaces that are common sites for interaction with people. Thus, inclusion of the *W* compartment allows us to investigate the degree to which the environment is infectious at any given point, and its impact on the transmission dynamics of SARS CoV-2. We will sometimes use the term “environment” and “surfaces” or “objects” interchangeably.

Model parameters are described in detail in Table 1. The system of equations in the proposed mathematical model corresponding to these dynamics are defined in equations 1-6:

$$\frac{dS}{dt} = \mu(N - S) - \left(\frac{\beta_A I_A + \beta_S I_S}{N} + \beta_W W\right)S \tag{1}$$

$$\frac{dE}{dt} = \left(\frac{\beta_A I_A + \beta_S I_S}{N} + \beta_W W\right)S - (\epsilon + \mu)E \tag{2}$$

$$\frac{dI_A}{dt} = \epsilon E - (\omega + \mu)I_A \tag{3}$$

$$\frac{dI_S}{dt} = (1 - p)\omega I_A - (\nu + \mu_S)I_S \tag{4}$$

$$\frac{dR}{dt} = p\omega I_A + \nu I_S - \mu R \tag{5}$$

$$\frac{dW}{dt} = \left(\frac{\sigma_A I_A + \sigma_S I_S}{N}\right)(1 - W) - kW. \tag{6}$$

Infection trajectories. Our model also deviates from the traditional SEIR form by splitting the infectious compartment into an I_A-compartment (*A* for asymptomatic), and an I_S-compartment (*S* for symptomatic). The former represents an initial infectious stage (following the non-infectious, exposed stage), from which individuals will either move on to recovery directly (representing those individuals who experienced mild to no symptoms) or move on to the I_S-compartment (representing those with a more severe response). Finally, individuals in the I_S-compartment will either move on to recovery or death due to the infection. This splitting of the traditional infectious compartment is motivated by mounting evidence of asymptomatic transmission of SARS CoV-2 (17, 33-36). Thus, we consider two trajectories for the course of the disease, similar to those employed by (15): (1) $E \rightarrow I_A \rightarrow R$ and (2) $E \rightarrow I_A \rightarrow I_S \rightarrow R$ *(or death)*. More precisely, once in the *E* state, an individual will transition to the infectious state I_A, at a per-person rate of ε. A proportion *p* will move from I_A to the recovered state *R* (at a rate of p). A proportion $(1 - p)$ of individuals in the I_A state will develop more severe systems and transition to I_s (at a rate of $(1 - p)$). Individuals in the I_s state recover at a per-person rate of ν or die at a per-person rate μ_S. In each state, normal mortality of the individual occurs at the per-person rate μ and newly susceptible (*S*) individuals enter the population at a rate μN. The important differences between these two trajectories are in how likely an individual is to move down one path or another, how infectious individuals are (both for people and for the environment), how long individuals spend in each trajectory, and how likely death is along each trajectory.

Interactions between the environment and people. The model couples the environment and people in two ways: (1) people can deposit the infectious virus to surfaces (infecting the environmental reservoirs), and (2) people can become infected by interacting with these surfaces (infecting the people). Surfaces infect people through the β_W term (equations 1 and 2), a proxy for a standard transmission coefficient, corresponding specifically to the probability of successful infectious transmission from the environment to a susceptible individual (the full rate term being $\beta_W W \cdot S$). Hence, the β_W factor is defined as the fraction of people who interact with the environment daily, per fraction of the environment, times the probability of transmitting infection from surfaces

to people. The factor $\beta_W W$ (where W is the fraction of surfaces infected) represents the daily fraction of people that will interact with the infected portion of the environment and become infected themselves. The full term $\beta_W W \cdot S$ is thus the total number of infections caused by the environment per day.

In an analogous manner, we model the spread of infection *to* the environment with the two terms $\sigma_A I_A \cdot (1 - W) / N$ and $\sigma_S I_S \cdot (1 - W) / N$ representing deposition of infection to the environment by asymptomatic individuals, in the former, and symptomatic individuals, in the latter. In this case, σ_A (and analogously for σ_S) gives the fraction of surfaces/objects that interact with people at least once per day, times the probability that a person (depending on whether they are in the I_A or the I_S compartment) will deposit an infectious viral load to the surface/object. Thus, $\sigma_A I_A / N$ and $\sigma_S I_S / N$ (where N is the total population of people) represent the daily fraction of the environment that interacts with asymptomatic and symptomatic individuals, respectively. Lastly, the additional factor of $(1 - W)$ gives the fraction of surfaces/objects in the environment that have the potential for becoming infected, and so $\sigma_A I_A \cdot (1 - W) / N$ (and analogously for I_S) gives the fraction of the environment that becomes infected by people each day. We use W to represent a fraction of the environment, although we could have just as well multiplied the W equation by some value representing the total number of surfaces/objects in the environment (expected to remain constant throughout the course of the epidemic, assuming no intervention strategies).

Parameter values estimation. Table 1 displays information on the population definitions and initial values in the model**.** Tables 2 and 3 contain the fixed and estimated values and their sources (respectively). Because this model iteration is relatively underexplored with regards to COVID-19, we have worked to justify its use in various ways. The model's estimated parameters are based on model fits to 17 countries with the highest cumulative COVID-19 cases (of the 181 total countries affected) as of 03/30/2020, who have endured outbreaks that had developed for at least 30 days following the first day with ≥10 cumulative infected cases within each country

(37) (See Supplemental Tables S1 – S3). In addition, we compare country fits of the SEIR-W model to fits with a standard SEIR model. Lastly, we compare how various iterations of these mathematical models compare to one another with regards to the general model dynamics. For additional details, see the Supplemental Information.

Estimation of fixed parameters. There are 6 fixed parameters, 6 fitted parameters, and one parameter (□) dependent on the values of one of the fixed parameters (η) and one of the fitted parameters ($\boldsymbol{\varepsilon}$). These fixed parameters are η, μ, μ_S, ν, k, & p. The first, η, is the incubation period) (6, 38), and we assume that the expected time in the *E* state ($1/\boldsymbol{\varepsilon}$) and the expected time in the I_A state (1/□) sums to η, i.e. $\eta = 1/\boldsymbol{\varepsilon} + 1/$□. Fixing η constrains one of the two parameters, $\boldsymbol{\varepsilon}$ or □, and the other can be fitted; we choose to fit $\boldsymbol{\varepsilon}$ and therefore constrain □. The second fixed parameter μ, the normal death rate, was calculated by taking the reciprocal of the average life expectancy (in days) of the 17 countries sampled, weighted by population size. We calculated a value of 80.3 years, based on data from individual countries (39). The third parameter μ_S is the sum of the normal death rate and an additional death rate due to a more severe form of the infection. We assumed a death rate of 3.8% (38) and that death follows after initial symptoms between 3 and 4 weeks (38). Thus $\mu_S = \mu + 0.038/(3.5 * 7)$, where we use the average of 3 and 4 weeks and we convert to days with the factor of 7. The fourth fixed parameter ν, the recovery rate once in the symptomatic state, was assumed to be the reciprocal of the average of 3 and 6 weeks (the range of recovery times) (4, 38) times the fraction of individuals in the symptomatic state that do not die, i.e. 1 - 0.038, so $\nu = (1 - 0.038)/(4.5 * 7)$. The fifth fixed parameter k, the rate of viral decay in the environment, is the reciprocal of the average time that SARS-CoV-2 is expected to survive in the environment across a set of physical surfaces, based on the survival times across a handful of materials (19). The sixth fixed parameter *p*, the fraction of individuals in the I_A state that move on to recovery without experiencing severe symptoms, was taken to be 0.956 (2).

We fit our model variations to the *daily new cases data* provided (See: Supplemental Information) starting on the day when there were ≥10 cumulative infected cases in that

region. We choose the starting point of 10 cumulative cases in order to allow the outbreak to settle into a more consistent doubling time while also providing enough of an early-on window to capture the dynamics relevant to the $\mathcal{R}_0$ and force of infection estimations.

We calculate the number of daily new infections in our model by numerically integrating the influx rate of new symptomatic infections over the course of a single day (i.e. $\int (1 - p) \cdot I_A \, dt$). We perform this calculation for each of 30 consecutive days and fit these values to the daily new cases data. We use the influx rate of symptomatic infections, as opposed to the total rate of new infection (including asymptomatic individuals), as we expect that the large majority of reported cases in the early COVID-19 outbreak to be symptomatic. And we expect that—in the outbreaks—almost all asymptomatic cases go unreported (17).

Initial conditions. For each country, we use the first cumulative count that is ≥10 as a proxy for the initial number of *active* symptomatic cases I_{S0}. We can justify this by proposing that, given that the doubling time is expected to fall between 3 and 6 days (40) then the exponential growth rate parameter of the infection (*r in exp(rt)*) would fall between 0.231 days^{-1} and 0.116 days^{-1} respectively. And, assuming 1 initial infected individual, the time to reach 10 cases for the former rate would be about 10.0 days (~log(10)/0.231) and the time for the latter would be 19.8 days (~log(10)/0.116). Thus, since recovery of symptomatic individuals, which takes between 3 and 6 weeks (36, 38) exceeds this interval, we expect that at the point when 10 cases have accumulated, all cases are still active. Lastly, in fitting the data to our model, we initialize all fitting parameters to a value of 1.5, in whatever units are appropriate for that parameter (expected to be close to the true value for most of the fitting parameters).
As an estimate for the initial number active *asymptomatic* cases, we take $I_{A0} = I_{S0}$. That is, we expect that there are approximately as many asymptomatic cases as symptomatic cases early on. This assumption appears to be consistent with empirical findings. For example, data from the Diamond Princess cruise liner (11, 41), where all passengers were tested, revealed that approximately half of positively-testing cases

were asymptomatic. Lastly, we assumed that the initial number of exposed individuals was approximately $\mathcal{R}_0 \cdot (I_{A0} + I_{S0})$, based on the supposition that each of the initially infectious individuals $(I_{A0} + I_{S0})$ will have exposed the infection to approximately $\mathcal{R}_0$ other individuals. We take the value of $\mathcal{R}_0$ in this case to be 2.5, based on prior studies (15, 17). The *R* population is assumed to be 0 in the early stage of the outbreak, given the (average) 3 to 6-week recovery delay of COVID-19. The *W* population is assumed to be 1%.

Symbols	Initial Values	Units	Definitions	Sources
S_0	Varies by country	people	Susceptible individuals	Vary
E_0	$\mathcal{R}_0 \cdot (I_{A0} + I_{S0})$	people	Exposed individuals	Deduced
I_{A0}	I_{S0}	people	Asymptomatic individuals	Deduced
I_{S0}	Varies	people	Symptomatic individuals	Deduced
Rec_0	0	people	Recovered individuals	Deduced
W_0	1%	unitless	% of viruses in environment	Deduced

Table 1. *Model population definitions and initial values denoted with subscript 0 for each state variable*. Here we present definitions for the population groups represented by each compartment as well as their initial values. The initial value of the *S* and I_S populations vary by country, as shown in Table 2. We take the initial value of the I_A population to be the same as the initial value of symptomatic individuals as a conservative estimate. The initial value of the *E* population is computed by assuming that all initially-infected people ($I_{A0} + I_{S0}$) have exposed the virus to approximately $\mathcal{R}_0$ (≈ 2.5) other people.

Symbols	Values	Units	Definitions	Sources
μ	(80.3 x 365)	1/day	Natural Death Rate (Reciprocal of the average life expectancy of 17 countries sampled)	(37, 42)
μ_S	0.00159	1/day	Infected death rate	(4, 32)
η	5.5	days	SARS-CoV-2 Incubation Period	(38)
1/	η - ε^{-1}	days	Expected time in the asymptomatic state	Fitted and dependent on η
ν	0.031	1/day	Recovery rate (Average of 3 to 6 weeks)	(38)
p	0.956	unitless	raction that move along the “mild” recovery track	(2)
k	0.649	1/day	Viral decay rate in environment using average of all material values, wood, steal, cardboard, plastic)	(12)

Table 2. *Fixed parameter values estimated based on available published literature.* These estimated values derived from the existing COVID-19 and SARS-CoV-2 literature.

Symbols	Average values (SEIR-W)	Standard Deviation (SEIR-W)	Average values (SEIR)	Standard Deviation (SEIR)	Units	Definitions
β_A	0.550	0.345	0.429	0.751	1/day	(Contact rate of people with people) x (transmission probability of people to people by an asymptomatic person)
β_S	0.491	1.260	8.019	5.972	1/day	(Contact rate of people with people) x (transmission probability of people to people by *I*-person)
β_W	0.031	0.039	0.0	--	1/day	(Contact rate of person with environment) x (transmission probability of environment to people)
σ_A	3.404	6.662	0.0	--	1/day	(Contact rate of person with environment) x (probability of shedding by asymp.-person to environment)
σ_S	13.492	18.849	0.0	--	1/day	(Contact rate of person with environment) x (probability of shedding by symp.-person to environment)
$1/\varepsilon$	2.478	1.325	2.381	2.249	days	Average number of days before infectious

Table 3. *Estimated parameter values, averaged across countries.* Here we provide a table of the average values of the fitted parameters used in this model. These averages are taken across all of the selected 17 countries. See Supplementary Information for more details on country data and parameter estimation.

***Basic reproductive ratios* ($\mathcal{R}_0$).** We can express the $\mathcal{R}_0$ (eq. 7) in a form that makes explicit the contributions from the environment and from person-to-person interactions. In this way, the full $\mathcal{R}_0$ is observed to comprise two $\mathcal{R}_0$ sub-components: one the number of secondary infections caused by a single infected person through person-to-person contact alone (R_p) and the other is the number of secondary infections caused by exchanging infection with the environment (R_e).

$$\mathcal{R}_0 = \frac{R_p + \sqrt{R_p^2 + 4\,R_e^2}}{2} \tag{7}$$

where R_p and R_e are defined in equations 8a and 8b.

$$R_p = \frac{\varepsilon\,(\beta_A\,(\mu_S + \nu) + \beta_S(1-p)\,\omega)}{(\mu+\varepsilon)(\mu+\omega)(\mu_S+\nu)},\ \ R_e^2 = \frac{\varepsilon\,\beta_W\,(\sigma_A\,(\mu_S+\nu) + \sigma_S(1-p)\,\omega)}{k\,(\mu+\varepsilon)(\mu+\omega)(\mu_S+\nu)} \tag{8a, 8b}$$

Note that when $R_p = 0$, the $\mathcal{R}_0$ reduces to R_e and when $R_e = 0$, the $\mathcal{R}_0$ reduces to R_p. Thus, when person-to-person transmission is set to zero, the $\mathcal{R}_0$ consists only of terms associated with transmission from the environment, and when transmission from the environment is set to zero, the $\mathcal{R}_0$ consists only of infection directly between people. When both routes of transmission are turned on, the two $\mathcal{R}_0$-components combine in the manner in equation 7.

While R_e represents the component of the $\mathcal{R}_0$ formula associated with infection from the environment, the square of this quantity R_e^2 represents the expected number of *people* who become infected in the two-step infection process: people → environment → people, representing the flow of infection from people to the environment, and then from the environment to people. Thus, while R_p gives the expected number of people infected by a single infected person when the environmental transmission is turned off, R_e^2 gives the expected number of *people* infected by a single infected person by way of the environmental route exclusively (no direct person-to-person transmission).

Elaboration on formulas 8a-b—and associated derivation-discussions—appear in the Supplemental Information.

RESULTS

Evidence for environmental transmission in different countries. Using the Akaike information criterion (AIC), SEIR models with an environmental compartment (SEIR-W) provide a strong relative fit to country incidence data. As discussed in the **Methods**, we compared the performance of models with (SEIR-W) and without (SEIR) environmental transmission across multiple countries to assess the role of environmental transmission in different contexts. Using the fitted parameters provided in Tables 1-3, and S1-S3, we calculate AIC values for the two mechanistic models: the standard SEIR model and the SEIR-W model. Table S4 displays the summary of the AIC values for each model-type fit to the first 30 days after the first day with total counts ≥10. In 10/17 countries (including 9/11 European countries), the SEIR-W model provided a better fit to the country data. In Figure 2, we display the comparative individual country fit results for 4 of the countries with the fastest 30-day case growth rates—Spain, Italy, Iran, and Switzerland. The SEIR-W variant provides a better fit (significantly lower AIC score) than the standard SEIR model for all of these. Note that, as features of independent country epidemics are myriad and difficult to disentangle, several aspects independent of the model structure could explain the superior fit of the SEIR-W models. Results for additional country fits can be found in the Supplemental Information, Figure S1.

Environmental transmission modulates COVID-19 epidemiology. Partial Rank Correlation Coefficient (PRCC) analyses for the four examined features of the outbreak—(i) $\mathcal{R}_0$, (ii) total number of infected individuals after 30 days, (iii) time to peak number of infected individuals, and (iii) size of peak number of infected individuals. Figure 3 demonstrates the PRCC calculations for all four of these outbreak characteristics. For $\mathcal{R}_0$, we observe that the model was strongly sensitive to several aspects related to virus transmission— β_A, β_S, β_W——as well as the rate at which asymptomatic individuals develop symptoms (?), the rate of recovery (ν) and SARS-CoV-2 free-living survival rate (k).

One can also observe how some parameters are better suited to modify the peak of the infection, such as the recovery rate (ν; which includes in it the swiftness of diagnosing and treating the virus). Others modulate the timing of the peak, such as ε, the rate of leaving the "exposed" compartment (or equally well, the reciprocal of the average time spent in the exposed compartment). Note that across all features, the fraction of cases that move along the "mild" route (p), from $E \rightarrow I_A \rightarrow R$, has a powerful influence on all factors.

The SARS-CoV-2 $\mathcal{R}_0$ comprises person-to-person and environmental transmission. In the **Methods**, we described how the $\mathcal{R}_0$ is composed of two sub-$\mathcal{R}_0$ components, corresponding to different infectious interactions: person to person (R_p), person to environment and environment to person (R_e). Tornado plots were constructed that demonstrate how the $\mathcal{R}_0$-components have their own architecture and sensitivity (Figure 4).

In Figure 5, we observe how variation in free-living survival ($1/k$) influences four characteristics of a SARS-CoV-2 outbreak: $\mathcal{R}_0$, total number of infected individuals after 30 days, time to peak number of infected and symptomatic individuals, and maximum number of symptomatic individuals in the first 30 days. Note the annotations on the figure that highlight where the empirically-determined survival times of SARS-CoV-2 on a range of surfaces (copper, plastic, cardboard, stainless steel) (12). Also note that the quantitative relationships between $1/k$ and various outbreak features are slightly different. For example, the $\mathcal{R}_0$ increases more gradually across a wider range of free-living survival values than some of the other features (Figure 5).

Surface composition modulates outbreak dynamics. Figure 6 (simulations) depict the results of "surface world" simulations, where the k values correspond to those from a 2020 study highlighting the survival of SARS-CoV-1 and SARS-CoV-2 on different physical surfaces (12). The summary of these simulations (Figure 6) highlights that the surface composition of a setting has a meaningful impact on several features of

outbreak dynamics. “Copper world” (Figure 5a) takes the longest amount of time (88.4 days) to rise to the peak number of infected-symptomatic individuals, indicating an outbreak which is slower to develop. Relatedly, the $\mathcal{R}_0$ values are much different in the different “surface world” scenarios: The “copper world” simulation has an $\mathcal{R}_0$ of 2.4, while the plastic “surface world” simulation has an $\mathcal{R}_0$ of 3.18 (Figure 7 and Table S5). In addition, the total number of individuals infected after 30 days of the outbreak, and the total number dead after 30 days are both significantly lower in the “copper world” setting (Figure 7 and Table S5). The peak value of infected individuals is not dramatically different across “surface worlds.” That is, while many features associated with severity differ greatly across “surface world” settings, we observed significantly less variation in the peak of the epidemic as compared with the time to the peak of the epidemic (Table S5). Maybe the most noteworthy of the differences is the vast disparity in the number of deaths in the first 30 days of the outbreak, where the “plastic world” setting has more than 30 times the number of deaths as the “copper world” scenario (1,814 vs. 55, respectively; Figure 7 and Table S5).

DISCUSSION

In any emerging infectious disease, the degree of environmental transmission should be the focus of early inquiry. In this study, we introduce a model for SARS-CoV-2 transmission that rigorously incorporates environmental transmission, labeled SEIR-W. We demonstrate that the SEIR-W model is often superior to SEIR models with regards to fitting the empirical data for the early trajectory of outbreaks across the world. That the SEIR-W model fits certain country data relative to others may be the consequence of many characteristics of an epidemic (e.g. quality of data, testing capacity), none of which represents anything meaningful about the mechanism of an outbreak. However, given that features of environmental transmission can influence central properties of disease dynamics, we should consider the possibility that variation in environmental transmission may contribute to variation in aspects of disease dynamics. The model is strongly sensitive to several aspects related to virus transmission, including the rate at which both symptomatic and asymptomatic

individuals transmit infection to susceptible hosts, the rate at which asymptomatic individuals develop symptoms, the rate of recovery (ν) and SARS-CoV-2 decay rate (k).

Deconstructing the basic reproductive number ($\mathcal{R}_0$) into subcomponent reveals the role of environmental transmission. By deconstructing the basic reproductive number into components, we can better understand how variation in the $\mathcal{R}_0$—by setting, time, or geography—may reside in how these contexts are driven by environmental transmission. Many of these effects may be (as they are in this study) localized to one component of the $\mathcal{R}_0$, labeled R_e in this study. Notably, the R_e component is highly sensitive to the transmission interaction between people and the environment ($\boldsymbol{\beta}_w$), and the decay rate of virus in the environment (k). Interestingly, the R_e is relatively robust to the rate of infectious virus shed into the environment from the asymptomatic infected individuals (the parameter called σ_A in this model).

SARS-CoV-2 dynamics in different "surface world" settings resemble essentially different outbreaks. Analysis of the $\mathcal{R}_0$ and its subcomponents highlights that many aspects of outbreak dynamics are sensitive to the parameter associated with environmental decay rate (k in the model presented in this study). Analysis of hypothetical settings purely comprising substances of a certain kind ("surface world") fortifies the significance of free-living survival on physical surfaces and environmental transmission in outbreak dynamics. While our findings cannot speak to the outbreak dynamics in any particular setting in the real world, they do reveal that the surface composition of a setting can significantly influence the behavior of an outbreak. For example, The SARS-CoV-2 $\mathcal{R}_0$ in the "plastic world" simulation ($\mathcal{R}_0 = 3.18$) is over 1.3 times the $\mathcal{R}_0$ in the "copper world" simulation ($\mathcal{R}_0 = 2.4$). Many other differences between these outbreaks come as a consequence of the different $\mathcal{R}_0$ values. For example, the "plastic world" simulation reaches a peak number of symptomatic infectious individuals almost 1.7 times faster than the "copper world" simulation, and kills over 30 times more people in the first 30 days (1,814 deaths in "plastic world" vs. 55 deaths in the "plastic world").

These differences are so significant that they might be naively interpreted as completely different outbreaks early on. Note, however, that the maximum value of the infected-symptomatic populations are roughly equivalent across “surface worlds,” and so the influence of SARS-CoV-2 survival on physical surfaces (mediated by difference in free-living survival) doesn’t affect all aspects of outbreak dynamics equally.

As of April 15, 2020, the scientific community remains in the fact-finding phase of SARS-CoV-2 biology and COVID-19 understanding. A significant source of fear and speculation in the pandemic involves the plausibility that SARS-CoV-2 has undergone local adaptation in certain settings, translating to different epidemiological properties. While there is no currently convincing molecular or clinical support for local adaptation in SARS-CoV-2, our findings highlight how easy it is to conflate an environmental (or ecological) difference for a genetic one: the same virus, spreading in populations of identical size and behavior, differing only in the composition of physical surfaces where the virus can be transmitted through the environment, can have $\mathcal{R}_0$ values between 2.4 and 3.18, with early death rates 30 times apart.

These disparities may underlie the difficulty in predicting outbreaks of a single type from setting to setting, a property that has recently been captured by a concept called permutation entropy (43). Even more, settings composed of certain physical surfaces (plastic-like in our model) may be associated with phenomenon resembling a “superspreading” event, where individual variation in contagiousness can drive unusually large numbers of infections (44). Perhaps a better understanding of how, and on what surfaces, viral populations survive may one day improve the predictability of outbreak trajectories.

Public Health Implications. Our findings suggest that the effect of social distancing in ameliorating pandemics can be amplified by limiting interactions between susceptible individuals and surfaces, rather than solely interactions between individuals. Said differently, it is not enough that individuals remain separated from other individuals during a pandemic like COVID-19, but also that individuals remain protected from

surfaces where other infectious individuals may have interacted. Highlighting the importance of measures like mask wearing in stemming COVID-19 outbreaks.

ACKNOWLEDGMENTS
The authors would like to thank W. Turner, L. Smolin, S. Ramachandran, and D. Weinreich for seminar invitations where iterations of the ideas in this manuscript were discussed. ALM would like to acknowledge NIH U54GM115677 for funding support.

DATA AND CODE AVAILABILITY
Data are either available or the source is referenced in the main text and supplemental information. The code used for the analyses in this study is publicly available on Github: https://github.com/OgPlexus/Copperland

REFERENCES

1. Cowling B. J., and G. M. Leung, 2020 Epidemiological research priorities for public health control of the ongoing global novel coronavirus (2019-nCoV) outbreak. Eurosurveillance 25.
2. Ferguson N., D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, *et al.*, 2020 Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand
3. Li Q., X. Guan, P. Wu, X. Wang, L. Zhou, *et al.*, 2020a Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N. Engl. J. Med.
4. Guan W., Z. Ni, Y. Hu, W. Liang, C. Ou, *et al.*, 2020 Clinical characteristics of 2019 novel coronavirus infection in China. MedRxiv.
5. Wu J. T., K. Leung, M. Bushman, N. Kishore, R. Niehus, *et al.*, 2020 Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. Nat. Med. 26: 506–510. https://doi.org/10.1038/s41591-020-0822-7
6. Backer J. A., D. Klinkenberg, and J. Wallinga, 2020 Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. Eurosurveillance 25: 2000062.
7. Lauer S. A., K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, *et al.*, 2020 The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann. Intern. Med.
8. Linton N. M., T. Kobayashi, Y. Yang, K. Hayashi, A. R. Akhmetzhanov, *et al.*, 2020 Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. J. Clin. Med. 9: 538. https://doi.org/10.3390/jcm9020538

9. Tindale L., M. Coombe, J. E. Stockdale, E. Garlock, W. Y. V. Lau, *et al.*, 2020 Transmission interval estimates suggest pre-symptomatic spread of COVID-19. medRxiv 2020.03.03.20029983. https://doi.org/10.1101/2020.03.03.20029983
10. Liu Y., R. M. Eggo, and A. J. Kucharski, 2020 Secondary attack rate and superspreading events for SARS-CoV-2. The Lancet 395: e47.
11. Mizumoto K., K. Kagaya, A. Zarebski, and G. Chowell, 2020 Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Eurosurveillance 25: 2000180.
12. Van Doremalen N., T. Bushmaker, D. H. Morris, M. G. Holbrook, A. Gamble, *et al.*, 2020 Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. N. Engl. J. Med.
13. Santarpia J. L., D. N. Rivera, V. Herrera, M. J. Morwitzer, H. Creager, *et al.*, 2020 Transmission Potential of SARS-CoV-2 in Viral Shedding Observed at the University of Nebraska Medical Center. medRxiv.
14. Cai J., W. Sun, J. Huang, M. Gamber, J. Wu, *et al.*, 2020 Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China, 2020. Emerg. Infect. Dis. 26. https://doi.org/10.3201/eid2606.200412
15. Kissler S. M., C. Tedijanto, M. Lipsitch, and Y. Grad, 2020 Social distancing strategies for curbing the COVID-19 epidemic. medRxiv.
16. Prem, K., Liu, Y., Russell, T.W., Kucharski, A.J., Eggo, R.M., Davies, N., Flasche, S., Clifford, S., Pearson, C.A., Munday, J.D. and Abbott, S., 2020. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet Public Health*.
17. Li R., S. Pei, B. Chen, Y. Song, T. Zhang, *et al.*, 2020b Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science.
18. Kraemer M. U., C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, *et al.*, 2020 The effect of human mobility and control measures on the COVID-19 epidemic in China. Science.
19. Ferretti, Luca, et al. "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing." *Science* (2020).
20. Abad F. X., R. M. Pinto, and A. Bosch, 1994 Survival of enteric viruses on environmental fomites. Appl Env. Microbiol 60: 3704–3710.
21. Walther B. A., and P. W. Ewald, 2004 Pathogen survival in the external environment and the evolution of virulence. Biol. Rev. 79: 849–869.
22. Boone S. A., and C. P. Gerba, 2007 Significance of fomites in the spread of respiratory and enteric viral disease. Appl Env. Microbiol 73: 1687–1696.
23. Weber T. P., and N. I. Stilianakis, 2008 Inactivation of influenza A viruses in the environment and modes of transmission: a critical review. J. Infect. 57: 361–373.

24. Li S., J. N. Eisenberg, I. H. Spicknall, and J. S. Koopman, 2009 Dynamics and control of infections transmitted from person to person through the environment. Am. J. Epidemiol. 170: 257–265.
25. Bani-Yaghoub M., R. Gautam, Z. Shuai, P. Van Den Driessche, and R. Ivanek, 2012 Reproduction numbers for infections with free-living pathogens growing in the environment. J. Biol. Dyn. 6: 923–940.
26. Zhao J., J. E. Eisenberg, I. H. Spicknall, S. Li, and J. S. Koopman, 2012 Model analysis of fomite mediated influenza transmission. PloS One 7.
27. Breban R., 2013 Role of environmental persistence in pathogen transmission: a mathematical modeling approach. J. Math. Biol. 66: 535–546.
28. Cortez M. H., and J. S. Weitz, 2013 Distinguishing between indirect and direct modes of transmission using epidemiological time series. Am. Nat. 181: E43–E52.
29. Brouwer A. F., M. H. Weir, M. C. Eisenberg, R. Meza, and J. N. Eisenberg, 2017 Dose-response relationships for environmentally mediated infectious disease transmission models. PLoS Comput. Biol. 13: e1005481.
30. Miller-Dickson Miles D., Meszaros Victor A., Almagro-Moreno Salvador, and Brandon Ogbunugafor C., 2019 Hepatitis C virus modelled as an indirectly transmitted infection highlights the centrality of injection drug equipment in disease dynamics. J. R. Soc. Interface 16: 20190334. https://doi.org/10.1098/rsif.2019.0334
31. Meszaros V. A., M. D. Miller-Dickson, F. B.-A. Junior, S. Almagro-Moreno, and C. B. Ogbunugafor, 2020 Direct transmission via households informs models of disease and intervention dynamics in cholera. PLOS ONE 15: e0229837. https://doi.org/10.1371/journal.pone.0229837
32. Wang W., J. Tang, and F. Wei, 2020 Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. J. Med. Virol. 92: 441–447. https://doi.org/10.1002/jmv.25689
33. Bai Y., L. Yao, T. Wei, F. Tian, D.-Y. Jin, *et al.*, 2020 Presumed asymptomatic carrier transmission of COVID-19. Jama.
34. Rothe C., M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, *et al.*, 2020 Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. N. Engl. J. Med. 382: 970–971.
35. Tao Y., P. Cheng, W. Chen, P. Wan, Y. Chen, *et al.*, 2020 High incidence of asymptomatic SARS-CoV-2 infection, Chongqing, China
36. He X., E. H. Y. Lau, P. Wu, X. Deng, J. Wang, *et al.*, 2020 Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat. Med. 1–4. https://doi.org/10.1038/s41591-020-0869-5
37. Roser M., H. Ritchie, E. Ortiz-Ospina, and J. Hasell, 2020 Coronavirus Disease (COVID-19) – Statistics and Research. Our World Data.

38. World Health Organization. (2020). *Report of the WHO-China Joint Mission on coronavirus disease 2019 (COVID-19).* Retrieved from https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf
39. Central Intelligence Agency. (2020). The World Factbook. Retrieved from https://www.cia.gov/library/publications/the-world-factbook/geos/xx.html
40. Bar-On Y. M., A. Flamholz, R. Phillips, and R. Milo, 2020 SARS-CoV-2 (COVID-19) by the numbers, (M. B. Eisen, Ed.). eLife 9: e57309. https://doi.org/10.7554/eLife.57309
41. Russell T. W., J. Hellewell, C. I. Jarvis, K. van Zandvoort, S. Abbott, *et al.*, 2020 Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020. Eurosurveillance 25: 2000256. https://doi.org/10.2807/1560-7917.ES.2020.25.12.2000256
42. European Centre for Disease Prevention and Control (ECDC) 2020. COVID-19 Geographical Distribution Worldwide. Retrieved from https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide
43. Scarpino, Samuel V., and Giovanni Petri. "On the predictability of infectious disease outbreaks." *Nature communications* 10, no. 1 (2019): 1-8.
44. Lloyd-Smith J. O., S. J. Schreiber, P. E. Kopp, and W. M. Getz, 2005 Superspreading and the effect of individual variation on disease emergence. Nature 438: 355–359.
45. S. Cobey, Modeling infectious disease dynamics. *Science* (2020), doi:10.1126/science.abb5659.
46. E. T. Lofgren, M. E. Halloran, C. M. Rivers, J. M. Drake, T. C. Porco, B. Lewis, W. Yang, A. Vespignani, J. Shaman, J. N. Eisenberg, Opinion: Mathematical models: A key tool for outbreak response. *Proc. Natl. Acad. Sci.* **111**, 18095–18096 (2014).
47. O. Diekmann, J. A. P. Heesterbeek, M. G. Roberts, The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface*. **7**, 873–885 (2010).
48. J. J. Moré, in *Numerical analysis* (Springer, 1978), pp. 105–116.
49. H. T. Banks, M. L. Joyner, AIC under the framework of least squares estimation. *Appl. Math. Lett.* **74**, 33–45 (2017).
50. D. M. Hamby, A review of techniques for parameter sensitivity analysis of environmental models. *Environ. Monit. Assess.* **32**, 135–154 (1994).
51. M. D. McKay, R. J. Beckman, W. J. Conover, Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*. **21**, 239–245 (1979).

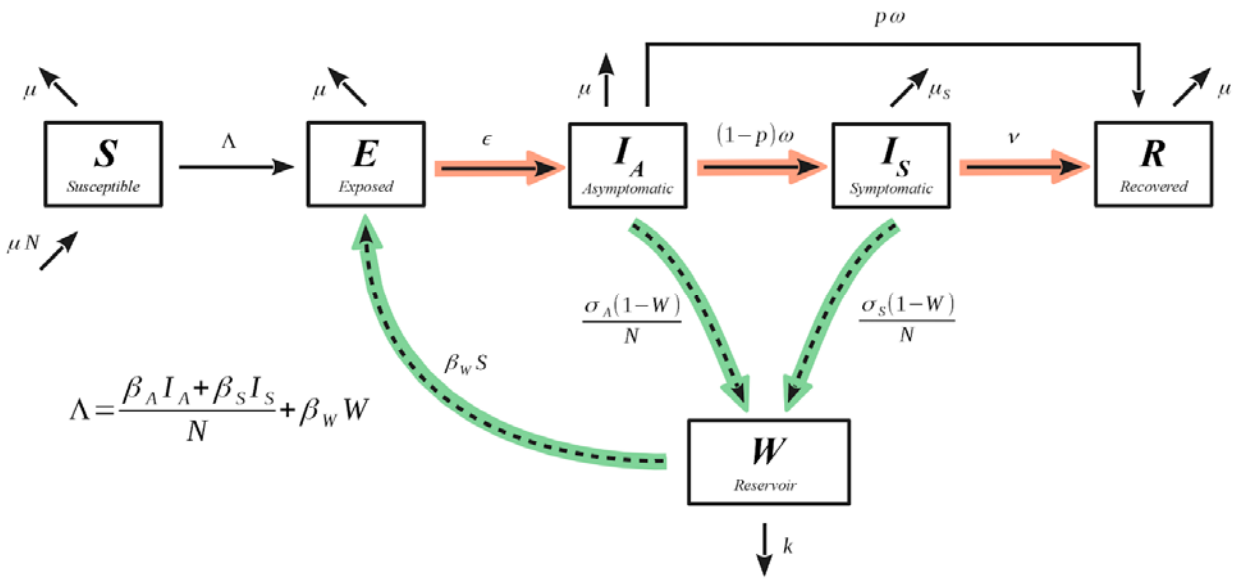

Figure 1. *SEIR-W model for COVID-19*. Compartmental diagram with dynamic information, where the green highlighted arrows represent how the infection couples with the environment, and the red highlighted arrows represent the progression of the infection through individuals. The model is referred to as SEIR-*W* throughout the text.

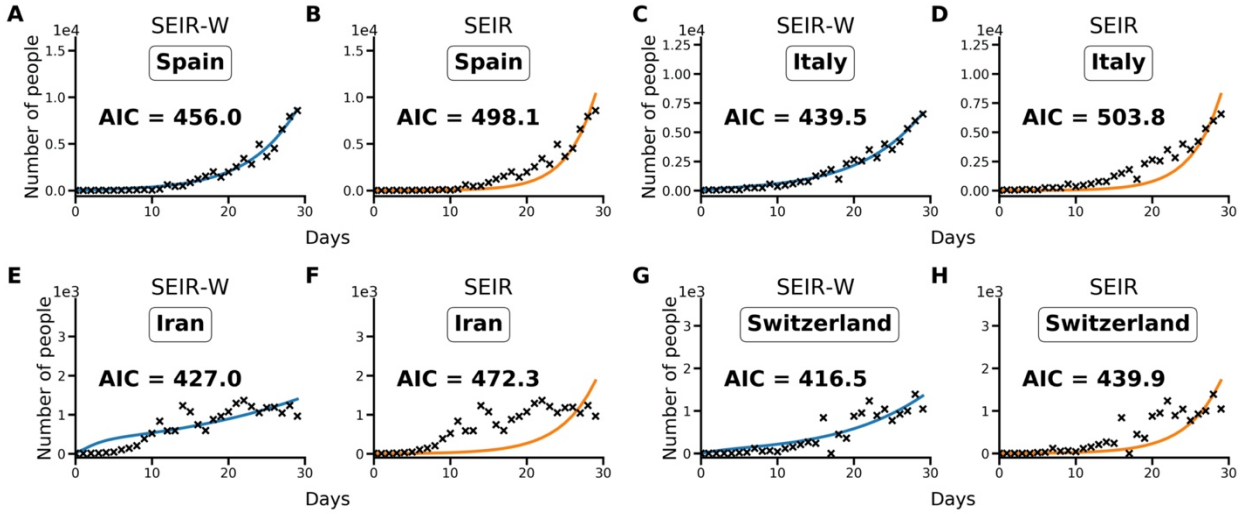

Figure 2. *Model fit comparisons for SEIR-W and standard SEIR to case counts in early windows of the outbreak*. The model fits are comparable across four countries with the largest early epidemics. These were chosen based having the highest cumulative number of infected cases after 30 days, following the first day when case counts were greater than or equal to 10. The four countries are (a,b) Spain, (c,d) Italy, (e,f) Iran and (g,h) Switzerland. These constitute a subset of 17 countries that had the highest number of cumulative COVID-19 cases (of the 181 total countries affected) as of March 30, 2020. Data come from the European Centre for Disease Control and Prevention, and from ourworldindata.org (37, 42).

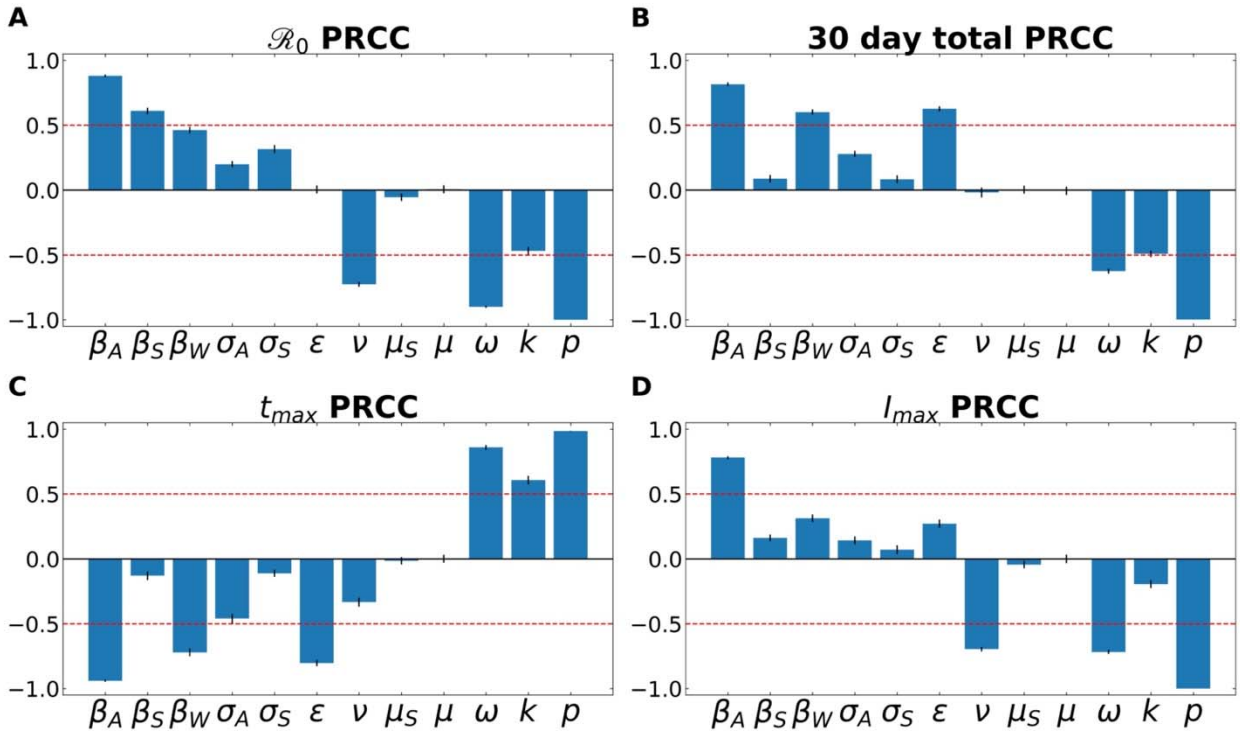

Figure 3. A Partial Rank Correlation Coefficient (PRCC) sensitivity analysis was performed with respect to (A) $\mathcal{R}_0$, (B) total number of infected (and symptomatic) after 30 days of outbreak, (C) time to peak number of symptomatic individuals (t_{max}), and (D) peak number of symptomatic individuals. This analysis highlights the intercorrelated sensitivities of each of the model parameters. The blue bars show the mean value of each PRCC, with error bars at one standard deviation. This analysis was performed by sampling over uniform distributions of 4.5% around the nominal model parameter values. Parameters correspond to the fixed ones in Table 3, and the average fitted parameters values in Table S5. The red line marks PRCC values of +/- 0.50 and helps identify parameters that are more influential (greater than 0.50 or less than - 0.50). See Supplemental Information for more details.

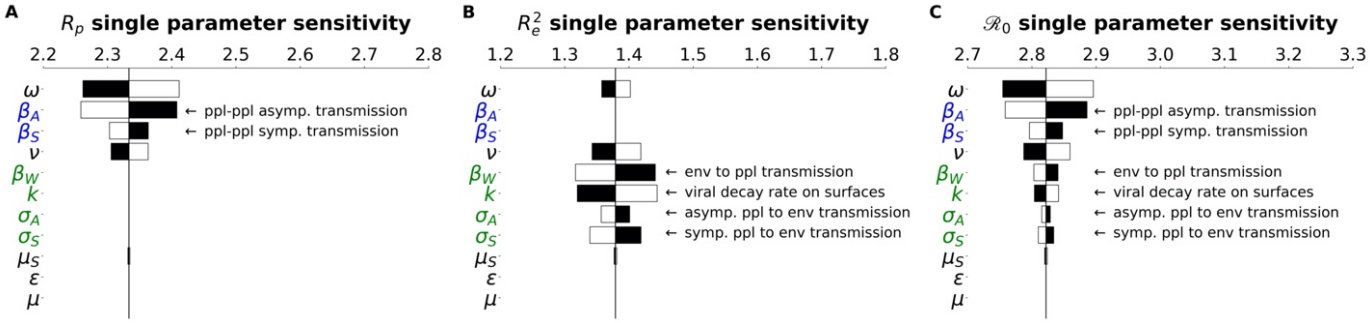

Figure 4. *$\mathcal{R}_0$ subcomponents have different parameter architecture*. We compare the parameter architecture for the two $\mathcal{R}_0$ components that compose the full $\mathcal{R}_0$ expression, (a) R_p, (b) R_e^2 and (c) $\mathcal{R}_0$. Parameters are colored according to their relation with the environment or people: green parameters refer to the environment, blue parameters strictly refer to people, and black parameters are neutral in this regard. Black bars show the extent to which the component after changed when the parameter values are *increased* by 4.5%, The white bars show the same except for a *decrease* of 4.5%. For clarity, the single parameter that most influences the $\mathcal{R}_0$ and its subcomponents is the faction of cases that move through the mild route (*p*) has been removed. For more details on how this parameter influences the $\mathcal{R}_0$ and other features of the outbreak, see the PRCC analysis as discussed in the Methods and Supplemental Information.

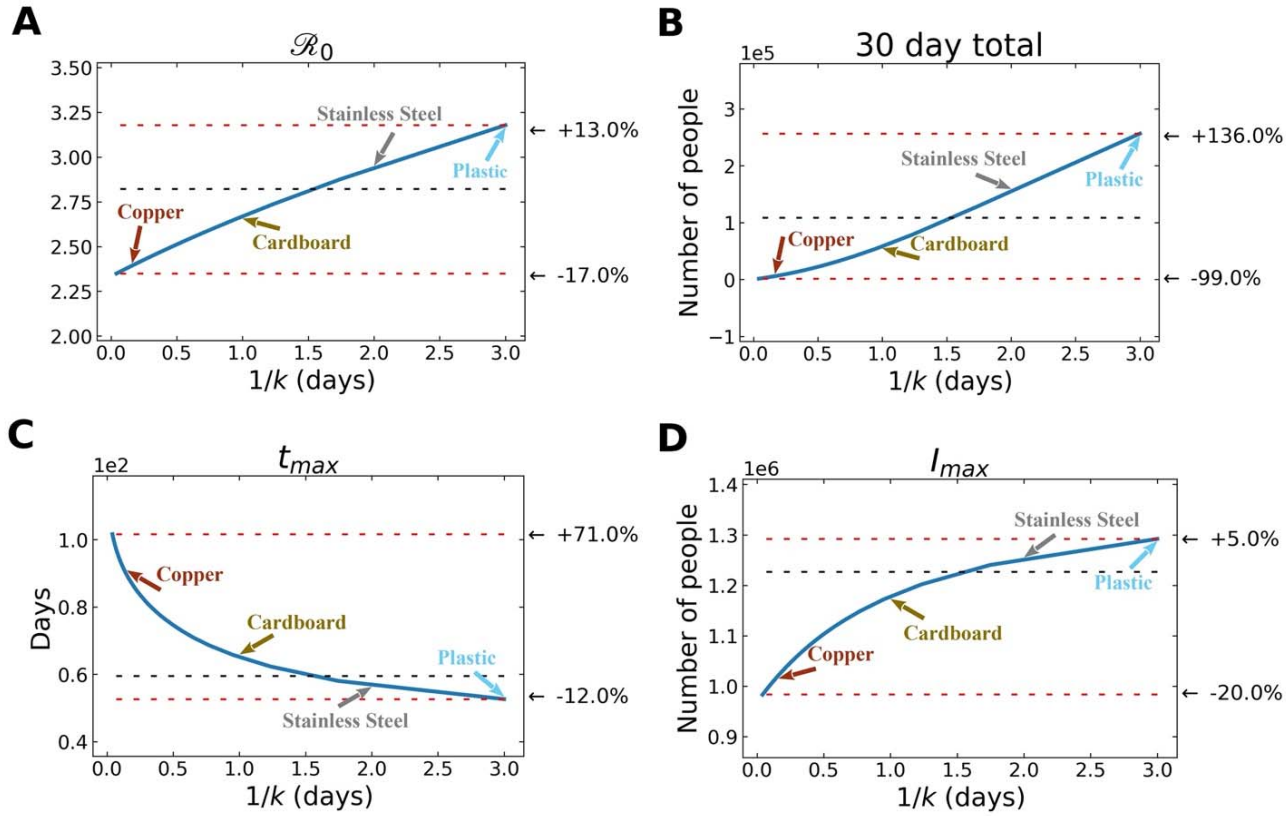

Figure 5. *Various features of an outbreak change as a function of* $1/k$ (where k is the rate of decay of SARS-CoV-2 survival in the environmental compartment): to (A) $\mathcal{R}_0$, (B) total number of infected (and symptomatic) after 30 days of outbreak, (C) time to peak number of symptomatic individuals (t_{max}), and (D) peak number of symptomatic individuals. The black dashed lines show the value of the respective plotted value at the average value of $1/k$ (~ 1.5 days), used in the fits from above. The top red line shows the maximum of plotted value for either the smallest value of $1/k$ chosen (= *1 hr)* or the largest value of $1/k$ chosen (= *3 days)*, depending on whether the plotted value decreases or increases with $1/k$, and the bottom red line shows the plotted value at the other extreme of $1/k$.

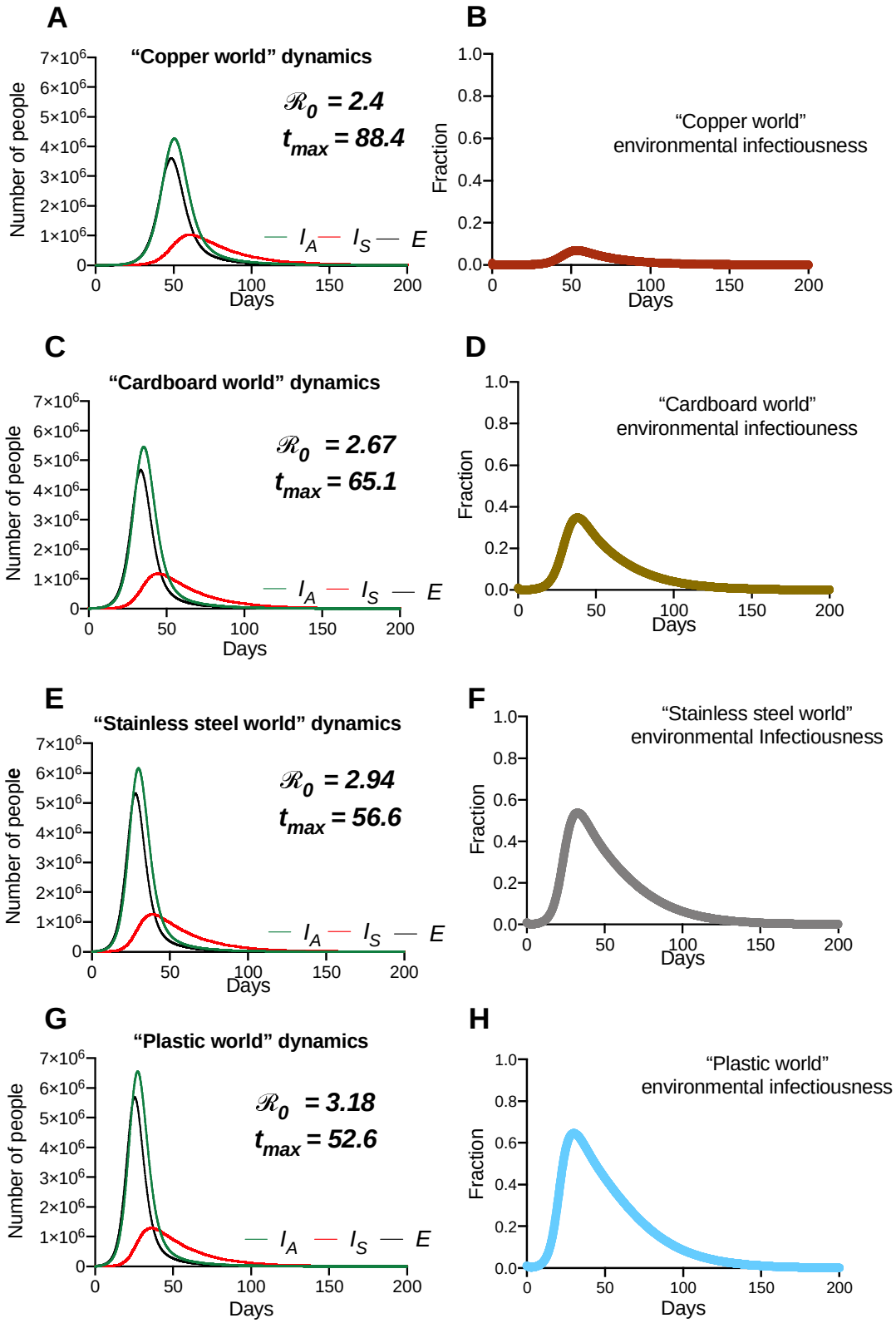
A
"Copper world" dynamics
$\mathcal{R}_0$ = 2.4
t_{max} = 88.4
Number of people
Days
I_A I_S E
B
"Copper world" environmental infectiousness
Fraction
Days
C
"Cardboard world" dynamics
$\mathcal{R}_0$ = 2.67
t_{max} = 65.1
Number of people
Days
I_A I_S E
D
"Cardboard world" environmental infectiouness
Fraction
Days
E
"Stainless steel world" dynamics
$\mathcal{R}_0$ = 2.94
t_{max} = 56.6
Number of people
Days
I_A I_S E
F
"Stainless steel world" environmental Infectiousness
Fraction
Days
G
"Plastic world" dynamics
$\mathcal{R}_0$ = 3.18
t_{max} = 52.6
Number of people
Days
I_A I_S E
H
"Plastic world" environmental infectiousness
Fraction
Days

Figure 6. *Hypothetical "surface world" simulations* feature differing dynamics. Population and environmental dynamics of SEIR-W model COVID-19 outbreaks in hypothetical settings composed of pure substances where SARS-CoV-2 can survive and be transmitted. (A,B) "copper world," (C,D) "cardboard world," (E,F) "stainless steel world" and (G,H) "plastic world." Environment infectiousness corresponds to the proportion of the environment that contains infectious SARS-CoV-2. Note that the surface where the viral decay is strongest (Copper), the peak of the epidemic is pushed farthest from the origin. Also note the $\mathcal{R}_0$ values graphs A, C, E, and G, which highlight that the different "surface worlds" behave like fundamentally different outbreaks in several ways.

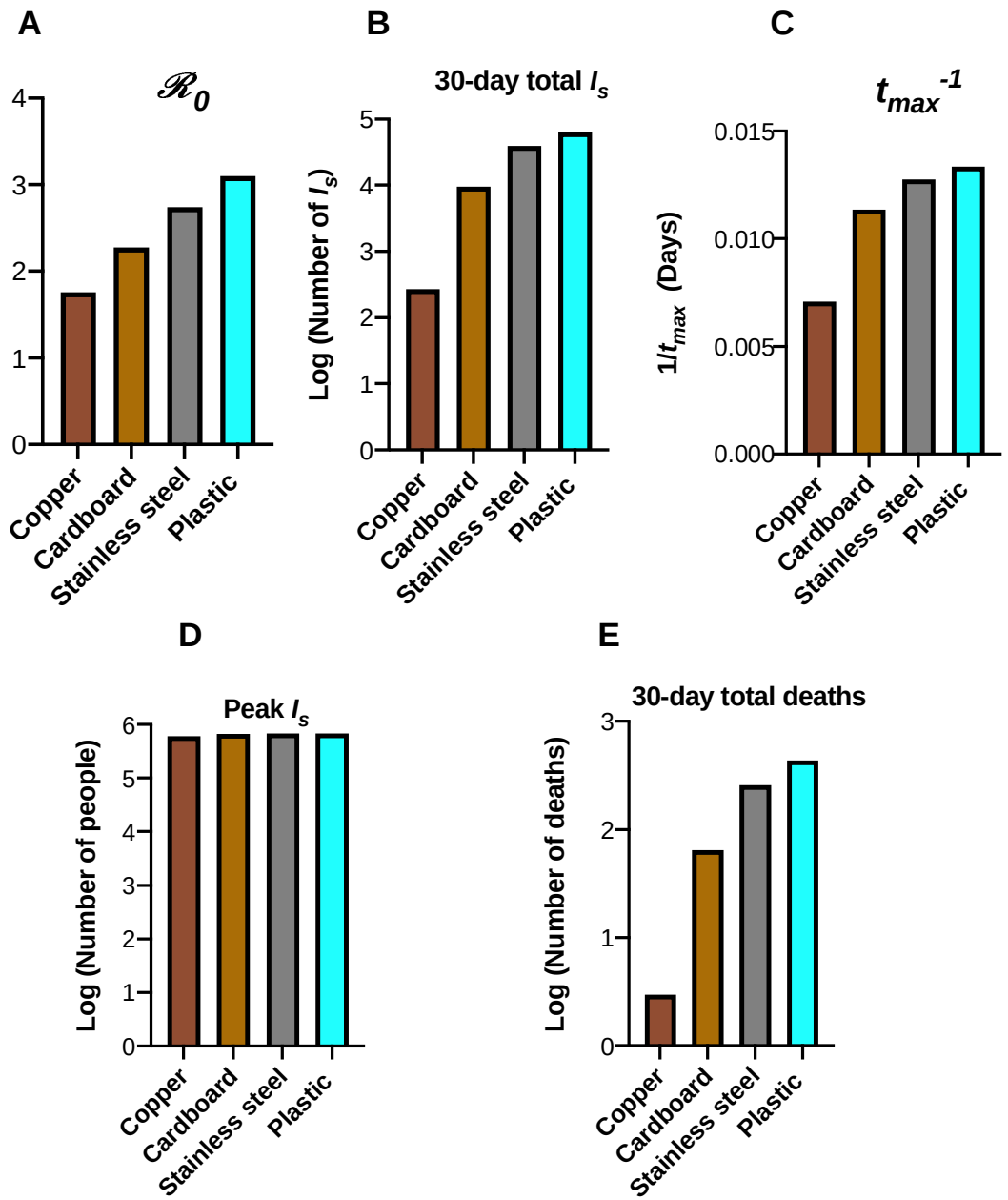

Figure 7. *Summary of the "surface world" outbreak intensity measures. Graphs* correspond to the attributes of simulated epidemics where environments are entirely composed of a given physical surface, and larger values correspond to various aspects of outbreak intensity. A) $\mathscr{R}_0$, (B) total number of infected (and symptomatic) after 30 days of outbreak, (C) the inverse time to peak number of symptomatic individuals (t_{max}^{-1}; larger values = shorter times to reach peak), (D) peak number of symptomatic individuals, and (E) deaths after 30 days. Note the log scales on the y-axis in (B), (D) and (E).