

Article

# The Epidemiological Signature of Pathogen Populations That Vary in the Relationship between Free-Living Parasite Survival and Virulence

Lourdes M. Gomez <sup>1,2,†</sup> , Victor A. Meszaros <sup>2,†</sup>, Wendy C. Turner <sup>3,‡</sup>  and C. Brandon Ogbunugafor <sup>1,2,\*</sup> 

<sup>1</sup> Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA; lourdes.gomez@yale.edu

<sup>2</sup> Department of Ecology and Evolutionary Biology, Brown University, Providence, RI 02912, USA; victor\_meszaros@brown.edu

<sup>3</sup> Department of Biological Sciences, University at Albany–State University of New York, Albany, NY 12222, USA; wcturner@albany.edu

\* Correspondence: brandon.ogbunu@yale.edu

† These authors contributed equally to this work.

‡ Current address: U.S. Geological Survey, Wisconsin Cooperative Wildlife Research Unit, Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI 53706, USA.

Received: 1 August 2020; Accepted: 17 September 2020; Published: 22 September 2020



**Abstract:** The relationship between parasite virulence and transmission is a pillar of evolutionary theory that has implications for public health. Part of this canon involves the idea that virulence and free-living survival (a key component of transmission) may have different relationships in different host–parasite systems. Most examinations of the evolution of virulence–transmission relationships—Theoretical or empirical in nature—Tend to focus on the evolution of virulence, with transmission being a secondary consideration. Even within transmission studies, the focus on free-living survival is a smaller subset, though recent studies have examined its importance in the ecology of infectious diseases. Few studies have examined the epidemic-scale consequences of variation in survival across different virulence–survival relationships. In this study, we utilize a mathematical model motivated by aspects of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) natural history to investigate how evolutionary changes in survival may influence several aspects of disease dynamics at the epidemiological scale. Across virulence–survival relationships (where these traits are either positively or negatively correlated), we found that small changes (5% above and below the nominal value) in survival can have a meaningful effect on certain outbreak features, including  $R_0$ , and on the size of the infectious peak in the population. These results highlight the importance of properly understanding the mechanistic relationship between virulence and parasite survival, as the evolution of increased survival across different relationships with virulence may have considerably different epidemiological signatures.

**Keywords:** virulence; free-living survival; disease dynamics; evolution of infectious disease

## 1. Introduction

Interactions between the life history of a pathogen and the environment in which it is embedded drive the evolution of virulence. These interactions thus dictate both the experience of disease at the individual host level and the shape of disease dynamics in host populations [1,2]. The nature of the interaction between virulence and transmission has been the object of both theoretical and empirical examination [2–8]. Free-living survival, here defined as the ability of a pathogen to persist outside

of its host, is one of many transmission life-history traits associated with virulence. The relationship between the two varies between host–pathogen types and different environments [4,8–10].

Several hypotheses serve as the canon in the evolution of virulence, theorizing its relationship with transmission traits. The Curse of the Pharaoh hypothesis—Named after a tale about a mythical curse that torments individuals who dig up tombs of Egyptian pharaohs [11]—Suggests that, if a parasite has high free-living survival, then it is far less dependent on its host for transmission and, consequently, will have no evolutionary incentive to decrease virulence [2,4,12]. The potential negative fitness consequences of killing hosts rapidly (being highly virulent) can be counteracted by persisting in the environment until the arrival of new susceptible hosts. Any presumptive selection on beneficence may be relaxed: parasites can detrimentally affect the health of hosts at no cost to transmission because most of their life cycle is spent outside of a host. Previous studies support a positive correlation between free-living survival and mortality per infection (a common proxy for virulence) [13].

Alternatively, the “tradeoff” hypothesis suggests that there is some intermediate level of parasite virulence [3,6,14] that is optimal for a given setting. In this scenario, too high a virulence kills the host and parasite and too low a virulence leads to failure to transmit. Applying this hypothesis specifically to free-living survival would suggest that selection for increased free-living survival should come at the expense of virulence (producing a pathogen that is less harmful to the host). Mechanistically, as a consequence of increased adaptation to a nonhost environment, a virus may be less fit to replicate inside a host [9,15]. For example, a more robust viral capsid may help to survive harsh environmental conditions but may make it more difficult to package RNA/DNA [15]. More generally, the tradeoff hypothesis can be framed in the context of a life-history tradeoff: investment in certain parts of the life cycle often comes at the expense of others [2,16].

Theoretical studies have explored varying evolutionary relationships between heightened virulence and extreme pathogen longevity [4,5,12,17–19]. One critical component of these studies revolves around whether virulence evolves independently of free-living survival. For example, some models have argued [4] that pathogen virulence is independent of survival under a set of conditions: when the host–pathogen system is at an equilibrium (evolutionary and ecological), if host density fluctuates around an equilibrium, or if turnover of the infected host population is fast relative to the pathogen in the environment. However, if the host–pathogen system is at disequilibrium and if the dynamics of propagules in the environment are fast compared to the dynamics of infected hosts, then virulence is, as hypothesized, an increasing function of propagule survival [4]. Kamo and Boots [17] examined this hypothesis by incorporating a spatial structure in the environment using a cellular, automata model and found that, if virulence evolution is independent of transmission, then long-lived infective stages select for higher virulence. However, if there is a tradeoff between virulence and transmission, there is no evidence for the Curse of the Pharaoh hypothesis, and in fact, higher virulence may be selected for by shorter rather than long-lived infectious stages. Further, the evolution of high virulence does not have to occur solely through a transmission–virulence tradeoff. Day [18] demonstrated how pathogens can evolve high virulence and even select for traits to kill the host (e.g., toxins) if pathogen transmission and reproductive success are decoupled. These studies emphasized the context-dependence of virulence–survival relationships. Understanding where in the relationship between virulence and survival a given pathogen population exists may allow one to understand how virus evolution will manifest at the level of epidemics.

In this study, we examine the epidemic consequences of different virulence–survival relationships—Positive and negative correlation—In a viral disease with an environmental transmission component. In order to measure how pathogen survival influences disease dynamics, we included an environmental compartment in our model, which represents contaminated environments that act as a reservoir for persisting pathogens, causing disease spread when they come in contact with susceptible individuals (infection via “fomites”) [20,21].

We find that the identity of the virulence–free-living survival relationship (e.g., positive vs. negative) has distinct implications for how an epidemic will unfold. Some, but not all, features of an

outbreak are dramatically influenced by the nature of the underlying virulence–survival relationship. This indicates that signatures for evolution (adaptive or other) in a pathogen population will manifest more conspicuously in certain features of an outbreak. We reflect on these findings in light of their theoretical implications on the evolution and ecology of infectious disease and for their potential utility in public health interventions.

## 2. Materials and Methods

### 2.1. Model Motivation and Application

The mathematical model explored in this study is adapted from a recent one developed to investigate environmental transmission of SARS-CoV-2 during the early-stage outbreak dynamics of coronavirus disease 2019 (COVID-19), with parameter values based on fits to actual country outbreak data [22]. In this study, we utilize this model to examine questions about the evolution of free-living survival. While the phenomenon we examine is a very relevant one that manifests in the real world, we want to emphasize that none of the methods or results in this study are intended to be applied to the current COVID-19 pandemic (as of September, 2020). This study is an attempt at responsible theoretical biology, with data-informed models and inferences that are germane to the natural world. However, neither do we support the extrapolation of these findings to any particular aspect of COVID-19 nor should they inform a policy or intervention. The model applies to a number of scenarios that include outbreaks in a naïve host population. This describes situations such as the evolution of novel viral lineages, viral spillover events, or host shifts, where a virus with a preexisting relationship between virulence and survival emerges in a population of new hosts. Another such scenario where this model applies is one where a virus has already emerged but evolves in a subpopulation in the novel hosts before a migration event of some kind introduces the evolved virus population to a fully susceptible population of hosts.

### 2.2. Model Description

The model is implemented via a set of ordinary differential equations, defined by Equations (1)–(6). It implements viral free-living survival via the “Waterborne Abiotic or other Indirect Transmission (WAIT)” modelling framework, coupling individuals and the pathogen within the environment [23,24].

Within the model, the  $\beta_w$  term allows for individuals to become infected via viral pathogen deposited in the environment and terms  $\sigma_A$  and  $\sigma_I$  allow asymptomatic and symptomatic individuals to deposit pathogens into the environment, respectively. Adapted from the more traditional SEIR (susceptible-exposed-infected-recovered) model, the SEAIR-W (susceptible-exposed-asymptomatic-infected-recovered-WAIT) model interrogates the consequences of the two hypotheses outlined above while representing the dynamics of a very relevant disease system (SARS-CoV-2) that includes an asymptomatic infectious population. While the importance of asymptomatic transmission was debated early in the pandemic, many studies have affirmed its role in the spread of disease [25–27]. Though environmental transmission of SARS-CoV-2 remains a controversial topic, it is plausible that asymptomatic individuals may spread disease through frequent contact with the environment, thus increasing the proportion of virus that is free-living [28]. We acknowledge that mathematical models of epidemics can be limited by “identifiability,” which can obfuscate the relative importance of some routes of transmission. In models that have both indirect and direct routes of transmission, it can be very difficult to conclude that one route is predominant [29–31].

$$\frac{dS}{dt} = \mu(N - S) - \left( \frac{\beta_A A + \beta_I I}{N} + \beta_W W \right) S \quad (1)$$

$$\frac{dE}{dt} = \left( \frac{\beta_A A + \beta_I I}{N} + \beta_W W \right) S - (\varepsilon + \mu) E \quad (2)$$

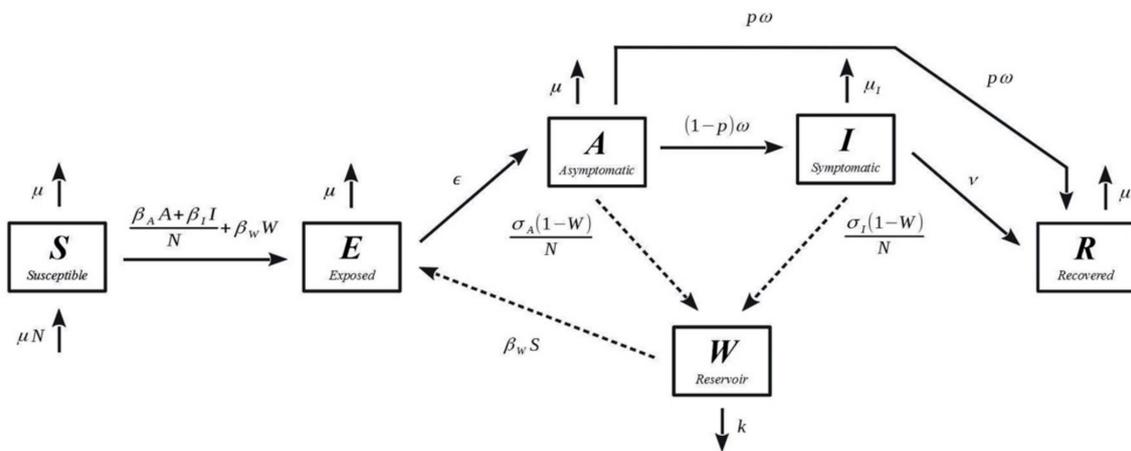
$$\frac{dA}{dt} = \epsilon E - (\omega + \mu)A \tag{3}$$

$$\frac{dI}{dt} = (1 - p)\omega A - (v + \mu_I)I \tag{4}$$

$$\frac{dR}{dt} = p\omega A + vI - \mu R \tag{5}$$

$$\frac{dW}{dt} = \left(\frac{\sigma_A A + \sigma_I I}{N}\right)(1 - W) - kW \tag{6}$$

Figure 1 depicts the compartmental diagram for the model. The direction of the arrows corresponds to the flow of the individuals and the pathogen through the system. Note that individuals can move directly from the asymptotically infected compartment to the recovered compartment (bypassing the symptomatic compartment) via what we call a “mild track”. The dashed arrows represent WAIT coupling to the environment. The model is inspired by one developed to interrogate environmental transmission of SARS-CoV-2 [22].



**Figure 1.** Compartmental diagram of the SEAIR-W (susceptible-exposed-asymptomatic-infected-recovered) version of a-WAIT (Waterborne Abiotic or other Indirect Transmission)) model: this is based on a previously developed mathematical model used to interrogate environmental transmission of SARS-CoV-2 (see [22]).

### 2.3. Simulations of Outbreaks

The system was numerically integrated using the “odeint” solver in the Scipy 1.4—Python scientific computation suite [32]. The simulations track the populations for each of the compartments listed in Figure 1. Each model run occurred over 250 days, which amounts to over 8 months of the epidemic or 5× the peak of the outbreak. This length of time is consistent with the antecedent SARS-CoV-2 model [22], long enough for the dynamics of the system to manifest. Note however that, for this study, we are especially interested in the early window of an outbreak: the first 30 days. We focus on this window because this is the time frame that best captures the underlying physics of an epidemic, as 30 days is often before populations are able to adjust their individual behaviors. The code constructed for the analysis in this study is publicly available on github: <https://github.com/OgPlexus/Pharaohlocks>.

### 2.4. Population Definitions and Parameter Values

Table 1 outlines the definitions of each population and provides the initial population values used for all simulations conducted in this study. The nominal parameter values used are defined in Table 2. The initial values are drawn from the aforementioned COVID-19 outbreak study, derived from empirical findings and country-level outbreak data [22].

**Table 1.** Definitions and initial values for the populations represented by each compartment: the values here are the averages of the model values across all countries developed in a prior COVID-19 study [22].

| Symbols | Value               | Units    | Definitions                        |
|---------|---------------------|----------|------------------------------------|
| $S_0$   | $57.05 \times 10^6$ | people   | Susceptible individuals            |
| $E_0$   | 66.50               | people   | Exposed individuals                |
| $A_0$   | 13.30               | people   | Asymptomatic individuals           |
| $I_0$   | 13.30               | people   | Symptomatic individuals            |
| $Rec_0$ | 0                   | people   | Recovered individuals              |
| $W_0$   | 1%                  | unitless | % of viral pathogen in environment |

**Table 2.** Definitions for the nominal parameter values used in this study: parameter values were developed from empirical findings and country-level data, as discussed in another study [22].

| Symbols                              | Values                    | Units   | Definitions  |
|--------------------------------------|---------------------------|---------|--|
| $\mu$                                | $1/(80.3 \times 365)$     | 1/day   | Natural death rate (reciprocal of the upper bound of average human lifespan)   |
| $\mu_I$                              | 0.00159                   | 1/day   | Infected death rate (natural death rate + disease-induced death rate)  |
| $\eta = (\omega - \varepsilon^{-1})$ | 5.5                       | days    | SARS-CoV-2 incubation period   |
| $1/\omega$                           | $\eta - \varepsilon^{-1}$ | days    | Expected time in the asymptomatic state  |
| $\nu$                                | 0.0305                    | 1/day   | Recovery rate (average of 3 to 6 weeks)  |
| $p$                                  | 95.6%                     | percent | Percent that moves along the “mild” recovery track   |
| $k$                                  | 0.649                     | 1/day   | Waning virus rate in the environment (using the average of all material values, wood, steel, cardboard, and plastic) |
| $\beta_a$                            | 0.550                     | 1/day   | Contact rate of people with people $\times$ transmission probability of people to people by A-person                 |
| $\beta_I$                            | 0.491                     | 1/day   | Contact rate of people with people $\times$ transmission probability of people to people by I-person                 |
| $\beta_W$                            | 0.031                     | 1/day   | Contact rate of person with environment $\times$ transmission probability of environment to people                   |
| $\sigma_a$                           | 3.404                     | 1/day   | Contact rate of person with environment $\times$ probability of shedding by A-people to environment                  |
| $\sigma_I$                           | 13.492                    | 1/day   | Contact rate of person with environment $\times$ probability of shedding by I-people to environment                  |
| $1/\varepsilon$                      | 2.478                     | days    | Average number of days before infectiousness   |

### 2.5. Virulence Definition

In this study, we define virulence as the capacity to cause a disease. In order to measure it, we utilize a set of parameters that uniformly increase the rate or probability of causing symptomatic disease or the severity of those symptoms (including death). Our definition is more comprehensive than many other models of parasite virulence (e.g., [4,13]), which tend to focus on a single aspect of the natural history of disease associated with harm to a host (e.g., the fitness consequences of an infection on the host population or the case fatality rate). Instead of having to justify a definition built around a single term (e.g., the term associated with fatality), we took a collective approach to defining virulence through all terms that foment the viral-induced onset of symptomatic disease and death. This definition allows for the reality of pleiotropic effects in viral pathogens, where adaptations can have multiple effects on the natural history of disease [2,33]. Our definition of virulence emphasizes terms that influence host wellness and/or are symptoms of disease. The iteration of virulence used in this study also undermines the potential for overly weighting only one or a small number of parameters under a large umbrella of virulence. Because so many varying definitions exist for virulence, we have also performed calculations according to a different definition of virulence, one that exclusively considers terms that have a detrimental direct effect on the host and neither of the terms that reflect symptoms of severe disease ( $\sigma_a$  and  $\sigma_I$ ). These calculations can be found in the Supplementary Materials.

The collection of parameters that we use to define virulence are as follows: the infected population death rate ( $\mu_I$ ), the incubation period of SARS-CoV-2 ( $\eta$ ), the rate of transfer from asymptomatic to symptomatic ( $1/\omega$ ), the infected population recovery rate ( $\nu$ ), the percent of individuals that move from the asymptomatic to the recovered compartment without showing symptoms (the “mild” recovery track,  $p$ ), the contact rate of people with people  $\times$  the transmission probability of people to people by

an asymptomatic individual ( $\beta_A$ ), the contact rate of people with people  $\times$  the transmission probability of people to people by an asymptotically infected person ( $\beta_I$ ), the contact rate of people with the environment  $\times$  the probability of shedding by an asymptomatic individual to the environmental ( $\sigma_A$ ), the contact rate of people with the environment  $\times$  the probability of symptomatically infected individuals shedding in the environment ( $\sigma_I$ ), and the average number of days before infection ( $1/\varepsilon$ ).

Table 3 outlines the direction in which each of the virulence-associated parameters are modulated as virulence decreases or increases. An up arrow ( $\uparrow$ ) indicates the parameter increases (by an equivalent percent) when the percent virulence is changed. A down arrow ( $\downarrow$ ) indicates the parameter decreases (by an equivalent percent) when the percent change in virulence is applied. Changes in virulence are then defined, in this study, as an equivalent uniform (percent) change in each of the parameters listed above. For the purposes of our study, we modify virulence by changing all parameters associated with virulence by 5%. One could also disambiguate virulence into changes in individual subcomponents; however, that is not the focus of this current study.

**Table 3.** Virulence parameters: this is a list of uniformly modulated parameters and the direction in which they change when virulence is increased or decreased. When virulence changes, an up arrow ( $\uparrow$ ) indicates the parameter is increased (by an equivalent percent) and a down arrow ( $\downarrow$ ) indicates the parameter is decreased (by an equivalent percent).

| Symbols                              | Definition  | Virulence Increased | Virulence Decreased |
|--------------------------------------|---|---------------------|---------------------|
| $\mu_I$                              | Infected death rate (natural death rate + disease induced death rate)                                 | $\uparrow$          | $\downarrow$        |
| $\eta = (\omega - \varepsilon^{-1})$ | SARS-CoV-2 incubation period  | $\downarrow$        | $\uparrow$          |
| $1/\omega$                           | Expected time in the asymptomatic state   | $\uparrow$          | $\downarrow$        |
| $\nu$                                | Recovery rate (average of 3 to 6 weeks)   | $\downarrow$        | $\uparrow$          |
| $p$                                  | Percent that moves along the “mild” recovery track  | $\downarrow$        | $\uparrow$          |
| $\beta_A$                            | Contact rate of people with people $\times$ transmission probability of people to people by A-person  | $\uparrow$          | $\downarrow$        |
| $\beta_I$                            | Contact rate of people with people $\times$ transmission probability of people to people by I-person  | $\uparrow$          | $\downarrow$        |
| $\sigma_A$                           | Contact rate of person with environment $\times$ (probability of shedding by A-people to environment) | $\uparrow$          | $\downarrow$        |
| $\sigma_I$                           | Contact rate of person with environment $\times$ probability of shedding by I-people to environment   | $\uparrow$          | $\downarrow$        |
| $1/\varepsilon$                      | Average number of days before infectious  | $\downarrow$        | $\uparrow$          |

## 2.6. Survival Definition

Survival is defined as the set of parameters that, when uniformly modulated, increases the pathogen’s probability of surviving the outside environment and successfully infecting a new host [2]. In our model, this includes both the waning virus rate in the environment ( $k$ ) and the contact rate of an individual with the environment  $\times$  the transmission probability of the environment to people ( $\beta_w$ ). Table 4 outlines the direction (increasing or decreasing) in which these parameters are modulated when survival is decreased or increased. Within both models, a (percent) change in survival is defined as an equivalent uniform (percent) change in the survival parameters.

Throughout this study, the impact of changes in virulence and survival (and the relationship between these traits) are assessed with respect to the following four epidemic metrics: the number of infected individuals (asymptomatic and symptomatic) at the maximum (when the outbreak is at its most severe), the rate at which the peak infected population is reached ( $t_{\max}^{-1}$ ), the total infected population after 30 days, and the basic reproductive ratio ( $R_0$ ). Importantly, among these signatures, the basic reproductive ratio is the most frequently used in epidemiology and benefits from familiarity and mathematical formalism (see Section 2.7). The other signatures are determined through simulations of an epidemic for a given set of parameter values. Nonetheless, this study’s inclusion of multiple features of the epidemic allows us to examine how variation in virus life-history traits may influence different aspects of an epidemic in peculiar ways.

**Table 4.** Survival parameters: this is a list of parameters that are uniformly modulated and the direction in which they change when survival is increased or decreased. An up arrow (↑) indicates the parameter is increased by some percent when the equivalent (percent) change in survival is applied. A down arrow (↓) indicates the parameter is decreased by some percent when the equivalent (percent) change in survival is applied.

| Symbols   | Definition   | Survival Increased | Survival Decreased |
|-----------|--|--------------------|--------------------|
| $k$       | Waning virus rate in environment (using the average of all material values, wood, steel, cardboard, and plastic) | ↓                  | ↑                  |
| $\beta_w$ | Contact rate of person with environment × transmission probability of environment to people                      | ↑                  | ↓                  |

### 2.7. Basic Reproductive Ratio

Equations (7)–(9) give the analytic expression of the basic reproductive ratio ( $R_0$ ) for the model used in this study. This expression for  $R_0$  can be deconstructed into two components. Equation (8) only contains parameters associated with person to person transmission ( $R_p$ ), while Equation (9) solely contains parameters associated with transmission from the environment ( $R_e$ ). In the Supplementary Materials, we provide additional information on these terms and their derivations. Applying the parameters values in Table 2, the numerical value of the basic reproductive ratio is given as  $R_0 \sim 2.82$ .

$$R_0 = \frac{R_p \sqrt{R_p^2 + 4R_e^2}}{2} \quad (7)$$

$$R_p = \frac{\varepsilon(\beta_A(\mu_I + v) + \beta_I(1-p)\omega)}{(\mu + \varepsilon)(\mu + \omega)(\mu_I + v)} \quad (8)$$

$$R_e^2 = \frac{\varepsilon\beta_w(\sigma_A(\mu_I + v) + \sigma_I(1-p)\omega)}{k(\mu + \varepsilon)(\mu + \omega)(\mu_I + v)} \quad (9)$$

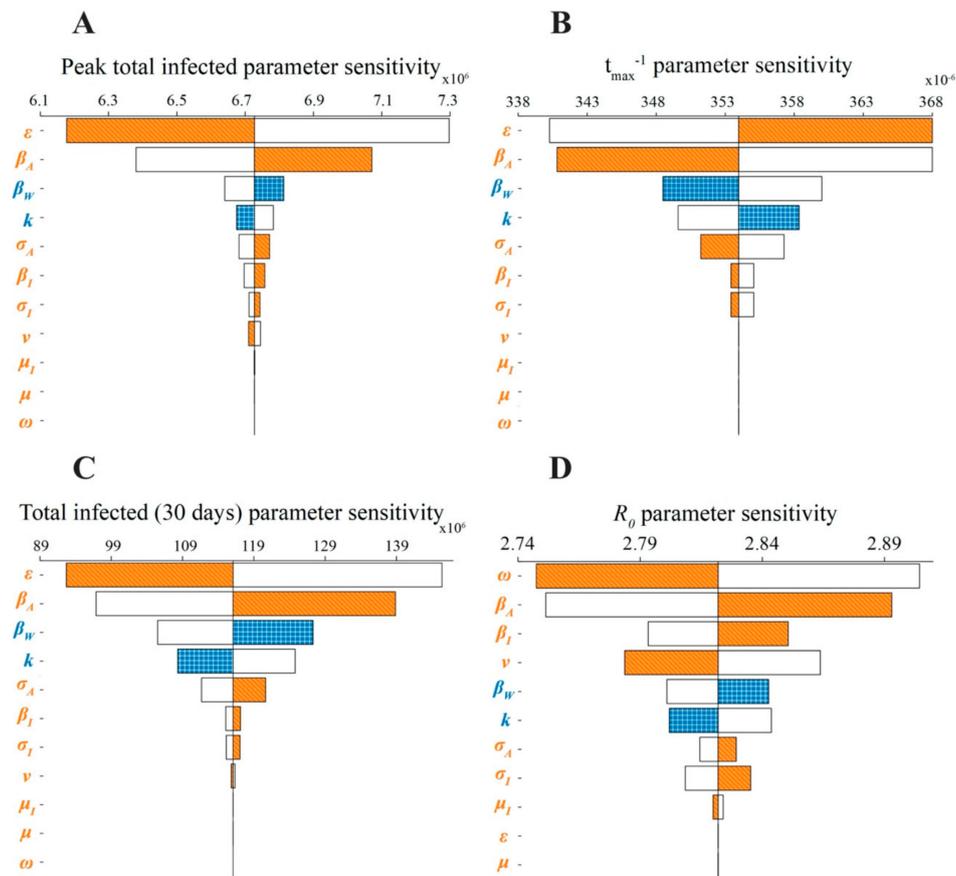
## 3. Results

### 3.1. Model Sensitivity Analysis

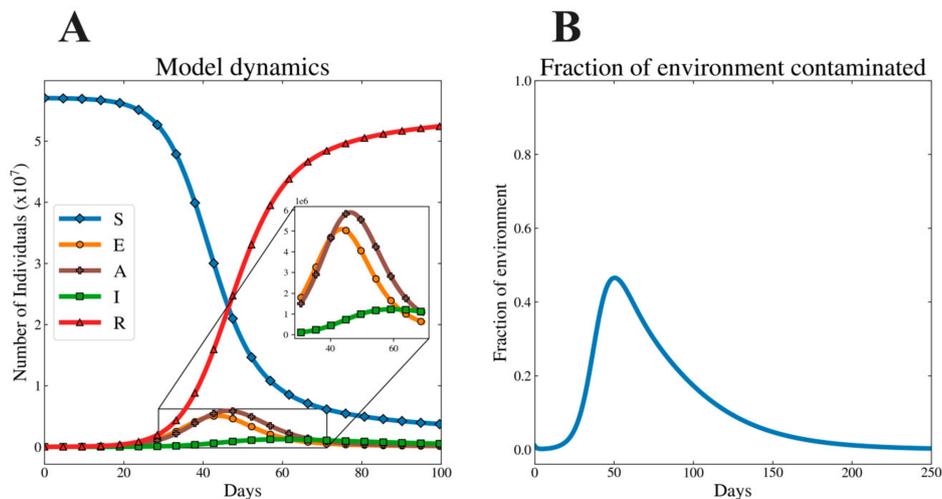
Figure 2 depicts a tornado plot that communicates the sensitivity of the model to permutations in parameters. Across features, the model is most sensitive to parameters that are considered virulence-associated (Table 3) and is relatively less sensitive to survival-associated parameters (Table 4). Similar to other features,  $R_0$  (Figure 2D) of the model is most sensitive to the parameters  $\omega$ ,  $\beta_A$ , and  $v$ . The sensitivity of  $R_0$  to changes in  $\omega$  reflects the importance of the rate of conversion to the symptomatic state on model dynamics. In addition,  $\beta_A$  has a very important influence on the model, consistent with other findings for COVID-19 that have emphasized the importance of asymptomatic transmission in disease spread [25–27].

### 3.2. Illustrative Dynamics of Model System

Based on the parameter values in Table 2, Figure 3A demonstrates the base dynamics of the model playing out over the first 100 days while Figure 3B shows the dynamics within the environment over the course of 250 days. In these dynamics, the population begins to be fixed for susceptible hosts. The disease dynamics manifest in the shapes of the curves corresponding to exposed, asymptomatic, and symptomatic individuals. Note the long tail of the curve corresponding to contamination by the environment. The environment remains infectious even after the infected populations have declined in number. The length and shape of this tail are influenced by the free-living survival of the virus.



**Figure 2.** Tornado plot showing the sensitivity of epidemic properties to individual parameter changes: (A) the number of infected individuals (asymptomatic and symptomatic) at the epidemic peak; (B) the rate at which the epidemic peak is reached,  $t_{\max}^{-1}$ ; (C) the total infected population after 30 days; and (D) the basic reproductive ratio ( $R_0$ ). Filled bars indicate the value of the epidemic feature when the associated parameter is increased by 5.0% from its nominal value. White bars indicate the value of a feature when the associated parameter is decreased by 5.0%. Blue coloring with checkered patterning indicates a parameter associated with survival, and orange coloring with lined patterning indicates a parameter associated with virulence.

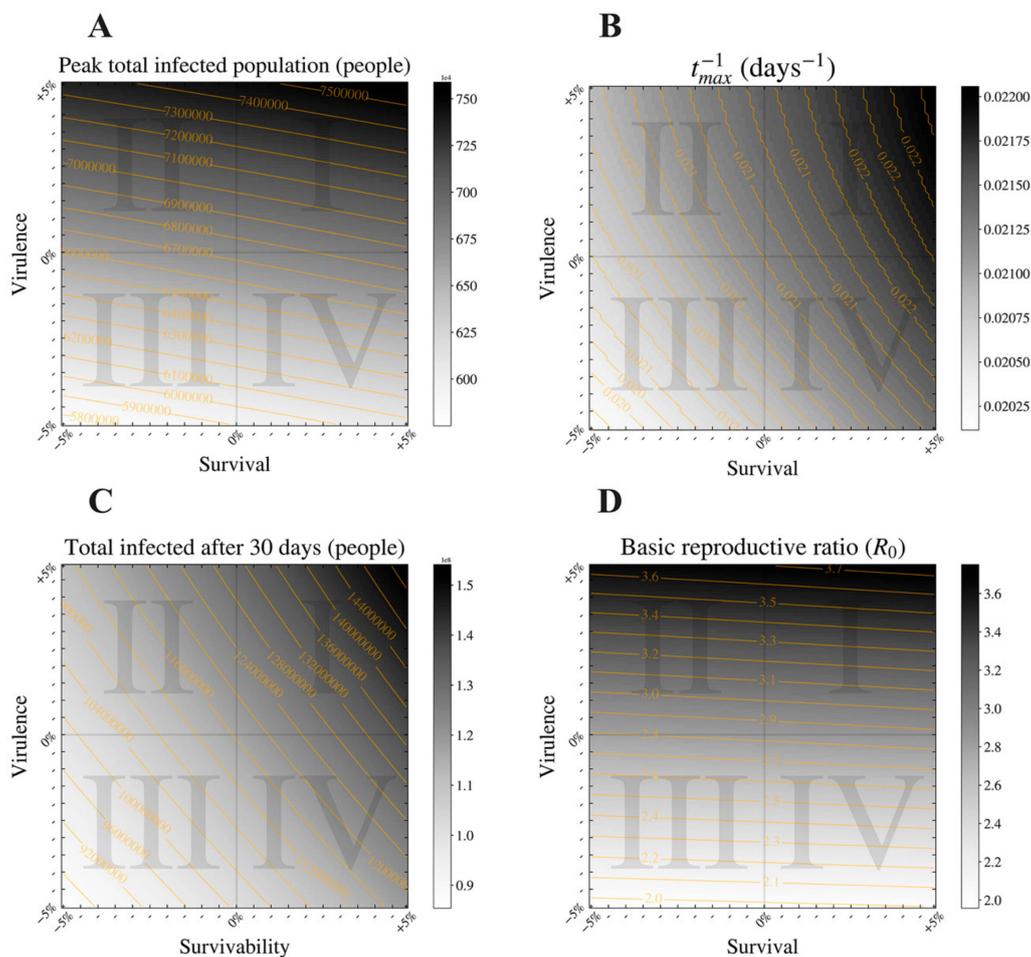


**Figure 3.** Sample dynamics for the model system: (A) the dynamics for all host compartments within the model and (B) the fraction of environmental reservoirs in a setting that are contaminated with infectious virus.

### 3.3. The Epidemic Consequences of Varying Virulence and Survival

In the next analysis, we examine the epidemic consequences of varying traits associated with survival and virulence. One can consider this as a scenario where we compare the endpoints of evolution of different virus populations (corresponding to combinations of values of survival and virulence) and calculating how these evolved populations manifest in epidemic terms.

In Figure 4, we observe how dynamics of the outbreak are influenced across a space of combinations of traits altering virulence (see Table 3 for a list of virulence-associated parameters) and survival (see Table 4 for a list of survival-associated parameters), changed by  $\pm 5\%$  (10% overall). In Figure 4D, we demonstrate how changes in virulence and free-living survival traits influence  $R_0$ , with variation in virulence-related traits having the largest effect on  $R_0$ . Of note is how the range in  $R_0$  values varies widely across virulence–survival values, from nearly 2.0 to 3.7 (Figure 4D).



**Figure 4.** Heatmap describing the impact of varying virus virulence and survival trait values assessed across four key epidemic metrics: these heatmaps express the change in (A) the number of infected individuals (asymptomatic and symptomatic) at the epidemic peak; (B) the rate at which the epidemic peak is reached,  $t_{max}^{-1}$ ; (C) the total infected population after 30 days; and (D) the basic reproductive ratio ( $R_0$ ) when virulence and survival are modulated by  $\pm 5\%$  within the model. Contour lines are available for clarity.

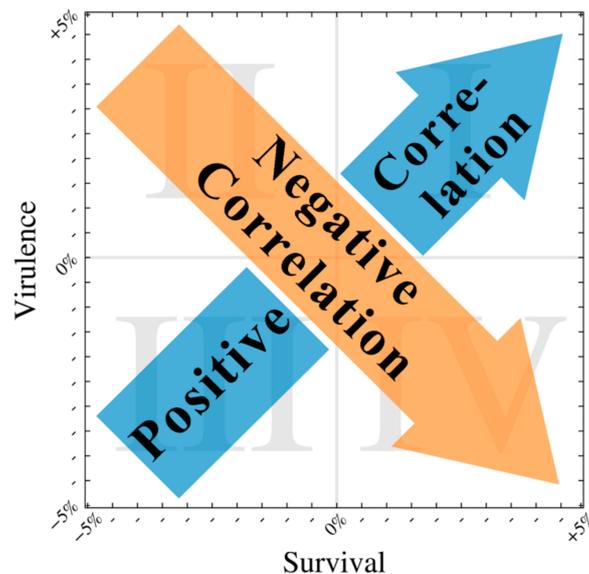
### 3.4. Implications of Virulence–Survival Relationships at Their Relative Extremes

Having observed how outbreak dynamics are influenced by variation in traits that alter virulence–survival phenotypes, we then examined how each outbreak metric is influenced by the extreme ( $\pm 5\%$ ) values of the trait combinations considered. Specifically, we assess how a change in

pathogen survival affects outbreak dynamics, based on two expected relationships between survival and virulence traits.

### 3.5. Positive Correlation Between Survival and Virulence

In a positive correlation scenario, high values for survival would be associated with high values for virulence [4,13]. Because the correlations we observe are often not exactly linear, we utilize quadrants to express a trend, allowing for some variance around the expected “line”. In Figure 5, the positive correlation scenario can be represented by combinations of virulence and survival residing in quadrants I and III.

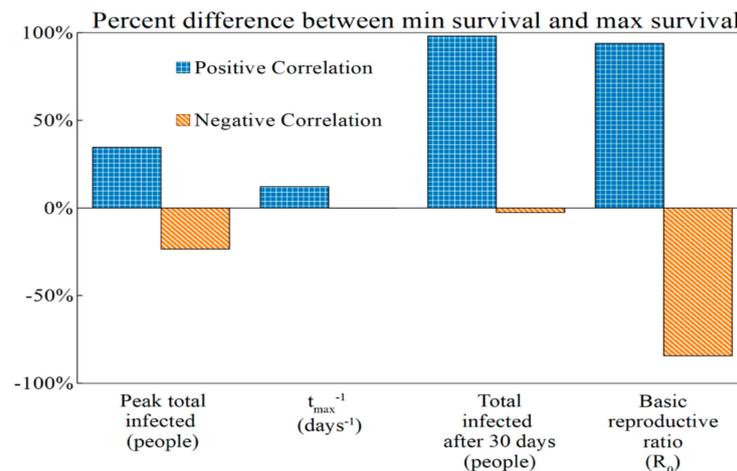


**Figure 5.** The expected effect of increasing survival on virulence for the two correlation models considered: here, we present a schematic of how the different hypotheses for the relationship between virulence and survival manifest on a map with a structure similar to the heat maps shown in Figure 4. The directions of the arrows depict how increasing survival would affect virulence under the two hypotheses: the blue arrow indicates the flow of an increasing positive correlation dynamic, while the direction of the orange arrow indicates an increasing negative correlation dynamic.

If host–pathogen evolution proceeds according to a positive correlation scenario, all outbreak metrics would show an increase in severity as both survival and virulence increase. Across the range of variation in virulence and survival traits considered (5% above and below the nominal value), the peak number of infected individuals increases by approximately 35%, the rate at which the peak is reached increases by approximately 16%, the total number of infected individuals after 30 days increases by approximately 98%, and  $R_0$  increases by approximately 94% (Figure 6 and Table 5).

**Table 5.** Positive correlation scenario: comparing epidemic metrics under low survival/low virulence versus high survival/high virulence scenarios (as in the positive correlation scenario). For each metric analyzed, these are the heatmap values for the bottom left (at “coordinates” (Vir, Sur)  $\rightarrow$  (−5%, −5%)) and top right (at “coordinates” (Vir, Sur)  $\rightarrow$  (+5%, +5%)) corners.

| Epidemic Metric                    | min Virulence,<br>min Survival | max Virulence,<br>max Survival | % Difference between<br>min Survival and max Survival |
|------------------------------------|--------------------------------|--------------------------------|---|
| Peak total infected (people)       | $5.68 \times 10^6$             | $7.64 \times 10^6$             | +34.51%   |
| $t_{\max}^{-1}$ (days $^{-1}$ )    | $1.99 \times 10^{-2}$          | $2.23 \times 10^{-2}$          | +12.06%   |
| Total after 30 days (people)       | $8.18 \times 10^7$             | $1.62 \times 10^8$             | +98.04%   |
| Basic reproductive ratio ( $R_0$ ) | 1.95                           | 3.78                           | +93.84%   |



**Figure 6.** The percent change in SEAIR-W outbreak metrics as survival increases from  $-5\%$  to  $+5\%$  for different virulence–survival relationships (positive correlation and negative correlation). For each metric analyzed, we present the percent difference between the minimum and maximum survival values given the two hypotheses tested: (i) positive correlation between survival and virulence (comparing low virulence/low survival to high virulence/high survival) and (ii) negative correlation (high virulence/low survival to high survival/low virulence). The bars here correspond to the values (percent) in the third columns of Tables 5 and 6, which denote the differences between the minimum and maximum values.

**Table 6.** Negative correlation scenario: comparing epidemic metrics under low survival/low virulence versus high survival/high virulence scenarios (as in the negative correlation scenario). For each metric analyzed, these are the global heatmap values for the top left (at “coordinates” (Vir, Sur)  $\rightarrow (+5\%, -5\%)$ ) and bottom right corners (at “coordinates” (Vir, Sur)  $\rightarrow (-5\%, +5\%)$ ).

| Epidemic Metric                       | max Virulence,<br>min Survival | min Virulence,<br>max Survival | % Difference between<br>min Survival and max Survival |
|---------------------------------------|--------------------------------|--------------------------------|---|
| Peak total infected (people)          | $7.39 \times 10^6$             | $5.99 \times 10^6$             | $-23.47\%$  |
| $t_{\max}^{-1}$ (days <sup>-1</sup> ) | $2.10 \times 10^{-2}$          | $2.23 \times 10^{-2}$          | $-0.15\%$   |
| Total infected after 30 days (people) | $1.16 \times 10^8$             | $1.13 \times 10^8$             | $-2.68\%$   |
| Basic reproductive ratio ( $R_0$ )    | 3.67                           | 1.99                           | $-84.39\%$  |

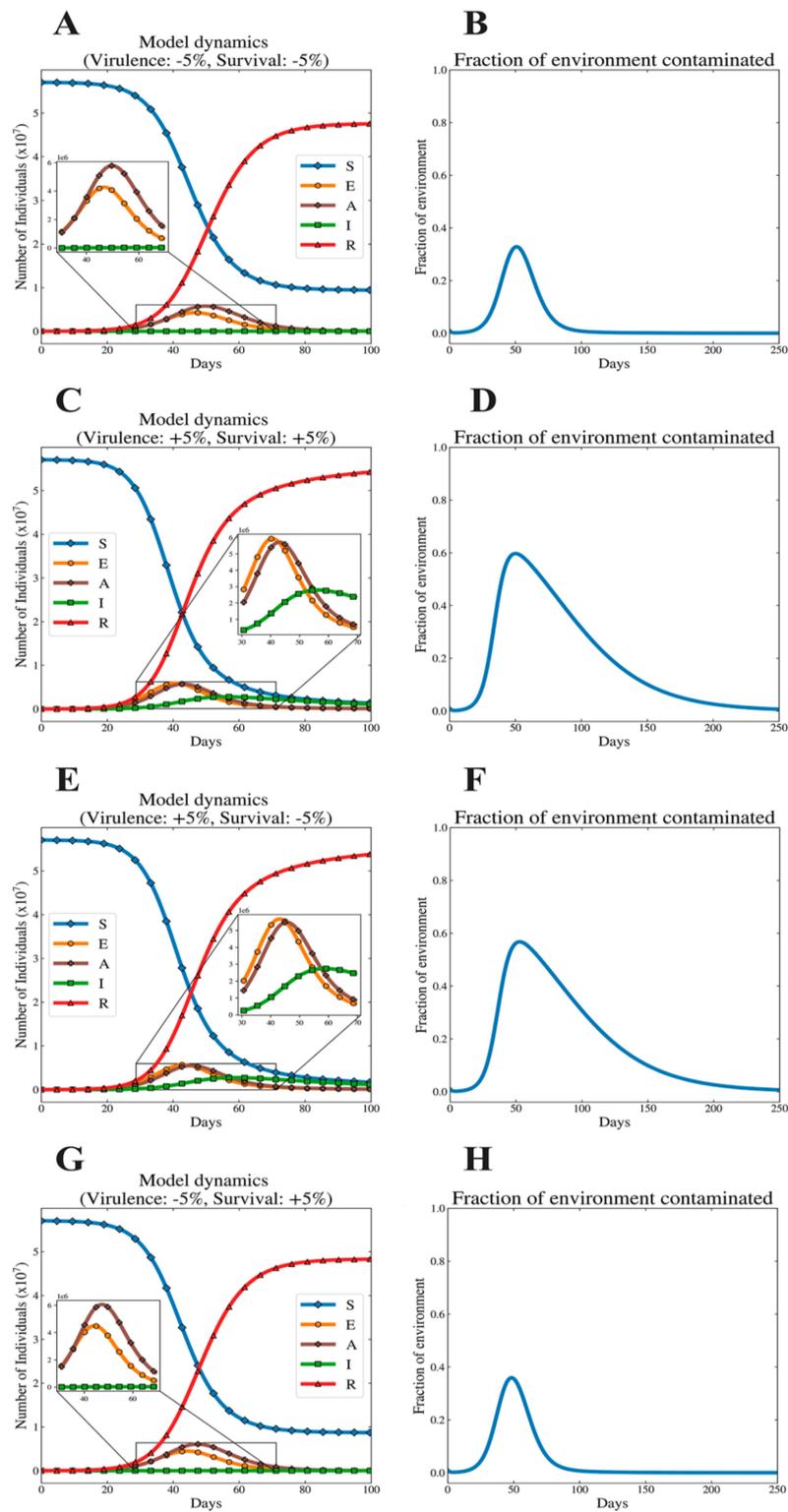
### 3.6. Negative Correlation Between Survival and Virulence

In a negative correlation scenario, high values for survival would be associated with low values for virulence [2,9,15] and a low peak in total infected population. Pathogens with a life history that exhibits negative virulence–survival associations would likely appear in quadrants II and IV in Figure 5.

Under negative correlation, outbreak severity decreases as survival increases. Across the measured range of variation in virulence–survival traits, the peak number of infected individuals decreases by approximately 23%, the rate at which the epidemic peak is reached decreases by 0.15%, the total number of infected individuals decreases by 3%, and  $R_0$  decreases by approximately 84% (Figure 6 and Table 6). Across all metrics considered, the effects of increased viral survival on outbreak dynamics is more extreme under the positive correlation than the negative correlation scenarios (Figure 6).

### 3.7. Dynamics of Epidemics at Extreme Values for Virus Free-Living Survival

In Figure 7, we observe the disease dynamics at extreme values for survival and the dynamics corresponding to the fraction of the environment that is contaminated with infectious virus. Consistent with the data represented in Figure 6, we observe that minimum and maximum simulations differ more substantially for extreme survival scenarios in the positive correlation scenario than for the negative correlation scenario.



**Figure 7.** Virus outbreak dynamics for the extreme values of virulence and free-living survival considered for the two different relationships (positive or negative) between virulence and survival traits: these plots are similar to the illustrative dynamics in Figure 1. Here, we observe the dynamics of disease corresponding to the extreme values presented in Tables 5 and 6. Subfigures (A,C,E,G) depict disease dynamics, and (B,D,F,H) depict the dynamics of contaminated environments. (A–D) correspond to the parameter values considered for the positive correlation scenario, while (E–H) correspond to the negative correlation scenario.

The feature of different outbreaks that varies most ostensibly between the correlation scenarios is the time needed to reach the peak number of infected individuals. In positive correlation simulations, one can observe that the low virulence, low survival scenario (Figure 7A,B) takes longer to reach the peak number of infected individuals. Most notably, however, the low virulence, low survival setting has a far smaller peak of environmental contamination and shorter tail relative to its high virulence, high survival counterpart (Figure 7D). Similarly, intriguing findings exist in the comparison between the simulation sets corresponding to extremes in the negative correlation setting (Figure 7E–H). Especially notable is the difference in the length of the tail of the environmental contamination for the high virulence, low survival combination (Figure 7F) vs. the low virulence, low survival combination variant (Figure 7H). The explanation is that, in this model, higher virulence influences (among many other things) the rate at which the virus is shed into the environment from either the asymptomatic ( $\sigma_A$ ) or symptomatic ( $\sigma_I$ ) host. We observe how the high virulence, low survival simulation (Figure 7E) features a symptomatic peak that is larger in size and is prolonged relative to the lower virulence counterpart (Figure 7G). This relatively large symptomatic population sheds infectious virus into the environment for a longer period of time, contributing to the long tail of contaminated environments observed in Figure 5F.

#### 4. Discussion

The virulence–survival relationship drives the consequences of virus evolution on the trajectory of an outbreak. In this study, we examined how different virulence–survival relationships may dictate different features of outbreaks at the endpoints of evolution (according to the positive or negative correlation scenarios). When the parameter space for virulence and survival is mapped, we find that certain outbreak metrics are more sensitive to change in free-living survival and virulence than others and that the nature of this sensitivity differs depending on whether survival and virulence are positively or negatively correlated.

For the positive correlation scenario, when free-living survival varies between 5% below and above the nominal value, we observed a dramatic change in the total number of infected individuals in the first 30 days (98% increase from minimum survival to maximum survival; Table 5), and  $R_0$  nearly doubles (94% increase; Table 5). These two traits are, of course, connected: the theoretical construction of the  $R_0$  metric specifically applies to settings where a pathogen spreads in a population of susceptible hosts [34,35], an early window that is captured in the first 30 days.

When survival and virulence are negatively correlated, different outbreak dynamics emerge: while the  $R_0$  difference between minimum and maximum survival is significant (approximately 84% decrease), the total number of infected individuals only changes by roughly 3% (Table 6). This large difference between  $R_0$  at higher and lower survival values also does not translate to a difference in the total number of infected individuals in the first 30 days of an infection (the early outbreak window). In a scenario where survival and virulence are negatively correlated, a highly virulent and less virulent virus population can have similar signatures on a population with respect to the number of infected individuals in the first month. Thus, simply measuring the number of infected individuals in the first month of an outbreak is unlikely to reveal whether a pathogen population has undergone adaptive evolution or has evolved in a manner that meaningfully influences the natural history of disease.

Notably, for scenarios where survival and virulence are both positively and negatively correlated, the time that it takes for an epidemic to reach its maximum number of infected individuals changes little across extreme values of survival (12% in the positive correlation scenario; 0.15% in the negative correlation scenario; see Tables 5 and 6). That is, the time that it takes for an epidemic to reach its peak (however high) is not especially sensitive to evolution in virulence or survival.

#### *Practical Implications for the Understanding of Outbreaks Caused by Emerging Viruses*

That different features of an outbreak are differentially influenced by the endpoints of viral life-history evolution highlights how epidemiology should continue to consider principles in the

evolution and ecology of infectious disease in its analyses and predictions. As not all features of an epidemic are going to be equally reliable signatures of virus evolution, we should carefully consider the data on how the dynamics of an epidemic change when making inferences about whether a pathogen population is essentially different from prior iterations (e.g., prior outbreaks of the same virus type). The results of this study suggest that carefully constructed, mechanistically sound models of epidemics are important, both for capturing the dynamics of an outbreak and for abetting our efforts to understand how evolution of survival and virulence influences disease dynamics.

For example, the potential for adaptive evolution of SARS-CoV-2 has emerged as a possible explanation for different COVID-19 dynamics in different countries. We suggest that such interpretations should be considered with caution and that they require very specific types of evidence to support them. As of 1 July 2020, any conclusion that widespread SARS-CoV-2 evolution is an explanation for variation in disease patterns across settings (space and/or time) is premature.

The practical process of interpreting the evolutionary consequences of signals of virus evolution should encompass several discrete steps. Firstly, we should determine whether molecular signatures exist for adaptive evolution. Adaptive evolution would manifest in observable differences in genotype and phenotype and, perhaps, in the natural history of disease. Secondly, we should aim to attain knowledge of the underlying mechanistic relationship between survival and virulence. This knowledge is not necessarily easy to attain (it requires extensive laboratory studies) but would allow added biological insight: we may be able to extrapolate how changes in some traits (e.g., those that compose survival) influence others (e.g., those that influence virulence).

More generally, our findings suggest that the ability to detect the consequences of virus evolution would depend on which feature of an outbreak an epidemiologist measures: from our analysis,  $R_0$  is most impacted by changes in virulence and survival. In addition, the total number of infected individuals in the early window and the size of the infected “peak” would each be impacted most readily by changes in virulence–survival traits. The rate at which the epidemic peak was reached, on the other hand, showed relatively little change as survival increased or between the two correlation scenarios. Consequently, it would not serve as a useful proxy for virus evolution.

While the stochastic, sometimes entropic nature of epidemics renders them very challenging to predict [36], we suggest that canons such as life-history theory and the evolution of virulence provide useful lenses that can aid in our ability to interpret how life-history changes in virus populations will manifest at the epidemiological scale. We propose that, in an age of accumulating genomic and phenotypic data in many pathogen–host systems, we continue to responsibly apply or modify existing theory in order to collate said data into an organized picture for how different components of the host–parasite interaction influence the shape of viral outbreaks of various kinds.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1999-4915/12/9/1055/s1>. Table S1. Fixed parameters, and their references. Table S2. Alternative virulence definition containing only items that directly affect host well-being (fitness). Figure S1. Alternative virulence definition: The impact of varying virus virulence and survival assessed on four key epidemic metrics. Figure S2. Alternative virulence definition: The percent change in SEAIR-W outbreak metrics as survival increases from −5% to +5% for different virulence–survival relationships (positive correlation and negative correlation). Figure S3. Alternative virulence definition: Virus outbreak dynamics for the extreme values of virulence and free-living survival considered for the two hypotheses. Table S3. Alternative virulence definition: Comparing epidemic metrics under low survival/low virulence versus high survival/high virulence scenarios. Table S4. Alternative virulence definition: Comparing epidemic metrics under low survival/low virulence versus high survival/high virulence scenarios (as in the negative correlation scenario). Figure S4. Side by side comparison of the main text virulence definition ( $A_1 \rightarrow D_1$ ) and alternative definition of virulence ( $A_2 \rightarrow D_2$ ). Figure S5. Side by side comparison of the main text virulence definition (left) and alternative definition of virulence (right). Figure S6. Side by side comparison of the main text definition ( $A_1 \rightarrow H_1$ ) and supplemental definition of virulence ( $A_2 \rightarrow H_2$ ).

**Author Contributions:** Conceptualization, L.M.G., W.C.T. and C.B.O.; methodology, V.A.M. and C.B.O.; software, V.A.M.; validation, V.A.M. and C.B.O.; formal analysis, L.M.G., V.A.M., W.C.T. and C.B.O.; investigation, L.M.G., V.A.M., W.C.T. and C.B.O.; resources, C.B.O.; data curation, L.M.G. and V.A.M.; writing—Original draft preparation, L.M.G., V.A.M., W.C.T. and C.B.O.; writing—Review and editing, L.M.G., V.A.M., W.C.T. and C.B.O.; visualization, V.A.M.; supervision, C.B.O. and W.C.T.; project administration, C.B.O.; funding acquisition, N/A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors would like to thank members of the OG Plexus for helpful discussions and input on all aspects of the manuscript. In particular, we thank M.M.D. and A.L.M. for especially helpful interactions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ewald, P.W. Host-parasite relations, vectors, and the evolution of disease severity. *Annu. Rev. Ecol. Syst.* **1983**, *14*, 465–485. [[CrossRef](#)]
2. Goldhill, D.; Turner, P.E. The evolution of life history trade-offs in viruses. *Curr. Opin. Virol.* **2014**, *8*, 79–84. [[CrossRef](#)] [[PubMed](#)]
3. Anderson, R.; Roy, M.; Robert, R.M. Trends and perspectives: Coevolution of hosts and parasites. *Parasitology* **1982**, *85*, 411–426. [[CrossRef](#)] [[PubMed](#)]
4. Bonhoeffer, S.; Lenski, R.; Ebert, D. The curse of the pharaoh: The evolution of virulence in pathogens with long living propagules. *Proc. R. Soc. B Biol. Sci.* **1996**, *263*, 715–721. [[CrossRef](#)]
5. Gandon, S. The curse of the pharaoh hypothesis. *Proc. R. Soc. B Biol. Sci.* **1998**, *265*, 1545–1552. [[CrossRef](#)] [[PubMed](#)]
6. Alizon, S.; Hurford, A.; Mideo, N.; van Baalen, M. Virulence evolution and the trade-off hypothesis: History, current state of affairs and the future. *J. Evol. Biol.* **2008**, *22*, 245–259. [[CrossRef](#)]
7. Cressler, C.E.; McLeod, D.V.; Rozins, C.; Hoogen, J.V.D.; Day, T. The adaptive evolution of virulence: A review of theoretical predictions and empirical tests. *Parasitology* **2015**, *143*, 915–930. [[CrossRef](#)]
8. Rafaluk-Mohr, C. The relationship between parasite virulence and environmental persistence: A meta-analysis. *Parasitology* **2019**, *146*, 897–902. [[CrossRef](#)]
9. Ogbunugafor, C.B.; Alto, B.W.; Overton, T.M.; Bhushan, A.; Morales, N.M.; Turner, P.E. Evolution of increased survival in RNA viruses specialized on cancer-derived cells. *Am. Nat.* **2013**, *181*, 585–595. [[CrossRef](#)]
10. Rafaluk, C.; Jansen, G.; Schulenburg, H.; Joop, G. When experimental selection for virulence leads to loss of virulence. *Trends Parasitol.* **2015**, *31*, 426–434. [[CrossRef](#)]
11. Corelli, M. *The Curse of Osiris*; Daily News: London, UK, 1923; p. 1.
12. Roche, B.; Drake, J.M.; Rohani, P. The curse of the Pharaoh revisited: Evolutionary bi-stability in environmentally transmitted pathogens. *Ecol. Lett.* **2011**, *14*, 569–575. [[CrossRef](#)] [[PubMed](#)]
13. Walther, B.A.; Ewald, P.W. Pathogen survival in the external environment and the evolution of virulence. *Biol. Rev.* **2004**, *79*, 849–869. [[CrossRef](#)] [[PubMed](#)]
14. Frank, S.A. Models of parasite virulence. *Q. Rev. Biol.* **1996**, *71*, 37–78. [[CrossRef](#)] [[PubMed](#)]
15. De Paepe, M.; Taddei, F. Viruses' life history: Towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS Biol.* **2006**, *4*, 1248–1256. [[CrossRef](#)] [[PubMed](#)]
16. Stearns, S.C. Trade-offs in life-history evolution. *Funct. Ecol.* **1989**, *3*, 259. [[CrossRef](#)]
17. Kamo, M.; Boots, M. The curse of the pharaoh in space: Free-living infectious stages and the evolution of virulence in spatially explicit populations. *J. Theor. Biol.* **2004**, *231*, 435–441. [[CrossRef](#)]
18. Day, T. Virulence evolution via host exploitation and toxin production in spore-producing pathogens. *Ecol. Lett.* **2002**, *5*, 471–476. [[CrossRef](#)]
19. Caraco, T.; Wang, I.-N. Free-living pathogens: Life-history constraints and strain competition. *J. Theor. Biol.* **2008**, *250*, 569–579. [[CrossRef](#)]
20. Barker, J.; Stevens, D.; Bloomfield, S. Spread and prevention of some common viral infections in community facilities and domestic homes. *J. Appl. Microbiol.* **2001**, *91*, 7–21. [[CrossRef](#)]
21. Boone, S.A.; Gerba, C.P. Significance of fomites in the spread of respiratory and enteric viral disease. *Appl. Environ. Microbiol.* **2007**, *73*, 1687–1696. [[CrossRef](#)]
22. Ogbunugafor, C.B.; Miller-Dickson, M.; Meszaros, V.A.; Gomez, L.M.; Murillo, A.L.; Scarpino, S.V. The intensity of COVID-19 outbreaks is modulated by SARS-CoV-2 free-living survival and environmental transmission. *medRxiv* **2020**. [[CrossRef](#)]
23. Miller-Dickson, M.D.; Meszaros, V.A.; Almagro-Moreno, S.; Ogbunugafor, C.B. Hepatitis C virus modelled as an indirectly transmitted infection highlights the centrality of injection drug equipment in disease dynamics. *J. R. Soc. Interface* **2019**, *16*, 20190334. [[CrossRef](#)] [[PubMed](#)]

24. Meszaros, V.A.; Miller-Dickson, M.D.; Baffour-Awuah, F.; Almagro-Moreno, S.; Ogbunugafor, C.B. Direct transmission via households informs models of disease and intervention dynamics in cholera. *PLoS ONE* **2020**, *15*, e0229837. [[CrossRef](#)] [[PubMed](#)]
25. Bai, Y.; Yao, L.; Wei, T.; Tian, F.; Jin, D.-Y.; Chen, L.; Wang, M. Presumed asymptomatic carrier transmission of COVID-19. *JAMA* **2020**, *323*, 1406. [[CrossRef](#)] [[PubMed](#)]
26. Park, S.W.; Cornforth, D.M.; Dushoff, J.; Weitz, J.S. The time scale of asymptomatic transmission affects estimates of epidemic potential in the COVID-19 outbreak. *Epidemics* **2020**, *31*, 100392. [[CrossRef](#)] [[PubMed](#)]
27. Tao, Y.; Cheng, P.; Chen, W.; Wan, P.; Chen, Y.; Yuan, G.; Chen, J.; Huo, D.; Guan, G.; Sun, D.; et al. High incidence of asymptomatic SARS-CoV-2 infection, Chongqing, China. *SSRN Electron. J.* **2020**. [[CrossRef](#)]
28. Santarpia, J.L.; Rivera, D.N.; Herrera, V.L.; Morwitzer, M.J.; Creager, H.M.; Santarpia, G.W.; Crown, K.K.; Brett-Major, D.M.; Schnaubelt, E.R.; Broadhurst, M.J.; et al. Aerosol and surface contamination of SARS-CoV-2 observed in quarantine and isolation care. *Sci. Rep.* **2020**, *10*, 1–8. [[CrossRef](#)]
29. Eisenberg, M.C.; Robertson, S.; Tien, J.H. Identifiability and estimation of multiple transmission pathways in cholera and waterborne disease. *J. Theor. Biol.* **2013**, *324*, 84–102. [[CrossRef](#)]
30. Cortez, M.H.; Weitz, J.S. Distinguishing between indirect and direct modes of transmission using epidemiological time series. *Am. Nat.* **2013**, *181*, E43–E52. [[CrossRef](#)]
31. Roosa, K.; Chowell, G. Assessing parameter identifiability in compartmental dynamic models using a computational approach: Application to infectious disease transmission models. *Theor. Biol. Med. Model.* **2019**, *16*, 1. [[CrossRef](#)]
32. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovsk, E.; Petersom, P.; Wekesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)] [[PubMed](#)]
33. Duffy, S.; Turner, P.E.; Burch, C.L. Pleiotropic costs of niche expansion in the RNA bacteriophage  $\Phi$  6. *Genetics* **2005**, *172*, 751–757. [[CrossRef](#)] [[PubMed](#)]
34. Jones, J.H. Notes on R0. *Calif. Dep. Anthropol. Sci.* **2007**, *323*, 1–19.
35. Diekmann, O.; Heesterbeek, H.; Roberts, M.G. The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface* **2009**, *7*, 873–885. [[CrossRef](#)]
36. Scarpino, S.V.; Petri, G. On the predictability of infectious disease outbreaks. *Nat. Commun.* **2019**, *10*, 898. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).