Infectious Disease Modelling 6 (2021) 560–583

Contents lists available at ScienceDirect

Infectious Disease Modelling

journal homepage: www.keaipublishing.com/idm

Effects of environmental variability on superspreading transmission events in stochastic epidemic models

Nika Shakiba [a, *], Christina J. Edholm [b], Blessing O. Emerenini [c, h], Anarina L. Murillo [d], Angela Peace [e], Omar Saucedo [f], Xueying Wang [g], Linda J.S. Allen [e]

[a] School of Biomedical Engineering, University of British Columbia, Vancouver, BC, Canada
[b] Department of Mathematics, Scripps College, Claremont, CA, USA
[c] Department of Mathematics, Oregon State University, Corvallis, OR, USA
[d] Department of Pediatrics and Center for Statistical Sciences, Brown University, Providence, RI, USA
[e] Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, USA
[f] Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
[g] Department of Mathematics and Statistics, Washington State University, Pullman, WA, USA
[h] School of Mathematical Sciences, Rochester Institute of Technology, Rochester, NY, USA

ARTICLE INFO

Article history:
Received 16 January 2021
Received in revised form 2 March 2021
Accepted 2 March 2021
Available online 18 March 2021
Handling editor: Dr. Daihai He

2020 MSC:
60H30
60J28
92D30

Keywords:
Superspreader
MERS
Ebola
Environmental variability
Stochastic models

ABSTRACT

Superspreaders (individuals with a high propensity for disease spread) have played a pivotal role in recent emerging and re-emerging diseases. In disease outbreak studies, host heterogeneity based on demographic (e.g. age, sex, vaccination status) and environmental (e.g. climate, urban/rural residence, clinics) factors are critical for the spread of infectious diseases, such as Ebola and Middle East Respiratory Syndrome (MERS). Transmission rates can vary as demographic and environmental factors are altered naturally or due to modified behaviors in response to the implementation of public health strategies. In this work, we develop stochastic models to explore the effects of demographic and environmental variability on human-to-human disease transmission rates among superspreaders in the case of Ebola and MERS. We show that the addition of environmental variability results in reduced probability of outbreak occurrence, however the severity of outbreaks that do occur increases. These observations have implications for public health strategies that aim to control environmental variables.

1. Introduction

Infectious diseases are one of the biggest threats to human life and a major concern to public health (Tatem et al., 2006). The spreading of emerging and re-emerging diseases, such as Middle East Respiratory Syndrome (MERS) and Ebola virus disease (Ebola), have been attributed to population mobility and the interconnectedness of the world's heterogeneous

* Corresponding author.
E-mail address: nika.shakiba@ubc.ca (N. Shakiba).
Peer review under responsibility of KeAi Communications Co., Ltd.

https://doi.org/10.1016/j.idm.2021.03.001

population (Tatem et al., 2006). Superspreader (SS) is a term used to informally describe high-risk individuals who disproportionately spread infection, because they have increased contacts with susceptible individuals, they are carriers of a more contagious strain, and/or experience a longer infection time (Leavitt, 1996). There are several factors that enable superspreading of infectious diseases, including immunological, physiological, and behavioural differences in individuals, the emergence of new pathogens, or exposure to a new environment (Stein, 2011). Findings from studies of Ebola underscore the importance of identifying SS individuals as quickly as possible to avoid an outbreak or epidemic (Althaus, 2015; Lau et al., 2017). However, the exact characteristics of SS and their impact on disease dynamics are not easy to identify given the variability of how each infection presents.

Although superspreading features prominently in first-hand narratives of disease transmissions, the characteristics and the dynamics of SS have not been systematically characterized, hindering refinements of future epidemic predictions and explorations of targeted interventions. In our previous study, we investigated the role of SS versus nonsuperspreaders (NS) on disease dynamics using both deterministic and stochastic epidemic models that incorporate demographic host heterogeneity (Edholm et al., 2018). A continuous-time Markov chain (CTMC) model was developed and parameterized for two case studies, MERS and Ebola. Branching process approximations and numerical simulations were used to compare disease epidemics initiated by SS and NS. Extensive numerical investigations were conducted in Edholm et al. (2018) to identify disease characteristics associated with the initial number of infected SS or the proportion of SS in the population, including probability of an outbreak, time to outbreak, peak number of infections, time to peak infection, and number of deaths. We defined an outbreak as the sum of exposed, asymptomatic, and infectious individuals ≥ 50. Our prior study demonstrated that the probability of an outbreak increases and the time to an outbreak decreases with the initial number of infected SS and the proportion of SS in the population. But after an outbreak occurred, the peak number of infections, and the time to peak infection did not change significantly with the proportion of SS when initiated either by one infected SS or one infected NS.

While our previous work explored the impact of demographic heterogeneity on disease dynamics, the field continues to lack an understanding of the impact of environmental variability on outbreak probability and severity. To date, models have focused on single environments to characterize point-source outbreaks or hospital-related transmission (Nardell et al., 1991; Pankhurst et al., 2012). Environmental variability has not been broadly factored into these models, in part due to logistical difficulties in characterizing contact patterns in various environments.

As a step toward improved assessment of environmental interventions, there is a need for models that in parallel investigate the effect of environmental and demographic heterogeneity on disease spread. The impact of environmental variability on the transmission rates of SS and NS will provide insights into theoretical contexts for identifying superspreaders as well as a framework for interpreting environmental data to inform environmental interventions (Li et al., 2009). A report from the World Health Organization identified some environmental factors that influence the spread of infectious diseases, including water supply, sanitation facilities, food, and climate (WHO, 2019). Eisenberg et al. found that the coupling of environmental and disease transmission processes provides a much-needed construct for furthering our understanding of both specific and general relationships between environmental change, infectious disease and their spread (Eisenberg et al., 2007). Hence, it is established that the environment can be a major factor in superspreading events. For example, crowding, poor ventilation, improper isolation procedures, unnecessary movement of the infectious, and misdiagnosis have all been identified as environmental factors that contributed to MERS (Chan et al., 2015) and Ebola (Glynn et al., 2017) outbreaks.

Building on our prior research, we extend our stochastic models for MERS and Ebola to investigate the effect of environmental variability in transmission rates on the probability of initiating an outbreak as well as the severity of outbreaks (Edholm et al., 2018). Accounting for this variability results in temporally correlated transmission parameters that vary randomly and continuously with the environment.

In the following sections, we review briefly some background on MERS and Ebola, before we formulate two new stochastic models that include demographic and environmental variability. We discuss advantages and disadvantages of the two stochastic modeling approaches. In addition, we summarize the results of our extensive numerical simulations comparing model outputs on designated measures: probability of an outbreak, time to the outbreak, number of deaths, peak number of infections reached in the outbreak, and time to the peak infection.

2. Background on MERS and Ebola

MERS coronavirus belongs to a family of enveloped, single-stranded RNA viruses (genus *Betacoronavirus*), that infect a number of different species, including humans, by predominantly causing mild self-limiting upper respiratory tract infections and other related illnesses. Other well-known viruses from the same genus include severe acute respiratory syndrome coronavirus (SARS-CoV) and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which cause the two diseases SARS and COVID-19, respectively (Peeri et al., 2020). MERS was first identified in 2012 from an outbreak in Saudi Arabia (Chan et al., 2015). MERS transmission arises from human-to-human interactions, where infected individuals can be asymptomatic or symptomatic (WHO, 2017). An outbreak in South Korea in 2015 with 166 cases was identified to be driven by three SS. The initial infected individual was a SS responsible for 29 secondary infections, of which two individuals were also SS who were responsible for 106 subsequent infections (WHO, 2017).

Ebola virus is a member of the *Filoviridae* virus family, known as one of the emerging and re-emerging zoonotic pathogens, causing acute hemorrhagic fever with high case fatality rate in humans (Beeching et al., 2014; Nii-Trebi, 2017). Ebola was first reported in 1976 in Africa during the Ebola outbreak in the Democratic Republic of the Congo. The transmission of Ebola virus

results from direct contact with bodily fluids containing virus particles from infected patients or diseased individuals, as well as from contact with contaminated objects. Epidemiological studies have revealed that family members of infected persons and associated health care workers are potential spreaders of the virus since they are likely to form a link between the infected persons and susceptible individuals. Though people can recover from Ebola infection the mortality rate is still very high (WHO, 2020).

3. Mathematical models

We extend our previous stochastic SEAIR model, which included demographic variability, to incorporate environmental variability in the disease transmission rate (Edholm et al., 2018). We did this by assuming an average transmission rate and applying a mean-reverting stochastic differential equation for both MERS (Fig. 1(a)) and Ebola (Fig. E1(a) in Appendix E). This led to the development of two novel models: a time-nonhomogenous stochastic process (NHP) model with discrete random variables, and a stochastic differential equation (SDE) model with continuous random variables. As with our previously published CTMC model (Edholm et al., 2018), these new models enabled us to track several key measures of disease dynamics, including: time to the outbreak, number of deaths, peak number of infections reached in the outbreak, time to the peak infection, as well as probability of an outbreak (when total infections in the population exceeds an outbreak threshold of 100) (Fig. 1(b)). Our models do not account for the emergence of new pathogens, which would have allowed for NS individuals to become SS depending on the pathogenic variance. Rather, they focus on physiological and behavioral causes of superspreading.

3.1. MERS models

The variables S_i, E_i, A_i, I_i and R_i denote the number of susceptible, exposed, asymptomatic, infected and recovered individuals of class i, respectively, where $i = 1$ is NS (nonsuperspreader) and $i = 2$ is SS (superspreader). The underlying ODE framework for MERS can be seen in the compartmental flow diagram in Fig. 1(a). The differential equations for the ODE model are given in Appendix A. Parameter N is the total population size and $\beta_i(t)$ is the time-dependent transmission rate for $i = 1, 2$. If $\beta_i(t) = \beta_i$ is constant, then the basic reproduction number for the ODE and CTMC models is

$$\mathcal{R}_0 = \underbrace{\frac{\beta_1 \frac{N_1}{N}(\gamma_1 + \delta_1 + \mu_{I_1})}{(\delta_1 + \mu_{A_1})(\gamma_1 + \mu_{I_1})}}_{NS} + \underbrace{\frac{\beta_2 \frac{N_2}{N}(\gamma_2 + \delta_2 + \mu_{I_2})}{(\delta_2 + \mu_{A_2})(\gamma_2 + \mu_{I_2})}}_{SS},$$

where $N_1 + N_2 = N$. A description of the parameters and their values are summarized in Table 1. We assume the basic reproductive number is $\mathcal{R}_0 = 2.5$ and use this to parameterize the mean transmission rate for superspreaders β_2. We note that the value of $\mathcal{R}_0 = 2.5$ is in the absence of control measures, in contrast to commonly reported *effective* reproductive numbers of 0.3–0.5 for MERS outbreaks (Majumder et al., 2014) and 1.3 to 1.9 for Ebola outbreaks, e.g., (Chowell et al., 2004; Khan et al., 2015). In addition, we assume a fixed proportion of SS and NS (20% SS and 80% NS out of a total population size of 2000). A full exploration of the effects of initial conditions and proportion of SS in the population in our designated model outputs was presented in our previous paper focusing on demographic variability (Edholm et al., 2018). Our previous stochastic modeling with MERS and Ebola found that the SS proportion did not significantly affect time to peak infection or peak number of infections when initiated either by one infected SS or one infected NS (Edholm et al., 2018).

First, we describe how demographic variability is incorporated in the MERS NHP and SDE models in Sections 3.1.1 and 3.1.2, respectively. Second, we describe how environmental variability is incorporated through the transmission rates in Section 3.1.3.

3.1.1. NHP model

The MERS NHP model is based on the 12 discrete changes in Table 2 and their associated transition probabilities which form the basis for demographic variability in the MERS NHP model. Because the process with variable $\beta_i(t)$ is time-nonhomogeneous, a simple Gillespie-type algorithm for numerical simulation of sample paths does not apply (Gillespie, 1977). Instead we approximate the NHP through a Monte Carlo simulation. For each time step, at most one of the 12 discrete changes, if any, occurs. For more details see Appendix D. An example of two sample paths, illustrating the total number of infected individuals $I_1 + I_2$ over time are graphed in Fig. 1(b).

3.1.2. SDE model

The MERS SDE model is based on a diffusion approximation, where the first two moments of the process are approximated from the transition rates in Table 2 (Allen, 2007; Allen et al., 2008; Kurtz, 1970, 1971, 1972). The result is a system of Itô SDEs. In particular, the SDE model has the general form:

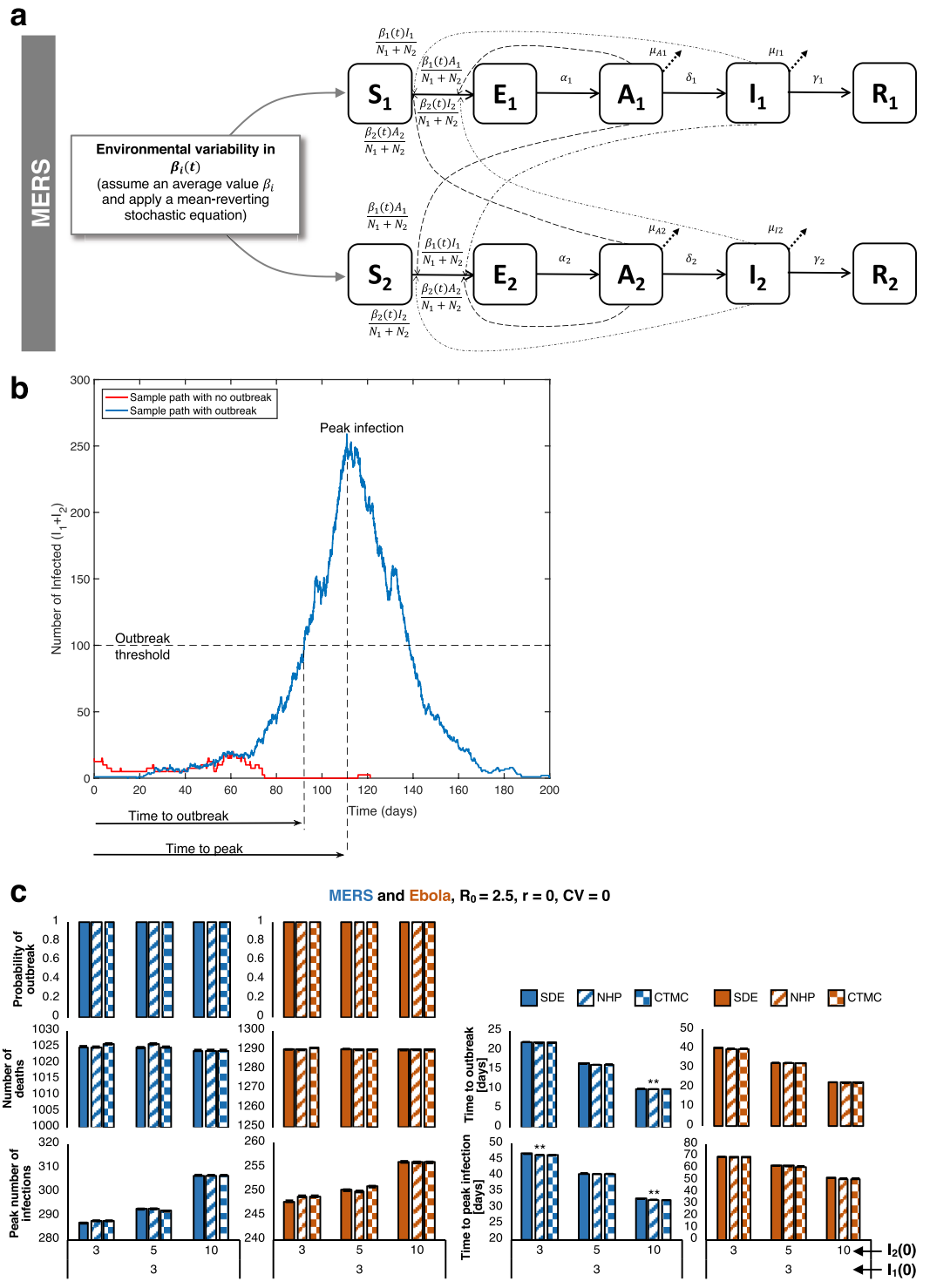

Fig. 1. A time-nonhomogeneous process that incorporates environmental variability in transmission rates. (a) A schematic representation of our compartmental model for MERS disease dynamics in which individuals pass through the susceptible (S_i), exposed (E_i), asymptomatic (A_i), infected (I_i), and recovered (R_i) classes, where $i = 1$ corresponds to non-superspreaders and $i = 2$ corresponds to superspreaders. The model incorporates demographic variability, as before Edholm et al. (2018), and environmental variability is incorporated through the use of a mean-reverting stochastic differential equation in which the disease transmission rate ($\beta_i(t)$) changes continuously over time. (b) Example simulated sample paths, demonstrating a case where an outbreak was observed (blue) and no outbreak was observed (red). Key disease severity metrics used throughout our study are labelled on the graph. The outbreak threshold is defined as a total number of infections surpassing 100 individuals, while time to outbreak is the time at which the threshold is reached. Peak number of infections are the maximal infected population size, with time to peak infection also shown. (c) Comparisons of disease severity predictions between the SDE, NHP and CTMC model (previously published), as quantified by the number of deaths, peak number of infections, time to peak infection, time to outbreak, and probability of outbreak. Multiple initial conditions are shown. Bar graph data shows median values, error bars show standard error of the median ($n = 10,000$ simulations). In (c), one-way analysis of variance (ANOVA) tests were used to assess for mean differences in all outcomes except for probability of extinction based on model predictions with statistical significance accepted when p-value $< 0.05/30$ to adjust for multiple comparisons (** indicates significance).

Table 1
Description of the parameter values for the MERS and Ebola models. Subscripts are $i = 1, 2$ and $j = I, A$. We assume that $\alpha_1 = \alpha_2$, $\delta_1 = \delta_2$, and $\mu_{A1} = \mu_{A2} = \mu_{I1} = \mu_{I2}$. Estimated parameters are marked by ** for assumed $\mathcal{R}_0 = 2.5$ (Edholm et al., 2018). †For a total population size of $N = 2000$, $N_1 = 0.8N$, and $N_2 = 0.2N$ are assumed.

-	Parameter	MERS	Ebola
β_1	NS mean transmission rate	0.06 (Rivers et al., 2014)	0.128 (Chowell & Nishiura, 2014)
β_2	SS mean transmission rate	1.6428**	1.3172**
α_i^{-1}	latent period	6.3 (Cowling et al., 2015)	10 Chowell and Nishiura (2014)
δ_i^{-1}	duration of asymptomatic stage	0.4 (Cowling et al., 2015)	0 (Chowell & Nishiura, 2014)
μ_{Ji}	disease induced death rate	0.08 (Majumder et al., 2015)	1/10.38 (CDC, 2016)
γ_i	recovery rate	0.075 (Jeong et al., 2016)	1/20 (Rivers et al., 2014; Velásquez et al., 2015)
N_1	NS population size	1600†	1600†
N_2	SS population size	400†	400†

Table 2
State transitions and rates for the MERS NHP model with Poisson probabilities $a(t)\Delta t + o(\Delta t)$.

Event	Description	Transition	Rate, $a(t)$
1, 2	Infection of S_i	$S_i \to S_i - 1 E_i \to E_i + 1$	$\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)(A_k + I_k)$
3, 4	Transition of E_i	$E_i \to E_i - 1 A_i \to A_i + 1$	$\alpha_i E_i$
5, 6	Transition of A_i	$A_i \to A_i - 1 I_i \to I_i + 1$	$\delta_i A_i$
7, 8	Death of A_i	$A_i \to A_i - 1$	$\mu_{A_i} A_i$
9, 10	Death of I_i	$I_i \to I_i - 1$	$\mu_{I_i} I_i$
11, 12	Recovery of I_i	$I_i \to I_i - 1 R_i \to R_i + 1$	$\gamma_i I_i$

$$dX(t) = \underbrace{F(X(t),t)}_{drift}\,dt + \underbrace{G(X(t),t)}_{diffusion}\,dW(t),\ X(0) \geq 0, \tag{1}$$

where X is the vector of random variables. The drift vector F agrees with the underlying ODE model, the diffusion matrix G accounts for demographic variability and W is a vector of independent Wiener processes. In the derivation, to order Δt, $F(X(t),t)\Delta t \approx \mathbb{E}[\Delta X(t)]$ and $G(X(t),t)G(X(t),t)^T\Delta t \approx \mathbb{E}[\Delta X(t)\Delta X(t)^T]$ with $\Delta X(t) = X(t+\Delta t) - X(t)$ (Allen, 2007; Allen et al., 2008). As $\Delta t \to 0$, the SDE takes the form of model (1).

In the MERS SDE model, there are 10 state variables and 12 distinct events. We use the same notation for the variables as we did for the ODE and NHP models, specifically

$$X(t) = (S_1, E_1, A_1, I_1, R_1, S_2, E_2, A_2, I_2, R_2)^T.$$

Also, $W(t) = (W_{11}(t), \ldots, W_{16}(t), W_{21}(t), \ldots, W_{26}(t))^T$ is a vector of 12 independent one-dimensional Wiener processes. The first index i of $W_{ij}(t)$ represents NS ($i = 1$) or SS ($i = 2$) and the second index j is one of the six different events $j = 1, \ldots, 6$ for NS or SS. The drift vector $F(X(t), t)$ has dimension 10×1 and is given by the rates in the ODE model. The diffusion matrix G can be expressed as a 10×12 matrix ($10 =$ number of random variables and $12 =$ number of events) (Allen, 2007; Allen et al., 2008):

$$G = \begin{pmatrix} C_{NS} & O \\ O & C_{SS} \end{pmatrix},$$

where O is a 5×6 zero matrix and C_{NS} is a 5×6 matrix with the following form:

$$\begin{pmatrix} -B_1 & 0 & 0 & 0 & 0 & 0 \\ B_1 & -\sqrt{\alpha_1 E_1} & 0 & 0 & 0 & 0 \\ 0 & \sqrt{\alpha_1 E_1} & -\sqrt{\delta_1 A_1} & -\sqrt{\mu_{A_1} A_1} & 0 & 0 \\ 0 & 0 & \sqrt{\delta_1 A_1} & 0 & -\sqrt{\mu_{I_1} I_1} & -\sqrt{\gamma_1 I_1} \\ 0 & 0 & 0 & 0 & 0 & \sqrt{\gamma_1 I_1} \end{pmatrix},$$

where all variables are evaluated at t and

$$B_1 \equiv B_1(t) = \sqrt{\frac{S_1(t)}{N(t)}\sum_{k=1}^{2}\beta_k(t)(A_k(t) + I_k(t))}.$$

Matrix C_{SS} has a similar form but with the role of NS and SS interchanged (subscript 1 is replaced by 2). The order of the six different events for NS is the same as that described in Table 2.

The explicit Itô SDE model for demographic stochasticity in a MERS epidemic consists of a system of 10 equations:

$$\begin{aligned}
dS_i &= -\left(\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)(A_k+I_k)\right)dt - B_i(t)\,dW_{i1}(t),\\
dE_i &= \left(\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)(A_k+I_k) - \alpha_i E_i\right)dt + B_i(t)\,dW_{i1}(t) - \sqrt{\alpha_i E_i}\,dW_{i2}(t),\\
dA_i &= (\alpha_i E_i - \delta_i A_i - \mu_{Ai}A_i)dt + \sqrt{\alpha_i E_i}\,dW_{i2}(t) - \sqrt{\delta_i A_i}\,dW_{i3}(t) - \sqrt{\mu_{A_i}A_i}\,dW_{i4}(t),\\
dI_i &= (\delta_i A_i - \mu_{Ii}I_i - \gamma_i I_i)dt + \sqrt{\delta_i A_i}\,dW_{i3}(t) - \sqrt{\mu_{I_i}I_i}\,dW_{i5}(t) - \sqrt{\gamma_i I_i}\,dW_{i6}(t),\\
dR_i &= \gamma_i I_i dt + \sqrt{\gamma_i I_i}\,dW_{i6}(t),
\end{aligned} \tag{2}$$

for $i = 1, 2$. For simplicity, the dependence on t has been omitted from the state variables, e.g., $S_i \equiv S_i(t)$. To ensure that solutions are nonnegative and that the dynamics follow that of the CTMC when solutions are near zero, we modify this model slightly. We make the assumptions that all transition rates given in Table 2 are nonnegative and that specific terms in the diffusion matrix are zero when the random variables are close to zero (Allen et al., 2020). See details in Appendix B.

The description of the NHP and SDE models for Ebola are similar to those for the MERS models, with the exception that there is no asymptotic stage for Ebola. The Ebola models are described in Appendix C. The similarity between the overall dynamics of the two diseases is consistent with the epidemiological compartmental models which have been previously studied, focusing on different parameters to differentiate the diseases (Chowell et al., 2016; Wong et al., 2015). In terms of compartmental structure the difference between the two diseases is the presence of an asymptomatic class for MERS, which from previous models is not present for Ebola (Agusto, 2017; Bonyah et al., 2016; Chowell et al., 2014; Edholm et al., 2018; Webb et al., 2015). The incorporation of environmental variability in the time-dependent transmission rates $\beta_i(t)$ for the stochastic MERS and Ebola models is described next.

3.1.3. Environmental variability

Environmental variability affects model parameters in a variety of ways (Allen et al., 2006; Cai et al., 2018; Marion et al., 2000; Truscott & Gilligan, 2003; Varughese & Fatti, 2008). For example, environmental fluctuations have been included in the carrying capacity of a zoonotic reservoir (Allen et al., 2006), the death rate in an epidemic model (Cai et al., 2018), and the parasite egg development in a helminth infection (Marion et al., 2000). The parameters in these models, carrying capacity, death rate, and egg development, were modeled by a mean-reverting Ornstein-Uhlenbeck (OU) process. A disadvantage of the OU process is that the parameter governed by an OU process has an asymptotic normal distribution, and therefore, nonnegativity of the parameter is not preserved (Allen, 2016; Iacus, 2009). Several alternative mean-reverting processes have been proposed that are more realistic than the OU process and preserve nonnegativity (Allen, 2016).

The Cox-Ingersoll-Ross (CIR) process is used to model the continuous and random environmental effects on the transmission parameters (Iacus, 2009). The CIR process takes the form of the following SDE:

$$d\beta_i(t) = r_i(\bar{\beta}_i - \beta_i(t))) + \sigma_i\sqrt{\beta_i(t)}dW_i(t), \quad \beta_i(0) = \bar{\beta}_i,\ i = 1, 2. \tag{3}$$

The mean and variance of $\beta_i(t)$ are

$$m_i(t) = \bar{\beta}_i \ \text{ and } \ var_i(t) = \frac{\bar{\beta}_i\sigma_i^2}{2r_i}\left(1 - \exp(-2r_i t)\right) \le \frac{\bar{\beta}_i\sigma_i^2}{2r_i}.$$

The mean value of β_i was the constant transmission parameter used for NS or SS, respectively, in the CTMC models for MERS and Ebola by Edholm et al. (2018) (See Table 1).

The advantages of the CIR process over the OU process are that the CIR process has solutions that are nonnegative and has a well-known and commonly-used asymptotic distribution, a gamma distribution with constant mean $\bar{\beta}_i$ and variance $\bar{\beta}_i\sigma_i^2/(2r_i)$. The coefficient of variation CV (standard deviation divided by the mean) for the asymptotic gamma probability density function (pdf) equals

$$CV = \frac{\sigma_i}{\sqrt{2r_i\bar{\beta}_i}} \ \text{ or } \ \sigma_i = CV\sqrt{2r_i\bar{\beta}_i}. \tag{4}$$

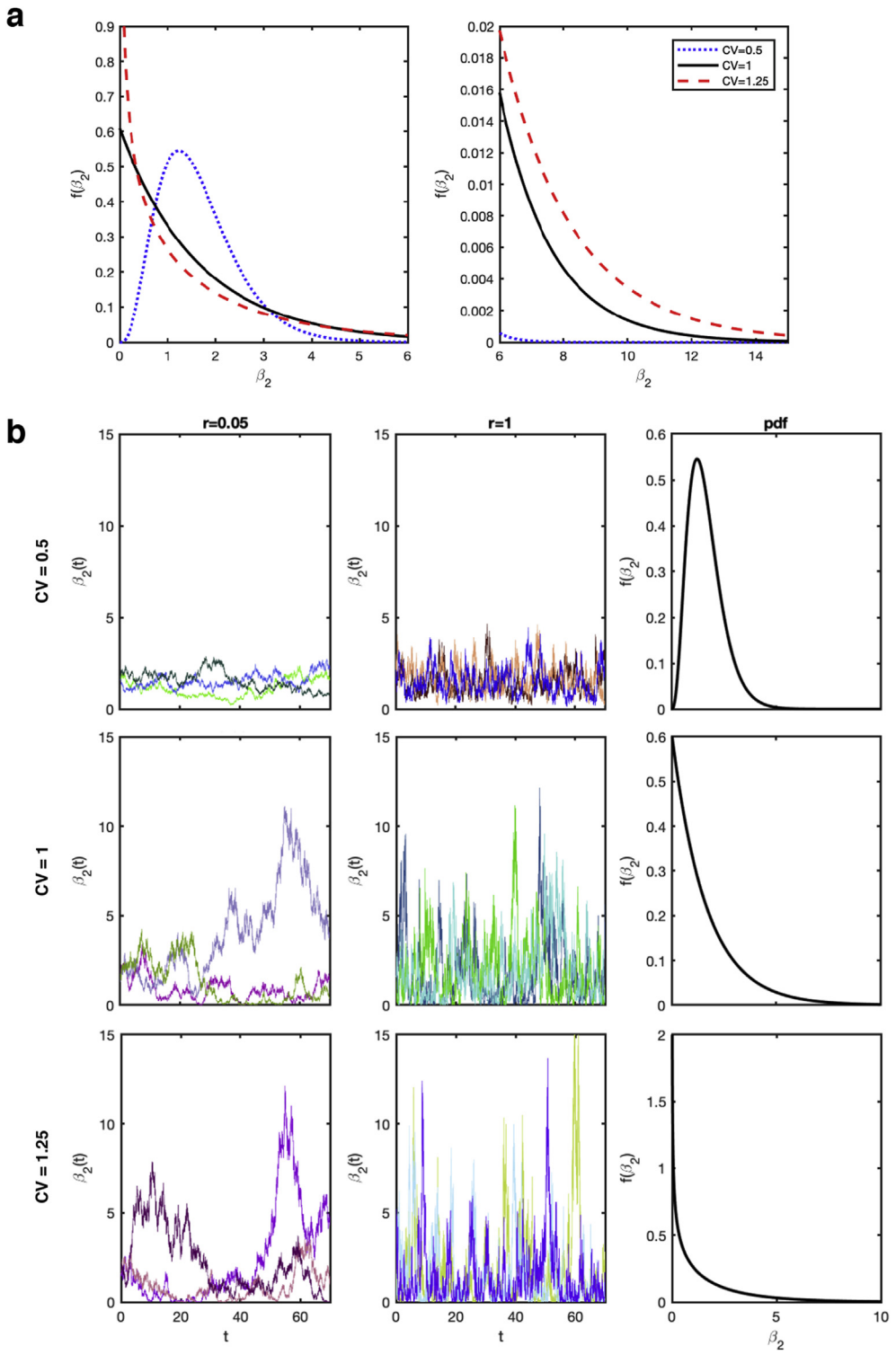

Fig. 2. Impact of environmental variability on SS transmission rate. (a) The probability density function is shown for varied *CV* values over the range [0,6] (left) and range [6,14] (right). (b) Example sample paths for the MERS transmission rate $\beta_2(t)$ over time, with examples for varied *r* and *CV* values shown. The colors indicate different sample paths. The corresponding probability density function (pdf) of the transmission rate is shown on the right.

The asymptotic pdf depends on the parameters, β_i, r_i, and *CV*. As the value of *CV* increases ($CV > 1$) the tail of the distribution becomes fatter, meaning that there is a greater probability for the transmission rate to have large values. See

Fig. 2(a). Formally, the parameter r_i is the rate of return to equilibrium β_i (Iacus, 2009). More specifically, the parameter r_i is related to the fluctuations in $\beta_i(t)$ over time. For a small value of r_i there is greater temporal correlation between the transmission values at time t and $t + \Delta t$ so that the transmission rate changes slowly even when the *CV* is large. Conversely, for a large value of r_i the temporal correlation between two consecutive times is much smaller, and the two values can have widely differing values when the *CV* is large. See Fig. 2(b) which illustrates several sample paths and the asymptotic pdf. Understanding the amount of temporal correlation associated with environmental variability is important as it affects the entire population.

We make the assumption that environmental variability in the transmission rates has the same impact on NS and SS, relative to their mean values. That is, the coefficient of variation, *CV*, and the rate of return to equilibrium, $r_i = r$, are the same for NS and SS. The two SDEs in (3) differ in their mean values β_i, $i = 1, 2$. To evaluate the impact of environmental variability on the epidemic dynamics, the values of r and *CV* are varied, $r = 0.05, 0.2, 0.5, 1$ and $CV = 0.5, 1, 1.25$, where *CV* is based on the formulas for the asymptotic mean and variance of the gamma distribution. Once r and *CV* are selected, then the value of σ_i is defined in formula (4).

In summary, the two SDEs in equation (3) coupled with the NHP and SDE models discussed in the previous two sections account for demographic and environmental variability in MERS and Ebola epidemics. Numerical methods for simulation of the NHP and SDE models are discussed in Appendix D.

4. Results

4.1. Incorporating demographic and environmental stochasticity in MERS and Ebola disease models

We tested the functionality of the SDE and NHP models and found that both models generated predictions of outbreak severity that were in agreement with our established CTMC model, in the absence of environmental variability (Figs. 1(c) and E1(b)–(c) in Appendix E). Though minor differences were observed in the time to outbreak and time to peak infection for some initial conditions, the magnitude of their differences were not practically problematic.

While the NHP model is expected to generate the most accurate predictions for sufficiently small time steps (because it is based on probabilities for individual events), the SDE offers the advantage of significantly lower algorithm run times because a larger step size retains the same accuracy (Fig. 3(a)). The NHP run time, on the other hand, scales with population size, becoming increasingly slow as larger populations are simulated. The SDE loses prediction accuracy for outbreaks when initial infection numbers are close to zero (when differences between discrete and continuous random variables are most evident), a common feature of diffusion approximations in general, e.g., (Ovaskainen & Meerson, 2010). Better agreement between these two models is reached as the total number of infected individuals increases (Fig. 3(b)). An example of the discrepancies between the models at low initial conditions and in the presence of environmental variability is shown in Fig. E2(a) in Appendix E, which compares SDE and NHP predictions in the number of deaths, peak number of infections, and time to peak infection. This is further exemplified by comparisons of SDE versus NHP model predictions for low initial conditions, with varied environmental stochasticity (Fig. E2(b)–(c) in Appendix E).

4.2. Exploring the impact of environmental variability on outbreak probability and severity

We next utilized our model to explore the impact of environmental variability on MERS and Ebola outbreak probability and severity, given outbreaks that begin with a single infected SS, NS, or one of each infected individual. Given the low initial conditions used, we utilized our NHP model based on the observations that the SDE model was less accurate at low initial conditions, as shown in Figs. 3(b) and E2 in Appendix E.

We were able to tune the level of environmental variability using r and *CV*, where r impacts the speed at which the disease transmission rate reverts to the mean value and *CV* impacts the degree to which fluctuations in the transmission rate deviate from the mean (Fig. 2). Therefore, over a fixed period of time, the environmental variability increases with r but the magnitude of this effect is controlled by *CV*. On the other hand, decreasing r results in deviation from the mean for a longer duration of time. With this in mind, we conducted a sweep of r and *CV* values and measured the impact on observed outbreaks, where $(r, CV) = (0, 0)$ corresponds to no environmental variability (demographic variability only). As before, we observed that when disease transmission is initiated by a single infected NS, rather than SS, a lower probability of disease outbreak results for both MERS (Fig. 4(a)) and Ebola (Fig. 5(a)), regardless of the degree of environmental variability. Further, we noted that small values of r and large values of *CV* led to an outbreak probability that is below the levels seen when environmental variability is not included. This was true for both diseases, particularly for scenarios in which a single infected SS initiates transmission. On the other hand, of the outbreaks that were achieved, including environmental variability can lead to more severe outbreaks for both MERS and Ebola. This was evident in the increase in the number of deaths (Figs. 4(b) and 5(b)), faster time to outbreak (Figs. 4(c) and 5(c)), increase in the total number of infections (Figs. 4(d) and 5(d)), increase in peak number of infections (Figs. 4(e) and 5(e)), and a faster time to peak infections (Figs. 4(f) and 5(f)). Example distributions of measures of MERS outbreak severity are shown in the violin plots in Fig. E3 in Appendix E. It is interesting to note that while disease scenarios with environmental variability result in fewer outbreaks, outbreaks that do occur display a higher severity than those without the presence of environmental (and only demographic) variability. For a small value of $r = 0.05$, varying the *CV* value influences the time to outbreak (Figs. 4(c) and 5(c)) and time to peak infection (Figs. 4(f) and

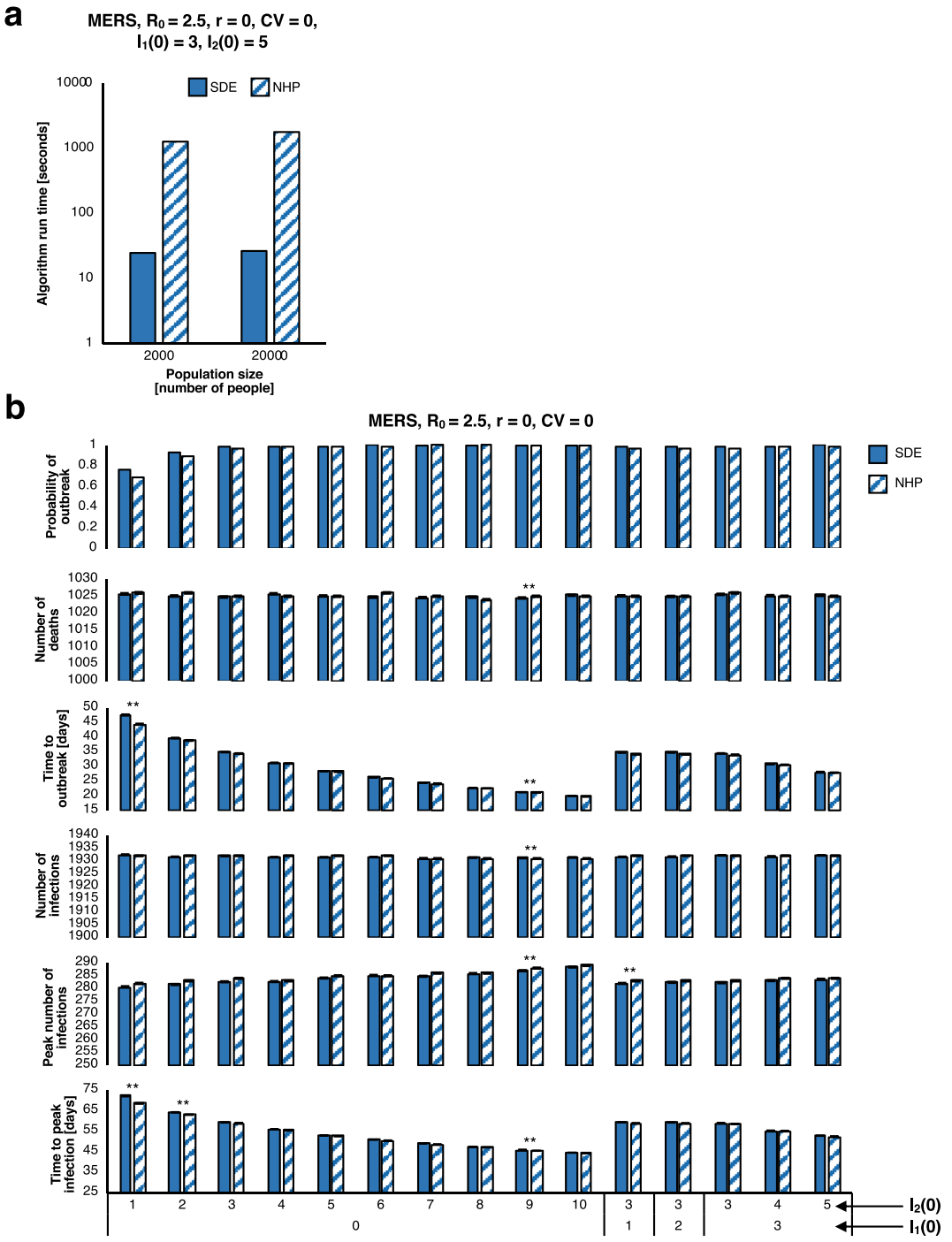

Fig. 3. Comparing the performance of the NHP and SDE models as a function of population size. (a) Exploring the impact of increasing total population size on run times for the NHP and SDE models. Simulations (n = 10,000) were run for MERS, using the largest step size possible while maintaining model prediction accuracy (SDE step size = 0.05, NHP step size = 0.0005). (b) Comparisons of disease severity predictions between the SDE and NHP, as quantified by the number of deaths, peak number of infections, time to peak infection, time to outbreak, number of infections, and probability of outbreak. Results are compared for various initial conditions with demographic variability only (r, CV) = (0, 0). Bar graph data shows median values, error bars show standard error of the median (n = 10, 000 simulations). In (b), one-way analysis of variance (ANOVA) tests were used to assess for mean differences between SDE and NHP model predictions with statistical significance accepted when p-value < 0.05/90 to correct for multiple comparisons (** indicates significance).

5(f)) for both MERS and Ebola. Whereas with a large value of $r = 1$, these variations do not have a large impact on the time to outbreak (Figs. 4(c) and 5(c)) and time to peak infection (Figs. 4(f) and 5(f)) for both MERS and Ebola.

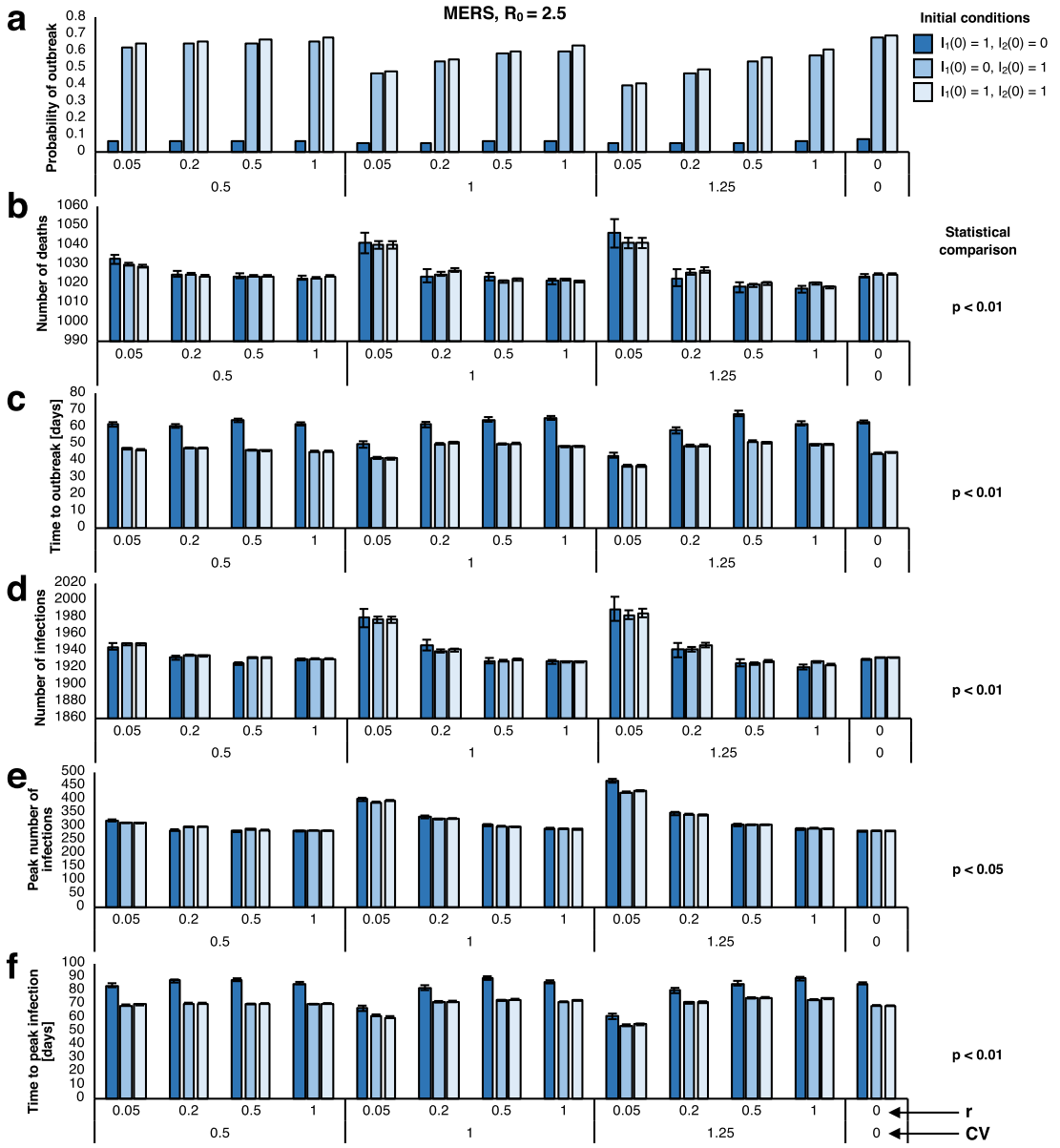

Fig. 4. NHP model prediction exploring the impact of environmental variability on MERS disease dynamics with low initial conditions. Predictions of the MERS NHP model with three initial conditions ("1 NS" corresponds to $I_1(0) = 1$, $I_2(0) = 0$; ("1 SS" corresponds to $I_1(0) = 0$, $I_2(0) = 1$; "1 NS, 1 SS" corresponds to $I_1(0) = 1$, $I_2(0) = 1$) is shown for varied r and CV values. Bar graph data shows median values, error bars show standard error of the median ($n = 10,000$ simulations). Several measures are shown, including (a) probability of outbreak, as well as measures of disease severity for instances in which an outbreak is experienced: (b) number of deaths, (c) time to outbreak, (d) total number of infections, (e) peak number of infections, and (f) time to peak infection. Model predictions were compared in (a)–(f) for values of $(r, CV) = \{(0.05,0.5),(0.05,1.25),(1,0.5),(1,1.25),(0,0)\}$ for $I_1(0) = 1$, $I_2(0) = 1$. Significant differences (p-value < 0.05/30, adjusting for multiple comparisons) were observed in all cases using one-way analysis of variance (ANOVA) tests (p-values indicated beside the graph).

In order to better understand the impact of introducing environmental variability on disease dynamics, we reasoned that this variability introduces fluctuations in the disease transmission rate over time. Any decreases in the transmission rate below the mean value reduced the chance of observing outbreaks, as quantified by the probability of outbreaks (Figs. 4(a) and 5(b)) and demonstrated in example sample paths (Fig. 6(a)). On the other hand, increases in the transmission rate above the mean value contribute to the number of outbreaks observed, given the all-or-none nature of the outbreak definition. Further, these outbreaks experience increased severity due to their enhanced transmission rate, which can be better seen by tracking the transmission rate over time and comparing it to the number of infections observed in example sample paths. For example, the peak infection can reach large values and remain large for a long time in the case $r = 0.05$ and $CV = 1.25$, resulting in the

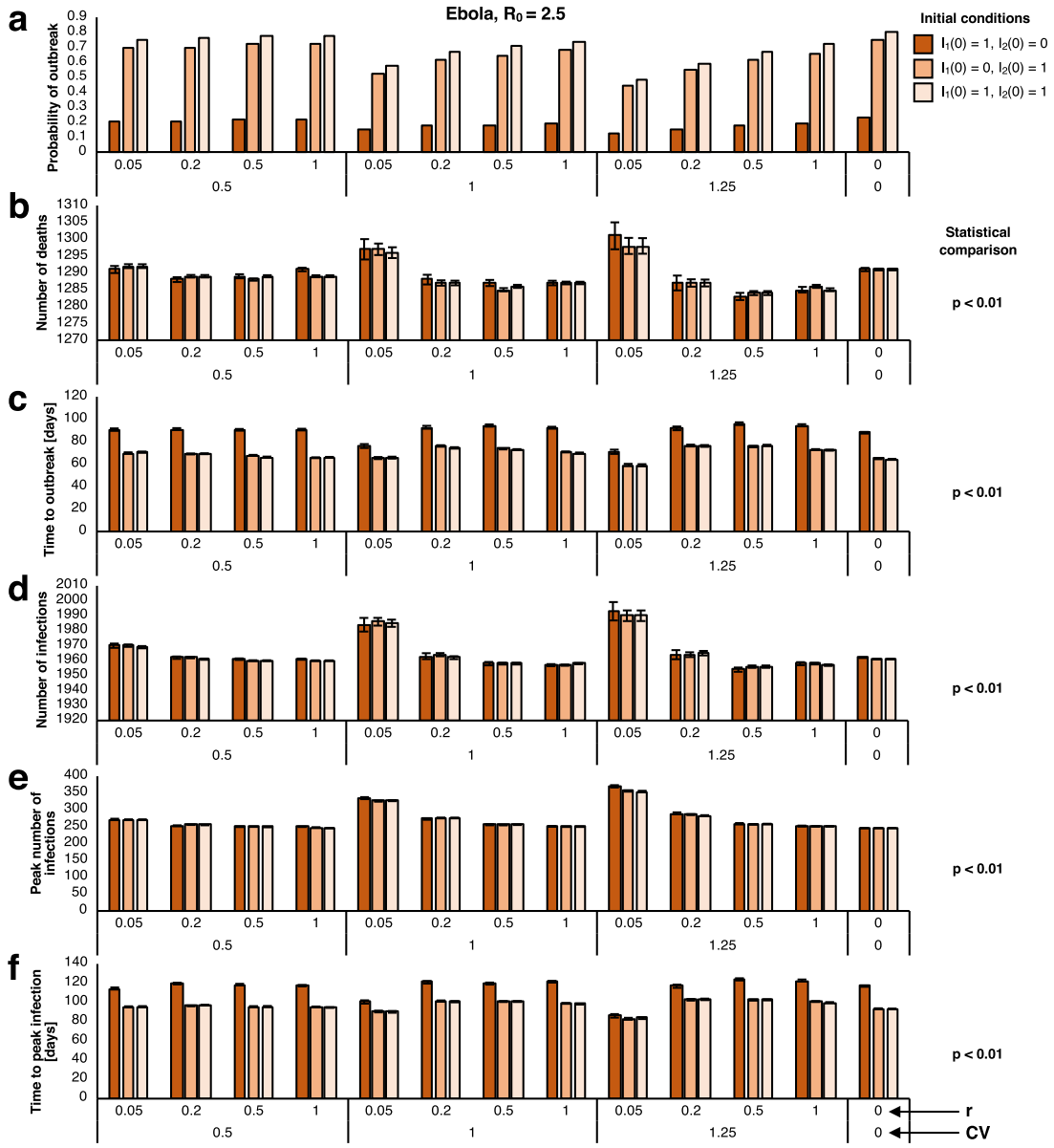

Fig. 5. NHP model prediction exploring the impact of environmental variability on Ebola disease dynamics with low initial conditions. Predictions of the NHP Ebola model with three initial conditions ("1 NS" corresponds to I1(0) = 1, I2(0) = 0; "1 SS" corresponds to $I_1(0) = 0$, $I_2(0) = 1$; "1 NS, 1 SS" corresponds to $I_1(0) = 1$, $I_2(0) = 1$) is shown for varied r and CV values. Bar graph data shows median values, error bars show standard error of the median ($n = 10,000$ simulations). Several measures are shown, including (a) probability of outbreak, as well as measures of disease severity for instances in which an outbreak is experienced: (b) number of deaths, (c) time to outbreak, (d) total number of infections, (e) peak number of infections, and (f) time to peak infection. Model predictions were compared in (a)–(f) for values of $(r, CV) = \{(0.05,0.5),(0.05,1.25),(1,0.5),(1,1.25),(0,0)\}$ for $I_1(0) = 1$, $I_2(0) = 1$. Significant differences (p-value < 0.05/30, adjusting for multiple comparisons) were observed in all cases using one-way analysis of variance (ANOVA) tests (p-values indicated beside the graph).

accumulation of more infections (Fig. 6(a)). These demonstrate that spikes in the transmission rate above the mean level often trigger corresponding increases in the number of infections over time (Fig. 6(b)). We also explored the effect of initial NS and SS transmission rates on model predictions, either using the mean value $\overline{\beta}_i$, $i = 1, 2$, or a random value from the asymptotic gamma distribution at the first time step. We found that the trends observed were conserved across model predictions using these two rules for transmission rate initialization, validating our model assumptions (Fig. E4 in Appendix E). We did note differences in the predicted probability of an outbreak, time to outbreak, and time to peak infection when these two transmission rate initialization strategies were used for small values of r, though this was not of notable issue within the range of r values we explored.

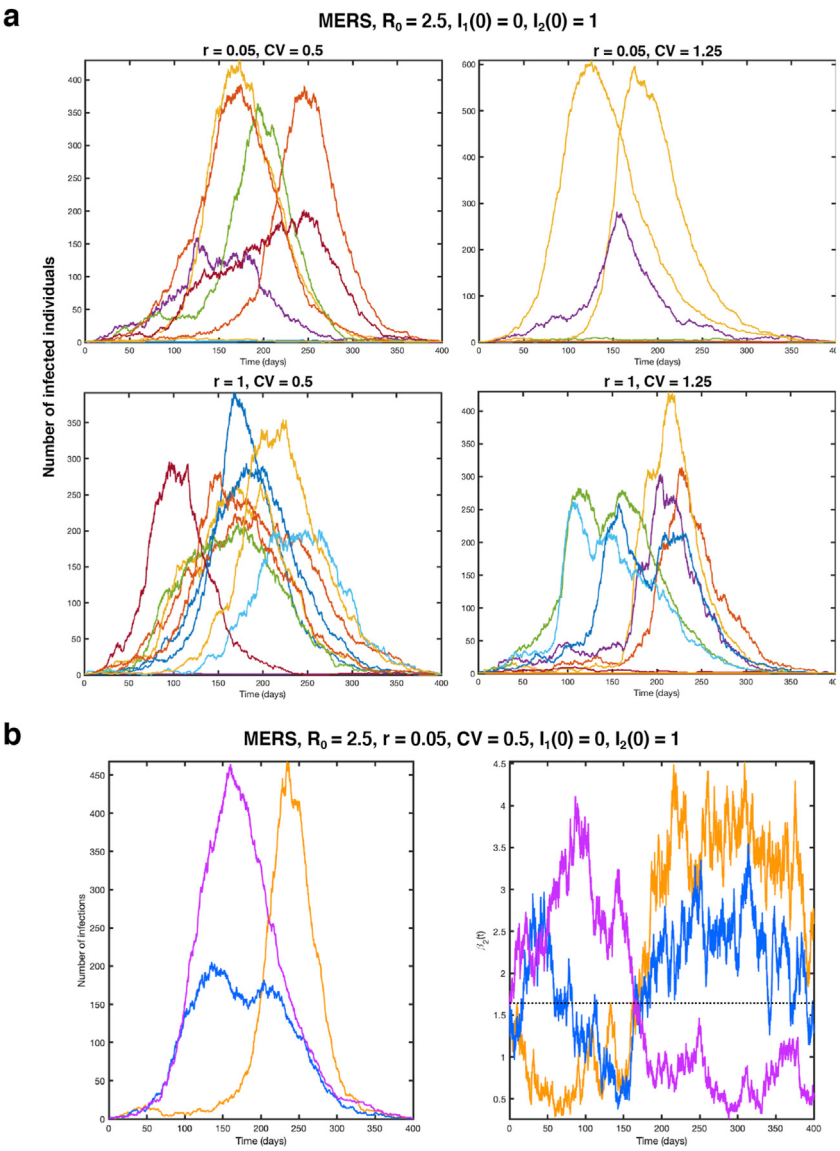

Fig. 6. Sample paths for MERS model simulations with changing environmental variability. (a) Ten example sample paths are shown for MERS NHP model simulation with changing r and CV values for a disease scenario in which a single infected SS individual starts the disease transmission ($I_1(0) = 0$, $I_2(0) = 1$). (b) Three example sample paths are shown for MERS NHP with environmental variability, $(r, CV) = (0.05, 0.5)$, for a disease scenario in which a single infected SS individual starts the disease transmission ($I_1(0) = 0$, $I_2(0) = 1$). Corresponding sample paths for SS disease transmission rate over time is shown to the right, with the mean transmission rate indicated by the dotted black line. The colors indicate different sample paths.

We next sought to explore the effect of sweeping r and CV values in disease scenarios with high initial numbers of infected individuals, all introduced to the population at the same time, in which the SDE model is known to give rise to reliable predictions (Fig. E2(a) in Appendix E). In fact, we validated that the SDE model predictions agreed with those of the NHP with environmental variability (Fig. E5 in Appendix E), with only minor differences that do not impact the trends observed. As with simulations conducted with low initial conditions, we observed that the probability of an outbreak increased with increasing r values, especially with large CV values (Fig. 7(a)). Furthermore, we observed that while environmental variability (by increasing CV) reduces the chance of both MERS and Ebola outbreaks, those that do occur experience higher severity in terms of number of deaths (Fig. 7(b)), time to outbreak (Fig. 7(c)), number of infections (Fig. 7(d)), and peak number of infections (Fig. 7(e)). Interestingly, however, we observed a non-monotonic relationship between r and measures of outbreak severity. This is in contrast to our previous observation that decreasing the r value increases the outbreak severity (Figs. 4 and 5). In fact, we found that a medium r value resulted in the slowest time to outbreak (Fig. 7(c)) and peak infection (Fig. 7(f)) for both MERS and Ebola. We also found that these trends were observed for conditions with equal numbers of NS and SS susceptible individuals.

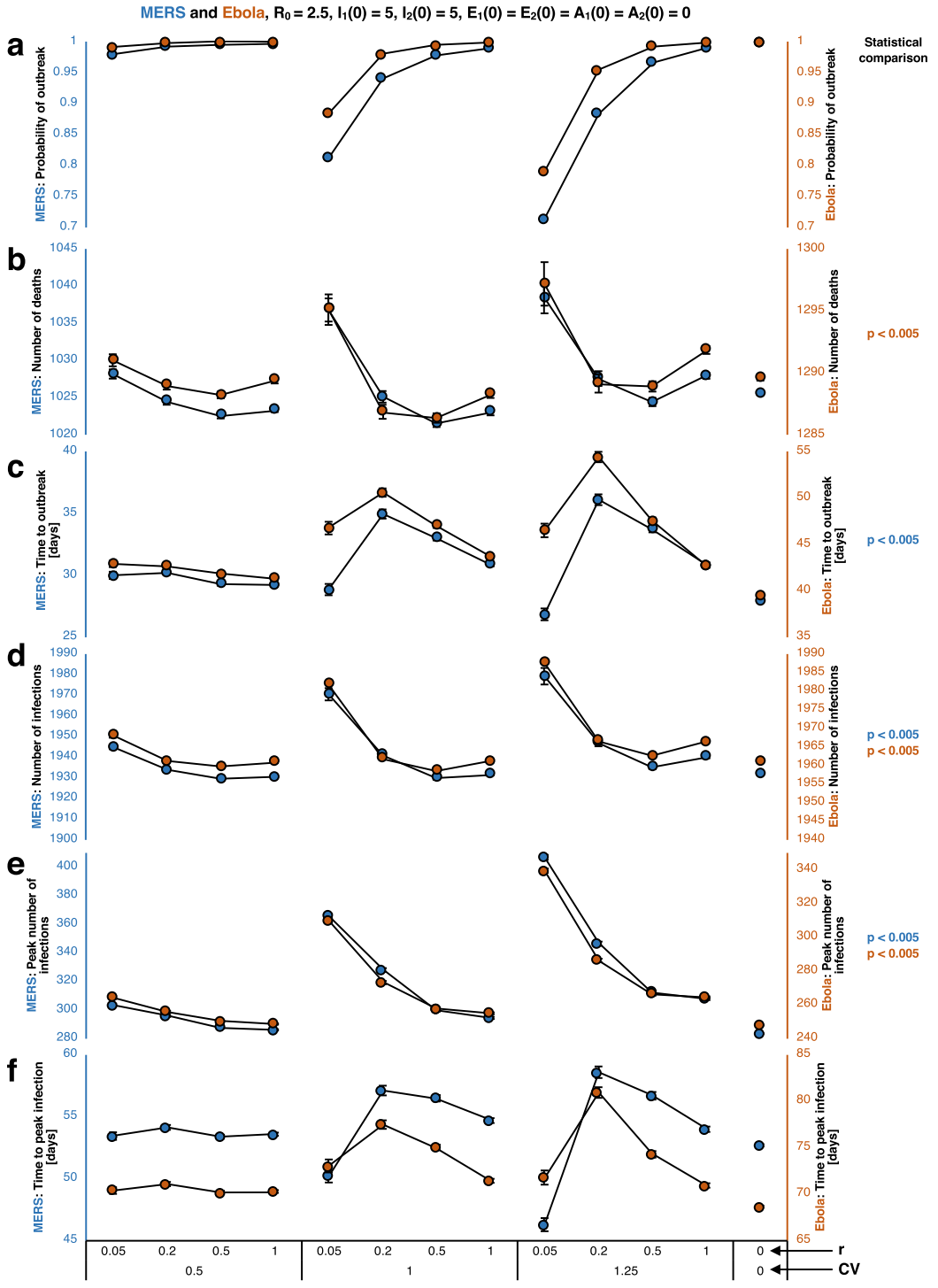

Fig. 7. SDE model predictions exploring the impact of environmental variability on MERS and Ebola disease dynamics with high initial conditions. Predictions of the MERS and Ebola SDE models with high initial condition ($I_1(0) = 5$, $I_2(0) = 5$, $E_1(0) = E_2(0) = A_1(0) = A_2(0) = 0$) is shown for varied r and CV values. Line graph data shows median values standard error (n = 10,000 simulations). Several measures are shown, including (a) probability of outbreak, as well as measures of disease severity for instances in which an outbreak is experienced: (b) number of deaths, (c) time to outbreak, (d) total number of infections, (e) peak number of infections, and (f) time to peak infection. Blue shows MERS results (values accompanied by the y-axis on the left side of the plot) and orange shows Ebola results (values accompanied by the y-axis on the right side of the plot). Model predictions were compared in (a)–(f) for values of (r, CV) = {(0.05,0.5),(0.05,1.25),(1,0.5),(1,1.25),(0,0)} for $I_1(0) = 1$, $I_2(0) = 1$. Significant differences (p-value < 0.05/30, adjusting for multiple comparisons) were observed in all cases using one-way analysis of variance (ANOVA) tests (p-values indicated beside the graph).

5. Discussion

Environmental changes over time, such as social behavior, temperature fluctuations, or medical quality, influence contact between individuals and successful disease transmission (Brauer & Castillo-Chavez, 2012). Our simulations reveal that deviations in disease transmission rates, both above and below the mean, drive significant changes in outbreak dynamics. This aligns with our previously reported sensitivity analysis predictions, which uncovered the transmission rate as particularly impactful on disease outcomes (Edholm et al., 2018). Importantly, environmental variability reduces the probability of outbreak occurrence since temporal decreases in the transmission rate reduces the chance that the outbreak threshold will be reached. On the other hand, once an outbreak is triggered, variability in the disease transmission rate increases the severity of the disease dynamics by increasing the number of deaths and total infections, increasing the peak number of infections reached, and reducing the time to the outbreak and peak number of infections. This is a result of the fact that any spikes in the transmission rate has compounding effects as the number of infectious individuals increases. These methods can be used to investigate the effects of environmental variability on transmission or other model parameter values in epidemic or ecological settings.

While this study focuses on modeling transmission dynamics rather than evaluating specific public health strategies, our findings highlight the importance of altering environmental factors to mitigate disease spread. Masking and social distancing are two such factors that have been shown to affect the transmission rate (Eikenberry et al., 2020; Lai et al., 2012). For example, public health interventions that are applied, later rescinded but then applied again, such as the lockdowns with respect to COVID-19 (Acuña-Zegarra et al., 2020; Della Rossa et al., 2020; Dropkin, 2020), could be viewed as contributing to a high amplitude and low frequency of variability (small r and large CV). In this case, our models predict that such interventions may lead to more severe epidemics. On the other hand, with high frequency in variability (large r) compared to the time scale of the disease dynamics, the outcomes may be similar to those with no environmental variability. Consequently, high variability in public health interventions may result in their reduced effectiveness. Hence, findings of this study indirectly support the potential for modifiable environmental factors, influenced by public health strategies, to control disease transmission. Further work is needed to fully evaluate intervention programs and their impact on disease prevalence.

We also made the interesting observation that the severity of outbreaks does not vary monotonically as a function of the degree of environmental variability. Indeed, we found that for outbreaks that are started by a number of SS and NS individuals introduced at the same time, rather than a single SS or NS individual, there is an optimal level of environmental variability (determined by r and CV values) that results in a delayed time to outbreak and time to peak infection. This suggests that modeling could play a role in predicting an optimized distribution for transmission rates that can be attained by tuning social measures in such a way as to reduce outbreak probability and severity while minimizing socioeconomic impacts of social distancing.

Our study analyzes the dynamics of two diseases, MERS and Ebola, with the same $\mathcal{R}_0$ value. Qualitatively, the trends in outcome measures shown in Figs. 3–5 appear similar as the initial conditions, r, and CV parameters are varied. In Figs. 4 to 5, the number of deaths, the number infected, and peak number infected are similar among the three cases of initial conditions but are higher when r is small and CV is large for both diseases. The probability of an outbreak was higher for both diseases when the initial conditions included both a superspreader and a non-superspreader. Further, the time to outbreak and time to peak infection were more sensitive to increases in r when CV was higher. While the results between the two diseases appear similar despite differences in compartments, the quantitative or estimated outcomes differ between MERS and Ebola. In Fig. 7, the ranges of values are higher in Ebola in comparison to MERS for the three outcome measures including averages of: the total number of deaths (Ebola: 1285 to 1300 vs. MERS: 1020 to 1045), time to outbreak (Ebola: 35 to 55 vs. MERS: 25 to 40), and time to peak infections (Ebola: 45 to 60 vs. MERS 65 to 85). In contrast, the average peak number of infections was lower in Ebola (240–340) in comparison to MERS (280–400). The probability of outbreak ranged between 0 and 1 for both diseases and the number of total infected in the population approximately ranged between 1930 and 1990 for both diseases.

In using Ebola and MERS as a case study for the impact of environmental variability on disease transmission dynamics, our theoretical investigations lay the foundation for understanding the impact of randomly varying transmission rates on outbreak probability and severity. However, the amount of variability observed in a particular epidemic (r and CV) depends on the variability and heterogeneity in social contacts, as well as factors affecting successful transmission during the epidemic. Quantifying and identifying the factors that account for this variability present a challenging problem. Nevertheless, the modeling strategy and trends we observed, which are closely in agreement between Ebola and MERS despite differences in asymptomatic disease states, are expected to apply more broadly to other communicable diseases, such as COVID-19. Overall, our results suggest that for a set mean transmission rate, environmental variability can reduce the probability of an outbreak yet increase the severity of those outbreaks that occur. Thus, tighter control over environmental variables such as adherence to hospital disease transmission-reduction guidelines, social distancing in the general population, and other control measures may be warranted.

It is worth mentioning that, given the large number of random variables, time-nonhomogeneity, and nonlinearity, exact analytical results for our models are not feasible. Certainly, in the case of time-homogeneous stochastic processes such as continuous-time Markov chain and SDE models, methods to approximate the mean and higher-order moments of the random variables, the final size, the duration until disease extinction, and the variability about the endemic state have been applied, e.g., Britton (2010); Ekanayake and Allen (2010); Krishnarajah et al. (2005); Lloyd (2004); Ovaskainen and Meerson (2010); Van Kampen (1992). Unfortunately, many of these methods do not extend to the time-nonhomogeneous models with mean-

reverting processes. However, in simple SIR and SIS models, various outcomes from the time-nonhomogeneous processes can be more easily compared to the analytical approximations from the corresponding time-homogeneous processes. An interesting topic for future research is to apply our methods and the preceding methods in simpler settings to investigate and compare the effects of demographic and/or environmental variability and the amount of temporal correlation with mean-reverting processes on disease outcomes. Another future extension of our models could seek to include the contributions of novel pathogenic variants as well as the behavioral aspects of SS individuals, which may make them more susceptible to infection by NS and SS individuals.

Author contributions

L.J.S.A conceived the study. N.S., C.J.E, B.O.E, A.L.M, A.P., O.S., X.W., L.J.S.A contributed to model formulation and analysis, simulations, interpretations of results, and manuscript writing.

Declaration of competing interest

None.

Acknowledgements

The authors acknowledge the support of an American Institute of Mathematics SQuaREs grant. A.P. was supported by NSF grant DMS-1815750. N.S. was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Post-doctoral Fellowship. In addition, we thank Texas Tech University for hosting our third WAMB group meeting and the P.W. Horn Professorship of L.J.S.A. for providing financial support. We thank the two referees for their helpful suggestions.

Appendix A. ODE model for MERS

The underlying ODE model for MERS is

$$\begin{aligned}
\frac{dSi}{dt} &= -\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)(I_k + A_k),\\
\frac{dE_i}{dt} &= \frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)(I_k + A_k) - \alpha_i E_i,\\
\frac{dA_i}{dt} &= \alpha_i E_i - \delta_i A_i - \mu_{Ai} A_i,\\
\frac{dI_i}{dt} &= \delta_i A_i - \mu_{Ii} I_i - \gamma_i I_i,\\
\frac{dR_i}{dt} &= \gamma_i I_i.
\end{aligned}$$

Appendix B. SDE Derivation

The MERS and Ebola SDE models with vector random variables $X = (X_1, \ldots, X_n)^T$, drift vector $F(X, t) = F = (F_1, \ldots, F_n)^T$ and diffusion matrix $G(X, t) = G = [G_{ij}]$, $i = 1, \ldots, n$ and $j = 1, \ldots, m$ (n = number of random variables, m = number of events) are modified slightly to form new SDE models to ensure (1) the rates in the SDE are nonnegative and (2) the SDE sample paths are nonnegative (Allen et al., 2020; Cresson & Sonner, 2018).

To ensure that the SDE dynamics follow closely those of the NHP, the transition rates defined in Tables 2 and C1 (Appendix C) are assumed to be nonnegative (Allen et al., 2020). In particular, the transition rates $a(t_n)$ at time t_n are replaced by

$$\hat{a}(t_n) = \begin{cases} a(t_n), & \text{if } a(t_n) \geq 0,\\ 0, & \text{if } a(t_n) < 0. \end{cases}$$

With this change the SDE models have the form $dX = \hat{F}(X,t)\,dt + \tilde{G}(X,t)\,dW$.

Next to ensure that the SDE solutions are nonnegative some of the diffusion terms are modified slightly when a random variable is close to zero. For the SDE to have nonnegative sample paths it must be the case that each term $\tilde{G}_{ij}$ in the diffusion matrix satisfy $\tilde{G}_{ij|X_i=0} = 0$ (Cresson & Sonner, 2018). If this condition is not satisfied for a fixed i and j, then the SDE term $\tilde{G}_{ij}$ is modified as follows:

$$\hat{G}_{ij} = \begin{cases} \tilde{G}_{ij}, & X_i \geq \varepsilon, \\ \tilde{G}_{ij}\sqrt{X_i/\varepsilon}, & 0 \leq X_i < \varepsilon, \end{cases}$$

for $0 < \varepsilon \leq 1$ (Allen et al., 2020). A value $\varepsilon = 0.1$ was chosen after testing the numerical code to ensure that reducing the value of ε does not change the results. Applying the previous modifications, the SDE models applied to MERS and Ebola have nonnegative transition rates and nonnegative sample paths (Allen et al., 2020):

$$dX = \hat{F}(X,t)\,dt + \hat{G}(X,t)\,dW, \; X(0) \geq 0.$$

Appendix C. Ebola models

Unlike MERS, the asymptotic stage for Ebola is short (Chowell & Nishiura, 2014) and thus it is omitted. The underlying Ebola model is a system of 8 ODEs, four equations for SS and four for NS:

$$\begin{aligned} \frac{dS_i}{dt} &= -\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)I_k, \\ \frac{dE_i}{dt} &= \frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)I_k - \alpha_i E_i, \\ \frac{dI_i}{dt} &= \alpha_i E_i - \mu_{Ii} I_i - \gamma_i I_i, \\ \frac{dR_i}{dt} &= \gamma_i I_i, \end{aligned}$$

for $i = 1, 2$. If $\beta_i(t) \equiv \beta_i$ is constant, then the basic reproduction number for the Ebola ODE model is

$$\mathcal{R}_0 = \underbrace{\frac{\beta_1 \frac{N_1}{N}}{\gamma_1 + \mu_{I_1}}}_{NS} + \underbrace{\frac{\beta_2 \frac{N_2}{N}}{\gamma_2 + \mu_{I_2}}}_{SS}.$$

The corresponding Itô SDE SEIR model for Ebola has 8 Wiener processes, corresponding to the 8 events in Table C1.

Table C.1
State transitions and rates for the NHP model with Poisson probabilities $a(t)\Delta t + o(\Delta t)$.

Event	Description	Transition	Rate, a
1, 2	Infection of S_i	$S_i \to S_i - 1 E_i \to E_i + 1$	$\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)I_k$
3, 4	Transition of E_i	$E_i \to E_i - 1 I_i \to I_i + 1$	$\alpha_i E_i$
5, 6	Death of I_i	$I_i \to I_i - 1$	$\mu_{I_i} I_i$
7, 8	Recovery of I_i	$I_i \to I_i - 1 R_i \to R_i + 1$	$\gamma_i I_i$

The Ebola SDE model with demographic variability can be derived in a similar manner as for the MERS SDE model:

$$
\begin{aligned}
dS_i &= -\left(\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)I_k\right)dt - \sqrt{\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)I_k}\,dW_{i1}(t),\\
dE_i &= \left(\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)I_k - \alpha_i E_i\right)dt + \sqrt{\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)I_k}\,dW_{i1}(t) - \sqrt{\alpha_i E_i}\,dW_{i2}(t),\\
dI_i &= (\alpha_i E_i - \mu_{Ii}I_i - \gamma_i I_i)dt + \sqrt{\alpha_i E_i}\,dW_{i2}(t) - \sqrt{\mu_{I_i} I_i}\,dW_{i3}(t) - \sqrt{\gamma_i I_i}\,dW_{i4}(t),\\
dR_i &= \gamma_i I_i dt + \sqrt{\gamma_i I_i}\,dW_{i4}(t),
\end{aligned}
\tag{C.1}
$$

for $i = 1, 2$. For simplicity, the dependence on t has been omitted from the state variables, e.g., $S_i \equiv S_i(t)$. The Ebola SDE model is modfied slightly, similar to the MERS SDE model, to ensure that the rates in Table C1 are nonnegative and that the solutions are nonnegative. (See Appendix B)

Appendix D. Numerical

For the MERS NHP model, the sum of all of the rates in the right-hand column of Table 2 is

$$\Gamma(t) = \sum_{i=1}^{2}\left[\frac{S_i}{N}\sum_{k=1}^{2}\beta_k(t)(A_k + I_k) + \alpha_i E_i + \cdots + \gamma_i I_i\right]$$

In general, for well-defined and integrable time-varying parameters, given an event at time t_n, the time until the next event τ can be found by inverting the following cumulative distribution function:

$$H_X(t) = 1 - \exp\left(-\int_0^t \Gamma(t_n + \tau)\,d\tau\right), \tag{D.1}$$

where the transmission rates $\beta_i(t)$ are functions of t but each of the random variables in $\Gamma(t)$ are fixed at the time of the last event t_n, e.g., $S_i(t_n)$, $E_i(t_n)$, etc. (Ross, 2014). In particular, the time until the next event is $T = H_X^{-1}(U)$, where U is a uniform random variable on $[0, 1]$ and $X = (S_1, E_1, A_1, I_1, R_1, S_2, E_2, A_2, I_2, R_2)^T$ is the discrete random vector for the stochastic process. For constant β_i, this method simplifies to the Gillespie algorithm (Gillespie, 1977). Calculation via inversion of equation (D.1) is time-consuming and complex. Also, as $\beta_i(t)$ is a random variable, this method is not possible for our models. Instead we apply a Monte Carlo method by incrementing time by a sufficiently small positive time step Δt to ensure that each of the 12 probabilities and $\Gamma(t)\Delta t$ are nonnegative and less than one and that only one event, if any, occurs in a given time interval Δt. A similar method is applied to the Ebola NHP model.

It is straightforward to include the two SDE equations for $\beta_i(t)$ in equation (3) in the NHP and SDE models for MERS and Ebola. The two SDES in equation (3) are approximated at a discrete set of points $t = 0, \Delta t, 2\Delta t, \ldots,$

$$\beta_i(t + \Delta t) = \beta_i(t) + r_i\Delta t(\bar{\beta}_i - \beta_i(t)) + \sigma_i\sqrt{\beta_i(t)\Delta t}\,\eta,$$

where Δt is sufficiently small and η is a random value from a standard normal distribution. Each of the 10^4 sample paths are approximated at the discrete set of points.

For the NHP models, the time step is chosen sufficiently small to ensure the probabilities and their sum in Tables 2 and C1 (Appendix C) are nonnegative in the Monte Carlo simulation. For the SDE models, the Euler-Maruyama method is applied. The random vector for 10^4 sample paths is updated each time step. The numerical methods are checked for accuracy by reducing the time steps sequentially by 1/2 to ensure there are no significant differences in the output measures with smaller time steps. The time step chosen for the numerical simulation of the NHP models was $\Delta t = 5 \times 10^{-4}$ and for the SDE models was $\Delta t = 5 \times 10^{-2}$.

Appendix E. Figures

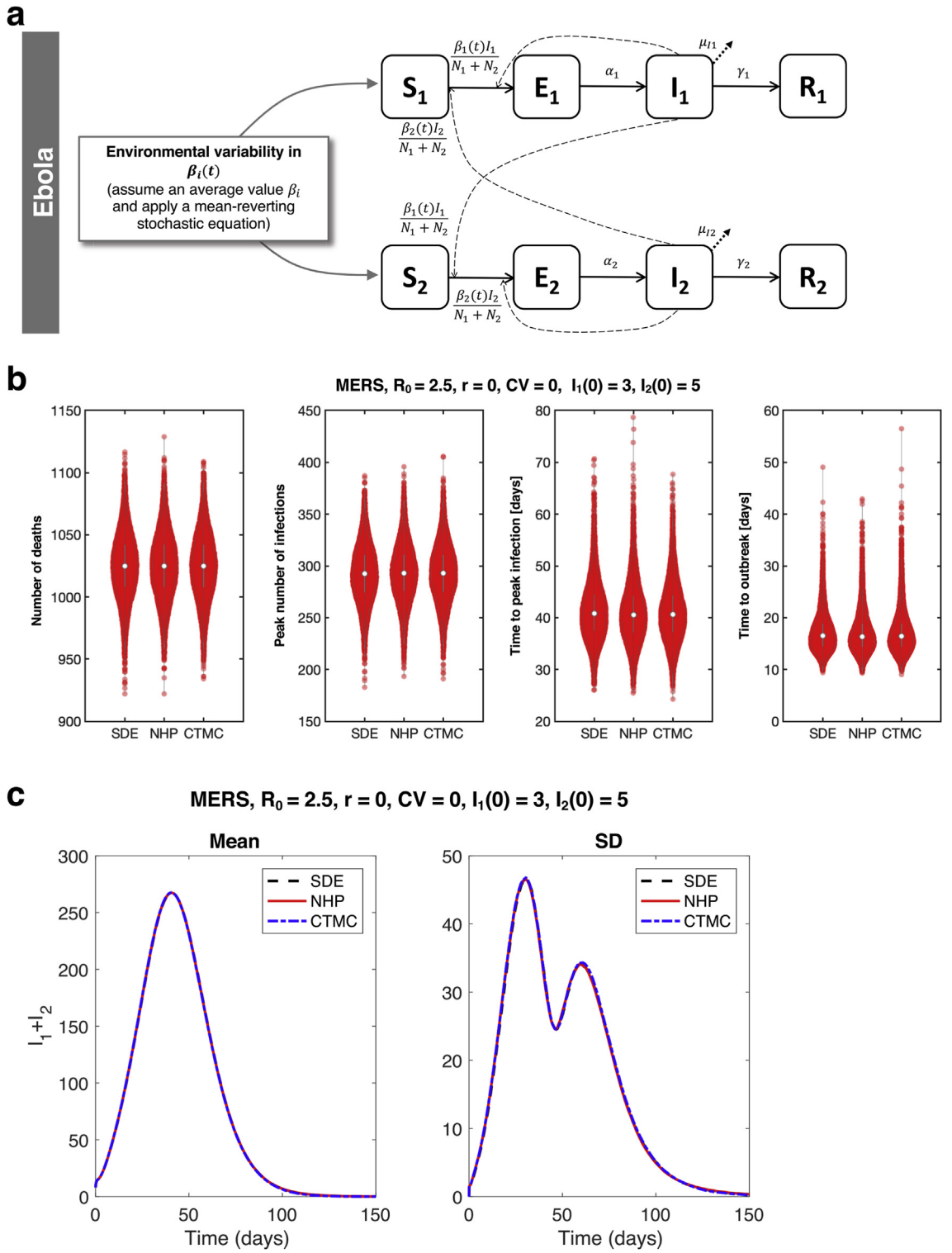

Fig. E.1. A time-nonhomogeneous process that incorporates environmental variability in an Ebola disease model. (a) A schematic representation of our compartmental model for Ebola disease dynamics in which individuals pass through the susceptible (S_i), exposed (E_i), asymptomatic (A_i), infected (I_i), and recovered (R_i) classes, where $i = 1$ corresponds to non-superspreaders and $i = 2$ corresponds to superspreaders. The model incorporates demographic variability,

as before Edholm et al. (2018), and environmental variability is incorporated through the use of a mean-reverting stochastic differential equation in which the disease transmission rate changes over time. (b) Comparisons of disease severity predictions between the SDE, NHP and CTMC model in violin plots (previously published), as quantified by the number of deaths, peak number of infections, time to peak infection, and time to outbreak. The MERS model is used to demonstrate the model predictions for the initial conditions $I_1(0) = 3$, $I_2(0) = 5$, $E_1(0) = 3 = A_1(0)$, and $E_2(0) = 5 = A_2(0)$. (c) Mean and standard deviation (SD) of total number of infected individuals ($I_1 + I_2$) shown comparing SDE, NHP, and CTMC model predictions for the case of no environmental variability, $(r, CV) = (0, 0)$, and high initial conditions ($I_1(0) = 3$, $I_2(0) = 5$, $E_1(0) = 3 = A_1(0)$, and $E_2(0) = 5 = A_2(0)$).

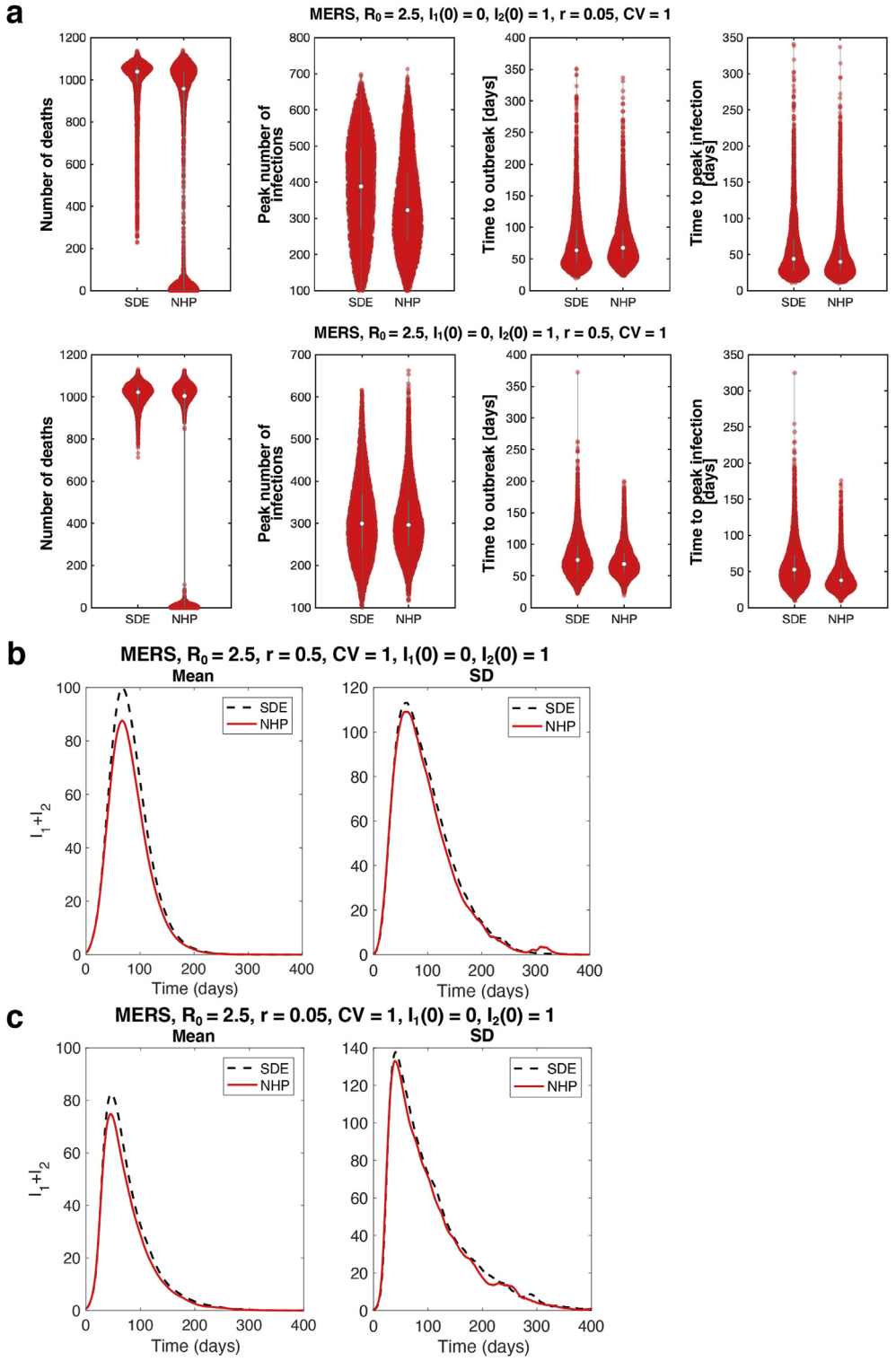

Fig. E.2. Comparisons of SDE versus NHP model performance as a function of initial conditions. (a) Comparisons of NHP versus SDE model predictions in violin plots for an example MERS disease scenario in which a single infected SS individual begins the disease transmission. Two environmental variability conditions are included for comparison: $(r, CV) = (0.05, 1)$ and $(r, CV) = (0.5, 1)$. Displayed as violin plots. (b) Mean and standard deviation (SD) of total number of infected individuals $(I_1 + I_2)$ shown comparing: SDE and NHP model predictions for the case of environmental variability, $(r, CV) = (0.5, 1)$, and low initial conditions $(I_1(0) = 0, I_2(0) = 1, E_1(0) = 0 = A_1(0)$, and $E_2(0) = 0 = A_2(0))$. (c) SDE and NHP model predictions for the case of environmental variability, $(r, CV) = (0.05, 1)$, and low initial conditions $(I_1(0) = 0, I_2(0) = 1, E_1(0) = 0 = A_1(0)$, and $E_2(0) = 0 = A_2(0))$.

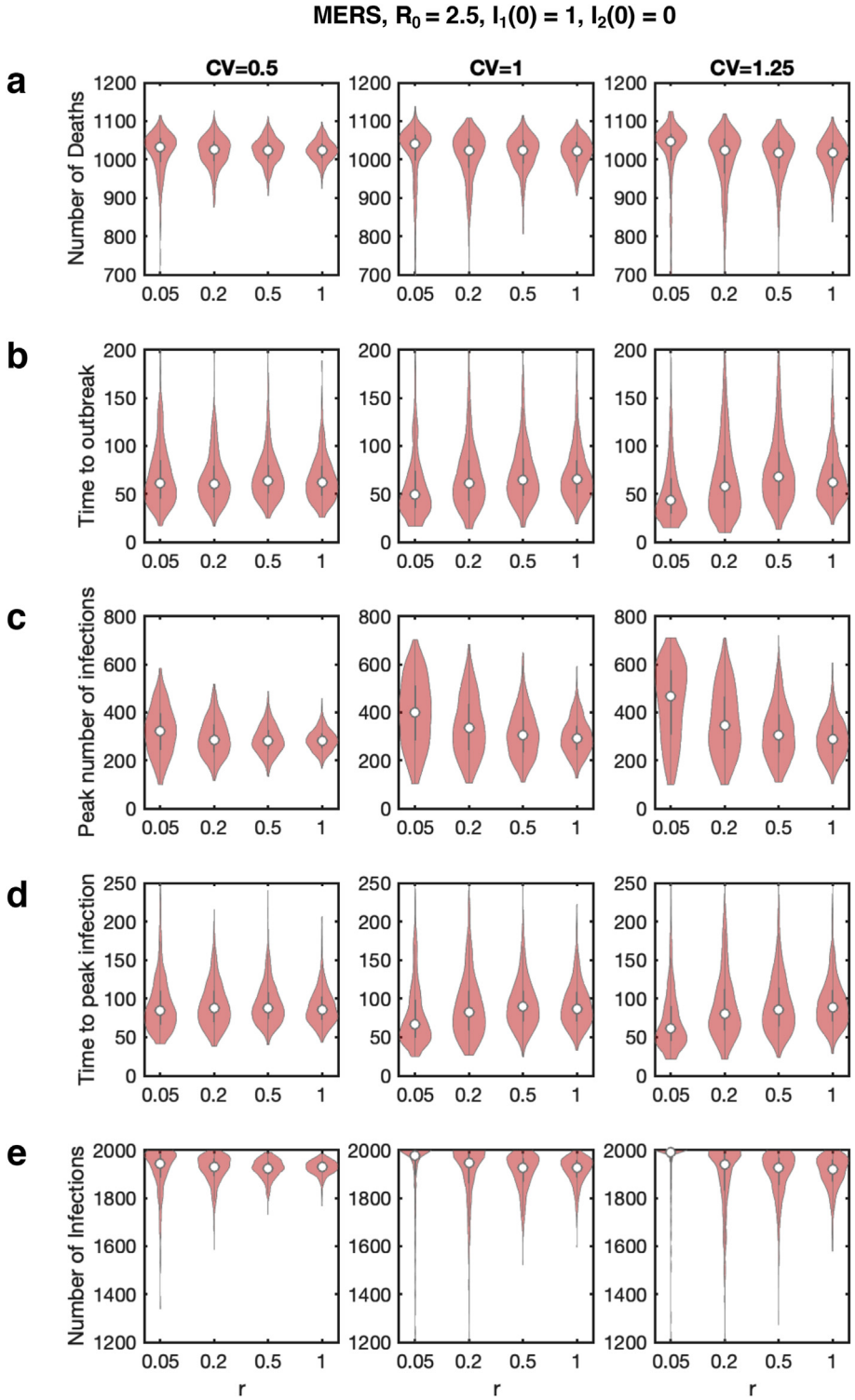

Fig. E.3. Violin plots showing impact of degree of environmental variability on MERS disease severity. Example violin plots predicted by the MERS NHP model with one initial infected NS ($I_1(0) = 1$, $I_2(0) = 0$) is shown for the following measures of disease severity for instances in which an outbreak occurs: (a) number of deaths, (b) time to outbreak, (c) peak number of infections, (d) time to peak infection, and (e) total number of infected individuals.

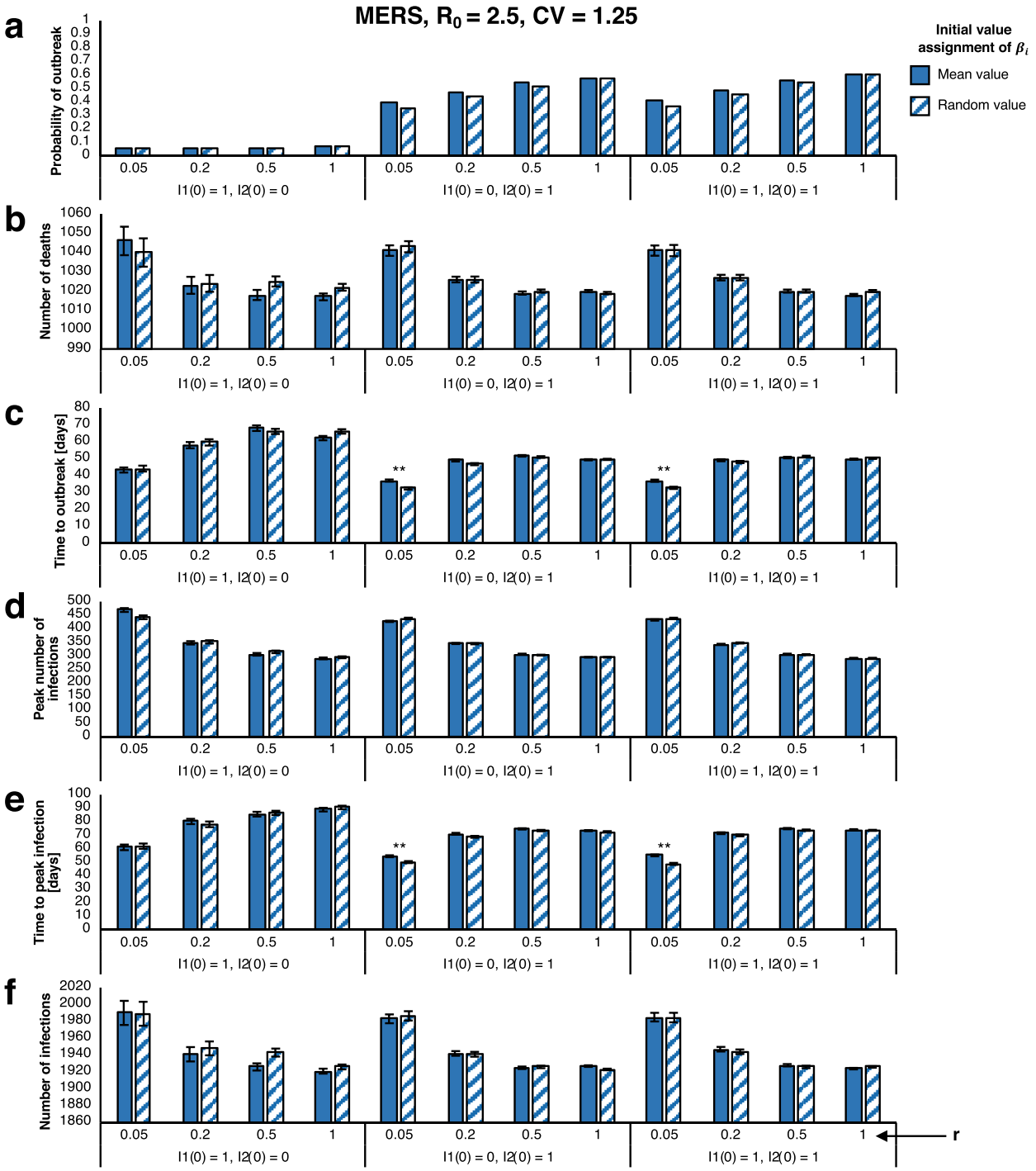

Fig. E.4. Exploring the effect of initial SS transmission rate on MERS NHP model prediction results. Predictions of the MERS NHP model with low initial conditions, a single infected NS or SS or one infected NS and SS ($I_1(0) = 1$, $I_2(0) = 0$; $I_1(0) = 0$, $I_2(0) = 1$; $I_1(0) = 1$, $I_2(0) = 1$) is shown for varied r values and $CV = 1.25$. The effect of using an initial transmission rate () set at the mean transmission rate (shown in sold blue) versus a randomly selected rate from the asymptotic gamma distribution of β_i (shown in striped blue) are compared. Bar graph data shows median values, error bars show standard error of the median ($n = 10,000$ simulations). Several measures are shown, including (a) probability of outbreak, as well as measures of disease severity for instances in which an outbreak is experienced: (b) number of deaths, (c) time to outbreak, (d) total number of infections, (e) peak number of infections, and (f) time to peak infection (** indicates significance).

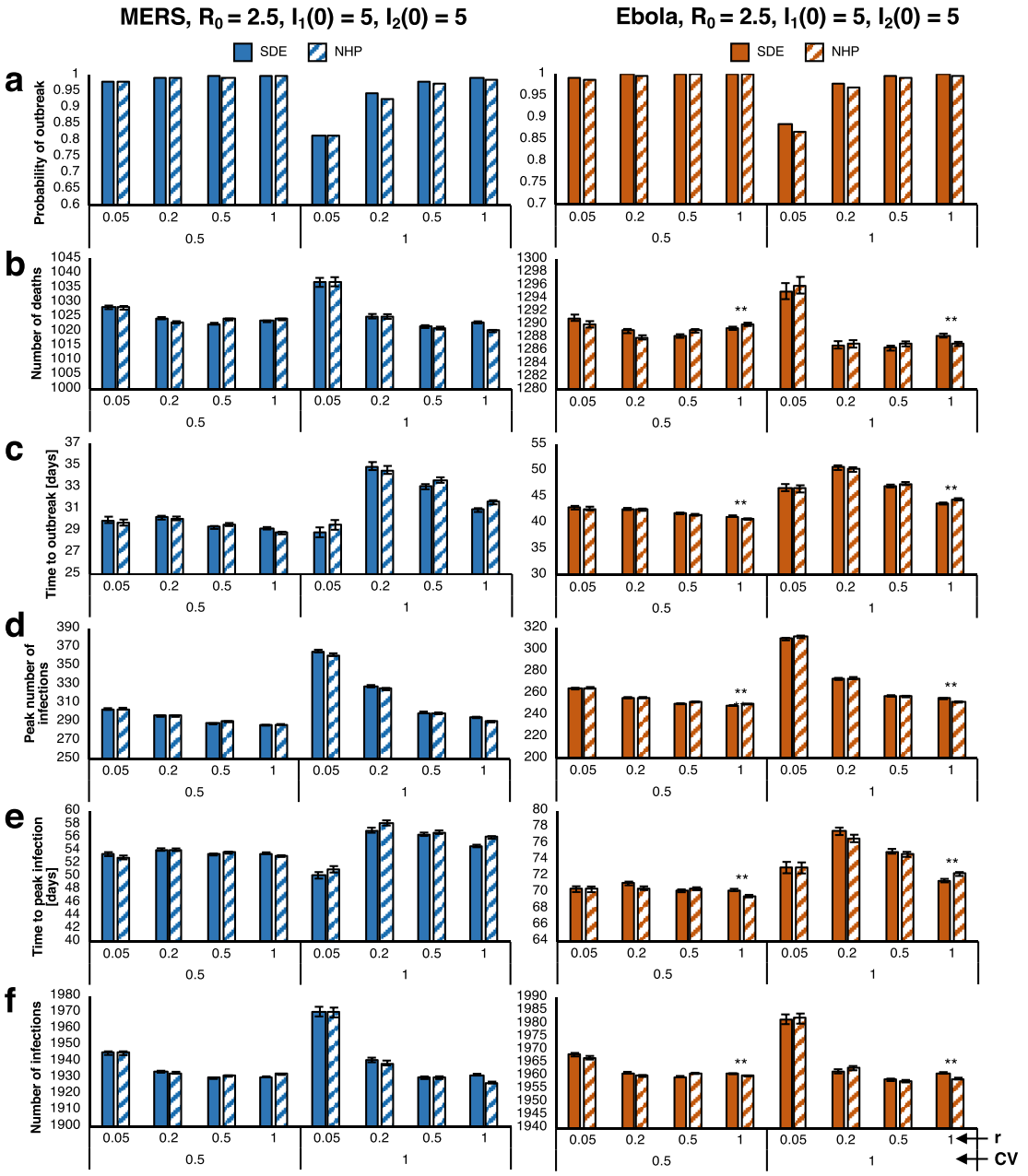

Fig. E.5. Comparing the performance of the NHP and SDE models with varied environmental variability and high initial conditions. Comparing predictions of the SDE versus NHP models for both MERS (blue) and Ebola (orange) with high initial condition ($I_1(0) = 5, I_2(0) = 5, E_1(0) = E_2(0) = A_1(0) = A_2(0) = 0$) for varied r and CV values. Bar graph data shows median values, error bars show standard error of the median ($n = 10,000$ simulations). Several measures are shown, including (a) probability of outbreak, as well as measures of disease severity for instances in which an outbreak is experienced: (b) number of deaths, (c) time to out, (d) total number of infections, (e) peak number of infections, and (f) time to peak infection (** indicates significance).

References

Acuña-Zegarra, M. A., Santana-Cibrian, M., & Velasco-Hernandez, J. X. (2020). Modeling behavioral change and COVID-19 containment in Mexico: A trade-off between lockdown and compliance. *Mathematical Biosciences, 325*, 108370.
Agusto, F. (2017). Mathematical model of Ebola transmission dynamics with relapse and reinfection. *Mathematical Biosciences, 283*, 48–59.
Allen, E. (2007). *Modeling with Itô stochastic differential equations* (Vol. 22). Springer Science & Business Media.
Allen, E. (2016). Environmental variability and mean-reverting processes. *Discrete & Continuous Dynamical Systems-B, 21*, 2073–2089.
Allen, E. J., Allen, L. J. S., Arciniega, A., & Greenwood, P. E. (2008). Construction of equivalent stochastic differential equation models. *Stochastic Analysis and Applications, 26*, 274–297.

Allen, L. J. S., Allen, E. J., & Jonsson, C. B. (2006). The impact of environmental variation on hantavirus. In *Mathematical studies on human disease dynamics: Emerging paradigms and challenges: AMS-IMS-SIAM joint summer research conference on modeling the dynamics of human diseases: Emerging paradigms and challenges, july 17-21, 2005* (Vol. 410, pp. 1–15). Snowbird, Utah: American Mathematical Soc.
Allen, E. J., Allen, L. J. S., & Smith, H. (2020). On real-valued SDE and non-negative SDE population models with demographic variability. *Journal of Mathematical Biology, 81*, 487–515.
Althaus, C. L. (2015). Ebola superspreading. *The Lancet Infectious Diseases, 15*, 507–508.
Beeching, N. J., Fenech, M., & Houlihan, C. F. (2014). Ebola virus disease. *British Medical Journal, 349*, g7348.
Bonyah, E., Badu, K., & Asiedu-Addo, S. K. (2016). Optimal control application to an Ebola model. *Asian Pacific Journal of Tropical Biomedicine, 6*, 283–289.
Brauer, F., & Castillo-Chavez, C. (2012). *Mathematical models in population biology and epidemiology* (Vol. 2). Springer.
Britton, T. (2010). Stochastic epidemic models: A survey. *Mathematical Biosciences, 225*, 24–35.
Cai, Y., Jiao, J., Gui, Z., Liu, Y., & Wang, W. (2018). Environmental variability in a stochastic epidemic model. *Applied Mathematics and Computation, 329*, 210–226.
CDC. (2016). Ebola (Ebola virus disease) 2014-2016 Ebola outbreak in west Africa. Available at https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/index.html (2020-05-13.
Chan, J. F., Lau, S. K., To, K. K., Cheng, V. C., Woo, P. C., & Yuen, K.-Y. (2015). Middle East respiratory syndrome coronavirus: Another zoonotic betacoronavirus causing SARS-like disease. *Clinical Microbiology Reviews, 28*, 465–522.
Chowell, G., Blumberg, S., Simonsen, L., Miller, M. A., & Viboud, C. (2014). Synthesizing data and models for the spread of MERS-CoV, 2013: Key role of index cases and hospital transmission. *Epidemics, 9*, 40–51.
Chowell, G., Cleaton, J. M., & Viboud, C. (2016). Elucidating transmission patterns from internet reports: Ebola and middle east respiratory syndrome as case studies. *Journal of Infectious Diseases, 214*, S421–S426.
Chowell, G., Hengartner, N. W., Castillo-Chavez, C., Fenimore, P. W., & Hyman, J. M. (2004). The basic reproductive number of Ebola and the effects of public health measures: The cases of Congo and Uganda. *Journal of Theoretical Biology, 229*, 119–126.
Chowell, G., & Nishiura, H. (2014). Transmission dynamics and control of Ebola virus disease (EVD): A review. *BMC Medicine, 12*, 196.
Cowling, B. J., Park, M., Fang, V. J., Wu, P., Leung, G. M., & Wu, J. T. (2015). Preliminary epidemiological assessment of MERS-CoV outbreak in South Korea, May to June 2015, eurosurveillance 20. URL: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=21163.
Cresson, J., & Sonner, S. (2018). A note on a derivation method for SDE models: Applications in biology and viability criteria. *Stochastic Analysis and Applications, 36*, 224–239.
Della Rossa, F., Salzano, D., Di Meglio, A., De Lellis, F., Coraggio, M., Calabrese, C., Guarino, A., Cardona-Rivera, R., De Lellis, P., Liuzza, D., et al. (2020). A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nature Communications, 11*, 1–9.
Dropkin, G. (2020). COVID-19 UK lockdown forecasts and R0. *Frontiers in Public Health, 8*, 256.
Edholm, C. J., Emerenini, B. O., Murillo, A. L., Saucedo, O., Shakiba, N., Wang, X., Allen, L. J. S., & Peace, A. (2018). Searching for superspreaders: Identifying epidemic patterns associated with superspreading events in stochastic models. In *Understanding complex biological systems with Mathematics* (pp. 1–29). Springer.
Eikenberry, S. E., Mancuso, M., Iboi, E., Phan, T., Eikenberry, K., Kuang, Y., Kostelich, E., & Gumel, A. B. (2020). To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infectious Disease Modelling, 5*, 293–308.
Eisenberg, J. N., Desai, M. A., Levy, K., Bates, S. J., Liang, S., Naumoff, K., & Scott, J. C. (2007). Environmental determinants of infectious disease: A framework for tracking causal links and guiding public health research. *Environmental Health Perspectives, 115*, 1216–1223.
Ekanayake, A. J., & Allen, L. J. S. (2010). Comparison of Markov chain and stochastic differential equation population models under higher-order moment closure approximations. *Stochastic Analysis and Applications, 28*, 907–927.
Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry, 81*, 2340–2361.
Glynn, J. R., Bower, H., Johnson, S., Turay, C., Sesay, D., Mansaray, S. H., Kamara, O., Kamara, A. J., Bangura, M. S., & Checchi, F. (2017). Variability in intra-household transmission of Ebola virus, and estimation of the household secondary attack rate. *The Journal of Infectious Diseases, 217*, 232–237.
Iacus, S. M. (2009). *Simulation and inference for stochastic differential equations: with R examples*. Springer Science & Business Media.
Jeong, D., Lee, C. H., Choi, Y., & Kim, J. (2016). The daily computed weighted averaging basic reproduction number $R_{0, k, \omega}^{n}$ for MERS-CoV in South Korea. *Physica A: Statistical Mechanics and Its Applications, 451*, 190–197.
Khan, A., Naveed, M., Dur-e Ahmad, M., & Imran, M. (2015). Estimating the basic reproductive ratio for the Ebola outbreak in Liberia and Sierra Leone. *Infectious Diseases of Poverty, 4*, 13. https://doi.org/10.1186/s40249-015-0043-3. URL.
Krishnarajah, I., Cook, A., Marion, G., & Gibson, G. (2005). Novel moment closure approximations in stochastic epidemics. *Bulletin of Mathematical Biology, 67*, 855–873.
Kurtz, T. G. (1970). Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of Applied Probability, 7*, 49–58.
Kurtz, T. G. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *Journal of Applied Probability, 8*, 344–356.
Kurtz, T. G. (1972). The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics, 57*, 2976–2978.
Lai, A., Poon, C., & Cheung, A. (2012). Effectiveness of facemasks to reduce exposure hazards for airborne infections among general populations. *Journal of The Royal Society Interface, 9*, 938–948.
Lau, M. S., Dalziel, B. D., Funk, S., McClelland, A., Tiffany, A., Riley, S., Metcalf, C. J. E., & Grenfell, B. T. (2017). Spatial and temporal dynamics of superspreading events in the 2014–2015. West Africa Ebola epidemic. *Proceedings of the National Academy of Sciences, 114*, 2337–2342.
Leavitt, J. W. (1996). *Typhoid mary: Captive to the public's health*. Beacon Press.
Li, S., Eisenberg, J. N., Spicknall, I. H., & Koopman, J. S. (2009). Dynamics and control of infections transmitted from person to person through the environment. *American Journal of Epidemiology, 170*, 257–265.
Lloyd, A. L. (2004). Estimating variability in models for recurrent epidemics: Assessing the use of moment closure techniques. *Theoretical Population Biology, 65*, 49–65.
Majumder, M. S., Kluberg, S. A., Mekaru, S. R., & Brownstein, J. S. (2015). Mortality risk factors for Middle East respiratory syndrome outbreak, South Korea, 2015. *Emerging Infectious Diseases, 21*, 2088.
Majumder, M. S., Rivers, C., Lofgren, E., & Fisman, D. (2014). Estimation of MERS-coronavirus reproductive number and case fatality rate for the spring 2014 Saudi Arabia outbreak: Insights from publicly available data. *PLoS Currents, 6*. https://doi.org/10.1371/currents.outbreaks.98d2f8f3382d84f390736cd5f5fe133c. URL.
Marion, G., Renshaw, E., & Gibson, G. (2000). Stochastic modelling of environmental variation for biological populations. *Theoretical Population Biology, 57*, 197–217.
Nardell, E. A., Keegan, J., Cheney, S. A., & Etkind, S. (1991). Airborne infection, Theoretical limits of protection achievable by building ventilation. *American Review of Respiratory Disease, 144*, 302–306.
Nii-Trebi, N. I. (2017). Emerging and neglected infectious diseases: Insights, advances, and challenges. *BioMed Research International, 2017*. https://doi.org/10.1155/2017/5245021. URL.
Ovaskainen, O., & Meerson, B. (2010). Stochastic models of population extinction. *Trends in Ecology & Evolution, 25*, 643–652.
Pankhurst, L. J., Anaraki, S., & Lai, K. M. (2012). Combining environmental assessment and contact investigations to make tuberculosis screening decisions. *International Journal of Tuberculosis & Lung Disease, 16*, 1023–1029.
Peeri, N. C., Shrestha, N., Rahman, M. S., Zaki, R., Tan, Z., Bibi, S., Baghbanzadeh, M., Aghamohammadi, N., Zhang, W., & Haque, U. (2020). The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: What lessons have we learned? *International Journal of Epidemiology, 49*, 717–726.

Rivers, C. M., Lofgren, E. T., Marathe, M., Eubank, S., & Lewis, B. L. (2014). Modeling the impact of interventions on an epidemic of Ebola in Sierra Leone and Liberia. *PLoS Currents, 6*.
Ross, S. M. (2014). *Introduction to probability models*. Academic Press.
Stein, R. A. (2011). Super-spreaders in infectious diseases. *International Journal of Infectious Diseases, 15*, e510–e513.
Tatem, A. J., Rogers, D. J., & Hay, S. I. (2006). Global transport networks and infectious disease spread. *Advances in Parasitology, 62*, 293–343.
Truscott, J., & Gilligan, C. (2003). Response of a deterministic epidemiological system to a stochastically varying environment. *Proceedings of the National Academy of Sciences, 100*, 9067–9072.
Van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry* (Vol. 1). Elsevier.
Varughese, M., & Fatti, L. (2008). Incorporating environmental stochasticity within a biological population model. *Theoretical Population Biology, 74*, 115–129.
Velásquez, G. E., Aibana, O., Ling, E. J., Diakite, I., Mooring, E. Q., & Murray, M. B. (2015). Time from infection to disease and infectiousness for Ebola virus disease, a systematic review. *Clinical Infectious Diseases, 61*, 1135–1140.
Webb, G., Browne, C., Huo, X., Seydi, O., Seydi, M., & Magal, P. (2015). A model of the 2014 Ebola epidemic in West Africa with contact tracing. *PLoS Currents, 7*.
WHO. (2017). Middle East respiratory syndrome coronavirus (MERS-CoV). Available at https://www.who.int/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov) (2020-05-15.
WHO. (2019). Environmental health in emergencies. Environmental factors influencing the spread of communicable diseases. Available at https://www.who.int/environmental_health_emergencies/disease_outbreaks/communicable_diseases/en/(2020-05-15.
WHO. (2020). Ebola virus disease. Available at https://www.who.int/news-room/fact-sheets/detail/ebola-virus-disease (2020-05-15.
Wong, G., Liu, W., Liu, Y., Zhou, B., Bi, Y., & Gao, G. F. (2015). MERS, SARS, and Ebola: The role of super-spreaders in infectious disease. *Cell Host & Microbe, 18*, 398–401.