

Modelling Ebola within a community

R. N. LEANDER*, W. S. GOFF, C. W. MURPHY AND S. A. PULIDO

Department of Mathematical Sciences, Middle Tennessee State University, Murfreesboro, TN, USA

*Received 25 September 2015; Final revision 10 February 2016; Accepted 24 February 2016;
first published online 28 March 2016*

SUMMARY

The 2014 Ebola epidemic was the largest on record. It evidenced the need for improved models of the spread of Ebola. In this research we focus on modelling Ebola within a small village or community. Specifically, we investigate the potential of basic Susceptible-Exposed-Infectious-Recovered (SEIR) models to describe the initial Ebola outbreak, which occurred in Meliandou, Guinea. Data from the World Health Organization is used to compare the accuracy of various models in order to select the most accurate models of transmission and disease-induced responses. Our results suggest that (i) density-dependent transmission and mortality-induced behavioural changes shaped the course of the Ebola epidemic in Meliandou, while (ii) frequency-dependent transmission, disease-induced emigration, and infection-induced behavioural changes are not consistent with the data from this epidemic.

Key words: Ebola virus, modelling.

INTRODUCTION

Ebola is a viral disease that is transmitted through contact with bodily fluids [1]. The 2014 Ebola epidemic, which began in the village of Meliandou, Guinea on 26 December 2013 [2] was by far the largest Ebola outbreak on record, and it evidenced the need for improvements in modelling Ebola. Even some complex models of the epidemic, which incorporate control measures, overestimated the epidemic's final size and duration. An example is provided by the Susceptible-Exposed-Infectious-Recovered (SEIR) model presented in [3] which considered the efficacy of various control measures on mitigating the epidemic. Here it was predicted that the most effective control scenario, which included 100% contact tracing and a 75% reduction in hospital

transmission, could not halt the spread of the disease. Moreover, in the absence of additional controls, the CDC estimated the epidemics in Liberia and Sierra Leone would include >500 000 cases by 20 January 2015 [4]. In contrast, as of 31 March 2015, there were <10 000 suspected cases in Liberia and <26 000 suspected cases in Guinea, Liberia, and Sierra Leone combined [5].

A common feature of the models above is that they seek to describe the mature Ebola epidemic at a national scale. However, at this large scale and late stage, spatial effects, inhomogeneities within the population, and evolving controls complicate efforts to model the epidemic. In order to improve the accuracy of Ebola forecasting and better understand the mechanisms through which this virus spreads, we evaluate models of an early Ebola outbreak in a small village. Our investigation employs data from the initial outbreak, centred in Meliandou, Guinea. Because this outbreak developed with limited external intervention and within a small village, it is well suited for testing the ability of basic

* Author for correspondence: Professor R. N. Leander, Department of Mathematical Sciences, Middle Tennessee State University, Murfreesboro, TN 37132, USA.
(Email: rachel.leander@mtsu.edu)

Table 1. *Parameters*

Parameter	Explanation	Value	Range	Ref.
$1/\alpha$	Average latent period*	10.7 day ⁻¹	2–19.4	[1]
ρ	Probability of death	0.707	0.667–0.743	[1]
$1/\lambda$	Average time to recovery†	16.3 day ⁻¹	10.2–22.4	[1]
R_0	Reproductive number	1.71	1.44–2.01	[1]
N	Population size	300	233–720	[8–10]
D	Population density	$\frac{1}{5.5}$	$\frac{1}{3.5} - \frac{1}{12.6}$	[7]

* The average latent period is taken as that for single-day exposures.

† The average time to recovery is taken as the average time from symptom onset to hospital discharge.

SEIR models, which assume a population is well mixed and homogeneous, to describe the spread of Ebola. We compare the accuracy of various models in order to select mechanisms of Ebola transmission and the population’s response. Our results led us to select density-dependent transmission and mortality-induced behavioural changes as important for shaping the course of an Ebola outbreak.

METHODS

Data from the Meliandou outbreak were taken from the World Health Organization’s (WHO) website [2]. These data include dates of onset and death for victims in the Meliandou chain, when available. From these data we extracted the total number of Ebola cases (measured as the total number of infectious, dead, and recovered) through time for comparison against model simulations. Due to incomplete data, the date of onset for two patients in generation 5 of the epidemic is unknown. The date of onset for one of the patients was assumed to be after 7 February. This estimate is based on the average incubation period of the disease, the average duration of the disease, and the fact that the individuals this person infected died in early April. The date of onset of the second patient was assumed to be after the date of onset of the patient that infected him.

Standard SEIR models, with both frequency- and density-dependent transmission are constructed according to [6].

Model equations

$$\frac{dS}{dt} = -\lambda I, \tag{1}$$

$$\frac{dE}{dt} = \lambda I - \alpha E, \tag{2}$$

$$\frac{dI}{dt} = \alpha E - \frac{\gamma}{1 - \rho} I, \tag{3}$$

$$\frac{dR}{dt} = \gamma I, \tag{4}$$

$$\frac{dD}{dt} = \frac{\gamma \rho}{1 - \rho} I. \tag{5}$$

Here S , E , I , and R denote the number of susceptible, exposed, infectious, recovered and dead but no longer infectious individuals in the population, respectively. The form of λ depends on the type of transmission. In particular, for frequency-dependent transmission $\lambda = \beta(S/N)$, and for density-dependent transmission $\lambda = \beta(S/A)$. Where A denotes the area over which the epidemic occurs, $N = S + E + I + R$, and the value of the parameter β is model-specific. Because of the relatively short time over which the epidemic evolves, we do not consider natural deaths and births in these models. It is also important to note that these models are considerably simpler than many other models of Ebola. In particular, they differ from more complex models of Ebola in that they do not include a separate class for those who are both dead and infectious. Instead the infectious class may be considered to include both living and dead infectious or, alternatively, the model may be considered to neglect infection through contact with the dead. In these models, the time to leave the infectious class is determined by the time to recover ($1/\gamma$) and the probability of death (ρ). For the baseline parameter values from Table 1, this leads to an average time to death of about 6.8 days. For comparison, the time from symptom onset to death was reported as 6.4 ± 5.3 days in Guinea [1].

To parameterize the model with density-dependent transmission, we take the most relevant measure of population density as floor area per individual within a family dwelling. This parameter is estimated using

[7], which reports the floor area per individual in several African nations. Since no parameter value is available for Guinea, we take the baseline value of this parameter from that of the closest nation for which data is available, Ghana, i.e. 5.5 [7].

Estimates of the size of the population of Meliandou vary. In particular, Vatican Radio reports that Meliandou was home to approximately 700 individuals before the epidemic and approximately 400 individuals after the epidemic [8], while a detailed map produced by the European Commission estimates that the entire region of interest, which includes two other villages of a similar size, was inhabited by approximately 719 individuals as of January 2014 [9]. In retrospect, the WHO has reported that the village includes only 31 households [10]. Assuming 10 members per household, the population size is then around 300, which is consistent with the data from the European Commission. Hence, we take the baseline population of Meliandou as 300.

The model parameters α , γ , R_0 and ρ are taken from a WHO publication [1]. These parameters are then used to estimate the value of β , as follows. In case of frequency-dependent transmission, we use

$$R_0 = \frac{\beta(1 - \rho)}{\gamma}.$$

In case of density-dependent transmission, we use

$$R_0 = \frac{\beta(1 - \rho)d}{\gamma},$$

where d denotes the initial density of the population in terms of individuals/m² of floor space.

In addition to the basic SEIR models described above, we consider models with disease-induced emigration. Although this mechanism is specific to small, well-defined populations, it has the effect of limiting the access of the infectious to the susceptible and so is, in some ways, similar to hospitalization. In developing these models we assume that the rate of emigration depends on either the number of infectious or the number of dead individuals. Specifically, in the case of density-dependent transmission, a term of the form $-u(I/A)X$ or $-u(D/A)X$ is added to the differential equation for variable X , for $X = S, E, I, R$, in order to model emigration in response to infection or mortality, respectively. Analogous terms are added to the equations of the basic frequency-dependent model in order to create the corresponding models with frequency-dependent emigration. In fitting these

models to the data, only the parameter u is varied. The other parameters, including β , remain fixed.

Finally, we consider the potential of behavioural changes to impact the course of an epidemic: Individuals may change their behaviour in response to the epidemic in an attempt to avoid infection. As was the case for the models with emigration, we consider models in which individuals change their behaviour in response to either the infectious or the dead. Furthermore, we assume that these behavioural changes lower the basic reproductive ratio of the disease according to either

$$R_0 = \frac{R_m}{kI/A + 1} \quad \text{or} \quad R_0 = \frac{R_m}{kD/A + 1},$$

in case of density-dependent transmission, and either

$$R_0 = \frac{R_m}{kI/N + 1} \quad \text{or} \quad R_0 = \frac{R_m}{kD/N + 1},$$

in case of frequency-dependent transmission. In these models, R_m represents the reproductive ratio of the disease in a population that is unaware of the presence of either an infectious or deadly pathogen, i.e. in the absence of any extra precautionary measures, and k determines the sensitivity of the population's response to the presence of the infectious or dead. In simulating these models, R_0 is recalculated and then used to update β at each time step.

Because estimates of the parameters R_m , k , and u are not available, these parameters are chosen to fit the data with the stipulation that the optimal parameter values must be biologically meaningful. Since the model parameters are uncertain, in some simulations the parameters for which estimates are available are also chosen to fit the data; however, these parameters are constrained to lie in a given range. For the parameters provided in [1], parameter ranges are determined by 95% confidence intervals (in the case of R_0) or mean values ± 1 standard deviation (in the case of γ and α). The other parameter ranges are delineated by the maximal and minimal values of the available parameter estimates. The optimal parameters are chosen to minimize the squared error between the model and the data (Matlab, lsqnonlin). The data is a time series of cumulative cases derived from data provided by the WHO, augmented with an additional data point at 500 days, in order to penalize parameterizations that overestimate the final size of the epidemic. The models are simulated in Matlab,

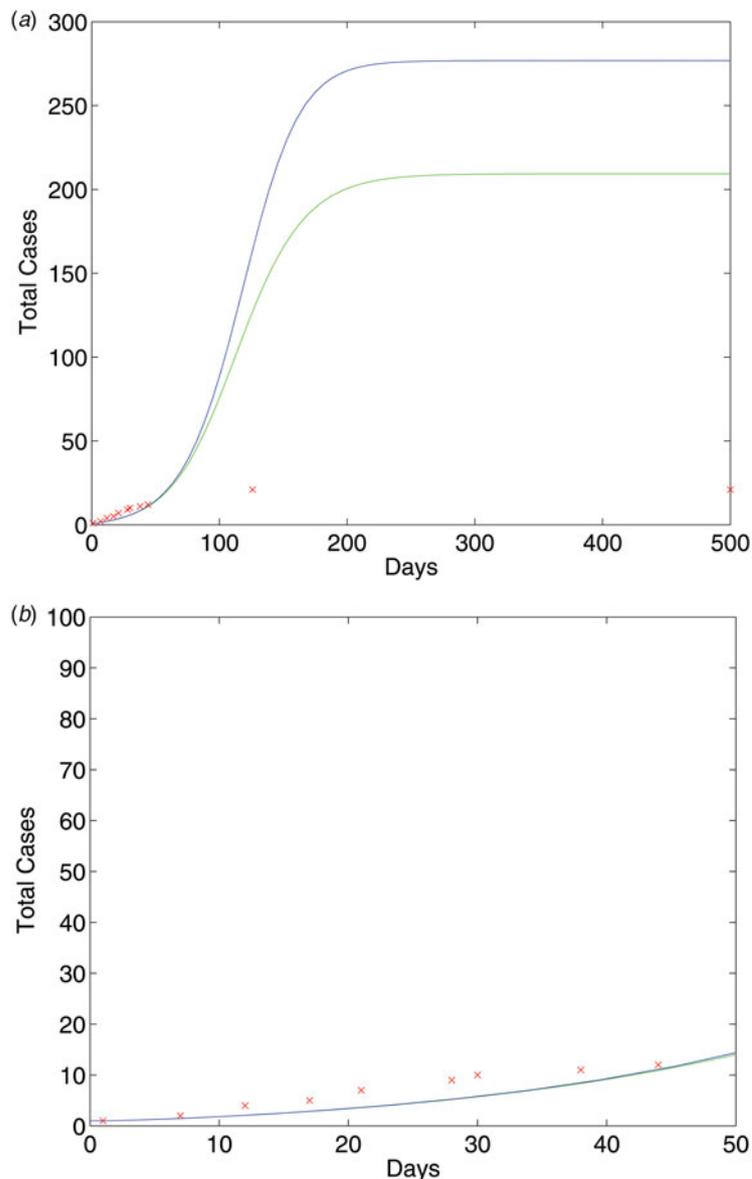


Fig. 1. (a) Simulations of the basic SEIR models with either density-dependent (green) or frequency-dependent (blue) transmission. The red \times s denote the data. (b) The early epidemic.

and model predictions of the extent and duration of the epidemic in Meliandou are compared to the data.

RESULTS

The initial Ebola outbreak that began in the village of Meliandou, Guinea on 26 December 2013 [2] is reported to have ended in April 2014 with a total of about 21 cases [1, 2, 11]. Simulations of this epidemic using the basic SEIR models with either frequency- or density-dependent transmission reveal that both models overestimate the size and duration of the epidemic. Indeed, when parameterized with the baseline

parameter values, the model with frequency-dependent transmission predicts that the epidemic will include approximately 277 cases which represents approximately 92% of the population, while the model with density-dependent transmission predicts that the epidemic will include approximately 209 cases, representing about 70% of the population (see Fig. 1a). However, both models provide a reasonably good fit to the initial epidemic (see Fig. 1b). One explanation for the poor fits of the models to the data is that the baseline parameter values are not representative of the epidemic in Meliandou. However, optimizing the parameters within the given parameter ranges, does not significantly

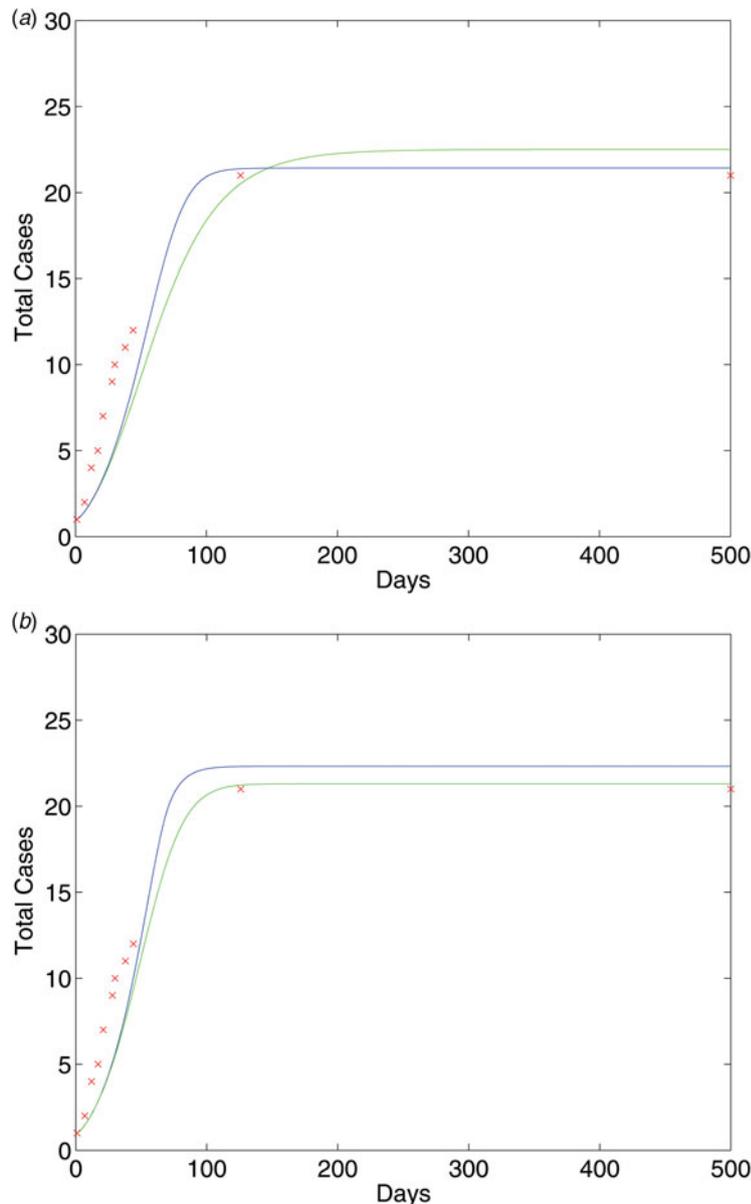


Fig. 2. (a) Simulations of the models with infection-induced emigration. For the density-dependent model $u^* = 15.8764$ per individual per day, the error is $E = 116.6211$, and the final population size is $N_f = 108$ individuals. For the frequency-dependent model $u^* = 4.8833$ per individual per day, $E = 99.3366$, and $N_f = 0$ individuals. (b) Simulations of the models with mortality-induced emigration. For the density-dependent model $u^* = 3.7825$ per individual per day, the error is $E = 86.4491$, and the final population size is $N_f = 0$ individuals. For the frequency-dependent model $u^* = 1$ per individual per day, $E = 82.4230$, and $N_f = 0$ individuals.

improve the fit to the data (not shown). Moreover, in optimizing the model parameters, the fit to the initial epidemic is lost.

A significant portion of Meliandou's population is reported to have fled in response to the epidemic [1, 11], and such a removal of individuals from the population could be important for shaping the course of an Ebola epidemic within a small village or community. Hence we augment the basic SEIR models to account

for disease-induced emigration. In so doing, we consider both infection- and mortality-induced emigration (see Methods section). Simulations of the resulting models, with the parameter, u , that determines the rate of emigration optimized, reveal that, after taking emigration into account, all models can accurately predict the epidemic's size and duration (see Fig. 2). However, the frequency-dependent model with either type of disease-induced emigration also predicts that

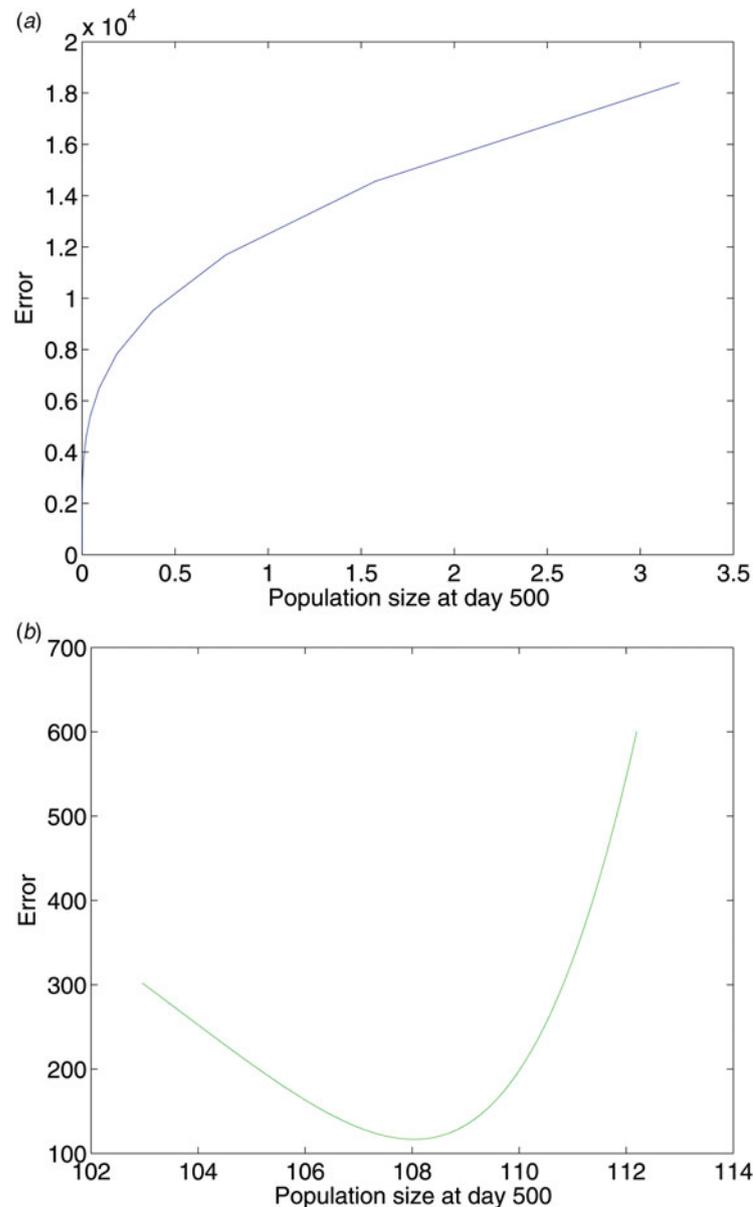


Fig. 3. Final population size vs. mean squared error for (a) frequency-dependent and (b) density-dependent models of transmission, with infection-induced emigration.

Meliandou will be abandoned. The density-dependent model with mortality-induced emigration also predicts that Meliandou will be abandoned. In contrast, the density-dependent model with infection-induced emigration predicts approximately 36% of the population remains. Thus, of the models tested here, the model with density-dependent transmission and infection-induced emigration appears the best able to match the dynamics of the epidemic while preserving the village of Meliandou. In order to further explore the relationship between the rate of fleeing, the mean squared error, and the final population size for the models

with infection-induced emigration, we vary the parameter u between $0.1u^*$ and u^* , where u^* denotes the optimal value of u . The results are shown in Figure 3, which plots the final population size vs. the mean squared error for each model. Looking at Figure 3 we see that the frequency-dependent model with infection-induced emigration is unable to simultaneously preserve the population of Meliandou and describe the epidemic's dynamics.

Although the model with infection-induced emigration and density-dependent transmission provides a good approximation of the duration and extent of

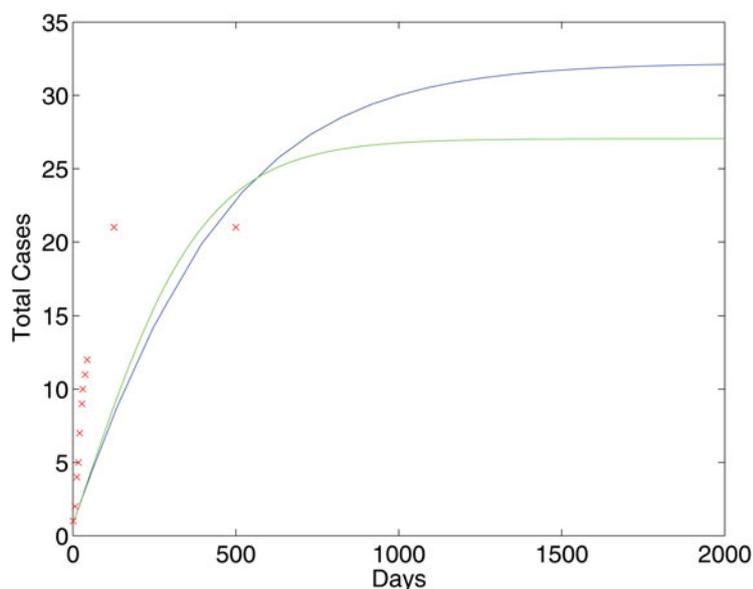


Fig. 4. Simulations of the models with R_0 dependent on the infectious. For the density-dependent model the optimal value of k is $k^* = 5.1977 \times 10^{-5}$, the maximal basic reproductive ratio is $R_m^* = 1.0109$, the error is $E = 422.7983$, and the final population size is $N_f = 281$ individuals. For the frequency-dependent model: $k^* = 0.0133$, $R_m^* = 0.9873$, $E = 440.6276$, and $N_f = 277$ individuals. The frequency-dependent model is shown in blue, the density-dependent model is shown in green.

the Meliandou outbreak, it does not capture the initial rapid increase in cases. Hence we develop additional models to test the hypothesis that disease-induced alterations in behaviour can explain these dynamics. We formulate models in which the basic reproductive number, and hence the transmission rate, falls as the frequency or density of the infectious or dead rises.

The models in which the transmission rate varies with the frequency or density of *infectious* do not fit the data (see Fig. 4). Indeed, in this case, the model with frequency-dependent transmission predicts the epidemic will slowly die out leaving a final population of about 277 individuals. The results are similar in the case of density-dependent transmission.

In contrast, the models in which the transmission rate varies with *mortality* provide an excellent fit to the data (see Fig. 5). In fact, these models successfully capture both the initial rapid increase in cases and the size and duration of the epidemic. The error of the model with density-dependent transmission is somewhat lower than that with frequency-dependent transmission, $E = 8.3390$ vs. $E = 33.2964$. The optimal values of R_m for the models are similar; $R_m = 5.8479$ for the frequency-dependent model and $R_m = 3.3924$ for the density-dependent model. This parameter represents the basic reproductive number for a population that is free of disease-induced mortality, i.e. it is the number of secondary infections that would result from a single infectious individual being

introduced into an entirely susceptible population that is unaware of the presence of a fatal disease. Although these values of R_m are somewhat larger than the estimated basic reproductive number of Ebola, they are not unreasonably large. Indeed, in the absence of any precautions it is reasonable that a single infectious individual could infect 3–5 others through the course of family-provided care and traditional funeral.

DISCUSSION

In this paper, we investigated the ability of simple SEIR-type models to describe the 2014 Ebola outbreak in the village of Meliandou in order to identify fundamental mechanisms that may shape the course of a local Ebola outbreak. We tested the accuracy of two types of transmission, (i) density-dependent transmission and (ii) frequency-dependent transmission, two types of response, (i) emigration and (ii) behavioural changes, and two response triggers, (i) the infectious and (ii) the dead. As is the case for the large-scale epidemic, we found that basic SEIR models, which do not include a population's response to the epidemic, greatly overestimate the size of the Meliandou epidemic, while SEIR models that describe how individuals respond to the epidemic can simultaneously capture both the epidemic's final size and dynamics. Specifically, models that include *mortality-induced* behavioural changes

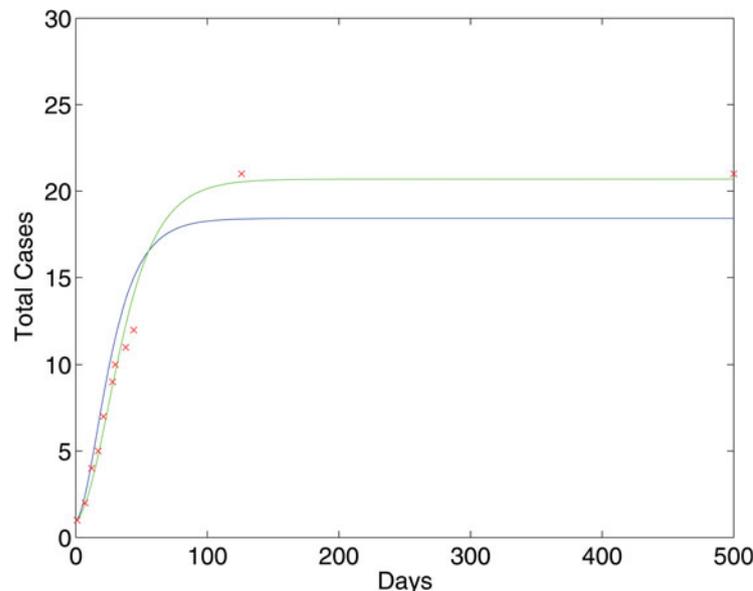


Fig. 5. Simulations of the models with R_0 dependent on mortality. For the density-dependent model the optimal value of k is $k^* = 817.1867$, the maximal basic reproductive ratio is $R_m^* = 3.3924$, the error is $E = 8.3390$, and the final population size is $N_f = 285$ individuals. For the frequency-dependent model: $k^* = 401.7542$, $R_m^* = 5.8479$, $E = 33.2964$, and $N_f = 287$ individuals. The frequency-dependent model is shown in blue, the density-dependent model is shown in green.

are able to closely approximate the rapid initial spread, duration, and final size of the epidemic. Our experiments also suggest that density-dependent transmission may provide a more accurate model of Ebola transmission. This finding is of particular interest since the form of transmission can impact control. For example, if Ebola transmission is primarily density-dependent, a successful vaccination campaign would be aimed at lowering the density, rather than the frequency, of susceptible individuals. The appropriateness of frequency- or density-dependent transmission for modelling disease transmission has been a topic of debate [12]. Although frequency-dependent transmission is often used to model human diseases [6], we hypothesize that frequency-dependent transmission may be inappropriate for modelling diseases like Ebola that are spread through close contact and have a high mortality rate. Indeed, while frequency-dependent transmission assumes a constant contact rate, the average contact rate in a population undergoing an Ebola-like epidemic is likely to fall. Because Ebola transmission requires close contact, which occurs primarily between family members, contacts that are lost on the death of an Ebola victim cannot be regenerated on the timescale over which the epidemic evolves. Although behavioural changes can also result in falling contact rates, the fact that in each of our simulations the frequency-dependent model is either excluded from consideration because it predicts the population

will fall to zero or is outperformed by the density-dependent model supports the idea that Ebola transmission is density-dependent.

Although we have gained some insights into the spread of Ebola at a local scale, it remains to develop a large-scale model of an Ebola epidemic. The model developed here is certainly not intended to describe Ebola transmission at this scale; however, its large-scale predictions may help guide future modelling efforts. Hence, taking the values of k and R_m from our study of Meliandou, we simulated an Ebola epidemic in the nation of Guinea. For this simulation, we took the total population size as 11 948 726, which represents this nation's population in 2013 [13]. The model predicts that Guinea would experience a rather short, though severe epidemic. This epidemic would end after approximately 250 days with a total of almost 800 000 cases, which represents almost 7% of the population. It is interesting to compare these predictions with those of the models presented in [3]. In particular, while the previous models tend to overestimate the duration of the epidemic, this model underestimates the duration of the epidemic. The speed with which this model predicts the disease will spread is likely overestimated due to the negligence of space. The large number of predicted cases can be considered as a worst-case approximation of the epidemic's final size, i.e. in the absence of a systematic

government-led effort to control the spread of the disease. While spatial effects are likely to be important for determining the speed of an epidemic, control measures, such as contact tracing and education, are likely to be important for determining the final size of an epidemic. In order to better describe the spread of the epidemic, the simple model selected here could be used to construct a meta-population model. Modelling control efforts, on the other hand, would require the introduction of additional parameters and/or compartments. For example, in a parsimonious model of contact tracing, individuals might exit the exposed class in order to enter a quarantine class. Here, the rate of exit could be determined by the probability of being traced before the onset of the infectious stage. The impact of education could be modelled in multiple ways. Education could prompt individuals to leave the infectious class for the quarantine class, or it could lower the transmission rate. In the latter case, the impact of education on the transmission rate could be included in the simple model above. For example, education could function to increase the sensitivity of the population to the death and/or infection of individuals.

In summary, we have used mathematical models to select mechanisms that are important for the local spread of Ebola. In the future these mechanisms may be important for designing control measures and designing large-scale models of Ebola in order to improve Ebola forecasting.

DECLARATION OF INTEREST

None.

REFERENCES

1. **World Health Organization Ebola Response Team.** Ebola virus disease in West Africa: the first 9 months of the epidemic and forward projections. *New England Journal of Medicine* 2014; **371**: 1481–1495.
2. **World Health Organization.** Ground zero in Guinea: the outbreak smoulders undetected for more than 3 months

- (<http://www.who.int/csr/disease/ebola/ebola-6-months/guinea/en/>). Accessed 16 March 2015.
3. **Rivers C, et al.** Modeling the impact of interventions on an epidemic of Ebola in Sierra Leone and Liberia. *PLOS Currents Outbreaks*. Published online: 16 October 2014. doi:10.1371/currents.outbreaks.fd38dd85078565450b0be3fcd78f5ccf.
 4. **Metzler M, et al.** Estimating the future number of cases in the ebola epidemic – Liberia and Sierra Leone, 2014–2015. 2014. *Morbidity and Mortality Weekly Report* 2014; **63**: 1–14.
 5. **Center for Disease Control (CDC).** 2014 Ebola outbreak in West Africa – case counts (<http://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/case-counts.html>). Accessed 20 February 2015.
 6. **Keeling M, Rohani P.** *Modeling Infectious Diseases in Humans and Animals*, 1st edn. Princeton, New Jersey: Princeton University Press, 2008, pp. 17.
 7. **United Nations, Department of Economic and Social Affairs, Population Division.** Annex table. Data for 12 indicators, by country. In: *Charting the Progress of Populations*. New York, NY, USA, 2000; ST/ESA/SER.R/151.
 8. **Faul M.** The village of Meliandou: Guinea's ground zero for the Ebola virus. *Vatican Radio*. Published online: 12 October 2014 (http://en.radiovaticana.va/news/2014/12/10/the_village_of_meliandou_guineas_ground_zero Ebola_/1114310). Accessed 20 June 2015.
 9. **Copernicus: The European Earth observation programme.** Meliandou-Guinea Other. 1 March 2014. Reference Map. Detail (http://emergency.copernicus.eu/mapping/system/files/components/EMSR076_01MELIANDOU_REFERENCE_DETAIL01_v2_300dpi.pdf). Accessed 20 June 2015.
 10. **World Health Organization (WHO).** Origins of the 2014 Ebola epidemic: one year into the Ebola epidemic. January 2015 (<http://www.who.int/csr/disease/ebola/one-year-report/virus-origin/en/>).
 11. **Beukes S.** Finding Ebola's patient zero. *Daily Maverick*. Published online: 28 October 2014 (<http://www.theguardian.com/world/2014/oct/28/ebola-virus-guinea-first-victim-patient-zero>). Accessed 16 March 2015.
 12. **McCallum H, Barlow N, Hone J.** How should pathogen transmission be modeled? *Trends in Ecology and Evolution* 2001; **16**: 295–300.
 13. **The World Bank.** Population, total (<http://data.worldbank.org/indicator/SP.POP.TOTL>). Accessed 5 February 2016.