A metapopulation model for the 2018 Ebola outbreak in Equateur province in the Democratic Republic of the Congo

Sophie R. Meakin[1, 2], Mike J. Tildesley[2], Emma Davis[1, 2], and Matt J. Keeling[2]

[1]EPSRC & MRC Centre for Doctoral Training in Mathematics for Real-World Systems, University of Warwick

[2]Zeeman Institute: SBIDER, Mathematics Institute and School of Life Sciences, University of Warwick

Abstract

Ebola is a viral haemorrhagic fever with high mortality that has caused a number of severe outbreaks in Central and West Africa, the largest of which was in 2014-16 and resulted in 11,325 deaths. Although most previous outbreaks have been relatively small in comparison, the result of managing outbreaks places huge strains on already limited resources. Mathematical models matched to early case reporting data can be used to identify outbreaks that are at high risk of spreading. Here we consider the Ebola outbreak in Equateur Province in the Democratic Republic of the Congo, which was declared on 8 May 2018. We use a simple stochastic metapopulation model to capture the dynamics in the three affected health zones (Bikoro, Iboko and Wangata) and the capital city Kinshasa. We are able to rapidly simulative a large number of realisations and use a likelihood-free method to determine parameters by matching between reported and simulated cases. This likelihood-free matching has a number of advantages over more traditional likelihood-based methods as it is less sensitive to errors in the data and is a natural extension to the prediction framework. Using data from 8 to 24 May 2018 we are able to capture the exponential increases in the number of cases in three locations (Bikoro, Iboko and Wangata) and the probability of transmission to capital city Kinshasa, although our estimated basic reproductive ratio of 4.02 is higher than for previous outbreaks. Using data until 28 June 2018 we are able to infer the decrease in transmission due to public-health intervention measures, such that the reproductive ratio is predicted to drop below one around 16 May 2018 leading to decreasing numbers of cases. We believe this method of fitting models to data offers a generic approach that can deliver rapid results in real time during a range of future outbreaks.

Author summary

Ebola is an infectious disease that, if left untreated, is often fatal. In addition, the consequence of managing Ebola outbreaks places huge strains on already limited healthcare resources in affected countries. Mathematical models can be useful in identifying high-risk outbreaks, and deciding where to allocate resources. However, many existing models of Ebola cannot capture the spatial spread of the outbreak, require highly detailed data, or are too complicated to be used in real-time during an outbreak.

In this paper we describe a framework that can capture the spatial spread of the 2018 Ebola outbreak in Equateur province in the Democratic Republic of the Congo, and is simple enough that it can be re-evaluated as new data become available. We used this framework to understand how transmission changed as a result of public-health interventions, and the risk of transmission to Kinshasa, which was a major public-health concern. This framework is highly flexible and can easily be adapted to new geographic regions, to use different data sources, or to answer pressing public-health questions. This research provides important insights into this Ebola outbreak, and has the potential to generate substantial impact on a range of pubic-health decision-making.

1 Introduction

1.1 Background

Ebola virus disease (herein referred to simply as Ebola) is a viral haemorrhagic fever that, if left untreated, is often fatal: the case fatality ratio is approximately 50%, but can range between 25% and 90% [1]. In addition to the high number of deaths, the consequence of managing Ebola outbreaks places huge strains on already limited healthcare resources and can reduce the control of other life-threatening diseases [2, 3, 4].

Fruit bats are thought to be the natural Ebola host [5], although non-human primates can also be infected. Animal-to-human transmission occurs through contact with the blood or organs of infected animals [1]; previous outbreaks have been linked to the handling and consumption of bushmeat [6]. Human-to-human transmission occurs via direct contact with blood or other bodily fluids of symptomatic individuals, or with contaminated materials such as bedding, clothing or needles [1]: as a result, healthcare workers and caregivers within the community are at high risk of infection [1, 2]. Transmission can also occur during traditional burial ceremonies that involve direct contact with the body [7], as individuals remain highly infectious even after death [1]. Funerals have been attributed to high levels of onward transmission during previous outbreaks [8].

The incubation period is typically between 2 and 21 days, during which time individuals are not infectious. Early symptoms include sudden onset of fever, fatigue, muscle pain, headache and sore throat, followed by vomiting, diarrhoea, rash, reduced kidney and liver function and, in some cases, internal and external bleeding [1]. Symptoms of Ebola are

non-specific, and typical of other diseases endemic to Central and West Africa such as malaria or typhoid fever. As a result, individuals may not seek immediate treatment and cases may be misdiagnosed, particularly at the beginning of an outbreak when surveillance is limited.

To date, outbreaks have occured in Central and West Africa, the largest of which was in 2014-16 and primarily located in Guinea, Liberia and Sierra Leone: the outbreak resulted in 11,325 deaths from 28,652 cases [9]. Other outbreaks have also been recorded in the Democratic Republic of the Congo (DRC), Uganda, South Sudan and Gabon, although each of these outbreaks have been at most 500 cases and located in predominantly rural areas.

1.2 2018 outbreak of Ebola virus disease in Equateur Province

The 2018 outbreak of Ebola in Equateur Province was the twenty-sixth outbreak globally and the ninth within the DRC. On 3 May 2018, 21 suspected cases of Ebola, including 17 deaths, were identified in the Ikoko-Impenge health area within Bikoro Health Zone in Equateur Province and reported to the DRC Ministry of Health. On 5 May, five active cases in hospitalised patients were identified and tested for Ebola by a team from the Ministry of Health, the WHO and Médicins Sans Frontières (MSF); of these five active cases, two tested positive for *Zaire ebolavirus*, one of six known ebolaviruses. The outbreak was officially declared by the World Health Organisation (WHO) on 8 May 2018, and concluded on 24 July 2018, 42 days (two maximum incubation periods) after the last confirmed case on 2 June 2018. At the time of writing a further unrelated outbreak of Ebola in North Kivu Province in the DRC is ongoing.

Data on the number of Ebola cases in each of the affected health zones was released daily by the DRC Ministry of Health [10] and less frequently by WHO [11]. Cases were classified as suspected, probable or confirmed according to WHO guidelines [12] and could undergo reclassification: for example, suspected or probable cases could be reclassified as confirmed positive or negative for Ebola. Cases were confirmed by laboratory testing by a method called reverse polymerase chain reaction; this technique requires access to specialised equipment and skilled technicians and so there is some delay between case detection and confirmation. A total of 52 cases were reported during the outbreak, of which 38 were classed as confirmed and 14 classed as probable [13].

Three health zones were affected over the course of the outbreak: Bikoro, where the outbreak was first detected; Iboko, a rural health zone to the east of Bikoro; and Wangata, located within Mbandaka, a city on the Congo river with a population of approximately 1,2000,000 people [14] (Figure 1). Of the 38 confirmed cases in the outbreak, 4 were reported in Wangata health zone, 10 in Bikoro and 21 in Iboko. Within Bikoro and Iboko health zones, the cases appear to be localised to a small number of remote villages [13]; however, data on the exact location of each case was not publicly available at the time of writing.

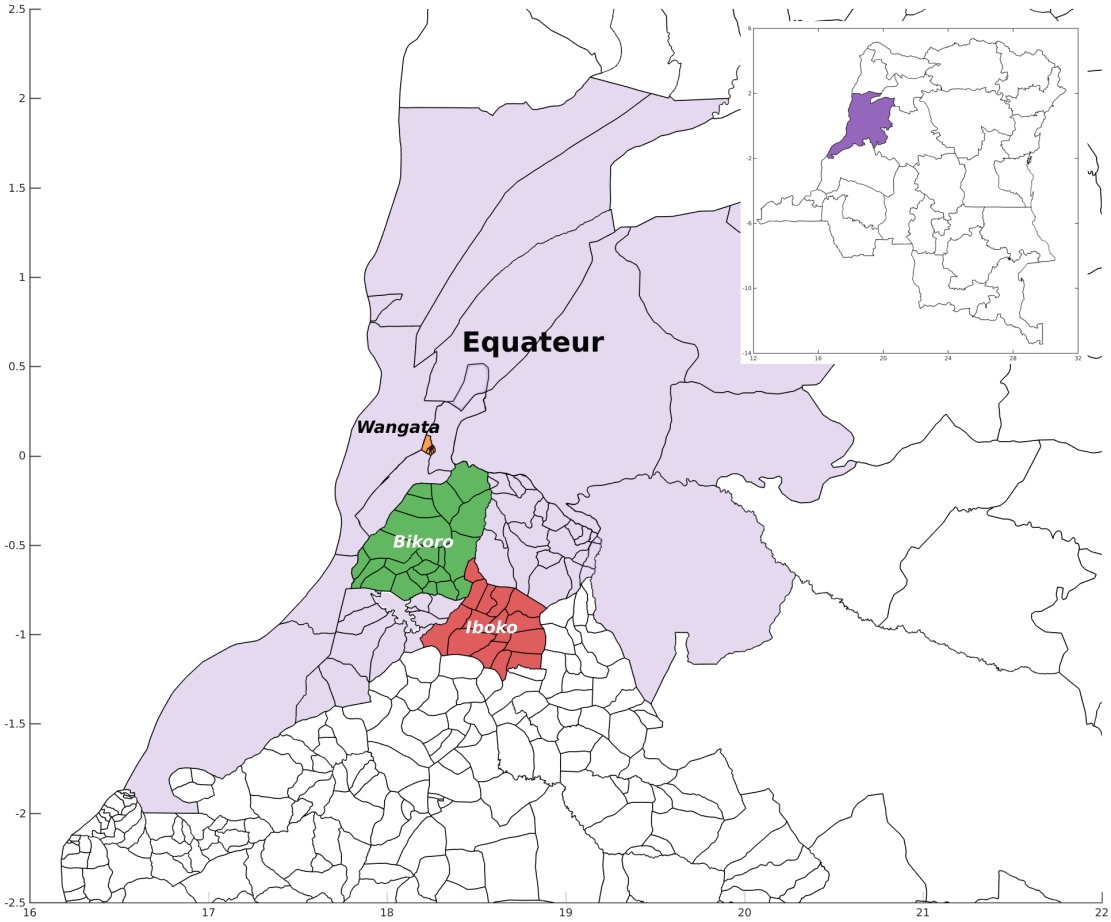

Figure 1. Location of the 2018 Ebola outbreak in Equateur Province (purple) in the Democratic Republic of the Congo, highlighting the three affected health zones: Iboko (red), Bikoro (green) and Wangata (yellow). Health zones are further subdivided into health areas. Data from Référentiel Géographique Commun (RGC), DRC Ministry of Health, OpenStreetMap, UCLA, and WHO. Borders do not necessarily reflect official legal boundaries.

Largely driven by experience and fears from the Ebola outbreak in West Africa, there was a rapid local and international response in an effort to prevent further transmission within or out of the DRC. In particular, there was considerable concern over the outbreak reaching the capital city Kinshasa, which has a population of approximately 11,000,000 people and serves international flights to countries in Africa and Europe. Organisations within the affected regions implemented various control measures, including: enhanced community surveillance; the use of rapid diagnostic tests; contact identification and tracing; safe and dignified burials (to reduce transmission during funerals); and vaccination [13]. Contact identification and tracing began on 5-6 May 2018 during the first field investigation. By 27 June 2018 all contacts had completed the 21-day follow-up period and no individuals were being actively monitored.

In addition to usual control measures, an experimental Ebola vaccine rVSV-ZEBOV was approved by the WHO for use as part of a targeted vaccination campaign [15]. This vaccine was studied in Phase 3 trials in Guinea during the 2014 outbreak in West Africa and was found to be highly protective against Ebola [16]; however, the exact efficacy of the vaccine is unknown. A ring vaccination strategy began on 21 May 2018 targeting contacts of confirmed Ebola cases, and contacts of these contacts, plus high-risk individuals such as local and international health-care workers in affected areas. By the end of the outbreak a total of 3,481 individuals had been vaccinated [13].

Real-time mathematical modelling of outbreaks has significantly increased in the last ten years, driven by developments in data collection and new statistical methods [17, 18, 19, 20, 21, 22], although efforts are constrained by limited or poor-quality data. This can be circumvented by using novel online data sources to augment traditional surveillance data [20, 22], or by using likelihood-free methods, such as approximate Bayesian computation (ABC) [18, 23]. ABC relies only upon efficient simulations of the epidemic and so is viable in low-resource settings where additional data are not available.

Despite the many challenges, mathematical modelling can provide important quantitative guidance to public health bodies about the likely scale of an outbreak, the efficacy of current intervention policies, and help to identify routes that can still sustain onward transmission. This information may be particularly beneficial in low-resource settings to target limited resources to regions of greatest need. Due to the severity of the disease and motivated by the need to allocate resources effectively, mathematical modelling of Ebola outbreaks has focussed predominantly on evaluating the efficacy of interventions [24, 25, 26, 27, 28, 29]. Modelling has also been used to assess the risk of international spread [30]; however, only a few were explicitly focussed on modelling in real-time [28, 30]. To date, and to the authors' knowledge, there has been no modelling to assess interventions in real-time during the outbreak in Equateur province; however, a recent epidemiological study of the outbreak [31] showed a decrease in the delay between symptom onset and hospitalisation and sample testing.

Compartmental models of Ebola infection are designed to capture possible transmission in different settings, namely the community, hospitals and funerals. The majority of results

(including those presented here) are based upon a six-compartment model introduced by Legrand et al. [32]; Kucharski et al. [24] use a variant of the six-compartment model that includes a compartment specifically for Ebola community care centres and Ebola treatment centres. Almost all existing Ebola modelling literature is concerned with the 2014-15 outbreak in West Africa, although some earlier analyses focus on the larger outbreaks in the DRC in 1995 [32, 25, 26] and Uganda in 2000 [32, 25].

Due to the close contact required for onward transmission and the severity of the symptoms, cases of Ebola are typically spatially clustered, with occasional long-distance transmission as a result of human movement. Despite this, only a few models allow for spatial dynamics [27, 30]. Merler et al. [27] develop an agent-based model for the 2014-16 outbreak in West Africa; however, parameterisation of agent-based models appears to require highly detailed sociodemographic data and computationally expensive Markov chain Monte Carlo simulations, limiting their usefulness during an outbreak. On the other hand, Gomes et al. [30] use a metapopulation modelling framework, whereby the population of the world is divided into sub-populations according to a Voronoi-like tessellation around transportation hubs. Metapopulation models require less detailed data and computational power than network or agent-based models, and so are an excellent framework to use in real-time outbreak modelling.

For the outbreak in Equateur province we are interested in assessing the early growth of cases, whether (and when) there were any notable change in the dynamics and the risk of Ebola reaching the capital city Kinshasa. To account for the spatial clustering of observed cases and to allow us to quantify the risk of case importation to Kinshasa, we used a metapopulation framework with four sub-populations representing the three affected health zones and Kinshasa. For the within-population model we adapted the widely-used six-compartment model [32], with a non-constant transmission parameter to allow for possible changes in transmission due to various interventions that were implemented.

2 Methods

2.1 Data

Over the course of the outbreak we obtained reports from the DRC Ministry of Health [10] and WHO [11], which were used to produce time series of the number of cases in each of Wangata, Bikoro and Iboko health zones. Due to the uncertainty of the true status of both suspected and probable cases, and the possibility of reclassification of these cases, in the following analysis we consider cumulative confirmed cases only. We stress that these data only give the date of laboratory confirmation, not the date of symptom onset or case detection, and therefore these dates are multiple steps removed from the underlying epidemiological dynamics. A time series of the cumulative data for the three health zones is given in Figure 2. We note that there is an error in the data for Bikoro on 17 May 2018, where the number of confirmed cases drops from 13 to 10. Although there is no explanation

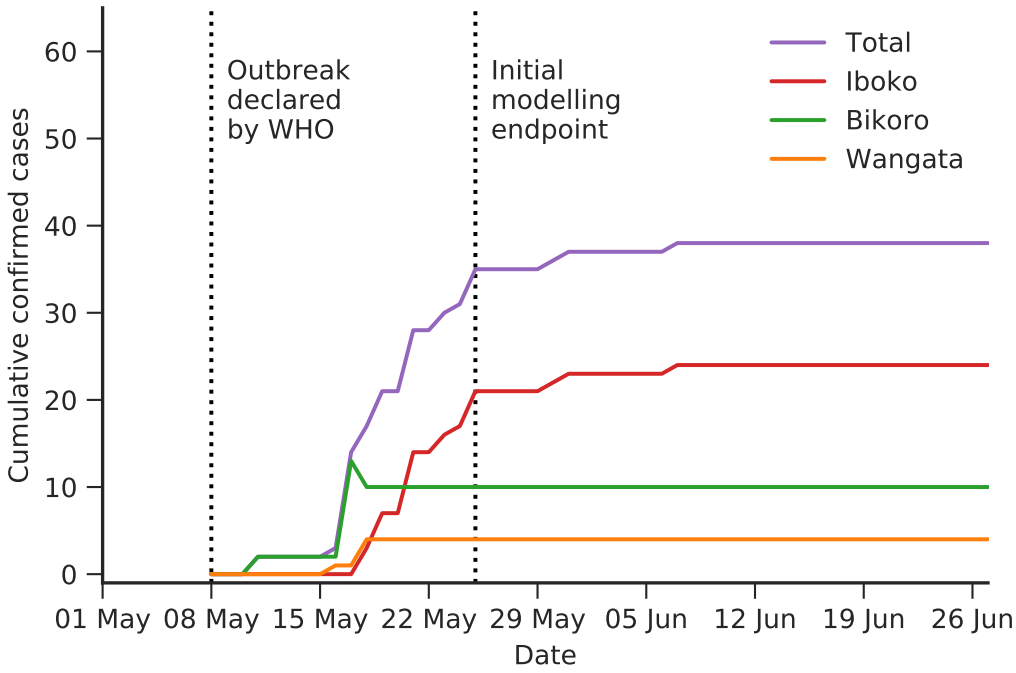

Figure 2. Time series of cumulative confirmed cases in the 2018 Ebola outbreak in Equateur Province in the DRC. The outbreak was officially declared by WHO on 8 May 2018. We also indicate the initial modelling endpoint on 25 May 2018.

from the DRC Ministry of Health, this error is most likely due to reporting before the case was officially confirmed by laboratory testing.

2.2 Model

Our model can be described in two parts: a compartmental model that describes the epidemiological dynamics of Ebola infection within a population [32]; and a metapopulation model that describes the spatial dynamics [33].

2.2.1 Epidemiological dynamics

We use a stochastic compartmental model introduced by Legrand et al. [32] to describe the dynamics of Ebola infection within a single isolated population, with the following 6 compartments: susceptible individuals (S), who can be infected after contact with infectious individuals; exposed individuals (E), who are infected but not yet infectious to others;

infectious individuals (I) within the community; hospitalised infectious individuals (H); dead individuals (F), who are still infectious and may transmit infection during burial; and removed individuals (R), who are either recovered or dead and safely buried. Individuals move between these compartments according to certain rates, determined by a set of ten parameters, including a scaling parameter $\hat{\beta} > 0$ that scales each of the transmission parameters β_I, β_H and β_F, associated with transmission in the community, hospitals and funerals. Figure 3 shows a schematic representation of the model, while Table 1 summarises the possible events and transition rates.

Event	Transition	Rate
Infection	$(S, E) \to (S-1, E+1)$	$\hat{\beta} S(\beta_I I + \beta_H H + \beta_F F)/N$
Symptom onset	$(E, I) \to (E-1, I+1)$	αE
Hospitalisation	$(I, H) \to (I-1, H+1)$	$\theta_1 \gamma_{IH} I$
Death (community)	$(I, F) \to (I-1, F+1)$	$\delta_1(1-\theta_1)\gamma_{IF} I$
Recovery (community)	$(I, R) \to (I-1, R+1)$	$(1-\delta_1)(1-\theta_1)\gamma_{IR} I$
Death (hospital)	$(H, F) \to (H-1, F+1)$	$\delta_2 \gamma_{HF} H$
Recovery (hospital)	$(H, R) \to (H-1, R+1)$	$(1-\delta_2)\gamma_{HR} H$
Burial	$(F, R) \to (F-1, R+1)$	$\gamma_{FR} F$

Table 1. A summary of possible transitions and their rates for the compartmental model used to describe Ebola dynamics within a population of size N. There are ten free parameters in the model and five additional parameters that are calculated from the first ten. Model parameters are: β_I community transmission rate; β_H hospital transmission rate; β_F funeral transmission rate; α^{-1} mean duration of incubation period; γ_{IH}^{-1} mean time from symptom onset to hospitalisation; γ_{IF}^{-1} mean time from symptom onset to death; γ_{IR}^{-1} mean time from symptom onset to end of infectious period for survivors; γ_{HF}^{-1} mean time from hospitalisation to death; γ_{HR}^{-1} mean time from hospitalisation to end of infectious period for survivors; γ_{FR}^{-1} mean time from death to traditional burial. θ_1 is calculated from θ such that θ% of infectious cases are hospitalised; δ_1 and δ_2 are calculated from δ such that the case fatality ratio is δ. Calculations are given in Appendix A.

2.2.2 Spatial dynamics

To describe the spatial dynamics of Ebola infection we use a metapopulation model, whereby the total population is split into K interacting sub-populations of sizes $N_i, i = 1, \ldots, K$. We define σ_{ij} to be the proportion of epidemiologically relevant contacts that individuals from population i have with individuals in population j, which we will simply refer to as the coupling from population i to population j. We naturally have that $\sum_{j=1}^{K} \sigma_{ij} = 1$, and so the within population coupling (which we expect to be close to one) can be expressed as $\sigma_{ii} = 1 - \sum_{j \neq i} \sigma_{ij}$. The force of infection in population i, the rate at

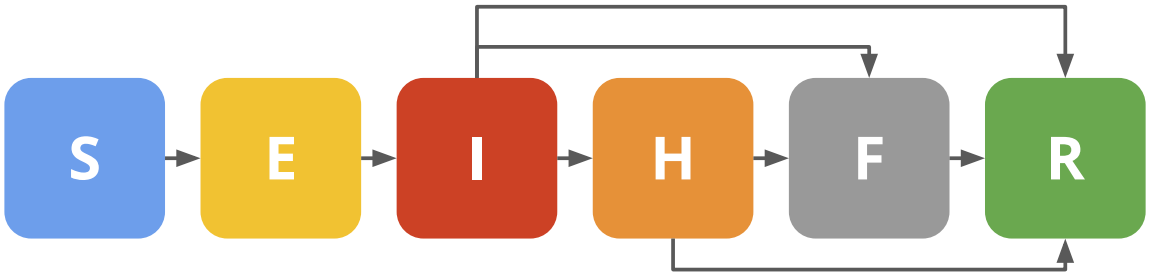

Figure 3. A schematic of the compartmental model used to describe Ebola dynamics within a population, where individuals can be classified as susceptible (S), exposed (E), infected (I), hospitalised (H), dead but not yet buried (F) or removed (R).

which susceptible individuals become infected, can then be written in terms of the coupling parameters as:

$$\hat{\beta}\left(\sum_{j=1}^{K}\frac{\sigma_{ij}}{N_j}\left(\beta_I I_j+\beta_H H_j+\beta_F F_j\right)\right). \tag{1}$$

As such the transmission is assumed to be due to the movement of healthy susceptible individuals visiting infected locations, such that the risk to individuals in population i is related to the coupling terms σ_{ij}.

The coupling is usually parameterised using mobility data [34, 35], or some suitable and available proxy, such as mobile phone data [36]. Such mobility data is not available for the DRC, and some remote regions (including some rural areas of Equateur province) are not covered by mobile phone networks. In the absence of such data, we use a generalised gravity model [37] to generate the coupling parameters.

We define v_{ij} to be the the number of visits from population i to j. According to the generalised gravity model this is proportional to $N_i^a N_j^b/d_{ij}^c$, where d_{ij} is the distance between populations i and j and the parameters a, b and c are to be inferred. The coupling, σ_{ij}, should then be proportional to the fraction of individuals from population i visiting population j, which is v_{ij}/N_i. In addition, we need to ensure that within population coupling $\sigma_{ii} = 1-\sum_{j\neq i}\sigma_{ij}$ is always positive, placing a limit on the maximum size of external couplings. To this end, we normalise the proportion of visits by the maximum proportion over all populations; however, from this definition we have $\min_i \sigma_{ii} = 0$. Therefore, we introduce an additional scaling parameter $A \in [0, 1]$ such that $\min_i \sigma_{ii} \geq 0$. Combining each of these elements we define the coupling $\sigma_{ij}, i \neq j$, to be:

$$\sigma_{ij} = A \frac{N_i^{a-1} N_j^b / d_{ij}^c}{\max_k \left(\sum_{l \neq k} N_k^{a-1} N_l^b / d_{kl}^c \right)}, \tag{2}$$

where a, b, c and A are parameters to be inferred from the epidemiological dynamics.

2.2.3 Parameters and inference

We use the above model to describe the dynamics of the Ebola outbreak in $K = 4$ populations: Wangata, Bikoro and Iboko health zones, and Kinshasa. We begin each realisation on 5 April 2018 (identified as the date of symptom onset for the first case) with a single infected individual in Bikoro and run the simulations until some later date T_1: we use 25 May 2018 as our initial modelling endpoint, and 29 June 2018 as the later modelling endpoint. These simulations are rapid: when $T_1 = 25$ May 2018 we can perform around 1,000,000 simulations an hour. Individual realisations of the model are simulated using the tau-leaping algorithm[38], with time intervals of $\tau = 1$ day, implemented in C++.

There is insufficient data resolution (in terms of individual dynamics or who-infected-who) or duration of outbreak to effectively infer all the epidemiological parameters needed for the simulation model. We therefore take the majority of the parameter values from [32] which were estimated from the 1995 Ebola outbreak in Kikwit city in the south-west of DRC (Table 2) and infer the transmission scaling parameter ($\hat{\beta}$) and four spatial parameters (A, a, b, c) from the data.

To parameterise the spatial component of the model, we need an estimate of both population sizes and pairwise distances between populations. We use estimates of population sizes from global health bodies [14, 44] and define pairwise distances to be great-circle distances between populations (Table 3).

We perform parameter inference using approximate Bayesian computation, a flexible likelihood free approach that is straightforward to compute. A traditional likelihood based methodology would require us to infer infection times and subsequent disease progression (together with associated times) for each case, and is likely to be affected by the temporal aggregation of observed cases as a result of detection and testing. Therefore for each realisation we calculate an error ϵ between the realised (C^{sim}) and observed (C^{obs}) cumulative confirmed cases from 11 May 2018 until some later date T_1. We define confirmed cases in our model to be individuals who have moved from the infected class to either the hospitalised or funeral class. We define the error as a weighted root mean square error, summed over the four sub-populations:

$$\epsilon(T_1) = \sum_{i=1}^{4} \sqrt{\sum_{T_0 \leq t \leq T_1} \frac{\left(C_i^{obs}(t) - C_i^{sim}(t)\right)^2}{\max(C_i^{obs}(T_1), 1)}}. \tag{3}$$

Parameter	Value	Reference
Community transmission rate, β_I	0.588 weeks^{-1}	[32]
Hospital transmission rate, β_H	0.794 weeks^{-1}	[32]
Funeral transmission rate, β_F	7.653 weeks^{-1}	[32]
Mean duration of incubation period, α^{-1}	7 days	[39, 40, 41]
Mean time from onset to hospitalisation, γ_{IH}^{-1}	5 days	[42]
Mean time from onset to death, γ_{IF}^{-1}	9.6 days	[42]
Mean time from onset to end of infectiousness for survivors, γ_{IR}^{-1}	10 days	[41, 43]
Mean time from death to traditional burial, γ_{FR}^{-1}	2 days	[32]
Proportion of cases hospitalised, θ	0.8	[42]
Case fatality ratio (CFR), δ	55%	[42]
Mean time from hospitalisation to death, γ_{HF}^{-1}	4.6 days	Appendix A
Mean time from hospitalisation to end of infectiousness for survivors, γ_{HR}^{-1}	5 days	Appendix A
Rate of transition from infectious to hospitalised, θ_1	0.67	Appendix A
Effective CFR in infected class, δ_1	54%	Appendix A
Effective CFR in hospitalised class, δ_2	53%	Appendix A

Table 2. A summary of the epidemiological parameters used in the compartmental model, estimated from the 1995 Ebola outbreak in Kikwit, DRC.

The denominator in this expression is motivated by considering a Poisson distribution. In a Poisson distribution the variance is equal to the mean, therefore we would normalise by dividing through by the observed value at each point; however, given the associated uncertainties in the data, this placed far too much emphasis on correctly matching to the early dynamics when the cases were low. We therefore normalise by the maximum of the observed cases in each location, providing some degree of normalisation between the different sized outbreaks. This approach would fail for Kinshasa where no cases were reported, so we take the normalisation constant to be one.

Parameter inference for the five parameters $(\hat{\beta}, A, a, b, c)$ is performed as a two step process. Firstly, parameter values are chosen from appropriate prior distributions reflecting our belief about their values: Uniform$(0, 1)$ for a, b, and A; and Uniform$(0, 3)$ for c and $\hat{\beta}$. From 10,000,000 random parameter choices and stochastic simulations we choose the 1,000 that have the lowest error, noting that the error is dependent on both the parameter values and the stochastic nature of any one simulation. Secondly, we use these 1,000 best parameter sets, and select nearby parameters to test (normally distributed about one of the best 1,000); we continually update the 1,000 best parameters to reflect newly tested parameters. In this manner, we can hone in on parameter sets that are associated with the lowest errors, and we are no-longer constrained by our initial choice of sample distribution.

Population	Pop. size	Location	Great-circle distance (km): Kinshasa	Wangata	Bikoro	Iboko
Kinshasa	1.1×10^7	04°18′05″S 15°19′05″E	0.0	583.4	520.9	530.2
Wangata	1.2×10^6	00°02′52″N 18°15′21″E		0.0	96.3	101.4
Bikoro	124,608	00°48′00″S 18°25′60″E			0.0	17.5
Iboko	80,027	00°49′38″S 18°35′18″E				0.0

Table 3. A summary of the demographic data used to parameterise the spatial model. For each population of interest, we give approximate population size and location, from which we can derive pairwise great-circle distances between populations. Population size estimates are from census data [44], or, where not available, through other online sources [14].

2.2.4 Model extension

Our simple model is likely to predict long-term exponential growth of infection until the susceptible population sizes become depleted. In practise sustained exponential growth is rarely observed as a range of control and mitigation measures are usually implemented to limit the spread of infection. We believe that the national and international response to the outbreak in Equateur Province is therefore likely to have substantially reduced the per-capita transmission and hence brought the outbreak under control.

We capture possible changes in transmission by including a step-change in $\hat{\beta}$ governed by two additional parameters. The transmission scaling, $\hat{\beta}$ now becomes a function of time such that:

$$\hat{\beta}(t) = \begin{cases} \hat{\beta}^* & \text{if } t < T_C \\ \delta\hat{\beta}^* & \text{otherwise} \end{cases} \tag{4}$$

where T_C defines the time at which control effects begin, and $(1 - \delta)$ determines the reduction in transmission from all infectious classes.

These two new parameter values are inferred as detailed above, using the associated errors. We start with the first 50 days of the outbreak (to 24 May 2018) and determine the best 1,000 parameter sets (for all 7 parameters) for 10,000,000 parameter updates; these are then fed into simulations for the first 55 days (to 29 May 2018) and we again iterate the fitting procedure for a further 10,000,000 steps. This incremental fitting processes mimics what would occur in real-time as more data becomes available on a daily basis and parameters are refined. The re-fitting procedure is easily completed by an over-night run. We perform incremental re-fitting until we reach 85 days (29 June 2018) by which point we have had three weeks without any additional confirmed cases in the region.

3 Results

3.1 Early time estimates

We run our simple model (without any change in transmission) until 25 May 2018 and retain the 1,000 realisations with the smallest total error. We obtain estimates of the five unknown parameters (Figure 4 and Figure B.1) and time series fit to the observed cumulative cases (Figure 5). We note that these time series represent stochastic realisations that accurately match to the data by minimising the error, rather than capturing the expected behaviour associated with the parameters.

From the model fitting process we obtain best-fit distributions of the five unknown parameters: the transmission rate scaling $\hat{\beta}$ and the four spatial coupling parameters a, b, c and A. The posterior distributions are summarised in Figure 4; individual posterior distributions for these parameters and the pairwise relationships between them are given in the Supplementary Information (Figure B.1). The parameter $\hat{\beta}$ scales the transmission rate from the 2018 outbreak in Equateur Province compared to parameters matched to the 1995 outbreak in Kikwit; the posterior distribution for $\hat{\beta}$ is given in Figure 4a. We get an estimate for $\hat{\beta}$ of 1.98 (95% credible interval (CI) [1.17, 2.95] – this is the interval that contains 95% of all the parameter values) and so an estimate for the basic reproductive ratio, R_0, of 5.34 (95% CI [3.17, 7.96]). We recombine the four spatial parameters to obtain meaningful distributions of the coupling between populations. In Figure 4b we include only the most relevant interactions: coupling between Bikoro and Iboko; from Bikoro and Iboko to Wangata; and from each of Wangata, Bikoro and Iboko to Kinshasa. From these results we observe that the coupling between populations is primarily dominated by distance, with the largest coupling between the closest populations, Bikoro and Iboko. In addition, we find that coupling is slightly larger towards the bigger population, so that the coupling from Iboko to Bikoro is larger than the from Bikoro to Iboko, since Bikoro is larger than Iboko.

From the model fitting process we also obtain time series fits to the observed cumulative cases and final distribution of cases corresponding to 25 May 2018 for the metapopulation as a whole (Figure 5a) and for the four sub-population separately (Figure 5b). For the time series we obtain a good qualitative fit to the observed confirmed cases both at the sub-population and metapopulation scale. The individual replicates (shown in grey) are tightly clustered and generally envelope the cumulative reported cases; however, our model noticeably overestimates the reported number of confirmed cases early in the outbreak (up to approximately 19 May); we believe that this overestimate is due to the process by which cases are classed as confirmed. In general, the model estimates improve over time. We compare the final distribution of the realised cumulative confirmed cases to the observed cumulative confirmed cases at the end of the simulation, corresponding to the 25 May 2018. Overall our model replicates the observed data well: our model estimates the total number of confirmed cases across the entire metapopulation to be 34 (mean 34.0, 95% CI [27, 41]),

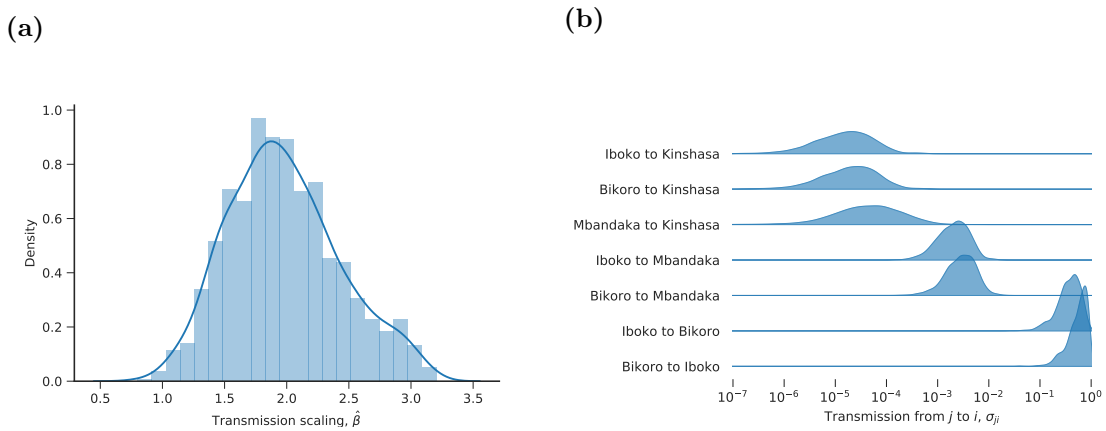

Figure 4. Posterior distributions determined by the 1,000 realisations with the smallest total error fitted to data from 5 April to 25 May 2018 for (a) the transmission scaling parameter $\hat{\beta}$ and (b) select between-population coupling. Here we refer to the transmission, which is the direction in which the disease moves; however the coupling is determined by the movement of susceptible individuals who bring infection back into their home population.

compared to 35 observed confirmed cases. We can also examine the distribution of cases in each of the four sub-populations. For Kinshasa and Wangata our model replicates the observed data extremely well: we estimate the number of confirmed cases to be 0 (mean 0.03, 95% CI [0, 1]) and 5 (mean 4.62, 95% CI [3, 7]) in Kinshasa and Wangata, respectively, compared to 0 and 4 observed confirmed cases. Our model is less accurate for Bikoro and Iboko, although the data still fall within our credible intervals: our model estimates the number of confirmed cases in Bikoro to be 13 (mean 12.54, 95% CI [9, 16]), compared to 10 actual observed cases; in Iboko our model estimates the number of confirmed cases to be 17 (mean 16.81, 95% CI [13, 22]), compared to 21 actual observed confirmed cases.

3.2 Effect of intervention

Our simple model assumes that the transmission rates remain constant over the course of the outbreak and would continue to predict exponential growth of cases if allowed to continue running forward from 25 May 2018; however, this contradicts the observed decline in the rate of observing new cases as various public-health measures come into effect. We therefore consider the step-change model described in Section 2.2.4 in which the transmission rate scaling $\hat{\beta}$ undergoes a discrete drop. This addition to the model introduces two additional unknown parameters to infer: the time at which the transmission rate changes T_C, and the percentage reduction in transmission 1-δ.

We run our step-change model for multiple endpoints: 24 May 2018 and in 5 day increments to 28 June 2018. At each end point we retain the best 1,000 realisations with

(a)

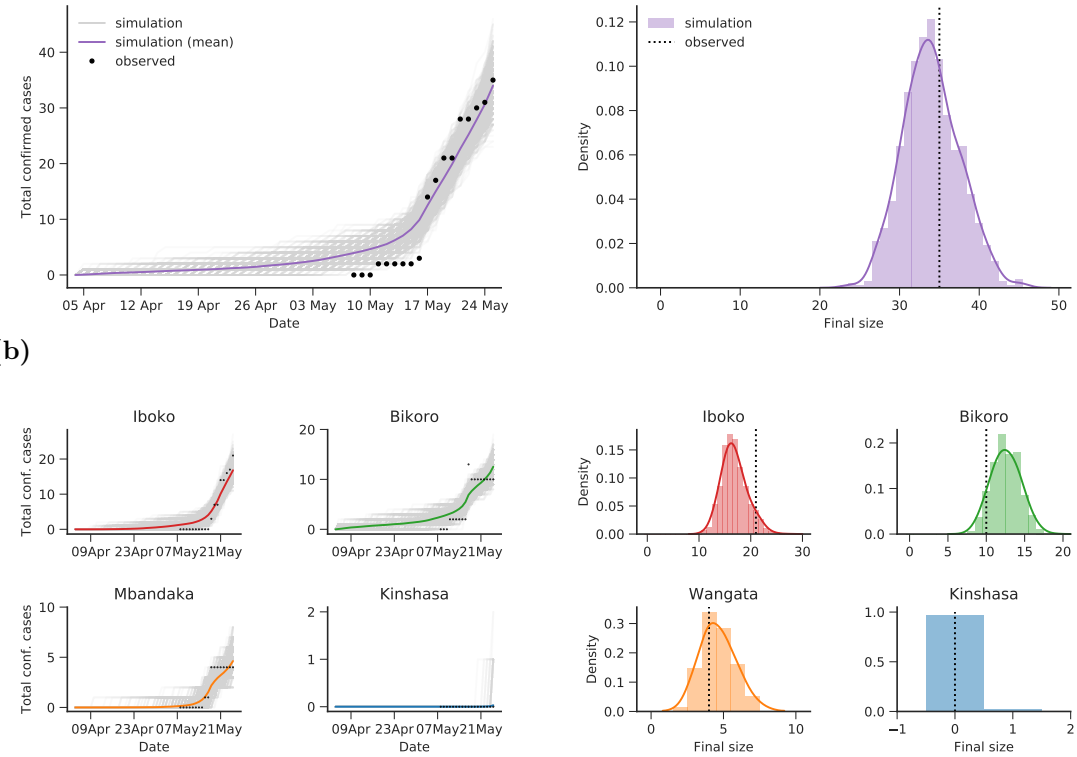

Figure 5. The epidemic curves and final distribution of cases for the 1,000 realisations with the smallest total error for our simple model (without any change in transmission) up to 25 May 2018. Results are shown for (a) the metapopulation, summed over all sub-populations, and (b) the four sub-populations separately. On the left-hand side we show individual realisations of the outbreak (shown in light grey) plotting the cumulative number of cases moving from the infected to either the hospitalised or funeral class; the mean across all 1,000 realisations is plotted in colour and solid points denote actual cumulative confirmed cases. On the right-hand side we show the final distribution together with the reported value (shown as a dashed vertical line).

the smallest total error and obtain estimates of the seven unknown parameters (Figure 6 and Figure B.2).

We estimate R_0 as more observed data are included in the model fitting procedure (Figure 6a). As additional data are observed with few or no new cases, our model estimates an earlier and larger decrease in R_0, providing increasing evidence for a change in transmission. Early in the outbreak (using data up to 24 May 2018) there is only a slight signature of a decline in R_0; between 29 May and 3 June 2018 there is sufficient change in transmission that R_0 drops below the threshold for continued transmission ($R_0 = 1$), and by 28 June 2018 R_0 has decreased to well below 1. In the remainder of the analysis we consider the best-fit distributions of the seven unknown parameters for the model run until 28 June 2018. Our estimates of the five parameters present in the simple model $(\hat{\beta}, a, b, c, A)$ change as a result of including more data. We now get an estimate for the transmission rate scaling $\hat{\beta}$ of 1.49 (95% CI [0.90, 2.33]) and so an initial estimate for R_0 of 4.02 (95% CI [2.43, 6.29]). Our mean estimate for R_0 is lower than for our simple model (4.02 compared to 5.34 previously), and the credible interval is narrower (range of 3.86 compared to 4.78 previously). Again, we recombine the four spatial parameters (a, b, c, A) to obtain meaningful distributions of the coupling (Figure 6e); our estimates are in broad agreement with values from our simple model, but again more tightly distributed. In addition to the five original parameters, we estimate the time at which there is a change in transmission T_C (Figure 6b) and the percentage reduction in transmission $1 - \delta$ (Figure 6c). We get an estimate for T_C of 15 May 2018 (95% CI [9 May 2018, 22 May 2018]) and an estimate for $1 - \delta$ of 94.8% (95% CI [85.4, 99.8]). Together with other parameters, we estimate that the reproductive ratio first drops below one on 16 May 2018 (95% CI [9 May 2018, 22 May 2018]), eight days after the outbreak was officially declared. At the end of the simulation on 28 June 2018, we have an estimate for R_0 of 0.21 (95% CI [0.13, 0.33]).

As before, we also obtain time series fits to the observed cumulative cases and the final distribution of cases corresponding to 28 June 2018 for the metapopulation as a whole (Figure 7a) and for the four sub-populations separately (Figure 7b). For the time series we now more robustly capture the bulk shape of the outbreak including the transition from exponential growth to disease eradication, although we still slightly overestimate the number of confirmed cases during the early stages of the outbreak. We compare the final distribution of the realised cumulative confirmed cases to the observed cumulative cases on 28 June 2018 (Figure 7, RHS). We now estimate the total number of confirmed cases to be 39 (mean 39.4, 95% CI [34, 45]), compared to 38 observed confirmed cases. At the subpopulation level, our mean estimates for the final size are very similar to the observed number of confirmed cases and are more tightly distributed than for the simple model. In Kinshasa we estimate the number of confirmed cases to be 0 (mean 0.005, 95% CI [0, 0]), which matches the 0 actual observed confirmed cases; in Wangata we estimate 4 confirmed cases (mean 4.18, 95% CI [3, 6]), which also matches the 4 observed cases; in Bikoro we estimate 11 confirmed cases (mean 11.06, 95% CI [9, 14]) compared to 10 observed cases; and in Iboko we estimate 24 confirmed cases (mean 24.14, 95% CI [20, 29]), which matches

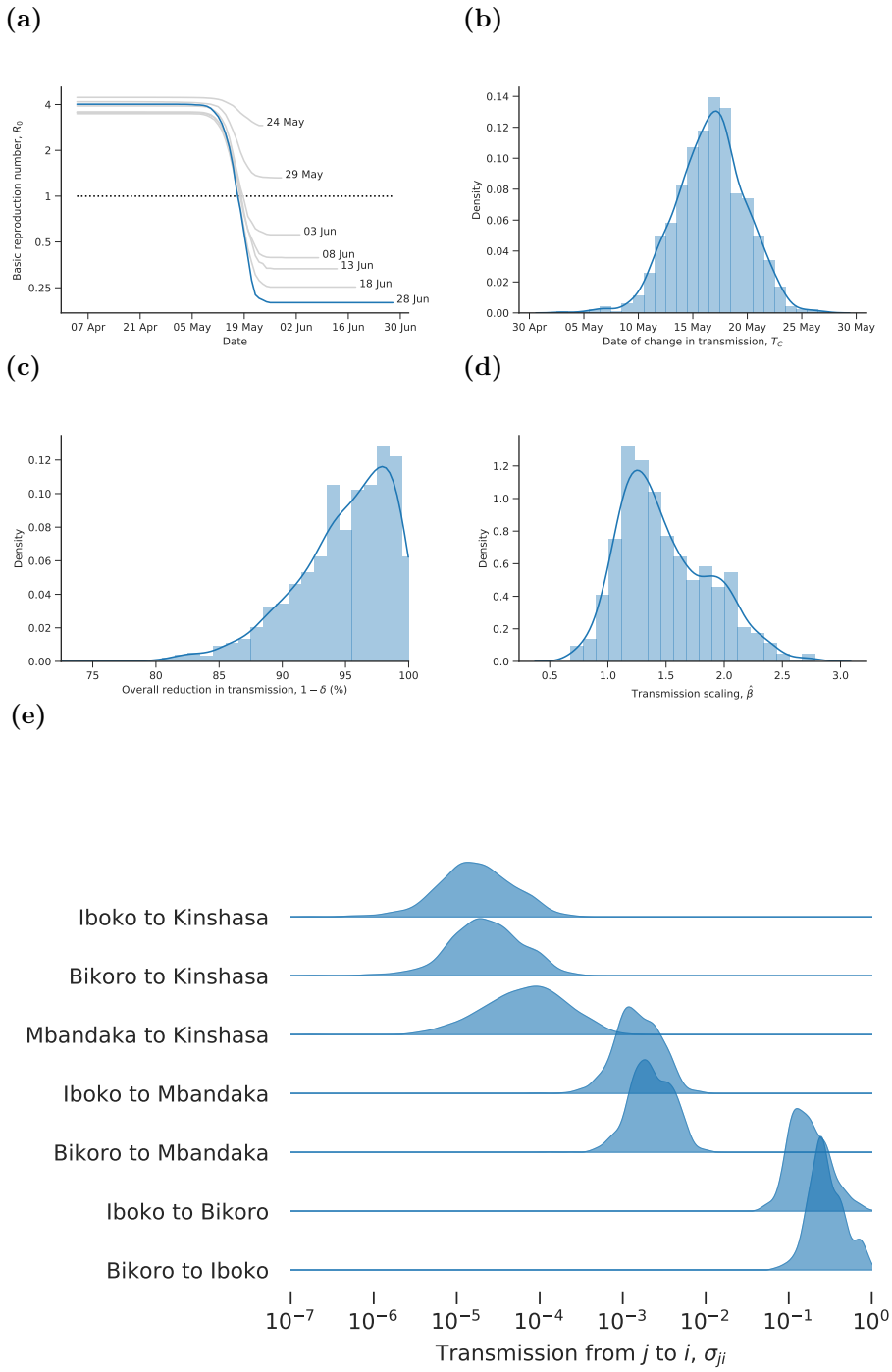

Figure 6. Results for the model with a change in transmission scaling. (a) the mean reproductive ratio from 1,000 best realisations fitted to differing lengths of data. (b-d) the final distribution from fitting to data up to 28 June 2018, showing the reduction in transmission scaling ($1-\delta$), the time at which the transmission rate changes (T_C) and the value of transmission scaling ($\hat{\beta}$) respectively. (e) distributions of select between-population coupling (σ_{ij}) from the model fitted to the full data set. Here we refer to the transmission, which is the direction in which the disease moves; however the coupling is determined by the movement of susceptible individuals who bring infection back into their home population.

the 24 observed cases.

4 Discussion

Ebola presents a significant burden to resource-poor countries in Central and West Africa. In addition to the high number of deaths that Ebola can cause, the result of managing outbreaks places huge strains on already limited resources and can reduce the control of other life-threatening diseases. Mathematical models can be used to effectively allocate resources and identify areas at risk of importing new cases. Here we develop a spatial modelling framework for the Ebola outbreak in Equateur Province in the DRC that can allow us to understand the risk of transmission to Kinshasa, using data available in real-time during the outbreak. Two models are fitted to observed confirmed cases: the first assumes no change in transmission and the second assumes a stepwise change in transmission at some time.

Our first model is fitted to the first three weeks of the outbreak. From these results we predict a low risk for the outbreak reaching the major population centre of Kinshasa (Supplementary Information). If we assume constant transmission then we would predict long-term exponential growth, which was not observed; however our second model is able to capture the exponential growth phase and disease eradication by early June (approximately four weeks after the outbreak was officially declared). We estimate a 94.8% reduction in transmission and that the basic reproductive ratio R_0 drops below one on 16 May 2018, indicating that the control measures are having their desired impact. The timing of the change in transmission is eight days after the outbreak was officially declared and approximately five days after intensive contact tracing had begun. The magnitude of the change is in qualitative agreement with WHO reports detailing the extensive local and international response to the outbreak.

Our model combines a well-established compartmental model for Ebola [32] with a spatial metapopulation structure. Metapopulation modelling has previously been used to assess the risk of international spread of Ebola during the 2014 outbreak in West Africa [30]; however, this approach relies upon the Global Epidemic and Mobility Model [45] and hence cannot be easily adapted for use at smaller scales. Given the temporal disconnect between the epidemiologically important infection times and the observed confirmation times, we adopt a likelihood-free approach where repeated stochastic simulations are used to minimise the error between model and data. Our framework has several practical advantages: it can be readily used with any form of stochastic model, and for small outbreak sizes many realisations can be generated quickly, such that many millions of simulations can be performed and analysed.

Although our model is able to capture the dynamics of the outbreak in Equateur Province, the inferred parameter values for $\hat{\beta}$, and thus our estimate for R_0, are somewhat surprising. For the simple model without change in transmission for the first three

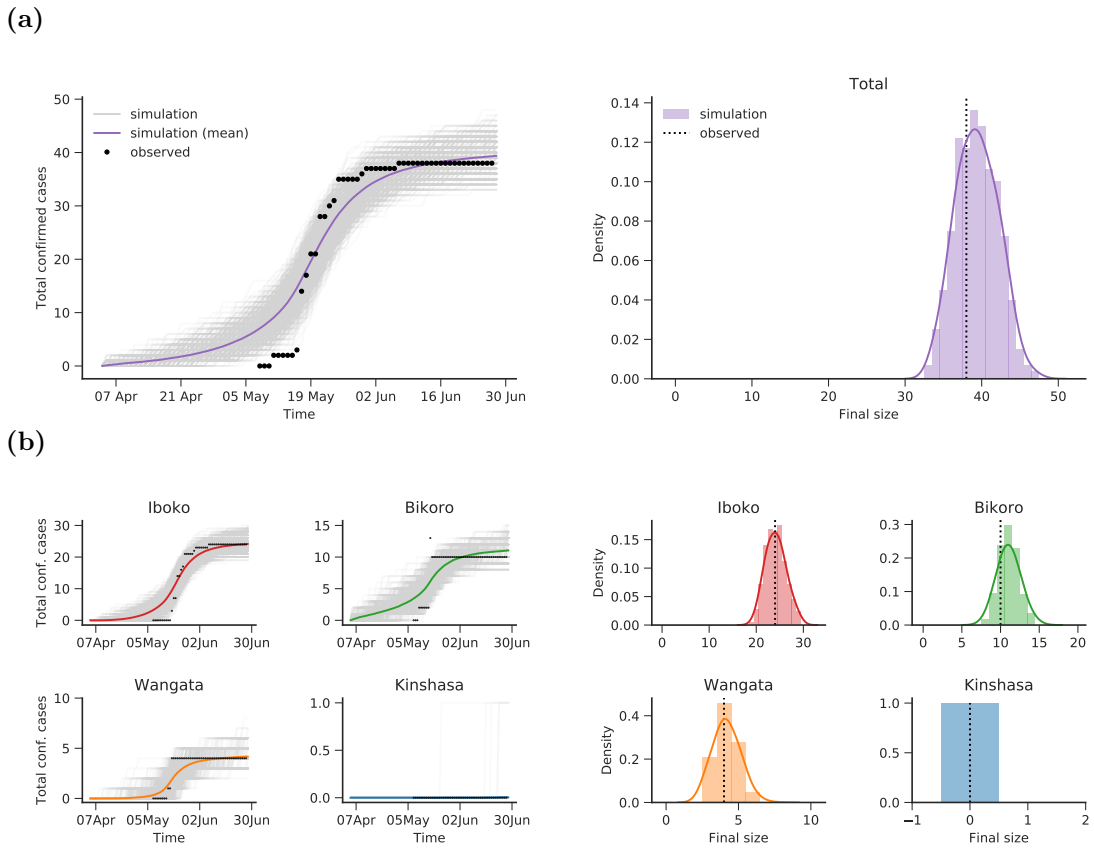

Figure 7. The epidemic curve and final distribution of cases for the 1,000 realisations with the smallest total error for our step-change model up to 28 June 2018. Results are shown for (a) the metapopulation, summed over all sub-populations, and (b) the four sub-populations separately. On the left-hand side we show individual realisations of the outbreak (shown in light grey) plotting the cumulative number of cases moving from the infected to either the hospitalised or funeral class; the mean across all 1,000 realisations is plotted in colour and solid points denote actual cumulative confirmed cases. On the right-hand side we show the distribution of total cases from the 1,000 best realisations together with the reported value (shown as a dashed vertical line).

weeks of data we get an estimate for R_0 of 5.34 [95% CI [3.17, 7.96]); for the step-change model using all the data, our new estimate of R_0 is 4.02 (95% CI [2.44, 6.28]). Both values are appreciably larger than other estimates of 1.03 [31], and larger than estimates for recent outbreaks: 1.38-3.65 for DRC, 1995; 1.34-2.7 for Uganda, 2000-1 [46]; 1.51-2.53 for West Africa, 2013-15 [47, 48]. We believe that this overestimate of R_0 is a result of the process by which cases are confirmed. Due to the delay between symptom onset and laboratory confirmation, confirmed cases are both spatially and temporally clustered, particularly early on in the outbreak. This delay is shown to be longer at the beginning of the outbreak when surveillance is low [31]. The effect of the delay between symptom onset and laboratory confirmation can be seen as large jumps in the number of confirmed cases in each of Wangata, Bikoro and Iboko. It is unlikely, however, that all cases confirmed on the same day share the same data of symptom onset, which is clear if we compare our data to Barry et al. [31]. This temporal clustering of confirmed cases distorts the data, and is likely the main factor that leads to our overestimate of R_0 compared to previous outbreaks. In principle, we could formulate a model that could mimic the temporal aggregation of cases, but this would generate an additional parameter that would need to be estimated and would place an extra layer of filtering between the epidemiology and results. Alternatively, the overestimate of R_0 could be addressed by modifying and refitting our model to data on the timing of symptom onset, if available.

Our modelling is motivated by the need for real-time analysis of Ebola outbreaks and interventions in a spatial setting. Our analysis is constrained by the quality and detail of the limited data publicly available during the outbreak in Equateur Province. In an attempt to minimise uncertainty around the true status of probable and confirmed cases we have restricted our analysis to confirmed cases only. However, even when we only consider confirmed cases the data we use contains at least one error: the number of confirmed cases in Bikoro drops from 13 to 10 on 17 May 2018. Due to the relatively small number of cases and the limited amount of data publicly available, we only infer parameters associated with the spatial component of the model and a single epidemiological parameter that scales the transmission rate; other epidemiological parameters are taken from the literature [32]. We believe that inference of all the epidemiological parameters of our model is not possible with the type of data that is publicly available; instead we would require information on the history and treatment of cases at the individual level.

Our quantitative results are also limited by assumptions and approximations made during the modelling process. To define the coupling we use the pairwise straight-line distances between populations; however, straight-line distances are likely a poor proxy for ease of travel between populations: in some remote areas (such as parts of Bikoro and Iboko health zones), it may take a significant amount of time to travel over short distances due to poor road infrastructure. In addition, it is not clear that each of the four sub-populations is acting as a homogeneously mixing population, and hence additional spatial structure may be acting on the dynamics, although without more detailed reporting this is impossible to assess. Modelling at a finer spatial scale would also require additional information on

population structure and would increase the computational power and time required for simulation and analysis, which is at odds with our aim to generate results in real-time in low-resource settings.

Our analysis has demonstrated that practically useful mathematical models can be matched to publicly available data early in an outbreak, especially if previous analysis has helped to set the time-course of disease progression. The likelihood-free method we have adopted is highly convenient, allowing us to quickly and easily perform matches between a rapid stochastic simulation model and available data. As such the modelling framework that we have described offers a template for early model inference to other outbreaks. In particular, the framework can easily be modified to accommodate different compartmental models, spatial scales, or data sources. Using this framework, and with very limited publicly-available data, we have been able to attribute a very low risk for the infection reaching Kinshasa which would exacerbate wider dissemination; we have also been able to rapidly identify changes in the transmission rate due to public-health interventions and predict that these interventions are sufficient to curtail the spread of infection. Obviously, as more data become available, especially individual-level data on cases, there is a desire to develop bespoke models fitted to the details of the ensuing outbreak; however, rapid early predictions before too much infection has arisen, such as outlined here, have the potential to generate a substantial impact on public-health decision-making.

References

[1] World Health Organisation. Ebola virus disease fact sheet, 2018. URL http://www.who.int/en/news-room/fact-sheets/detail/ebola-virus-disease.

[2] David K Evans, Markus Goldstein, and Anna Popova. Health-care worker mortality and the legacy of the Ebola epidemic, 2015. ISSN 2214109X. URL www.thelancet.com/lancetgh.

[3] Patrick G T Walker, Michael T White, Jamie T. Griffin, Alison Reynolds, Neil M Ferguson, and Azra C Ghani. Malaria morbidity and mortality in Ebola-affected countries caused by decreased health-care capacity, and the potential effect of mitigation strategies: A modelling analysis. *The Lancet Infectious Diseases*, 15(7):825–832, 2015. ISSN 14744457. doi: 10.1016/S1473-3099(15)70124-6. URL http://dx.doi.org/10.1016/.

[4] Alyssa S Parpia, Martial L Ndeffo-Mbah, Natasha S Wenzel, and Alison P Galvani. Effects of response to 2014-2015 ebola outbreak on deaths from malaria, HIV/AIDS, and tuberculosis, West Africa. *Emerging Infectious Diseases*, 22(3):433–441, 2016. ISSN 10806059. doi: 10.3201/eid2203.150977. URL http://dx.doi.org/10.3201/eid2203.150977.

[5] J.J. Muyembe-Tamfum, S. Mulangu, Justin Masumu, J.M. Kayembe, A. Kemp, and Janusz T. Paweska. Ebola virus outbreaks in Africa: Past and present. *Onderstepoort J Vet Res*, 79(2):1–8, 2012. ISSN 2219-0635. doi: 10.4102/ojvr.v79i2.451. URL http://ojvr.org/index.php/ojvr/article/view/451.

[6] Eric M Leroy, Alain Epelboin, Vital Mondonge, Xavier Pourrut, Jean-Paul Gonzalez, Jean-Jacques Muyembe-Tamfum, and Pierre Formenty. Human Ebola Outbreak Resulting from Direct Exposure to Fruit Bats in Luebo, Democratic Republic of Congo, 2007. *Vector-Borne and Zoonotic Diseases*, 9(6):723–728, 2009. ISSN 1530-3667. doi: 10.1089/vbz.2008.0167. URL https://www.liebertpub.com/doi/pdfplus/10.1089/vbz.2008.0167 http://www.liebertonline.com/doi/abs/10.1089/vbz.2008.0167.

[7] Barry S. Hewlett and Richard P. Amolat. Cultural contexts of Ebola in Northern Uganda. *Emerging Infectious Diseases*, 9(10):1242–1248, 2003. ISSN 10806040. doi: 10.3201/eid0910.020493.

[8] K R Victory, F Coronado, S O Ifono, T Soropogui, B A Dahl, Control Centers for Disease, and Prevention. Ebola transmission linked to a single traditional funeral ceremony - Kissidougou, Guinea, December 2014-January 2015. *MMWR Morbidity and Mortality Weekly Report*, 64(14):386–388, 2015. ISSN 01492195. doi: 10.15585/mmwr.mm6438a2. URL https://www.cdc.gov/mmwr/pdf/wk/mm6414.pdf http://www.ncbi.nlm.nih.gov/pubmed/25879897.

[9] Centre for Disease Control. Ebola (Ebola Virus Disease). 2014-2016 Ebola Outbreak in West Africa., 2017. URL https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/index.html.

[10] DRC Ministère de la Santé mailing list, 2018. URL https://us13.campaign-archive.com/home/?u=89e5755d2cca4840b1af93

[11] World Health Organisation. Ebola virus disease. Ebola situation reports: Democratic Republic of the Congo., 2018.

[12] World Health Organisation. Case definition recommendations for Ebola or Marburg virus diseases. Technical report, 2014. URL http://apps.who.int/iris/bitstream/handle/10665/146397/WHO_EVD_CaseDef_14.1_eng.pdf?sequence=1.

[13] World Health Organisation. Ebola virus disease. Democratic Republic of Congo. External Situation Report 17. Technical report, 2018. URL http://www.who.int/ebola/historical-outbreaks-drc/en/.

[14] World Health Organisation. Emergencies preparedness, response. Ebola virus disease- Democratic Republic of the Congo. Disease outbreak news., 2018. URL http://www.who.int/csr/don/14-may-2018-ebola-drc/en/.

[15] World Health Organisation. Meeting of the Strategic Advisory Group of Experts on immunization - conclusions and recommendations. Ebola vaccines. Technical report, 2018. URL http://www.who.int/csr/resources/publications/ebola/sage-vaccines-2017/en/.

[16] Ana Maria Henao-Restrepo, Anton Camacho, Ira M Longini, Conall H Watson, W. John Edmunds, Matthias Egger, Miles W Carroll, Natalie E Dean, Ibrahima Diatta, Moussa Doumbia, Bertrand Draguez, Sophie Duraffour, Godwin Enwere, Rebecca Grais, Stephan Gunther, Pierre Stéphane Gsell, Stefanie Hossmann, Sara Viksmoen Watle, Mandy Kader Kondé, Sakoba Kéïta, Souleymane Kone, Eewa Kuisma, Myron M Levine, Sema Mandal, Thomas Mauget, Gunnstein Norheim, Ximena Riveros, Aboubacar Soumah, Sven Trelle, Andrea S Vicari, John Arne Røttingen, and Marie Paule Kieny. Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!). *The Lancet*, 389(10068):505–518, 2017. ISSN 1474547X. doi: 10.1016/S0140-6736(16)32621-6. URL http://dx.doi.org/10.1016/.

[17] Maria D Van Kerkhove and Neil M Ferguson. Epidemic and intervention modelling a scientific rationale for policy decisions? Lessons from the 2009 influenza pandemic. *Bulletin of the World Health Organization*, 90(4):306–310, 2012. ISSN 0042-9686. doi: 10.2471/BLT.11.097949. URL http://www.who.int/entity/bulletin/volumes/90/4/11-097949.pdf.

[18] J Y Abrams, J R Copeland, R V Tauxe, K A Date, E D Belay, R. K. Mody, and E. D. Mintz. Real-time modelling used for outbreak management during a cholera epidemic, Haiti, 2010-2011. *Epidemiology and Infection*, 141(6):1276–1285, 2013. ISSN 09502688. doi: 10.1017/S0950268812001793. URL https://doi.org/10.1017/S0950268812001793.

[19] Anton Camacho, Adam Kucharski, Yvonne Aki-Sawyerr, Mark A. White, Stefan Flasche, Marc Baguelin, Timothy Pollington, Julia R. Carney, Rebecca Glover, Elizabeth Smout, Amanda Tiffany, W. John Edmunds, and Sebastian Funk. Temporal changes in ebola transmission in Sierra Leone and implications for control requirements: a real-time modelling study. *PLoS Currents*, 7(OUTBREAKS): 1–18, 2015. ISSN 21573999. doi: 10.1371/currents.outbreaks.406ae55e83ec0b5193e3085.

[20] Shihao Yang, Samuel C Kou, Fred Lu, John S Brownstein, Nicholas Brooke, and Mauricio Santillana. Advances in using Internet searches to track dengue. *PLoS Computational Biology*, 13(7), 2017. ISSN 15537358. doi: 10.1371/journal.pcbi.1005607. URL https://doi.org/10.1371/journal.pcbi.1005607.

[21] Matthew Graham, Jonathan E Suk, Saki Takahashi, C. Jessica Metcalf, A Paez Jimenez, Vladimir Prikazsky, Matthew J Ferrari, and Justin Lessler. Challenges and opportunities in disease forecasting in outbreak settings: A case study of measles in Lola Prefecture, Guinea. *American Journal of Tropical Medicine and Hygiene*, 98(5):1489–1497, 2018. ISSN 00029637. doi: 10.4269/ajtmh.17-0218. URL www.ajtmh.org.

[22] Sarah F. McGough, John S Brownstein, Jared B Hawkins, and Mauricio Santillana. Forecasting Zika Incidence in the 2016 Latin America Outbreak Combining Traditional Disease Surveillance with Search, Social Media, and News Report Data. *PLoS Neglected Tropical Diseases*, 11(1), 2017. ISSN 19352735. doi: 10.1371/journal.pntd.0005295. URL https://www.

[23] Jonathan K. Pritchard, Mark T. Seielstad, Anna Perez-Lezaun, and Marcus W. Feldman. Population Growth of Human Y Chromosomes: A Study of Y Chromosome Microsatellites. *Molecular Biology and Evolution*, 16(12):1791–1798, 1999. ISSN 0737-4038. URL https://academic.oup.com/mbe/article-abstract/16/12/1791/2925409.

[24] Adam J Kucharski, Anton Camacho, Francesco Checchi, Ron Waldman, Rebecca F Grais, Jean-Clement Cabrol, Sylvie Briand, Marc Baguelin, Stefan Flasche, Sebastian Funk, and W John Edmunds. Evaluation of the benefits and risks of introducing Ebola community care centers, Sierra Leone. *Emerging infectious diseases*, 21(3):393–9, 2015. ISSN 1080-6059. doi: 10.3201/eid2103.141892. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4344293&tool=pmcentrez&rendertype=abstract.

[25] G Chowell, N W Hengartner, C Castillo-Chavez, P W Fenimore, and J M Hyman. The basic reproductive number of Ebola and the effects of public health measures: The cases of Congo and Uganda. *Journal of Theoretical Biology*, 229(1):119–126, 2004. ISSN 00225193. doi: 10.1016/j.jtbi.2004.03.006. URL http://www.eebweb.arizona.edu/Courses/Ecol409_509/chowell04.pdf.

[26] Phenyo E Lekone and Bärbel F Finkenstädt. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics*, 62 (4):1170–1177, 2006. ISSN 0006341X. doi: 10.1111/j.1541-0420.2006.00609.x. URL https://pdfs.semanticscholar.org/a695/273fa9694c0d898e723c82a265f7122e8f3b.pdf.

[27] Stefano Merler, Marco Ajelli, Laura Fumanelli, Marcelo F.C. Gomes, Ana Pastore y. Piontti, Luca Rossi, Dennis L Chao, Ira M. Longini, M. Elizabeth Halloran, and Alessandro Vespignani. Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: A computational modelling analysis. *The Lancet Infectious Diseases*, 15(2):204–211, 2015. ISSN 14744457. doi: 10.1016/S1473-3099(14)71074-6. URL http://dx.doi.org/10.1016/.

[28] Caitlin Rivers, Eric Lofgren, Madhav Marathe, Stephen Eubank, and Bryan Lewis. Modeling the Impact of Interventions on an Epidemic of Ebola in Sierra Leone and Liberia. *PLoS Currents*, 1, 2014. ISSN 2157-3999. doi: 10.1371/currents.outbreaks.fd38dd85078565450b0be3fcd78f5ccf. URL https://github.com/cmrivers/ebola. http://currents.plos.org/outbreaks/?p=42673.

[29] Maria Vittoria Barbarossa, Attila Dénes, Gábor Kiss, Yukihiko Nakata, Gergely Röst, and Zsolt Vizi. Transmission dynamics and final epidemic size of ebola virus disease outbreaks with varying interventions. *PLoS ONE*, 10(7), 2015. ISSN 19326203. doi: 10.1371/journal.pone.0131398. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4510538/pdf/pone.0131398.pdf.

[30] Marcelo F. C. Gomes, Ana Pastore Piontti, Luca Rossi, Dennis Chao, Ira Longini, M Elizabeth Halloran, Alessandro Vespignani, Gomes Mfc, and Pastore A Piontti. Assessing the international spreading risk associated with the 2014 West African Ebola outbreak. *PLOS Currents*, 6, 2014. doi: 10.1371/currents.outbreaks.cd818f63d40e24aef769dda7df9e0da5.

[31] Ahmadou Barry, Steve Ahuka-Mundeke, Yahaya Ali Ahmed, Yokouide Allarangar, Julienne Anoko, Brett Nicholas Archer, Aaron Aruna Abedi, Jayshree Bagaria, Marie Roseline Darnycka Belizaire, Sangeeta Bhatia, Théophile Bokenge, Emanuele Bruni, Anne Cori, Ernest Dabire, Amadou Mouctar Diallo, Boubacar Diallo, Christl Ann Donnelly, Ilaria Dorigatti, Tshewang Choden Dorji, Aura Rocio Escobar Corado Waeber, Ibrahima Socé Fall, Neil M Ferguson, Richard Gareth FitzJohn, Gervais Leon Folefack Tengomo, Pierre Bernard Henri Formenty, Alpha Forna, Anne Fortin, Tini Garske, Katy AM Gaythorpe, Celine Gurry, Esther Hamblion, Mamoudou Harouna Djingarey, Christopher Haskew, Stéphane Alexandre Louis Hugonnet, Natsuko Imai, Benido Impouma, Guylain Kabongo, Oly Ilunga Kalenga, Emerencienne Kibangou, Theresa Min-Hyung Lee, Charles Okot Lukoya, Ousmane

Ly, Sheila Makiala-Mandanda, Augustin Mamba, Placide Mbala-Kingebeni, Franck Fortune Roland Mboussou, Tamayi Mlanda, Vital Mondonge Makuma, Oliver Morgan, Anastasie Mujinga Mulumba, Patrick Mukadi Kakoni, Daniel Mukadi-Bamuleka, Jean-Jaques Muyembe, Ndjoloko Tambwe Bathé, Patricia Ndumbi Ngamala, Roland Ngom, Guillaume Ngoy, Pierre Nouvellet, Justus Nsio, Kevin Babila Ousman, Emilie Peron, Jonathan Aaron Polonsky, Michael J Ryan, Alhassane Touré, Rodney Towner, Gaston Tshapenda, Reinhilde Van De Weerdt, Maria Van Kerkhove, Annika Wendland, Da Konan Michel Yao, Zabulon Yoti, Etienne Yuma, Guy Kalambayi Kabamba, Jean de Dieu Lukwesa Mwati, Gisele Mbuy, Leopold Lubula, Anny Mutombo, Oscar Mavila, Yyonne Lay, Emma Kitenge, and The Ebola Outbreak Epidemiology Team. Outbreak of Ebola virus disease in the Democratic Republic of the Congo, AprilMay, 2018: an epidemiological study. *The Lancet*, 392:213–221, 2018. ISSN 01406736. doi: 10.1016/S0140-6736(18)31387-4.

[32] J. Legrand, R. F. Grais, P. Y. Boelle, A. J. Valleron, and A. Flahault. Understanding the dynamics of Ebola epidemics. *Epidemiology and Infection*, 135(4):610–621, 2007. doi: 10.1017/S0950268806007217.

[33] Matt J Keeling and Pejman Rohani. Estimating spatial coupling in epidemiological systems: a mechanistic approach. *Ecology Letters*, 5(1):20–29, 2002. ISSN 1461023X. doi: 10.1046/j.1461-0248.2002.00268.x.

[34] C Viboud, O N Bjørnstad, D L Smith, L Simonsen, M A Miller, and B T Grenfell. Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science*, 312(5772):447–51, 2006. ISSN 1095-9203. doi: 10.1126/science.1125237.

[35] Duygu Balcan, Vittoria Colizza, Bruno Gonçalves, Hao Hu, José J Ramasco, and Alessandro Vespignani. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, 106(51):21484–21489, 2009.

[36] Michele Tizzoni, Paolo Bajardi, Adeline Decuyper, Guillaume Kon, Kam King, and Christian M Schneider. On the use of human mobility proxies for modeling epidemics. *PLoS Computational Biology*, 10(7), 2014. doi: 10.1371/journal.pcbi.1003716.

[37] Y Xia, O N Bjørnstad, and B T Grenfell. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *The American Naturalist*, 164(2):267–281, 2004. doi: 10.1086/422341.

[38] Daniel T Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *Journal of Chemical Physics*, 115(4):1716–1733, 2001. ISSN 00219606. doi: 10.1063/1.1378322. URL `http://jcp.aip.org/jcp/copyright.jsp`.

[39] Mpia A. Bwaka, Marie-José Bonnet, Philippe Calain, Robert Colebunders, Ann De Roo, Yves Guimard, Kasongo R. Katwiki, Kapay Kibadi, Mungala A. Kipasa, Kivudi J. Kuvula, Bwas B. Mapanda, Matondo Massamba, Kibadi D. Mupapa, JeanJacques MuyembeTamfum, Edouard Ndaberey, Clarence J. Peters, Pierre E. Rollin, and Erwin Van den Enden. Ebola Hemorrhagic Fever in Kikwit, Democratic Republic of the Congo: Clinical Observations in 103 Patients. *The Journal of Infectious Diseases*, 179(s1):S1–S7, 1999. ISSN 0022-1899. doi: 10.1086/514308. URL `https://academic.oup.com/jid/article-abstract/179/Supplement_1/S1/879641` `https://academic.oup.com/jid/article-lookup/doi/10.1086/514308`.

[40] Roger Ndambi, Philippe Akamituna, Marie-Jo Bonnet, Anicet Mazaya Tukadila, JeanJacques MuyembeTamfum, and Robert Colebunders. Epidemiologic and Clinical Aspects of the Ebola Virus Epidemic in Mosango, Democratic Republic of the Congo, 1995. *The Journal of Infectious Diseases*, 179(s1):S8–S10, 1999. ISSN 0022-1899. doi: 10.1086/514297.

URL https://academic.oup.com/jid/article-abstract/179/Supplement_1/S8/882632 https://academic.oup.com/jid/article-lookup/doi/10.1086/514297.

[41] Scott F. Dowell, Rose Mukunu, Thomas G. Ksiazek, Ali S. Khan, Pierre E. Rollin, and C. J. Peters. Transmission of Ebola Hemorrhagic Fever: A Study of Risk Factors in Family Members, Kikwit, Democratic Republic of the Congo, 1995. *The Journal of Infectious Diseases*, 179(s1):S87–S91, 1999. ISSN 0022-1899. doi: 10.1086/514284. URL https://academic.oup.com/jid/article-abstract/179/Supplement_1/S87/882673 https://academic.oup.com/jid/article-lookup/doi/10.1086/514284.

[42] Ali S Khan, F Kweteminga Tshioko, David L Heymann, Bernard Le Guenno, Pierre Nabeth, Barbara Kerstiëns, Yon Fleerackers, Peter H Kilmarx, Guenael R Rodier, Okumi Nkuku, Pierre E Rollin, Anthony Sanchez, Sherif R Zaki, Robert Swanepoel, Oyewale Tomori, Stuart T Nichol, C J Peters, Thomas G Ksiazek, and Edgar Allan Poe. The Reemergence of Ebola Hemorrhagic Fever, Democratic Republic of the Congo, 1995. *The Journal of Infectious Diseases*, page 1842, 1999. URL https://academic.oup.com/jid/article-abstract/179/Supplement_1/S76/882543.

[43] Alexander K. Rowe, Jeanne Bertolli, Ali S. Khan, Rose Mukunu, J. J. MuyembeTamfum, David Bressler, A. J. Williams, C. J. Peters, Luis Rodriguez, Heinz Feldmann, Stuart T. Nichol, Pierre E. Rollin, and Thomas G. Ksiazek. Clinical, Virologic, and Immunologic FollowUp of Convalescent Ebola Hemorrhagic Fever Patients and Their Household Contacts, Kikwit, Democratic Republic of the Congo. *The Journal of Infectious Diseases*, 179(s1):S28–S35, 1999. ISSN 0022-1899. doi: 10.1086/514318. URL https://academic.oup.com/jid/article-abstract/179/Supplement_1/S28/882033 https://academic.oup.com/jid/article-lookup/doi/10.1086/514318.

[44] OCHA DR Congo. DR Congo- Health Zones. URL https://data.humdata.org/dataset/dr-congo-health-0.

[45] Duygu Balcan, Bruno Gonçalves, Hao Hu, José J Ramasco, Vittoria Colizza, and Alessandro Vespignani. Modeling the spatial spread of infectious diseases: The global epidemic and mobility computational model. *Journal of Computational Science*, 1(3):132–145, 2010. ISSN 18777503. doi: 10.1016/j.jocs.2010.07.002. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3056392/pdf/nihms225903.pdf.

[46] A. Camacho, A.J. Kucharski, S. Funk, J. Breman, P. Piot, and W.J. Edmunds. Potential for large outbreaks of Ebola virus disease. *Epidemics*, 9:70–78, dec 2014. doi: 10.1016/j.epidem.2014.09.003. URL http://linkinghub.elsevier.com/retrieve/pii/S1755436514000528.

[47] Maria D. Van Kerkhove, Ana I. Bento, Harriet L. Mills, Neil M. Ferguson, and Christl A. Donnelly. A review of epidemiological parameters from Ebola outbreaks to inform early public health decision-making. *Scientific Data*, 2:1–10, 2015. ISSN 20524463. doi: 10.1038/sdata.2015.19.

[48] Christian L. Althaus. Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa. pages 1–9, 2014. ISSN 2157-3999. doi: 10.1371/currents.outbreaks.91afb5e0f279e7f29e7056095255b288. URL http://arxiv.org/abs/1408.3505%0Ahttp://dx.doi.org/10.1371/currents.outbreaks.91afb5e0f279e7f29e7056095255

A Formulation of parameters in epidemic model

There are ten free parameters in the model and five additional parameters that are calculated from the first ten. The model parameters that can be derived from the free parameters are: γ_{HF}^{-1}, the mean time from hospitalisation to death; γ_{HR}^{-1}, the mean time from hospitalisation to end of infectious period for survivors; θ_1, the rate of transition from infectious to hospitalised, which is calculated from θ such that $\theta\%$ of infectious cases are hospitalised; δ_1 and δ_2, the effective case fatality ratios in the infected and hospitalised class, respectively, both which are calculated from δ such that the case fatality ratio is δ. These five parameters are calculated as follows:

$$\gamma_{HF}^{-1} = \gamma_{IF}^{-1} - \gamma_{IH}^{-1} \tag{A.1}$$

$$\gamma_{HR}^{-1} = \gamma_{IR}^{-1} - \gamma_{IH}^{-1} \tag{A.2}$$

$$\theta_1 = \frac{\theta\left((1-\delta_1)\gamma_{IR} + \delta_1\gamma_{IF}\right)}{\theta\left((1-\delta_1)\gamma_{IR} + \delta_1\gamma_{IF}\right) + (1-\theta)\gamma_{IH}} \tag{A.3}$$

$$\delta_1 = \frac{\delta\gamma_{IR}}{\delta\gamma_{IR} + (1-\delta)\gamma_{IF}} \tag{A.4}$$

$$\delta_2 = \frac{\delta\gamma_{HR}}{\delta\gamma_{HR} + (1-\delta)\gamma_{HF}} \tag{A.5}$$

B Posterior distributions of individual coupling parameters

Figure B.1 and B.2 show the pairwise relationships and posterior distributions for the free parameters in the simple model and the step-change model, respectivly. In the simple model, the parameters are the transmission scaling parameter ($\hat{\beta}$) and four spatial parameters (A, a, b, c). In the step-change model, we also infer the percentage reduction in transmission ($1 - \delta$), the time at which the transmission rate changes (T_C). Parameters are inferred using an approximate Bayesian likelihood method; full details of this method are given in Section 2.2.3.

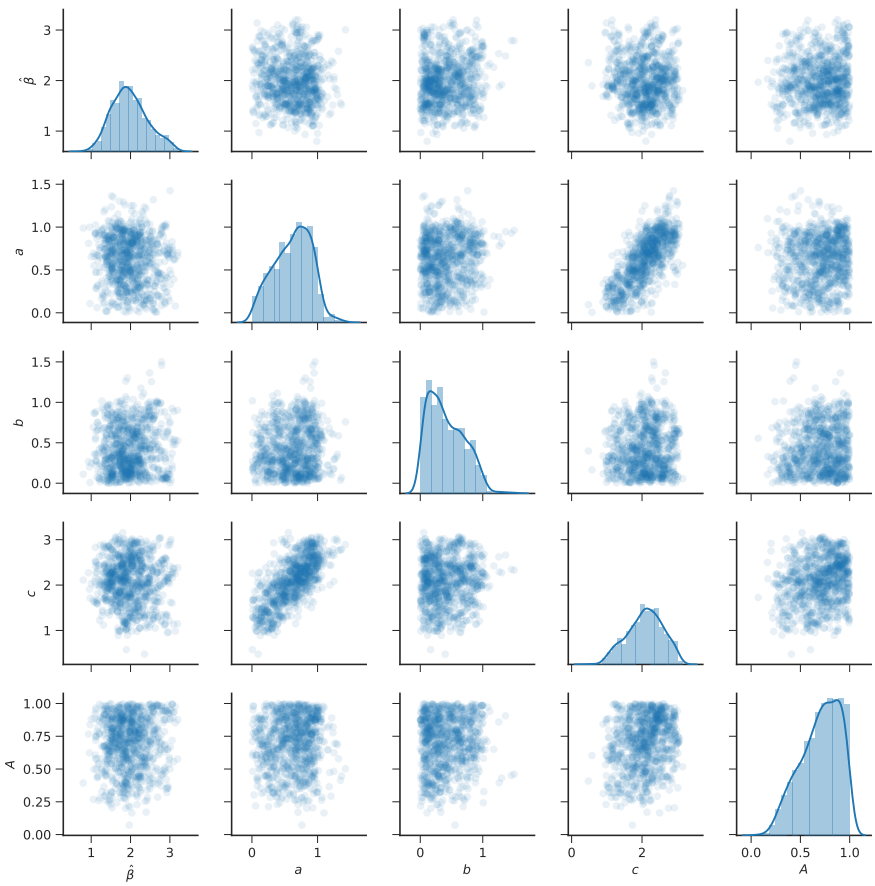

Figure B.1. Pairwise scatter plots of each of the six unknown parameters in the simple model (without change in transmission). Diagonal plots show histograms of the posterior distribution of each parameter; off diagonal plots show pairwise relationship between parameters.

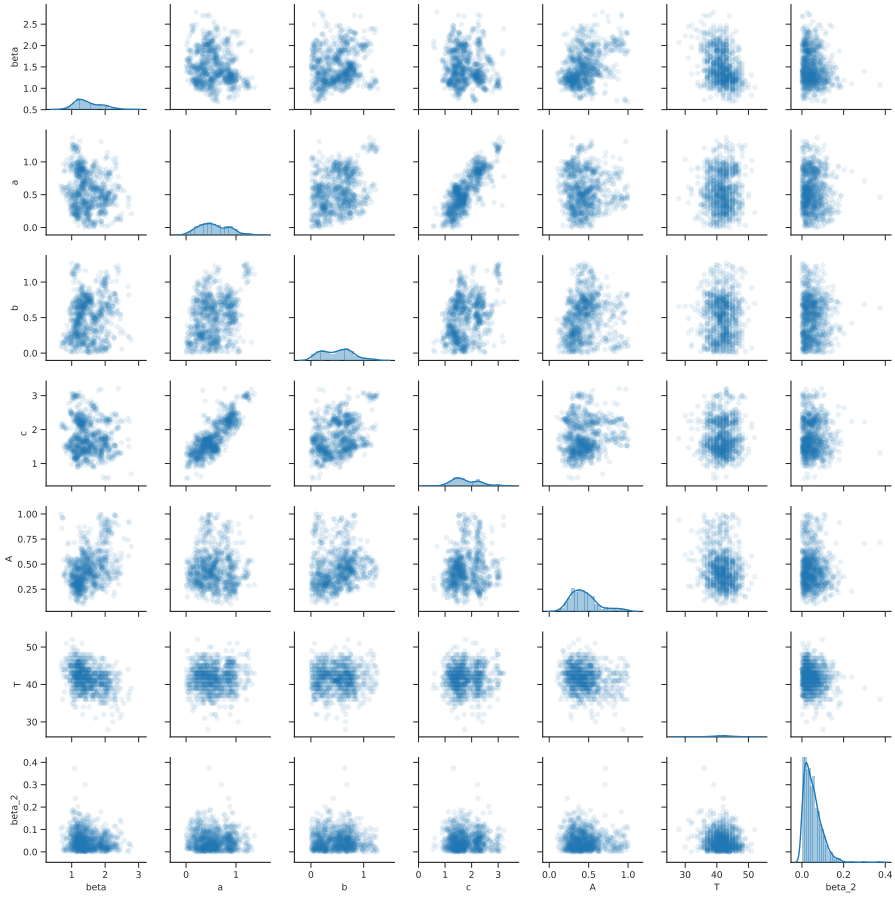

Figure B.2. Pairwise scatter plots of each of the six unknown parameters in the step-change model. Diagonal plots show histograms of the posterior distribution of each parameter; off diagonal plots show pairwise relationship between parameters.

C Force of infection at the subpopulation level

Figure C.1 shows the cumulative force of infection over time in each of the four subpopulations. The force of infection at population i at time t is

$$\begin{cases} \hat{\beta} \sum_{j=1}^{K} \frac{\sigma_{ij}}{N_j} C_j^{obs}(t) & \text{for } t < T_C, \\ \delta \hat{\beta} \sum_{j=1}^{K} \frac{\sigma_{ij}}{N_j} C_j^{obs}(t) & \text{for } t \geq T_C \end{cases} \tag{C.1}$$

where $C^{obs}(t)$ is the cumulative observed confirmed cases in population j at time t, $\hat{\beta}$ is the transmission scaling parameter, $(1-\delta)$ is the reduction in transmission, T_C is the time at which the transmission rate changes, and σ_{ij} is the coupling from population i to population j.

In all four sub-populations the force of infection increases as new cases are reported, and decreases around mid-May as transmission is reduced as a result of intervention measures. The force of infection plateaus once no new cases are observed.

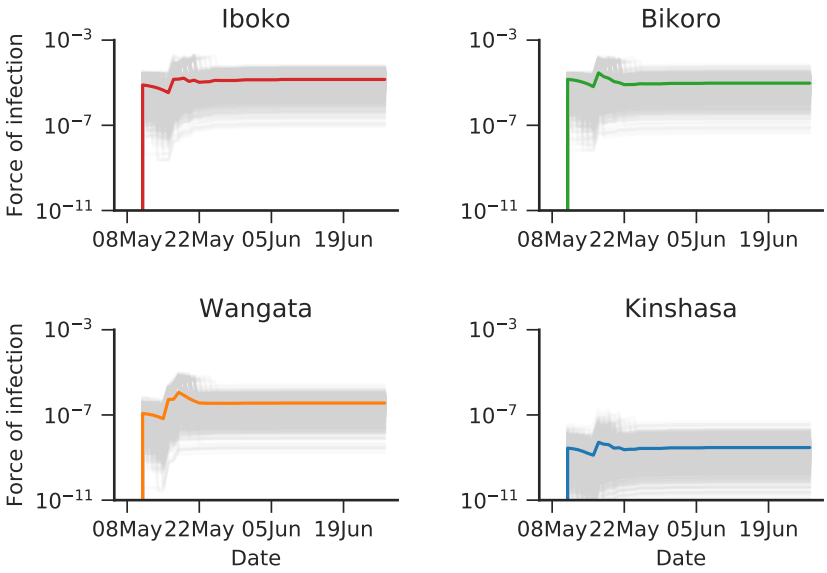

Figure C.1. Cumulative force of infection over time in each of the four sub-populations. The mean force of infection taken over all 1,000 realisations is shown as a thick coloured line; individual realisations are shown in light grey.