Multiscale modeling to explore Ebola vaccination strategies

Van Kinh Nguyen¹,², Rafael Mikolajczyk²,³,⁴,⁵ and Esteban A. Hernandez-Vargas¹,²

¹Frankfurt Institute for Advanced Studies, Ruth-Moufang-Strasse 1, 60438, Frankfurt am Main, Germany
²Helmholtz Centre for Infection Research, Braunschweig, Germany
³German Centre for Infection Research, Germany
⁴Hannover Medical School, Hannover, Germany
⁵Martin-Luther University, Halle-Wittenberg, Germany

The power to assess potential outcomes and intervention strategies is critical for epidemic preparedness. But emerging and mutated pathogens always challenge our current knowledge, pleading for fresh approaches to explore their epidemic potentials up front.

This paper coupled a within-host viral dynamics model and a between-host network model of Ebola virus (EBOV) infection showing that its transmission characteristics can be faithfully recapitulated.

Based on this multiscale model, EBOV’s incubation period is predicted in the range from 2.6 to 12.4 days, while infected subjects can remain infectious until day 17. The predicted basic reproductive number (R0) differs by age-groups: the overall is 1.4 and the highest is 4.7 for the 10-14 years old. Random vaccination strategies can reduce R0 and case-fatality rate, eliminate the possibility of large outbreaks, but the effect depends on timing and coverage.

A random vaccination program can reduce R0 below one if 85% coverage is achieved, and if it was conducted during the period from five months before to one week after the start of an epidemic. A vaccination coverage of 33% can reduce the epidemic size by ten to hundred times compared to a non-intervention scenario. Altogether, infection characteristics and epidemic mitigation approaches could be assessed using experimental data. An early, age-group specific, and high coverage vaccination program is the most beneficial.
1. Introduction

Epidemics of infectious diseases are listed among the potential catastrophes and can be potentially misused as mass destruction weapons [1]. Overwhelming research efforts have been developed to early predict the danger of the epidemics but their crisis nature left scientists no better option than learning from the past [1,2]. However, confronting outbreaks of emerging infections requires swift responses and thus the ability to evaluate quickly and early potential outcomes [1]. As such, computer simulations of epidemic models undoubtedly hold the potential as the first-aid toolbox for decision making amid the crisis [1,3].

A majority of epidemic modelling studies has exclusively relied on the availability of outbreak data [4–6]. This approach requires that sufficient incidence data are available; for example, data at the end of an epidemic or at least until its peak [7]. As such, it has limited applicability to newly emerging epidemics. Moreover, mechanistic models based on outbreak data are often oversimplified [8]. For example, the effective transmission probability [6] has been usually simplified as a single parameter that reflects collective effects of the contact rate with the infectious, the infectivity of the infectious, and the susceptibility of the susceptible. As a result, these key processes in the disease transmission are lost, especially the transient nature of the infection course [9]. In reality, the within-host infection process determines key parameters in the disease transmission [9–13]. In an infected subject, interactions between the viruses and immune responses shape the viral load dynamics that ultimately defines the incubation period, the transmission potential, and the recovery rate [11,14]. It is also evident that susceptibility to infection is not the same for all the susceptible but, among others, it is highly correlated with a subject’s age due to the aging of the immune systems [15,16]. Differences in the within-host infection profile as well as the susceptibility to infection complicate greatly epidemic models but at the same time underline their influential roles in determining epidemics features.

The interplays between within-host infection and between hosts transmission (Fig. 1) led to arising attempts connecting the two levels [9,11,17–21], but the approach is still at its infancy [10]. Most of these models are conceptual and theoretical [10], and rely on assumptive or previously obtained parameter estimates [14,22]. Thus, based on our thorough assessments of modelling within-host EBOV infection previously [12,13,23], we attempt to simulate EBOV transmission fitness at the population level using explicitly experimental and epidemiological data.
In particular, we embedded a within-host infection model of EBOV infection directly into a network transmission model at population level to simulate epidemic trajectories. Both the used models were derived based on empirical data of EBOV infection and human contact networks. Parameters obtained from simulations were then compared to those estimated based on actual outbreak data and empirical observations. The results showed that using with-host infection model not only uncovered faithful estimate of the transmission parameters, but also allowed the evaluations of realistic vaccination effects. In that capacity, epidemic consequences can be evaluated ahead of time once within-host viral dynamics are available.

Material and Methods

In an EBOV-infected subject, different immune systems components dynamically evolve in response to the viral replication dynamic. As a result, a series of events is triggered determining infection outcomes such as infectious status, symptoms, recovery, or death [24–26]. Therefore, the EBOV replication dynamics within a host were used in this paper to infer transmission parameters.

Within-host model

Using viral dynamics and immune responses data within a host, mathematical relations can be defined to test hypothesized infection mechanisms [12,27]. In this context, non-human primates (NHPs) are the standard animal model for developing EBOV’s therapeutics and vaccines in humans [28,29] which recently has been used to develop an effective vaccine against EBOV [30]. Epidemiological and pharmacological studies reported that a viral load higher than $10^6$ copies/mL [29,31] is associated with a higher mortality rate, whereas observations on experimental data in NHPs showed that a viral load level higher than $10^8$ TCID$_{50}$ was fatal [24,25]. Here the viral load dynamics were simulated based on the model as follows [13]:

$$\frac{dV}{dt} = r_V V \left(1 - \frac{V}{K_V}\right) \left(\frac{V}{I_n + V}\right) \left(1 - \frac{Ab}{K_{Ab}}\right)$$

(1.1)

where $r_V$, $K_V$, and $I_n$ denote the replication rate, the host’s carrying capacity, and a constraint threshold expressing the lag-phase growth of the virus. The parameter $K_{Ab}$ represents the strength of the immune system at which the antibody titre inhibits the viral net growth rate [13]. The model parameters were obtained previously [13] using two experimental datasets on NHPs [24,25]. The antibody dynamic ($Ab$) was fitted previously [13] to data of NHPs vaccinated with vesicular stomatitis virus (VSV-EBOV) vaccine [25]. The VSV-EBOV has recently showed efficacy in human [30]. Detailed of model fitting and data can be found in [13].

Simulated subject-specific infection course

To simulate subject-specific infection course, the antibody response strength $K_{Ab}$ was varied from a normal level approximately $10^{2.5}$ [25,32] to the highest observed level of $10^{1.5}$ [25]. This value was assumed to vary based on individual’s age, i.e., a U-shaped function of age with larger values for the infant and the elderly [15]. As infective dose can alter the course of infection [33], the initial condition $V(0)$ of model Eq. (1.1) was varied depending on from whom a subject acquires the infection, i.e., equals the lethal dose ($V_c = 10^{0.15}$ [15]) times the transmission potential of whom transmits the disease. Here we assumed a direct relation [10] between the transmission potential and the viral load at the time of infection, i.e., the transmission potential $p_{\text{Trans}}(t) = V(t)/K_V$. Note that $p_{\text{Trans}}(t) = 1$ does not guarantee a successful transmission, but it was considered collectively with its contacts susceptibility and with the existence of such a contact.
Table 1. Definitions of transmission parameters based on viral load and epidemics outcomes based on network model.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Incubation period</td>
<td>the interval between exposure to a pathogen and initial occurrence of symptoms [4] was defined from the infection time to the first time the viral load crosses over the detectable threshold (Fig. S2).</td>
</tr>
<tr>
<td>B Time from symptom onset to recovery [4]</td>
<td>defined as the interval between the first day of detectable viral load and the first day the viral load goes undetectable (Fig. S2).</td>
</tr>
<tr>
<td>C Time from symptom onset to death [4]</td>
<td>defined as the interval between the first day with detectable viral load and the day the area under the viral load curve (AUC) crosses the reference threshold $AUC_7$ (Fig. S2).</td>
</tr>
<tr>
<td>D Basic reproductive number (R0)</td>
<td>calculated based on the network of infected subjects at the end of an epidemic. In terms of network models, this equals the mean degree distribution of the infected network, considering a directed network without loops (e.g., Fig. S3). The R0 by age-group was also calculated in the same fashion based on the assigned age-attribute. Note that in epidemics with intervention, the R0 is called the effective reproductive number (Re).</td>
</tr>
<tr>
<td>E Final infected fraction</td>
<td>the proportion of infected nodes at the end of the epidemic simulations.</td>
</tr>
<tr>
<td>F Case-fatality rate</td>
<td>the proportion of nodes died as a result of EBOV infection calculated as the end of epidemics.</td>
</tr>
</tbody>
</table>

Infection outcomes definitions

Empirical observations from both EBOV infected human and NHPs showed that the time from symptom onset to death is approximately one week [24,25,34]. Based on this and the viral load, we used the total area under the viral load curve (AUC) seven days post-infection in the subjects that died as a threshold above which the infection is lethal, i.e., $AUC_7 = \int_0^7 V(t)dt$. Otherwise, infected subjects were assumed recovered once the viral load was no longer detectable (Fig. S2). Depending on the infective dose and the adaptive immune response strength, an infection will manifest different viral dynamics. Based on that, we defined the transmission parameters as in Table 1A-C.

The network model

The European’s contact patterns survey data [35] were used to generate a network model reflecting the number of contacts, the mixing patterns among age-groups, and a specific population age-structure. The age-distribution of the city Freetown in Sierra Leon was used as the reference [36]. A detailed description of the implementation can be found in Supplemental 1. Because EBOV spreads through direct contacts with infectious subjects [33], and that the highest risk of infection is contacting with blood, faeces, and vomit [37], we used only the data of physical contacts and excluded those contacts with a duration less than five minutes. To account for the transmission route through funeral practices in EBOV outbreaks [2], we considered deceased EBOV-infected subjects infectious until they were buried. During the last epidemics in Sierra Leone, the time from death to burial was one to two days on average but can be a week [38]. This number was randomly assigned using a truncated normal distribution at zero and seven with unit mean and variance.
Transmission outcomes definitions
To obtain EBOV’s epidemics metrics, the within-host infection model was embedded into network model. Simulations of EBOV epidemic are detailed in Supplemental 2. In short, a network of ten thousand nodes was generated. Scenarios in which the population was randomly vaccinated during one-week vaccination programs were tested and compared to a control simulation without interventions. For each scenario, one thousand simulations were performed, each of which started with a single random index case. Each time when a contact occurs, the viral load at the time point was extracted to determine the transmission potential. Next, the susceptibility of the contact persons were computed as a function of their age [15]. A Bernoulli trial was then used to determine if the contact results in an infection given the overall transmission probability. If the transmission succeeds, the newly infected subject has its own infection profile computed. Based on simulation outputs, the epidemic outcomes were determined as in Table 1D-F.

Results

Basic transmission characteristics
Simulations of the outcomes of the within-host infection model showed a highly skewed distribution of the basic transmission parameters (Fig. 2). The incubation period derived from viral load dynamics ranged from 2.6 to 12.4 days (median: 3.8) compared to the previous estimates based on actual outbreak data ranging from 3.35 to 12.7 days [4]. The delay time from infection to recovery ranged from 6.9 to 17.6 days (median: 9.7). Previous estimates of this interval ranged from 2 to 26 days (median: 10) [4]. The time from infection to death ranged from 8.1 to 15.1 days (median: 9) compared to previous estimates ranged from 3 to 21 (median: 9–10) [4].

Basic reproductive number (R0)
Simulation results showed that the overall estimate of the R0 was 1.43 (Fig. 3). However, the estimates differed by age-groups with the highest of 4.7 for the group of 10-14 years of age. Generally, the age-groups with a higher contact rate had also a higher R0. Simulations of epidemics with varied intervention strategies showed that the Re can be reduced below one if the vaccination program with 85% coverage were deployed as far as five months before the
introduction of the index case (time zero) or as late as one week after that (Fig. S4). This coverage threshold was tested as it is the highest vaccine coverage currently achieved worldwide for some diseases, e.g. Hepatitis B, measles, and polio [39]. Late initiations of similar interventions from one to five months after the time zero gradually shift the Re to the outbreak domain.

A lower vaccination coverage of 33% appeared not protective and posed a potential of outbreak regardless the time of vaccination program (Fig. S4). This coverage was tested as it is a theoretical protective threshold, i.e., 1-1/R0 [40]. Note that the tested time window of five months before the appearance of the index case was chosen based on the windows of opportunity for EBOV vaccination [13]. As of now, no data are available on the secondary antibody responses to EBOV; it was assumed that secondary responses are similar to the primary responses.

Case-fatality rate

Simulations showed that the case-fatality rate in the absence of intervention is 90.93% (Fig. S5) which falls in the range of literature estimates of 0.4 to 0.91 [4]. Furthermore, simulation results showed that all the intervention strategies mentioned previously can reduce the case-fatality rate. These results highlight a benefit of vaccination programs even they are late, i.e., they can reduce the disease severity in newly infected subjects after the vaccination program. As such, relying solely on R0 as the indicator for evaluating intervention programs could have overlooked this life-saving aspect.

Epidemic final size

Theoretical analyses of epidemic models showed when the R0 is larger than one, the final size of an epidemic will converge to a two points distribution: either the epidemic dies out with a small number of infected cases or the epidemic takes off and converges to a normal distribution [40]. Simulation results confirmed this epidemic behavior (Fig. 4). The results showed that without intervention, EBOV had approximately 50% to infect more than half the population. The introduction of vaccination programs at both the coverage thresholds previously mentioned and at any vaccination time points under assessments were able to scale down the epidemic

---

Figure 3. Estimates of the basic reproductive number without any intervention, overall and by age-groups. Simulations of a network of size ten thousand during a period of one year. One thousand simulations were run, each time with a random index case. At the end of each simulation, networks of infected nodes were extracted to compute the average number of secondary infections.
size (Fig. 4). The two points epidemics size distribution gradually converged to a uni-modal distribution centring at a low infected fraction when the vaccination programs were deployed earlier. The high vaccine coverage strategy can effectively eliminate the possibilities of having a major outbreak infecting a large proportion of the population. This can be achieved when the vaccination programs were deployed any time from a week to five months before time zero.

A random vaccination program covering 33% of the population one week before the epidemic can reduce the final size by more than 100 times compared to a no intervention scenario. However, the low coverage strategy still showed a small probability that epidemics can become major whereas the high coverage strategies did not. Vaccination programs deployed during the epidemics can also substantially reduce the epidemics size. The intervention conducted one month after time zero can also reduce the final size by more than ten times. These interventions not only able to reduce the final size, but they can also increase the epidemics extinction probability.

Discussion

Epidemic modelling aims to obtain generalized solutions to questions such as whether or not a substantial population fraction is getting infected? how large would the outbreak spread? and how can the outbreak be mitigated with certain intervention approaches [6,40]. Answering those questions requires the use of assumptive parameters as well as actual outbreak data [6,14,22,40]. Our results showed that using information on within-host infection dynamics allows the identification of those key characteristics in the disease transmission.

Estimates of the incubation period suggest a contact tracing period of three weeks for Ebola epidemics, matching the current WHO’s recommendation of 21 days [42]. Estimates of the delay distributions agreed that EBOV infected subjects can be infectious from day 3 up to three weeks post infection [4]. Understanding of these delay distributions is critical in both clinical and
epidemiological perspectives [43]. These distributions, however, are most often only partially observed in practice: it is difficult to know the exact time of exposure to the pathogen or to have complete outbreak data [5,44]. As such, parameter estimation of these distributions have been relied on testing and comparing different distributional assumptions [44]. In this paper, mechanistically generated transmission characteristics using viral dynamics remarkably resemble literature estimates of Ebola. This approach is thus promising and practical given the accumulating experimental data on varieties of pathogens, notably, the one that as yet unknown in epidemic contexts.

To determine infection outcomes, the threshold AUC was chosen based on suggestions from empirical data in humans [34] and non-human primates [24,25]. Simulations of the epidemics using this threshold revealed faithful estimates of the EBOV case-fatality rate (Fig. S5), supporting the use of the total viral load (AUC) as a criterion for determining infection outcomes. Although a more precise threshold criterion is desirable, it might not be feasible to obtain in practice considering inherent ethical reasons. Thus a similar criterion could be considered when adapting this approach to other infectious diseases, but ideally with dedicated experimental data.

Different classes of network models have been proposed, but they cannot reproduce properties observed in real world networks [45]. Thus, a network model obeying empirical data provides a more solid ground for epidemic simulations. Apart from mimicking the contact data properties, our network model can express age-related infection traits via the assigned age attributes. It was used here to express individual differences in the susceptibility to viral infection - an important element in a realistic disease transmission. Although contact data might not be available for a certain target area, the assortative patterns of human contacts and the highly skewed distribution of the number of contacts might hold true across regions and cultures [35,46]. Thus, the model demonstrates a simple way to bring empirical data into epidemic modelling studies.

Previous R0 estimates based on EBOV outbreak data were diverse, depending on model choices and assumptions [4]. Our estimate of R0 was 1.4 which is within the range of the previous estimates, ranging from 1.2 to 2.6, with some exceptional estimates up to 4.7 and 8.3 [4]. Notably, the estimates differed by age-groups with the highest of 4.7 for the group of 10-14-years of age. Although these estimates depend on Sierra Leon’s age-structure, the differences of R0 estimate stress the role of the age-structure and contact patterns in the estimation of R0, prompting that age-specific intervention strategies should be considered. The estimates by sub-groups single out the effort required to control the epidemic [7]. With current assumptions, targeting interventions to the group 5-20-years of age could be an effective strategy. Note that the differences of R0 by age-group could explain the wide variation of the previous estimates where different samples were employed [4].

The following assumptions were used in the paper given the lack of specific experimental data, but further efforts to produce data are needed to refine and to adapt to other settings: (i) Secondary antibody responses are the same as primary responses: This underestimates the effect of the vaccination strategies conducting before the epidemics. Experimental studies on secondary immune responses to EBOV infection are needed, especially those with a longer follow-up period. (ii) The transmission potential is directly related to viral load: This is although simple and reasonable, but different types of relationship, such as non-linear, might exist [10]. Dedicated animal experiments to define the exact relationship between the viral load the ability to transmit the virus are needed. (iii) The contact pattern is the same between European countries and Sierra Leone: Although the contact patterns seemed similar across countries [35], a more sociable population would increase the estimate of R0. (iv) Infection statuses have no influences on the network structure, except those were buried. This could overestimate R0 [47]. Taking people’s behaviour changes into epidemic modelling remains a grand challenge [47]. (v) Susceptibility to EBOV infection is similar to a general viral infection disease: Studies on susceptibility functions are lacking and require more attentions of the infection research community.

Throughout this paper, we showed the possibilities to investigate practical and intriguing questions using a within-host viral dynamic model. The advantage of this approach is the
availability of experimental data and the possibility of conducting experiments to characterize epidemic transmission. Therefore, *in vitro* and *in vivo* studies of infectious agents could be seamlessly integrated into studies of between-host transmission, promoting evidence-based public health practices.

**Funding**

This work was supported by the Alfons und Gertrud Kassel-Stiftung and iMed- the Helmholtz Initiative on Personalized Medicine. VKN has been supported by the Presidents Initiative and Networking Funds of the Helmholtz Association of German Research Centres (HGF) under contract number VH-GS-202.

**Authors’ contributions**

EAHV supervised the project. VKN designed the modelling and performed the simulations. VKN, EAHV and RM provided and analysed the data. VKN, EAHV and RM discussed and wrote the manuscript. All authors reviewed the manuscript.

**Competing interests**

The authors declare that they have no any competing interests.

**References**


36. Statistics Sierra Leone (SSL) and ICF International. *Sierra Leone Demographic and Health Survey 2013*. Freetown, Sierra Leone and Rockville, Maryland, USA: SSL and ICF International., 2014.


41. A. Bowman and A. Azzalini. *sm: nonparametric smoothing methods* (version 2.2-5.4).


