

Temporal Variations in the Effective Reproduction Number of the 2014 West Africa Ebola Outbreak

September 18, 2014 · Research

Authors

[Sherry Towers](#) [Oscar Patterson-Lomba](#) [Carlos Castillo-Chavez](#)

Abstract

Background

The rapidly evolving 2014 Ebola virus disease (EVD) outbreak in West Africa is the largest documented in history, both in terms of the number of people infected and in the geographic spread. The high morbidity and mortality have inspired response strategies to the outbreak at the individual, regional, and national levels. Methods to provide real-time assessment of changing transmission dynamics are critical to the understanding of how these adaptive intervention measures have affected the spread of the outbreak.

Methods

In this analysis, we use the time series of EVD cases in Guinea, Sierra Leone, and Liberia up to September 8, 2014, and employ novel methodology to estimate how the rate of exponential rise of new cases has changed over the outbreak using piecewise fits of exponential curves to the outbreak data.

Results

We find that for Liberia and Guinea, the effective reproduction number rose, rather than fell, around the time that the outbreak spread to densely populated cities, and enforced quarantine was imposed on several regions in the countries; this may indicate that enforced quarantine may not be an effective control measure.

Conclusions

If effective control measures are not put in place, and the current rate of exponential rise of new cases continues, we predict 4400 new Ebola cases in West Africa during the last half of the month of September, with an upper 95% confidence level of 6800 new cases.

Funding Statement

This publication was made possible by grant number 1R01GM100471-01 from the National Institute of General Medical Sciences (NIGMS) at the National Institutes of Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS. OPL additionally gratefully acknowledges the support of NIH training grant T32AI007358-26. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Introduction

Ebola virus disease (EVD) is a severe, often fatal viral infection in humans, with a case fatality risk (CFR) of up to 90% in previous outbreaks ^{1,2,3,13,14}. There have been 13 recorded outbreaks since 2000, associated with at least two distinct strains of the Ebola virus (Ebola-Zaire and the Ebola-Sudan) ²³, but the current West African outbreak has spread over the largest geographic area and caused the greatest number of infections and deaths ^{3,15,16,22,23}. Critically, unlike previous epidemics, this outbreak has spread to densely populated areas, increasing the risk of international spread ²²; the problem is compounded by lack of resources for effective quarantine and isolation in the under-developed countries that have been affected, and the high mobility of the population in a region with porous borders.

The spread of EVD requires direct contact with infected blood, tissue, or body fluids of either the living ill or the recently deceased, and the disease is particularly prone to transmission in unprotected homecare settings and during traditional burials ^{24,25}. Symptoms typically appear after 2–21 days, and the disease quickly progresses and kills most infected patients within a few days ^{13,14}. No licensed vaccine or specific treatment is currently available ²⁶, leaving improved hygiene, quarantine, isolation, and social distancing as the only potential interventions. In an attempt to quell the outbreak, on August 1st the governments of Guinea, Liberia, and Sierra Leone announced plans to impose a military-enforced mass quarantine of entire regions and villages (referred to as *cordons sanitaire*), in an attempt to prevent spread of the disease to other areas ¹⁷. The spread of disease within the isolated areas has typically been allowed to run its course ^{19,20}.

Assessing the net positive or negative impact of such attempted intervention measures involves assessment of a complex adaptive system, where the system dynamics are intimately connected to the impact of multiple feedback mechanisms that evolve over time ⁹. Because the situation in West Africa is rapidly evolving in time, and attempts at intervention measures are far ranging and also rapidly evolving, understanding how best to control the outbreak, and/or assessing whether or not current attempts to contain the spread are working, poses unique challenges.

In this analysis, we examine current outbreak incidence data for Guinea, Sierra Leone, and Liberia up to September 8, 2014, as estimated by the World Health Organization (WHO) ^{3,4}. Using novel methodology, we assess the impact of adaptive response measures by fitting piecewise exponential curves along the data time series to estimate the evolving rate of exponential rise (or decline) in cases. We then use a Susceptible, Exposed, Infected, Recovered (SEIR) model to estimate the temporal patterns of the effective reproduction number for the outbreak in each country. In the following sections, we describe our sources of data, and the statistical and modeling methods used to estimate the effective reproduction number.

Methods and Materials

Data

As part of the HealthMap project ⁴, epidemiologists and researchers at the Boston Children's Hospital have compiled from WHO reports of the time series of confirmed and probable EVD case counts during the 2014 West African outbreak ³. They have made the compilation publicly available online at

www.healthmap.org/ebola. The HealthMap data do not include separate information regarding the confirmed and probable cases. For this analysis, the time series of EVD cumulative confirmed and probable cases and deaths in Guinea, Sierra Leone, and Liberia were obtained from HealthMap up to September 4, 2014, and supplemented with more recent data from the WHO website.

The estimated daily average of the new case incidence data for the three countries are shown in Figure 1. Updates to the WHO case counts occur irregularly in time (situation updates have occurred every 2 to 9 days since July 1st), thus to estimate the daily average incidence data between two updates, we take the new incidence accumulated between updates, and divide by the time between updates.

Statistical Methods

Cumulative incidence data are highly inter-correlated from point to point, and thus are inappropriate for fitting with standard least squares or likelihood fitting methods, which have as an underlying assumption that the observations are independent^{5,6}. Fitting to cumulative incidence data, instead of the time series of number of new cases, can thus lead to dramatically under-estimated confidence intervals on the model parameters, and also potential biases in the fit central values, since points further along the time series (which are nearly 100% correlated) are given too much weight in the fit. In this analysis we thus fit to the time series of newly identified cases for each country.

Using a Negative Binomial likelihood fit to account for over-dispersion in the data⁷, for each country we fit piecewise exponential curves along the incidence time series, in which ten contiguous points are fit at a time, and the local exponential rate of rise, ρ_{local} , is estimated for each. If we have M incidence data points Y_i^{obs} , recorded between times t_i^{min} to t_i^{max} , where $i=1, \dots, M$, then the predicted incidence for each point under an exponential rise hypothesis with some rate, ρ_{local} , is⁸:

$$Y_i^{\text{pred}} \sim \int_{t_i^{\text{min}}}^{t_i^{\text{max}}} dt \exp(\rho_{\text{local}} t)$$

which simplifies to

$$Y_i^{\text{pred}} \sim \exp(\rho_{\text{local}} t_i^{\text{max}}) - \exp(\rho_{\text{local}} t_i^{\text{min}}) \quad \text{Eqn 1}$$

where the Y_i^{pred} are normalized such that $\sum_i M Y_i^{\text{pred}} = \sum_i M Y_i^{\text{obs}}$.

Because of the irregular updates in the data, the time spanned by ten contiguous points ranges from 21 to 34 days. Sensitivity analyses indicate the conclusions of this analysis do not depend on the exact number of contiguous points used for the fits.

We also fit an exponential curve to the entire time series for each country, from July 1st onwards (after July 1st the WHO EVD case count updates for each country became more regular, occurring on average every three days).

Estimation of the effective reproduction number from the local rates of exponential rise

The basic reproduction number, \mathcal{R}_0 , of an infectious disease is the average number of secondary cases a case generates during the course of its infectious period, in a completely susceptible population. The basic reproduction number is affected not only by the ease at which an infection is transmitted upon

contact between two hosts, but also by the contact rate between members of the population. Because of heterogeneity in susceptibility, infectiousness, and differing social dynamics, the estimates of \mathcal{R}_0 for an infectious disease tend to be different across populations. For example, rural and densely populated urban areas are likely to have substantial differences in contact rates¹⁰. Traditionally \mathcal{R}_0 has been computed under simplifying assumptions, like the existence of a constant contact structure, or under the assumption that interventions like quarantine and isolation are uniformly applied from the very beginning of the epidemic. However, particularly in the case of a disease with high morbidity and mortality, it has been shown that human behaviors adapt to limit spread of the disease⁹, thus limiting the usefulness of \mathcal{R}_0 as a means to assess the efficacy of evolving intervention strategies, and to forecast the future progression of the outbreak.

Given these limitations of \mathcal{R}_0 , in this analysis we thus instead assess the time evolution of the effective reproduction number, \mathcal{R}_{eff} , which is a dynamic estimate of the average number of secondary cases per infectious case in a population composed of both susceptible and non-susceptible individuals during the course of an outbreak. The effective reproduction number is a function of, among other things, the probability of transmission upon contact of the disease, the contact rate within the population, and the number of susceptible individuals in the population, all of which may be time dependent.

In an epidemic, given a short enough period of time, the value of \mathcal{R}_{eff} during that time period can be considered to be approximately constant. Given a mathematical model for the spread of infectious disease, such as a classic Susceptible, Exposed, Infected, Recovered (SEIR) model, the equations of the mathematical model can be linearized about this temporary “equilibrium”, to determine the predicted local rate of exponential rise of the epidemic curve, ρ_{local} . For the SEIR model, this is related to \mathcal{R}_{eff} by

$$\mathcal{R}_{\text{eff}} = (1 + \rho_{\text{local}}/\gamma)(1 + \rho_{\text{local}}/\kappa) \quad \text{Eqn 2}$$

where $1/\kappa$ and $1/\gamma$ are the average incubation and infectious periods of the disease, respectively¹².

With estimates of ρ_{local} from piecewise exponential rise fits to the incidence data from an outbreak (along with estimates of the incubation and infectious periods of the disease), we can thus use Equation 2 to obtain estimates of the temporal behavior of \mathcal{R}_{eff} , in essence approximating the temporal behavior with a piecewise step-function.

Typically, if the transmissibility and contact rates in the population remain constant, one expects to see the effective reproduction number decrease in time as the fraction of susceptible individuals in the population decreases. However, if \mathcal{R}_{eff} increases, it can signal that, for instance, contact rates within the population have increased, the virus has become more transmissible, and/or that attempted intervention strategies are actually making the outbreak worse.

While we use an SEIR model and Equation 2 in this analysis to estimate \mathcal{R}_{eff} , it needs to be stressed that our methodology for estimating the temporal variation in the effective reproduction number from the local rates of exponential rise in incidence can be applied with any mathematical model appropriate to simulate the spread of a particular disease.

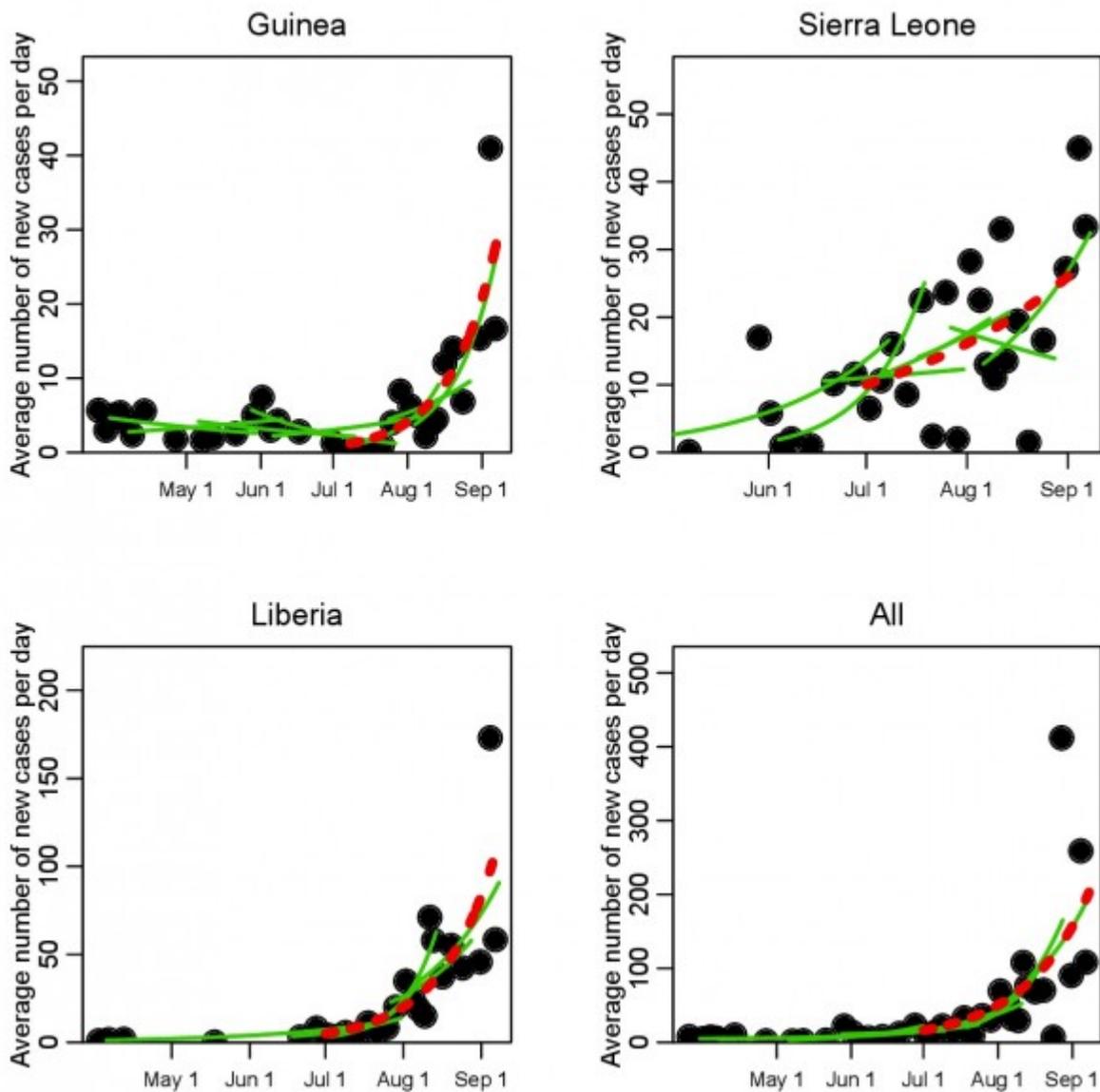
Results

In Figure 1 we show the time series of the recorded daily average incidence data for Guinea, Sierra

Leone, and Liberia. Overlaid in green, we show a selection of the piecewise exponential fits along the incidence time series, in which ten contiguous points are fit at a time, and the local exponential rate of rise, ρ_{local} , is estimated. For clarity of presentation, we only show every third fit. The fit of an exponential to the time series from July 1st onwards is shown overlaid with a red dotted line.

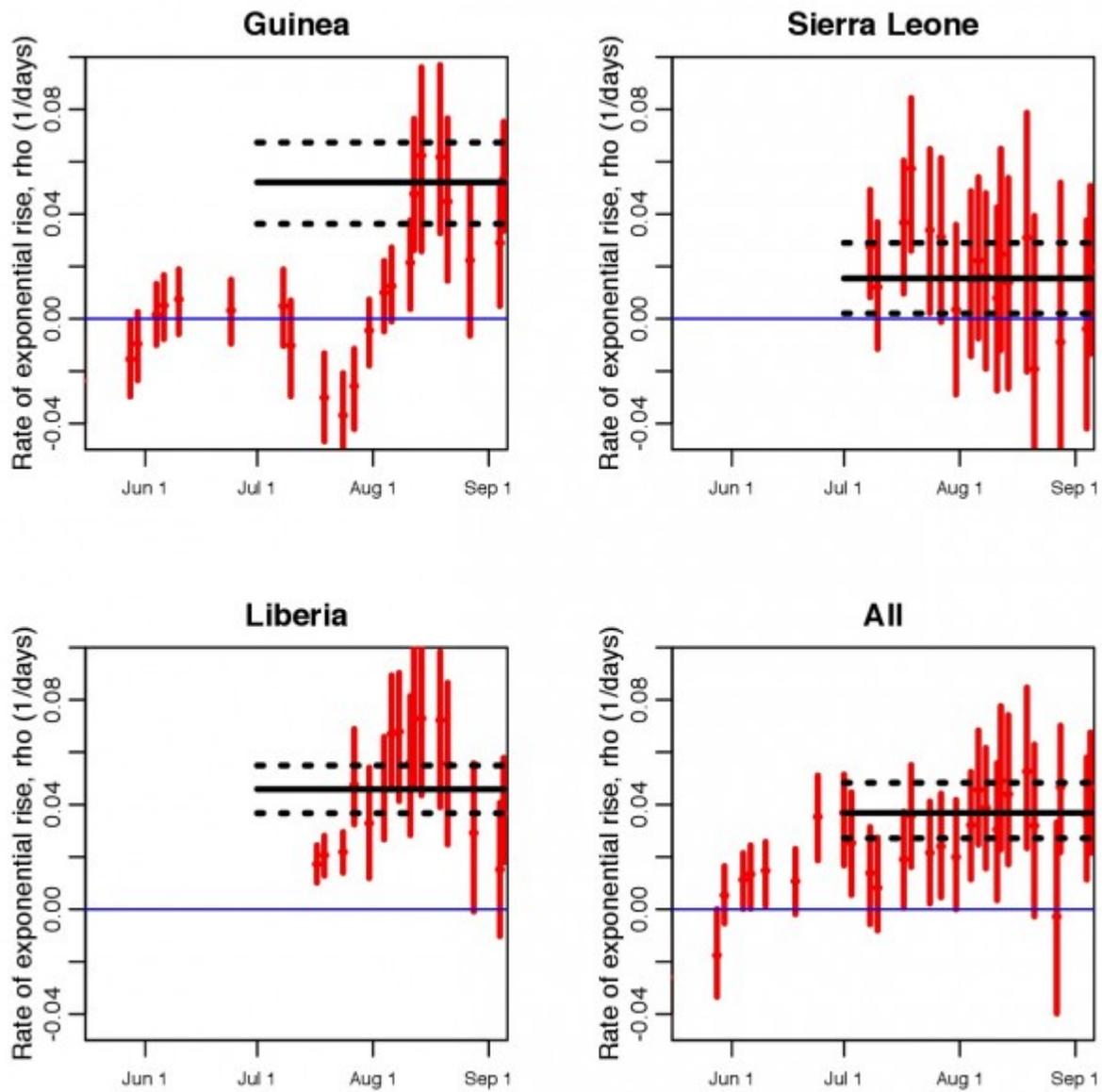
In Figure 2 we show the time series of the estimates of the local exponential rise ρ_{local} (along with the 95% confidence intervals obtained from the Negative Binomial likelihood fit) for each of the piecewise exponential fits for each country. We also indicate the estimate of the the exponential rise from a fit to the time series from July 1st onwards for each country.

Note that the rates of ρ_{local} are significantly correlated from point to point.



F
1
T
S
C
i
l
C
E
C
i
V
A
T
S
C
r
e
r
C
r
E
C
P
C
C
t
i
P
C
t
2
V
A
C
f
C
S

Leone, and Liberia (dots). The green lines show a selection of the piecewise exponential fits to the data (not all fits are shown to clarify the presentation); a moving window of groups of 10 contiguous points are taken at a time, and the rate of exponential rise estimated for those 10 points. The results for the estimations of the exponential rise for the full set of piecewise fits are shown in Figure 2. Shown in red is the fitted exponential rise from July 1st onwards.



F
2
E
r
c
e
r
f
e
f
E
n
c
e
r
f
i
t
e
f
i
t
t
e
c
E
i
c
a
s
i
F
1
e
n
v
c
c
1
c
i
c

time series points are taken at a time, and the rate of exponential rise estimated for those 10 points. The dates shown on the x axis are last date in each contiguous set of 10 points, and the vertical error bars denote the 95% confidence interval. The horizontal black line shows the estimated rate of rise of an exponential fit to the incidence time series from July 1st to Sep 8th, with the black dotted lines indicating the 95% confidence interval.

Discussion

For each country in Figure 2 the solid and dotted black lines denote the central value and 95% CI, respectively, of the rate of exponential rise of case incidence data between July 1st and September 8th. The estimates of the rates of exponential rise between July 1st and September 8th are $\rho=0.052 \pm 0.008$, $\rho=0.015 \pm 0.008$, and $\rho=0.046 \pm 0.005$ for Guinea, Sierra Leone, and Liberia, respectively. The value estimated from the combined data samples is $\rho=0.037 \pm 0.006$. The values are statistically consistent for the Guinea and Liberia data (one-sided Z test $p = 0.28$). However, the estimate for the Sierra Leone data is significantly lower than the estimates obtained from the Guinea and Liberia data (one-sided Z test $p = 0.0005$, and $p = 0.0003$, respectively). In Table 1 we show the values of the effective reproduction number, R_{eff} , as estimated from Equation 2 under several hypotheses for the incubation and infectious periods (the incubation and infectious periods for this outbreak are poorly known, and have shown wide apparent variation in previous outbreaks ^{13,14}). The reproduction numbers estimated by this analysis are in broad agreement with the reproduction numbers estimated from this and previous outbreaks, which range from approximately 1.3 to 2.7 ^{13,14,22}.

Table 1: Estimates of the effective reproduction number between July 1 to present

Estimates of the effective reproduction number for the 2014 West African EBV outbreak data, for an SEIR model as obtained from Equation 2 under several hypotheses for the incubation and infectious periods. The estimates and the 95% confidence intervals are derived from the rate of rise of an exponential fit to the new case incidence data for the various countries between July 1st and September 8th (the beginning of July is when the World Health Organization began frequent updates of the new case information for all three countries).				
Hypothesized incubation and infectious periods, respectively (days)	Guinea	Liberia	Sierra Leone	All
5 and 5	1.6 [1.4,1.8]	1.5 [1.4,1.6]	1.2 [1.0,1.3]	1.4 [1.3,1.5]
7 and 5	1.7 [1.5,2.0]	1.6 [1.5,1.8]	1.2 [1.0,1.4]	1.5 [1.3,1.7]
10 and 5	1.9 [1.6,2.3]	1.8 [1.6,2.0]	1.2 [1.0,1.5]	1.6 [1.4,1.8]
5 and 7	1.7 [1.5,2.0]	1.6 [1.5,1.8]	1.2 [1.0,1.4]	1.5 [1.3,1.7]
7 and 7	1.9 [1.6,2.2]	1.8 [1.6,1.9]	1.2 [1.0,1.5]	1.6 [1.4,1.8]
10 and 7	2.1 [1.7,2.5]	1.9 [1.7,2.2]	1.3 [1.0,1.6]	1.7 [1.5,2.0]
5 and 10	1.9 [1.6,2.3]	1.8 [1.6,2.0]	1.2 [1.0,1.5]	1.6 [1.4,1.8]
7 and 10	2.1 [1.7,2.5]	1.9 [1.7,2.2]	1.3 [1.0,1.6]	1.7 [1.5,2.0]
10 and 10	2.3 [1.8,2.8]	2.1 [1.9,2.4]	1.3 [1.0,1.7]	1.9 [1.6,2.2]

Two recent papers have attempted to estimate the effect of control measures on the spread of EBV during the West African outbreak. The first (Althaus, 2014) employs a Poisson likelihood fit to the

cumulative incidence under the assumption of an SEIR model with control that begins the day the first case appears in the population, and causes the transmission rate to exponentially decay¹⁵. A second recent paper (Fisman *et al* , 2014) also fits to the cumulative incidence curve, with a model that assumes that the incidence over time is a function of the basic reproduction number, and a control parameter that causes the incidence to decay¹⁶. The Althaus and Fisman *et al* estimates of the reproduction number for the Guinea and Liberia outbreaks are within our corresponding 95% confidence intervals under the same hypotheses for the incubation and infectious periods. However, both analyses predict significantly higher reproduction numbers for the Sierra Leone outbreak. The discrepancy is likely due to the fact that both analyses employ models that assume that control measures are immediately employed and effective; the results thus inherently attribute the low effective reproduction number to a high basic reproduction number being modulated by highly effective control measures. Neither analysis provides a comparison of the model fit results to those of a reduced model where control is excluded, so it is unknown what the estimates for the reproduction number would be if control was not inherently assumed.

We stress again that our analysis makes no assumptions about whether or not control measures have been employed, and/or when they were employed. We simply examine the local rates of exponential rise to estimate how the effective reproduction number appears to be changing in time. We note that up until mid-August, the values of ρ_{local} for Guinea are significantly below the average central value obtained from the fit the the data from July 1st onwards, and were rising (unless otherwise noted, all subsequent statistical tests are Bonferroni corrected one-sided Z tests, with rejection of the null hypothesis occurring when $p < 0.05/k$, where k is the number of values being compared). The dip in the two recent Guinea data points compared to the central value is not significant. In addition, the dip in the two recent estimates of ρ_{local} from the Liberia data are also not significant. However, the estimates of ρ_{local} for the three earliest points in the Liberia local time series are significantly lower than the central value, and rising. None of the estimates of ρ_{local} from the Sierra Leone data and combined data for all countries are significantly different than the central value.

It is unclear why the transmission rate of the disease apparently rose for both Guinea and Liberia between mid July to mid August, and why the transmission rate in Sierra Leone is systematically lower, although it is important to note that the WHO data are obtained from rudimentary surveillance systems in under-developed countries, under the stress of a rapidly evolving outbreak situation. The temporal patterns we observe may thus partly be due to variations in surveillance during the outbreak, under-reporting, and/or reporting delays. In addition, serial passage of the disease as the outbreak progresses may be leading to increased pathogenicity, and a subsequent increasingly larger rate of increase in case counts . However, it also must be considered that otherwise well intentioned attempts at intervention may in fact be making the situation worse, at least in some regions; in a joint meeting of officials from Guinea, Liberia, and Sierra Leone on Aug 1st , it was announced that *cordons sanitaire* would be implemented, to seal off the villages and regions worst hit by the outbreak, in an attempt to limit its spread outside those areas¹⁷. In addition, at that time Liberia closed all schools and non-essential government offices¹⁸, and two weeks later imposed a military-enforced *cordon sanitaire* on the West Point slum of Monrovia, sparking riots¹⁹. It must be stressed that *cordons sanitaire* do not generally attempt to limit disease spread within the quarantined areas, and the implementation of these measures in West Africa has received criticism due to the fact that the quarantined areas are at risk for crowding, lack of medical and basic services, and poor sanitation, potentially increasing the spread of

disease within those areas²⁰. It was around the period of time that the *cordons sanitaire* were first imposed, and when the outbreak moved to the densely populated cities of Conakry and Monrovia, that the exponential rise in new cases in Guinea and Liberia increased, rather than went down. Again, it is unclear why the rate of exponential rise in Sierra Leone was apparently unaffected by these events.

Based on our estimates of the exponential rise in cases between July 1st to the beginning of September, if this rise is to continue unabated, there will be approximately 4400 new EVD cases in West Africa during the last half of the month of September (95% CI [3000, 6800]), 500, 900, and 3000 of which will be in Sierra Leone, Guinea, and Liberia, respectively.

Summary

We have presented an analysis where piecewise exponential fits are employed to estimate the local rate of exponential rise in cases in the 2014 West Africa Ebola outbreak. We have shown that these local rates of exponential rise can be used in conjunction with a mathematical model to estimate the temporal changes in the effective reproduction number (in the case of this analysis, we employed an SEIR model). Our analysis indicates that the spread of the disease to densely populated cities, and/or the imposition of *cordons sanitaire* in West Africa may have accelerated the spread of the disease in some regions. Our methodology is novel, and is applicable to any outbreak, and any mathematical disease model.

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgements

The authors wish to acknowledge useful discussions with Gerardo Chowell regarding these studies.

References

1. Sanchez, Anthony. Ebola Viruses. John Wiley & Sons, Ltd, 2001.
2. WHO. Ebola Virus Disease in Guinea, Mar 23 2014.
[REFERENCE LINK](#)
3. World Health Organization. Global Alert and Response situation reports: Ebola response roadmap.
[REFERENCE LINK](#)
4. Brownstein, John S., et al. "Surveillance Sans Frontieres: Internet-based emerging infectious disease intelligence and the HealthMap project." PLoS medicine 5.7 (2008): e151.
5. Hamilton, James Douglas. Time series analysis. Vol. 2. Princeton: Princeton university press, 1994.
6. Casella, George, and Roger L. Berger. Statistical inference. Vol. 2. Pacific Grove, CA: Duxbury, 2002.

7. Lloyd-Smith, James O. "Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases." *PloS one* 2.2 (2007): e180.
8. Chowell, Gerardo, et al. "The influence of climatic conditions on the transmission dynamics of the 2009 A/H1N1 influenza pandemic in Chile." *BMC infectious diseases* 12.1 (2012): 298.
9. Fenichel, Eli P., et al. "Adaptive human behavior in epidemiological models." *Proceedings of the National Academy of Sciences* 108.15 (2011): 6306-6311.
10. de Jong, Mart CM, Odo Diekmann, and Hans Heesterbeek. "How does transmission of infection depend on population size." *Epidemic models: their structure and relation to data* 5.2 (1995): 84-94.
11. Wallinga, Jacco, and Marc Lipsitch. "How generation intervals shape the relationship between growth rates and reproductive numbers." *Proceedings of the Royal Society B: Biological Sciences* 274.1609 (2007): 599-604.
12. Yan, Ping, and Zhilan Feng. "Variability order of the latent and the infectious periods in a deterministic SEIR epidemic model and evaluation of control effectiveness." *Mathematical biosciences* 224.1 (2010): 43-52.
13. Legrand, J., et al. "Understanding the dynamics of Ebola epidemics." *Epidemiology and infection* 135.04 (2007): 610-621.
14. Chowell, Gerardo, et al. "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 (2004): 119-126.
15. Christian Althaus. "Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa." *PLoS currents*, 1, Sep 2, 2014.
16. Fisman, David, Khoo, Edwin, and Tuite, Ashleigh. "Early epidemic dynamics of the West African 2014 Ebola outbreak: Estimates derived with a simple two-parameter model." *PLoS currents*, 1, Sep 8, 2014.
17. Agence France-Presse. "Ebola-hit African states seal off outbreak epicentre." Aug 1 2014.
[REFERENCE LINK](#)
18. Agence France-Presse. "Liberia shuts schools in battle against Ebola: president", Jul 30 2014.
[REFERENCE LINK](#)
19. The Guardian. "Ebola quarantine in Liberia's capital sparks violence in slum", Aug 21 2014.
[REFERENCE LINK](#)
20. Washington Post. "Quarantining an entire liberian slum to fight Ebola is a recipe for disaster", Aug 28 2014.
[REFERENCE LINK](#)
21. Agence France-Presse. "Sierra Leone to visit every home to track down Ebola dead", Sep 8 2014.
[REFERENCE LINK](#)
22. Gomes MFC, Pastore y Piontti A, Rossi L, Chao D, Longini I, Halloran ME, Vespignani A.

Assessing the International Spreading Risk Associated with the 2014 West African Ebola Outbreak. PLOS Currents Outbreaks. 2014 Sep 2. Edition 1.
doi:10.1371/currents.outbreaks.cd818f63d40e24aef769dda7df9e0da5.

23. United States Centers for Disease Control and Prevention "Ebola Outbreaks 2000-2014", Sep 10, 2014

[REFERENCE LINK](#)

24. Weyer, Jacqueline, Lucille H. Blumberg, and Janusz T. Paweska. "Ebola virus disease in West Africa-an unprecedented outbreak: editorial." South African Medical Journal 104.8 (2014): 555-556.

25. World Health Organization "Media centre: Ebola virus disease fact sheet". April, 2014

[REFERENCE LINK](#)

26. United States Centers for Disease Control and Prevention "Ebola Hemorrhagic Fever", Sep 2014

[REFERENCE LINK](#)

27. Leroy, E. M., et al. "Human asymptomatic Ebola infection and strong inflammatory response." The Lancet 355.9222 (2000): 2210-2215.

1 Comment

PLOS Currents

September 19, 2014

Correction

The confidence intervals for expected number of cases have been corrected following the discovery of an error by the authors. The correct 95% confidence interval on the expected number of cases by the end of the month is [3000,6800]. This has been updated in the text of the manuscript.
