

# Early transmission dynamics of Ebola virus disease (EVD), West Africa, March to August 2014

H Nishiura (nishiurah@m.u-tokyo.ac.jp)<sup>1</sup>, G Chowell<sup>2,3</sup>

1. Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

2. School of Human Evolution and Social Change, Arizona State University, Tempe, Arizona, United States

3. Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, Maryland, United States

## Citation style for this article:

Nishiura H, Chowell G. Early transmission dynamics of Ebola virus disease (EVD), West Africa, March to August 2014. *Euro Surveill.* 2014;19(36):pii=20894. Available online: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20894>

Article submitted on 23 August 2014 / published on 11 September 2014

The effective reproduction number,  $R_t$ , of Ebola virus disease was estimated using country-specific data reported from Guinea, Liberia and Sierra Leone to the World Health Organization from March to August, 2014.  $R_t$  for the three countries lies consistently above 1.0 since June 2014. Country-specific  $R_t$  for Liberia and Sierra Leone have lied between 1.0 and 2.0.  $R_t < 2$  indicate that control could be attained by preventing over half of the secondary transmissions per primary case.

## Introduction

The largest and first regional outbreak of Ebola virus disease (EVD) has been unfolding in West Africa since approximately December 2013, with the first cases traced back to southern Guinea [1]. However, the outbreak was not recognised until March 2014 [1], which facilitated the spread to neighbouring Sierra Leone and Liberia through porous borders as well as Nigeria via a commercial airplane on 20 July [2]. The World Health Organization (WHO) declared this EVD epidemic a Public Health Emergency of International Concern on 8 August 2014 [3]. According to phylogenetic analyses, the causative Ebola virus strain is closely related to a strain associated with past EVD outbreaks in Central Africa, and could have been circulating in West Africa for about a decade [4].

A total of 3,707 cases (including 2,106 confirmed, 1,003 probable and 598 suspected cases, respectively) and 1,848 deaths (concerning 1,050 confirmed and 557 probable cases, as well as 241 suspected cases and deaths, respectively) have been reported in Guinea, Sierra Leone, Liberia, Nigeria, and Senegal as of 31 August 2014 [5]. The total number of cases in Guinea, Sierra Leone, Liberia, Nigeria and Senegal have been 771, 1,216, 1,698, 21 and one, respectively. By contrast, the great majority of past outbreaks have been associated with small numbers of reported cases and have been confined to isolated rural areas in Central Africa. For reference, the largest outbreaks in Central

Africa generated 315 cases in Congo in 1976 and 425 cases in Uganda in 2000 [6,7].

The effective reproduction number,  $R_t$ , which measures the average number of secondary cases generated by a typical primary case at a given calendar time, can be helpful to understand the EVD transmission dynamics over time in affected countries as well as gauge the effect of control interventions [8]. Values of  $R_t < 1$  indicate that the epidemic is in a downward trend. By contrast, an epidemic is in an increasing trend if  $R_t > 1$ . The mean reproduction number for EVD has been estimated at 1.83 for an outbreak in Congo in 1995 and 1.34 in Uganda in 2000 prior to the implementation of control interventions [9]. Here we sought to estimate the  $R_t$  in real time in order to assess the current status of the evolving outbreak across countries affected in 2014. We also compare our estimates of the reproduction number for the current outbreak with those previously published for the largest outbreaks in Central Africa and discuss our findings from a public health perspective.

## Methods

### Case data

We analysed the cumulative case counts reported by the WHO [10] as of 26 August 2014. Case counts are classified into three categories, i.e. confirmed, probable and suspected cases. Confirmed cases are laboratory diagnosed by polymerase chain reaction (PCR), positive IgM antibody or viral isolation while suspected cases correspond to individuals presenting fever ( $\geq 38.5^\circ\text{C}$  ( $101^\circ\text{F}$ )) and no favourable response to treatment for usual causes of fever in the area, and at least one of the following clinical signs: bloody diarrhoea, bleeding from gums, bleeding into skin (purpura), bleeding into eyes and urine. Probable cases are suspected cases of EVD with an epidemiological link to a confirmed EVD case [11]. We analysed two different sets of grouped data, i.e. (i) confirmed plus probable

cases and (ii) the total number of reported cases (i.e. confirmed, probable and suspected cases).

Because case counts were reported in irregular time intervals, we estimated daily incidence curves of EVD cases in order to estimate  $R_t$ . For this purpose, we first fit a smoothing spline to country-specific cumulative curves of reported cases. Next we took the daily difference of the cumulative counts to obtain daily incidence time series. Of note, the cumulative case series reflects the diagnostic process (among suspected and probable cases) and sometimes declined as a function of time (e.g. 5 April and 12 July in Guinea and Sierra Leone, respectively). When the difference was negative, we replaced it by 0. The smoothing spline was chosen to obtain a coefficient of determination  $R^2$  at 0.995. Data from Nigeria and Senegal have been omitted due to a limited number of cases recorded in these countries thus far.

### Mathematical model

We employed mathematical modelling together with time- and country-specific incidence data to estimate the  $R_t$ . Thus, here we model the transmission dynamics of EVD using a country-specific next-generation matrix  $\{k_{ij,t}\}$  representing the average number of secondary cases in country  $i$  at time  $t$  generated by a single primary case in country  $j$ . Let  $g_\tau$  represent the probability density function of the generation time of length  $t$  days for EVD. Hence, the expected value of EVD incidence in country  $i$  at time  $t$  is modelled as

$$E(c_{i,t}) = \sum_j k_{ij,t} \sum_{\tau=1}^{\infty} c_{j,t-\tau} g_\tau$$

The univariate version of Equation 1 has been employed by White and Pagano [12,13] in order to jointly estimate  $R_0$  and the generation time distribution of EVD. Assuming that EVD incidence follows a Poisson distribution, the likelihood to estimate  $\{k_{ij,t}\}$  is

$$\prod_t \prod_i \frac{(\sum_j k_{ij,t} \sum_{\tau=1}^{\infty} r_{j,t-\tau} g_\tau)^{r_{i,t}} \exp[-(\sum_j k_{ij,t} \sum_{\tau=1}^{\infty} r_{j,t-\tau} g_\tau)]}{r_{i,t}!}$$

where  $r_{i,t}$  is the estimated daily incidence in country  $i$  on day  $t$  derived from the difference of the smoothing spline fit to the cumulative data as explained above.

Each element of the next-generation matrix is interpreted as the average number of secondary cases generated by a single primary case at time  $t$ . We assume that the per-contact probability of infection and the average generation time do not differ by country. Thus, the contact matrix regulates the relative difference between each pair of entries of the next-generation matrix, and because the contact patterns within and between countries cannot be directly observed, we made a qualitative assumption for the matrix  $\{k_{ij,t}\}$  to approximately capture the pattern of (domestic and transnational) transmission [14], i.e.

$$M_t = \begin{pmatrix} k_{g,t} & \alpha & \alpha \\ \alpha & k_{s,t} & \alpha \\ \alpha & \alpha & k_{l,t} \end{pmatrix}$$

The matrix  $M_t$  qualitatively assumes that there are more frequent within-country transmissions (denoted by  $k_{g,t}$ ,  $k_{s,t}$  and  $k_{l,t}$ , where the subscripts  $g$ ,  $s$  and  $l$  represent Guinea, Sierra Leone and Liberia, respectively) compared with transnational spread. The transnational spread is modelled by a single parameter  $\alpha$ . We employed a piecewise constant model and change the parameters for the above-mentioned elements every seven days. Maximum likelihood estimates of the parameters were obtained by minimising the negative logarithm of Equation 2. Using the most recent incidence estimate  $i_0$  and the exponential growth rate  $r$  as calculated from  $r=(R-1)/12$  (where  $R$  is the most recent reproduction number and 12 is the mean generation time), the expected number of additional cases in 2014 was calculated as

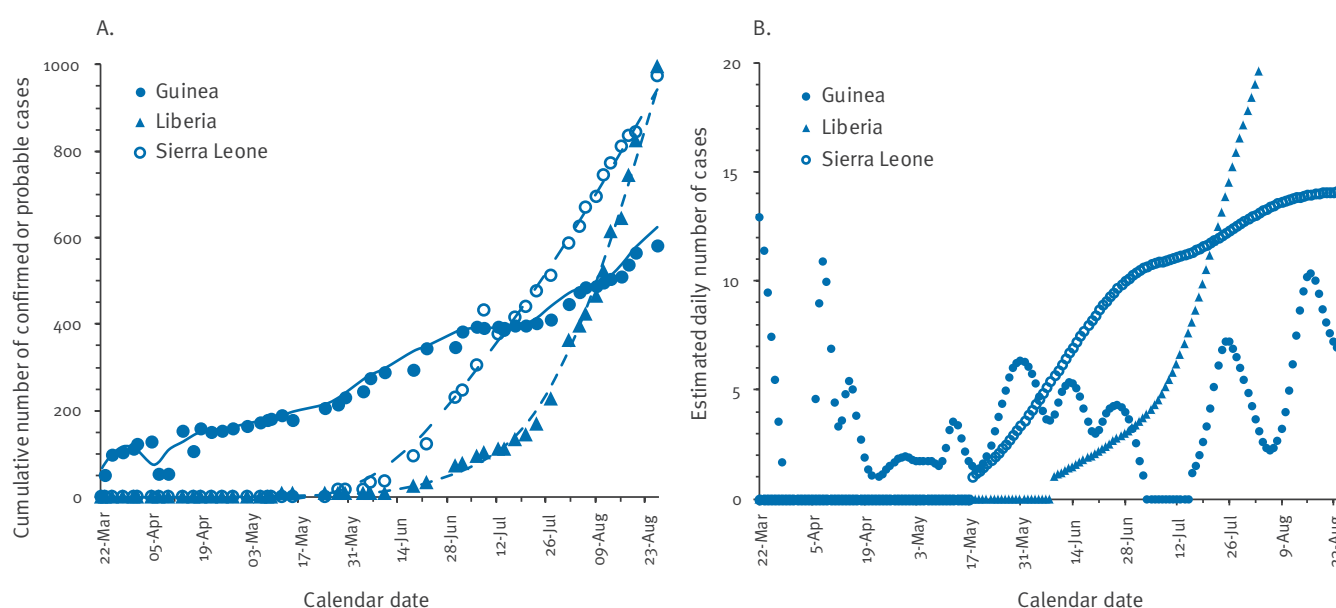
$$I = i_0 \int_0^{120} \exp(rt) dt$$

. The expected cases represent a ‘worst-case’ scenario based on the current situation by assuming a fixed reproduction number  $R$  for the remainder of the year (i.e. approximately 120 days remaining in 2014).

We also computed the  $R_t$  for all countries (hereafter referred to as the ‘global’ estimate of the reproduction number) by calculating the dominant eigenvalue of the estimated next-generation matrices. Moreover, we calculated column sums of the matrices to estimate the average number of secondary transmissions arising in and from a specific country and also extracted estimates of  $2\alpha$ , the value that governs the transnational spread generated by a single primary case. Although White and Pagano achieved the joint estimation of  $R_0$  and generation time distribution [12,13], we assumed that the generation time is known, because our analysis relies solely on the cumulative number of reported cases with irregular reporting intervals. The generation time was assumed to follow an exponential distribution with a mean of 12 days [15], which is known to be close to the mean incubation period [16]. Based on empirical data of the serial interval distribution [15], we also carried out a sensitivity analysis of reproduction numbers by varying the mean generation time between nine and 15 days. The 95% confidence intervals of the  $R_t$  can be computed via bootstrapping methods. However, our study focused on examining model uncertainty associated with the transnational mixing patterns and the mean generation time as model uncertainty in our study is likely more influential on  $R_t$  compared to uncertainty relating to measurement error. In sensitivity analyses, we also examined the impact of varying specified time interval on  $R_t$ . For this purpose, we also

**FIGURE 1**

Cumulative and daily epidemic curves of Ebola virus disease (EVD) in Guinea, Liberia, and Sierra Leone, 23 March–26 August 2014



A) Cumulative number of confirmed or probable cases of EVD reported to the World Health Organization [10]. Solid lines are the smoothing spline fits to cumulative curves for each country with a coefficient of variation  $R^2$  at 0.995.

B) Estimated daily incidence curves based on the smoothing spline model. Data from Nigeria and Senegal have been omitted due to the limited number of cases recorded in these countries thus far.

analysed the piecewise constant model for every six and eight days instead of seven days.

## Results

Figure 1 illustrates the process of deriving daily EVD incidence curves by country from cumulative curves of reported cases. Multiple fluctuations are evident from the incidence curve for Guinea (Figure 1). In Liberia, the early transmission phase did not appear to exhibit sustained growth and was probably driven by case importations during first epidemic month. Exponential growth was subsequently seen, reflecting self-sustaining transmission. Similarly, the incidence curve for Sierra Leone also displayed steady growth since early June. Most recent EVD incidence data for Guinea also showed an increasing pattern.

Our weekly maximum likelihood estimates of the  $R_t$  for each affected country and for the global system in West Africa are displayed in Figure 2. Results indicate that the reproduction number for all countries reached levels below unity in April and May, but has appeared to be continuously above one since early June (Figure 2A). This pattern was robust when using two different datasets (including and excluding suspected cases). Estimates of  $R_t$  using total case reports from June to July 2014, a period during which exponential growth of cases has been observed in Sierra Leone and Liberia, ranged from 1.4 to 1.7, respectively. In the hypothetical worst-case scenario that the current situation with an

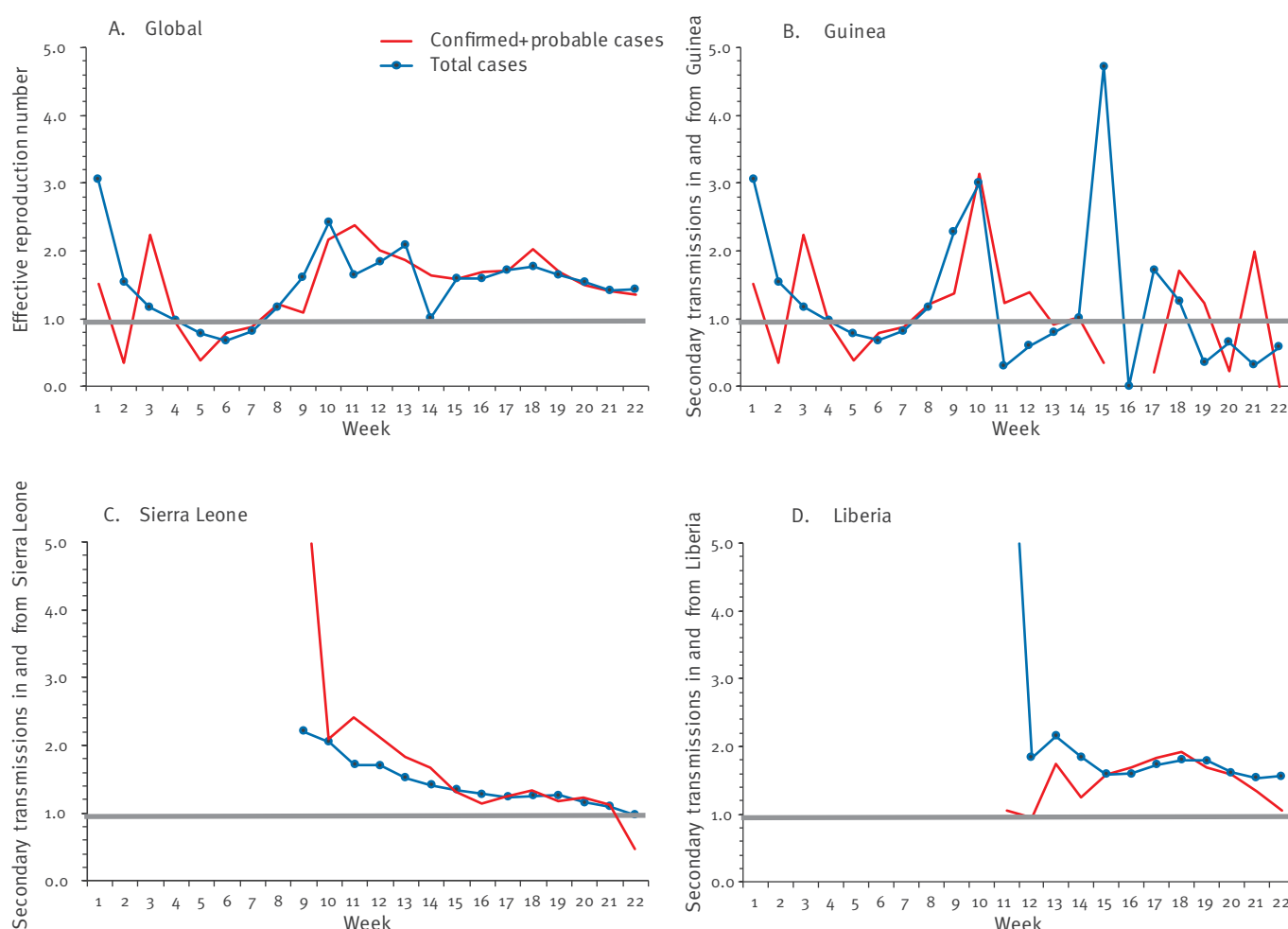
estimated reproduction number  $R$  ranging from 1.4 to 1.7 continues for the remainder of the year, we would expect to observe a total of 77,181 to 277,124 additional cases within 2014.

Maximum likelihood estimates of  $R_t$  in Guinea appeared to have fluctuated around 1.0 (Figure 2B), which reflects the observed variation in the corresponding incidence curve. Importantly,  $R_t$  in this country has not been continuously below 1.0, which supports the view that in this country the outbreak is not yet under control. Estimates of  $R_t$  in Sierra Leone and Liberia appeared to be consistently above 1.0 up to week 22 (i.e. the week starting on 18 August) (Figure 2C and 2D). Although  $R_t$  in Sierra Leone has been declining with the highest estimates obtained for early June,  $R_t$  has not been consistently below 1.0 in this country, including estimates for the latest reporting week (Figure 2). The pattern of  $R_t$  in Liberia shows values well above 1.0 since July 2014. In this country, the estimates of  $R_t$  reaching values up to 2.0 indicate that the outbreak could only be brought under control if more than half of secondary transmissions per primary case were prevented.

Figure 3A shows the estimated average number of transnational transmissions per single primary case as a function of time (calculated by 2a).  $\alpha$  has been high in early June, but has declined dramatically since late June. Nevertheless, most recent model estimates still suggest a non-negligible number of cross-border

**FIGURE 2**

Effective reproduction number of Ebola virus disease (EVD) estimated for Guinea, Sierra Leone, Liberia, and for the global system in West Africa, 23 March–26 August 2014



A) Global (maximum likelihood) estimates of the effective reproduction number of EVD based on data from all affected countries (Guinea, Sierra Leone and Liberia) were derived from the dominant eigenvalue of the next generation matrix.

B-D) The average number of secondary transmissions arising from Guinea, Sierra Leone and Liberia, was calculated from the corresponding column sum of the next generation matrix. The horizontal grey solid line indicates the reproduction number at 1.0 for reference, below which the epidemic follows a declining trend. Estimates were derived using either confirmed cases plus probable cases or the total reported case counts (confirmed, probable plus suspected cases). Data from Nigeria and Senegal have been omitted due to limited number of cases recorded thus far. Epidemic week 0 corresponds to the week that includes 22 March 2014.

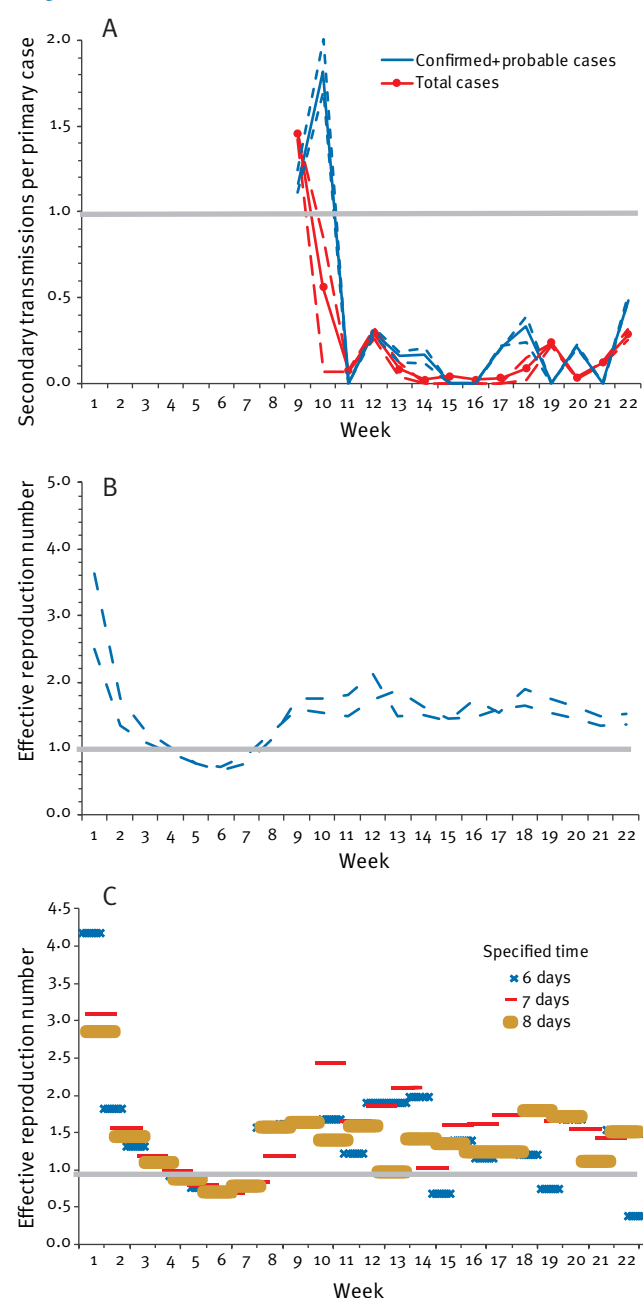
transmissions. Figure 3B examines the sensitivity of  $R_t$  for all countries to changes in the mean generation time. Although the absolute values of  $R_t$  are positively correlated with the mean generation time, the above-mentioned qualitative patterns of  $R_t$  are preserved, which indicates that the ongoing EVD epidemic has yet to be brought under control. Figure 3C examines the sensitivity of  $R_t$  to a specified time interval of the piecewise constant model. Perhaps not surprisingly, as the interval is shortened, fluctuations in  $R_t$  tend to increase, perhaps due to stochastic effects. Nevertheless, all models roughly provide qualitatively similar patterns in  $R_t$ .

## Discussion

We have derived global and country-specific estimates of the  $R_t$  of EVD for the ongoing outbreak in West Africa. Our global estimates of the  $R_t$  appear to be continuously above one since early June, indicating that the epidemic has been steadily growing and has not been brought under control as of 26 August 2014. The country-specific estimates for Sierra Leone and Liberia were also above one, perhaps reflecting the increasing trend in cases in these countries since June. Our estimated reproduction numbers, broadly ranging from one to two, are consistent with published estimates from prior outbreaks in Central Africa [9,17]. Our estimates of  $R_t < 2$  indicate that the outbreak could

**FIGURE 3**

Sensitivity analysis of the effective reproduction number of Ebola virus disease (EVD), West Africa, 23 March–26 August 2014



- A) The estimated average number of secondary cases per single primary case arising from transnational spread. Solid lines represents estimates derived from the mean generation time of 12 days, while dashed lines correspond to estimates derived using nine and 15 days as the mean generation time.
- B) Upper and lower bounds of the effective reproduction number ( $R_t$ ) for the global dynamics in West Africa are shown assuming a mean generation time of EVD ranging from nine to 15 days. The horizontal grey line is shown as a reference for the reproduction number at 1.0 below which the epidemic follows a declining trend.
- C) Sensitivity of  $R_t$  to varying specified time intervals of the piecewise constant model. Estimates in B and C were derived using the total number of reported EVD cases (confirmed, probable plus suspected cases). Epidemic week 0 corresponds to 22 March 2014. Of note, estimates overlap at week 9 as these were derived from epidemiological data for a single country (i.e. Guinea).

be brought under control if more than half of secondary transmissions per primary case are prevented.

Our statistical analysis of the reproduction number of EVD in West Africa has demonstrated that the continuous growth of cases from June to August 2014 signalled a major epidemic, which is in line with estimates of the  $R_t$  above 1.0. Moreover, the timing of  $R_t$  reaching levels above one is in line with a concomitant surge in cases in Sierra Leone and Liberia. In a worst-case hypothetical scenario, should the outbreak continue with recent trends, the case burden could gain an additional 77,181 to 277,124 cases by the end of 2014. Although such numbers must be interpreted with caution (as they rest on an assumption of continued exponential growth within 2014, which is unlikely), our study supports the notion that the ongoing EVD epidemic must be regarded as a Public Health Emergency of International Concern [3]. This finding also implies that transnational spread of EVD might have hindered control efforts, suggesting that preparedness plans for potential case introductions is critical particularly for countries at high risk of EVD case importations [18] with suboptimal public health systems. The transnational spread per person appears to have been reduced over time, but our most recent model estimates still suggest a non-negligible number of secondary cases arising from transnational spread. Uncontrolled cross-border transmission could fuel a major epidemic to take off in new geographical areas (e.g. as seen in Liberia). Unaffected countries at risk of transnational spread should be on high alert for potential EVD introductions and be ready to launch comprehensive and timely containment responses to avert outbreaks.

Our analysis is not exempted of limitations. First, the epidemic is ongoing in multiple geographical locations, and no simple mixing matrix can capture the complex geographical patterns of spread in the region. Second, cases may be under-ascertained, and hence reported cases may represent only a portion of the total number of infected individuals. However, our estimates of the reproduction number are not affected whenever the diagnosis and reporting rates have not dramatically changed over time. Third, the reporting delays are known to induce a downward bias in incidence in the latest observation, which can complicate real-time analyses. Several studies have successfully addressed this bias [19–22], but we were unable to incorporate this delay into our analyses due to a lack of empirical data to characterise the reporting delay distribution.

Despite the above-mentioned limitations, we believe that our findings are useful to demonstrate that the cases have been steadily growing in the last three months with an  $R_t$  above one. Close monitoring of this evolving epidemic should continue in order to assess the status of the outbreak in real time and guide control interventions in the region. Reviewing possible countermeasures for countries at risk of transnational



spread [18] would be of utmost importance to confront the ongoing propagation of cases over time and space.

## Acknowledgements

HN received funding support from the Japanese Society for the Promotion of Science (JSPS) KAKENHI Grant Number 26700028, St Luke's Life Science Institute Research Grant for Clinical Epidemiology Research 2014 and the Commissioned Research program of the Ministry of Health, Labour and Welfare, Japan (H26-ShinkoJitsuyoka- General-016). GC acknowledges financial support from the NSF grant 1414374 as part of the joint NSF-NIH-USA Ecology and Evolution of Infectious Diseases program, UK Biotechnology and Biological Sciences Research Council grant BB/M008894/1, the Fogarty International Center, US National Institutes of Health and the Visiting Foreign Scholar program of the Heiwa Nakajima Foundation for supporting a sabbatical research visit at the University of Tokyo.

## Conflict of interest

None declared.

## Author contributions

HN conceived mathematical modeling method and analyzed the data. HN and GC drafted and revised the manuscript.

## References

- Baize S, Pannetier D, Oestereich L, Rieger T, Koivogui L, Magassouba N, et al. Emergence of Zaire Ebola Virus Disease in Guinea - Preliminary Report. *N Engl J Med*. 2014. Apr 16. Forthcoming.
- World Health Organization (WHO). Ebola virus disease, West Africa – update on 27 July 2014. Geneva: WHO; 27 Jul 2014. [Accessed 23 Aug 2014]. Available from: [http://www.who.int/csr/don/2014\\_07\\_27 Ebola/en/](http://www.who.int/csr/don/2014_07_27 Ebola/en/)
- World Health Organization (WHO). WHO Statement on the Meeting of the International Health Regulations Emergency Committee Regarding the 2014 Ebola Outbreak in West Africa. Geneva: WHO; 8 Aug 2014. [Accessed 10 Sep 2014]. Available from: <http://www.who.int/mediacentre/news/statements/2014/ebola-20140808/en/>
- Gire SK, Goba A, Andersen KG, Sealfon RSG, Park DJ, Kanneh L, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*. 2014. Forthcoming. <http://dx.doi.org/10.1126/science.1259657>
- World Health Organization (WHO). Ebola virus disease outbreak – West Africa. Geneva: WHO; 4 Sep 2014. [Accessed 10 Sep 2014]. Available from: [http://www.who.int/csr/don/2014\\_09\\_04 Ebola/en/](http://www.who.int/csr/don/2014_09_04 Ebola/en/)
- Heymann DL, Weisfeld JS, Webb PA, Johnson KM, Cairns T, Berquist H. Ebola hemorrhagic fever: Tondala, Zaire, 1977-1978. *J Infect Dis*. 1980;142(3):372-6. <http://dx.doi.org/10.1093/infdis/142.3.372>
- Okware SI, Omaswa FG, Zaramba S, Opio A, Lutwama JJ, Kamugisha J, et al. An outbreak of Ebola in Uganda. 2002;7(12):1068-75.
- Nishiura H, Chowell G, Heesterbeek H, Wallinga J. The ideal reporting interval for an epidemic to objectively interpret the epidemiological time course. *J R Soc Interface*. 2010;7(43):297-307. <http://dx.doi.org/10.1098/rsif.2009.0153>
- Chowell G, Hengartner NW, Castillo-Chavez C, Fenimore PW, Hyman JM. The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. *J Theor Biol*. 2004;229(1):119-26. <http://dx.doi.org/10.1016/j.jtbi.2004.03.006>
- World Health Organization (WHO). Ebola virus disease update – west Africa. Geneva: WHO; 28 Aug 2014. [Accessed 10 Sep 2014]. Available from: [http://www.who.int/csr/don/2014\\_08\\_28 Ebola/en/](http://www.who.int/csr/don/2014_08_28 Ebola/en/)
- World Health Organization (WHO). Case definition recommendations for Ebola or Marburg Virus Diseases. Geneva: WHO; 9 Aug 2014. [Accessed 10 Sep 2014]. Available from: <http://www.who.int/csr/resources/publications/ebola/ebola-case-definition-contact-en.pdf>
- White LF, Pagano M. A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. *Stat Med*. 2008;27(16):2999-3016. <http://dx.doi.org/10.1002/sim.3136>
- White LF, Pagano M. Transmissibility of the influenza virus in the 1918 pandemic. *PLoS One* 2008;3(1):e1498. <http://dx.doi.org/10.1371/journal.pone.0001498>
- Glass K, Mercer GN, Nishiura H, McBryde ES, Becker NG. Estimating reproduction numbers for adults and children from case data. *J R Soc Interface*. 2011;8(62):1248-59. <http://dx.doi.org/10.1098/rsif.2010.0679>
- Francesconi P, Yoti Z, Declish S, Onok PA, Fabiani M, Olango J, et al. Ebola hemorrhagic fever transmission and risk factors of contacts, Uganda. *Emerg Infect Dis*. 2003;9(11):1430-7. <http://dx.doi.org/10.3201/eid0911.030339>
- Eichner M, Dowell SF, Firese N. Incubation period of ebola hemorrhagic virus subtype zaire. *Osong Public Health Res Perspect*. 2011;2(1):3-7. <http://dx.doi.org/10.1016/j.phrp.2011.04.001>
- Legrand J, Grais RF, Boelle PY, Valleron AJ, Flahault A. Understanding the dynamics of Ebola epidemics. *Epidemiol Infect*. 2007;135(4):610-21. <http://dx.doi.org/10.1017/S0950268806007217>
- Gomes MFC, Piontti AP, Rossi L, Chao D, Longini I, Halloran ME, et al. Assessing the International Spreading Risk Associated with the 2014 West African Ebola Outbreak. *PLoS Currents Outbreaks*. 2014; Sep 2. Edition 1. [Accessed 10 Sep 2014]. Available from: <http://currents.plos.org/outbreaks/article/assessing-the-international-spreading-risk-associated-with-the-2014-west-african-ebola-outbreak/>
- Donker T, van Boven M, van Ballegooijen WM, Van't Klooster TM, Wielders CC, Wallinga J. Nowcasting pandemic influenza A/H1N1 2009 hospitalizations in the Netherlands. *Eur J Epidemiol*. 2011;26(3):195-201. <http://dx.doi.org/10.1007/s10654-011-9566-5>
- White LF, Wallinga J, Finelli L, Reed C, Riley S, Lipsitch M, et al. Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the USA. *Influenza Other Respir Viruses*. 2009; 3(6):267-76. <http://dx.doi.org/10.1111/j.1750-2659.2009.00106.x>
- Cauchemez S, Boëlle PY, Donnelly CA, Ferguson NM, Thomas G, Leung GM, et al. Real-time estimates in early detection of SARS. *Emerg Infect Dis*. 2006; 12(1):110-3. <http://dx.doi.org/10.3201/eid1201.050593>
- Cowling BJ, Lau MS, Ho LM, Chuang SK, Tsang T, Liu SH, et al. The effective reproduction number of pandemic influenza: prospective estimation. *Epidemiology*. 2010;21(6):842-6. <http://dx.doi.org/10.1097/EDE.0b013e3181f20977>