Epidemiology of the outbreak of Ebola Virus, Democratic Republic of the Congo, April to May 2018

Authors: Ebola Response Team*

Corresponding author: Dr Oly Ilunga Kalenga (secretariat.dep@minisanterdc.cd) , Dr. Peter Salama (salamap@who.int)

* Ministry of Health, Democratic Republic of the Congo; World Health Organization; and Imperial College London, the WHO Collaborating Center for Infectious Disease Modelling.

Acknowledgements: A large number of organizations are involved in the outbreak response, including Médecins Sans Frontières, the International Federation of the Red Cross and Red Crescent Societies, Centers for Disease Control and Prevention, UNICEF, the World Food Programme, UNOCHA, and MONUSCO. Acknowledgements are listed at the end.

Background: On 8 May 2018, the Government of the Democratic Republic of the Congo (DRC) reported an outbreak of Ebola Virus Disease (EVD) in Equateur Province in the northwest of the country. The remoteness of most affected communities and the involvement of an urban centre connected to the capital city and neighbouring countries makes this outbreak the most complex and high risk ever experienced by the DRC.

Methods: Epidemiological investigations of cases were conducted to obtain demographic characteristics, determine possible exposures, collect information about signs and symptoms, and identify contacts to be followed up for 21 days. Cases were classified as suspected, probable or confirmed case using the national EVD case definitions. The reproduction number and projected number of cases for the four week period 25 May to 21 June were estimated.

Results: Update as of 30 May, 50 cases (37 confirmed, 13 probable) of *Zaire ebolavirus*, were reported across Bikoro (42% of cases), Iboko (50% of cases) and Wangata (8% of cases) health zones. Wangata is part of Mbandaka, the urban capital of Equateur Province connected to major national and international transport routes. By 30 May, 25 deaths had been reported, giving a case fatality ratio (CFR) of 56% (95% CI: 39% - 72%) after adjustment for censoring. This CFR is consistent (p=0.427) with estimates for the 2013-15 West African epidemic. The median age of cases was 40 years (range: 8-80 years) and 30 (60%) were male. The most common reported signs and symptoms included fever (95%), fatigue (90%) and loss of appetite (90%). Gastrointestinal symptoms were common and 32% cases reported haemorrhagic signs. Time from illness onset or hospitalisation to specimen testing decreased over time. On 30 May, 734 contacts had been identified, of which 69% had been followed up. The estimated reproduction number is 1.03 (95%CI 0.83–1.37) and the cumulative case incidence

for the outbreak by 21 June is projected to be 78 cases (95% CI: 37 to 281). The initial source of the outbreak is still under investigation.

Conclusions:

The current Ebola virus outbreak has similar epidemiological features to previous Ebola outbreaks. Rapid case isolation, contact tracing and the ongoing vaccination programme is expected to stop the outbreak. The forecast of the number of cases does not exceed the current capacity to respond, if the epidemiological situation does not change.

Introduction

On 3 May 2018, the Ministry of Health of the Democratic Republic of the Congo (DRC) received a notification from the Health Division of Equateur Province of 21 cases of fever with haemorrhagic signs, including 17 community deaths, from the Ikoko Impenge Health Area, Bikoro Health Zone, which is approximately 125 km south of the provincial capital of Mbandaka. An investigation team, composed of members of the Ministry of Health, Médecins Sans Frontières (MSF) and the World Health Organization (WHO), travelled to Bikoro Health Zone from 5 to 6 May 2018. Blood samples were collected from five hospitalised cases and transported to the National Institute of Biological Research (INRB) in Kinshasa for laboratory testing on 6 May 2018. Of these, two were positive for *Zaire* e*bolavirus* by reverse transcription polymerase chain reaction (RT-PCR). In line with the International Health Regulation (IHR) requirements, the Ministry of Health notified WHO of the confirmed cases and declared the outbreak on 8 May 2018. Further investigation found cases in neighbouring Wangata and Iboko health zones.

Ebola virus is a filovirus with five sub-species (Zaire, Bundibugyo, Sudan, Reston and Taï Forest). It causes Ebola virus disease (EVD) which has a case fatality ratio (CFR) of between 25% and 90%[1]. The Zaire strain is the most fatal with an overall CFR ranging from 69% to 88%[2]. EVD is transmitted primarily through contact with the body fluids of symptomatic patients, most commonly to adults of 17-44 years, with relative sparing of children under the age of 16 years[3,4,5,6]. Transmission can be stopped by early diagnosis, patient isolation and care, infection control, safe and dignified burial of the remains of cases, rigorous tracing of contacts and more recently, targeted vaccination.

DRC has recorded eight previous EVD outbreaks since 1976 (Figure 1)[7,8]. The last outbreak occurred in May 2017 in a remote area in the north-east of the country, Likati Health Zone in the Bas-Uele Province, causing a total of eight cases with four deaths. Most of the previous outbreaks have been confined to remote rural areas with the exception an outbreak in Kikwit, a town with a population of just under 400,000 that resulted in 315 cases and 250 deaths[9]. The response to this outbreak includes use of traditional measures such as early identification, isolation and care of cases, contact tracing, safe and dignified burials, culturally appropriate community mobilisation. These traditional measures are being supplemented by use of the recombinant vesicular stomatitis virus–Zaire Ebola virus (rVSV-ZEBOV) vaccine, with vaccination of first and second line contacts. This paper is the first in a series on the latest outbreak in DRC. It provides an early overview of the descriptive epidemiology using best data available from field teams working to response to the epidemic.

Methods

Case investigation: Cases were classified as suspected, probable or confirmed according to the EVD case definitions of the Ministry of Health (Table 1 below)[10]. Confirmation of cases required detection of Ebola RNA in blood or body fluids by RT-PCR. Information on all cases was recorded using the Ministry of Health case investigation form and entered into an electronic database. Case investigations were conducted to record demographic characteristics, determine possible exposures, document information on illness onset and signs and symptoms, and to identify potentially exposed contacts. For cases who had recovered or died before 5 May 2018, retrospective case classification was through review of medical records at health facilities in the affected locations. For cases alive or newly ill since declaration of the outbreak, information was collected prospectively at the time of case investigation. Our analysis included probable and confirmed cases as of 30 May 2018.

Contact tracing: Contacts were identified during the case investigation process for each case. Contact tracers are required to visit all contacts once a day (in Mbandaka city, contacts are visited twice daily) for 21 days following the last date of contact with an infectious suspected, probable or confirmed case. Information on their health status and the development of any EVD like symptoms is collected[11].

Data analysis: Data analyses were performed using R (version 4.3). Missing/unknown data were excluded. Confidence intervals were calculated assuming symptom occurrence was binomially distributed. Spatial locations of cases were analysed in ArcGIS (ESRI, version 10.5) using area boundaries developed by a range of partners, including WHO, in consultation with the Ministry of Health. Cases were plotted to village, health area and overall health zone in Bikoro, Iboko and Wangata, respectively. Boundaries are subject to confirmation.

Case fatality ratio: The observed deaths by 30 May were used to obtain a naïve CFR estimate, which was then adjusted by the proportion of deaths among the cases in the database that would have been expected by 30 May 2018, based on their dates of illness onset and the illness-onset-to-death delay distribution estimated using the data from the West African Ebola epidemic[5]. In addition, the age-dependent CFR[12] and illness-onset-to-death distributions[5] from the West African Epidemic were used to predict the numbers of deaths expected among the cases in the current outbreak by 30 May, based solely on the ages of cases and dates of illness onset. This predicted number of deaths was compared with the observed number of deaths by 30 May by calculating the two-sided p-value, 2×Poisson($X \leq x \mid \lambda$) where x is the observed number of deaths by 30 May and λ is the predicted number of deaths by 30 May.

Time from Illness onset to first hospitalisation and sample testing were calculated for cases with available data and with dates of onset after 30 April, in line with the “trusted period” defined with the reproduction number estimates. A simple linear regression was fitted against dates of case illness onset or hospitalisation to assess trends over time.

Reproduction number estimates: Due to the delay between illness onset and notification, the most recent cases are likely not yet reported. Based on the confirmed cases only, we defined a "trusted period" where we estimated that the recorded incidence of confirmed cases with dates of illness onset between 30 April and 24 May (inclusive) were at least 95% complete (potentially relative to an unknown but constant level of overall under-reporting). The analyses of the reproduction number (*R*) and onward projections were therefore based only on the confirmed cases with illness onset during this period (Annex 2). The analysis used an approach similar to those previously described[12,13] , using a Poisson process or renewal equation to approximate the daily incidence and assumed: i) a serial interval distribution inferred for the West African Ebola Epidemic[6] (Annex 1: Figure A3); and ii) constant transmissibility throughout the trusted period.

Forward Projections: Using the estimates of the reproduction number obtained above, we projected incidence for the 4-week period, 25 May to 21 June, following the end of the trusted period (Annex 1). Those forward projections assumed that transmissibility and reporting rates remained the same as during the trusted period. Two transmission assumptions were explored: i) homogeneous transmission among cases (no super-spreading) approximated using a Poisson process and ii) heterogeneous transmission among cases (with super-spreading) using a negative binomial distribution which incorporates additional variability in the number of secondary cases. This level of heterogeneity was assumed to be similar to that seen during the West African Ebola epidemic[6].

95% Credible Intervals: The reproduction number estimates and the forward projections were estimated in a Bayesian framework using an MCMC (Monte Carlo Markov Chain) approach. Therefore, the uncertainty is reported here as 95% credible intervals (95% CI), obtained by taking the 2.5% and 97.5% quantiles of the posterior distribution.

Results

As of 30 May 2018, a total of 50 EVD cases (37 confirmed, 13 probable), including 25 deaths (unadjusted CFR of 50% [95% CI: 36% - 64%] by 30 May), have been identified in Equateur Province, with illness onsets of the cases between 5 April and 28 May (Figure 2). After adjustment for censoring, the CFR is 56% (95% CI: 39% - 72%), ignoring uncertainty in the illness-onset-to-death distribution. Based on the recorded ages of cases in the current outbreak, their recorded dates of illness onset, and epidemiological parameters estimated from the West African Ebola epidemic[9], we would have expected 29.9 deaths by 30 May 2018, with a total of 33.0 deaths eventually expected among these 50 cases. Thus, the fatalities seen in the current outbreak by 30 May are, after adjustment for censoring, consistent with CFR estimates seen in the West African Ebola epidemic (p-value = 0.427)[9]. The median age of cases was 40 years (range: 8–80 years) and 30 (60%) were male (Figure 3). Cases were reported in northern areas of Iboko (n=25; 23 confirmed, two probable), southern areas of Bikoro (n=21; 10 confirmed, 11 probable) and Wangata (n=4; all confirmed) health zones (Figure 4).

Of 50 confirmed and probable cases, 45 had at least one reported symptom. The most frequently reported symptoms were: fever (n=40/42), loss of appetite (n=37/ 41) and intense general fatigue (n=37/ 41), followed by diarrhoea (n=23/32), abdominal pain (n=22/ 35) and nausea/vomiting (n=22/35) (Figure 5). Haemorrhagic signs were observed in 14 of 43 cases. The symptom profile of confirmed and probable cases was statistically similar. The overall median time from illness onset to first hospitalisation was 1 day (range: 0–10 days) with no evidence of a reduction over time (, p=0.54) (Figure 6). However, marked reductions in the time from illness onset to specimen sampling (data not shown) and illness onset to sample testing were apparent (p<0.0001, overall median 6 days, range 1–13 days). Similarly, time from first hospitalization to sample testing improved over time (p=0.0004, overall median 11 days, range 0–13 days).

Five health care workers, two of whom died, were among the cases. Other commonly affected occupational groups included farmers (n=14), students (n=5), household workers (n=5) and religious leaders (n=4). The most common exposure risks were having contact with another sick person (29 of 41 cases) and participation in a funeral (24 of 40 cases) (Table 2).

As of 30 May, 1458 epidemiological contacts had been identified of which 746 remained under active follow-up. Of the 504 first and second line contacts eligible for vaccination, 496 had been vaccinated by teams of trained vaccinators.

The estimated reproduction number in the period 30 April to 24 May was 1.03 (95% confidence interval (CI): 0.83 to 1.37), an estimate that was robust to assumptions about the serial interval distribution and the trusted period. The projected cumulative number of confirmed cases on 21 June 2018 is on average 76 (95% CI: 54 - 109) assuming a homogeneous transmissibility (Poisson) model, and 78 (95% CI: 37 to 281) assuming a

heterogeneous transmissibility (negative binomial) model. The resulting projected incidence patterns are shown in Figure 7.

Discussion:

Our analysis shows that the epidemiological features of the current outbreak in DRC such as demographic characteristics and signs and symptoms of cases are consistent with previous outbreaks of EVD[14,15,16]. Contact with other cases and participation in a funeral are the most commonly reported exposures among cases, similar to previous EVD outbreaks, reinforcing the importance of community engagement and implementation of safe and dignified burials for outbreak control. The CFR is similar to that seen in previous outbreaks in DRC and elsewhere, but higher than was seen towards the end of the 2014-2016 West Africa outbreak, where there was greater access to Ebola Treatment Units (ETUs)[13,17]. With the rapid installation of ETUs in the affected areas, the CFR is expected to decrease[18,19,20] The reduction in the time from illness onset to isolation and testing is encouraging because prompt isolation and testing minimizes exposure and transmission of Ebola virus to other people. It is concerning that five of 50 cases are health care workers, again highlighting the risk for clinical staff and the importance of providing sufficient training and equipment for health care workers to protect themselves. Moreover, that nearly half of the cases reported hospitalization or contact with a hospitalized patient prior to their Ebola infection is a clear reminder that health care facilities with inadequate infection control procedures can amplify Ebola outbreaks [5,21,22,23,24].

The EVD outbreak in DRC currently remains geographically limited to three health zones in Equateur Province. Two of the affected communities are in remote areas, which whilst reducing the risk of widespread expansion of the outbreak, creates serious logistical barriers for a rapid response including the follow up of contacts each day. The response teams have had to overcome major infrastructure challenges in multiple sites across a wide geographic area, such as the lack of electricity for essential laboratory and clinical equipment, absence of communications networks for transmitting data, very limited road access for contact tracers to travel on, and absence of accommodation for responders. The complexity of the context also makes it extremely difficult to collate and analyse epidemiological and response data for analysis and operational planning. In addition to these challenges, the spread of transmission to the provincial capital, Mbandaka, an urban area of nearly one million people, raises concerns about an urban Ebola outbreak. Even more concerning is that Mbandaka is a port city on the Congo River and is a major transportation hub – to the capital Kinshasa with nearly 10 million inhabitants, and also to neighbouring countries such as the Republic of the Congo and the Central African Republic. The proximity of this outbreak to major national and international transportation routes underpins WHO`s assessment that the public health risk from this outbreak is very high for DRC and high for other neighbouring countries. The risk internationally remains low [25].

At present the source of the outbreak is unknown. Investigations are ongoing, but one hypothesis is that this outbreak is linked to a cluster reported in February 2018 of 15 persons who had a febrile illness that occurred in Ingende and Bikoro health zones of Equateur Province. Of those 15 cases, 11 had haemorrhagic signs, of whom eight died. According to the investigation report, the first case died on 20 December 2017. The aetiology of that

cluster has not been confirmed. While a link between the two clusters cannot be ruled out, the long period of time between these two events without identified chains of transmission calls into question whether they were causally linked. However, there are epidemiological links between the ongoing clusters in the different locations, which underscores the potential for geographic spread, even in remote areas. Ongoing field investigations are being conducted to describe the chains of transmission that link the identified cases and information on transmission chains will be published online as it becomes available. In addition, further information on contact tracing and the proportion of cases emerging from contact lists will also be made available.

Statistical forward projections suggest that if interventions remain as effective as they were between 30 April and 24 May, possibly twice as many cases may occur by 21 June. Even under this pessimistic scenario, the current isolation capacity available in the affected communities would be sufficient. Nonetheless, considering that a period of 42 days after the last cases is required before the outbreak can be considered over, the ongoing occurrence of cases would mean that the response will need to continue for at least the next three months or more. Furthermore, is not possible to rule out further expansion of the outbreak if there is exportation of cases to new areas or if there are ongoing but hitherto unrecognized chains of transmission. It is also possible that a new chain of transmission may occur following sexual transmission of the virus from a male survivor, if appropriate services and counselling are not provided, again requiring an even longer response.

As for all outbreak investigations, some data are collected retrospectively and some data are incomplete. Data on signs and symptoms for some patients were collected retrospectively from medical records, which may have resulted in errors or missing data. An analysis of a subset of patients with prospectively collected data results in a similar frequency of signs and symptoms. Detailed information about chains of transmission is being compiled by field investigation teams and are not available currently. The dynamic nature of outbreaks and response means that some numbers are revised as additional information becomes available.

A major sustained response is therefore needed to ensure ongoing case identification, contact tracing, isolation, and other control measures. Implementation of WHO`s Early Warning Alert and Response System, (EWARS), a data collection system that uses handheld devices, represents a major improvement for data collection compared with the 2014-2016 West Africa outbreak. However, this information system is not optimally designed for contact tracing. Collecting, managing, and analysing epidemiological data in real-time continues to be a significant challenge in the field. Nonetheless, the analysis presented in this paper shows that real-time data collection and epidemiological analysis for the control of complex Ebola outbreaks is achievable.

The epidemiology of the current Ebola virus outbreak in DRC has similar features to previous Ebola outbreaks, which indicates that early detection of the outbreak combined with tried-and-tested interventions including early isolation and treatment, contact tracing, safe burials and community engagement currently being implemented, along with the additional benefit of targeted vaccination, should be sufficient to control this

239 outbreak. However the combination of remote communities and spread to an urban centre that is connected to
240 the capital city and neighbouring countries, makes this outbreak the most complex and high risk ever
241 experienced by the DRC.

242

Tables and figures:

Figure 1: Previous outbreaks of Ebola virus disease in the Democratic Republic of the Congo, 1976–2018. Boundaries are subject to confirmation and locations are approximate. The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

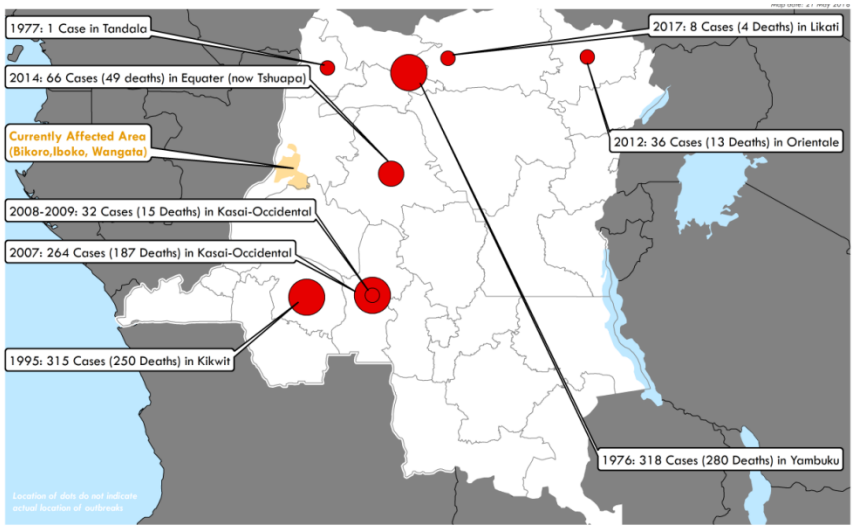


Table 1: Ebola Virus Disease Case and contact definitions.

Suspected case	Any living person having or having had a high fever with a sudden onset, with an epidemiological link to: • a suspected, probable or confirmed case of Ebola • a dead or sick animal OR Any deceased person having or having had a high fever with a sudden onset, and who has been in contact with: • a suspected or probable case of Ebola • a dead or sick animal OR Anyone with a high fever with a sudden onset and at least three of the following symptoms: headache, severe fatigue, anorexia / loss of appetite, difficulty swallowing, abdominal pain, difficulty breathing, vomiting, hiccups, diarrhoea; muscle or joint pain OR Anyone with unexplained bleeding; OR Anyone with sudden and unexplained death
Probable case	Any suspected case evaluated by a clinician; OR Any suspect case that has died (and for which it has not been possible to obtain biological samples for laboratory confirmation) with an epidemiological link to a confirmed case
Confirmed case	Any suspected or probable case with a positive laboratory result for viral RNA by reverse transcription polymerase chain reaction (RT-PCR), or for retrospective diagnosis, antibodies against Ebola.
Contacts	Any person having had contact with a confirmed, probable or suspected EVD case by: • sleeping in the same house as the case in the month before illness onset • Having direct physical contact during the cases illness or with the body of a deceased case • Having shared the same transport vehicle as a case during their illness • Having touched any bodily fluids of a case during their illness • Having handled any clothes or linen of a case during their illness • Having been breastfed by a case.

Figure 2: Confirmed and probable EVD cases by date of illness onset and classification, Democratic Republic of the Congo, data as of 30 May 2018 (n=50)

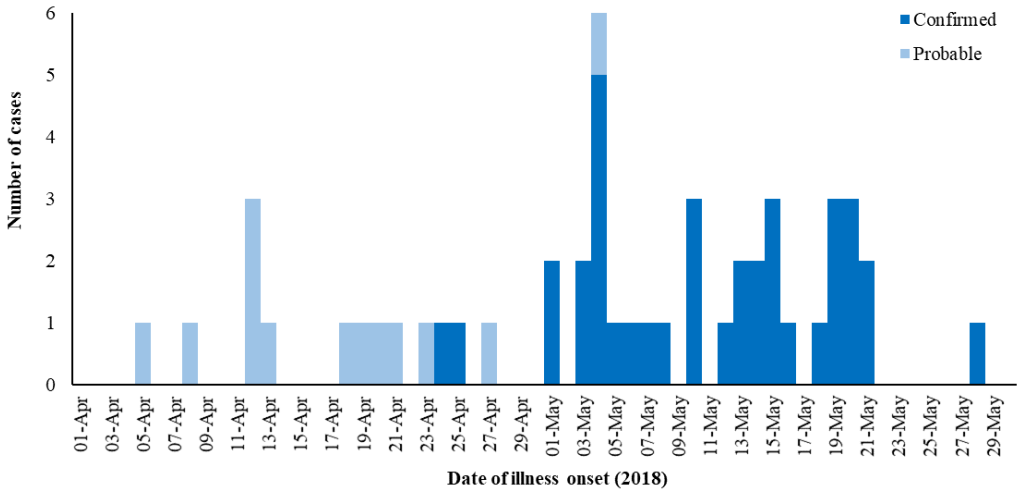


Figure 3: Confirmed and probable EVD cases by age and sex, Democratic Republic of the Congo, data as of 30 May 2018 (n=49). Age was unknown for n=1 female case.

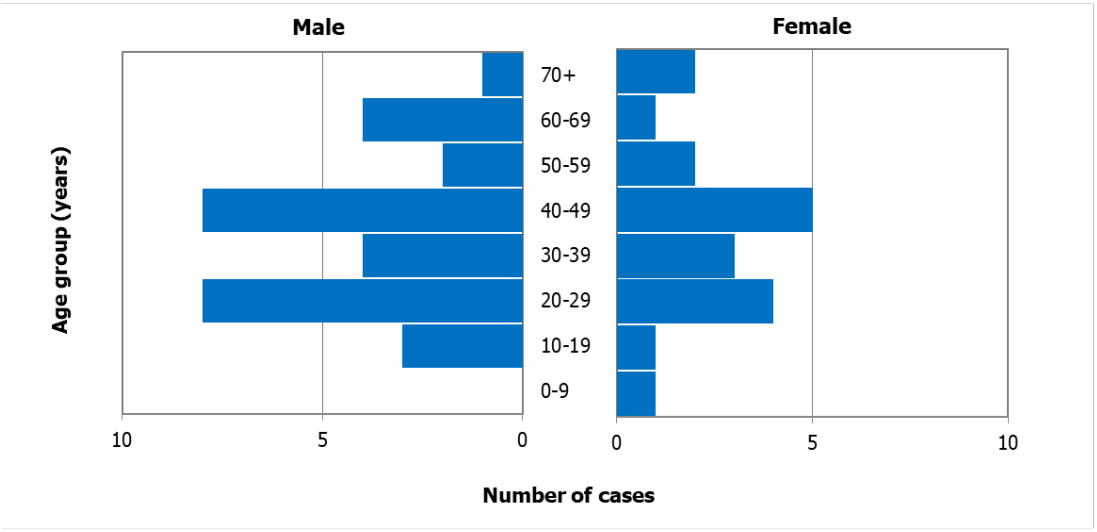
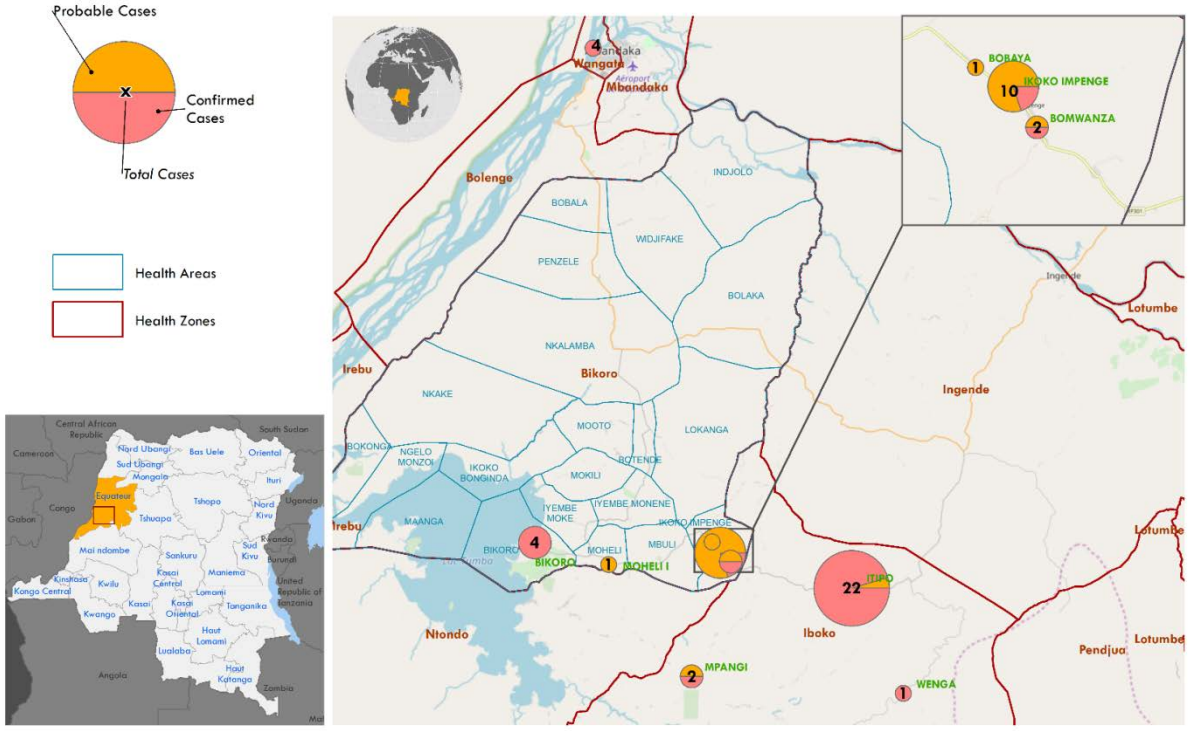


Figure 4: Confirmed and probable Ebola virus disease cases by approximate place of residence, Democratic Republic of the Congo, reported as of the 30 May 2018. Cases only displayed where location can be determined at the scale of this map. Other cases not indicated. Boundaries are subject to confirmation and locations are approximate. The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.



269

Figure 5: Frequency distribution of the most common symptoms reported for confirmed and probable Ebola virus disease cases, Democratic Republic of Congo, data as of 30 May 2018. Bars denote binomial 95% confidence interval. Additional symptoms reported in less than 25% of cases not shown.

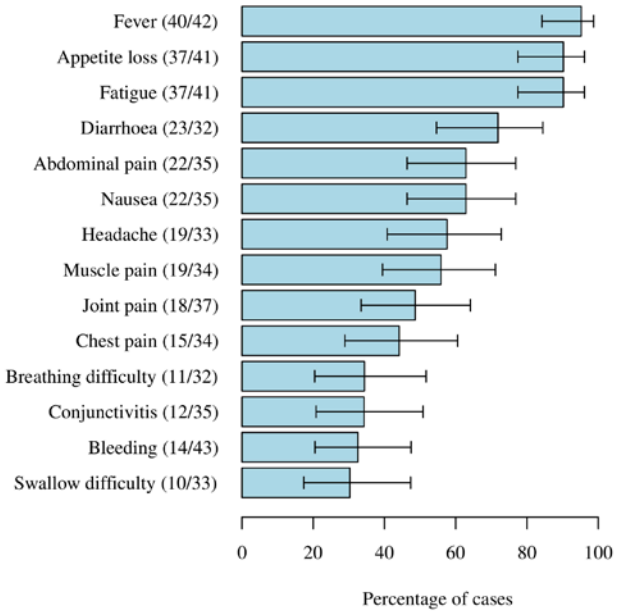
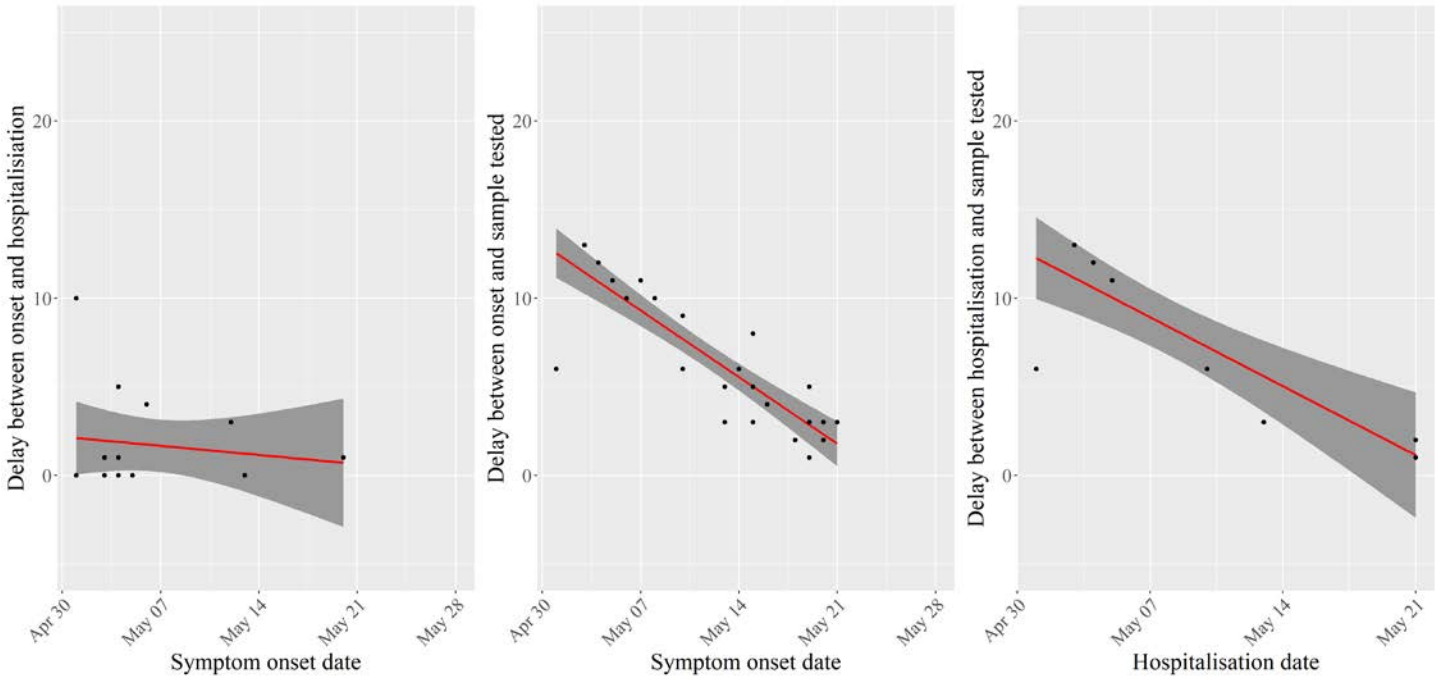


Figure 6: Simple linear regressions showing the delay from illness onset to first reported hospitalization (n=16) and sample testing (n=30), and hospitalization to sample testing (n=12), confirmed and probable Ebola virus disease cases with date of onset after 30 April, Democratic Republic of the Congo, data as of 30 May 2018. Points: case observation; line: linear model mean predicted value; shaded area: 95% confidence interval around the mean.


Delay between onset and hospitalisation
Delay between onset and sample tested
Delay between hospitalisation and sample tested
20
10
0
Apr 30
May 07
May 14
May 21
May 28
Symptom onset date
Symptom onset date
Hospitalisation date

279
280

Table 2: Exposures prior to onset of illness reported for confirmed and probable Ebola virus disease cases, Democratic Republic of the Congo, data as of 30 May 2018.

	Proportion of cases reporting exposure:		
Exposure	**Confirmed**	**Probable**	**Total**
Contact with other cases/sick persons in month before illness	22/31 (71%)	7/10 (70%)	29/41 (71%)
Funeral participant	18/31 (58%)	6/9 (67%)	24/40 (60%)
Travel outside of home village/town	11/27 (41%)	1/8 (13%)	12/35 (34%)
Prior hospitalization	11/28 (39%)	2/7 (29%)	13/35 (37%)
Visited traditional healer	2/26 (8%)	1/7 (14%)	3/33 (9%)
Direct contact with animals/raw meat	1/21 5(%)	0/5 (0%)	1/26 (4%)

*Missing and inconclusive responses excluded.

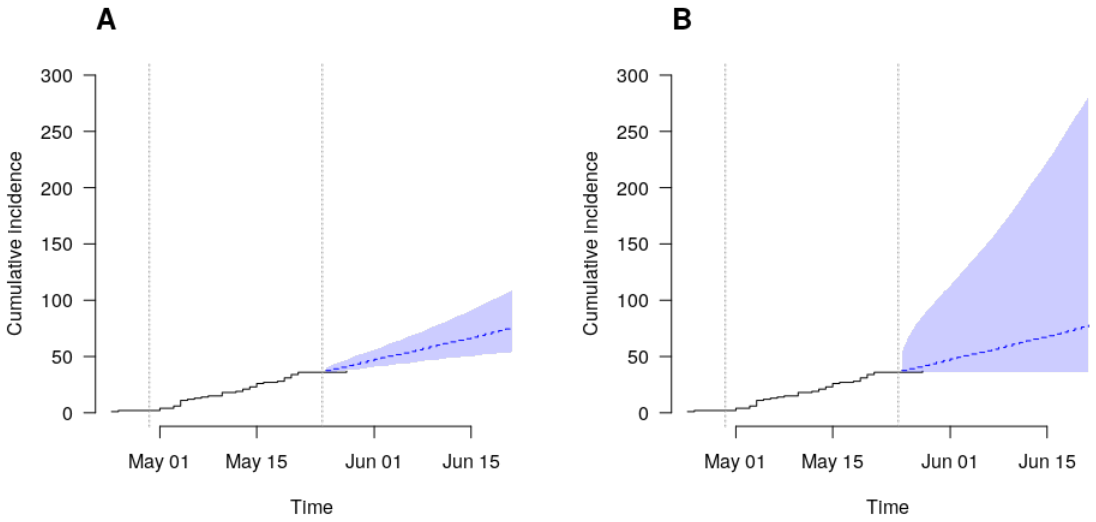


Figure 7: Observed and projected cumulative incidence of illness onset, over time, using a (A) homogeneous transmissibility (Poisson) model and a (B) heterogeneous transmissibility (negative binomial) model. The black solid lines show the observed cumulative incidence of confirmed cases over time. The blue dashed line shows the mean projected cumulative incidence and the shaded area the 2.5% and 97.5% quantiles of the projected cumulative incidence. The vertical dotted lines delineate the trusted period.

Annex 1: Supplemental information about reproduction number estimates, case projections

Case Fatality Ratio (CFR)

The line list dataset received 1 June 2018 includes three variables relevant to the case fatality ratio (CFR): i) status at time information was collected; status as time of notification; and; final status. For status at time of collection, there were 29 "Alive" and 21 "Dead"; for status at time of notification, there were 23 "Alive", 25 "Dead" and 2 "NA"; and for final status, there were 6 "Alive", 13 "Dead" and 31 "NA". Of the 18 possible combinations of these levels, there were 6 combinations observed (as presented in Table A1).

Table A1. The number of confirmed and probable cases by status at time information was collected, status at time of notification, and final status.

Status at time information was collected	Status at time of notification	Final status	Number	Status used in current analyses
Alive	Alive	Alive	6	Alive
Alive	Alive	NA	17	Alive
Alive	Dead	Dead	4	Dead
Alive	NA	NA	2	Alive
Dead	Dead	Dead	9	Dead
Dead	Dead	NA	12	Dead

On the basis of these variables, we conclude that there were 25 deaths and 25 people alive up to 30 May 2018 (the most recent date variable recorded in the variables: date of illness onset, date of hospitalisation, date of notification, and date of death).

We calculated the expected individual-level probability of having observed death by 30 May 2018, among those that would eventually die over the course of their illness based on the estimated gamma distribution fitted to the onset-to-death observations among confirmed and probable cases in the West African Ebola epidemic (shape = 1.651 and rate = 0.202 giving a mean and standard deviation of 8.17 days and 6.36 days, respectively.[5] The average probability was 0.900. Thus, we take the observed CFR by 30 May 2018 and its exact 95% binomial confidence interval: 50% (95% CI: 36% - 64%) and obtain the adjusted CFR by dividing each of these numbers by 0.900 to obtain an estimate of CFR adjusted for censoring of 56% (95% CI: 39% - 72%), ignoring uncertainty in the illness-onset-to-death distribution.

We estimated individual-level CFRs based on age (which was recorded in years for 49 of the 50 cases) using the equation:

$$CFR(age) = \frac{\exp(-0.0350 - 0.0820\,\text{age.child} + 0.0288\,\text{age.adult})}{1 + \exp(-0.0350 - 0.0820\,\text{age.child} + 0.0288\,\text{age.adult})}$$

where *age.child* = min(age – 15, 0) and *age.adult* = max(age – 15, 0)

where the parametric form and parameter estimates were estimated from data on confirmed and probable cases during the 2014-16 West African epidemic[12] (Figure A1). Note that the average individual-level CFR (66.1%) observed in the database was assumed for the single case without recorded age. The mean of these individual CFRs did not vary substantially between those that were recorded as having died (67.7%) and those still alive (64.5%).

The expected individual-level probability of death by 30 May 2018 is the product of the estimated individual-level CFR and the estimated individual-level probability of having observed death by 30 May 2018, among those that would eventually die over the course of their illness. The mean of these individual probabilities of having observed death by 30 May 2018 varied substantially between those that died (65.9%) and those still alive (53.7%). Summing these probabilities over all 50 cases in the case database, we find that we would have expected 29.9 deaths by 30 May 2018 out of a total of 33.0 deaths expected among these 50 confirmed and probable cases over the course of their illness. Thus, the fatality data observed from the current outbreak are consistent with what are predicted based on the West African Ebola epidemic (p-value = 0.427).

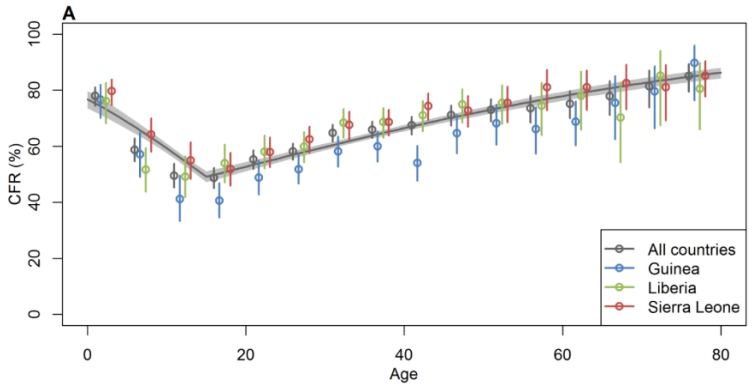


Figure A1. The estimated mean case fatality ratio (CFR %) as a function of age (in years) as estimated for confirmed and probable cases in ref 24[12]. The line shows the mean and the shaded area the 95% prediction interval. Data are shown with 95% confidence interval by age group, by country and overall.

Key Delays

We investigated the delays between the dates of illness onset and death and notification, respectively, and fitted gamma distributions to the observed delays using maximum likelihood. One case had a negative onset to notification delay recorded and was removed from the analysis. The summary statistics of the observed delays are shown alongside the equivalent estimates from the fitted gamma distributions and the distributions' parameters with 95% confidence intervals. The observed and fitted values match very well for both fitted distributions.

Table A2. Summary statistics of mean delays and parameters of the fitted gamma distributions.

	observed			fitted			
Onset to	**n**	**mean**	**sd**	**mean (95% CI)**	**sd (95% CI)**	**shape (95%**	**rate (95% CI)**

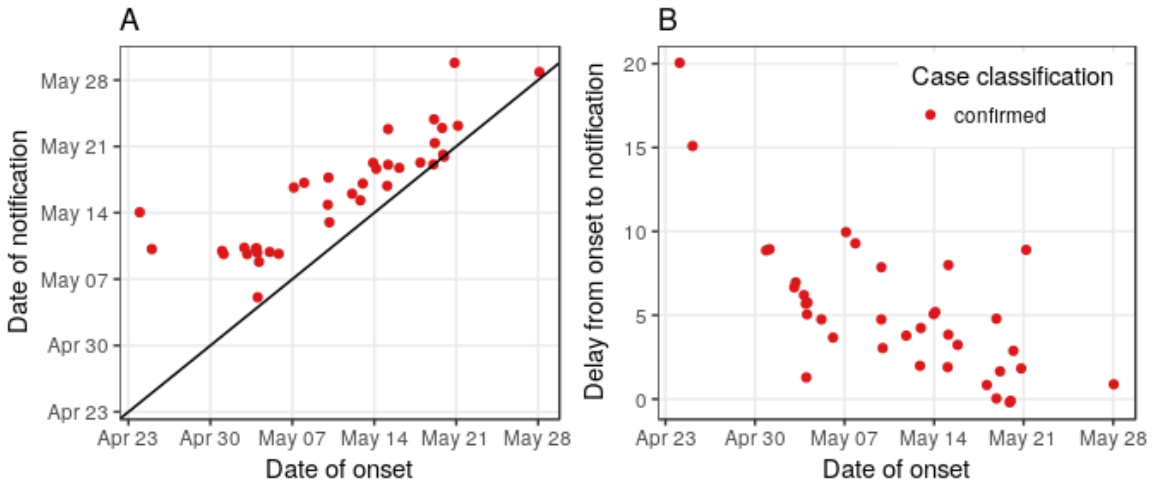
		(range)				CI)	
death	14	9.3 (2 - 27)	7.2	9.3 (6.6 - 13.6)	6 (4 - 11)	2.4 (1.1 - 4.5)	0.3 (0.1 - 0.5)
notification	49	10 (0 - 38)	10.4	10 (7.6 - 13.7)	10.4 (7.6 - 15.2)	0.9 (0 - 1.3)	0.1 (0.1 - 0.1)

342

343 **Defining the trusted period**

344 Although the earliest case illness onset was 5 April, the first case notification for a confirmed case was not until 5
345 May. As a result, the cases with the earliest illness onset have the longest illness-onset-to-notification delays.
346 We must remain cautious until this delay distribution has stabilized. However, the short illness-onset-to-
347 notification delays observed for the most recent illness onset cases suggests that we could generally expect
348 short delays for cases with illness onset dates in the recent past and into the near future. We base all following
349 analyses in this section on the confirmed cases only.

350 The linear relationship between illness-onset-to-notification delay and date of illness onset among confirmed
351 cases is shown in Figure A2. A linear regression model fitted to data with dates of illness onset from 30 April
352 onwards showed that date of illness onset explained 33% of the variation in the illness onset-to-notification
353 delay.



354

355 **Figure A2: A) Dates of notification and illness onset. B) Delay from illness onset to notification against**
356 **date of illness onset**

357

358 The regression model fitted to these data implies that for cases with illness onset on day d_o, the mean delay to
359 notification is given by $\Delta_{n-o} = ad_o + b$, where $a = -0.23$ and $b = 4149.57$ (given in days since 1 January
360 1970) are the slope and intercept of the linear model fitted above, respectively. Hence the expected date of
361 notification d_n for a case with illness onset on day d_o is $d_o + \Delta_{n-o}$, with the actual values being normally
362 distributed around this mean, with a standard deviation defined by the residual standard error of the regression,
363 sd=2.42. The simple regression fitted to the confirmed cases appears a good description of the data, implying
364 that the variance of this normal distribution is independent of the illness onset date. This means that we would

expect x% of cases with illness onset on d_o to have been reported by day $d_{n,x} = d_o + \Delta_{n-o}(d_o) + q_x$, where q_x is the inverse cumulative distribution of the normal distribution with mean 0 and standard deviation sd=2.42. Substituting Δ_{n-o}, we can resolve this to give the critical illness onset date as

$$d_o(x) = \frac{d_n - q_x - b}{a + 1}.$$

This estimated linear relationship allows us to estimate what proportion of the cases which experienced illness onset on a particular date (from 30 April onwards) have already been included in the dataset. On this basis we estimate that 90% of cases with illness onset on 25 May will have been included and 95% of those with illness onset on 24 May will have been included. Note: The latest date of any sort included in the analysed dataset was 30 May. Thus, having estimated that the recorded incidence of cases with dates of illness onset between 30 April and 24 May (inclusive) were at least 95% complete (potentially relative to an unknown but constant level of overall under-reporting), we consider this interval to be our 'trusted period' from the point of view of estimating incidence trends, and thereby predicting future incidence, for confirmed cases (Figure 1 main text).

Estimating the Reproduction Number *R*

We use an approach similar to those previously described[13,6] to quantify transmissibility from the incidence time series during the trusted period of the current epidemic, assuming a certain distribution for the serial interval (the time between illness onset in a case and illness onset in their infector). Here, we assumed the serial interval distribution inferred for the West African Ebola epidemic[6], namely a gamma distributed serial interval with mean 15.3 days and standard deviation 9.1 days.

The distribution of the serial interval used in our analyses is shown in Figure A3.

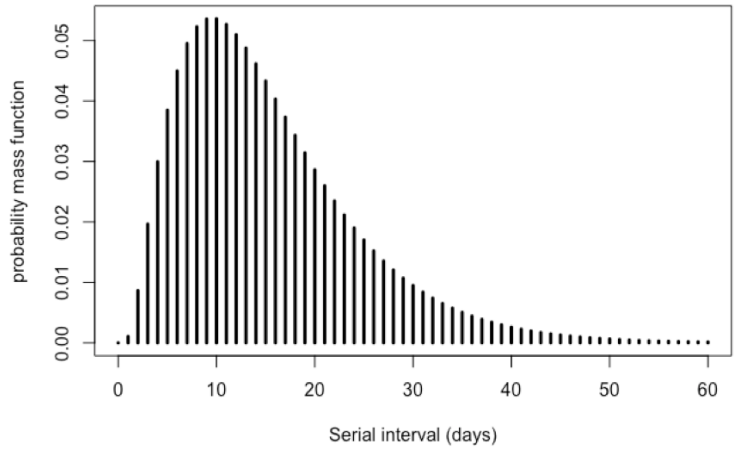


Figure A3: Distribution of the serial interval (the time between illness onset in a case and illness onset in their infector), assuming a gamma distributed serial interval with mean 15.3 days and standard deviation 9.1 days, as estimated during the West African Ebola epidemic[6].

We assumed that transmissibility was constant throughout the trusted period, and estimated the reproduction number, *R*, defined as the average number of secondary cases infected by an infected individual. The estimate of *R* is informative as if *R* is above the threshold value 1, and remains above 1, the outbreak is likely to grow further, whereas if *R* is below 1, and remains below 1, the outbreak will die out.

Given uncertainty surrounding the epidemiological situation before the trusted period, we only used incidence data during the trusted period, and reconstructed the incidence before the trusted period whilst estimating R[11]. Our method assumes that the daily incidence can be approximated by a Poisson process using the so-called renewal equation:

$$I_t \sim Poisson(R_t \sum_{s=1}^{t} I_{t-s} w_s) \quad (1)$$

where I_t is the incidence on day t, R_t is the reproduction number on day t, and w is the probability mass function of the serial interval.

Sensitivity analyses

Sensitivity analyses were performed

- using an alternative distribution of the serial interval, with mean 16.1 days and standard deviation 4.4 days as estimated during a previous Ebola outbreak in DRC[26].
- Changing the end of the trusted period, bringing it forward or backward by one day.
- Changing the start of the trusted period, to keep only a week-long trusted period.

The estimates of R obtained in sensitivity analyses were:

Sensitivity analysis	Median R estimate	95% Credible Interval
Main analysis	1.03	0.829-1.37
R estimated over trusted period minus 1 day	1.05	0.834-1.41
R estimated over trusted period plus 1 day	1.01	0.817-1.3
R estimated over last week of trusted period	1.03	0.786-1.62
Alternative serial interval distribution from a previous outbreak in DRC	1.03	0.818-1.41

Forward Projections

We used the renewal equation (equation 1) to project the incidence forward, given a back-calculated early incidence curve, an estimated reproduction number, and the observed incidence over the trusted period. We sampled 200 sets of back-calculated early incidence curves and reproduction numbers from the posterior distribution obtained in the estimation process. For each of these sets, we simulated 2000 stochastic realisations

of the renewal equation starting from the end of the trusted period; leading to a total of 400,000 projected incidence trajectories.

Projections were made on a 4-week horizon (25 May to 21 June). The projections assume that the transmissibility remains constant over this 4-week horizon. If transmissibility were to decrease as a result of additional control interventions and/or changes in behaviour over this time period, we would predict a lower number of cases; similarly, if transmissibility were to increase over this time period, we would predict a higher number of cases. We limited our projection to 4 weeks only as assuming constant transmissibility over longer time horizons seemed unrealistic.

Super-spreading has been shown to be an important characteristics of Ebola transmission[27]. To account for this characteristic, we considered an alternative projection method, assuming that secondary cases are generated according to a negative binomial distribution:

$$I_t \sim \text{NegBin}\left(R_t \sum_{s=1}^{t} I_{t-s} w_s, z\right)$$

The value of the overdispersion parameter, z, was taken from analyses of exposure patterns during the West African Ebola epidemic[27].

Figure A4 shows the 4-week projected daily incidence and cumulative incidence from the end of the trusted period (25 May to 21 June).

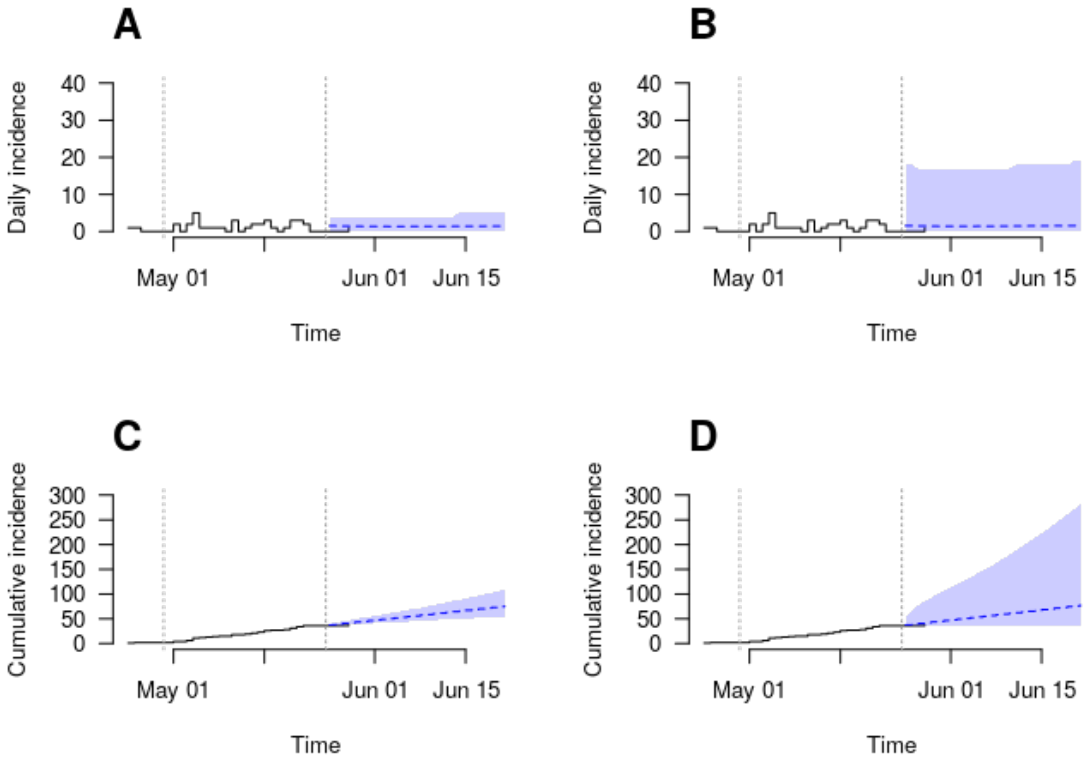


Figure A4: Observed and projected incidence (A-B) and cumulative incidence (C-D) of illness onset, over time, using the homogeneous transmissibility (or Poisson) model (A, C) and the heterogeneous transmissibility (or negative binomial) model (B, D). The black solid lines show the observed incidence of confirmed cases over time. The blue dashed lines show the mean and the shaded area the 2.5% and 97.5% quantiles of the projected incidence. The vertical dotted lines show the trusted period. Note the y-axis scale on panels A and B differ to that of panels C and D.

438 **References**

1 WHO. Ebola Virus Disease Factsheet. www.who.int/news-room/fact-sheets/detail/ebola-virus-disease (accessed May 31 2018)

[2] Van Kerkhove M, Bento AI, Mills HL et al. A review of epidemiological parameters from Ebola outbreaks to inform early public health decision- making. Nature, 2015; Scientific Data, 2, art. No. 150019

[3] WHO Ebola haemorrhagic fever in Zaire. 1976. Report of an International Commission. Bulletin of the World Health Organization. 1978;56(2):271-293.

[4] WHO Ebola Virus Disease- The Democratic Republic of the Congo. http://www.who.int/csr/don/13-may-2017-ebola-drc/en/ (accessed 24 May 2018)

[5] WHO Ebola Response Team. West African Ebola Epidemic after One Year — Slowing but Not Yet under Control. N Engl J Med. 2015;372: 584–587

[6] WHO Ebola Response Team. Ebola Virus Disease in West Africa — The First 9 Months of the Epidemic and Forward Projections. N Engl J Med. 2014;371: 1481–1495.

7 CDC. Years of Ebola Virus Disease Outbreaks. www.cdc.gov/vhf/ebola/history/chronology.html (accessed May 24 2018)

8 WHO Ebola Virus Disease- The Democratic Republic of the Congo. http://www.who.int/csr/don/13-may-2017-ebola-drc/en/ (accessed 24 May 2018)

9 Khan AS, Tshioko FK, Heymann DL, et al. The Reemergence of Ebola Hemorrhagic Fever, Democratic Republic of the Congo, 1995. Journal of Infectious Diseases. 1999;179:S76-S86.

10 WHO. Case definition recommendations for Ebola or Marburg virus diseases. 9 April 2014. http://apps.who.int/iris/bitstream/handle/10665/146397/WHO_EVD_CaseDef_14.1_eng.pdf?sequence=1

[11] WHO. Implementation and management of contact tracing for Ebola virus disease. Sept 2015 http://apps.who.int/iris/bitstream/handle/10665/185258/WHO_EVD_Guidance_Contact_15.1_eng.pdf?sequence=1

[12] Garske T, Cori A, Ariyarajah A, Blake IM, Dorigatti I, Eckmanns T, et al. Heterogeneities in the case fatality ratio in the West African Ebola outbreak 2013-2016. Philos Trans R Soc B. 2017;372. pii: 20160308

13 Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. Am J Epidemiol. 2013;178: 1505–1512.

14 Bower H, Johnson J, Bangura MS, Kamara AJ, Kamara O et al, Exposure-Specific and Age-Specific Attack Rates for Ebola Virus Disease in Ebola-Affected Households, Sierra Leone Emerg Infect Dis. 2016 Aug; 22(8): 1403–1411.

15 Okware SI, Omaswa FG, Zaramba S, et al. An outbreak of Ebola in Uganda. Tropical Medicine and International Health. 2002;7(12):1068-1075.

16 World Health Organization. Outbreak of Ebola haemorrhagic fever in Yambio, south Sudan, April-June 2004[159 KB, 8 pages]. Weekly Epidemiological Record. 2005;80(43):370-375.

17 Shultz J, Espinel Z, Espinola M, and Rechkemmer A Distinguishing epidemiological features of the 2013–2016 West Africa Ebola virus disease outbreak. Disaster Health. 2016; 3(3): 78–88.

18 Hunt L, Gupta-Wright A, Simms V, Tamba F, Knott V, Tamba K, et al. Clinical presentation, biochemical, and haematological parameters and their association with outcome in patients with Ebola virus disease: an observational cohort study. Lancet Infect Dis. 2015;15:1292–9.

19 Rudolf F, Damkjær M, Lunding S, Dornonville de la Cour K, Young A, Brooks T, et al. Influence of Referral Pathway on Ebola Virus Disease Case-Fatality Rate and Effect of Survival Selection Bias. Emerg Infect Dis. 2017;23(4):597-600. https://dx.doi.org/10.3201/eid2304.160485

20 Crowe SJ, Maenner MJ, Kuah S, Erickson BR, Coffee M, Knust B, et al. Prognostic indicators for Ebola patient survival. Emerg Infect Dis. 2016;22:217–23.

21 Shears P, O'Dempsey T. Ebola virus disease in Africa: epidemiology and nosocomial transmission J of Hosp Inf 90 (2015) 1e9

22 Dunn AC, Walker TA, Redd J, Sugerman D et al. Nosocomial transmission of Ebola virus disease on pediatric and maternity wards: Bombali and Tonkolili, Sierra Leone, 2014. Am J of Inf Control. 2016;44 (3): 269-272

23 Forrester JD, Hunter JC, Pillai SK et al. Cluster of Ebola Cases Among Liberian and U.S. Health Care Workers in an Ebola Treatment Unit and Adjacent Hospital — Liberia, 2014. MMWR. 2014; 63(41);925-929

24 WHO. Health worker Ebola infections in Guinea, Liberia and Sierra Leone. Preliminary report. May 2015. http://www.who.int/csr/resources/publications/ebola/health-worker-infections/en/

25 WHO. Ebola Virus Disease- the Democratic Republic of the Congo. 17 May 2018. http://www.who.int/csr/don/17-may-2018-ebola-drc/en/ (accessed 31 May 2018)

[26] Maganga GD, Kapetshi J, Berthet N, Kebela Ilunga B, Kabange F, Mbala Kingebeni P, et al. Ebola Virus Disease in the Democratic Republic of Congo. N Engl J Med. 2014;371: 2083–2091.

[27] International WHO Response Team. Exposure Patterns Driving Ebola Transmission in West Africa: A Retrospective Observational Study. PLOS Med. 2016;13:e1002170.