Simulating familial Ebola transmission in urban environments

Team # 38723

February 10, 2015

Abstract

We present an efficient way to implement a limited medicine store so as to minimize casualties, specifically considering the best way to distribute a hypothetical cure for Ebola virus disease (EVD) in an urban environment. To account for the high likelihood of EVD spreading within households, we break the population of a city into family units of various sizes and treat the possibilities of intra- and inter-familial spread separately. After this approach has been outlined, we will focus our analysis on Sierra Leone, the country currently worst affected by the 2014 outbreak. We will compare the time taken to eradicate the disease and the number of deaths in the process without the cure and with the cure dispensed according to several different schemes.

Contents

1   Introduction

The 2014 outbreak of Ebola Virus Disease (EVD) has been the deadliest and most widespread in the history of the disease. While staying below many predictions, the total number of deaths from this outbreak is greater than the total of all deaths in previous outbreaks. Since the beginning of the outbreak in December 2013, there have been about 22500 total reported cases with over 9000 fatalities [1]. It is estimated that these numbers underrepresent reality by a factor of three.

Although EVD is relatively hard to transmit, the poverty, poor hospital procedures, and burial practices of West Africa are the perfect confluence of factors for its virulent spread. Moreover, the lack of infrastructure in this region of the world means that optimally distributing a potential cure would be a difficult challenge. It is therefore vital that we are prepared to implement a hypothetical EVD cure, which would likely be in scarce supply, in an efficient and optimal manner.

1.1   Introduction to Ebola

EVD is a fatal and quick-acting disease of humans. When contracted, it first manifests itself through sudden influenza-like symptoms, particularly fever, followed by diarrhea and vomiting. In the later stages of the disease, patients often experience the internal and external bleeding that gives the "hemorrhagic fever" its name, which frequently lead to death by fluid loss. If this does not occur, the patient recovers within about 7 days following, although residual symptoms may persist for several weeks [13]. As seen in the above chart, the time from contraction to death or recovery is no more than 21 days. Therefore, the total number of Ebola cases reported in the last 21 days is an upper bound to the total number of Ebola cases currently existing.

Despite the large amount of media coverage portraying Ebola as a quickly spreading epidemic, the number of current cases is quite small and certainly treatable. According to the World Health Organization (WHO), there have been about 368 reported cases of Ebola is the last 21 days. Virtually all new cases have occurred in one of three West African countries: Sierra Leone, Guinea, and Liberia. We make the following series of assumptions to reduce this problem to an manageable position.
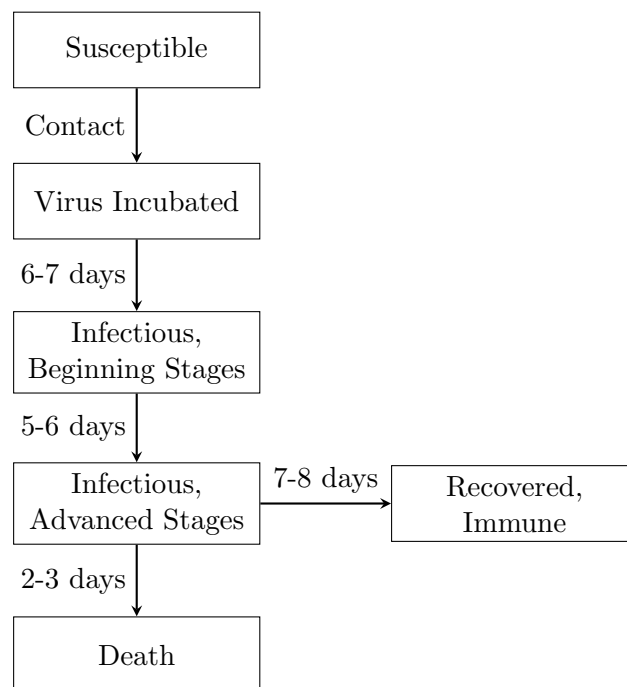
3

Figure 1: A simplified progression of EVD symptoms

1.2   Assumptions and Simplifications

1. Restatement and clarification of the problem:
   The developed medicine works to stop Ebola by curing patients whose diseases has not yet progressed to advanced stages. The medicine is fast-acting and cured patients become immune to the disease. Examination of similar Ebola outbreaks (Congo 1995 and Uganda 2000) have seen that recovered Ebola victims are indeed immune to reinfection [5].

   The non-advanced stage of Ebola is defined above as the period of 5–6 days after symptoms develop, so this is the period in which patients will be classified as curable. This is generally before the symptoms of external and internal bleeding set in, but while vomiting and diarrhea are already present.

   The medicine will most likely be intravenous, and therefore can only be delivered at cities with significant healthcare infrastructure already in place.

2. We expect error margins to be large.
   Many previous projections of Ebola cases have had error margins of greater than a factor of 10. The potentially exponential growth rate of the Ebola virus makes models extremely sensitive to conditions that are often beyond their ability to accurately measure. Therefore, we cannot expect linear accuracy and thus we can discard "edge" cases.

   Similarly, we only consider recorded data, that is, cases that are confirmed, probable, or suspected. This is obvious as there is nothing to be done about data which we do not have.

3. We only consider the three main countries: Sierra Leone, Guinea, and Liberia.
   Of the 22500 total cases, all but 35 occurred in these three countries. The six other countries – Nigeria, Mali, the US, the UK, Senegal, and Spain − have confirmed that they have not received a case since 12/17/2014. Thus it is justfied to ignore the other countries.

   In fact, over the last 21 days, there have been 368 new cases of Ebola. Of those, only 17 have been from Liberia, which is less than 5% of the total cases. From the above assumption, Liberia is not considered in our upcoming model. Rather, the cities there can be thought of as smaller-scale equivalents of populations in Sierra Leone or Guinea.

   Our model ignores rural areas and focuses our on urban population centers. This is for three reasons: par It is less feasible to distribute

medicine and identify cases in the rural areas in an efficient manner. Therefore, it is simpler to consider the application of medicine in urban areas, and we will focus our analysis on the effects of the medicine on these areas.

EVD in rural areas is less likely to affect the macroscopic, long term evolution of the epidemic, as West African customs and economics imply that most care of the sick in rural areas is done inside the families' homes, and thus the spread of EVD in rural areas is often limited to its spread in families and close friends. This is supported by evidence from the 1976 Sudan and 1995 Congo epidemics, where the disease eventually died out in the more rural locales, and mostly propagated in the densely populated hospitals [13, 5].

4. On a related note, the presence of EVD in urban centers poses a greater threat of spreading the disease to a wider area, through travel.

5. Transportation between cities and other densely populated areas is negligible.

During epidemics, travel is often restricted or at least carefully monitored. Sources show that this was indeed the case in West Africa during the height of the 2014 EVD epidemic [12]. In any case, it is enough to assume that the density of Ebola patients remains roughly homogeneous given any traveling.

2   Modeling Different Types of Interaction Within Cities

The Ebola virus, although highly dangerous when contracted, differs from other epidemic diseases in its relatively limited means of communication. The disease is not airborne, and while it can be spread through contact, the most common way it is transmitted is though contact with the bodily fluids of the infected [6]. As a result, infections are significantly more likely to spread within households (in which family members must care for the infected in the early stages of the disease) and hospitals [18]. This suggests a natural, if simplified, division of the ways people can be infected by contacts:

1. Within households, usually in the early stages of the infection.

2. By acquaintances outside of one's household; this is less common, but must occur for the disease to become a threat.

   3. In hospitals, where the former two methods of transmission can still occur, but are augmented by the possibility of transmission through healthcare workers.

   While the spread of the disease through hospitals was an important and dangerous factor in its early stages, we assumed that the current increased awareness of the disease and the safety precautions it requires would reduce hospital-transmitted cases to a negligible amount, or at least to the point where we could effectively combine it with other cases transmitted to non-family contacts. The dynamics of how Ebola is transmitted in healthcare settings are complex and would be the next priority in a further investigation.

2.1   Variations on R_0

One of the most important statistics relevant to any outbreak of a disease is the *basic reproduction number*, denoted R_0. Roughly, R_0 is defined as the average number of people one person will infect while they are infectious. So $R_0 > 1$ implies that a disease can spread and become an epidemic, while $R_0 < 1$ implies that it will eventually die out, through exponential growth and decay, respectively.

   In past and current Ebola outbreaks, models have been used to estimate the value of R_0, usually yielding results between 1.5 and 2.0 [2, 5, 15]. Averaging the various values predicted in the literature for the Zaire strain of Ebola virus gives a basic reproduction number of 1.7, which is what we used. This is a fairly low number, suggesting that strategic application of medication could eradicate the disease quickly.

   In order to apply the simple value R_0 to the model outlined, here, it is helpful to estimate related values: R_{0fam}, the average number of family members or similarly close contacts infected by one person, and $R_{0friend}$, the average number of others infected. As the specific means of transmission of Ebola are less thoroughly studied, and are dependent on cultural context as opposed to being properties of the disease alone, these numbers are more difficult to get good values for. However, certain rough estimation techniques, along with the assumption that R_{0fam} is going to be significantly higher than $R_{0friend}$, can be applied.

   In the case of a smaller outbreak of the virus in South Sudan, the WHO was able to get very detailed statistics on how the disease spread among 5 families it affected [4]. In particular, the authors traced the number of cases in each generation of the disease's spread, until it died down after 7

7

generations. After the index case, the disease spread almost entirely within 5 families, meaning that the $R_{0\text{fam}}$ value for this specific case could be obtained by choosing a cut-off point in the middle of the disease's progression by generations (before its spread was curbed), counting the number of infections up to and including that generation, subtracting the 4 infections that occurred between families, and dividing by the number of infected up to the previous generation.

In this case, by the 5th of 7 generations, 27 people had been infected, not including the index case, while 18 people had been infected by the 4th generation. This gives an $R_{0\text{fam}}$ value of $\frac{23}{18} = 1.28$. As this value also seemed intuitively reasonable, we based our investigation upon an $R_{0\text{fam}}$ of 1.3.

3   Model Sans Medicine

3.1   Parameters

Our model has the following parameters:

1. N, the total population.

2. I_{init} initially infected.

3. M, medicine units available per day.

3.2   Other Necessary Data

Currently, the center of the epidemic is in Sierra Leone, although it is still also present in the two other countries. So, to determine the size of an average family for the purposes of the model, we consulted statistics on the average size of a Sierra Leonean family, produced by the World Food Program [11]. This gave the average size of a household, defined as being "a group of people who eat from the same pot and are responsible to the same head", as 10 people. From the data given on the average household sizes for each of the country's districts, we estimated a standard deviation of 2.9.

Just as important were statistics on the progression of Ebola. The incubation period, the time from initial infection to manifestation of symptoms, was highly variable, with official recommendations giving only the time interval of 2–21 days. However, the general consensus seemed to fall upon a mean incubation period of 6–7 days [5, 13]. The progression of the disease after this was taken from a WHO report on its symptoms published in the

wake of the first major outbreak [13]: initial fever and pains would last for 2 days, with the gastrointestinal issues such as diarrhea and vomiting setting in on the 3rd day, and bleeding occurring around day 6. Finally, it was assumed that, with death frequently occurring on days 8 and 9, and with recovery generally beginning around day 15, the average infectious period was 10 to 11 days.

Finally, we took the basic R_0 values estimated above and converted them into probabilities that the disease would be transmitted within and outside of families. If p_{fam} is the probability that a specific family member living in a household with some sick family member gets sick on any given day, then assuming a person has an average of 10 family members and remains infectious for 10 days, they will infect an average of $10 \times 10 \times p_{\text{fam}}$ family members; this must equal $R_{0\text{fam}}$. From our estimated $R_{0\text{fam}}$ value of 1.3, this gives a p_{fam} of 0.013.

Although similar in appearance, the "friendship coefficient" p_{other} is defined slightly differently from p_{fam} in that it refers to the probability that a specific infected person will infect any other person outside of their family on a given day. Using the infectious period of 10 days again, this implies that $10 \times p_{\text{other}} = R_{0\text{friend}}$. In this model, since $R_{0\text{fam}} + R_{0\text{friend}} = R_0 = 1.7$, we have that $R_{0\text{friend}} = 0.4$, and $p_{\text{other}} = 0.04$.

3.3   Constructing the Initial State

We pair each person into a family, generating the family size using our estimated distribution of family sizes in Sierra Leone. Once everyone has been placed into families, we randomly choose I_{init} of them to become infected, generating the stage of their EVD based upon an expected total time of infection of 16 days (above) using a uniform distribution to get the values. Once this has been achieved, our model is ready to go - we have N people divided into families with I_{init} of them infected with Ebola.

3.4   The Algorithm Without Medicine

Our medicine-free algorithm evolves the system with a time interval of 1 day. For each person i that is infected with Ebola, the following happens:

1. For each not yet infected family member j of i, i infects j with probability p_{fam}.

2. i infects some non-family member j with probability p_{other}.

9

3. If i has been infected for more than 3 days, then using a normal distribution centered at 14 days we either kill/recover i or let them continue to survive with EVD. For example, if i was infected for 14 days then he or she would have a 1/2 chance of either recovering/dying or continuing to suffer from EVD.

For the purposes of this model, as callous as it may sound, it is actually not particularly relevant whether the infected die or recover. Assuming a constant death rate, the number killed will be directly proportional to the number of infected, so the latter is the only quantity we need to focus on optimizing.

4 Optimizing Medicine Implementation

4.1 Schemes and Medication Assumptions

We built our model with different Medicine Implementation plans, or Schemes. This allowed us to make an educated choice for the best possible implementation with different initial conditions, with the parameter of most import being the amount of medicine available per day.

Our biggest assumption was with respect to the medicines applicability time-frame: namely that the medicine is only effective after the patient is symptomatic (6 days since initial infection, see Figure 1) and is no longer effective once the patient is displaying advanced symptoms such as external bleeding, (12 days since initial infection, see Figure 1). We call EVD in this range curable-EVD. Hence in the schemes detailed below we are only applying medicine to those who have curable-EVD.

4.2 The Schemes

Our schemes were as follows:

Scheme 1: This was the most naive scheme of all. Here, given some medicine M to distribute per day, we hand it out on a first come, first serve, basis. That is, as we update our model, we randomly choose people with curable-EVD and cure them. This corresponds to the physical situation of aid workers having M initial units of medicine in the beginning of each day and then finding people with curable-EVD, and handing out the medicine then.

Scheme 2: Since we think that the primary mode of EVD transmission is intra-familial contact, we supposed it might be best to do as follows: order

10

all families by size, and then going from largest families to smallest families, check to see if at least one of the family members has curable-EVD, and then hand out medicine to the *entire* family (even if they aren't all suffering from curable-EVD). Then, if any of them currently have curable-EVD they can immediately take the medicine, and if they don't currently have it, if they ever develop it they can take the medicine and be cured. This way, since the family members have likely already caught it, we can stop the spread in their family and hopefully slow its overall spread. We could implement this in reality by requiring everyone who desires medication to petition for medication, and submit basic statistics including family size as part of their petition. Once armed with this data we could distribute the medicine accordingly.

Scheme 3: This is the same as Scheme 2, except that when we find a family with at least one sick member we only give out the medicine to the currently sick members (i.e. those who currently can take the medication and have it cure them; so those who have had the infection for between 6 and 12 days). This could be implemented physically in the same way as Scheme 2.

Scheme 4: This is a variant of Scheme 3 where we rank each family with at least one member with curable-EVD not just by family size but also by number suffering from EVD. Then we distribute medicine accordingly: priority is given to those families with the greatest size, and then to the families with the most members suffering from EVD, giving medicine only to those family members who are currently suffering from EVD and are curable. This could also be implemented physically by requiring everyone desirous of the medication to submit a request (or to have their family members submit one if they are incapable) wherein they list the current number of their family members who have EVD and their family size.

Scheme 5: Here we vary Scheme 3 by ordering families with at least one member suffering from curable-EVD not by size but by number infected. Then we distribute medicine solely to those in each family currently have curable-EVD. This again could be accomplished with a pamphlet system.

11

4.3   Comparing the Schemes

Now we will run the model with the different schemes and compare their performance at different medicine allocation levels and a test population of $N = 100000$. (The actual value of N doesn't really matter as our model is only sensitive to local connections which is, in our case, dominated by the family units. Since the family units don't change, on average, with the overall population N, it follows our model isn't sensitive to the value of N for reasonable initial conditions. This is reasonable because EVD is only transmitted via extremely close contacts, and the number of such contacts for any given person is almost certainly invariant under the total population of their environment, except in extreme cases.)

The data comparing the Schemes is shown on the next page. (Here ∞ means that the disease was not eradicated and continued growing indefinitely.) As we can see, with low amounts of medicine (i.e. $M < 10$), Scheme 3 and Scheme 4 are the best two, with Scheme 3 just barely performing better than Scheme 4. With larger amounts of medicine and more initial infected, Scheme 4 performs better than Scheme 3. Hence our model suggests that we should use Scheme 4 in situations with more infected and more medicine available, and Scheme 3 in those with not very much medicine available. This is what we shall do subsequently.

4.4   Running the Model on Freetown

Using the assumption that travel between cities is negligible, we implemented our algorithm for Freetown. Freetown has a population of about 1.61 million people, with 73 current cases of Ebola [9, 3]. The results of our model in that specific case are plotted in Figure 5. We see from Figure 5 that our model predicts we could end the entire outbreak in Freetown in about 10 days given 10 units of medicine per day. Even if we only had 1 unit available per day, our model predicts the outbreak could be ended in about a month and a half.

5   Prescriptions for Medicine Implementation

As we found in §4.3, the optimal implementation is Scheme 4 if there are more than about 10 units of medicine per day, and Scheme 3 if not. These schemes are detailed in §4.1 for your reference.

Scheme, #	Total Cases	Days Until Eradicated	Medicine Required
1	204	28	84
2	327	90	270
3	156	20	60
4	173	22	66
5	183	39	117

Figure 2: Scheme Performance with $M = 3$ per day, $I_{init} = 100$, and $N = 100000$

Scheme, #	Total Cases	Days Until Eradicated	Medicine Required
1	204	38	380
2	∞	∞	∞
3	137	16	160
4	123	13	130
5	146	23	230

Figure 3: Scheme Performance with $M = 10$ per day, $I_{init} = 100$, and $N = 100000$

Scheme, #	Total Cases	Days Until Eradicated	Medicine Required
1	433	27	1080
2	∞	∞	∞
3	394	13	520
4	372	11	440
5	393	14	560

Figure 4: Scheme Performance with $M = 3$ per day, $I_{init} = 100$, and $N = 100000$

Medicine Per Day Given	Total Cases	Days Until Eradicated
1	297	45
3	228	26
10	177	11
20	158	9
50	155	9

Figure 5: Our Model of Freetown with $N = 1.61 \cdot 10^6$, $I_{init} = 67$

6   Directions For Further Investigation

One of the most distinctive features of Ebola's spread was its tendency to explode within hospitals, especially when the disease had not been conclusively identified [18]. Without very specific knowledge of healthcare practices in the affected countries, it's difficult to model this component, but a more developed form of our model could include additional, larger-scale groups of people in hospitals, similar to the family model, to account for this.
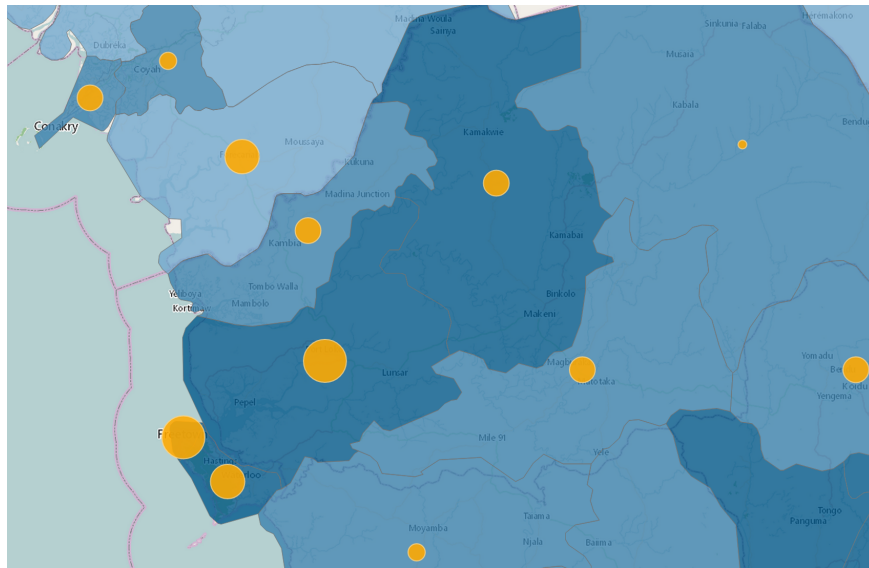
Transportation is another important factor in any epidemic, and is particularly necessary in considering how to stop a later resurgence of the disease. Our model could potentially be expanded to take account of transportation by considering connections between cities and the likelihood of cities transmitting the disease to each other through these connections, just as we have considered the simpler connections between family members.

While we made an effort to collect information on Ebola from a wide variety of sources, some basic statistics used by our model were difficult to verify. Given additional detailed accounts of the beginnings of Ebola outbreaks among families and in other small communities, we could improve our estimates for the values of $R_{0\text{fam}}$ and $R_{0\text{friend}}$, and perhaps diversify the latter to refer to different types of contact.

7   Letter for the World Medical Association

Recently, the World Medical Association has announced the development of an effective medicinal treatment for the Ebola virus. The medicine is able to completely eliminate Ebola and its symptoms from patients who are not yet in the advanced stages of the disease. Once cured, a patient is no longer at risk of contracting Ebola.

We plan to begin distributing the medicine immediately to the people in West Africa. Currently, there are approximately 400 cases of Ebola known, with possibly more undiagnosed. Our top priority will be the heavily populated cities within the region. We look to specifically target Sierra Leone and Guinea, as the most heavily hit were the regions of Freetown and Port Loko in Sierra Leone. Also in need of assistance are the Forecariah and Conakry districts of Guinea.

(Image courtesy of the World Health Organization.) Fortunately, preexisting healthcare infrastructure in these population regions, along with previous preparations for Ebola, means that the delivery of the medicine can be efficient and effective. We are unfortunately unable to provide to every individual case of Ebola with our current resources. However, models suggest that if a reasonable proportion of Ebola victims are successfully cured, we believe this will be sufficient to eradicate the virus.

We are cautiously optimistic about the future. Already, a carefully coordinated response on the part of various health organizations has curtailed

Ebola's spread to well below the projections made at the outbreak's height. However, this is still a dangerous time to be complacent, as any lapse in care or security protocols could bring the disease back. With this new cure, we can not only cut off the current epidemic but prepare to quickly put a stop to future outbreaks before they reach dangerous levels.

References

[1] 2014 Ebola Outbreak in West Africa - Case Counts. Center for Disease Control and Prevention, February 6, 2015

[2] Althaus, Christian. Estimating the reproduction number of Ebola virus (EBOV) During the 2014 Outbreak in West Africa. PLOS Currents Outbreaks, Ed. 1, September 2014.

[3] Balk, Deborah. Global rural-urban mapping project (GRUMP). Socioeconomic Data and Applications at NASA, 2010.

[4] Baron, Roy, Joseph McCormick, and Osman Zubeir. Ebola virus disease in southern Sudan: hospital dissemination and intrafamilial spread. Bulletin of the WHO, Vol. 61, No. 6, 1983.

[5] Chowell, G. et. al. The basic reproductive number of Ebola and the effects of public health measures. Journal of Theoretical Biology, March 2004.

[6] Dowell, Scott F. et. al. Transmission of Ebola hemorrhagic fever. The Journal of Infectious Diseases, No. 179, February 2004.

[7] Kermack, W. and A. McKendrick. The problem of endemicity. Contributions to the Mathematical Theory of Epidemics II, October 1932.

[8] Liu, Su-Yu, Andrea Baronchelli, and Nicola Perra. Contagion dynamics in time-varying metapopulation networks. Physical Review E87, Boston MA, March 2013.

[9] Rivers, Caitlin M. Data for the 2014 Ebola outbreak in West Africa. Github, December 2014.

[10] Scarpino, Samuel et. al. Epidemiological and viral genomic sequence analysis of the 2014 Ebola outbreak reveals clustered transmission. Clinical Infectious Diseases, January 2015.

[11] Sierra Leone: Household Food Security Survey in Rural Areas. World Food Program, Vulnerability Analysis and Mapping Branch, November 2008.

[12] Sierra Leone quarantines over one million. Aljazeera, September 2014.

[13] Simpson, D. Marburg and Ebola virus infections. World Health Organization, Department of Medical Microbiology, No. 36, 1977.

[14] WHO Ebola Response Team. Ebola virus disease in West Africa −
the first 9 months of the epidemic and forward projections. The New
England Journal of Medicine, Vol. 371, No. 16, October 2014.

[15] −. Ebola virus disease in West Africa − Supplementary Index 1. The
New England Journal of Medicine, Vol. 371, No. 16, October 2014.

[16] −. Cumulative Ebola cases and past 21 days new cases, February 4 2015.
World Health Organization.

[17] −. Factors that contributed to undetected spread of the Ebola virus and
impeded rapid containment. World Health Organization, January 2015.

[18] WHO/International Study Team. Ebola haemorrhagic fever in Sudan,
1976. World Health Organization, 1978.