The SIR Model and the 2014 Ebola Virus Disease Outbreak in Guinea, Liberia and Sierra Leone

Collins Bekoe
Department of Mathematics and Computer Science
Faculty of Science and Technology
Prince of Songkla University
Muang Pattani, 94000, Thailand

bekoedarkocollins@ymail.com

Abstract
This research presents a mathematical model aimed at understanding the spread of the 2014 Ebola Virus Disease (EVD) using the standard \textit{SIR} model. In modelling infectious disease dynamics, it is necessary to investigate whether the disease spread could attain an epidemic level or it could be wiped out. Data from the 2014 Ebola Virus Disease outbreak is used and Guinea where the outbreak started is considered in this study. A three dimensional non-linear differential equation is formulated and solved numerically using the Runge-Kutta 4\textsuperscript{th} order method in the Vensim Personal Learning Edition Software. It is shown from the study that, with public health interventions, the effective reproductive number can be reduced making it possible for the outbreak to die out. It is also shown mathematically that the epidemic can only die out when there are no new infected individuals in the population.

Keywords: Mathematical Model, Epidemiology, Epidemic, Susceptible, Infected, Recovered, SIR, Ebola Virus Disease.

1. INTRODUCTION
Epidemiology is a branch of science that investigates the risks factors responsible for causing diseases through retrospective and prospective observations, a complete history of disease, and the frequency of occurrence or transmission mechanisms of disease in populations and explores preventive and therapeutic control measures [1].

While epidemiology is mainly about diseases, the free dictionary puts the definition of disease as an abnormal condition of a part, organ, or system of an organism resulting from various causes, such as infection, inflammation, environmental factors, or genetic defect, and characterized by an identifiable group of signs, symptoms or both. That is an illness that affects a person, animal or plant (Merrian - Webster Dictionary). In many cases, terms such as disease, disorder, morbidity and illness are used interchangeably. Disease classification can be either airborne, that is any disease caused by pathogens and transmitted through the air or infectious. These are also known as transmissible diseases. This can be further broken down into contagious (an infection such as influenza or the common cold, that commonly spreads from one person to another) and communicable. We can also talk about non-communicable disease which is non-transmissible. Non-communicable diseases cannot be spread directly from one person to another. Heart diseases and cancers are some examples of non-communicable diseases.

Infectious diseases are caused by pathogenic microorganisms, such as bacteria, viruses, parasites or fungi (WHO; Infectious Diseases). These diseases can be spread, directly or indirectly from one person to the other. Many organisms live in and on our bodies. Some of the infectious diseases can be transmitted by bites from insects or animals. Others are acquired by ingesting contaminated food or water or being exposed to harmful organisms in the environment.
The study of epidemics with its long history has come up with an astonishing number and variety of models and explanations for the spread and cause of epidemic outbreaks [2].

The Ebola Virus Disease, which the World Health Organisation declared an outbreak on 23rd March 2014, first occurred in Yambuku, Zaire and surrounding areas in 1976. The disease was spread by close personal contact and by use of contaminated needles and syringes in hospitals/clinics. This outbreak was the first recognition of the disease [15]. Since then, there have been many outbreaks with the latest in Guinea, Liberia and Sierra Leone. Now the Ebola Virus Disease which began in December, 2013 and killed a two year old boy in Guinea, was the beginning of an outbreak that killed over 5000 people in a year, leaving hundreds of children orphaned and affecting thousands more [16].

More than 18 months on from the first confirmed case recorded on 23 March 2014, at least 11,312 people have been reported as having died from the disease in six countries; Liberia, Guinea, Sierra Leone, Nigeria, the US and Mali. The total number of reported cases is more than 28,457 [10]. As the disease spread is through contacts and at a mortality rate of 70% [9], there is the need to understand the epidemiology and patterns of the disease. Therefore, the Susceptible (S), Infected (I) and Recovered generally called the SIR model first developed in 1927 by W. O. Kermack and A. G. McKendrick [21], is used in this study to model the spread of the Ebola Virus Disease (EVD) mathematically to know if the epidemic would ever die out and also understand the behavior of the transmission and recovery incidence of the disease to help Policy makers, Governments and Public Health experts in containing future outbreaks.

2. RELATED WORK
Mathematical models and its associated literature have been well studied and researched for some time now. Modelling and simulation of the pattern and spread of infectious disease in a particular location to provide vital and early information to public health experts and medical personnel has had sufficient grounds in recent times especially in Africa.

A model is an appropriate representation of a given part of the material world, the study or understanding of whose part it facilitates, Oko, 1998. An epidemic model is a simplified means of describing the transmission of communicable disease through individuals [21].

In [21], it is shown that, W.O kermarck and A.G Mckendrick created a model in which a fixed population was divided into compartments with some assumptions about the nature and time rate of transfer of disease from one compartment to the other. When dealing with locations with large number of people, deterministic compartmental mathematical models are used [21].

Koriko and Yusuf, 2008 formulated a mathematical model to simulate Tuberculosis disease population dynamics. In the model, it is stated that tuberculosis disease population dynamics depends more on the number of actively infected individuals in the population at the initial time and also on the disease incidence transmission rate at a given time [5].

The effects of the Ebola outbreak in Congo in 1995 and Uganda in 2000 was modelled by [4] using a modified variation of the original SIR model. In the study, they quantified the effectiveness of interventions in terms of the reproductive number and they were able to conclude that the interventions were successful in controlling the epidemic.

This work presents the application of the standard SIR model to understand the 2014 Ebola Virus Disease outbreak in Guinea, Liberia and Sierra Leone with further mathematical analysis.

3. MODEL FORMULATION
The Susceptible(S), Infected/Infective (I) and Recovered/Removed(R) model is used in epidemiology to show the interaction between the number of people who are susceptible to a
disease, those who are currently with the infection and those who have recovered or died at a
given time in a population. This can be represented schematically as below.

Susceptible (S)  \[\rightarrow\]  Infected (I)  \[\rightarrow\]  Recovered (R)

3.1 Assumptions of the Model
- The disease spread or outbreak is severe
- Everyone who gets infected is removed from the population either through recovery or
death.
- The population is large, fixed in size and it is confined geographically.

From the assumptions, we can write

\[\text{Susceptible (S)} \xrightarrow{\alpha} \text{Infected (I)} \xrightarrow{\beta} \text{Recovered (R)}\]

In this model, two parameters \(\alpha\) and \(\beta\) would be used to depict the transmission and recovery rate respectively.

The transmission varies jointly as the susceptible and infective (as they are capable of
transmitting the disease) population as the infection depends on the interaction between the
susceptible class and infective class.

The infection rate of a disease is dependent on the mortality rate \(M\) and the number of people
who are susceptible to the disease. This is the rate at which the disease is transmitted from an
infected person to a susceptible individual. The transmission rate \(\alpha\) always lies between 0 and 1.

Therefore, we have

\[\alpha = \frac{M}{S}.\quad (1)\]

3.2 The SIR model
A two dimensional system of autonomous differential equations with state variables \(S, I, R\) is
presented.

\[
\begin{align*}
\frac{dS}{dt} &= -\alpha SI \\
\frac{dI}{dt} &= \alpha SI - \beta I \\
\frac{dR}{dt} &= \beta I.
\end{align*}
\]

From equations (2), (3) and (4), \(S(t) + I(t) + R(t) = N\), where \(N\) is the total size of the
population. The \(S, I\) and \(R\) are all bounded above by \(N\) [2].

Suppose that each infected individual has \(k\) number of contacts from the susceptible
compartment per unit time, then we have

\[kS/N\quad (5)\]

Also if a fraction of the contacts made, results in transmission of the Ebola Virus Disease (EVD),
then each infected individual infects \(kzS/N\) individuals per unit time. That is
\[ \alpha = \frac{a}{N} \]  \hspace{1cm} (6) \] where \( a = k z \) and \( z \) is the transmissibility of infectious disease.

4. MODEL ANALYSIS

We now analyse the model and state some theorems.

4.1 Defining the Existence of Long Term Limits

From equation (2), the only way an individual can leave the susceptible compartment is by coming into contact with somebody who has the EVD and because of this, the right hand side of (2) is negative. Therefore, we have

\[ \frac{dS}{dt} \leq 0. \]

From equation (4), the only way a person is removed through recovery or death is by being infected with the EVD. Therefore, we have

\[ \frac{dR}{dt} \geq 0. \]

Since \( 0 \leq S(t) \leq S(0) \leq N \) and \( R(0) \leq R(t) \leq N \), we have

\[ S(\infty) = \lim_{t \to \infty} S(t), \]  \hspace{1cm} (7) \]

\[ R(\infty) = \lim_{t \to \infty} R(t) \] and \hspace{1cm} \[ I(\infty) = \lim_{t \to \infty} I(t) = N - S(\infty) - R(\infty). \]

From the model, the total population \( N \) is given by \( N = S + I + R \). Therefore, there exists a long term.

4.2 Defining the Epidemic Threshold Theorem

The effective reproductive number \( R_e \), which estimates the average number of secondary cases per infectious case in a population made up of both susceptible and non-susceptible hosts, is given by \( R_e = (S(0)/N) a/\beta \) and the basic reproductive number is also denoted by \( R_o \). Although \( R_o \) is only defined at the time of invasion, \( R_e \) is defined at all times. For most models, the contact number remains constant as the infection spreads, so it is always equal to the basic reproduction number \( R_o \) [6].

If the entire population is at the beginning susceptible, then we can have

\[ S(0) = N - 1, \]  \hspace{1cm} (10) \]

\[ I(0) = 1, \]  \hspace{1cm} (11) \]

\[ R(0) = 0. \]  \hspace{1cm} (12) \]

Then \( R_e = ((N - 1)/N) a/\beta \) is approximately equal to the basic reproductive number \( R_o \).

Modelling infectious diseases like EVD demands that we investigate whether the disease spread could cause an epidemic or it could be wiped out. Therefore, we will assume that the term
\( (N-1)/N \) is equal to 1 and show that \( R_e \), the effective reproductive number is the threshold to determine the epidemiology of the EVD. A population will rarely be totally susceptible to an infection in the real world. We therefore state this theorem

4.3 Theorem 1

a) If \( R_e \leq 1 \), then \( I(t) \) decreases monotonically to zero as \( t \to \infty \).

b) If \( R_e > 1 \), then \( I(t) \) starts increasing, reaches its maximum and then decreases to zero as \( t \to \infty \). This is called an epidemic [3].

Therefore an epidemic can occur in a population if the basic reproductive number \( R_e > 1 \).

Proof: from equation (3) and section 4.1, we have

\[
\frac{dI}{dt} = (\alpha S - \beta)I \leq (\alpha S(0) - B)I
\]

\[= \beta(R_e - 1)I \leq 0, \text{ for } R_e < 1\]

Again from equation (3)

\[
\frac{dI}{dt}(0) = \beta(R_e - 1)I(0) > 0, \text{ for } R_e > 1
\]

That is \( I(t) \) increases at \( t = 0 \)

Equation (3) implies that \( I(t) \) has only one non-zero critical point. Hence Theorem 1a is proven[3].

4.4 The SIR Model and Data from the 2014 Ebola Outbreak

Although these West African Countries (Guinea, Liberia, Sierra Leone) have been severely affected by the outbreak, parameter values for the model will be from Guinea as it is where the outbreak first begun in 2014 and also to better understand the analysis, it will be applied to the other countries later in the study.

The table below shows the total number of people living in these 3 West African countries as stated on the website of the World health Organisation (WHO) [17].

<table>
<thead>
<tr>
<th>Country</th>
<th>Population Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guinea</td>
<td>11745000</td>
</tr>
<tr>
<td>Liberia</td>
<td>4294000</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>6092000</td>
</tr>
</tbody>
</table>

**TABLE 1**: The total number of people living in the 3 countries as at 2013.

The WHO officially declared Ebola Virus Disease on 23rd March, 2014 [19], but two confirmed and probable cases had been reported in Guinea on 5th January, 2014 [14].

Therefore assigning parameter values to the model, we have

\[
N = 11745000, \\
I = 2, \\
R = 0.
\]
So the susceptible individuals in the population at the time was $N - I + R = 11744998$.

The WHO, puts the incubation (the time interval from infection with the EVD to onset of symptoms) period to 2 to 21 days [18]. Therefore, it could be estimated that the EVD has a duration of $D = 10$.

$$\beta = \frac{1}{10} = 0.1$$

The WHO also puts the mortality rate ($M$) Of the EVD at 0.7 [9]. That is, the disease kills seven out of 10 patients.

Therefore from equation (1) we have

$$\alpha = \frac{0.7}{11744998} = 5.9599 \times 10^{-8}$$

To better understand the outbreak of the EVD from March 2014 till the time of this study, it would be appropriate to solve the non-linear system of equations (2), (3) and (4). The equations cannot be solved completely with an explicit formula solution. A numerical approach is then used. 

The Vensim PLE software is used over a three month period.

5. RESULTS AND DISCUSSION

The Vensim PLE Software was used to solve the model numerically and also for plotting the outcome of the numerical results. Alpha and beta in figure 1 is the transmission and recovery rates, respectively.

![Modelling The 2014 Ebola outbreak with the standard SIR Model using the Vensim PLE Software](image)

**FIGURE 1:** Modelling the 2014 Ebola outbreak.

Figures 2, 3, 4 and 5 below show the graphs of the susceptible, Infective and removed individuals and their interactions with respect to time.
From figure 2 it is observed that the number of individuals susceptible to the EVD remains the same for the first 18 to 22 days. As these individuals come into contact with infected people and begin to show signs or symptoms of the disease, they are taken to the infective compartment. However, as the infection continues, there is a steep decline in the susceptible compartment. Towards the end of the 3 months’ period. There is little or no change in the number of susceptible people as a result of a decline in the number of infected individuals. This is because everyone infected is removed through recovery or death.

From figure 3 the number of infected individuals increase from the first 15 to 30 days probably because of the uncertainty of the EVD which allowed a greater transmission rate. The maximum number of infected individuals is seen on the 30th day and this indicates the number ever to be infected and after this point the infection rate decreases.
FIGURE 4: The graph of the Recovered compartment.

The graph of the recovery/removed compartment increases as the number of infected individuals decreases. This is mostly because of increased awareness of the disease and other medical support given by stakeholders and public health experts.

FIGURE 5: The graph of all three compartments.

The graph of the interaction between all three compartments is shown in figure 5 for Guinea and in figure 6 and 7 for Liberia and Sierra Leone, respectively. This gives a better understanding into the behavior of the spread of the disease.
5.1 Further Analysis. Why will the Disease Die Out?

To show that the disease dies out, we have \( I(\infty) = 0 \).

For all initial conditions,

if \( I(\infty) \neq 0 \), then equation (4) implies that for sufficiently large \( t \),

\[
\frac{dR}{dt} > \beta I(\infty)/2 > 0.
\]

\( \therefore R(\infty) = \infty \), which is a contradiction.

5.2 The maximum number of infected individuals

A solution for \( I_{\text{max}} \) can be obtained even though the system of differential equations cannot be solved explicitly.

Dividing equations (2) by (3), we have
\[ \frac{dS}{dl} = -\frac{\alpha SI}{\alpha S - \beta I} \]  

(14)

This is separable, since \( I > 0 \)

\[ \int \frac{\alpha S - \beta}{\alpha S} ds = -\int dl. \]

Hence

\[ -I - S + \frac{\beta}{\alpha} \log S = C. \]

Therefore

\[ I(t) + S(t) - \frac{\beta}{\alpha} \log s(t) = I(0) + S(0) - \frac{\beta}{\alpha} S(0). \]  

(15)

Then we can have \( I_{max} \) when \( \frac{dl}{dt} = 0. \) This can occur when \( S = \frac{\beta}{\alpha}, \) then equation (15) will become

\[ I_{max} + \frac{\beta}{\alpha} - \frac{\beta}{\alpha} \log \beta / \alpha = I(0) + S(0) - \frac{\beta}{\alpha} \log S(0). \]  

(16)

We can also show that for an initially susceptible population, the maximum fraction of infected individuals is a function of the basic reproductive number \( R_o, \)

\[ I_{max} / N = 1 - \frac{1}{R_o} (1 + \log R_o). \]  

(17)

### 5.3 Will the Epidemics Ever End?

When there is an outbreak of an infectious disease, public health professionals always look out for the time the epidemic would end. Does the epidemic end because there are no susceptible individuals in the population? If this is true then we have \( S(0), \) but the epidemic rather ends because there are no new individuals with infection but not because of the lack of susceptible individuals. We show this by stating this proposition.

### 5.4 Proposition: The Limiting Number of Susceptible Individuals

We have

\[ S(\infty) \geq S(0) \exp(-R_o) > 0 \]

Proof: Dividing equation (2) by (4)

We have

\[ \frac{dS}{dR} = -\frac{\alpha}{\beta I} = -\frac{\alpha S}{\beta} \]

This is separable since \( S > 0. \)

\[ \int \frac{1}{S} ds = \int -\frac{\alpha}{\beta} dR, \]

\[ S(t) = S(0) \exp(-\alpha(R(t) - R(0))/\beta). \]
Since $0 \leq R(t) - R(0) \leq N$, it follows that
\[
S(t) \geq S(0) \exp(-\alpha N / \beta) \quad \text{and} \\
S(\infty) \geq S(0) \exp(-\alpha N / \beta) = S(0) \exp(-R_0) > 0
\] (18)

As the outbreak is ongoing and the epidemic is high, the susceptible compartment decreases and the infected compartment increases. Eventually, $S(t)$ drops below $\beta/\alpha$ and this makes the rate at which individuals recover exceed the rate at which new individuals get infected.

This is a fundamental fact which cannot be claimed from data but rather from a mathematical model.

6. CONCLUSION
In this study, the general epidemiology of the Ebola Virus Disease is considered and modelled using the SIR model. The non-linear systems of differential equation formulated were solved numerically using the RK4 method over time period of 120 days in the Vensim PLE Software. For simplicity, Guinea as one of the three West African Countries that was hit severely by EVD in 2014, was used for this study. As stated earlier, modelling infectious disease dynamics demands that we investigate whether the disease spread could attain an epidemic level or it could be wiped out. The Ebola Virus Disease is spread largely through contacts between an infected person and a non-infected person [18]. Infected people cannot be entirely separated from the rest of the large population. The separation can only be done when an infected person visits a health center.

However, Public Health experts can prevent infection from developing into an epidemic by reducing the effective reproductive number ($R_0$) in theorem 1. This can be done by reducing $S(0)$ by giving vaccines to those who are susceptible to the disease. The contact rate $k$, can be reduced by self-isolation of susceptible individuals and also the transmissibility $z$ can be reduced by encouraging the washing of hands and the use of hand sanitizers.

However, the model can be valid in providing information to governments and organisations like the WHO and the CDC and also disease prevention agencies when more information with regards to data from ground zero of the outbreak is obtained accurately. The transmission ($\alpha$) and recovery ($\beta$) rates are often inaccurate. According to the World Bank’s estimate/statistics, the patient to doctor ratio in the 3 West African countries is 10000 in Guinea, 70000 in Liberia and 45000 in Sierra Leone [16]. Hence, there are inadequate resources to keep track of the number of patients who are still living with or recovered from an infection.

On the other hand, the model does not separate the individuals who have recovered and those who have died.

The “Identify, Isolate and inform” strategy by the CDC for Emergency department (ED) staff is ideally good for such disease outbreaks [20]. According to WHO’s Ebola situation report on 14th October, 2015, no confirmed cases of Ebola virus disease (EVD) were reported in the week to 11th October. It states that, it is the second consecutive week with zero confirmed cases. However, 150 registered contacts remain under follow-up in Guinea of which 118 are high risk and an additional 259 contacts remain untraced. There also remains a near-term risk of further cases among both registered and untraced contacts. In Sierra Leone, 2 high-risk contacts associated with the 2 most recently active chains of transmission in the country were lost to follow-up and have not yet been found [12]. This confirms the analysis made in section 5.1 that, the outbreak can only die out when $I(\infty) = 0$ and also under section 5.3, the epidemic can only end when there are no new infected individuals in the population.
According to the model, more than half of the population got infected with the disease in the 120 day period but in the real world situation, 3800 confirmed, probable and suspected EVD cases were reported in Guinea of which 2534 individuals died [13]. This is because of the Public health intervention and strategies employed by WHO health officials and other local health professionals in containing the disease. A graph of the total suspected, probable and confirmed cases is visually represented in [11] which gives the current situation of the EVD as it is been contained.

Considering previous global health crises like the 2009 Influenza pandemic, the SARS epidemic, the 2014 Ebola outbreak and others, mathematical models are proving to be essential and necessary to forecast the expected total number of infections and deaths during and at the end of such outbreaks [3, 4, 5, 6, 7, 8]. A fact is that a timely identification of an outbreak remains of paramount importance in controlling the spread of the Ebola Virus Disease [8] and other infectious.

Modelling the outbreak of the EVD mathematically can help scientist to analyse epidemics and help prevent further outbreaks which might be fatal in the future. The general trend of the epidemic has been presented and this can be modified to account for other parameters for better understanding and analyzing future outbreaks.

7. RECOMMENDATIONS FOR FURTHER STUDY
What has been achieved through simulation is the ability to be able to predict how the disease can spread through a population. The reality is that some individuals living with the disease do not visit health centers. As a result, they are not accounted for and sometimes the model can over estimate or underestimate the spread of the infection. What should be understood is that, in modelling infectious disease dynamics, anybody living with the disease is important when accounting for the number of people living with disease as one infected person can spread the disease through the population. With this in mind, the model can be modified to add a parameter to account for people who are not recoded as infected or removed (through recovery or death). Also, with almost accurate parameter values, section 5.1 through to 5.3 can be analyzed to show the maximum number of people who can be infected and when the epidemic would die out.

8. ACKNOWLEDGEMENT
This work is supported by the Higher Education Research Promotion and the Thailand’s Education Hub for Southern Region of Asean Countries Project Office of Higher Education Commission.

9. REFERENCES


