Simulation of the Third Wave of COVID 19 Infections in Zambia using the SIR Model

Moses Mwale (mkmwale@hotmail.com)
University of Johannesburg  https://orcid.org/0000-0002-3414-3523

Boniface Kanjere
National Food and Nutrition Commission

Chuncky Kanchele
The Road Transport and Safety Agency

Solomon Mukosa
Lusaka Apex Medical University

Research Article

Keywords: COVID-19, SIR Model, Infectious Disease Modeling, Basic Reproductive Number, Simulations, Zambia, Forecasting

Posted Date: February 17th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1337105/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Background

Various mathematical modeling approaches are used to provide a robust framework for understanding the transmission dynamics of infectious diseases in human populations. In an epidemic, models can be used for the analysis of the spread of a disease, forecasting, identifying trends and making parameter estimates which can be used for planning and implementing intervention measures.

Methods

This study utilizes the classical Susceptible - Infected - Recovered (SIR) model to analyze the evolution of COVID-19 in Zambia during the third wave of infections. The model is fitted to actual COVID-19 data for Zambia for the third wave of the pandemic obtained from the Zambia National Public Health Institute (ZNPHI). The transmission and recovery rates are estimated by minimizing the error between the fitted curve and the real data using the least square approach.

Results

Model simulations indicate that the basic reproductive number \(R_0\) for Zambia is 1.31 meaning that, on average, 1.31 persons are infected for each infected person. At the worst point of the epidemic, we expect that 591,743 people will contract the virus and 7,144 fatalities will be recorded. To prevent the spread of infection, the model estimates that at least 24 percent of the population will need to be vaccinated. With the country’s projected population of about 18.91 million, this translates to roughly 4.5 million people.

Conclusion

The severity of COVID-19 infection, hospitalizations, and deaths in Zambia is substantially higher than national testing data suggests, according to model projections. More modeling work is needed to acquire accurate estimates of the disease burden in society, to inform resource allocation, and to aid mitigation planning, especially in countries that may lack adequate national surveillance systems.

Background

Coronaviruses (CoVs) are a broad group of viruses that cause respiratory diseases in humans which include common cold, bronchitis, pneumonia, Middle East respiratory syndrome (MERS) and severe acute respiratory syndrome (SARS) [1, 2]. The current COVID-19 pandemic is caused by a variation of the coronavirus called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [3]. On February 14, 2020, Egypt announced the first cases of COVID-19 on the African continent [4]. Zambia recorded its first two confirmed COVID-19 cases on 18th March 2020 [5]. The government moved quickly and enacted
regulation aimed at limiting the spread of the virus within communities [6, 7]. The public health safety measures implemented by the Ministry of Health (MOH) include; screening of travelers entering the country; limitations of public gatherings; restriction of all non-essential travel; and closure of schools, religious institutions, nightclubs, bars, cinemas, casinos and gyms [8]. MOH further actively promoted frequent hand hygiene practices, physical distancing, stay at home and wearing of masks as precautionary measures to be exercised by the public as recommended by the World Health Organization [9]. As of 10th July 2021, Zambia was experiencing a 3rd wave of the COVID-19 infections and had recorded a total of 174,789 and 2,016 deaths [10].

In many Low-middle income countries (LMICs), determining the scope of the pandemic is a huge challenge in Zambia. In comparison with high income countries, the majority of LMICs are characterized by poorly resourced diagnostic laboratories resulting in inadequate testing capacity [11]. This situation has resulted in a lack of adequate data and information on the true number of incidents of COVID-19 making it difficult to determine the epidemiology of the virus in these communities [12, 13].

Epidemiologists have used mathematical models extensively to research the evolution of an outbreak and to advise public health strategies. Models for infectious diseases enable us to estimate key parameters such as threshold values, transmission rates and mortality rates; and help to evaluate the sensitivity of changes to these parameters on the spread of the pandemic [14]. This knowledge on infectious diseases can greatly aid in developing better strategies for managing transmission of the virus in communities. The analysis of an infectious disease using statistical models will help us address important questions about the disease, such as when it may appear, how it may spread, its severity, identification of immunization target groups, how likely are many cycles of resurgence, and what mitigation measures to put in place to stop the outbreak. In the absence of high quality and detailed disease data, modelling is especially useful for decision-making [15, 16]. Models are also used to optimize the choice of a vaccination programme, allocation of scarce resources as well as the best combination of control techniques or eradication methods.

This study utilizes the classical susceptible-infected-recovered (SIR) modeling approach to analyze the evolution of COVID-19 in Zambia during the third wave of infections. We specifically evaluate the basic reproductive number $R_0$ of the virus over this period, determine the peak of the infections and determine the maximum number of people who need to be vaccinated in order for the disease to stop. This paper is structured as follows; Section 2 presents the classical SIR model presenting the model formulation and assumptions and defines the basic reproductive number $R_0$ and presents some limitations of the modeling process. In section 3, we present the results of the analysis and simulations, and these results are discussed in section 4. Section 5 concludes by summarizing the important results.

**Methods**

**The SIR Model**
The standard approach to epidemic modeling is the use of compartmental models, that divides a population in different groups (compartments) in relation to diseases and assigns rules for transfer of individuals between the groups [17, 18]. The modelling approach is based on non-linear differential equations and is predominantly deterministic in nature. At any given time, each individual is considered to be in one compartment but can move to another compartment depending on the model parameters [19]. The most widely used compartmental model is the Susceptible–Infected–Recovered (SIR) model which assumes that people become immune once they recover. The classical SIR model was developed and presented in a series of publications by Kermack and McKendrick [20, 21, 22] almost a century ago. At any point of time, the SIR model is used to estimate the number of people who are susceptible to the diseases, are infected or have recovered from the disease [19].

**Model Formulation**

Let the number of people susceptible, infected and removed from the population at any given time \( t \) be represented by \( S(t) \), \( I(t) \), and \( R(t) \) respectively. Further, the total population at any given time \( t \) is given by

\[
N(t) = S(t) + I(t) + R(t).
\]

1

At the start of the epidemic, we have the initial conditions \( S(0) = S_0 = N - 1 \), \( I(0) = I_0 = 1 \) and \( R(0) = R_0 = 0 \) meaning that initially, only one person is infected in the population. Keeping all other variables constant, the number of susceptible persons decrease symmetrically by

\[
\frac{dS}{dt} = -\frac{\beta}{N} S(t) I(t)
\]

2

The parameter \( \beta \) represents the effective contact rate of transition from the susceptible group to the infected group. As the epidemic evolves, the number of individuals in the infected population is given as

\[
\frac{dI}{dt} = \frac{\beta}{N} S(t) I(t) - \lambda I(t)
\]

3

The rate of recovery and mortality is represented by the parameter \( \lambda \), the rate of transition from the Infected group to the recovered group. This results in the following dynamical system which represents the classical SIR model.
Two parameters specify the SIR model: 1) the effective contact or infection rate $\beta$ and 2) the recovery rate $\lambda$. If the infection rate $\beta$ exceeds the recovery rate $\lambda$, the result will be an accumulation of people in the infected group. An epidemic is caused by a continual rise in the number of people in the infected compartment i.e. $\frac{dI}{dt} > 0$.

At the beginning of the epidemic, everyone is susceptible to the diseases and thus we approximate $\frac{S(t)}{N}$ to 1 and get

$$\frac{\beta}{\lambda} \times \frac{S(t)}{N} > 1$$

$$\frac{\beta}{\lambda} = R_0 > 1$$

The SIR model satisfies the following properties

1. The susceptible function $S(t)$ decreases to

$$S_\infty \exp \left( -\frac{R_0}{N} \left( R_\infty - R_0 \right) \right) \text{ as } t \to \infty$$

2. The recovered function $R(t)$ increases up to

$$R_\infty = N - S_0 \exp \left( -\frac{R_0}{N} \left( R_\infty - R_0 \right) \right) \text{ as } t \to \infty$$
3. If $R_0 > \frac{N}{S_{0}}$, then $I(t)$ increases up to a maximum value

$$I_{max} = I_0 + S_0 - \left(1 + \ln\frac{R_0S_0}{N}\right)\frac{N}{R_0}$$

4. If $R_0 < \frac{N}{S_{0}}$, then $I(t)$ decreases to 0

At the start of an epidemic, everyone (apart from the infected cases) is susceptible and so we say that $S_0 \approx N$ then we replace $N/S_0$ by 1. As the virus progresses over time in the population, the number of susceptible (S) decreases as individuals are infected and move to the infectious group. The number of people in the infectious (I) compartment either recover or die and move to the recovered (R) group.

**Model Assumptions**

It is important to make certain assumptions about the means of transmission in order to describe a mathematical model for the spread of an infectious disease. Numerous researchers believe that SARS-CoV-2, the virus which causes COVID-19 is spread mainly when an infected person is in contact with another person [23]. The SIR model is based on the following assumptions

1. The total population remains constant in time i.e. the population is closed with no births, deaths or migrations.

2. The infected and susceptible populations are homogeneously mixed, which means that all individuals within the population have an equal chance of coming into contact with each other.

**The Parameter $R_0$**

The basic reproductive number (universally denoted as $R_0$) is a central idea to mathematical epidemiology. $R_0$ is defined as the number of cases of the disease transmitted by a typical infected person over the full course of infection [23]. This number is an epidemiological metric used to measure the infectious potential of a disease. The threshold $R_0 = 1$ is the divisive line between the end of the outbreak and the beginning of an epidemic. We assume that for a single wave of infection, the time period is short enough to ignore demographic shifts, and that those who recover gain full immunity against reinfection.

If $R_0 > 1$, the epidemic increases exponentially as on average, one infected individual infects more than one other individual. If $R_0 \leq 1$, there will be no epidemic and the disease will die out with out affecting a large proportion of the population [23, 24]. $R_0$ is calculated as
\[ R_0 = \frac{\beta}{\lambda} \]

8

Limitations

The parameters estimated using the model should be taken with some caution as the model formulation is based on unrealistic assumptions (e.g, fixed reproduction number \( R_0 \), no public health interventions etc.). The classic SIR model also does not take into account the interventions which may be implemented to curb the spread of the virus and mutations of the virus.

The Data

The Data used to model the COVID-19 was obtained from the Zambia National Public Health Institute which is under the Ministry of Health. ZNPHI publishes country situational reports which provide statistics on the evolution of COVID-19 in Zambia. Figure 1 displays an overview of the daily number of cases and death recorded in Zambia from 18th March 2020 to 10th June 2021.

Results

Estimating Parameters of the Model

The SIR model developed was fitted to the data on COVID-19 in Zambia using a ODE solver and optimizer packages in R programming language. We minimized the sum of squared differences between \( I(t) \) and the corresponding number of cases predicted by our model to obtain the residual sum of squares (RSS)

\[ RSS(\beta, \lambda) = \Sigma (I(t) - \hat{I}(t))^2 \]

The daily cumulative incidence rates for Zambia were computed from the number of active cases recorded during the period 5th May to 16th June 2021. This period corresponded to the third wave of COVID-19 infections. The predicted incidences from the SIR model was then compared to the data with the actual incidence. On 5th May the country had a total of 808 active cases and this figure was used as the initial number of cases in the model. It is estimated that Zambia has population of around 18,920,651 for the year 2021 [25].

Figure 2 (a) shows that the number of observed cases during the period of interest follows the number of confirmed cases simulated by our model over the observation period. Figure 2 (b) is similar to 3 (a) except that the cumulative incidence given by the y-axis is measures on a log scale. Transforming the number of confirmed cases makes it easier to see the differences between the observed and the expected number of cases.
The optimized parameters, $\beta$ and $\lambda$ are acquired by curve fitting the daily cumulative incidence rates for Zambia. The SIR model was fitted to the country data by finding the parameter values of $\beta$ and $\lambda$ that minimize the residual sum of squares between the observed incidence and the predicted cumulative incidence (as predicted by our model). The best parameter estimates give $\beta = 0.568$ and $\lambda = 0.432$. Model parameters were checked for convergence.

Figure 3 (b) shows epidemic's peak will be reached within 60 day. At this point, approximately 591,743 people would have contracted the virus representing 3.1% of the country's total population. At the end of the epidemic, 43.7% percent of the population will remain susceptible (see Fig. 3 (a)).

**Reproduction Number**

The basic reproductive number, $R_0$, indicates how many secondary illnesses each infectious person causes on average. The $R_0$ was computed from the estimated value of the parameters $\beta$ and $\lambda$, as 1.31 (with 95% Confidence interval, 1.312541 – 1.316890) meaning that, on average in Zambia, 1.31 persons are infected for each infected person. Table 1 below gives published estimates of $R_0$ from other countries.

<table>
<thead>
<tr>
<th>Study</th>
<th>Location</th>
<th>Study date</th>
<th>$R_0$ estimates (average)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferguson <em>et al.</em> [28]</td>
<td>UK</td>
<td>March 2020</td>
<td>2.4</td>
<td>2.0–2.6</td>
</tr>
<tr>
<td>Nguemdjo <em>et al.</em> [29]</td>
<td>Cameroon</td>
<td>6 March to 9 April 2020</td>
<td>1.567</td>
<td>1.536–1.597</td>
</tr>
<tr>
<td>Lounis &amp; Bagal [30]</td>
<td>Algeria</td>
<td>Feb 25th to Aug 12th 2020</td>
<td>1.23</td>
<td>-</td>
</tr>
</tbody>
</table>

The effective reproduction number $R_{eff}$ denotes the number of people in the population who can be infected by a single person at any given time. It is useful in determining the virus's transmissibility at any particular time and is influenced by factors such as the number of infectives (I), the number of
susceptible people (S), disease control measures, and immunization. Figure 4 below shows the effective reproductive number for Zambia.

Effects of Vaccination

Vaccines have been regarded as vital tools for facilitating social and economic activity, hence reducing the disease's impact on society. The proportion of the population that needs to be vaccinated or immunized to prevent disease spread is known as the threshold for heard immunity in classic epidemiological models. This threshold has to be greater than \( 1 - \left(\frac{1}{R_0}\right) \) \cite{32}. Given the basic reproductive number of 1.31 calculated by our model, 24% of the population should be immunized to stop the spread of the infection.

Discussion

We employed the classic SIR model to estimate parameters for COVID-19 in Zambia during the third wave of the pandemic. The basic reproductive number, \( R_0 \) for the country during this period was calculated to be 1.31, meaning that each primary case of the disease resulted in 1.31 other cases. Getting a good estimate of \( R_0 \) is key to answering questions such as; How fast will the disease spread? How many hospital bed-spaces and ventilators will the country need? When should we go into lockdown? Will the second wave be worse than the first? The \( R_0 \) for COVID-19 was initially predicted by the World Health Organization (WHO) to be between 1.4 and 2.4. \( R_0 \) is however tricky to calculate because it depends on; 1) the model used and its underlying assumptions, 2) biological traits of the virus and 3) an understanding of how people come into contact with each other. This makes it vary greatly among countries, cities and neighborhoods.

At the peak of the third wave, our model estimates that 591,743 people will catch the disease and the country will record 7,144 fatalities. These model projections are much higher than the national official figures which are obtained from surveillance and case management data. The model also predicts that the country's vaccination program must reach at least 24 percent of the population in order to develop herd immunity. With the country's population of about 18.91 million \cite{25}, this translates to roughly 4,529,213 people.

Various works have shown that the current data as reported by many countries does not provide an accurate depiction of the health burden of the disease \cite{33, 34, 35}. For example, \cite{36, 37} estimate that the overall number of excess deaths caused by COVID-19 is two to four times higher than the recorded number of confirmed deaths. Variations in testing, inadequate testing capacity, variances in the application of regulations for attribution of death and certification, delays in reporting, data gaps, and insufficient health information systems in many countries are among the reasons for the disparity between reported and estimated figures \cite{37, 33}. 


Estimates predicted by our model can be used by public health institutions across the country to simulate future scenarios and aid in planning for crucial components of the response, such as the number of people who may require hospitalization and ICU treatment. Estimates of the disease's actual burden can also assist governments make decisions about how to prioritize between routine and emergency health services. The SIR model's dynamics also suggest that reducing the contact ratio, or the proportion of the population that comes into contact with an infected person is crucial for slowing the spread of the disease. Measures which have been put in place by governments all over the world such as hand hygiene practices, staying at home, the wearing of face masks and social distancing are all aimed at reducing the contact ratio and slowing the spread of the virus while more effective interventions such as production, procurement and distribution of a vaccine are being put in place.

**Conclusions**

Mathematical modelling of infectious disease is critical to obtaining estimates of the prevalence of the disease in a population. This paper aimed at formulating an SIR model to study the evolution of COVID-19 during the third wave of the pandemic in Zambia. Even the most basic of models as we have formulated here, can provide us with useful estimates for planning public health measures. The transmission and recovery rates for the model are estimated using data for the daily reported cases from 24th May to 16th June 2021. A value of $R_0$ for the country was then calculated as 1.31. The model estimated that 3.1% of the total population was infected with the virus and the vaccination campaign needs to target 24% of the population to prevent sustained spread of the disease. Our results have added to the state of current knowledge on the actual burden or impact of the COVID-19 pandemic in Zambia. These model projections reveal that the severity of COVID-19 infection, hospitalizations, and deaths in Zambia is substantially higher than national testing data indicates. The results of the study can help provide insights and estimates to adequately respond to the COVID 19 pandemic. More modeling work is needed to obtain accurate estimates of the burden of disease especially in countries which may not have adequate national surveillance systems. These predictions also highlight the importance of the enforcement of strict public health measures and restrictions by the government and the urgency for citizens to follow these guidelines in order to lessen the impact of subsequent waves of the pandemic and to relieve pressure on our health care systems.

**List Of Abbreviations**
Declarations

Ethics approval and consent to participate
Not Applicable

Consent for publication
Not Applicable

Availability of data and materials
The data and code used in this study are available upon request from the corresponding author

Competing interests
The authors declare that they have no competing interests.

Funding
This research did not receive any specific grant from any funding sources.

Authors Contributions
MM and BK contributed to conceptualization of the idea. SM, BB and CK worked on the study design and the methodology. MM run the formal analysis, simulations and visualizations. MM, SM, BB and CK contributed in the interpretation of results and writing of the manuscript. All authors have read and approved the manuscript

Acknowledgements
We wish to thank Farai Nyabadza, Biomathematics Professor at the University of Johannesburg for providing guidance and feedback throughout this project. We would also like to thank Dr Muzala Kapina,
Dr Nkomba Kayeyi, Dr Masauso Phiri, Ms Anita Enane and Mr Kelvin Mwangilwa for taking time to review the paper.

References


**Figures**

![New COVID-19 Cases and Deaths](image)

**Figure 1**

The number of confirmed COVID-19 cases in Zambia from March 2021 to July 2021

**Figure 2**

Fitted vs Observed Cumulative incidence

**Figure 3**
Interactions between the Susceptible, Infected and Recovered Populations

Figure 4

Tracking the Effective Reproductive number $R_{eff}$