

Continuous Time Markov Chain Model for Cholera Epidemic Transmission Dynamics

Yohana Maiga Marwa^{1,2}, Isambi Sailon Mbalawata³ & Samuel Mwalili⁴

¹ Mathematics & Statistics Department, Pan African University, Institute for Basic Sciences, Technology and Innovation, Nairobi, Kenya

² Tanzania Public Service College, Singida, Tanzania

³ African Institute for Mathematical Sciences, Dar es Salaam, Tanzania

⁴ Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Correspondence: Yohana Maiga Marwa, Mathematics & Statistics Department, Pan African University, Institute for Basic Sciences, Technology and Innovation, Nairobi, Kenya.

Received: February 27, 2019 Accepted: April 8, 2019 Online Published: April 18, 2019

doi:10.5539/ijsp.v8n3p32 URL: <https://doi.org/10.5539/ijsp.v8n3p32>

Abstract

This paper is concern with modeling cholera epidemic. Despite the advances made in understanding this disease and its treatment, cholera continues to be a major public health problem in many countries. Deterministic and stochastic models emerged in modeling of cholera epidemic, in order to understand the mechanism by which cholera disease spread, conditions for cholera disease to have minor and large outbreaks. We formulate a continuous time Markov chain model for cholera epidemic transmission from the deterministic model. The basic reproduction number (R_0) and the extinction thresholds of corresponding cholera continuous time Markov chain model are derived under certain assumptions. We find that, the probability of extinction (no outbreak) is 1 if $R_0 < 1$, but less than 1 if $R_0 > 1$. We also carry out numerical simulations using Gillespie algorithm and Runge–Kutta method to generate the sample path of cholera continuous time Markov chain model and the solution of ordinary differential equation respectively. The results show that the sample path of continuous time Markov chain model fluctuates within the solution of the ordinary differential equation.

Keywords: cholera epidemic, basic reproduction number, branching process, probability of extinction, Gillespie algorithm

1. Introduction

Cholera is a virulent disease that affects both adults and children, and can kill within few hours if remain untreated. It is a severe water borne infectious disease caused by the vibrio cholerae bacterium (World Health Organization, 2010). Cholera bacterium is found in water or food sources that have been contaminated by faeces from a person who is infected with cholera. Cholera is most likely to be found and spread in places with inadequate clean water, poor sanitation, and inadequate hygiene. Therefore, one can get cholera by drinking water or eating food that are contaminated with the cholera bacterium (Ali et al, 2012).

Cholera has been declared by World health organization (WHO) as the public health problem. Hence, there is a need of finding better ways of dealing with cholera so as to reduce cases of cholera in different countries worldwide. According to WHO report, about 1.4 to 4.3 million cases of cholera are reported each year worldwide and more than 140,000 deaths per year are reported due to cholera (World Health Organization, 2016). For example, in Tanzania the most recent information on cholera suggest that between August 2015 to April 2016, cholera resulted in over 24,000 cases and caused 378 deaths (World Health Organization, 2016).

There are several measures which have been suggested by (World Health Organization, 2010) to prevent cholera. These measures are environmental sanitation, provision of clean water, provision of education on the effect of cholera (World Health Organization, 2006). Despite various efforts made against cholera, the disease still persist in the community and is considered as a very threatening disease in many countries.

A number of studies have been done to study the spread of infectious diseases in deterministic and stochastic models setting (Dimi et al, 2016, Edward & Nyerere, 2015, Einsenberg et al, 2013, Khan et al, 2015, King et al, 2008, Longini et al, 2007, Neilan et al, 2010, Njagarah & Nyabadza, 2015, Panja & Mondal, 2014, Tuite et al, 2013, Marwa et al, 2018). Some of the studies consider compartment I for infected individuals as a single compartment instead of splitting it into two heterogeneous groups that is symptomatic and asymptomatic infected individuals to study the transmission dynamics of cholera epidemic. Also, continuous time Markov chain models, were ignored in their study.

Continuous time Markov chain models help to determine the probability of an outbreak or disease extinction through the use of branching process theory. In continuous time Markov chain we can study the stochastic thresholds for the disease extinction or major outbreak (Allen & Lahodny, 2012). The stochastic thresholds in continuous time Markov chain models are similar to the basic reproduction number, but these stochastic thresholds usually depend on the initial number of infectious individuals and the initial level of the environmental contamination. However, the stochastic thresholds are usually studied from the theory of multitype branching process. Berriman et al. (2013) developed a multi-group semi-stochastic model to describe Salmonella dynamics. In their study did not consider the probability of a major outbreak or host demography. The continuous time Markov chain for cholera epidemic developed in (Lahodny et al, 2015). The model accounted for host demography, direct and indirect transmission, replication of free-living pathogens in the environment and removal of free-living pathogens by natural death or environmental decontamination. The limitations to their model, the infected individuals were termed as a homogeneous group and also, they did not account for removal of free-living pathogens from the environment due to ingestion by hosts.

In this paper we, thus extend the deterministic model developed in (Marwa et al, 2018) by formulating an equivalent cholera continuous time Markov chain model. The formulated cholera continuous time Markov chain (CTMC) model will be analyzed using branching process approximation and probability of extinction. Also, numerical simulation will be carried using Gillespie algorithm. The advantage of using continuous time Markov chain (CTMC) model over deterministic model in epidemic modelling is that the CTMC model yields an estimate of the probability of extinction and outbreaks of cholera disease (Zevika & Soewono, 2018). Also, the measured trajectories behave as predicted. Therefore, the overall purpose of this paper is to study the transmission dynamics of cholera using continuous time Markov chain.

In the next section, we present a deterministic cholera epidemic model with water treatment, which is a model formulated by (Marwa et al, 2018). In section 3, the corresponding cholera continuous time Markov chain model is formulated basing on the assumptions of the ODE model. In section 4, the multitype branching process is discussed. In section 5, the numerical results from the deterministic and stochastic simulations are presented, discussed and compared. Finally, the last section concludes our paper

2. Cholera Deterministic Model

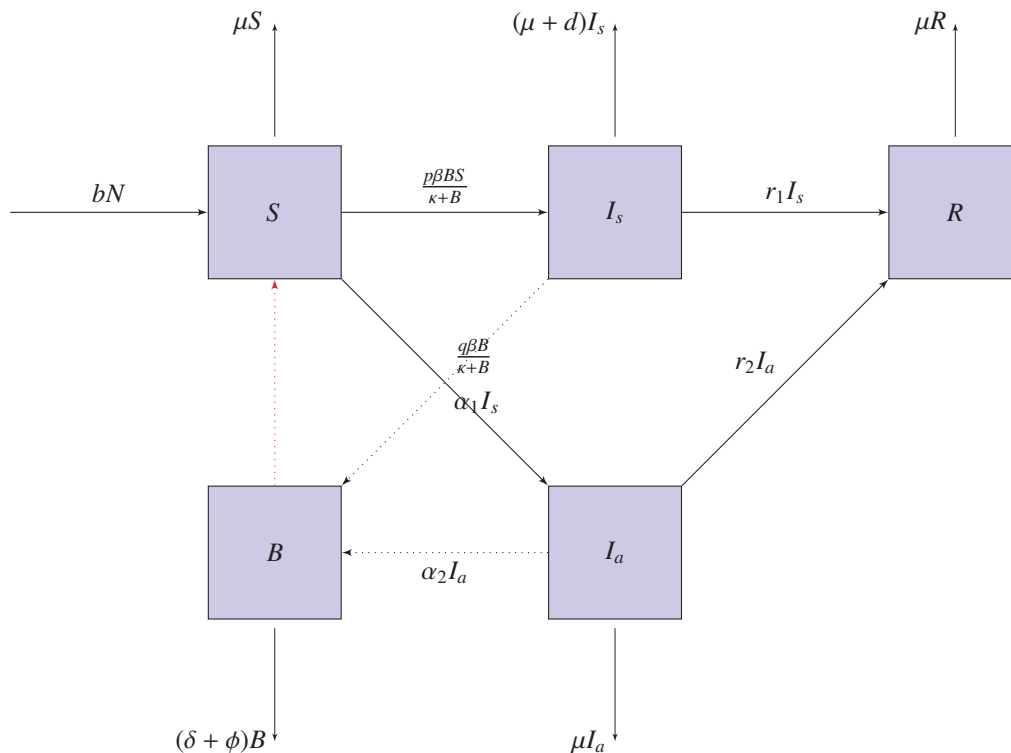


Figure 1. The interaction of cholera epidemic transmission dynamics between compartments are shown

This section, introduce a deterministic $SI_sI_aR - B$ cholera model formulated by (Marwa et al, 2018). In their model, $S(t)$ denotes number susceptible individuals, $I_s(t)$ denotes the number of symptomatic infected individuals, $I_a(t)$ denotes

Table 1. Parameters of the cholera model

Parameters	Description
b	Birth rate or Recruitment rate
β	Rate of exposure to contaminated water
κ	Concentration of vibrio cholerae in water
α_1	rate of shedding bacteria into the environment from $I_s(t)$
α_2	rate of shedding bacteria into the environment from $I_a(t)$
d	Disease induced death rate
δ	Death rate of vibrio cholerae due to water treatment
ϕ	Mortality rate for bacteria including phage degradation
r_1	Recovery rate of symptomatic infected individuals
r_2	Recovery rate of asymptomatic infected individuals
p	Prob. of new infected from $S(t)$ to be symptomatic
q	Prob. of new infected from $S(t)$ to be asymptomatic
μ	Natural death rate

the number of asymptomatic infected individuals, $R(t)$ denotes the number of recovered individuals and $B(t)$ denotes the pathogen concentration vibrio cholerae in environment at time t . The parameters used in the model formulation are summarized in Table 1. From the description of the dynamics of cholera as shown in Figure 1, we have the following system of nonlinear ordinary differential equations (Marwa et al, 2018).

$$\begin{aligned}
 \frac{dS(t)}{dt} &= bN - \frac{\beta B(t)S(t)}{\kappa + B(t)} - \mu S(t), \\
 \frac{dI_s(t)}{dt} &= \frac{p\beta B(t)S(t)}{\kappa + B(t)} - (\mu + r_1 + d)I_s(t), \\
 \frac{dI_a(t)}{dt} &= \frac{q\beta B(t)S(t)}{\kappa + B(t)} - (\mu + r_2)I_a(t), \\
 \frac{dR(t)}{dt} &= r_1I_s(t) + r_2I_a(t) - \mu R(t), \\
 \frac{dB(t)}{dt} &= \alpha_1I_s(t) + \alpha_2I_a(t) - (\delta + \phi)B(t),
 \end{aligned} \tag{2.1}$$

2.1 Reproduction Number, R_0

The key parameter in epidemiology is the basic reproduction number R_0 , which is defined as the average number of secondary infectious cases transmitted by a single primary infectious cases introduced into a whole susceptible population (Cohn et al, 1991). This parameter is useful because it helps determine whether an infectious disease will spread within the population or not. To compute R_0 , the next generation matrix approach is used as in (Van den Driessche & Watmough, 2002). It is obtained by taking the largest (dominant) eigenvalue value (spectral radius) of

$$\left[\frac{\partial \mathcal{F}_i(E_0)}{\partial \mathcal{X}_i} \right] \left[\frac{\partial \mathcal{V}_i(E_0)}{\partial \mathcal{X}_i} \right]^{-1},$$

where \mathcal{F}_i is the rate of appearance of new infection in compartment i , \mathcal{V}_i is the net transition between compartments, E_0 is the disease free equilibrium and \mathcal{X}_i stand for the terms in which the infection is in progression i.e., $I_s(t)$, $I_a(t)$ and $B(t)$ in the model (2.1). Hence, the basic reproduction number obtained in (Marwa et al, 2018) was as follows:

$$R_0 = \left[\frac{\alpha_1\beta\mu p + \alpha_2\beta q r_1 + \alpha_1\beta p r_2 + \alpha_2\beta d q + \alpha_2\beta\mu q}{D} \right] \frac{bN}{\mu}, \tag{2.3}$$

where

$$\begin{aligned}
 D &= \delta\kappa\mu d + \delta\kappa\mu^2 + \delta\kappa\mu r_1 + \delta\kappa d r_2 + \delta\kappa\mu r_2 + \delta\kappa r_1 r_2 + \delta\kappa\mu\phi + \phi\kappa\mu^2 \\
 &+ \kappa\mu\phi r_1 + \kappa\phi d r_2 + \kappa\mu\phi r_2 + \kappa\phi r_1 r_2.
 \end{aligned}$$

3. Cholera Continuous Time Markov Chain Model

From the deterministic model (2.1), a cholera continuous time Markov chain (CTMC) model is formulated. The cholera continuous time Markov chain is a time-homogeneous process with the Markov property. By considering the spread

of epidemic disease, clearly it is more realistic to keep track of where the epidemic disease is at minor, large outbreak or endemic at any continuous time $t \geq 0$ as opposed to only where the epidemic disease is after n steps under various conditions.

The cholera CTMC model gives an indication to the stochastic variability inherent in birth, death, and transmission which are tracked in the deterministic model. However, cholera CTMC model provide enough information for the given population dynamics when the population sizes are small. Also, these birth, death, transmission and other transitions are expressed in terms of probabilities.

The cholera CTMC epidemic processes are defined on a continuous time scale, $t \in [0, \infty)$, but the states

$S(t), I_s(t), I_a(t), R(t) \in \{0, 1, 2, 3, \dots, N\}$ are discrete random variables. Also, in cholera deterministic model there is $B(t)$ that denotes the concentration of vibrio cholerae in the environment at time t . However, $B(t)$ is not a compartment occupancy as the other one are so its transition is not considered in the formulation of transitions. Now suppose that, the population is homogeneously mixed so that each susceptible individual has equal chance of acquiring cholera through consuming water which are contaminated with vibrio cholerae in the reservoir or environment $B(t)$. The susceptible individuals change their classification and belong to class $I_a(t)$ or $I_s(t)$. Class $I_a(t)$ and $I_s(t)$ change to $R(t)$, where $I_a(t)$, $I_s(t)$ will remain indefinitely, independent of the past. The length of time spent in state $I_a(t)$, $I_s(t)$ or $R(t)$ is continuously and strictly positive.

In order to develop the transition probability matrix $p_s(u)$ of possible states for cholera CTMC model, we denote the change for each variable during time Δt as

$$\begin{aligned}\Delta S(t) &= S(t + \Delta t) - S(t), & \Delta I_s(t) &= I_s(t + \Delta t) - I_s(t), & \Delta I_a(t) &= I_a(t + \Delta t) - I_a(t), \\ \Delta R(t) &= R(t + \Delta t) - R(t).\end{aligned}$$

The lower case (s, i_s, i_a, r) denotes the values of the discrete random variables from the set $\{0, 1, 2, 3, \dots, N\}$. The transition probabilities together with stochastic process are defined for a small time interval of length $\Delta t > 0$ as

$p_s((s, i_s, i_a, r), (s + e_1, i_s + e_2, i_a + e_3, r + e_4))(\Delta t) = (s + e_1, i_s + e_2, i_a + e_3, r + e_4) | (S(t), I_s(t), I_a(t), R(t)) = (s, i_s, i_a, r)$. The transition probabilities depend on the time between events Δt but not on the specific time t , known as a time-homogeneous process. However, when given the current state of the process at time t , the future state of the process at time $t + \Delta t$, for any $\Delta t > 0$, does not depend on time prior to t , this is termed as Markov property (Allen, 2017). Table 2, shows the changes, $\Delta S(t) = S(t + \Delta t) - S(t)$, $\Delta I_s(t) = I_s(t + \Delta t) - I_s(t)$, $\Delta I_a(t) = I_a(t + \Delta t) - I_a(t)$, $\Delta R(t) = R(t + \Delta t) - R(t)$, associated with four events, birth, death, infection and recovery.

Table 2. $S I_s I_a R - B$ CTMC model assumptions

change	probability	description
$S \rightarrow S + 1$	$bN\Delta t + o(\Delta t)$	birth of susceptible
$S \rightarrow S - 1$	$\mu S \Delta t + o(\Delta t)$	death of susceptible people
$S \rightarrow S - 1, I_s \rightarrow I_s + 1$	$\frac{p\beta BS}{\kappa + B} \Delta t + o(\Delta t)$	new infection resulted by I_s due excretion
$S \rightarrow S - 1, I_a \rightarrow I_a + 1$	$\frac{q\beta BS}{\kappa + B} \Delta t + o(\Delta t)$	new infection resulted by I_a due excretion
$I_s \rightarrow I_s - 1$	$dI_s \Delta t + o(\Delta t)$	death of I_s due to disease
$I_s \rightarrow I_s - 1$	$\mu I_s \Delta t + o(\Delta t)$	natural death of symptomatic people I_s
$I_a \rightarrow I_a - 1$	$\mu I_a \Delta t + o(\Delta t)$	natural death of asymptomatic people I_a
$R \rightarrow R - 1$	$\mu R \Delta t + o(\Delta t)$	natural death of recovered people
$I_s \rightarrow I_s - 1, R \rightarrow R + 1$	$r_1 I_s \Delta t + o(\Delta t)$	rate of recovery of I_s
$I_a \rightarrow I_a - 1, R \rightarrow R + 1$	$r_2 I_a \Delta t + o(\Delta t)$	rate of recovery of I_a

Basing on the comparison purposes, the transition probabilities are defined in terms of the rates in the cholera model (2.1)

as follows.

$$p_s((s, i_s, i_a, r), (s + e_1, i_s + e_2, i_a + e_3, r + e_4))(\Delta t) = \begin{cases} bN\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (1, 0, 0, 0) \\ \frac{p\beta B(t)s(t)}{\kappa + B(t)}\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (-1, 1, 0, 0) \\ \frac{q\beta B(t)s(t)}{\kappa + B(t)}\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (-1, 0, 1, 0) \\ di_s(t)\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (0, -1, 0, 0) \\ \mu s(t)\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (-1, 0, 0, 0) \\ r_1 i_s(t)\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (0, -1, 0, 1) \\ \mu i_s(t)\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (0, -1, 0, 0) \\ \mu i_a(t)\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (0, 0, -1, 0) \\ r_2 i_a(t)\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (0, 0, -1, 1) \\ \mu r(t)\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (0, 0, 0, -1) \\ 1 - \Omega\Delta t + o(\Delta t), & \text{if } (e_1, e_2, e_3, e_4) = (0, 0, 0, 0) \\ o(\Delta t), & \text{otherwise} \end{cases} \quad (2.4)$$

where

$$\Omega = bN + \frac{p\beta B(t)s(t)}{\kappa + B(t)} + \frac{q\beta B(t)s(t)}{\kappa + B(t)} + di_s(t) + \mu s(t) + r_1 i_s(t) + \mu i_s(t) + \mu i_a(t) + r_2 i_a(t) + \mu r(t).$$

But $p_s(s, i_s, i_a, r)(t + \Delta t)$ can be expressed in terms of probabilities at time t as follows.

$$\begin{aligned} p_s(s, i_s, i_a, r)(t + \Delta t) &= p_s(s - 1, i_s, i_a, r)(t)bN\Delta t \\ &+ p_s(s + 1, i_s - 1, i_a, r)(t)\frac{(p - 1)\beta B(s + 1)}{\kappa + B}\Delta t \\ &+ p_s(s + 1, i_s, i_a - 1, r)(t)\frac{(q - 1)\beta B(s + 1)}{\kappa + B}\Delta t \\ &+ p_s(s, i_s + 1, i_a, r)(t)d(i_s + 1)\Delta t \\ &+ p_s(s + 1, i_s, i_a, r)(t)\mu(s + 1)\Delta t \\ &+ p_s(s, i_s + 1, i_a, r - 1)(t)(r_1 - 1)(i_s + 1)\Delta t \\ &+ p_s(s, i_s + 1, i_a, r)(t)\mu(i_s + 1)\Delta t \\ &+ p_s(s, i_s, i_a + 1, r)(t)\mu(i_a + 1)\Delta t \\ &+ p_s(s, i_s, i_a + 1, r - 1)(t)(r_2 - 1)(i_a + 1)\Delta t \\ &+ p_s(s, i_s, i_a, r + 1)(t)(r + 1)\Delta t \\ &+ p_s(s, i_s, i_a, r)(t)(1 - \Omega)\Delta t + o(\Delta t). \end{aligned} \quad (2.5)$$

3.1 Kolmogorov Differential Equations

From Equation (2.5), we can derive differential equations for the transition probabilities. These differential equations can be referred to as the backward or forward Kolmogorov differential equations (Allen, 2017). The forward Kolmogorov differential equations, are also known as the Master equations, which are used to predict the future dynamics (Allen, 2017). Therefore, from Equation (2.5), we compute $\frac{p_s(s, i_s, i_a, r)(t + \Delta t) - p_s(s, i_s, i_a, r)(t)}{\Delta t}$, and letting $\Delta t \rightarrow 0$, leads

to the system of forward Kolmogorov differential equations:

$$\begin{aligned}
 \frac{dp_s(s, i_s, i_a, r)(t)}{dt} = & p_s(s-1, i_s, i_a, r)(t) bN \\
 & + p_s(s+1, i_s-1, i_a, r)(t) \frac{(p-1)\beta B(s+1)}{\kappa+B} \\
 & + p_s(s+1, i_s, i_a-1, r)(t) \frac{(q-1)\beta B(s+1)}{\kappa+B} \\
 & + p_s(s, i_s+1, i_a, r)(t) d(i_s+1) \\
 & + p_s(s+1, i_s, i_a, r)(t) \mu(s+1) \\
 & + p_s(s, i_s+1, i_a, r-1)(t) (r_1-1)(i_s+1) \\
 & + p_s(s, i_s+1, i_a, r)(t) \mu(i_s+1) \\
 & + p_s(s, i_s, i_a+1, r)(t) \mu(i_a+1) \\
 & + p_s(s, i_s, i_a+1, r-1)(t) (r_2-1)(i_a+1) \\
 & + p_s(s, i_s, i_a, r+1)(t) (r+1) \\
 & + p_s(s, i_s, i_a, r)(t) (1-\Omega) \\
 & - p_s(s, i_s, i_a, r)(t) \Omega.
 \end{aligned} \tag{2.6}$$

It should be noted that both b and β may be constant or time dependent. Then by assuming that T_j is the time of the j^{th} jump, then the time of the $(j+1)^{th}$ jump is,

$$T_{j+1} = T_j + W_j,$$

where W_j is the interevent time. Now, for $x(T_j) = (e_1, e_2, e_3, e_4)^T$, and $n = e_1 + e_2 + e_3 + e_4$, we get the interevent time as

$$W_j = -\frac{\ln(U)}{\alpha(e_1, e_2, e_3, e_4)}, \tag{2.7}$$

where U is a uniformly distributed random variable on the interval $[0,1]$ (Allen, 2010),

$$\alpha(e_1, e_2, e_3, e_4) = bn + \mu e_1 + \frac{p\beta B e_1}{\kappa+B} + \frac{q\beta B e_1}{\kappa+B} + \mu e_2 + r_1 e_2 + d e_2 + \mu e_3 + r_2 e_3 + \mu e_4. \tag{2.8}$$

and $\frac{1}{\alpha(e_1, e_2, e_3, e_4)}$, is the mean for the exponentially distributed random variable W_j .

4. Multitype Branching Process

The multitype branching process helps to approximate the dynamics of disease of nonlinear continuous time Markov chain model near the disease free equilibrium (dorman et al, 2004). In this paper, we use the idea of multitype branching process to determine the stochastic threshold.

4.1 Branching Process Approximations

In this paper, we also study the stochastic behaviour near the disease free equilibrium. This helps to know whether an epidemic lead to major outbreak or minor outbreak for cholera CTMC model when few infectious individuals are introduced into the population. This is possible by using branching process theory and other techniques from the probability generating functions. It should be noted that when states $I_s(t)$, $I_a(t)$ and $B(t)$ reaches an absorbing states, then $I_s(t) = I_a(t) = 0$, and $B(t) = 0$. This implies that, the disease transmission stops (Allen, 2017).

However, the branching process is termed as the linear approximation of the SIR stochastic process when near the disease free equilibrium. Hence, when there is few initial infectious individuals, the branching process tend to hit zero or grows exponentially (Allen, 2017). The main two features which are captured in the branching process approximation for the continuous time Markov chain model near the disease free equilibrium are:

1. The major outbreak occurs when infectious individuals increases to a large number of cases and
2. The minor outbreak occurs when there is few additional cases above the initial number of cases.

Furthermore, it should be noted that the branching process is the best approximation for continuous time Markov chain model when the number of susceptible individuals are large (Allen, 2017).

The branching process is a birth and death process for $I_s(t)$, $I_a(t)$ and $B(t)$, therefore the variable $R(t)$ will not be considered in the branching process approximation and the susceptible individuals are assumed to be at the disease free state as follows

$$S^f = \frac{bN}{\mu}.$$

Both infectious individuals $I_s(t)$ and $I_a(t)$ can produce secondary infectious hosts through shedding vibrio cholerae to the environment and hence, susceptible individuals will be at high risk of getting infection. Consider the cases; $r_1 I_s(t)$ be recovery rate of symptomatic infected individuals, $d I_s(t)$ death rate of symptomatic infected individuals due to disease, $r_2 I_a(t)$ recovery rate of asymptomatic infected individuals, and bN the birth rate or recruitment rate. The process start with few infectious individuals (i.e., symptomatic and asymptomatic infected individuals).

The following are the three important assumptions to underlie the branching process approximation (Allen, 2017).

1. Each infectious individual behaviour is independent from other infectious individual,
2. Each infectious individual has the same probability of transmitting an infection and same probability of recovery,
3. The susceptible population is very large.

From the assumptions with see that assumption (1) is feasible if a number of infectious individuals introduced into a large homogeneously mixed population is small. Where as assumption (2) is said to be feasible in a homogeneously mixed population with constant transmission and recovery rates, β , r_1 and r_2 (Allen, 2017).

4.2 Probability of Extinction

The probability of extinction can be studied by using two probability generating functions (pgfs).

1. The first probability generating function applies to each infectious individual known as the offspring pgf, and the second one
2. applies to the entire infectious class $I_s(t)$ and $I_a(t)$ at time t .

The term offspring, in this paper will help to describe susceptible individuals that become infectious due to indirect transmission, that is through the focal-oral route of contaminated food or water caused by poor sanitation (World Health Organization, 2010). The offspring pgf has the form:

$$f(s) = \sum_{k=0}^{\infty} p_k s^k, s \in [0, 1], \quad (2.9)$$

where p_k is the probability of one individual generating k new individuals of the same type i.e., one infectious individual generate k infectious individuals and s is the extinction probability.

The probability generating function has some properties that are useful in the analysis, when $f(1) = 1$, then the mean number of offspring generated from one individual is defined as

$$f'(1) = \sum_{k=1}^{\infty} k p_k.$$

Referring to the nature of our model, the approximation of the continuous time Markov chain near the disease free equilibrium leads to the multi type branching process in the three dependent variables $I_s(t)$, $I_a(t)$ and $B(t)$. The three offspring pgfs are defined; two for infected individuals and one for vibrio cholerae concentration in the environment, where each offspring has the general pgf form:

$$\sum_{j,l} p_s((I_s, B) = (j, l)) s_1^j s_2^l, s_i \in [0, 1], i = 1, 2.$$

For each infected individual we have two events, that is infection and recovery. Therefore, the offspring pgf for $I_s(t)$ i.e., $I_s(0) = 1$ and $B(0) = 0$ is given by

$$f_1(s_1, s_2) = \frac{r_1}{r_1 + \alpha_1} + \left(\frac{\alpha_1}{\alpha_1 + r_1} \right) s_1 s_2, \quad (3.0)$$

From Equation (3.0), each symptomatic infectious individuals recovers with the probability $\frac{r_1}{r_1 + \alpha_1}$ and each symptomatic infected host shed vibrio cholerae on the environment with the probability $\left(\frac{\alpha_1}{\alpha_1 + r_1}\right)$, where the term s_1 implies that each symptomatic infected individual generates an individual with vibrio cholerae (s_2 raised to the power of one) and they remains infectious (s_1, s_2 raised to the power one).

However, the offspring pgf for $B(t)$ i.e., $B(0) = 1$ and $I_s(0) = 0$ is given by

$$f_2(s_1, s_2) = \frac{\phi + \delta}{\phi + \delta + \lambda_1 + \alpha_1} + \left(\frac{\lambda_1 + \alpha_1}{\phi + \delta + \lambda_1 + \alpha_1}\right) s_1 s_2, \quad (3.1)$$

From Equation (3.1), one vibrio cholerae dies with the probability $\frac{\phi + \delta}{\phi + \delta + \lambda_1 + \alpha_1}$ due to water treatment or natural death. Also infects symptomatic individuals with the probability

$$\frac{\lambda_1 + \alpha_1}{\phi + \delta + \lambda_1 + \alpha_1},$$

where

$$\lambda_1 = \frac{p\beta B}{\kappa + B}.$$

To find the probability of extinction (no outbreak), we compute the fixed points of the system $(w_1, w_2) \in [0, 1]$ of the offspring pgfs for the infectious classes, that is, we solve $f_1(s_1, s_2) = s_1$ and $f_2(s_1, s_2) = s_2$, e.g., (Allen & Van den Driessche, 2013). The two solutions for this system are $(1, 1)$ and w_1, w_2 ,

where

$$w_1 = \frac{r_1}{r_1 + \alpha_1} + \left(\frac{\alpha_1}{\alpha_1 + r_1}\right) \left(\frac{1}{R_{01}}\right),$$

$$w_2 = \frac{\phi + \delta}{\phi + \delta + \lambda_1 + \alpha_1} + \left(\frac{\lambda_1 + \alpha_1}{\phi + \delta + \lambda_1 + \alpha_1}\right) \left(\frac{1}{R_{01}}\right),$$

and

$$R_{01} = \frac{p\beta\alpha_1 bN}{\kappa\phi\mu^2 + \kappa\phi d\mu + \kappa\phi\mu r_1 + \kappa\delta\mu^2 + \kappa\delta d\mu + \kappa\mu\delta r_1}.$$

For the case of asymptomatic infected $I_a(t)$, To find the probability of extinction (no outbreak), we compute the fixed points of the system $(w_1, w_2) \in [0, 1]$ of the offspring pgfs for the infectious classes, that is, we solve $f_3(s_3, s_4) = s_3, f_4(s_3, s_4) = s_4$, e.g., (Allen & Van den Driessche, 2013). We have the offspring pgf for $I_a(t)$ i.e., $I_a(0) = 1$ and $B(0) = 0$.

Hence

$$f_3(s_3, s_4) = \frac{r_2}{r_2 + \alpha_2} + \left(\frac{\alpha_2}{\alpha_2 + r_2}\right) s_3 s_4, \quad (3.2)$$

In Equation (3.2), each asymptomatic infectious individuals recovers with the probability $\frac{r_2}{r_2 + \alpha_2}$ and each asymptomatic infected shed vibrio cholerae on the environment with the probability $\left(\frac{\alpha_2}{\alpha_2 + r_2}\right)$, where the term s_3 implies that each asymptomatic infected individual generates an individual with vibrio cholerae (s_4 raised to the power of one) and they remains infectious (s_3, s_4 raised to the power one).

However, the offspring pgf for $B(t)$, given $B(0) = 1$ and $I_a(0) = 0$. We have

$$f_4(s_3, s_4) = \frac{\phi + \delta}{\phi + \delta + \lambda_2 + \alpha_2} + \left(\frac{\lambda_2 + \alpha_2}{\phi + \delta + \lambda_2 + \alpha_2}\right) s_3 s_4, \quad (3.3)$$

In Equation (3.3), one vibrio cholerae dies with the probability $\frac{\phi + \delta}{\phi + \delta + \lambda_2 + \alpha_2}$ due to water treatment or natural death. Also infects individuals with the probability

$$\frac{\lambda_2 + \alpha_2}{\phi + \delta + \lambda_2 + \alpha_2},$$

where

$$\lambda_2 = \frac{q\beta B}{\kappa + B}.$$

The two solutions for this system are $(1, 1)$ and w_3, w_4 , where

$$w_3 = \frac{r_2}{r_2 + \alpha_2} + \left(\frac{\alpha_2}{\alpha_2 + r_2} \right) \left(\frac{1}{R_{02}} \right),$$

$$w_4 = \frac{\phi + \delta}{\phi + \delta + \lambda_2 + \alpha_2} + \left(\frac{\lambda_2 + \alpha_2}{\phi + \delta + \lambda_2 + \alpha_2} \right) \left(\frac{1}{R_{02}} \right),$$

and

$$R_{02} = \frac{q\beta\alpha_2bN}{\kappa\delta\mu^2 + \kappa\delta r_2\mu + \kappa\delta\mu\phi + \kappa\phi\mu^2 + \kappa\mu\phi r_2}.$$

It should be noted that, the fixed point (w_1, w_2, w_3, w_4) exists in the unit square provided that $R_{01}, R_{02} > 1$. Also, w_1 and w_3 have the biological interpretation, that is from the beginning of one infectious hosts, there is no outbreak if the infectious individuals recovers with the probabilities $\frac{r_1}{r_1 + \alpha_1}$ and $\frac{r_2}{r_2 + \alpha_2}$ or there is no successful transmission to susceptible cholera with probability $\left(\frac{\lambda_1 + \alpha_1}{\phi + \delta + \lambda_1 + \alpha_1} \right) \left(\frac{1}{R_{01}} \right)$ and $\left(\frac{\lambda_2 + \alpha_2}{\phi + \delta + \lambda_2 + \alpha_2} \right) \left(\frac{1}{R_{02}} \right)$ respectively.

Then, the probability of no outbreak is 1, if $R_{01} < 1$ and $R_{02} < 1$ respectively. Also less than 1, if $R_{01} > 1$ and $R_{02} > 1$ respectively. Given that $I_s(0) = i_s$, $I_a(0) = i_a$, and $B(0) = \bar{b}$. From the independent assumptions of branching process approximation, the probability of major and minor outbreak for symptomatic infected individuals are given by

$$\mathbb{P}_{\text{minor outbreak}} = \begin{cases} w_1^{i_s} w_2^{\bar{b}}, & R_{01} > 1 \\ 1, & R_{01} < 1, \end{cases}$$

and

$$\mathbb{P}_{\text{major outbreak}} = \begin{cases} (1 - w_1^{i_s} w_2^{\bar{b}}), & R_{01} > 1 \\ 0, & R_{01} < 1, \end{cases}$$

where

$$i_s = I_s.$$

Also, for asymptomatic infected individuals we have

$$\mathbb{P}_{\text{minor outbreak}} = \begin{cases} w_3^{i_a} w_4^{\bar{b}}, & R_{02} > 1 \\ 1, & R_{02} < 1, \end{cases}$$

and

$$\mathbb{P}_{\text{major outbreak}} = \begin{cases} (1 - w_3^{i_a} w_4^{\bar{b}}), & R_{02} > 1 \\ 0, & R_{02} < 1, \end{cases}$$

where

$$i_a = I_a.$$

The branching process theory when applied to infectious disease provide a clear method for the prediction of disease outbreaks. Whittle (1955) as one of the investigator of this theory showed that, the basic reproduction number R_0 , computed from the deterministic model yields the same threshold as the stochastic Markov chain SIR model, when given an initial infectious individuals i related to the population size N . Whittle (1955) also, stated that when $R_0 < 1$, it means no outbreak, while when $R_0 > 1$ there is an outbreak. Thus, the probability of a minor outbreak is $\left(\frac{1}{R_0} \right)^i$ and the probability of major outbreak is $1 - \left(\frac{1}{R_0} \right)^i$. These estimates for extinction are more accurate for a large susceptible population size N and when there is few infectious individuals i . Thus

$$\mathbb{P}_{\text{minor outbreak}} = \begin{cases} \left(\frac{1}{R_0} \right)^i, & R_0 > 1 \\ 1, & R_0 < 1 \end{cases}$$

and

$$\mathbb{P}_{\text{major outbreak}} = \begin{cases} \left(1 - \frac{1}{R_0} \right)^i, & R_0 > 1 \\ 0, & R_0 < 1 \end{cases}$$

where

$$i = i_s + i_a.$$

5. Numerical simulation and Discussion

The model formulated in this paper involves multivariate processes, so it is difficult to find analytical solution for the transition probabilities from the forward Kolmogorov differential equation. Therefore, a numerical method for simulating the sample path of CTMC model formulated in this paper is to use Gillespie algorithm. This method was proposed by Gillespie (Gillespie & Marshall, 1977). Also, in order to numerically simulate the change in state, we need to have this uniform random numbers, $u_i \in U[0, 1]$ which are required for each change. Here u_i stands for the particular interevent time and event. While i takes the rates for birth, death and movement between compartments.

Furthermore, to numerically compute the sample path of cholera continuous time Markov chain (CTMC) model, we need to employ a fact that interevent time has an exponential distribution. This is deduced from the Markov property, due to this an exponential distribution has a property called memoryless (Allen, 2008). The formulation of cholera CTMC model from the deterministic model (2.1) in this study is based on three rates: birth, death and movement between compartments. The birth, death and movement rates are defined as follows:

$i \rightarrow i + 1$ at a rate of $\lambda_b(i)$, $i \rightarrow i - 1$ at a rate of $\lambda_d(i)$, $k \rightarrow k - 1$ is the leaving compartment at a rate of $\lambda_{m1}(k)$ and $k \rightarrow k + 1$ is the entering compartment at a rate of $\lambda_{m2}(k)$. The rates $\lambda_b(i)$, $\lambda_d(i)$, $\lambda_{m1}(k)$ and $\lambda_{m2}(k)$ denotes birth, death and movement of individuals respectively. The total rate λ is given by

$$\lambda = \lambda_b(i) + \lambda_d(i) + \lambda_{m1}(k) + \lambda_{m2}(k). \quad (3.4)$$

When an event occurs, it will be a birth with probability

$$p\{i \rightarrow i + 1\} = \frac{\lambda_b(i)}{\lambda_b(i) + \lambda_d(i) + \lambda_{m1}(k) + \lambda_{m2}(k)}. \quad (3.5)$$

While an event dies with probability

$$p\{i \rightarrow i - 1\} = \frac{\lambda_d(i)}{\lambda_b(i) + \lambda_d(i) + \lambda_{m1}(k) + \lambda_{m2}(k)}. \quad (3.6)$$

The movement between individuals occur with probability

$$p\{k \rightarrow k - 1\} = \frac{\lambda_{m1}(k)}{\lambda_b(i) + \lambda_d(i) + \lambda_{m1}(k) + \lambda_{m2}(k)}. \quad (3.7)$$

and

$$p\{k \rightarrow k + 1\} = \frac{\lambda_{m2}(k)}{\lambda_b(i) + \lambda_d(i) + \lambda_{m1}(k) + \lambda_{m2}(k)}, \quad (3.8)$$

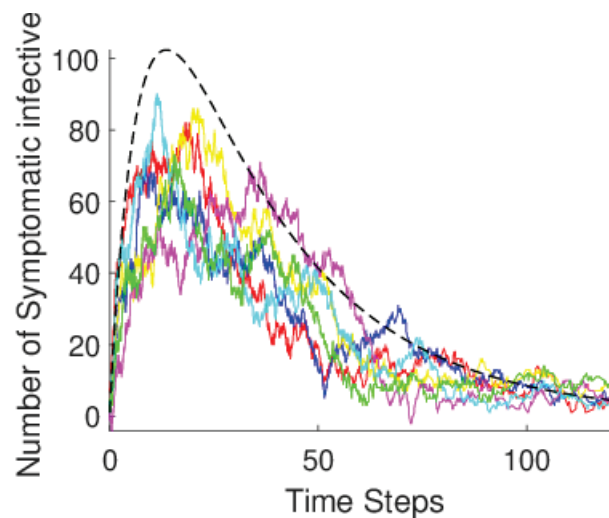
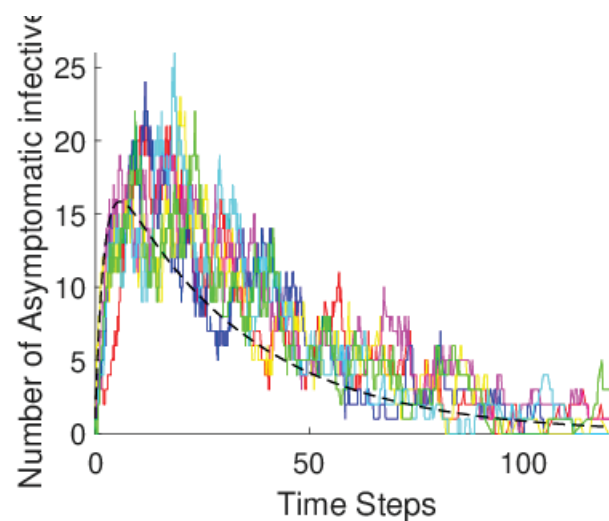
Also, we apply a uniform random number on $[0,1]$ for numerical computation of the interevent time. Now, let U be a uniform random variable on $[0,1]$. Then the interevent time $T_{i,k}$ satisfies

$$T_{i,k} = -\frac{\ln(U)}{\lambda_b(i) + \lambda_d(i) + \lambda_{m1}(k) + \lambda_{m2}(k)}. \quad (3.9)$$

From the idea of Equations (3.4) to (3.9), four sample paths of the $SI_s I_a R - B$ continuous time Markov chain model are graphed in Figures 2 to 5. Also, we numerically solve the ordinary differential equations (ODEs) model (2.1) using the fourth order Runge-Kutta method. A stable endemic equilibrium exists in the ordinary differential equations (ODEs) model if $R_0 > 1$ and the sample paths of $SI_s I_a R - B$ continuous time Markov chain model fluctuates around this endemic equilibrium. The parameter values used in plotting the sample path of Figures 2 to 5 are described in Table 3, with the initial conditions $S(0) = 997$, $I_s(0) = 1$, $I_a(0) = 1$, $R(0) = 1$, $N = 1000$, $B = 1$, and $\Delta t = 0.01$. The results show that the sample path of cholera continuous time Markov chain (CTMC) model fluctuates within the solution of ordinary differential equations as seen in Figures 2 to 5.

Table 3. Parameter values for cholera epidemic as used in (Marwa et al, 2018)

Parameters	Value	Source
b	0.000072	(Akman et al, 2016)
β	0.35	(Njagarah & Nyabdza, 2015)
κ	1000000	(Njagarah & Nyabdza, 2015)
α_1	1.5	(Akman et al, 2016)
α_2	0.1	(Akman et al, 2016)
d	0.00065	(Neilan et al, 2010)
δ	0.05	(Marwa et al, 2018)
ϕ	0.025	(Akman et al, 2016)
r_1	0.14	(Akman et al, 2016)
r_2	0.5	(Akman et al, 2016)
p	0.7	(Marwa et al, 2018)
q	0.3	(Marwa et al, 2018)
μ	0.000044	(Akman et al, 2016), Khan et al, 2015)

Figure 2. The sample path of the symptomatic infected individuals for CTMC model is graphed with the deterministic solution (*dashed curve*). The values are $\Delta t = 0.01$, $I_s(0) = 1$, $N = 1000$ Figure 3. The sample path of asymptomatic infected individuals for CTMC model is graphed with the deterministic solution (*dashed curve*). The values are $\Delta t = 0.01$, $I_a(0) = 1$, $N = 1000$

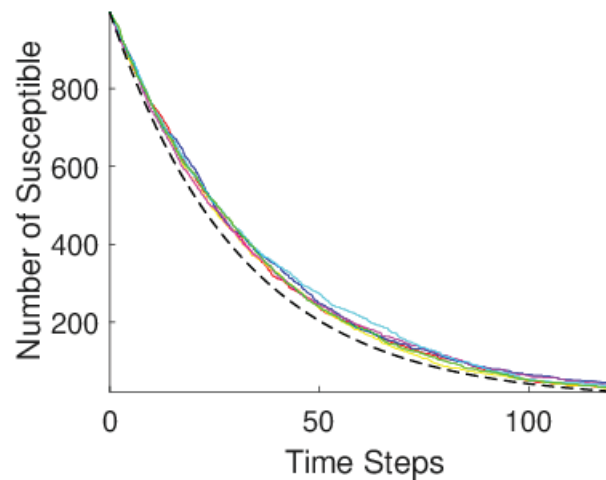


Figure 4. The sample path of the susceptible individuals for CTMC model is graphed with the deterministic solution (dashed curve). The values are $\Delta t = 0.01$, $S(0) = 997$, $N = 1000$

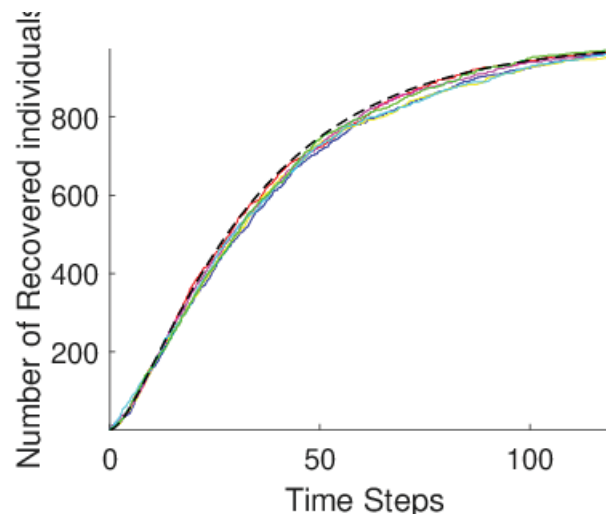


Figure 5. The sample path of the recovered individuals for CTMC model is graphed with the deterministic solution (dashed curve). The values are $\Delta t = 0.01$, $R(0) = 1$, $N = 1000$

6. Conclusion

The intent of this paper is to formulate a cholera continuous time Markov chain model from the deterministic model. Many papers of relevance to stochastic epidemic modelling omitted continuous time Markov chain. This paper provide enough information on cholera CTMC model and other topics in CTMC epidemic modelling. We apply the theory of multitype branching processes in order to estimate the probability of disease extinction or major outbreak. Also, we illustrate the effects of infectious individuals shedding vibrio cholerae into the environment and the environmental decontamination on the probability of disease extinction or major outbreak. From this study we find that, the probability of extinction (no outbreak) is 1, if $R_0 < 1$, but less than 1, if $R_0 > 1$ and when $I_s = I_a = 0$, and $B = 0$ the epidemic stops. This implies that an absorbing state has been reached and the process stops. However, numerical simulation is carried using Gillespie algorithm and Runge–Kutta method to simulate the sample path of $SI_sI_aR - B$ continuous time Markov chain model and the solution of ordinary differential equation respectively. We find that, the sample path of $SI_sI_aR - B$ continuous time Markov chain model fluctuates within the solution of the $SI_sI_aR - B$ ordinary differential equation model.

Acknowledgements

The authors wish to thank Pan African University and Tanzania Public Service College for their financial support and the reviewers for the insightful comments that made this manuscript better.

References

- Akman, O., Corby, M. R., & Schaefer, E. (2016). Examination of models for cholera: insights into model comparison methods. *Letters in Biomathematics*, 3(1), 93-118. <https://doi.org/10.1080/23737867.2016.1211495>
- Allen, L. J. (2017). A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. *Infectious Disease Modelling*, 2(2), 128-142. <https://doi.org/10.1016/j.idm.2017.03.001>
- Allen, L. J., & van den Driessche, P. (2013). Relations between deterministic and stochastic thresholds for disease extinction in continuous-and discrete-time infectious disease models. *Mathematical biosciences*, 243(1), 99-108. <https://doi.org/10.1016/j.mbs.2013.02.006>
- Allen, L. J. (2010). *An introduction to stochastic processes with applications to biology*. Chapman and Hall/CRC.
- Allen, L. J. (2008). An introduction to stochastic epidemic models. In *Mathematical epidemiology*, 81-130. Springer, Berlin, Heidelberg.
- Allen, L. J., & Lahodny Jr, G. E. (2012). Extinction thresholds in deterministic and stochastic epidemic models. *Journal of Biological Dynamics*, 6(2), 590-611. https://doi.org/10.1007/978-3-540-78911-6_3
- Ali, M., Lopez, A. L., You, Y., Kim, Y. E., Sah, B., Maskery, B., & Clemens, J. (2012). The global burden of cholera. *Bulletin of the World Health Organization*, 90, 209-218. <https://doi.org/10.2471/BLT.11.093427>
- Berriman, A. D., Clancy, D., Clough, H. E., & Christley, R. M. (2013). Semi-stochastic models for Salmonella infection within finishing pig units in the UK. *Mathematical biosciences*, 245(2), 148-156. <https://doi.org/10.1016/j.mbs.2013.06.004>
- Castillo-Chavez, C., Feng, Z., & Huang, W. (2002). On the computation of R_0 and its role on. *Mathematical approaches for emerging and reemerging infectious diseases: an introduction*, 1, 229.
- Cohn, J. N., Johnson, G., Ziesche, S., Cobb, F., Francis, G., Tristani, F., ... & Bhat, G. (1991). A comparison of enalapril with hydralazine/isosorbide dinitrate in the treatment of chronic congestive heart failure. *New England Journal of Medicine*, 325(5), 303-310. <https://doi.org/10.1056/NEJM199108013250502>
- Dimi, J. L., Bissila, J. F., & Mbaya, T. (2016) *Some Stochastic Properties of Cholera Model*.
- Dorman, K. S., Sinsheimer, J. S., & Lange, K. (2004). In the garden of branching processes. *SIAM review*, 46(2), 202-229. <https://doi.org/10.1137/S0036144502417843>
- Edward, S., & Nyerere, N. (2015). A mathematical model for the dynamics of cholera with control measures. *Applied and Computational Mathematics*, 4(2), 53-63. <https://doi.org/10.11648/j.acm.20150402.14>
- Eisenberg, M. C., Kujbida, G., Tuite, A. R., Fisman, D. N., & Tien, J. H. (2013). Examining rainfall and cholera dynamics in Haiti using statistical and dynamic modeling approaches. *Epidemics*, 5(4), 197-207. <https://doi.org/10.1016/j.epidem.2013.09.004>
- Gelman, A., & Rubin, D. B. (1996). Markov chain Monte Carlo methods in biostatistics. *Statistical methods in medical research*, 5(4), 339-355. <https://doi.org/10.1177/096228029600500402>
- Gillespie, J. M., & Marshall, R. C. (1977). Proteins of the hard keratins of echidna, hedgehog, rabbit, ox and man. *Australian journal of biological sciences*, 30(5), 401-410. <https://doi.org/10.1071/BI9770401>
- Hill, D. R., Ford, L., & Lalloo, D. G. (2006). Oral cholera vaccines: use in clinical practice. *The Lancet infectious diseases*, 6(6), 361-373. [https://doi.org/10.1016/S1473-3099\(06\)70494-7](https://doi.org/10.1016/S1473-3099(06)70494-7)
- Khan, M. A., Ali, A., Dennis, L. C. C., & Gui, T. (2015). Dynamical behavior of cholera epidemic model with non-linear incidence rate. *Appl. Math. Sci*, 9(20), 989-1002. <https://doi.org/10.12988/ams.2015.41166>
- King, A. A., Ionides, E. L., Pascual, M., & Bouma, M. J. (2008). Inapparent infections and cholera dynamics, *Nature*, 454(7206), 877. <https://doi.org/10.1038/nature07084>
- Lahodny Jr, G. E., Gautam, R., & Ivanek, R. (2015). Estimating the probability of an extinction or major outbreak for an environmentally transmitted infectious disease. *Journal of biological dynamics*, 9(sup1), 128-155.
- Longini Jr, I. M., Nizam, A., Ali, M., Yunus, M., Shenvi, N., & Clemens, J. D. (2007). Controlling endemic cholera with oral vaccines. *PLoS medicine*, 4(11), e336. <https://doi.org/10.1371/journal.pmed.0040336>
- Martinez, J. G., Liang, F., Zhou, L., & Carroll, R. J. (2010). Longitudinal functional principal component modelling via Stochastic Approximation Monte Carlo. *Canadian Journal of Statistics*, 38(2), 256-270.

<https://doi.org/10.1002/cjs.10062>

- Marwa, Y. M., Mwalili, S., & Mbalawata, I. S. (2018). Markov chain Monte Carlo analysis of cholera epidemic. *Journal of Mathematical and Computational Science*, 8(5), 584-610.
- Neilan, R. L. M., Schaefer, E., Gaff, H., Fister, K. R., & Lenhart, S. (2010). Modeling optimal intervention strategies for cholera. *Bulletin of mathematical biology*, 72(8), 2004-2018.
- Njagarah, J. B., & Nyabadza, F. (2015). Modelling optimal control of cholera in communities linked by migration. *Computational and mathematical methods in medicine*, 2015.
- Panja, P., & Mondal, S. K. (2014). A mathematical study on the spread of Cholera. *South Asian J. Math*, 4(2), 69-84.
- Tuite, A. R., Tien, J., Eisenberg, M., Earn, D. J., Ma, J., & Fisman, D. N. (2011). Cholera epidemic in Haiti, 2010: using a transmission model to explain spatial spread of disease and identify optimal control interventions. *Annals of internal medicine*, 154(9), 593-601.
- Van den Driessche, P., & Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1-2), 29-48.
- Whittle, P. (1955). The outcome of a stochastic epidemic? a note on Bailey's paper. *Biometrika*, 42(1-2), 116-122.
- World Health Organization. (2010). *The world health report: health systems financing: the path to universal coverage: executive summary* (No. WHO/IER/WHR/10.1). Geneva: World Health Organization.
- World Health Organization. (2010). Cholera vaccines: who position paper, *Weekly Epidemiological Record*, 85(13), 117-128.
- World Health Organization. (2016). *Prevention and control of cholera outbreaks*.
- Zevika, M., & Soewono, E. (2018). *Deterministic and stochastic CTMC models from Zika disease transmission*. In AIP Conference Proceedings (Vol. 1937, No. 1, p. 020023). AIP Publishing.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).