

## RESEARCH ARTICLE

## Determination of critical community size from an HIV/AIDS model

Sarmistha Das<sup>1</sup>, Pramit Ghosh<sup>2</sup>, Sandip Banerjee<sup>3</sup>, Saumyadipta Pyne<sup>4,5</sup>, Joydev Chattopadhyay<sup>6</sup>, Indranil Mukhopadhyay<sup>1\*</sup>

**1** Human Genetics Unit, Indian Statistical Institute, Kolkata, West Bengal, India, **2** Deben Mahata Government Medical College & Hospital, Purulia, West Bengal, India, **3** Department of Mathematics, Indian Institute of Technology, Roorkee, Uttarakhand, India, **4** Public Health Dynamics Lab and Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America, **5** Health Analytics Network, Pittsburgh, Pennsylvania, United States of America, **6** Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata, West Bengal, India

\* [indranil@isical.ac.in](mailto:indranil@isical.ac.in)

## OPEN ACCESS

**Citation:** Das S, Ghosh P, Banerjee S, Pyne S, Chattopadhyay J, Mukhopadhyay I (2021) Determination of critical community size from an HIV/AIDS model. PLoS ONE 16(1): e0244543. <https://doi.org/10.1371/journal.pone.0244543>

**Editor:** Eric Forgoston, Montclair State University, UNITED STATES

**Received:** February 3, 2020

**Accepted:** December 11, 2020

**Published:** January 28, 2021

**Copyright:** © 2021 Das et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data from website of UNAIDS: <https://www.unaids.org/en/regionscountries/countries/madagascar> <http://cfs.fivci.org/country-factsheet.html>. The authors had no special access privileges others would not have.

**Funding:** This work was supported by the NIH Fogarty International Center Award #R25 TW009717 and #R25 TW009717-02S1. IM received this award. NO - The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

After an epidemic outbreak, the infection persists in a community long enough to engulf the entire susceptible population. Local extinction of the disease could be possible if the susceptible population gets depleted. In large communities, the tendency of eventual damp down of recurrent epidemics is balanced by random variability. But, in small communities, the infection would die out when the number of susceptible falls below a certain threshold. Critical community size (CCS) is considered to be the mentioned threshold, at which the infection is as likely as not to die out after a major epidemic for small communities unless reintroduced from outside. The determination of CCS could aid in devising systematic control strategies to eradicate the infectious disease from small communities. In this article, we have come up with a simplified computation based approach to deduce the CCS of HIV disease dynamics. We consider a deterministic HIV model proposed by Silva and Torres, and following Näsell, introduce stochasticity in the model through time-varying population sizes of different compartments. Besides, Metcalf's group observed that the relative risk of extinction of some infections on islands is almost double that in the mainlands i.e. infections cease to exist at a significantly higher rate in islands compared to the mainlands. They attributed this phenomenon to the greater recolonization in the mainlands. Interestingly, the application of our method on demographic facts and figures of countries in the AIDS belt of Africa led us to expect that existing control measures and isolated locations would assist in temporary eradication of HIV infection much faster. For example, our method suggests that through systematic control strategies, after 7.36 years HIV epidemics will temporarily be eradicated from different communes of island nation Madagascar, where the population size falls below its CCS value, unless the disease is reintroduced from outside.

## Introduction

Periodicity in the recurrence of local epidemics has attracted mathematicians and epidemiologists for centuries. Soper [1] pointed out that the time to extinction of infection and the size of

**Competing interests:** The authors have no competing interest related to this work.

the susceptible population are possibly the responsible candidates behind the periodicity. The persistence of a disease in a community, often measured by the time for an epidemic to die out, is a major concern to date. After an epidemic outbreak, the infection persists in the community long enough to engulf the entire susceptible population. Local extinction of the disease could be possible if the susceptible population gets depleted. The success of Hamer-Soper model [1] in predicting periodicity of epidemics from simple assumptions, led Bartlett [2–4] to observe a threshold for the susceptible population size. If the susceptible population is equal to that threshold, there is a 50% chance that the infection would cease to exist after the corresponding mean fade-out time unless the disease is reintroduced from outside. He analyzed this threshold using a semi-stochastic version of the Hamer-Soper model and established time to extinction of the epidemic using quasi stationarity. Bartlett's model sufficiently explained the undamped epidemic outbreaks and its extinction. Bartlett applied the semi-stochastic model for the recurrent measles outbreak in England and Wales. He observed that in large communities, the tendency of eventual damp down of recurrent epidemics is balanced by random variability; but, in small communities, the infection would die out when the size of the susceptible population falls below a certain threshold. These led Bartlett to define Critical Community Size (CCS) as a threshold size of the susceptible population in a small community, at which infection is as likely as not to die out, after a major epidemic, unless reintroduced from outside. The idea is that, if the size of susceptible population drops below that threshold, the pathogen might cease to exist in the population.

Apart from the works by Andersson and Britton [5], scarcely significant mathematical contributions were made in this direction until Näsell [6] reformulated CCS for Bartlett's Susceptible-Infected (SI) model using quasi stationarity under more realistic assumptions. Näsell [6, 7] used a fully stochastic model to study the time to extinction of an epidemic for a number of disease models. The model and method proposed by Näsell [6] is based on a standard SI model that involves a system of two differential equations. This model can easily be extended to an SIR model where the population size remains constant at any time, as this SIR model can be reduced to a two-equation model and the above method is applicable without any significant modification. Näsell [6] also focused to unleash the potential of CCS as a control strategy for the epidemic outbreak. Later, Metcalf et al. [8] mentioned this potential of CCS. But, calculation of CCS is limited mostly to models based on a two-equation system of differential equations or measles type diseases. One predominant reason is that the analytical formulation of CCS for complex infectious diseases is quite cumbersome. The absence of any method to calculate CCS for higher-order models motivated us to develop a cost (in terms of time) effective technique to evaluate CCS. In this paper, we propose a simplified computation based approach for rapid calculation of CCS, that could emphasize prompt formulation of control strategies and/or policies at the time of disease outbreak.

Technically, this paper determines the time to extinction of an epidemic outbreak theoretically and also develops a computation based approximate formulation of CCS using the diffusion approximation. The application of our method is free from any restriction on the number of variables in the system that defines the disease dynamics. To demonstrate our method we consider an HIV model by Silva, Torres, and Djordjevic [9–11] that describes the disease dynamics using a system of four differential equations. Since infectious diseases mostly have a stochastic nature [12–14], we consider a stochastic version of the HIV disease dynamics. Following Näsell, we introduce stochasticity through time-varying population sizes of different compartments in the deterministic HIV model [9]. Next, we calculate the mean time to extinction of the HIV epidemic and compute approximate value of CCS using diffusion approximation based on the Ornstein-Uhlenbeck process.

In this paper, we have cited the HIV epidemic case of Madagascar, in Sub-Saharan Africa because of a special reason. In spite of less remarkable progress in the 90-90-90 HIV treatment target of the Joint United Nations Programme on HIV and AIDS (UNAIDS), Madagascar has shown a much lower prevalence and incidence rate of HIV infection, compared to some other African countries. It is interesting to note that, despite being in the part of Africa where HIV is in a state of epidemic, it has much lower rates of prevalence and incidence. Incidentally, Madagascar is situated in a geographically isolated location, separated by 60 km of water from the other eastern and southern countries that lie in the mainland of the African continent. This observation echoes a fact from Metcalf et al. [8], that says the relative risk of extinction of some infectious diseases on islands is almost double that of those in the mainland. They also suggest that persistence of the infection within a region in the mainland is fuelled constantly by recolonization and infections cease to exist at significantly higher rates in islands where access to and from the mainland is much less. This fascinating phenomenon led us to postulate that, in conjunction with UNAIDS controlling strategies if it is possible to keep the HIV susceptible population of Madagascar small or rather below CCS, effective and rapid eradication of HIV might be possible there, after the mean extinction time of the disease. This is primarily based on the fact that the isolated geographical location of Madagascar might lead to lesser chances of re-introduction of the infection.

We, therefore, applied our method on the demographic facts and figures of Madagascar to obtain the CCS value and mean time to extinction of the disease. Our calculations suggest that, unless the disease is reintroduced from outside, after 7.36 years HIV epidemics might fade out from different cities or communes of the island nation where the susceptible population size drops below 4585, the CCS value of Madagascar. This time period of 7.36 years could appear astonishing as HIV has become a long-standing epidemic situation in the countries of the Sub-Saharan AIDS belt. But there is a strong possibility that the idea of CCS would propel the controlling strategies of UNAIDS to eradicate the disease in the near future. This work suggests that the systematic spreading of control measures could lead to accelerated progress to achieve the 90-90-90 target of UNAIDS for HIV eradication. Such progress shall prevent the HIV epidemic in Madagascar unless the infection is reintroduced from outside. Moreover, countries like Uganda, Malawi (located in Eastern and Southern Africa) are separated from their neighboring countries by water bodies such as fresh water lakes. These countries have lower rates of prevalence and incidence compared to landlocked countries such as Eswatini (in Southern Africa). Notably, these rates are still much higher than that of Madagascar.

To date, CCS has largely been a part of mathematical theory only. But implementing CCS as controlling strategies of epidemics might prove beneficial in the future. Therefore, to utilize the potential that CCS holds, it is important to determine its value promptly for any disease dynamics. This motivated us to calculate CCS (approximately) using simplified computations. Our method will help in finding country-specific CCS and time to extinction of the disease, using real demographic facts and figures for HIV infection in different countries as parameter values. We emphasize the potential utility of CCS to guide public health practitioners to develop appropriate area-specific control strategies and interventions especially when they are limited, before or after an epidemic outbreak.

## Materials and methods

### Model formulation

We work with the model proposed by Silva and Torres [9] that explains the dynamics of HIV infection assuming homogenous mixing within communities. Here the entire population is divided into four compartments. These compartments are mutually exclusive in the sense that

no person can belong to more than one compartment at any time point. The four compartments are: susceptible individuals ( $S$ ), individuals infected with HIV but having no clinical symptoms of AIDS ( $I$ ), individuals infected with AIDS and having ART treatment as they are supposed to belong to chronic stage ( $C$ ), and individuals infected with HIV having clinical symptoms of AIDS ( $A$ ). Note that a person belonging to  $I$  may transmit HIV to other individuals although he/she has no symptoms of AIDS. Under this situation, the model we consider is,

$$\frac{dS}{dt} = \Lambda - \beta(I(t) + \eta_C C(t) + \eta_A A(t)) \frac{S(t)}{N} - \mu S(t) \tag{1}$$

$$\frac{dI}{dt} = \beta(I(t) + \eta_C C(t) + \eta_A A(t)) \frac{S(t)}{N} - (\rho + \phi + \mu)I(t) + \omega C(t) + \alpha A(t) \tag{2}$$

$$\frac{dC}{dt} = \phi I(t) - (\omega + \mu)C(t) \tag{3}$$

$$\frac{dA}{dt} = \rho I(t) - (\alpha + \mu + d)A(t) \tag{4}$$

Here  $\beta$  represents the contact rate for HIV transmission,  $\eta_A (\geq 1)$  is the relative infectiousness of individuals with AIDS symptoms compared to those infected with HIV but no AIDS symptoms, and  $\eta_C (\leq 1)$  is the rate of partial restoration of immune function of HIV infected individuals who are correctly treated under ART.  $\rho$  and  $\phi$  are the rates at which individuals in class  $I$  move to  $A$  and  $C$  respectively.  $\alpha$  and  $\omega$  are the respective rates at which individuals from class  $A$  and class  $C$  move to  $I$ . We assume that  $d$  is the HIV induced death rate while  $\mu$  is the natural death rate in the population.

We consider the total population as  $N(t) = S(t) + I(t) + A(t) + C(t)$  implying that  $\frac{dN(t)}{dt} = \Lambda - \mu N(t) - dA(t)$ . It can be easily shown that  $N(t) \leq \frac{\Lambda}{\mu}$ .

### Basic reproduction number

We use the approach of next generation matrix [15] to calculate  $R_0$ . Note that the rate of new infections coming in the compartments  $I$ ,  $C$  and  $A$  are  $\beta(I + \eta_C C + \eta_A A) \frac{S}{N}$ , 0 and 0 respectively. On the other hand, new infections moving out of the compartments  $I$ ,  $C$ , and  $A$  are  $(\rho + \phi + \mu)I - \omega C - \alpha A$ ,  $(\omega + \mu)C - \phi I$  and  $(\alpha + \mu + d)A - \rho I$  respectively. Thus following [9, 15], we have basic reproduction number ( $R_0$ ) as:

$$R_0 = \frac{S^0 \beta}{N} \frac{\xi_2 (\xi_1 + \rho \eta_A) + \eta_C \phi \xi_1}{\mu (\xi_2 + (\rho + \xi_1) + \phi \xi_1 + \rho d) + \rho \omega d} = \frac{S^0 \zeta}{N \gamma} \approx \frac{\zeta}{\gamma} \tag{5}$$

where  $\zeta = \beta(\xi_2(\xi_1 + \rho\eta_A) + \eta_C \phi \xi_1)$ ,  $\gamma = \mu(\xi_2 + (\rho + \xi_1) + \phi \xi_1 + \rho d) + \rho \omega d$ ,  $\xi_1 = \alpha + \mu + d$  and  $\xi_2 = \omega + \mu$ .

### Stochastic model and quasi-stationarity

Evaluation of CCS depends on the time to extinction of disease based on the given model. In order to find the time to extinction, we now construct the fully stochastic version of the model (1)–(4). Next, we write the Kolmogorov forward equation for this stochastic model. This would be the key step for further downstream analysis.

First, we note the nature of transitions and the respective transition rates from one compartment to another (Table 1).

**Table 1. Chart for transition rates.**

Event	Transition	Transition rate
Immigration of susceptibles (S)	$(s, i, c, a) \rightarrow (s + 1, i, c, a)$	$\lambda_1(s, i, c, a) = \mu N$
Death of susceptibles (S)	$(s, i, c, a) \rightarrow (s - 1, i, c, a)$	$\lambda_2(s, i, c, a) = \mu s$
Susceptible (S) to Infected (I)	$(s, i, c, a) \rightarrow (s - 1, i + 1, c, a)$	$\lambda_3(s, i, c, a) = \beta si/N$
Susceptible (S) to Infected under treatment (C)	$(s, i, c, a) \rightarrow (s - 1, i, c + 1, a)$	$\lambda_4(s, i, c, a) = \beta \eta_C sc/N$
Susceptible (S) to Infected with AIDS (A)	$(s, i, c, a) \rightarrow (s - 1, i, c, a + 1)$	$\lambda_5(s, i, c, a) = \beta \eta_A sa/N$
Infected (I) to Infected under treatment(C)	$(s, i, c, a) \rightarrow (s, i - 1, c + 1, a)$	$\lambda_6(s, i, c, a) = \phi i$
Infected (I) to Infected with AIDS (A)	$(s, i, c, a) \rightarrow (s, i - 1, c, a + 1)$	$\lambda_7(s, i, c, a) = \rho i$
Death of Infected (I)	$(s, i, c, a) \rightarrow (s, i - 1, c, a)$	$\lambda_8(s, i, c, a) = \mu i$
Infected under treatment (C) to infected (I)	$(s, i, c, a) \rightarrow (s, i + 1, c - 1, a)$	$\lambda_9(s, i, c, a) = \omega c$
Death of Infected under treatment(C)	$(s, i, c, a) \rightarrow (s, i, c - 1, a)$	$\lambda_{10}(s, i, c, a) = \mu c$
Infected with AIDS symptoms (A) to Infected (I)	$(s, i, c, a) \rightarrow (s, i + 1, c, a - 1)$	$\lambda_{11}(s, i, c, a) = \alpha a$
Death of Infected with AIDS symptoms (A)	$(s, i, c, a) \rightarrow (s, i, c, a - 1)$	$\lambda_{12}(s, i, c, a) = (\mu + d)a$

<https://doi.org/10.1371/journal.pone.0244543.t001>

Here,  $s = S/N, i = I/N, c = C/N, a = A/N$ . Then the Kolmogorov forward equation becomes,

$$\begin{aligned}
 p'_{s,i,c,a}(t) &= \lambda_1(s - 1, i, c, a)p_{s-1,i,c,a} + \lambda_2(s + 1, i, c, a)p_{s+1,i,c,a} \\
 &+ \lambda_3(s + 1, i - 1, c, a)p_{s+1,i-1,c,a} + \lambda_4(s + 1, i, c - 1, a)p_{s+1,i,c-1,a} \\
 &+ \lambda_5(s + 1, i, c, a - 1)p_{s+1,i,c,a-1} + \lambda_6(s, i + 1, c - 1, a)p_{s,i+1,c-1,a} \\
 &+ \lambda_7(s, i + 1, c, a - 1)p_{s,i+1,c,a-1} + \lambda_8(s, i + 1, c, a)p_{s,i+1,c,a} \\
 &+ \lambda_9(s, i - 1, c + 1, a)p_{s,i-1,c+1,a} + \lambda_{10}(s, i, c + 1, a)p_{s,i,c+1,a} \\
 &+ \lambda_{11}(s, i - 1, c, a + 1)p_{s,i-1,c,a+1} + \lambda_{12}(s, i, c, a + 1)p_{s,i,c,a+1} - \kappa(s, i, c, a)p_{s,i,c,a}
 \end{aligned} \tag{6}$$

where  $\kappa(s, i, c, a) = \sum_{i=1}^{12} \lambda_i(s, i, c, a)$  and  $p_{s,i,c,a}(t) = P[S(t) = s, I(t) = i, C(t) = c, A(t) = a]$ .

Now, we find quasi-stationarity by conditioning on non-extinction of the disease. So we write,

$$q_{s,i,c,a}(t) = P[S(t) = s, I(t) = i, C(t) = c, A(t) = a | (I(t), C(t), A(t)) \neq (0, 0, 0)] = \frac{p_{s,i,c,a}(t)}{1 - p_{\bullet 000}(t)}$$

where  $p_{\bullet 000}(t) = \sum_{s=0}^{\infty} P[S(t) = s, I(t) = 0, C(t) = 0, A(t) = 0] = \sum_{s=0}^{\infty} p_{s,0,0,0}(t)$ .

Differentiating  $q_{s,i,c,a}(t)$  with respect to  $t$ , we have,

$$q'_{s,i,c,a}(t) = \frac{p'_{s,i,c,a}(t)}{1 - p_{\bullet 000}(t)} + \frac{p_{s,i,c,a}(t)}{(1 - p_{\bullet 000}(t))^2} \cdot p'_{\bullet 000}(t). \tag{7}$$

Now, from Eq (6), we have, after simplification,

$$\begin{aligned}
 p'_{\bullet 000}(t) &= \sum_{s=0}^{\infty} \left[ \mu N p_{s-1,0,0,0}(t) + \mu(s + 1)p_{s+1,0,0,0}(t) + \mu(0 + 1)p_{s,1,0,0}(t) \right. \\
 &+ \left. \mu(0 + 1)p_{s,0,1,0}(t) + (\mu + d)(0 + 1)p_{s,0,0,1}(t) \right] - (\mu N + \mu s) \sum_{s=0}^{\infty} p_{s,0,0,0}(t) \\
 &= \mu p_{\bullet 1 0 0}(t) + \mu p_{\bullet 0 1 0}(t) + (d + \mu)p_{\bullet 0 0 1}(t) = p_{\bullet}^{(d,\mu)}(t) \text{ (say)}.
 \end{aligned} \tag{8}$$

From Eqs (7) and (8) we have,

$$\begin{aligned}
 q'_{s,i,c,a}(t) &= \frac{p'_{s,i,c,a}(t)}{1 - p_{\bullet 0 0 0}(t)} + \frac{p_{s,i,c,a}(t)}{(1 - p_{\bullet 0 0 0}(t))^2} \cdot p'_{\bullet 0 0 0}(t) \\
 &= \frac{p'_{s,i,c,a}(t)}{1 - p_{\bullet 0 0 0}(t)} + \frac{p_{s,i,c,a}(t)}{1 - p_{\bullet 0 0 0}(t)} \cdot \frac{p_{\bullet}^{(d,\mu)}(t)}{1 - p_{\bullet 0 0 0}(t)} \\
 &= \frac{p'_{s,i,c,a}(t)}{1 - p_{\bullet 0 0 0}(t)} + \frac{p_{s,i,c,a}(t)}{1 - p_{\bullet 0 0 0}(t)} \cdot q_{\bullet}^{(d,\mu)}(t), \quad (\text{say})
 \end{aligned}
 \tag{9}$$

$$\text{where } q_{\bullet}^{(d,\mu)}(t) = \frac{p_{\bullet}^{(d,\mu)}(t)}{1 - p_{\bullet 0 0 0}(t)}
 \tag{10}$$

Now, following Näsell [6, 7] we have,

$$\begin{aligned}
 q'_{s,i,c,a}(t) &= 0 \\
 \Rightarrow p'_{s,i,c,a}(t) &= -\frac{p_{s,i,c,a}(t)}{(1 - p_{\bullet 0 0 0}(t))} \cdot q_{\bullet}^{(d,\mu)}(t)(1 - p_{\bullet 0 0 0}(t)) = -q_{\bullet}^{(d,\mu)}(t)p_{s,i,c,a}(t) \\
 \therefore p_{s,i,c,a}(t) &= ce^{-q_{\bullet}^{(d,\mu)} \cdot t} = q_{\bullet}^{(d,\mu)} e^{-q_{\bullet}^{(d,\mu)} \cdot t}
 \end{aligned}
 \tag{11}$$

Let  $\tau_Q$  be the time to extinction when the initial distribution is equal to the quasi-stationary distribution. Therefore, we have,

$$E(\tau_Q) = \frac{1}{q_{\bullet}^{(d,\mu)}}.
 \tag{12}$$

### Equilibrium points

The disease-free equilibrium is obtained as:  $\Sigma_0 = (S^0, I^0, C^0, A^0) = (\frac{\Delta}{\mu}, 0, 0, 0)$ . To find the other endemic equilibrium, if exists, we put  $N = N(0)$ ,  $x_1(t) = S(t)/N$ ,  $x_2(t) = I(t)/N$ ,  $x_3(t) = C(t)/N$ , and  $x_4(t) = A(t)/N$ . The equilibrium point is obtained by equating the first differentiation in Eqs (1)–(4) to zero [9, 10], i.e.

$$x'_1(t) = \mu - \beta(x_2(t) + \eta_C x_3(t) + \eta_A x_4(t))x_1 - \mu x_1(t) = 0
 \tag{13}$$

$$x'_2(t) = \beta(x_2(t) + \eta_C x_3(t) + \eta_A x_4(t))x_1 - (\rho + \phi + \mu)x_2(t) + \omega x_3(t) + \alpha x_4(t) = 0
 \tag{14}$$

$$x'_3(t) = \phi x_2(t) - (\omega + \mu)x_3(t) = 0
 \tag{15}$$

$$x'_4(t) = \rho x_2(t) - (\alpha + \mu + d)x_4(t) = 0
 \tag{16}$$

For simplicity we use the notations:  $x_j(t) = x_j$  for  $j = 1, \dots, 4$ ,  $\xi_1 = \alpha + \mu + d$ ,  $\xi_2 = \omega + \mu$ ,  $\gamma = \xi_1 \xi_2(\rho + \phi + \mu) - \omega \phi \xi_1 - \rho \alpha \xi_2$ ,  $\zeta = \beta(\xi_1 \xi_2 + \eta_C \phi \xi_1 + \eta_A \rho \xi_2)$ .

Then solving Eqs (13)–(16), we have the endemic equilibrium as:

$$\hat{x}_1 = \frac{\gamma}{\zeta} = \frac{1}{R_0}
 \tag{17}$$

$$\hat{x}_2 = \frac{(\zeta - \gamma)}{\zeta\gamma} \cdot \mu \zeta_1 \zeta_2 = (R_0 - 1) \frac{\mu}{\zeta} \cdot \zeta_1 \zeta_2 \tag{18}$$

$$\hat{x}_3 = \frac{(\zeta - \gamma)}{\zeta\gamma} \cdot \mu \phi \zeta_1 = (R_0 - 1) \frac{\mu}{\zeta} \cdot \phi \zeta_1 \tag{19}$$

$$\hat{x}_4 = \frac{(\zeta - \gamma)}{\zeta\gamma} \cdot \mu \rho \zeta_2 = (R_0 - 1) \frac{\mu}{\zeta} \cdot \rho \zeta_2 \tag{20}$$

### Diffusion approximation

We consider a diffusion approximation to the stochastic version of our model (1)–(4). This would allow us to approximately evaluate the quasi-stationary distribution using a multivariate normal distribution when  $N$  is large and  $R_0$  is greater than 1. First we find the multivariate normal distribution corresponding to the state variables. Let the changes in the scaled state variables  $x_1, x_2, x_3,$  and  $x_4$  during the time interval  $(t, t + \delta t)$  be denoted by  $\delta x_1, \delta x_2, \delta x_3,$  and  $\delta x_4$  respectively, where  $\delta x_i(t) = x_i(t + \delta t) - x_i(t), 1, 2, 3, 4.$

Under the assumptions of the original process on the sequence of transitions, we evaluate the mean vector and covariance matrix for  $\delta x_i (i = 1, 2, 3, 4)$  during the time interval  $(t, t + \delta t)$  as follows. At first, we assume that we are in the state  $(S, I, C, A)$ . So, the possible transitions from this state are:

- (a).  $S$  increases by 1 at the rate  $\mu$
- (b).  $S$  decreases by 1 at the rate  $\mu$
- (c).  $S$  decreases by 1 and  $I$  increases by 1 at the rate  $\beta SI/N$
- (d).  $S$  decreases by 1 and  $I$  increases by 1 at the rate  $\beta \eta_C SC/N$
- (e).  $S$  decreases by 1 and  $I$  increases by 1 at the rate  $\beta \eta_A SA/N$
- (f).  $I$  decreases by 1 and  $C$  increases by 1 at the rate  $\phi I$
- (g).  $I$  decreases by 1 and  $A$  increases by 1 at the rate  $\rho I$
- (h).  $I$  decreases by 1 at the rate  $\mu I$
- (i).  $C$  decreases by 1 and  $I$  increases by 1 at the rate  $\omega C$
- (j).  $C$  decreases by 1 at the rate  $\mu C$
- (k).  $A$  decreases by 1 and  $I$  increases by 1 at the rate  $\alpha A$
- (l).  $A$  decrease by 1 at the rate  $(\mu + d)A.$

The random variable  $\delta x_1$  equals  $\frac{1}{N}$  in cases (a),  $-\frac{1}{N}$  in cases (b), (c), (d), (e), and 0 in other cases. Similarly,  $\delta x_2$  equals  $\frac{1}{N}$  in cases (c), (d), (e), (i), (k),  $-\frac{1}{N}$  in cases (f), (g), (h), and 0 in other cases.  $\delta x_3$  equals  $\frac{1}{N}$  in cases (f),  $-\frac{1}{N}$  in cases (i), (j), and 0 in other cases.  $\delta x_4$  equals  $\frac{1}{N}$  in cases (g),  $-\frac{1}{N}$  in cases (k), (l), and 0 in other cases. Denoting  $\mathbf{x} = (x_1, x_2, x_3, x_4)'$  we have,

$$E(\delta \mathbf{x}) = b(\mathbf{x})\delta t + o(\delta t)$$

$$\text{where } b(\mathbf{x}) = \begin{pmatrix} \mu - \beta(x_2 + \eta_C x_3 + \eta_A x_4)x_1 - \mu x_1 \\ \beta(x_2 + \eta_C x_3 + \eta_A x_4)x_1 - (\rho + \phi + \mu)x_2 + \omega x_3 + \alpha x_4 \\ \phi x_2 - (\omega + \mu)x_3 \\ \rho x_2 - (\alpha + \mu + d)x_4 \end{pmatrix} \tag{21}$$

Now to derive the covariance matrix we need to find the Jacobian matrix at point  $\mathbf{x}$ , which is obtained from  $b(\mathbf{x})$  as,

$$B(\mathbf{x}) = \frac{\partial b(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} -\beta(x_2 + \eta_C x_3 + \eta_A x_4) - \mu & -\beta x_1 & -\beta \eta_C x_1 & -\beta \eta_A x_1 \\ \beta(x_2 + \eta_C x_3 + \eta_A x_4) & \beta x_1 - (\rho + \phi + \mu) & \beta \eta_C x_1 + \omega & \beta \eta_A x_1 + \alpha \\ 0 & \phi & -(\omega + \mu) & 0 \\ 0 & \rho & 0 & -(\alpha + \mu + d) \end{pmatrix}$$

Then we approximate  $B(\mathbf{x})$  at equilibrium point  $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)'$  by  $B(\hat{\mathbf{x}})$ . So, covariance matrix of  $\delta \mathbf{x} = (\delta x_1, \delta x_2, \delta x_3, \delta x_4)'$  is,

$$V(\delta \mathbf{x}) = \frac{1}{N} S(\mathbf{x}) \delta t + o(\delta t) \text{ where,}$$

$$S(\mathbf{x}) = \begin{pmatrix} \beta(x_2 + \eta_C x_3 + \eta_A x_4)x_1 & -\beta x_1 x_2 & -\beta \eta_C x_1 x_3 & -\beta \eta_A x_1 x_4 \\ \mu' + \mu x_1 & & & \\ -\beta x_1 x_2 & \beta(x_2 + \eta_C x_3 + \eta_A x_4)x_1 & -\phi x_2 - \omega x_3 & -\rho x_2 - \alpha x_4 \\ & +(\rho + \phi + \mu)x_2 + \omega x_3 + \alpha x_4 & & \\ -\beta \eta_C x_1 x_3 & -\phi x_2 - \omega x_3 & \phi x_2 + (\omega + \mu)x_3 & 0 \\ -\beta \eta_A x_1 x_4 & -\rho x_2 - \alpha x_4 & 0 & \rho x_2 \\ & & & +(\alpha + \mu + d)x_4 \end{pmatrix}$$

We again approximate  $S(\mathbf{x})$  by  $S(\hat{\mathbf{x}})$ , where  $\hat{\mathbf{x}}$  is the equilibrium point. For large  $N$ , the process  $\sqrt{N}(\mathbf{x}(t) - \hat{\mathbf{x}})$  is approximated by a multivariate Ornstein-Uhlenbeck (O-U) process [16], with a local drift matrix  $B(\hat{\mathbf{x}})$  and local covariance matrix  $S(\hat{\mathbf{x}})$ .

The stationary distribution of this O-U process approximates the quasi-stationary distribution. It is approximately normal with mean zero and covariance matrix  $\Sigma$ , where  $\Sigma$  is obtained by solving

$$B(\hat{\mathbf{x}})\Sigma + \Sigma B'(\hat{\mathbf{x}}) = -S(\hat{\mathbf{x}}). \tag{22}$$

Exact analytical solution for  $\Sigma$  is not straightforward [5]. However, since we are interested in calculating the CCS, we can easily solve the Eq (22) numerically using the parameter values and the equilibrium point.

Let  $\sigma_{ij}$  be the solution of the  $(i, j)$ th element of  $\Sigma$ , where  $i, j = 1, \dots, 4$ . The diffusion approximation guides us to consider appropriate mean and variance-covariance matrix for the joint



distribution of  $x(t)$ . Thus we have,

$$\sqrt{N}(\mathbf{x}(t) - \hat{\mathbf{x}}) \sim N_4(\mathbf{0}, \Sigma), \text{ with } \mathbf{x}(t) = (x_1(t), x_2(t), x_3(t), x_4(t))', \hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)'$$

$$\text{and } \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{pmatrix} \tag{23}$$

The approximation for quasi-stationary distribution is obtained by using conditional truncated distributions of the above multivariate normal distribution. Thus in order to evaluate  $p_{\bullet 000}, p_{\bullet 100}, p_{\bullet 010}$ , and  $p_{\bullet 001}$  we need to use a result from conditional truncated multivariate normal distribution, which is given below as theorem 1 (for proof see [S1 Appendix](#)). Even then, it is extremely difficult, if not impossible [5], to get an exact expression for  $q_{\bullet}(d, \mu)$ . So, we evaluate an approximate expression of  $q_{\bullet}(d, \mu)$  using Result 1 (for proof see [S1 Appendix](#)).

**Theorem 1** Let  $Y \sim N_p(\mu, \Sigma)$  and write  $\mathbf{y} = (y_1', y_2')'$ ,  $\mu = (\mu_1', \mu_2')'$ , and  $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ . Suppose instead of  $\mathbb{R}^p$ ,  $Y$  is defined only on a truncated support  $\mathbf{c} < \mathbf{y} < \mathbf{d}$ . Consider the partitions  $\mathbf{c} = (\mathbf{c}'_1, \mathbf{c}'_2)'$  and  $\mathbf{d} = (\mathbf{d}'_1, \mathbf{d}'_2)'$ . Then, the conditional distribution of  $Y_1$  given  $y_2$  is given by

$$f^*(y_1|y_2) = \frac{f(y_1|y_2)}{\int_{c_1}^{d_1} f(y_1|y_2) dy_1} \tag{24}$$

where  $f(y_1|y_2)$  is the conditional probability density function of  $Y_1$  given  $y_2$  i.e.

$$Y_1|y_2 \sim N_q(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(y_2 - \mu_2), \Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$$

**Result 1** Let  $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$  and  $\Phi(x) = \int_{-\infty}^x \phi(t) dt$  for any  $x \in (-\infty, \infty)$ . Then, an approximate expression of  $q_{\bullet}(d, \mu)$  is given as,

$$q_{\bullet}(d, \mu) = \frac{\mu p_{\bullet 100} + \mu p_{\bullet 010} + (d + \mu) p_{\bullet 001}}{1 - p_{\bullet 000}} \tag{25}$$

where

$$p_{\bullet 100} \approx \frac{1}{2N\sqrt{\sigma_{22}^*}} \frac{\phi\left(\frac{\hat{x}_2^*}{\sqrt{\sigma_{22}^*}}\right)}{\Phi\left(\frac{\hat{x}_2^*}{\sqrt{\sigma_{22}^*}}\right)} \cdot \frac{1}{2N\sqrt{\sigma_{33}^*}} \frac{\phi\left(\frac{\hat{x}_3^*}{\sqrt{\sigma_{33}^*}}\right)}{\Phi\left(\frac{\hat{x}_3^*}{\sqrt{\sigma_{33}^*}}\right)} \cdot \frac{1}{2N\sqrt{\sigma_{44}^*}} \frac{\phi\left(\frac{\hat{x}_4^*}{\sqrt{\sigma_{44}^*}}\right)}{\Phi\left(\frac{\hat{x}_4^*}{\sqrt{\sigma_{44}^*}}\right)}$$

$$p_{\bullet 010} \approx \frac{1}{2N\sqrt{\sigma_{33}^{**}}} \frac{\phi\left(\frac{\hat{x}_3^{**}}{\sqrt{\sigma_{33}^{**}}}\right)}{\Phi\left(\frac{\hat{x}_3^{**}}{\sqrt{\sigma_{33}^{**}}}\right)} \cdot \frac{1}{2N\sqrt{\sigma_{22}^{**}}} \frac{\phi\left(\frac{\hat{x}_2^{**}}{\sqrt{\sigma_{22}^{**}}}\right)}{\Phi\left(\frac{\hat{x}_2^{**}}{\sqrt{\sigma_{22}^{**}}}\right)} \cdot \frac{1}{2N\sqrt{\sigma_{44}^*}} \frac{\phi\left(\frac{\hat{x}_4^*}{\sqrt{\sigma_{44}^*}}\right)}{\Phi\left(\frac{\hat{x}_4^*}{\sqrt{\sigma_{44}^*}}\right)}$$

$$p_{\bullet 001} \approx \frac{1}{2N\sqrt{\sigma_{44}^{***}}} \frac{\phi\left(\frac{\hat{x}_4^{***}}{\sqrt{\sigma_{44}^{***}}}\right)}{\Phi\left(\frac{\hat{x}_4^{***}}{\sqrt{\sigma_{44}^{***}}}\right)} \cdot \frac{1}{2N\sqrt{\sigma_{22}^{***}}} \frac{\phi\left(\frac{\hat{x}_2^{***}}{\sqrt{\sigma_{22}^{***}}}\right)}{\Phi\left(\frac{\hat{x}_2^{***}}{\sqrt{\sigma_{22}^{***}}}\right)} \cdot \frac{1}{2N\sqrt{\sigma_{33}^*}} \frac{\phi\left(\frac{\hat{x}_3^*}{\sqrt{\sigma_{33}^*}}\right)}{\Phi\left(\frac{\hat{x}_3^*}{\sqrt{\sigma_{33}^*}}\right)}$$

$$\tag{26}$$

where  $\hat{x}_i^*, \sigma_{ii}^*$  for  $i = 2, 3$ ,  $\hat{x}_i^{**}, \sigma_{ii}^{**}$  for  $i = 2, 3$ ,  $\hat{x}_i^{***}, \sigma_{ii}^{***}$  for  $i = 2, 4$  are obtained from the truncated conditional distribution of multivariate normal distribution as given in (23).

Once we find  $q_*^{(d,\mu)}$ , we have an expression for expected time to extinction  $\hat{E}(\tau_Q)$  using Eq (12). However, this involves  $N$ , which is unknown. Now, our aim is to find the above expression in another way so that we can equate them and solve for  $N$ . So, following [6], if we can find the quasi-period ( $\hat{T}_0$ ), we may use the relation  $\hat{E}(\tau_Q) \log 2 = \hat{T}_0$  to get the value of  $N$ , which is the critical community size.

We determine the quasi-period of the oscillation about the critical point using linearisation method [17]. Note that for our model, the linearised system about the equilibrium point  $\hat{x} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)'$  can be written as:

$$\frac{d\mathbf{x}^*}{dt} = \begin{pmatrix} -\beta\left(1 + \eta_C \frac{\phi}{\xi_2} + \eta_A \frac{\rho}{\xi_1}\right)\hat{x}_2 - \mu & -\frac{\beta}{R_0} & -\frac{\beta\eta_C}{R_0} & -\frac{\beta\eta_A}{R_0} \\ \beta\left(1 + \eta_C \frac{\phi}{\xi_2} + \eta_A \frac{\rho}{\xi_1}\right)\hat{x}_2 & \frac{\beta}{R_0} - (\rho + \phi + \mu) & \frac{\beta\eta_C}{R_0} + \omega & \frac{\beta\eta_A}{R_0} + \alpha \\ 0 & \phi & -(\omega + \mu) & 0 \\ 0 & \rho & 0 & -(\alpha + \mu + \alpha) \end{pmatrix} \mathbf{x}^* \quad (27)$$

where  $\mathbf{x}^* = \mathbf{x} - \hat{\mathbf{x}}$ .

Now we can find the eigenvalues of the matrix in Eq (27) and hence the angular frequency, provided there are imaginary roots. Using Eq (27) and putting the values of the parameters, we can also find the angular frequency ( $\theta$ ) numerically. Hence the quasi-period ( $\hat{T}_0$ , say) is obtained as  $\hat{T}_0 = \frac{2\pi}{\theta}$ , which is independent of  $N$ . Now noting the relation  $\hat{E}(\tau_Q) \log 2 = \hat{T}_0$  and using Result 1, we can solve for  $N$ .

Since we are dealing with a system consisting of four equations, it is not possible to obtain an explicit or compact expression for CCS (i.e.  $N$ ) in terms of the model parameters. Hence we find an approximate value of CCS numerically.

Our approach is not restricted to only two [6] or three variables [5]. It generalizes the numerical calculation of CCS for any finite number of variables. Thus, our method for finding CCS is a general one and is applicable to any number of variables. Let  $n$  be the number of variables and  $m$  be the number of equations in the system of differential equations that explains the disease dynamics. An algorithm to calculate CCS for such a process is given below.

- **Step I:** Calculate basic reproduction number and endemic equilibrium for the system of differential equations i.e. the model.
- **Step II:** Write down a fully stochastic model and the joint distribution of the state variables conditioning on non-extinction of the disease.
- **Step III:** Write the expression for expected time to extinction based on quasi-stationary distribution.
- **Step IV:** Introduce diffusion approximation to obtain the local drift matrix and local covariance matrix from multivariate Ornstein-Uhlenbeck process.
- **Step V:** Using conditional truncated multivariate normal distribution, calculate expected time to extinction ( $\hat{E}(\tau_Q)$ , say)
- **Step VI:** Using linearisation of the model, calculate the angular frequency ( $\hat{T}_0$ , say).
- **Step VII:** Merge expected time to extinction with angular frequency by the relation  $\hat{E}(\tau_Q) \log 2 = \hat{T}_0$ , and solve for  $N$ , which is the CCS with respect to the model.

Note that CCS thus calculated is an approximate one as we have made some approximations at few stages. However, this algorithm is a general one and may be applied to any model involving a system of differential equations that explains the disease dynamics. However, the basic reproduction number ( $R_0$ ) must be greater than 1 in order to calculate CCS using our method.

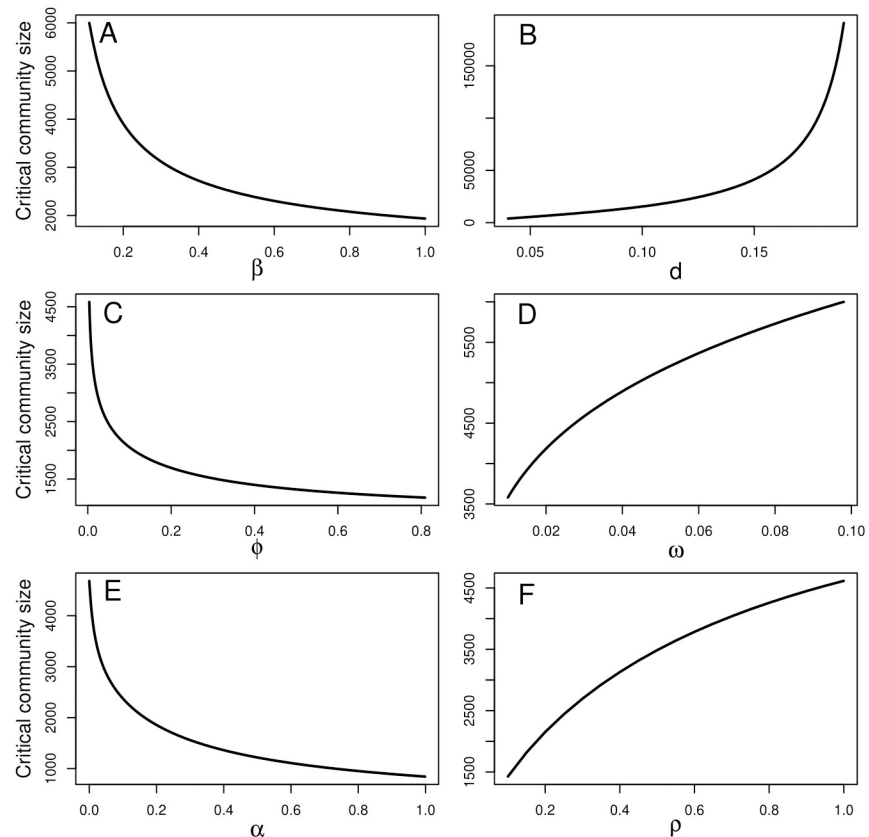
## Results

### Case study

For HIV transmission, we can calculate the critical community size for any region or community. Note that this value of CCS and time to extinction are approximate as we have made some approximation while applying diffusion approximation using conditional truncated multivariate normal distribution. All calculations are based on the assumed model (1)-(4). We suppose that an individual in the susceptible compartment/class (S) either moves to the infected (I) compartment or infected under treatment (C) compartment or infected with AIDS (A) compartment; while an individual from infected (I) class may receive treatment and move to infected under treatment (C) class or get infected with AIDS and move to infected with AIDS (A) class; while an individual after treatment may have reduced viral load and thus move from class C to the infected (I) class; similarly, an individual infected with AIDS (A) may have reduced viral load and move to infected (I) class. Death may occur in each class while immigration occurs only in the susceptible class.

It is clear that in order to calculate CCS for a region or community, we need to know the values of the parameters in the model. These parameter values are usually estimated based on different studies. As an illustration, we consider the values of the parameters based on the data available for Madagascar. However, the values of all parameters are not available in the literature. In such cases, we have assumed a few values considering the meaning and range of the parameters. Based on the available data for 2018 from the UNAIDS ([www.unaids.org/en/regionscountries/countries/madagascar](http://www.unaids.org/en/regionscountries/countries/madagascar)), we see that the number of people living with HIV who are on ART is 3500 whereas the number of adults and children living with HIV is 39000. The number of adults and children newly infected with HIV is 6100 and the number of deaths due to AIDS is 1700. Based on this information and following Näsell [6] we have considered  $\beta = 0.156 \times 365 = 56.94$  and  $d = 0.044 \times 365 = 16.06$ . Based on the available reports (<http://cfs.hivci.org/country-factsheet.html>) from the World Health Organisation (WHO), we have  $\omega = 0.03 \times 365 = 10.95$ ,  $\phi = 0.0035 \times 365 = 1.2775$ ,  $\eta_C = 0.000105$ , and  $\mu = 0.015$ . Other parameters such as  $\alpha = 0.001 \times 365 = 0.365$ ,  $\eta_A = 1.1$ , and  $\rho = 0.98 \times 365 = 357.7$  are based on available literature and intuitive meaning of the parameters. Given the poor socio-economic conditions of Madagascar, the 90-90-90 HIV treatment target initiative by UNAIDS has not performed remarkably well. At severe stages of HIV infection, available medication has very little effect. This information has led us to opt for such a high value of  $\rho$  and small value for  $\alpha$ . Here,  $\beta$  is the contact rate (per year) of an HIV-infected person with the susceptible pool of individuals and  $\rho$  signifies the rate of progress (per year) of HIV-infected individuals to become AIDS patient. Similarly, we define other mentioned rates.

For this set of parameters, the CCS is obtained as 4585. So we can say that if the population in a commune drops below 4585, the infection will die out automatically. Obviously, the range of CCS for the HIV epidemic is far less than comparatively more infectious diseases such as measles that has much higher incidence rate. UNICEF reports that there were 244,607 cases of measles and 1,080 deaths in Madagascar from August 2018 to November 2019 ([www.unicef.org/press-releases/measles-outbreaks-continue-unabated-five-countries-accounted-nearly-half-all-measles](http://www.unicef.org/press-releases/measles-outbreaks-continue-unabated-five-countries-accounted-nearly-half-all-measles)). Naturally, Näsell [6] obtained CCS as 561,000 (with  $R_0 = 14$ ) and Bartlett [4]



**Fig 1. Relation of CCS with different parameters.**

<https://doi.org/10.1371/journal.pone.0244543.g001>

estimated CCS as 250,000 to 300,000 in England and Wales for measles. Note that, using the above parameters for Madagascar and assuming homogenous mixing within communities, we obtain  $R_0$  as 4.059, which essentially means that in a completely susceptible population, a single infection may produce 4.059 secondary infections in the duration of infectiveness. Our calculation also indicates that temporary eradication or fade-out could happen after 7.365 years unless the infection is reintroduced from outside. This fade-out tends to happen in local spatial regions until the reintroduction of the disease occurs from other areas. However, this does not imply that a person already infected with HIV will be removed from the population within this period. In fact, she/he may continue to exist for more than 7.365 years and we assume that no new infection will be spread by her/him.

Next, we study the nature of CCS by varying one parameter and keeping others fixed at the above values. For this we consider the most important parameters ( $\beta$ ,  $d$ ,  $\phi$ ,  $\omega$ ,  $\alpha$ , and  $\rho$ ) in our model. Fig 1 indicates the relation of CCS with different parameters. Here,  $\beta$  is the contact rate for HIV transmission. An increase of  $\beta$  signifies an increment of infection circulation in the population. So the susceptible population will get infected at a much higher rate. In order to stop the infection, the population size should be very small. So if the number of the susceptible population falls below some threshold, the chance of the spread of infection will diminish. This indicates that smaller the population size, lesser would be the chance of spreading the infection. Thus, it is expected that if  $\beta$  increases, CCS will decrease (Fig 1A).  $d$  is the rate of death from class A. So as  $d$  increases, the spread of infection will decrease. For large values of  $d$ , a large susceptible population will remain unaffected. Hence the CCS value will be high for

large  $d$  (Fig 1B).  $\phi$  and  $\omega$  respectively are the rates at which individuals in class  $I$  move to  $C$  and individuals in class  $C$  move to  $I$ . So, as  $\phi$  increases, the number of individuals receiving treatment increases. It is expected that individuals under treatment will spread less infection than those who are infected but not treated. So, as more individuals go to class  $I$  from class  $C$ , the chance of spreading infection in the susceptible population decreases. But, it is known that the infection stays in a community long enough to engulf the entire susceptible population. Therefore, the infection might be eradicated if the susceptible population decreases. This would mean that, under the condition that  $\phi$  increases, eradication of infection is expected to happen as the CCS value decreases (Fig 1C). On the contrary, as  $\omega$  increases, CCS also increases (Fig 1D). Two other parameters viz.  $\alpha$  and  $\rho$  represent the rates at which individuals in class  $A$  move to  $I$  and individuals in class  $I$  move to  $A$  respectively. For similar reasons, as  $\alpha$  increases, CCS decreases (Fig 1E) but as  $\rho$  increases, CCS increases (Fig 1F).

## Discussion

Bartlett [3, 4] describes fade-out tendency of measles epidemic for small populations. He noticed that for such populations as the size of the susceptible pool of individuals is equal to CCS, measles has a 50% chance of fading out. Bartlett observed that for some cities in the United States where the population size is below CCS, the epidemic ceases to exist in absence of reintroduction of the infection from outside. But for larger cities, the tendency of damping down of successive epidemics is offset by random variability. The idea of epidemic fade-out in smaller spatial locations provides a clue to systematically treat epidemic rather than concentrating treatment for an entire continent or a large community at once.

Shocking statistics ([www.prb.org/thestatusofthehivaidsepidemicinsubaharanafrica](http://www.prb.org/thestatusofthehivaidsepidemicinsubaharanafrica)) reveal that 16 countries in Sub-Saharan Africa constitute nearly 4% of the world's population but account for more than 50% of HIV infections worldwide. In some of these countries, the prevalence rate is up to 30%. However, efforts to control HIV epidemic situations have resulted in a decrease in the disease prevalence in some African countries. Total eradication of highly prevalent diseases such as HIV infection in the context of countries in Sub-Saharan Africa is almost impossible in a very short time period. So we may instill the idea of CCS to systematically eradicate HIV. If we concentrate on the treatment of HIV in a more focussed manner in smaller spatial regions, it may accelerate the HIV annihilation process unless the disease is reintroduced from outside.

When we look at the geographical locations of Uganda, Malawi, Eswatini, and Madagascar, interestingly we find that being at the heart of the Great Lakes region, Uganda is surrounded by three lakes among which one is a fresh water lake. These lakes, to some extent separate Uganda from the rest of the African countries. Malawi again is separated by a freshwater lake (Lake Malawi) that comprises of 25% of its area. But Eswatini is a landlocked country in Southern Africa. We observe that, even with quite impressive progress towards achieving the 90-90-90 target, countries that are more landlocked have higher rates of incidence and prevalence compared to other countries that are isolated to varying extent. Although being in the AIDS belt of Africa, Madagascar has shown remarkably lower rates of incidence and prevalence even with little success of the 90-90-90 target. It is possibly because of the isolated location of Madagascar that separates it from the mainland of the African continent to a great extent.

Metcalf et al. [8] pointed out with respect to some childhood infections that the relative risk of extinction of infections on islands is almost double that in the mainland. Such infections cease to exist at a significantly higher rate in islands where access to/from the mainland is much less. But the persistence of infection within a region in the mainland is fuelled constantly by recolonization. Hence, although the epidemic infection should apparently die out because

of the natural stochastic extinction property (damping down) of itself after invading the total susceptible subpopulation in the mainland, recolonization introduces infection and hinders the damping down process. Thus, our observations are consistent with the views of this paper.

Madagascar has 111 large and small cities or districts with 39.6% of them have population size less than 20000, 30.6% with population size somewhere between 20000 – 30000, 16.2% has 30000 – 40000 and 5.4% have population more than 100000 (<http://worldpopulationreview.com/countries/madagascar-population/cities/>). Further, each district is comprised of several communes of an even smaller population. So our idea is that if the population size falls below the CCS of Madagascar which is 4585 in each spatially separated population, HIV infection may cease to exist in different spatial locations, after a mean time of 7.365 years, provided infection is not re-introduced from outside. But this will definitely not happen without continuing the controlling measures like educating people about the disease and taking necessary preventive measures/treatment available. Through this paper, we are emphasizing that the systemic spreading of control measures could lead to an acceleration in achieving the 90-90-90 goal of UNAIDS.

We observe from our case study that in addition to the existing controlling measures of HIV epidemics in Madagascar, employing the idea of CCS might accelerate the eradication of the infection from smaller populations after the corresponding mean extinction time. Besides, CCS could be looked upon as an index of performance of epidemic controlling strategies in different countries.

## Conclusions

Combating pandemic and epidemic outbreaks require efficient controlling strategies, as the available treatment facilities and/or vaccines become limited in such situations. But, the infection persists in the community long enough to engulf the entire susceptible population. Hence, in large communities, the infection recurs in the subsequent timeline. Interestingly, Bartlett observed that in smaller communities the infection dies out if the susceptible population drops below a threshold size (CCS). Later, Näsell formulated CCS for infectious diseases having Susceptible-Infected model dynamics. But the analytical calculation of CCS for more complex disease dynamics or higher-order models are quite cumbersome. So, we propose a simplified computation based approach for rapid calculation of the approximate value of CCS, for HIV disease dynamics with four differential equations. Since our method is free from any restriction on the number of variables in the system, it could be applied to other diseases with higher-order model dynamics. We envisage CCS for devising control strategies during epidemic outbreaks such as COVID-19 [18]. Because, during epidemic outbreaks with limited treatment facilities, keeping the susceptible population below CCS by enforcing appropriate containment plans might control the disease substantially.

So to sum up, through this paper, firstly, we have generalized the approach for CCS calculation for a system of differential equations having more than two variables. Our method for a system containing four variables may be extended to higher dimensions for any infectious disease dynamics. This is because our method is able to tackle multivariate normal distribution for calculating expected time to extinction as a function of CCS (Result 1). Secondly, we have developed a method to unleash the potential that CCS holds in devising control strategies by promptly determining its approximate numerical value from any disease dynamics using a simplified computation approach. In particular, we believe that our method will provide a suitable measure for accelerated extermination of epidemic infections by aiding policymakers in making decisions on control strategies. Thirdly, we hypothesize that, if we start treating spatially isolated regions or smaller communes and then progress to the district level, and after

that to larger regions systematically, the annihilation of HIV infection could be accelerated. Although it is way difficult to reach out to smaller communes with treatments and other controlling measures (such as educating them about the infection etc.), CCS could be worth giving a trial, as this could utilize the natural force of damping down of epidemic oscillations to eradicate HIV infection from smaller populations.

## Supporting information

**S1 Appendix.**  
(PDF)

## Author Contributions

**Conceptualization:** Sarmistha Das, Sandip Banerjee, Joydev Chattopadhyay, Indranil Mukhopadhyay.

**Data curation:** Sarmistha Das, Pramit Ghosh.

**Formal analysis:** Sarmistha Das, Indranil Mukhopadhyay.

**Funding acquisition:** Saumyadipta Pyne, Indranil Mukhopadhyay.

**Investigation:** Indranil Mukhopadhyay.

**Methodology:** Sarmistha Das, Indranil Mukhopadhyay.

**Project administration:** Indranil Mukhopadhyay.

**Supervision:** Indranil Mukhopadhyay.

**Writing – original draft:** Sarmistha Das, Indranil Mukhopadhyay.

**Writing – review & editing:** Sarmistha Das, Pramit Ghosh, Sandip Banerjee, Saumyadipta Pyne, Joydev Chattopadhyay, Indranil Mukhopadhyay.

## References

1. Soper HE. The interpretation of periodicity in disease prevalence. *Journal of the Royal Statistical Society*. 1929; 92(1):34–73. <https://doi.org/10.2307/2341437>
2. Bartlett M. S. 1956 Deterministic and stochastic models for recurrent epidemics. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (ed. J. Neyman); 1956. p. 81–109.
3. Bartlett MS. Measles periodicity and community size. *Journal of the Royal Statistical Society Series A (General)*. 1957; 120(1):48–70. <https://doi.org/10.2307/2342553>
4. Bartlett M. The critical community size for measles in the United States. *Journal of the Royal Statistical Society: Series A (General)*. 1960; 123(1):37–44. <https://doi.org/10.2307/2343186>
5. Andersson H, Britton T. Stochastic epidemics in dynamic populations: quasi-stationarity and extinction. *Journal of mathematical biology*. 2000; 41(6):559–580. <https://doi.org/10.1007/s002850000060>
6. Nåsell I. A new look at the critical community size for childhood infections. *Theoretical population biology*. 2005; 67(3):203–216. <https://doi.org/10.1016/j.tpb.2005.01.002> PMID: 15808337
7. Nåsell I. Stochastic models of some endemic infections. *Mathematical biosciences*. 2002; 179(1):1–19. [https://doi.org/10.1016/S0025-5564\(02\)00098-6](https://doi.org/10.1016/S0025-5564(02)00098-6) PMID: 12047919
8. Metcalf CJE, Hampson K, Tatem AJ, Grenfell BT, Bjørnstad ON. Persistence in epidemic metapopulations: quantifying the rescue effects for measles, mumps, rubella and whooping cough. *PloS one*. 2013; 8(9):e74696. <https://doi.org/10.1371/journal.pone.0074696> PMID: 24040325
9. Silva CJ, Torres DF. Global stability for a HIV/AIDS model. *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*. 2018; 67:93–101.
10. Silva CJ, Torres DF. A TB-HIV/AIDS coinfection model and optimal control treatment. *Discrete & Continuous Dynamical Systems A*. 2015; 35(9):4639–4663. <https://doi.org/10.3934/dcds.2015.35.4639>

11. Djordjevic J, Silva CJ, Torres DF. A stochastic SICA epidemic model for HIV transmission. *Applied Mathematics Letters*. 2018; 84:168–175. <https://doi.org/10.1016/j.aml.2018.05.005>
12. Arif MS, Raza A, Rafiq M, Bibi M. A reliable numerical analysis for stochastic hepatitis B virus epidemic model with the migration effect. *Iranian Journal of Science and Technology, Transactions A: Science*. 2019; 43(5):2477–2492. <https://doi.org/10.1007/s40995-019-00726-0>
13. Arif MS, Raza A, Rafiq M, Bibi M, Fayyaz R, Naz M, et al. A reliable stochastic numerical analysis for typhoid fever incorporating with protection against infection. *Computers, Materials & Continua*. 2019; 59(3):787–804. <https://doi.org/10.32604/cmc.2019.04655>
14. Raza A, Arif MS, Rafiq M. A reliable numerical analysis for stochastic dengue epidemic model with incubation period of virus. *Advances in Difference Equations*. 2019; 2019(1):32. <https://doi.org/10.1186/s13662-019-1958-y>
15. Heffernan JM, Smith RJ, Wahl LM. Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*. 2005; 2(4):281–293. <https://doi.org/10.1098/rsif.2005.0042> PMID: 16849186
16. Näsell I. On the time to extinction in recurrent epidemics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 1999; 61(2):309–330. <https://doi.org/10.1111/1467-9868.00178>
17. Dietz K. Transmission and control of arbovirus diseases. *Epidemiology*. 1975; 104:121.
18. Das S, Ghosh P, Sen B, Pyne S, Mukhopadhyay I. Critical community size for COVID-19: A model based approach for strategic lockdown policy. *Statistics and Applications*. 2020; 18(1):181–196.