

## Regional variation and epidemiological insights in malaria underestimation in Cameroon

Sarafa A. Iyaniwura<sup>1</sup>, Qing Han<sup>2,3</sup>, Ngem Bede Yong<sup>6</sup>, Ghislain Rutayisire<sup>4</sup>, Agnes Adom-Konadu<sup>5</sup>, Okwen Patrick Mbah<sup>7</sup>, David Poumo Tchouassi<sup>8</sup>, Kingsley Badu<sup>9</sup>, and Jude D. Kong<sup>2,3\*</sup>

**1** Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

**2** Laboratory for Industrial and Applied Mathematics (LIAM), Department of Mathematics and Statistics, York University, Toronto, M3J 1P3, Ontario, Canada.

**3** Africa-Canada Artificial Intelligence and Data Innovation Consortium (ACADIC), York University, Toronto, M3J 1P3, Ontario, Canada.

**4** Department of Mathematics and Statistics, York University, Toronto, M3J 1P3, Ontario, Canada.

**5** Department of Mathematics, University of Cape Coast, Cape Coast, Central Region, Ghana.

**6** Department of Public Health, Regional Delegation of Public Health North West Region, Bamenda, Cameroon.

**7** Effective Basic Services Africa (eBASE), Research and Development, Bamenda, North West Region, Cameroon.

**8** International Centre of Insect Physiology and Ecology, Nairobi, Kenya.

**9** Department of Theoretical and Applied Biology, Kwame Nkrumah University of Science and Technology, University Post Office, Kumasi AK 039, Ghana.

 These authors contributed equally to this work.

\*Corresponding author: [jdkong@yorku.ca](mailto:jdkong@yorku.ca) (JK)

### Abstract

Malaria, caused by *Plasmodium* parasites and transmitted by female *Anopheles* mosquitoes, is most common in tropical regions, especially in Sub-Saharan Africa. Despite significant global effort to control and eradicate the disease, many cases and deaths are still reported yearly. These efforts are hindered by several factors, including the severe underestimation of cases and deaths, especially in Africa, making it difficult to assess the disease burden accurately. We used a mathematical model of malaria, incorporating the underestimation of cases and seasonality in mosquito biting rate, to study the disease dynamics in Cameroon. Using a Bayesian inference framework, we calibrated our model to the monthly reported malaria cases in ten regions of Cameroon from January 2019 to December 2021 to quantify the underestimation of cases and estimate other important epidemiological parameters. We performed Hierarchical Clustering on Principal Components analysis to understand regional disparities, looking at underestimation rates, population sizes, healthcare personnel, and healthcare facilities per 1,000 people. We found varying levels of underestimation of cases across regions, with the East region having the lowest underestimation (14%) and the Northwest region with the highest (70%). The mosquito biting rate peaks once every year in most of the regions, except in the Northwest region where it peaks every 6.02 months and in Littoral every 15 months. We estimated a median mosquito biting rate

of over five bites per day for most of the regions with Littoral having the highest (9.86 bites/day). Two regions have rates below five bites per day: Adamawa (4.78 bites/day) and East (4.64 bites/day). The notably low estimation of malaria cases in Cameroon underscore the pressing requirement to bolster reporting and surveillance systems. Regions in Cameroon display a range of unique features, which may contribute to the differing levels of malaria underestimation. These distinctions should be considered when evaluating the efficacy of community-based interventions.

## Author summary

- i. We used a deterministic mathematical model of malaria that incorporated the underestimation of cases and seasonality in the biting rate of mosquitoes to retroactively study the dynamics of the disease in Cameroon from January 2019 to December 2021.
- ii. We found varying levels of underestimation of malaria cases across regions in Cameroon, with the East region having 14% underestimation and the Northwest region having 70%.
- iii. We found consistent malaria-induced death rates and natural immunity duration across Cameroon. We estimated that the mosquito biting rate for the Northwest region oscillated with a period of 6.02 months, while those of the remaining regions had a period of 12 months or more. Most regions had median mosquito biting rates exceeding five bites per day, with the Littoral having the highest (9.86 bites/day). In comparison, two regions had rates below five bites per day: Adamawa (4.78 bites/day) and East (4.64 bites/day).
- iv. We clustered the ten regions into four major groups using the case underestimation rate, population size, total healthcare human resources per 1,000, and total healthcare facilities per 1,000.

## Introduction

Malaria is a mosquito-borne disease caused by *plasmodium* parasites and transmitted by female *Anopheles* mosquitoes. It is common in sub-Saharan Africa and other tropical regions of the world [1–3]. There are 123 known species of the genus *Plasmodium*. However, only five of them are known to cause infections in human: *P. falciparum*, *P. malariae*, *P. ovale*, *P. vivax*, and *P. knowlesi* [7–11]. The most prevalent and dangerous malaria parasite is *Plasmodium falciparum* [12]. It is responsible for a significant portion of malaria-related fatalities [7, 8, 13]. *Plasmodium* parasites are mainly spread through mosquito bites, but can also be spread from a malaria-infected pregnant individual to their fetus through the placenta, and through blood transfusion [14]. Individuals infected with malaria may have symptoms like fever, rigors, and chills, in the case of uncomplicated malaria [15]. However, severe malaria mostly in children under five may be presented as fever, impaired consciousness, severe anaemia, respiratory distress, convulsions, and hypoglycemia, among other symptoms [8, 16]. Symptoms usually start 10 to 15 days after the initial mosquito bite, however for infections involving some strains of *plasmodium vivax*, symptoms could delay for 8 to 10 months or longer [17].

About half of the world's population, the majority of whom reside in Africa, are susceptible to malaria and dealing with its economic hardships. According to the WHO estimates, there were 247 million malaria cases and 619,000 deaths in 2021, where 234 million (~ 95%) of the cases and 593,000 (~ 96%) of the deaths were in Africa [12]. In

21 addition,  $\sim 79\%$  of the deaths were in children under five years [12]. Despite recent  
22 advancements in the fight against malaria, the disease is still widespread with over 2  
23 million reported cases in Cameroon yearly, posing a serious public health  
24 concern [18–22]. It is identified as the third leading cause of death in healthcare centers  
25 in Cameroon, accounting for 30% of the mortality [23].

26 The burden of malaria differs geographically depending on the local malaria  
27 transmission intensity [24]. Underestimation of cases can be in terms of  
28 under-ascertainment (when not all patients seek healthcare), under-diagnosis (when  
29 cases seek healthcare but are not properly diagnosed) and/or under-reporting or  
30 under-notification (when diagnosed cases are not adequately reported and notified) [25].  
31 The burden of asymptomatic malaria, i.e., infection with *P. falciparum* in the absence  
32 of overt clinical symptoms, is huge [3]. Individuals may harbor malaria parasites but  
33 due to the absence of symptoms will not report to any facility for diagnosis or treatment  
34 and thus will not be counted. In addition, some individuals may self-treat, either with  
35 local herbs or through medication from undocumented sources. These may lead to the  
36 under-ascertainment of cases. Furthermore, infected individuals may visit hospitals but  
37 their cases go undetected due to low parasite load and/or the sensitivity of the testing  
38 device [26]. Recently, Afriyie et al. [26] found that up to 40% of individuals diagnosed  
39 as negative in the lab by microscopy and rapid diagnostic tests (RDT) in the Ashanti  
40 Region of Ghana, were identified as positive under research settings when tested with  
41 molecular techniques with higher sensitivity. Such individuals with undetected infection  
42 bring concern to malaria elimination in Africa, as they present as a reservoir for  
43 mosquitoes to pick up infections and spread to others.

44 One of the earliest mathematical models used to study malaria transmission is the  
45 "Ross-Macdonald model" developed in 1915 [27]. Over the years, this model has led to  
46 the development of other deterministic mathematical models used to gain insight into  
47 the transmission dynamics of malaria [7, 28, 29, 29–31]. Of recent, Ndamuzi et al. [7]  
48 proposed a deterministic model for studying the dynamics of malaria parasite in  
49 mosquito and human populations. They demonstrated the need for effective control  
50 measures to decrease the number of mosquito bites per individual per unit time and the  
51 population of mosquitoes, among other factors to adequately control the spread of  
52 malaria. Another malaria dynamics model with age-structure in mosquito population  
53 was developed by Bakary et al. [10]. The model divides the human population into  
54 non-immune and semi-immune, where the non-immune are the most vulnerable, while  
55 the semi-immune are the least vulnerable individuals. They analyzed the model  
56 mathematically and studied the stability properties of its steady-states. The model of  
57 Collins et al. [32] was used to study the dynamics of malaria in Nigeria. The model  
58 incorporates the treatment of malaria, drug resistance, and the usage of mosquito nets  
59 as a preventive measure. Their finding suggests that unless better control efforts are  
60 directed at the dominant resistant strain of the disease, treatment is improved, and  
61 mosquito nets are used widely, malaria is likely to remain endemic in Nigeria.

62 The underestimation of malaria cases has been reported in some regions around the  
63 world [46–49, 51, 52]. In an India forest community, Chourasia et al. [46] estimated that  
64 malaria cases maybe have been underestimated by up to 20% and 22.8% in 2013 and  
65 2014, respectively, due to asymptomatic cases. A cross-sectional study by Nankabirwa et  
66 al. [49] indicated that up to 39.9% of malaria cases in children under five years old are  
67 undiagnosed in Uganda. The under-reporting of malaria cases in Europe was estimated  
68 to vary between 20% to 59% by Legros and Danis [51]. This estimate is consistent with  
69 the 33.33% rate estimated for the Netherlands in 1996 by Van Hest et al. [52]. Overall,  
70 there are limited studies on the underestimation of malaria cases in the literature. To  
71 our knowledge only two studies have looked at the underestimation of malaria cases in  
72 Africa [48, 49]. To better understand the magnitude of malaria epidemic in Africa, it is

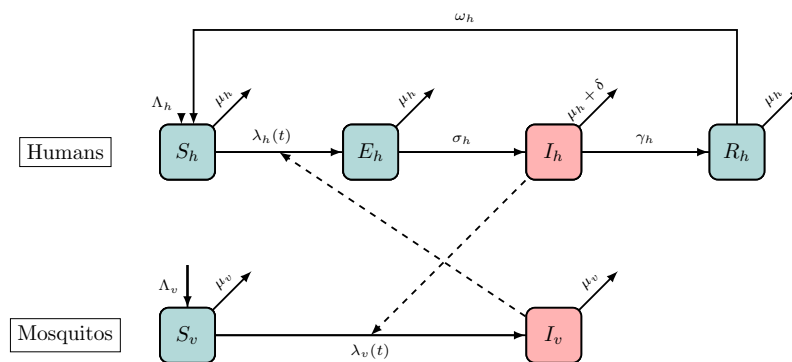
essential to understand the degree of underestimation of cases and deaths. 73

In this study, we used a deterministic mathematical model to study the dynamics of malaria in Cameroon. Our model incorporates the human and mosquito population, and a seasonal mosquito biting rate. Using a Bayesian inference framework, we fitted our model to the reported cases of malaria in each of the ten regions of Cameroon, from January 2019 to December 2021, from which we estimated the underestimation of malaria cases in each region, and other key epidemiological parameters. In addition, through Hierarchical Clustering on Principal Components analysis (HCPC), we clustered the ten regions based on the quantified underestimation rates and other regional characteristics. 74  
75  
76  
77  
78  
79  
80  
81  
82

## Methods 83

### Mathematical model 84

We used a deterministic compartmental model to study the dynamics of malaria in the ten geographical regions of Cameroon, following the framework of [10, 33–36]. Our model considers the human and mosquito population in each region, dividing the human population into four compartments: susceptible ( $S_h$ ), exposed ( $E_h$ ), infectious ( $I_h$ ), and recovered ( $R_h$ ), and the mosquito population into two compartments: susceptible ( $S_v$ ) and infected ( $I_v$ ) (see Fig 1). Although many of the malaria models in the literature explicitly consider the population of mosquitoes in the exposed stage of the disease [35, 36], these mosquitoes are included in the infected compartment ( $I_v$ ) in our model. 85  
86  
87  
88  
89  
90  
91  
92  
93



**Fig 1. Schematic illustration of the compartmental model.** An illustration of our model incorporating human and mosquito populations. The human population is divided into susceptible ( $S_h$ ), exposed ( $E_h$ ), infectious ( $I_h$ ), and recovered ( $R_h$ ) populations, while the mosquito population is divided into susceptible ( $S_v$ ) and infected ( $I_v$ ) mosquitoes. Solid black arrows show the transition of humans and mosquitoes through the different stages of malaria at the rate beside the arrows. Dashed black arrows show malaria transmission from humans to mosquitoes and vice versa. 94  
95  
96  
97  
98  
99  
100  
101

In our model, a susceptible human becomes infected with malaria at a probability  $\beta_h$  following a bite from an infected mosquito. Similarly, a susceptible mosquito becomes infected with malaria at a probability  $\beta_v$  upon biting an infected human. After a successful bite and malaria transmission from a mosquito to a human, the malaria parasite in this individual undergoes an incubation period during which the infected humans cannot transmit the parasite to susceptible mosquitoes [37]. The incubation period lasts for a couple of days depending on the initial dose of the *plasmodium* parasite injected into the human by the mosquito [37]. This period is captured by the 94  
95  
96  
97  
98  
99  
100  
101

exposed compartment ( $E_h$ ) in our model. After this period, the infected individual becomes infectious and can transmit the parasite to susceptible mosquitoes. We assumed a constant influx of humans into each regions through birth and/or immigration at the rate  $\Lambda_h$ , and similarly for mosquitoes at the rate  $\Lambda_v$ . We did not explicitly model mosquito habitat and assumed that all the mosquitoes considered are the female *Anopheles* mosquitoes that can be infected with and transmit malaria [38, 39]. In addition to the natural human death rate  $\mu_h$ , we considered malaria-induced deaths in the human population at rate  $\delta$ . To capture the seasonality of the malaria epidemic in each region, we assumed a periodic change in the biting rate of mosquitoes. We also incorporated the waning of infection-acquired temporary immunity in humans at rate  $\omega_h$ , which we assumed varies between the regions. The differential equations of the model are given by

$$\begin{aligned}
 \frac{dS_h}{dt} &= \Lambda_h - \lambda_h S_h + \omega_h R_h - \mu_h S_h, \\
 \frac{dE_h}{dt} &= \lambda_h S_h - \sigma_h E_h - \mu_h E_h, \\
 \frac{dI_h}{dt} &= \sigma_h E_h - \gamma_h I_h - (\mu_h + \delta) I_h, \\
 \frac{dR_h}{dt} &= \gamma_h I_h - \omega_h R_h - \mu_h R_h, \\
 \frac{dS_v}{dt} &= \Lambda_v - \lambda_v S_v - \mu_v S_v, \\
 \frac{dI_v}{dt} &= \lambda_v S_v - \mu_v I_v,
 \end{aligned} \tag{1}$$

where  $\lambda_h \equiv \lambda_h(t)$  and  $\lambda_v \equiv \lambda_v(t)$  are the time-dependent forces of infection for the human and mosquito populations, respectively, given by

$$\lambda_h(t) = b \beta_h \frac{I_v}{N_h} \quad \text{and} \quad \lambda_v(t) = b \beta_v \frac{I_h}{N_h}. \tag{2}$$

In Eq 2,  $N_h$  denotes the human population size,  $b$  denotes the biting rate of one mosquito per unit time, and  $\beta_h$  ( $\beta_v$ ) denotes the probability of malaria transmission to human (to a mosquito) upon one infectious bite. We assumed that one mosquito bites the same human only once during its lifetime.

To account for the seasonality in the biting rate of mosquitoes, we used a periodic function for  $b \equiv b(t)$ , given by

$$b(t) = b_0 \left[ 1 + \varepsilon \cos \left( 2\pi\omega(t - t_c) \right) \right], \tag{3}$$

where  $b_0$  is the average biting rate of mosquitoes,  $\varepsilon$  is the degree of variation around the average biting rate,  $1/\omega$  is the period for seasonal malaria epidemic and  $t_c$  is the phase shift in the periodicity. All model parameters, their descriptions, and values are presented in Table 1.

Using the next-generation matrix approach [43], we derived the time-dependent effective reproduction number  $\mathcal{R}_e(t)$  of malaria for our model. Details of the derivation and the  $\mathcal{R}_e(t)$  computed for each region using the median estimated parameters are presented in the S1 File.

## Data

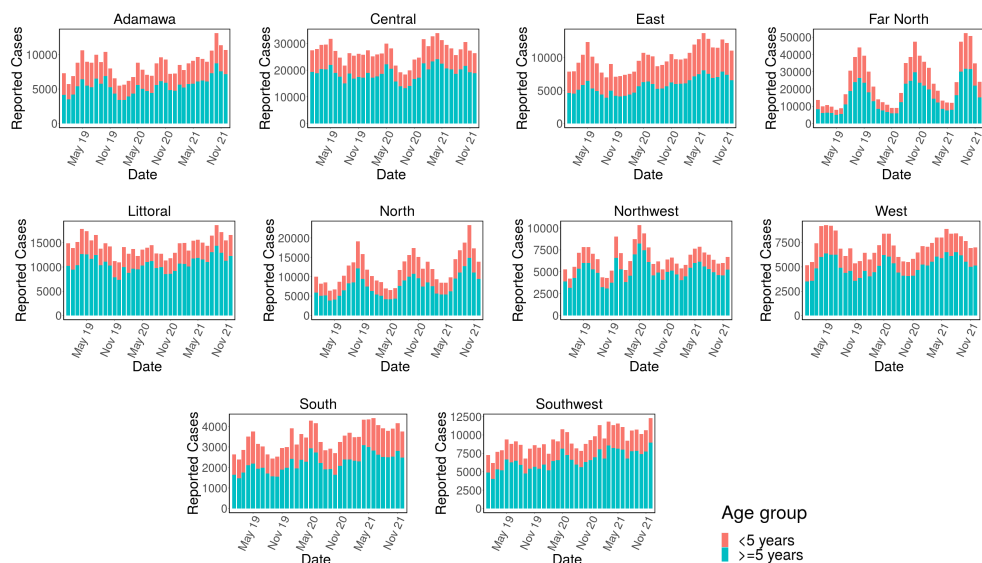
We considered the monthly reported malaria cases in the ten geographical regions of Cameroon: Adamawa, Central, East, Far North, Littoral, North, Northwest, West,

**Table 1. Descriptions and values for model variables and parameters.**

Variable	Description	Initial value	Source
$S_h$	Susceptible human population	$S_h(0) = N_h(0) - (E_h(0) + I_h(0))$	Assumed
$E_h$	Exposed human population	See S2 File	Fitted
$I_h$	Infectious human population	See S2 File	Fitted
$R_h$	Recovered human population	$R_h(0) = 0$	Assumed
$N_h$	Total human population	See S2 File	[40]
$S_v$	Susceptible mosquito population	$S_v(0) = 5,000$	Assumed
$I_v$	Infected mosquito population	See S2 File	Fitted
Parameter	Description	Value	Source
$\lambda_h(t)$	Force of infection for human	See Eq 2	-
$\lambda_v(t)$	Force of transmission for mosquitoes	See Eq 2	-
$b(t)$	Time-dependent mosquito biting rate	See Eq 3	-
$\Lambda_h$	Human birth/immigration rate	See S2 File	[40]
$\Lambda_v$	Mosquito birth/immigration rate	See S2 File	Fitted
$\beta_h$	Probability of malaria transmission to human per bite	0.022	[41]
$\beta_v$	Probability of malaria transmission to mosquito per bite	0.48	[41]
$\omega_h$	Region-specific waning rate of malaria natural immunity in human	See S2 File	Fitted
$\mu_h$	Human death rate	0.0014 month <sup>-1</sup>	[42]
$\mu_v$	Mosquito death rate	0.033 day <sup>-1</sup>	[41]
$\sigma_h$	Disease progression rate from exposed to infected for human	0.1 day <sup>-1</sup>	[41]
$\gamma_h$	Human recovery rate	0.0035 day <sup>-1</sup>	[41]
$\delta$	Human death rate due to malaria	See S2 File	Fitted
$p$	Fraction of undetected malaria cases	See S2 File	Fitted
$b_0$	Average biting rate of a mosquito	See S2 File	Fitted
$\varepsilon$	Degree of variation around the average biting rate of mosquitoes	See S2 File	Fitted
$\omega$	Frequency of seasonality in mosquito biting rate	See S2 File	Fitted
$t_c$	Phase shift in the periodicity of mosquito biting rate	See S2 File	Fitted

South, and Southwest, from January 2019 to December 2021, as obtained from DHIS2 [44]. The data is stratified into two age groups: < 5 years old and ≥ 5 years old (see Fig 2). However, in this study, we combined the data for the two age groups to obtain the total estimated cases of malaria for each region, which is used in our model calibration.

Fig 2 shows the bar plots of the monthly reported cases of malaria for each region for the < 5 years old (red bars) and those ≥ 5 years old (green bars). The seasonality in the total cases of malaria in each region can easily be seen. Many of the regions have only one (usually large) outbreak of malaria each year. However, the Northwest region has multiple smaller outbreaks yearly.



**Fig 2. Monthly reported cases of malaria.** The monthly reported cases of malaria in different geographical regions of Cameroon from January 2019 to December 2021, for different age groups, obtained from DHIS2 [44]. Red bars represent the reported cases for < 5 years age-group, and the green bars is for the  $\geq 5$  years age-group.

## Bayesian inference

We fitted our model in Eq (1) to the monthly reported cases of malaria in each region independently using a Bayesian inference framework and the RStan package in R version 3.6.3 [45]. Although the data are stratified by age groups when presented in Fig 2, we used the total reported cases (without age stratification) in our fit. In our calibration, we defined the likelihood function for each region as

$$\text{ReportedCases}(t) \sim \text{NegBin}(\text{incidence}(t)(1 - p), \phi), \quad (4)$$

where  $\text{NegBin}(\cdot)$  is the negative binomial distribution,  $\text{ReportedCases}(t)$  is the total monthly reported cases of malaria for the region obtained from the data in Fig 2,  $\text{incidence}(t)$  is the monthly incidence of malaria computed using our model Eq 1 (it is computed as the number of exposed individuals transitioning to the infectious compartment each month),  $p$  is the fraction of cases undetected (i.e.,  $1 - p$  is the fraction of malaria cases that was detected), and  $\phi$  is the dispersal parameter.

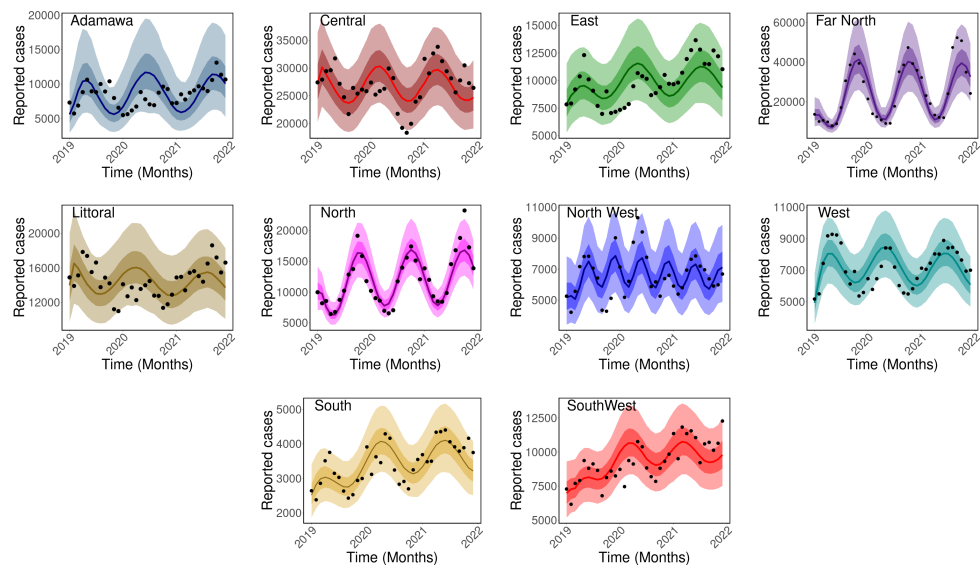
The Bayesian inference framework gives us the flexibility to incorporate our prior knowledge into the model parameters and the ability to evaluate probabilistic statements of the data based on the model. We implemented uninformative priors when fitting our model to the case data. To ensure that our model is coded accurately in the Stan and also for validation, we generated synthetic case data using our model with known parameter values. We then calibrated our model to the synthetic dataset, and inspected the resulting posterior distributions for biases and coverage of the true parameter values used to generate the data. We also checked for the identifiability of the parameters. We used the adaptive Hamiltonian Monte Carlo method No-U-Turn sampling (NUTS) as implemented in RStan with 5,000 iterations and four chains for our model calibrations. From our fit, for each region, we estimated the fraction of underestimated cases of malaria ( $p$ ), the average biting rate of a mosquito ( $b_0$ ), the malaria-induced death rate for humans ( $\delta$ ), and other important model parameters (see Results Section). These parameters are used to compute the time-dependent effective

reproduction number  $\mathcal{R}_e$  of malaria for each region (See S1 File).

## Results

### Data fitting

We calibrated our model in Eq (1) to the monthly reported cases of malaria in each of the ten regions of Cameroon (see Fig 2) independently using a Bayesian inference framework. The results of our inference showing the monthly reported malaria cases (black dots), together with the median predicted cases, obtained from our model fit (solid lines), are presented in Fig 3. The narrower (darker) bands are 50% credible intervals (CrI), while the wider (lighter) bands are 90% CrI.



**Fig 3. Observed and estimated monthly malaria cases.** The monthly reported malaria cases (black dots) and median predicted cases (solid lines) for each region of Cameroon. The darker bands are the 50% CrI, while the lighter bands are the 90% CrI.

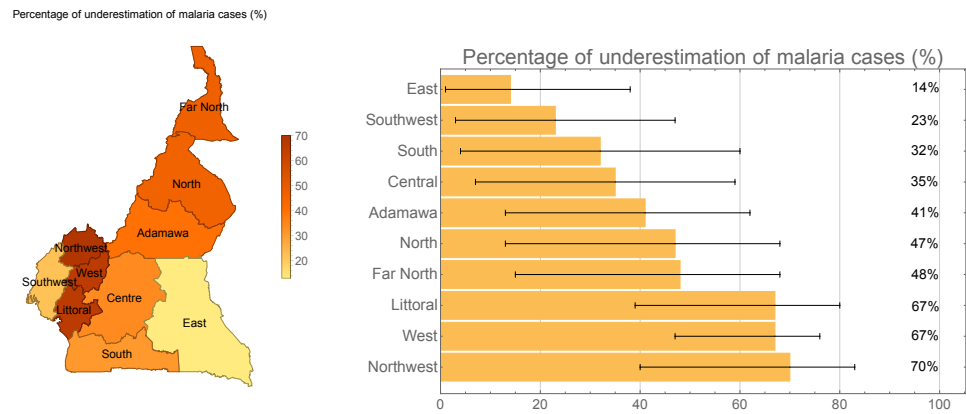
### Quantifying the underestimation of malaria cases

From our calibration, we estimated the percentage of underestimated cases of malaria in each region, which varies significantly between the regions (Fig 4). The smallest percentage was estimated for the East region with a median of 14% (90% CrI: 1% - 38%), while the largest was estimated for the Northwest region, 70% (90% CrI: 40% - 83%). Following the Northwest region is the Littoral and West regions, with similar rates, 67% (90% CrI: 39% - 80%) and 67% (90% CrI: 47% - 76%), and then Far North, 48% (90% CrI: 15% - 68%) and North, 47% (90% CrI: 13% - 68%).

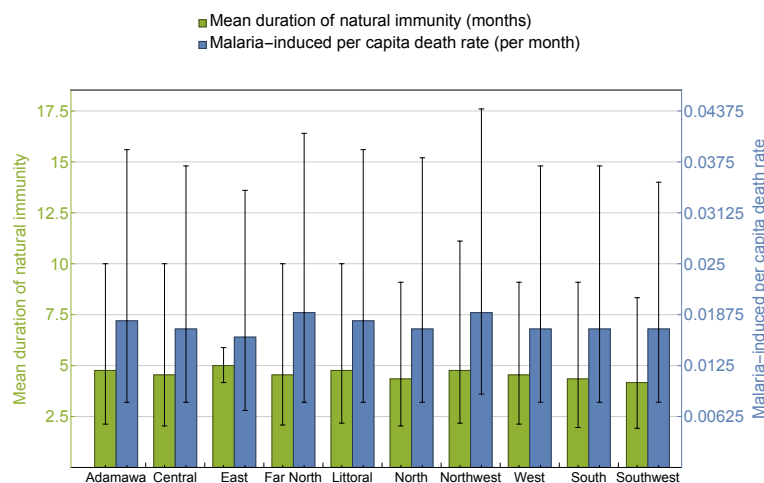
We present five key malaria parameters estimated from our fit in Fig 5 and Fig 6: the mean duration of natural immunity ( $1/\omega_h$ ) in months, malaria-induced per capita death rate ( $\delta$ ), the period of seasonal mosquito biting rate ( $1/\omega$ ) in months, the baseline (or average) mosquito biting rate ( $b_0$ ) per day, and the degree of variation around baseline biting rate ( $\varepsilon$ ) in percentage.

The estimated malaria-induced death rate is similar for the regions ranging from a median of 0.016/month (90% CrI: 0.007-0.034) for the East region to 0.019/month for





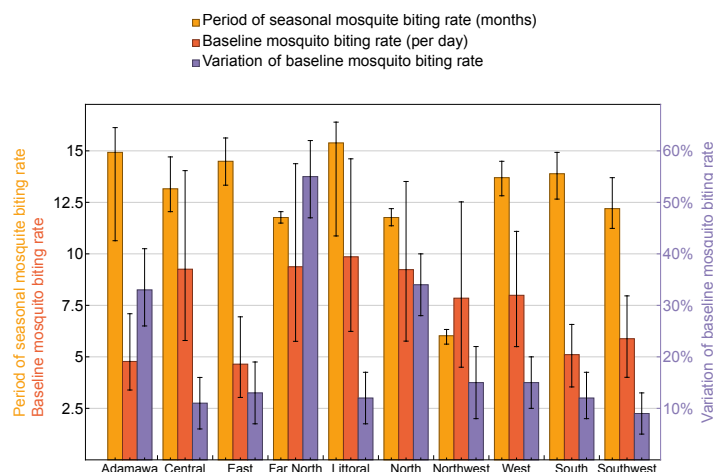
**Fig 4. Percentage of malaria cases underestimation.** The percentage of underestimation of malaria cases estimated for each region. Left panel: map visualization with the median underestimation percentage indicated with color intensity: high (reddish) and low (yellowish). Right panel: barplots showing the percentage of underestimation for each region with 90% CrI (in ascending order).



**Fig 5. Estimated region-specific malaria parameters: mean duration of natural immunity in months and malaria-induced per capita death rate per month.** The bars are the median estimated values and error bars are for the 90% CrI.

Far North (90% CrI: 0.008-0.041) and Northwest (90% CrI: 0.009-0.044). The mean duration of natural immunity is also consistent among the regions (see Fig 5). The East region has a natural immunity duration of 5 months (90% CrI: 4.17-5.88) and the Southwest region has the shortest duration of 4.17 months (90% CrI: 1.92-8.33). It is important to note that the East region has the highest estimation rates of malaria cases while it has the lowest malaria-induced death rates and longest natural immunity duration.

We estimated that the mosquito biting rate for the Northwest region oscillated with a period of 6.02 months (90% CrI: 5.62-6.33), while those of the other regions oscillated with a period of at least 12 months, with the biting rate of Littoral having the largest periodicity of 15.38 months (90% CrI: 10.87-16.39). The low periodicity of mosquito biting rate in the Northwest region may have contributed to the high underestimation of malaria cases estimated for this region. Most of the regions have an estimated mosquito



**Fig 6. Estimated region-specific malaria parameters: period of oscillation of mosquito biting rate in months, baseline mosquito biting rate per day and variation of baseline mosquito biting rate in percentage.** The bars represent the median estimated values and error bars are for the 90% CrI.

biting rate above 5 bite/day, with the highest estimated for the Littoral region with a median of 9.86 bites/day (90% CrI: 6.24-14.61). Two regions have biting rates below 5 bites/day, and they are the Adamawa region with a median of 4.78 bites/day (90% CrI: 3.39-7.09), and the East region with 4.64 bites/day (90% CrI: 3.03-6.94). The degree of variation around baseline mosquito biting rate was estimated to be mostly between 9% and 35%, with Far North having the only exceptionally large variation of 55% (90% CrI: 47%-62%), and Southwest having the least variation of 9% (90% CrI: 5%-13%).

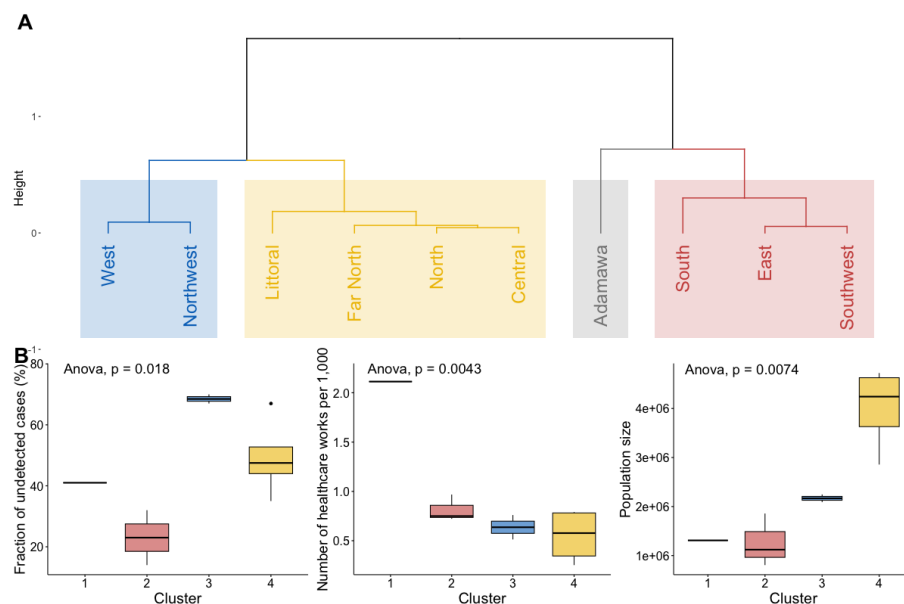
### Clustering of regions

We used the estimated malaria underestimation rates, the population sizes, total numbers of healthcare human resources per 1,000, and total numbers of healthcare facilities per 1,000 to cluster the ten Cameroon regions through Hierarchical Clustering on Principal Components (HCPC).

HCPC resulted in four major clusters as shown in Fig 7A, indicated by different colors. Cluster 1 includes Adamawa alone, cluster 2 includes the South, East and Southwest regions, cluster 3 contains the West and Northwest regions, and cluster 4 has Littoral, Far North, North and Central. Among the four factors used for this clustering, the fraction of undetected cases ( $p=0.018$ ), number of healthcare human resources ( $p=0.0043$ ) and population size ( $p=0.0074$ ) are significantly different across clusters from ANOVA test, as shown in Fig 7B. However, these clusters do not differ in number of healthcare facilities ( $p=0.19$ ).

### Discussion

We have used a deterministic mathematical model to quantify the underestimation of malaria cases in ten regions of Cameroon. The research aimed to estimate essential epidemiological parameters and reveal the extent of underestimation in malaria dynamics across the country. Our findings showed significant variations in the degree of underestimation, ranging from 14% in the East region to a substantial 70% in the Northwest region. We observe from our model fit that the Northwest region has the



**Fig 7. Results from Hierarchical Clustering on Principal Components (HCPC).** A: Dendrogram showing the clustering of the Cameroon regions. B: Significant variables characterizing the clusters, with p-values from ANOVA test.

largest number of smaller epidemics that occurred during our study period of January 2019 to December 2021, among the regions. Although these epidemics do not result in high number of malaria cases like those reported in other regions (Adamawa, Central, East, Far North, Littoral North, and Southwest), the frequency of the epidemics may have resulted in the high percentage of underestimation obtained for this region.

Furthermore, the research indicated a consistent malaria-induced death rate across Cameroon, with the East region reporting the lowest rate at 0.016/month, while the Far North and Northwest regions recorded the highest at 0.019/month. The mean duration of natural immunity displayed uniformity among regions, with all regions showing a duration of less than five months. Notably, the East region had a relatively longer natural immunity duration of 5 months, while the Southwest region had the shortest at 4.17 months. Moreover, the study revealed that only the Northwest region exhibited a minor period of oscillation for the mosquito biting rate (6.02 months), while all other regions demonstrated oscillation periods greater than or equal to 12 months. The Littoral region exhibited the longest periodicity at 15.38 months, indicating substantial variation in mosquito activity across Cameroon.

Regarding mosquito biting rates, most regions had estimated rates exceeding 5 bites per day, with the Littoral region topping the list at a median of 9.86 bites per day. In contrast, only two regions reported biting rates below 5 bites per day, specifically the Adamawa region (4.78 bites/day) and the East region (4.64 bites/day).

To gain insights into regional disparities, we conducted Hierarchical Clustering on Principal Components analysis, considering undetected fractions, total population size, total healthcare human resources per 1000, and total healthcare facilities per 1000. This analysis resulted in four major clusters: Cluster 1 consisted of Adamawa alone, Cluster 2 included the South, East, and Southwest regions, Cluster 3 contained the West and Northwest regions, and Cluster 4 comprised the Littoral, Far North, North, and Central regions. Significantly, the fraction of undetected cases ( $p=0.018$ ), the number of healthcare human resources ( $p=0.0043$ ), and population size ( $p=0.0074$ ) exhibited notable differences across these clusters, as determined by the ANOVA test. However,

the number of healthcare facilities did not display significant variation ( $p=0.19$ ) among the clusters. 262

In the current academic literature, there have been a limited number of efforts to assess the hidden impact of malaria in both malaria-prone and non-malaria-prone regions worldwide. For instance, in malaria-prone areas, Chourasia et al. [46] estimated the burden of asymptomatic malaria among a tribal population in a forested village in central India. They used peripheral blood smears from 134 and 159 individuals in the village during 2013 and 2014. The authors found that the prevalence rates of asymptomatic malaria were 20% and 22.8% for the two years, respectively. Yadav et al. [47] conducted a retrospective exploration of surveillance data and health records from major public and private health facilities in Ahmedabad city, India. They observed that the detected malaria incidence (37,431 cases) was nine times higher than what was officially reported (4,119 cases), meaning that only 11% of the total malaria cases were reported. In the Ifanadiana District in Madagascar, passive surveillance was estimated to have reported 1 in every 5 malaria cases among all individuals (20%), and 1 in every 3 cases among children under five (33.33%) from 2014 to 2017 [48]. Nankabirwa et al. [49] conducted a cross-sectional study that revealed 39.9% of children under five years old in Uganda are under-diagnosed for malaria. Breman [50] observed that less than 20% of malaria incidence cases in sub-Saharan Africa reach the formal health system. In non-malaria-prone regions of the world, Legros and Danis [51] estimated the under-reported rate of malaria in Europe varies from 20% to 59%. This is consistent with Van Hest et al. [52] estimating that approximately one-third (33.33%) of cases in the Netherlands in 1996 went unreported and Cartcath et al. estimating that 34% of cases in England went unreported from 2003-2004. Based on these studies, there isn't much disparity in malaria underestimation between malaria-prone regions and non-malaria-prone regions. Our study reveals that the mean undetected malaria cases across Cameroon is 44.4% (S.D. 19.24%), which aligns with the undetected values in the studies above. 263-289

The studies mentioned were identified through searches in two major scholarly electronic databases in the biomedical field, namely, MEDLINE via its openly accessible interface, PubMed, and Scopus. To ensure we captured all relevant studies that hadn't been indexed, we also conducted searches on Google Scholar. The following search strings were utilized: ("malaria" OR "malarial fever") AND( ("Underestimate" OR "Reporting gap" OR "Hidden cases" OR "Under-declared" OR "Unrecorded" OR "Missed reports" OR "Under-ascertainment" OR "Under ascertainment" OR "Reporting bias" OR "Undercount" OR "under reporting" OR "underreporting" OR "underascertainment") OR ("misdiagnosis" OR "Missed diagnosis" OR "False negative" OR "Incorrect diagnosis" OR "Underdiagnosis" OR "Failed detection" OR "Diagnostic error" OR "Diagnostic bias" OR "Undetected" OR "Under-assessment") OR ("Under-assess" OR "Under-rate" OR "underrate" OR "Under-ascertain" OR "Underestimate bias" OR "Inaccurate assessment" OR "under-estimation" OR "Underestimation")OR ("Non-reporting" OR "Unreported " OR "Under-declared " OR "Underdeclared" OR "Missed notification" OR "Notification deficit" OR "Silent reporting" OR "Notification bias" OR "Notification-bias" OR "Underdocumented" OR "Under-registered" OR "Underestimated"))AND ("mathematical model" OR "statistical model" OR "model-based study" OR "modeling study" OR "model\*"). 290-307

We found that in the existing scholarly literature, there is no mechanistic model for estimating the underestimated proportions of malaria. As a result, our model is the first mechanistic model designed for this purpose. The observation that underestimated fractions and the number of healthcare personnel are variables defining clusters is quite significant and raises some important points for investigation. The significance of underestimated fractions within clusters alongside a low number of healthcare personnel 308-313

suggests a potential correlation. It may imply that regions with fewer healthcare professionals are more likely to under-estimate cases. 314 315

Our study has several limitations that should be acknowledged, including the limited variables chosen for clustering analysis and the omission of an assessment of the factors responsible for underestimation. Future studies should explore the impact of a wide array of socio-economic, climatic, epidemiological, health system-related, demographic, and clinical public health variables on the underestimation of malaria. Further research is warranted to investigate the impact of climatic factors, such as rainfall and temperature on the dynamics of malaria across Cameroon. Seasonal variations in these factors can affect mosquito populations and biting habits, which, in turn, may influence the spread of malaria 316 317 318 319 320 321 322 323 324

In conclusion, factors like self-administered treatment, the availability of healthcare facilities, the healthcare workforce, the presence of asymptomatic malaria cases, the capacity of medical laboratories, and the quality of sanitation infrastructure all have the potential to influence people's inclination to reveal their malaria status or to seek medical help. Recognizing these variations is essential for designing customized and comprehensive strategies that can efficiently tackle underestimation, ultimately resulting in more significant and enduring improvements in clinical public health. The notably low reported rates in Cameroon indicate a distinct necessity for the enhancement of reporting and surveillance systems. 325 326 327 328 329 330 331 332 333

## Supporting information 334

**S1 File. Effective reproduction numbers.** 335

**S2 File. Supplementary tables.** 336

## Data availability 337

The monthly reported cases of malaria in Cameroon were obtained from DHIS2 [44], and are available within the article and its Supplementary material. DHIS2 is an open source, web-based software platform for data collection, management, and analysis in all government hospitals in the countries where aggregation of health data from both government and non-government hospitals is collected. The Bayesian inference code used for our model calibration is available on GitHub at [https://github.com/iyaniwura/Malaria\\_Model\\_Fit\\_Code](https://github.com/iyaniwura/Malaria_Model_Fit_Code). 338 339 340 341 342 343 344

## Funding 345

This research is funded by Canada's International Development Research Centre (IDRC) (Grant No. 109981). 346 347

## Acknowledgments 348

JDK acknowledges support from New Frontier in Research Fund-Exploratory (Grant No. NFRFE-2021-00879) and NSERC Discovery Grant (Grant No. RGPIN-2022-04559). Portions of this work were performed at the Los Alamos National Laboratory under the auspices of the US Department of Energy contract 89233218CNA000001 and supported by NIH grant R01-OD011095. 349 350 351 352 353

## References

1. Robert V, Macintyre K, Keating J, Trape JF, Duchemin JB, Warren M, et al. Malaria transmission in urban sub-Saharan Africa. *The American journal of tropical medicine and hygiene*. 2003;68(2):169–176.
2. De Silva PM, Marshall JM, et al. Factors contributing to urban malaria transmission in sub-Saharan Africa: a systematic review. *Journal of tropical medicine*. 2012;2012.
3. Lindblade KA, Steinhart L, Samuels A, Kachur SP, Slutsker L. The silent threat: asymptomatic parasitemia and malaria transmission. *Expert review of anti-infective therapy*. 2013;11(6):623–639.
4. Colón-González FJ, Sewe MO, Tompkins AM, Sjödin H, Casallas A, Rocklöv J, et al. Projecting the risk of mosquito-borne diseases in a warmer and more populated world: a multi-model, multi-scenario intercomparison modelling study. *The Lancet Planetary Health*. 2021;5(7):e404–e414.
5. Gainor EM, Harris E, LaBeaud AD. Uncovering the burden of dengue in Africa: considerations on magnitude, misdiagnosis, and ancestry. *Viruses*. 2022;14(2):233.
6. Gallup JL, Sachs JD. The economic burden of malaria. *CID Working Paper Series*. 2000;.
7. Ndamuzi E, Gahungu P. Mathematical modeling of Malaria transmission dynamics: Case of burundi. *Journal of Applied Mathematics and Physics*. 2021;9(10):2447–2460.
8. Beretta E, Capasso V, Garao DG. A mathematical model for malaria transmission with asymptomatic carriers and two age groups in the human population. *Mathematical biosciences*. 2018;300:87–101.
9. Faniran T, Falade A, Ogunsanwo J. Sensitivity Analysis of an Untreated Liver-Stage Malaria. *Journal of Computer Science & Computational Mathematics*. 2020;10:61–68.
10. Bakary T, Boureima S, Sado T. A mathematical model of malaria transmission in a periodic environment. *Journal of biological dynamics*. 2018;12(1):400–432.
11. Tabue RN, Njeambosay BA, Zeukeng F, Esemu LF, Fodjo BA, Nyonglema P, et al. Case definitions of clinical malaria in children from three health districts in the north region of Cameroon. *BioMed research international*. 2019;2019.
12. World Health Organization. Malaria (accessed October 16, 2023). Available from: <https://www.afro.who.int/health-topics/malaria>
13. World Health Organization. Global technical strategy for malaria 2016-2030. 2015.
14. Sinan M, Ahmad H, Ahmad Z, Baili J, Murtaza S, Aiyashi M, et al. Fractional mathematical modeling of malaria disease with treatment & insecticides. *Results in Physics*. 2022;34:105220.
15. Grobusch M, Kremsner P. Uncomplicated malaria. *Malaria: Drugs, disease and post-genomic biology*. 2005; p. 81–104.

16. Bartoloni A, Zammarchi L. Clinical aspects of uncomplicated and severe malaria. *Mediterranean journal of hematology and infectious diseases*. 2012;4(1).
17. Garnham PCC, Bray R, Bruce-Chwatt L, Draper C, Killick-Kendrick R, Sergiev P, et al. A strain of *Plasmodium vivax* characterized by prolonged incubation: morphological and biological characteristics. *Bulletin of the World Health Organization*. 1975;52(1):21.
18. Antonio-Nkondjio C, Ndo C, Njiokou F, Bigoga JD, Awono-Ambene P, Etang J, et al. Review of malaria situation in Cameroon: technical viewpoint on challenges and prospects for disease elimination. *Parasites & vectors*. 2019;12:1–23.
19. Massoda Tonye SG, Kouambeng C, Wounang R, Vounatsou P. Challenges of DHS and MIS to capture the entire pattern of malaria parasite risk and intervention effects in countries with different ecological zones: the case of Cameroon. *Malaria journal*. 2018;17(1):1–14.
20. Severe Malaria Observatory. Malaria facts (accessed October 1, 2023). Available from: <https://www.severemalaria.org/countries/cameroon>
21. Antonio-Nkondjio C, Sonhafouo-Chiana N, Ngadjeu C, Doumbe-Belisse P, Talipouo A, Djamouko-Djonkam L, et al. Review of the evolution of insecticide resistance in main malaria vectors in Cameroon from 1990 to 2017. *Parasites & vectors*. 2017;10(1):1–14.
22. Feufack-Donfack LB, Sarah-Matio EM, Abate LM, Bouopda Tuedom AG, Ngano Bayibéki A, Maffo Ngou C, et al. Epidemiological and entomological studies of malaria transmission in Tibati, Adamawa region of Cameroon 6 years following the introduction of long-lasting insecticide nets. *Parasites & vectors*. 2021;14(1):1–14.
23. MINSANTE. Plan stratégique National de lutte contre le paludisme au Cameroun 2014–2018; 2016.
24. Snow RW, Marsh K. The consequences of reducing transmission of *Plasmodium falciparum* in Africa. 2002;.
25. Gibbons CL, Mangen MJJ, Plass D, Havelaar AH, Brooke RJ, Kramarz P, et al. Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. *BMC public health*. 2014;14(1):1–17.
26. Afriyie SO, Addison TK, Gebre Y, Mutala AH, Antwi KB, Abbas DA, et al. Accuracy of diagnosis among clinical malaria patients: comparing microscopy, RDT and a highly sensitive quantitative PCR looking at the implications for submicroscopic infections. *Malaria Journal*. 2023;22.
27. Ross R. The prevention of malaria. 1911. London: Murray. 1911;2.
28. Macdonald G. Epidemiological basis of malaria control. *Bull World Health Organization*. 1956;15:613–626.
29. Anderson RM, May RM. *Infectious diseases of humans: dynamics and control*. Oxford university press; 1992.
30. Yang HM, Ferreira MU. Assessing the effects of global warming and local social and economic conditions on the malaria transmission. *Revista de saude publica*. 2000;34(3):214–222.

31. Pampana E. A Textbook of Malaria Eradication. Oxford University Press, London. 1969;.
32. Collins O, Duffy K. A mathematical model for the dynamics and control of malaria in Nigeria. *Infectious Disease Modelling*. 2022;7(4):728–741.
33. Mandal S, Sarkar RR, Sinha S. Mathematical models of malaria-a review. *Malaria journal*. 2011;10(1):1–19.
34. Ibrahim MM, Kamran MA, Mannan MMN, Kim S, Jung IH. Impact of Awareness to Control Malaria Disease: A Mathematical Modeling Approach. *Complexity*. 2020;2020.
35. Eikenberry SE, Gumel AB. Mathematical modeling of climate change and malaria transmission dynamics: a historical review. *Journal of mathematical biology*. 2018;77:857–933.
36. Olaniyi S, Obabiyi O. Mathematical model for malaria transmission dynamics in human and mosquito populations with nonlinear forces of infection. *International journal of pure and applied Mathematics*. 2013;88(1):125–156.
37. Matteelli A, Castelli F, Caligaris S. Life cycle of malaria parasites. *Handbook of Malaria infection in the tropics 1st edn Bologna: Associazione Italiana” Amici di R Follereau*. 1997; p. 17–23.
38. Tuteja R. Malaria- an overview. *The FEBS journal*. 2007;274(18):4670–4679.
39. Beier JC. Malaria parasite development in mosquitoes. *Annual review of entomology*. 1998;43(1):519–543.
40. National Institute of Statistics. Projections démographiques et estimations des cibles prioritaires des différents programmes et interventions de santé; 2016. Available from: <https://ins-cameroun.cm/statistique/projections-demographiques-et-estimations-des-cibles-prioritaires-des-differents-programmes-et-interventions-de-sante/>
41. Chitnis N, Hyman JM, Cushing JM. Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model. *Bulletin of mathematical biology*. 2008;70(5):1272–1296.
42. The World Bank. Death rate, crude (per 1,000 people) - Cameroon (accessed May 8, 2023). Available from: <https://data.worldbank.org/indicator/SP.DYN.CDRT.IN?locations=CM>
43. Van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*. 2002;180(1-2):29–48.
44. DHIS2. Available from: <https://dhis2.org>
45. Stan Development Team. RStan: the R interface to Stan; 2020. Available from: <http://mc-stan.org/>
46. Chourasia M, Raghavendra K, Bhatt R, Swain D, Valecha N, Kleinschmidt I. Burden of asymptomatic malaria among a tribal population in a forested village of central India: a hidden challenge for malaria control in India. *Public health*. 2017;147:92–97.



47. Yadav R, Bhatt R, Kohli V, Sharma V. The burden of malaria in Ahmedabad city, India: a retrospective analysis of reported cases and deaths. *Annals of Tropical Medicine & Parasitology*. 2003;97(8):793–802.
48. Hyde E, Bonds MH, Ihantamalala FA, Miller AC, Cordier LF, Razafinjato B, et al. Estimating the local spatio-temporal distribution of malaria from routine health information systems in areas of low health care access and reporting. *International Journal of Health Geographics*. 2021;20(1):1–17.
49. Nankabirwa J, Zurovac D, Njogu JN, Rwakimari JB, Counihan H, Snow RW, et al. Malaria misdiagnosis in Uganda—implications for policy change. *Malaria journal*. 2009;8:1–9.
50. Breman JG. Ears of the hippopotamus: manifestations, determinants, and estimates of the malaria burden. *The Intolerable Burden of Malaria: A New Look at the Numbers: Supplement to Volume 64 (1) of the American Journal of Tropical Medicine and Hygiene*. 2001;.
51. Legros F, Danis M, Board EE, et al. Surveillance of malaria in European Union countries. *Eurosurveillance*. 1998;3(5):45–7.
52. Van Hest N, Smit F, Verhave J. Underreporting of malaria incidence in The Netherlands: results from a capture–recapture study. *Epidemiology & Infection*. 2002;129(2):371–377.