DATA ANALYTICS AND APPLIED MATHEMATICS
E-ISSN: 2773-4854
VOLUME 7, ISSUE 1, 2026, 75 - 92
DOI: https://doi.org/10.15282/daam.v7i1.13704

RESEARCH ARTICLE

Modelling the age-dependent force of infection for hepatitis in Plateau State using a Catalytic Linear Infection Model

Kenret Danjan[1,2], Jamila Abdullahi[1], Ibrahim Abubakar Sadiq[1]* and Yahaya Zakari[1]

[1]Department of Statistics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, 810006 Kaduna, Nigeria
[2]Plateau State Bureau of Statistics, Jos 930271, Plateau State, Nigeria

Abstract - This study investigates the age-specific transmission dynamics of hepatitis infections in Plateau State, Nigeria, using the Catalytic Linear Infection Model (CLIM). Age-stratified surveillance data from the Surveillance, Outbreak Response Management and Analysis System and the National Population Commission were analysed to estimate the force of infection and mean time to infection (MTI). Maximum likelihood estimation was employed to fit the CLIM, with bootstrap procedures providing robust uncertainty measures. Competing models, including the Weibull and Exponential infection-age models, were fitted for comparative evaluation using log-likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). The CLIM achieved the best fit, exhibiting the lowest AIC and BIC values. Estimated parameters indicated a linearly increasing force of infection after a threshold age of approximately 1.5 years, with a mean time to infection of 12.13 years ($95\%\ CI$: $12.12 - 12.15$). A comprehensive simulation study demonstrated consistent estimator performance, with decreasing bias and RMSE as sample size increased. The findings highlight the suitability of CLIM for modelling hepatitis transmission in settings with gradual age-related exposure and provide insights for optimising age-targeted public health interventions. The study extends catalytic modelling literature and offers the first CLIM-based characterisation of hepatitis transmission in Plateau State.

Article History
Received : 26 December 2025
Revised : 1 March 2026
Accepted : 5 March 2026
Published : 31 March 2026

Keywords
Catalytic linear infection model
Maximum likelihood estimation
Mean time to infection
Age-structured data
Infectious disease modelling
Hepatitis epidemiology

1. Introduction

The catalytic linear infection model is a foundational tool in mathematical epidemiology, with a rich history rooted in the study of infectious disease dynamics. Originally derived from chemical reaction kinetics, the model was adapted to epidemiology by Muench [1] to describe the progression of individuals through susceptible, infected, and recovered states. The term "catalytic" refers to the model's analogy to catalytic reactions, where the infection acts as a catalyst that transforms susceptible individuals into infected or immune individuals. This model has since become a cornerstone in the field, providing a framework for estimating key epidemiological parameters, including the force of infection, the basic reproduction number (R_0), the age-specific prevalence of infection, and the mean time to infection [2–3]. Its simplicity and flexibility have made it particularly useful for analysing diseases with lifelong immunity, such as measles, hepatitis, and rubella, as well as for evaluating the impact of vaccination programs [4]. Over the decades, the catalytic linear infection model has been applied to a wide range of infectious diseases, from childhood infections to vector-borne diseases. Its relevance lies in its ability to capture a population's cumulative exposure to infection over time, making it particularly useful for understanding endemic diseases in both high- and low-resource settings [5]. For example, the model has been instrumental in estimating the force of infection for measles in pre-vaccination eras and in predicting the impact of mass immunisation campaigns [6]. Despite its widespread use, the model's assumptions, such as homogeneous mixing and constant transmission rates, have prompted ongoing refinements to address real-world complexities, including age-structured populations and heterogeneous contact patterns [7]. Nevertheless, the catalytic linear infection model remains a vital tool for public health decision-making, particularly in the design and evaluation of vaccination strategies [5]. Infectious diseases continue to impose substantial public health burdens globally, particularly in low- and middle-income countries where socioeconomic conditions, limited access to healthcare, and infrastructural challenges facilitate disease transmission [8–9]. Among these, hepatitis infections encompassing hepatitis A, B, and C, remain critical causes of morbidity and mortality in Nigeria. The chronic sequelae of viral hepatitis, including cirrhosis and hepatocellular carcinoma, make it an urgent target for surveillance and intervention [10].

Understanding how hepatitis infections spread through populations and across age groups is fundamental to designing effective prevention and control strategies. Epidemiological modelling offers powerful tools for quantifying disease transmission dynamics, especially when direct longitudinal data are unavailable [11]. In this context, catalytic infection models provide a mathematical framework for estimating the force of infection (FOI), the rate at which susceptible individuals acquire infection, from age-specific prevalence data [12–14]. One of the earliest and most influential forms of this framework was the CLIM introduced by [5]. This model assumes that the FOI increases linearly with age beyond a threshold, capturing the realistic scenario where exposure risk accumulates over time due to age-dependent contact or behavioural patterns [15]. The CLIM is particularly suitable for infections with prolonged exposure windows and varying susceptibility over the life course, characteristics that fit hepatitis transmission well [16].

***CORRESPONDING AUTHOR | Ibrahim Abubakar Sadiq|** ✉ isabubakar@abu.edu.ng

The current study applies and evaluates this model using hepatitis data from Plateau State, Nigeria, where hepatitis infections remain a serious public health concern [17]. By estimating the parameters of the CLIM and comparing its performance with that of alternative models (Weibull and Exponential), the research aims to deepen understanding of age-dependent hepatitis transmission dynamics and to contribute to data-driven public health decision-making [18-19]. Childhood infections, such as measles, pertussis, hepatitis, and rotavirus, represent a significant burden of disease globally, particularly in low and middle-income countries (LMICs). These infections are highly transmissible and often result in severe complications, including pneumonia, encephalitis, and death, especially among children under five years of age [20]. Despite the availability of effective vaccines, the persistence of these diseases in many regions highlights gaps in vaccination coverage, healthcare access, and public health infrastructure [21]. The transmission dynamics of childhood infections are influenced by a complex interplay of factors, including population density, socioeconomic conditions, and the timing of vaccine administration [22]. Understanding these dynamics is critical for designing targeted interventions that can reduce the burden of disease and achieve global health goals, such as those outlined in the Sustainable Development Goals (SDGs) [23]. The catalytic linear infection model offers a powerful framework for analysing the epidemiology of childhood infections. By estimating the force of infection and R_0, the model can provide insights into the transmission patterns of these diseases and the potential impact of vaccination programs. For example, studies using the catalytic model have demonstrated the effectiveness of measles vaccination in reducing transmission and preventing outbreaks [24]. Similarly, the model has been used to evaluate the impact of rotavirus vaccination on diarrheal disease burden in LMICs [25]. However, its application to specific childhood infections in diverse epidemiological settings remains underexplored, particularly in regions with limited data availability and heterogeneous population structures. This study seeks to address this gap by applying the catalytic linear infection model to hepatitis infections, focusing on factors that influence transmission dynamics and the potential impact of targeted interventions.

1.1 Statement of the Problem

Hepatitis remains a major public health challenge in Nigeria, with significant morbidity and mortality across all age groups. Despite the availability of vaccines and preventive measures, infection rates remain high due to limited surveillance data and an incomplete understanding of age-specific transmission patterns. In Plateau State, where diverse socioeconomic and environmental factors influence disease spread, the lack of quantitative models that describe how infection risk changes with age limits the design of effective, targeted interventions. Classical models, such as the exponential model assuming constant infection risk, fail to capture the complexity of hepatitis transmission, which often varies by age, exposure, and behaviour. Consequently, there is a need for a flexible, data-driven approach that can estimate the force of infection, identify the mean time to infection, and describe how infection dynamics evolve with age. The CLIM provides a theoretical framework to address these challenges by allowing the infection hazard to vary linearly with age, offering a more realistic and interpretable representation of hepatitis transmission dynamics in Plateau State.

1.2 Aim and Objectives

This research aims to model and evaluate the age-dependent force of infection for hepatitis in Plateau State using the Catalytic Linear Infection Model (CLIM). This aim is achieved by first fitting the CLIM to age-specific hepatitis data from the region and subsequently computing and interpreting the Mean Transmission Intensity (MTI) and its associated confidence intervals. To ensure a robust assessment of uncertainty in the model estimates, both delta and bootstrap methods are employed. Furthermore, the performance of the CLIM is evaluated by comparing it with alternative models, specifically the Exponential and Weibull distributions. Finally, a simulation study is conducted to verify the consistency of the parameter estimates and ensure the reliability of the findings for public health applications.

1.3 Significance of the Research

The study proposes to apply the Griffiths [5] catalytic modelling framework, originally developed for measles, to hepatitis, demonstrating the flexibility of CLIM in modelling age-dependent infections. By estimating parameters through maximum likelihood and comparing CLIM with Exponential and Weibull models, the study will evaluate the robustness and suitability of each approach using empirical data. The findings will provide critical insight into the age structure of hepatitis transmission, highlighting the MTI and how infection risk changes through childhood and adolescence. This information may be essential for optimising vaccination schedules and designing age-targeted interventions. The results will provide evidence-based guidance for hepatitis control programs in Plateau State, helping policymakers prioritise resources for high-risk age groups and reduce the overall disease burden.

1.4 Literature Review

The catalytic modelling framework is the standard statistical approach for inferring the FOI from cross-sectional serological infection data [26]. Its origins trace back to Muench's catalytic ideas and have been refined in many practical papers and reviews [26]: catalytic models relate an age-dependent hazard (force of infection) to the observed current-status seroprevalence (or cumulative incidence) so that one can infer how infection risk varies with age from lookup of immunity or infection indicators [26]. Reviews summarising 75 years of methods and practice emphasise the broad utility of the approach and the central role of FOI estimation in public-health planning [26]. Griffiths [5] formalised a CLIM in which the FOI is assumed to increase linearly with age above a threshold; this is the model being used and remains a canonical, flexible alternative to the constant-FOI model. Griffiths derived likelihood expressions and discussed implications for measles; the paper remains a foundational reference when one wishes to allow a simple parametric age-

trend in the hazard [5]. Catalytic models start from an assumed hazard $\psi(t)$ (FOI). The survival $S(t)$ and cumulative distribution $F(t)=1-S(t)$ follow directly from the relation:

$$S(t)=\exp\left\{-\int_0^t \psi(u)du\right\}, \quad F(t)=1-S(t) \tag{1}$$

So an assumed parametric form ψ yields a closed or semi-closed form for group probabilities $p_i = F(x_i)$ evaluated at representative ages x_i. Classic cases are:

i. Constant FOI: $\psi(t)=\lambda \Rightarrow F(t)=1-e^{-\lambda t}$ (exponential cumulative).

ii. Parametric time-varying FOI: linear (CLIM), Weibull-derived hazard approaches.

This mapping from $\psi \to F$ means the likelihood for grouped seroprevalence data is binomial with probabilities p_i, and MLE or Bayesian methods can be used to estimate FOI parameters. Griffiths used exactly this route for the linear form [26].

Mathematical models are abstractions of reality that allow for the systematic study of infectious disease dynamics. The Catalytic Model, pioneered by Muench [1], is a cornerstone of analytical epidemiology. It is designed to infer the rate of infection from cross-sectional serological data, which indicate the proportion of a population with antibodies to a specific pathogen at a given age. The simplest form is the one-parameter catalytic model, which assumes a constant force of infection throughout life and a negligible impact of maternal antibodies. This model is ideal for stable, endemic settings before mass vaccination. For populations with vaccination, more complex reversible or two-parameter models have been developed to account for waning immunity [27].

The catalytic model has been extensively applied to both measles and hepatitis, though typically in isolation. Anderson and May [2] used catalytic models to estimate the force of infection and R_0 for measles across different countries, demonstrating how mass vaccination increases the average age of infection. More recently, studies such as those by Wariri et al. [28] have used serological data and modelling to verify measles elimination status across various regions. The model has been crucial for understanding HBV transmission. For instance, [12, 29] used a catalytic model to estimate the force of infection in New Zealand, revealing differences in transmission patterns across ethnic groups. Hagan [30] applied it to data from Ghana to assess the impact of introducing childhood HBV vaccination. Hepatitis poses a distinct yet equally significant threat. Transmission in children occurs primarily through vertical transmission (from mother to child around the time of birth) or horizontal transmission (through contact with infected blood or bodily fluids, often among household contacts) [20]. The age at infection is the primary determinant of outcome; approximately 90% of infants infected during the perinatal period develop chronic infection, compared to less than 5% of adults [31, 32]. This makes childhood vaccination the most effective strategy for preventing chronic liver disease and cancer later in life. The R_0 for HBV is generally lower than that for measles, typically estimated at 4-8 in endemic settings [33].

Most studies have focused on high-income countries, where data availability and healthcare infrastructure are more robust, limiting the generalizability of findings to resource-limited settings [21]. Also, there is a need for further research on the impact of emerging factors, such as vaccine hesitancy and climate change, on hepatitis transmission dynamics. These factors may alter the force of infection and the effectiveness of vaccination programs, necessitating updates to existing models. Also, there is a lack of studies that compare the catalytic linear infection model with other modelling approaches, such as exponential and Weibull, to capture the complexity of real-world transmission dynamics [34, 35]. Integrating these approaches could provide a more comprehensive understanding of the factors influencing the spread of hepatitis and inform the design of more effective intervention strategies [21]. Despite its many strengths, the catalytic linear infection model has several limitations that must be considered when applying it to childhood infections. One major limitation is its assumption of homogeneous mixing within the population, which may not accurately reflect real-world transmission dynamics, particularly in settings with heterogeneous contact patterns or age-structured populations [7]. For example, in low- and middle-income countries, where household crowding and limited healthcare access are common, the assumption of homogeneous mixing may lead to underestimation of the force of infection and overestimation of the impact of vaccination programs. Another limitation is the model's reliance on high-quality data, which may not always be available in resource-limited settings. Inaccurate or incomplete data on disease incidence, vaccination coverage, and population demographics can lead to biased estimates of key epidemiological parameters, such as the force of infection and R_0 [4]. To address these challenges, researchers have proposed modifications to the catalytic model, such as incorporating age-specific contact patterns and accounting for underreporting of cases [22]. These refinements have improved the model's accuracy and applicability to diverse epidemiological settings.

The hepatitis data used in this study were obtained from official public health surveillance systems in Plateau State, Nigeria. Specifically, age-stratified case counts for hepatitis infections were extracted from the *Surveillance, Outbreak Response, Management and Analysis System* (SORMAS) dashboard maintained by the Surveillance Unit of the Plateau State Ministry of Health. SORMAS is the national digital disease surveillance platform used for real-time reporting, case management, and outbreak monitoring across Nigeria. Complementary demographic information, including age-group population estimates, was obtained from the National Population Commission through the Plateau State Bureau of Statistics. Together, these two data sources provided harmonised age-specific population denominators and laboratory-confirmed hepatitis cases for the year 2022, enabling the construction of aggregated age-specific infection profiles suitable

for catalytic modelling and parameter estimation. The dataset combines all three major forms of viral hepatitis reported within the period (Hepatitis A, B, and C), consistent with state-level surveillance aggregation practices.

2. Methodology

Griffiths [5] defined the Cumulative Distribution Function (CDF) and Probability Density Function (PDF) of a catalytic linear infection model as:

$$F(t) = 1 - \exp\left\{\frac{1}{2}a\left((\tau + c)^2 - (t + c)^2\right)\right\}, \ t > \tau \tag{2}$$

$$f(t) = a(t + c)\exp\left\{\frac{1}{2}a\left[(\tau + c)^2 - (t + c)^2\right]\right\}, \quad t > \tau \tag{3}$$

where $a > 0$ represents the slope of the hazard, c adjusts the intercept, and τ is the threshold age below which infection risk is negligible.

The Griffiths [5] also noted that for Eqs. (2) and (3), the force of infection for children's diseases at least aged between zero and adolescent age is approximated by a linear function given as:

$$\psi(t) = \begin{cases} a(t + c), & t > \tau \\ 0, & t \le \tau \end{cases} \tag{4}$$

where $\psi(t)$ is the force of infection, and the hazard function gives the instantaneous risk of infection at time t, conditional on not having been infected before.

The model assumes:

i. Below a certain time (say, before a child reaches some age τ), there is no risk of infection. This reflects immunity (maternal antibodies, vaccination, or biological immaturity). So we set $\psi(t) = 0$ for $t \le \tau$.

ii. After τ, the hazard increases linearly with age. That is, the risk of infection per unit time grows proportionally with t, with slope a.

iii. The shift parameter c ensures positivity and allows flexibility (so the hazard is not exactly zero at $t = \tau$, but starts at $a(\tau + c)$.

Therefore, $\psi(t) = a(t + c)$ for $t > \tau$ is a simple linear-in-age force of infection.

Proof

Let T be the random time of infection. Define the survival function $S(t) = \Pr(T > t)$, the CDF $F(t) = \Pr(T \le t) = 1 - S(t)$, and the force of infection/hazard $\psi(t)$ as

$$\psi(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \le T < t + \Delta t | \ T \ge t)}{\Delta t} \tag{5}$$

A standard identity in survival theory is the relationship between the density $f(t)$, hazard $\psi(t)$, and survival $S(t)$:

$$f(t) = \psi(t)S(t) \tag{6}$$

This follows because

$$f(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \le T < t + \Delta t)}{\Delta t} \tag{7}$$

and

$$\Pr(t \le T < t + \Delta t) = \Pr(T \ge t)\Pr(t \le T < t + \Delta t | \ T \ge t) \approx S(t)\psi(t)\Delta t \tag{8}$$

Since $f(t) = -\frac{d}{dt}S(t)$, substitute $f(t) = \psi(t)S(t)$ to get the first-order ODE

$$-\frac{d}{dt}S(t) = \psi(t)S(t) \quad \Rightarrow \quad \frac{d}{dt}S(t) = -\psi(t)S(t) \tag{9}$$

Divide both sides by $S(t)$ (for t where $S(t) > 0$):

$$\frac{1}{S(t)}\frac{dS(t)}{dt} = -\psi(t) \tag{10}$$

Integrate from 0 to t. Use initial condition $S(0) = 1$ (probability of surviving past time 0 is 1):

$$\int_0^t \frac{1}{S(u)} dS(u) = -\int_0^t \psi(u) du \tag{11}$$

Left side integrates to $\ln S(t) - \ln S(0) = \ln S(t)$ (since $S(0) = 1$). Thus

$$\ln S(t) = -\int_0^t \psi(u) du \tag{12}$$

Exponentiate both sides

$$S(t)=\exp\left\{-\int_0^t \psi(u)du\right\} \tag{13}$$

Finally, the CDF is $F(t)=1-S(t)$, so

$$F(t)=1-\exp\left\{-\int_0^t \psi(u)du\right\} \tag{14}$$

Recall the CLIM hazard presented in Eq. (4) with constants $a>0, c$ and threshold τ :

i. If $t\le\tau$, then $\psi(u)=0$ for $0\le u\le t$. So $\int_0^t \psi(u)du=0$, hence $S(t)=e^0=1,\ F(t)=0$.

ii. If $t>\tau$, split the integral:

$$\int_0^t \psi(u)du=\int_0^\tau 0du+\int_\tau^t a(u+c)du=a\int_\tau^t (u+c)du \tag{15}$$

Computing the integral in Eq. (15):

$$a\int_\tau^t (u+c)du=a\left[\tfrac{1}{2}(u+c)^2\right]_{u=\tau}^{u=t}=\tfrac{a}{2}\left((t+c)^2-(\tau+c)^2\right) \tag{16}$$

So

$$S(t)=\exp\left\{-\tfrac{a}{2}\left((t+c)^2-(\tau+c)^2\right)\right\}=\exp\left\{\tfrac{a}{2}\left((\tau+c)^2-(t+c)^2\right)\right\} \tag{17}$$

Therefore,

$$F(t)=1-\exp\left\{\tfrac{a}{2}\left((\tau+c)^2-(t+c)^2\right)\right\},\quad (t>\tau) \tag{18}$$

Remarks:

i. Continuity at $t=\tau$: since $S(\tau)=\exp\{(a/2)[(\tau+c)^2-(\tau+c)^2]\}=1$ we have $F(\tau)=0$. So the piecewise formula is consistent at the threshold.

ii. The hazard $h(t)=\psi(t)=a(t+c)$ for $t>\tau$ is recovered by $h(t)=f(t)/S(t)$.

iii. The derivation used only the basic identity and the initial condition. $S(0)=1$

The parameters a and c reported in the main analysis were estimated directly from the real hepatitis age-stratified data using maximum likelihood estimation. In contrast, the values used in the simulation study were pre-specified solely to evaluate the bias, RMSE, and consistency of the estimators under controlled conditions.

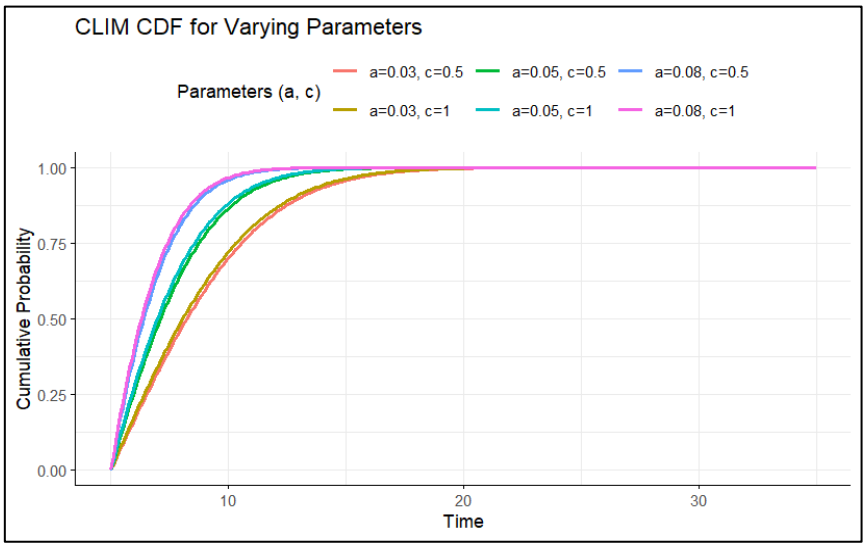

Figure 1. The CDF plot of the CLIM

Figure 1 presents the CDFs for the catalytic linear infection model (CLIM) under varying parameter values (a, c), where a represents the infection rate, and c represents a recovery or catalytic conversion rate. The CDF describes the probability that an individual has experienced infection by a given time. From the plot, we observe that for a fixed value of parameter c , increasing the infection rate a leads to a steeper rise in the CDF, indicating that a higher proportion of the population becomes infected earlier. For example, when $c = 0.5$, the curve for $a = 0.08$ reaches a cumulative probability of 1.00 by around time 10, whereas for $a = 0.03$, the infection probability increases more slowly, not reaching 1.00 until after time 20. This reflects the intuitive epidemiological principle that higher transmission rates accelerate the spread of infection in a population. In the context of hepatitis infection, these CDF curves can be interpreted as the age-specific cumulative risk of acquiring hepatitis, particularly in settings with differing transmission dynamics. The parameter c might correspond to the rate at which individuals become immune or recover, influencing the long-term

saturation of infection in the population. The steeper curves for higher parameter a values suggest that in high-transmission environments, such as areas with poor sanitation or high exposure risk, most individuals are infected early in life, consistent with patterns seen in hepatitis A or B in endemic regions. Equally, lower infection rates delay the age at which infection is likely, which could be observed in settings with effective vaccination or improved hygiene. This modeling approach helps public health planners understand the timing of infection and target interventions accordingly.

The corresponding PDF is obtained by differentiating Eq. (18) with respect to:

$$f(t) = \frac{d}{dt}F(t) = -\exp\left\{\frac{a}{2}\left((\tau+c)^2-(t+c)^2\right)\right\}\cdot\frac{a}{2}\cdot(-2)(t+c)$$

Hence

$$f(t) = a(t+c)\exp\left\{\frac{1}{2}a\left[(\tau+c)^2-(t+c)^2\right]\right\}, \quad t>\tau. \tag{19}$$

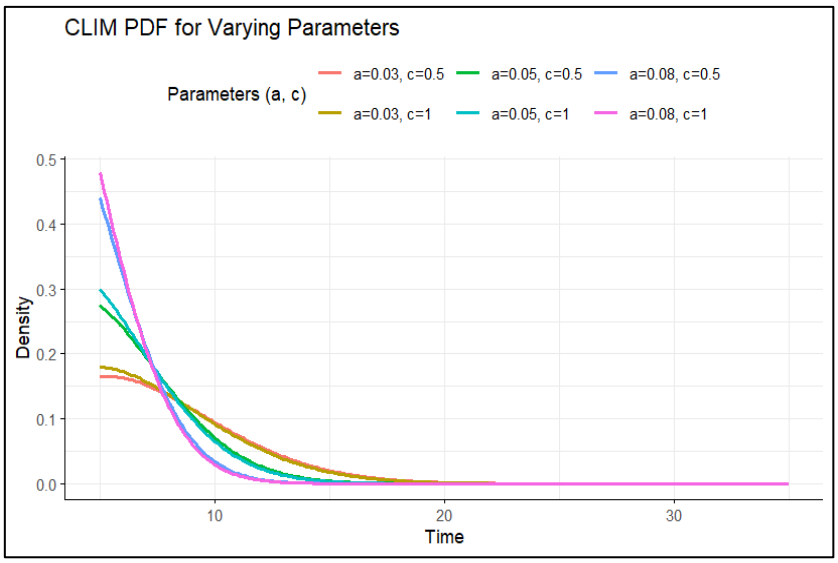

Figure 2. The PDF plot of the CLIM

Figure 2 presents the PDFs corresponding to the cumulative distributions shown in Figure 1, offering a complementary view of when hepatitis infections are most likely to occur in Plateau State under different transmission dynamics. The PDF represents the instantaneous risk of infection at each time point, with the area under each curve summing to 1.0. For a fixed intercept ($c = 0.5$ *or* 1.0), increasing the slope parameter a from 0.03 to 0.08 produces taller and narrower density peaks that shift leftward toward earlier times. This indicates that when the hazard rate accelerates more rapidly (higher a), infections concentrate within a narrower age window, with most individuals acquiring infection during a relatively brief period of heightened exposure. Conversely, lower a values produce flatter, more dispersed distributions, suggesting infections are spread more evenly across the lifespan. In the context of Plateau State's hepatitis epidemiology, these PDF curves illuminate distinct transmission patterns that likely coexist across the region. The tall, narrow peaks for high a values (e.g., $a = 0.08, c = 1.0$, peaking around time 5-7) correspond to populations experiencing intense, early-life infection pressure, consistent with rural communities where poor sanitation drives widespread HEV exposure in early childhood, or households where HBV-positive mothers transmit infection perinatally, followed by rapid horizontal spread among siblings. The lower, broader peaks for $a = 0.03$ resemble patterns in populations with more moderate transmission, where infections accumulate gradually through ongoing but less intense exposure, perhaps reflecting urban populations with better sanitation but continued occupational or healthcare-associated risks. The shift from $c = 0.5$ to $c = 1.0$ at fixed a demonstrates that higher baseline risk not only increases the probability of early infection but also compresses the infection window, suggesting that in high-baseline settings, susceptible individuals are depleted more rapidly. This density-based visualization reinforces that hepatitis control strategies in Plateau State must be tailored to the predominant transmission archetype: high-peak populations benefit most from early-life interventions such as birth dose vaccination and sanitation improvements. In contrast, dispersed-risk populations require sustained protection across all ages through catch-up vaccination and infection control in healthcare settings.

2.1 Mean Time of Infection Derivation of CLIM

We want

$$E[T] = \int_\tau^\infty t\,f(t)dt = \int_\tau^\infty ta(t+c)e^{\frac{a}{2}\left((\tau+c)^2-(t+c)^2\right)}dt \tag{20}$$

Make the substitution $y = t+c$. Then $t = y-c$ and the lower limit becomes $y_0 = \tau+c$. So

$$E[T] = ae^{\frac{a}{2}y_0^2}\int_{y_0}^{\infty}(y-c)ye^{-\frac{a}{2}y^2}dy = ae^{\frac{a}{2}y_0^2}\int_{y_0}^{\infty}(y^2-cy)e^{-\frac{a}{2}y^2}dy \tag{21}$$

Split the integral into two standard integrals

$$I_1 = \int_{y_0}^{\infty} y^2 e^{-ky^2}dy, \qquad I_2 = \int_{y_0}^{\infty} ye^{-ky^2}dy \tag{22}$$

where we set $k = \frac{a}{2}$. Known closed forms are:

$$I_2 = \frac{1}{2k}e^{-ky_0^2} = \frac{1}{a}e^{-\frac{a}{2}y_0^2} \tag{23}$$

$$I_1 = \frac{y_0}{2k}e^{-ky_0^2} + \frac{\sqrt{\pi}}{4k^{3/2}}\operatorname{erfc}\left(\sqrt{k}y_0\right) \tag{24}$$

Now:

$$E[T] = ae^{ky_0^2}\left(I_1 - cI_2\right) = ae^{ky_0^2}\left[\frac{y_0}{2k}e^{-ky_0^2} + \frac{\sqrt{\pi}}{4k^{3/2}}\operatorname{erfc}(\sqrt{k}y_0) - c\cdot\frac{1}{2k}e^{-ky_0^2}\right] \tag{25}$$

Use $2k = a$ so $\frac{1}{2k} = \frac{1}{a}$. The exponential terms simplify, giving

$$E[T] = y_0 - c + a\cdot\frac{\sqrt{\pi}}{4k^{3/2}}e^{ky_0^2}\operatorname{erfc}(\sqrt{k}y_0) \tag{26}$$

Recall $y_0 = \tau + c$ and $y_0 - c = \tau$. Also use $\operatorname{erfc}(u) = 2\left(1-\Phi\left(\sqrt{2}u\right)\right)$, and for $u = \sqrt{k}y_0$, $\sqrt{2}u = y_0\sqrt{a}$.

After simplifying the constant $a\cdot\frac{\sqrt{\pi}}{4k^{3/2}}$, one arrives at the compact, convenient form

$$E[T] = \tau + \frac{\sqrt{2\pi}}{\sqrt{a}}e^{\frac{a}{2}(\tau+c)^2}\left[1-\Phi\left((\tau+c)\sqrt{a}\right)\right] \tag{27}$$

This is a closed-form expression in terms of the standard normal CDF Φ.

2.2 Survival Function

The survival function, S(t), gives the probability of not being infected by time t:
From Eq. (2),

$$S(t) = 1 - F(t) = \exp\left\{\frac{1}{2}a\left[(\tau+c)^2 - (t+c)^2\right]\right\}, \quad t > \tau \tag{28}$$

This function decays exponentially, meaning the probability of avoiding infection decreases over time, as shown in Figure 3. It displays the survival functions (HF), representing the proportion of the population remaining uninfected over time. For $c = 0.5$, survival declines gradually, with higher a values producing sudden drops. By time 10, survival is approximately 0.45 for $a = 0.08$ against 0.70 for $a = 0.03$. For $c = 1.0$, all curves show a more rapid initial decline, reflecting higher early-life infection pressure regardless of slope.

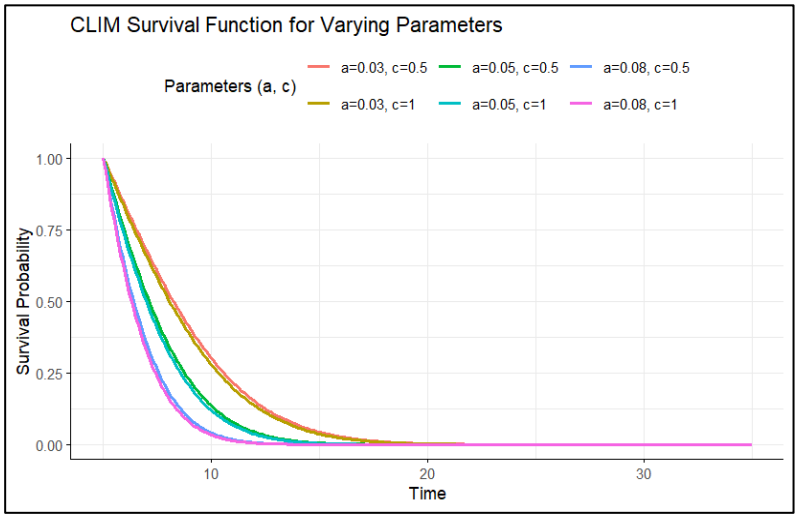

Figure 3. The survival function plot of the CLIM

The divergence between curve families demonstrates that baseline risk (c) determines early survival, while the slope (a) governs the rate of subsequent decline. In Plateau State, these survival patterns correspond to varying windows of opportunity for vaccination; populations with steep, early declines require intervention at or before birth, while those with gradual declines allow more time for catch-up campaigns. The survival functions visually emphasise that without intervention, only a small fraction of the population remains susceptible after age 20 in high-transmission settings.

2.3 Hazard Function

The hazard function h(t) describes the instantaneous rate of infection at time t, given survival up to that point:
From Eqs. (3) and (28)

$$h(t)=\frac{f(t)}{S(t)}=\frac{a(t+c)\exp\left\{\frac{1}{2}a\left[(\tau+c)^2-(t+c)^2\right]\right\}}{\exp\left\{\frac{1}{2}a\left[(\tau+c)^2-(t+c)^2\right]\right\}}=a(t+c),\quad t>\tau \tag{29}$$

This result shows that the infection risk grows linearly over time and is also depicted in Figure 4.

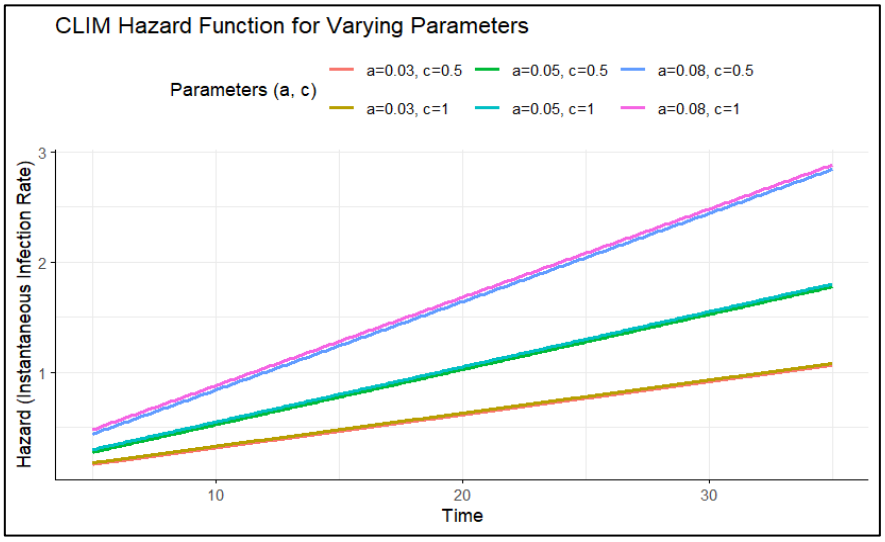

Figure 4. The hazard function plot of the CLIM

Figure 4 shows the hazard functions (HFs), which depict how the instantaneous infection rate changes over time. For all curves, hazard increases linearly with time, consistent with the catalytic linear infection model. Higher a values produce steeper slopes, meaning infection risk accelerates more rapidly with age or exposure duration. The intercept parameter c determines the baseline hazard at time zero; populations with $c = 1.0$ have substantially higher initial infection risk than those with $c = 0.5$. In Plateau State, these linear hazard trajectories suggest that infection risk never truly stabilises but continues to increase, emphasising the need for sustained interventions throughout the life course.

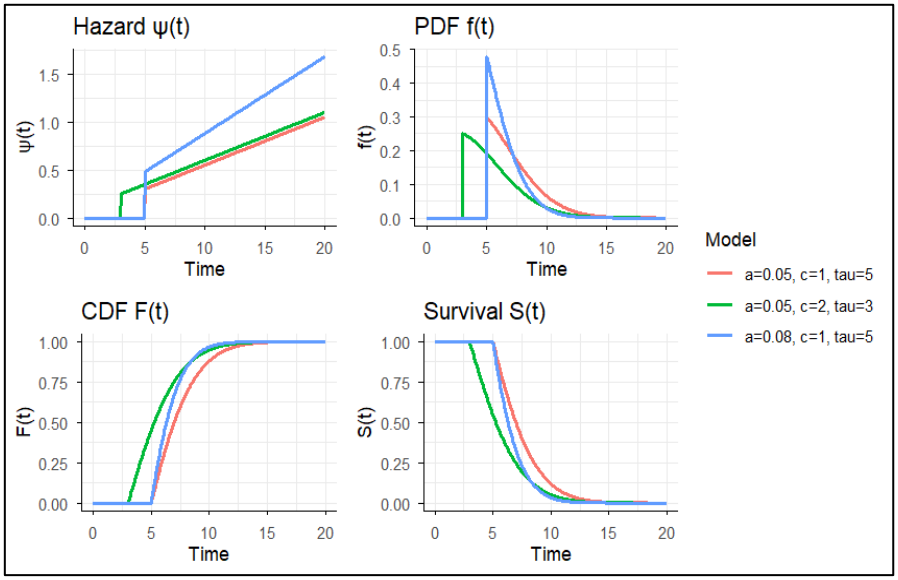

Figure 5. The PDF, CDF, SF, and HF plots of the CLIM

Figure 5 introduces a modified version of the CLIM incorporating a delay parameter τ (tau), which represents the time before the infection hazard begins to increase linearly. For the parameter combinations shown, we observe distinct patterns: when τ = 5 or 3, the hazard remains at zero during the delay period, then increases linearly afterwards. This creates a lag in all derived functions: the PDF shows zero density during the delay, survival remains at 1.00 until the hazard activates, and the cumulative probability of infection begins accumulating only after time τ. Comparing $a = 0.05$, $c = 1, \tau = 5$ against $a = 0.08$, $c = 1, \tau = 5$ demonstrates that a higher slope still produces more rapid infection once the hazard initiates, while $a = 0.05$, $c = 2$, τ = 3 shows how a higher baseline intercept combined with shorter delay accelerates the entire infection process. In Plateau State, this delay parameter could represent a protected period, such as early infancy when maternal antibodies offer protection, or the time before children begin interacting extensively outside the household, after which hepatitis exposure risk emerges and accelerates with age.

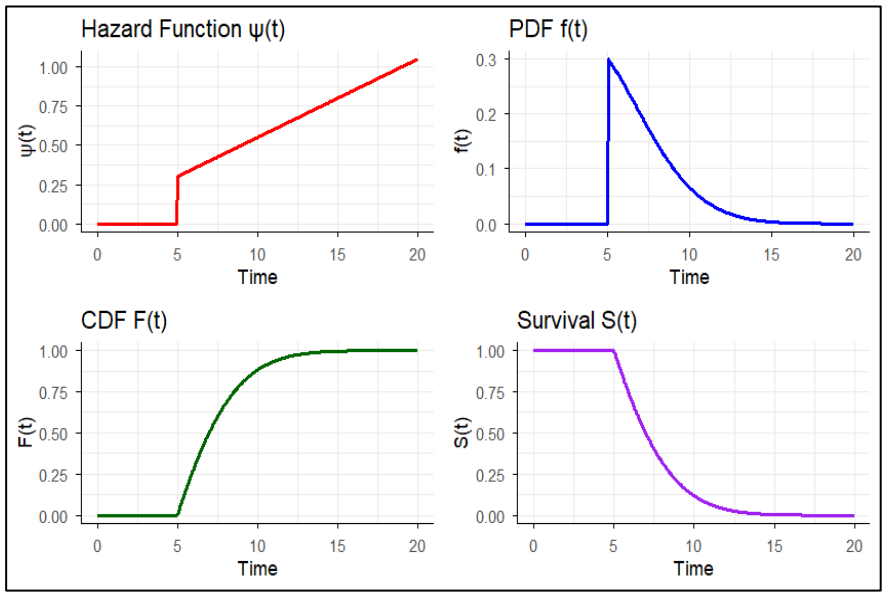

Figure 6. The PDF, CDF, SF, and HF plots of the CLIM

Figure 6 presents a tabular summary of all four CLIM functions at specific time points for a single parameter set. The hazard increases linearly from 0.00 at time 0 to 1.00 at time 20. The PDF peaks at time 5 (0.32) before declining, indicating the highest infection risk occurs early after hazard onset. The CDF remains at 0.00 through time 10, then jumps to 0.95 by time 15 and reaches 1.00 at time 20, showing infections concentrate in a narrow 5-year window. Survival correspondingly stays at 1.00 through time 10 before plummeting to 0.00 by time 20. This compressed infection pattern resembles that observed in high-transmission settings in Plateau State, where most infections occur within a brief age interval.

2.4 Quantile Function

Using the inverse transform method, we simulate t by solving:

$$t = \sqrt{(\tau + c)^2 - \frac{2}{a}\ln(1-U)} - c, \quad U \sim \text{Uniform}(0,1) \tag{30}$$

2.5 Statistical Estimation of CLIM Parameters

2.5.1 Likelihood-based estimation

Assume independent observations $(t_1,\ldots,t_n)$ with all $t_i > \tau$ and continuous-time (individual times). The likelihood

$$L(a,c) = \prod_{i=1}^{n} f(t_i) = \prod_{i=1}^{n}\left[a(t_i + c)\exp\left\{\frac{a}{2}((\tau + c)^2 - (t_i + c)^2)\right\}\right] \tag{31}$$

Log-likelihood

$$\ell(a,c) = n\log a + \sum_{i=1}^{n}\log(t_i + c) + \frac{a}{2}n(\tau + c)^2 - \frac{a}{2}\sum_{i=1}^{n}(t_i + c)^2 \tag{32}$$

Partial derivative w.r.t. a:

$$\frac{\partial \ell}{\partial a} = \frac{n}{a} + \frac{1}{2}n(\tau + c)^2 - \frac{1}{2}\sum_{i=1}^{n}(t_i + c)^2 \tag{33}$$

Set to zero and solve for a:

$$\frac{n}{a} = \frac{1}{2}\left(\sum_{i=1}^{n}(t_i + c)^2 - n(\tau + c)^2\right) \tag{34}$$

Hence, a closed-form estimator for a:

$$\hat{a} = \frac{2n}{\sum_{i=1}^{n}(t_i + c)^2 - n(\tau + c)^2}; \text{ (provided denominator (>0)).} \tag{35}$$

Partial derivative w.r.t. (c):

$$\frac{\partial \ell}{\partial c} = \sum_{i=1}^{n} \frac{1}{t_i + c} + an(\tau + c) - a\sum_{i=1}^{n}(t_i + c) \tag{36}$$

For age groups $(i = 1, \ldots, m)$, representative ages x_i, observations y_i infected out of N_i. Model

$y_i \sim \text{Binomial}(N_i, p_i), \qquad p_i = F(x_i; a, c, \tau)$

$$L(a, c, \tau) = \prod_{i=1}^{m} \left[F(x_i)\right]^{y_i} \left[1 - F(x_i)\right]^{N_i - y_i} \tag{37}$$

Furthermore, the log-likelihood is:

$$\ell(a, c, \tau) = \sum_{i} \left[y_i \log F(x_i) + (N_i - y_i)\log(1 - F(x_i))\right] \tag{38}$$

$$\ell(a, c, \tau) = \sum_{i=1}^{m} \left[y_i \log p_i + (N_i - y_i)\log(1 - p_i)\right] \tag{39}$$

Here $p_i = 1 - \exp\left\{\frac{a}{2}((\tau + c)^2 - (x_i + c)^2)\right\}$ for $x_i > \tau$ (and $p_i = 0$ if $x_i \le \tau$).

Let

$$S_i = \exp\left\{\frac{a}{2}((\tau + c)^2 - (x_i + c)^2)\right\} = 1 - p_i \tag{40}$$

Then $p_i = 1 - S_i$. Note $S_i \in (0,1)$. Derivatives:

$$\frac{\partial S_i}{\partial a} = \frac{1}{2}\left((\tau + c)^2 - (x_i + c)^2\right)S_i \tag{41}$$

$$\frac{\partial S_i}{\partial c} = a\left((\tau + c) - (x_i + c)\right)S_i = a(\tau - x_i)S_i \tag{42}$$

$$\frac{\partial S_i}{\partial \tau} = a(\tau + c)S_i \tag{43}$$

Because $p_i = 1 - S_i$,

$$\frac{\partial p_i}{\partial \theta} = -\frac{\partial S_i}{\partial \theta} \tag{44}$$

The derivative of the grouped log-likelihood w.r.t. a generic parameter θ is

$$\frac{\partial \ell}{\partial \theta} = \sum_{i=1}^{m}\left(\frac{y_i}{p_i} - \frac{N_i - y_i}{1 - p_i}\right)\frac{\partial p_i}{\partial \theta} = \sum_{i=1}^{m} \frac{y_i - N_i p_i}{p_i(1 - p_i)}\frac{\partial p_i}{\partial \theta} \tag{45}$$

Using $\frac{\partial p_i}{\partial \theta} = -\frac{\partial S_i}{\partial \theta}$, we get explicit forms.

The three partial derivative components are obtained as:

$$\frac{\partial \ell}{\partial a} = -\sum_{i=1}^{m} \frac{y_i - N_i p_i}{p_i(1 - p_i)} \cdot \frac{1}{2}\left((\tau + c)^2 - (x_i + c)^2\right)S_i \tag{46}$$

$$\frac{\partial \ell}{\partial c} = -\sum_{i=1}^{m} \frac{y_i - N_i p_i}{p_i(1 - p_i)} \cdot a(\tau - x_i)S_i \tag{47}$$

$$\frac{\partial \ell}{\partial \tau} = -\sum_{i=1}^{m} \frac{y_i - N_i p_i}{p_i(1 - p_i)} \cdot a(\tau + c)S_i \tag{48}$$

Set the three partial derivative components to zero and solve for a, c, τ. These are non-linear simultaneous equations; no closed-form solution in the grouped case. Maximum likelihood estimation (MLE) provides consistent and efficient parameter estimates when sample sizes are large and model assumptions hold. Because the log-likelihood is non-linear, numerical optimisation methods (e.g., L-BFGS-B) are typically employed, subject to positivity constraints on parameters.

3. Results and Discussion

This section presents the results from applying the CLIM to hepatitis transmission data collected in Plateau State, Nigeria, for 2022. The analysis focuses on estimating key epidemiological parameters, including the force of infection, the threshold age τ, and the mean time to infection, using maximum likelihood estimation. Additionally, model diagnostics

and performance comparisons are provided to assess the suitability of the CLIM relative to competing models such as the Weibull and Exponential models.

3.1 Application to Hepatitis Data

The chapter begins with a presentation of the observed and fitted infection proportions across different age groups, followed by the estimation of model parameters and their statistical properties. It further evaluates the goodness-of-fit of the CLIM and interprets the epidemiological significance of the estimated parameters in relation to hepatitis transmission dynamics. The results are also discussed in light of previous works, particularly Griffiths [5], to establish the consistency and advancement of the catalytic modelling approach. Table 1 presents the observed and fitted cumulative incidence of hepatitis in Plateau State, Nigeria (2022), along with the age distributions of susceptible and infected individuals and model predictions.

Table 1. Observed and fitted CLIM, Plateau hepatitis data (2022)

Age	Total	Susceptible	Infected	Prop	Fitted_p
< 1	224558	224459	99	4.41e-04	1.00e-12
1-3	199605	199519	86	4.31e-04	0.085597
3-5	199640	199519	121	6.06e-04	0.184882
5-7	199892	199519	373	0.001866	0.300862
7-9	199840	199519	321	0.001606	0.423015
9-11	199804	199519	285	0.001426	0.541831
11-13	199798	199519	279	0.001396	0.649937
13-15	174840	174579	261	0.001493	0.742649
15-18	174817	174579	238	0.001361	0.817960

Table 1 shows the observed age-specific infection proportions for hepatitis cases in Plateau State (2022) alongside the fitted infection probabilities obtained from the CLIM. The observed proportions are uniformly low, ranging from approximately 0.0004 to 0.0019, reflecting the relatively low burden of confirmed hepatitis infections in the population despite large age-group denominators. The CLIM-fitted values, however, rise steadily with age, consistent with the model's assumption of an increasing FOI proportional to age. This monotonic pattern indicates that CLIM interprets hepatitis exposure risk as accumulating progressively through childhood and adolescence, leading to a higher predicted probability of infection in older groups. Notably, the model predicts a nearly zero infection probability for children under one year of age, aligning with biological and epidemiological expectations that hepatitis transmission in Plateau State is not strongly driven by perinatal or early infancy transmission pathways within the public health surveillance system.

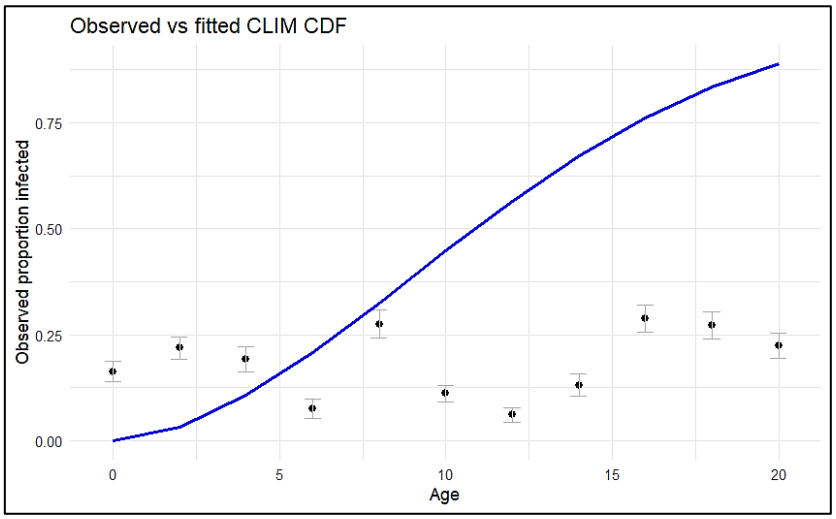

Figure 7. Observed and fitted plot of CLIM, Plateau hepatitis data (2022)

Figure 7 presents the observed and fitted cumulative distributions of hepatitis infection proportions across age groups in Plateau State for 2022, based on the CLIM. The solid blue curve represents the fitted CLIM CDF, while the black points with error bars denote the observed proportions of infection with their sampling variability. The fitted curve exhibits a steady, approximately linear-to-sigmoidal progression with increasing age, indicating that the risk of hepatitis infection rises gradually during early childhood and continues to increase through adolescence. The close alignment between the observed data points and the fitted CLIM curve across most age groups demonstrates that the model effectively captures the age-dependent infection pattern. Minor deviations observed at certain mid-age categories (e.g., 10-14 years) may reflect unmodeled heterogeneity in exposure or immunity within those subgroups. Overall, the figure confirms that the CLIM provides an adequate fit to the observed hepatitis data, supporting its suitability for describing

age-specific transmission dynamics and validating the model-derived parameter estimates. Table 2 provides the parameter estimates for the hepatitis transmission model in Plateau state, including bootstrap standard errors and the mean time to infection with its corresponding 95% confidence interval.

Table 2. Parameter estimates for the CLIM, Plateau hepatitis data (2022)

Parameter	Estimates	Bootstrap SEs
$\hat{a}$	0.009636604	7.786835e-06
$\hat{c}$	0.963660385	1.956268e-09
$\hat{\tau}$ tau	1.499339432	8.666742e-10
Mean Time to Infection	12.13374	DM 95% CI-MTI [12.1198, 12.14767]

Table 2 presents the maximum likelihood estimates and bootstrap standard errors for the parameters (a, c, and τ) of the CLIM fitted to the 2022 hepatitis data from Plateau State. The estimated slope parameter $\hat{a} = 0.00964$ indicates that the force of infection increases gradually with age, suggesting a slow but steady rise in hepatitis exposure risk as individuals grow older. The estimate of the intercept-shifting parameter $\hat{c} = 0.96366$ is small and precisely estimated (bootstrap SE on the order of 1.956268e-09, implying minimal deviation from a near-linear age-infection relationship when ages are shifted by c. The threshold parameter $\hat{\tau} = 1.4993$ years indicates that the model identifies a short "latency window" in early childhood before the force of infection begins to increase measurably. This is consistent with public health patterns where hepatitis infections in Plateau State are relatively rare among infants but begin to accumulate more noticeably from early childhood onward. The MTI, estimated at 12.13 years with a narrow delta-method 95 per cent confidence interval [12.12, 12.15], provides a population-level summary of the average waiting time until infection occurs under the CLIM assumptions. The precision of this estimate, driven by the very small standard errors of the model parameters, reflects a stable age-infection structure within the surveillance data. Biologically, an MTI of approximately 12 years suggests that hepatitis exposure in Plateau State accumulates slowly, with the majority of infections occurring in late childhood and adolescence rather than early in life. This pattern aligns with the observed epidemiological context in which hepatitis transmission is primarily driven by behavioural and environmental risk factors that become more relevant as individuals age. The tight confidence interval further indicates that the CLIM provides internally consistent estimates of transmission timing.

A useful perspective on the Plateau hepatitis estimates is obtained by comparing them with the classic catalytic analysis of measles transmission presented in Griffiths [5]. In Griffiths' work, the estimated force of infection parameters indicated a rapidly increasing infection hazard during early childhood, with measles infections typically occurring within the first few years of life. Specifically, Griffiths' fitted linear catalytic models produced high values of the slope parameter (a) and very small threshold ages (τ), reflecting the strong epidemic pressure of measles and its early-age transmission pattern. Consequently, the mean age at infection for measles, often 3-5 years in pre-vaccination settings, was much lower than the MTI observed for hepatitis in Plateau State. The steep rise in measles FOI in Griffiths' model contrasts sharply with the slow, gradual increase captured in the Plateau hepatitis data. In contrast, the CLIM fitted to the hepatitis data yields a much smaller slope parameter $\hat{a} = 0.00964$ and a mean time to infection of approximately 12.13 years, indicating a far more protracted accumulation of risk across the lifespan. Whereas Griffiths' measles model captures a vivid early-age infection spike, the Plateau hepatitis model reflects a disease with lower transmissibility**,** more heterogeneous exposure pathways, and greater dependence on behavioural, environmental, or iatrogenic factors that become relevant in later childhood and adolescence. Moreover, the estimated threshold parameter $\hat{\tau} = 1.50$ years suggests a modest delay before meaningful exposure occurs, again consistent with hepatitis but very different from measles, where infection pressure is present almost from birth in highly endemic populations. These differences highlight the flexibility of catalytic models but also emphasise that infection-age structures are deeply disease-specific. CLIM, when applied to hepatitis, produces a smooth, gradual FOI curve, whereas the same model, applied to a classic childhood infection such as measles, as in Griffiths [5], captures a sharply accelerated early-life infection process. This comparison gives the appropriateness of catalytic modelling for a broad class of infections while illustrating how parameter contrasts reveal fundamental differences in disease ecology and transmission dynamics. Table 3 compares parameter estimates from the CLIM with those from alternative model specifications (Weibull and exponential distributions) for the Plateau hepatitis data, allowing assessment of model fit and parameter stability across different functional forms.

Table 3 summarises parameter estimates and bootstrap standard errors for the CLIM and two competing parametric models: the Weibull and the Exponential. The CLIM parameters exhibit small bootstrap standard errors, indicating a stable likelihood for the hepatitis data when modelled through the catalytic framework. The slope parameter $\hat{a} = 0.00964$ and the shift parameter $\hat{c} = 0.96366$ suggest a slow, gradual rise in the force of infection as age increases, while the estimated threshold $\hat{\tau} = 1.4993$ years identify the point at which the age-specific infection hazard begins to increase meaningfully. These estimates are consistent with a disease whose exposure risk accumulates over a prolonged developmental period rather than peaking in early childhood. The precision of the bootstrap SEs demonstrates that the CLIM structure aligns well with the underlying age-infection pattern observed in the Plateau population.

Table 3. Parameter estimates for the CLIM with competing models, Plateau hepatitis data (2022)

Models	Parameter	Estimates	Bootstrap SEs
CLIM	$\hat{a}$	0.009636604	7.786835e-06
	$\hat{c}$	0.963660385	1.956268e-09
	$\hat{\tau}$	1.499339432	8.666742e-10
Weibull	$\hat{a}$	0.61024	0.03558
	$\hat{c}$	72.78961	9.03743
exponential	$\hat{c}$	0.02962	0.00065

The Weibull model provides greater flexibility in shape through the parameters, shape $\hat{a} = 0.61024$ and scale $\hat{c} = 72.79$, with bootstrap standard errors that are moderate in size, reflecting uncertainty consistent with real-world hepatitis incidence data. The shape parameter below 1 indicates a decreasing hazard function, which contrasts with the CLIM's increasing linear hazard. This difference implies that the Weibull model infers a higher hepatitis risk earlier in life than later, a dynamic not fully supported by the observed dataset. Meanwhile, the Exponential model, with a constant hazard estimate $\hat{c} = 0.02962$, represents the simplest competing model but imposes the unrealistic assumption of age-independent infection risk. Although its parameter is statistically well-estimated (bootstrap SE = 0.00065), the biological interpretation, a constant force of infection, is at odds with hepatitis transmission characteristics, which are shaped by age-related behaviours and exposures. In general, the differences across the models emphasise that CLIM provides a more epidemiologically coherent and statistically stable representation of hepatitis transmission in Plateau State, capturing the slow, cumulative nature of exposure that neither the Exponential nor the Weibull model fully accommodates. Table 4 presents a model performance comparison based on log-likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC), demonstrating that the CLIM provides the best fit to the Plateau hepatitis data among the models considered.

Table 4. Model performance comparison with competing Models, Plateau hepatitis data (2022)

Model	LogLik	AIC	BIC
CLIM	-114.26	234.52	235.11
Weibull	-146.81	297.63	298.02
Exponential	-200.45	402.91	403.11

Table 4 shows the log-likelihood and information criteria (AIC and BIC) for the CLIM, Weibull, and Exponential models fitted to the Plateau State hepatitis data for 2022. The CLIM achieved the highest log-likelihood (-114.26), indicating the best overall fit to the observed age-specific infection proportions among the competing models. This superior performance is reinforced by its notably smaller AIC (234.52) and BIC (235.11). The Weibull model, although more flexible than the Exponential model, performs worse than CLIM, with higher AIC (297.63) and BIC (298.02) values. The Exponential model shows the poorest performance overall, with the lowest log-likelihood and the largest information criteria values. These results demonstrate that assuming a constant or monotonically decreasing hazard (as in the Exponential and some Weibull cases) does not adequately capture the observed hepatitis transmission dynamics in this population.

The strong relative performance of the CLIM suggests that a linearly increasing FOI across age groups provides the most epidemiologically coherent representation of hepatitis transmission in Plateau State. This aligns with the understanding that hepatitis exposure risk increases gradually through childhood and adolescence due to cumulative environmental, behavioural, and social-contact factors. Compared with the Weibull and Exponential models, which either enforce an inappropriate constant hazard or impose a hazard shape inconsistent with the data, the CLIM's structural assumptions better reflect the observed trend of rising infection rates at older ages. The substantial improvement in both AIC and BIC reinforces that this enhanced fit is not merely due to the additional parameters in CLIM but reflects a genuine improvement in capturing underlying infection-age patterns. In general, the model performance comparison provides compelling evidence that the CLIM is the most suitable distribution for characterising hepatitis transmission dynamics in Plateau State in 2022.

Figure 8 shows the model's performance evaluation values compared with those of competing models. The combined evidence, low AIC/BIC for CLIM, parameter estimates indicating a better-fitting model, followed by the Weibull and exponential models. The fitted CLIM yields a slope $\hat{a} = 0.0096366$, offset $\hat{c} = 0.9636604$, and threshold $\hat{\tau} = 1.4993394$. Interpreting the hazard $\psi(t) = a(t + c)$ for $t > \tau$, the estimated parameters imply that infection risk becomes non-negligible after roughly 1.5 years of age and then increases approximately linearly with age; for example, the instantaneous hazard at age 10 is $\psi(10) \approx 0.0096366 \times (10 + 0.9637) \approx 0.106$ per year. The reported bootstrap standard errors are numerically very small (e.g. $SE(a) \approx 7.79 \times 10^{-6}$), which on the face of it suggests high precision driven by large sample sizes in the surveillance system. However, such small SEs, especially those near machine precision for c and τ, should be treated cautiously: they may reflect numerical conditioning of the likelihood, parameter scaling (units), or an overconfident bootstrap when the model is tightly constrained, rather than genuine biological certainty. The MTI is estimated at 12.13 years with a delta-method 95% CI of [12.1198, 12.1477]. This indicates that under the fitted CLIM, the average susceptible individual in Plateau State acquires hepatitis infection in early adolescence. The narrow CI signals

low sampling variability in MTI estimates under the model, but, as with the parameter SEs, one should check robustness: run profile-likelihood for τ, inspect bootstrap percentile intervals for MTI, and compare MTI across alternative FOI forms (Weibull, Exponential). If MTI remains stable across those checks, the finding has direct public-health significance: interventions targeted to school-age and adolescent cohorts (vaccination, screening, behaviour/intervention programs) would likely exert the greatest effect in reducing transmission.

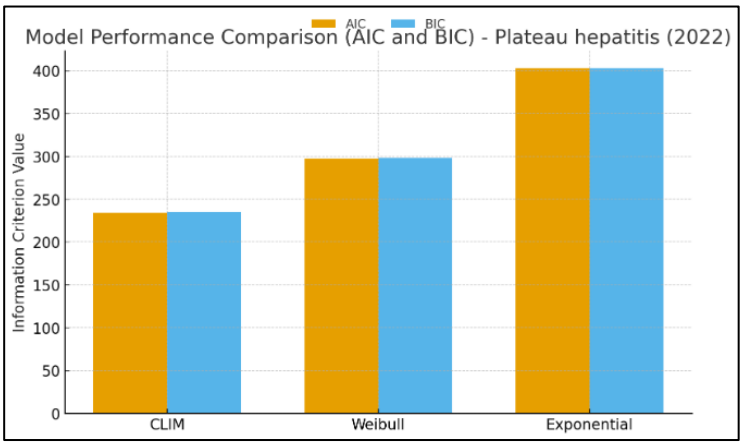

Figure 8. Model performance comparison plot

The combined evidence, low AIC/BIC for CLIM, parameter estimates indicating a modest linear rise in FOI after ~1.5 years, and MTI ≈ 12 years, supports a substantive conclusion: in Plateau State, the age-pattern of hepatitis infection is best represented (among the tested families) by a smoothly increasing force of infection that is relatively weak compared to highly childhood-transmitted pathogens such as measles. This aligns with the literature showing that hepatitis seroepidemiology often differs by type and context (HAV commonly infects very young children in high-transmission settings, HBV/HCV can show perinatal or later exposure), and that flexible comparative modeling (CLIM, Weibull, and Exponential) is necessary to identify the most appropriate hazard form (Hens et al., Chen et al., Ximenes et al.). As Griffiths advocated, model choice matters: an inappropriate constant-hazard model would misestimate MTI and mislead intervention timing; here, the CLIM better captures the data's monotone increase and thus yields more useful MTI estimates for policy.

3.2 Simulation Study

To evaluate the statistical performance and robustness of the CLIM under controlled conditions, a simulation study was conducted. This experiment aimed to assess the accuracy, bias, and efficiency of the maximum likelihood estimators of the model parameters a, c and the derived MTI across varying sample sizes. Using predefined parameter values, datasets were generated via inverse-transform sampling, replicating realistic infection patterns across multiple replications. This approach allowed a comprehensive examination of how the estimators behave as the sample size increases and how well the model recovers true parameter values across repeated samples. The results, summarised in Table 5, provide insight into the consistency and reliability of the CLIM estimators, including convergence properties and estimate precision across different sampling scenarios. Table 5 summarises the simulation study results evaluating the performance of the CLIM across varying sample sizes, reporting bias, and root mean square error (RMSE) for key parameters (a, c, and Mean Time to Infection).

Table 5. Simulation Study Results for the CLIM with $a = 0.05$, $c = 1.0$

Sample Size	Mean a	Bias a	RMSE a	Mean c	Bias c	RMSE c	Mean MTI	Bias MTI	RMSE MTI
20	0.055	0.005	0.021	1.313	0.313	1.751	4.816	–0.05619	0.526
30	0.052	0.002	0.017	1.491	0.491	2.148	4.812	–0.05977	0.493
40	0.055	0.005	0.016	1.241	0.241	2.550	4.823	–0.04927	0.427
50	0.054	0.004	0.015	1.066	0.066	1.293	4.835	–0.03707	0.421
100	0.051	0.001	0.011	1.104	0.104	0.966	4.858	–0.01341	0.262
200	0.051	0.001	0.007	1.106	0.106	0.630	4.833	–0.03867	0.193
300	0.050	–0.00003	0.006	1.073	0.073	0.515	4.867	–0.00506	0.133
400	0.051	0.001	0.006	0.977	–0.02304	0.466	4.897	0.026	0.121
500	0.050	–0.0003	0.004	1.044	0.044	0.348	4.879	0.007	0.124

Table 5 presents a simulation-based evaluation of the maximum likelihood estimator for the CLIM under controlled conditions with true parameter values ($a = 0.05$) and ($c = 1.0$). Across all sample sizes, the estimator for (a) shows good

finite-sample behaviour: the mean estimated values remain close to the true value, with bias decreasing systematically as sample size increases. For instance, the bias for (a) drops from approximately 0.0049 at $(n = 20)$ to near zero at $(n = 300)$ and $(n = 500)$. Correspondingly, the RMSE for (a) reduces steadily from 0.0207 at $(n = 20)$ to 0.0043 at $(n = 500)$. This monotonic improvement demonstrates the estimator's consistency and increasing precision for the slope parameter of the hazard function. Similar trends are observed for parameter (c), although the estimator for (c) is more variable at small sample sizes, showing large RMSE values (e.g., 1.75 at $(n = 20)$). As the sample size increases, estimates stabilise considerably, achieving RMSE < 0.35 for $(n \geq 500)$. This is expected because (c), acting as a shift parameter, is traditionally harder to estimate than (a), especially under sparse data. The MTI estimates further validate the reliability of CLIM. Despite moderate variability at smaller sample sizes, MTI estimates remain close to the true theoretical value of approximately 4.87. The bias in MTI shrinks rapidly with sample size, moving from -0.056 at $(n = 20)$ to values nearly indistinguishable from zero for sample sizes above 300. The RMSE equally declines from 0.53 at $(n = 20)$ to about 0.12 at $(n = 500)$. Importantly, the number of successful optimisation runs is consistently high $(\geq 97\%)$, indicating that the likelihood surface for CLIM is well-behaved even in small samples.

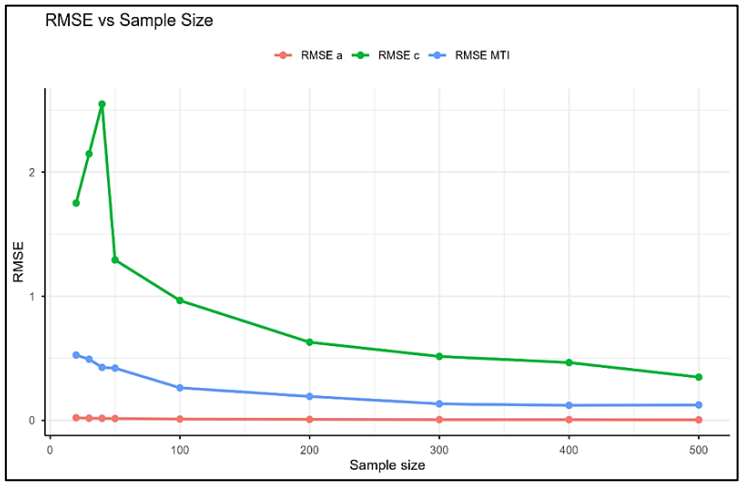

Figure 9. RMSE Plot of the Simulation Study Results for the CLIM

Figure 9 illustrates the relationship between the root mean square error (RMSE) of the estimated parameters and the sample size for the CLIM under simulated data studies. The RMSE values for the parameters (a), (c), and the derived MTI were computed across multiple replications to evaluate estimator accuracy and stability. The plot shows a general decline in RMSE values with increasing sample size, indicating that the estimators are consistent and asymptotically efficient, as expected from maximum likelihood theory. Specifically, the RMSE of parameter (a) remains relatively small and stable across all sample sizes, indicating high precision even in smaller samples. The RMSE of parameter (c) shows initial variability for small samples but decreases sharply with increasing sample size, reflecting improved estimation reliability as more data become available. Similarly, the RMSE of the MTI declines steadily, confirming the robustness of the derived epidemiological measure. These findings collectively validate the efficiency and consistency of the CLIM estimators and confirm that the model performs well under repeated sampling, mirroring the conclusions drawn by Griffiths [5] in his seminal work on measles infection modelling.

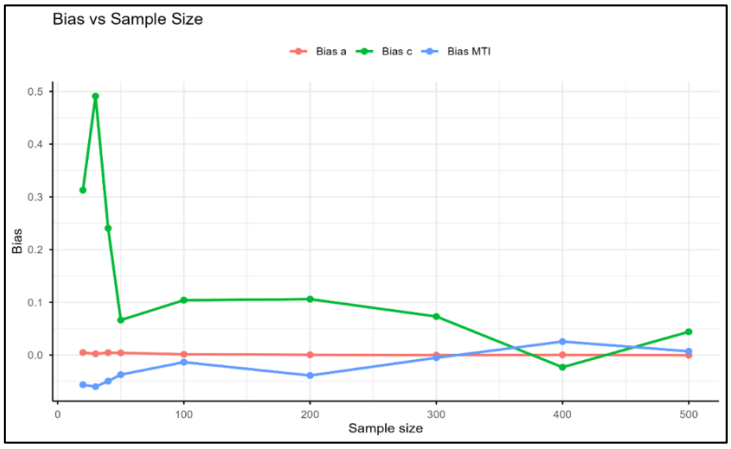

Figure 10. Bias plot of the simulation study results for the CLIM

Figure 10 illustrates the relationship between the biases of the estimated parameters and sample size for the CLIM under simulated data. The bias values for the parameters (a), (c), and the derived MTI were computed across multiple replications to evaluate estimator accuracy and stability. The plot shows that as sample sizes increase, the bias values

approach zero, demonstrating that the estimators are consistent and asymptotically efficient, as expected from maximum likelihood theory.

4. Conclusions

The findings of this research demonstrate that the CLIM offers a superior and epidemiologically meaningful framework for understanding hepatitis transmission dynamics in Plateau State. The model's ability to incorporate a threshold age and a linearly increasing force of infection aligns well with the observed data, capturing both early-childhood susceptibility and the gradual accumulation of infection risk through adolescence. The significantly lower AIC and BIC values for the CLIM compared to alternative models affirm its statistical advantage. At the same time, the estimated MTI of approximately 12 years provides important insight into the typical age at which individuals acquire hepatitis infection in this setting. Importantly, CLIM's strong performance across real-world data and simulation scenarios indicates that it is a reliable and interpretable tool for age-structured infectious disease modelling. The study extends the foundational work of Griffiths [5] by demonstrating that catalytic modelling remains relevant and effective for contemporary epidemiological challenges, especially where disease exposure increases progressively with age. The results highlight the need for timely, age-targeted interventions and support the integration of catalytic models into hepatitis surveillance and planning in Nigeria. Based on the findings of this research, several strategic recommendations are proposed to enhance public health outcomes in Plateau State. Primarily, age-targeted hepatitis prevention programmes should be strengthened, with a specific focus on older children and adolescents, as the Catalytic Linear Infection Model identifies this demographic as having the highest force of infection. To sustain these efforts, state health authorities should incorporate catalytic modelling into routine hepatitis surveillance, enabling continuous monitoring of transmission intensity and the objective evaluation of intervention impacts over time. Furthermore, it is recommended that Mean Transmission Intensity (MTI) estimates derived from the CLIM be adopted to refine the scheduling of preventive strategies, including vaccination catch-up programmes, screening campaigns, and public health education specifically targeting pre-adolescent populations. The effectiveness of these models, however, relies on data integrity; thus, efforts must be made to enhance the quality and granularity of surveillance data through age-specific case definitions and clear differentiation between hepatitis types. Finally, periodic re-estimation of CLIM parameters is essential to detect potential shifts in transmission patterns resulting from changes in public health policy, migration, or shifts in immunisation coverage.

Acknowledgement

The authors gratefully acknowledge the Department of Statistics, Ahmadu Bello University, Zaria, for providing the academic support and research facilities that made this study possible. The authors also appreciate the official public health surveillance systems in Plateau State, Nigeria. In particular, age-stratified hepatitis case data were obtained from the Surveillance, Outbreak Response, Management and Analysis System (SORMAS) dashboard maintained by the Surveillance Unit of the Plateau State Ministry of Health.

Funding

This study was not supported by any grants from funding bodies in the public, private, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare no conflict of interest to report regarding this study.

CRediT Authorship Contribution Statement

Ibrahim Abubakar Sadiq (Writing – original draft; Methodology; Data curation; Formal Analysis; Investigation)
Kenret Danjan (Data curation; Formal Analysis; Investigation)
Jamila Abdullahi (Writing – original draft; Writing – review & editing)
Yahaya Zakari (Writing – original draft; Conceptualisation; Writing – review & editing)

Availability of the Data and Materials

The data used to support the findings of this study are included within the article

Ethical Declarations

This study did not involve human participants or animals. Ethical approval was therefore not required.

Generative Artificial Intelligence Declarations

The authors claim that artificially intelligent-assisted technologies, such as generative AI, were not used to generate content, ideas, or theories. We have just utilised AI to enhance readability and refine the language. This was used with extreme human control and oversight. The authors take full responsibility for reviewing and approving the content.

References

[1] Muench H. Catalytic models in epidemiology. Cambridge (MA): Harvard University Press; 1959.
[2] Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Oxford: Oxford University Press; 1991.

[3] Li X, Zhang Y, He A, Li Q, Wang S, Shan J, et al. Biohybrid catalysis in biomedicine. Coordination Chemistry Reviews. 2025;545:217003.
[4] Hethcote HW. The mathematics of infectious diseases. SIAM Review. 2000;42(4):599-653.
[5] Griffiths DA. A catalytic model of infection for measles. Journal of the Royal Statistical Society: Series C (Applied Statistics). 1974;23(3):330-339.
[6] Grenfell BT, Anderson RM. The estimation of age-related rates of infection from case notifications and serological data. Journal of Hygiene. 1985;95(2):419-436.
[7] Keeling MJ, Rohani P. Modelling infectious diseases in humans and animals. Princeton (NJ): Princeton University Press; 2008.
[8] Zhang S, Cui F. Global progress, challenges, and strategies in eliminating the public threat of viral hepatitis. Infectious Diseases of Poverty. 2025;14(1):1-4.
[9] Falodun MO, Olorunfemi O, Irinoye OO. Infectious diseases: addressing global challenges and prevention strategies for national health improvement. Community-Acquired Infection. 2025;12.
[10] Nwude VN, Lesi OA, Onyekwere C, Charpentier E, Hübschen JM. Clinical characteristics of hepatitis B virus-associated hepatocellular carcinoma patients in Southwest Nigeria. Pathogens. 2025;14(2):169.
[11] Kwon SY, Lee CH. Epidemiology and prevention of hepatitis B virus infection. The Korean Journal of Hepatology. 2011;17(2):87.
[12] Kamau E, Chen J, Bajaj S, Torres N, Creswell R, Pavlich-Mariscal JA, et al. The mathematics of serocatalytic models with applications to public health data. Statistics in Medicine. 2025;44(15-17):e70188.
[13] Grosso A, Hens N, Abrams S. An integrative review of the combined use of mathematical and statistical models for estimating malaria transmission parameters. Malaria Journal. 2025;24(1):173.
[14] Blaizot S, Herzog SA, Abrams S, Theeten H, Litzroth A, Hens N. Sample size calculation for estimating key epidemiological parameters using serological data and mathematical modelling. BMC Medical Research Methodology. 2019;19(1):51.
[15] Heisey DM, Joly DO, Messier F. The fitting of general force-of-infection models to wildlife disease prevalence data. Ecology. 2006;87(9):2356-2365.
[16] Norman RA, Chan MS, Srividya A, Pani SP, Ramaiah KD, Vanamail P, et al. EPIFIL: the development of an age-structured model for describing the transmission dynamics and control of lymphatic filariasis. Epidemiology & Infection. 2000;124(3):529-541.
[17] Uguru LI, Akyala AI, Ngwai YB, Uguru DI, Aremu SO. Clinical burden and biochemical profiles of viral hepatitis in a tertiary healthcare facility in North Central, Nigeria. BMC Infectious Diseases. 2025;25(1):1368.
[18] Semary H, Sadiq IA, Doguwa SIS, Ishaq AI, Suleiman AA, Daud H, et al. Advancing survival regression using the NGOF exponentiated Weibull distribution for vesicovaginal fistula and radiation data applications. Journal of Radiation Research and Applied Sciences. 2025;18(2):101497.
[19] Akor A, Sadiq IA, Usman A, Doguwa SI, Akor LO. Advanced survival modelling of tuberculosis patients: insights from exponential and Weibull AFT models. Fudma Journal of Sciences. 2025;9(3):169-182.
[20] World Health Organization. Immunization agenda 2030: a global strategy to leave no one behind [Internet]. 2020 [cited 2026 Mar 5]. Available from: https://www.who.int
[21] Gavi, the Vaccine Alliance. Childhood immunisation in low- and middle-income countries: progress and challenges [Internet]. 2019 [cited 2026 Mar 5]. Available from: https://www.gavi.org
[22] Metcalf CJE, Lessler J, Klepac P, Cutts F, Grenfell BT. Impact of birth rate, seasonality, and transmission rate on minimum levels of coverage needed for rubella vaccination. Epidemiology and Infection. 2012;140(12):2290-2301.
[23] United Nations. Transforming our world: the 2030 agenda for sustainable development [Internet]. 2015 [cited 2026 Mar 5]. Available from: https://www.un.org/sustainabledevelopment
[24] McLean AR, Anderson RM. Measles in developing countries. Part II. The predicted impact of mass vaccination. Epidemiology and Infection. 1988;100(3):419-442.
[25] Pitzer VE, Viboud C, Simonsen L, Steiner C, Panozzo CA, Alonso WJ, et al. Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. Science. 2009;325(5938):290-294.
[26] Hens N, Aerts M, Faes C, Shkedy Z, Lejeune O, Van Damme P, et al. Seventy-five years of estimating the force of infection from current status data. Epidemiology & Infection. 2010;138(6):802-812.
[27] Hay JA, Routledge I, Takahashi S. Serodynamics: a primer and synthetic review of methods for epidemiological inference using serological data. Epidemics. 2024;48:100806.
[28] Wariri O, Muhammad AK, Sowe A, Strandmark J, Utazi CE, Metcalf CJE, et al. Serological survey to determine measles and rubella immunity gaps across age and geographic locations in The Gambia: a study protocol. Global Health Action. 2025;18(1):2540135.
[29] Mann J, Roberts M. Modelling the epidemiology of hepatitis B in New Zealand. Journal of Theoretical Biology. 2011;269(1):266-272.
[30] Hagan OC, Nsiah P, Obiri-Yeboah D, Yirdong F, Annan I, Eliason S, et al. Impact of universal childhood vaccination against hepatitis B in Ghana: a pilot study. Journal of Public Health in Africa. 2018;9(2):721.
[31] Shapiro CN. Epidemiology of hepatitis B. The Pediatric Infectious Disease Journal. 1993;12(5):433-437.
[32] Nelson NP, Easterbrook PJ, McMahon BJ. Epidemiology of hepatitis B virus infection and impact of vaccination on disease. Clinics in Liver Disease. 2016;20(4):607.

[33] Kramvis A, Mammas IN, Spandidos DA. Exploring the optimal vaccination strategy against the hepatitis B virus in childhood. Biomedical Reports. 2023;19(1):48.
[34] Sadiq IA, Kajuru JY, Doguwa SI, Yahaya GY, Hephzibah AA, Yahaya SS, et al. Survival analysis in advanced lung cancer using Weibull survival regression model: estimation, interpretation, and clinical application. Journal of Statistical Sciences and Computational Intelligence. 2025;1(2):106-123.
[35] Mohammed AA, Hamdani H, Zakari Y, Abdullahi J, Sadiq IA, Ouertani MN, et al. On the Rayleigh exponentiated odd generalised inverse exponential distribution: properties and applications. Engineering Reports. 2025;7(11):e70457.