frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 14 November 2024
DOI 10.3389/fpubh.2024.1437678

OPEN ACCESS

EDITED BY
Sana Tamim,
National Institute of Health, Pakistan

REVIEWED BY
Shewafera Wondimagegnhu Teklu,
Debre Berhan University, Ethiopia
Ali Raza,
The University of Chenab, Pakistan

*CORRESPONDENCE
Sangil Kim
✉ sangil.kim@pusan.ac.kr

RECEIVED 24 May 2024
ACCEPTED 15 October 2024
PUBLISHED 14 November 2024

CITATION
Abbas W, Masud MA, Parveen S, Lee H and Kim S (2024) Data-driven analysis of the effect of screening and treatment on the spread of HIV in developing and developed countries. *Front. Public Health* 12:1437678. doi: 10.3389/fpubh.2024.1437678

Data-driven analysis of the effect of screening and treatment on the spread of HIV in developing and developed countries

Wasim Abbas[1], M. A. Masud[2], Sajida Parveen[1], Hyojung Lee[3] and Sangil Kim[1,4]*

[1]Department of Mathematics and Institute of Mathematical Sciences, Pusan National University, Busan, Republic of Korea, [2]Integrated Mathematical Oncology Department, H Lee Moffitt Cancer Center and Research Institute, Tampa, FL, United States, [3]Department of Statistics, Kyungpook National University, Daegu, Republic of Korea, [4]Institute for Future Earth, Pusan National University, Busan, Republic of Korea

Introduction: In this study, we used a mathematical epidemic model to explore the status of the HIV epidemic in the USA and Pakistan. In addition to studying the dynamics of the model, we fitted the model with recent data to estimate the parameters describing the epidemic in both countries.

Results: Our estimation shows that in the USA, the reproduction number is 0.9688 (0.9684, 0.9694); if the reproduction number is maintained at this level, it would take a long time to eradicate HIV entirely. Meanwhile, it is 2.2599 (2.2556, 2.2656) in Pakistan, which is due to a lack of awareness in the confirmed group and a lower rate of maintained treatment. We also estimated the rate of vertical transmission, which plays a significant role in Pakistan but not in the USA.

Discussion: We conclude that improving the screening rate and educating people would be effective for controlling HIV in Pakistan, whereas improved screening rate in the USA can eradicate HIV faster.

KEYWORDS

HIV/AIDS, horizontal and vertical transmission, maximum likelihood method, screening, treatment

1 Introduction

Human immunodeficiency virus (HIV) attacks the body's immune system, leading to acquired immunodeficiency syndrome (AIDS) in the chronic stage. AIDS is among the most devastating diseases in human history. The dynamics of HIV transmission are influenced by both horizontal (e.g., sexual contact) and vertical (mother-to-child) pathways, and the impact of these pathways varies widely by region. However, awareness, proper treatment, and care can help to control the spread of the disease. Cutting-edge treatment of an infected person may increase lifespan, improve health, and reduce the risk of both types of transmission (1).

The incidence of HIV/AIDS is very high in most developing nations, with some countries experiencing worsening condition daily (2). For instance, HIV cases in Pakistan have increased rapidly in recent years (3). The first HIV-infected person in Pakistan was detected in 1987 (4, 5). The prevalence rate of HIV is still not high in Pakistan, but in recent years, the HIV-infected population has greatly increased in Pakistan. According to UNAIDS (6), from 2010 to 2018, the number of new HIV infections increased from 14,000 to 22,000, while AIDS-related deaths surged by over 350% since 2010.

In contrast, developed countries like the USA have seen a decline in new infections due to consistent efforts to improve awareness and treatment. As of 2018, ∼84% of the HIV population in the USA had been diagnosed, though one in seven individuals do not know the status of their infection (7). In Pakistan, only 14% of the HIV population knew the status of their infection, and a mere 10% were receiving treatment (6). This stark difference in screening and treatment rates between the USA and Pakistan underscores the need for targeted intervention strategies.

Mathematical modeling plays an important role in understanding epidemic diseases such as HIV. The pioneering mathematical model of HIV was proposed by Anderson and May in 1986 (8–10); this model has been further refined by many researchers (11–22). For instance, in Busenberg et al. (13), the role of sex workers was investigated, and dilution by increasing the number of sex workers was shown to be an effective measure for decreasing HIV incidence. However, this was later criticized, and a decrease in the sex industry with the use of condoms was proven more effective (21). A differential infectivity model stressed the importance of identifying super spreaders through contact tracing besides the use of condoms (15). Contact tracing was also emphasized in the study of De Arazoza and Lounes (23). Another model on a homosexual cohort with differential infectivity showed that reducing the number of partners is key to reducing the incidence of HIV (11).

In addition to horizontal transmission, HIV can also be transmitted vertically to newborns from infected parents. According to UNAIDS (24), 160,000 juveniles (aged 0–14 years) became infected globally in the year 2018 alone. Pakistan is also confronted with the devastating impact of the epidemic on its young population. The number of infected juveniles in Pakistan increased continuously from 2010 to 2018 (25). This may divert health and welfare resources. In Naresh et al. (26), the authors modeled horizontal and vertical transmission explicitly. They considered that a fraction of newborns were HIV-infected and grouped into the infectious class. However, newborns are not infectious, as they do not transfer the disease either horizontally or vertically until adulthood. This was further addressed in López et al. (27), where the authors proposed a two-age group model considering horizontal and vertical transmission, and only the adult infected group was responsible for either type of transmission. In a recent study, the role of vertical transmission was modeled by moving a fraction (proportional to the fraction of infected) of the new recruitment into the infectious class (28). Testing of pregnant women has been recommended to reduce vertical transmission (29).

The role of awareness in controlling HIV and AIDS was modeled in Kaur et al. and Kaymakamzade et al. (30, 31) by considering the transmission rate as an asymptotically decreasing function of the number of infected individuals. The roles of screening and treatment have been recognized by many authors (23, 32–36). The combined role of condom usage, screening, and treatment was investigated using optimal control analysis (33). As treatment increases the lifespan of an infected individual, it may also increase the incidence (27). However, treatment implemented with sufficient awareness is effective in reducing the incidence of HIV (33).

As the disease may be transmitted both horizontally and vertically, treatment should impact the transmission in both ways. Therefore, it is important to also understand the impact of treatment on vertical transmission to deal with the growing number of infected juveniles. However, existing models with vertical transmission are tailored with minor limitations regarding non-infectious juveniles becoming infected (26, 28, 29). In the two-age group model proposed by López et al. (27), only the adult class of infected individuals is considered responsible for transmitting the infection. The prevailing focus of some studies, including those conducted by the authors in Olaniyi et al. (37) and Alhassan et al. (38), is on general horizontal and vertical transmission, overlooking the inclusion of juvenile populations or real-world data, and the absence of country-specific analysis. On the other hand, recent studies conducted by Khan et al. (39) and Teklu and Mekonnen (40) highlighted the significance of incorporating awareness strategies and real-world data into epidemic models. In a similar vein, Teklu and Rao (41) emphasize the importance of identifying key transmission pathways to inform interventions that are effective, especially in settings where resources are limited. In the context of HIV/AIDS, researchers have also investigated delay strategies and stochastic models to achieve insights into the impact of delayed treatment on transmission dynamics. These studies underscore the significance of timely interventions in controlling the epidemic (42, 43). Our study expands upon these existing foundations by integrating awareness and treatment interventions into a compartmental model to determine the effects of these strategies on both horizontal and vertical transmission routes.

Our study further develops these understandings by providing quantitative estimations of the contributions made by both horizontal and vertical transmission pathways to the spread of the HIV epidemic in two distinct socioeconomic contexts—the USA and Pakistan. Based on the information available to us, this study is the first to quantify the relative impacts of these pathways and propose country-specific recommendations grounded in real-world data.

We also modeled the role of treatment as an extension of infectious life, which resulted in a "treat or not to treat" dilemma. The available cutting-edge treatments not only increase life span but also reduce the probability of transmission both vertically and horizontally. Furthermore, by incorporating both horizontal and vertical transmission pathways and distinguishing between treated and untreated infected populations, this work provides a detailed and practical framework for policymakers to optimize interventions suited to their national circumstances.

2 Mathematical modeling

We incorporated the heterogeneity in transmission from infected people, and people assumed that people maintained/continued their treatment in the model as proposed by López et al. (27). We denote the size of the total population at time t by $N(t)$, which is divided into a juvenile class $J(t)$ and adult class $A(t)$. The juvenile class is further divided into two sub-classes: susceptible juvenile population $J_1(t)$ and infected juvenile population $J_2(t)$. The adult class is divided into three

sub-classes: susceptible adult population $A_1(t)$, infected adult population not in treatment $A_2(t)$, and infected adult population in treatment $A_3(t)$.

We consider both horizontal and vertical transmission. Horizontal transmission occurs when an individual in class A_1 acquires infection from individuals in class A_2 or A_3 by risky acts, such as unprotected sexual intercourse, sharing needles, etc., with a transmission rate v_1 or v_2, respectively, and this individual moves to class A_2. We assume that juveniles are sexually inactive and can only be infected through vertical transmission. Vertical transmission occurs during birth from individuals in A_2 or A_3 with probabilities r and ϵ, respectively; the birth rate is β_2 for both classes. Therefore, the recruitment rate into class J_2 is $\beta_2(rA_2 + \epsilon A_3)$. The remaining births from A_2 and A_3, $\beta_2((1-r)A_2 + (1-\epsilon)A_3)$, are recruited to class J_1. Moreover, susceptible adults give birth to $\beta_1 A_1$ susceptible juveniles. Juveniles mature and move to A_1 and A_3 from J_1 and J_2, respectively, at a maturation rate η. Individuals from class A_2 learn their status of infection and start maintaining treatment at a rate θ and accordingly move to class A_3. It is important to note that while specific data related to testing pregnant women are utilized for certain parameters (e.g., $((r)$ and $(\epsilon))$, $(\theta))$ encompasses treatment scenarios for all infected individuals, considering various diagnosis routes. The natural death rates of the juvenile classes (J_1, J_2) and adult classes (A_1, A_2, A_3) are μ_1 and μ_2, respectively. We assume that the infected individuals in A_3 maintain treatment, and as a result, the disease-related death rate is negligible. However, the disease-related death rate for individuals in J_2 and A_2 is α.

We propose juveniles play no significant role in HIV transmission, a claim that is supported by the extensive South African Petra Study (44). The study effectively illustrated that individuals under the age of 15 handled a mere 2% of all HIV transmissions, underscoring the limited autonomous contribution of juveniles in fueling the HIV epidemic. Our modeling approach aligns closely with the guidelines established by the World Health Organization on Mother-to-Child Transmission (MTCT) (45), which emphasize interventions to prevent HIV transmission from infected mothers to their children. This perfectly aligns with our modeling assumption, emphasizing the significance of maternal transmission. Infected juveniles are assumed to transition into infected adults who receive treatment upon reaching adulthood. This assumption aligns with our specific scenario and contributes to our study. The density-dependent death rate is mN for all classes. A schematic diagram of the model is shown in Figure 1.

Following the aforementioned assumptions, we may express the model as follows:

$$\frac{dJ_1}{dt} = \beta_1 A_1 + \beta_2((1-r)A_2 + (1-\epsilon)A_3) - (\eta + \mu_1 + mN)J_1 \tag{1}$$

$$\frac{dA_1}{dt} = \eta J_1 - \frac{(v_1 A_1 A_2 + v_2 A_1 A_3)}{A} - (\mu_2 + mN)A_1 \tag{2}$$

$$\frac{dJ_2}{dt} = \beta_2(rA_2 + \epsilon A_3) - (\alpha + \eta + \mu_1 + mN)J_2 \tag{3}$$

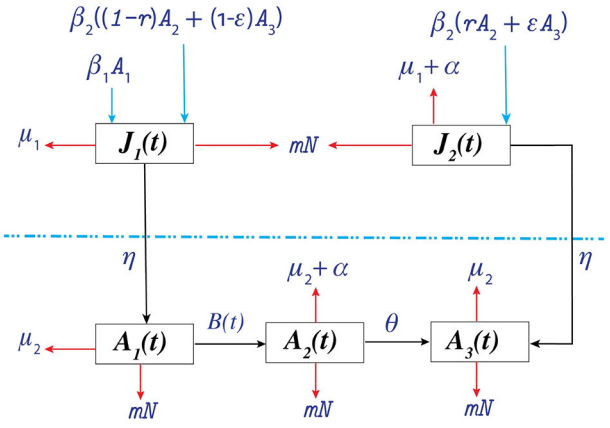

FIGURE 1
Schematic diagram of the proposed model. $B(t) = \frac{v_1 A_1 A_2 + v_2 A_1 A_3}{A}$. The juvenile and adult populations are separated by the dotted line.

$$\frac{dA_2}{dt} = \frac{(v_1 A_1 A_2 + v_2 A_1 A_3)}{A} - (\theta + \mu_2 + \alpha + mN)A_2 \tag{4}$$

$$\frac{dA_3}{dt} = \eta J_2 + \theta A_2 - (\mu_2 + mN)A_3 \tag{5}$$

where $A = A(t) = A_1(t) + A_2(t) + A_3(t)$ and $N = N(t) = J_1(t) + J_2(t) + A_1(t) + A_2(t) + A_3(t)$. We observe that $B(t)$ is continuous at $(A_1, A_2, A_3) = (0, 0, 0)$ if we define $B(0, 0, 0) = 0$. We also notice that $B \leq 2\max(A_1, A_2, A_3)$, and is Lipschitzian for $A_1 \geq 0$, $A_2 \geq 0$ and $A_3 \geq 0$. However, $B(t)$ is clearly not differentiable at $(0, 0, 0)$.

In our mathematical model, we incorporate a density-dependent death rate (mN) for all classes, a simplification commonly utilized in epidemiological models. This assumption serves to represent the impact of population density on mortality rates and imposes bounded growth. It is important to note that such modeling choices involve simplifications for mathematical tractability.

3 Model analysis

We performed a mathematical analysis of the model Equations 1–5 to understand its dynamics from an epidemiological perspective.

3.1 Equilibria

Our model exhibits one trivial equilibrium $(0, 0, 0, 0, 0)$, a disease-free equilibrium $(J_1^*, A_1^*, 0, 0, 0)$, a susceptible extinction equilibrium $(0, 0, J_2^*, 0, A_3^*)$, and an endemic equilibrium $(J_1^{**}, A_1^{**}, J_2^{**}, A_2^{**}, A_3^{**})$. As $B(t)$ is not differentiable at the trivial equilibrium, stability analysis at this equilibrium is unfeasible using standard linearization techniques. The criteria for the existence of the remaining equilibria are summarized in the following propositions.

10.3389/fpubh.2024.1437678

Proposition 1. *For the system of* Equations 1–5, *if* $\kappa_1 = \frac{\beta_1\eta}{\mu_2(\eta+\mu_1)} > 1$, *then the system has a unique disease-free equilibrium* $\left(J_1^*, A_1^*, 0,\ 0,\ 0\right)$, *where*

$$J_1^* = \frac{a_2 N_1^*}{\eta + a_2}, \quad A_1^* = \frac{\eta N_1^*}{\eta + a_2}$$

and

$$N_1^* = \frac{-(\eta+\mu_1+\mu_2) + \sqrt{(\eta+\mu_1+\mu_2)^2 - 4\mu_2(\eta+\mu_1)(1-\kappa_1)}}{2m}$$

Proposition 2. *For the system of* Equations 1–5, *if* $r = 1, \epsilon = 1$ *and*$\kappa_2 = \frac{\beta_2\eta}{\mu_2(\alpha+\eta+\mu_1)} > 1$, *then this system has a unique susceptible extinction equilibrium* $\left(0,\ 0, J_2^*, 0, A_3^*\right)$. *Here,*

$$J_1^* = \frac{a_2 N_1^*}{\eta + a_2}, \quad A_1^* = \frac{\eta N_1^*}{\eta + a_2}$$

$$J_2^{**} = \frac{a_4 a_5 (T_1 + T_2 T_5)(\Omega_1 - (1-T_2))A^{**}}{\Omega_1\Omega_2\eta},$$

$$A_2^{**} = \frac{a_5(1-T_2)(\Omega_1 - (1-T_2))A^{**}}{\Omega_1\Omega_2},$$

and

$$N_2^* = \frac{-(\eta+\mu_1+\mu_2+\alpha) + \sqrt{\kappa_3}}{2m}$$

with

$$\kappa_3 = (\eta+\mu_1+\mu_2+\alpha)^2 - 4\mu_2(\alpha+\eta+\mu_1)(1-\kappa_2).$$

Proposition 3. *For the system of* Equations 1–5, *if* $T_2 < 1$ *and* $R_0 > 1$, *then the system has a unique endemic equilibrium* $\left(J_1^{**}, A_1^{**}, J_2^{**}, A_2^{**}, A_3^{**}\right)$, *where*

$$J_1^{**} = \frac{a_5(1-T_2)(\Omega_2 + a_4(\Omega_1 - (1-T_2)))A^{**}}{\Omega_1\Omega_2\eta},$$

$$A_1^{**} = \frac{(1-T_2)A^{**}}{\Omega_1},$$

and

$$A_3^{**} = \frac{a_4(T_1+T_5)(\Omega_1 - (1-T_2))A^{**}}{\Omega_1\Omega_2}$$

with

$$A = A_1 + A_2 + A_3,\ \Omega_1 = T_3(1-T_2)$$

$+T_4(T_1+T_5)$, *and* $\Omega_2 = a_5(1-T_2) + a_4(T_1+T_5)$.

T_1, T_2, T_3, T_4, T_5, a_1, a_2, a_3, a_4, and a_5 are defined in **Section 3.2**.

Parameter (κ_1) is a critical factor in Proposition 1, influencing the existence of a unique disease-free equilibrium. Parameter κ_1: offspring number of susceptible adults.

Parameter (κ_2) plays a crucial role in Proposition 2, determining the existence of a unique susceptible extinction equilibrium. Parameter κ_2: basic offspring number of infected adults.

3.2 Basic reproduction number R_0

The basic reproduction number (R_0) is the expected number of infections produced by an infected individual in an entirely susceptible population throughout their entire infectious lifetime.

Following the approach introduced by Van den Driessche and Watmough (46), the next-generation matrix is given by

$$K = \begin{pmatrix} \frac{\epsilon\beta_2\eta}{a_3a_5} & \frac{\beta_2(ra_5+\epsilon\theta)}{a_4a_5} & \frac{\epsilon\beta_2}{a_5} \\ \frac{\eta\nu_2}{a_3a_5} & \frac{\nu_1}{a_4} + \frac{\nu_2\theta}{a_4a_5} & \frac{\nu_2}{a_5} \\ 0 & 0 & 0 \end{pmatrix}$$

where $\eta+\mu_1+mN = a_1$, $\mu_2+mN = a_2$, $\alpha+\eta+\mu_1+mN = a_3$, $\theta+\mu_2+\alpha+mN = a_4$, and, $\mu_2+mN = a_5 = a_2$.

The basic reproduction number is the spectral radius $B_1 = \begin{pmatrix} -\left(a_1+mJ_1^*\right) & \beta_1 - mJ_1^* \\ \eta - mA_1^* & -\left(a_2+mA_1^*\right) \end{pmatrix}$ of K. Thus,

$$R_0 = \frac{1}{2}\left((T_2+T_3+T_4T_5) + \sqrt{(T_2-T_3-T_4T_5)^2 + 4T_4(T_2T_5+T_1)}\right)$$

where $T_1 = \frac{r\beta_2\eta}{a_3a_4}$, $T_2 = \frac{\epsilon\beta_2\eta}{a_3a_5}$, $T_3 = \frac{\nu_1}{a_4}$, $T_4 = \frac{\nu_2}{a_5}$, and $T_5 = \frac{\theta}{a_4}$. Although the expression for R_0 is too complex to interpret, $T_1, T_2, \ldots, T_5$ produce a meaningful interpretation. The average infectious lifetimes of an infected individual in A_2 and A_3 are $1/a_4$ and $1/a_5$, respectively. The average lifetime of an infected juvenile individual is $1/a_3$; thus, the maturation probability of an infected juvenile is η/a_3. Hence, one infected individual from A_2 transmits the disease vertically to $r\beta_2 \times \frac{\eta}{a_3} \times \frac{1}{a_4} = T_1$ individuals on average throughout their infectious lifetime. Similarly, one infected individual from A_3 transmits the infection to $\epsilon\beta_2 \times \frac{\eta}{a_3} \times \frac{1}{a_5} = T_2$ individuals vertically. An infected individual belonging to A_2 transmits the disease to $\frac{\nu_1}{a_4} = T_3$ individuals on average horizontally. In addition, an infected individual belonging to A_3 transmits the disease to $\frac{\nu_2}{a_5} \times \frac{\theta}{a_4} = T_4 \times T_5$ individuals horizontally over their infectious lifetime. Thus, R_0 consists of the reproduction numbers associated with the two transmission pathways from each of the two infectious classes A_2 and A_3.

Another threshold of interest is $R_1 = \frac{\beta_1\eta}{a_1a_2}$, which is the average number of new $J_1(t)$ produced by each $A_1(t)$ during adulthood multiplied by the probability of the survival of each $J_1(t)$ during juvenility.

3.3 Stability analysis

In this subsection, we discuss the stability analysis of our system of Equations 1–5. We describe the local and global stability results for the aforementioned non-trivial equilibria. All threshold values are defined in Table 1, and the conditions are biologically meaningful. We present a series of key theorems that establish the stability properties of the system of Equations 1–5 under various conditions.

TABLE 1 All results proved for the system of Equations 1–5.

	Condition	Result
i.	$\kappa_1 > 1$ and $R_0 < 1$	Disease-free equilibrium is LAS
ii.	$\kappa_1 \leq 1 \Rightarrow R_1 < 1$ and $R_0 \geq 1$ with $r = \epsilon = 1$	Susceptible extinction equilibrium is LAS
iii.	$\kappa_1 > \frac{(\eta+\mu_1)(\nu_1+\nu_2+\mu_2)}{\eta\mu_2}$ and $T_1 + T_2 + T_3 + T_4 + T_5 < 1$	Disease-free equilibrium is GS
iv.	$r = \epsilon = 1, \kappa_1 \leq 1 \Rightarrow R_1 < 1$ and $\kappa_2 > \frac{\eta+\mu_1+\alpha}{\eta}$	Susceptible extinction equilibrium is GS
v.	$r = \epsilon = 1, \kappa_1 \leq 1 \Rightarrow R_1 < 1$ and $R_0 < 1$	Trivial equilibrium is GS

Here, LAS means locally asymptotically stable, and GS means globally stable.

3.3.1 Local stability analysis

Theorem 1. *If $\kappa_1 > 1$ and $R_0 < 1$, we demonstrate that the disease-free equilibrium is locally asymptotically stable within $\mathbb{R}^5_+$.*

Proof. If $\kappa_1 > 1$, then Proposition 1 ensures the uniqueness of the disease-free equilibrium. Now, for stability analysis, the Jacobian of the mathematical model represented by Equations 1–5 and evaluated at the disease-free equilibrium is:

$$B = \begin{pmatrix} B_1 & B_3 \\ O & B_2 \end{pmatrix}$$

where

$$B_2 = \begin{pmatrix} -a_3 & r\beta_2 & \epsilon\beta_2 \\ 0 & \nu_1 - a_4 & \nu_2 \\ \eta & \theta & -a_5 \end{pmatrix}$$

$$B_3 = \begin{pmatrix} -mJ_1^* & (1-r)\beta_2 - mJ_1^* & (1-\epsilon)\beta_2 - mJ_1^* \\ -mA_1^* & -\left(\nu_1 + mA_1^*\right) & -\left(\nu_2 + mA_1^*\right) \end{pmatrix}$$

and $O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$.

The characteristic polynomial of B_1, denoted by *charpoly* (B_1), is given as

$$\text{charpoly}\,(B_1) = \lambda^2 + b_1\lambda + b_2,$$

where $b_1 = a_1 + a_2 + mA_1^* + mJ_1^*$ and $b_2 = a_1a_2 - \beta_1\eta + a_1mA_1^* + \beta_1mA_1^* + a_2mJ_1^* + \eta mJ_1^*$.

From Proposition 1 and Equations 1–5, we note that $\eta\beta_1 = a_1a_2$. This implies that all the coefficients of *charpoly* (B_1) are positive. Hence, by the Routh–Hurwitz stability criteria, all eigenvalues of B_1 have negative real parts.

Moreover, *charpoly* $(B_2) = \lambda^3 + c_1\lambda^2 + c_2\lambda + c_3$, where $c_1 = a_3 + a_5 + a_4(1 - T_3)$, $c_2 = a_3a_4(1 - T_3) + a_3a_5(1 - T_2) + a_4a_5(1 - T_3 - T_4T_5)$, and $c_3 = a_3a_4a_5(T_2T_3 + 1 - T_2 - T_3 - T_1T_4 - T_4T_5)$.

Because $R_0 < 1$ by assumption, we have $c_1 > 0$, $c_3 > 0$, and $c_1c_2 > c_3$. Thus, according to the Routh–Hurwitz stability criteria, all the eigenvalues have negative real parts. This implies that the disease-free equilibrium is locally asymptotically stable.

The local asymptotic stability of the disease-free equilibrium suggests that the disease will eventually be eliminated from the population sufficiently close to disease-free equilibrium when $R_0 < 1$ and $\kappa_1 > 1$. Within this context, the transmission of HIV is at a minimal level, and the implemented control measures, encompassing treatment, preventive interventions, and public health strategies, demonstrate sufficient efficacy in averting the prolonged presence of the virus. The system reverts to a state devoid of disease, signifying the successful eradication of HIV within the population.

Theorem 2. *Under specific conditions involving $R_1 < 1$ and $\kappa_2 > 1$ ($\Rightarrow R_0 \geq 1$) with $r = \epsilon = 1$, we establish the local asymptotic stability of the susceptible extinction equilibrium in the domain $\mathbb{R}^5_+$.*

Proof. If $\kappa_2 > 1$, then Proposition 2 ensures the uniqueness of the susceptible extinction equilibrium. For the stability analysis of this equilibrium, the Jacobian of the mathematical model (Equations 1–5) evaluated at the susceptible extinction equilibrium is:

$$D = \begin{pmatrix} D_1 & D_3 \\ O & D_2 \end{pmatrix}$$

.

Here,

$$D_1 = \begin{pmatrix} -a_1 & \beta_1 \\ \eta & -(\nu_2 + a_2) \end{pmatrix}$$

$$D_2 = \begin{pmatrix} -\left(a_3 + mJ_2^*\right) & \beta_2 - mJ_2^* & \beta_2 - mJ_2^* \\ 0 & -a_4 & 0 \\ \eta - mA_3^* & \theta - mA_3^* & -\left(a_5 + mA_3^*\right) \end{pmatrix},$$

$$D_3 = \begin{pmatrix} -mJ_2^* & -mJ_2^* \\ 0 & \nu_2 \\ -mA_3^* & -mA_3^* \end{pmatrix}$$

and $O = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

The characteristic polynomial of D_1 denoted by *charpoly*(D_1) is given as

$$\text{charpoly}\,(D_1) = \lambda^2 + d_1\lambda + d_2,$$

while $d_1 = a_1 + a_2 + \nu_2$ and $d_2 = a_1a_2 - \beta_1\eta + a_1\nu_2$.

From supposition $R_1 < 1$, this implies that $a_1a_2 - \eta\beta_1 > 0$. Thus, all the coefficients of *charpoly*(D_1) are positive.

Moreover, *charpoly* $(D_2) = f_1(\lambda) \cdot f_2(\lambda)$.

Here,

$$f_1(\lambda) = \lambda + a_4$$

and

$$f_2(\lambda) = \lambda^2 + \left(a_3 + a_5 + mA_3^* + mJ_2^*\right)\lambda + (a_3 + \beta_2)\, mA_3^* + (a_5 + \eta)\, mJ_2^* + a_3a_5 - \eta\beta_2$$

From Proposition 2, we notice that $\eta\beta_2 = a_3a_5$; therefore all the coefficients of $f_2(\lambda)$ are positive. Thus, according to the Routh–Hurwitz stability criteria, all the eigenvalues have negative real parts. This implies that the susceptible extinction equilibrium is locally asymptotically stable.

The local stability analysis of the susceptible extinction equilibrium indicates that, given certain conditions ($R_1 < 1, \kappa_2 > 1$), the susceptible individuals will be eradicated from the population, leaving only the infected individuals. This scenario illustrates the worst possible outcome, in which HIV becomes endemic in the population due to inadequate control measures, resulting in unrestricted virus transmission.

3.3.2 Global stability analysis

Theorem 3. *For cases where* $\kappa_1 > \frac{(\eta+\mu_1)(\nu_1+\nu_2+\mu_2)}{\eta\mu_2}$ *and* $R_0 < 1$, *then* $N(t) > 0$ *(Trivial equilibrium is a Repeller) and the disease-free equilibrium is a global attractor in* $\mathbb{R}^5_+$.

Proof. To prove this theorem, we claim that $N(t) > 0$; on the contrary, suppose that $N(t) = 0$. Let us consider a function $x = A_1 + \xi_1 J_1$, where $0 < \xi_1 < 1$. Differentiating x with respect to t, along with the solution of the system of Equations 1–5, we obtain,

$$\begin{aligned} x' &= (\eta J_1 - a_2A_1) + \xi_1(\beta_1A_1 - a_1J_1) \\ &\geq \left(\tfrac{\eta}{\xi_1} - (\eta+\mu_1)\right)\xi_1J_1 + \left(\tfrac{\xi_1\beta_1\eta}{\eta+\mu_1} - \nu_1 - \nu_2 - \mu_2\right)A_1 - mxN \\ &= \left(\tfrac{\eta}{\xi_1} - (\eta+\mu_1)\right)\xi_1J_1 + (\xi_1\mu_2\kappa_1 - \nu_1 - \nu_2 - \mu_2)A_1 - mxN. \end{aligned}$$

Because $\kappa_1 > \frac{(\eta+\mu_1)(\nu_1+\nu_2+\mu_2)}{\eta\mu_2}$, $\exists\, \epsilon \in (0,\ 1)$ such that

$$\eta\mu_2\kappa_1 > (\eta + \mu_1 + \epsilon)(\nu_1 + \nu_2 + \mu_2 + \epsilon)$$

.

Let us choose ξ_1 such that,

$$\frac{(\nu_1 + \nu_2 + \mu_2 + \epsilon)}{\mu_2\kappa_1} < \xi_1 < \frac{\eta}{\eta + \mu_1 + \epsilon}$$

.

Thus, $x' > \epsilon\xi_1J_1 + \epsilon A_1 - mxN = \epsilon x - mxN$.

Now, if $N(t) = 0$, then we observe that $x \to \infty$. This contradicts the fact that x is bounded. Thus, $N(t) > 0$.

Next, we consider the positive definite function,

$$V(X) = \sigma_1J_2 + \sigma_2A_2 + \sigma_3A_3,$$
$$X = (J_1, A_1, J_2, A_2, A_3) \in \mathbb{R}^5_+ \sigma_1 = \frac{\eta\nu_2}{a_3a_5} = \frac{\eta T_4}{a_3}$$

with, $\sigma_2 = 1 - \frac{\epsilon\beta_2\eta}{a_3a_5} = 1 - T_2$, and $\sigma_3 = \frac{\nu_2}{a_5} = T_4$.

Because $R_0 < 1$, we derive that $T_2 < 1 \Rightarrow 1 - T_2 > 0$. Thus,

$$\begin{aligned} \dot{V} &= \tfrac{\eta T_4}{a_3}(r\beta_2A_2 + \epsilon\beta_2A_3 - a_3J_2) \\ &\quad + (1 - T_2)\left(\tfrac{\nu_1A_1A_2 + \nu_2A_1A_3}{A} - a_4A_2\right) \\ &\quad + T_4(\eta J_2 + \theta A_2 - a_5A_3) \\ &\leq \left(\tfrac{r\beta_2\eta T_4}{a_3} + \nu_1(1 - T_2) + \theta T_4 - a_4(1 - T_2)\right)A_2 \\ &\quad + \left(\tfrac{\epsilon\beta_2\eta T_4}{a_3} + \nu_2(1 - T_2) - a_5T_4\right)A_3 \\ &= a_4(T_1T_4 + T_3(1 - T_2) + T_4T_5 - (1 - T_2))A_2 \\ &\quad + a_5(T_2T_4 + T_4(1 - T_2) - T_4)A_3 \\ &= a_4(T_1T_4 + T_3 - T_2T_3 + T_4T_5 - 1 + T_2)A_2 \\ &= a_4(T_1T_4 + T_2 + T_3 + T_4T_5 - T_2T_3 - 1)A_2. \end{aligned}$$

Based on our supposition that $R_0 < 1 \Rightarrow T_1T_4 + T_2 + T_3 + T_4T_5 < 1 + T_2T_3$, we have $\dot{V} \leq 0$. This implies that V is a Lyapunov function with selected σ_1, σ_2, and σ_3. Thus, according to the Lyapunov–LaSalle invariance principle,

$$E = \left\{X \in \mathbb{R}^5_+ \mid \dot{V} = 0\right\} = \left\{(J_1, A_{1,}0, 0, 0) \mid J_1 \geq 0, A_1 \geq 0\right\}$$

is an invariant set. Consequently, the largest invariant set is E itself. Therefore, all positive solutions of the system of Equations 1–5 tend to the set E.

Let us consider the limiting system,

$$J_1^{'} = \beta_1A_1 - (\eta + \mu_1)J_1 - (J_1 + A_1)mJ_1 = F_1(J_1, A_1)$$

$$A_1^{'} = \eta J_1 - \mu_2A_1 - mA_1(J_1 + A_1) = F_2(J_1, A_1). \qquad (6)$$

Let $f(J_1, A_1) = \frac{1}{J_1A_1}$. Then, $\frac{\partial(fF_1)}{\partial J_1} + \frac{\partial(fF_2)}{\partial A_1} < 0\ \forall\ (J_1, A_1) \in \mathbb{R}^2_+$. By applying the Bendixson–Dulac criteria, we can say that there are no non-trivial periodic orbits in $\mathbb{R}^2_+$. Solutions of this limited system (6) are bounded, and there is only one non-trivial positive equilibrium point. Because trivial equilibrium is repellent in the system of Equations 1–5, it is also repellent in the system (6). In addition, Theorem 1 shows the local stability of the positive steady state of the system (6). Thus, from these two results, we conclude that the ω-limit set of bounded positive solutions of the system (6) is only (J_1^*, A_1^*). Hence, for the system of Equations 1–5, the disease-free equilibrium $(J_1^*, A_1^*, 0,\ 0,\ 0)$ is a global attractor in $\mathbb{R}^5_+$.

The global stability of the disease-free equilibrium reveals that, irrespective of disparate initial infection levels, the population will ultimately attain a disease-free state, if $R_0 < 1$ and κ_1 exceeds a certain threshold. From a biological standpoint, it can be concluded that the implementation of comprehensive HIV control measures, including extensive testing, treatment, and prevention strategies, holds the potential for effectively eradicating HIV within the population. This global stability result underscores the importance of maintaining strong public health interventions.

Theorem 4 *If* $r = \epsilon = 1$, $R_1 < 1$, *and* $\kappa_2 > \frac{\eta+\mu_1+\alpha}{\eta}$, we establish that *the susceptible extinction equilibrium is a global attractor in* $\mathbb{R}^5_+$.

Proof. First, we claim that $N(t) > 0$ under the aforementioned assumptions. On the contrary, assume that $N(t) = 0$. Let us suppose a function $y = A_3 + \xi_2 J_2$ with $0 < \xi_2 < 1$. The derivative y along with the solution of the system, Equations 1–5, with respect to t will be,

$$\begin{aligned} y' &= (\eta J_2 - a_5 A_3) + \xi_2 (\beta_2 A_3 - a_3 J_2) \\ &= (\eta J_2 - \mu_2 A_3) + \xi_2 (\beta_2 A_3 - (\eta + \mu_1 + \alpha) J_2) - myN \\ &\geq \left(\frac{\eta}{\xi_2} - (\eta + \mu_1 + \alpha)\right) \xi_2 J_2 + \left(\frac{\xi_2 \beta_2 \eta}{\eta + \mu_1 + \alpha} - \mu_2\right) A_3 - myN. \end{aligned}$$

Because $\kappa_2 > \frac{\eta+\mu_1+\alpha}{\eta} = \frac{\mu_2(\eta+\mu_1+\alpha)}{\eta\mu_2}$, $\exists\, \tilde{\epsilon} \in (0, 1)$ such that

$$\eta\mu_2\kappa_2 > (\eta + \mu_1 + \alpha + \tilde{\epsilon})(\mu_2 + \tilde{\epsilon}).$$

Let ξ_2 be chosen such that,
$\frac{(\mu_2+\tilde{\epsilon})}{\mu_2\kappa_2} < \xi_2 < \frac{\eta}{(\eta+\mu_1+\alpha+\tilde{\epsilon})} < 1$.
Then, $y' > \tilde{\epsilon}\xi_2 J_2 + \tilde{\epsilon} A_3 - myN = \tilde{\epsilon} y - myN$. If $N(t) = 0$, then $y \to \infty$, which contradicts the fact that $y = A_3 + \xi_2 J_2 < N$. Therefore, $N(t) > 0$.

Now, we consider the positive definite function $V(X) = \varrho J_1 + A_1 + A_2, X = (J_1, A_1, J_2, A_2, A_3) \in \mathbb{R}^5_+$.

Setting $\varrho = \frac{\eta}{a_1}$, we have

$$\begin{aligned} \dot{V}(X) &= \tfrac{\eta}{a_1}(\beta_1 A_1 - a_1 J_1) + \left(\eta J_1 - \tfrac{\nu_1 A_1 A_2 + \nu_2 A_1 A_3}{A} - a_2 A_1\right) \\ &\quad + \left(\tfrac{\nu_1 A_1 A_2 + \nu_2 A_1 A_3}{A} - a_4 A_2\right) \\ &= \left(\tfrac{\beta_1 \eta}{a_1} - a_2\right) A_1 - a_4 A_2 \\ &= a_2 (R_1 - 1) A_1 - a_4 A_2. \end{aligned}$$

By assumption $R_1 < 1$, $\dot{V} \leq 0$. This implies that $V(X)$ is a Lyapunov function with $\varrho = \frac{\eta}{a_1}$. Now, we apply the Lyapunov–LaSalle invariance principle. This indicates that the positive solutions tend to

$$E = \left\{X \in \mathbb{R}^5_+ \mid \dot{V} = 0\right\} = \left\{(0, 0, J_2, 0, A_3) \mid J_2, A_3 \geq 0\right\}$$

.

Again, by applying the Bendixson–Dulac criteria to the limited J_2A_3-system of Equations 1–5, we observe that the non-trivial equilibrium is $\left(J_2^*, A_3^*\right)$. Thus, the susceptible extinction equilibrium $\left(0,\ 0, J_2^*, 0, A_3^*\right)$ is a global attractor of the system (Equations 1–5) in the domain $\mathbb{R}^5_+$.

The global stability of the susceptible extinction equilibrium implies the eventual elimination of all susceptible individuals, leaving only infected individuals. This potential outcome may arise if control measures are inadequately implemented or if the disease may propagate without restraint.

Theorem 5. *In the scenario of* $r = \epsilon = 1$, $R_1 < 1$, *and* $R_0 < 1$, *we show that all solutions of the system* (Equations 1–5) *tend to* (0, 0, 0, 0, 0) *as time approaches infinity.*

Proof. Let us consider the positive definite function,

$$\begin{aligned} V(X) &= \rho_1 J_1 + A_1 + \rho_2 J_2 + A_2 + A_3, \\ X &= (J_1, A_1, J_2, A_2, A_3) \in \mathbb{R}^5_+. \end{aligned}$$

Setting $\rho_1 = \frac{\eta}{a_1}$ and $\rho_2 = \frac{\eta}{a_3}$, we have

$$\begin{aligned} \dot{V} &= \tfrac{\eta}{a_1}(\beta_1 A_1 - a_1 J_1) + (\eta J_1 - B(t) - a_2 A_1) \\ &\quad + \tfrac{\eta}{a_3}(\beta_2 A_2 + \beta_2 A_3 - a_3 J_2) + (B(t) - a_4 A_2) + (\eta J_2 + \theta A_2 - a_5 A_3) \\ &= \left(\tfrac{\beta_1 \eta}{a_1} - a_2\right) A_1 + \left(\tfrac{\beta_2 \eta}{a_3} + \theta - a_5 - \alpha - \theta\right) A_2 + \left(\tfrac{\beta_2 \eta}{a_3} - a_5\right) A_3 \\ &= a_2 \left(\tfrac{\beta_1 \eta}{a_1 a_2} - 1\right) A_1 + a_5 \left(\tfrac{\beta_2 \eta}{a_3 a_5} - \tfrac{\alpha}{a_5} - 1\right) A_2 + a_5 \left(\tfrac{\beta_2 \eta}{a_3 a_5} - 1\right) A_3 \\ &\leq a_2 (R_1 - 1) A_1 + a_5 (T_2 - 1) A_2 + a_5 (T_2 - 1) A_3. \end{aligned}$$

By our assumption $R_1 < 1$, $R_0 < 1 (\Rightarrow T_2 < 1)$; therefore, $\dot{V}(X) \leq 0$. This shows that V is a Lyapunov function with the selected ρ_1 and ρ_2. Applying the Lyapunov–LaSalle invariance principle,

$$E = \left\{X \in \mathbb{R}^5_+ \mid \dot{V} = 0\right\} = \{(0, 0, 0, 0, 0)\}$$

where E is an invariant set. Thus, using the aforementioned principle, all positive solutions of the system (Equations 1–5) tend to E as $t \to 0$.

It is also critical to see that all threshold values in our model are biologically meaningful. Table 1 summarizes all our major mathematical results for the system of Equations 1–5.

Theorem 1 proves the local stability of the disease-free equilibrium, i.e., if a ratio derived from the basic offspring number κ_1 is >1, then we obtain a unique disease-free equilibrium.

Moreover, if the basic reproduction number R_0 is <1 and κ_1 is greater than the ratio $\frac{\eta\mu_2}{(\eta+\mu_1)(\nu_1+\nu_2+\mu_2)}$, which is clearly >1, i.e., if κ_1 is greater than the inverse of the product of the probability of surviving in J_1 and probability of dying in A_1, then Theorem 3 shows that all infected classes will gradually cease to exist and only susceptible classes will approach a positive constant value.

If there is no recovery under any initial conditions and $R_0 \geq 1$, then Theorem 2 gives a unique susceptible extinction equilibrium. Moreover, if we consider all newborns from infected mothers to be infected, offspring number is <1, and $\kappa_2 > \frac{\eta+\mu_1+\alpha}{\eta}$, then Theorem 4 shows that susceptible classes will eventually vanish over time and infected classes will approach constant values. Here, κ_2 is derived from $T_2 = 1$ and $\kappa_2 > \frac{\eta+\mu_1+\alpha}{\eta}$. This means that κ_2 is greater than the inverse of the probability of survival of J_2.

By Theorem 5, if the basic offspring number is <1 and the sum of all ratios used in horizontal and vertical transmission of an infected individual is <1, then the total population will vanish over time. This means that if growth in susceptible and infected classes is not sufficient, then the whole population will become extinct.

However, as $r = 1$ and $\epsilon = 1$ is not practically feasible, it is not possible to observe a susceptible extinction equilibrium. Meanwhile, if a sufficient growth rate is maintained, the solution to the system will not approach a trivial equilibrium; rather, it will approach a disease-free equilibrium provided that the basic reproduction number is below one. However, owing to several non-linearities in the model, we skip the tedious stability analysis of the endemic equilibrium; from the present analysis, it is straightforward to conclude that the basic reproduction number is the threshold for an outbreak. We implemented this model to explore the recent HIV scenarios in the USA and Pakistan.

4 Model implementation

For the experimental setup, we first describe the parameters and their values in Tables 2, 3. To integrate the system of differential Equations 1–5, some parameters were estimated with the help of data; the remaining ones were estimated using the maximum likelihood method.

4.1 Parameterization for the USA

In the USA (47), the annual live infected births per thousand from 2014 to 2018 were recorded as 12.4, 12.2, 11.8, 11.6, and 11.6, respectively. Calculating the average value of β_1 over this period yields

$$\beta_1 = \frac{11.92}{1000} = 0.01192.$$

According to the World Health Organization, if infected mothers are unaware of their infection, then the transmission rate of HIV infection from mother to newborn is between 15 and 45%. With awareness and treatment, this can be reduced to 5%. Therefore, we considered the values of $r = 0.4$ and $\epsilon = 0.1$ for our model simulation.

To determine the transmission rate (v_2) due to diagnosed infected adults, we used the estimated incidence rate for the USA in 2014, which was 14.3 per 100,000 (7). We also know that the diagnosed adult population of the USA was 926,209 in 2014, and the estimated total number of infected adults in 2010 was 1,085,100. The total population (adults) was 256,498,913; therefore, the susceptible adult population in 2014 was 255,413,813, and we calculated the value of transmission rate due to the diagnosed infected adults as follows:

$$\frac{v_2 A_1\,(2014)\,A_3\,(2014)}{A\,(2014)} = \text{incidence rate in 2014.}$$

$$\text{This implies } v_2 = \frac{14.3 \times (256,\,498,\,913)}{(100,\,000)\,(255,\,413,\,813)\,(926,\,209)} = 1.55 \times 10^{-10}.$$

The average infection period of HIV in developed countries like the USA is reported to be between 8.6 and 19 years (27, 48). It was shown that the average infection period of HIV is $8.6 - 19$ years. Given the excellent health facilities, we considered an average infection period of 19 years, resulting in a mean infection duration (MID) for the USA of $\alpha = \frac{1}{MID} = \frac{1}{19}$.

The UN Inter-agency Group for Child Mortality Estimation (49) provided the probabilities of dying in the first 5 years of life in the USA from 2014 to 2018. The average values (JDR_{0-4}) for these years were calculated as:

$$\text{JDR}_{0-4} = \frac{0.00688 + 0.0068 + 0.00673 + 0.00666 + 0.00658}{5} = 0.00673.$$

Similarly, the estimated probabilities of dying between the ages of 5 and 15 for the same period were used to compute the average (JDR_{5-14}), resulting in:

$$\text{JDR}_{5-14} = \frac{0.0129 + 0.0131 + 0.0133 + 0.0135 + 0.0137}{5} = 0.0133.$$

Consequently, the juvenile natural mortality rate (μ_1) in the USA was determined as

$$\mu_1 = \frac{1}{13}\,(0.00673 + 0.0133) = 0.001545$$

[see (27)]

The UN Inter-agency Group for Child Mortality Estimation (49) provided the probabilities of dying in the first 5 years of life in the USA from 2014 to 2018. The average values (JDR_{0-4}) for these years were calculated as:

$$\text{JDR}_{0-4} = \frac{0.00688 + 0.0068 + 0.00673 + 0.00666 + 0.00658}{5} = 0.00673.$$

According to World Bank data, the average life expectancy (ALE) in the USA from 2014 to 2018 was 78.57 (50). The natural death rate μ is the inverse of ALE; therefore,

$$\text{Natural mortality rate} = \mu = \frac{1}{\text{ALE}} = \frac{1}{78.57},$$

and

$$\text{Natural mortality rate in adults} = \mu_2 = \frac{1}{78.57 - 13} = \frac{1}{65.57}$$

We consider the maximum age of the juvenile population as 13 years because the data were collected from the Centers for Disease Control and Prevention (CDC), which uses this cutoff (7, 51) therefore,

$$\text{Maturation rate of the juvenile population} = \eta = \frac{1}{13}$$

For the estimation of carrying capacity C, we considered the rate of change of the total population with respect to time as:

$$g\,(N) = N' = pN\,(C - N)$$

Here, p is the kinetic parameter. Differentiating g with respect to N, we get

$$\frac{dg}{dN} = p\,(C - 2N)$$

Because the function $g\,(N)$ is bounded and increasing, by the first-order derivative criteria, we can obtain the maximum value of $N_{\max} = C/2$ for this function. Thus,

$$g_{\max} = p\frac{C}{2}\left(C - \frac{C}{2}\right) = p\frac{C^2}{4}$$

and with the help of Table 4, we get $g_{\max} = 2,334,155$, where

TABLE 2 Parameters in our proposed model.

Symbol	Description	Value (USA)	Value (Pakistan)	Unit
β_1	Per capita birth rate of susceptible adults	0.01192	0.029	yr^{-1}
v_2	Transmission rate due to A_3	1.55×10^{-10}	0.0142	yr^{-1}
r	Probability of vertical transmission from A_2	0.40	0.40	D.L
ϵ	Probability of vertical transmission from A_3	0.1	0.1	D.L
α	Disease-induced death rate	1/19	1/13	yr^{-1}
μ_1	Natural juvenile mortality rate	0.001541	0.0109	yr^{-1}
μ_2	Natural adult mortality rate	1/65.57	1/51.7536	yr^{-1}
η	Maturation rate of juvenile population	1/13	1/15	yr^{-1}
K	Carrying capacity of population	645.42	407.39	Million
μ	Natural mortality rate	1/78.57	1/66.7536	yr^{-1}
m	Density-dependent death rate	$\frac{\beta_1-\mu}{K}$	$\frac{\beta_1-\mu}{K}$	yr^{-1}

All parameters were either estimated using data or taken from different references. Here, D.L denotes that the unit is dimensionless.

TABLE 3 Estimated parameters for the data of Pakistan and the USA.

Symbol	Description	Value (CI) for the USA	Value (CI) for Pakistan	Unit
β_2	Per capita birth rate of infected adults	0.0011 (0.000923, 0.00130)	0.0274 (0.02637, 0.02851)	yr^{-1}
v_1	Transmission rate due to A_2	0.3045 (0.30104, 0.30794)	0.2559 (0.25446, 0.25738)	yr^{-1}
θ	Rate of screening and treatment	0.2495 (0.24604, 0.25286)	0.0182 (0.01760, 0.01872)	yr^{-1}

These parameters were estimated using the maximum likelihood method. Here, CI represents the confidence interval.

$g_{\max} = \max \left| difference\ of\ two\ consecutive\ years \right|$.

The average total population of the USA (2014–2018) was $N = 322,710,104$.

$$g_{\max} = pN(C - N)$$
$$\text{Implies } 2,334,155 = p(322,710,104)(C - 322,710,104)$$
$$p = \frac{4,318,344}{(322,710,104)(C - 322,710,104)}.$$
$$\text{Moreover, } g_{\max} = \frac{pC^2}{4}.$$
$$\text{Implies } C^2 = \frac{4g_{max}}{p} = \frac{4(4,318,344)(322,710,104(C - 322,710,104))}{4,318,344}$$
$$= 4(322,710,104(C - 322,710,104)).$$
$$\Rightarrow C^2 - 4(322,710,104)C + 4(322,710,104)^2 = 0$$

This implies that the average value of C is 645.42 million.

Initial values for the simulation were given as follows:

$J_1(0) = Total\ juvenile\ population - J_2(0) = 61,802,095 - (2,477 + 2,477 \times 0.14)$ is the approximate initial susceptible juvenile population in 2014 in the USA (52, 53).

$A_1(0) = A(0) - Total\ adult\ infected\ population = 256,498,913 - 10,851,00$ is the approximate initial susceptible adult population in 2014 in the USA (7, 52, 53).

$J_2(0) = 2,477 + 2,477 \times 0.14$ is the initial HIV-infected juvenile population in 2014 in the USA (51).

$A_2(0) = 1,085,100 - 926,209$ is the initial HIV-infected adult population not in treatment in 2014 in the USA (7, 51).

$A_3(0) = 926,209$ is the initial HIV-infected adult population in treatment in 2014 in the USA (51).

4.2 Parameterization for Pakistan

In the study of the World Bank (47), the numbers of live-infected births each year (per thousand) from 2014 to 2018 in Pakistan were 29.318, 29.124, 28.888, 28.599, and 28.25, respectively. We consider the same values of r and ϵ for Pakistan as we did for the USA.

From UNAIDS (25), we estimated that the average number of newly infected adults was 18,000 per year from 2014 to 2018. The average percentage of people in treatment was 10% (6) and the average total infected during this time interval per year in Pakistan was 126,800 (54). We assume the same percentage of newly aware infected adults in treatment. Therefore, we use the formula given by López et al. (27) as follows:

$$v_2 = \frac{\text{Average total new infected adults(per year)} \times 10\%}{\text{Average total infected population}}$$
$$= \frac{(18,000)(0.1)}{126,800} \approx 0.0142$$

Once an individual is infected with HIV, the average infection period of HIV is 8.6–19 years (27, 48). Because Pakistan is a developing country, we consider the average life expectancy after infection of 13 years as the MID for our model. Thus,

$$\alpha = \frac{1}{\text{MID}} = \frac{1}{13}$$

According to the UN Inter-agency Group for Child Mortality Estimation, the probabilities of dying in the first 5 years of life

from 2014 to 2018 were 0.0783, 0.07605, 0.07381, 0.07158, and 0.0694, respectively.

Therefore, the average value (JDR_{0-4}) is:

$$\begin{aligned}\text{JDR}_{0-4} &= \frac{0.0783 + 0.07605 + 0.07381 + 0.07158 + 0.0694}{5} \\ &= 0.073828.\end{aligned}$$

In addition, the probabilities of dying between the ages of 5 and 15 from 2014 to 2018 were 0.0939, 0.0918, 0.0897, 0.0875, and 0.0856, respectively (55).

Therefore, the average value (JDR_{5-14}) is:

$$\begin{aligned}\text{JDR}_{5-14} &= \frac{0.0939 + 0.0918 + 0.0897 + 0.0875 + 0.0856}{5} \\ &= 0.0897.\end{aligned}$$

Thus, the juvenile natural mortality rate is:

$$\mu_1 = \frac{1}{15}(0.073828 + 0.0897) = 0.0109.$$

The average life expectancy in Pakistan from 2011 to 2018 was 66.7536 (50). The natural death rate μ is the inverse of this; therefore,

$$\text{Natural mortality rate} = \mu = \frac{1}{66.7536},$$

and

$$\text{Natural mortality rate of adults} = \mu_2 = \frac{1}{66.7536 - 15} = \frac{1}{51.7536}.$$

Considering the maximum age of the juvenile population as 15 years (27, 48),

$$\text{Maturation rate of the juvenile population} = \eta = \frac{1}{15}.$$

For carrying capacity C, we used the same estimation approach as for the US data.

As Table 4 shows, $g_{\max} = 43, 183, 44$. From the same table, the average total population of Pakistan (2014-2018) was $N = 203, 694, 557.8$. Therefore,

$$\begin{aligned}C^2 &= \frac{4(4,318,344)((203,694,557.8)(C-203,694,557.8))}{4,318,344} \\ &= 4\left((203,694,557.8)\left(C - 203,694,557.8\right)\right) \\ \Rightarrow C^2 &- 4\left(203,694,557.8\right)C + 4(203,694,557.8)^2 = 0.\end{aligned}$$

This implies that the average value of C is 407.39 million.

Initial values for the simulation were given as follows:

$J_1(0) = \textit{Total juvenile population} - J_2(0) = 70,797,567 - 3,500$ is the approximate initial susceptible juvenile population in 2014 in Pakistan (25, 53).

$A_1(0) = A(0) - A_2(0) - A_3(0) = 124,509,258 - 90,500$ is the approximate initial susceptible adult population in 2014 in Pakistan (25, 52–54).

$J_2(0) = 3,500$ is the initial HIV-infected juvenile population in 2014 in Pakistan (25).

$A_2(0) = 90,500 - 6,292$ is the initial HIV-infected adult population not in treatment in 2014 in Pakistan (54).

$A_3(0) = 6,292$ is the initial HIV-infected adult population in treatment in 2014 in Pakistan (54).

4.3 Maximum likelihood fitting

In this subsection, we explain the maximum likelihood method and estimation of the remaining parameters. Because β_2 the birth rate and ν_1 the transmission rate due to infected adults, A_2 and the status of infection are not always known, estimating them directly from data of newborns and the newly infected is not convenient. Instead, we estimate the parameters β_2, ν_1, and θ using the maximum likelihood method. We have time-series data of juvenile infected, HIV-positive adults in treatment and infected population not in treatment (total infected – infected in treatment; see Tables 5, 6).

Let $X = (J_1, A_1, J_2, A_2, A_3) \in \mathbb{R}_+^5$ be the vector of state variables and $\mathbb{F}$ be the right side of the system (Equations 1–5). $\mathbb{P} = (\beta_2, \nu, \theta)$ is the vector of parameters to be estimated. $\mathbb{Y}(t, \mathbb{P}) = (A_3(t, \mathbb{P}), J_2(t, \mathbb{P}), A_2(t, \mathbb{P}))$ is the vector of observables. $\mathbb{Y}^0(t, \mathbb{P})$ represent the observed data at $t = 2014, \cdots, 2018$. We assume that all $\mathbb{Y}^0(t, \mathbb{P})$ are independent and Poisson distributed with mean $\mathbb{Y}(t, \mathbb{P})$. Thus, the Poisson maximum likelihood function will be:

$$\text{L}\left(\mathbb{Y}^0(t, \mathbb{P}) \mid \mathbb{Y}(t, \mathbb{P})\right) = \prod_{j=1}^{3}\prod_{i=1}^{5} \frac{y_j(t_i)^{y_{ji}^0} \cdot e^{-y_j(t_i)}}{y_{ji}^0!}$$

.

Therefore, the negative log-likelihood function (NLF) reduces to

$$\begin{aligned}\text{NLF} &= -\ln\left(\prod_{j=1}^{3}\prod_{i=1}^{5} \frac{y_j(t_i)^{y_{ji}^0} . e^{-y_j(t_i)}}{y_{ji}^0!}\right) \\ &= -\sum_{j=1}^{3}\sum_{i=1}^{5} y_{ji}^0 \ln\left(y_j(t_i)\right) + \sum_{j=1}^{3}\sum_{i=1}^{5} y_j(t_i) + \sum_{j=1}^{3}\sum_{i=1}^{5} \ln\left(y_{ji}^0!\right).\end{aligned}$$

Because the last term in the aforementioned equation is constant, it will remain unchanged as the parametric values are varied. Therefore, we can minimize only the first two terms of the equation. Hence, the fitting problem can be expressed as

$$\min(\text{NLF}) = \min\left(-\sum_{j=1}^{3}\sum_{i=1}^{5} y_{ji}^0 \ln\left(y_j(t_i)\right) + \sum_{j=1}^{3}\sum_{i=1}^{5} y_j(t_i)\right)$$

subject to

$$\begin{aligned}&\frac{d}{dt}X(t, \mathbb{P}) = F(X, \mathbb{P}, t) \\ &\mathbb{Y}(t, \mathbb{P}) = (A_3(t, \mathbb{P}), J_2(t, \mathbb{P}), A_2(t, \mathbb{P})) \\ &X(0) = (J_1(2014), A_1(2014), J_2(2014), A_2(2014), A_3(2014)) \\ &X(t, \mathbb{P}), \mathbb{P} \geq 0.\end{aligned} \tag{7}$$

TABLE 4 Total populations of the USA and Pakistan from 2014 to 2018.

Year	2014	2015	2016	2017	2018
Total population of USA	318,301,008	320,635,163	322,941,311	324,985,539	326,687,501
Total population of Pakistan	195,306,825	199,426,964	203,627,284	207,896,686	212,215,030

These time-series data for the US and Pakistan populations were taken from the World Bank (52).

TABLE 5 HIV-diagnosed adult and juvenile populations and total adult infected populations in the USA from 2014 to 2018.

Year	2014	2015	2016	2017	2018
HIV-positive in treatment	926,209	951,346	976,097	1,000,062	1,023,832
Juvenile diagnosed HIV+	2,477	2,344	2,226	2,072	1,912
Total adult HIV+	1,085,100	1,108,400	1,131,100	1,152,500	1,173,900

Because the USA is a developed country, we assume that all diagnosed infected in the USA are in treatment. This time-series data for HIV-diagnosed adult population are from the CDC (51). Because 14% of HIV-infected population do not know their status of infection (7), we add this population to the diagnosed juvenile population and assume that this is the total infected juvenile population in the USA. The time-series data for HIV-diagnosed juvenile population are from the CDC (51). The time-series data for total estimated adult HIV-infected population are from the CDC (7).

The aforementioned minimization problem will provide the desired fitting with practically feasible values of parameters if $\mathbb{P}$ is identifiable, which can be confirmed in two steps. First, we examine the structural identifiability, and then we confirm the practical identifiability.

Structural identifiability refers to the existence of a unique solution to $X(t,\mathbb{P})$ for each $\mathbb{P}$ under an initial condition. If any component of $\mathbb{P}$ is implicitly related, different values of $\mathbb{P}$ may yield the same solution $X(t,\mathbb{P})$ for a given initial condition, which might hinder the unique estimation of parameters $\mathbb{P}$ for the data.

We use the Fisher information matrix (FIM) to confirm the structural identifiability. We have a set of observations at five distinct points, a system represented by a five-dimensional state vector, and a three-dimensional vector of parameters. Thus, the sensitivity matrix S consists of five time-dependent 5×3 blocks $\mathbb{A}(t_n)$:

$$S = \begin{bmatrix} \mathbb{A}(t_1) \\ \mathbb{A}(t_2) \\ \vdots \\ \mathbb{A}(t_5) \end{bmatrix}$$

where

$$\mathbb{A}_{jk}(t_n) = \frac{\partial x_j(t_n,\mathbb{P})}{\partial P_k}, \quad n = 1,\cdots,5,\ k = 1,2,3 \text{ and } j = 1,\cdots,5.$$

The 3×3 FIM is $M = S^T S$. The rank of the matrix M counts the number of identifiable parameter conditions in $\mathbb{P}$, and the parameters in $\mathbb{P}$ are structurally identifiable if and only if M has rank 3.

According to the aforementioned definition, the FIM for our problem has three columns, which correspond to the parameters to be estimated. Let us denote the parameter estimates by $\hat{\beta}_2, \hat{\nu}_1$, and $\hat{\theta}$. To approximate the FIM numerically, we perturb $\hat{\beta}_2$ to the new values $\hat{\beta}_2^+ = (1+0.001)\hat{\beta}_2$ *and* $\hat{\beta}_2^- = (1-0.001)\hat{\beta}_2$, for which we integrate the model for each observation time. Then, we numerically approximate the derivatives, $\mathbb{A}_{j1}(t_n) = \frac{\partial x_j\left(t_n,\hat{\beta}_2,\hat{\nu}_1,\hat{\theta}\right)}{\partial \hat{\beta}_2}, n = 1,\cdots,5, j = 1,\cdots,5$. This gives the first column. Meanwhile, the other two variables $\hat{\nu}_1$ and $\hat{\theta}_2$ remain fixed. We repeat the same procedure for $\hat{\nu}_1$ and $\hat{\theta}_2$ to obtain the second and third columns, respectively. Then, we check the rank of matrix M, which is 3. This ensures the structural identifiability of the parameters.

Practical identifiability or "estimableness" refers to the sufficiency of available observations, as too few observations might be insufficient for fitting. To investigate the practical identifiability, we compute the profile likelihood of the parameters β_2, ν_1, and θ_2. Profile likelihood reveals the dependency of the NLF on individual parameters, which helps us to find finite confidence intervals for each parameter; otherwise, practical non-identifiability is proved. The related profile likelihoods can be defined as

$$PL_{\beta_2}(\beta_2) = \min_{\nu_1,\theta} \{\text{NLF}(\beta_2,\nu_1,\theta)\}$$

, $PL_{\nu_1}(\nu_1) = \min_{\beta_2,\theta} \{\text{NLF}(\beta_2,\nu_1,\theta)\}$, and $PL_{\theta}(\theta) = \min_{\beta_2,\nu_1} \{\text{NLF}(\beta_2,\nu_1,\theta)\}$.

where $\beta_2 \in \left[\widehat{\beta}_2(1-0.001), \widehat{\beta}_2(1+0.001)\right]$, $\nu_1 \in [\widehat{\nu}_1(1-0.001), \widehat{\nu}_1(1+0.001)]$, and

$$\theta \in \left[\widehat{\theta}(1-0.001), \widehat{\theta}(1+0.001)\right]$$

To determine the confidence interval, we have $2\left(\text{NLF}(\mathbb{P}) - \text{NLF}\left(\widehat{\mathbb{P}}\right)\right) \tilde{} \chi_3^2$. Therefore, the NLF threshold for a 95% confidence interval is $\text{NLF}\left(\widehat{\mathbb{P}}\right) + 7.815/2$.

The maximum likelihood fitting of the time-series data for the USA and Pakistan is shown in Figure 2; the corresponding estimates of the parameters are shown in Table 3 along with the 95% confidence intervals. The fittings for both the USA and Pakistan appear relatively good at a glance. Moreover, the estimated parameter values are biologically meaningful. Because the birth rate of an infected population is less than that of a susceptible adult, we observe that $\beta_1 > \hat{\beta}_2$ in

TABLE 6 HIV-positive (aware and in treatment) adult, juvenile infected, and total infected population in Pakistan from 2014 to 2018.

Year	2014	2015	2016	2017	2018
HIV-positive in treatment	6,292	6,500	8,900	12,046	15,821
Juvenile HIV+	3,500	4,000	4,500	5,000	5,500
Total HIV+	94,000	100,000	130,000	150,000	160,000

This time-series data for HIV-diagnosed adult population are from UNAIDS and the AIDS Data Hub (54). The time-series data for HIV juvenile population are from UNAIDS estimation (2), and the time-series data for estimated total HIV-infected population are from UNAIDS and the AIDS Data Hub (54).

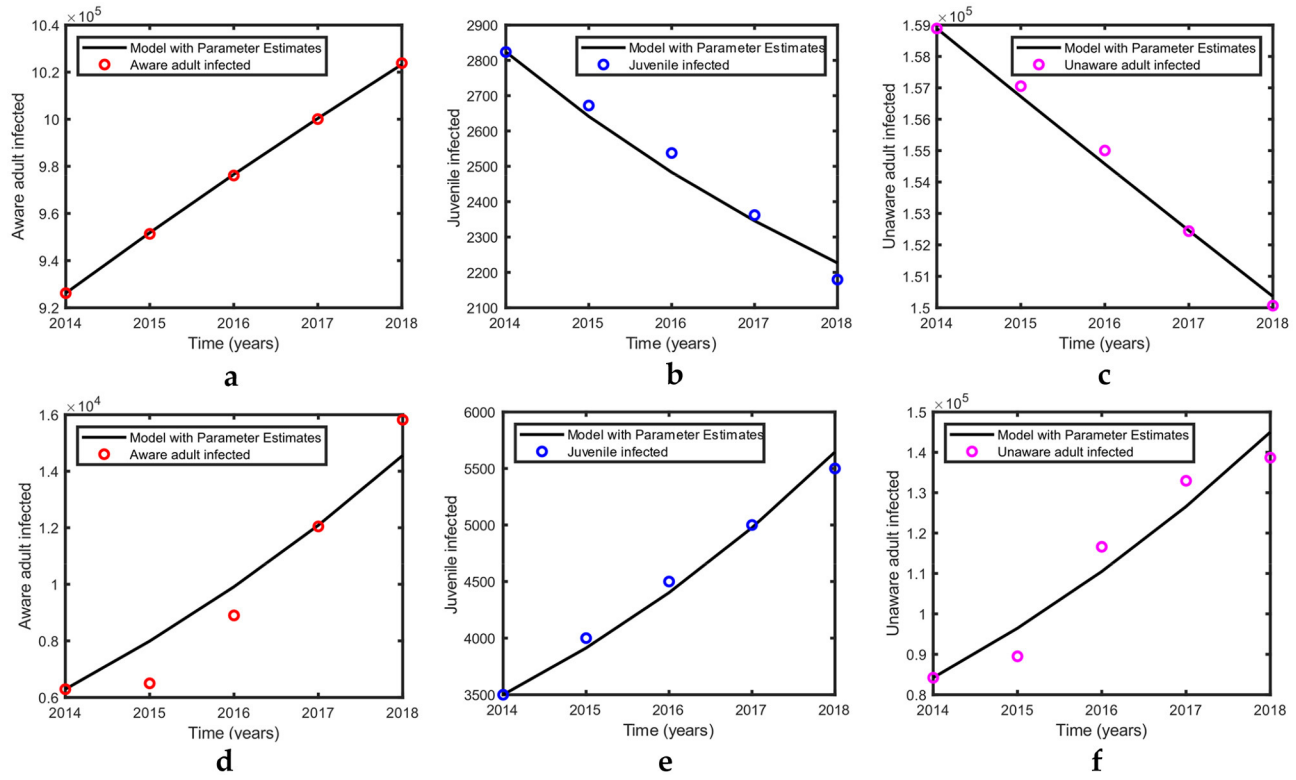

FIGURE 2
Simulation results for US and Pakistan data. Because the USA is a developed country, we assume that all diagnosed infected in the USA are in treatment. The parametric values for the USA in Tables 2, 3 were used in this simulation, whereas the initial values are defined in Section 4.1. **(a–c)** show the curves fitted with US data. Parametric values for Pakistan in Tables 2, 3 were used in this simulation, whereas the initial values are defined in Section 4.2. **(d–f)** show the curves fitted with Pakistan data.

the estimation. In addition, $\hat{\nu_1} > \nu_2$ due to awareness and treatment. When a person is aware of their infection, the transmission rate may be reduced because of treatment and precautions.

Profile likelihoods for the USA and Pakistan are shown in Figure 3, revealing the minimization of the NLF at the estimated values of the parameters while also confirming the practical identifiability. The solid black lines indicate the cost as a function of β_2, ν_1, and θ. For each cost function, we change the value of one parameter using the minimization algorithm; the others vary on the parallel axis over the interval shown. This pattern of the cost function verifies the identifiability criteria of our estimated parameters. The red cross on each curve shows the best-fitted value for the parameters β_2, ν_1, and θ.

4.4 Sensitivity analysis

Sensitivity indices were used to compute how minor variations in the parameters of interest cause variability in the quantities of interest (56). With the help of the normalized sensitivity index (NSI) of R_0 regarding the parameters, we can determine which parameters have a greater impact on disease transmissibility. The NSI of the basic reproduction number R_0 to the parameter ϕ is defined as

$$\frac{\partial R_0}{\partial \phi} \times \frac{\phi}{R_0},$$

where R_0 is the quantity of interest and ϕ is our parameter of interest. Table 7 lists the NSIs

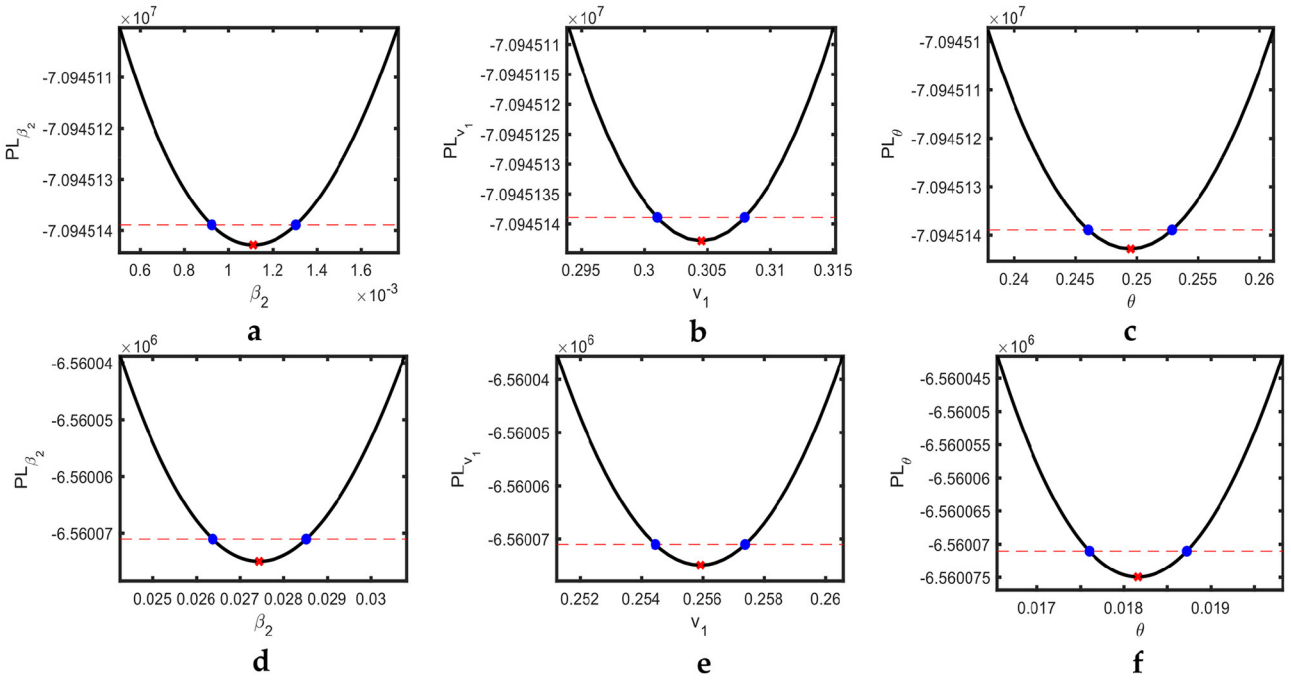

FIGURE 3
Estimated parameters versus profile likelihood function. **(a–c)** Show graphs of profile likelihood (cost) functions and estimated parameters β_2, v_1, and θ for the USA, respectively, and **(d–f)** show the same for Pakistan. Red crosses indicate the estimated values of the parameters, whereas blue dots show the confidence intervals.

TABLE 7 NSIs of basic reproduction number.

Parameters	NSI (USA)	NSI (Pakistan)
β_1	-0.0333	-0.1907
r	1.1485×10^{-11}	0.0046
ϵ	5.9286×10^{-11}	0.0010
μ_1	6.8130×10^{-4}	0.0250
η	6.8328×10^{-4}	-0.0321
β_2	7.0771×10^{-11}	0.0057
v_1	1	0.9483
v_2	1.0511×10^{-8}	0.0461
μ_2	-0.0067	-0.0443
θ	-0.7938	-0.1110
α	-0.1675	-0.6469
m	0	0

This table lists the sensitivity indices of the basic reproduction number R_0 to the parameters used in the mathematical model (Equations 1–5) at the parametric values in Tables 2, 3. Here, NSI stands for local sensitivity index.

of R_0 with respect to the parameters for our current estimate.

According to our analysis, the parameters with comparatively higher sensitivity are β_1, v_1, θ, and α. Negative signs of local sensitivity indices indicate that if we increase the value of these parameters of interest, then our quantity of interest will decrease, and vice versa. In short, R_0 reduces with an increase in awareness and treatment rate θ.

5 Results

According to our estimation, the values of R_0 for the USA and Pakistan are 0.9688 and 2.2599, respectively, which reveals that according to our analysis, HIV is approaching eradication in the USA. This is in agreement with the recent trend of decreasing incidence in the USA (Figure 2). Meanwhile, in Pakistan, the recent trend of increasing incidence (Figure 2) is reflected in the high estimate of the reproduction number.

We have two infectious classes, A_2 and A_3, which each transmit the disease vertically and horizontally, constituting four transmission pathways. We split the contributions of these pathways to R_0 in Figure 4.

This shows that horizontal transmission from A_2 plays a vital role in both countries. In the case of the USA, $T_3 \approx R_0$, which means the incidences in the USA are almost entirely due to adults being unaware of infection, whereas transmission due to other pathways is negligible. However, in the case of Pakistan, although horizontal transmission from A_2 plays a major role, transmissions through the other three pathways also play a noticeable role. Transmissions through these three pathways are several times greater than in the USA. Moreover, the value of v_2 is higher in Pakistan than in the USA. That is, individuals who know their status of infection and maintain treatment in the USA are more vigilant than those in the same class in Pakistan. On the contrary, T_3 is higher and v_1 is lower in Pakistan than in the USA, which

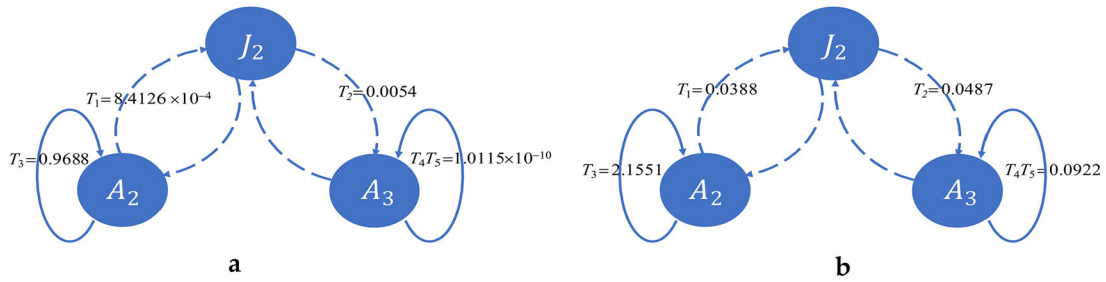

FIGURE 4
Contributions of horizontal and vertical transmissions to R_0. Solid lines indicate horizontal transmission, and dotted lines indicate vertical transmission. Horizontal transmission from A_2 plays a major role in both cases. **(a)** Shows transmission in the USA, whereas **(b)** shows transmission in Pakistan.

indicates that individuals who are unaware of their infection status in Pakistan transmit less than those in the USA.

R_0 has a negative sensitivity to the parameter θ (Table 7), which is related to awareness and treatment. This is also portrayed in Figure 5 for the estimated values of the parameters for both the USA and Pakistan. Here, the red crosses show the estimated values of θ, and the blue lines show the dependency of R_0 on θ. According to Figure 5, R_0 is below one if θ is above a threshold, which is 0.2396 and 0.4298 in the USA and Pakistan, respectively.

In the case of the USA, we observe that the estimated value of θ is 0.2495, which means that the disease is diagnosed within 4 years ($\frac{1}{0.2495} \approx 4$) of infection. This is above the threshold and has recently resulted in a declining prevalence in the USA. At present, R_0 is very close to 1 in the USA, which implies that the disease will be eradicated, but it will take a long time. In the case of Pakistan, the estimated value of θ is below the threshold (Figure 5), and as a result, the prevalence has an increasing trend (Figure 2).

If the unaware infected population learn their status of infection and they are in treatment within ~2 years ($\frac{1}{\theta} \leq \frac{1}{0.4302}$) after infection, then the disease will gradually be eradicated.

To clarify the role of θ in disease prevalence, we simulated our model for different values of θ while keeping other parameters fixed within a time interval of 10 years, and the results are shown in Figure 6. For the US case, we show the simulation results for the estimated value of θ and $\theta = 0.5$ (these values and corresponding R_0 are shown in Figure 5 by a cross and dot, respectively). For the estimated value of θ, the number of individuals in the J_2 and A_2 classes decrease monotonically, whereas it increases for A_3. For $\theta = 0.5$, the number of individuals in all infected classes decreases monotonically except for A_3.

A_3 shows an increasing trend up to ~5 years due to increasing awareness and seeking treatment, and then it starts to decrease. In the case of Pakistan, we show the simulation results for the estimated value of θ, where $\theta = 0.25, 0.6, 0.9$ (these values and corresponding R_0 are shown in Figure 5 by a cross and dots, respectively). For the estimated value of θ, the number of individuals in all infected classes increases monotonically. If $\theta = 0.6$, a decreasing trend in all the infected classes could be achieved after ~9 years. A decreasing trend could be achieved in ~5 years if θ is as high as 0.9.

If we increase the coverage of treatment, the pace of disease eradication will also increase. Therefore, screening and treatment can help control the spread of the disease in Pakistan and accelerate eradication in the USA.

6 Discussion

The HIV epidemic presents a distinct contrast between the USA and Pakistan. While the USA has made significant strides in controlling the spread of HIV, with continuous efforts in treatment and screening. Pakistan has seen a concerning in HIV in recent years. However, because of consistent improvement in healthcare infrastructure and availability of antiretroviral therapy (ART), infected people can lead an almost normal life with careful medication and a punctilious lifestyle (39). This shift in treatment efficacy addresses the concern raised decades ago that extending the lifespan of HIV patients could exacerbate the epidemic by increasing the basic reproduction number (R_0) (27). To understand the role of treatment with improved medication in today's world, we investigated the recent HIV trends in the USA and Pakistan using a mathematical epidemic model.

HIV treatment brings the infected person closer to a normal life from several perspectives. For instance, it (1) reduces the fatality of the infection and increases lifespan, (2) reduces the probability of transmission to partners, and (3) reduces the probability of vertical transmission to children. To consider all these aspects, we assumed a two-age group model. The mathematical analysis of our model reveals that the basic reproduction number is the key threshold for an outbreak. The model is defined by several parameters and estimating all of them by fitting with time-series data of infected individuals leads to identifiability issues. Hence, we determined the values of most parameters from available literature and estimated the key parameters, birth rate of infected adults, transmission rate for infected not in treatment, and rate of screening and treatment, by fitting with time-series data.

The central finding of our research emphasizes the significance of the basic reproduction number (R_0) as a key determinant in assessing the severity of an outbreak. The model parameters, which include birth rates, transmission rates, and treatment initiation rates, are crucial in predicting disease progression. Identifying all parameters from data is fraught with challenges, primarily stemming from issues of identifiability. To address this issue, we employed a combination of literature-based

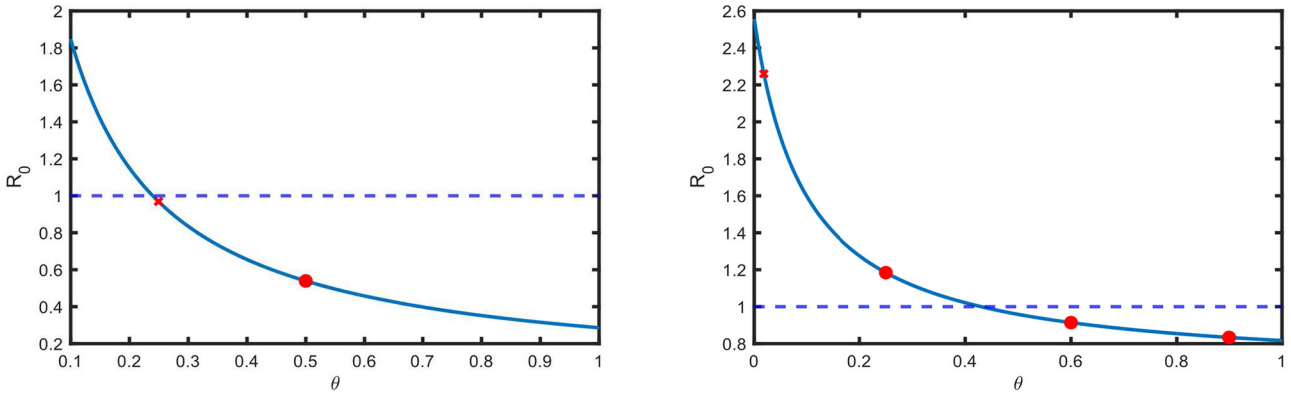

FIGURE 5
Basic reproduction number versus screening rate in the USA and Pakistan. The red crosses indicate the estimated value of the screening and treatment rate θ. Other red circles indicate that if we increase θ, then R_0 decreases.

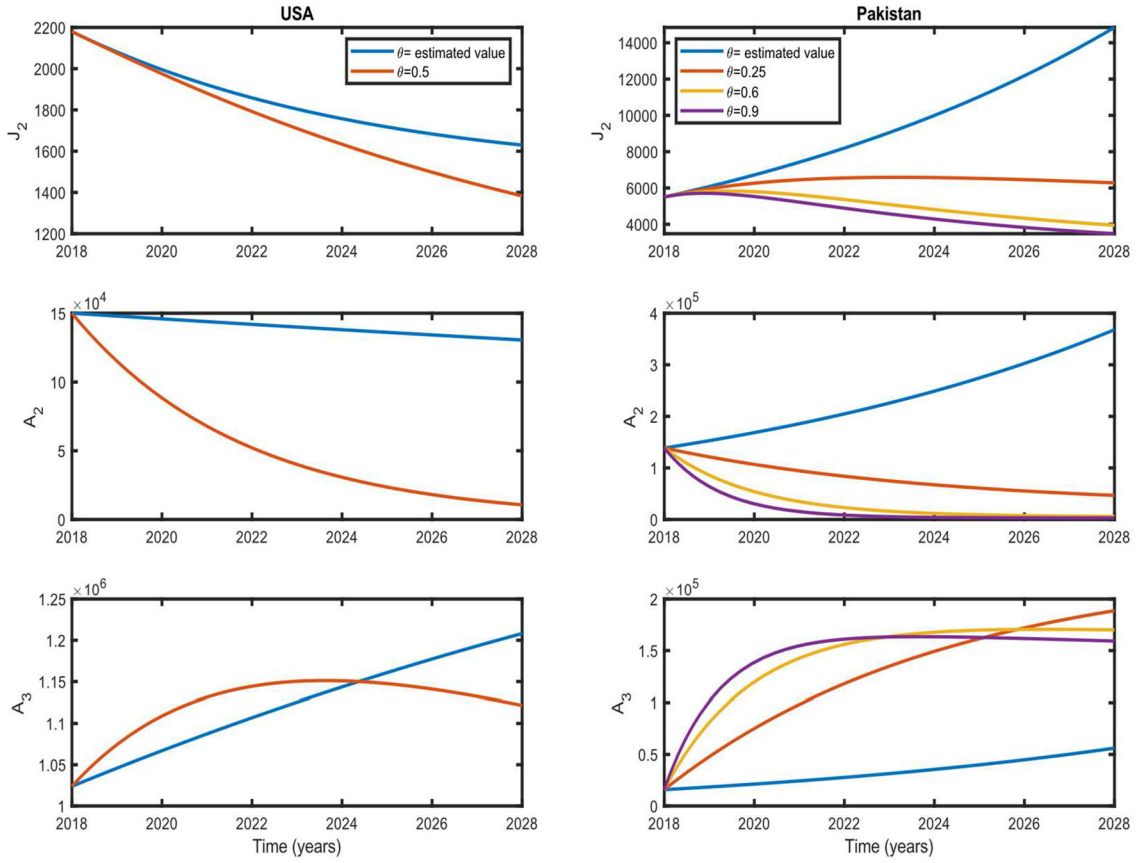

FIGURE 6
Simulation results with different values of treatment rate. These results show the situation of the infected population for the next 10 years with different values of θ. Initial values are US and Pakistan data from 2018 (see Tables 4–6).

parameters and estimation methods to determine crucial variables, such as the infection rate of adults, transmission rates among untreated individuals, and the rates of screening and treatment initiation. Based on our findings, the USA should focus on expanding treatment coverage, while Pakistan should emphasize the promotion of preventive strategies and raising awareness.

First, we ensured the *structural identifiability* of these key parameters with respect to the model set by checking the rank of the FIM. Next, using the profile likelihood of the estimates, we confirmed the *practical identifiability* of the parameters for the time-series data of the infected. According to our estimates, the birth rate of infected individuals is lower than that of

susceptible individuals. In addition, our estimation shows that treatment reduces the probability of both vertical and horizontal transmission, which indicates the role of treatment in HIV infection in both countries.

Furthermore, the reproduction number (R_0) shows high sensitivity to transmission rate from infected individuals not in treatment and the rate of undergoing treatment following screening. This highlights the critical importance of *screening and awareness programs*. As infected individuals not in treatment are mostly not aware of their infection, screening is a potential measure to impede the incidence. This is also evident from the transmission pathway-wise splitting of the reproduction number; transmission from infected individuals not in treatment is the dominant factor in both the USA and Pakistan. However, infected individuals in treatment seem more mindful in the USA than in Pakistan. Moreover, the screening rate is better in the USA than in Pakistan.

Given these findings, we emphasize that increased screening and awareness campaigns could play an important role in controlling the HIV epidemic in Pakistan. Although the USA has made substantial progress, improving the rate of screening further could still yield significant reductions in HIV incidence. Our model suggests that by identifying infected individuals and putting them in treatment within ∼2 years in the USA and 1.1 years in Pakistan, it is possible to achieve a declining trend within ∼5 years in all infected classes.

7 Conclusion

In this article, we proposed a two-age group model to differentiate between the transmissibility of a juvenile and two adult infected classes. We fitted it with time-series data of infected individuals from the USA and Pakistan. According to our estimation, the basic reproduction numbers for the USA and Pakistan are 0.9688 and 2.2599, respectively. In addition, the estimated values of the parameters show that the birth rate of infected individuals is lower than that of susceptible individuals, and the transmission rate due to the infected not undergoing treatment is higher than that of infected individuals in treatment, which demonstrates the consistency of our approach. The basic reproduction number is most sensitive to the transmission rate from infected individuals not in treatment and the rate at which these individuals are screened and given treatment. Furthermore, by splitting the roles of the different transmission pathways, it became apparent that the infected group not in treatment contributes the most to transmission in both countries. Therefore, screening and treatment will reduce the most transmissible group and work in a bi-directional manner to reduce the incidence rate, which we also confirmed through appropriate simulations of our model.

A comparison of the parameter values for the USA and Pakistan showed that the infected individuals in treatment in the USA are more inclined to follow the guidelines or take better care than those in Pakistan. On the contrary, infected individuals not in treatment transmit the disease in the USA more than those in Pakistan. Although the USA already has a healthy rate of screening and treatment, improvement in the rate is necessary for prompt eradication of the disease. However, besides screening and treatment, mass education and awareness are crucial in Pakistan.

Our study inevitably had some limitations because of certain assumptions. Different stages of HIV infection have different degrees of transmissibility. Therefore, modeling stage-dependent infectiousness may help identify the most effective target group for screening. Furthermore, considering a two-sex model would facilitate understanding and reduce vertical transmission, as the probability of mother-to-child transmission is much higher than that of father-to-child transmission. Future research considering the aforementioned factors could be useful for further enhancement of control strategies.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/ participants or patients'/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

WA: Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing, Data curation, Formal analysis. MM: Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing, Data curation, Supervision. SP: Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing. HL: Methodology, Writing – original draft, Writing – review & editing, Formal analysis. SK: Conceptualization, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIP) (2022R1A5A1033624, 2021R1A2B5B03087097) and Global—Learning and Academic research institution for Master's·PhD students, and Postdocs (LAMP) Program of the National Research

Foundation of Korea (NRF) grant funded by the Ministry of Education (No. RS-2023-00301938).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Centers for Disease Control and Prevention. *HIV basics*. Available at: https://www.cdc.gov/hiv/basics/whatishiv.html (accessed September 27, 2020).

2. UNAIDS. *HIV estimates with uncertainty bounds 1990-2019*. (2020). Available at: https://www.unaids.org/en/resources/documents/2020/HIV_estimates_with_uncertainty_bounds_1990-present

3. Kh A. *Strategic Framework for Prevention of Parent to Child Transmission (PPTCT) of HIV in Pakistan*. (2017). Available at: https://www.aidsdatahub.org/resource/strategic-framework-prevention-parent-child-transmission-hiv-pakistan (accessed January 17, 2020).

4. Khanani RM, Hafeez A, Rab SM, Rasheed S. Human immunodeficiency virus-associated disorders in Pakistan. *AIDS Res Hum Retroviruses*. (1988) 4:149–54. doi: 10.1089/aid.1988.4.149

5. Abdul MS, Hashmi M. A study of HIV-antibody in sera of blood donors and people at risk. *J Pak Med Assoc*. (1988) 38:221.

6. UNAIDS. *Pakistan Country Data*. (2019). Available at: https://www.aidsdatahub.org/resource/pakistan-country-data (accessed September 15, 2020).

7. Centers for Disease Control and Prevention. *Estimated HIV incidence and prevalence in the United States, 2014–2018*. (2020). Available at: https://www.cdc.gov/hiv/library/reports/hiv-surveillance.html

8. Anderson R, Medley GF, May RM, Johnson AM. A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *IMA J Math Appl Med Biol*. (1986) 3:229–63. doi: 10.1093/imammb/3.4.229

9. Anderson RM. The role of mathematical models in the study of HIV transmission and the epidemiology of AIDS. *J Acquir Immune Defic Syndr*. (1988) 1:241–56.

10. May RM, Anderson RM. Transmission dynamics of HIV infection. *Nature*. (1987) 326:137–42. doi: 10.1038/326137a0

11. Massad E. Ahomogeneously mixing population model for the AIDS epidemic. *Math Comput Model*. (1989) 12:89–96. doi: 10.1016/0895-7177(89)90448-2

12. Hethcote HW, Van den Driessche P. Some epidemiological models with nonlinear incidence. *J Math Biol*. (1991) 29:271–87. doi: 10.1007/BF00160539

13. Busenberg S, Cooke K, Hsieh Y-H. A model for HIV in Asia. *Mathe Biosci*. (1995) 128:85–210. doi: 10.1016/0025-5564(94)00072-8

14. Doyle M, Greenhalgh D. Asymmetry and multiple endemic equilibria in a model for HIV transmission in a heterosexual population. *Math Comput Model*. (1999) 29:43–61. doi: 10.1016/S0895-7177(99)00029-1

15. Hyman JM, Li J, Stanley EA. The differential infectivity and staged progression models for the transmission of HIV. *Math Biosci*. (1999) 155:77–109. doi: 10.1016/S0025-5564(98)10057-3

16. Hsieh Y-H, Sheu S-P. The effect of density-dependent treatment and behavior change on the dynamics of HIV transmission. *J Math Biol*. (2001) 43:69–80. doi: 10.1007/s002850100087

17. Greenhalgh D, Doyle M, Lewis F. A mathematical treatment of AIDS and condom use. *IMA J Math Appl Med Biol*. (2001) 18:225–62. doi: 10.1093/imammb18.3.225

18. Moghadas SM, Gumel AB. Global stability of a two-stage epidemic model with generalized non-linear incidence. *Math Comput Simul*. (2002) 60:107–18. doi: 10.1016/S0378-4754(02)00002-2

19. Manfredi P, Salinelli E. Population-induced oscillations in blended SI–SEI epidemiological models. *Math Med Biol*. (2002) 19:95–112. doi: 10.1093/imammb/19.2.95

20. Gielen J. A framework for epidemic models. *J Biol Syst*. (2003) 11:377–405. doi: 10.1142/S0218339003000919

21. Hsieh Y-H, Chen CH. Modelling the social dynamics of a sex industry: its implications for spread of HIV/AIDS. *Bull Math Biol*. (2004) 66:143–66. doi: 10.1016/j.bulm.2003.08.004

22. Naresh R, Omar S, Tripathi A. Modelling and analysis of HIV/AIDS in a variable size population. *Far East J Appl Math*. (2005) 18:345–60.

23. De Arazoza H, Lounes R. A non-linear model for a sexually transmitted disease with contact tracing. *Math Med Biol*. (2002) 19:221–34. doi: 10.1093/imammb19.3.221

24. UNAIDS. *UNAIDS Data 2019*. (2019). Available at: https://www.unaids.org/en/resources/documents/2019/2019-UNAIDS-data (accessed March 15, 2020).

25. UNAIDS. *HIV estimates with uncertainty bounds 1990-2019*. (2020). Available at: https://www.unaids.org/en/resources/documents/2019/HIV_estimates_with_uncertainty_bounds_1990-present (accessed November 12, 2020).

26. Naresh R, Tripathi A, Omar S. Modelling the spread of AIDS epidemic with vertical transmission. *Appl Math Comput*. (2006) 178:262–72. doi: 10.1016/j.amc.2005.11.041

27. López R, Kuang Y, Tridane A. A simple SI model with two age groups and its application to US HIV epidemics: to treat or not to treat? *J Biol Syst*. (2007) 15:169–84. doi: 10.1142/S021833900700212X

28. Aldila D, Aprilliani RR, Malik M. Understanding HIV spread with vertical transmission through mathematical model. In: *AIP Conference Proceedings*. Melville, NY: AIP Publishing LLC (2018). doi: 10.1063/1.5054546

29. Wang JJ, Reilly KH, Han H, Peng ZH, Wang N. Dynamic characteristic analysis of HIV mother to child transmission in China. *Biomed Environ Sci*. (2010) 23:402–8. doi: 10.1016/S0895-3988(10)60082-7

30. Kaur N, Ghosh M, Bhatia S. Modeling the spread of HIV in a stage structured population: effect of awareness. *Int J Biomath*. (2012) 5:1250040. doi: 10.1142/S1793524511001829

31. Kaymakamzade B. Sanlidag T, Hinçal E, Sayan M, Tijjani Sa'ad F, Baba IA. Role of awareness in controlling HIV/AIDS: a mathematical model. *Qual Quant*. (2018) 52:625–37. doi: 10.1007/s11135-017-0640-2

32. Tripathi A, Naresh R, Sharma D. Modeling the effect of screening of unaware infectives on the spread of HIV infection. *Appl Math Comput*. (2007) 184:1053–68. doi: 10.1016/j.amc.2006.07.007

33. Okosun K, Makinde O, Takaidza I. Impact of optimal control on the treatment of HIV/AIDS and screening of unaware infectives. *Appl Math Model*. (2013) 37:3802–20. doi: 10.1016/j.apm.2012.08.004

34. Naresh R, Tripathi A, Sharma D. Modelling and analysis of the spread of AIDS epidemic with immigration of HIV infectives. *Math Comput Model*. (2009) 49:880–92. doi: 10.1016/j.mcm.2008.09.013

35. Naresh R, Tripathi A, Sharma D. A nonlinear AIDS epidemic model with screening and time delay. *Appl Math Comput*. (2011) 217:4416–26. doi: 10.1016/j.amc.2010.10.036

36. Huo H-F, Chen R, Wang X-Y. Modelling and stability of HIV/AIDS epidemic model with treatment. *Appl Math Modell*. (2016) 40:6550–9. doi: 10.1016/j.apm.2016.01.054

37. Olaniyi S, Kareem GG, Abimbade SF, Chuma FM, Sangoniyi SO. Mathematical modelling and analysis of autonomous HIV/AIDS dynamics with vertical transmission and nonlinear treatment. *Iran J Sci*. (2024) 48:181–92. doi: 10.1007/s40995-023-01565-w

38. Alhassan C, Aondoakaa M, Amobeda E. *Vertical Transmission And The Dynamics Of HIV/AIDS In A Growing Population*. Researchjournali's Journal of Mathematics. (2017) 4.

39. Khan SU, Ullah S, Li S, Mostafa AM, Bilal Riaz M, AlQahtani NF, et al. A novel simulation-based analysis of a stochastic HIV model with the

time delay using high order spectral collocation technique. *Sci Rep*. (2024) 14:7961. doi: 10.1038/s41598-024-57073-3

40. Teklu SW, Mekonnen TT. HIV/AIDS-pneumonia coinfection model with treatment at each infection stage: mathematical analysis and numerical simulation. *J Appl Math*. (2021) 2021:5444605. doi: 10.1155/2021/5444605

41. Teklu SW, Rao KP. HIV/AIDS-pneumonia codynamics model analysis with vaccination and treatment. *Comput Math Methods Med*. (2022) 2022:3105734. doi: 10.1155/2022/3105734

42. Raza A, Ahmadian A, Rafiq M, Salahshour S, Naveed M, Ferrara M, et al. Modeling the effect of delay strategy on transmission dynamics of HIV/AIDS disease. *Adv Differ Equat*. (2020) 2020:663. doi: 10.1186/s13662-020-03116-8

43. Raza A, Rafiq M, Baleanu D, Shoaib Arif M, Naveed M, Ashraf K, et al. Competitive numerical analysis for stochastic HIV/AIDS epidemic model in a two-sex population. *IET Syst Biol*. (2019) 13:305–15. doi: 10.1049/iet-syb.2019.0051

44. Efficacy of three short-course regimens of zidovudine and lamivudine in preventing early and late transmission of HIV-1 from mother to child in Tanzania, South Africa, and Uganda (Petra study): a randomised, double-blind, placebo-controlled trial. *Lancet*. (2002) 359:1178–86. doi: 10.1016/s0140-6736(02)08214-4

45. Qu S, Wang Q, Wang X, Qiao Y, Yao J, Li Z, et al. Recommend guideline on prevention of mother-to-child transmission of HIV in China in 2020. *Infect Dis Immunity*. (2023) 3:52–9. doi: 10.1097/ID9.0000000000000083

46. Van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math Biosci*. (2002) 180:29–48. doi: 10.1016/S0025-5564(02)00108-6

47. World Bank. *Birth rate, crude (per 1,000 people) - Pakistan, United States*. (2010–2018). Available at: https://data.worldbank.org/indicator/SP.DYN.CBRT.IN?end=2018andlocations=PK-USandstart=2010 (accessed March 15, 2020).

48. López-Cruz R. *Structured SI Epidemic Models with Applications to HIV Epidemic*. Tempe, AZ: Arizona State University (2006).

49. UN Inter-agency Group for Child Mortality Estimation. *Childe mortality rate, crude (per 1,000 people) - USA*. (2014–2018). Available at: https://childmortality.org/data/United%20States%20of%20America (accessed September 16, 2020).

50. World Bank. *Life expectancy at birth, total (years) - Pakistan, United States*. (2014–2018). Available at: https://data.worldbank.org/indicator/SP.DYN.LE00.IN?end=2018andlocations=PK-USandstart=2010 (accesed 2020).

51. Centers for Disease Control and Prevention. *HIV Surveillance Report, 2018 (Updated)*. (2020). Available at: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html

52. The World Bank. *Population all ages, total - Pakistan, United States*. (2014–2018). Available at: https://data.worldbank.org/indicator/SP.POP.TOTL?end=2018andlocations=PK-USandstart=2014 (accessed January 26, 2020).

53. The World Bank. *Population ages 0-14, total - Pakistan, United States*. (2014–2018). Available at: https://data.worldbank.org/indicator/SP.POP.0014.TO?end=2018andlocations=PK-USandstart=2014 (accessed January 26, 2020).

54. UNAIDS. *Country snapshots and fact sheets*. (2014–2018). Available at: https://www.aidsdatahub.org/search/content?keywords=Pakistan (accessed September 12, 2020).

55. UN Inter-agency Group for Child Mortality Estimation. *Childe mortality rate, crude (per 1,000 people) - Pakistan*. (2014–2018) (accessed September 17, 2020).

56. Arriola LM, Hyman JM. Being sensitive to uncertainty. *Comput Sci Eng*. (2007) 9:10–20. doi: 10.1109/MCSE.2007.27