**Methods**

**Data generation**

To simulate the temporal dynamics of HIV transmission, we implemented a compartmental susceptible-infected (SI) model with four stages of infection (I1 to I4), reflecting progressive disease stages. The population is stratified into compartments: susceptible individuals (S), four sequential infected stages (I1–I4), cumulative incidence (CI), and cumulative deaths (CD).

**Table 1**. Initial conditions

|  |  |  |
| --- | --- | --- |
| Parameters | Value/expression | Label |
| initial\_prevalence | Exp (-9.5) | Initial prevalence of the disease |
| time\_sequence | Seq (1976, 2025, by = 1) | Time sequence for the simulation |
| initial\_conditions | * S = 1 - initial\_prevalence * I1 = initial\_prevalence * I2 = I3 = I4 = CI = CD = 0 |  |
| infection\_states | paste0('I', 1:4) | Infection states for the model |

The model incorporates a declining transmission rate as a function of HIV prevalence, represented by an exponential function of the form:

**Table 2**. Global disease parameters

|  |  |  |
| --- | --- | --- |
| Parameters | Value | Label |
| Beta ( | 0.6 | transmission coefficient when prevalence is 0 |
| Alpha ) | 3.5 | for transmission coefficient: decline with prevalence |
| Lambda | λ(t) | Transmission rate adjusted by prevalence |
| progRt | 1/15 | progression rate through each of the Infection compartment (I1-I4) |
| birthRt | 0.03 | birth rate, 3% of people give birth per year |
| deathRt | 1/60 | 60 years natural life expectancy |
| cMax | 0.7 | maximum control effect, 70% reduction in transmission |
| cRate | 0.5 | rate of control effect increase, 0.5 years to reach cMax |
| cStart | 1998 | year when control effect started |

The intervention effect was modeled using an exponential decay function starting in 1998, with maximum impact (cMax= 0.7) and rate of increase (cRate = 0.5). The model was solved over the period 1976–2025 using the `lsoda` solver in R. The force of infection is modulated by a time-varying control function C(t), reflecting the implementation of interventions (e.g., ART, prevention programs). The control function was modeled as:

*Control\_effect [C(t)] = pmin (1, cMax+(1-cMax)\*exp(-(time-cStart)\*cRate))*

From the simulated epidemic curve, we generated two sampling scenarios:

**Table 3**. Sampling scenario to generate datasets

|  |  |  |  |
| --- | --- | --- | --- |
| Scenario | Datasets generated | Sample size | Polling step |
| 1 | 20 | 250 (uniform) | Each 2 years from 1985 to 2024 |
| 2 | 20 | 200 to 50 (varing) | Each 5 years from 1985 to 2024 |

In each case, observed HIV prevalence was simulated by binomial sampling from the model-based prevalence. Exact 95% confidence intervals were computed using the Clopper–Pearson method. All simulations and visualizations were conducted in R (version X.X), using the `deSolve`, `dplyr`, and `ggplot2` packages.