Theoretical Biology and Medical Modelling

REVIEW

Open Access

Modeling methods for estimating HIV incidence: a mathematical review

Xiaodan Sun[1], Hiroshi Nishiura[2*] and Yanni Xiao[1]

Abstract

Estimating HIV incidence is crucial for monitoring the epidemiology of this infection, planning screening and intervention campaigns, and evaluating the effectiveness of control measures. However, owing to the long and variable period from HIV infection to the development of AIDS and the introduction of highly active antiretroviral therapy, accurate incidence estimation remains a major challenge. Numerous estimation methods have been proposed in epidemiological modeling studies, and here we review commonly-used methods for estimation of HIV incidence. We review the essential data required for estimation along with the advantages and disadvantages, mathematical structures and likelihood derivations of these methods. The methods include the classical back-calculation method, the method based on CD4+ T-cell depletion, the use of HIV case reporting data, the use of cohort study data, the use of serial or cross-sectional prevalence data, and biomarker approach. By outlining the mechanistic features of each method, we provide guidance for planning incidence estimation efforts, which may depend on national or regional factors as well as the availability of epidemiological or laboratory datasets.

Keywords: statistical estimation, HIV/AIDS, CD4, Biomarker, Mathematical model

Background

Since the first patient with acquired immunodeficiency syndrome (AIDS) was reported in 1981 [1], its causative agent, human immunodeficiency virus (HIV), has led to 77 million HIV infections globally and remains a major public health issue [2]. To strategically assess the impact of interventions and to guide policy makers in achieving improved control of HIV/AIDS, it is critical to quantify the dynamics of HIV epidemics accurately and reliably. HIV incidence (i.e., the transient number of new infections) and prevalence (i.e., the fraction of infected individuals at a given point in time) are two major indicators that are used to assess and interpret the transmission dynamics of HIV. HIV incidence and prevalence have been estimated using mathematical and statistical modeling approaches by many academic and governmental research groups. For instance, the Joint United Nations Program on HIV/AIDS (UNAIDS) regularly provides updates of national and global estimates, indicating that 1.8 million people were newly infected with HIV and 940,000 deaths occurred in the year 2017 [2].

Unlike many acute infectious diseases, HIV infection progresses slowly in vivo and has a complex natural history. During the first 2-4 weeks following infection, the virus replicates rapidly and this period is referred to as the acute stage [3, 4]. Thereafter, viral loads are greatly reduced and reach a quasi-steady state, which is called the asymptomatic stage. During the asymptomatic stage, the viral load reflects the steady state achieved between high rates of viral replication and virus clearance, and is maintained at a remarkably stable level (i.e., the viral load set point) over a number of years. If untreated, the median length of asymptomatic stage may range from 8-11 years. Infected individuals in the asymptomatic stage do not show overt symptoms but can transmit HIV infection through high-risk behaviors. Subsequently, the viral load increases slowly, resulting in the onset of AIDS [5–7]. Because their immune systems are severely damaged, individuals with AIDS experience a number of opportunistic infections and are at high risk of death without treatment.

Owing to the lengthy asymptomatic stage without symptoms, many individuals do not realize that they are

*Correspondence: nishiurah@med.hokudai.ac.jp
[2] Graduate School of Medicine, Hokkaido University, Kita 15 Jo Nishi 7 Chome, Kitaku, 0608638 Sapporo, Japan
Full list of author information is available at the end of the article

infected for a number of years. Moreover, through sexual contact and intravenous drug use, infections often remain undetected due to the reliance on voluntary testing following those high risk exposures[8, 9]. This issue both leads to increased HIV transmission and complicates modeling exercises, increasing the difficulty of explicitly quantifying the epidemiological dynamics of HIV/AIDS. Furthermore, owing to the widespread use of antiretroviral therapy (ART), prevalence estimation is controversial: even where prevalence can be estimated, this estimate may not reflect the current dynamics of HIV epidemics and may reflect only the degree of spread from many years in the past [10]. It is generally recognized that estimation of HIV incidence can provide greater insights into the real-time evaluation of HIV epidemics. Nevertheless, the long asymptomatic stage also causes challenges in estimating HIV incidence.

Starting in the 1980s, a large number of modeling studies have aimed to estimate HIV incidence, and a variety of useful methods have been proposed for this purpose. These diverse methods have played important roles in HIV incidence estimation in different parts of the world. However, only brief comparative notes have been published elsewhere [10–12], aiming for improvement in practical estimation settings. In this review, we comprehensively describe the major methods that have been used for HIV incidence estimation, including (i) the classical back-calculation method, (ii) the method based on CD4+ T-cell depletion, (iii) the use of HIV case reporting data, (iv) the use of cohort study data, (v) the use of serial or cross-sectional prevalence data, and (vi) biomarker approach. We focus on the structural mechanisms of modeling as well as the mathematical derivation of likelihood functions, and compare the advantages and disadvantages of existing methods. Our review is targeted to a general audience in theoretical biology. Finally, we summarize important implications for future development of estimation methods for HIV incidence.

Back-calculation

Back-calculation, one of the most widely-used statistical modeling approaches, exploits the distribution of incubation periods of AIDS. The back-calculation method uses epidemiological surveillance data to reconstruct HIV infections over time. The basic idea of the method can be described as follows. If the rate of incident HIV infections at time s is $I(s)$, and the probability density function of the incubation period $f(s)$ is known, then AIDS incidence at time t, denoted by $A(t)$, can be described by

$$A(t) = \int_0^t I(t-\tau)f(\tau)d\tau. \tag{1}$$

Conversely, if the dataset for $A(t)$ is available from surveillance data and $f(s)$ can be determined from the literature, HIV incidence can be "back-calculated" by rearranging (1) to

$$A(t) = \int_0^t I(s)f(t-s)ds. \tag{2}$$

If $F(t)$ denotes the cumulative distribution function of the incubation period, one can describe the expected number of AIDS diagnoses over the time interval $[\,T_{i-1}, T_i]$, denoted by X_i, as

$$\begin{aligned} E(X_i) &= \textstyle\int_{T_{i-1}}^{T_i} \int_0^t I(s)f(t-s)dsdt \\ &= \textstyle\int_0^{T_{i-1}} \int_{T_{i-1}}^{T_i} I(s)f(t-s)dtds \\ &+ \textstyle\int_{T_{i-1}}^{T_i} \int_s^{T_i} I(s)f(t-s)dtds \\ &= \textstyle\int_0^{T_{i-1}} I(s) \int_{T_{i-1}}^{T_i} f(t-s)dtds \\ &+ \textstyle\int_{T_{i-1}}^{T_i} I(s) \int_s^{T_i} f(t-s)dtds \\ &= \textstyle\int_0^{T_{i-1}} I(s)[\,F(T_i-s) - F(T_{i-1}-s)]\,ds \\ &\quad + \textstyle\int_{T_{i-1}}^{T_i} I(s)[\,F(T_i-s) - F(0)]\,ds \\ &= \textstyle\int_0^{T_i} I(s)[\,F(T_i-s) - F(T_{i-1}-s)]\,ds. \end{aligned} \tag{3}$$

Here, the last equality holds because $F(T_{i-1}-s) = F(0) = 0$ for $T_{i-1} - s \leq 0$. Then, we have

$$E(X_i) = \int_0^{T_i} I(s)[\,F(T_i-s) - F(T_{i-1}-s)]\,ds. \tag{4}$$

The classic method using AIDS data

The back-calculation method was first proposed by Brookmeyer et al. [13–15] who used AIDS incidence data to estimate discrete HIV incidence using the maximum likelihood estimation method. Let $T_0, T_1, \cdots, T_L$ denote discrete times, N denote the total number of infections before T_L, and X_i denote the number of diagnosed AIDS cases in the ith time interval $[\,T_{i-1}, T_i]$. Then, N is the sum of all infected cases that have been diagnosed, $X. = \sum_{i=1}^{L} X_i$, and those infected before T_L but have not been diagnosed are indicated by $X_{L+1} = N - X.$. Suppose that $X = (X_1, X_2, \cdots, X_L, X_{L+1})$ follows a multinomial distribution with sample size N, where probabilities $(p_1, p_2, \cdots, p_L, 1 - p_.)$ can be calculated according to Eq. (3), and $p_. = \sum_{j=1}^{L} p_j$. In fact, $p_i = \int_{T_0}^{T_i} i(s)[\,F(T_i - s) - F(T_{i-1} - s)]\,ds$, where $i(s)$ is the probability density function for these N individuals being infected at time s. Denoting the observed AIDS incidence in each time interval as $x_1, x_2, \cdots, x_L$, the likelihood function can be described as follows:

$$\frac{N!}{x_1!\,x_2! \cdots x_L!\left(N - \sum_{i=1}^{L} x_i\right)!} p_1^{x_1} p_2^{x_2} \cdots p_L^{x_L} (1-p_.)^{N-\sum_{i=1}^{L} x_i}.$$

The back-calculation method can estimate the historical incidence of infection that was already diagnosed and also the number of infections that have yet to be diagnosed.

Becker et al. [16] proposed a non-parametric approach to this method using the discrete version of Eq. (2). Let the number of HIV infections in the ith time interval be I_i

and the probability mass function of the incubation period be f_d. Then, the expected number of AIDS diagnoses in interval i can be described as

$$E(X_i|I_1, I_2, \cdots, I_i) = \sum_{j=1}^{i} I_j f_{i-j}. \tag{5}$$

Let $\mu_i = E(X_i)$ and $\lambda_j = E(I_j)$. Then,

$$\mu_i = \sum_{j=1}^{i} \lambda_j f_{i-j}.$$

Assuming that HIV infections are generated by a non-homogeneous Poisson process, X_i $(i = 1, \cdots, L)$ would follow Poisson distributions with means μ_i. Then, the likelihood function is

$$\prod_{i=1}^{L} \left(\sum_{j=1}^{i} \lambda_j f_{i-j} \right)^{x_i} \exp\left(- \sum_{j=1}^{i} \lambda_j f_{i-j} \right), \tag{6}$$

where x_i is the observed frequency of AIDS cases.

It should be noted that this method assumes that the distribution of the incubation period does not vary over time. In fact, it is easy to modify Eq. (2) to $A(t) = \int_0^t I(s) f(t-s \mid s) ds$, where $f(t - s \mid s)$ is the probability density function for an individual who was infected at time s and diagnosed at time t. Thus, $f(t - s \mid s)$ describes the time-dependent distribution of incubation period. Similarly, the discrete version of Eq. (5) becomes $E(X_i|I_1, I_2, \cdots, I_i) = \sum_{j=1}^{i} I_j f_{i-j,j}$ with $f_{i-j,j}$ representing the probability for an individual infected during time interval $[t_{j-1}, t_j)$ and diagnosed during time interval $[t_{i-1}, t_i)$.

In this way, the mathematical expression of the back-calculation method is straightforward, but the estimation of I_i using this method is challenging [17] because of the high dimension of I_i which leads to instability. To estimate the incidence of HIV I_i, several published studies [18, 19] have used either $A(t)$ or $I(t)$ as flexible parametric functions. Rosenberg et al. [20] estimated the infection curve $I(s)$ directly, assuming that $I(s)$ is a member of the general family $G = \{g_1(s), \cdots, g_I(s)\}$, where $g_i(s)$ are integrable real functions. That is,

$$I(s) = \Sigma_{i=1}^{I} g_i(s) \beta_i.$$

Specifically, this method includes splines and step functions.

It should be noted that for models involving spline and step functions, another weakness is the potential of overfitting and the ill-posed inverse problem. Overfitting arises when too many knots in the spline are applied. The ill-posed problem arises when the step function is too discrete and when the estimated HIV incidence becomes overly sensitive to temporal fluctuations of data points. Moreover, when the HIV epidemic has just started and the trend has not been stable, the back-calculated incidence in the most recent years would be more uncertain than that based on long-lasting epidemic dynamics. This is caused by small number of diagnosed infections in recent observed times, yielding substantial uncertainties. However, in many existing settings in developed countries, the HIV epidemic has continued for substantial number of years, and in such an occasion, the uncertainties in the estimated recent infections are not as large as that estimated in the early epidemic phase with dramatic peaks and troughs, as shown by Yan and Zhang [21]. In addition, this method has been criticized because the estimation strongly depends on the distribution of the incubation period, which needs to be determined from other cohorts. Estimation of the incubation period encountered critical challenges in the 1990s, as the introduction of ART extended the length of the incubation period, inevitably changing this distribution. To account for the effect of ART, extended methods were proposed [22–26].

Using both HIV and AIDS diagnoses

In addition to AIDS incidence, the frequency of diagnosed HIV infections has become available as part of epidemiological surveillance data, greatly assisting researchers to extend the back-calculation method [27–40]. Early studies used only HIV diagnoses of individuals who later progressed to AIDS [27–32]. Subsequently, several other methods were proposed to incorporate all HIV diagnoses, including infected individuals who have not yet developed AIDS [33–40]. Yan et al. [41] proposed an approach which uses the number of new HIV diagnoses to back-calculate historical HIV incidence, partially aided by supplementary data from the old AIDS case surveillance system in populations where there were such system in the 1980s. The estimate is also calibrated with supplementary data based on "recent infections", that is, the proportion among newly diagnosed HIV that is recently infected according to enhanced surveillance or laboratory assays. This method was used to estimate HIV incidence among men who have sex with men in Australia [42, 43]. Adding information on HIV diagnoses to the back-calculation method enables estimation of HIV incidence in recent years and reduces the uncertainty associated with this estimate to some degree. Moreover, the method enables joint estimation of HIV diagnosis rate [44]. However, challenges associated with estimating or assuming a time from infection to HIV diagnosis remain.

Including CD4+ T-cell counts at diagnosis

Upon HIV diagnosis, CD4+ T-cell count data has now become widely available. Various studies have defined HIV/AIDS progression based on CD4+ T-cell counts and employed Markov process models [33, 38] to estimate HIV incidence. Birrell et al. [45] formulated a CD4-stage structured model to use CD4+ T-cell counts at diagnosis.

CD4+ T-cell count data represented the first CD4 count recorded within three months of HIV diagnosis. The model included a total of five stages: CD4+ T-cell counts of $[500, \infty)$, $[350, 500)$, $[250, 300)$, $[0, 200)$, and the AIDS stage. Infected individuals are assumed to experience progressive decline in CD4+ T-cell counts and proceed through the five stages before they are diagnosed with AIDS. Let $\mathbf{d_j} = (d_{1,j}, d_{2,j}, d_{3,j}, d_{4,j})$, and $d_{k,j}, k = 1, 2, 3, 4$ denote the probability of diagnosis in the *k*th CD4 count stage during time interval *j*. Let $\mathbf{e_j} = (e_{1,j}, e_{2,j}, e_{3,j}, e_{4,j})$, where $e_{k,j}, k = 1, 2, 3, 4$ denotes the expected number of undiagnosed infections in the *k*th CD4 count stage during time interval *j*. Suppose that the expected number of diagnosed HIV and AIDS cases during time interval *j* is $\mu_j^{HIV} and \mu_j^{AIDS}$, respectively, then

$$\begin{aligned} \mu_j^{HIV} &= \mathbf{e_{j-1}} \cdot \mathbf{d_j}^T, and \\ \mu_j^{AIDS} &= e_{4,j}(1 - d_{4,j})\rho_{4,5} \end{aligned} \tag{7}$$

where

$$\mathbf{e_j} = \mathbf{P_j^T} \mathbf{e_{j-1}} + (\lambda_j, 0, 0, 0)^T.$$

λ_j is the expected number of new HIV infections in time interval *j*, and $\mathbf{P_j}$ is the transition matrix describing the proportion of individuals transitting between different stages during time interval *j*. Then,

$$(\mathbf{P_j})_{k,l} = \begin{cases} (1 - d_{k,j})(1 - \rho_{k,k+1}) & k = l, \\ (1 - d_{k,j})\rho_{k,k+1} & k = l - 1, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

where $\rho_{k,k+1}$ is the transition probability from stage *k* to $k + 1$. Let X_j and $Y_j (j = 1, \cdots, L)$ denote AIDS and HIV diagnoses during the time interval *j*, respectively, which are assumed to follow independent Poisson distributions with means μ_j^{HIV} and μ_j^{AIDS}, respectively. Then, the likelihood function for HIV and AIDS diagnoses can be calculated as

$$L_1(\mathbf{X}, \mathbf{Y}; \mathbf{h}, \mathbf{d}) \propto \prod_{j=1}^{L} \left(\mu_j^{AIDS}\right)^{X_j} \exp\left(-\mu_j^{AIDS}\right) \times \left(\mu_j^{HIV}\right)^{Y_j} \exp\left(-\mu_j^{HIV}\right).$$

CD4+ T-cell count data at diagnosis is also available for a subset of the above-diagnosed HIV-positive individuals. The CD4+ T-cell count data at diagnosis are divided into four sets: $[500, \infty)$, $[350, 500)$, $[250, 300)$, and $[0, 200)$. Let $\mathbf{C_j} = (C_{1,j}, C_{2,j}, C_{3,j}, C_{4,j})$ or $C_{k,j} (k = 1, 2, 3, 4)$ be the number of HIV-positive individuals whose CD4 counts fall into the *k*th CD4 stage during the *j*th time interval, and $N_j = \Sigma_{k=1}^{4} C_{k,j}$. That is, N_j individuals are diagnosed with HIV during the time interval *j* with the state variable, CD4-at-diagnosis data. We assume that these N_j HIV-positive individuals with CD4 data are multinomially distributed as

$$\mathbf{C_j} \sim Multinomial(N_j, \mathbf{r_j}),$$

where

$$\mathbf{r_j} = \{r_{k,j} : k = 1, 2, 3, 4\}, r_{k,j} = \frac{e_{k,j-1} d_{k,j}}{\mu_j^{HIV}}, j = 1, 2, \cdots, L.$$

Then, the likelihood of observing CD4-at-diagnosis data can be given as

$$L_2(\mathbf{C}|\mathbf{D}; \mathbf{h}, \mathbf{d}) \propto \prod_{j=1}^{L} \prod_{k=1}^{4} r_{k,j}^{C_{k,j}}.$$

The full likelihood is the product of L_1 and L_2:

$$L(\mathbf{X}, \mathbf{Y}, \mathbf{C}; \mathbf{h}, \mathbf{d}) = L_1(\mathbf{X}, \mathbf{Y}; \mathbf{h}, \mathbf{d}) L_2(\mathbf{C}|\mathbf{D}; \mathbf{h}, \mathbf{d}).$$

This method can make full use of all the available data, including HIV and AIDS diagnoses as well as CD4+ T-cell counts at diagnosis. Using this method, one cannot only estimate the incidence of HIV infections but also the diagnosis rates at different CD4 stages and during different time intervals, providing insightful information to comprehensively evaluate the epidemiology of HIV/AIDS. Using this model, Birrell et al. estimated HIV incidence in England and Wales [46], and found that the mean time to diagnosis had shortened from 2001 to 2010 owing to expansion of HIV testing. However, this method is also highly dependent on the progression rate between different stages. Moreover, the quantities requiring estimation have much higher dimensions yielding additional difficulties. Birrell et al. [45] employed the Bayesian estimation technique, ensuring the stability of estimates.

Using CD4+ T-cell count data at diagnosis based on a CD4+ T-cell depletion model

In addition to the back-calculation method, another major HIV incidence estimation method is to jointly use HIV diagnosis data and the first CD4 count data while employing the CD4+ T-cell depletion model [47–49]. This method first estimates the distribution of diagnosis delays (i.e., the time from infection to diagnosis), and then estimates the incidence of HIV from the depletion of CD4+ T-cells [49]. Here, HIV incidence refers to the number of new infections during each time interval, including both diagnosed and undiagnosed infections by the end of the study period. The CD4+ T-cell depletion model that was adopted by Lodi et al. and Touloumi et al. [50, 51] can be expressed as

$$\sqrt{CD4(t)} = a_i + (b_i \times t) + e_{it},$$

where *t* denotes the time from infection to the date of the first CD4+ T-cell count determination. Then, the time from date of infection to CD4 testing for an individual *i* can be estimated by

$$T_i = \frac{\sqrt{\text{first}CD4} - a_i}{b_i}.$$

a_i and b_i are assumed to follow a bivariate normal distribution $N[\,(a,b),(\sigma_a,\sigma_b),\rho]$, and are variable from person to person. Using standard survival analysis techniques, the diagnosis delay probability $P(x)$ was estimated, which is the probability that an infected person would be diagnosed within x time units after infection. To statistically estimate undiagnosed infections, the authors further defined the diagnosis delay weight as $W(x) = 1/P(x)$.

Let t_0 and t_N be the start and end times of the study period. The estimated infection time for each diagnosed individual may be either before or after t_0. Suppose the estimated number of infections in the ith year after t_0 is n_i (using the CD4+ T-cell depletion model), where $i = 1, 2, \cdots, N$, and the time of infection for each case is DI_j, $j = 1, 2, \cdots, n_i$. Then, the number of new infections in the ith year after t_0 can be estimated as

$$\lambda_i = \sum_{j=1}^{n_i} W(t_N - DI_j).$$

A certain number of individuals remain to be infected but are not diagnosed before t_0. Let U denote the number of such individuals. These individuals may either be diagnosed between t_0 and t_N, or not diagnosed until the end of the study period t_N. Let $u_i, i = 1, 2, \cdots$, be the number of newly diagnosed cases among these persons in the ith year after t_0. Then, $U = \sum_{i\geq 1} u_i$ is the total number of diagnoses observed during the study period. In addition, u_i for $i > N$ are cases who are not diagnosed until the end of the study period. H_i is further defined as the total number of cases diagnosed during the ith year after t_0 (including those infected before and after t_0), where $i = 1, \cdots, N$. Thus, $r_i = u_i/H_i$ is the proportion of new diagnosed cases in the ith year after t_0 who are infected before t_0. Both H_i and r_i are treated as linear regression functions of time t, so H_i and r_i for $i > N$ can then be predicted, and u_i for $i > N$ can at last be calculated as $u_i = H_i \times r_i$. For persons who are infected but not diagnosed before t_0, another diagnosis delay weight is defined as $W = U/\sum_{i=1}^{N} u_i$. Suppose the estimated number of infections in the ith year before t_0 is m_i (using the CD4 depletion model). Then, the number of new infections in the ith year before t_0 can be estimated as

$$\nu_i = m_i W.$$

In fact, the method based on a CD4+ T-cell depletion model is also a kind of back-calculation method (it is sometimes referred to as the extended back-calculation method) because it also uses HIV/AIDS or CD4 T-cell counts at diagnosis to 'back-calculate' the time of infection among infected individuals. In the classical back-calculation method, only the total number of HIV/AIDS cases is required. In the extended back-calculation method, CD4+ T-cell counts at diagnosis are required at the individual level. For non-experts, the extended back-calculation method is easier to carry out owing to its low computational complexity compared with the classical back-calculation method. Nevertheless, similar to the classical back-calculation method, the validity of the extended back-calculation method is highly dependent on the CD4+ T-cell depletion model. In many countries and geographic areas, the empirical data required to estimate parameters of the CD4+ T-cell depletion model are extremely scarce. In China for example, after the test-and-treat policy became widespread, it became much more difficult to empirically observe CD4+ T-cell count data during natural infection in the absence of ART.

Simple method using HIV case reporting data

In 2017, Xia et al. [52] proposed a very simple novel method by which even non-experts can estimate HIV incidence using HIV case reporting data. The method assumes that HIV incidence and case finding are stable within each 3-year period. The timeframe of interest is broken down into overlapping 3-year periods (e.g., $2002 - 2004, 2003 - 2005, \cdots, 2008 - 2010$). The HIV incidence for the second year of each 3-year period can be estimated by solving the following equations:

$$R = \frac{D_1}{U_1 + I_1} = \frac{D_2}{U_2 + I_2} = \frac{D_3}{U_3 + I_3},$$

and

$$U_2 = U_1 + I_1 - D_1,$$
$$U_3 = U_2 + I_2 - D_2,$$
$$I_2 = I_1 + \varepsilon_1, \text{ where } \varepsilon_1 \text{ is small},$$
$$I_3 = I_2 + \varepsilon_2, \text{ where } \varepsilon_2 \text{ is small}.$$

R is the case finding rate in a year, D_i is the number of new diagnoses in year i, U_i is the number of undiagnosed cases at the beginning of year i, and I_1 is the HIV incidence in year i ($i = 1, 2, 3$). Then,

$$I_2 \approx \frac{D_1 D_3 - D_2 D_2}{D_1 - 2D_2 + D_3}.$$

This method is simple enough for non-experts. Moreover, it is very easy to carry out, requiring only HIV case reporting data. However, the method is applicable only if both incidence and diagnosis rates are stable over three years.

Cohort studies

Another strategy for estimation of HIV incidence is to use cohort studies of uninfected individuals [53]. Since it is difficult to follow sufficient individuals at the national level, a cohort study design is employed for estimating incidence among subpopulations [11]. This method enables researchers to directly measure HIV incidence in the sample population, but biases are introduced when estimating incidence by cohort. These biases are mainly caused by two sources of error [11]. First, individuals who receive follow-up visits may not be representative of the population. Second, individuals who adhere to the follow-up visits may obtain counseling repeatedly, and thus, their knowledge of HIV may improve over time which could affect risk of acquiring HIV.

Prevalence data

Incidence and prevalence are two important metrics for evaluating HIV epidemics. In fact, these two measures are related to one another. Two different types of prevalence data have been used to estimate HIV incidence: serial prevalence and cross-sectional prevalence [54–59]. In this section, we review two different incidence estimation methods using serial and cross-sectional prevalence data.

Estimating incidence from serial prevalence surveys

UNAIDS has developed an Estimation and Projection Package (EPP) which can be used to obtain HIV prevalence and projections [57]. Another software program, SPECTRUM [58], internally linked with EPP, can be employed to calculate the HIV incidence using the AIDS Impact Model (AIM) module. Here, we summarize the simplified methodology implemented in SPECTRUM. Let $H_{a,t}, A_{a,t}$ and $P_{a,t}$ denote the number of HIV infections, the total number of adults in the population and HIV prevalence of individuals aged a at time t, respectively. Thus,

$$H_{a,t} = A_{a,t} \times P_{a,t}.$$

New HIV infections ($I_{a,t}$) are the number of HIV infections in year t minus the number of survived infections from year $t-1$. The number of survived infections from year $t-1$ can be further calculated as the number of HIV infections in year $t-1$ minus deaths among HIV infected individuals (including deaths caused by AIDS, $D^{A}_{a,t}$, or deaths from other reasons, $D^{NA}_{a,t}$) in year $t-1$:

$$I_{a,t} = H_{a,t} - \left(H_{a-1,t-1} - D^{A}_{a-1,t-1} - D^{NA}_{a-1,t-1}\right).$$

AIDS deaths in year t are calculated as the convolution of the number of new infections in year $t-i$ and the proportion of deaths caused by AIDS i years after infection (r_i):

$$D^{A}_{a,t-1} = \sum_{i=0}^{20}(I_{a-i,t-i} \times r_i).$$

SPECTRUM has been updated several times since its initial 2004 release [60–63], and the last update took place in 2017 [64]. Other studies using the similar modeling approach have been conducted to estimate HIV epidemic [65, 66]. Hallett et al. [10] indicated that this method can estimate HIV incidence from the earliest stages of the epidemic, which is helpful to evaluate HIV epidemics over time. However, if large amounts of data are available, the estimate will involve a large uncertainty as the variation range of the incidence curve is very large. Since SPECTURM need to use EPP to generate the prevalence estimate and projections, and subsequently estimate the incidence of new HIV infections, any change in the incidence can only be detected through prevalence changes that may be observed over several years in later time. An additional disadvantage of this method is the difficulty in choosing an appropriate dataset from which prevalence is estimated. The estimation of HIV incidence could be significantly biased if the prevalence for the entire population is not estimated properly. For the long time, epidemiologists have used the data from antenatal clinics to estimate the prevalence in the entire population in sub-Saharan Africa [57]. As the HIV prevalence then appeared to be greater than that of the general population, national population-based household HIV surveys data are additionally used to calibrate overall population prevalence [67–69]. In fact, EPP began to include such household survey data in the estimation [70, 71]. Besides, household survey could miss a large part of the population that was affected by the HIV epidemic, and may on the other hand yield substantially small estimate of the prevalence. Synthesizing the use of different datasets over time could act as a cause of biased estimation.

Calculating incidence from cross-sectional prevalence

Hallett et al. [59] proposed a method to estimate the age-specific incidence of HIV from cross-sectional prevalence data. They first estimated incidence based on cohort mortality rates of infected individuals as well as survival distributions following HIV infection, then calculated age-specific incidence according to the relationship between these two measures.

In the following, we first summarize the relationship between age-specific incidence and cohort incidence. The age group i is defined as individuals aged from $a_i - \frac{r}{2}$ to $a_i + \frac{r}{2}$. Thus, the age group i is centered at a_i with a width of r years. The total number of individuals and HIV-infected individuals in age group i at time j are denoted by $N_{i,j}$ and $H_{i,j}$, respectively. Then, the prevalence is $p_{i,j} = H_{i,j}/N_{i,j}$.

We assume that cross-sectional prevalence is measured with an interval of T years in such age groups. Thus, age cohorts can be constructed as aged $a_i - \frac{r}{2}$ to $a_i + \frac{r}{2}$ at the start and $a_i - \frac{r}{2} + T$ to $a_i + \frac{r}{2} + T$ at the end of each interval. Now the cohort incidence, which is denoted by $\tilde{\lambda}_i$, can be illustrated by diagonal parallelogram (regions A and B in Fig. 1). The conventional age-specific incidence

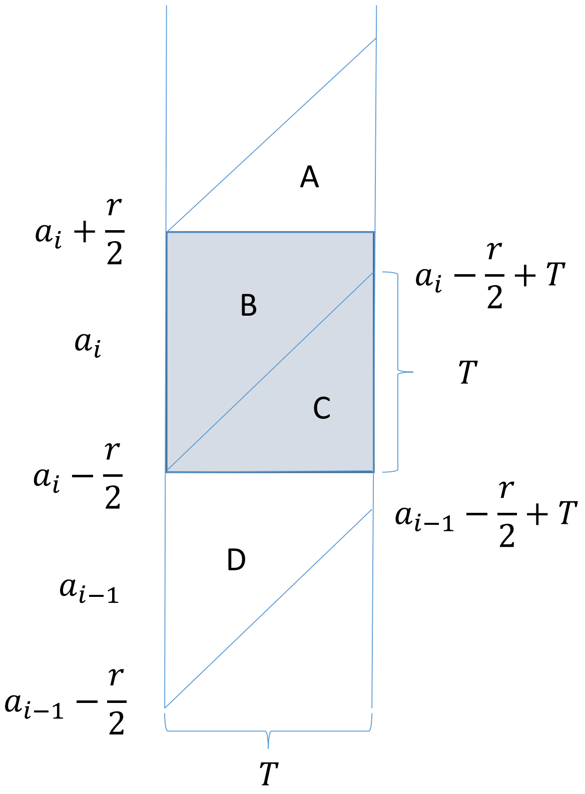

Fig. 1 Diagram of age cohort experience of incidence and conventional age-specific incidence

rate for age-group i, which is denoted by λ_i, is illustrated by regions B and C in Fig. 1. As Fig. 1 shows, region C can be seen as part of the incidence of cohort $i-1$. Denote the areas of regions A and B as S_A, S_B. The total area for the diagonal parallelogram is Tr (that is, $S_A + S_B$). The fractions contributed by cohort i and $i-1$ are $1 - T/2r$ $(S_B/(S_A + S_B))$ and $T/2r$ $(S_A/(S_A + S_B))$, respectively. Then, the conventional age-specific incidence rate can be calculated using the following equation:

$$\lambda_i = \left(1 - \frac{T}{2r}\right)\tilde{\lambda}_i + \left(\frac{T}{2r}\right)\tilde{\lambda}_{i-1}.$$

The derivation of the above formula assumes that $T \leq r$. When $T > r$, a similar method can be used for deriving a different formula, which is omitted here.

In the following, the methodological background of the cohort incidence estimation $\tilde{\lambda}_i$ is described. Let $\tilde{\pi}_i$ be the fraction of infected individuals in the ith age-group who survive from the start to the end of the interval, and $\tilde{\mu}_i$ be the mortality rate during this interval for individuals in the ith age-group who are uninfected. Then, the number of seroconverting individuals in age group i during the interval T can be approximated as

$$H_{i,T} - \tilde{\pi}_i H_{i,0},$$

and the number of person-years spent by age-group i during the interval T is approximated as

$$T\frac{(N_{i,0} - H_{i,0}) + (N_{i,T} - H_{i,T})}{2}.$$

Then, the cohort incidence can be derived as

$$\tilde{\lambda}_i = \frac{\text{seroconversions}}{\text{person-years}} = \frac{2(Q_i p_{i,T} - \tilde{\pi}_i p_{i,0})}{T(1 - p_{i,0} + Q_j(1 - p_{i,T}))},$$

where Q_j denotes the change in the size of the cohort over the time interval T.

$$Q_j \approx 1 - (1 - \tilde{\pi}_i)p_{i,0} - (1 - exp(-\tilde{\mu}_i T))(1 - p_{i,0}).$$

The authors further defined age cohort 0, calculating the prevalence at the start and end of the interval, and subsequently, the cohort incidence for this age cohort. $\tilde{\pi}_i$ can either be estimated based on the age-specific cohort mortality rates of infected individuals, or estimated using the distribution of survival time after infection, although we omit the details in this review. To use this method, age-specific cross-sectional prevalence data are required. Moreover, the duration between two cross-sectional measurements of prevalence should be small to ensure that the incidence and prevalence do not change significantly during this time interval. Because people with long survival time are preferably included in cohorts, the time-length bias is inevitable with this method. Both methods using serial and cross-sectional prevalence could be affected by the increasing coverage of ART [10, 59].

Biomarker approach for cross-sectional incidence estimation

It is widely recognized that recent infection rates are difficult to estimate using the back-calculation method owing to the long incubation period of AIDS, while cohort studies have difficulty following a sufficient number of high-risk uninfected persons. To complicate these issues, ART can considerably extend the incubation period, adding complexity to the majority of estimation methods mentioned above. As a possible alternative, a biomarker-based approach using cross-sectional incidence estimation was proposed and has clear advantages in estimation of recent infections [72–86]. This approach uses biomarkers from biological samples collected in cross-sectional studies to identify recent HIV infections.

Using diagnostic tests for the p24 antigen during the pre-seroconversion period.

In 1995, Brookmeyer et al. [72] proposed a simplistic modeling approach that uses diagnostic tests for HIV-1 p24 antigen to determine the prevalence of individuals who are p24 antigen-positive among HIV-seronegative individuals. Let μ be the mean duration of the p24 antigen-positive period before seroconversion, I be the infection risk per unit time for each uninfected individual (that is, the current incidence rate), and p be the expected proportion of individuals who are p24 antigen-positive among individuals whose HIV-antibody test results are negative or indeterminate. Then, p can be approximated as $I\mu$, and I can be estimated as

$$I = p/\mu.$$

Here, μ is referred to as the window period during which infected individuals have not yet seroconverted, but are still identifiable using biomarker(s). Supposing that $\tilde{p}$ is the number of individuals who are p24 antigen-positive during the window period, and n is the total number of individuals in the cross-sectional survey whose HIV antibody tests results are negative or indeterminate (i.e., $p = \frac{\tilde{p}}{n}$). Then, we have

$$I = \frac{1}{\mu} \cdot \frac{\tilde{p}}{n}.$$

The confidence interval for the incidence rate can be further estimated by assuming that $\tilde{p}$ follows a Poisson distribution with expectation $nI\mu$.

Using HIV enzyme immunoassay (EIA), antibody avidity index or genetic diversity

For the method proposed by Brookmeyer et al. [72], all individuals whose HIV antibody tests are negative need to undertake diagnostic testing for p24 antigen. Since the duration of the p24 antigen-positive pre-seroconversion period (window period) μ is very short (mean duration 22.5 days [72]), a large number of individuals need to be tested in situations where I (the population incidence rate) is high or n (the number of individuals that can be tested) is large. Janssen et al. [73] developed a new method to employ a testing algorithm based on either a sensitive assay (3A11) or a less-sensitive assay (3A11-LS). For a given cohort study, let T be the mean duration between seroconversion for the two assays (i.e., the window period), n be the number of individuals who are 3A11 reactive and 3A11-LS non-reactive, and N be the number of individuals who are HIV-negative or 3A11 reactive/3A11-LS non-reactive. Then, the incidence rate is

$$I = \frac{n}{N} \cdot \frac{1}{T}.$$

The window period using the sensitive/less sensitive assay testing algorithm is longer (i.e., 129 days) [73]. However, the algorithm does not perform well in populations infected with non-B HIV-1 subtypes [74]. Parekh et al. [75] proposed a subtype-independent assay called BED capture EIA (BED-CEIA; named after HIV subtypes B, E, and D), which can be used for detecting recent infections in populations infected by multiple HIV-1 subtypes. The mean BED window period is 156 days. Using the BED assay, Karon et al. [76] further proposed a method which can take into account information on history of HIV testing. Here, testing history refers both to whether an individual has undertaken HIV testing prior to HIV infection as well as the testing frequency. Since the antibody avidity index is always low during early infection, another method for estimation of recent infections based on the avidity index was proposed [77]. Genetic diversity of HIV has also been used as a biomarker to estimate HIV incidence [78–81], since it changes as the disease progresses. Other published studies [78, 79] identified recent HIV-1 infections based on data from traditional or next-generation DNA sequencing. Another research team [80, 81] developed a method based on a high-resolution melting (HRM) diversity assay to determine HIV diversity without sequencing.

Multiassay algorithms (MAAs)

The above serological assays have limitations because of their low accuracy in distinguishing recent from chronic infections. Some chronic infections may be misclassified as recent infection, and thus these methods may overestimate HIV incidence [82, 83]. Laeyendecker et al. [82, 83] demonstrated that factors such as low viral loads, low CD4+ T-cell counts, and $>$ 2 years of ART were associated with misclassification by the BED-CEIA. Avidity assays, which identify recent infections by studying the maturity of the antibody response against HIV, also have difficulties in distinguishing recent infections for HIV-1 incidence estimation [87, 88]. Laeyendecker et al. [84] and Brookmeyer [85] developed a MAA to estimate HIV incidence. The MAA integrates data from BED-CEIAs, antibody avidity assays, HIV viral loads and CD4+ T-cell counts. These algorithms are described in Fig. 2a and b, respectively.

All biomarker approaches estimate incidence at a time prior to sample collection, and the concept of the shadow describes the lag-time [85, 89–91]. Shadow and mean window period are two distinct but important concepts for evaluating the statistical accuracy of current HIV incidence estimates. Estimation approaches with large mean window periods will have smaller standard errors, and those with small shadows can better estimate more recent incidence [85, 89–91]. Thus, estimation approaches involving a larger mean window period and a smaller shadow are desirable.

The difference between the MAAs proposed by Laeyendecker et al. [84] and Brookmeyer [85] is that they use different cut-offs for CD4+ T-cell counts, BED-CEIAs, avidity and viral loads. Thus, the two algorithms have different mean window periods (141 days, 95% confidence interval (CI) (94, 150) vs. 159 days, 95% CI (134, 186), respectively) and shadows (128 days vs. 184 days, respectively). As Fig. 2a and b show, both of these algorithms

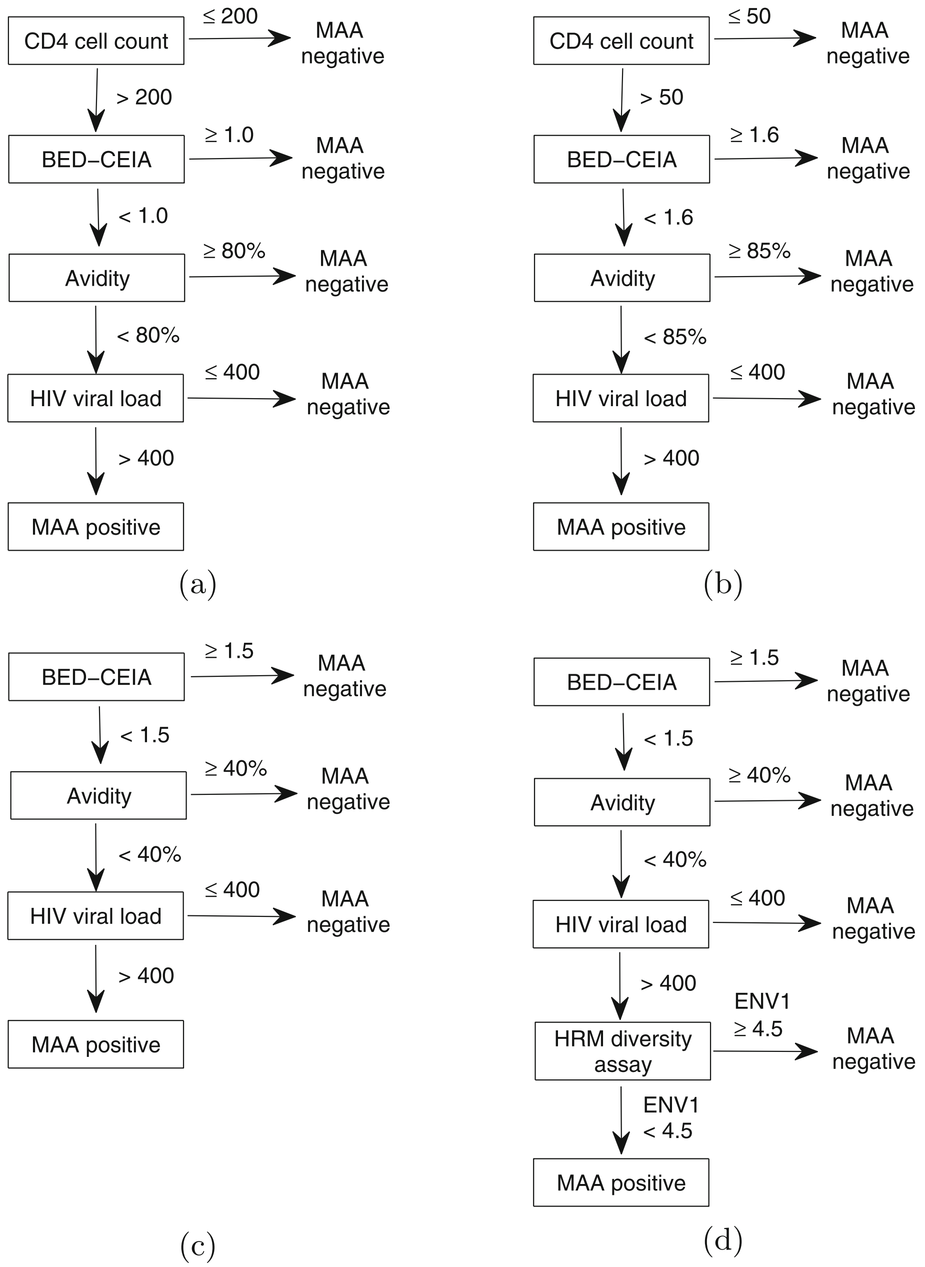

Fig. 2 Multi-assay algorithms (MAAs) for cross-sectional HIV incidence estimation (**a**) and (**b**) MAAs using CD4+ T-cell counts with different cut-off values. (**c**) MAA using only three biomarkers. (**d**) MAA using HRM diversity assay

require CD4+ T-cell counts, which are difficult to obtain in some settings. Thus, Laeyendecker et al. [85] developed another MAA using only three biomarkers (BED, avidity, and viral load) as shown in Fig. 2c. This three-biomarker-assay does not require CD4+ T-cell count data, and thus is less expensive. However, the mean window period for the three-biomarker-assay is 58 days shorter than that of the four-biomarker-assay. Hence, to achieve the same incidence standard error, the three-biomarker-assay requires a sample size about 57% larger.

Cousins et al. [86] proposed a new MAA in which a HRM diversity assay is used in place of CD4+ T-cell count data, as shown in Fig. 2d. The mean window period and shadow for the HRM-based MAA are 154 days (95% CI 128, 180 days) and 179 days (95% CI 135, 243 days), respectively. The performance of the HRM-based MAA was shown to be nearly identical to that of the MAA including CD4+ T-cell count data.

For all MAAs, HIV incidence is calculated using the following equation:

$$I = \frac{n}{N} \cdot \frac{1}{T},$$

where n is the number of MAA-positive subjects, N is the total number of individuals who are HIV seronegative, and T is the mean window period.

Several narrative reviews have been published describing incidence estimation approaches that use biomarker data [10, 11, 74, 88, 91–93]. Technical challenges of biomarker approach include misclassification of chronic infections as recent infections and a large variation in testing results between individuals [10], although the accuracy of recent infection estimate has been markedly improved by using MAAs. As reviewed by Murphy et al. [93], biomarker-based incidence could achieve high precision if false recency ratio is sufficiently close to zero. Moreover, as the biomarker approaches include a variety of biomarkers, the complexity to identify recent infections has become more and more complex over time, which may sometimes even require specialized equipments. Early treatment and the use of pre- and post-exposure prophylaxis also bring new challenges to the biomarker approaches. In recent years, the incidence in some populations or sub-populations have been estimated by using biomarker approaches [94–96], and sometimes the biomarker method was combined with other existing modelling approaches [97, 98]. Because of financial constraints, insufficient coordinated action among funding bodies, governments and developers could also act as a hazard for propagating this approach [93], frequently involving problems in purchasing agreement and limited financial support for quality control and training.

Discussion

In this review, six major methods for estimating HIV incidence were briefly described. These included the back-calculation method, methods using CD4+ T-cell depletion models, methods using HIV case reporting data, methods based on cohort studies, methods using prevalence data, and biomarker-based approaches. Back-calculation methods can be divided into three subgroups according to the data used: (i) AIDS diagnosis data only, (ii) both HIV and AIDS diagnosis data, and (iii) HIV/AIDS diagnosis data as well as CD4+ T-cell counts at diagnosis. Similarly, methods using prevalence data can be further divided into methods based on serial and cross-sectional data. Our primary foci were the background mechanism of estimation, the required data types, the scope of application, the model formulation, the derivation of the maximum likelihood function, and the advantages and disadvantages of applying each method in practice.

Back-calculation methods are widely used to estimate the incidence and prevalence of HIV in various parts of the world [43, 46]. These methods were initially developed using AIDS diagnosis data only, but were later extended to use both HIV and AIDS diagnoses, and then to further account for CD4+ T-cell counts at diagnosis. The back-calculation method has also been modified to include the effect of ART on the distribution of the incubation period. Back-calculation methods have clear advantages and disadvantages compared with other methods [99]. First, the back-calculation method requires only data from case reporting systems, and does not necessarily require laboratory testing and individual-level data. However, the incidence estimate in recent years tends to be unstable, especially where the HIV epidemic has just started, and accuracy of the estimate is influenced by the distribution of the incubation period (or the progression rate) as well as the testing rate.

Compared with back-calculation method, it would be easier to implement the statistical estimation using CD4+ T-cell depletion among non-experts. However, it assumes that the distribution of delays in diagnosis does not change over time. Thus, it may overestimate HIV incidence if HIV testing rates increase over time. As mentioned above, cohort studies have many difficulties and may introduce some biases when incidence is directly estimated among high-risk populations with close follow-up. For methods using prevalence data, both methods using serial and cross-sectional prevalence data are associated with uncertainties in evaluating HIV prevalence and AIDS deaths. Moreover, both methods are strongly influenced by the use of ART. Methods using cross-sectional prevalence data further assume that HIV incidence during the time interval between two prevalence surveys is constant, which is only true for very short time intervals. Biomarker-based approaches, which uses biomarkers in biological samples collected in cross-sectional studies to identify recent HIV infections, can avoid the difficulties associated with follow-up of high-risk uninfected persons in cohort studies as well as difficulties in estimating the distribution of long incubation periods. Biomarker-based methods can better estimate more recent HIV incidence compared with the back-calculation method. As laboratory testing techniques progress, MAAs have become available at low cost, which could minimize the effort and cost involved in incidence estimation in the future. Nevertheless, minimizing the ‘false recency ratio’ (FRR) at a sufficiently low level remains to be a challenge. Biomarker approaches also involve other technical difficulties in quality control, training and evaluation of assays.

The required data are, at the moment, divided into four different categories: (i) epidemiological data including AIDS diagnoses and HIV diagnoses, (ii) CD4 T-cell counts at diagnosis, (iii) prevalence data, and (iv) biomarker testing data. Prevalence data may be further divided into serial prevalence and cross-sectional prevalence data. It must be noted that definitions of HIV incidence are not uniform across different methods. For the back-calculation method, methods using CD4+ T-cell depletion models, methods using cohort studies and methods using serial prevalence data, HIV incidence is defined as the number of new HIV infections per unit time (year) or the instantaneous incident infections occurring at time *t*. However, for methods using cross-sectional prevalence data, HIV incidence is defined as the average hazard of new infections occurring during the interval. For the biomarker approach, an HIV incidence rate is estimated, which is defined as the infection risk per unit time for each uninfected individual (except for the method using BED-CEIA [76], which estimates conventional incidence instead). Obviously, conventional incidence and incidence rates can be converted as long as the total number of uninfected individuals is known. In addition to different incidence definitions, there is also another difference among these methods. The back-calculation method, methods using CD4+ T-cell depletion model, methods using cohort studies and methods using serial prevalence data can estimate serial incidence (i.e., the incidence year-over-year). However, the method using cross-sectional prevalence data and the biomarker approach estimate the cross-sectional incidence or the HIV incidence at a time prior to collection of samples. Thus, different methods estimate HIV incidence with variable time frames.

Conclusion

A variety of methods exist to estimate HIV incidence from different data types and scopes, and it is difficult to conclude which method perform best. Rather, it should be remembered that HIV incidence estimation itself described what cannot be directly validated, as the estimated quantity is not directly observable in natural settings. Thus, a new method should be regarded as way to mitigate uncertainty with respect to the estimates of another method, and analyzing HIV data from multiple standpoints and sources is one way to overcome such uncertainty. As the methods for HIV incidence estimation have different scopes and different advantages and disadvantages, we hope that this review will be useful for determining which datasets need to be collected to estimate HIV incidence in a comprehensive manner. Should a surveillance system be improved to collect multiple types of datasets as described above, it would be feasible to cross-validate different methodologies and see how different methods can complement each other so that an objective assessment of the HIV/AIDS epidemic will be eventually achieved.

Abbreviations

AIDS: Acquired immunodeficiency syndrome; ART: Antiretroviral therapy; CI: Confidence interval; HIV: Human immunodeficiency virus; UNAIDS: Joint united nations programme on HIV/AIDS

Acknowledgements

We thank Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Authors' contributions

XS and HN jointly developed the idea for the review. XS drafted the early version of the manuscript, and HN assessed and discussed the mathematical formulation and analysis of the methods. YX provided comments on the revised manuscript. XS drafted the figures. All authors provided comments on the revised manuscript and approved the final version of the manuscript.

Funding

XS was supported in Japan by the Japan Society for the Promotion of Science (JSPS) Program for Advancing Strategic International Networks to Accelerate the Circulation of Talented Researchers. HN received funding support from the Japan Agency for Medical Research and Development (JP18fk0108050), the Japan Society for the Promotion of Science KAKENHI (Grant Numbers 16KT0130, 17H04701, 17H05808 and 18H04895), the Health and Labour Sciences Research Grant (H28-AIDS-General-001), the Inamori Foundation, the Telecommunication Advancement Foundation; and the Japan Science and Technology Agency (JST) CREST program (JPMJCR1413). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that co-author H. Nishiura is the Editor-in-Chief of Theoretical Biology and Medical Modelling. This does not alter the authors' adherence to all the Theoretical Biology and Medical Modelling policies on sharing data and materials.

Author details

[1] Department of Applied Mathematics, Xi'an Jiaotong University, No 28, Xianning West Road, Xi'an, 710049 Shaanxi, China. [2] Graduate School of Medicine, Hokkaido University, Kita 15 Jo Nishi 7 Chome, Kitaku, 0608638 Sapporo, Japan.

Received: 14 December 2018 Accepted: 24 December 2019
Published online: 22 January 2020

References

1. Merson MH. The HIV-AIDS pandemic at 25-the global response. N Engl J Med. 2006;354(23):2414–7.
2. UNAIDS. Fact sheet - World AIDS Day. 2019. Available from: https://www.unaids.org/en/resources/fact-sheet. Accessed 10 Jan 2020.
3. Pratt RD, Shapiro JF, McKinney N, Kwok S, Spector SA. Virologic characterization of primary human immunodeficiency virus type 1 infection in a health care worker following needlestick injury. J Infect Dis. 1995;172(3):851–4.
4. Quinn TC. Acute primary HIV infection. JAMA. 1997;278(1):58–62.
5. Henrard DR, Phillips JF, Muenz LR, Blattner WA, Wiesner D, Eyster ME, et al. Natural history of HIV-1 cell-free viremia. JAMA. 1995;274(7):554–8.
6. O'brien TR, Blattner WA, Waters D, Eyster ME, Hilgartner MW, et al. Serum HIV-1 RNA levels and time to development of AIDS in the Multicenter Hemophilia Cohort Study. JAMA. 1996;276(2):105–10.
7. Wong MT, Dolan MJ, Kozlow E, Doe R, Melcher GP, Burke DS, et al. Patterns of virus burden and T cell phenotype are established early and are correlated with the rate of disease progression in human immunodeficiency virus type 1-infected persons. J Infect Dis. 1996;173(4): 877–87.
8. Scott HM, Vittinghoff E, Irvin R, Sachdev D, Liu A, et al. Age, race/ethnicity, and behavioral risk factors associated with per-contact risk of HIV infection among men who have sex with men in the United States. J Acquir Immune Defic Syndr. 2014;65(1):115.
9. Del Romero J, Marincovich B, Castilla J, Garcia S, Campo J, et al. Evaluating the risk of HIV transmission through unprotected orogenital sex. AIDS. 2002;16(9):1296–7.

10. Hallett TB. Estimating the HIV incidence rate: recent and future developments. Curr Opin HIV AIDS. 2011;6(2):102.
11. Brookmeyer R. Measuring the HIV/AIDS epidemic: approaches and challenges. Epidemiol Rev. 2010;32(1):26–37.
12. Working Group on Estimation of HIV Prevalence in Europe. HIV in hiding: methods and data requirements for the estimation of the number of people living with undiagnosed HIV. AIDS. 2011;25(8):1017–23.
13. Brookmeyer R, Gail M. Minimum size of the acquired immunodeficiency syndrome (AIDS) epidemic in the United States. Lancet. 1986;328(8519): 1320–22.
14. Brookmeyer R, Gail MH. A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. J Am Stat Assoc. 1988;83(402):301–8.
15. Brookmeyer R, Damiano A. Statistical methods for short-term projections of AIDS incidence. Stat Med. 1989;8(1):23–34.
16. Becker NG, Watson LF, Carlin JB. A method of non-parametric back-projection and its application to AIDS data. Stat Med. 1991;10(10): 1527–42.
17. Brizzi F. Estimating HIV incidence from multiple sources of data. Cambridge: Doctoral dissertation, University of Cambridge; 2018. https://www.repository.cam.ac.uk/handle/1810/273803. Accessed 10 Jan 2020.
18. Isham V. Estimation of the incidence of HIV infection. Philos T R Soc B. 1989;325(1226):113–21.
19. Day NE, Gore SM, McGee MA, South M. Predictions of the AIDS epidemic in the UK: the use of the back projection method. Philos T R Soc B. 1989;325(1226):123–34.
20. Rosenberg PS, Gail MH. Backcalculation of flexible linear models of the human immunodeficiency virus infection curve. J R Stat Soc Ser C Appl Stat. 1991269–82. https://doi.org/10.2307/2347592.
21. Yan P, Zhang F. A case study of nonlinear programming approach for repeated testing of HIV in a population stratified by subpopulations according to different risks of new infections. Oper Res Health Care. 2018;19:120–33.
22. Solomon PJ, Wilson SR. Accommodating change due to treatment in the method of back projection for estimating HIV infection incidence. Biometrics. 19901165–70.
23. Brookmeyer R, Liao J. Statistical modelling of the AIDS epidemic for forecasting health care needs. Biometrics. 19901151–63. https://doi.org/10.2307/2532455.
24. Brookmeyer R. Reconstruction and future trends of the AIDS epidemic in the United States. Science. 1991;253(5015):37–42.
25. Longini IM, Byers RH, Hessol NA, Tan WY. Estimating the stage-specific numbers of HIV infection using a Markov model and back-calculation. Stat Med. 1992;11(6):831–43.
26. Rosenberg PS, Goedert JJ, Biggar RJ. Effect of age at seroconversion on the natural AIDS incubation distribution. Multicenter Hemophilia Cohort Study and the International Registry of Seroconverters. AIDS. 1994;8(6): 803–10.
27. Aalen OO, Farewell VT, De Angelis D, Day NE. The use of human immunodeficiency virus diagnosis information in monitoring the acquired immune deficiency syndrome epidemic; 1994, pp. 3–16. https://doi.org/10.2307/2983501.
28. Marschner IC. Using time of first positive HIV test and other auxiliary data in back-projection of AIDS incidence. Stat Med. 1994;13(19–20):1959–74.
29. Farewell VT, Aalen OO, Angelis DD, MRC ND. Estimation of the rate of diagnosis of HIV infection in HIV infected individuals. Biometrika. 1994;81(2):287–94.
30. Dietz K, Seydel J, Schwartländer B. Back-projection of German AIDS data using information on dates of tests. Stat Med. 1994;13(19-20):1991–2008.
31. Raab GM, Fielding KL, Allardice G. Incorporating HIV test data into forecasts of the AIDS epidemic in Scotland. Stat Med. 1994;13(19-20): 2009–20.
32. De Angelis D, Gilks WR, Day NE. Bayesian projection of the acquired immune deficiency syndrome epidemic. J R Stat Soc Ser C Appl Stat. 1998;47(4):449–98.
33. Aalen OO, Farewell VT, De Angelis D, Day NE, Nöel Gill O. A Markov model for HIV disease progression including the effect of HIV diagnosis and treatment: application to AIDS prediction in England and Wales. Stat Med. 1997;16(19):2191–210.
34. Bellocco R, Marschner IC. Joint analysis of HIV and AIDS surveillance data in back-calculation. Stat Med. 2000;19(3):297–311.
35. Cui J, Becker NG. Estimating HIV incidence using dates of both HIV and AIDS diagnoses. Stat Med. 2000;19(9):1165–77.
36. Becker NG, Lewis JJC, Li Z, McDonald A. Age-specific back-projection of HIV diagnosis data. Stat Med. 2003;22(13):2177–90.
37. Chau PH, Yip PSF, Cui JS. Reconstructing the incidence of human immunodeficiency virus (HIV) in Hong Kong by using data from HIV positive tests and diagnoses of acquired immune deficiency syndrome. J R Stat Soc Ser C Appl Stat. 2003;52(2):237–48.
38. Posner SJ, Myers L, Hassig SE, Rice JC, Kissinger P, Farley TA. Estimating HIV incidence and detection rates from surveillance data. Epidemiology. 2004;15(2):164–72.
39. Sommen C, Alioum A, Commenges D. A multistate approach for estimating the incidence of human immunodeficiency virus by using HIV and AIDS French surveillance data. Stat Med. 2009;28(11):1554–68.
40. An Q, Kang J, Song R, Hall HI. A Bayesian hierarchical model with novel prior specifications for estimating HIV testing rates. Stat Med. 2016;35(9): 1471–87.
41. Yan P, Zhang F, Wand H. Using HIV diagnostic data to estimate HIV incidence: method and simulation. Stat Commun Infec Dis. 2011;3(1):. https://doi.org/10.2202/1948-4690.1011.
42. Wand H, Wilson D, Yan P, Gonnermann A, McDonald A, Kaldor J, et al. Characterizing trends in HIV infection among men who have sex with men in Australia by birth cohorts: results from a modified back-projection method. J Int AIDS Soc. 2009;12(1):19.
43. Wand H, Yan P, Wilson D, McDonald A, Middleton M, Kaldor J, et al. Increasing HIV transmission through male homosexual and heterosexual contact in Australia: results from an extended back-projection approach. HIV Med. 2010;11(6):395–403.
44. Nishiura H. Estimating the incidence and diagnosed proportion of HIV infections in Japan: a statistical modeling study. PeerJ. 2019;7:e6275.
45. Birrell PJ, Chadborn TR, Gill ON, Delpech VC, De Angelis D. Estimating trends in incidence, time-to-diagnosis and undiagnosed prevalence using a CD4-based Bayesian back-calculation. Stat Commun Infec Dis. 2012;4(1):. https://doi.org/10.1515/1948-4690.1055.
46. Birrell PJ, Gill ON, Delpech VC, Brown AE, Desai S, Chadborn TR, et al. HIV incidence in men who have sex with men in England and Wales 2001-10: a nationwide population study. Lancet Infect Dis. 2013;13(4):313–8.
47. Hall HI, Song R, Szwarcwald CL, Green T. Brief report: Time from infection with the human immunodeficiency virus to diagnosis, United States. J Acquir Immune Defic Syndr. 2015;69(2):248–51.
48. Szwarcwald CL, Pascom ARP, Souza J. Estimation of the HIV incidence and of the number of people living with HIV/AIDS in Brazil, 2012. J AIDS Clin Res. 2015;6(3):. https://doi.org/10.4172/2155-6113.1000430.
49. Song R, Hall HI, Green TA, Szwarcwald CL, Pantazis N. Using CD4 data to estimate HIV incidence, prevalence, and percent of undiagnosed infections in the United States. J Acquir Immune Defic Syndr. 2017;74(1): 3–9.
50. Lodi S, Phillips A, Touloumi G, Geskus R, Meyer L, Thiébaut R, et al. Time from human immunodeficiency virus seroconversion to reaching CD4+ cell count thresholds ¡200, ¡350, and ¡500 cells/mm^3: assessment of need following changes in treatment guidelines. Clin Infect Dis. 2011;53(8): 817–825.
51. Touloumi G, Pantazis N, Pillay D, Paraskevis D, Chaix ML, Bucher HC, et al. Impact of HIV-1 subtype on CD4 count at HIV seroconversion, rate of decline, and viral load set point in European seroconverter cohorts. Clin Infect Dis. 2013;56(6):888–97.
52. Xia Q, Teixeira-Pinto A, Forgione LA, Wiewel EW, Braunstein SL, Torian LV, Estimated HIV. incidence in the United States, 2003-2010. J Acquir Immune Defic Syndr. 2017;74(1):10–14.
53. Karon JM, Fleming PL, Steketee RW, De Cock KM. HIV in the United States at the turn of the century: an epidemic in transition. Am J Public Health. 2001;91(7):1060–8.
54. Sakarovitch C, Alioum A, Ekouevi DK, Msellati P, Leroy V, Dabis F. Estimating incidence of HIV infection in childbearing age African women using serial prevalence data from antenatal clinics. Stat Med. 2007;26(2): 320–35.

55. Williams B, Gouws E, Wilkinson D, Karim SA, Estimating HIV. incidence rates from age prevalence data in epidemic situations. Stat Med. 2001;20(13):2003–16.
56. White RG, Vynnycky E, Glynn JR, Jahn A, Mwaungulu F, Mwanyongo O, et al. HIV epidemic trend and antiretroviral treatment need in Karonga District, Malawi. Epidemiol Infect. 2007;135(6):922–32.
57. Ghys PD, Brown T, Grassly NC, et al. The UNAIDS Estimation and Projection Package: a software package to estimate and project national HIV epidemics[J]. Sex Transm Infect. 2004;80(suppl 1):i5–i9.
58. Stover J. Projecting the demographic consequences of adult HIV prevalence trends: the Spectrum Projection Package. Sex Transm Infect. 2004;80(suppl 1):i14–8.
59. Hallett TB, Zaba B, Todd J, Lopman B, Mwita W, Biraro S, et al. Estimating incidence from prevalence in generalised HIV epidemics: methods and validation. PLoS Med. 2008;5(4):e80.
60. Stover J, Walker N, Grassly NC, et al. Projecting the demographic impact of AIDS and the number of people in need of treatment: updates to the Spectrum projection package. Sex Transm Infect. 2006;82(suppl 3): iii45–50.
61. Stover J, Johnson P, Zaba B, et al. The Spectrum projection package: improvements in estimating mortality, ART needs, PMTCT impact and uncertainty bounds. Sex Transm Infect. 2008;84(Suppl 1):i24–30.
62. Stover J, Johnson P, Hallett T, et al. The Spectrum projection package: improvements in estimating incidence by age and sex, mother-to-child transmission, HIV progression in children and double orphans. Sex Transm Infect. 2010;86(Suppl 2):ii16–21.
63. Stover J, Andreev K, Slaymaker E, Gopalappa C, Sabin K, Velasquez C, et al. Updates to the spectrum model to estimate key HIV indicators for adults and children. AIDS. 2014;28(Suppl 4):S427–34.
64. Stover J, Brown T, Puckett R, et al. Updates to the Spectrum/Estimations and Projections Package model for estimating trends and current values for key HIV indicators. AIDS. 2017;31(1):S5–11.
65. Hallett TB, Gregson S, Mugurungi O, Gonese E, Garnett GP. Assessing evidence for behaviour change affecting the course of HIV epidemics: a new mathematical modelling approach and application to data from Zimbabwe. Epidemics. 2009;1(2):108–17.
66. Gregson S, Gonese E, Hallett TB, Taruberekera N, Hargrove JW, Lopman B, et al. HIV decline in Zimbabwe due to reductions in risky sex? Evidence from a comprehensive epidemiological review. Int J Epidemiol. 2010;39(5):1311–23.
67. Garcia-Calleja JM, Gouws E, Ghys PD. National population based HIV prevalence surveys in sub-Saharan Africa: results and implications for HIV and AIDS estimates. Sex Transm Infect. 2006;82(suppl 3):iii64–70.
68. Gouws E, Mishra V, Fowler TB. Comparison of adult HIV prevalence from national population-based surveys and antenatal clinic surveillance in countries with generalised epidemics: implications for calibrating surveillance data. Sex Transm Infect. 2008;84(Suppl 1):i17–23.
69. Alkema L, Raftery AE, Clark SJ. Probabilistic projections of HIV prevalence using Bayesian melding. Ann Appl Stat. 2007;1(1):229–48.
70. Bao L, Salomon JA, Brown T, et al. Modelling national HIV/AIDS epidemics: revised approach in the UNAIDS Estimation and Projection Package 2011. Sex Transm Infect. 2012;88(Suppl 2):i3–10.
71. Brown T, Bao L, Eaton JW, et al. Improvements in prevalence trend fitting and incidence estimation in EPP 2013. AIDS (London, England). 2014;28(4):S415.
72. Brookmeyer R, Quinn TC. Estimation of current human immunodeficiency virus incidence rates from a cross-sectional survey using early diagnostic tests. Am J Epidemiol. 1995;141(2):166–72.
73. Janssen RS, Satten GA, Stramer SL, Rawal BD, O'brien TR, Weiblen BJ, et al. New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes. JAMA. 1998;280(1):42–48.
74. Parekh BS, Hu DJ, Vanichseni S, Satten GA, Candal D, Young NL, et al. Evaluation of a sensitive/less-sensitive testing algorithm using the 3A11-LS assay for detecting recent HIV seroconversion among individuals with HIV-1 subtype B or E infection in Thailand. AIDS Res Hum Retrovir. 2001;17(5):453–8.
75. Parekh BS, Kennedy MS, Dobbs T, Pau CP, Byers R, Green T, et al. Quantitative detection of increasing HIV type 1 antibodies after seroconversion: a simple assay for detecting recent HIV infection and estimating incidence. AIDS Res Hum Retrovir. 2002;18(4):295–307.
76. Karon JM, Song R, Brookmeyer R, Kaplan EH, Hall HI. Estimating HIV incidence in the United States from HIV/AIDS surveillance data and biomarker HIV test results. Stat Med. 2008;27(23):4617–33.
77. Suligoi B1, Galli C, Massi M, Di Sora F, Sciandra M, Pezzotti P, et al. Precision and accuracy of a procedure for detecting recent human immunodeficiency virus infections by calculating the antibody avidity index by an automated immunoassay-based method. J Clin Microbiol. 2002;40(11):4015–20.
78. Kouyos RD, von Wyl V, Yerly S, Böni J, Rieder P, Joos B, et al. Ambiguous nucleotide calls from population-based sequencing of HIV-1 are a marker for viral diversity and the age of infection. Clin Infect Dis. 2011;52(4):532–9.
79. Yang J, Xia X, He X, Yang S, Ruan Y, Zhao Q, et al. A new pattern-based method for identifying recent HIV-1 infections from the viral env sequence. Sci China Life Sci. 2012;55(4):328–35.
80. Cousins MM, Laeyendecker O, Beauchamp G, Brookmeyer R, Towler WI, Hudelson SE, et al. Use of a high resolution melting (HRM) assay to compare gag, pol, and env diversity in adults with different stages of HIV infection. PLoS ONE. 2011;6(11):e27211.
81. Cousins MM, Swan D, Magaret CA, Hoover DR, Eshleman SH. Analysis of HIV using a high resolution melting (HRM) diversity assay: automation of HRM data analysis enhances the utility of the assay for analysis of HIV incidence. PLoS ONE. 2012;7(12):e51359.
82. Laeyendecker O, Rothman RE, Henson C, Horne BJ, Ketlogetswe KS, Kraus CK, et al. The effect of viral suppression on cross sectional incidence testing in the Johns Hopkins hospital emergency department. J Acquir Immune Defic Syndr. 2008;48(2):211.
83. Laeyendecker O, Brookmeyer R, Oliver AE, Mullis CE, Eaton KP, Mueller AC, et al. Factors associated with incorrect identification of recent HIV infection using the BED capture immunoassay. AIDS Res Hum Retrovir. 2012;28(8):816–22.
84. Laeyendecker O, Brookmeyer R, Cousins MM, Mullis CE, Konikoff J, Donnell D, et al. HIV incidence determination in the United States: a multiassay approach. J Infect Dis. 2012;207(2):232–9.
85. Brookmeyer R, Konikoff J, Laeyendecker O, Eshleman SH. Estimation of HIV incidence using multiple biomarkers. Am J Epidemiol. 2013;177(3): 264–72.
86. Cousins MM, Konikoff J, Laeyendecker O, Celum C, Buchbinder SP, Seage GR3rd, et al. HIV diversity as a biomarker for HIV incidence estimation: including a high-resolution melting diversity assay in a multiassay algorithm. J Clin Microbiol. 2014;52(1):115–21.
87. Murphy G, Parry JV. Assays for the detection of recent infections with human immunodeficiency virus type 1. Euro Surveill. 2008;13(36):18966.
88. Guy R, Gold J, Calleja JM, Kim AA, Parekh B, Busch M, et al. Accuracy of serological assays for detection of recent infection with HIV and estimation of population incidence: a systematic review. Lancet Infect Dis. 2009;9(12):747–59.
89. Kaplan EH, Brookmeyer R. Snapshot estimators of recent HIV incidence rates. Oper Res. 1999;47(1):29–37.
90. Brookmeyer R. On the statistical accuracy of biomarker assays for HIV incidence. J Acquir Immune Defic Syndr. 2010;54(4):406–14.
91. Brookmeyer R, Laeyendecker O, Donnell D, Eshleman SH. Cross-sectional HIV incidence estimation in HIV prevention research. J Acquir Immune Defic Syndr. 2013;63:S233–9.
92. Mastro TD, Kim AA, Hallett T, Rehle T, Welte A, Laeyendecker O, et al. Estimating HIV incidence in populations using tests for recent infection: issues, challenges and the way forward. J HIV AIDS Surveill Epidemiol. 2010;2(1):1–14.
93. Murphy G, Pilcher CD, Keating SM, Kassanjee R, Facente SN, et al. Moving towards a reliable HIV incidence test-current status, resources available, future directions and challenges ahead. Epidemiol Infect. 2017;145(5):925–41.
94. Rehle T, Johnson L, Hallett T, Mahy M, Kim A, et al. A comparison of South African national HIV incidence estimates: A critical appraisal of different methods. PLoS ONE. 2015;10(7):e0133255.
95. Aghaizu A, Tosswill J, De Angelis D, Ward H, Hughes G, et al. HIV incidence among sexual health clinic attendees in England: First estimates for black African heterosexuals using a biomarker, 2009-2013. PLoS ONE. 2018;13(6):e0197939.
96. Laeyendecker O, Konikoff J, Morrison DE, Brookmeyer R, Wang J, et al. Identification and validation of a multi-assay algorithm for cross-sectional HIV incidence estimation in populations with subtype C infection. J Int AIDS Soc. 2018;21(2):e25082.

97. Giardina F, Romero-Severson E, Axelsson M, Svedhem V, Leitner T, et al. Getting more from heterogeneous HIV-1 surveillance data in a high immigration country: estimation of incidence and undiagnosed population size using multiple biomarkers. bioRxiv. 2018345710. https://doi.org/10.1101/345710.
98. Grebe E, Welte A, Johnson LF, Cutsem G, Puren A, et al. Population-level HIV incidence estimates using a combination of synthetic cohort and recency biomarker approaches in KwaZulu-Natal, South Africa. PLoS ONE. 2018;13(9):e0203638.
99. Mallitt KA, Wilson DP, McDonald A, Wand H. Is back-projection methodology still relevant for estimating HIV incidence from national surveillance data? Open AIDS J. 2012;6:108–11.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.