Mathematical Biosciences and Engineering

MBE, 22 (8): 2105–2119.
DOI: 10.3934/mbe.2025077
Received: 02 April 2025
Revised: 07 June 2025
Accepted: 18 June 2025
Published: 02 July 2025

https://www.aimspress.com/journal/MBE

Research article

On the rate of clinical AIDS on diagnosis: The mathematical interpretation and goal for the successful control of HIV/AIDS

Seiko Fujiwara and Hiroshi Nishiura

School of Public Health and Center for Health Security, Graduate School of Medicine, Kyoto University, Kyoto, Japan

* **Correspondence:** Email: nishiura.hiroshi.5r@kyoto-u.ac.jp; Tel: +81-075-753-4456.

Abstract: The most widely used measurement of transmission dynamics in real time is the effective reproduction number $R(t)$. However, in the context of human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS), $R(t)$ has not been used frequently, possibly because of the slowly progressing nature of HIV infection that limits the knowledge of recent infection events. Gaining deeper insights into the practically used epidemiological metrics of HIV/AIDS is therefore vital. Notably, in many high-income countries, including Japan, the rate of clinical AIDS on diagnosis, $Q(t)$, has been routinely measured by calculating the proportion of newly diagnosed AIDS cases out of all new HIV infections that are diagnosed at a given calendar time. However, there has been no clear indication of whether the control of HIV/AIDS is effective in relation to this metric in Japan. In this study, we formulated the rate of clinical AIDS on diagnosis using a mathematical model and offered interpretations of it using the hazard rate of diagnosis among previously undiagnosed HIV-infected individuals. We showed that by taking the inverse of the odds of $Q(t)$ and multiplying it by the inverse of the mean incubation period, we obtained $\alpha(t)$, which is the hazard rate of diagnosis among undiagnosed HIV-infected individuals. We also showed that $\alpha(t)$ can be related to the goal of the diagnosed proportion P_0 among all people living with HIV. In addition to the rate of clinical AIDS on diagnosis $Q(t)$, $\alpha(t)$ can be calculated using a simplistic equation and can potentially act as a practical epidemiological metric for monitoring during surveillance.

Keywords: diagnosis; testing; human immunodeficiency virus; public health; control; elimination

1. Introduction

Numerous evaluation metrics for transmission dynamics have been formulated from remarkable scientific advances in the mathematical modeling of infectious diseases. The most widely used measurement of transmission is the effective reproduction number, $R(t)$, as a function of calendar time t [1], which is interpreted as the average number of secondary cases generated by a single primary case at time t. It explicitly provides information on whether an epidemic is in a growing or declining trend. During the course of the COVID-19 pandemic, an ad hoc innovation was the prediction of $R(t)$ using promising predictors, including human mobility, risk awareness, and environmental variables in real time [2–4].

$R(t)$ has been used less frequently in the context of human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) compared with COVID-19. The $R(t)$ of HIV infection has been estimated in specific instances, but this has not been as common as estimations for infectious diseases with an acute course of illness. A possible reason is that HIV infection is progressing slowly in its natural history, and from the observed dataset of AIDS cases plus the diagnosed fraction of HIV infections, it is difficult to understand effective contact in the present day in a timely manner. As a possible shortcut to overcome the problem, Amundsen et al. [5] attempted to use the incidence-to-prevalence ratio; however, maintaining the precise threshold property using that alternative metric was later questioned [6]. Another reason is that, compared with using transmission, a more practical evaluation of HIV control places greater stress on the effort of testing and diagnosis rather than entire epidemiological dynamics. To further complicate the matter, the epidemiological metrics for testing effort (e.g., the number of HIV antibody tests per unit time) have not been connected successfully to a component of $R(t)$. In recent years, the modeling of HIV/AIDS dynamics has incorporated increasingly realistic features, such as treatment strategies, risk heterogeneity, and stochastic effects. For instance, Luo et al. [7] analyzed the roles of antiretroviral therapy (ART) and pre-exposure prophylaxis (PrEP) in a stochastic HIV/AIDS model. Zhang et al. [8] incorporated media influence and spatial heterogeneity via a reaction-diffusion framework. Hao et al. [9] considered commercial sexual activity and random fluctuations using an Ornstein–Uhlenbeck process. These studies highlight the growing complexity and realism in HIV models.

Thus, we believe that greater insights into practically used epidemiological metrics for HIV/AIDS are vital. Particularly, it is noteworthy that in many high-income countries, the rate of clinical AIDS on diagnosis has been measured routinely. It is calculated as the proportion of newly diagnosed AIDS cases out of all newly diagnosed individuals with HIV. In Japan, the nationwide value of the rate of AIDS on diagnosis was above 40% before 2000, but has gradually decreased to approximately 30% during the last decade. In Tokyo and Osaka, where diagnostic performance is known to outperform that in other prefectures [10], the rate is smaller than Japan as a whole (e.g., approximately 20% and 25%, respectively, for2015–2024). In other prefectures, the rate has been between 35% and 40%. It is implicitly clear that the proportion should be small for diagnosed individuals to have a better clinical prognosis; however, there has been no indication of whether the control of HIV/AIDS in the present day is sufficiently good in relation to this metric.

The purpose of the present study is to formulate the rate of clinical AIDS on diagnosis using a mathematical model and offer an interpretation using the hazard rate of diagnosis among previously undiagnosed HIV-infected individuals. We contrast the rate of diagnosis with the effective reproduction number so that we can understand if the control of HIV currently in place is sufficiently effective.

2. Materials and methods

2.1. Epidemiological data

According to the Infectious Disease Control Law, HIV/AIDS is designated as a category V notifiable disease in Japan. Accordingly, it is the legally mandatory practice of diagnosing physicians to report all diagnosed infections to the government via local public health centers. For the exposition of our theory, we analyzed published surveillance data up to the end of 2023, which are openly announced by the Committee of AIDS Trends at the Ministry of Health, Labor and Welfare. Every 3 months, newly diagnosed HIV infections and AIDS cases are reported by region and nationality (i.e., Japanese or non-Japanese). We analyzed data from Japanese nationals because analyzing data from non-Japanese nationals would have required us to account for mobility. By the end of 2023, among both Japanese and non-Japanese nationals, there was a cumulative total of 24,532 HIV infections and 10,849 AIDS cases.

HIV infection was diagnosed using the antibody screening method (i.e., enzyme-linked immunosorbent assay, particle aggregation, and immunochromatography) followed by antibody confirmatory testing (i.e., western blot method and immunofluorescence assay) or antigen detection testing, including virus isolation and polymerase chain reaction. AIDS diagnosis was completed when individuals met the clinical diagnostic criteria (i.e., confirmed HIV infection and the presence of one of 23 indicator diseases representing opportunistic infections or tumors). From the surveillance records, we analyzed the yearly incidence dataset of HIV diagnoses and AIDS cases from 2008 to 2023.

2.2. Extended backcalculation model

First, we describe the competing risk model of HIV infection that we formulated in a previous study [11]. The compartmental illustration is provided in Figure 1. As a function of calendar time t, new HIV infections occur at the rate $\lambda(t)$. Then the rate of diagnosing HIV infections $\alpha(t)$ and hazard rate of developing AIDS $\rho(s)$ compete to absorb $h(t, s)$, which is the number of undiagnosed HIV infections at calendar time t, where s denotes the time since infection (i.e., infection age), in an independent manner. This function describes the distribution of undiagnosed HIV infections over time and infection age, and serves as the core state variable of our extended backcalculation model. We note that the rate of diagnosis is a function of calendar time, whereas the hazard of AIDS depends on infection age, thereby reflecting the incubation period of AIDS. This process is described by the following McKendrick partial differential equation:

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial s}\right)h(t,s)=-\big(\alpha(t)+\rho(s)\big)h(t,s), \tag{1}$$

with the boundary condition $h(t,0) = \lambda(t)$. See Table 1 for detailed descriptions of time-dependent variables and functions. Equation (1) is the first-order linear partial differential equation, and its solution can be obtained analytically.

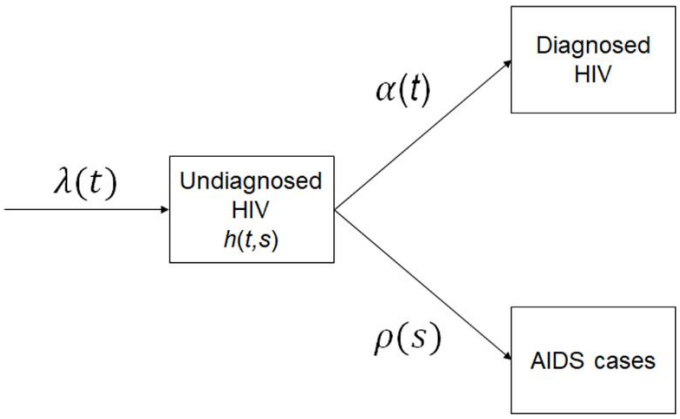

Figure 1. Competing risks of undiagnosed HIV infections to allow an explicit quantification of the system.

It is well known that Equation (1) can be integrated along the characteristic line [12]. Along the characteristic line defined by $ds/dt = 1$, we integrate the cumulative effects of diagnosis and AIDS onset over time. Together with the boundary condition. we obtain

$$h(t,s) = \lambda(t-s)exp\left(-\int_{t-s}^{t}\alpha(y)dy - \int_{0}^{s}\rho(x)dx\right), \tag{2}$$

for t-s > 0. Clearly, the incidence of HIV diagnoses and newly developed AIDS cases can be written as

$$b(t) = \alpha(t)\int_{0}^{t}h(t,y)dy \tag{3}$$

and

$$a(t) = \int_{0}^{t}\rho(x)h(t,x)dx \tag{4}$$

respectively. See Table 1 for the descriptions of $a(t)$ and $b(t)$.

2.3. Clinical AIDS on diagnosis

The so-called clinical AIDS on diagnosis $Q(t)$ represents the proportion (or relative size) of $a(t)$ out of all new HIV diagnoses at calendar time t, $a(t)$+$b(t)$:

$$Q(t) = \frac{a(t)}{a(t)+b(t)} = \frac{\int_0^t \rho(x)h(t,x)dx}{\int_0^t \rho(x)h(t,x)dx + \alpha(t)\int_0^t h(t,y)dy}. \tag{5}$$

The quantity $Q(t)$ has been monitored throughout the course of surveillance, but it has not had an explicit epidemiological interpretation.

The hazard rate of developing AIDS solely depends on infection age s. Thus, $\alpha(t)$ is the only

potential metric for yielding time-dependent transmission dynamics. Therefore, we rearrange Equation (5) with respect to $\alpha(t)$. That is, multiplying both sides of Equation (5) by the right-hand side denominator and isolating $\alpha(t)$, we obtain

$$\alpha(t) = \frac{1 - Q(t)}{Q(t)} \frac{\int_0^t \rho(x) h(t,x) dx}{\int_0^t h(t,y) dy}, \tag{6}$$

which links the observable proportion $Q(t)$ to the diagnosis rate $\alpha(t)$. The first division of the right-hand side of Equation (6) is the odds of AIDS on diagnosis, whereas the second division is the ratio of AIDS incidence to the prevalence of undiagnosed HIV infections. If we assume that the incubation period of AIDS is exponentially distributed with the mean $1/\rho_0$ (e.g., $1/\rho_0$ = 10 years in the case in which we assume an average of 10 years since infection to develop AIDS), Equation (6) is further simplified to

$$\alpha(t) = \rho_0 \frac{1 - Q(t)}{Q(t)}. \tag{7}$$

Thus, the hazard rate of HIV diagnosis before developing AIDS can be derived using the odds of the rate of clinical AIDS on diagnosis and a constant hazard rate of developing AIDS. Additionally, if $\alpha(t)$ is independent of time and a constant rate α_0, then

$$Q(t) = \frac{\rho_0}{\rho_0 + \alpha_0}. \tag{8}$$

Given $Q(t)$ = 0.3 and ρ_0 = 0.1 per year, α_0 has to be 0.233. We argue that, rather than continuously monitoring $Q(t)$ in epidemiological surveillance, it may be more insightful if we monitor the time-dependent diagnosis of HIV infections $\alpha(t)$, for example, using Equation (7). This is also because understanding $\alpha(t)$ would lead to an explicit interpretation of the control of HIV/AIDS (see the next subsection).

To better understand how $\alpha(t)$ behaves along with time-dependent changes in $Q(t)$, we numerically illustrated $Q(t)$ from epidemiological surveillance data for Japan from 1985 to 2024 and computed a simplified version of $\alpha(t)$, that is, the value calculated using Equation (7).

Table 1. Model variables and parameters for the diagnostic process of HIV infection.

Parameter / Variable	Description
$\lambda(t)$	Incidence rate of new HIV infections
$\alpha(t)$	Hazard rate of HIV diagnosis among previously undiagnosed individuals
$\rho(s)$	Hazard rate of AIDS onset s years after infection
$h(t,s)$	Number of undiagnosed HIV-infected individuals at time t and infection-age s
$a(t)$	Number of new AIDS cases at time t
$b(t)$	Number of newly diagnosed HIV cases (without AIDS) at time t
$Q(t)$	Proportion of clinical AIDS on diagnosis
$X(t)$	Total number of undiagnosed HIV infections at time t
$P(t)$	Proportion of diagnosed individuals among all HIV-infected persons
$P'(t)$	Alternative diagnosed proportion including AIDS cases

2.4. Goal of controlling HIV/AIDS

Presently, the control of HIV/AIDS adopts the test-and-treat strategy [13,14], in which substantial effort is spent diagnosing HIV infections at an early stage and bringing infected individuals under treatment. Expanding diagnoses not only contributes to improving clinical outcomes for infected individuals but also reduces secondary transmission from undiagnosed HIV-infected individuals, thereby helping HIV/AIDS to be gradually brought under control at a population level. As a slogan for intensifying the test-and-treat policy, the target 95-95-95 by 2030 was set by the Joint United Nations Program on HIV/AIDS (UNAIDS). The slogan aims to achieve 95% of people living with HIV/AIDS knowing their HIV status, 95% of people who know their HIV status receiving ART, and 95% of people on ART enjoying a suppressed viral load. Of these, the abovementioned model provides information on whether a country has achieved 95% diagnosis out of all HIV infections. Let $P(t)$ be the fraction of diagnosed HIV infections out of all infections:

$$P(t) = \frac{\int_0^t b(u)du}{X(t) + \int_0^t b(w)dw}, \tag{9}$$

where $X(t)$ is the total number of undiagnosed HIV infections:

$$X(t) = \int_0^t h(t,y)dy. \tag{10}$$

Sometimes, the calculation process of $P(t)$ may include AIDS cases both in the denominator and numerator, for example,

$$P'(t) = \frac{\int_0^t a(u)du + \int_0^t b(w)dw}{X(t) + \int_0^t a(w)dw + \int_0^t b(w)dw}, \tag{11}$$

which we denote by $P'(t)$ to distinguish it from $P(t)$. In the present study, we focus on $P(t)$ rather than $P'(t)$, but the same arguments that are made below can also be applied to $P'(t)$. In the public health field of HIV/AIDS, statistical estimates are routinely derived for $P(t)$ from empirical data (e.g., [10,15]) so that it can be compared with $P_0 = 0.95$, for instance.

To allow a practical interpretation of $P(t)$, we consider a specific hypothetical scenario in which the rate of HIV diagnosis $\alpha(t)$ is a constant α_0. The assumption slightly simplifies Equation (9) to

$$P(t) = \frac{\alpha_0 \int_0^t X(u)du}{X(t) + \alpha_0 \int_0^t X(w)dw}. \tag{12}$$

To connect α_0 to the transmission dynamics, we assume that the prevalence of undiagnosed HIV infections ($X(t)$) is in a linearly (i.e., exponentially) increasing state, that is, $X_0 \exp(rt)$, where X_0 is the initial value of $X(t)$ and r is the exponential growth rate of HIV infections, which satisfies the following relationship between r and R [16]:

$$R = \frac{1}{\int_0^\infty \exp(-rz) f(z)dz}. \tag{13}$$

where R is the reproduction number of HIV and $f(z)$ is the probability density of the generation time of HIV infection. Simplifying (12), we obtain

$$P(t) = \frac{\alpha_0 \int_0^t X_0 \exp(ru)du}{X_0 \exp(rt) + \alpha_0 \int_0^t X_0 \exp(rw)dw} = \frac{\frac{\alpha_0}{r}[\exp(rt) - 1]}{\exp(rt) + \frac{\alpha_0}{r}[\exp(rt) - 1]}. \quad (14)$$

For $P(t) > 0.95 = P_0$, we need

$$\frac{\frac{\alpha_0}{r}[\exp(rt) - 1]}{\exp(rt) + \frac{\alpha_0}{r}[\exp(rt) - 1]} > P_0 \quad (15)$$

to be satisfied. The inequality (15) can be rearranged as

$$\alpha_0 > \frac{rP_0}{1 - P_0} \frac{\exp(rt)}{\exp(rt) - 1}. \quad (16)$$

For the exposition of our theoretical arguments, we modeled α_0 as a function of the reproduction number of HIV/AIDS. Because our exercise was intended to offer an illustration, we simplified Equation (13), assuming that the generation time is a constant with mean T_g = 10 years, so that $R = \exp(rT_g)$. For the range of R from 1 to 5, we calculated the minimum rate of diagnosis α_0 to ensure that the diagnosed proportion P_0 exceeded 90% and 95%.

3. Results

Figure 2A shows the long-term trend of $Q(t)$, which is the percentage of clinical AIDS on diagnosis in Japan. During the relatively early period of the HIV/AIDS epidemic (e.g., 1980s and 1990s), the values of $Q(t)$ tended to be higher than in the present day and demonstrated substantial fluctuations. This implied that the diagnostic scenario was unstable, and the epidemic dynamics, in addition to the clinical prognosis of individual cases, drastically varied with time. Despite this, $Q(t)$ was gradually brought to approximately 30% in the present century, and in recent decades, $Q(t)$ has tended to be maintained at approximately 30%. Figure 2B shows $\alpha(t)$ values converted from the value of $Q(t)$, which represents the rate of diagnosing HIV infections among previously undiagnosed infected individuals. A simple interpretation of α = 0.25 is that the yearly probability of diagnosis is approximately 25% and it may take 1/0.25 = 4 years on average for newly infected individuals to be diagnosed. As was the case for $Q(t)$, the value of $\alpha(t)$ has gradually stabilized in recent decades.

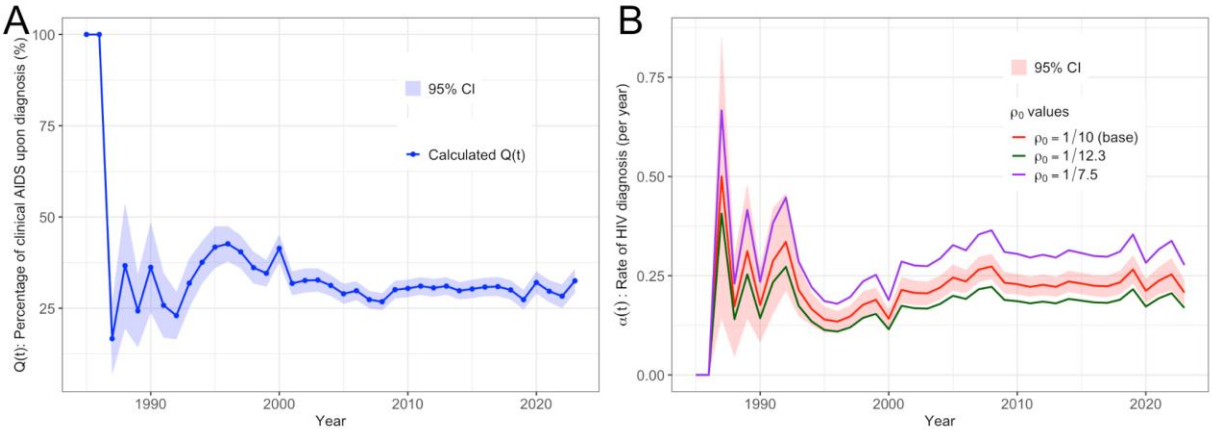

Figure 2. Epidemiological metrics of HIV/AIDS in Japan from 1985 to 2023. **A.** Percentage of clinical AIDS on diagnosis $Q(t)$. **B.** Rate of diagnosing HIV infections among previously undiagnosed infected individuals $\alpha(t)$. Data for $Q(t)$ from 1985 to 2023 were extracted from publicly available HIV/AIDS surveillance reports in Japan. The values of $\alpha(t)$ were computed using Equation (7), assuming an exponentially distributed incubation period with mean 10 years ($\rho_0 = 0.1$). The 95% confidence interval for $\alpha(t)$ was computed using the delta method, assuming a binominal variability in $Q(t)$. The standard error of $\alpha(t)$ was approximated by $SE_{\alpha(t)} = \frac{\rho_0}{Q(t)^2} SE_{Q(t)}$, where $SE_{Q(t)} = \sqrt{Q(t)(1-Q(t))/(a(t)+b(t))}$ and $\rho_0 = 1/10$. We then constructed the confidence interval as $\alpha(t) \pm 1.96SE_{\alpha(t)}$. The confidence interval should also reflect the uncertainty of $Q(t)$ by resampling $Q(t)$. In addition, we conducted a sensitivity analysis for $\alpha(t)$ using alternative values of ρ_0 ranging from 1/12.3 to 1/7.5, which correspond to mean incubation period of 12.3 years and7.5 years, respectively [17, 18].

Figure 3 shows the required minimum diagnostic effort for successfully controlling HIV, given an effective reproduction number up to 5, as computed using Equation (16). The required rate of diagnosing HIV is not very sensitive to $R(t)$ but rather highly dependent on P_0, which is the goal of the diagnosed proportion of HIV-infected individuals to be achieved. Even when the reproduction number is high, the target rate for diagnosing HIV infections does not change abruptly. During an exponentially increasing phase of the HIV epidemic, rather than $R(t)$, $\alpha(t)$ is determined by the odds of $P(t)$. Despite this, when the goal is too high (e.g., more than 80%), $a(t)$ must be elevated to an unrealistically high value. We assumed a mean generation time of 10 years [19], and we carried out sensitivity analyses using 5 and 15 years as the alternative mean generation time. For $P_0 = 0.8$, the rate of diagnosing HIV among previously undiagnosed individuals varied from 4.22 per year (when $T_g = 15$), 4.33 (when $T_g = 10$), to 4.67 (when $T_g = 5$), assuming the effective reproduction number at 5. Similarly, for $P_0 = 0.3$, the required rate of diagnosing HIV among previously undiagnosed individuals varied (i.e., 0.45 per year when $T_g = 15$, 0.46 when $T_g = 10$, and 0.50 when $T_g = 5$) under the same assumption.

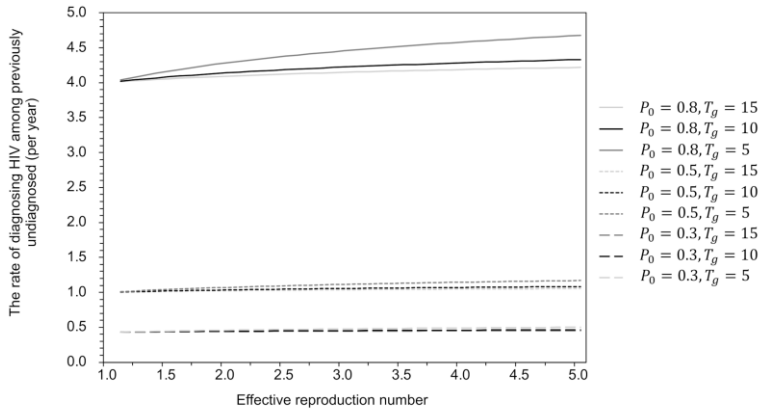

Figure 3. Relationship between the required minimum rate α_0 of diagnosing HIV among previously undiagnosed individuals and the effective reproduction number. The vertical axis represents the annual hazard rate of HIV diagnosis α_0, which is interpreted as the yearly probability of diagnosis for an undiagnosed individual. For example, $\alpha_0 = 0.25$ corresponds to an average time to diagnosis of 4 years. Extremely large values (e.g., $\alpha_0 > 10$) would imply near-immediate diagnosis and are practically unachievable. The minimum required diagnosis rate α_0 for achieving the target diagnosed proportion P_0 was computed using Equation (16). The generation time was assumed to be a constant (and thus, following a delta function). In addition to the baseline value of 10 years, sensitivity analyses were carried out assuming the mean generation time at 5 and 15 years. The number of years *t* to approximate exponential growth was set to be 1 year.

When the epidemiological dynamics are closer to equilibrium so that $X(t) = k$, where k is a constant, the resulting interpretation of α_0 changes to

$$P_0 < \frac{\alpha_0 \int_0^t k du}{k + \alpha_0 \int_0^t k dw} = \frac{\alpha_0 t}{1 + \alpha_0 t}. \tag{17}$$

Maintaining t, which can take any value, we obtain

$$\alpha_0 > \frac{1}{t} \frac{P_0}{1 - P_0}. \tag{18}$$

For instance, assuming that the prevalence of undiagnosed individuals can be approximated as a constant for $t = 1$ year, and if P_0 is as large as 0.90 or 0.95, α_0 must be greater than 9 or 19 per year, respectively, which are values that are too large to achieve. Despite this, when the epidemic is gradually brought under control and if t is greater (e.g., 10 years or 100 years), α_0 is downscaled by a factor of 10 or 100, and we obtain an α_0 that is realistically achievable. For reference, α_0=1 corresponds to an average time to diagnosis at 1 year, and α_0=0.23, which has been estimated from the surveillance data in Japan, corresponds to an average of approximately 4.3 years. Assuming that the prevalence of undiagnosed HIV infections remains approximately constant for a period of $t = 30$ years, and aiming to achieve a diagnosed proportion of $P_0 = 0.95$, Equation (18) leads to

$$\alpha_0 > \frac{1}{30} \cdot \frac{0.95}{1 - 0.95} = \frac{0.95}{1.5} \approx 0.0317.$$

That is, a diagnosis rate o over 0.03 per year would suffice to reach the 95% diagnosis goal over 30 years. This level is considerably lower than the estimate in Japan (see Figure 2B), suggesting that the goal of 95% diagnosis is practically attainable.

Equations (7)–(8) assumed that ρ_0 is a constant and thus the hazard of illness onset was exponentially distributed. To be more realistic, the incubation period distribution is right-skewed and may be better captured by employing a Weibull distribution. Using the scale parameter η and the shape parameter k and denoting the age since infection as s, the hazard rate of AIDS onset $\rho(s)$ can be expressed as

$$\rho(s) = \frac{k}{\eta}\left(\frac{s}{\eta}\right)^{k-1}. \tag{19}$$

To explore how impactful for having the realistic incubation period distribution on ב, we would also have to know the distribution of undiagnosed HIV infections, $h(t,s)$. Imposing a simplifying assumption that h is in a steady state $h_0(s)$,
Equations (3) and (4) are then

$$b(t) = \alpha_0 \int_0^t h(y)dy, \tag{20}$$

$$a(t) = \int_0^t \rho(x)h(x)dx. \tag{21}$$

Then, Q(t) is

$$Q(t) = \frac{a(t)}{a(t)+b(t)} = \frac{\int_0^t \rho(x)h(x)dx}{\int_0^t \rho(x)h(x)dx + \alpha_0 \int_0^t h(y)dy}. \tag{22}$$

Additionally, assuming that $h_0(x)$ is an exponentially decreasing function (e.g. as it may be the case for present-day Japan), it can be written as $h_0(x) = Ce^{-\beta x}$ using a constant C and the rate of decline β. Substituting this into Equation (22) leads to

$$Q(t) = \frac{\int_0^t \rho(x)Ce^{-\beta x}dx}{\int_0^t \rho(x)Ce^{-\beta x}dx + \alpha_0 \int_0^t Ce^{-\beta y}dy} = \frac{M(-\beta)}{M(-\beta) - \frac{\alpha_0 t}{\beta}}. \tag{23}$$

where $M(-\beta)$ is the moment-generating function of the incubation period hazard given the rate of decrease in prevalence β. Figure 4 compares the relationship between $Q(t)$ and α_0 for different $\rho(s)$. As shown in Figure 4C, assuming a Weibull hazard yields slightly higher values of $Q(t)$ compared to assuming a constant ρ_0.

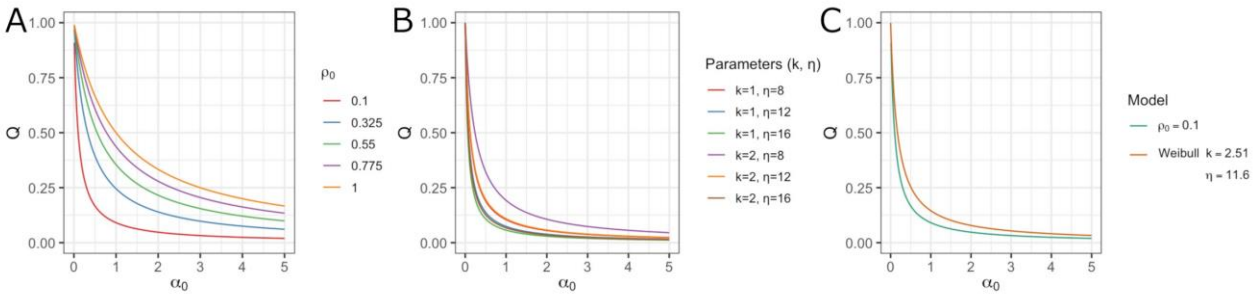

Figure 4. Variation of $Q(t)$ as a function of α_0 with different hazard of illness onset. **A.** Relationship between $Q(t)$ and α_0 while ρ is assumed as a constant. **B.** Relationship between $Q(\mathrm{t})$ and α_0 when $\rho(s)$ follows a Weibull distribution. Assuming that the mean incubation period is 10 years, we performed simulations using $\beta = 0.1$ [19]. **C.** Comparison between assuming an exponential distribution with $\rho = 0.1$ and a Weibull hazard with parameters $k = 2.41$ and $\eta = 11.59$, corresponding to the mean of 10-year.

4. Discussion

In many practical settings in Japan, the proportion of new AIDS cases out of all newly diagnosed HIV infections has been viewed as reflecting diagnostic capacity as part of HIV control. Thus, this readily calculable proportion has been monitored routinely. Despite this, there has not been an explicit interpretation of the rate of AIDS on diagnosis. Therefore, in this study, we attempted to transform the proportion to a more interpretable measure $\alpha(t)$ using a mathematical model that mimics the competing risk model. Because the rate of HIV diagnosis $\alpha(t)$ is also not as clear as the reproduction number for determining the goal of HIV control, we contrasted $\alpha(t)$ with the slogan 95-95-95 by UNAIDS using the same model. We demonstrated that the goal of achieving $P_0 = 0.90$ or 0.95 can be translated to the value of $\alpha(t)$, particularly during the exponentially increasing phase of the epidemic and endemic state of HIV/AIDS. It is possible to consider a realistic range of $\alpha(t)$ for control when the length of the endemic period continues for a few decades.

There are two important lessons to be learned from the present study. First, the rate of AIDS on diagnosis $Q(t)$ is simply a proportion of all new diagnoses. Rather than using $Q(t)$, we took the odds of $Q(t)$ and multiplied it by the inverse of the mean incubation period and obtained $\alpha(t)$, which can be used as a simple assessment of the hazard rate of HIV diagnosis among previously undiagnosed HIV-infected individuals. Practically, using $\rho_0(1 - Q(t))/Q(t)$ as a monitoring metric in addition to $Q(t)$ is our suggestion for epidemiological monitoring. In Japan, the proportion of people diagnosed with clinical AIDS on diagnosis has remained at approximately 30% since 2021, as shown in Figure 2A. If we use $Q(t) = 0.3$ and $\rho = 1/10$, then we obtain approximately $\alpha = 0.23$. This result is close to the estimate α =16.5% (95% CI: 14.9%, 18.1%) by Nishiura et al. up to the end of 2022 [15], supporting the validity of our discussion in the present study with reference to the surveillance data in Japan. Moreover, we demonstrated that α is associated with the goal of the diagnosed proportion out of all people living with HIV. Using (18) and dividing the odds of P_0 by the number of years that can be regarded as approximately stable, we can explicitly identify the goal of $\alpha(t)$ according to P_0 (e.g., for $t = 30$ years, $\alpha(t) > 0.3$ per year would be the goal to maintain $P_0 = 0.9$).

The theoretical advancement in HIV/AIDS modeling in this study is as follows: Using the formulated extended backcalculation model, we demonstrated that not only $P(t)$ but also $Q(t)$, which

is a more frequently discussed metric in the empirical setting, can be formulated explicitly. With such a formulation and parametric models for each rate function, we derived simplified epidemiological metrics for monitoring equations such as Equations (7) and (18). In both equations, $\alpha(t)$ acts as a key parameter (or a function) that determines the required effort for diagnosing HIV infections prior to the illness onset of AIDS.

Recent studies have explored HIV/AIDS dynamics incorporating stochasticity, media effects, and behavioral heterogeneity [7–9], but few studies have focused on the real-time interpretability of diagnostic rate, such as $\alpha(t)$. Thus, provided that $\alpha(t)$ characterizes the control effort, the question is how it compares between countries. In the United States of America, people who start treatment with a CD4 count below $200/\mu L$ or who have AIDS-related illnesses, regardless of CD4 count, are defined as having late presentation with advanced HIV disease (LPAD). In 2021, 21.6% of new HIV diagnoses were LPAD cases [20]. People who were more likely to be LPAD cases included men (adjusted prevalence ratio [aPR]: 1.22, 95% CI: 1.12–1.33), those aged 55 or older (aPR: 1.76, 95% CI: 1.62–1.92), and Black people (aPR: 1.02, 95% CI: 1.01–1.03). In Canada, the definition of late diagnosis is similar to that in the USA: it uses CD4 < $200/\mu L$ or having opportunistic infections. In 2022, there were 84 cases of AIDS (4.6%) [21]. In 2009, the late diagnosis rate was 8.8%, which is much lower than in the USA [22]. The Canadian rate was also lower than in European countries such as France (15.2%) and Italy (14.5%), and Australia (18.8%). In the UK, 1,136 out of 3,292 people diagnosed with HIV (34.5%) were categorized as late diagnosis in 2022 [23]. In Asia, a report covering 2003 to 2012 found that 72% of newly diagnosed people in the Asia-Pacific region were in advanced stages of HIV (CD4 < 200 $cells/mm^3$ or had an AIDS-defining event within 3 months of diagnosis) [24]. A large study in China using a similar definition (AIDS, WHO stage 3 or 4, or CD4 < 200 at diagnosis) showed that 34% of 528,230 people diagnosed from 2006 to 2014 were categorized as late diagnosis [25]. In Singapore, 58.4% (122 out of 209) of new diagnoses in 2023 were late-stage cases [26]. Even though definitions in Europe use higher cutoffs, such as CD4 < 350, the rate of late presentation remains high in comparison across Asia. Reports from many countries suggest that clinical AIDS on diagnosis is affected by race, geographic location, age, and gender, and early diagnosis must be encouraged [27].

In the following, we explain future work. Exploring the discussed metrics, including $\alpha(t)$, over geographic space (e.g., urban vs. remote areas) may yield practically identifiable differences in diagnostic capacities. Such a study could potentially explain possible differential diagnostic efforts or key differences in access or consultation of testing among various geographic areas. Another direction of analysis is to explore the heterogeneity of $\alpha(t)$ as a function of chronological age and sex. Various diagnostic trends by age can possibly be visualized using age-dependent $\alpha(t)$, in addition to the scenario in which men that have sex with men dominate infected individuals, and the analysis of $\alpha(t)$ among women to provide an indication of the epidemiological scenario of HIV/AIDS via heterosexual intercourse. Moreover, incorporating detailed demographic dynamics in the formulation of an extended backcalculation model would help to quantify the model system.

In Japan, the number of AIDS patients has restarted to increase for the first time in the last four years [28]. This may be because of fewer opportunities to visit clinics during the COVID-19 pandemic and the heavy burden on public health centers. It is important to continue monitoring the number of AIDS patients, the rate of AIDS on diagnosis, and the number of new HIV diagnoses each year.

The present study had three limitations. First, the model that we presented is a single population model without heterogeneity. We intend to explore $\alpha(t)$ over age and geographic space in the future.

Second, we modeled $\alpha(t)$ as solely time-dependent and not independent of infection age. The analytical expression of the proposed metric would be further complicated with such a dependence structure. Third, our model does not explicitly describe the secondary transmission process as part of the formulation. Thus, elevating diagnostic effort did not influence the overall transmission dynamics. In reality, herd immunity is partly attained by diagnosing HIV infections early and another formulation that connects $\alpha(t)$ with $R(t)$ would be required in the future.

5. Conclusions

Despite a number of limitations, we believe that we have successfully formulated the rate of clinical AIDS on diagnosis by translating it to $\alpha(t)$ (i.e., the rate of HIV diagnosis among previously undiagnosed HIV-infected individuals) and demonstrating that $\rho_0(1-Q(t))/Q(t)$ can lead to $\alpha(t)$ and act as a potential monitoring metric. $\alpha(t)$ can also be associated with the proportion of diagnosed HIV infections, $P(t)$, in line with the UNAIDS slogan. Finally, a realistically achievable range of $\alpha(t)$ should be discussed as part of HIV/AIDS control.

Acknowledgments

S. F. was supported by JST SPRING (grant number JPMJSP2110). H. N. received funding from Health and Labour Sciences Research Grants (grant numbers 20CA2024, 21HB1002, 21HA2016, 23HA2005and 24HB1001)), the Japan Agency for Medical Research and Development (grant numbers JP23fk0108612 and JP23fk0108685), JSPS KAKENHI (grant numbers 21H03198 and 22K19670), the Environment Research and Technology Development Fund (grant number JPMEERF20S11804) of the Environmental Restoration and Conservation Agency of Japan, the Daikin GAP Fund of Kyoto University, Japan Science and Technology Agency SICORP program (grant numbers JPMJSC20U3 and JPMJSC2105), the CREST program (grant number JPMJCR24Q3), and the RISTEX program for Science, Technology, and Innovation Policy (grant number JPMJRS22B4). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. J. Wallinga, P. Teunis, Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures, *Am. J. Epidemiol*, 160 (2004), 509–516. https://doi.org/10.1093/aje/kwh255

2. P. Nouvellet, S. Bhatia, A. Cori, K. E. C. Ainslie, M. Baguelin, S. Bhatt, et al., Reduction in mobility and COVID-19 transmission, *Nat. Commun*, 12 (2021), 1090. https://doi.org/10.1038/s41467-021-21358-2
3. S. M. Jung, A. Endo, A. R. Akhmetzhanov, H. Nishiura, Predicting the effective reproduction number of COVID-19: Inference using human mobility, temperature, and risk awareness, *Int. J. Infect. Dis.*, **113** (2021), 47–54. https://doi.org/10.1016/j.ijid.2021.10.007
4. Y. Okada, S. Yamasaki, A. Nishida, R. Shibasaki, H. Nishiura, Night-time population consistently explains the transmission dynamics of coronavirus disease 2019 in three megacities in Japan*, Front. Public Health*, **11** (2023), 1163698. https://doi.org/10.3389/fpubh.2023.1163698
5. E. J. Amundsen, H. Stigum, J. A. Røttingen, O. O. Aalen, Definition and estimation of an actual reproduction number describing past infectious disease transmission: application to HIV epidemics among homosexual men in Denmark, Norway and Sweden, *Epidemiol. Infect.*, **132** (2004), 1139–1149. https://doi.org/10.1017/s0950268804002997
6. H. Nishiura, Correcting the actual reproduction number: A simple method to estimate R(0) from early epidemic growth data, *Int. J. Environ. Res. Public Health*, **7** (2010), 291–302. https://doi.org/10.3390/ijerph7010291
7. Y. Luo, J. Huang, Z. Teng, Q. Liu, Role of ART and PrEP treatments in a stochastic HIV/AIDS epidemic model, *Math. Comput. Simul.*, **221** (2024), 337–357. https://doi.org/10.1016/j.matcom.2024.03.010
8. X. Zhang, T. Zheng, Y. Luo, P. F. Liu, Analysis of a reaction-diffusion AIDS model with media coverage and population heterogeneity, *Electron. Res. Arch.*, **33** (2025), 513–536. https://doi.org/10.3934/era.2025024
9. Y. Hao, Y. Luo, J. Huang, L. Zhang, Z. D. Teng, Analysis of a stochastic HIV/AIDS model with commercial heterosexual activity and Ornstein–Uhlenbeck process, *Math. Comput. Simul.,* 234 (2025), 50–72. https://doi.org/10.1016/j.matcom.2025.02.020
10. H. Nishiura, S. Fujiwara, A. Imamura, T. Shirasaka, Regional variations in HIV diagnosis in Japan before and during the COVID-19 pandemic, *Infect. Dis. Model.*, **10** (2024), 40–49. https://doi.org/10.1016/j.idm.2024.08.004
11. H. Nishiura, Estimating the incidence and diagnosed proportion of HIV infections in Japan: A statistical modeling study, *PeerJ*, **7** (2019), e6275. https://doi.org/10.7717/peerj.6275
12. B. L. Keyfitz, N. Keyfitz, The McKendrick partial differential equation and its uses in epidemiology and population study, *Math. Comput. Model.,* **26** (1997), 1–9. https://doi.org/10.1016/S0895-7177(97)00165-9
13. R. M. Granich, C. F. Gilks, C. Dye, K. M. De Cock, B. G. Williams, Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: A mathematical model, *Lancet*, **373** (2009), 48–57. https://doi.org/10.1016/S0140-6736(08)61697-9
14. K. Nah, H. Nishiura, N. Tsuchiya, X. Sun, Y. Asai, A. Imamura, Test-and-treat approach to HIV/AIDS: a primer for mathematical modeling, *Theor. Biol. Med. Model.,* **14** (2017), 16. https://doi.org/10.1186/s12976-017-0062-9
15. H. Nishiura, S. Fujiwara, A. Imamura, T. Shirasaka, HIV incidence before and during the COVID-19 pandemic in Japan, *Math. Biosci. Eng.,* **21** (2024), 4874–4885. https://doi.org/10.3934/mbe.2024215
16. J. Wallinga, M. Lipsitch, How generation intervals shape the relationship between growth rates and reproductive numbers, *Proc. Biol. Sci.,* **274** (2007), 599–604.

https://doi.org/10.1098/rspb.2006.3754
17. J. L. Boldson, J. L. Jensen, J. Sogaard, M. Sorensen, On the incubation time distribution and the Danish AIDS data, *J. R. Stat. Soc. A*, **151** (1988), 42–43. https://doi.org/10.2307/2982183
18. A. Muñoz, J. Xu, Models for the incubation of AIDS and variations according to age and period, *Stat. Med.,* **15** (1996), 2459–2473. https://doi.org/10.1002/(SICI)1097-0258(19961130)15:22<2459::AID-SIM453>3.0.CO;2-Q
19. R. Brookmeyer, J. J. Goedert, Censoring in an epidemic with an application to haemophilia-associated AIDS, *Biometrics*, **45** (1989), 325–335. https://doi.org/10.2307/2532057
20. F. Shi, J. Zhang, S. Chen, X. Yang, Z. Li, S. Weissman, et al., Multi-level Factors Associated with HIV late presentation with advanced disease and delay time of diagnosis in South Carolina, 2005–2019, *AIDS Behav.*, **28** (2024), 3205–3216. https://doi.org/10.1007/s10461-024-04414-y
21. Z. V. Rueda, L. Arroyave, M. Herrera, A. E. Singh, S. Skinner, C. Spence, et al., Evolution and possible explanations for the trends in new HIV diagnoses in Alberta, Saskatchewan, and Manitoba, compared to the rest of Canada, 1985–2022, *J. Assoc. Med. Microbiol. Infect. Dis. Can.*, **10** (2025). https://doi.org/10.3138/jammi-2024-0026
22. H. I. Hall, J. Halverson, D. P. Wilson, B. Suligoi, M. Diez, S. Le Vu, et al., Late diagnosis and entry to care after diagnosis of human immunodeficiency virus infection: A country comparison, *PLoS One*, **8** (2013), e77763. https://doi.org/10.1371/journal.pone.0077763
23. National AIDS Trust, UK HIV statistics, National AIDS Trust (2024). https://nat.org.uk/about-hiv/hiv-statistics/
24. S. J. Jeong, C. Italiano, R. Chaiwarith, O. T. Ng, S. Vanar, A. Jiamsakul, et al., Late Presentation into Care of HIV Disease and Its Associated Factors in Asia: Results of TAHOD, *AIDS Res. Hum. Retroviruses*, **32** (2016), 255–261. https://doi.org/10.1089/AID.2015.0058
25. H. Tang, Y. Mao, W. Tang, J. Han, J. Xu, J. Li, Late for testing, early for antiretroviral therapy, less likely to die: Results from a large HIV cohort study in China, 2006–2014, *BMC Infect. Dis.,* **18** (2018), 272. https://doi.org/10.1186/s12879-018-3158-x
26. Ministry of Health Singapore, Update on the HIV/AIDS situation in Singapore, 2023 (January–October 2023), Ministry of Health, (2024). https://www.moh.gov.sg/newsroom/update-on-hiv-aids-situation-in-singapore-2023-%28december-2023%29
27. C. S. Wong, L. Wei, Y. S. Kim, HIV Late Presenters in Asia: Management and Public Health Challenges, *AIDS Res. Treat.,* **2023** (2023), 9488051. https://doi.org/10.1155/2023/9488051
28. Committee of AIDS Trends, Ministry of Health, Labour and Welfare, Quarterly report 2024, AIDS Prevention Information Network (2024). https://api-net.jfap.or.jp/status/japan/index.html

AIMS Press