Jpn. J. Infect. Dis., 56, 2003

Laboratory and Epidemiology Communications

Arithmetic Estimation of the HIV-Infected in Japan

Hiroshi Yoshikura*

National Institute of Infectious Diseases, Tokyo 162-8640

Communicated by Hiroshi Yoshikura

(Accepted August 7, 2003)

How many people are HIV-infected in Japan? This simple question is difficult to answer. HIV infection in its early stage is usually without symptoms. Many infected people may not visit doctors until they develop AIDS.

Regular reporting of HIV/AIDS patients in Japan started in 1984. In the reporting system, individuals found to be HIV-infected for the first time are reported as "HIV-infected" and those diagnosed as having AIDS for the first time are reported as "AIDS". AIDS patients previously reported as "HIV-infected" are not counted as "AIDS" but are reported separately as "progression to AIDS". Being not obligatory, reporting of "progression" cases is irregular and unusable as a statistic (1).

*Corresponding author: Mailing address: National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan. Fax: +81-3-5285-1193, E-mail: yoshikura@nih.go.jp

The data in Table 1 extracted from the above document (1) are the only available pertinent official data. With such limitations, any estimation model has drawbacks. Therefore, even a simple arithmetic calculation may be of use if the shortcomings of such a method are well considered. I present here one such example. For this, I made several assumptions, i.e.:

1. All HIV-infected individuals visit clinics once they develop AIDS, and all of them are reported.
2. All the patients have exactly the same latency of 10 years, though the actual latency is reportedly several to 10-odd years (2).

Arithmetic calculation is as follows: In the calculation I excluded individuals with non-Japanese nationality.

Let the current HIV-infected population be Y. Then, $Y = N_9 + N_8 + N_7 \cdots + \cdots + N_i + \cdots + N_2 + N_1 + N_0$, where N_i is the number of individuals infected i years ago. Let the annual increase of N_i be constantly c, which is the case for the two parameters used below. Then, as $N_i = N_{10}[1 + c(10\text{-}i)]$,

Table 1. The data used for calculation

	"AIDS" [1)]	"HIV-infected" [1)]	HIV-positives/100,000 blood donors
1990	21	37	0.336
1991	24	69	0.359
1992	37	124	0.441
1993	58	124	0.486
1994	100	166	0.545
1995	119	166	0.730
1996	171	230	0.762
1997	182	268	0.900
1998	168	297	0.912
1999	224	424	1.042
2000	260	368	1.140
2001	245	525	1.368
2002	252	521	1.418

[1)]: The figure indicates number of reports.

$Y = N_{10}[10 + c(1 + 2 + \cdots\cdots + 10)] = N_{10}[10 + 55c]$.

Based on the two assumptions made above, the population of the HIV-infected 10 years ago, N_{10}, should be "AIDS" cases in 2002, A_{10}, which is 252, plus the population infected by HIV in 1992 and registered as "HIV-infected" in 1992 or later, H_{10}. H_{10} is, however, difficult to calculate. This is because the individuals infected in 1992 can be reported as "HIV-infected" at any time between 1992 and 2002. The number of individuals reported as "HIV-infected" in 1992, which is 124, does not represent H_{10} but a population reported as such in 1992 who had been infected in 1992 or in the preceding years but had not developed AIDS. Therefore, I make here another assumption that all the HIV-infected individuals are reported at the mid-point of latency, 5 years, i.e., H_{10} equals the "HIV-infected" in 1997, which is 268.

The co-efficient c, the annual increase rate of HIV infection, can be estimated by using either statistic, the seropositives among blood donors or the reported number of the "HIV-infected" (Table 1, second and third lines). If blood donor data are used, as HIV positives among 100,000 donors were 0.441 in 1992 and 1.418 in 2002, the estimated annual increase c is $[(1.418 - 0.441)/0.441]/10 = 0.22$. If the data of "HIV-infected" are used, as the reported HIV-infected patients were 124 in 1992 and 521 in 2002, the annual increase c is $[(521 - 124)/124]/10 = 0.32$.

Therefore, Y is $(252 + 268)(10 + 55 \times 0.22) = 11{,}492$ or $(252 + 268)(10 + 55 \times 0.32) = 14{,}352$. Namely, this simple arithmetic leads us to an estimation of 11,000 - 14,000 HIV-infected individuals (excluding those developing AIDS) in 2002 in Japan.

The following is a more formal description. Let the present number of the HIV-infected be y. Then,

$y = N_{x-1} + N_{x-2} + N_{x-3} + \cdots + N_i + \cdots + N_2 + N_1 + N_0$, (I)

where N_i is the number of HIV-infected patients i years ago, and x is latency in a year.

$N_x = A_0 + H_x$, (II)

where A_0 is the number of "AIDS" cases now and H_x the number of individuals infected x years ago and registered as "HIV-infected" in the same year or later.

$N_i = (A_0 + H_x)[1 + c(x - i)]$, (III)

where c is the annual increase of the HIV-infected population. Then,

$y = (A_0 + H_x)[x + c(1 + 2 + \ + x)] = (A_0 + H_x)[cx^2/2 + (1 + c/2)x]$. (IV)

Equation (IV) is a parabola; y increases as a quadric function of x, i.e., the longer the latency, the higher the estimated number of HIV-positives.

Equation (III) can be used to estimate the number of HIV-infections in a year. For example in 2002, by using the same parameters as above, i.e., $c = 0.22$ or 0.32, $A_0 = 252$ and $H_{10} = 268$, $(A_0 + H_{10})(1 + 10c)$ is $520 \times 3.2 = 1{,}664$ or $520 \times 4.2 = 2{,}184$. Therefore, the actual number of people infected in 2002 may be three- to four-fold higher than the reported number, 521, of the "HIV-infected" in the same year.

REFERENCES

1. Committee of AIDS Trends, Ministry of Health, Labour and Welfare (2003): Trend of AIDS incidence in Japan, Annual Report 2002 (in Japanese).
2. Braunwald, E., Fauci, A., Kasper, D., Hauser, S., Longo, D. and Jameson, J. (eds.) (2001): Harrison's Principles of Internal Medicine. 15th ed. McGraw-Hill Pub., Co., Berkshire.