REV CUBANA MED TROP 2003;55(1):30-7

FACULTAD DE MATEMÁTICA Y COMPUTACIÓN, UNIVERSIDAD DE LA HABANA
UNIVERSITÉ RÉNÉ DESCARTES, FRANCE
INSTITUTO DE MEDICINA TROPICAL "PEDRO KOURÍ"

What percentage of the Cuban HIV-AIDS Epidemic is known?

Dr. Héctor de Arazoza,[1] Dr. Rachid Lounes,[2] Dr. Jorge Pérez[3] y Dr. Thu Hoang[2]

SUMMARY

The data for the Cuban HIV-AIDS epidemic from 1986 to 2000 were presented. With the purpose of evaluating the efficiency of the HIV detection system, two methods were used to estimate the size of the HIV-infected population, backcalculation and a dynamical model. From these models it can be estimated that in the worst scenario 75 % of the HIV-infected persons are known and in the best case 87 % of the total number of persons that have been infected with HIV have been detected by the National Program. These estimates can be taken as a measure of the efficiency of the detection program for HIV-infected persons.

Subject headings: HIV; ADQUIRED IMMUNODEFICIENCY SYNDROME/diagnosis; ADQUIRED IMMUNODEFICIENCY SYNDROME/mortality; NATIONAL HEALTH PROGRAMS; CUBA.

The first AIDS case was diagnosed in Cuba in April of 1986, this signaled the start of the AIDS epidemic in the country. Some HIV seropositives had been detected at the end of 1985. Earlier the Cuban Government had started taking preventive measures to try to contain the possible outbreak of the epidemic by creating a National Program to fight HIV infection.[1-15] Among these measures were a total ban on the import of blood, and blood byproducts. Once the first cases were confirmed, a program based on the experience with other sexually transmitted diseases was started. This program had among other measures, the tracing of sexual contacts of known HIV seropositives (HIV+), to prevent the spreading of the virus. This was done by the Epidemiology Departments at all levels of the Cuban Health System through Partner Notification and Interviews. The Cuban HIV--AIDS epidemic has a mixed pattern, it started as a mainly heterosexual epidemic, later becoming mainly homobisexual and now with a very mixed pattern (Pérez *et al*. 1993 and Pérez *et al*. 1996). Our aim is not only to present the data on the epidemic but also to make predictions for HIV+ using backcalculation and a dynamical population model. All the information we present is up to December 2000, and the source is the Epidemiology Department at the Sanatorium for HIV+/AIDS in Santiago de Las Vegas, Havana, Cuba.

DESCRIPTION OF DATA

The number of AIDS cases in Cuba is 701 with 525 of males and 176 females. Of the males 74.3 % are homobisexuals (we consider the group of homobisexuals to be formed by men that have sex with men). There have been 482 deaths due to AIDS (4 died due to other causes). Through the program a total of 1 831 HIV+ individuals have

[1] Doctor en Ciencias Matemáticas. Profesor Titular. Facultad de Matemática y Computación. Universidad de La Habana, Cuba.
[2] Doctor en Ciencias Matemáticas. Maitre des Conference. Université Réné Descartes, PARIS-FRANCE.
[3] Maestro en Ciencias. Profesor Auxiliar. Instituto de Medicina Tropical "Pedro Kourí".

been found, 477 female and 1 354 males. Of the males 75.3 % are homobisexuals. 35.1 % of the HIV+ have been found through contact tracing. Table 1 gives the new HIV+/AIDS cases detected and the number of deaths due to AIDS by year.

TABLE 1. HIV+, AIDS cases and deaths due to AIDS by year Cuba: 1986-2000

Year	HIV+	AIDS	Death due to AIDS
1986	99	5	2
1987	75	11	4
1988	93	14	6
1989	121	14	5
1990	140	28	23
1991	183	37	17
1992	175	70	32
1993	102	82	59
1994	123	101	62
1995	118	110	77
1996	234	99	93
1997	363	125	99
1998	362	150	98
1999	493	176	122
2000	545	248	141

Figures 1 and 2 show the annual incidence rates respectively for AIDS cases and the detected HIV+ cases from the beginning of the epidemic. The peak observed in the HIV+ cases in the period 1990-1992 is due to the detection of a highly promiscuous group that was detected at that time. The Cuban epidemic is a small one, with a population of around 11 millions we have a cumulative incidence rate for AIDS of 63.5 per million. The main part of the epidemic is found in 5 provinces: the City of Havana with the neighboring Havana Province account for 51.1 % of the HIV+ cases, Pinar del Rio 7.6 %, Villa Clara 14.7 % and Sancti Spiritu 5.0 %. The first three are in the western part of the island, the others in the centre of the island. Figure 3 shows the HIV+ rates for the three regions of the country and the Havana area.

Figures 4 and 5 show the distribution by age groups at the time of detection and at the time of AIDS onset. The most affected age groups at the time of detection are from 15 to 35 years of age (mean 27 years old). For the AIDS cases the mean age is 29.7 years.

The epidemic started as a mainly heterosexual one but the homosexual component has increased in time. Vertical transmission has been reduced (6 cases) as all pregnant women are tested at the beginning of their pregnancy and given the option of an abortion. There are only 10 cases of infection by transfusions in Cuba, and 2 hemophiliacs, these are the result of the early measures taken by the Cuban Government. Figure 6 shows a Kaplan-Meier estimation for the survival to AIDS. The mean is 2.29 years (sd 0.13). The 95 % confidence interval for the median is [1.08, 1.33] (fig. 6).

The incubation period is also estimated by the Kaplan-Meier method, the mean is 9.2 years (sd 0.27). The 95 % confidence interval for the median is [7.58, 8.92] (fig. 7).

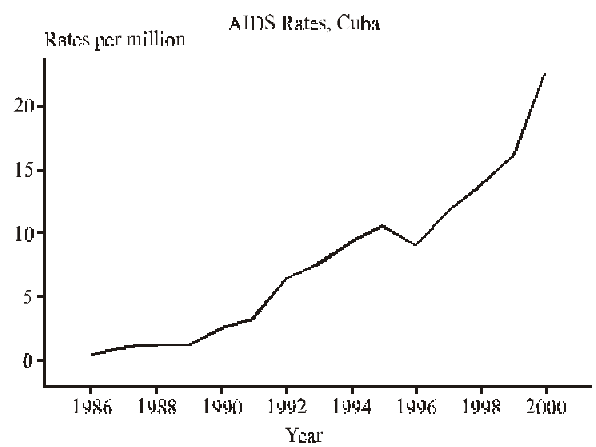


Fig. 1. Yearly incidence rates for AIDS.

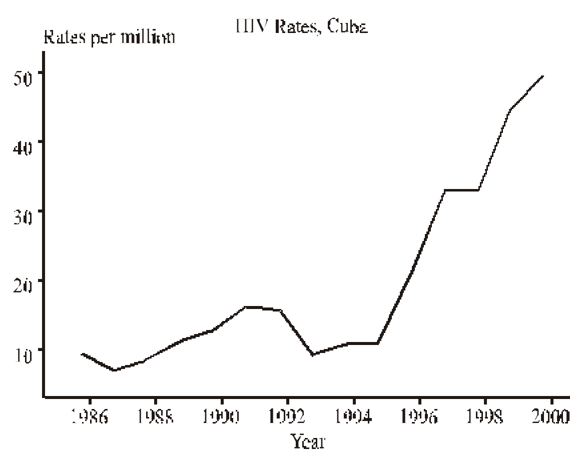


Fig. 2. Yearly incidence rates for detected HIV+.

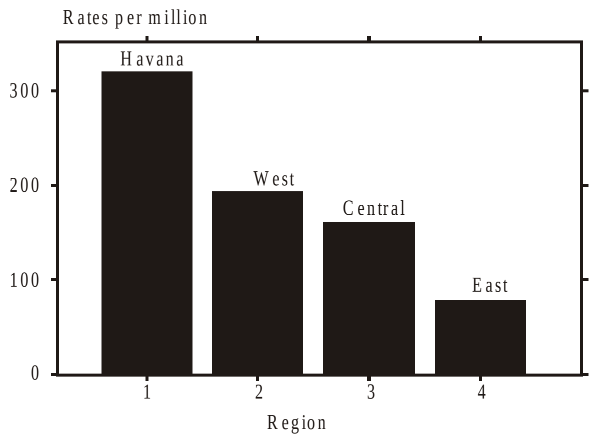


Fig. 3. Distribution by province of residence.

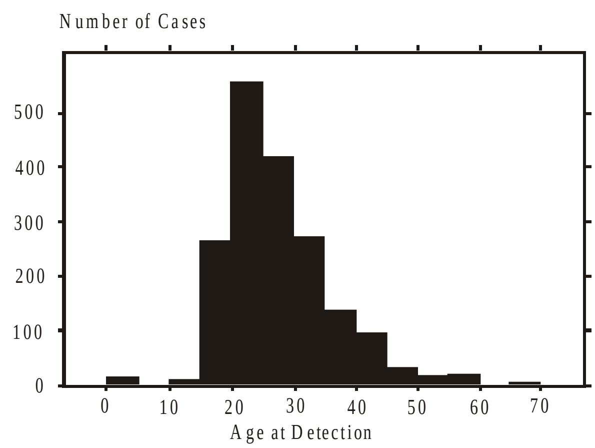


Fig. 4. Age distribution at the time of detection.

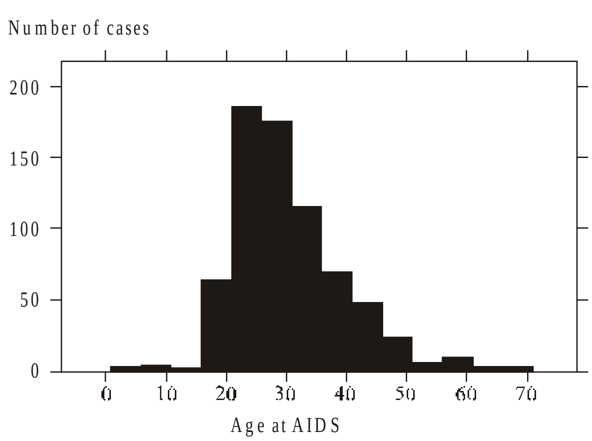


Fig. 5. Age distribution at the time of AIDS.

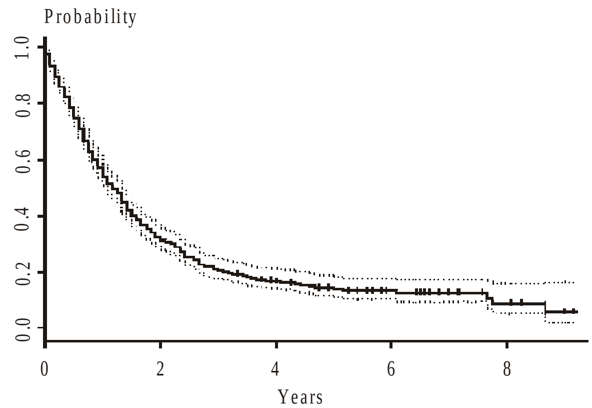


Fig. 6. K-M for Survival to AIDS.

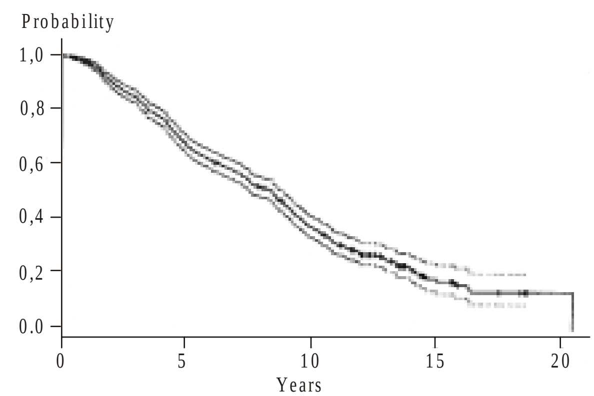


Fig. 7. K-M for Incubation period.

BACKCALCULATION

We consider the point process of new HIV+ cases and, we assume the lengths of the incubation periods (from infection to AIDS diagnosis) are independent, identically distributed random variables with probability density function f. Backcalculation is based on the underlying relation between the number of new AIDS cases in time t to $t + dt$ which we will denote $a(t)$, and the number of new HIV infections $h(s)$ at time s since the start of the epidemic at time (s=0). Let u be the time between infection and AIDS diagnosis and $f(u)$ be the probability density function of this incubation time u.

$$a(t) = \int_0^\infty h(t-u)\, f(u)\, du \qquad (*)$$

Thus if we know *h* and *f* we could find the distribution of the number of new AIDS diagnoses in any period up to time *t*. Unfortunately information on *h* is unavailable in most cases, and we can not use the equation directly. However we can obtain *a*(*t*) by adjusting a parametric model to the new AIDS cases. Using this model we can predict the future AIDS cases. On the other hand we can use equation (*) with *a* and *f* and by inverting the equation we obtain *h*, this is the backcalculation method introduced by *Brookmeyer* and *Gail* (1986, 1988). First we propose a parametric models of *a*(*t*) for the whole epidemic, and predict the number of new AIDS cases up to the year 2000. Secondly we define the probability density function of the incubation period, and finally, we use the backcalculation to predict the number of HIV+ cases for all groups up to the year 2000, and compare what is known until 1997 with the values given by the method. In this section we will use the data up to 1995 to obtain the models and parameters for the calculations.

MODEL FOR *a*(*t*) AND PREDICTION OF NEW AIDS CASES

Consider the following parametric model for the rate *a*(*t*) at which AIDS are diagnosed

$$a(t) = \exp(a_o + a_1 t - a_2 t^2)$$

We will use the data for the period 1986-1995 to obtain the parameter values. For the Cuban AIDS incidence data from 1986 to 1995 the fitted parameters are

$$a_o = 1.69,\ a_1 = 0.47,\ a_2 = 0{,}014$$

with multiple correlation coefficient for the model of 0 .97 (fig. 8)

Based on this model of the intensity of the AIDS point process *a*(*t*), we make predictions for the Cuban AIDS epidemic, by taking the value given by

$$A_i = \int_{i-1986}^{i-1985} a(t)dt$$

for each year (denoted *i*) which follows 1986 (*Isham*, 1989). Table 2 gives these values up to the year 2000.

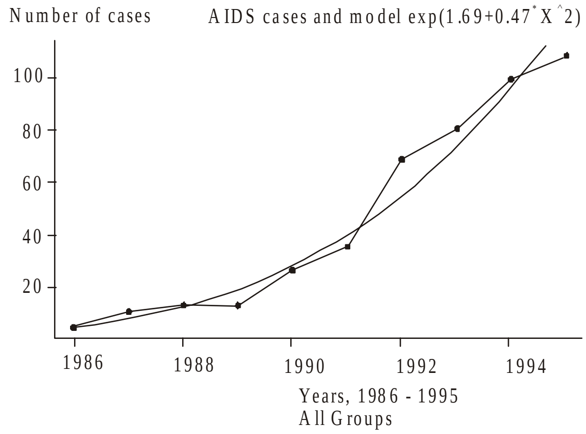


Fig. 8. AIDS Cases and exponential model.

TABLE 2. Expected Annual Incidence of AIDS Diagnoses, A_i

Year-i	All
1987	11
1988	16
1989	24
1990	34
1991	48
1992	65
1993	86
1994	110
1995	138
1996	168
1997	199
1998	230
1999	258
2000	282

ESTIMATION OF THE TIME TO AIDS CURVE

We use a Kaplan-Meier estimation for the time from infection to AIDS, using the data available until 1995. The median incubation time is 7.4 years. (95 % confidence interval [6.67, 8.42]) and the mean incubation time is 8.2 years. This is obtained from the 642 HIV+ patients with a documented probable date of infection.

We fit a gamma curve of scale parameter 1 to the Kaplan-Meier curve (table 3). When estimating the shape parameter by the inverse of the median (i.e., 0.14), we find a better fit than in the case of estimating by the inverse of the mean (i.e., 0.12).

The mean survival time to AIDS is 2.2 years (SD= 0.143 years) and the median survival time to AIDS is 1.29 years (95 % confidence: [1.07, 1.44]). The fitted curve Γ (1,0.45) with the shape parameter estimated by the inverse of the mean survival time 2.2 will be used. These curves will be used for backcalculation

TABLE 3. Coefficients for the Gamma distribution

Group	a	mean	b=a/mean	median	b=a/median
All Groups	1	8.17	0.12	7.42	0.13

BACKCALCULATION AND PREDICTION FO THE HIV EPIDEMIC IN CUBA

As in the first subsection we can compute the values for the new HIV+ cases in Cuba for each year by using

$$H_i = \int_{i-1986}^{i-1985} h(t)dt$$

as in Isham (1989). Table 4 gives these values up to the year 2000.

TABLE 4. Expected Annual Incidence of HIV Diagnoses, H_i

Year-i	H_i	Real
1986	21	99
1987	32	75
1988	48	93
1989	69	121
1990	96	140
1991	129	183
1992	169	175
1993	214	102
1994	262	123
1995	310	118
1996	355	239
1997	392	363
Subtotal	2 097	1 831
1998	417	362
1999	428	493
2000	420	545
Total	3 362	3 231

The predictions do not adjust well to what was found in Cuba each year. However if we look at the total figures up to 1997, we see that there should have been 2 097 HIV+ persons in Cuba and the Health System had found 1831. Therefore we can estimate that there were 266 unknown seropositives at that time. In another work, *Arazoza* and *Lounes* (1996) estimate the number of unknown seropositives for the Cuban HIV/AIDS epidemic, in the year 1995, to be 263 which is consistent with the new estimate. Following this model we can expect an epidemic with a total number of 3 362 HIV+ by the year 2000. The total number detected at the end of the year 2000 was 3 231. This would give us a total number of unknown HIV infected of 131. On the other hand the number of new infections detected after 1998 has increased sharply. This has been looked into by one of the authors and Dr. Y. Hsieh (*Hsieh et al.* 2001). There seems to be a change in the pattern of detection after 1997. This will be the subject of further studies.

Another way of using the back projection is by discretizing the equation

$$a(t) = \int_0^{\infty} h(t-u) f(u) du$$

as in *Day et al*. (1989), to obtain a matrix equation $H_i = F_i^{-1} A_i$ where $H_i = (H_1,...,H_i)$ and $A_i = (A_1,...,A_i)$, are, respectively, discrete versions of $a(t)$ and $h(t)$ and F_i is a triangular matrix obtained from $f(t)$. Using the observed values of A_i and Γ for the probability density function f of the incubation period time, we obtain values for H_i for the years 1986 to 1995 as in table 5.

TABLE 5. Expected Annual Incidence of HIV Diagnoses, H_i

Year-i	H_i	Real
1986	43	99
1987	57	75
1988	37	93
1989	14	121
1990	136	140
1991	106	183
1992	323	175
1993	174	102
1994	247	123
1995	179	118
1996	399	239
1997	438	363
Total	2 153	1 831

Again we can see that the values do not adjust properly by year, but the total gives us a difference of 322 unknown HIV+ cases in Cuba for the year 1995 which is in the same order as the one found above and in *Arazoza* and *Lounes* (1996). To make predictions for the years 1996-2000 we have to take for A_i the values given by the model found in the previous subsection for all groups, this gives us the following values. The total HIV+ epidemic for the period 1986-2000 is 3558 HIV+ cases (table 6). This result is similar to the one found before because the method is essentially the same, the only difference is that one is done with the exact calculation of the function $h(t)$ and the second one is by an approximation of the integral.

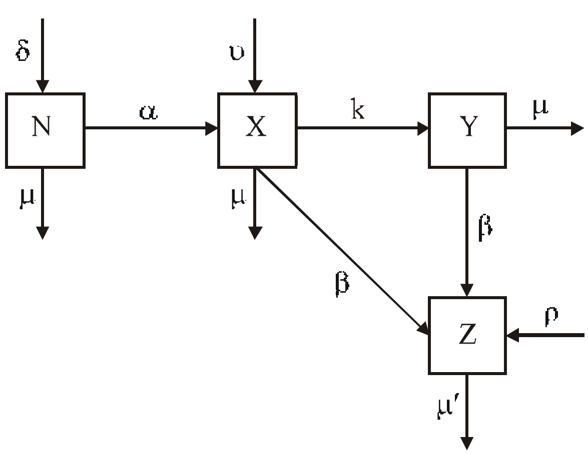
TABLE 6. Expected Annual Incidence of Future HIV Diagnoses, H_i

Year-i	H_i	Real
1998	465	362
1999	474	493
2000	465	545
Total	1 404	1 400

DYNAMICAL MODEL

We consider the following variables for the Cuban HIV epidemics:

N the population of sexually active persons or susceptibles,
X the number of HIV positives that do not know they are infected,
Y the number of HIV positives that know they are infected,
Z the number of AIDS cases.
With transitions between the variables given by the following diagram:



where α is the rate that new HIV positives are produced by the individuals in class *X*, *k* the rate at which the HIV positives are detected, β the rate at which the HIV positives develop AIDS, μ the mortality rate of the sexually active population, μ' the mortality rate of the AIDS population, δ the recruitment into the class of susceptibles, ν the immigration of HIV positives, ρ the immigration of AIDS cases.

We represent in the diagram that the infection is carried out only by the HIV positives that do not know that they are infected.

From the infinitesimal transition probabilities of this model we can obtain a linear system for the differential equations for the expectations of $X(t)$, $Y(t)$ and $Z(t)$ together with their variances and covariances. In this case a new parameter λ is considered which can be taken as, $\alpha = \frac{\lambda}{N}$. We solve these equations and fit the formula for the known HIV-seropositives [$Y(t)$] to the Cuban Data in the following way. We take the following values for the parameters that can be estimated:

Initial time will be the year the program started, 1986.

$X(0)$= 230, estimated from the number of HIV positives that were found after 1986 and were already infected at that time.

$Y(0)$= 94, number of HIV positives, that were alive at the end of 1986.

$Z(0)$= 3, number of AIDS cases that were alive at the end of 1986.

μ =0.0053, yearly mortality rate for the HIV+ cases for 1991-1997, (*SD*=0.00254),

μ' = 0.3027, yearly mortality rate for the AIDS cases for 1986-1997, (*SD*=0.07),

ν = 20, number of imported HIV-positive persons (average for 1986-1987),

β = 0.10870, obtained from the average incubation period,

ρ = 0, there are practically no imported AIDS cases.

With these values we fit the formula for the population of known HIV positives [$Y(t)$] to the values of the HIV positives known to the Cuban Health System, and we obtain the following values for the remaining parameters of the systems λ =0.5594 and *k*=0.4554. Using the formulas obtained for the variances to compute an interval for the expectation of X [$EX(t)$] obtained by the model, we find that the number of unknown HIV positives in Cuba for 1997 is within the limits (342,486).

TABLE 7. Variation of the parameters for the dynamical model

Parameter	Mean - 2SD	Mean - SD	Mean	Mean + SD	Mean + 2SD
β	0.1026	0.1052	0.1087	0.1116	0.1156
μ	0.00022	0.00276	0.0053	0.0078	0.01038
μ'	0.1627	0.2327	0.3027	0.3727	0.4427
λ	0.5353	0.5464	0.5594	0.5709	0.58494
k	0.4487	0.4517	0.4554	0.4587	0.4629
EX (1997)	394	403	414	424	435

We can also use the standard deviation of β, μ and μ' to get an alternative estimate for *EX(t)* as follows.

For the model used, the relation

$$\frac{\alpha\delta}{\mu(k+\mu+\beta)} \qquad (1)$$

is the basic reproduction number for the epidemic.

If this quantity is less than 1, the epidemic does not grow. How valid is this condition for the Cuban AIDS data? Assume that $\alpha = \frac{\lambda}{N}$, where *N* is the sexually active population of age 15 to 45 estimated to be 6 millions. Take the census value of 50 000 per year for δ,[1] the migration rate into the class of sexually active persons, 1.7 per thousand for μ, the mortality rate of the sexually active Cuban population. Then (1) is of the order of 4.7. Therefore according to the model there is an HIV-AIDS epidemic in Cuba. Furthermore the epidemic is autonomous. The future endemic level given by the model could be as high as 8 percent of the population infected with HIV or living with AIDS.

In order that the condition be satisfied, *k* should be increased 5 folds. It seems that there was no epidemic 3 years before, because at that time the value of (1) was close to 1. We may interpret this shift in 3 years as being due to the dip in the HIV+ detection rate from 1993 to 1995 (fig. 2) when difficult economic conditions impeded the Health System and the partner notification interview facilities. At that time the HIV+ not detected by the system, and not informed of their condition, may have continued to transmit the virus therefore accelerating the epidemic.

Parameter *k* plays an important role in the dynamics of the epidemic. Other parameters in the model can not be changed in a short term. On the other hand, *k* depends on the efficiency of the contact tracing system and can vary according to how well the different levels of the Cuban AIDS program do their job, $1/k$ is the mean time from infection to detection, therefore a smaller *k* means that the individuals stay more time undetected and, following the model, can infect more susceptibles, a larger *k* means that the infected persons are detected faster and can infect fewer susceptibles. The following table presents predictions based on the models for the HIV epidemic in Cuba in the year 2000 according to the values of *k* (table 8).

TABLE 8. Effect of the variation of *k*

Change in *k*	Total HIV+ in the year 2000	Change in %
0 %	1 683	
+ 10 %	1 641	– 3.5
+ 20 %	1 603	– 4.7
+ 30 %	1 568	– 6.8
+ 40 %	1 536	– 8.7
+ 50 %	1 507	– 10.4
– 30 %	1 834	+ 9
– 50 %	1 962	+ 16.5

By giving more resources to the program a 50 % increase of *k* could achieve a reduction of the epidemic of HIV by 10.4 % in the year 2000.

DISCUSSION

The Cuban HIV-AIDS epidemic is not very large when compared to other countries. There could be several reasons for this. We think that one of the elements that play a key factor on the reduced transmission is the contact tracing program developed as part of the Cuban Program to Fight the HIV-AIDS epidemic. The dynamical model presented in the last section gives evidence that a reduction of the effort in locating the new infections results in a short term increase of the epidemic, and if the efficiency in locating sexual partners of known HIV positive persons, is increased, this will lead to a reduction of the size of the epidemic.

From these models it can be estimated that in the worst case scenario, 75 % of the HIV-infected persons are known and in the best case 87 % of the total number of persons that have been infected with HIV have been detected by the National Program. These estimates can be taken as a measure of the efficiency of the detection program for HIV-infected persons.

There has been an increase in the number of cases detected after 1997. This is the reason for using the information up to 1996 for the estimates presented. The effects of the decrease in detection in the period 1993-1996, and the increase that follows will be the subject of further studies.

RESUMEN

Se presentan los datos sobre la epidemia cubana VIH-SIDA entre 1986 y 2000. Con el objetivo de evaluar la eficiencia del sistema de detección del VIH se usaron 2 métodos para calcular el tamaño de .la población infectada por el VIH, el cálculo regresivo y el modelo dinámico. A partir de estos modelos se pudo determinar que en el peor de los casos se conoce 75 % de las personas infectadas por el VIH y en el mejor, 87 % del número total de individuos que han sido infectados con VIH han sido detectados por el Programa Nacional. Estos cálculos pueden tomarse como una medida de la eficiencia del programa de detección para las personas infectadas con VIH.

DeCS: VIH; SINDROME DE INMUNODEFICIENCIA ADQUIRIDA/diagnóstico; SINDROME DE INMUNO-DEFICIENCIA ADQUIRIDA/mortalidad; PROGRAMAS NACIONALES DE SALUD; CUBA.

REFERENCES

1. Anuario Estadístico, República de Cuba, Ministerio de Salud Pública. 1996.
2. How Cuba controls HIV infection. Lancet 1991;337(8753):1340-1.
3. Arazoza H, Lounes R. A two-sex model for the AIDS-epidemic. Application to the Cuban national program on HIV-AIDS. Investigación Operacional. 1996;17(3-4):95-100.
4. Arazoza H, Lounes R, Hoang T, Interian Y. Modeling the HIV Epidemic‘under contact tracing- the Cuban case. J Theoret Med 2000;2:267-74.
5. Brookmeyer R, Gail MH. The minimum size of the AIDS epidemic in the United States. Lancet 1986:1320-2.
6. ————————. Method for obtaining short term projections and lower bounds on the size of the AIDS epidemic. J Am Statist Ass 1988;83:301-8.
7. Burr C. Havana. Assessing Cuba’s approach to contain AIDS and HIV. Lancet 1997;350(9078):647.
8. Day NE, Gore SM, McGee MA, South M. Prediction of AIDS Epidemic in the U.K.: The use of the back projection method. Phil Trans R Soc Lond 1989;325:123-34.
9. Hsieh YH, Chen CWS, Lee SM, Arazoza H. On the recent sharp increase in HIV detection in Cuba. AIDS 2001;15(3):426-8.
10. V. Isham. Estimation of incidence of HIV infection. Phil Trans R Soc Lond 1989;325: 113-21.
11. Lounes R, de Arazoza H. A two-type model for the Cuban national programme on HIV/AIDS. IMA. J Math Appl Med Biol 1999;16:143-54.
12. Pérez J, Torres R, Joanes J, Lantero M, de Arazoza H. HIV control in Cuba. Biomed Pharmacother 1991;50:216-9.
13. Pérez J, Torres R, Joanes J, de Arazoza H. in AIDS und HIV infeftionen . Ecomed 1993;1-9.
14. Swanson JM. Comprehensive care and the sanatoria: Cuba’s response to HIV/AIDS. J Assoc Nurses AIDS Care 1995;6(1):33-41.
15. Terry Molinert H, Galbán García E, Rodríguez Cruz R. Prevalence of infection with human immunodeficiency virus in Cuba. Bull Pan Am Health Organ 1989. 23(1-2): p.62-70. HIV. Lancet. 1997;350(9078):647.

Recibido: 8 de agosto de 2002. Aprobado: 12 de noviembre de 2002.

Dr. *Héctor de Arazoza*. Departamento de Ecuaciones Diferenciales. Facultad de Matemática y Computación, Universidad de La Habana, San Lázaro y L, Habana 4, Cuba. Correo electrónico: arazoza@matcom.uh.cu Tel: (53 7) 879 5771, Fax:(53 7) 863 7480