Mathematical models for the analysis of Hepatitis B and AIDS epidemics

Manuel Alfonseca, María T. Martínez-Bravo, José L. Torrea

Universidad Autónoma de Madrid

Manuel.Alfonseca@ii.uam.es, Teresa.Martinez@uam.es, Joseluis.Torrea@uam.es

Acknowledgement: *This paper has been partially sponsored by the Spanish Interdepartmental Commission of Science and Technology (CICYT), project number TIC-96-0723-C02-01, and by project number 2472DP of the Consejería de Educación y Cultura, Comunidad Autónoma de Madrid.*

Keywords: Continuous simulation, Epidemics, Hepatitis B, AIDS.

Abstract

Continuous simulation mathematical models are proposed for two different epidemic situations: Hepatitis B in a cohort of newborns followed for life, and one of the danger groups in the current AIDS epidemic. The paper describes the rational behind the systems of differential equations used to model both situations, and the way to test alternative policies, such as vaccination, preventive measures, or the effects of new drugs on AIDS.

Introduction

The use of mathematical models in Medicine goes back to Daniel Bernoulli, who in 1760 used a mathematical method to study techniques of protection against smallpox. In 1927, W.O.Kermack and A.G.McKendrick [1] used differential equations to understand a cholera epidemic. Since then, many authors have developed mathematical models of epidemics. With the emergence of AIDS, this discipline has had a renaissance.

In this paper, we will model two different epidemics using differential equations: Hepatitis B among a cohort of the total population, and AIDS restricted to a homosexual population.

To build the models, we have made use of differential equations, based on a variation of the basic Kermack and McKendrick vaccination model, where the parameters can be viewed as the mathematical expectations of random variables subject to a Poisson distribution. The systems of differential equations we have developed have been solved numerically by means of the OOCSMP language, designed by the first author [2]. The associated simulation environment makes it possible to adjust the parameters of the models with many degrees of freedom. Its use makes it very easy to try and test different scenarios and "what if" situations.

For Hepatitis B, we discuss different strategies of universal vaccination. The model is nonlinear and provides an estimate of the indirect effect of vaccination. It is a standard vaccination model, what is new is the way in which it displays clearly the *herd immunity effect*: the fact that vaccination also protects unvaccinated individuals in an indirect way, because of the smaller number of candidates to spread infection [3]. To validate the model we have used data obtained from studies of prevalence of Hepatitis B markers [4].

In the case of AIDS [5] [6] we have built a similar model adjusted to the real data, that shows the effect of behavioral prevention and predicts the effect of the new anti-AIDS drugs introduced in 1997, our main contribution to the field, compared to previous models of the same disease [7] [8]. Two factors are considered: the maximum effect of the treatment, and the rate of its spread. In this case, no validation is possible, since real data are not yet available, but the model can be used for prediction purposes [include here a note about validation by real data published after the paper was written].

Building the basic model.

Since Kermack and McKendrick established their well-known mathematical model [1] for a Cholera epidemic, the model has been used in many other situations [3]. We have used this model as starting point, therefore it is summarily explained in this section. However, we have introduced several enhancements to take in consideration the special cases of Hepatitis-B and AIDS epidemics.

In some epidemics, a closed population may be classified into three different disjoint subsets: Susceptible (S), Infectious (I), and Resistant (R). The applicable model consists of the following equations:

$$dS/dt = - r(t)\, S(t)I(t)$$
$$dI/dt = r(t)\, S(t)I(t) - ß\, I(t)$$
$$dR/dt = ß\, I(t)$$

The coefficient r(t) is a measure of the proportion of new infections arising from meetings between susceptible and infectious. This coefficient is unknown and must be estimated if the model is to describe the epidemic. In general, it is a function of time (t) and is affected by medical prevention measures, social behavior and the virulence of the infection.

The β coefficient is the proportion of infectious people that pass to the Resistant class every unit of time. This coefficient can be also viewed as the inverse of the mean of a random variable X with a Poisson distribution, that represents the time needed for an infectious person to pass to the Resistant class. Observe that, as the mathematical expectation of X is 1/β, the probability of staying in the infectious class after time t, assuming that a person has been infected at t=0, is:

$$P(X>t) = \exp(-ß\, t)$$

Therefore, the number of infectious people after time t, for an initial number I(0), will be:

$$I(t) = I(0)\ \exp(-ß\, t) + \int_{[0-t]} r(x)S(x)I(x)\exp(-ß\,(t-x))dx$$

We derivate this expresion with respect to t, and get:

$$dI/dt = -ß\, I(0)\exp(-ßt) + r(t)S(t)I(t)\exp(-ßt) - ß\exp(-ßt)\int_{[0-t]} r(x)S(x)I(x)\exp(-ßx)dx = -ß\, I(t) + r(t)S(t)I(t)$$

i.e. the second equation in the Kermack-McKendrick model.

Simulation of a Hepatitis B epidemic in Spain

The Hepatitis B virus (HBV) infects people everywhere in the world. In general, the outcome of HBV infection depends on the immune response of the individual. An adequate immune response will lead to the production of antibodies that clear the infection and give lifelong immunity. An inadequate immune response allows continued viral replication; if maintained for six months or longer, such viral production usually persists indefinitely, and the infected individual becomes an asymptomatic carrier of infection. Such carriers are at risk to develop different diseases later in life and can transmit HBV infection.

Transmission itself is via blood or blood products, and can occur during sexual intercourse, or in the sharing of needles by drug abusers.

Infection by HBV is a major public health concern in Spain. 60000 new infections are estimated to occur every year, although only about 20% are symptomatic [9]. There is serological evidence of previous contact with the virus in at least 20% of the adult population. Moreover, about half a million people are chronic virus carriers, the prevalence being especially high in young adults. The seroepidemiological surveys conducted in special collectives and in the general population indicate, in our environment, an epidemiological pattern of intermediate endemicity, similar to that of other Mediterranean countries.

Vaccination is the main method currently available to reduce the morbidity associated with HBV. In 1986, a new Hepatitis B vaccine was obtained by genetic engineering. Its efficacy (85-95%) [10] is identical to that of the previous vaccine derived from human plasma, and its side-effects and adverse reactions are mild or benign.

The aim of this section is to discuss different strategies of vaccination, such as universal immunization of teenagers and/or newborns. We first describe the natural history of Hepatitis B and deduce the appropriate equations. As population, we will consider a cohort of newborns to be followed for life, which is divided into several classes:

- Susceptible (X), containing those people who have had no contact with the virus.
- Subclinic infected (Is), those who have had contact with the virus, but don't need medical assistance.
- Clinic infected (Ic), those who do need that assistance.
- Carriers of low replication (Pb), defined by the presence of the viral marker HBsAg and the absence of virus DNA.
- Carriers of high replication (Pa), defined by the presence of the two markers HBsAg, HBeAg and virus DNA.
- Immunes (Im) defined as the class of subjects who had in the past contact with the virus and got so many antibodies that they cannot be infected. This class also includes successfully vaccinated people.

- Dead from Hepatitis B (DHB). This groups includes those who die during the infection phase, plus those who become high replication carriers and die at that stage. We assume that the second cause starts taking effect at about 30 years of age.

The following set of equations can be used to describe the course of the illness:

$$dX/dt = -\lambda(t).X.(Is/4+Ic/8+Pb+Pa/4)$$
$$dIs/dt = 0.8\ \lambda(t).X.(Is/4+Ic/8+Pb+Pa/4) - 2Is$$
$$dIc/dt = 0.2\ \lambda(t).X.(Is/4+Ic/8+Pb+Pa/4) - 2Ic$$
$$dPb/dt = 0.075\ (2Is+2Ic) + 0.1Pa - 0.02Pb$$
$$dPa/dt = 0.025\ (2Is+2Ic) + 0.01Pb - 0.1Pa - 0.05K[30...)Pa$$
$$dIm/dt = 0.9\ (2Is+2Ic) +0.02Pb$$
$$dDHB/dt= 0.001Ic + 0.05K[30...)Pa$$

where K[30...) is zero for time less than 30 years and one otherwise, representing the Hepatitis B mortality in the class of high replication carriers. More details about the medical data can be found in references [4], [11] and [12]. Table 1 summarizes those data for Catalonia (Spain).

Age(years)	Carriers(%)	Immune(%)
6-11	0.2	1.5
13-14	0.8	1.7
15-24	1.2	9.3
25-44	2.4	16.7
45-64	1.9	19.9
>64	-	23.2

Table 1: Data on Hepatitis B in Catalonia (Spain.) The age groups are heterogeneous because they represent age states (infancy, adolescent, teenagers...) The Carriers column represents the sum of both types of carriers (Pa+Pb.)

The mathematical model has been translated into the OOCSMP language, an object-oriented extension to the old CSMP III language on which we have been working for some time [2]. The OOCSMP compiler automatically generates C++ or JAVA code that includes a graphic simulation environment that makes it very easy to test different alternatives, to refine the values of the parameters, and to obtain the most appropriate approximation for the behavior of the affected populations, compatible with the existent data.

To make the results of the simulation runs agree with the data, we had to approximate a function that gives the number of infecting encounters between susceptible and infected people, which is not known directly.

We decided to define a reasonable piecewise function $\lambda(t)$ and use parameter approximation techniques to fix the turning points. The best approximation was attained with the following set of points: [(0,1), (11,1), (22,3), (35,0.3), (120, 0.3)], which correspond to a constant danger for children, a linearly growing danger for teenagers (reaching a triple rate of infection at age 22), a progressively smaller danger for young adults until 35, and a later stabilization at a risk 30 per cent of the initial risk for children. This adjustment looks reasonable in theory and similar to what can be expected from people behavior: transmission takes place mainly through sexual contacts and among drug-abusers, thus we can expect the infection to grow for teenagers and young adults, and reach a minimum for individuals 35 or older. A few infections are also vertically transmitted by the mother at childbirth.

This is a normal way to do parameter adjustments (different examples can be seen in [13].) It is possible to build an automatic parameter adjustment program that tries every combination of values of interest and compares the results to the data to be adjusted with some mathematical criterion, but in our case the number of parameters was small enough that the adjustment could be done by manual trial and error. Thus, the general shape of the piecewise function was derived theoretically, but the actual values at the turning points were manually adjusted to get the best fit.

Figure 1 shows the number of infected people in three different situations: an unprotected population, a population where all the babies are vaccinated at birth, and a third where vaccination was delayed till age 11. Roughly, in the case of no vaccination, the infected population corresponds to the derivative of the sum of both columns in table 1. It can easily be seen that this derivative reaches its maximum value at an age of about 25, in agreement with our results. It can also be seen that vaccination at 10 provides almost the same protection as vaccination at birth.

Figure 2 shows the herd immunity effect (see [3] and the explanation of the effect at the introduction of this paper) by comparing the results of the simulation of two different vaccination campaigns. In the first, 100% of the children are vaccinated at age 11, in the second only 80% of them. The figure shows the number of infected people in the cohort. It can be seen that the second option is almost as successful as the first, compared to the absence of vaccination (see figure 1), where the number of infected people is more than an order of magnitude larger. The fact that 80% of the population is vaccinated reduces the

probability of infection for unprotected people, and provides for a larger than proportional protection for the population as a whole.

Simulation of AIDS evolution in a homosexual population

AIDS is very similar to Hepatitis B in its way of transmission, which happens mainly through sexual contacts and sharing of needles among drug-abusers. We have tried to reproduce the evolution of AIDS among the homosexual risk group in Spain, using the data provided by the Spanish Ministry of Health [14], which represent the number of new AIDS positives diagnosed every year and is shown in table 2. When this paper has been written, the published data only reach to 1996.

1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
1	1	4	11	37	87	207	338	508	586	687	778	779	878	927	739

Table 2: Historical data on AIDS cases among homosexuals in Spain. The numbers correspond to people diagnosed with symptomatic AIDS, corresponding to variable A in our model.

The population under consideration may be divided into the following sections:

- Susceptible (S), containing those people in the population who have had no contact with the virus. This section increases through natural population growth (PN) and decreases by contagion (DS), going to the next two sections.
- Carriers without development (CN), for those who have had contact with the virus but never develop AIDS. Medical estimations [15] indicate that 8% of people subject to contagion may be immune and come into this group. The number of people in this group would decrease through deaths by other causes, which we have not considered, as this segment of the population is small and we are not interested in its evolution.
- Carriers with development (CY), for those who have been infected with the HIV virus but have not developed AIDS symptoms, although they will develop AIDS, unless they receive a successful preventive treatment. About 92% of people subject to contagion would come into this group, according to the same medical estimations. After a certain stay in this state, the members of this group would go to the next section of the population (develop symptomatic AIDS.)

- AIDS sick people (A), for those members of the population who have developed fully symptomatic AIDS and exhibit specific clinical features, in opposition to both types of carriers, which are asymptomatic. This group receives new arrivals from the preceding section and loses members by death.
- Dead because of AIDS (DA).

We have used the following set of equations to describe the course of the illness:

$$dS/dt = PN - DS$$
$$dCN/dt = 0.08 \cdot DS$$
$$dCY/dt = 0.92 \cdot DS - ADP3 \cdot CY$$
$$dA/dt = ADP3 \cdot CY - DAP3 \cdot A$$
$$dDA/dt = DAP3 \cdot A$$

where the following considerations have been taken into account:

- PN (natural growth of the homosexual population) has been taken as a constant and estimated at 100 per year.
- DS (contagion) has been computed using the following formula:

```
DS = TASA . PREV . S . (CN + CY + K.A)
```

where we assume that both carriers and sick people may transmit AIDS. However, most of the sick people will be under treatment and subject to special preventive measures, represented in the formula by constant K. In our simulations, we have set K to zero, indicating that contagion by AIDS sick people is negligible as compared to contagion by carriers.

In the formula, TASA is a constant of proportionality, while PREV is a term we call the "prevention factor" that embodies the application of general preventive measures against AIDS contagion. This factor decreases as the prevention measures increase. These measures depend strongly on social behavior, were practically nonexistent during the first years of the AIDS epidemic, and are widely spread now, at least in the homosexual population. We have modeled this term by means of the following equation:

$$PREV = MIN (9, PV1 + (12-PV1) \cdot EXP (-PV0 \cdot TIME))$$

Figure 3 shows the evolution of the prevention factor between 1980 and 1996. PV1 represents the minimum contagion rate attainable by current preventive measures. We have used a value about 100 times smaller than the maximum contagion rate. PV0 describes the speed of application of the preventive measures. Notice that the smallest the prevention factor, the less will AIDS be transmitted from carriers to susceptible persons.

- ADP3 (rate of pass from carrier to sick) is a constant during most of the period under study, with a value of 0.1, which corresponds to an average of 10 years delay from contagion to the development of full AIDS. However, we have made this term time dependent, substituting for it a piecewise linear function that attains a better adjustment of the model during the first 10 years (1980-1990).
- DAP3 (death rate for AIDS sick people) is a constant during most of the period under study, approximately inverse to the mean life expectancy.

Additional magnitudes of interest are:

dAD/dt = ADP3 . CY
ADiag = AD(TIME)-AD(TIME-1)

where AD is the cumulative number of diagnosed AIDS carriers, and ADiag is the number of carriers diagnosed per year. This is the magnitude that we have used to validate the model, by comparing its results to the real data shown in table 2.

Table 3 shows the results of a simulation of our model, where we have added the real data as an additional column.

YEAR	AD	DA	CY	A	ADiag	Real Data
1980	1	0	9	1	1	
1981	2	0	31	1	1	1
1982	4	1	105	3	2	1
1983	11	2	291	9	7	4
1984	29	6	671	22	18	11
1985	71	16	1319	55	42	37
1986	165	40	2260	126	94	87
1987	345	91	3442	255	180	207
1988	646	188	4749	459	301	338
1989	1098	354	6030	745	452	508
1990	1718	611	7143	1107	620	586
1991	2474	978	8019	1496	756	687
1992	3306	1452	8660	1854	832	778

1993	4193	2024	9066	2169	887	779
1994	5109	2678	9256	2431	916	878
1995	6037	3401	9263	2636	928	927
1996	6958	4174	9122	2784	921	739
1997	7859	4983	8866	2877	901	
1998	8731	5811	8527	2920	872	
1999	9566	6646	8132	2920	835	
2000	10360	7477	7702	2884	795	

Table 3. Simulation of AIDS development among the Spanish homosexual population. Each column shows the results of the execution of the model for the corresponding variable, as described in the text.

The last two columns in the table show a strong correlation. In fact, the last piece of real data (1996) is unreliable, due to delayed declaration of the illness, which means that the actual number for this year will not be known until some time in the future. Both the model and the real data show the number of new diagnosed cases reaching a maximum in 1995 and starting a slow descent, as a result of the use of preventive measures. The real data include changes in the definition for AIDS along time. These changes have not been considered in our model, which is tailored to the current definition, which may explain some of the detailed differences. For a discussion on the effects of a change of definition, see [16].

To attain these results (which can be seen in graphical form in figure 4), we have adjusted several of the parameters in the model, mainly TASA, PV1, PV0 and A2D.

Simulation of the effect of the new AIDS treatment

Very recently, some new anti-AIDS drugs have been introduced with promising results. Although they are currently under experimentation, their effect, if used during the carrier phase, seems to be a longer duration of the contagion/sickness delay, while they apparently increase the life expectancy of AIDS sick people once the illness is fully declared. It is possible that the use of these drugs may "cure" some of the sick, taking them back to the carrier phase, but this is not fully established.

Although there are no current statistical data on the effect of the drugs, which means that we cannot validate our simulations, we have included in our model the first two improvements mentioned, for prediction purposes. The longest contagion/sickness delay is represented by an adjustment to the ADP3

term. For time greater than 1996 (the new drugs were introduced in 1997), this term is replaced by the function:

1 / (10 + EMMed . (1 - EXP (EfTP . (MedD-TIME))))

where MedD corresponds to 1996, EMMed is the maximum increase (in years) of the stay expectancy in the carrier phase, and EfTP is a parameter that regulates the speed with which the new treatments will be adopted. For instance, for EMMed=10 and very large EfTP, the stay expectancy in the carrier phase would suddenly jump from 10 to 20 years in 1997, while for EfTP=0 the stay expectancy would remain at 10 years. Intermediate values of EfTP correspond to a gradual increase from 10 to 20 years (see figure 5).

In a similar way, the longest life expectancy of AIDS sick people can be represented by adjusting the DAP3 term, which, for time greater than 1996, is replaced by the function:

1 / (A2D + EMMed1 . (1 - EXP (EfTP1 . (MedD-TIME))))

where EMMed1 is the maximum increase (in years) of the life expectancy of AIDS sick people, and EfTP1 regulates the speed with which the new treatments will be adopted. A2D is the expectancy constant prior to the introduction of the new drugs.

With this modification, the results of our simulation for the last four years (1997-2000) become those in table 4 (see them also in graphical form in figure 6). At the time this paper is written, there are no real data for these years, therefore our results are a prediction. Figure 7 shows a graphic comparison between predicted and real data for new AIDS diagnostics.

YEAR	AD	DA	CY	A	ADiag
1997	7653	4742	9077	2910	695
1998	8171	5150	9110	3022	518
1999	8648	5531	9107	3118	477
2000	9110	5909	9053	3201	462

Table 4. Simulation of the effect of the new anti-AIDS drugs. This table replaces the last four columns in table 3 when the predicted effect of the new drugs is introduced in the model. The drugs were used since 1996, and their effect could not be tested until 1997.

Comparing both series of results, we can estimate the expected effect of the use of the drugs:

- The decrease on the number of new diagnosed cases will accelerate substantially.

- The number of deaths per year will decrease.
- The number of carriers will increase initially, due to the longer stay of infected people in this phase and the increase of infection due to the larger number of carriers.
- The number of AIDS sick people will also increase, due to their longer life expectancy.

Other scenarios may be tested quite easily using the OOCSMP simulation environment. For instance, we can adjust the model to the real data as soon as they are available, by changing the values of the EMMed, EfTP, EMMed1 and EfTP1 parameters.

Conclusions

The use of a set of differential equations has proved useful to model the development of different epidemics and to test the effects of different policies for prevention and treatment. Continuous simulation is a powerful tool that can be used to compare alternatives and estimate their costs and effectiveness.

The procedure has been applied to two very different situations: endemic Hepatitis B in a very special population (a cohort) and AIDS in a real situation (the homosexual danger group). In the first case, the effects of different vaccination campaigns has been simulated, with a clear demonstration of the herd immunity effect. In the latter case, a good correlation with real data has been attained. The model was easily extended to estimate the possible effects of the recently discovered drugs against AIDS.

The simulation environment we have used (the OOCSMP language and compiler) has proved very appropriate and flexible for validation and parameter adjustment.

References

[1] Kermack, W.O. and McKendrick, A.G. 1927. *A contribution to the mathematical theory of epidemics*, Proc. R. Soc. Edinburgh, Vol. 115, p. 700-721.

[2] Alfonseca, M., Pulido, E., Lara, J. and Orosco, R. 1997. *OOCSMP: An Object-Oriented Simulation Language*, Proc. 9th European Simulation Symposium ESS97, SCS Int., Erlangen, pp. 44-48.

[3] Anderson, R.M., May, R.M., 1991. *Infectious diseases of humans. Dynamics and control*. Oxford University Press, Oxford.

[4] Salleras, L., Bruguera, M., Vidal, J. et al. 1992. *Prevalence of Hepatitis B markers in the population of Catalonia (Spain)*, Eur. J. Epidemiology Vol. 8, p. 640-644.

[5] Jager, J.C. and Ruitenberg, E.J., ed. 1992. *AIDS Impact Assessment. Modelling and Scenario Analysis*. Elsevier, Amsterdam.

[6] Brookmeyer, R. and Gail, M.H. 1994. *AIDS Epidemiology: A Quantitative Approach*, Oxford University Press, New York.

[7] Dangerfield, B.C. and Roberts, C.A. 1996. *Relating a transmission model of AIDS spread to data: some international comparisons*. In: Isham, V. and Medley, G. (eds): *Models for Infectious Human Diseases: their structure and relation to data*. Cambridge University Press, pp.473-476.

[8] Jager, J.C. and Ruitenberg, E.J. 1997. *Multinational Scenario Analysis Concerning Epidemiological Social and Economic Impact of HIV/AIDS on Society*. Final Report of the EC Concerted Action, Contract Number BMH1-CT94-1723.

[9] Bruguera, M and Sanchez Tapias, J.M. 1990. *Epidemiología de la hepatitis B en España*. Medicina Clinica (Barcelona) Vol. 95, p. 470-475.

[10] Lee,C.Y. et al. 1989. *Low dose hepatitis B vaccine*. Lancet, Vol. 1989:2, p. 860-861.

[11] Torrea, J.L. and Garuz, R. 1994. *Efectividad de la vacunación universal frente al virus Hepatitis B. Simulación con un modelo matemático*, Gaceta Sanitaria Vol.8, p. 294-303.

[12] Garuz, R., Torrea, J.L., Arnal, J.M. et al. 1997. *Vaccination against hepatitis B virus in Spain: a cost-effectiveness analysis*, Vaccine Vol 15:15, p. 1652-1660.

[13] Bailey, N.T.J.; Heisterkamp, S.H. 1995. *Preliminary review on the mathematical Models and Data used for the baseline analysis 1995*, EU Concerted Action, Multinational Scenario Analysis Concerning Epidemiological Social and Economic Impacts of HIV/AIDS on Society, report BMH1-CT-941723, which is part of the Biomedical and Health Research Programme (BIOMED) of the European Union.

[14] SIDA. 1995. *Epidemiología del SIDA en España. Situación en el contexto mundial*, Ministerio de Sanidad y Consumo, Secretaría del Plan Nacional sobre el SIDA.

[15] Rowley, J.T. and Anderson, R.M. 1994. *Modeling the impact and cost-effectiveness of HIV prevention efforts*. AIDS, Vol.?8, p. 539-548.

[16] Jager, J.C., Heisterkamp, S.H. and Brookmeyer, R. 1993. *AIDS surveillance and prediction of the HIV and AIDS epidemic: methodological developments*. AIDS Vol 7:suppl 1, p. S67-S71.

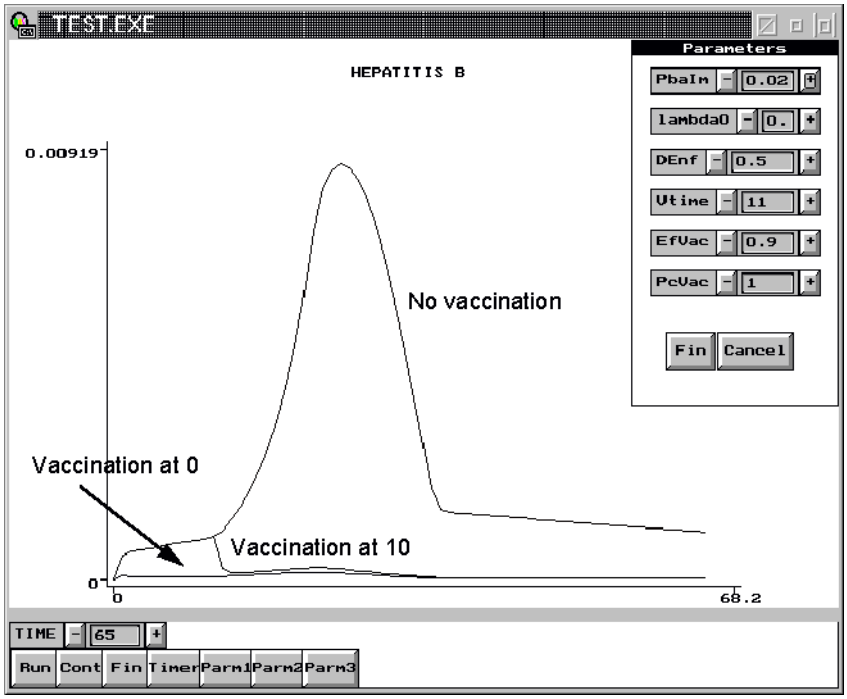

Figure 1: Three different strategies for Hepatitis B vaccination

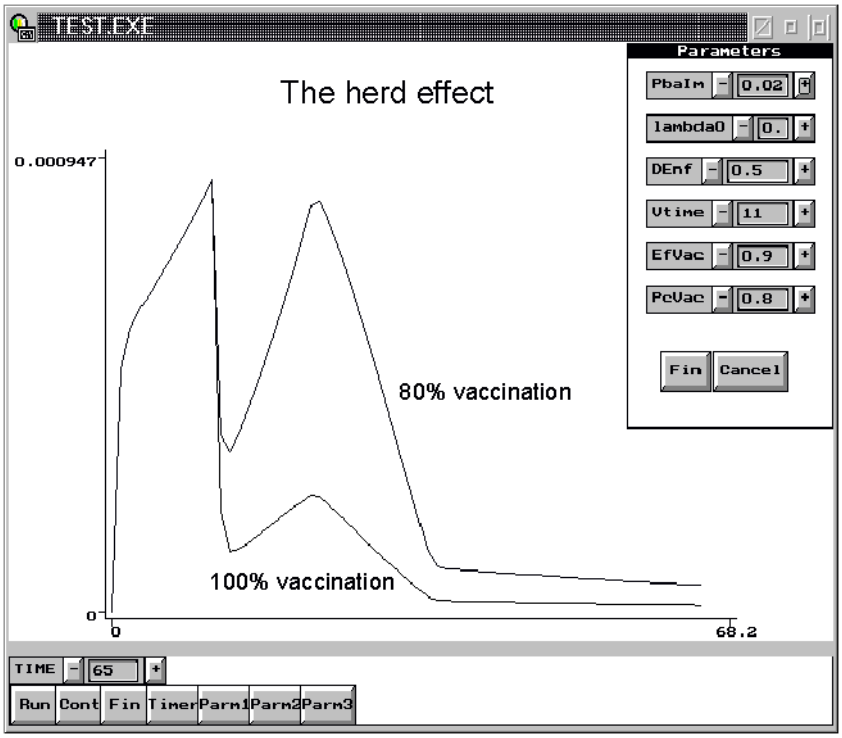

Figure 2: The herd effect in Hepatitis B vaccination

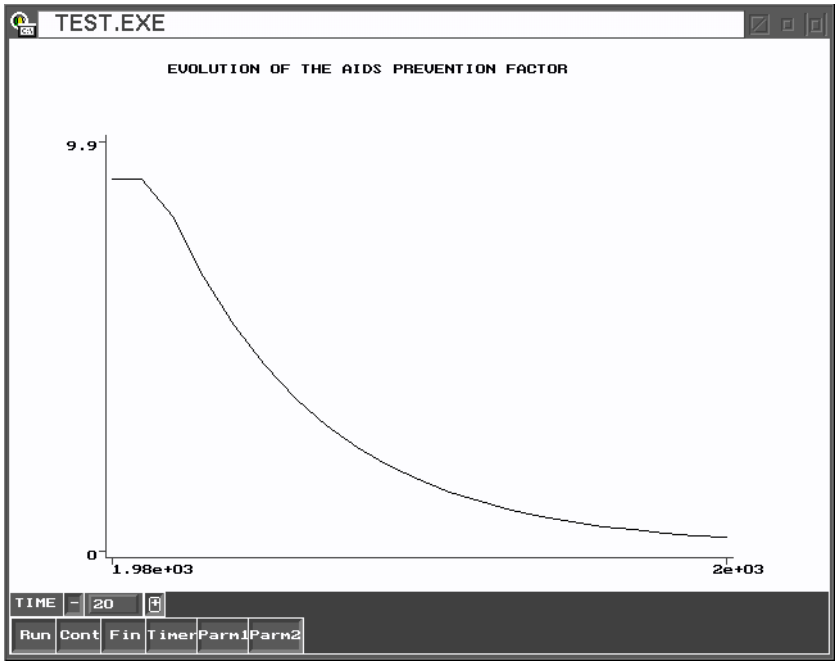

Figure 3: Evolution of the effect of AIDS prevention on homosexuals

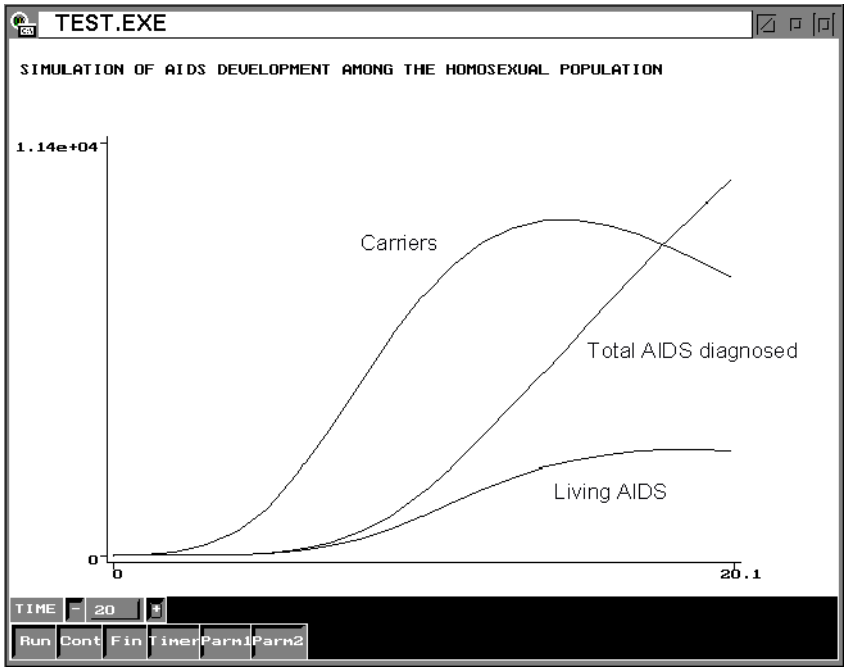

Figure 4: Simulation of the AIDS epidemic among homosexuals

Figure 5: Estimated effect of the new drugs on AIDS

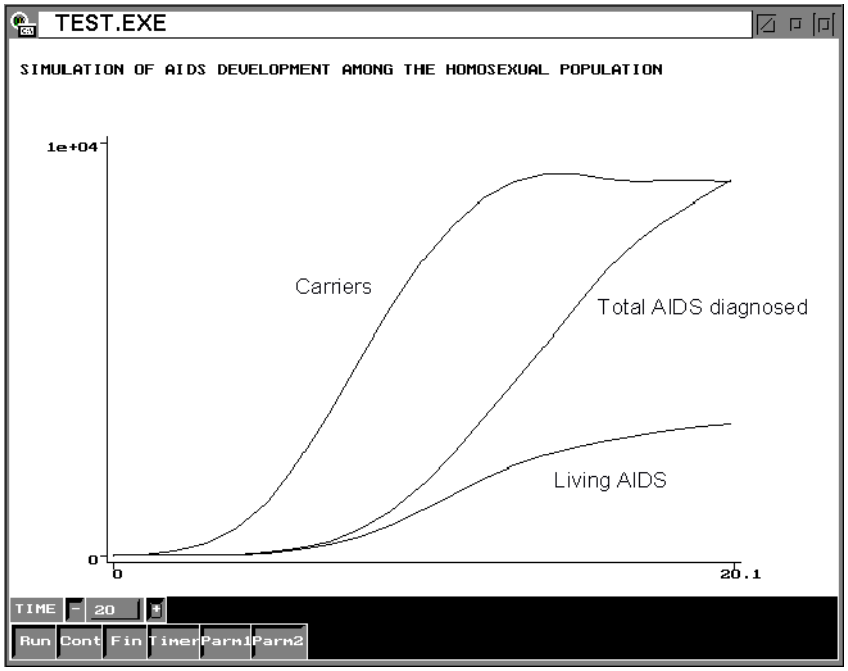

Figure 6: Prediction of the effect of the new drugs on AIDS

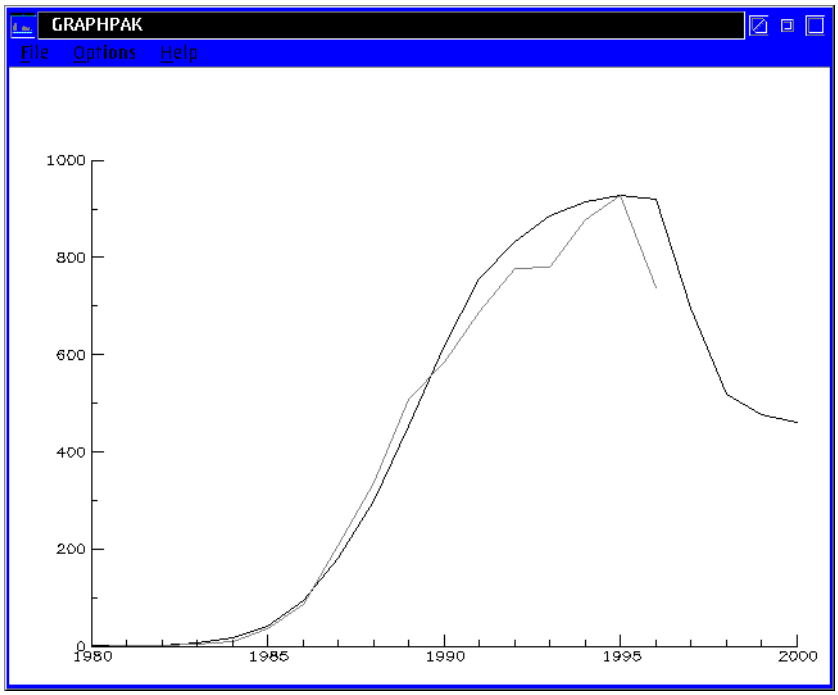

Figure 7: New AIDS diagnosed cases (simulation results vs. real data)