Monte Carlo simulation of the transmission of measles: Beyond the mass action principle

Nouredine Zekri[1,2,*] and Jean Pierre Clerc[2,†]
[1]*USTO, Département de Physique, LEPM, Boite Postale 1505 El M'Naouar, Oran, Algéria*
[2]*IUSTI, Technopôle Château Gombert, Université de Provence, Marseille, France*
(Received 13 September 2001; published 25 March 2002)

We present a Monte Carlo simulation of the transmission of measles within a population sample during its growing and equilibrium states by introducing two different vaccination schedules of one and two doses. We study the effects of the contact rate per unit time ξ as well as the initial conditions on the persistence of the disease. We found a weak effect of the initial conditions while the disease persists when ξ lies in the range $1/L-10/L$ (L being the latent period). Further comparison with existing data, prediction of future epidemics and other estimations of the vaccination efficiency are provided. Finally, we compare our approach to the models using the mass action principle in the first and another epidemic region and found the incidence independent of the number of susceptibles after the epidemic peak while it strongly fluctuates in its growing region. This method can be easily applied to other human, animal, and plant diseases and includes more complicated parameters.

DOI: 10.1103/PhysRevE.65.046108 PACS number(s): 02.50.Ng, 87.53.Wz, 02.50.−r, 89.65.−s

I. INTRODUCTION

The mathematical investigation of disease transmission was initiated by Bernoulli three centuries ago [1], but this field started to grow considerably only at the begining of the twentieth century when Hamer, in his attempt to understand the recurrence of measles epidemics, assumed that the incidence (rate of new cases) depends on the product of the densities of susceptibles and infectives [2]. This assumption is now the basis of the modern mathematical epidemiology and is known as the *mass action principle*. Epidemiological models using this principle are widely reviewed in the literature [3–5]. Basically, they formulate the flow patterns between three classes of population: the susceptibles (S), the infectives (I), and the recovered (R) that are immune either by a vaccination or by the disease itself. Some more complicated models increase the number of classes to five by including the class of passively immune individuals due to maternal antibodies (M) and that of exposed ones (E) during the latent period [4,5]. The mathematical formulation of a classical SIR (susceptibles-infectives-recovered) model yields a set of coupled first-order differential equations:

$$\frac{dI}{dt}=\beta I\frac{S}{N}-\gamma I-\mu I, \tag{1}$$

$$\frac{dS}{dt}=-\beta I\frac{S}{N}+m-\mu S, \tag{2}$$

$$\frac{dR}{dt}=\gamma I-\mu R. \tag{3}$$

Here N denotes the total population number, β the transmission rate, γ the recovery rate, μ the death rate, and m the number of births per unit time. When the population reaches its equilibrium, the rate of births becomes equal to that of deaths ($m=\mu N$). The transmission rate depends, in general, on both the infective and infected ages and becomes a continuous matrix with elements $\beta_{aa'}$ describing the probability by unit time of infecting a susceptible of age a by an infective of age a'. It is in practice impossible to solve the above equations with such a continuous matrix that is reduced in general to a 5×5 matrix within cohorts of age called ''who acquires infection from whom'' (WAIFW) [4,5]. This rate depends also on genetic and spatial heterogeneities. In the case of measles there is no genetic heterogeneity but the spatial dependence has been shown by some geometrical models such as small world networks [6] and cellular automata [7] to affect sensitively the dynamics of epidemics. Some empirical models were proposed within the framework of the SIR equations [4,5] to take into account the spatial heterogeneity either by including an N dependence of β with an exponent v or proposing more complicated algebraical expressions to β, but this rate depends in fact on a complicated distribution of the number of acquaintances between individuals [6,8]. Models based on the previous equations are widely used both to predict epidemics [9] and to optimize the vaccination schedule [10], or even to find the critical coverage to eradicate the disease [4].

However, the mass action principle neglects the fluctuations and its validity was shown recently to be limited in many cases [11]. Furthermore, since only infectives and susceptibles contribute to the incidence, the N dependence of this principle is questionable at first glance. In Eqs. (1)–(3) the S dependence is linear and is easier to handle, but this could lead to an underestimation of the data as found in New Zealand [9], where some epidemics were observed before the predicted dates. Also the rate β (which is fixed in the above equations) should change asymptotically depending on whether the incidence is much larger or much smaller than the number of receptives S.

On the other hand, the estimation of the vaccination efficiency should take into account the distribution of susceptibles among the cohorts of vaccination age. For example,

*Email address: zekri@mail.univ-usto.dz
†Email address: clair@iusti.univ-mrs.fr

if in a given population all susceptibles are within ages below that at which the vaccination holds, it will not be efficient. The number of susceptibles among cohorts of ages is impossible to measure for the whole population because they need an extensive serological investigation that is very expensive. Therefore, it seems necessary to simulate by a new approach the dynamics of the infection in order to examine the temporal behavior of the disease and also to check the cases where the mass action principle is not applicable.

This is the aim of the present paper, where we use a Monte Carlo simulation of measles propagation in a population for a period of 250 years (from 1850 to 2100) in order to take into account the growing periods in the steady state and the equilibrium (N constant). We investigate the effect of the initial conditions and the contact rate on the temporal dependence of the infection as well as the vaccination efficiency (we have introduced two different vaccination schedules). We compare our numerical data to the existing one in Oran (Algeria) and our incidence to that of Eqs. (1) and (2) based on the mass action principle.

II. DESCRIPTION OF THE METHOD

The present algorithm is inspired by that used in particle physics (GEANT) where particles are followed within different detectors to estimate their geometrical acceptance [12]. We start increasing time by steps of one day from the date 0 (i.e., January 1, 1850) towards 250 years. In each step we generate m new births and attribute to them maternal antibodies with random lifetimes following an exponentially decreasing distribution. In order to take into account the fact that only 20% remain naturally immune at 9 months while they lose all their antibodies after 15 months [4,13], the distribution of remaining maternal antibodies (P_{ab}) reads

$$P_{ab}(t)=\begin{cases}\exp(-0.345t), & t\leqslant 15 \text{ months}\\ 0 \text{ otherwise}\end{cases}. \tag{4}$$

When losing their natural antibodies they change to the susceptible class and the vector element $\mathrm{S}(a)$ of age a (in days) is incremented. In the same step we generate m times the death age a' from a distribution centered at the life expectancy and remove the corresponding susceptibles from the S vector. We consider m constant ($m=100$) in this paper, leading, in the absence of infectives, to a linear growth with time T of the number of susceptibles ($S=mT$). We assume also a delta-peak distribution of deaths at 100 yr in order to ensure a constant total population at equilibrium and show the dependence on N of the incidence in this case. This choice will affect only the equilibrium date, since the infection holds mainly below 20 years of age. After a time T, we introduce an external infective individual (coming from another sample) assumed to be at the end of its latent period that is removed after its infection period (7 days). The time T allows us to adjust the number of initial susceptibles before the infection. This infective *attacks* ξ susceptibles per day (ξ is the contact rate per unit time and per infective and corresponds to $\beta S/N$ in the above equations). The new infected persons are removed from the corresponding elements of the vector **S** and increment the number of infected I, and after the latent period ($L=7$) they become infective and increment the number of infectives I_1 and remove them from I before being removed from I_1 at the end of their infection period, so that the recovery period γ^{-1} is 14 days. During their infection period each new infective *attacks* an average of ξ susceptibles per day and so on. The age of the susceptibles to be infected is generated randomly with a probability distribution

$$P(a_i)=\begin{cases}25\%;\ a_i\leqslant 5 \text{ y}\\ 45\%;\ 5<a_i\leqslant 10 \text{ y}\\ 20\%;\ 10<a_i\leqslant 15 \text{ y}.\\ 9\%;\ 15<a_i\leqslant 20 \text{ y}\\ 1\%;\ a_i>20 \text{ y}\end{cases} \tag{5}$$

We vary this age within ± 2 years until finding a susceptible having an age in this range. If not found, we do not increment I. We neglect the spatial heterogeneity by considering an average contact rate (it is possible to use a distribution), and assume the incidence independent of the age of infectives that corresponds to the case WAIFW 3 (it is possible to include the age of infectives by using in this algorithm I and I_1 as vectors). The probability distribution $P(a)$ was chosen from the average distribution of infected persons within the same cohort of age in developing countries where the infection seems to have an average age in the scholar ages [14]. In developed countries this average age is slightly higher.

We introduce two vaccination programs in periods corresponding to the case of Algeria. The first vaccination schedule in 1970 provides one dose at 9 months and the second one in 1995 providing two doses at 9 months and 2 years (other ages of the second dose administration are discussed below to compare with the data of Algeria). Note here that the first dose fails for all persons losing their maternal antibodies after 9 months. The coverage is fixed at 80% in the rest of the paper. These parameters simulate the average situation in the countries administrating a double-dose vaccination schedule. The quantities S, I, and I_1 are recorded by steps of 14 days while $I^{-1}S^{-1}dI/dS$ (the incidence per infective and per infected) is recorded by steps of 1 day to compare our incidence with Eqs. (1) and (2). In order to show the persistence of the disease within the whole period of the simulation, we need a sufficiently high contact rate between susceptibles enabling the first epidemic peak. In a recent paper, we found within the framework of a small world network the number of connections between susceptibles for a childhood disease randomly distributed between 1 and 6 at the threshold concentration of susceptibles leading to an epidemic state [see Fig. 3(b) of [8]]. Therefore, since the infection period is 7 days, an epidemic situation holds if we choose $\xi=0.5$ corresponding to an average number of acquaintance of about 3. The value of ξ is fixed for the rest of the paper except when it is varied. This algorithm takes an average computation time of about 5 h in Pentium III 600 MHz personal computers. It is obviously possible to introduce in this algorithm complicated distributions of ξ, birth and death rates, vaccination age, and latent period, but we fix them in this paper.

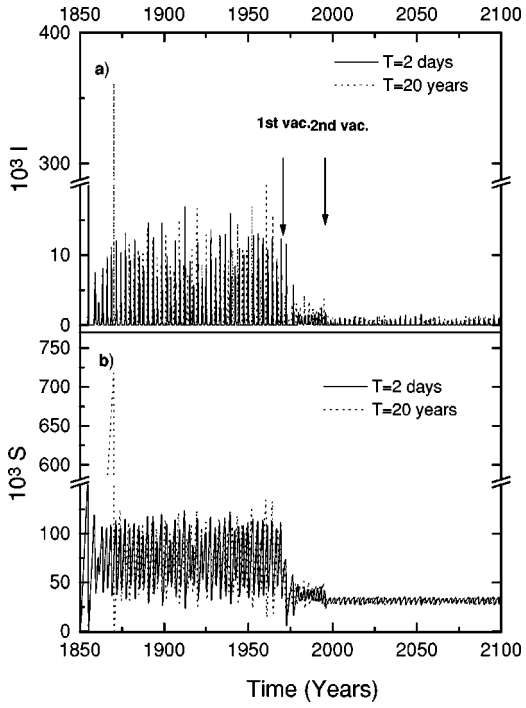

FIG. 1. Predicted temporal evolution (between 1850 and 2100) of the cases (a) and susceptibles (b) by thousands for $\xi=0.5$ and two different first infection dates: $T=2$ days and $T=20$ years. The arrows show the dates of introduction of the first and second vaccination schedules.

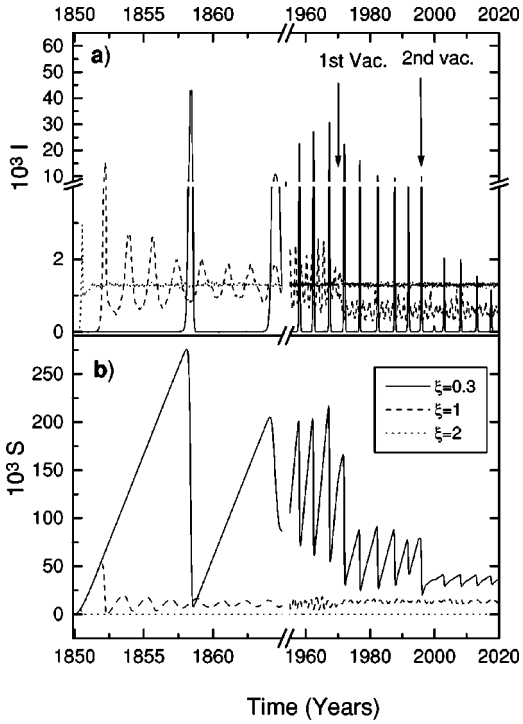

FIG. 2. Predicted temporal evolution (from 1850 to 2020) of the cases (a) and susceptibles by thousands for three different values of ξ: 0.3, 1, and 2. The arrows show the dates of introduction of the first and second vaccination schedules.

III. RESULTS AND DISCUSSION

In Fig. 1, we show the temporal dependence of the number of infected and susceptible persons for two different situations for the initial date of infection: $T=2$ days and 20 years, to compare the cases of low and high density of susceptibles with respect to the total population number. For the demographic situation described in this paper (m constant), the first case corresponds to only 100 susceptibles, while in the second one we have 730 000 susceptibles in the sample. After the first strong epidemic peak whose amplitude depends on the number of susceptibles, the number of infected (and susceptibles), which oscillates as a function of time in the steady state, shows a similar amplitude for the two initial conditions with a slight variation of the interepidemic period (due to the nonlinear aspect of the infection process), which can influence the measure of the number of predicted cases. On the other hand, since measles appeared before the 19th century the first case ($T=2$ days) appears more realistic. The persistence of the disease concerns the whole period of the simulation [in Fig. 1(a)], although we introduced a two-dose vaccination schedule. The one- and two-dose schedules seem to decrease the average numbers of susceptibles and infected peaks by about 60% and 90%, respectively. A simple estimation taking into account the coverage and the vaccination failure due to its annihilation with maternal antibodies yields rates slightly higher (64% for the one dose and 94% for two doses). In fact, this estimation assumes that the elements of the vector **S** are filled (100 susceptibles) at the age of the vaccination that is not the case here.

The effect of the contact rate ξ in the case $T=2$ days is shown in Fig. 2. This parameter seems to affect significantly both the incidence and the interepidemic period. The amplitudes of the infection peaks (and the average number of susceptibles) decrease as ξ increases. The first epidemic peak is stronger than the other ones and occurs after a period decreasing as ξ increases (the same behavior holds for the interepidemic period). Indeed, initially the number of susceptibles is much smaller than the incidence that decreases the effective contact rate $\xi_{\mathrm{eff}}=P(a)\xi$ below the recovery rate ($\xi_{\mathrm{eff}}<\gamma$), while the number of susceptibles grows linearly with time in this period. As a consequence, the propagation remains endemic until the time T_1 (at which the first epidemic peak occurs) satisfying the following equation:

$$\xi P(T_1)-2\gamma=0. \tag{6}$$

The factor 2 means that the first term concerns the infectives I_1 while the second one is the rate of all infected I that are twice I_1, it corresponds to $1/L_\gamma$ in a general disease. If T_1 is within the same class, it behaves as ξ^{-1} with a threshold $\xi_{\min}=2\gamma$. The number of susceptibles [Fig. 2(b)] reaches its maximum when the total incidence corresponds to the birth rate ($\xi_{\mathrm{eff}}PI_1=m$), while after the epidemic peaks $\xi_{\mathrm{eff}}=\xi_{\min}$ and S reaches its minimum. The interepidemic period seems to decrease slightly during the vaccination periods but the amplitude of the peaks depends on the number of susceptibles at the age of the vaccination that decreases as ξ increases. This can be seen in Fig. 2(a) for $\xi=0.3$, where both the first and second vaccination schedules decrease sensi-

TABLE I. Comparison of the simulation for $\xi=0.5$ and a vaccination at 6 years with the data of Oran (Algeria) during the period (1994–1997).

Year	1994	1995	1996	1997
Simulation data	90	773	219	816
Data of Oran [14]	86	655	179	286

tively the epidemic peaks, while if ξ increases this effect becomes weaker up to a critical contact rate where all susceptibles have smaller ages than the vaccination one. This is shown for $\xi=2$, where the first epidemic peak is as strong as all susceptibles are infected and the effective value of this rate is sufficiently large to infect all new susceptibles [see Fig. 2(a)], so that $S=0$ and the number of infected remains constant $I=m/\gamma$ during the whole period. This is the case if the minimum value of ξ_{eff} is larger than 2γ. Indeed, since the age of infection varies within a range of ± 2 years, even if there are susceptibles having only 1 day, from Eq. (4) they are infected with a probability of 10% and the corresponding effective contact rate is $\xi/10$. Consequently, the critical contact rate above which all the new susceptibles are infected is $\xi_c=20\gamma=10/L$. Note in this case that neither vaccination schedule affects the number of infected or susceptibles since all the infected persons are within their first year of age. Therefore, the range of the contact rate necessary to ensure a persistent disease propagation is

$$1/L<\xi<10/L. \tag{7}$$

In the asymptotic limit of vanishing L the disease does not persists, while if L is very large any small contact rate makes the disease persistent.

Now let us compare our results with the existing data of Oran (Algeria) in the period 1994–1997 [15]. We use the second dose at 6 years; that is the case in this country. A good agreement is shown for the period 1994–1996, while in 1997 the number of cases is significantly different (as shown in Table I). We think that the contact rate ξ used here simulates a real one but the birth rate could change leading to an increase of the interepidemic period. Indeed, the birth rate significantly decreased in this country at the end of the eighties due to different social events. The vaccinated children in 1996 were born in 1990 and their rate decreased, implying an increase of the interepidemic period that was initially about 2 years. Therefore, a more accurate simulation should take into account the real variation of the demographic situation.

The effect of the age of the second dose on the incidence of the virus is shown in Fig. 3, where we compare in the period 1990–2020 the evolution of the number of infected for two different ages (2 and 6 years). The vaccination at 2 years decreases significantly the infected peaks in comparison with those at 6 years, while it increases the interepidemic period. From this figure, we deduce a reduction of the predicted numbers of infected of about 87%, 77%, 78%, and 79% for the periods between 1995 and 2000, 2005, 2010, and 2020, respectively. We expect from 1995 to 2020 about 12 800 cases if the vaccination is at 6 years and about 2740 cases at 2 years. We expect also from the above comparison with the data in Oran an epidemic peak in this city in 2002 if the birth rate does not change.

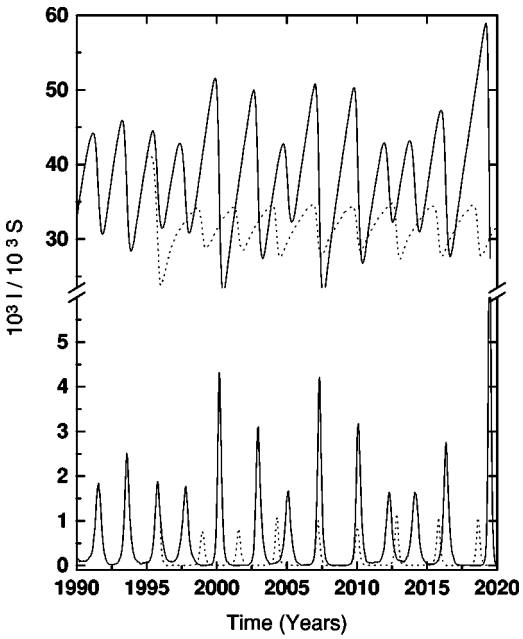

FIG. 3. Predicted temporal evolution (from 1990 to 2020) of the cases (lower curves) and susceptibles (upper curves) by thousands for $\xi=0.5$ and two different ages of administration of the second vaccination dose: 6 years (solid curves) and 2 years (dotted curves).

Let us now compare our results to those using the mass action principle in Eqs. (1) and (2). Already from Figs. 1 and 2 we deduce that both the incidence and the serological situation (number of susceptibles) does not depend on the total population number since the oscillations remain with the same amplitude and period (interepidemic period) in both the growing period (before 1950) and the equilibrium one (after this date N is constant). In Fig. 4, we show the variation of $I^{-1}S^{-1}dI/dt$ with the number of susceptibles in two different regions: the two sides of the first epidemic peak [Figs 4(a) and 4(b)] and those of another one at equilibrium [Figs. 4(c) and 4(d)]. We see in the regions where I increases [Figs. 4(a) and 4(c)] strong fluctuations of this quantity with S with a power law variation in a part of Fig. 4(c), in the case of a small number of susceptibles (due to an endemic peak). In the regions of decreasing I [Figs. 4(b) and 4(d)] this quantity decreases as a power law with S with an exponent close to unity, indicating the independence of the incidence in Eqs. (1) and (2) on the number of susceptibles. This discrepancy with the mass action principle is due to the saturation effects shown in our approach based on the contact rate ξ. Indeed, the behavior in the decreasing regions (after the peak) is due to the fact that the number of susceptibles is much smaller than the incidence ξI_1 (I_1 is large in this region) which varies in this case independently of S (the decrease is mainly governed by the recovery rate γ), while in the increasing regions, the fluctuations are due to the stochastic behavior of the distribution $P(a)$ in Eq. (4). Therefore, the expression of the incidence in the mass action principle should take into

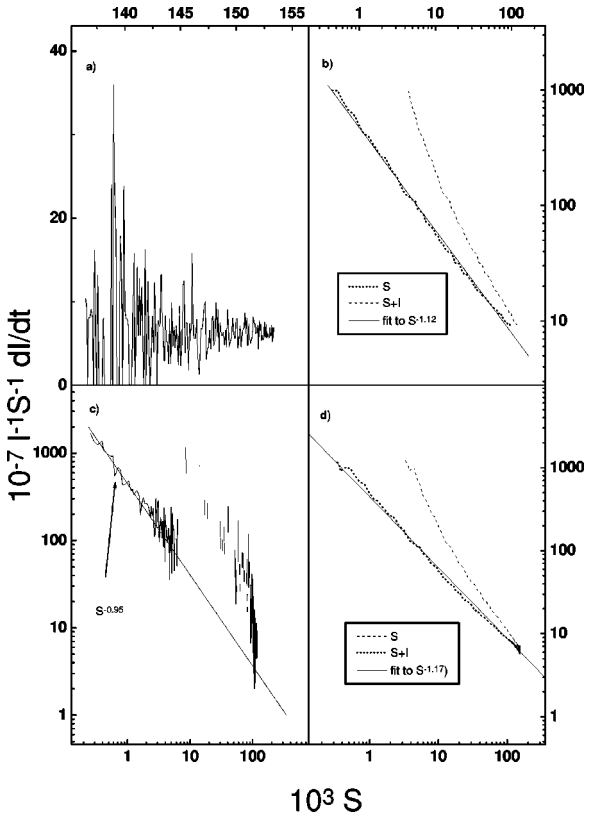

FIG. 4. Incidence normalized to the infected and susceptibles ($x10^{-7}$ per day) as a function of the number of susceptibles by thousands for $\xi=0.5$ in four different situations of the propagation of the disease: (a) before the first epidemic peak, (b) after the first epidemic peak, (c) before an epidemic peak chosen in the equilibrium, and (d) after this epidemic peak. The solid lines show linear fits.

account these asymptotic situations based on the contact rate for a realistic prediction of disease propagation.

IV. CONCLUSION

We presented a Monte Carlo simulation of measles transmission in a population having a constant birth rate within 250 years. In this simulation, we introduced realistic parameters such as distribution of the lifetime of the maternal antibodies, the latent and infection periods, and two different vaccination schedules. We fixed the life expectancy at 100 years to ensure a constant population number. Although this value is high and not realistic, it does not change the general behavior of our results. We found the disease to be persistent for a contact rate ξ between $1/L$ and $10/L$ with interepidemic periods depending on ξ and m, while the disease propagation is only slightly affected by the initial conditions, although we expected a nonlinear behavior predicting an unstable evolution. This investigation allowed us to predict both the efficiency of the vaccination and future epidemic peaks. We compared our simulation results with existing data (from 1994 to 1997) in the city of Oran (Algeria) which has a similar demographic situation as well as approximately identical periods of the vaccination as used in this paper. We found a good agreement for the first 3 years while the number of simulated cases is higher for the fourth one, probably due to the variation of the birth rate at the beginning of the last decade in this country. We showed also that a second-dose vaccination at 2 years reduces the number of cases by more than 75% within the next two decades in comparison with a vaccination at 6 years (occurring in Algeria). This suggests that 75% of the children between 2 and 6 years are naturally infected. Finally, we compared our results to the incidence suggested by the mass action principle and found a discrepancy due to the asymptotic behavior of the incidence when the number of susceptibles is small compared to the incidence.

Therefore, this algorithm seems to be more realistic and can be easily extended to other human, animal, and plant diseases only by introducing the corresponding parameters and distributions. Since the spatial heterogeneity was characterized by a distribution of connections in a recent work on the small world network [8] and the disease propagation depends mainly on the connections, we can introduce this heterogeneity only by including the right distribution. It is also possible to combine this algorithm with the small world network to take into account the spatial heterogeneity.

ACKNOWLEDGMENTS

One of the authors (N.Z.) would like to thank the Arab Fund for Economic and Social Development for financial support. We thank Professor A. M. Dykhne for fruitful discussions on the subject.

[1] D. Bernoulli, in *Mémories de Mathématiques et de Physique* (Académie Royale des Sciences, Paris, 1760), pp. 1–45.
[2] W. H. Hamer, Lancet **1**, 733 (1906).
[3] N. T. Bailey, *The Mathematical Theory of Infectious Disease*, 2nd ed. (Hafner, New York, 1975).
[4] R. M. Anderson and R. M. May, *Infectious Diseases of Humans, Dynamics, and Control* (Oxford University Press, Oxford, 1991).
[5] H. W. Hethcote, SIAM Rev. **42**, 599 (2000).
[6] D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness* (Princeton University Press, New Jersey, 1999); D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998); S. H. Strogatz, *ibid.* **410**, 268 (2001).
[7] N. Boccara, K. Cheong, and M. Oram, J. Phys. A **27**, 1585 (1994).
[8] N. Zekri and J. P. Clerc, Phys. Rev. E **64**, 056115 (2001).
[9] M. G. Roberts and M. I. Tobias, Epidemiol. Infect. **124**, 279 (2000); G. N. Becker and A. Bahrampour, Math. Biosci. **142**, 63 (1997).
[10] D. Lévy-Bruhl, J. Maccario, S. Richardson, and N. Guérin, Bull. Epidémiologique Hebdomadaire **29**, 133 (1997); N. G. Becker and V. Rouderfer, Math. Biosci. **131**, 81 (1996).
[11] N. M. Shnerb, E. Bettelheim, Y. Louzoum, O. Agam, and S. Solomon, Phys. Rev. E **63**, 021103 (2001); E. Ahmed and H.

N. Aziga, Physica A **253**, 347 (1998); E. Ahmed and A. S. Elgazzar, *ibid.* **296**, 529 (2001); M. J. Keeling, P. Rohani, and B. Grenfell, Physica D **148**, 317 (2001).

[12] GEANT, CERN software library; N. Zekri, Ph.D. thesis, Université de Savoie, 1988.

[13] www.cdc.gov/publication.htm

[14] A. R. Mc Lean and R. M. Anderson, Epidemiol. Infect. **100**, 419 (1988).

[15] A. Benthabet (private communication).