ANZIAM J. 44(2003), 501–512

SEASONALITY AND CRITICAL COMMUNITY SIZE FOR INFECTIOUS DISEASES

R. M. CULLEN[1], A. KOROBEINIKOV[2] and W. J. WALKER[1]

(Received 16 August, 2000; revised 21 November, 2001)

Abstract

The endemicity of infectious diseases is investigated from a deterministic viewpoint. Sustained oscillation of infectives is often due to seasonal effects which may be related to climatic changes. For example the transmission of the measles virus by droplets is enhanced in cooler, more humid seasons. In many countries the onset of cooler, more humid weather coincides with the increased aggregation of people and the seasonal effect can be exacerbated. In this paper we consider non-autonomous compartmental epidemiological models and demonstrate that the critical community size phenomenon may be associated with the seasonal variation in the disease propagation. This approach is applicable to both the prevaccination phenomenon of critical community size and the current goal of worldwide elimination of measles by vaccination.

1. Introduction

For some years it was observed that measles has intriguing behaviour. Notably, in an unvaccinated community of larger than half a million people measles persists endemically, exhibiting recurrent outbreaks (epidemics) with regular periods, while in a community of less than a quarter of a million people measles vanishes completely after a few regular outbreaks [3, 4]. For example, in small island communities such as Iceland the occasional epidemics that occurred were triggered by travellers and the virus vanished from the island between epidemics [2].

The empirical data has been reviewed by Anderson and May [2, pp. 81–86], who distinguish between epidemic fadeout and endemic fadeout. Epidemic fadeout occurs when a small number of infectives is introduced into a totally susceptible population. Regardless of the population size the infectives increase rapidly then decline to zero.

[1]Department of Mathematics, University of Auckland, Private Bag 92019, Auckland, New Zealand.
[2]Centre for Mathematical Biology, Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, UK; e-mail: korobeinikov@math.ox.ac.uk.

Influenza and the common cold exhibit this type of behaviour, completely vanishing from a population each year after an epidemic. Endemic fadeout occurs for populations of less than a critical size (prior to mass vaccination it was a quarter to half a million people for measles) and irrespective of the number of susceptibles present. Anderson and May [1] demonstrated that endemic fade-out can be produced by stochastic models which, however, generally overestimate the critical community size. Keeling and Grenfell [10] obtain improved estimates with an age-structured stochastic model constructed to fit prevaccination data in England and Wales.

In this paper we develop a deterministic explanation of the measles critical community size phenomenon. Elimination by vaccination has been achieved for smallpox and is imminent for the polio virus. Worldwide elimination of measles is also a goal of the World Health Organisation. A deterministic model which is applicable to the prevaccination problem of critical community size should also be applicable to the problem of global elimination in the vaccination era.

Numerous deterministic epidemiological models have been developed for infectious diseases which have asymptotically stable equilibria and consequently exhibit damped oscillation towards an equilibrium. This stability of deterministic models is in striking contradiction to observations which demonstrate that if an infection persists in a population endemically then it maintains self-sustained oscillations [2, 3, 4, 6]. Fundamental to a deterministic explanation of the critical size phenomenon is a mechanism for the reversal of the stability of the equilibria in these deterministic models for infectious diseases. Further it is desirable that this mechanism should be directly applicable to measles, the infectious disease which most readily demonstrates the critical community size phenomenon.

Measles is spread by droplets. The virus remains viable outside the host and inside the droplet for longer when the droplet maintains its integrity for longer (for example in cooler, more humid seasons). In many but not all countries the onset of cooler, more humid weather coincides with the increased aggregation of people and the seasonal effect can be exacerbated. This provides a biological basis for incorporating seasonal variation of the basic reproductive rate into the model. We obtain seasonal models which exhibit sustained oscillations of infectives.

Our deterministic investigation also explores the possibility of other explanations in an autonomous system with a constant reproductive rate. For instance an alternative mechanism for the critical size phenomenon is the presence of a limit cycle usually associated with a Hopf bifurcation (or with its discrete-time analogue, a Neimark-Sacker bifurcation) in the infectives – susceptibles phase space. Assume that a measles endemicity model has a fixed point which is asymptotically stable as the population size N tends to infinity. Further assume that as the population decreases to some critical value of N a supercritical bifurcation occurs and a stable limit cycle appears while the fixed point becomes unstable. As the population decreases further the stable

limit cycle expands until at some value of N it crosses the infected population level with one infective. Thus the disease cannot persist with less than one infective and there is a critical community size phenomenon.

In Section 3 we explore the possibility of such a bifurcation for a general class of autonomous models and in particular a negative binomial model for nonhomogeneous mixing. We find that spatial models incorporating heterogeneity do not reverse the stability of the equilibrium. A geometrical approach demonstrates that a reversal arises most naturally by relaxing the condition for an autonomous system.

2. The basic difference equation models

In this paper we will consider discrete-time epidemiological models. Discrete-time models are not new for mathematical epidemiology — difference equations have been used by Soper [14], Bartlett [3], Hoppensteadt [8, 9] and others. The main advantage of discrete-time compared with continuous-time models is their natural time scaling. The natural time scale for epidemiological processes is a parasite generation. By the term "parasite generation" we mean the average time interval which commences when a susceptible is exposed to an infective dose, includes the period during which the host passes infection and ends when the host is fully recovered or dead. The choice of a parasite generation as a time interval has the important consequence that all infected recover after a definite period of time. Whereas for continuous-time models, unless we use integro-differential equations or time delay, we postulate that "continuous recovery" from the infected to the recovered class is at a constant rate. This is rarely realistic while discrete time with a parasite generation as a time interval allows a natural interpretation of all the model parameters.

We consider an SIR compartmental model which has been reported previously in [5]. We assume a human population of size N partitioned into susceptible S, infected I and recovered R compartments. We assume that recovery implies life-long immunity, that is, no return from the recovered compartment into the susceptibles compartment is possible. The number of susceptibles and infectives in the tth parasite generation are denoted by S_t, I_t and R_t respectively. By the basic reproductive rate R_0 we mean the average number of people (susceptible, infected or recovered) that an infective comes into infective contact with.

The population is assumed to be homogeneously mixing and its size is constant. We assume that all hosts live to age L then die ("Type I mortality" in ecologists' terminology [2, p. 62]). We further assume that there are k parasite generations each year and that to keep the population size constant the number of births in a parasite generation $c = N/kL = bN$ $(b = 1/kL)$. All births are into the susceptibles' compartment; all deaths are from the recovered compartment. Measles has about 24

parasite generations per year [2] and we will use $k = 24$ year^{-1} from here onwards.

The principle of mass action takes into account that an infective comes into infective contact with and might infect R_0 people, some of whom may be already infected or recovered. The equations for mass action are

$$\begin{aligned} I_{t+1} &= R_0 I_t \frac{S_t}{N}, \\ S_{t+1} &= S_t - R_0 I_t \frac{S_t}{N} + c. \end{aligned} \tag{2.1}$$

The constant population size assumption permits us to disregard the third equation,

$$R_{t+1} = R_t + I_t - c,$$

for the recovered compartment R.

The system has a fixed point $\bar{I} = c$, $\bar{S} = N/R_0$. For endemic persistence of measles the long term average of the susceptibles is taken to be N/R_0.

If $R_0 < 4kL$ (the case $R_0 > 4kL$ is biologically infeasible) the usual linear analysis at the fixed point yields a matrix with eigenvalues which satisfy $|\lambda| = 1$. Graphically the orbits in the phase space are closed loops corresponding to invariant sets. These closed orbits are structurally unstable but they give a description of average behaviour and a reasonable prediction for inter-epidemic periods. The phenomenon of critical community size will be demonstrated by adding two refinements to the mass action model.

The first refinement is motivated by the observation that in the mass action model one susceptible may be infected by more than one infective. Assuming that infective contacts are Poisson distributed we come to a system of equations introduced in [5],

$$\begin{aligned} I_{t+1} &= S_t - S_t \exp(-R_0 I_t/N), \\ S_{t+1} &= S_t \exp(-R_0 I_t/N) + c. \end{aligned} \tag{2.2}$$

The term $\exp(-R_0 I_t/N)$ is the zeroth term in a Poisson distribution with mean $\mu = R_0 I_t/N$ and represents the probability that a susceptible in the tth generation escapes infection. It is shown in [5] that the system has an attracting fixed point and orbits gently spiral in to the fixed point.

The second refinement stems from the observation that measles propagation varies with season. This variation may be captured assuming that R_0 is not a constant but a periodic function with a one-year, or k generation, period. Our results do not depend on a particular formula for R_0 but we shall leave a discussion of this aspect until Section 4.

https://doi.org/10.1017/S144618110001289X Published online by Cambridge University Press

3. Nonhomogeneous mixing

The mass action principle assumes homogeneous mixing. We now examine the impact of spatial heterogeneity whether due to demographic, social or geographical factors. Let x be a random variable having a a Poisson distribution $\mu^q e^{-\mu}/q!$ $(q = 0, 1, 2, \dots)$, which is the probability that a susceptible has q infective contacts. Inhomogeneity can be captured with a probability density function for μ,

$$P(\mu) = \frac{\alpha^m}{\gamma(m)}\mu^{m-1}e^{-\alpha\mu}, \quad \mu > 0,$$

where $m, \alpha > 0$ are parameters. The probability that x takes the value q is

$$\begin{aligned} Q(k) &= \int_0^\infty \frac{\mu^q e^{-\mu}}{q!} P(\mu)\, d\mu \\ &= \int_0^\infty \frac{\mu^q e^{-\mu}}{q!}\frac{\alpha^m}{\gamma(m)}\mu^{m-1}e^{-\alpha\mu}\, d\mu = \left(\frac{\alpha}{1+\alpha}\right)^m \binom{-m}{q}\frac{(-1)^q}{(1+\alpha)^q}. \end{aligned}$$

This is the negative binomial distribution which has mean m/α and variance $m(1+\alpha)/\alpha^2$ (see [13]). The parameter α can be eliminated by taking the mean to be the average number of contacts per susceptible so that $m/\alpha = R_0 I_t/N$. Then the probability of escaping infection is

$$Q(0) = \left(1 + \frac{1}{\alpha}\right)^{-m} = \left(1 + \frac{R_0 I_t}{mN}\right)^{-m}$$

which leads to the equations

$$\begin{aligned} I_{t+1} &= S_t\left[1 - \left(1 + \frac{R_0 I_t}{mN}\right)^{-m}\right], \\ S_{t+1} &= c + S_t\left(1 + \frac{R_0 I_t}{mN}\right)^{-m}. \end{aligned} \tag{3.1}$$

The negative binomial distribution has been used to describe variation in the environment, diversity leading to a qualitative change in the behaviour of a model. The classical example is the stabilisation of the Nicholson-Bailey equations for a host-parasitoid system [7]. In that example it is the stabilisation itself which is important rather than the bifurcation which is not a Neimark-Sacker bifurcation (the Neimark-Sacker bifurcation is equivalent to a Hopf bifurcation for discrete-time systems leading to the appearance of a closed invariant curve surrounding a fixed point [11]).

With the particular system (3.1) as an example, we consider a general model of infection

$$\begin{aligned} I_{t+1} &= f(I_t, S_t, N), \\ S_{t+1} &= S_t - f(I_t, S_t, N) + c, \end{aligned} \tag{3.2}$$

https://doi.org/10.1017/S144618110001289X Published online by Cambridge University Press

where the function $f(I, S, N)$ represents reproduction of the infection and depends on N as a parameter. We investigate the possibility of a supercritical Neimark-Sacker bifurcation and assume the natural conditions

$$f(0, S, N) = f(I, 0, N) = 0, \quad f(I, S, N) > 0, \qquad \frac{\partial f}{\partial I} > 0, \quad \frac{\partial f}{\partial S} > 0 \tag{3.3}$$

for all $I, S, N > 0$. We assume also that, for all $a > 0$, $\lim_{S\to\infty} f(a, S, N) = R_0 a$.

The system (3.2) has an unique fixed point $(\bar{I}, \bar{S})$ given by the equations

$$\bar{I} = c, \quad f(\bar{I}, \bar{S}, N) = c. \tag{3.4}$$

The linear approximation of the system (3.2) near the fixed point is

$$\begin{aligned} I_{t+1} &= \frac{\partial f(\bar{I}, \bar{S},)}{\partial I} I_t + \frac{\partial f(\bar{I}, \bar{S})}{\partial S} S_t, \\ S_{t+1} &= -\frac{\partial f(\bar{I}, \bar{S})}{\partial I} I_t + \left(1 - \frac{\partial(\bar{I}, \bar{S}}{\partial S}\right) S_t. \end{aligned} \tag{3.5}$$

The characteristic equation of the system (3.5) is $\lambda^2 - \delta\lambda + \tau = 0$, where

$$\delta(N) = 1 + \frac{\partial f(\bar{I}, \bar{S}, N)}{\partial I} - \frac{\partial f(\bar{I}, \bar{S}, N)}{\partial S}, \quad \tau(N) = \frac{\partial f(\bar{I}, \bar{S}, N)}{\partial I}.$$

The system (3.5) has a pair of complex conjugate eigenvalues $\lambda, \bar{\lambda}$ with $|\lambda| = 1$ when $-2 < \delta(N) < 2$ and $\tau(N) = 1$.

A Neimark-Sacker bifurcation occurs in the system (3.2) when as the parameter N varies the two complex eigenvalues of the linearised system (3.5) cross the unit circle $|\lambda| = 1$. If the eigenvalues cross from the inside to the outside it is a supercritical bifurcation and a stable fixed point bifurcates to a stable closed orbit, otherwise it is a subcritical bifurcation [11].

THEOREM 3.1. *If there is a supercritical Neimark-Sacker bifurcation of the system* (3.2) *then there is an open interval* J *such that for each* $N \in J$ *there exists* $I_0 \in (0, \bar{I})$ *such that* $\partial^2 f(I_0, \bar{S}, N)/\partial I^2 > 0$.

PROOF. If a Neimark-Sacker bifurcation occurs then there is an open interval J such that for $N \in J$

$$\tau(N) \equiv \frac{\partial f(\bar{I}, \bar{S}, N)}{\partial I} > 1. \tag{3.6}$$

https://doi.org/10.1017/S144618110001289X Published online by Cambridge University Press

Since $f(0, \bar{S}, N) = 0$ and $f(\bar{I}, \bar{S}, N) = c$ for all N, see (3.3) and (3.4), by applying the mean value theorem we obtain the existence of a point $I_1 \in (0, \bar{I})$ such that

$$\frac{\partial f(I_1, \bar{S}, N)}{\partial I} = \frac{f(\bar{I}, \bar{S}, N) - f(0, \bar{S}, N)}{\bar{I} - 0} = \frac{c}{c} = 1.$$

Now using this result and (3.6) and by applying the mean value theorem to the function $g(I) = \partial f(I, \bar{S}, N)/\partial I$ we obtain $I_0 \in (I_1, \bar{I})$ such that

$$\frac{\partial^2 f(I_0, \bar{S}, N)}{\partial I^2} = \frac{dg(I_0)}{dI} = \frac{g(\bar{I}) - g(I_1)}{\bar{I} - I_1} = \frac{\partial f(\bar{I}, \bar{S}, N)/\partial I - \partial f(I_1, \bar{S}, N)/\partial I}{\bar{I} - I_1} > 0.$$

The theorem is proven.

This theorem can now be applied to the system (3.1) for which

$$\frac{\partial^2 f}{\partial I^2} = -R_0^2 \frac{m+1}{m} \frac{S}{N^2} \left(1 + \frac{R_0 I}{Nm}\right)^{-(m+2)} < 0 \quad \text{for all } I, S > 0.$$

We are able to conclude that nonhomogeneous mixing does not lead to a Neimark-Sacker bifurcation of the system. Indeed the theorem indicates that it is more appropriate to consider a non-autonomous system and seasonal variation since it is difficult to imagine a biological reason for $\partial^2 f/\partial I^2 > 0$ in an autonomous system, while a condition $\partial^2 f/\partial I^2 < 0$ may arise as a consequence of saturation effects. In fact it is easy to see that for discrete-time models the latter condition must hold to avoid a possibility of multiple infection of one susceptible.

4. The seasonal reproductive rate and critical community size phenomenon

In this section we assume that the infective contacts are Poisson distributed and consider (2.2). We assume also that the reproductive rate R_0 is a periodic function of parasite generation with a one-year period. Here we take

$$R_0(t) = \begin{cases} 14 + 4(t-1)/11 & \text{for } 1 < t \leq 12, \\ 18 - 4(t-13)/11 & \text{for } 13 < t \leq 24. \end{cases}$$

The numerical average of R_0 is 16, which is the reproductive rate for measles [2]. It is remarkable that our results do not depend on this particular formula which is taken as an example only. The average host's life span L is taken to be 60 years which fits pre-vaccination data.

To investigate endemic persistence and endemic fade-out we fix the initial values I_1 and S_1 and vary the population size N. The simulations are indicative of a trend of

https://doi.org/10.1017/S144618110001289X Published online by Cambridge University Press

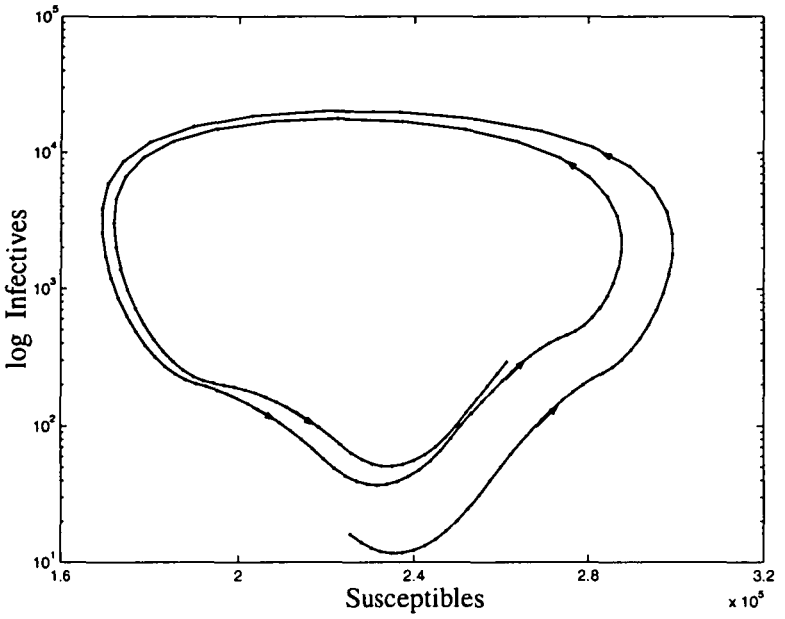

FIGURE 1. Inward spiralling. Here $N = 3.6$ million, $S_1 = N/16 = 225,000$ and $I_1 = 16$.

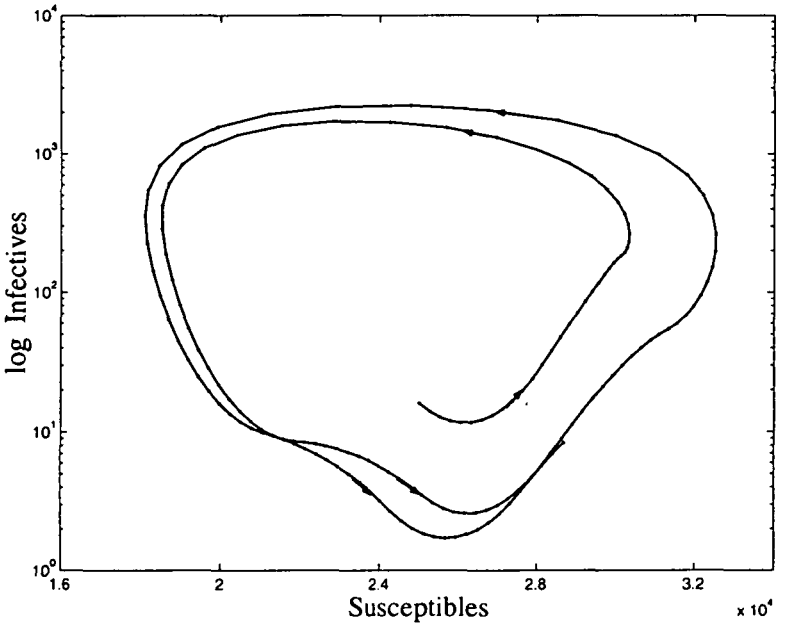

FIGURE 2. Outward spiralling. Here $N = 400,000$, $S_1 = N/16 = 25,000$ and $I_1 = 16$.

https://doi.org/10.1017/S144618110001289X Published online by Cambridge University Press

spiralling in for $N \geq 500,000$ and spiralling out for $N \leq 400,000$. Figures 1 and 2 clearly indicate the N-dependence. We have chosen these two particular diagrams since they have no points of self-intersection. We have used a $\log_{10}$ scale for I_t which shows clearly the orbits' behaviour. The orbit in Figure 1 spirals inwards similar to the behaviour exhibited with constant R_0. Figure 2 has been obtained by taking $N = 400,000$ and $S_1 = N/16 = 25,000$. The orbit now spirals out.

Table 1 demonstrates the minimal number of infectives $I_{\min}$ for different populations for endemically persisting measles. The data was obtained by running the model for 10 years (240 parasite generations) for each population N with initial values of $I_1 = 16$ and $S_1 = N/16$. In the table we ignored transient behaviour in the first few generations of the simulation. The table clearly indicates that for this model a population of 0.2 million is below the critical community size and a population of 0.4 million is at the edge of it.

TABLE 1.

N (million)	$I_{\min}$
0.2	0.50
0.4	1.00
0.6	1.64
0.8	2.04
1.0	2.52
2.0	5.00
3.6	11.6

The seasonal $R_0(t)$ of the form used here has induced spiralling in of orbits for $N \geq 500,000$ and spiralling out for $N \leq 400,000$. The question now arises, how dependent is this phenomenon on the particular form of the periodic function $R_0(t)$? The model was extensively tested for a variety of different periodic functions including non-symmetric functions. Surprisingly we found that the average value of R_0 was not as important as the magnitude of its variation, $R_0(t)_{\max} - R_0(t)_{\min}$. Periodic functions with magnitudes of variation of three to four including non-symmetric functions produced the same type of N-dependence as the abovementioned function. A magnitude of variation of two or less gave spiralling in, while a magnitude of six gave spiralling out for all considered N and for all considered periodic functions. For a smaller average value of R_0 and for a magnitude of variation of four the critical community size effect could still be observed.

This goes against intuition which says that the critical size phenomenon in the case of measles is very much due to the large R_0. We believe that it is a consequence of the spread of measles by droplets which leads to considerable seasonal variation of R_0. Polio in contrast is spread by a fecal oral route and mumps by saliva — these methods

https://doi.org/10.1017/S144618110001289X Published online by Cambridge University Press

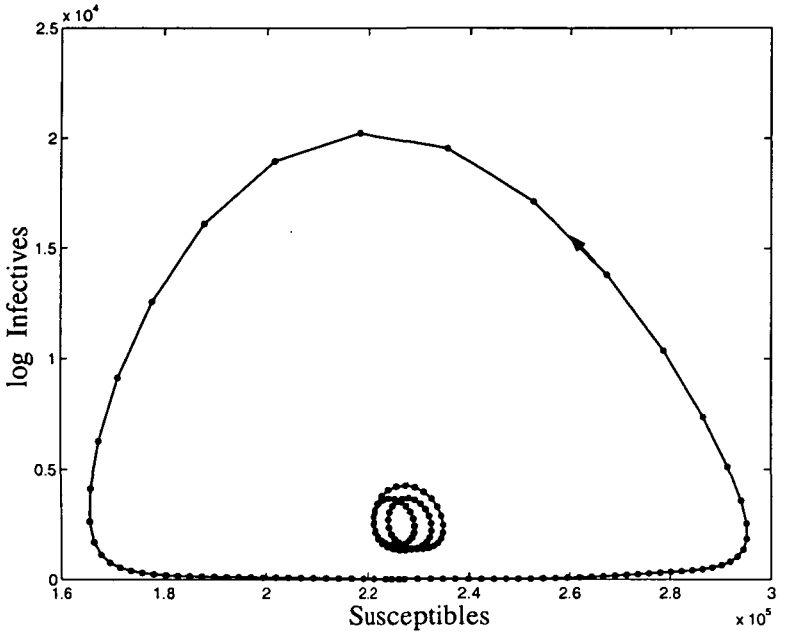

FIGURE 3. Large orbit and central tight invariant curves. Here $N = 3.6$ million and $S_1 = N/16$; for the large orbit $I_1 = 25$ and for the tight orbit $I_1 = 2040$.

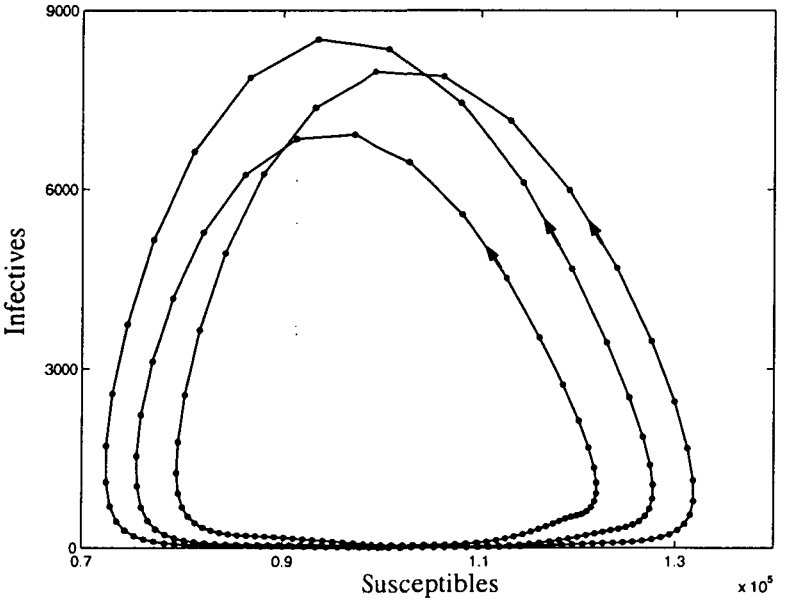

FIGURE 4. Self-intersecting steady state; here $N = 1.6$ million, $S_1 = N/16 = 100,000$ and $I_1 = 23$.

https://doi.org/10.1017/S144618110001289X Published online by Cambridge University Press

of transmission mean that there will not be the required degree of seasonal variation in R_0 and hence there is no critical community size phenomenon.

5. The seasonal model and endemicity

For values of $N > 400,000$ the long term behaviour is of interest. Because the system is non-autonomous it may take upwards of sixty years or more to reach a steady state. One possible steady state is a single large closed invariant curve with a comparatively low minimum value of $I_{\min} > 1$ as seen in Figure 3. (We use a $\log_{10}$ scale here for the infectives axis.) The closed orbit in Figure 3 has an epidemic cycle of about 74–75 parasite generations and an epidemic itself lasts for about 18 parasite generations which is approximately as observed [2, p. 83]. (By contrast the closed orbit of mass action has an epidemic cycle and an epidemic length of 79 and 20 parasite generations respectively.) The lower estimate for the epidemic cycle corresponds better to available data but if anything it is still too high. This may be because a population of this size is likely to exhibit clumping and consequently chaotic behaviour [12].

Another possible steady state is a tight closed invariant orbit in the vicinity of the steady state of the corresponding autonomous system (see Figure 3; here $N =$ 3.6 million, $I_1 = 2040$ and $S_1 = N/16$).

The third steady state is the most interesting. Figure 4 represents a simple self-intersecting invariant orbit which is an icon of the prevaccination era for measles in a population of $N = 1.6$ million, such as was the case in New Zealand at the time. There are in effect two epidemic cycles of a period of three years and one of a period of two years. The average epidemic period here is 2.75 years, whereas in New Zealand the average epidemic cycle between 1949 and the commencement of mass immunisation in 1969 was 2.86 years [6].

If N is fixed and I_1 is gradually increased, each of the three steady states exists over a number of windows of I_1 values. The order of the states seems unpredictable although as I_1 approaches the value associated with the fixed point (of the mass action equations) the tight orbit predominates. We offer the caveat that it may take an orbit a very long time to reach a steady state particularly when I_1 is close to a transition value.

6. Conclusions

In this paper the critical community phenomenon for measles was investigated for populations in the range $100,000 \leq N \leq 3,000,000$, that is, populations which operate as a single unit. The final process of elimination at a low level of infectives will

be subject to stochastic effects. However our approach has been to seek a deterministic explanation. It was found that endemic fade-out can occur if the basic reproductive rate of the parasite exhibits sufficient seasonal variation. We suggest that measles is the only common infectious disease with the required seasonal variation in the basic reproductive rate. Comparable fade-out might be observed for a less infectious disease which also exhibited sufficient seasonal variation.

Acknowledgement

We would like to thank S. Adi Purnomo, Y.-J. Tsai and J. M. Wilson for assistance with computer simulations.

References

[1] R. M. Anderson and R. M. May, "The invasion, persistence and spread of infectious diseases within animal and plant communities", *Phil. Trans. Roy. Soc.* **B314** (1986) 533–570.
[2] R. M. Anderson and R. M. May, *Infectious diseases in humans—dynamics and control* (Oxford University Press, Oxford, 1991).
[3] M. S. Bartlett, "The critical community size for measles in the United States", *J. Roy. Statist. Soc.* **123** (1960) 37–44.
[4] F. L. Black, "Measles endemicity in insular populations: critical community size and its implications", *J. Theor. Biol.* **11** (1966) 207–211.
[5] R. M. Cullen, N. D. Ellis and W. J. Walker, "A model of measles endemicity", *Nonlinear Anal.* **35** (1999) 191–198.
[6] R. M. Cullen and W. J. Walker, "Measles epidemics 1949–91: the impact of mass immunisation in New Zealand", *N. Z. Med. J.* **109** (1996) 400–402.
[7] M. P. Hassell, *The dynamics of arthropod predator-prey systems* (Princeton University Press, Princeton, 1978).
[8] F. C. Hoppensteadt, *Mathematical theories of populations, demographics, genetics and epidemics* (SIAM, Philadelphia, 1975).
[9] F. C. Hoppensteadt, *Mathematics in medicine and the life sciences* (Springer, New York, 1992).
[10] M. J. Keeling and B. T. Grenfell, "Disease extinction and community size: Modeling the persistence of measles", *Science* **275** (1997) 65–67.
[11] Y. A. Kuznetsov, *Elements of applied bifurcation theory* (Springer, New York, 1995).
[12] R. M. May, "Necessity and chance: Deterministic chaos in ecology and evolution", *Bull. Amer. Math. Soc.* **32** (1995) 291–308.
[13] E. C. Pielou, *An introduction to mathematical ecology* (Wiley-Interscience, New York, 1969).
[14] H. E. Soper, "Interpretation of periodicity in disease-prevalence", *J. Roy. Statist. Soc.* **92** (1929) 34–73.