EPIDEMIC MODELS: THRESHOLDS AND POPULATION REGULATION

The complexity of epidemic processes is often obscured by the relatively simple way in which transmission is characterized in some epidemic models. A large part of current deterministic epidemic theory extends the S (susceptible host density), I (infective host density), and N (total host population density) differential equation model first proposed by Ross (cf. Ross 1916; Ross and Hudson 1917). Most of these models predict that the density of susceptibles (S) must exceed a threshold (N_T) before an infectious disease can become epidemic (for a review of the theory, see Serfling 1952; Bailey 1975). Here we show that a seemingly minor change in the way transmission is modeled can significantly alter conclusions as to whether such a threshold exists and to whether an infectious disease can regulate its host population.

In a recent series of papers Anderson and May (Anderson 1978, 1979*a*, 1979*b*, 1980, 1982; Anderson and May 1978, 1979, 1980, 1981, 1982; May and Anderson 1978, 1979) focused considerable attention, using Ross-type differential equation models, on the role of infectious diseases in the observed dynamics of host populations including invertebrate hosts. They examined the role of parasites—broadly defined to include viruses, bacteria, protozoans and helminths—in biological control (Anderson 1979*a*, 1982; Anderson and May 1980), and they determined measures necessary for the control and eradication of pathogens, especially viral diseases such as rabies, whooping cough, and measles (Anderson et al. 1981; Anderson and May 1982). In most cases, their models predict the existence of a threshold $N_T > 0$. This threshold is a consequence of the assumption that the rate of disease transmission is proportional to the number of random encounters between infectives (of density I) and susceptibles (of density S) in the population, i.e., the transmission rate is βSI, where β is the transmission parameter. More complex characterizations of transmissions are common and have been discussed at length (Anderson 1979*b*, 1980, 1982; Bailey 1975; Anderson and May 1981; Yorke et al. 1979).

Generally, the rate of disease transmission will depend on the rate at which propagules are produced by the infected individuals within a population, the dynamics of propagule transfer (e.g., vector dynamics), and the rate at which propagules invade susceptible hosts. For our purposes we assume that propagule production is a linear function of the number of infectives and either that the vector density remains sufficiently high so that transmission is essentially independent of their dynamics or that vectors are not required for transmission (e.g., venereal diseases). Then, if we assume that the rate at which propagules invade susceptible hosts is dependent on the proportion of individuals at risk in the host

Am. Nat. 1983. Vol. 121, pp. 892–898.

population as well as on host density, a characterization of the transmission rate is

$$\text{transmission rate} = \beta I f(S/N) g(N). \tag{1}$$

The susceptibility function $f(S/N)$ may simply be taken as S/N; although in situations in which a vector preferentially bites an infected individual or in sexually transmitted diseases when carriers are on average more promiscuous, the function will be nonlinear. For our purposes it will be assumed that

$$f \text{ is a continuous, strictly increasing function that satisfies } f(0) = 0 \text{ and } f(1) = 1. \tag{2}$$

The density-dependent response function $g(N)$ could, in some situations, be characterized by the one-parameter family of curves $g(N) = N^k, k \geq 0$ (R. M. May, personal communication). An alternative form for $g(N)$ that could be appropriate for certain vector-borne diseases is

$$g(N) = N/(N_0 + N) \tag{3}$$

in which N_0 is a constant characterizing vector efficiency (cf. MacDonald 1961).

Anderson and May (1979) consider the basic model

$$\frac{dN}{dt} = (a - b)N - \alpha I \tag{4}$$

$$\frac{dI}{dt} = \beta I(N - I) - (\alpha + b + \nu)I \tag{5}$$

in which a, b, α, and ν are the per capita birth, death, disease-induced mortality and infective-recovery rates. The transmission rate in this model is equivalent to assuming that $f(S/N) = S/N$ and $g(N) = N$. This model, as discussed by Anderson and May (1979), implicitly defines a threshold density $N_T = (\alpha + b + \nu)/\beta$ such that I only increases if $S = N - I > N_T$ (see fig. 1).

If $g(N)$ has the form expressed in (3), then model (4) and (5) provides an approximate analysis for the case: $N_0 \gg N(t)$ for all t over which the process is considered. Alternatively, if $N_0 \ll N(t)$, i.e., $g[N(t)] \simeq 1$ for all t, then equation (5) can be replaced by

$$\frac{dI}{dt} = \beta I f[(N - I)/N] - (\alpha + b + \nu)I, \tag{6}$$

in which f is a susceptibility function that satisfies (2). In this case, as is known for the special case when f is linear (May and Anderson 1979), a threshold no longer exists and the infective class can grow at low population densities, i.e., $N_T = 0$ (see fig. 2). In addition, host population regulation is generally no longer possible. The fundamental difference between (5) and (6) is that in (5) the transmission rate is directly related to the absolute density of susceptibles, while in (6) it is related to their relative density. The former assumption may be reasonable for airborne disease transmission processes, as discussed by Anderson and May (1978, 1980; May and Anderson 1978) but may be inadequate for characterizing the transmission of certain vector-borne and venereal diseases.

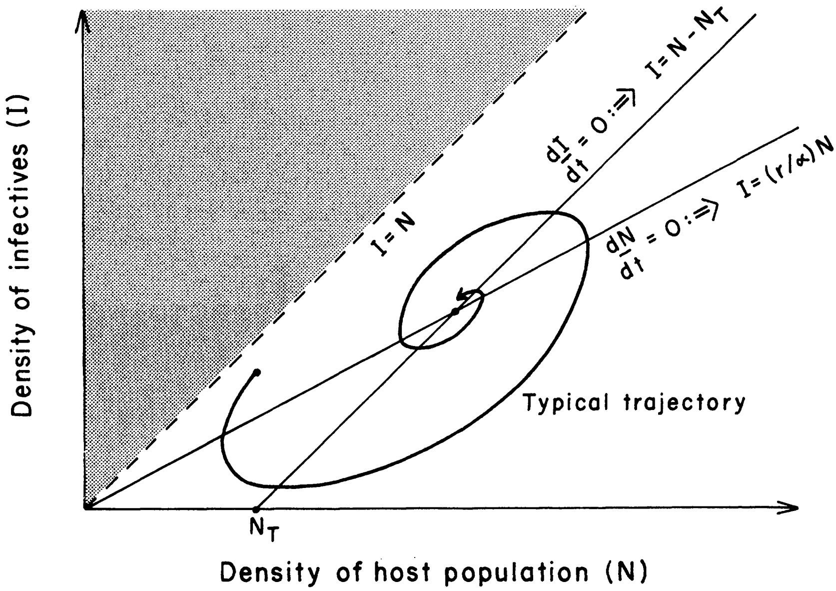

FIG. 1.—Phase plane diagram of an epidemic process in which the transmission rate is proportional to the absolute density of the infectives (I) and susceptibles ($N - I$). Exponentially increasing populations are regulated provided the disease-induced mortality rate (α) exceeds the hosts' intrinsic growth rate ($r = a - b$). Whenever the density of susceptibles is below the threshold density (N_T), the density of infectives will decrease. This type of transmission permits the disease to regulate the host population.

The following questions are thus raised by the analysis of equations (4) and (6). Can host populations be regulated by vector-borne or venereal diseases in the absence of density-dependent processes not related to the disease? And, does a threshold really exist for a particular disease under consideration? For example, a threshold concept proved to be inappropriate for smallpox, and the disease was eventually eradicated using a surveillance-containment strategy.

Although parasites cannot regulate host population density when transmission is a function of the relative density of susceptibles, they may suppress it below a carrying capacity determined by other factors. For example, consider the case in which the death rate b increases linearly with population density N (cf. Anderson 1980; Anderson and May 1981), i.e.,

$$b = b_0 + b_1 N. \tag{7}$$

Then from equations (4), (6), and (7) we obtain for the linear case $f[(N - I)/N] = (N - I)/N$,

$$\frac{dN}{dt} = rN(1 - N/K) - \alpha I \tag{8}$$

$$\frac{dI}{dt} = \rho I - I[\beta(I/N) + r(N/K)], \tag{9}$$

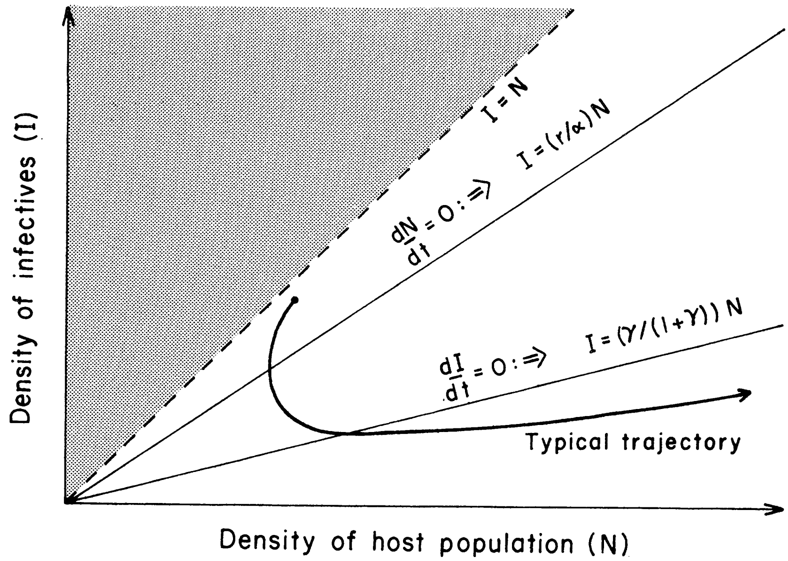

FIG. 2.—Phase plane diagram of an epidemic process in which the transmission rate is proportional to the absolute density of infectives (I) and the relative density of susceptibles $[(N - I)/N]$. If the slope of the N isocline is greater than that of the I isocline, the host population (as portrayed by the typical trajectory) grows without bound. If the relationship between the slopes is reversed, the disease causes the host to approach extinction. The diagram holds for all susceptibility functions $f[(N - I)/N]$ satisfying condition (2) in the text. Note that $\gamma = f^{-1}[(\alpha + \beta + \nu)/\beta]$ in which f^{-1} is the inverse function to f, i.e., $\gamma = f^{-1}[f(\gamma)]$. This type of transmission does not permit the diseases to regulate the host population.

in which $r = a - b_0$ is the intrinsic rate of increase of the host population,

$$\rho = \beta - (\alpha + b_0 + \nu) \tag{10}$$

is the intrinsic rate of increase of infectives, and $K = (a - b_0)/b_1$ is the environmental carrying capacity for the host population. Equation (8) is a classical logistic growth model with the addition of a disease-induced mortality term. Equation (9) implies that the rate of growth of the infected class decreases both as relative parasite prevalence (I/N) increases and as host density approaches its carrying capacity K. If ρ and r are positive and

$$1 < \rho/r < \beta/\alpha \tag{11}$$

then and only then do equations (8) and (9) imply the coexistence of the host-parasite system around a stable equilibrium (see fig. 3), where the equilibrium population $\bar{N}$ and the equilibrium endemic level $(\bar{I}/\bar{N})$ satisfy

$$\bar{N} = K\left[1 - \left(\frac{\rho}{r} - 1\right)\Big/\left(\frac{\beta}{\alpha} - 1\right)\right] \tag{12}$$

and

$$(\bar{I}/\bar{N}) = 1 - (a + \nu)/(\beta - \alpha). \tag{13}$$

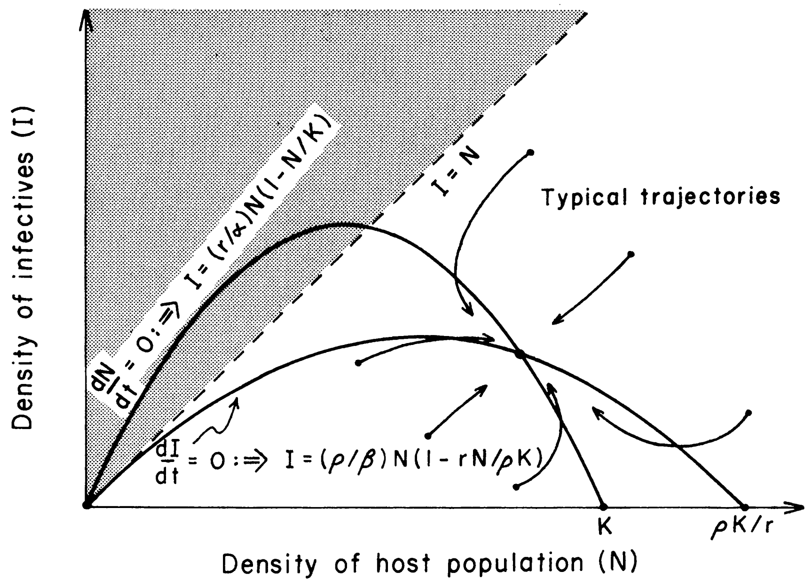

FIG. 3.—Phase plane diagram of the epidemic process depicted in fig. 2 except that density-dependent mortality is included. The equilibrium (as depicted) is stable only if the slope of the host population isocline is steeper than the slope of the infectives' isocline at the origin (i.e., $[r/\alpha] > [\rho/\beta]$) and the isoclines intersect the N axis as shown, i.e., $(\rho/r) > 1$. In this case the host population is regulated by a density-dependent mortality factor but is suppressed below its environmental carrying capacity K by the disease.

For ρ/r satisfying inequality (11), it follows that $\bar{N} \rightarrow K$ as $\rho/r \rightarrow 1$ and $\bar{N} \rightarrow 0$ as $\rho/r \rightarrow \beta/\alpha$, i.e., the relative position of ρ/r on the interval $[1, \beta/\alpha]$ is a direct measure of the level of population suppression below the carrying capacity. For example, if ρ/r is at the center of this interval, i.e., $\rho/r = [(\beta/\alpha) - 1]/2$, then $\bar{N} = K/2$. Because of the biological interpretation of r and ρ (cf. text circa [9] and [10]), inequality [11] neatly summarizes conditions for improving parasite or host control in terms of the ratio between the intrinsic rate of increase of infectives and the intrinsic rate of increase of the host population. The parasite is controlled by reducing ρ (i.e., from expression [10] by reducing the transmission rate β, by increasing the disease-induced mortality rate α, or by increasing the recovery rate ν) or by increasing the intrinsic growth rate of the host. On the other hand the host population is best controlled biologically by increasing ρ: specifically by reducing the transmission to disease-induced mortality ratio (β/α). Because both parameters in the latter ratio are likely to be increasing functions of parasite reproductive rates within hosts (Anderson and May 1981; Bremermann and Pickering 1983), it is important to determine how a particular management action will affect both the difference $(\beta - \alpha)$ and the ratio (β/α) of the parameters β and α.

From the parasites' point of view, a low recovery rate ν and a large difference between the transmission and disease-induced mortality rates, i.e., $(\beta - \alpha)$, ensures persistence at a high endemic level (i.e., $\bar{I}/\bar{N}$ close to unity); although a large value for ρ resulting from (β − α) being large (cf. expression [10]) can ultimately destroy a host population. In general, selection may strive to maximize the actual transmission rate $\beta\bar{I}(\bar{N} - \bar{I})/\bar{N}$, but these questions are pursued elsewhere (cf. Bremermann and Pickering 1983).

Endemic levels and levels of host suppression are central to the problem of using pathogens as biological control agents (Anderson 1982). Thus questions relating to the characteristics of transmission at different host densities are important in evaluating the possible success of a particular biological control program. Each epidemic process is unique, and thus the applicability of the results discussed above are limited. It could be claimed that certain qualitative concepts remain valid, even though the models presented here and elsewhere oversimplify the disease process. In a final analysis, however, broad principles may be misleading and a detailed study of the specific host and parasite biologies will be required before the most basic qualitative properties, such as the existence of a threshold, can be accepted as an intrinsic part of a particular epidemic process.

ACKNOWLEDGMENTS

We thank R. M. Anderson, R. M. May, G. F. Oster, and an anonymous reviewer for comments on earlier drafts.

LITERATURE CITED

Anderson, R. M. 1978. The regulation of host population growth by parasitic species. Parasitology 76:119–157.

———. 1979*a*. Parasite pathogenicity and the depression of host population equilibria. Nature 279:150–152.

———. 1979*b*. The persistence of direct life cycle infectious diseases within populations of hosts. Pages 1–66 *in* S. Levin, ed. *Lectures on mathematics in the life sciences*. Vol. 12. Am. Math. Soc. Providence, R.I.

———. 1980. Depression of host population abundance by direct life cycle macroparasites. J. Theor. Biol. 82:283–311.

———. 1982. Theoretical basis for the use of pathogens as biological control agents of pest species. Parasitology 84:3–33.

Anderson, R. M., H. C. Jackson, R. M. May, and A. M. Smith. 1981. Population dynamics of fox rabies in Europe. Nature 289:765–771.

Anderson, R. M., and R. M. May. 1978. Regulation and stability of host parasite population interactions. I. Regulatory processes. J. Anim. Ecol. 47:219–249.

———. 1979. Population biology of infectious diseases. Part I. Nature 280:361–367.

———. 1980. Infectious diseases and population cycles of forest insects. Science 210:658–661.

———. 1981. The population dynamics of microparasites and their invertebrate hosts. Philos. Trans. R. Soc. Lond. Ser. B 291:451–524.

———. 1982. Directly transmitted infectious diseases: control by vaccination. Science 215:1053–1060.

Bailey, N. T. J. 1975. The mathematical theory of infectious disease. 2d ed. Macmillan, New York.

Bremermann, H. J., and J. Pickering. 1983. A game-theoretical model of parasite virulence. J. Theor. Biol. 100.

MacDonald, G. 1961. Epidemiological models in studies of vector-borne diseases. Public Health Reports, Washington, D.C. 76:753–764.

May, R. M., and R. M. Anderson. 1978. Regulation and stability of host parasite population interactions. II. Destabilizing processes. J. Anim. Ecol. 47:249–268.

———. 1979. Population biology of infectious diseases. Part II. Nature 280:455–461.

Ross, R. 1916. An application of the theory of probabilities to the study of a priori pathometry. Part I. Proc. R. Soc. Lond. Ser. A 92:204–230.

Ross, R., and H. P. Hudson. 1917. An application of the theory of probabilities to the study of a priori pathometry. Parts II and III. Proc. R. Soc. Lond. Ser. A. 93:212–240.

Serfling, R. E. 1952. Historical review of epidemic theory. Hum. Biol. 24:145–165.

Yorke, J. A., N. Nathanson, G. Pianigiani, and J. Martin. 1979. Seasonality and the requirements for perpetuation and eradication of viruses in populations. Am. J. Epidemiol. 109:103–123.

WAYNE M. GETZ*
JOHN PICKERING†

DIVISION OF BIOLOGICAL CONTROL*
DEPARTMENT OF PLANT PATHOLOGY*
DEPARTMENT OF ENTOMOLOGICAL SCIENCES*†
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

Submitted May 11, 1982; Accepted November 11, 1982