

Changes in parasite aggregation with age: a discrete infection model

R. J. QUINNELL¹, A. GRAFEN² and M. E. J. WOOLHOUSE³

¹ *Department of Medical Parasitology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT*

² *Department of Plant Sciences and* ³ *Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS*

(Received 29 July 1994; revised 11 May 1995; accepted 11 May 1995)

SUMMARY

We present a discrete time model for age-related changes in the mean and variance of the number of helminth parasites per host. We assess the degree of aggregation as the negative binomial parameter, k , and use the model to examine the influence of various factors on changes in aggregation with host age: discrete versus continuous infection; the degree of predisposition to infection; infection rate; parasite survival rate; and the variance in exposure to infective stages. The model can produce both increases and decreases in k with host age. However, with parameter values typical of many human helminth infections, a monotonic increase in k (decrease in aggregation) with age is predicted. With an age-dependent infection rate, convex relationships between k and age are possible. These predictions are consistent with data from field studies, but differ from those of previous models which have suggested that k is independent of host age in the absence of density dependence in parasite population dynamics. Differences between the models, and some difficulties in the interpretation of field data, are discussed.

Key words: aggregation, density-dependence, helminth, mathematical models, negative binomial.

INTRODUCTION

Most theoretical studies of helminth infection have assumed an immigration-death process with continuous infection (Anderson & May, 1985*a*), where, in the simplest case, hosts have a constant probability of infection, and parasites have a constant *per capita* probability of dying per unit time. If the expected rate of infection is uniform across hosts, then such a model predicts a random distribution of parasites among hosts (Anderson & Gordon, 1982). However, when heterogeneity between hosts is introduced, for instance, when rate of exposure or susceptibility varies between hosts, distributions of parasites become aggregated. Aggregated (or, equivalently, overdispersed) distributions are characterized by high variance-to-mean ratios. Theoretical studies have suggested that one commonly used index of the degree of aggregation, the negative binomial parameter, k , will, under some circumstances, be constant with host age (Pacala & Dobson, 1988; Anderson & May, 1991; but see also Haderl & Dietz, 1983; Anderson, 1985; Grenfell, Dietz & Roberts, 1995). A decrease in aggregation with host age in field data has thus been interpreted as a possible indicator of the action of density-dependent processes, such as parasite-induced host mortality or acquired immunity (Pacala & Dobson, 1988). This prediction has been widely, though critically, discussed in relation to field data (Srividya *et al.* 1991; Bratney & I-Hsun Ni, 1992; Dobson *et al.* 1992;

Gregory, 1992; Quinnell, 1992; Woolhouse, Ndamba & Bradley, 1994).

In this paper we develop a more general model for changes in the mean and variance of parasite burden with host age in the absence of density-dependent processes. The model allows consideration of the effects of several realistic complications of simple immigration-death processes.

First, we consider both continuous infection and discrete infection events. Short periods when infection can occur, separated by long intervals with no infection have been reported for several host-parasite systems in the field, perhaps most dramatically for *Pseudodiploorchis americanus*, a monogean parasite of the anuran *Scaphiopus couchii* where transmission only occurs during a spawning period of a few hours every year (Tinsley & Jackson, 1988). Seasonal variations in rates of infection have been widely reported, e.g. for human schistosomes (Wilkins *et al.* 1984; Chandiwana, Christensen & Frandsen, 1987) and for human hookworms (Chandiwana, 1990). Marked temporal, though not necessarily seasonal, variations also occur, e.g. for human hookworms (Hominick, Dean & Schad, 1987).

Second, we consider the variance in the numbers of parasites acquired per infection event, representing this feature in a different way than Pacala & Dobson (1988) and not assuming any particular form of the frequency distribution. Data providing direct estimates of this variance for naive hosts in the field are not available. However, several studies have

shown aggregated spatial distributions of infective stages, e.g. for hookworm larvae (Hominick *et al.* 1987) and for snails releasing schistosome cercariae (Woolhouse & Chandiwana, 1989; Udonsi, 1990). Conversely, data for schistosome cercariae densities (as measured by hamster immersions) have suggested an underdispersed (variance less than mean) distribution (Chandiwana, 1987). An indirect indicator of the overdispersion of infection rates can be obtained from the distribution of parasite burdens in the youngest hosts (see below).

Third, we allow variable predisposition to infection. Predisposition is a measure of the correlation between numbers of parasites acquired by individual hosts at different infection events. In practice, predisposition is normally assessed as the correlation between parasite burdens pre- and post-treatment with anthelmintics. Significant positive correlations have been reported for a number of different human helminth infections, but this correlation is typically much less than one (see Keymer & Pagel, 1990).

Within a framework allowing these complications we examine the effects of discrete infection events and varying infection rate, parasite survival, aggregation of infective stages and degree of predisposition on the relationship between mean parasite burden, parasite aggregation and host age. Model predictions can then be compared with the patterns observed in the field, with special reference to human helminth infections, for which estimates of the necessary parameter values are most available.

MODEL DEVELOPMENT

The definitions of the variables and parameters used in the models developed in this section are summarized, for ease of reference, in Table 1. We consider a large cohort of hosts. An individual host acquires X_i parasites at each time interval i for $i = 0, 1, 2, \dots, t$ where t is the current time. The number of parasites acquired at time i that survive to time t has a binomial distribution with parameters $n = X_i$ and $p = S^{t-i}$ where S is the probability that a parasite survives one time interval, and expected mean life-expectancy is $-1/\ln(S)$. The expected value of X_i , which is equivalent to the mean number of parasites acquired per host at any time point i , is $E(X_i) = \mu$ with variance $Var(X_i) = \sigma^2$. The total number of parasites surviving to time t in one host is denoted Y_t . The expected values of Y_t and its variance are:

$$E[Y_t] = \mu P_t \quad (1)$$

$$Var[Y_t] = \mu P_t + (\sigma^2 - \mu) Q_t + \rho \sigma^2 (P_t^2 - Q_t) \quad (2)$$

with

$$P_t = \sum_{i=0}^t S^i \quad (3)$$

$$Q_t = \sum_{i=0}^t S^{2i} \quad (4)$$

and ρ is the correlation between X_i and X_j where i and j represent different time intervals. ρ is therefore

a measure of predisposition. The derivation of equations (1)–(4) is given in full in the Appendix.

We characterize the distribution of X_i by the moment estimate of the negative binomial parameter, $k = \mu^2/(\sigma^2 - \mu)$, and focus on the behaviour of k_t , the moment estimate of the aggregation parameter for the distribution of Y_t s. We obtain the following expression:

$$\frac{1}{k_t} = \frac{1}{k} \frac{Q_t}{P_t^2} + \rho \left(\frac{1}{k} + \frac{1}{\mu} \right) \left(1 - \frac{Q_t}{P_t^2} \right). \quad (5)$$

Note that the moment estimate of k_t takes this form although the frequency distribution is not expected to be negative binomial (see Anderson, 1974; Grafen & Woolhouse, 1993). This expression for the moment estimate of k_t is obtained for any frequency distribution of the numbers of parasites acquired per host at a given time-point where the mean and the variance are known (see Appendix).

Equation (5) has an interesting special case where:

$$k = \frac{\mu(1-\rho)}{\rho}, \quad (6)$$

which gives $k_t = k$ for all t .

If we consider the behaviour of the model for the limit $t \rightarrow \infty$ we obtain the following expressions for the mean and aggregation of parasites among hosts:

$$E(Y_\infty) = \frac{\mu}{1-S} \quad (7)$$

with

$$\frac{1}{k_\infty} = \frac{1}{k} \frac{(1-S)}{(1+S)} + \rho \left(\frac{1}{k} + \frac{1}{\mu} \right) \left(1 - \frac{1-S}{1+S} \right). \quad (8)$$

We also consider the behaviour of the model as the time interval is subdivided into T time units. As $T \rightarrow \infty$, i.e. we move from discrete to continuous infection, in place of equations (1)–(2) we now have:

$$E[Y_t] = t\mu P_t \quad (9)$$

$$Var[Y_t] = t\mu P_t + t(\sigma^2 - \mu) Q_t + t\rho\sigma^2(tP_t^2 - Q_t), \quad (10)$$

where

$$P_t = \frac{1 - e^{-\alpha t}}{\alpha t} \quad (11)$$

$$Q_t = \frac{1 - e^{-2\alpha t}}{2\alpha t} \quad (12)$$

with $\alpha = -\ln(S)$. The derivation of equations (9)–(12) is given in the Appendix. Equation (9) is equivalent to the standard solution to the immigration-death equation in continuous time (Anderson & May, 1991). If the distribution of X_i is characterized by the constant parameter k , as before, we have the following expression for k_t :

$$\frac{1}{k_t} = \rho \left(\frac{1}{k} + \frac{1}{\mu} \right) + \frac{\alpha}{2} \frac{(1 + e^{-\alpha t})}{(1 - e^{-\alpha t})} \left(\frac{1}{k} - \rho \left(\frac{1}{k} + \frac{1}{\mu} \right) \right), \quad (13)$$

with the limit as $t \rightarrow 0$ then $k_t \rightarrow 0$. As $t \rightarrow \infty$ we obtain:

$$\frac{1}{k_\infty} = \frac{1}{k} \frac{\alpha}{2} + \rho \left(\frac{1}{k} + \frac{1}{\mu} \right) \left(1 - \frac{\alpha}{2} \right), \quad (14)$$

Table 1. Definitions of variables and model parameters

$E[Y_t]$ = expected (mean) number of parasites/host at age t .
$Var[Y_t]$ = variance of expected number of parasites/host at age t .
k_t = negative binomial parameter for the distribution of parasites among hosts at age t .
μ = expected (mean) number of parasites acquired/host/time interval.
λ = rate of change in μ with age, such that $\mu_t = \mu_0 \lambda^t$.
σ^2 = variance of number of parasites acquired/time-interval.
k = negative binomial parameter for the distribution of number of parasites acquired/time-interval, i.e. aggregation of infection.
S = probability of parasite survival/year.
ρ = least squares correlation coefficient between numbers of parasites acquired by hosts at any two different ages t and $t+s$, i.e. predisposition.

except where the special case of equation (6) applies, giving $k_t = k$ for all t .

Finally, we consider a more complex case in discrete time where the expected number of parasites acquired at time i , $E(X_i) = \mu_i$, declines exponentially over time, so that $\mu_i = \mu_0 \lambda^i$ where $\lambda < 1$. If X_i are distributed negative binomially with constant parameter k then $Var(X_i) = \sigma_i^2 = \lambda^{2i}(\sigma_0^2 - \mu_0) + \mu_0 \lambda^i$. As before, the number of parasites acquired at time i that survive to time t has a binomial distribution with $n = X_i$ and $p = S^{t-i}$. The behaviour of this model can be explored by numerical analysis making use of equations (A 1) and (A 2) in the Appendix with these substitutions for μ_i and σ_i^2 .

MODEL APPLICATION

The behaviours of the models are examined for a range of different values of parasite survival (S), mean infection rate (μ), predisposition (ρ) and aggregation of infection (k). We concentrate on biologically meaningful values of these parameters, particularly for the common human helminths, *Ascaris lumbricoides*, *Trichuris trichiura*, hookworms (*Necator americanus* and *Ancylostoma duodenale*) and schistosomes (especially *Schistosoma mansoni* and *S. haematobium*). Some estimates of survival rates, degrees of predisposition to infection and infection rates are available in the literature. The degree of aggregation of infection, k , incorporates aggregation of infective stages and differences in exposure rates and susceptibility to infection. Some information is available on infective stages and exposure (see above), but an estimate of k can also be obtained from field studies as the degree of aggregation of parasite burdens in the very youngest hosts. Typical values of S , μ , ρ and k are given (with references) in Table 2. The time-step is 1 year, corresponding to brief periods of seasonal transmission.

RESULTS

Mean parasite burden and host age

The expected mean parasite burden rises with host age, from $E[Y_0] = \mu$ to an asymptote in older hosts at $E(Y_\infty) = \mu/(1-S)$. This pattern is similar to that arising from continuous infection models (see below) except that the rise is stepped, mean numbers of parasite rising instantaneously at each infection event and then falling exponentially until the next.

Aggregation and host age

Examples of the relationships between k_t and age are shown in Fig. 1. Although k_t , like $E[Y_t]$, changes with age in a stepped fashion, for convenience all graphs are plotted smoothed. Varying the degree of predisposition (ρ), infection rate (μ) and degree of aggregation (k) can produce 3 general patterns for the relationship between k_t and host age. For values of ρ , μ and k such that $\rho < \mu/(\mu+k)$, k_t increases monotonically with age to a plateau in the older age classes; however, if $\rho > \mu/(\mu+k)$ then k_t decreases monotonically with host age, and for the special case (from equation 6) where $\rho = \mu/(\mu+k)$, k_t is constant with respect to host age. This special case is equivalent to Pacala & Dobson's (1988) model (see Appendix). These effects are illustrated by Figs 1A–C. For most typical values of ρ , μ and k , k_t increases with host age, and the magnitude of this increase is most marked for low ρ , high μ and low k . As ρ and k increase, or μ decreases, the magnitude of the change in k_t is reduced, and where ρ and k are very high, or μ is very small, declines in k_t with age are seen. The special case of constant k_t is not illustrated (in Fig. 1A the lower line shows a very slight decline). The effect of parasite survival (S) on the relationship between k_t and host age is shown in Fig. 1D. As S decreases, changes in k_t become less apparent; in the limit case where $S = 0$, k_t is constant for any value of ρ , μ or k .

Aggregation in the oldest age classes

The effects of varying ρ , μ , S and k can also be shown by examining their effect on k_∞ , which will approximate the degree of aggregation in adult hosts. Again, where $\rho < \mu/(\mu+k)$, k_∞ is greater than k ($= k_0$). For $\rho = \mu/(\mu+k)$, $k_\infty = k$, and if $\rho > \mu/(\mu+k)$, $k_\infty < k$. These relationships are illustrated in Fig. 2. As ρ increases k_∞ declines, this decline being most marked for low ρ (Fig. 2A). As μ increases, k_∞ increases asymptotically, except for $\rho = 0$, when k_∞ is independent of μ (Fig. 2B); k_∞ is linearly related to k when $\rho = 0$; for $\rho > 0$ k_∞ increases less rapidly than k (Fig. 2C). As S increases the difference between k_∞ and k increases, especially if ρ is large (Fig. 2D). In the limit case $\rho = 0$ and $S = 1$ k_∞ is infinite, i.e. the distribution becomes random.

Table 2. Some estimates of the survival rate (S), rate of infection (μ), degree of predisposition to infection (ρ) and aggregation of infection (k) for the major human helminths

Species	S (/year)	μ (/year)	ρ	k
<i>Ascaris lumbricoides</i>	0.37–0.61	5–22	0.2–0.5	0.02
<i>Trichuris trichiura</i>	0.37–0.61	90	0.6	< 0.02
Hookworms	0.61–0.78	1–25	0.3	0.04
Schistosomes	0.75–0.90	50	0.3	0.1

Sources: Anderson (1980); Anderson & May (1991); Bradley & McCullough (1973); Bundy *et al.* (1987, 1988); Croll *et al.* (1982); Elkins *et al.* (1986); Haswell-Elkins *et al.* (1988); Keymer & Pagel (1990); Quinnell *et al.* (1993); Sweet (1925); Warren *et al.* (1974); Wilkins *et al.* (1984); Woolhouse, Hasibeder & Chandiwana (unpublished observations).

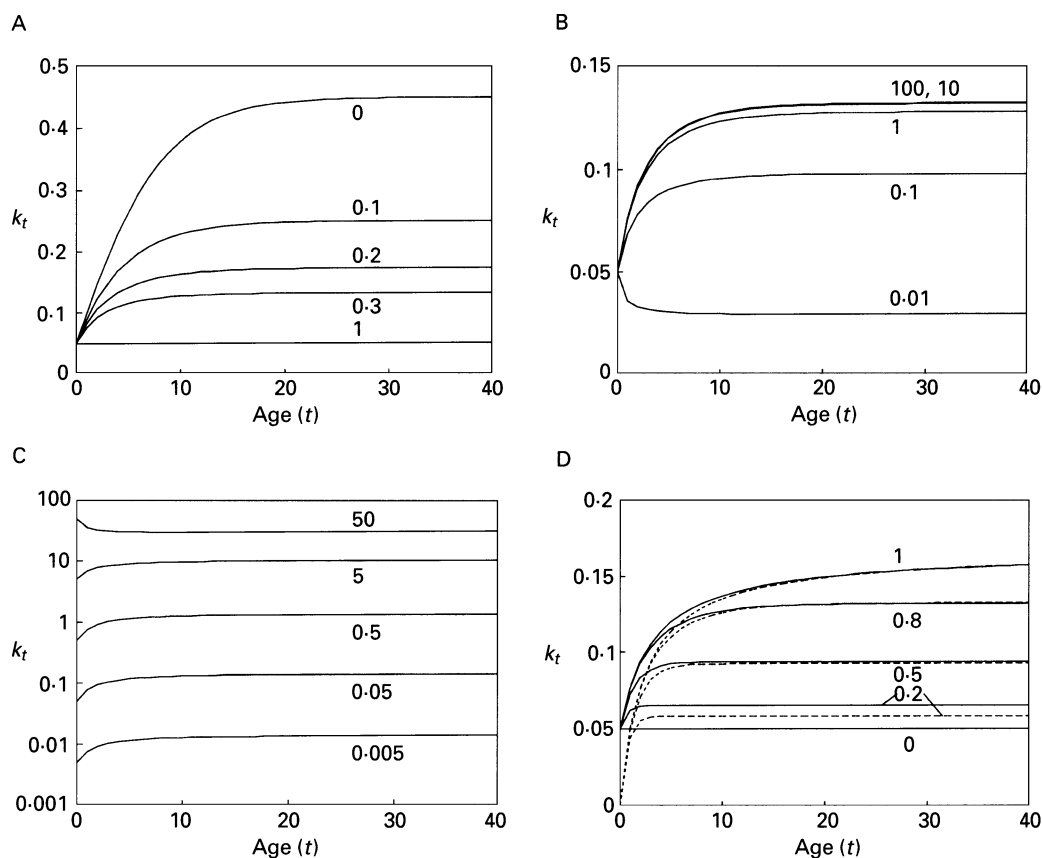


Fig. 1. Changes with age in the aggregation of parasite burdens, indexed by the negative binomial parameter, k_t , as predicted by equation (5) in the main text. Baseline parameter values are $\rho = 0.3$, $\mu = 10$, $S = 0.8$ and $k = 0.05$ with a time-step of 1 year (see Table 2). Effects are shown of varying each parameter value in turn (values shown): (A) degree of predisposition, ρ ; (B) mean rate of infection, μ ; (C) aggregation of infection, k ; (D) parasite survival rate, S , with results for the continuous time model of equation (13) compared (dashed lines). Note the different scales on the vertical axes.

Discrete versus continuous infection

Changing the infection process from discrete to continuous infection does not qualitatively change the relationship between k_t and host age. Again, for $\rho < \mu/(\mu + k)$, k_t increases monotonically with host age, and in the special case where $\rho = \mu/(\mu + k)$, k_t is constant with respect to host age. Values of ρ , μ and

k such that $\rho > \mu/(\mu + k)$ are not valid in the discrete model, since infection is defined to be with integral numbers of parasites (see Appendix). A continuous infection process does, however, quantitatively affect k_t . This effect depends on the value of survivorship (S); when $S = 1$, k_t is the same for continuous and discrete infection; as S decreases, the magnitude of k_t is reduced in the continuous relative to the discrete

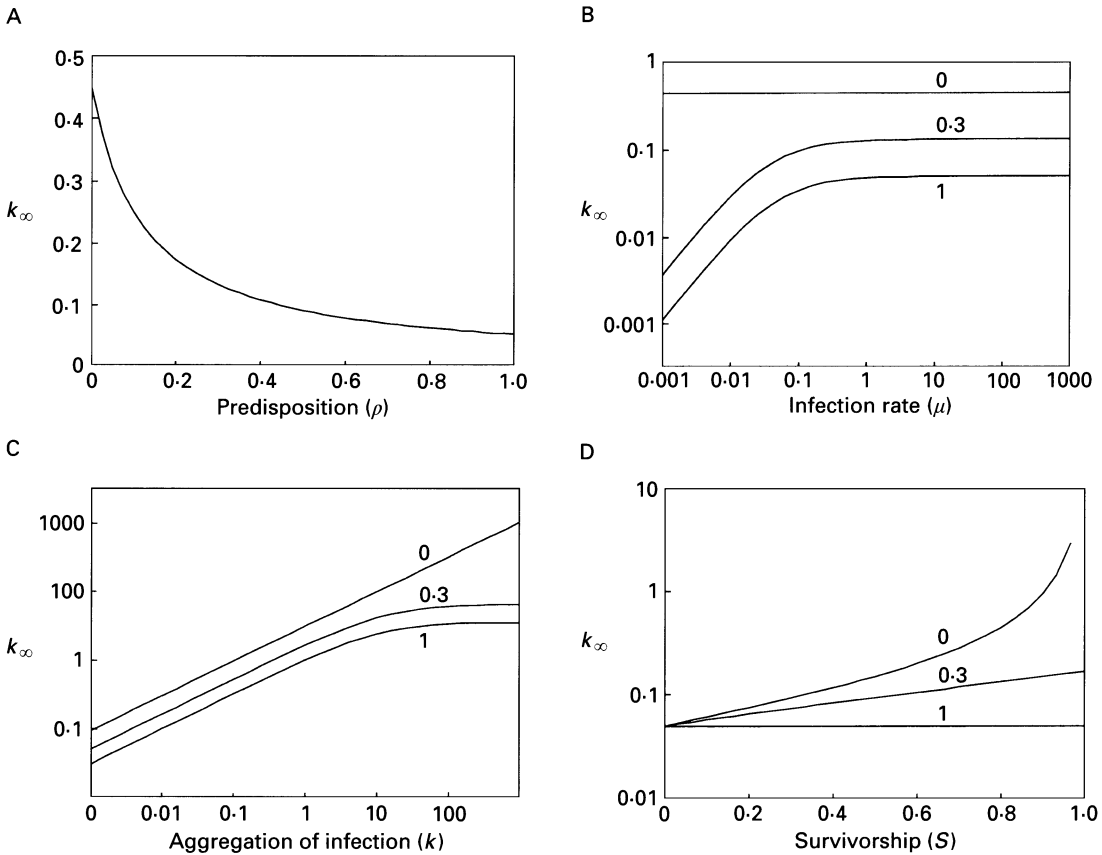


Fig. 2. Aggregation of parasite burdens in the oldest age class, indexed by the negative binomial parameter k_{∞} , as predicted by equation (8) in the main text. Baseline parameter values are $\mu = 10$, $S = 0.8$ and $k = 0.05$ (see Table 2). Relationships are shown of k_{∞} with (A) degree of predisposition, ρ ; (B) mean rate of infection μ ; (C) aggregation of infection, k ; (D) parasite survival rate, S . For (B)–(D) the relationships are compared for 3 different degrees of predisposition, $\rho = 0, 0.3$ and 1. Note the different scales on the vertical axes.

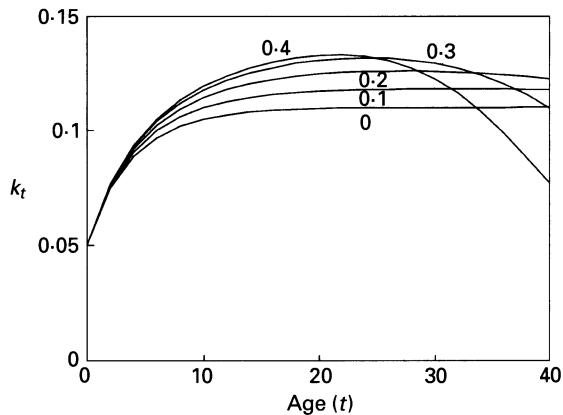


Fig. 3. Changes with age in the aggregation of parasite burdens, indexed by the negative binomial parameter k_t , as predicted by a model allowing age-related changes in the mean rate of infection (see Appendix). Parameter values are $\rho = 0.3$, $\mu_0 = 20$, $S = 0.64$ and $k = 0.05$ with a time-step of 2 years (see Table 2). Effects are shown of varying λ , the rate of decline in infection rate with age.

case (Fig. 1D). The value of k_0 also differs for continuous infection: in the discrete case, the first infection is assumed to occur instantaneously at $t = 0$, hence $k_0 = k$; with continuous infection k_t tends to zero as t tends to zero, but is undefined at $t = 0$.

Age-dependent rate of infection

The effect of age-dependent exposure is illustrated in Fig. 3. As the rate of infection declines more rapidly with host age (λ increases), the magnitude of the initial rise in k_t increases. More importantly, as λ increases the relationship between k_t and age changes from monotonic to convex, with k_t declining in the oldest hosts. The peak value of k_t increases as λ increases.

DISCUSSION

The discrete-infection model used here can produce a wide range of patterns of changes in parasite aggregation with host age (see also Dietz, 1988). In contrast, the special case of the model corresponding to Pacala & Dobson's (1988) model has a much narrower range of behaviour. With no predisposition and a constant rate of infection (and, trivially, non-zero parasite survival), the more general model predicts a monotonic rise in the negative binomial parameter k_t with increasing host age. This prediction is not qualitatively affected by changes in infection rate or the aggregation of exposure, although these parameters do affect the magnitude

and rate of the rise in k_i . Increases in aggregation with age may also be expected if parasite mortality rates differ between hosts rather than infection rates (Anderson & May, 1991). In contrast, when there is very high predisposition, our model predicts that k_i will fall with host age, although the magnitude of this fall is very slight and would not be detectable in practice. This prediction is again robust to changes in other parameters. With intermediate degrees of predisposition k_i typically rises with age, although for very low infection rates or aggregation of infection (high k) falls in k_i are possible. In one special case, the model produces k_i constant with respect to age.

Qualitatively similar behaviour is predicted if infection occurs continuously rather than at discrete intervals, although the constraint that infection occurs with integral numbers of parasites means that falls in k_i with age are not possible. Except in the special case above, the continuous model predicts a monotonic rise in k_i with host age, but the magnitude of this increase is smaller than with discrete infection. Thus, even with no predisposition, the continuous model predicts aggregated parasite distributions through time, resulting from the aggregation of infective stages.

With age-dependent rates of infection the behaviour of the model becomes more complicated. If rates of infection fall, at least over some age range, then the aggregation of parasite burdens may be lowest (highest k_i) in intermediate age classes, with increasing aggregation among older age classes.

Given this wide range of possible patterns, it is relevant to examine the patterns produced using appropriate parameter estimates from field studies. Some degree of predisposition to infection has been found in nearly all studies of helminth epidemiology (Keymer & Pagel, 1900), but the degree of predisposition, ρ , is usually quite low. With moderate predisposition, and typical (and constant) values of μ , the rate of infection, S , the parasite survival rate, and k , the aggregation of infection, the model predicts a monotonic increase in k_i with host age. The magnitude of this increase is sensitive to changes in the values of, especially, S and ρ , and less so changes in the values of μ and k .

Patterns of changes in aggregation of parasite burden with host age will, of course, be affected by the index of aggregation used. We have expressed results in terms of the moment estimate of the negative binomial parameter, denoted k_i . The general model, however, does not assume a negative binomial distribution of parasites, and the expressions we derive are valid for any distribution. We use the negative binomial parameter for two reasons: its widespread use as a descriptive variable for parasite frequency distributions; and its suggested use as a suitable variable with which to investigate density-dependence (Pacala & Dobson,

1988). Use of the negative binomial distribution to examine changes in aggregation within and between populations has been criticized (Taylor, Woiwod & Perry, 1979; Fulford *et al.* 1992), but inspection of equations (1) and (2) does not suggest any simple relationship between variance and mean that is conserved through time; aggregation, however, quantified, will vary. If this result is robust, the analysis of the relationship between aggregation and host age to reveal underlying population processes such as density dependence may not be possible. In the absence of an ideal measure of aggregation, what is important is to understand how the relationship between the mean and variance of parasite burdens varies under the conditions of interest.

The predictions of the model developed here differ from those of Pacala & Dobson (1988), who suggested that, in the absence of density-dependence, k_i would be constant with host age. The difference between the two models arises from an assumption implicit in the latter that hosts only acquire a single parasite at each infection event. Our model is more general in that more than one parasite can be acquired at any one instant of time. Within our model framework this assumption also has consequences for the degree of predisposition: Pacala & Dobson's (1988) model implicitly constrains predisposition to a function of the mean and variance in parasite acquisition whereas our model is more flexible in allowing some variability in predisposition. A more formal comparison of the two models is given in the Appendix. The precise number of parasites acquired by a host at any given instant is not known. However, infective stages may be highly aggregated in the environment (Hominick *et al.* 1987; Woolhouse & Chandiwana, 1989), as is the amount of contact of the host with the contaminated environment (Wong *et al.* 1988; Chandiwana & Woolhouse, 1991). It thus seems reasonable to suppose that hosts may indeed acquire more than one parasite at a time.

Accurate measurement of the variation in the parameter k_i with host age in the field requires large sample sizes to allow fine stratification of the data by age (Pacala & Dobson, 1988; Gregory & Woolhouse, 1993). The number of suitable studies is small, with the interpretation of data often complicated by the use of excreted egg counts as an index of parasite burdens (De Vlas *et al.* 1992). Some examples of age-related patterns for k_i are shown in Fig. 4. For several different human helminths k_i is initially very low, rising to a peak in young adults. Thereafter k_i may plateau or fall in the older age classes. Both types of behaviour are reproduced by the model presented here. Exposure to schistosomes is known to be strongly age dependent, but exposure to geohelminths may be less so (see references above), and this may account for the different patterns observed for k_i . Although neither the precise form of

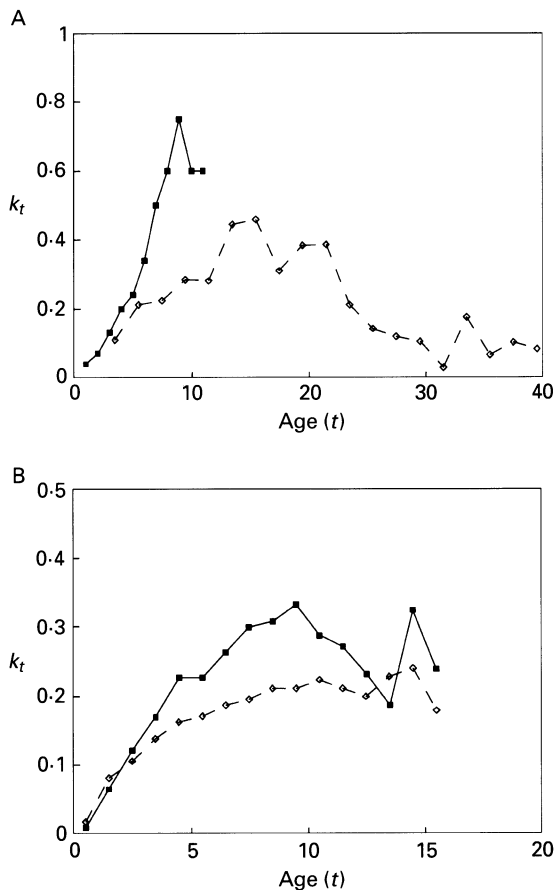


Fig. 4. Changes with age in the aggregation of parasite burdens, indexed by moment estimates of the negative binomial parameter k_t , as observed in field studies of human helminth infections. (A) *Necator americanus* (—), data from Anderson (1980), and *Schistosoma haematobium* in Tanzania (---), data from Bradley & McCullough (1973). (B) *Trichuris trichiura* (—) and *Ascaris lumbricoides* (---) in Malaysia, redrawn from Bundy *et al.* (1988). All parasites show an initial increase in k_t . A decline in k_t among older age classes is apparent for the schistosomes (see also De Vlas *et al.* 1992; Woolhouse *et al.* 1994), but less so for the geohelminths.

the relationship between rate of infection and age used to obtain the results shown in Fig. 3, nor the assumption that the aggregation of numbers of parasites acquired per infection event, k , is invariant with age, are likely to be accurate, the model does serve to demonstrate that the observed patterns can be generated by known processes other than density dependence or acquired immunity.

The model predicts that the magnitude of changes in the aggregation of parasite burdens should covary with mean rate of infection, parasite life-expectancy and predisposition. Unfortunately, there are not sufficient data available to test these hypotheses. It would also be interesting to examine age-related variation in the mean and variance of parasite burdens following reinfection after chemotherapy. The model predicts that all ages show the same aggregation of reinfection parasite burdens.

This prediction may be affected by age-dependent exposure or susceptibility or by density-dependent effects or acquired immunity. Again, large sample sizes would be needed to estimate aggregation indices from reinfection data across a wide age range; such data are not easily available.

In summary, a large number of factors can produce changes in aggregation of parasite burdens with host age. In addition to the effects described here, sampling biases (Pacala & Dobson, 1988; Gregory & Woolhouse, 1993), age-dependent aggregation of exposure (Woolhouse *et al.* 1994) and density dependence, including the effects of acquired immunity (Anderson & Gordon, 1982; Pacala & Dobson, 1988; Woolhouse, 1992), and heterogeneities in parameters other than exposure (Anderson, 1985; Anderson & May, 1985*b*, 1991) may all be involved. In view of this, there is a need for cautious interpretation of field observations.

We would like to thank Hamish McCallum for useful discussions. M. E. J. W. is supported by The Royal Society.

APPENDIX

The purpose of this appendix is to derive equations in the text of the paper, numbers (1)–(4) for the discrete time model, and (9)–(12) for continuous time; and to show what choice of parameter values corresponds to the special case of no clumping in parasite acquisition.

Notes on method

We take a random variable approach, using $E(X)$ and $Var(X)$ to denote the expectation (mean) and variance of a random variable X , and $Cov(X, Y)$ for the covariance between random variables X and Y . One piece of non-standard notation is $Bin(X, p)$, which applies to a random variable X taking values that are non-negative integers and a probability p . $Y = Bin(X, p)$ is shorthand for saying that conditional on X , Y is an independent binomial random variable with parameters X and p .

The results are expressed as propositions. The only non-obvious feature is the use of the following two lemmas relating to the notation $Bin(X, p)$. Proofs of the lemmas and propositions are omitted, as they are very easy.

Lemma 1. Let $X_i, i = 1, 2, \dots, n$ be a sequence of random variables, with joint probability generating function $f(s_1, s_2, \dots, s_n)$. Let $Y = Bin(X_k, p)$. Then the joint probability generating function of $(X_1, \dots, X_{k-1}, Y, X_{k+1}, \dots, X_n)$ is given by g where

$$g(s_1, \dots, s_{k-1}, t, s_{k+1}, \dots, s_n) = f(s_1, \dots, s_{k-1}, 1 - p + pt, s_{k+1}, \dots, s_n).$$

$$\text{Lemma 2. (i) } Var[Bin(X, p)] = p^2 Var[X] + p(1 - p) E[X]$$

$$(ii) Cov[Bin(X, p), Bin(Y, q)] = pq Cov[X, Y].$$

Discrete time model

Proposition 1. Let $X_i, i = 0, 1, \dots$ be random variables over the non-negative integers with means μ_i and variances σ_i^2 ,

and suppose $\mathbf{Cov}[X_i, X_j]/\sigma_i\sigma_j = \rho$ for $i \neq j$. Let $U_i = \text{Bin}(X_i, p_i)$, where the $p_i, i = 0, 1 \dots$ are arbitrary probabilities. Let $Y = \sum_{i=0}^t U_i$.

Then

$$\begin{aligned} \mathbf{E}[Y] &= \sum_{i=0}^t p_i \mu_i \\ \mathbf{Var}[Y] &= \sum_{i=0}^t p_i \mu_i + (1-\rho) \sum_{i=0}^t p_i^2 \sigma_i^2 - \sum_{i=0}^t p_i^2 \mu_i + \rho \left(\sum_{i=0}^t p_i \sigma_i \right)^2. \end{aligned}$$

Corollary. In the special case that all the means are equal to μ , all the variances are equal to σ^2 , and there is a constant chance of survival S from one time-interval to the next, we can substitute $\mu_i = \mu$, $\sigma_i^2 = \sigma^2$, and $p_i = S^i$ to obtain

$$\begin{aligned} \mathbf{E}[Y] &= \mu P \\ \mathbf{Var}[Y] &= \mu P + (\sigma^2 - \mu) Q + \rho \sigma^2 (P^2 - Q), \end{aligned}$$

where $P = \sum_{i=0}^t S^i$, and $Q = \sum_{i=0}^t S^{2i}$.

These are equations (1)–(4) of the text, as required.

Remark. The assumption that the X_i are discrete random variables implies a minimum possible variance for any given mean. If $[x]$ represents the fractional part of x , then $\sigma_i^2 \geq [\mu_i](1 - [\mu_i])$. A similar but more complex restriction applies to the covariance between two discrete random variables. There are consequent restrictions on the values that ρ may take, given the μ_i and σ_i^2 . In particular, ρ can equal 1 only if certain equations hold between the μ_i and σ_i^2 . These equations are unlikely to hold true except in the case of equality of the μ_i and equality of the σ_i^2 .

Continuous time model (as limit of discrete model)

We construct a continuous time process as a limit of discrete time processes, which divide the time interval up into increasingly fine regular partitions. The time intervals in each discrete time process are assumed to have equal means and equal variances.

Proposition 2. Let ${}_T X_i, i = 1 \dots T, T = 1, 2 \dots$ be random variables over the non-negative integers with means μ_T and variances σ_T^2 independent of i . Suppose

$$\mathbf{Cov}[{}_T X_i, {}_T X_j] = \gamma_T \quad \text{for } i \neq j, 1 \leq i, j \leq T, T > 1.$$

Let ${}_T U_i = \text{Bin}({}_T X_i, \exp(-\alpha(i-1)t/T))$, for some $\alpha > 0$. Let ${}_T Y = \sum_{i=1}^T {}_T U_i$ for $T = 1, 2 \dots$. Suppose

$$\begin{aligned} \mu_T &= \mu_1/T \\ \gamma_T &= \gamma_1/T^2 \text{ for some constant } \gamma_1 \\ \sigma_T^2 &= (\sigma_1^2 - (1-1/T)\gamma_1)/T. \end{aligned}$$

Then the processes indexed by T converge to a continuous process on the time interval $[0, t]$ in the sense that for $0 \leq a \leq b \leq c \leq d \leq t$, as $T \rightarrow \infty$,

- (1) $\mathbf{E}\left(\sum_{i:a \leq \frac{i-1}{T}t < b} {}_T X_i\right) \rightarrow \mu_1 \frac{b-a}{t}$.
- (2) $\mathbf{Var}\left(\sum_{i:a \leq \frac{i-1}{T}t < b} {}_T X_i\right) \rightarrow \sigma_1^2 \frac{b-a}{t} - \gamma_1 \frac{b-a}{t} \left(1 - \frac{b-a}{t}\right)$.
- (3) $\mathbf{Cov}\left(\sum_{i:a \leq \frac{i-1}{T}t < b} {}_T X_i, \sum_{i:c \leq \frac{i-1}{T}t < d} {}_T X_i\right) \rightarrow \gamma_1 \left(\frac{b-a}{t}\right) \left(\frac{d-c}{t}\right)$.

Further, taking limits as $T \rightarrow \infty$,

$$\begin{aligned} \text{Lim } \mathbf{E}[{}_T Y] &= \mu_1 P_t \\ \text{Lim } \mathbf{Var}[{}_T Y] &= \mu_1 P_t + (\sigma_1^2 - \mu_1) Q_t + \gamma_1 (P_t^2 - Q_t), \end{aligned}$$

where $P_t = \frac{1}{t} \int_{\tau=0}^t \exp(-\alpha\tau) d\tau = \frac{1 - \exp(-\alpha t)}{\alpha t}$
and $Q_t = \frac{1}{t} \int_{\tau=0}^t \exp(-2\alpha\tau) d\tau = \frac{1 - \exp(-2\alpha t)}{2\alpha t}$.

These are equations (11) and (12) in the main text.

Let μ, σ^2, γ and ρ be the mean, variance, covariance and correlation coefficient for time intervals of unit length. Then

$$\begin{aligned} \mu &= \mu_1/t \\ \sigma^2 &= \frac{\sigma_1^2}{t} - \gamma_1 \frac{1}{t} \left(1 - \frac{1}{t}\right) \\ \gamma &= \frac{\gamma_1}{t^2} \\ \rho &= \frac{\gamma}{\sigma^2} \end{aligned}$$

and

$$\begin{aligned} \text{Lim } \mathbf{E}[{}_T Y] &= t\mu P_t \\ \text{Lim } \mathbf{Var}[{}_T Y] &= t\mu P_t + t((1-\rho)\sigma^2 - \mu)Q_t + \rho\sigma^2 t^2 P_t^2. \end{aligned}$$

These are equations (9) and (10) of the text.

Remark. There is a further result which follows from the discreteness of the ${}_T X_i$. We know that

$$T\sigma_T^2 = \sigma^2 \left(1 - \frac{T-1}{T}\rho\right) \quad \text{and} \quad T\mu_T = \mu.$$

But a discrete random variable with mean $\nu < 1$ and variance τ^2 must satisfy $\tau^2 \leq \nu(1-\nu)$. It follows that $\text{Lim } \sigma_T^2 \geq \text{Lim } \mu_T$, and so in view of the above equalities,

$$\sigma^2(1-\rho) \geq \mu \quad \text{and so substituting using } k = \frac{\mu^2}{\sigma^2 - \mu},$$

$$\text{we obtain } \rho \leq \frac{\mu}{\mu + k}.$$

The first inequality represents the fact that the combination of symmetric time intervals cannot produce sub-Poisson variance. The next subsection formalizes the idea of individual predisposition, and shows we can interpret ρ as the correlation across individuals between the numbers of parasites acquired in two non-overlapping time periods of unit length. This correlation is lower, the higher the within-individual within-time-period variance. The second inequality above represents the fact that the combination of symmetrical time intervals cannot produce sub-Poisson variance at the individual level, implying that the within-individual within-time period variance cannot be less than the corresponding mean. The consequent bound for ρ is attained when that variance is Poisson, an important special case to which we now turn.

The case of absence of clumping

Pacala & Dobson's (1988) model is more general than our continuous case in allowing arbitrarily varying acquisition and death rates. However, our model is more general than

Pacala & Dobson's in allowing clumping of parasite acquisition. Our model follows Pacala & Dobson's assumption of lack of clumping when

$$(1 - \rho)\sigma^2 = \mu.$$

The implications are dramatic. The expressions for k , which in general depend on t , become simple functions of the parameters and so independent of t when this assumption is made. The exact functions are given in the text. We conclude that the 'common k ' result of Pacala & Dobson (1988) relies crucially on their assumption that parasite acquisition is not clumped, or in other words, that at the level of an individual organism, parasite acquisition is a Poisson process.

The $(1 - \rho)\sigma^2 = \mu$ result is proved as follows. Let X_1 and X_2 be random variables representing the numbers of parasites acquired by a random individual in two equal non-overlapping time intervals. Let Y be the expected number of parasites acquired for each individual, which is the same in both time intervals because they are equal. $\text{Var}(Y)$ measures the extent of variation in individual predisposition, and so $\text{Var}(Y) = 0$ would correspond to the case where all individuals are the same. The Poisson property implies that, given the mean, an individual's parasite acquisition is Poisson within each time interval and independent between time intervals. Formally,

$$E(X_i|Y) = Y, \text{Var}(X_i|Y) = Y, \text{Cov}(X_1, X_2|Y) = 0,$$

and so by definition of Var and Cov ,

$$E(X_i^2|Y) = Y^2 + Y, E(X_1 X_2|Y) = E(X_1|Y)E(X_2|Y) = Y^2.$$

We can now calculate the unconditional variances and covariances using

$$\begin{aligned} E(X_i) &= E(Y) \\ E(X_i^2) &= E(Y^2) + E(Y) \\ \text{Var}(X_i) &= E(X_i^2) - E(X_i)^2 = E(Y) + E(Y^2) - E(Y)^2 = \\ &= E(Y) + \text{Var}(Y) \\ \text{Cov}(X_1, X_2) &= E(X_1 X_2) - E(X_1)E(X_2) = E(Y^2) - E(Y)^2 \\ &= \text{Var}(Y). \end{aligned}$$

This allows us to calculate the correlation coefficient ρ between X_1 and X_2 as

$$\rho = \frac{\text{Cov}(X_1, X_2)}{\sqrt{[\text{Var}(X_1)\text{Var}(X_2)]}} = \frac{\text{Var}(Y)}{E(Y) + \text{Var}(Y)}.$$

Writing the mean in each time interval as $\mu = E(X_i)$, and the variance as $\sigma^2 = \text{Var}(X_i)$, it follows as required that

$$(1 - \rho)\sigma^2 = \frac{E(Y)}{E(Y) + \text{Var}(Y)}(E(Y) + \text{Var}(Y)) =$$

$$E(Y) = \mu.$$

This result applies to both the discrete and continuous cases, as in each μ , σ^2 and ρ are defined as applying to equal time intervals. Eliminating σ^2 from $(1 - \rho)\sigma^2 = \mu$ in favour of k defined (when $\sigma^2 > \mu$) as $\mu^2/(\sigma^2 - \mu)$ yields equation (6) of the main text.

REFERENCES

- ANDERSON, R. M. (1974). Mathematical models of host-helminth parasite interactions. In *Ecological Stability* (ed. Usher, M. B. & Williamson, M. H.), pp. 43-69. London: Chapman and Hall.
- ANDERSON, R. M. (1980). The dynamics and control of direct life cycle helminth parasites. *Lecture Notes in Biomathematics* **39**, 278-322.
- ANDERSON, R. M. (1985). Mathematical models in the study of the epidemiology and control of ascariasis in man. In *Ascariasis and its Public Health Importance* (ed. Crompton, D. W. T., Nesheim, M. C. & Pawlowski, Z. S.), pp. 39-67. London: Taylor & Francis.
- ANDERSON, R. M. & GORDON, D. M. (1982). Processes influencing the distribution of parasite numbers within host populations with special emphasis on parasite-induced host mortalities. *Parasitology* **85**, 373-98.
- ANDERSON, R. M. & MAY, R. M. (1985a). Helminth infections of humans: mathematical models population dynamics and control. *Advances in Parasitology* **24**, 1-101.
- ANDERSON, R. M. & MAY, R. M. (1985b). Herd immunity to helminth infection and implications for parasite control. *Nature, London* **315**, 493-6.
- ANDERSON, R. M. & MAY, R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford: Oxford University Press.
- BRADLEY, D. J. & McCULLOUGH, F. S. (1973). Egg output stability and the epidemiology of *Schistosoma haematobium*. Part II. An analysis of the epidemiology of *S. haematobium*. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **67**, 491-500.
- BRATTEY, J. & I-HSUN NI (1992). Ascaridoid nematodes from the stomach of harp seals, *Phoca groenlandica*, from Newfoundland and Labrador. *Canadian Journal of Fisheries and Aquatic Science* **49**, 956-66.
- BUNDY, D. A. P., COOPER, E. S., THOMPSON, D. E., ANDERSON, R. M. & DIDIER, J. M. (1987). Age-related prevalences and intensity of *Trichuris trichiura* infection in a St. Lucian community. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **81**, 85-94.
- BUNDY, D. A. P., KAN, S. P. & ROSE, R. (1988). Age-related prevalence, intensity and frequency distribution of gastrointestinal helminth infection in urban slum children from Kuala Lumpur Malaysia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **82**, 289-94.
- CHANDIWANA, S. K. (1987). Analysis of the Dynamics of Transmission of Human Schistosomiasis in Two Rural Communities in Zimbabwe. Ph.D. thesis, University of Zimbabwe.
- CHANDIWANA, S. K. (1990). Hookworm population ecology in Zimbabwe. In *Hookworm Disease: Current Status and New Directions* (ed. Schad, G. A. & Warren, K. S.), pp. 165-176. London: Taylor & Francis.
- CHANDIWANA, S. K., CHRISTENSEN, N. O. & FRANSEN, F. (1987). Seasonal patterns in the transmission of *Schistosoma haematobium*, *S. mattheei* and *S. mansoni* in the highveld region of Zimbabwe. *Acta Tropica* **44**, 433-44.
- CHANDIWANA, S. K. & WOOLHOUSE, M. E. J. (1991). Heterogeneities in water contact patterns and the epidemiology of *Schistosoma haematobium*. *Parasitology* **103**, 363-70.

- ROLL, N. A., ANDERSON, R. M., GYARKOS, T. W. & GHADIRIAN, E. (1982). The population biology and control of *Ascaris lumbricoides* in a rural community in Iran. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **76**, 187–97.
- DE VLAS, S. J., GRYSSELS, B., VAN OORTMARSEN, G. J., POLDERMAN, A. M. & HABBEMA, J. D. F. (1992). A model for variations in single and repeated egg counts in *Schistosoma mansoni* infections. *Parasitology* **104**, 451–60.
- DIETZ, K. (1988). Mathematical models. In *Malaria*, Vol. 2 (ed. Wernsdorfer, W. H. & McGregor, I.), pp. 1091–134. Edinburgh: Churchill-Livingstone.
- DOBSON, A. P., PACALA, S. V., ROUGHGARDEN, J. D., CARPER, E. R. & HARRIS, E. A. (1992). The parasites of anolis lizards in the northern Lesser Antilles. 1. Patterns of distribution and abundance. *Oecologia* **91**, 110–17.
- ELKINS, D. B., HASWELL-ELKINS, M. R. & ANDERSON, R. M. (1988). The importance of host age and sex to patterns of reinfection with *Ascaris lumbricoides* following mass anthelmintic treatment in a South Indian fishing community. *Parasitology* **96**, 171–84.
- FULFORD, A. J. C., BUTTERWORTH, A. E., STURROCK, R. F. & OUMA, J. H. (1992). On the use of age-intensity data to detect immunity to parasitic infections with special reference to *Schistosoma mansoni* in Kenya. *Parasitology* **105**, 219–28.
- GRAFEN, A. & WOOLHOUSE, M. E. J. (1993). Does the negative binomial distribution add up? *Parasitology Today* **9**, 475–7.
- GREGORY, R. D. (1992). On the interpretation of host-parasite ecology: *Heligmosomoides polygyrus* (Nematoda) in wild wood mouse (*Apodemus sylvaticus*) populations. *Journal of Zoology* **226**, 109–21.
- GREGORY, R. D. & WOOLHOUSE, M. E. J. (1993). Quantification of parasite aggregation: a simulation study. *Acta Tropica* **54**, 131–9.
- GRENFELL, B. T., DIETZ, K. & ROBERTS, M. G. (1995). Modelling the immuno-epidemiology of macroparasites in naturally-fluctuating host populations. In *Ecology of Infectious Diseases in Natural Populations* (ed. Grenfell, B. T. & Dobson, A. P.), pp. 362–383. Cambridge: Cambridge University Press.
- HADELER, K. P. & DIETZ, K. (1983). Nonlinear hyperbolic partial differential equations for the dynamics of parasite populations. *Computers and Mathematics with Applications* **9**, 415–30.
- HASWELL-ELKINS, M. R., ELKINS, D. B., MANJULA, K., MICHAEL, E. & ANDERSON, R. M. (1988). An investigation of hookworm infection and reinfection following mass anthelmintic treatment in the South Indian fishing community of Vairavankuppam. *Parasitology* **96**, 565–77.
- HOMINICK, W. M., DEAN, C. G. & SCHAD, G. A. (1987). Population biology of hookworms in West Bengal: analysis of infective larvae recovered from damp pads applied to the soil surface at defaecation sites. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **81**, 978–86.
- KEYMER, A. E. & PAGEL, M. (1990). Predisposition to helminth infection. In *Hookworm Disease: Current Status and New Directions* (ed. Schad, G. A. & Warren, K. S.), pp. 177–209. London: Taylor & Francis.
- PACALA, S. W. & DOBSON, A. P. (1988). The relation between the number of parasites/host and host age: population dynamic causes and maximum likelihood estimation. *Parasitology* **96**, 197–210.
- QUINNELL, R. J. (1992). The population dynamics of *Heligmosomoides polygyrus* in an enclosure population of wood mice. *Journal of Animal Ecology* **61**, 669–79.
- QUINNELL, R. J., SLATER, A. F. G., TIGHE, P., WALSH, E. A., KEYMER, A. E. & PRITCHARD, D. I. (1993). Reinfection with hookworm after chemotherapy in Papua New Guinea. *Parasitology* **106**, 379–85.
- SRIVIDYA, A., KRISHNAMOORTHY, K., SABESAN, S., PANICKER, A. N., GRENFELL, B. T. & BUNDY, D. A. P. (1991). Frequency distribution of *Brugia malayi* microfilariae in human populations. *Parasitology* **102**, 207–12.
- SWEET, W. C. (1925). Average egg count per gm faeces per female hookworm in Ceylon. *Journal of Parasitology* **12**, 39–42.
- TAYLOR, L. R., WOIWOD, I. P. & PERRY, J. N. (1979). The negative binomial as a dynamic ecological model for aggregation and the density dependence of *k*. *Journal of Animal Ecology* **48**, 289–304.
- TINSLEY, R. C. & JACKSON, H. C. (1988). Pulsed transmission of *Pseudodiploorchis americanus* (Monogenea) between desert hosts (*Scaphiopus couchii*). *Parasitology* **97**, 437–53.
- UDONSI, J. K. (1990). Human community ecology of urinary schistosomiasis in relation to snail vector bionomics in the Igwun river basin of Nigeria. *Tropical Medicine and Parasitology* **41**, 131–5.
- WARREN, K. S., MAHMOUD, A. A. F., CUMMINGS, P., MURPHY, D. J. & HOUSER, D. B. (1974). Schistosomiasis mansoni in Yemeni in California: duration of infection, presence of disease and therapeutic management. *American Journal of Tropical Medicine and Hygiene* **23**, 902–9.
- WILKINS, H. A., GOLL, P. H., MARSHALL, T. F. DE C. & MOORE, P. J. (1984). Dynamics of *Schistosoma haematobium* infection in a Gambian community. III. The acquisition and loss of infection. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **78**, 227–32.
- WONG, M. S., BUNDY, D. A. P. & GOLDEN, M. H. N. (1988). Quantitative assessment of geographic behaviour as a potential source of exposure to geohelminth infection. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **82**, 621–5.
- WOOLHOUSE, M. E. J. (1992). A theoretical framework for the immunoepidemiology of human helminth infection. *Parasite Immunology* **14**, 563–78.
- WOOLHOUSE, M. E. J. & CHANDIWANA, S. K. (1989). Spatial and temporal heterogeneity in the population dynamics of *Bulinus globosus* and *Biomphalaria pfeifferi* and in the epidemiology of their infection with schistosomes. *Parasitology* **98**, 21–34.
- WOOLHOUSE, M. E. J., NDAMBA, J. & BRADLEY, D. J. (1994). On the interpretation of intensity of infection data for *Schistosoma haematobium*. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **88**, 520–6.