

Patterns of macroparasite aggregation in wildlife host populations

D. J. SHAW^{1*}, B. T. GRENFELL¹ and A. P. DOBSON²

¹*Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK*

²*Department of Ecology and Evolutionary Biology, Eno Hall, Princeton University, Princeton, New Jersey, USA*

(Received 27 March 1998; revised 22 June 1998; accepted 22 June 1998)

SUMMARY

Frequency distributions from 49 published wildlife host–macroparasite systems were analysed by maximum likelihood for goodness of fit to the negative binomial distribution. In 45 of the 49 (90%) data-sets, the negative binomial distribution provided a statistically satisfactory fit. In the other 4 data-sets the negative binomial distribution still provided a better fit than the Poisson distribution, and only 1 of the data-sets fitted the Poisson distribution. The degree of aggregation was large, with 43 of the 49 data-sets having an estimated k of less than 1. From these 49 data-sets, 22 subsets of host data were available (i.e. host data could be divided by either host sex, age, where or when hosts were sampled). In 11 of these 22 subsets there was significant variation in the degree of aggregation between host subsets of the same host–parasite system. A common k estimate was always larger than that obtained with all the host data considered together. These results indicate that lumping host data can hide important variations in aggregation between hosts and can exaggerate the true degree of aggregation. Wherever possible common k estimates should be used to estimate the degree of aggregation. In addition, significant differences in the degree of aggregation between subgroups of host data, were generally associated with significant differences in both mean parasite burdens and the prevalence of infection.

Key words: negative binomial, aggregation, macroparasites, wildlife.

INTRODUCTION

The basic measure of parasite burden is the frequency distribution of parasites within a host population. If all host–parasite encounters were random and there were no host effects on parasite numbers or parasite-induced host mortality, and parasite burdens quickly reach equilibrium with different host ages, then parasite burdens would be expected to follow a Poisson distribution, where the variance is equal to the mean ($\sigma^2 = \mu$). In practice, burdens are generally aggregated (that is most of the hosts have very low parasite burdens, and a few hosts have very high burdens) (e.g. Crofton, 1971; Pennycuik, 1971*b*; Anderson & May, 1978; Anderson & Gordon, 1982; Dietz, 1982; Dobson, 1985; Grenfell, Smith & Anderson, 1986; Pacala & Dobson, 1988; Guyatt & Bundy, 1991; Shaw & Dobson, 1995). The degree of aggregation can be measured in a number of ways. One statistical distribution that is extensively used to describe parasite aggregation, is the negative binomial distribution (NBD) (e.g. Pennycuik, 1971*a*; May & Anderson, 1978; Dietz, 1982; Keymer, 1982; Kennedy, 1984; Scott, 1987; Guyatt *et al.* 1990).

* Corresponding author: Centre de Recherche sur l'Épidémiologie des Désastres (CRED), Ecole de Santé Publique, Université Catholique de Louvain, 30–34, Clos Chappelle aux Champs, 1200 Bruxelles, Belgium. Tel/Fax: +32 2 764 3368/3441.

This general use may bring about the situation where parasite burdens are *de facto* assumed to fit the NBD. In human parasite infections, this assumption has been examined in detail and found to be justified (e.g. Anderson & May, 1982; Bundy & Medley, 1992; Maizels *et al.* 1993). However, no comparable review of animal–parasite infections has been undertaken. Such an analysis is the aim of this paper.

The NBD is described by 2 parameters, μ – the mean, and k – an inverse measure of aggregation (Fisher, 1941). The parameter k can be estimated crudely using either prevalence of infection, or s^2 and \bar{x} . A more accurate estimate can be obtained by maximum likelihood techniques applied to a frequency distribution of parasites within a host population (Bliss & Fisher, 1953; Anderson & Gordon, 1982; Dietz, 1982; Lawless, 1987; Pacala & Dobson, 1988).

We use 49 host–parasite frequency distributions obtained from a literature review of over 250 parasite infections of wildlife (Shaw & Dobson, 1995), and make maximum likelihood estimates of the degree of parasite aggregation. In all cases, the fit of the data-sets to the NBD was compared to the fit to the Poisson distribution (in effect acting as a null random distribution).

For a given infection, there are often differences in the sex or age of hosts and where and when the hosts were sampled. That is, the degree of aggregation is not likely to be homogeneous in all host groups. In

fact, most of the mechanisms that have been proposed to generate the observed aggregated distribution of parasites depend on heterogeneities in infection rates between particular groups of hosts. If encounters with parasites were Poisson distributed within each host group, this would lead to a compound Poisson distribution; where the Poisson mean varies between groups. The NBD is a flexible compound Poisson distribution, with gamma-distributed means (Feller, 1943; Anderson, 1976*a*). A number of authors (Crofton, 1971; Anderson & Gordon, 1982; Pacala & Dobson, 1988; Grenfell, 1988, 1992; Grenfell *et al.* 1986, 1990, 1995*a*) have proposed biological explanations for such a distribution. (1) A series of random infections, with different densities of infectious stages (Feller, 1943; Anderson, 1976*a*). (2) Host individuals vary in susceptibility to infection, e.g. with sex, or age (Pennycuik, 1971*b*; Anderson, 1974; Anderson, 1976*b*; Anderson & May, 1978; Anderson & Gordon, 1982; Guyatt *et al.* 1990). (3) Non-random distribution of infective stages in the habitat, e.g. laid in clutches, or aggregated in the vicinity of host individuals (Quenouille, 1949; Keymer & Anderson, 1979; Anderson & Gordon, 1982). Given the above diversity in mechanisms which all can generate heterogeneity, an analysis of variations in aggregation *within* host populations needs to be undertaken (Anderson & Gordon, 1982; Dietz, 1982; Scott, 1988; Pacala & Dobson, 1988). Fourteen of the 49 data-sets could also be split according to the following host characteristics (Table 2). (1) Host age (Chubb, 1963, 1964; James & Llewellyn, 1967; Folstad *et al.* 1991), (2) host sex (Folstad *et al.* 1991; James & Llewellyn, 1967), (3) where hosts were sampled (Folstad *et al.* 1991; Stromberg, Toussant & Dubey, 1978), (4) when hosts were sampled (Boxshall, 1974; Stromberg & Crites, 1975; Skorping, 1981; Evans, 1983; Bratley, 1988; Folstad *et al.* 1991; Kennedy, 1984). These 14 host-parasite systems were therefore subdivided according to the host characteristics provided for each infection (sometimes more than 1 host characteristic was present), with the net result 22 sets of data containing host subsets (Table 2).

We then require a method for testing the significance of heterogeneities in aggregation between host groups. This problem is complicated by the fact that the mean burden, as well as k , may vary with host sex, age etc (Pacala & Dobson, 1988). As described below, we can express this method as an extension of analysis of variance, in which the degree of aggregation may vary between groups as well as the mean. As a shorthand, we call this procedure 'analysis of dispersion'.

In the following sections, we begin by introducing the data-sets and methods of analysis. The first part of the results is concerned with whether NBD provides the best fit to the observed frequency

distributions. Analysis of dispersion is then applied to the 22 data-sets outlined above. In addition, in order to help the interpretation of the results, we also analysed patterns of prevalence and means, which influence estimates of aggregation (Scott, 1987). Finally, all the results are then brought together in the discussion.

MATERIALS AND METHODS

Data-sets

All but 1 of the 49 data-sets are taken from published papers on wildlife host-parasite systems. Only parasite counts were used – no egg count data were considered. The criteria for selection was a frequency distribution of the number of parasites/host that was sufficiently complete to apply maximum likelihood techniques. The data-sets (Table 1) were divided by host type as follows. (1) Five invertebrate-parasite systems ([1] *Orconectes rusticus* (Stromberg *et al.* 1978), [2] *Herpetocypris reptans* (Evans, 1983), [15] *Pisidium subtruncatum*, and [16] *Valvata piscinalis* (Evans, Whitfield & Dobson, 1981), and [17] *Culicoides crepuscularis* (Schmid & Robinson, 1972)). (2) Fifteen fish-parasite systems ([3–5] *Perca fluviatilis* (Bratley, 1988; Chubb, 1964; Skorping, 1981), [6] *Esox lucius* (Chubb, 1963), [7] *Dorosoma cepedianum* (Stromberg & Crites, 1975), [8] *Pleuronectes platessa* (Boxshall, 1974), [9] *Anguilla anguilla* (Kennedy, 1984), [18] *Sebastes marinus*, and [19] *Sebastes mentella* (Williams, 1963), [20] *Tristopterus luscus* (Evans *et al.* 1983), [21] *Carassius auratus* (Breyev, 1973), [22] *Rutilus rutilus* (Evans, 1978), and [23–25] *Salvelinus alpinus* (Henricson, 1977; Halvorsen & Andersen, 1984)). (3) Four amphibian-parasite systems ([26] *Rana nigromaculata*, [27, 28] *Bufo bufo gargazins* (Li & Hsu, 1951) and [29] *Notophthalmus viridescens* (Jarroll, 1979)). (4) Thirteen bird-parasite systems ([10, 38] *Sturnus vulgaris* and [11, 39] *Turdus musicus* (James & Llewellyn, 1967), and [30–37] *Gallus gallus* (Hodasi, 1969; Li & Hsu, 1951)). (5) Twelve mammal-parasite systems ([13] *Pedetes capensis* (Butynski, 1982), [14] *Rangifer tarandus* (Folstad *et al.* 1991), [40] *Peromyscus maniculatus* (Ignoffo, 1959), [41, 42] *Mus mus*, [43, 44] *Rattus rattus*, and [46–48] *Canis canis* (Li & Hsu, 1951), [45] *Mustela vison* (Nuorteva, 1966) and [49] *Ovis ovis* (Milne, 1943)).

In a number of the host-parasite systems (‡ in Tables 1 and 2) parasite categories were lumped (e.g. hosts with 10–20 parasites were counted as 1 observed frequency count), and the maximum likelihood fits to the NBD were calculated from these lumped categories.

The negative binomial distribution

Only a relatively brief outline of the fitting of the NBD is presented here, but see Bliss & Fisher

Table 1. Summary of fits to the NBD

(*P* values for goodness of fit were calculated as described in the text; \hat{k} – maximum likelihood estimate of *k*; *N* – host sample size; %inf – prevalence of infection; \bar{x} – arithmetic mean burden; PD – *P* values for goodness of fit of data to a Poisson distribution.)

Host species	Parasite species	NBD	\hat{k}	<i>N</i>	%inf	\bar{x}	PD
1 <i>Orconectes rusticus</i>	<i>Paragonimus kellicotti</i>	0.006	0.676	796	58.7	1.775	< 0.001
2 <i>Herpetocypris reptans</i>	<i>Hymenolepis tenerrima</i>	0.072	0.431	2271	10.3	0.125	< 0.001
3 <i>Perca fluviatilis</i>	<i>Triaenophorus nodulosus</i>	0.010	6.014	490	57.6	0.873	< 0.001
4 <i>Perca fluviatilis</i>	<i>Acanthocephalus lucii</i>	0.985	0.542	412	79.6	7.964	< 0.001
5 <i>Perca fluviatilis</i> ‡	<i>Bunodera luciopercae</i>	0.547	0.228	212	46.2	3.266	< 0.001
6 <i>Esox lucius</i>	<i>Triaenophorus nodulosus</i>	0.019	0.650	99	75.8	9.000	< 0.001
7 <i>Dorosoma cepedianum</i>	<i>Camallanus oxycephalus</i>	0.270	2.959	238	56.4	0.979	0.011
8 <i>Pleuronectes platessa</i>	<i>Lepeophtheirus pectoralis</i>	0.152	0.792	2261	63.2	1.993	< 0.001
9 <i>Anguilla anguilla</i>	<i>Acanthocephalus clavula</i>	0.407	0.502	140	60.0	2.450	< 0.001
10 <i>Sturnus vulgaris</i>	<i>Porrocaecum ensicaudatum</i>	0.605	0.356	122	27.0	0.500	< 0.001
11 <i>Turdus musicus</i>	<i>Aploparakis dukardinii</i>	0.159	0.252	159	22.6	0.453	< 0.001
12 <i>Turdus musicus</i> ‡	<i>Prosthorhynchus cylindraceus</i>	0.025	0.329	159	49.1	2.327	< 0.001
13 <i>Pedetes capensis</i> †	<i>Physaloptera capensis</i>	0.290	0.124	404	42.1	10.983	< 0.001
14 <i>Rangifer tarandus</i> ‡	<i>Hypoderma tarandi</i>	0.253	1.083	2597	98.8	59.077	< 0.001
15 <i>Pisidium subtruncatum</i>	<i>Echinoparyphium recurvatum</i>	0.547	0.414	374	40.7	1.441	< 0.001
16 <i>Valvata piscinalis</i>	<i>Echinoparyphium recurvatum</i>	0.324	0.843	954	71.2	3.038	< 0.001
17 <i>Culicoides crepuscularis</i>	<i>Chandlerella quiscali</i>	0.305	0.471	143	59.0	2.594	< 0.001
18 <i>Sebastes marinus</i>	<i>Chondracanthopsis nodusus</i>	0.104	0.223	585	15.2	0.244	< 0.001
19 <i>Sebastes mentella</i>	<i>Chondracanthopsis nodusus</i>	< 0.001	0.607	1380	31.9	0.527	< 0.001
20 <i>Tristopterus luscus</i>	<i>Lernaocera lусi</i>	0.121	0.774	293	34.8	0.573	< 0.001
21 <i>Carassius auratus</i>	<i>Digamma interrupta</i>	0.300	0.118	102	11.8	0.225	< 0.001
22 <i>Rutilus rutilus</i> ‡	<i>Asymphylogora kubanicum</i>	0.059	0.294	153	62.1	7.164	< 0.001
23 <i>Salvelinus alpinus</i> ‡	<i>Diphyllobothrium dendriticum</i>	0.948	2.102	50	98.0	12.934	< 0.001
24 <i>Salvelinus alpinus</i> ‡	<i>Diphyllobothrium ditremum</i>	0.483	1.366	50	98.0	60.379	< 0.001
25 <i>Salvelinus alpinus</i> ‡	<i>Diphyllobothrium ditremum</i>	< 0.001	0.521	235	80.0	7.525	< 0.001
26 <i>Rana nigromaculata</i> ‡	<i>Spiroxys japonica</i>	0.544	0.038	690	12.6	1.298	< 0.001
27 <i>Bufo bufo gargarzins</i> ‡	<i>Baerietta baeri</i>	0.035	0.085	317	15.8	0.575	< 0.001
28 <i>Bufo bufo gargarzins</i> ‡	<i>Oswaldocruzia peipingensis</i>	0.003	0.258	162	62.3	11.181	< 0.001
29 <i>Notophthalmus viridescens</i> ‡	<i>Bothriocephalus rarus</i>	0.028	0.283	272	55.5	4.490	< 0.001
30 <i>Gallus gallus</i>	<i>Raillietina echinobothrida</i>	0.197	0.653	108	85.2	11.021	< 0.001
31 <i>Gallus gallus</i>	<i>Raillietina tetragona</i>	0.964	0.266	108	58.3	7.639	< 0.001
32 <i>Gallus gallus</i>	<i>Ascaridia galli</i>	0.021	0.167	108	26.9	0.926	< 0.001
33 <i>Gallus gallus</i>	<i>Cheilospirura hamulosa</i>	0.023	0.186	108	25.0	0.704	< 0.001
34 <i>Gallus gallus</i>	<i>Tetrameres fissipina</i>	0.823	0.130	108	32.4	2.722	< 0.001
35 <i>Gallus gallus</i> ‡	<i>Heterakis gallinae</i>	0.010	0.208	183	55.7	9.111	< 0.001
36 <i>Gallus gallus</i> ‡	<i>Ascaridia galli</i>	0.011	0.246	201	58.7	10.109	< 0.001
37 <i>Gallus gallus</i> ‡	<i>Cheilospirura hamulosa</i>	0.012	0.188	254	42.1	3.532	< 0.001
38 <i>Sturnus vulgaris</i>	<i>Brachylaemus fuscatus</i>	0.037	0.100	122	20.5	0.869	< 0.001
39 <i>Turdus musicus</i>	<i>Brachylaemus fuscatus</i>	< 0.001	0.126	167	31.7	2.701	< 0.001
40 <i>Peromyscus maniculatus</i>	<i>Hoplopleura hesperomydis</i>	0.066	0.242	188	48.9	3.205	< 0.001
41 <i>Mus mus</i> ‡	<i>Syphacia obvelata</i>	0.276	0.271	150	68.0	24.117	< 0.001
42 <i>Mus mus</i> ‡	<i>Aspiculum tetraptera</i>	0.041	0.257	133	74.4	39.843	< 0.001
43 <i>Rattus rattus</i> ‡	<i>Heterakis spumosa</i>	0.023	0.297	165	60.6	8.477	< 0.001
44 <i>Rattus rattus</i> ‡	<i>Cysticercus fasciolaris</i>	< 0.001	0.640	358	46.1	0.938	< 0.001
45 <i>Mustela vison</i>	<i>Corynosoma semerne</i>	0.547	1.958	190	75.3	2.242	< 0.001
46 <i>Canis canis</i> ‡	<i>Ancylostoma caninum</i>	0.001	0.233	202	46.0	3.446	< 0.001
47 <i>Canis canis</i> ‡	<i>Dirofilaria immitis</i>	0.009	0.315	697	51.9	3.017	< 0.001
48 <i>Canis canis</i> ‡	<i>Diplyidium caninum</i>	0.001	0.307	101	63.4	5.548	< 0.001
49 <i>Ovis ovis</i>	<i>Ixodes ricinus</i>	0.998	0.590	492	88.2	16.364	< 0.001

‡ Probabilistic frequency distributions fitted to ‘lumped’ categories.

(1953), Elliott (1977), Pielou (1969), or Southwood (1966) for a more detailed background. Fitting the NBD required estimates of the mean, μ , and *k*; *k* is an inverse measure of aggregation. If *k* approaches infinity, the NBD converges to a Poisson (i.e. $\sigma^2 = \mu$), but if *k* approaches zero, then the NBD converges

to the logarithmic series distribution (Fisher, Corbet & Williams, 1943; Elliott, 1977). The most accurate way to calculate *k* from observed frequency data is by ML. Bliss & Fisher (1953) proposed an iterative method for this; however, here we achieve this by maximizing the log likelihood directly. Neglecting

Table 2. Summary of analysis of dispersion for various subsets of host data

(N_s – the number of subsets host data divided into; k_c – estimate of common k when differences between mean burdens is taken into account; P value – significance level associated with the common k analysis test; \bar{x} – significance level associated with the common mean analysis test; % – significance level associated with prevalence of infection in subsets.)

Parasite species	Subset	N_s	k_c	P value	\bar{x}	%	
<i>Orconectes rusticus</i>	<i>Paragonimus kellicotti</i>	Place	2	0.852	< 0.001	< 0.001	< 0.001
<i>Herpetocypris reptans</i>	<i>Hymenolepis tenerrima</i>	Time	12	0.807	0.003	< 0.001	< 0.001
<i>Perca fluviatilis</i>	<i>Triaenophorus nodulosus</i>	Age	3	6.800	< 0.001	0.161	< 0.001
<i>Perca fluviatilis</i>	<i>Acanthocephalus luci</i>	Time	6	0.632	0.030	0.001	0.004
<i>Perca fluviatilis</i> ‡	<i>Bunodera luciopercae</i>	Time	3	0.289	0.005	< 0.001	< 0.001
<i>Esox lucius</i>	<i>Triaenophorus nodulosus</i>	Age	2	1.502	< 0.001	< 0.001	< 0.001
<i>Dorosoma cepedianum</i>	<i>Camallanus oxycephalus</i>	Time	3	4.145	0.168	< 0.001	< 0.001
<i>Pleuronectes platessa</i>	<i>Lepeophtheirus pectoralis</i>	Time (year)	2	0.810	< 0.001	< 0.001	< 0.001
<i>Pleuronectes platessa</i>	<i>Lepeophtheirus pectoralis</i>	Month (1972)	12	2.351	< 0.001	< 0.001	< 0.001
<i>Pleuronectes platessa</i>	<i>Lepeophtheirus pectoralis</i>	Month (1973)	12	2.922	0.112	< 0.001	< 0.001
<i>Anguilla anguilla</i>	<i>Acanthocephalus clavula</i>	Time	3	0.507	0.799	0.667	0.970
<i>Sturnus vulgaris</i>	<i>Porrocaecum ensicaudatum</i>	Age	2	0.364	0.106	0.986	0.441
<i>Turdus musicus</i>	<i>Aploparakis dujardinii</i>	Sex	2	0.261	0.337	0.318	0.233
<i>Turdus musicus</i> ‡	<i>Prosthorrhynchus cylindraceus</i>	Sex	2	0.361	0.876	0.005	0.087
<i>Pedetes capensis</i> ‡	<i>Physaloptera capensis</i>	Sex	2	0.127	0.874	0.002	0.364
<i>Rangifer tarandus</i> ‡	<i>Hypoderma tarandi</i>	Time (year)	2	1.217	< 0.001	< 0.001	< 0.001
<i>Rangifer tarandus</i> ‡	<i>Hypoderma tarandi</i>	Age (1985)	5	1.638	< 0.001	< 0.001	0.049
<i>Rangifer tarandus</i> ‡	<i>Hypoderma tarandi</i>	Sex (1985)	2	1.657	< 0.001	< 0.001	0.580
<i>Rangifer tarandus</i> ‡	<i>Hypoderma tarandi</i>	Herd (1985)	7	1.640	0.288	< 0.001	0.638
<i>Rangifer tarandus</i> ‡	<i>Hypoderma tarandi</i>	Age (1988)	5	0.973	< 0.001	< 0.001	< 0.001
<i>Rangifer tarandus</i> ‡	<i>Hypoderma tarandi</i>	Sex (1988)	2	0.967	< 0.001	< 0.001	< 0.001
<i>Rangifer tarandus</i> ‡	<i>Hypoderma tarandi</i>	Herd (1988)	9	1.000	< 0.001	< 0.001	< 0.001

‡ See Table 1 footnote.

constant terms, we minimize L , (minus the log likelihood):

$$L = - \sum_{i=0}^F f_i \cdot \log(e(k)_i),$$

where F is the maximum observed frequency, $e(k)_i$ is the expected relative frequency of hosts having i parasites as a function of k and f_i is the observed absolute frequency of hosts with i parasites. The overall goodness of fit can then be tested by calculating D_i (the deviance associated with the fit (Crawley, 1993)):

$$D_i = 2(L - L_{\text{prob}}),$$

where $L_{\text{prob}} = - \sum_{i=0}^F f_i \cdot \log\left(\frac{f_i}{N}\right)$, $f_i > 0$, N is the total number of hosts in the sample and D_i is approximately χ^2 distributed, with $B - 3$ degrees of freedom, where $B =$ the number of categories/bins that burdens are divided into (maximum parasite burden + 1 complete frequency distributions, and number of categories for lumped frequency distributions). This therefore provides a test of goodness of fit of the distribution (Crawley, 1993). However, this method tends to overestimate the deviance of the goodness of fit. In addition 100 goodness of fits were employed in the analysis and taking the standard P value of 0.05 to indicate significant differences, increased the possibility of Type I errors (incorrectly

ascribing significant differences between observed and expected frequency distributions). Therefore, using a Bonferroni correction, a more conservative P value of 0.001 was taken to indicate significance.

Analysis of dispersion

We use maximum likelihood fits to various models in order to test for significant variations in the mean and k . An interesting discussion of the possibilities and pitfalls of this approach is given by Pacala & Dobson (1988). The methodology is best explained via a simple example. If we have a worm burden frequency distribution, collected for male and female hosts, then in ANOVA terms this corresponds to a single factor with 2 levels. Given that the mean and/or k could vary significantly between the 2 groups, we have 4 possibilities. (1) Both the mean and k vary significantly between groups. Essentially we fit a separate distribution to each group. (2) The mean varies significantly, but k does not. In this case a ‘common k ’ can be assumed. (3) The mean is not significantly different (‘common mean’), but k varies between groups. (4) Neither the mean or k vary (‘common mean and k ’) – the distribution is completely homogenous with respect to sex.

Given estimates of the deviance associated with these 4 possibilities we can use an analysis of deviance (Venables & Ripley, 1995; Wilson, Grenfell & Shaw, 1996) to assess the significance of the various

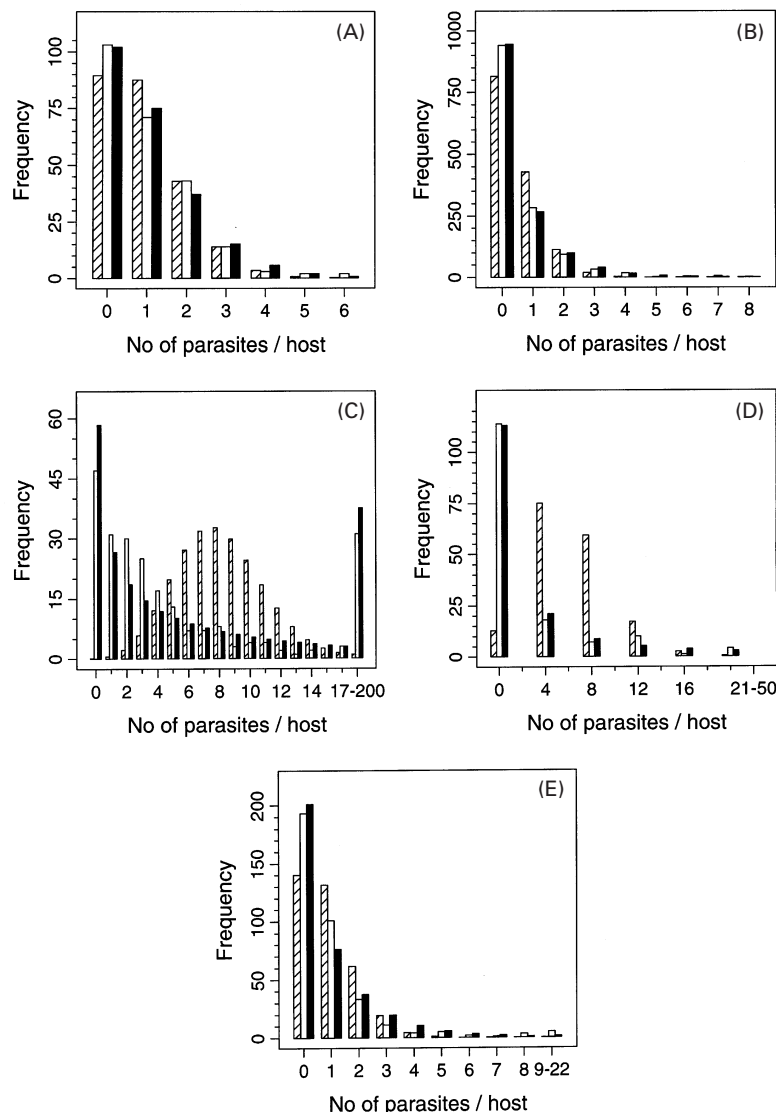


Fig. 1. Observed parasite frequency distributions in the host–parasite infections where the Poisson distribution fitted the data (A), and where the calculated negative binomial distribution was significantly different from the observed frequency distribution (B–E). In all cases the open bars (□) are the observed frequency distribution, the solid bars (■) the calculated negative binomial distribution, and the hatched bars (▨) the calculated Poisson distribution. (A) Shad, *Dorosoma cepedianum*, infected by the nematode *Camallanus oxycephalus* (Stromberg & Crites, 1975); (B) Redfish, *Sebastes mentella*, infected by the copepod *Chondracanthopsis nodus* (Williams, 1963); (C) Arctic charr, *Salvelinus alpinus*, infected by the cestode *Diphyllobothrium ditremum* (Halvorsen & Andersen, 1984); (D) the redwing, *Turdus turdus*, infected by the digenean *Brachylaemus fuscatus* (James & Llewellyn, 1967); (E) the rat, *Rattus rattus*, infected by the cestode *Cysticercus fasciolaris* (Li & Hsu, 1951).

heterogeneities. It does not matter in which order differences between k 's and μ 's are tested, as we are dealing with the subtraction of particular deviances.

In the present paper, we first test whether there are significant differences in the estimates of k between groups, by calculating:

$$D_a = D(\mu_m \mu_f k_c) - [D(\mu_m k_m) + D(\mu_f k_f)],$$

where m = male, f = female, c = common, and $D(\mu_m \mu_f k_c)$ = deviance associated with separate means for the male and female data, but a common k , and $D(\mu_m k_m)$ and $D(\mu_f k_f)$ = the deviances if separate distributions are fitted to the male and female data, respectively. D_a is approximately χ^2 distributed, and

the degree of freedom associated with this value is the number of host groups minus one, i.e. 1. If the resultant P value (P_a) is less than 0.001, then there are significant differences in the degree of aggregation between groups and a common k cannot be assumed.

If $P_a < 0.001$, then we test whether there are also significant differences in the mean burdens between groups, by calculating

$$D_b = D(\mu_c k_m k_f) - [D(\mu_m k_m) + D(\mu_f k_f)],$$

where $D(\mu_c k_m k_f)$ = deviance associated with separate k 's for the male and female data, but a common mean: and $D(\mu_m k_m)$ and $D(\mu_f k_f)$ = as for D_a . D_b is

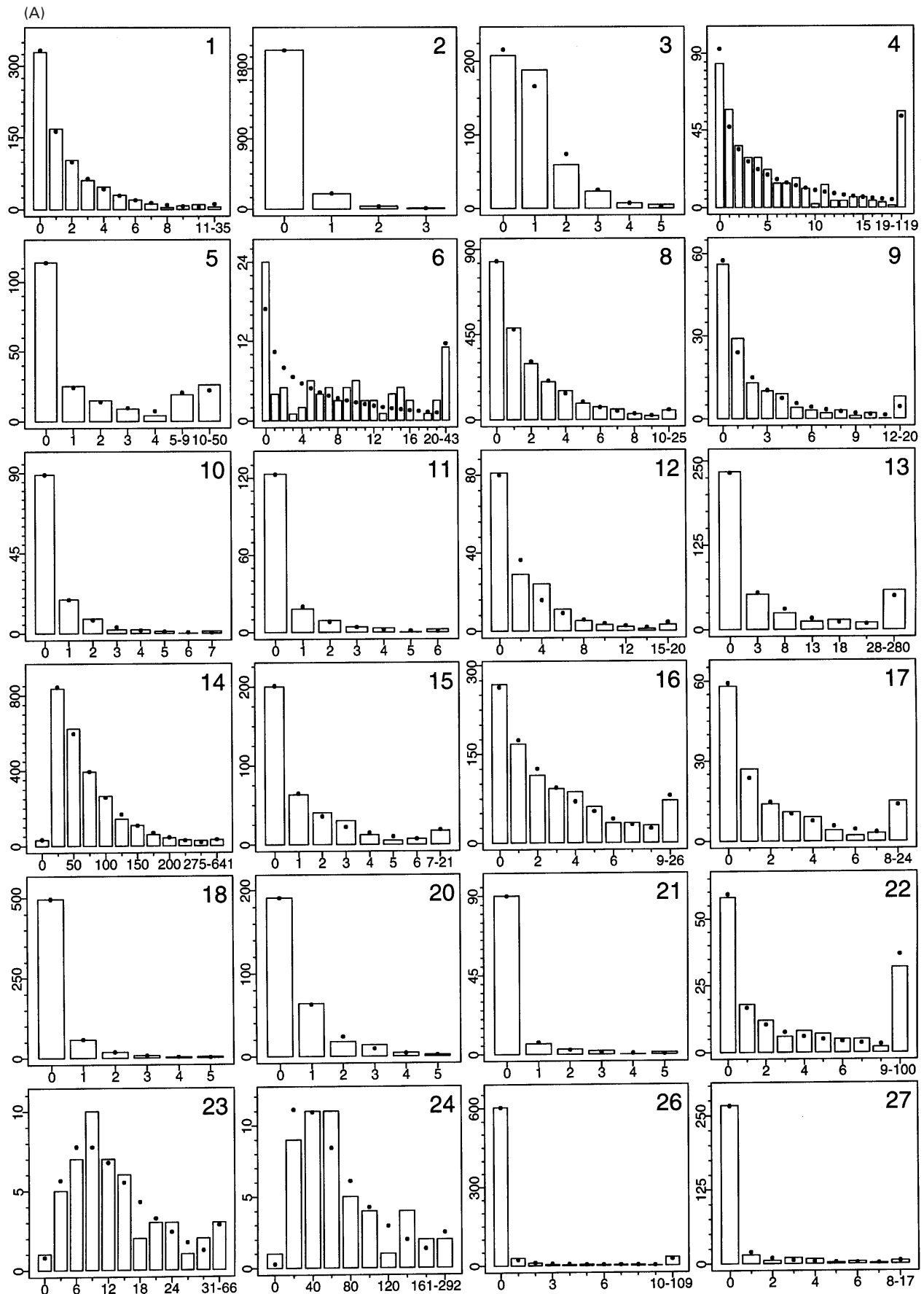


Fig. 2. For legend see opposite.

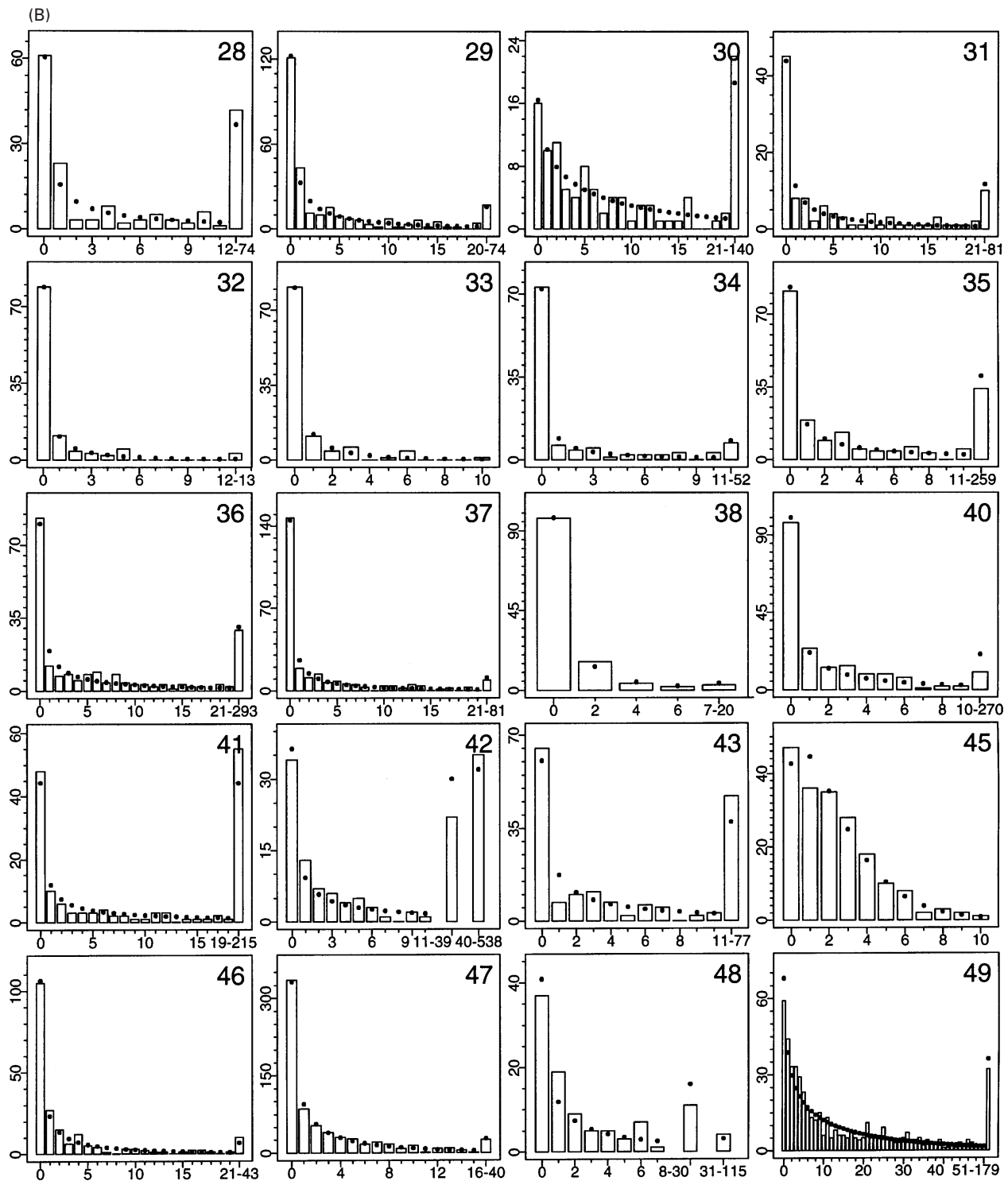


Fig. 2. (A and B) Observed parasite frequency distributions of those infections in Table 1 where the negative binomial provided a good fit to the data. In all cases the *x*-axis is the number of parasites per host, the *y*-axis the frequency of particular numbers of parasites per host, the bars (□) are the observed frequency distributions, and the points (●) the expected negative binomial distribution. The numbers in the top left of the figures correspond to the host-parasite systems in Table 1.

also approximately χ^2 distributed, and the degrees of freedom associated with this value is also the number of host groups minus one, i.e. 1. If the resultant *P* value (P_b) is less than 0.001, then there are significant differences in both the mean and *k* between groups and a common mean and *k* cannot be assumed.

Finally, if $P_a > 0.001$, then we test whether there are significant differences in the mean burdens between groups, by calculating

$$D_c = D(\mu_c k_c) - D(\mu_m \mu_f k_c),$$

where $D(\mu_c k_c)$ = deviance associated with both a

common mean and k , and $D(\mu_m, \mu_f, k_c) =$ as for D_a . D_c is also approximately χ^2 distributed, and the degrees of freedom associated with this value is again the number of host groups minus one, i.e. 1. If the resultant P value (P_c) is less than 0.001, then there are significant differences in the mean burden between groups, but not k and only a common k can be assumed.

This method is easily extended to examine heterogeneities as a function of age, season, etc. The analyses were carried out using the *S-Plus* statistical package (MathSoft, Inc., 1993; Seattle, USA). Differences in the prevalence of infection between subgroups were investigated by calculating Pearson χ^2 statistics for differences in the proportion of individuals positive in each subgroup. The degree of freedom was the number of subgroups minus 1, and a P value of < 0.01 was taken to indicate significant differences (to reduce Type I errors) in the prevalence of infection. Correlations between estimates of k , μ and prevalence of infection were calculated using Spearman Rank correlations.

RESULTS

NBD fits

Table 1 summarizes the 49 data-sets and the associated χ^2 P values testing the fit to both the NBD and Poisson distribution. Of 49 data-sets 45 (92%) fitted the negative binomial (Fig. 2), and only 1 fitted the Poisson (Fig. 1A). However, even with this data-set the NBD provided a better fit than the Poisson distribution, and Fig. 1B–E shows that in the 4 data-sets where the NBD did not fit statistically, the fit was always better than for the Poisson distribution.

This indicates that all but 1 of the parasite distributions show significant aggregation, compared to the Poisson distribution. The estimated degree of aggregation calculated is quite large – in 43/49 studies \hat{k} is less than 1. The variance/mean ratios (Southwood, 1966) are also significantly greater than 1 in all these cases (Shaw & Dobson, 1995). Mean burdens varied from 0.125 to over 60, with the prevalence of infection varying from 10 to 99%. There are significant positive correlations between the prevalence of infection, and \hat{k} ($\rho = 0.646$, $P < 0.001$; Fig. 3A), and mean burdens ($\rho = 0.787$, $P < 0.001$; Fig. 3B).

Heterogeneity in mean burdens and k

Table 2 summarizes the results from the common k and common mean analyses on the 22 data-sets with host subdivisions. One feature apparent from Table 2 is that the common k estimate (k_c) is always greater than the \hat{k} calculated by simply lumping the data together (see Table 1 for comparison of values).

In 11/22 of the data-sets there was significant

variation in the degree of aggregation between host subsets (Table 2). Of 22 data-sets 16 (73%) had significant variation in mean burdens between host subgroups, even after taking into account whether or not there was significant variation in the degree of aggregation. Finally, 14/22 (64%) of the data-sets had significant variation in the prevalence of infection between host subgroups.

In addition, of the 11 data-sets with significant differences in the degree of aggregation there were significant differences in mean burden in 10 (91%) data-sets and in prevalence of infection 9 (82%) data-sets (Table 2). Another general result from the analyses of dispersion (details not shown), was that both increases in mean burden and prevalence of infection in host subsets were generally associated with increases in \hat{k} estimates. To illustrate these latter results Fig. 4A and B show the change in \hat{k} estimates, mean and prevalence for 2 of the 14 data-sets: perch, *Perca fluviatilis*, infected by the cestode *Triaenophorus nodulosus* (Chubb, 1964); and plaice, *Pleuronectes platessa*, caught in 1972 infected by the copepod, *Lepeophtheirus pectoralis*, (Boxshall, 1974). As the figures show, a change in the value of 1 of the parameters is generally associated with the same direction of change in the other 2 parameters.

DISCUSSION

These results provide the broadest evidence to date for the almost ubiquitous fit of the negative binomial distribution to parasite burden data (Crofton, 1971; Anderson & May, 1978; May & Anderson, 1978; Anderson & Gordon, 1982; Keymer, 1982; Pacala & Dobson, 1988; Guyatt *et al.* 1990). This reflects the fundamentally aggregated distribution of parasites between hosts. Other probability distributions have also been proposed to quantify this aggregation – the Logarithmic series (Williams, 1964), Neyman (1939) Type A, Thomas (1949) and Pólya-Aeppli distributions (Pólya, 1931). However, these latter 3 distributions are limited in their biological application (Elliott, 1977), particularly in terms of parasite frequency data, and the logarithmic series distribution ignores zero counts – which Crofton (1971) argued often constitute the most important group of hosts. In addition, they are less flexible distributions and rarely provide better fits for parasite frequency data than the NBD.

This excellent fit to the NBD is despite the many differences between the systems: diversity of host types (invertebrates–mammals); host habitat (aquatic and terrestrial); and parasite taxonomic class (digeneans–ectoparasites). In addition, the high degree of aggregation estimated is in agreement with reviews on human–parasite infections (Anderson, 1978; Anderson & May, 1978; Bundy & Medley, 1992; Guyatt *et al.* 1990).

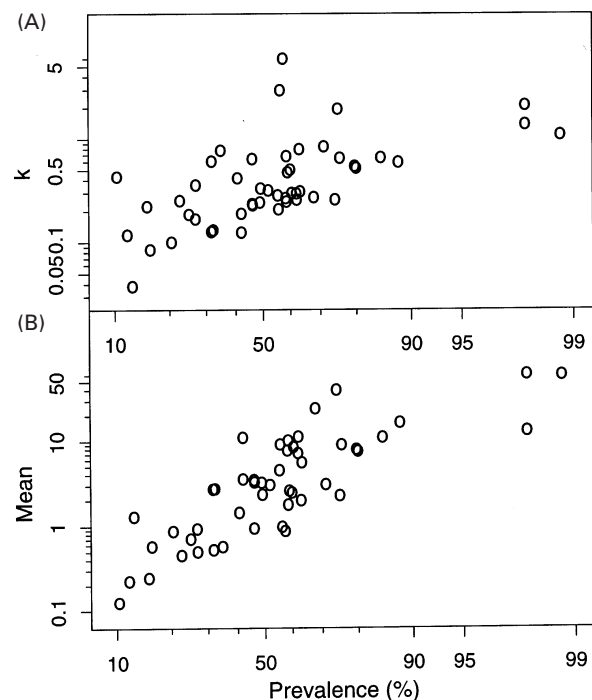


Fig. 3. Logit transformed prevalence of infection ($\log_{10}(\%/ (100-\%))$) of the studies in Table 1 against (A) the negative binomial parameter estimate, \hat{k} on a \log_{10} scale and (B) arithmetic mean burden on a \log_{10} scale.

Biological sources of aggregation

Crofton (1971) and Pacala & Dobson (1988) and Grenfell *et al.* (1995a) have proposed a number of biological conditions under which a NBD could arise. A key mechanism is provided by systematic variation in the acquisition or loss rate of parasites between different groups of hosts. This analysis has revealed many such patterns and we discuss these below. These patterns can be summarized as season, aggregation of infection stages, and host effects. In the present paper no specific experimental work was carried out for each system. Nevertheless, reasonable hypotheses can be made, based on population dynamic information.

(1) *Seasonality.* The occurrence of infectious stages of parasites in aquatic systems (whether they are present within invertebrate intermediate hosts or in water column) is highly seasonal (Gemmell, 1977; Heyneman, 1958; Keymer, 1982). In general, the developmental rate of aquatic-living parasitic stages and invertebrate hosts is temperature dependent. There are also likely to be differences in climatic regimes between years. Therefore when a host population is sampled, a cross-section of individuals which have experienced varying infection rates is removed, and the end result will be aggregation.

(2) *Aggregation of parasitic stages.* The distribution of infection stages of parasites within a habitat tends

to be heterogeneous (Keymer & Anderson, 1979). While hosts may encounter infectious stages at random, the burden associated with an encounter will be very different, resulting in aggregated burdens in hosts (Anderson & Gordon, 1982). The filarial nematode *Chandlerella quiscalis* provides an interesting illustration of this. There is a diurnal pattern of the density of nematodes circulating in definitive host blood (Hawking, 1967), and therefore gnats are likely to encounter different infection intensities depending on when they feed (Schmid & Robinson, 1972).

(3) *Host effects.* Finally, variation between hosts can generate aggregated parasite burdens. This variation can be behavioural, physiological or immunological. An example of behavioural differences is provided by the pike, *Esox lucius*, infected by *Triaenophorus nodulosus* (Chubb, 1963). Pike undergo marked dietary changes as they get older; younger pike eat minnows and loach (which are not thought to be principal intermediate hosts for *T. nodulosus*), whereas older pike eat perch and gwyniad (which are).

Host immune responses to parasite infection, can also alter observed frequency distributions (Anderson & Gordon, 1982). The strength of the response can be modulated by host stress (Christian & Davis, 1964). Male reindeer were associated with higher mean *Hypoderma tarandi* burdens and higher degrees of aggregation. A possible reason for this is a testosterone-induced immune depression at a crucial period in the parasite life-cycle; the migration of larvae towards the dorsal region of the reindeer (Folstad *et al.* 1989).

Exceptions to the NBD

Only 1 data-set fitted the Poisson: shad, *Dorosoma cepedianum* infected by a nematode, *Camallanus oxycephalus*, (Stromberg & Crites, 1975). Further empirical work is required to see whether a lack of aggregation is always present in this system, and whether such factors as little variation in transmission between hosts, or a relatively homogeneous distribution of infectious stages are causing the random distribution of parasites. However, even in this system the negative binomial provided a statistically better fit than the Poisson.

With the 4 observed distributions which did not fit the NBD, the distributions were all aggregated ($s^2/\bar{x} > 1.7$), and again the NBD always provided a better fit than the PD, emphasizing the plasticity of the NBD applied to overdispersed distributions.

Analysing dispersion

At present much of the analysis of parasite burdens concentrates on differences in mean burden, and only

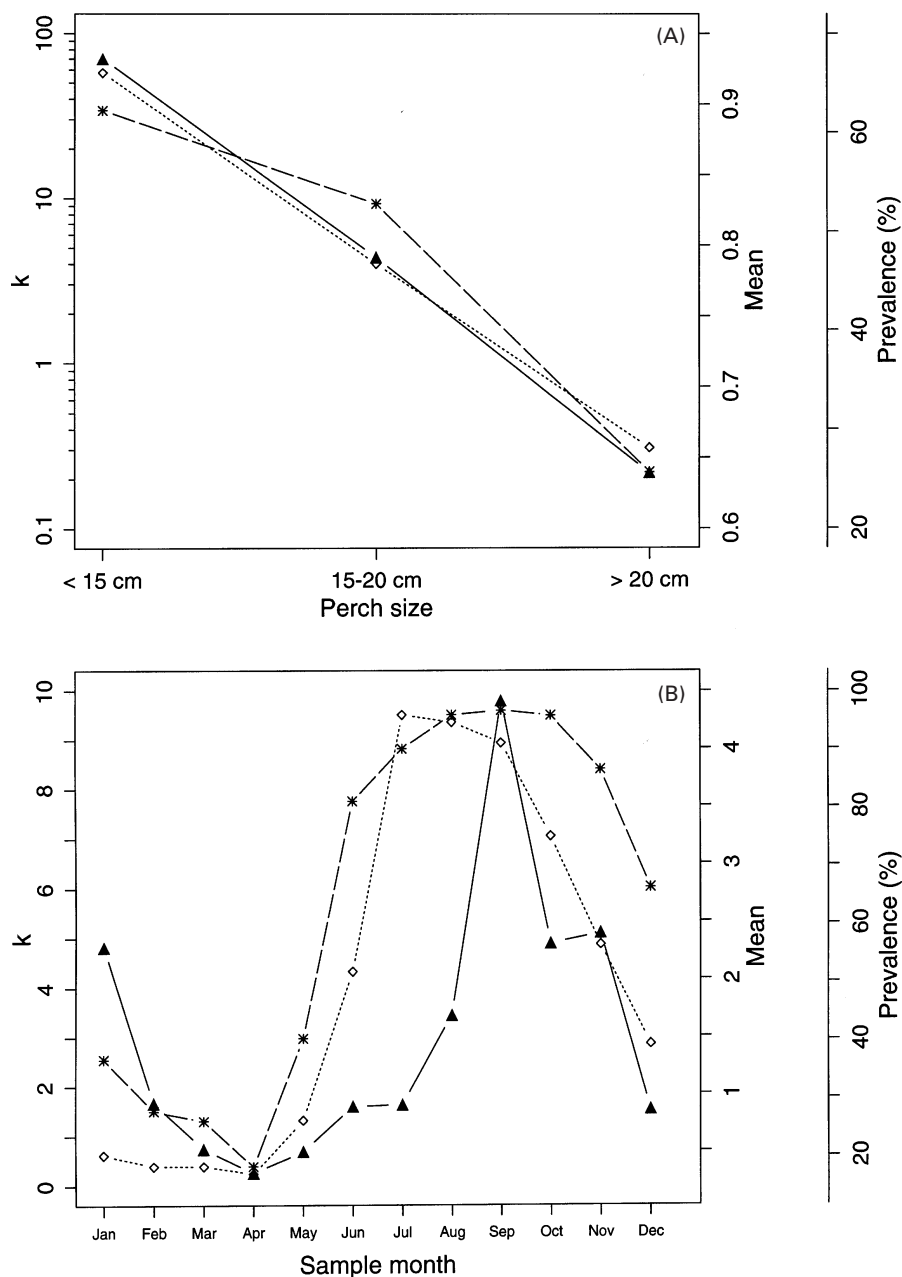


Fig. 4. Negative binomial parameter estimate, \hat{k} (\blacktriangle — \blacktriangle), arithmetic mean burden (\diamond ····· \diamond) and prevalence of infection (*—*—*) from (A) Perch, *Perca fluviatilis*, infected by the cestode, *Triaenophorus nodulosus*, and subdivided into 3 age-groups, based on size (Chubb, 1964); (B) plaice, *Pleuronectes platessa*, caught in 1972 infected by the copepod, *Lepeophtheirus pectoralis*, and subdivided into the month that the plaice were caught (Boxshall, 1974).

occasionally is account taken of the underlying aggregated distribution of parasites (Wilson *et al.* 1996). However, even when authors do take into account aggregation (e.g. general linear models (GLM) with negative binomial errors (Venables & Ripley, 1994)), these GLM's assume a common k and therefore just test for differences in mean burdens. While differences in mean burdens are important in terms of determining relative levels of parasite load, of equal importance is how the parasite burden is distributed between hosts, particularly in terms of morbidity and the dynamics of transmission (Guyatt *et al.* 1990). What is required are statistical models which consider whether there are significant

differences in the degree of aggregation *and* mean burdens between hosts.

The analysis of dispersion technique used in this paper provides a useful and powerful methodology to examine differences in both the degree of aggregation and mean burdens between subgroups of hosts. As described above, for each of the 22 data sets with host groups we were able to establish whether there was just variation in the degree of aggregation, or mean burden or whether both parameters varied. Unfortunately, one major disadvantage is that the use of ML techniques are cumbersome and iteratively intense. Of course, a better approach would be to use a GLM which

considers both k and the mean, but at present these are not available (W. N. Venables, personal communication).

Observed patterns of dispersion

When differences in mean burdens between host groups were taken into account when estimating the degree of aggregation, then k_c was always larger than if all the host data were put together and \hat{k} estimated. That is, putting together host data from different categories (e.g. age group, sex, location, season), consistently exaggerates the actual degree of aggregation (see (Dietz, 1982; Pacala & Dobson, 1988)). In addition, as Grafen & Woolhouse (1993) have shown there are potential problems with lumping host data which consists of different NBD. Therefore, wherever possible, k_c estimates should be used to describe the degree of aggregation within a host–parasite system.

The observed positive relationships between the degree of aggregation, arithmetic mean burden and prevalence of infection, reveal another generality of our results: higher mean burdens and k 's are generally associated with higher prevalence of infection. However, Shaw & Dobson (1996) have demonstrated that very different mean burdens can be found with a similar prevalence of infection (and *vice versa*) – it depends which host–parasite systems are being compared. In other words, we cannot predict what k or the mean burden are likely to be purely on the basis of the prevalence of infection.

In over half of the data-sets where subsets of host data were available, significant differences in the degree of aggregation were observed. Such variations indicate that the degree of aggregation estimated from 1 sample, may not be that representative of a particular host–parasite system as a whole. Indeed changes in k with age are frequently observed (e.g. Fig. 4A; Pacala & Dobson, 1988; Quinnell, Grafen & Woolhouse, 1995), indicating that as large and representative a sample of a host population is required before reasonable estimates on the degree of parasite aggregation present can be made (Gregory & Woolhouse, 1993).

Overall, the observed \hat{k} estimates tended to vary with the mean burden for the subsets of host data; with 10/16 data-sets have significant variation in \hat{k} estimates, when the mean burdens were significantly different. Such 'tracking' of the mean by k has been observed in theoretical work on the relationships between infection parameters (Gregory & Woolhouse, 1993; Grenfell, Dietz & Roberts, 1995*b*), and is related to the fact that the dispersion parameter of the NBD – k – is not independent of the mean (Pennycuick, 1971*a*; Scott, 1987). In many cases, k tends to be roughly proportional to the mean (see Grenfell *et al.* 1995*a*), though in some cases, immunity or parasite-induced host mortality can

cause k to rise with age as the mean falls (Anderson & Gordon 1982; Dobson & Pacala 1988; Grenfell *et al.* 1995*a*). The correlation between the mean and k also suggests that only particular subsets of all the possible NBDs are actually observed in parasite frequency distributions. However, further work is required to actually see whether this is the case.

In conclusion, the NBD is a powerful and robust tool for describing how parasites are distributed amongst animal hosts. As illustrated above, a number of mechanisms are likely to be generating aggregated distributions. However, subdivisions of host data also reveal significant (and biologically explicable) differences in the degree of aggregation *within* the data-sets. In addition, the common k value assuming different group means (k_c), was always larger than the \hat{k} value calculated when the data are lumped together. That is, we potentially exaggerate the overall degree of aggregation of parasites by lumping data across important host dynamics (Dietz, 1982; Pacala & Dobson, 1988). Wherever possible, potential heterogeneities within a sample host population should be analysed explicitly. Prevalence of infection, mean burden and \hat{k} estimates were positively related, and changes in one of the parameters was generally associated with changes in the other two parameters, indicating the lack of independence between the parameters used to study infection (Grenfell *et al.* 1995*a*; Scott, 1987). This paper has also outlined a powerful method to test for differences in aggregation and worm burden between host groups. At present there are no specific GLM which test for differences in both the degree of aggregation and mean burdens, and therefore this analysis of dispersion, while cumbersome, could be extended to look at other host–parasite dynamics.

The authors would like to acknowledge Ivar Folstad and colleagues at the Department of Ecology/Zoology, Institute of Biology and Geology, University of Tromsø, Tromsø, Norway for the very generous use of the reindeer-wearable fly data. D. J. S. was supported by the SERC as part of a Ph.D. thesis, and B. T. G. by the NERC and BBSRC.

REFERENCES

- ANDERSON, R. M. (1974). Population dynamics of the cestode *Caryophyllaeus laticeps* (Pallas, 1781) in the bream (*Aramis brama* L.) *Journal of Animal Ecology* **43**, 305–321.
- ANDERSON, R. M. (1976*a*). Some simple models of the population dynamics of eucaryotic parasites. *Lecture Notes in Biomathematics* **11**, 16–57.
- ANDERSON, R. M. (1976*b*). Dynamic aspects of parasite population ecology. *Ecological Aspects of Parasitology* **21**, 431–462.
- ANDERSON, R. M. (1978). The regulation of host population growth by parasitic species. *Parasitology* **76**, 119–157.
- 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**, 373–398.
- ANDERSON, R. M. & MAY, R. M. (1978). Regulation and stability of host-parasite population interactions—I. Regulatory processes. *Journal of Animal Ecology* **47**, 219–247.
- ANDERSON, R. M. & MAY, R. M. (1982). Population dynamics of human helminth infections: control by chemotherapy. *Nature, London* **297**, 557–563.
- BLISS, C. I. & FISHER, R. A. (1953). Fitting the negative binomial distribution to biological data. *Biometrics* **9**, 176–199.
- BOXSHALL, G. A. (1974). The population dynamics of *Lepeophtheirus pectoralis* (Muller): dispersion pattern. *Parasitology* **69**, 373–390.
- BRATTEY, J. (1988). Life history and population biology of adult *Acanthocephalus lucii* (Acanthocephala: Echinorhynchidae). *Journal of Parasitology* **74**, 72–80.
- BREYEV, K. A. (1973). On the nature of relation between the incidence of infection in host population and the mean density of parasite. *Folia Parasitologica* **20**, 227–234.
- BUNDY, D. A. P. & MEDLEY, G. F. (1992). Epidemiology and population dynamics of helminth infection. In *Allergy and Immunity to Helminths: Common Mechanisms or Divergent Pathways?* (ed. Hoqbel, R.), pp. 17–37. Taylor & Francis, London.
- BUTYNSKI, T. M. (1982). Ecology of *Physaloptera capensis* (Nematoda: Spiruroidea) infecting the Springhare, *Pedetes capensis* (Mammalia) in Botswana. *Journal of Zoology, London* **196**, 233–254.
- CHUBB, J. C. (1963). Seasonal occurrence and maturation of *Triaenophorus nodulosus* (Pallas, 1781) (Cestoda: Pseudophyllidae) in Pike *Esox lucius* L. of Llyn Tegid. *Parasitology* **55**, 419–433.
- CHUBB, J. C. (1964). Observations on the occurrence of the plerocercoids of *Triaenophorus nodulosus* (Pallas, 1781) (Cestoda: Pseudophyllidae) in the perch *Perca fluviatilis* L. of Llyn Tegid (Bala Lake), Merionethshire. *Parasitology* **54**, 481–491.
- CRAWLEY, M. J. (1993). *GLIM for Ecologists*. Blackwell Scientific Publications, London.
- CHRISTIAN, J. J. & DAVIS, D. E. (1964). Endocrines, behaviour and population. *Science* **146**, 1550–1560.
- CROFTON, H. D. (1971). A quantitative approach to parasitism. *Parasitology* **62**, 179–193.
- DIETZ, K. (1982). Overall population patterns in the transmission cycle of infectious agents. In *Population Biology of Infectious Diseases* (ed. Anderson, R. M. & May, R. M.), pp. 87–102. Springer, Berlin.
- DOBSON, A. P. (1985). The population dynamics of competition between parasites. *Parasitology* **91**, 317–347.
- ELLIOTT, J. M. (1977). *Statistical Analysis of Samples of Benthic Invertebrates*. 2nd Edn. Freshwater Biological Association, Ambleside.
- EVANS, N. A. (1978). The occurrence and life history of *Asymphylogdora kubanicum* (Platyhelminths: Digenea: Monorchidae) in the Worcester-Birmingham canal, with special reference to the feeding habits of the definitive host, *Rutilus rutilus*. *Journal of Zoology, London* **184**, 143–153.
- EVANS, N. A. (1983). The population biology of *Hymenolepis tenerrima* (Linstow 1882) Fuhrmann, 1906 (Cestoda, Hymenolepididae) in its intermediate host *Herpetocypris reptans* (Ostracoda). *Zeitschrift für Parasitenkunde* **69**, 105–111.
- EVANS, N. A., WHITFIELD, P. J., BAMBER, R. N. & ESPIN, P. M. (1983). *Lernaocera lusci* (Copepoda: Pennellidae) on bib (*Tristopterus luscus*) from Southampton Water. *Parasitology* **86**, 161–173.
- EVANS, N. A., WHITFIELD, P. J. & DOBSON, A. P. (1981). Parasite utilisation of a host community: the distribution and occurrence of metacercarial cysts of *Echinoparyphium recurvatum* (Digenea: Echinostomatidae) in seven species of mollusc at Harting Pond, Sussex. *Parasitology* **83**, 1–12.
- FELLER, W. (1943). On a general class of ‘contagious’ distributions. *Annals of Mathematical Sciences* **14**, 389–400.
- FISHER, R. A. (1941). The negative binomial distribution. *Annals of Eugenics* **11**, 182–187.
- FISHER, R. A., CORBET, A. S. & WILLIAMS, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* **12**, 42–58.
- FOLSTAD, I., NILSSEN, A. C., HALVORSEN, O. & ANDERSEN, J. (1989). Why do male reindeer (*Rangifer t. tarandus*) have higher abundance of second and third instar larvae of *Hypoderma tarandi* than females? *Oikos* **55**, 87–92.
- FOLSTAD, I., NILSSEN, A. C., HALVORSEN, O. & ANDERSEN, J. (1991). Parasite avoidance: the cause of post-calving migrations in *Rangifer*? *Canadian Journal of Zoology* **69**, 2423–2429.
- GEMMELL, M. A. (1977). Taeniidae: modification to the life span of the egg and the regulation of tapeworm populations. *Experimental Parasitology* **41**, 314–328.
- GRAFEN, A. & WOOLHOUSE, M. E. J. (1983). Does the negative binomial distribution add up? *Parasitology Today* **9**, 475–477.
- GREGORY, R. D. & WOOLHOUSE, M. E. J. (1993). Quantification of parasite aggregation: a simulation study. *Acta Tropica* **54**, 131–139.
- GRENFELL, B. T. (1988). Gastrointestinal nematode parasites and the stability and productivity of intensive ruminant grazing systems. *Philosophical Transactions of the Royal Society of London, B* **321**, 541–563.
- GRENFELL, B. T. (1992). Parasitism and the dynamics of ungulate grazing systems. *The American Naturalist* **139**, 907–929.
- GRENFELL, B. T., DAS, P. K., RAJAGOPALAN, P. K. & BUNDY, D. A. P. (1990). Frequency distribution of lymphatic filariasis microfilariae in human populations: population processes and statistical estimation. *Parasitology* **101**, 417–427.
- GRENFELL, B. T., DIETZ, K. & ROBERTS, M. G. (1995b). Modelling the immuno-epidemiology of macroparasites in wildlife host-populations. In *Ecology of Infectious Diseases in Natural Populations* (ed. Grenfell, B. T. & Dobson, A. P.) Cambridge University Press, Cambridge.
- GRENFELL, B. T., SMITH, G. & ANDERSON, R. M. (1986). Maximum-likelihood estimates of the mortality and

- migration rates of the infective larvae of *Ostertagia ostertagi* and *Cooperia oncophora*. *Parasitology* **92**, 643–652.
- GRENFELL, B. T., WILSON, K., ISHAM, V., BOYD, H. E. G. & DIETZ, K. (1995a). Modelling patterns of parasite aggregation in natural populations: trichostrongylid nematode-ruminant interactions as a case study. *Parasitology* **111**, S135–S151.
- GUYATT, H. L. & BUNDY, D. A. P. (1991). Estimating prevalence of community morbidity due to intestinal helminths: prevalence of infection as an indicator of the prevalence of disease. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **85**, 778–782.
- GUYATT, H. L., BUNDY, D. A. P., MEDLEY, G. F. & GRENFELL, B. T. (1990). The relationship between the frequency distribution of *Ascaris lumbricoides* and the prevalence and intensity of infection in human communities. *Parasitology* **101**, 139–143.
- HALVORSEN, O. & ANDERSEN, K. (1984). The ecological interaction between arctic charr, *Salvelinus alpinus* (L.) and the plerocercoid stage of *Diphyllbothrium ditremum*. *Journal of Fish Biology* **25**, 305–316.
- HAWKING, F. (1967). On 24 hours periodicity of microfilariae: biological mechanisms responsible for its production and control. *Proceedings of the Royal Society, B* **169**, 59–76.
- HENRICKSON, J. (1977). The abundance and distribution of *Diphyllbothrium dendriticum* (Nitzsch) and *D. ditremum* (Creplin) in the charr *Salvelinus alpinus* (L.) in Sweden. *Journal of Fish Biology* **11**, 231–248.
- HEYNEMAN, D. (1958). Effect of temperature on the rate of development and viability of the cestode, *Hymenolepis nana* in its intermediate host. *Experimental Parasitology* **7**, 374–382.
- HODASI, J. K. M. (1969). Comparative studies on the helminth fauna of native and introduced domestic fowls in Ghana. *Journal of Helminthology* **43**, 35–52.
- IGNOFFO, C. M. (1959). Louse populations of some rodents of the Great Salt Lake Desert. *Parasitology* **49**, 511–518.
- JAMES, B. L. & LLEWELLYN, L. C. (1967). A quantitative analysis of helminth infestation in some Passerine birds found dead on the island of Skomer. *Journal of Helminthology* **41**, 19–44.
- JARROLL, E. L. JR. (1979). Population biology of *Bothriocephalus rarus* Thomas (1937) in the red-spotted newt, *Notophthalmus viridescens* Raf. *Parasitology* **79**, 183–193.
- KENNEDY, C. R. (1984). The dynamics of a declining population of the acanthocephalan *Acanthocephalus clavula* in eels *Anguilla anguilla* in a small river. *Journal of Fish Biology* **25**, 665–677.
- KEYMER, A. E. (1982). Tapeworm infections. In *Population Dynamics of Infectious Diseases: Theory and Applications* (ed. Anderson, R. M.), pp. 109–138. Chapman & Hall, London.
- KEYMER, A. E. & ANDERSON, R. M. (1979). The dynamics of infection of *Tribolium confusum* by *Hymenolepis diminuta*: the influence of infective-stage density and spatial distribution. *Parasitology* **79**, 195–207.
- LAWLESS, J. F. (1987). Negative binomial and mixed Poisson regression. *The Canadian Journal of Statistics* **15**, 209–225.
- LI, S. Y. & HSU, H. F. (1951). On the frequency distribution of helminths in their naturally infected hosts. *Journal of Parasitology* **37**, 32–41.
- MAIZELS, R. M., BUNDY, D. A. P., SELKIRK, M. E., SMITH, D. F. & ANDERSON, R. M. (1993). Immunological modulation and evasion by helminth parasites in human populations. *Nature, London* **365**, 797–805.
- MAY, R. M. & ANDERSON, R. M. (1978). Regulation and stability of host-parasite population interactions II. Destabilising processes. *Journal of Animal Ecology* **47**, 248–267.
- MILNE, A. (1943). The comparison of sheep-tick populations (*Ixodes ricinus* L.). *Annals of Applied Biology* **30**, 240–253.
- NEYMAN, J. (1939). On a new class of ‘contagious’ distributions, applicable in entomology and bacteriology. *Annals of Mathematical Statistics* **10**, 35–57.
- NUORTEVA, P. (1966). *Corynosoma strumosum* (Rud.) and *C. semerne* (Forssell) (Acanthocephala) as pathogenic parasites of farmed minks in Finland. *Journal of Helminthology* **40**, 77–80.
- 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.
- PENNYCUICK, L. (1971a). Frequency distributions of parasites in a population of three-spined sticklebacks, *Gasterosteus aculeatus* L., with particular reference to the negative binomial distribution. *Parasitology* **63**, 389–406.
- PENNYCUICK, L. (1971b). Seasonal variation in the parasite infections in a population of three-spined sticklebacks, *Gasterosteus aculeatus* L. *Parasitology* **63**, 373–388.
- PIELOU, E. C. (1969). *An Introduction to Mathematical Ecology*. Wiley Interscience, New York.
- PÓLYA, G. (1931). Sur quelques points de la théorie des probabilités. *Annales de l'Institut Henri Poincaré* **1**, 117–161.
- QUENOUILLE, M. H. (1949). A relation between the logarithmic, Poisson, and negative binomial series. *Biometrics* **5**, 162–184.
- QUINNELL, R. J., GRAFEN, A. & WOOLHOUSE, M. E. J. (1995). Changes in parasite aggregation with age: a discrete infection model. *Parasitology* **111**, 635–644.
- SCHMID, W. D. & ROBINSON, E. J. (1972). The pattern of a host-parasite distribution. *Journal of Parasitology* **58**, 907–910.
- SCOTT, M. E. (1987). Temporal changes in aggregation: a laboratory study. *Parasitology* **94**, 583–595.
- SCOTT, M. E. (1988). The impact of infection and disease on animal populations: Implications for conservation biology. *Conservation Biology* **2**, 40–56.
- SHAW, D. J. & DOBSON, A. P. (1995). Patterns of macroparasite abundance and aggregation in wildlife populations: a quantitative review. *Parasitology* **111**, S111–S133.
- SKORPING, A. (1981). Seasonal dynamics in abundance, development and pattern of infection of *Bunodera luciopercae* (Muller) in perch, *Perca fluviatilis* L. from an oligotrophic lake in Norway. *Journal of Fish Biology* **18**, 401–410.
- SOUTHWOOD, T. R. E. (1966). *Ecological Methods*. 2nd Edn. Chapman & Hall, London.

- STROMBERG, P. C. & CRITES, J. L. (1975). Population biology of *Camallanus oxycephalus* Ward and Magath, 1916 (Nematoda: Camallanidae) in white bass in western Lake Erie. *Journal of Parasitology* **61**, 123–132.
- STROMBERG, P. C., TOUSSANT, M. J. & DUBEY, J. P. (1978). Population biology of *Paragonimus kellicotti* metacercariae in central Ohio. *Parasitology* **77**, 13–18.
- THOMAS, M. (1949). A generalisation of Poisson's binomial limit for use in ecology. *Biometrika* **36**, 18–25.
- VENABLES, W. N. & RIPLEY, B. D. (1994). *Modern Applied Statistics with S-Plus*. Springer-Verlag, New York.
- WILLIAMS, I. C. (1963). The infestation of the redfish *Sebastes marinus* (L.) and *S. mentella* Travin (Scleroparei: Scorpaenidae) by the copepods *Peniculus clavatus* (Müller), *Sphyrion lumpi* (Kroyer) and *Chondracanthopsis nodusus* (Müller) in the eastern North Atlantic. *Parasitology* **53**, 501–525.
- WILLIAMS, C. B. (1964). *Patterns in the Balance of Nature, and Related Problems in Quantitative Biology*. Academic Press, London.
- WILSON, K., GRENFELL, B. T. & SHAW, D. J. (1996). Analysis of aggregated parasite distributions: a comparison of methods. *Functional Ecology* **10**, 592–601.