International Journal for Parasitology 53 (2023) 763–776

Contents lists available at ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara

Current Opinion

Interrelationships and properties of parasite aggregation measures: a user's guide

A. Morrill [a,*], R. Poulin [b], M.R. Forbes [a]

[a] Biology Department, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada
[b] Department of Zoology, University of Otago, 340 Great King Street, Dunedin North, New Zealand

ARTICLE INFO

Article history:
Received 6 April 2023
Received in revised form 2 June 2023
Accepted 6 June 2023
Available online 17 July 2023

Keywords:
Aggregation
Hoover's index
Mean crowding
Negative binomial
Parasites
Patchiness
Poulin's *D*
Variance-to-mean ratio

ABSTRACT

Aggregation of macroparasites among hosts is nearly universal among parasite-host associations. Researchers testing hypotheses on origins of parasite aggregation and its importance to parasite and host population ecology have used different measures of aggregation that are not necessarily measuring the same thing, potentially clouding our understanding of underlying epidemiological processes. We highlight these differences in meanings by exploring properties and interrelationships of six common measures of parasite aggregation, and provide a "user's guide" to inform researchers' decisions regarding their application. We compared the mathematical expressions of the different measures of aggregation, and ran two series of simulations and analyses. The first simulations tested the effect of random removals of parasites on aggregation levels under different conditions, while the second explored interrelationships between the measures, as well as between other individual parasitological sample measures (i.e. mean abundance, prevalence) and aggregation. Results of simulations and analyses showed that the six measures of aggregation could be separated readily into three groups: the variance-to-mean ratio (VMR) together with mean crowding, patchiness with k of the negative binomial, and Poulin's D with Hoover's index. These three pairs of measures showed differing responses to random parasite removals and differing relations with mean abundance and/or prevalence, highlighting that metrics capture different variation in other sample measures and different attributes of aggregation. We used results of our simulations and analyses, and a literature review, to list the properties, advantages, and disadvantages of each aggregation metric. We provide a comprehensive exploration of what is assessed by each metric, as a guide to metric choice. We implore researchers to provide enough information such that aggregation measures from each group are reported or can be readily calculated. Such steps are needed to allow large-scale analyses of variation in degrees of aggregation within and among parasite-host associations, to uncover epidemiological processes shaping parasite distributions.

1. Introduction

Aggregation of macroparasites among hosts is a nearly universal phenomenon, and as such it is one of the few "laws" of parasite ecology (Poulin, 2007a). This widespread pattern of most hosts harbouring few or no infecting macroparasites (hereafter parasites), while a minority of hosts experience much higher intensity infections, has important consequences for the impacts of disease-causing organisms on hosts (Hudson et al., 1992; Poulin, 2007b), the probability and intensity of co-infection of hosts by different parasite strains or species (Morrill et al., 2017), and the stability of host-parasite associations (Anderson and May, 1978; Rosà and Pugliese, 2002). Such a common and consequential phenomenon has attracted much attention from researchers – principally ecologists, wildlife researchers and parasitologists – who have sought to measure and explain aggregation, as well as predict its consequences (Shaw and Dobson, 1995; Johnson and Hoverman, 2014; Cox et al., 2017). Understanding aggregation of parasites is a fundamental problem in the study of parasite evolutionary ecology, owing to its impact on, as examples, parasite population genetic structure (Cornell et al., 2003), infection intensity-dependent population regulation (Møller, 2005), and both parasite reproduction and mating systems (Criscione et al., 2005; Cox et al., 2017). It is an important problem in applied ecology and wildlife management as well, since treating only the small subset of heavily infected hosts can be an efficient and cheaper approach to parasite control

---

* Corresponding author at: André Morrill, 3600 Carleton Technology and Training Centre, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada.

E-mail address: andre_morrill@carleton.ca (A. Morrill).

https://doi.org/10.1016/j.ijpara.2023.06.004

in managed populations (Perkins et al., 2003). The advances in the study of parasite aggregation also have the potential to inform other branches of science (e.g. contaminant biology), where measures of aggregation or skew are used (Morrill et al., 2014).

Researchers may agree – or, at least, take it for granted that they agree – on a general definition of parasite aggregation, but have adopted and/or developed various measures for its quantification. When discussing measures of aggregation in ecology more broadly, Pielou (1977) offered the following observations on how these different methods highlight contrasting interpretations (quoted also by Hurlbert, 1990; McVinish and Lester, 2020):

> "... the phrase 'degree of aggregation' describes a vague, undefined notion that is open to several interpretations. If aggregation is to be measured, we must first choose from a number of possibilities some measurable property of a spatial pattern that is to be called its aggregation, and the method of measurement is then implicit in the chosen definition. Thus the several existing ways of measuring aggregation are not different methods of measuring the same thing: they measure different things."

It is incumbent on the research community using, and choosing between, various measures of aggregation to know that common measures can capture contrasting properties, and that these disparities point to differences in the meanings of aggregation between such measures. Having discussions about the underlying and general causes of aggregation, for example, are most fruitful if researchers are measuring the same thing, or at least know when they are discussing different attributes of aggregated distributions. Different measures of aggregation might be more or less amenable to specific research questions and analyses (as, for example, when researchers are interested in wholly different research questions such as intraspecific competition between parasites versus parasite-mediated effects on hosts).

A simple expectation is that measures less similar in their implicit meanings should be less strongly correlated with one another; they may also show distinctly different relationships with other sample-level parasitological indices such as prevalence (the proportion of hosts that are infected in a sample) and/or mean abundance (the mean number of parasite individuals per host, including uninfected hosts). Naturally, as the different measures of aggregation all attempt to encapsulate an emergent, population-level general property of the "clustering" of infecting parasites, they should be expected to correlate with one another to some degree. Nevertheless, departures from near perfect correlations, or losses of correlations in certain contexts, would help to identify differences between the various aggregation measures in their meanings. By extension, wherever near perfect correlations do arise, researchers can be more confident that the choice of one measure over another is of less consequence.

As other sample-level measures also describe how parasites are distributed among hosts (e.g. infection prevalence, mean abundance), it is not surprising that these too show varying levels of correlation with degree of parasite aggregation (Gregory and Woolhouse, 1993; Poulin, 1993; McVinish and Lester, 2020), and with each other. Pielou (1977) proposed an insightful question that highlights a simple, but informative, apparent dichotomy in measures of aggregation in terms of their relationship with the mean number of individuals per unit area/volume: should aggregation change or remain unchanged when a random proportion of individuals are removed from a distribution (e.g. random removals of parasites from a sample of hosts)? As she describes, both answers are reasonable. One could consider that random removals would be expected to result, simply by chance, in more parasites being removed from larger infrapopulations (i.e. groups of conspecific parasites infecting a single host) than smaller ones, thereby making those infrapopulations less "dense." From this perspective, as the overall number of parasites decreases, so too should aggregation. On the other hand, as removals would be random, the only resulting difference in the distribution would be in its overall "density," while its shape would otherwise remain unchanged; therefore, one could consider that the essential property of aggregation relating to the distribution's shape similarly should remain unchanged. Importantly, this simple thought experiment highlights two contrasting expected consequences of random removals: either aggregation should decrease, or remain the same.

The present study has three objectives. The first two are to address the related questions: how do the implicit meanings of parasite aggregation differ among the common methods used for its quantification? And how do these measures compare in terms of their relationship with mean abundance, prevalence, and each other? Such a synthesis is, to our knowledge, lacking in the study of distributions of infecting parasites among hosts, and where previous analyses have partially addressed these questions, conclusions have been inconsistent. Our third objective is to provide a "user's guide" to help researchers choose the measure(s) of aggregation most relevant to their studies, based on the properties and context-specific advantages and disadvantages identified in simulations, and in previous research. To accomplish these objectives, we first summarize the various commonly used measures of aggregation and consider explicitly their mathematical expressions. We then compare measures in terms of their responses to random parasite removals, following from the thought experiment posited by Pielou (1977). With a second series of simulations, we explore their relationships to one another and to mean abundance and prevalence. We propose a framework that simplifies and synthesizes these perspectives, clarifies previous results relating to the attributes of some measures, and highlights important and previously unrecognized properties of some of the popular aggregation metrics.

More specifically, we illustrate that random removals of infecting parasites results in three groupings of measures either showing no change in response to removals, a negative effect, or a positive effect on their estimate of aggregation. We then show strong correlations between certain common measures of aggregation, and that these pairs of measures align with the same three groups seen in the initial simulation. Importantly, the three pairs of measures also differ in their relationships with sample prevalence and mean abundance. This work will help guide approaches to, and interpretations of, studies of parasite aggregation, and help researchers report or calculate degrees of aggregation in such a way as to ensure inclusion of their studies in future syntheses.

2. Summary of considered aggregation measures

Given the many available measures of parasite aggregation, a problem faced by researchers is deciding which one(s) to choose to report or include in analyses. To aid researchers in their choices, our synthesis should encompass a variety of measures including those measures which are most popular, as well as those whose properties may make them particularly useful in specific contexts. However, the diversity of available measures – several of which are closely related to at least one other – means that we, too, are faced with a choice of only a subset to keep our analyses tractable. Our first inclusion criterion was that each measure can be applied directly to a sample of hosts variously infected by a single taxon of parasites at known abundances; i.e. the measure can be applied to a set of discrete counts of parasites infecting each sampled host. Secondly, we included measures if they were commonly used, only recently proposed, or if they were long-established indices of particular value in quantifying levels of aggregation from the parasite's perspective.

The first criterion results in the exclusion of measures which are based on multiple samples of the same parasite-host system, either across time or space. The established approach of measuring the slope of the log-variance–log-mean relationship based on multiple samples was therefore excluded (Shaw and Dobson, 1995; Poulin, 2013), as well as any related methods (e.g. the I_{10} index; Lester and Blomberg, 2021). While methods have been developed to extend some of these approaches to single samples via bootstrapping (e.g. Boag et al., 2001), we chose to follow our first criterion strictly and exclude those approaches, focusing on those measures most broadly applicable across past treatments and future studies of parasite distributions. Still, approaches based on Taylor's power law have had notable success in explaining determinants of parasite aggregation (Johnson and Hoverman, 2014; Johnson and Wilber, 2017). The first criterion also resulted in exclusion of indices used to differentiate patterns of aggregation across hierarchical levels of sampling, as the appeal of the approach must not depend on such stratification in the data (e.g. comparing aggregation levels of ticks infecting chicks between the within- or among-nests scales; Boulinier et al., 1996).

We therefore considered six measures of parasite aggregation: the variance-to-mean ratio (VMR), the parameter *k* of the negative binomial distribution, mean crowding, the patchiness index, Poulin's *D*, and Hoover's index (the R code used to calculate each of these measures from samples of parasites is included with the associated simulation scripts). The first four were described as general measures of aggregation in early and still valuable treatments of this topic (Pielou, 1977), although they later found varying degrees of popularity (Wilson et al., 2001). Poulin's *D* and Hoover's index are relatively more recent measures, both based on Lorenz curves (Poulin, 1993; McVinish and Lester, 2020). Together, these six measures encompass a variety of conceptualizations of parasite aggregation, both host- and parasite-centric, and are either expected or not expected to relate to mean abundance. We summarize each index below and explain why we excluded some other indices due to their close relations with one or more of these six. Formulas for each of the measures are provided in Table 1; alternative formulations and some additional mathematical details are provided in the Supplementary Data S1.

The first two measures, VMR and *k*, are arguably the most often used measures of parasite aggregation (Wilson et al., 2001; Poulin, 2007b). VMR, besides being easy to calculate, allows a comparison with a theoretical random (Poisson) distribution of parasites, as the variance in parasite abundances in such a case would be equal to the mean (i.e. VMR = 1). Note, however, that use of VMR as a test criterion for randomness versus as a sample statistic measuring aggregation should be distinguished, and that discrete distributions other than Poisson may have a VMR = 1 (Pielou, 1977; Hurlbert, 1990). When parasite distributions are more even than a random distribution, VMR is less than one; when they are more aggregated, VMR is greater than one. Given the strong relationship of VMR with mean abundance and lack of a direct biological interpretation, it is not useful for comparisons of aggregation across samples where the mean abundances differ (Morisita, 1962; Poulin, 1993).

The parameter *k* of the negative binomial distribution provides an inverse measure of aggregation: as aggregation increases, *k* decreases. The nearly universally aggregated distributions of parasites infecting hosts generally can be well approximated by the negative binomial (Shaw and Dobson, 1995), meaning that *k* is often an appropriate measure to quantify aggregation of parasites in natural systems. As *k* tends towards infinity, the negative binomial distribution converges with the Poisson; therefore, if the distribution is not aggregated – i.e. does not differ much from random – aggregation quantified via *k* becomes intractably variable, since minute changes in distribution shape may cause large fluctuations in *k*. The majority of empirical *k* values in natural systems are less than or around one (Shaw and Dobson, 1995). Notably, use of *k* as a measure of aggregation becomes inappropriate in those cases

Table 1

Relationships between different measures of aggregation and other parasitological measures (mean abundance and prevalence; top two rows) observed in simulations, as well as pairwise relationships between aggregation measures (bottom six rows). Relationships are described as either positive (+), negative (–), or absent (o); strong relationships are indicated with duplicated symbols (e.g. ++). Note that relationships listed as lacking (o) do not necessarily mean no correlation, but can signify that the relationship was only weak. Calculations for each of the aggregation measures are provided in the first column (in all cases, σ^2 is the sample variance in abundance, μ is the sample mean, N is the sample size, and x_j (j = 1, 2, ..., N) represents the number of parasites infecting host j). Alternative formulations for some of the measures (e.g. mean crowding can instead be calculated from the sample mean and variance), as well as details of some of the measures' properties (e.g. maxima and minima) are provided in Supplementary Data S1.

	VMR	*k*	Mean crowding	Patchiness	Poulin's *D*	Hoover's index
Related to mean abundance	++	o	++	o	–	–
Related to prevalence	+	+	+	–	– –	– –
VMR $\left(VMR = \frac{\sigma^2}{\mu}\right)$						
k $\left(k = \frac{(\mu^2 - \sigma^2/N)}{(\sigma^2 - \mu)}\right)$ [a]	o					
Mean crowding $\left(\overset{*}{m} = \frac{\sum x_j^2}{\sum x_j} - 1\right)$	++	o				
Patchiness $\left(P = \frac{\overset{*}{m}}{\mu}\right)$	o	– –	o			
Poulin's *D* $\left(D = 1 - \frac{2\sum_{i=1}^{N}\left(\sum_{j=1}^{i} x_j\right)}{\mu N(N+1)}\right)$ [b]	–	–	–	+		
Hoover's index $\left(H = \frac{\sum_{i=1}^{N} \lvert x_i - \mu \rvert}{2\sum_{i=1}^{N} x_i}\right)$	–	–	–	+	++	

VMR, variance-to-mean ratio.

[a] Note that this calculation for *k* assumes a very close fit of the data to a negative binomial distribution, and regardless that in many cases (e.g. extremely low *k*, low sample size, very high μ), it would instead be more accurate to estimate *k* via maximum likelihood approaches rather than using this formula.

[b] Importantly, this formula for Poulin's *D* requires that hosts in the sample be ordered from lowest to highest individual abundances prior to calculation.

where the parasite distribution is not well approximated by a negative binomial.

Mean crowding ($\overset{*}{m}$) is a parasite-centric measure that describes the average number of conspecifics co-occurring on/within a host from the perspective of any individual parasite. That is, $\overset{*}{m}$ quantifies the average intraspecific "crowding" experienced across the sampled parasites.

Patchiness is the mean crowding divided by the sample mean. It therefore describes the average conspecific "crowding" on/within a single host from the perspective of any one parasite, in units of the mean abundance. Alternatively, it can be thought of as how many times more "crowded" an average parasite is, compared with if the same parasites were distributed randomly among hosts (Lloyd, 1967).

Poulin's *D* and Hoover's index are both related to Lorenz curves, commonly used in economics to represent relative inequalities in distributions of wealth (McVinish and Lester, 2020), and range from a minimum of zero (all abundance values equal to the mean) to a theoretical maximum of one (in the parasitological context, all parasites infecting a single host). As they both quantify the deviation of the observed parasite distribution from a hypothetical even distribution towards all the parasites being concentrated in/on a single host, these measures attempt to quantify what many parasitologists naturally intuit as the meaning of aggregation, and allow for comparisons across samples or studies with different mean abundances. Unlike *D*, Hoover's index has an exact biological interpretation: it represents the proportion of parasites that would need to be redistributed to achieve an even distribution among hosts (e.g. a value of 0.7 indicates that 70% of the parasites would need to be redistributed in the sample to achieve evenness).

There are other measures of aggregation that fit most of our selection criteria, but that we chose to exclude because they were too closely related to other values, and not as frequently used. Ives' *J*, similarly to the patchiness index, describes the relative increase in the number of conspecifics co-occurring in a single host compared with a hypothetical random distribution, and is equal to the patchiness minus one (Ives, 1988). Patchiness is also nearly mathematically equivalent to Morisita's index of dispersion, the latter being more common in spatial ecology (Morisita, 1959; Taylor, 1984). The coefficient of variation (*CV*) has also been used as a measure of parasite aggregation, and is equal to $\frac{\sigma}{\mu}$ (the standard deviation divided by the mean abundance; Wilson et al., 2001). Its use is less common, and it is closely correlated with both Poulin's *D* and Hoover's index, although it lacks their useful constraint of ranging from zero to one, as well as a clear biological interpretation (McVinish and Lester, 2020).

We suspected these six selected measures covered the vast majority of indices used in parasitological studies where aggregation could be quantified in a single sample of hosts. To assess this, we first searched Web of Science (https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/webofscience-platform/) for research which included "parasite" (or related words; e.g. "parasitism") in the title, and both "parasite" and either "aggregation" or "overdispersion" in the abstract, as well as any of our considered aggregation measures in the abstract (search term: AB=(parasit* AND (aggregat* OR overdisper*) AND (VMR OR "variance-to-mean" OR k OR "mean crowding" OR patchiness OR "Poulin's D" OR "Hoover's")) AND TI=(parasit*)). We then found that adding Ives' J, Morisita's index, and the coefficient of variation to the abstract search terms (OR "Ives' J" OR Morisita OR CV or "coefficient of variation") only returned seven (8.75%) more records. We explored the properties of these six chosen measures and the interrelationships of these measures with results from our simulations.

3. Simulation 1: effects of random parasite removals on aggregation

3.1. Simulation methods

Our first series of simulations explored the effects of random parasite removals on the six measures of aggregation, testing whether changes in mean abundance which do not otherwise affect the general "shape" of the distribution of parasites among hosts influenced quantified aggregation levels. Here, probabilities of each individual parasite's removal across hosts were proportional to hosts' infrapopulation sizes. This simple comparison serves to categorize the different measures, highlighting how they relate aggregation to properties of the distribution of infection abundances.

To start, a single simulated distribution of parasites infecting 1,000 hosts was generated randomly based on a negative binomial with a mean of five (approximately 5,000 parasites total) and *k* equal to one. That parasite load distributions generally closely match negative binomials in natural systems ensures that this initial state reflects realistic conditions (Shaw and Dobson, 1995). Beginning with this distribution, 2,000 of the parasites were randomly removed one-by-one, with aggregation quantified for each of the six measures at each step (after each single parasite was removed). This process – which is analogous to binomial thinning (subsampling based on the binomial distribution), although implemented in a stepwise manner – was repeated, always beginning with the same initial distribution, over 100 trials. The effects of starting from an initial distribution with either higher or lower mean abundance (mean = 250 or 0.5) or aggregation (*k* = 0.1 or 2) were also explored (detailed results presented in the Supplementary Data S1; Supplementary Figs. S1–S4). In this analysis, we used the corrected moment estimate for quantifying *k* (Table 1) at each step, rather than using maximum likelihood approaches, to increase simulation speed. We were confident that the large sample size and pre-determined fit to the negative binomial distribution guaranteed that the estimation would be accurate, as the accuracy of this estimator depends on these criteria and the mean and *k* values not being intractably high or low, respectively (Bliss and Fisher, 1953; Elliot, 1977; Wilson et al., 2001).

All simulations, both for these random removals and the later correlation analyses, were conducted using the R programming language (version 4.2.1; R Core Team, 2022).

3.2. Results: three contrasting effects of random parasite removals

Three distinct patterns of change in estimated aggregation given random removals emerged across the six measures (Fig. 1). First, there are those measures – *k* and patchiness – that on average do not change with random parasite removals. While the stochastic nature of the simulations results in deviations from the average, these are quite random, and represent a random walk around the original aggregation values (note the very small range of values on the Y-axes in Fig. 1). That these measures should not change with random parasite removals agrees with mathematical descriptions from Pielou (1977); a somewhat analogous demonstration is given in discussion of the Bernoulli damage processes in Johnson et al. (2005). However, that *k* does not seem to be affected by changes in mean abundance does not agree with expectations commonly described in the literature (Scott, 1987; Poulin, 1993; Johnson and Wilber, 2017; Lester and Blomberg, 2021); this is a point worth exploring further.

The second pattern was a strong, linear decrease in aggregation, with little variation around the trend; this pattern was observed for mean crowding and VMR. These patterns again were not sur-

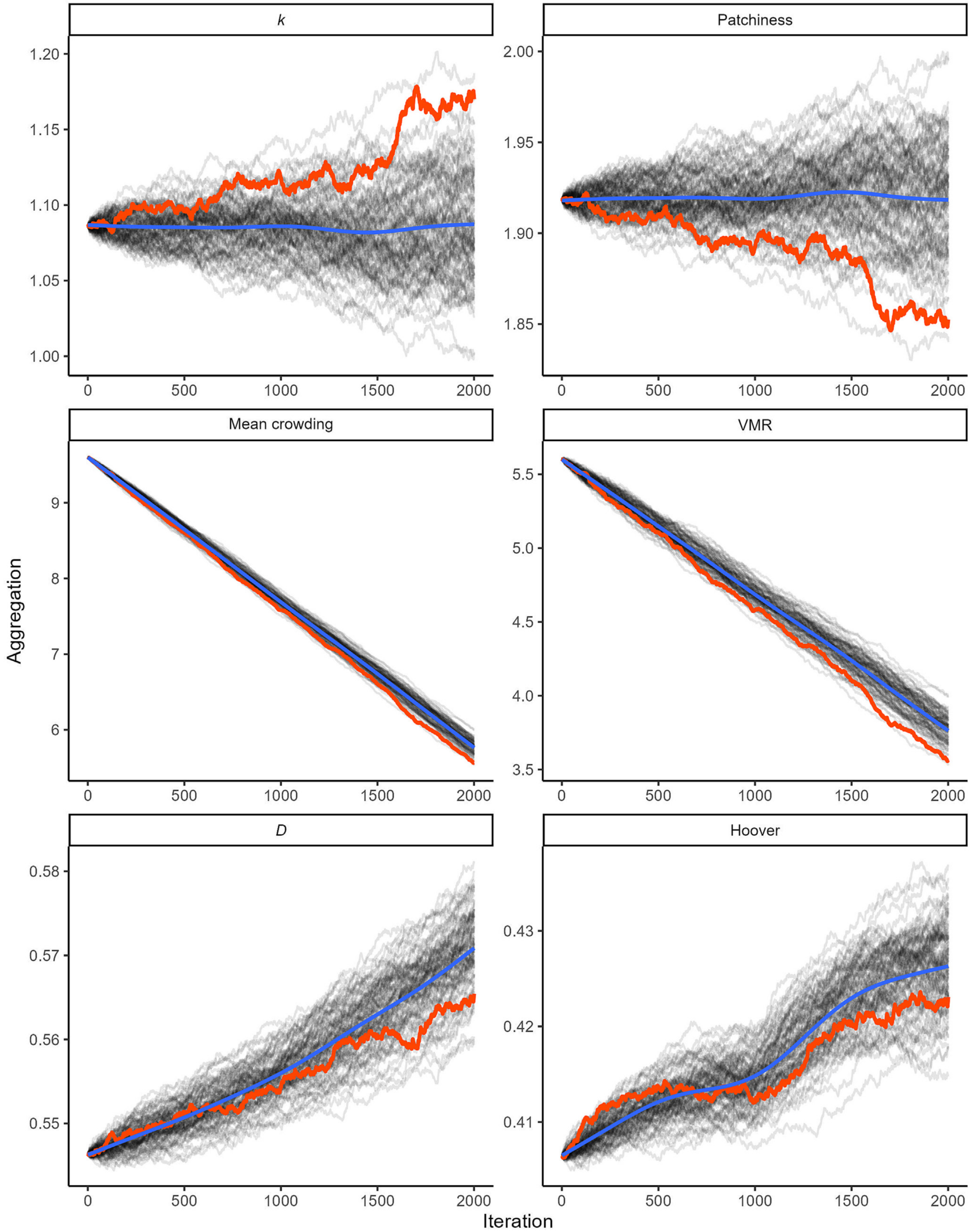

Fig. 1. Aggregation measured after each consecutive removal of a random parasite, repeated across 100 trials. Simulations all began with the same initial negative binomial distribution (mean = 5 and k = 1; n = 1,000). Effects of removals on aggregation demonstrated three contrasting trends, reflecting the measures' various relationships with distribution density: aggregation either fluctuated randomly around the initial value (k and patchiness), showed a strong linear decrease (mean crowding and the variance-to-mean ratio), or demonstrated a slow but significant slight increase (Poulin's D and Hoover's index). The thicker, fluctuating red–orange lines represent the same single trial across all six measures, while the thicker, smoother blue lines are smoothed averages across trials. VMR, variance-to-mean ratio.

prising, and agree with predictions and mathematical treatments by Pielou (1977), as both measures are highly related to distribution mean abundance (rationale for this pattern is discussed in more detail in the following sections considering measure interrelationships and correlations with mean abundance).

The third pattern, observed for Poulin's *D* and Hoover's index, was unexpected, given that there were only two hypothesized possible outcomes of the original thought experiment (i.e. no change in aggregation or a decrease with random removals). These two indices showed a gradual increase in aggregation with removals of parasites. In other words, decreasing the total number of parasites through random removals (eventually) resulted in higher aggregation when quantified using these two measures. This move to higher aggregation likely occurs because the removal of a parasite resulting in a reduction of a small infrapopulation (i.e. smaller than the mean abundance) – particularly when diminishing an infrapopulation to zero – must always have a larger impact on *D* or Hoover's in the positive direction than would the removal of one parasite reducing a large infrapopulation (i.e. larger than the mean abundance) on the measure in the negative direction. The observed pattern is perhaps not surprising in retrospect, given the strong negative relationship between prevalence and both *D* and Hoover's index (Poulin 1993; McVinish and Lester, 2020; prevalence decreased with random removals). These intermittent reductions of small parasite infrapopulations cause a slow, slight, but significant overall trend towards higher aggregation, even though many parasites are also being removed from relatively large infrapopulations; in other words, the positive effect is disproportionate. Additionally, McVinish and Lester (2020) demonstrated how aggregation measured with *D* or Hoover's index decreases as the probability of a parasite being included in the sample increases, using their "incomplete count" setting (binomial thinning); inversely, one could expect aggregation measured using these indices to generally increase as parasites are randomly removed. These explanations agree with patterns presented in the Supplementary Data S1, as the rate of increase in *D* and Hoover's with random removals decreases when the mean abundance of the initial parasite distribution is high (many large infrapopulations; few removals would result in decreases of infrapopulations to zero), and the rate appears to increase when initial mean abundance is low.

This simulation of random parasite removal highlights three categories of aggregation measures, each represented by two measures considered herein. This simulation adds another possible outcome to the hypothetical removal experiment proposed by Pielou (1977), that of aggregation increasing with removals. The qualitative patterns appear robust to changes in the initial parasite distribution; quantitatively, the rate of the increase in *D* and Hoover's appears dependent on mean abundance, and the slope of the relationship with VMR decreases as the initial distribution becomes less aggregated (see Supplementary Data S1).

In the next section, we consider interrelationships of measures more directly, as well as the relationships of measures with both sample prevalence and mean abundance.

4. Simulation 2: interrelationships of aggregation measures, and correlations with other parasitological measures

4.1. Simulation and analytical methods

The aim of the second series of simulations was to generate many parasite distributions of varying mean abundance and aggregation levels that reflected realistic samples of parasites in or on hosts, and then to explore the correlations between the resulting measures of aggregation, and between those measures and mean abundance and prevalence, using principal component analysis (PCA). As such, this analysis will identify interrelationships that are emergent and unavoidable statistical phenomena, as well as ones that arise due to some broadscale (not system-specific) biological constraints. Additionally, this series of simulations will help to further delineate the categorizations, as well as help define properties of the measures, thereby equipping the reader to evaluate the relevance and importance of such measures in their research.

To start, 1,000 parasite distributions for $n = 200$ hosts were randomly generated following negative binomial distributions. *k* values were sampled from a uniform distribution between 0.01 and 2.5. The review and analysis of parasitism of vertebrate host taxa by Shaw and Dobson (1995) suggested that most naturally observed *k*'s fall in this range. Mean abundances of parasites per host were sampled from a lognormal distribution, following the distribution described in the same review by Shaw and Dobson (1995; mean log = 0.4735, S.D. log inferred from other statistics, ≈ 2.754). While this omits any possible covariance between aggregation levels and mean abundance observed in natural systems, it ensures, besides that the means and *k*'s of the distributions follow realistic ranges, a large amount of variation for the correlation analysis. Aggregation was calculated for all 1,000 random distributions using the same six measures as in the previous section, providing the values for the PCA. *k* was estimated using maximum likelihood due to the lower computational cost of analyzing only 1,000 distributions, compared with the first simulations. We recognize that several of the aggregation measures are expected to be sufficiently correlated with one another that they may seem redundant in the analysis (e.g. *k* and patchiness), but we note that the PCA serves as an illustrative tool to demonstrate the potential for correlations between measures.

After simulating the data and measuring the different degrees of aggregation, trials were removed if there were errors calculating *k* (which sometimes occurred during maximum likelihood estimations), if $k > 10$ (which resulted from problematic estimations, or possibly chance occurrences of seemingly random distributions), or if less than three hosts were infected (prevalence < 0.015). Note that rare instances of high *k* values and erroneous calculations of *k* were possible despite the simulations restricting sampling of this parameter between 0.01 and 2.5 (*k* is estimated from the resultant simulated distribution), due to chance emergences of very low variation in abundances or extremely low prevalences. This filtering left 851 of the original 1,000 distributions. As some of the aggregation measures were positively skewed (sometimes extremely so), all those values (*k*, patchiness, mean crowding, and VMR) required log-transformation prior to running the analysis of correlations. Finally, all values (including *D* and Hoover's index) were standardized prior to the PCA, by subtracting means and then dividing by standard deviations. After performing the PCA, two additional variables not used to calculate the principal components (i.e. mean abundance (log-transformed) and prevalence) were projected onto the first two principal components as supplementary variables to evaluate with which of the aggregation measures they demonstrated strong correlations. To confirm that resultant relationships between the mean abundance and VMR were robust and not contingent on unimportant model design decisions, we include in Supplementary Data S2 results of an alternative parameterization of the simulation using randomly generated mean abundances and VMRs to produce the negative binomial distributions, rather than mean abundances and *k*'s. This alternative method also produced positive mean abundance-VMR correlations after the accompanying *k* values were restricted to the realistic range (0.01–2.5).

4.2. Results: three groups of related aggregation measures

The distributions and co-distributions of these transformed variables are visualized in Fig. 2, where several patterns become apparent. First, some pairs of measures are extremely correlated: VMR with mean crowding, *k* with patchiness, and *D* with Hoover's index. These strong correlations were expected from relationships between the measures' modes of calculation, and for the most part, were documented previously (see Pielou, 1977 for relationships between mean crowding, VMR, patchiness, and *k*). To illustrate, VMR should be equal to $\overset{*}{m} - \mu + 1$, where $\overset{*}{m}$ and μ are the mean crowding and mean abundance. Patchiness should be approximately equal to $1 + 1/k$ (assuming the distribution fits well to a negative binomial). As mentioned, both *D* and Hoover's index are related to Lorenz curves, and were expected to be related (McVinish and Lester, 2020).

There are regions highlighted in the co-distributions in Fig. 2 where values do not occur. For example, given a low value of *k* (high aggregation), there appears to be a corresponding lower limit on what values of *D* and Hoover's are possible (somewhat similar lower limits on *D* and Hoover's appear in relation to the other measures as well). Also, there are apparent bounds around the co-distribution of mean crowding and patchiness values, perhaps arising from limits given a certain number of hosts across mean abundances.

The six measures of aggregation are represented in relation to the first two principal components in Fig. 3. Three groups of highly correlated measures appear: Poulin's *D* with Hoover's index (more aligned with the first principal component), VMR with mean crowding (more aligned with the second principal component), and *k* with patchiness (negative correlation). The overall general degrees of correlations between the aggregation measures are summarized in Table 1. The projected supplementary variables in Fig. 3 demonstrate the different relationships between the aggregation measures and both mean abundance and prevalence. Naturally, as mean abundance and prevalence are themselves somewhat correlated, measures of aggregation which correlate with one may be expected to correlate somewhat with the other. However, three distinct patterns of relationships with these parasitological measures were evident. First, prevalence was strongly and negatively associated with *D* and Hoover's index (these two are less strongly, but nonetheless associated negatively, with mean abundance). Second, mean abundance was positively associated with mean crowding and VMR; and third, patchiness and *k* appear almost completely uncorrelated with mean abundance, though are somewhat correlated with prevalence (negatively and positively, respectively). Note that these three pairs of correlated variables with their corresponding contrasting relationships with prevalence and mean abundance match the same three contrasting responses to random parasite removals identified in the first simulations.

Aggregation of the simulated distributions as measured by Poulin's *D* and Hoover's index are plotted in relation to prevalence in Fig. 4. That both measures should be expected to relate negatively to prevalence is well established (Poulin, 1993; McVinish and Lester, 2020). However, the visualization shows an important difference between *D* and Hoover's index with respect to their rela-

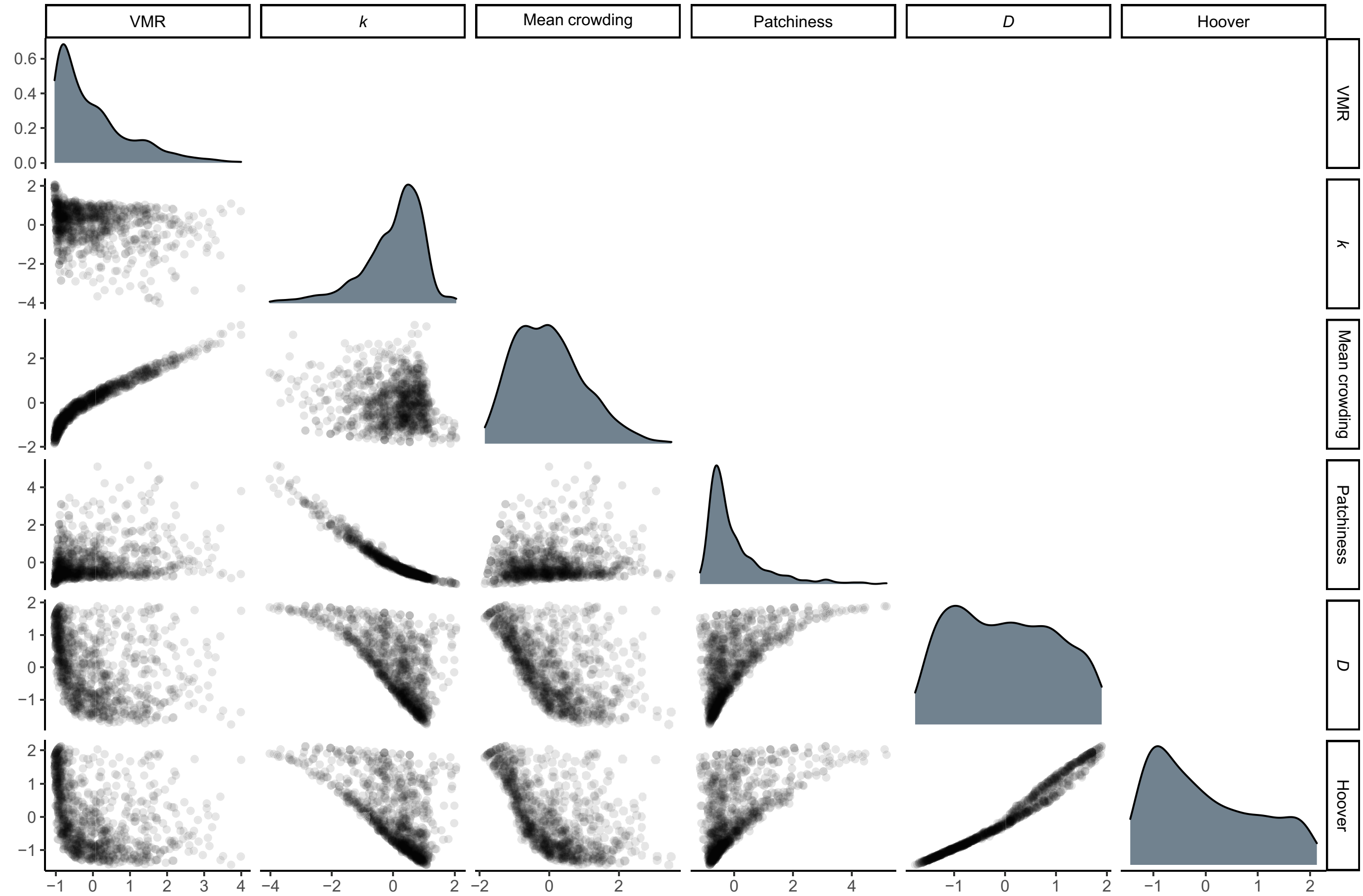

Fig. 2. Distributions (along the diagonal) and co-distributions of aggregation measures calculated from $n = 851$ parasite distributions simulated randomly based on negative binomials. All values except for Poulin's *D* and Hoover's index were $\log_{10}$-transformed. All values were standardized (mean subtracted; divided by standard deviation). VMR, variance-to-mean ratio.

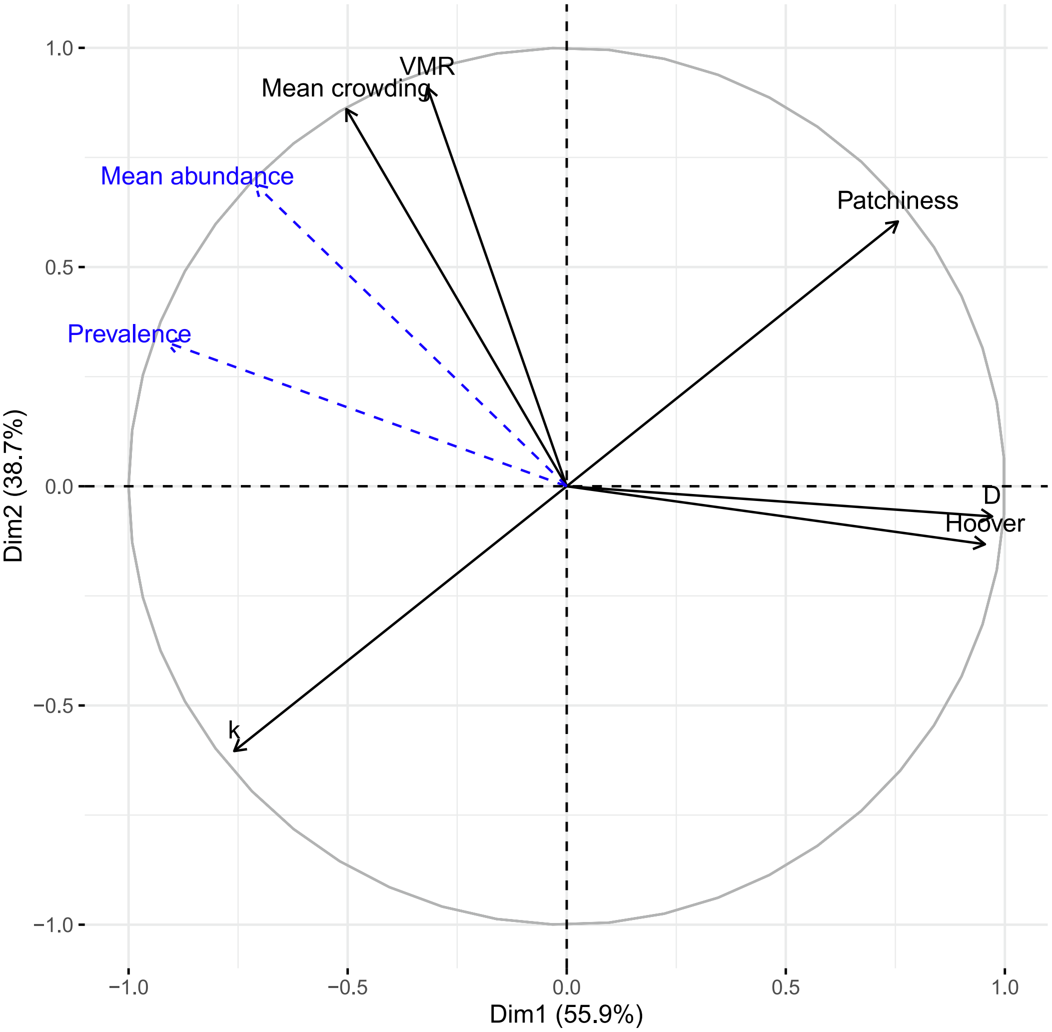

Fig. 3. Aggregation measures plotted on the first two axes of the principal component analysis, with mean abundance and prevalence projected as supplementary variables. Three pairs of correlated measures are apparent; note that *k* is inversely related to aggregation, and is therefore, in terms of aggregation, similar to patchiness. The three pairs of measures differ in their expected relationships with mean abundance and prevalence. The first two principal components explain 94.6% of the variation. VMR, variance-to-mean ratio.

tionship with prevalence: at lower prevalences, particularly those lower than 50%, many of the Hoover values are exactly equal to one minus the prevalence (this measure's theoretical minimal value). This arises from a necessary constraint: Hoover's index will exactly equal one minus the prevalence whenever all infected hosts harbour infrapopulations larger than or equal to the overall mean. When parasite distributions are aggregated, and particularly when prevalence and/or mean abundance is low, it becomes more likely that all infrapopulations will have abundances higher than the overall mean. A consequence is that Hoover's index becomes more and more constrained as prevalence decreases; the potential for variance in Hoover's index becomes extremely limited and is often zero. In the simulations, of those distributions with prevalence less than 25%, only 17% had Hoover indices not exactly equal to one minus the prevalence (i.e. 83% of distributions with prevalence less than 25% had zero variability in aggregation measured by Hoover's index, despite existing variation in degrees of parasite clustering). This constraint is relevant in natural systems as well: using a published dataset of 771 lesser snow geese (*Chen caerulescens*) in which individuals were categorized by age (subadult or adult) and sex, and where nine parasite species were enumerated (Dargent et al., 2017a, 2017b), we found that 27 of the 36 samples (75%) demonstrated Hoover's indices equal to exactly one minus the prevalence. *D*, on the other hand, while clearly negatively related to prevalence (Fig. 4), and while showing a sort of constraint on variation in that there seems to be a defined region of permitted values for certain degrees of prevalence, nonetheless shows much more variation at low prevalence.

The positive correlation of both mean crowding and VMR with sample mean abundance was expected (Fig. 5). While there is room for variation in mean crowding at any given sample mean abundance with varying aggregation, if mean abundance increases and there are more parasite individuals in total, then naturally the expected number of co-infecting parasites from any individual parasite's perspective should generally increase. Note that the minimum value for mean crowding is the mean abundance minus 1, so the lower limit on this measure of aggregation increases with the mean. The positive relationship between VMR and the mean abundance is also intuitive. While the lower limit of the VMR is always zero (given an even distribution), parasite distributions are not expected to be underdispersed; given a random distribution, the variance should equal the mean, and the VMR should be one. However, as parasite distributions are aggregated, the variance should be higher than the mean, and we do not expect the VMR to equal

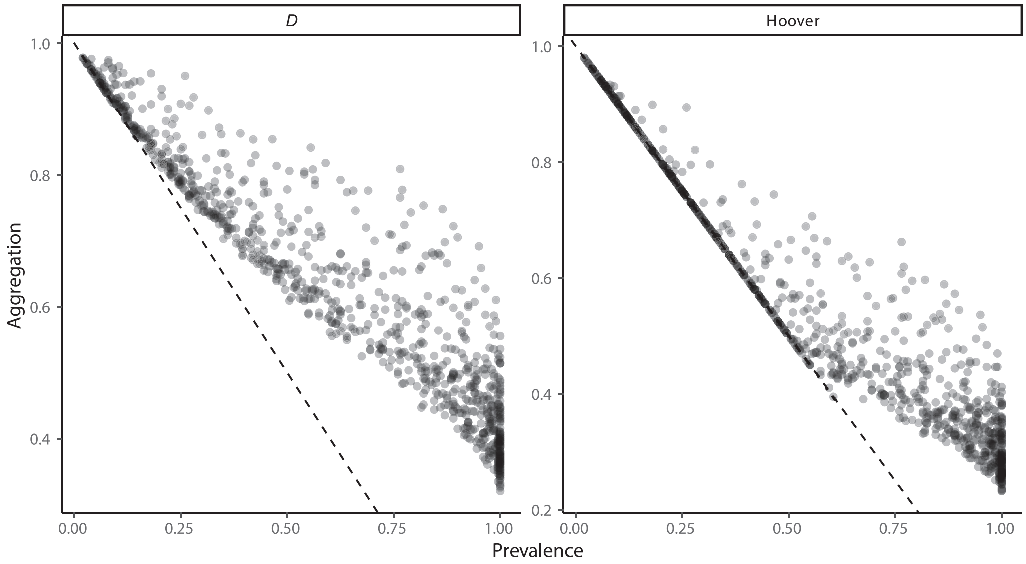

Fig. 4. Aggregation of simulated parasite distributions quantified using Poulin's *D* and Hoover's index in relation to prevalence. The dashed line represents 1 – prevalence (the theoretical minimum value for Hoover's). The lower the prevalence, the more frequently Hoover's index is exactly equal to its minimum, reducing observed variation.

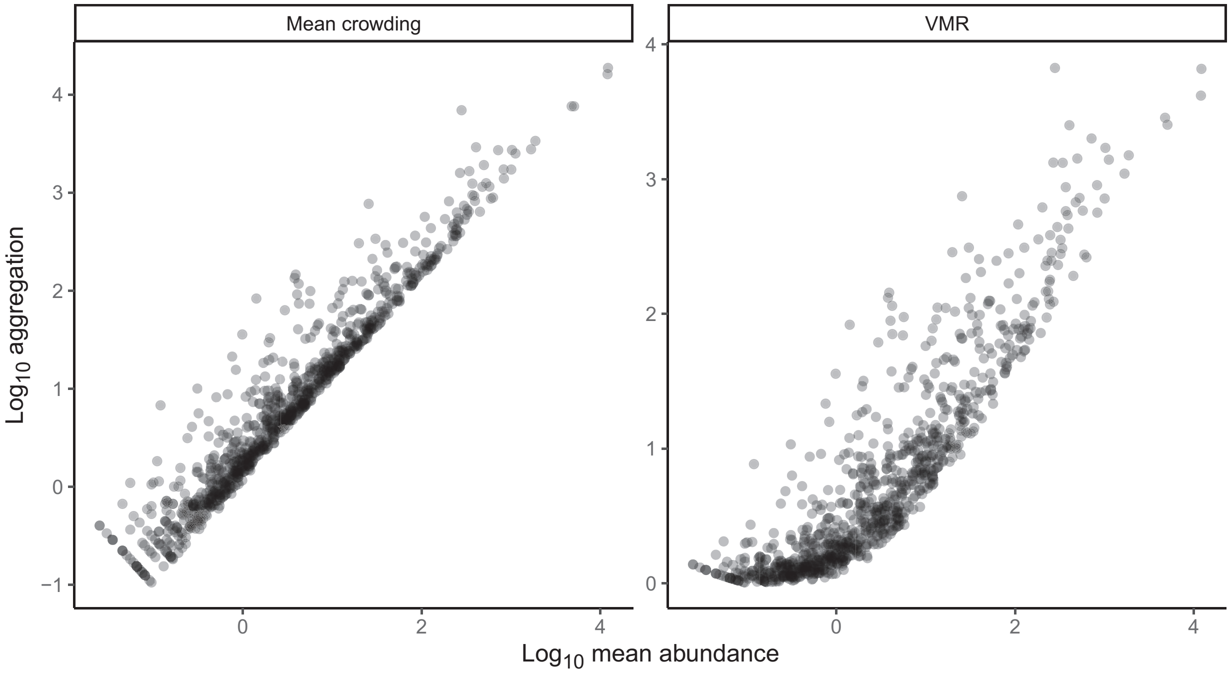

Fig. 5. Aggregation of simulated parasite distributions quantified using mean crowding and the variance-to-mean ratio (VMR), both $\log_{10}$-transformed, in relation to the $\log_{10}$-transformed mean abundance.

one. More importantly, given an aggregated distribution, we expect any increase in the mean to result in a larger increase in the variance (consider that the variance of a negative binomial distribution can be expressed as $\mu + \frac{\mu^2}{k}$, i.e. as a function of the mean added to the squared mean divided by k). In other words, we should expect that as the mean abundance increases, the magnitude of the variance relative to the mean should also increase. Also, while not a direct explanation for the pattern, the upper limit on the VMR is equal to the total number of parasites, and this upper limit naturally increases with increasing mean abundance given a certain number of hosts.

Notably, both patchiness and k appeared unrelated to sample mean abundance, while being only weakly related (negatively and positively, respectively) with prevalence (Fig. 6). The lack of clear relationship with mean abundance is contrary to an often-cited expectation for such a relationship, at least with respect to k (e.g. Scott, 1987; Poulin, 1993; Johnson and Wilber, 2017; Lester and Blomberg, 2021). However, patchiness is not expected to correlate with mean abundance, and (assuming a negative binomial), patchiness and k are directly related (patchiness = 1 + 1/k). Thus, the expectation that k is de facto related to mean abundance is incorrect. Naturally, there is no interdependency between the mean and k parameters in a negative binomial distribution, and therefore no expectation for any relationship emerging from an underlying statistical constraint; whether a relationship between the mean abundance and k, after controlling for prevalence, is common across samples from natural systems is another question. However, the biological interpretation of k, as discussed in the fol-

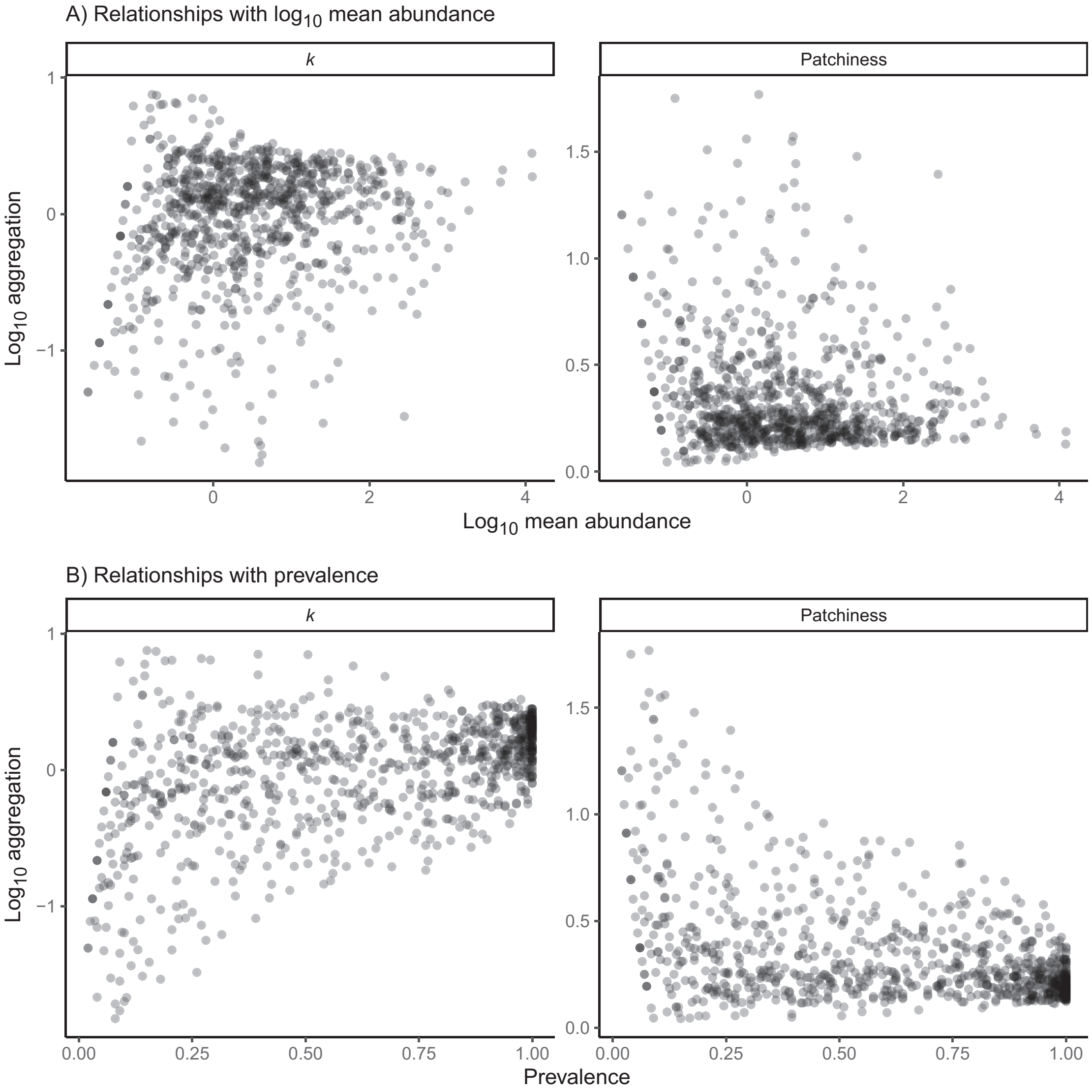

Fig. 6. Aggregation of simulated parasite distributions quantified using the parameter *k* of the negative binomial distribution and the patchiness index, in relation to (A) $\log_{10}$-transformed mean abundance and (B) prevalence.

lowing section, accounts for and encapsulates the sample mean abundance, which could make *k* appropriate for analyses of aggregation even in cases where mean abundance varies across samples. We also discuss further potential reasons for the misconstrued association between *k* and mean abundance in the following section. Both *k* and patchiness appear correlated, although weakly, with prevalence (positively and negatively). Either way, aggregation slightly decreased with increased prevalence. This relationship with prevalence is not nearly as strong as that for *D* and Hoover's index. In the bottom two panels in Fig. 6, the simulated parasite abundances appear to produce an approximately triangular distribution of points, perhaps outlining constraints on possible aggregation values given certain prevalence values. As prevalence decreases, "permitted" variation in *k* or patchiness appears to increase, so long as prevalence is not too low.

5. A user's guide to aggregation measures: considerations and recommendations

Differences between measures of parasite aggregation in their "meanings" were evident given differences in their properties that we observed. These dissimilarities between measures make them differentially appropriate for use in various research contexts, despite them all purporting to describe a common population-level natural phenomenon. Among the six widely used measures of parasite aggregation we considered, there exist three groups of paired measures with similar properties, i.e. similar both in terms of their expected correlations with each other and with mean abundance and prevalence. We suspect that the three identified groups represent general categories that capture similar, single-sample-appropriate aggregation measures. Other aggregation indices that could also be applied to single parasite taxon-single host taxon individual samples but that were not included in this research (e.g. Ives' *J*, or the coefficient of variation) are closely related to one or more of the measures considered herein. The identified interrelationships and correlations with other parasitological values, together with the potential and different biological interpretations of the aggregation measures, highlight the advantages and disadvantages of each, and suggest guiding principles for their use in different contexts.

The general advantages and disadvantages of each of the six aggregation measures are summarized in Table 2. These advantages and disadvantages can help researchers discuss and choose the most appropriate measure(s) for their study problem(s), based on context-specific answers to questions, and as such Table 2 may serve as a first reference point for those who are faced with the choice of how best to quantify parasite aggregation. For example, are researchers interested in how aggregation compares between samples/populations/species (in which case any strong, necessary relationship with mean abundance should be avoided), or is the assessment specific to quantifying aggregation in a single sample (e.g. asking if the distribution differs from random)? Care should be taken in considering the relevant null before inferences about aggregation are made; are deviations from a Poisson (random) or from an even distribution of parasites requiring explanation? Researchers also must be concerned with whether variation in aggregation across a collection of samples is sufficient to allow its use as an explanatory or response variable in analytical models.

Table 2

Potential advantages and disadvantages of the six considered measures of aggregation. Among other considerations, specific attention is given to each measure's ease of interpretation, whether or not they lend themselves to comparisons across studies, and to what degree they correlate with other parasitological measures such as mean abundance and prevalence. Any biological interpretations of the different measures are written in italics under the "Advantages" column. Note that the listed disadvantages of certain measures may be context-specific, and in certain circumstances may not actually be disadvantageous; for example, in studies interested in quantifying parasite intraspecific competition arising from aggregation, a measure such as mean crowding could be the most appropriate despite (or even due to) its strong relationship with mean parasite abundance.

Aggregation measure	Advantages	Disadvantages
Variance-to-mean ratio (VMR)	Easy to calculate Provides quick estimation of whether a parasite distribution may be more aggregated than random (VMR > 1) or less aggregated than random (more even; VMR < 1) The significance of a distribution's deviation from randomness can be estimated using tests based on the VMR (e.g. Sun and Hughes, 1994)	No direct biological interpretation, outside of a distribution's potential deviation from randomness Strongly positively correlated with sample mean abundance (less strongly, but still positively, correlated with prevalence) Inappropriate for comparisons between samples due to its direct relationship with mean abundance
k	Not necessarily correlated with mean abundance; only weakly correlated with prevalence Direct biological interpretation: *the reciprocal of* k *describes the proportion of the mean abundance by which the mean crowding exceeds the mean abundance* (e.g. if $k = 0.5$, or ½, the average number of conspecifics infecting with a given parasite is two × mean abundance greater than the mean abundance) Realistic samples of parasite loads can be simulated from negative binomial distributions based on biologically relevant values of *k*	The use of *k* becomes inappropriate when parasite distributions are not well approximated by a negative binomial; its applicability as a measure of aggregation, and its biological interpretation, are lost when this condition is not satisfied Weakly positively related to prevalence (i.e. aggregation decreases with increasing prevalence, as *k* is an inverse measure of aggregation)
Mean crowding	Direct biological interpretation: *the average number of conspecifics infecting together with a single parasite; i.e. the expected number of conspecifics sharing the same host*	Strongly positively correlated with sample mean abundance Not immediately interpretable in terms of the host-perspective experience of parasite aggregation Historically relatively less common in the parasitological literature
Patchiness	Not necessarily correlated with mean abundance; only weakly correlated with prevalence Direct biological interpretation: *the average amount of conspecific co-occurrence from the perspective of a single parasite, in units of the sample mean abundance* (mean crowding divided by the mean abundance). Alternatively, it is *how many times as crowded an individual parasite is, compared* with *a hypothetical Poisson distribution.* Can be directly converted to *k* for a historically more common measure when the distribution is well-approximated by a negative binomial ($k = 1 / [\text{patchiness} - 1]$)	Historically relatively less common in the parasitological literature Weakly negatively related to prevalence
Poulin's *D*	Interpretable with reference to what many parasitologists mean by aggregation: *the degree to which a given number of parasites infecting a sample of hosts deviates from an even distribution (*$D = 0$*) towards the most extreme aggregation possible (all parasites on a single host;* ($D = 1$*)* Owing to its interpretation and its constraint to falling between zero and one, *D* lends itself well to comparisons of aggregation between studies	Strongly negatively correlated with prevalence; weakly negatively correlated with mean abundance While capturing an arguably intuitive general meaning of aggregation, exact values of *D* have little direct biological interpretation (e.g. little can be said about the difference between a sample with $D = 0.75$ and one with $D = 0.70$, other than that the former appears more aggregated) Actual permitted maximum value only approaches one when given a sufficiently large sample size
Hoover's index	While similar to Poulin's *D* in that it theoretically falls between zero and one (i.e. from an even distribution to maximum possible aggregation), it has a more directly biologically interpretable meaning: *the proportion of parasites that would need to be moved to achieve an even distribution* Owing to its interpretation and its constraint to falling between zero and one, this measure lends itself well to comparisons of aggregation between studies	Permitted variation in Hoover's index can be limited, or even zero, particularly at low prevalences of infection (it is always equal to 1 – prevalence when all parasite loads are equal to or larger than the mean abundance) Strongly negatively correlated with prevalence, weakly negatively correlated with mean abundance Actual permitted maximum value only approaches one when given a sufficiently large sample size Only recently publicized as a measure of parasite aggregation (few examples in the literature; McVinish and Lester, 2020)

Of course, there is also the question of how interpretable the measure is, relevant to the biological inquiry. The relationships between the answers to these and similar questions, and the suitability of the different measures, are discussed throughout the rest of this section.

All six measures correlate to some degree with either or both of mean abundance and prevalence. As mentioned, this is expected because these two variables and aggregation all seek to quantify properties of the distribution of parasites among hosts, and therefore cannot be entirely independent. VMR and mean crowding – together constituting one of the three identified groups – demonstrated strong positive correlations with mean abundance and are therefore generally not suitable for comparisons between samples where means differ, as mentioned. The VMR can, however, provide a straightforward evaluation of whether a distribution may differ from random; one can statistically test the null hypothesis that VMR = 1 (i.e. the expected value for a Poisson distribution in a given sample; Pielou, 1977; Sun and Hughes, 1994). For example, one might be interested in experimentally testing whether aggregating factors remain (i.e. do parasite distributions remain aggregated?) after controlling for hypothesized explanatory effects relating to heterogeneity in exposure and/or host characteristics (Karvonen et al., 2004; Johnson and Hoverman, 2014); however, testing for differences in aggregation, rather than an absence of aggregation, may be more appropriate. Mean crowding is arguably the measure that most explicitly takes the parasite's perspective, and therefore may be of the highest value in parasite-centric studies (e.g. those assessing intensity-dependent effects on infrapopulation dynamics, or analyzing individual parasite growth or fecundity within infrapopulations).

All the considered measures correlated with prevalence, although inconsistently, as some correlations were positive while others were negative. The (weaker) positive correlations between prevalence and both VMR and mean crowding likely arose due to the strong relationship of those measures with mean abundance; in other words, they are likely emergent due to an underlying coincident relationship between prevalence and mean abundance (Poulin, 2007b). All the other indices demonstrated increasing aggregation with decreasing prevalence. This negative relationship makes sense, considering aggregation colloquially as the relative clustering of parasites among hosts: as prevalence decreases, the sampled parasites are more "concentrated" within a smaller proportion of hosts (Poulin, 1993).

Poulin's *D* and Hoover's index showed particularly strong, and expected, negative correlations with prevalence. As described above, our analyses revealed that potential variability in Hoover's index, unlike variability in *D*, may be extremely diminished at low prevalence. Hoover's index is invariably equal to one minus the prevalence when all infrapopulations are larger than, or equivalent to, the mean abundance. Intuitively, this constraint is revealed in the biological interpretation of this index, i.e. that it represents the proportion of the parasite population that would need to be redistributed to achieve an even distribution. When all infrapopulations are at least as large as the mean abundance, the only hosts that would receive parasites in a hypothetical redistribution to reach evenness would be those that are uninfected; the number of uninfected hosts would be $N_{hosts} \times (1 - \text{prevalence})$. The total number of parasites moved would be this quantity of hosts multiplied by the mean abundance (μ). Finally, to arrive at what proportion of the parasite population this represents, we would divide by $N_{hosts} \times \mu$; therefore, the only term not cancelled out is the one minus prevalence term. Hoover's index's use as a highly interpretable measure of aggregation is not lost in any contexts; however, our study suggests low prevalence, high aggregation, and/or low mean abundance can result in cases where it demonstrates extremely limited or no possible variation. As such, Hoover's index may not be appropriate in some comparative studies, or those studies involving analytical models with aggregation included as a variable, when explanatory/explainable variation is lacking, but otherwise might be preferred over *D* when biological interpretation of degree of aggregation is paramount.

In comparison, Poulin's *D* did not demonstrate a similar impediment, and showed far less of a constraint on variation given variable parasite prevalence. While *D* is particularly apt for comparisons of aggregation across samples and even across different host-parasite associations, this permitted variability also allows its use as a response variable in analytical models seeking to explain predictors of parasite aggregation, or as an independent variable in models testing for biological consequences of aggregation among hosts. For example, one may wish to compare whether degrees of parasite aggregation differ among host groups with variable expression of behavioural resistance mechanisms (cf. Horn et al., 2023), or whether aggregation predicts expected within-host genetic variation of parasites (cf. Eppert et al., 2002; Poulin 2007b). While the strong underlying relationship between *D* and prevalence may still limit its usefulness, researchers have successfully modeled *D* as a response variable, for example in beta regressions demonstrating the influence of parasite taxonomy on degree of aggregation in a common host species (Morrill et al., 2022).

The final pair of measures that group together are the parameter *k* of the negative binomial distribution and patchiness. *k* and patchiness demonstrated only weak correlations with prevalence in the second simulation (aggregation increasing with decreasing prevalence), and essentially no correlation with mean abundance. This lack of a necessary relationship with mean abundance, as well as their biological interpretations (Table 2), makes them particularly valuable for comparisons of aggregation across studies. As mentioned, the absence of any correlation between *k* and mean abundance in our analyses will seem unanticipated to some readers, as expectations for such an unavoidable relationship have been discussed (e.g. Scott, 1987; Lester and Blomberg, 2021; although see Gregory and Woolhouse, 1993). We expect that previously identified relationships with mean abundance were largely due to the relationship of *k* with prevalence. Such an occurrence could arise from the frequent correlation between the mean and prevalence, and thereby result in a spurious correlation between mean abundance and *k*. Also, if parasite distributions deviated from approximating negative binomials, poorly fitted values of *k* would result. Those poorly fitted *k* values may demonstrate unexpected behaviours, even suggesting a false relationship with mean abundance. Alternatively, confusion may have arisen from the fact that distributions with different mean abundances but the same values of *k* can have very different shapes (for example, prevalence estimates may be dissimilar). Here, distributions clearly differ but this is not reflected in values of *k*; the interpretation of *k* with respect to aggregation in those varying conditions does not change, as its reciprocal still represents the proportion of the mean abundance by which the mean crowding exceeds the mean abundance (e.g. if $k = ¼$, then mean crowding is four times in excess of the mean abundance, i.e. equal to five times the mean abundance).

A few other points regarding *k* are worth highlighting. Several studies state that *k* does not have a direct biological interpretation (Taylor, 1984; McVinish and Lester, 2020), but as demonstrated by Lloyd (1967), *k* does possess a clear interpretation in terms of mean crowding, i.e. as aggregation interpreted from the parasite's perspective. Returning to the previous purported relationship between *k* and mean abundance, aggregation as measured by *k* on average did not change following random removals of parasites in our simulations. We thus reaffirm *k*'s value as an interpretable aggregation measure that is not constrained to relate to mean abundance. *k* may also prove useful as a response or explanatory variable in models exploring correlates of aggregation. Comparisons of aggregation between groups of interest may also now be accomplished using modern implementations of generalized linear model-fitting procedures with a negative binomial response whereby different *k* parameters can be fit to each of the focal groups. This desirable approach was identified as impossible in previous decades due to technological limitations (Shaw et al., 1998; Morrill et al., 2022). The disadvantage of *k* as a measure of aggregation, on the other hand, remains that its relevance, ability to be meaningfully converted to other measures, and its biological interpretation are lost as soon as the distribution of parasite abundances is no longer well approximated by a negative binomial.

Patchiness arises as a potentially underappreciated measure of aggregation. Despite its infrequent use in the parasitological literature, patchiness has the advantages that it is not correlated with mean abundance, it does not require that the distribution of parasite loads fits well to a negative binomial, and it boasts a clear biological interpretation. Between these qualities and its ease of calculation, we submit patchiness as a measure to (re)consider. Several of its useful properties, and the calculation for its associated standard error, are presented in Lloyd (1967). Patchiness may also be appropriate as either an outcome variable or predictor in models testing relationships between aggregation and other ecological variables across collections of separate samples of hosts. When choosing between patchiness and *k*, the latter may be preferable when one wishes to generate simulated parasite abundance data (from a negative binomial distribution) based on observations, or where direct comparisons with previously reported levels of *k* are desired; otherwise, and certainly whenever the observed abundance data is not well-approximated by a negative binomial, patchiness may be preferred.

In light of the three contrasting groups of measures identified in our analyses, we add the recommendation that authors include at least one measure from each of the three groups (e.g. VMR or mean crowding, combined with *k* or patchiness, and *D* or Hoover's index). The expected mathematical relationships between some of the measures, at least when the distribution is satisfactorily approximated by a negative binomial, ensure that the values of still other indices may be calculated even if they are not listed. This may be beneficial for quantitative reviews and meta-analyses, for instance analyses seeking to uncover general epidemiological processes shaping the emergent distributions of parasites among hosts.

Regardless of which measures researchers select for their analyses, we repeat the recommendation of Reiczigel et al. (2019), that these aggregation values always be reported with some degree of the associated error, as this is not a common enough practice (e.g. bias-corrected and accelerated bootstrap confidence intervals for *D*; maximum likelihood estimation confidence intervals for *k*). Finally, we remind the reader that, due to the relative rarity of larger infrapopulations among hosts when abundances are positively skewed, aggregation measures tend to be biased towards underestimations at low host sample sizes (Gregory and Woolhouse, 1993; Poulin, 2007b).

6. Conclusions

Whether it is to quantify the distribution of parasites among hosts, or more generally that of individual organisms among discrete sampling units (e.g. quadrats, core samples, etc.), different measures of aggregation are commonly used. We considered six measures (VMR, *k*, mean crowing, patchiness, Poulin's *D*, and Hoover's Index) all purported to measure parasite aggregation, but shown to differ in what they depict relative to one another and relative to sample prevalence and mean abundance of parasites, based on simulations. These different measures are straightforward to calculate, and can be used to address various determinants of aggregation and its consequences to parasite and host populations (what is observable in nature).

The six measures are better characterized as three groups of paired measures (*k* and patchiness, VMR and mean crowding, Poulin's *D* and Hoover's index) with similarities between correlated members of a pair in terms of their response to random parasite removal and, relatedly, their relations to mean abundance and/or prevalence. We provide a list of advantages and disadvantages of each aggregation measure based on results of simulations and the range of variation to be explained (or used to explain), the past use of each measure, their ease of calculation and conversion to other measures, and their biological interpretability. This information is meant to guide researchers in the choice of aggregation measures best suited to their study questions while also encouraging researchers to provide information on other measures (perhaps as supplementary information) to allow across-study tests of determinants and consequences of parasite aggregation.

Acknowledgements

Funding was provided through a Natural Sciences and Engineering Research Council of Canada (NSERC) CGS-D scholarship to AM and an NSERC Discovery grant to MRF (no. 100118). We are grateful to Prof. Ken Wilson for comments on an earlier version of the manuscript. The scripts to run all the simulations that produce the data used in analyses are included as Supplementary Data S3.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpara.2023.06.004.

References

Anderson, R.M., May, R.M., 1978. Regulation and stability of host-parasite population interactions: I. Regulatory processes. J. Anim. Ecol. 47, 219–247. https://doi.org/10.2307/3933.

Bliss, C.I., Fisher, R.A., 1953. Fitting the negative binomial distribution to biological data. Biometrics 9, 176–200. https://doi.org/10.2307/3001850.

Boag, B., Lello, J., Fenton, A., Tompkins, D., Hudson, P.J., 2001. Patterns of parasite aggregation in the wild European rabbit (*Oryctolagus cuniculus*). Int. J. Parasitol. 31, 1421–1428. https://doi.org/10.1016/S0020-7519(01)00270-3.

Boulinier, T., Ives, A., Danchin, E., 1996. Measuring aggregation of parasites at different host population levels. Parasitology 112, 581–587. https://doi.org/10.1017/S0031182000066166.

Cornell, S., Isham, V., Smith, G., Grenfell, B., 2003. Spatial parasite transmission, drug resistance, and the spread of rare genes. Proc. Natl. Acad. Sci. U.S.A. 100, 7401–7405. https://doi.org/10.1073/pnas.0832206100.

Cox, R., Groner, M., Todd, C.D., Gettinby, G., Patanasatienkul, T., Revie, C., 2017. Mate limitation in sea lice infesting wild salmon hosts: The influence of parasite sex ratio and aggregation. Ecosphere 8, e02040.

Criscione, C.D., Blouin, M.S., 2005. Effective sizes of macroparasite populations: A conceptual model. Trends Parasitol. 21, 212–217. https://doi.org/10.1016/j.pt.2005.03.002.

[dataset] Dargent, F., Morrill, A., Alisauskas, R.T., McLaughlin, J.D., Shutler, D., Forbes, M.R., 2017a. Data from: Lesser snow goose helminths show recurring and positive parasite infection-diversity relations. https://doi.org/10.5061/dryad.407c5.

Dargent, F., Morrill, A., Alisauskas, R.T., McLaughlin, J.D., Shutler, D., Forbes, M.R., 2017b. Lesser snow goose helminths show recurring and positive parasite infection-diversity relations. Int. J. Parasitol. Parasites Wildl. 6, 22–28. https://doi.org/10.1016/j.ijppaw.2017.01.003.

Elliott, J.M., 1977. Statistical analysis of samples of benthic invertebrates. Freshwater Biological Association, Ambleside, England.

Eppert, A., Lewis, F., Grzywacz, C., Coura-Filho, P., Caldas, I., Minchella, D., 2002. Distribution of schistosome infections in molluscan hosts at different levels of parasite prevalence. J. Parasitol. 88, 232–236. https://doi.org/10.1645/0022-3395(2002)088[0232:DOSIIM]2.0.CO;2.

Gregory, R., Woolhouse, M., 1993. Quantification of parasite aggregation: A simulation study. Acta Trop. 54, 131–139. https://doi.org/10.1016/0001-706X(93)90059-K.

Horn, C.J., Liang, C., Luong, L.T., 2023. Parasite preferences for large host body size can drive overdispersion in a fly-mite association. Int J Parasitol 53, 327–332. https://doi.org/10.1016/j.ijpara.2023.03.003.

Hudson, P.J., Newborn, D., Dobson, A.P., 1992. Regulation and stability of a free-living host-parasite system: *Trichostrongylus tenuis* in red grouse. I. Monitoring and parasite reduction experiments. J. Anim. Ecol. 61, 477–486. https://doi.org/10.2307/5338.

Hurlbert, S.H., 1990. Spatial distribution of the montane unicorn. Oikos 58, 257–271. https://doi.org/10.2307/3545216.

Ives, A.R., 1988. Aggregation and the coexistence of competitors. Ann. Zool. Fenn. 25, 75–88.

Johnson, P.T., Hoverman, J.T., 2014. Heterogeneous hosts: How variation in host size, behaviour and immunity affects parasite aggregation. J. Anim. Ecol. 83, 1103–1112. https://doi.org/10.1111/1365-2656.12215.

Johnson, N.L., Kemp, A.W., Kotz, S., 2005. Stopped-sum distributions. Univariate Discrete Distributions. John Wiley & Sons Inc, Haboken, New Jersey.

Johnson, P.T., Wilber, M.Q., 2017. Biological and statistical processes jointly drive population aggregation: Using host–parasite interactions to understand Taylor's power law. Proc. R. Soc. B: Biol. Sci. 284, 20171388. https://doi.org/10.1098/rspb.2017.1388.

Karvonen, A., Hudson, P.J., Seppälä, O., Valtonen, E.T., 2004. Transmission dynamics of a trematode parasite: Exposure, acquired resistance and parasite aggregation. Parasitol. Res. 92, 183–188. https://doi.org/10.1007/s00436-003-1035-y.

Lester, R.J., Blomberg, S.P., 2021. Three methods to measure parasite aggregation using examples from Australian fish parasites. Methods Ecol. Evol. 12, 1999–2007. https://doi.org/10.1111/2041-210X.13668.

Lloyd, M., 1967. Mean crowding. J. Anim. Ecol. 36, 1–30.

McVinish, R., Lester, R., 2020. Measuring aggregation in parasite populations. J. R. Soc. Interface 17, 20190886. https://doi.org/10.1098/rsif.2019.0886.

Møller, A., 2005. Parasitism and the regulation of host populations. In: Thomas, F., Renaud, F., Guégan, J.-.-F. (Eds.), Parasitism and Ecosystems. Oxford University Press, New York, pp. 43–53.

Morisita, M., 1959. Measuring of interspecific association and similarity between communities. Memoirs of the Faculty of Science. Kyushu University. Series E (Biology) 3, 65–80.

Morisita, M., 1962. I σ-index, a measure of dispersion of individuals. Res. Popul. Ecol. 4, 1–7. https://doi.org/10.1007/BF02533903.

Morrill, A., Provencher, J., Forbes, M., 2014. Testing for dual impacts of contaminants and parasites on hosts: The importance of skew. Environ. Rev. 22, 445–456. https://doi.org/10.1139/er-2014-0026.

Morrill, A., Dargent, F., Forbes, M.R., 2017. Explaining parasite aggregation: More than one parasite species at a time. Int. J. Parasitol. 47, 185–188. https://doi.org/10.1016/j.ijpara.2016.11.005.

Morrill, A., Nielsen, Ó., Skírnisson, K., Forbes, M., 2022. Identifying sources of variation in parasite aggregation. PeerJ 10, e13763. https://doi.org/10.7717/peerj.13763.

Perkins, S.E., Cattadori, I.M., Tagliapietra, V., Rizzoli, A.P., Hudson, P.J., 2003. Empirical evidence for key hosts in persistence of a tick-borne disease. Int. J. Parasitol. 33, 909–917. https://doi.org/10.1016/S0020-7519(03)00128-0.

Pielou, E.C., 1977. The measurement of aggregation. In: Mathematical Ecology. Wiley-Interscience, New York.

Poulin, R., 1993. The disparity between observed and uniform distributions: A new look at parasite aggregation. Int. J. Parasitol. 23, 937–944. https://doi.org/10.1016/0020-7519(93)90060-C.

Poulin, R., 2007a. Are there general laws in parasite ecology? Parasitology 134, 763–776. https://doi.org/10.1017/S0031182006002150.

Poulin, R., 2013. Explaining variability in parasite aggregation levels among host samples. Parasitology 140, 541–546. https://doi.org/10.1017/S0031182012002053.

Poulin, R., 2007b. Parasite aggregation: Causes and consequences. In: Evolutionary Ecology of Parasites. Princeton University Press, Princeton, New Jersey.

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reiczigel, J., Marozzi, M., Fábián, I., Rózsa, L., 2019. Biostatistics for parasitologists – a primer to quantitative parasitology. Trends Parasitol. 35, 277–281. https://doi.org/10.1016/j.pt.2019.01.003.

Rosà, R., Pugliese, A., 2002. Aggregation, stability, and oscillations in different models for host-macroparasite interactions. Theor. Popul. Biol. 61, 319–334. https://doi.org/10.1006/tpbi.2002.1575.

Scott, M.E., 1987. Temporal changes in aggregation: A laboratory study. Parasitology 94, 583–595. https://doi.org/10.1017/S0031182000055918.

Shaw, D., Dobson, A., 1995. Patterns of macroparasite abundance and aggregation in wildlife populations: A quantitative review. Parasitology 111, S111–S133. https://doi.org/10.1017/S0031182000075855.

Shaw, D., Grenfell, B., Dobson, A., 1998. Patterns of macroparasite aggregation in wildlife host populations. Parasitology 117, 597–610. https://doi.org/10.1017/S0031182098003448.

Sun, P., Hughes, G., 1994. On a statistical test to detect spatial pattern. Biom. J. 36, 751–758. https://doi.org/10.1002/bimj.4710360616.

Taylor, L., 1984. Assessing and interpreting the spatial distributions of insect populations. Annu. Rev. Entomol. 29, 321–357. https://doi.org/10.1146/annurev.en.29.010184.001541.

Wilson, K., Bjørnstad, O., Dobson, A., Merler, S., Poglayen, G., Randolph, S., Read, A., Skorping, A., 2001. Heterogeneities in macroparasite infections: Patterns and processes. In: Hudson, P., Rizzoli, A., Grenfell, B., Heesterbeek, H., Dobson, A. (Eds.), The Ecology of Wildlife Diseases. Oxford University Press, New York, pp. 6–44.