International Journal for Parasitology 44 (2014) 437–445

Contents lists available at ScienceDirect

International Journal for Parasitology

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

The reliability of observational approaches for detecting interspecific parasite interactions: comparison with experimental results

Andy Fenton [a,*], Sarah C.L. Knowles [b], Owen L. Petchey [c], Amy B. Pedersen [d]

[a] Institute of Integrative Biology, University of Liverpool, Biosciences Building, Crown Street, Liverpool L69 7ZB, UK
[b] Department of Infectious Disease Epidemiology, Imperial College London, St. Mary's Campus, Norfolk Place, London W2 1PG, UK
[c] Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
[d] Institute of Evolutionary Biology and Centre for Immunity, Infection and Evolution (CIIE), School of Biological Sciences, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

ARTICLE INFO

Article history:
Received 7 December 2013
Received in revised form 27 February 2014
Accepted 2 March 2014
Available online 3 April 2014

Keywords:
Coinfection
Helminths
Polyparasitism
Cross-sectional and longitudinal analyses
Field study
Experimental perturbation
Small mammals

ABSTRACT

Interactions among coinfecting parasites have the potential to alter host susceptibility to infection, the progression of disease and the efficacy of disease control measures. It is therefore essential to be able to accurately infer the occurrence and direction of such interactions from parasitological data. Due to logistical constraints, perturbation experiments are rarely undertaken to directly detect interactions, therefore a variety of approaches are commonly used to infer them from patterns of parasite association in observational data. However, the reliability of these various approaches is not known. We assess the ability of a range of standard analytical approaches to detect known interactions between infections of nematodes and intestinal coccidia (*Eimeria*) in natural small-mammal populations, as revealed by experimental perturbations. We show that correlation-based approaches are highly unreliable, often predicting strong and highly significant associations between nematodes and *Eimeria* in the opposite direction to the underlying interaction. The most reliable methods involved longitudinal analyses, in which the nematode infection status of individuals at one month is related to the infection status by *Eimeria* the next month. Even then, however, we suggest these approaches are only viable for certain types of infections and datasets. Overall we suggest that, in the absence of experimental approaches, careful consideration be given to the choice of statistical approach when attempting to infer interspecific interactions from observational data.

1. Introduction

Interspecific parasite interactions are a major research focus in disease ecology. Most hosts, including humans in communities around the globe, are coinfected by many parasite species (Petney and Andrews, 1998; Cox, 2001). Numerous laboratory studies (Behnke et al., 1978; Christensen et al., 1987; Adams et al., 1989; Frontera et al., 2005) have shown coinfecting parasites can interact strongly, either positively or negatively (Griffiths et al., 2011), with important implications for disease progression, transmission and control. In particular, if strong interactions are present then targeted treatment may result in potentially unwanted responses in other, non-target parasite species (Lello et al., 2004; Pedersen and Fenton, 2007; Knowles et al., 2013; Pedersen and Antonovics, 2013). Clearly it is essential to know the occurrence and direction of such interactions in order to predict disease dynamics and the likely impact of control efforts.

Given the evidence for parasite interactions in the laboratory, there is great interest in evaluating their occurrence in nature. As is well known in community ecology, experimental perturbation (e.g., measuring responses to the removal or addition of other species) is the most reliable way to detect natural interspecific interactions (Bender et al., 1984). Unfortunately, such experiments are rarely undertaken on parasite communities (but see Ferrari et al., 2009; Knowles et al., 2013; Pedersen and Antonovics, 2013). Hence, our knowledge of the natural occurrence and significance of parasite interactions is based primarily on observational studies, with various papers reporting clear evidence of strong interspecific parasite interactions, in both animal and human populations (Lello et al., 2004, 2013; Telfer et al., 2010; Shrestha et al., 2013). However, other studies have found little evidence for interactions in natural populations, concluding they are insignificant in shaping parasite communities (Haukisalmi and Henttonen, 1993; Poulin,

* Corresponding author. Tel.: +44 151 7954473; fax: +44 151 7954408.
E-mail address: A.fenton@liverpool.ac.uk (A. Fenton).

http://dx.doi.org/10.1016/j.ijpara.2014.03.001
0020-7519/

1996; Behnke et al., 2005; Behnke, 2008). There is therefore great variation among studies from natural populations, and a disconnection between these observational results and the consistent interactions reported from laboratory experiments.

One explanation for this variability is that parasite interactions are indeed highly variable and context-dependent. This would be an important result, telling us that parasite interactions are only significant under certain conditions (e.g., dependent on timings of coinfection, or infection burdens; Fenton, 2013) or within certain subsets of the host population (e.g., immuno-compromised hosts, or varying with sex or age etc.); if so, and we can identify the conditions or individuals in which interactions are strongest, this may improve our ability to predict the implications of those interactions and target treatment appropriately. A second explanation is that there is genuine variation in the importance of parasite interactions between different study systems, such as the types of parasite communities considered. However, assessing these possibilities is confounded by the fact that different studies often use different statistical approaches to infer interactions, and so it is not clear whether the reported differences are due to differences in biology of the systems or differences in the techniques used. Clearly, if different studies use different methods that themselves vary in reliability then we may be getting an inaccurate picture of the extent of parasite interactions in natural populations. Before we can fully evaluate the occurrence of these interactions, we need to establish the reliability of the various techniques used to infer their presence, ideally within a single study system.

Observational approaches have been suggested to have limited ability to infer interspecific interactions in general (Schluter, 1984). We have previously assessed this theoretically for parasite interactions (Fenton et al., 2010), showing that some commonly-used approaches are limited in their ability to detect genuine interactions. However, that analysis was purely theoretical and ignored many of the complexities of natural systems that could prevent, or even enhance, the performance of different statistical tests. There is a clear need to test the reliability of these various approaches on genuine parasite infection data from natural systems; to do so requires independent measures of the occurrence of interspecific parasite interactions within a given system, against which the different analytical approaches can be compared. We have previously carried out perturbation experiments using targeted drug treatments on two different natural rodent parasite communities, and have found clear evidence of interspecific parasite interactions in both systems (Knowles et al., 2013; Pedersen and Antonovics, 2013). These provide an ideal opportunity to test the inferences made using standard analytical techniques applied to observational (unmanipulated) data from the same populations. We show that many of the standard approaches are unable to detect the experimentally-demonstrated interactions, and often report associations in the opposite direction to those found experimentally. Overall, we urge caution for the interpretation of observational data when inferring the occurrence of interspecific interactions, suggesting it is only feasible for certain types of analysis applied to certain datasets, and highlight the importance of using perturbation approaches where possible to measure the strength and occurrence of parasite interactions in wild animal and human systems.

2. Materials and methods

2.1. Summary of interspecific interactions determined via experimental perturbations

We previously conducted experimental manipulations of the natural parasite communities of two small mammal species: wood mice, (*Apodemus sylvaticus*) in the UK (Knowles et al., 2013) and a mixed population of white footed mice (*Peromyscus leucopus*), and deer mice (*Peromyscus maniculatus*) in the USA (Pedersen and Antonovics, 2013). Specific details of each study are given in the relevant papers and information about the data structure, parasite diversity and infection prevalences are given in Table 1. Both studies adopted similar longitudinal designs, whereby permanent sampling grids were regularly trapped (fortnightly in the *Peromyscus* study or monthly in the *Apodemus* study). All individuals caught were given a unique identification tag and biometric data (size,

Table 1
Summary of observational data available for the *Peromyscus* and *Apodemus* analyses in this study.

	Peromyscus	*Apodemus*
Sample size (experimental data)	270 individuals (453 captures)	146 individuals (312 captures)
Sample size (observational data)	235 individuals (363 captures)	362 individuals (653 captures)
Mean nematode		
Prevalence	35.5%	57.4%
Abundance in EPG (range)	68.86 (0–9087)	41.15 (0–1023)
Intensity in EPG (range)	193.8 (3.23–9087)	72.04 (0.91–1023)
Dominant species (prevalence)	*Aspicularus americana* (15.4%) *Capillaria americana* (15.7%)	*Heligmosomoides polygyrus* (52%) *Syphacia stroma* (8.2%) *Aonchotheca murissylvatici* (1.4%) *Aspiculuris sp* (0.8%)
Mean Eimeria		
Prevalence	64.7%	49.0%
Abundance[a] in EPG (range)	1847 (0–61350)	2402 (0–181000)
Intensity[a] in EPG (range)	2853 (2.66–61350)	4918 (1.25–181000)
Dominant species (prevalence)	*Eimeria delicata* (9.1%) *Eimeria arizoniensis* A (57.0%) *Eimeria arizoniensis* B (30.9%)	*Eimeria hugaryensis* (27.6%) *Eimeria apionodes* (14.2%) *Eimeria uptoni* (2.4%)
Covariates (levels)[b]	Species (*leucopus*, *maniculatus*) Sex (Male, Female) Age (Adult, Sub-adult, Juvenile) Year (2001, 2002, 2003[c], 2004) Trap session (1, 2, 3, 4)	Grid (A[c], B[c], C[c], D[c], E[c], F) Sex (Male, Female) Age (Adult, Sub-adult, Juvenile) Year (2009, 2010[c]) Trap month (8 levels, May–Dec)

[a] Abundance refers to data including uninfected hosts; intensity refers to data from infected hosts only.
[b] All covariates were coded as factors.
[c] Indicates the year and/or grids from which the experimental data were taken.

weight, age, reproductive condition) were taken at each capture. Gastrointestinal infection status was assessed at each capture by faecal examination for the presence of infective stages (eggs for helminths or oocysts for coccidial protozoa).

In both studies the parasite communities were perturbed by treating a subset of animals with the anthelmintic drug ivermectin to reduce their gastrointestinal nematodes (>90% reduction in nematode prevalence over a period of 4 weeks in *Peromyscus*, and 71% reduction in prevalence within 3 weeks of treatment in *Apodemus*). Only a subset of animals was treated on the treatment grids in both studies; the remaining animals were left as untreated controls. In addition, in both studies there were untreated control grids (two grids in the *Peromyscus* study and two grids in the *Apodemus* study) on which no animals received treatment; there was no evidence that infections in untreated mice on treatment grids differed from those of animals from the untreated control grids. Comparing non-target parasite infections at subsequent captures between treated and untreated mice showed, in both studies, that coccidial parasites from the genus *Eimeria* (a genus of directly-transmitted protozoa that inhabit the gastrointestinal tract of small mammals) increased following anti-nematode treatment. Specifically, in *Peromyscus*, *Eimeria* showed a 20% increase in prevalence post- treatment (Fig. 1A) whereas in *Apodemus*, *Eimeria* increased 15-fold in intensity post-treatment (Fig. 1B). These classic perturbation experiments therefore provide clear evidence of negative interactions between nematodes and *Eimeria* in two separate host-parasite systems. Given this, we can then ask whether any of the standard approaches used for inferring interspecific parasite interactions from observational data would suggest the presence of these interactions.

2.2. Analytical approaches to infer interspecific interactions from observational data

Data were analysed from the untreated mice in the *Peromyscus* and *Apodemus* studies using five standard approaches (plus variants) that cover the broad range of techniques typically used in such analyses, examining either qualitative (presence/absence) or quantitative (parasite intensity (excluding uninfected individuals) or abundance (including uninfected individuals) data, based on eggs or oocysts per gram of faeces; EPG) measures of infection. Note that we were restricted to estimating abundance and intensity data indirectly using EPG, since we used non-destructive sampling to allow longitudinal analyses of each individual's infection status. For each approach we considered whether we would reasonably infer the negative interactions between nematodes and *Eimeria* revealed by our experimental results. In all cases model assumptions for the analyses (e.g., normality of residuals, homoscedasticity etc.) were checked and found to be upheld. Analyses were conducted in the statistical package R (v. 3.0.1).

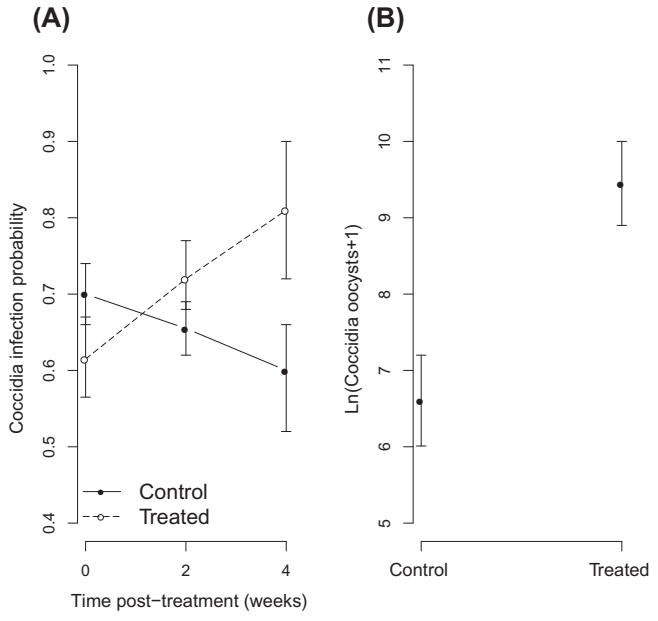

Fig. 1. Summary of experimental evidence for negative nematode-coccidia interactions in (A) the study by Pedersen and Antonovics (2013) (n = 270 individuals) and (B) the study by Knowles et al. (2013) (n = 146 individuals).

Four cross-sectional analyses were conducted on the observational data from both studies, asking: what is the association in contemporary infection levels between nematodes and *Eimeria* across the host population? Additionally, due to the number of individual recaptures in the *Apodemus* dataset, an additional longitudinal analysis was performed on these data, asking: what is the association between nematode infection one month and *Eimeria* infection the following month? Here we briefly describe each approach, leaving detailed descriptions for Supplementary Data S1.

2.2.1. Correlation approach (cross sectional)

This approach seeks correlations (Pearson's R, Spearman Rank or Kendall's tau) in parasite intensity (EPG in infected hosts) of the two parasite species. If this approach is reliable for inferring interactions we would expect a negative correlation between nematodes and *Eimeria*, to match the experimental results. Two forms of analysis were explored.

2.2.1.1. Analysis of raw data. A simple correlation of logged nematode and *Eimeria* intensity data (EPG counts among coinfected hosts).

2.2.1.2. Analysis of residuals controlling for potential confounding factors. Confounding variables (e.g., age, sex, sampling location) may create spurious associations between parasites. One way to control for these effects has been to conduct two ANOVA (or equivalent) analyses, one with each parasite species as the response variable, on parasite intensity data with potential confounders (see Table 1 for lists of covariates for each study) as explanatory variables (e.g., Behnke et al., 2005). A significant correlation between the residuals from each analysis is then used as evidence of an interspecific interaction independent of the confounding factors. It should be noted that using residuals in this way can result in biased parameter estimates and has been discouraged (Freckleton, 2002); an alternative approach is to conduct a Generalised Linear Model (GLM) which directly controls for covariates in the analysis (e.g., Analysis 4, below). This method now tends not to be used, but has been used previously, and was included here for completeness. This analysis was run on nematode and *Eimeria* intensity data (EPG counts among coinfected hosts).

2.2.2. T-test comparison (cross-sectional)

Here, an interaction is inferred from a significant difference in infection levels of one parasite between hosts infected and not infected by the other. Unpaired two sample Student's t-tests on logged *Eimeria* EPG from *Eimeria*-infected hosts (i.e., using *Eimeria* intensity data) were used to compare nematode-infected with -uninfected hosts. Based on our experimental results, we would expect nematode infected hosts to have significantly lower *Eimeria* EPG counts than hosts without nematode infections. Again, this analysis does not account for potential confounding covariates, and an alternative approach using GLMs to control for covariates is conducted later (Analysis 4). However, as with the correlation approach, it has been used previously (Chappell, 1969;

Hendrickson and Curtis, 2002), and was included here for completeness.

2.2.3. Pairwise association matrices (cross-sectional)

This approach compares the observed numbers of single- and coinfected hosts with those expected from a null model, based on the observed prevalences of the two parasite species; a significant departure (by χ^2) from the null model implies the parasites are associated independently from each other. Given our experimental results, we would predict coinfected hosts should occur less often than expected by chance, as nematodes suppress *Eimeria*, thereby reducing the number of coinfected hosts. Two suggested methods were explored for the construction of the null model.

2.2.3.1. Basic null model. Here the expected proportion of coinfected hosts is calculated from the prevalences of the two parasite species, assuming independent assortment. Hence, if parasite species i and j have decimal prevalences p_i and p_j, the expected proportion of hosts carrying both species is $p_i.p_j$, and the expected proportion carrying neither is $(1\text{-}p_i)(1\text{-}p_j)$, etc.

2.2.3.2. Modified null model accounting for species dominance. The above null model ignores the potential role of interactions in shaping the observed prevalences. Lafferty et al. (1994) suggested a method to alleviate this circularity by adjusting the prevalences to account for the potential influence of the 'dominant' species over the other (see Supplementary Data S1), which are then used to calculate the expected proportions of singly- and coinfected hosts as above.

2.2.4. Cross-sectional GLM

This approach uses GLM to assess whether infections by one parasite are influenced by infections of the other at the same time point, whilst controlling for covariates (Table 1). For the *Apodemus* study the following model was used:

$$\textit{Eimeria} \sim \text{Nematodes} + \text{Age} + \text{Sex} + \text{Trap Month} + \text{Year} + \text{Grid}$$

and for *Peromyscus* it was:

$$\textit{Eimeria} \sim \text{Nematodes} + \text{Host species} + \text{Age} + \text{Sex} + \text{Trap Session} + \text{Year}$$

In each case the infection status by both *Eimeria* and nematodes could either be qualitative (presence/absence) or quantitative (EPG), resulting in four possible combinations of variable types. When the response variable (*Eimeria*) was qualitative a binomial GLM was used, and when quantitative (log(*Eimeria* EPG), restricted to *Eimeria*-infected individuals) a Gaussian GLM was used. These models were simplified by backwards stepwise deletion (using the function '*step*' in R), until a minimal model was reached. For these results to match our experimental results, we would expect a significant negative relationship between nematodes and *Eimeria*.

Note that a further version of this analysis was explored, controlling for potential non-linear effects of host age, using body length and body mass as proxies (Fenton et al., 2010; Supplementary Data S1 but found it did not significantly change model log-likelihood. We therefore only present the results from the standard GLMs.

2.2.5. Longitudinal GLM

All analyses considered so far have been cross-sectional, examining the contemporary associations between *Eimeria* and nematodes. For the *Apodemus* study an additional, longitudinal analysis was carried out, using the same four baseline models as the cross-sectional GLMs above, but here the 'Nematodes' explanatory variable referred to infection status the previous month. Once again, model simplification was used to reach a minimal model, and we sought evidence that nematode infection one month reduced *Eimeria* infection the next.

2.3. Controlling for pseudoreplication arising from multiple captures of individuals

The full data included multiple captures of some individuals, which are not independent from each other (i.e., pseudoreplication at the level of the individual). In Supplementary Data S1 we describe a range of approaches we explored to control for this pseudoreplication. However, in all cases the results (the terms remaining in the minimal models, and effect sizes of those terms) were very similar for all three methods (Supplementary Data S1; Supplementary Fig. S1), presumably due to the relatively low numbers of recaptures in the data. We therefore concentrate on the results from the full datasets here.

2.4. Assessing the reliability of each analytical technique

The reliability in inferring the experimentally-revealed negative interaction between nematodes and *Eimeria* was assessed for each of the above approaches. Because our observational data may not be optimal for inferring interactions using a given technique (e.g., the sample size may be too small, or of insufficient temporal resolution), a broad approach was taken to assess reliability. First, we used a simple qualitative assessment, asking whether each technique predicted the correct direction (negative) of association between nematodes and *Eimeria*. We then used a quantitative assessment of the statistical significance ($P < 0.05$) of association between nematodes and *Eimeria*. Finally, we sought a quantitative measure of the magnitude of reported effects of nematodes on *Eimeria*. Since the various approaches return different statistical metrics, these metrics were converted to a common, standardised effect size, Hedge's g (Borenstein, 1994; Nakagawa and Cuthill, 2007). This allows effect sizes from different tests to be presented on the same scale, aiding comparison with the experimental results; a negative value of Hedge's g in these analyses implies a negative association between nematodes and *Eimeria*, matching the experimental results.

Note that, to maximise sample size for these analyses data were used from a wider range of years and study grids than were used in the experiments (Table 1). To check whether this explains any discrepancies between these analyses and the experimental results, we re-ran our analyses of the *Apodemus* data restricted to the same year as the experiment was conducted, and found the results were little affected (Supplementary Data S1; Supplementary Fig. S2). Finally, we emphasise that our results are only directly applicable to the host-parasite systems examined and the quality and resolution of data available to us (we return to this point in the Discussion). However we suggest that many of our conclusions are likely to be applicable to many other empirical systems where similarly structured data are used to infer the existence of interspecific parasite interactions.

3. Results

Here we summarise the reliability of each technique in comparison to our experimental results, leaving more detailed descriptions of each analysis in Supplementary Data S1.

Overall, there was considerable variation between the different approaches in the predicted association between nematodes and *Eimeria* (Fig. 2; Table 2), with relatively few tests matching the experimental results by returning negative associations (5/17 tests

for *Apodemus* (two significant at $P < 0.05$), and 0/13 tests for *Peromyscus*; Fig. 2). Indeed, the majority of tests returned positive associations between nematodes and *Eimeria* (12/17 for *Apodemus* and 13/13 for *Peromyscus*), the opposite direction of that seen with experimental manipulation.

The least reliable techniques were the correlation-based ones, where all variations reported positive associations between nematodes and *Eimeria* (Fig. 2 and Table 2). Furthermore, even the cross-sectional GLMs, which controlled as much as possible for potential confounders, fared poorly; one variant (where both nematode and *Eimeria* infections were analysed as EPG) resulted in the strongest positive effect size out of all tests for both *Apodemus* and *Peromyscus* (Fig. 2), and the other cross-sectional GLMs returned effect sizes around zero. This suggests that adding covariates into the analysis does not necessarily improve model accuracy. For example, the cross-sectional '*Eim*(EPG)∼Nematode(PA)' GLM is closely related to the *t*-test analysis (both have *Eimeria* EPG as the response variable and nematode presence/absence as the predictor), except that the GLM controls for covariates, whereas the *t*-test does not. However there was no evidence that the GLM performed any better (predicted effect sizes were not stronger, and confidence intervals were not narrower) than the *t*-test. Similarly, the cross-sectional '*Eim*(EPG)∼Nematode(EPG)' GLM is related to the standard correlations in the nature of the response and predictor variables, but there was no evidence that the GLM, which controls for covariates, performed any better than the correlation approach (Fig. 2).

Overall the most reliable methods tended to be longitudinally-based, which examined the association between nematode infections one month and *Eimeria* infections the following month; three out of four of these analyses predicted a negative association between nematodes and *Eimeria* and two were statistically significant (Fig. 2). However, the form of analysis that most closely matches the experimental result for *Apodemus*, in terms of the nature of the response and predictor variables (*Eimeria* EPG and nematode presence/absence), although predicting a negative effect size, had a wide confidence interval and was not statistically significant (Fig. 2 and Table 2).

4. Discussion

Few of the observation-based statistical approaches tested were successful at inferring the experimentally-revealed negative interaction between nematodes and *Eimeria*. In particular most cross-sectional approaches, particularly the correlation-based ones, performed extremely poorly, often returning highly significant but strongly positive associations between the parasites. This was particularly apparent for the *Peromyscus* dataset (Fig. 2), which had a lower sample size than the *Apodemus* dataset. These results match, and extend, our previous theoretical analyses which showed that correlation-based approaches can perform very poorly when attempting to detect negative interactions between parasites (Fenton et al., 2010). While we may expect detection of

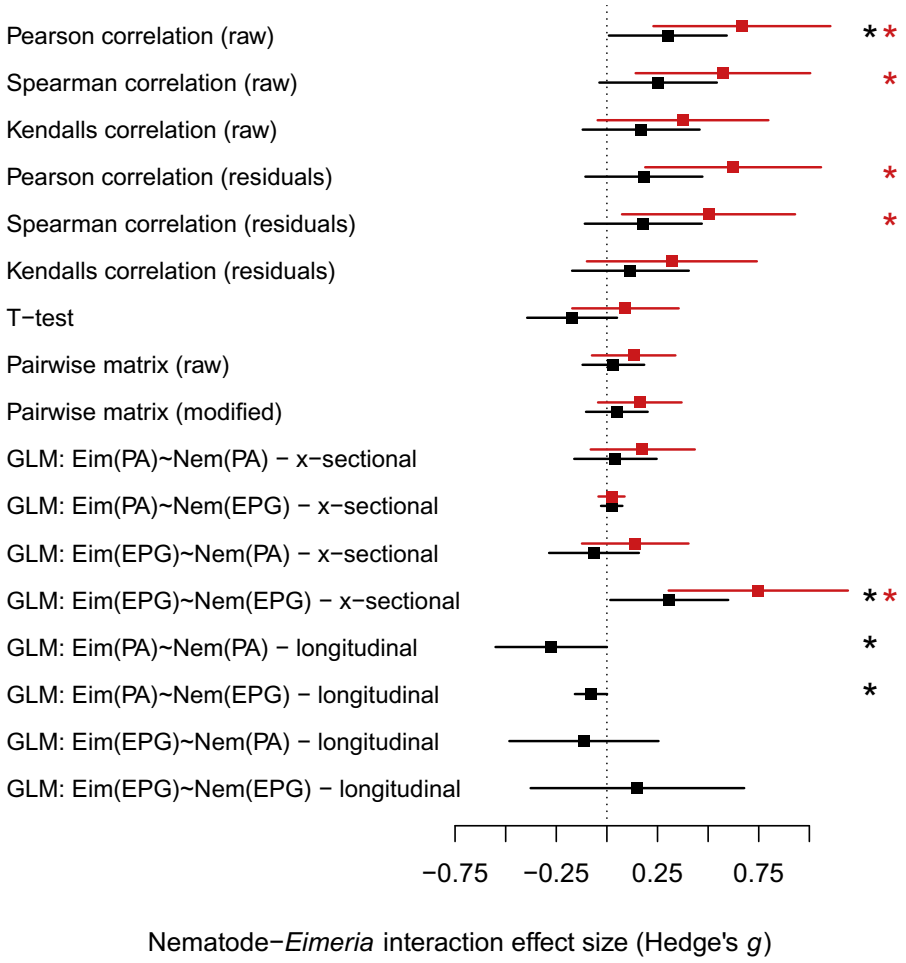

Fig. 2. Mean (±95% Confidence Intervals) effect sizes (Hedge's *g*) of the relationship between nematode and *Eimeria* infections for each analytical approach examined for the *Apodemus sylvaticus* (black) and *Peromyscus* spp. (red; grey) studies, using all data (including multiple captures per individual; corresponding results for the bootstrapped data are given in Supplementary Fig. S1). Asterisks indicate approaches where the relevant 95% Confidence Interval of effect size does not overlap with zero. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

genuine interactions in real-world data to be difficult, it is highly concerning that such frequently-used techniques can lead to the inference of significant associations in the opposite direction to the genuine interaction. Therefore, we strongly advise against using correlation-based approaches, even those that attempt to control for confounding factors, when seeking interspecific interactions from ecological data (parasitological or otherwise).

The most reliable approaches examined were longitudinally-based, which sought associations between nematodes one month and *Eimeria* the next. Intuitively this makes sense, as it reflects the cause and effect of the underlying interaction; *Eimeria* levels will decline following nematode infection. As such, we advocate the use of longitudinal methods where possible in the inference of interspecific interactions. However, this should be tempered with the recognition that, for our *Apodemus* data set, the longitudinal analysis that most closely matched the experimental data in terms of the nature of the response and predictor variables ['*Eim*(EPG)∼Nematode(PA)'], although matching the direction of the experimental results, did not predict a statistically significant association. Furthermore, the merits of longitudinal analyses will depend greatly on the time-scale of sampling relative to the interaction dynamics of the parasites. If the parasites interact very strongly, with one species responding rapidly to the other, then fine-scale sampling will be needed to detect the interaction. Notably, for the *Apodemus* experimental data, the interaction could only be detected within 1–3 weeks of treatment; by the fourth week nematodes and *Eimeria* had returned to pre-treatment levels (Knowles et al., 2013). Hence, fine-scale sampling would be needed to maximise the chances of detecting the interaction from observational data under such rapid dynamics. In addition, sample size will be an important consideration in the viability of longitudinal analyses. Restrictions in the number of recaptures per individual, as seen in the *Peromyscus* dataset, will severely limit the ability to perform these analyses for many systems. Even with the *Apodemus* data set, with the larger overall sample size, the requirement for data on the same individual caught over successive months greatly restricted the power of the analyses (the *Apodemus* dataset was reduced from 653 captures to 254 for the longitudinal analyses). It is notable that one study that found a particularly dense network of associations among coinfecting parasites (Telfer et al., 2010) used a longitudinal approach on a very large dataset (14,075 captures of 5,981 animals), which enabled the authors to look at how transitions in infection status (switching from uninfected to infected between captures) related to their coinfection status. Such approaches and data are likely to be beyond the scope of many studies of natural parasite communities (including those of humans), therefore great caution should be taken either when applying longitudinal approaches to more restricted datasets, or when having to resort to less desirable cross-sectional approaches.

Why then did so many of the other approaches perform so badly? There are several, not mutually exclusive, explanations. One possibility is that the methods may be reliable given the right data, but the datasets used here are lacking in terms of sample size, resolution (frequency of sampling) or type of data available (egg count data, which are an indirect and potentially unreliable proxy for parasite abundance). For these reasons we used liberal assessments of reliability, tending to base our conclusions on directions of effects, rather than strict statistical significance. Hence, if a given approach is reasonable, but our sample size was inadequate, we may expect a predicted effect in the correct direction even if it was not statistically significant. However, our results do not suggest a mere lack of statistical power, as the associations we found were not necessarily small or insignificant but were, in the majority of cases, in the opposite direction to the experimentally-observed interactions. It is certainly true that two of the most prominent observational studies reporting strong associations among coinfecting parasites (Lello et al., 2004; Telfer et al., 2010) had particularly large sample sizes. However, the sample sizes of our datasets were not particularly different from those used in many observational studies of parasite communities, and so it seems reasonable to suggest that if our datasets were inadequate for these statistical methods then the same may apply to other studies. An alternative explanation is that observational studies do not adequately control for confounding factors that either obscure genuine interactions, or generate spurious associations. In our GLM analyses we attempted to control for such effects as much as possible (e.g., host age, sex, sampling location, time-point etc.) but these analyses did not necessarily perform better than the equivalent analyses that did not control for covariates (e.g., correlations or *t*-tests). Clearly there may be other important factors that we did not account for (e.g., exposure, host genetic resistance/susceptibility, local spatial heterogeneity etc.), but the factors we controlled for are consistent with those used in many other studies. Thus these approaches may not be expected to be any more reliable for other, similar studies. Finally, it is possible that interactions among parasites are non-linear, meaning that linear statistical models, as are commonly used for inferring the existence of parasite interactions (and as were used here) are not adequate to detect the true relationships between parasites. In particular, if interactions are strongest at low infection intensities (i.e., between nematode-free hosts and those with light nematode infections) then the typically higher burdens seen in untreated individuals may reduce the ability of observational approaches to detect those interactions. To assess this possibility we re-ran our cross-sectional and longitudinal GLM analyses with a quadratic term for nematode EPG as the predictor variable (for both *Eimeria* EPG and presence/absence as the response variables). However in no cases did the quadratic term stay in the final model, suggesting there was no detectable non-linearity in the nematode-*Eimeria* interaction that could have caused the differences between our observational and experimental results.

An alternative explanation for the mismatch between experimental and observational analyses is that it is the experimental results are incorrect, while the observational results reflect the true interaction. For example, the administered drug (ivermectin, a broad-spectrum nematocidal drug) may directly affect *Eimeria*, generating the apparent interactions we saw experimentally. However, this seems highly unlikely, as ivermectin is one of the most widely used anthelmintic drugs for both medical and veterinary usage and has been tested multiple times and in a diverse array of systems, yet we have not found any reported direct effects on coccidia. Furthermore, if ivermectin did directly affect *Eimeria* it would have to have a positive effect (increase *Eimeria* infection status) to create the post-treatment effects that we found in both datasets. Ivermectin targets the nervous system, and it is hard to envisage how that mode of action would directly lead to an increase in the abundance of *Eimeria*. Alternatively, ivermectin may be affecting a nematode that is not detected in the observational samples (i.e., that does not pass eggs in the host's faeces), but that is affecting *Eimeria*. However, there is evidence that the negative interaction between nematodes and *Eimeria* in *Apodemus* occurs through a highly localised interaction between two species that are regularly passed in faeces: the nematode *Heligmosomoides polygyrus* (the most common nematode in this population), and a common species of *Eimeria* (*Eimeria hungaryensis*) that inhabits the same section of the gut as *H. polygyrus*; another *Eimeria* species found in lower parts of the gut (where *H. polygyrus* is absent) was unaffected by anti-nematode treatment (Knowles et al., 2013). It would be hard to explain these species-specific effects if the experimental results were mediated by an undetected nematode. It therefore seems more parsimonious to suggest that the nematodes we detect do indeed negatively affect *Eimeria*, but that many of the

Table 2
Summary of results in this study. *P*-values refer to associations between nematodes and *Eimeria*. 'All data' includes multiple captures per individual. Bootstrapped results from 100 random subsamples, each including one capture per individual, presenting the percentage of runs (Sig%) returning a significant association between nematodes and *Eimeria*. Cell shadings show analyses that may imply an association between nematodes and *Eimeria* (based on $P < \sim 0.05$ for all data, or >50% of runs were significant for bootstrapped data); blue, positive association; yellow, negative association.

Test	Variant	*Peromyscus*		*Apodemus*	
		All data	Bootstrapped	All data	Bootstrapped
1. Correlation [a]	Raw data	*P* = 0.002 *r* = 0.319	Sig% = 93% (all +ve)	*P* = 0.042 *r* = 0.149	Sig% = 46% (all +ve)
	Residuals	*P* = 0.004 *r* = 0.300	Sig% = 83% (all +ve)	*P* = 0.216 *r* = 0.091	Sig% = 20% (all +ve)
2. T-test		*P* = 0.494	Sig% = 0%	*P* = 0.128	Sig% = 13% (all -ve)
3. Pairwise	Basic null model	*P* = 0.207	Sig% = 2% (all +ve)	*P* = 0.681	Sig% = 0%
	Modified null model	*P* = 0.121	Sig% = 14% (all +ve)	*P* = 0.526	Sig% = 11% (all +ve)
4. Cross-sectional GLM	*Eim*(PA) ~ Nem(PA)	*P* = 0.195 [b]	Sig% = 2% (all +ve)	*P* = 0.684	Sig% = 1% (all +ve)
	Eim(PA) ~ Nem(EPG)	*P* = 0.487	Sig% = 0%	*P* = 0.375	Sig% = 9% (all +ve)
	Eim(EPG) ~ Nem(PA)	*P* = 0.286	Sig% = 0%	*P* = 0.563	Sig% = 2% (all -ve)
	Eim(EPG) ~ Nem(EPG)	*P* = 0.0007 OR= 1.684 [c] (1.263 - 2.273)	Sig% = 93% (all +ve)	*P* = 0.028 OR = 1.341 (1.019 - 1.765)	Sig% = 12% (all +ve)
5. Longitudinal GLM	*Eim*(PA) ~ Nem(PA)			*P* = 0.050 OR = 0.607 (0.367 - 0.999)	Sig% = 34% (all -ve)
	Eim(PA) ~ Nem(EPG)	Not undertaken for *Peromyscus*		*P* = 0.053 OR = 0.867 (0.748 - 1.000)	Sig% = 35% (all -ve)
	Eim(EPG) ~ Nem(PA)			*P* = 0.518	Sig% = 0%
	Eim(EPG) ~ Nem(EPG)			*P* = 0.510	Sig% = 1% (all +ve)

+ve, positive; −ve, negative.
[a]Pearson's correlations. *r* is the correlation coefficient. PA, the parasite is coded as present/absent (categorical variable); EPG, eggs per gram (continuous variable); *Eim, Eimeria* (the response variable); and Nem, nematodes (the predictor).
[b]*P*, value at which 'nematodes' drops out or is retained in the final model.
[c]OR, Odds Ratio (and 95% Confidence Intervals) for the effect of nematodes on *Eimeria* from models where 'nematodes' is retained.

observational analyses are unable to reliably detect this interaction.

Above we asked why many of the observational approaches failed to detect the underlying interactions. However it also seems reasonable to ask the reverse question: why would we expect them to be able to detect a true interaction? Real-world parasitological (or indeed any ecological) data are typically highly noisy with multiple sources of variation and confounding factors, generated by highly non-linear processes acting across multiple levels of biological organisation and spatial and temporal scales. It therefore seems unlikely that many statistical approaches will fare well when confronted with purely observational data, particularly if datasets are of limited size or resolution. A previous individual-based model showed that even with model-generated data, where the generating processes and confounding variables were completely known and could be statistically controlled for, most techniques failed to detect the underlying parasite interactions (Fenton et al., 2010). Furthermore, mathematical modelling shows that there may be highly non-linear, and even non-monotonic, relationships between individual-level helminth burdens and the population-level transmission of coinfecting pathogens; just by changing worm burden there may be a switch from a positive relationship between worms and coinfecting pathogens at the population level to a negative one, for the same system with the same underlying interaction between them (Fenton, 2008, 2013). Hence, samples from the same system taken from different times or different locations may suggest opposite relationships between parasites, simply by sampling at different points along the same curve (Fenton, 2013). As such, not only may the occurrence of interspecific interactions be highly context dependent, but so may our ability to detect them. We suggest that while having more data will generally be better than having less, it is equally important to have data that span the full range of possible infection levels, appropriate control for all sources of heterogeneity (spatial, temporal and individual level) in the data and account for potential non-linearities in their effects (e.g., non-linear relationships between worm burden, host age and interaction strength; Fenton et al., 2010; Lello et al., 2013). Furthermore, we strongly recommend using longitudinal approaches where possible, even though that requires non-destructive sampling and may involve indirect, and possibly unreliable, estimates of infection status (e.g., faecal egg counts as proxies of helminth burdens). It is possible that

destructive sampling, although constraining analyses to be cross-sectional, may prove more reliable than the results presented here, due to the ability to make direct measures of infection status (e.g., Behnke et al., 2005); however, we were not able to assess that here due to a lack of such data for our systems.

We emphasise that our results and conclusions relate specifically to the characteristics of our datasets and systems. As noted above, for example, the extent to which longitudinal analyses out-perform cross-sectional analyses will depend on having an appropriate time scale of sampling relative to the dynamics of the parasites. In addition, whether qualitative (presence/absence) or quantitative (EPG) analyses are preferable will depend on the system-specific accuracy of the assays for quantifying infection levels. Infection by microparasites (viruses, bacteria etc.) is often detected either directly by PCR or indirectly by immunity-based assays for antibodies or antigen (e.g. ELISA, immunofluorescence-based assays), whereas macroparasite (e.g., helminth) detection by non-destructive sampling is most often by faecal egg counts from microscopy. The reliability of the different qualitative and quantitative measures of infection status will vary between systems and so will the relative benefits of using qualitative or quantitative analyses.

Overall we have shown considerable variability among observation-based approaches in their ability to infer interspecific parasite interactions from natural parasite data. Based on these results, and previous theory (Fenton et al., 2010), we would strongly discourage the use of cross-sectional approaches, particularly correlation-based methods, for inferring interspecific interactions. If possible, we would recommend longitudinal approaches, although the size of dataset needed, the requirement for non-destructive sampling and the required frequency of sampling may restrict their applicability to only certain systems (e.g., Telfer et al., 2010). Ultimately, as is well known in free-living community ecology, we would suggest that experimental approaches, if possible, are the most direct way of detecting genuine interspecific interactions. In terms of human infectious diseases we see great, but currently under-exploited, potential for detecting parasite interactions following drug treatment programmes. There are many programmes underway in human communities around the globe that use specific drugs to target narrow groups of parasites, or broad spectrum drugs that differ in their efficacy against different parasite species (Basáñez et al., 2012), but often little attempt is made to follow non-target parasites (but see Blackwell et al., 2013). Often, these programmes show great variability in their benefits to host health (Taylor-Robinson et al., 2012), and one possibility is that coinfecting, non-target, parasites may be responding to suppression of the target parasites (Fenton, 2013). Such treatment programmes provide ideal perturbations (exactly like those used in our *Apodemus* and *Peromyscus* studies) which, if appropriate data can be collected, could provide great insight into how the targeted parasites interact with other members of the parasite community, and may help the design of more effective and sustainable treatment programmes.

Acknowledgements

We are very grateful to everyone who helped with fieldwork and microscopy work, and the relevant land owners, for the work carried out for the two field studies (see Acknowledgements in Pedersen and Antonovics (2013) and Knowles et al. (2013) for details). This work was funded by a Natural Environment Research Council (U.K.) grant to ABP, AF and OP (NE/G006830/1 and NE/G007349/1). ABP was supported by grants from several U.S. funding sources (NSF-DIG, Sigma Xi, Mountain Lake Biological Station, the American Society of Mammalogists), an Advanced Fellowship (Wellcome Trust U.K. Strategic grant to the CIIE, 095831) and a University of Edinburgh U.K., Chancellors Fellowship. We also thank all the technicians and field assistants who helped in the collection of data for these experiments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijpara. 2014.03.001.

References

Adams, D.B., Anderson, B.H., Windon, R.G., 1989. Cross-immunity between *Haemonchus contortus* and *Trichostrongylus colubriformis* in sheep. Int. J. Parasitol. 19, 717–722.
Basáñez, M.-G., French, M.D., Walker, M., Churcher, T.S., 2012. Paradigm lost: how parasite control may alter pattern and process in human helminthiases. Trends Parasitol. 28, 161–171.
Behnke, J.M., 2008. Structure in parasite component communities in wild rodents: predictability, stability, associations and interactions .... or pure randomness? Parasitology 135, 751–766.
Behnke, J.M., Gilbert, F.S., Abu-Madi, M.A., Lewis, J.W., 2005. Do the helminth parasites of wood mice interact? J. Anim. Ecol. 74, 982–993.
Behnke, J.M., Wakelin, D., Wilson, M.M., 1978. *Trichinella spiralis*: delayed rejection in mice concurrently infected with *Nematospiroide dubius*. Exp. Parasitol. 46, 121–130.
Bender, E.A., Case, T.J., Gilpin, M.E., 1984. Perturbation experiments in community ecology: theory and practice. Ecology 65, 1–13.
Blackwell, A.D., Martin, M., Kaplan, H., Gurven, M., 2013. Antagonism between two intestinal parasites in humans: the importance of co-infection for infection risk and recovery dynamics. Proc. R. Soc. Lond. B Biol. Sci. 280, 20131671-20131671.
Borenstein, M., 1994. Effect sizes for continuous data. In: Cooper, H.M., Hedges, L.V. (Eds.), The handbook of research synthesis, Russell Sage Foundation, pp. 221–235.
Chappell, L.H., 1969. Competitive exclusion between two intestinal parasites of the three-spined stickleback, *Gasterosteus aculeatus* L. J. Parasitol. 55, 775-&.
Christensen, N.O., Nansen, P., Fagbemi, B.O., Monrad, J., 1987. Heterologous antagonistic and synergistic interactions between helminths and between helminths and protozoans in concurrent experimental infection of mammalian hosts. Parasitol. Res. 73, 387–410.
Cox, F.E.G., 2001. Concomitant infections, parasites and immune responses. Parasitology 122, S23–S38.
Fenton, A., 2008. Worms and germs: the population dynamic consequences of microparasite-macroparasite co-infection. Parasitology 135, 1545–1560.
Fenton, A., 2013. Dances with worms: the ecological and evolutionary impacts of deworming on coinfecting pathogens. Parasitology 140, 1119–1132.
Fenton, A., Viney, M.E., Lello, J., 2010. Detecting interspecific macroparasite interactions from ecological data: patterns and process. Ecol. Lett. 13, 606–615.
Ferrari, N., Cattadori, I.M., Rizzoli, A., Hudson, P.J., 2009. *Heligmosomoides polygyrus* reduces infestation of *Ixodes ricinus* in free-living yellow-necked mice, *Apodemus flavicollis*. Parasitology 136, 305–316.
Freckleton, R.P., 2002. On the misuse of residuals in ecology: regression of residuals vs. multiple regression. J. Anim. Ecol. 71, 542–545.
Frontera, E., Alcaide, A., Dominguez-Alpizar, J.L., Boes, J., Reina, D., Navarrete, I., 2005. Evidence of interaction between *Ascaris suum* and *Metastrongylus apri* in experimentally infected pigs. Vet. Parasitol. 127, 295–301.
Griffiths, E.C., Pedersen, A.B., Fenton, A., Petchey, O.L., 2011. The nature and consequences of coinfection in humans. J. Infect. 63, 200–206.
Haukisalmi, V., Henttonen, H., 1993. Coexistence in helminths of the bank vole *Clethrionomys glareolus*.1. Patterns of co-occurrence. J. Anim. Ecol. 62, 221–229.
Hendrickson, M.A., Curtis, L.A., 2002. Infrapopulation sizes of co-occurring trematodes in the snail *Ilyanassa obsoleta*. J. Parasitol. 88, 884–889.
Knowles, S.C.L., Fenton, A., Petchey, O.L., Jones, T.R., Barber, R., Pedersen, A.B., 2013. Stability of within-host – parasite communities in a wild mammal system. Proc. R. Soc. Lond. B Biol. Sci. 280, 20130598.
Lafferty, K.D., Sammond, D.T., Kuris, A.M., 1994. Analysis of larval trematode communities. Ecology 75, 2275–2285.
Lello, J., Boag, B., Fenton, A., Stevenson, I.R., Hudson, P.J., 2004. Competition and mutualism among the gut helminths of a mammalian host. Nature 428, 840–844.
Lello, J., Knopp, S., Mohammed, K.A., Khamis, I.S., Utzinger, J., Viney, M.E., 2013. The relative contribution of co-infection to focal infection risk in children. Proc. R. Soc. Lond. B: Biol. Sci., 20122813.
Nakagawa, S., Cuthill, I.C., 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol. Rev. 82, 591–605.
Pedersen, A.B., Antonovics, J., 2013. Anthelmintic treatment alters the parasite community in a wild mouse host. Biol. Lett. 9, 20130205.
Pedersen, A.B., Fenton, A., 2007. Emphasising the ecology in parasite community ecology. Trends Ecol. Evol. 22, 133–139.
Petney, T.N., Andrews, R.H., 1998. Multiparasite communities in animals and humans: frequency, structure and pathogenic significance. Int. J. Parasitol. 28, 377–393.

Poulin, R., 1996. Richness, nestedness, and randomness in parasite infracommunity structure. Oecologia 105, 545–551.
Schluter, D., 1984. A variance test for detecting species associations, with some example applications. Ecology 65, 998–1005.
Shrestha, S., Foxman, B., Weinberger, D.M., Steiner, C., Viboud, C., Rohani, P., 2013. Identifying the interaction between influenza and pneumococcal pneumonia using incidence data. Sci. Trans. Med. 5, 191ra184.
Taylor-Robinson, D.C., Maayan, N., Soares-Weiser, K., Donegan, S., Garner, P., 2012. Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin and school performance. Cochrane Database of Systematic Reviews 7.
Telfer, S., Lambin, X., Birtles, R., Beldomenico, P., Burthe, S., Paterson, S., Begon, M., 2010. Species interactions in a parasite community drive infection risk in a wildlife population. Science 330, 243–246.