Epidemics 11 (2015) 48–55

Epidemics

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

Unknown age in health disorders: A method to account for its cumulative effect and an application to feline viruses interactions

Eléonore Hellard *, Dominique Pontier, Aurélie Siberchicot, Frank Sauvage, David Fouchet

Université de Lyon, Université Lyon1, CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, 43 Bd du 11 novembre 1918, F-69622, Villeurbanne, France

ARTICLE INFO

Article history:
Received 25 February 2014
Received in revised form 12 February 2015
Accepted 12 February 2015
Available online 24 February 2015

Keywords:
Felis silvestris catus
Multiple infections
Parametric boostrap
SI model
Serology

ABSTRACT

Parasite interactions have been widely evidenced experimentally but field studies remain rare. Such studies are essential to detect interactions of interest and access (co)infection probabilities but face methodological obstacles. Confounding factors can create statistical associations, i.e. false parasite interactions. Among them, host age is a crucial covariate. It influences host exposition and susceptibility to many infections, and has a mechanical effect, older individuals being more at risk because of a longer exposure time. However, age is difficult to estimate in natural populations. Hence, one should be able to deal at least with its cumulative effect. Using a SI type dynamic model, we showed that the cumulative effect of age can generate false interactions theoretically (deterministic modeling) and with a real dataset of feline viruses (stochastic modeling). The risk to wrongly conclude to an association was maximal when parasites induced long-lasting antibodies and had similar forces of infection. We then proposed a method to correct for this effect (and for other potentially confounding shared risk factors) and made it available in a new R package, *Interatrix*. We also applied the correction to the feline viruses. It offers a way to account for an often neglected confounding factor and should help identifying parasite interactions in the field, a necessary step towards a better understanding of their mechanisms and consequences.

1. Introduction

Epidemiological, clinical, or biological outcomes are often studied considering each parasite separately. Nevertheless, parasites rarely exist in isolation. Hosts are exposed to numerous parasites (from micro- to macro-parasites, pathogenic or not) simultaneously and multiple infections of hosts are more frequently encountered than infections by a single parasite (Petney and Andrews, 1998; Cox, 2001).

Interactions between members of the intra-host community of parasites, e.g., competition for resources or interactions mediated by host immune responses, have been widely evidenced in animal models and experimental conditions (e.g., (Cox, 2001; Behnke et al., 1978; Christensen et al., 1987; Frontera et al., 2005)), with a main focus on human pathogens such as HIV, tuberculosis, malaria, sexually transmitted infections, and helminths (e.g., (Bentwich et al., 1999; Corbett et al., 2002; Celum, 2004; Druilhe et al., 2005; Abu-Raddad et al., 2006)). There is now strong evidence that parasites are affected by the presence of other parasites, their interactions altering the rates of co-occurrence, levels of infection and disease severity (e.g., (Abu-Raddad et al., 2006; Read and Taylor, 2001; Weiss and McMichael, 2004; Telfer et al., 2010; Ives et al., 2011)), as well as the success of parasites management measures (Pedersen and Fenton, 2007; Harris et al., 2009; Koch and Schmid-Hempel, 2011) and possibly disease (re)emergence (Pontier et al., 2009; Keesing et al., 2010).

Numerous studies on parasite interactions have been led in experimental conditions but are much rarer in natural host populations, with some exceptions for macroparasites (e.g., helminthes) communities (e.g., (Telfer et al., 2010; Dezfuli et al., 2001; Lello et al., 2004; Behnke et al., 2005; Jolles et al., 2008)). Field studies are however essential because experimental systems are oversimplified and because lab studies require an existing suspicion of interaction between the parasites. In addition, only studies in natural populations can give access to infection and co-infection probabilities. In other words, before studying their mechanisms in the lab, interactions of interest must be identified in the field.

Main difficulties encountered with studies in natural populations are methodological. Numerous confounding factors can create

* Corresponding author. Present address: DST/NRF Centre of Excellence, Percy FitzPatrick Institute of African Ornithology, University of Cape Town, Private Bag X3, Rondebosch, South Africa. Tel.: +27 0 21 650 4008; fax: +27 0 21 650 3295.
E-mail addresses: eleonore.hellard@gmail.com (E. Hellard), dominique.pontier@univ-lyon1.fr (D. Pontier), aurelie.siberchicot@univ-lyon1.fr (A. Siberchicot), frank.sauvage@univ-lyon1.fr (F. Sauvage), david.fouchet@univ-lyon1.fr (D. Fouchet).

http://dx.doi.org/10.1016/j.epidem.2015.02.004

statistical associations, i.e., 'false interactions', between pathogens. For instance, as the risk to be infected by sexually transmitted parasites is strongly influenced by sexual behaviors (e.g., (Anderson et al., 1992)), those infections may appear associated even if they do not biologically interact, i.e., even if there is no 'true interaction'. In general, if several parasites share common risk factors the same host individuals are simultaneously at risk for those parasites and 'false interactions' are generated.

When confounding factors can be identified and recorded, true parasite interactions can be searched for using statistical tools such as log-linear models (e.g., (Behnke et al., 2005; Howard et al., 2001)), or modified chi-square analyses (e.g., (Kuris and Lafferty, 1994; Hellard et al., 2012)). However, some confounding factors are difficult to measure in the field. Host age is a striking example. Estimating the age of individuals is difficult for many wildlife species, leading to potentially strong errors in age determination or in the classification of individuals into age classes. Nonetheless, there are surprisingly very few statistical methods enabling to adjust for misclassification errors (Heisey et al., 2006; Conn and Diefenbach, 2007) and none in the framework of parasite interactions.

Host age is yet an absolutely crucial covariate to take into account as it is a risk factor for many infectious diseases. Among other examples, younger animals generally harbor fewer species of helminthes and lower worm burdens (Montgomery and Montgomery, 1989; Abu-Madi et al., 1998), and the seroprevalence increases with age for various parasites (e.g., Toxoplasmosis: (Jones et al., 2001); Hepatitis: (Murrill et al., 2002); Feline viruses: (Hellard et al., 2011)). This age-dependence is due to two additive phenomena. First, age has a 'biological' effect as host behaviors and immune defenses may evolve with age (e.g., (Anderson et al., 1992; Gasparoni et al., 2003; Levy, 2007; Bogaards et al., 2010)). Second, older individuals are more likely to be seropositive because of a longer exposure time, mechanically creating a cumulative effect of age.

When host age cannot be precisely estimated, it is crucial to deal at least with its mechanical effect. This is particularly important when studying microparasites in natural populations as such studies are most of the time cross-sectional and serological. Contrary to macroparasites whose follow up can be done quantitatively (i.e., fecal or blood counts), microparasites are often detected using indirect signs such as specific antibodies. Many infections are short (i.e., acute infections) and shedding times too brief (few hours or days) to make the search for the microparasites themselves efficient. This would require capturing hosts exactly when they are infectious. Most field data are thus limited to observed frequencies of seronegative, seropositive, and doubly seropositive individuals, with no information on the exact time of infection or on its intensity. In this context, the search for potential interactions between pairs of parasites consists in determining whether they are more often associated than expected by chance. The use of serological data reinforces the cumulative effect of age because antibodies may persist for months or years within the host. Aged individuals have a higher probability to have been exposed to the parasite and to have acquired the specific antibodies. The number of double seropositive individuals should also increase with age, thereby creating statistical associations between parasites.

In this paper, we address the problem of the detection of parasite interactions in cross-sectional data when there is no individual-specific age information or when only age-classes can be determined (e.g., juveniles *versus* adults). We focus on the impact of the cumulative effect of age in the search for interactions in serological data as this is an obvious application. However our approach is applicable for any disease, infectious or not (e.g., exposure to toxic chemicals, air pollutants, environmental agents, or toxic substances that may lead to health disorders later in life), as long as they are detectable for a certain period of time. We first use a dynamic model to investigate the impact of the cumulative effect of age in the detection of pairwise parasite interactions. Second, we propose a method to correct for this mechanical effect. We adapt a statistical method previously proposed to account for identified confounding factors in the search for parasite interactions ('the corrected chi-square', (Hellard et al., 2012)). This approach is applied to a real serological dataset of four feline viruses obtained in natural populations of domestic cats.

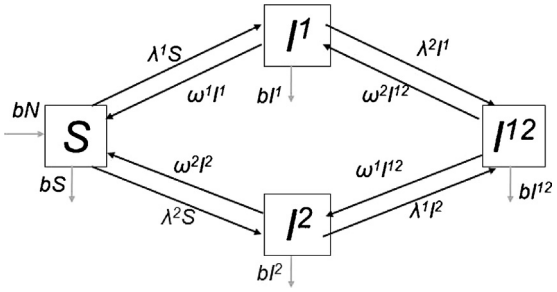

Fig. 1. Compartmental representation of the model of circulation of two independent parasites. The host population is divided in four compartments: susceptible (S), positive to parasite 1 (I^1), positive to parasite 2 (I^2) and coinfected (I^{12}) individuals. λ^x and ω^x are the force of infection and the antibodies' disappearance rate of parasite x, respectively. Birth and mortality rates (grey arrows) are considered to be constant, of equal value (b) and similar whether hosts are susceptible, singly or doubly positive.

2. Material and methods

2.1. The cumulative effect of age and false parasite interactions

A dynamic compartmental model in continuous time was used to model the circulation of two independent parasites sharing age as a common risk factor. In this model, the population is split into four classes, representing the frequency of individuals in each serological state: susceptible individuals, S, individuals positive to parasite 1, I^1, individuals positive to parasite 2, I^2, and doubly positive individuals, I^{12} (Fig. 1). Hosts are infected by parasite x with a force of infection λ^x, whereas the specific antibodies elicited after infection by parasite x disappear at a rate ω^x. Birth and mortality rates are assumed to be constant, both equal to b and not influenced by infections. Parasites are independent, i.e., the force of infection and antibodies' disappearance rate of a given parasite are independent from those of the other parasite.

2.2. Parasite characteristics favoring false interactions: a theoretical approach

The deterministic version of the model was used to determine how the cumulative effect of age can generate false interactions depending on the parasite characteristics. We considered the rate at which individuals acquire the infection (λ^x) and the antibodies' disappearance rate (ω^x) (Table 1).

Table 1
Parameters of the deterministic model and their values used in the theoretical approach.

Parameter	Definition	Value
b	Birth and mortality rate (both are considered equal)	0.25
ω^x	Antibodies' disappearance rate of parasite x	[0; 2]
λ^x	Force of infection of parasite x	[0.1; 20]

The dynamics of the two parasites within the host population are governed by the following set of ordinary differential equations:

$$\frac{dS}{dt} = b - \lambda^1 S - \lambda^2 S - bS + \omega^1 I^1 + \omega^2 I^2 \quad (1)$$

$$\frac{dI^1}{dt} = \lambda^1 S + \omega^2 I^{12} - \omega^1 I^1 - bI^1 - \lambda^2 I^1 \quad (2)$$

$$\frac{dI^{12}}{dt} = \lambda^2 I^1 + \lambda^1 I^2 - \omega^1 I^{12} - \omega^2 I^{12} - bI^{12} \quad (3)$$

$$\frac{dI^2}{dt} = \lambda^2 S + \omega^1 I^{12} - \omega^2 I^2 - bI^2 - \lambda^1 I^2 \quad (4)$$

$$S + I^1 + I^2 + I^{12} = 1 \quad (5)$$

To see whether the cumulative effect of age could generate false interactions, we used the Pearson's chi-square statistic (chi^2) as a measure of statistical association between the two parasites. It compares the observed frequencies given by the model to the theoretical frequencies expected if parasites are independent. The theoretical frequencies ($E_{k,l}$) are obtained under the null hypothesis that the joint distribution of the cell counts in the 2-dimensional contingency table is the product of the row and column marginals. The Pearson's chi^2 is then calculated as follows:

$$\chi^2 = \sum_{k=0}^{1}\sum_{l=0}^{1} \frac{(O_{k,l} - E_{k,l})^2}{E_{k,l}} \quad (6)$$

where k is the status to parasite 1 (0 for seronegative and 1 for seropositive) and l is the status to parasite 2. The chi^2 should be null if parasites are independent but should increase if statistical associations are generated by the cumulative effect of age.

False interactions are studied at the equilibrium, i.e. when new infections are exactly balanced by recovery and death of infected individuals. In that theoretical case, individuals suffer a constant force of infection, so we simply consider λ^x as constants, which renders the analysis simpler.

In our illustrative example, birth/mortality rate is taken equal to 0.25 (Table 1), corresponding to a four years life expectancy of the host.

2.3. Bias introduced in a real dataset: an empirical approach

The stochastic version of the model was then applied to a real serological dataset obtained in natural populations of domestic cats to study the influence of the cumulative effect of age in the search for interactions between feline viruses in the field. Data for four viruses were available: the Feline Immunodeficiency Virus (FIV), the Feline Herpesvirus (FHV), the Feline Calicivirus (FCV), and the Feline Parvovirus (FPV). Details on the studied viruses, on the protocols and on the model can be found in Appendix A1 and in previous papers (Hellard et al., 2011; Fouchet et al., 2009).

As the value of the antibodies' disappearance rates of the studied viruses varied a lot between studies, we tested different scenarios: (1) both viruses induce long lasting antibodies, (2) one virus induces long lasting antibodies and the other induces short-lasting antibodies, (3) both viruses induce short lasting antibodies (see Appendix A1 for more details).

The purpose of this section was to show how the cumulative effect of age can introduce a strong bias in the distribution of the Pearson's chi^2. To simplify, we did not introduce other risk factors (which is the purpose of the next Section 2.4) and consider no age-class for now.

2.4. Correcting for the cumulative effect of age

2.4.1. The correction

To correct for the cumulative effect of age, we extend the corrected chi^2 approach proposed in (Hellard et al., 2012) by including the change of the probability of infection with age as modeled in the mathematical model presented above (Eqs. (1)–(5), Fig. 1).

The first step of the approach consists in modeling the probability of infection of each individual. In this model, the rate at which individuals acquire the infection may differ between individuals. Let us call $X_1, \ldots X_N$ the set of covariates that we want to include in the model (i.e. the factors that may affect the individual probability of acquiring the infection and hence may generate false interactions, see (Hellard et al., 2012) for more details). The rate at which individual i gets infected by the parasite x at time t is given by:

$$\ln\left(\lambda_i^x(t)\right) = a_0^x + \sum a_k^x X_k(t) \quad (7)$$

where a_k^x are constant coefficients describing how the k-th risk factor X_k affects the rate of infection. Note that the only covariate that may vary over time is age-class. We assume that coefficients are constant, which implicitly means that we assume that the rate of infection is constant as long as the individuals remain in the same age-class. Since the transmission rate is age-class dependent and that age varies over time, we denote the transmission rate by individual i belonging to age-class A $\lambda_{i,A}^x$.

If we call ω the rate at which a (sero)positive individual becomes negative, then the probability (p_i) that individual i is observed negative at age a is given by the following equation:

$$\frac{dp_i^x}{da} = -\lambda_{(i,A)}^x(t) p_i^x(a) + \omega^x[1 - p_i^x(a)] \quad (8)$$

Since we do not know the exact age of individuals, the probability that individual i is observed negative is given by the average probability over the entire age-class:

$$P_i^x(A) = \int_{a1(A)}^{a2(A)} d(a) p_i^x(a) da \quad (9)$$

where *a1(A)* and *a2(A)* are the boundaries of age-class A (*a2* equals infinity for the last age-class). Note that the length of the age-classes can be different and that if no information on age is known, a single age-class with an infinite upper-bound can be used. *d(a)* is the age-distribution of individuals within age-class. Unfortunately, for most species we do not know this age-distribution. In this paper, we consider the simplified case where both the flow of newborns and the mortality rate within age-classes are constant, leading to a distribution of the density of individuals within age-classes exponentially decreasing with age (see Appendix A2 for computation details):

$$d(a) = \frac{m(A)e^{-m(A)a}}{e^{-a1(A)m(A)} - e^{-a2(A)m(A)}} \quad (10)$$

where *m(A)* is the mortality rate within the age-class A.

After solving Eq. (8), it can be easily shown that the probability of being negative for an individual belonging to age-class A is given by (see Appendix A2 for computation details):

$$P_i^x(A) = \left[Q_i^x(A) - \frac{\omega^x}{\omega^x + \lambda_{i,A}^x}\right] e^{\left[\omega^x + \lambda_{i,A}^x\right] a1(A)} \times \frac{e^{-\left[\omega^x + \lambda_{i,A}^x + m(A)\right] a1(A)} - e^{-\left[\omega^x + \lambda_{i,A}^x + m(A)\right] a2(A)}}{e^{-m(A)a1(A)} - e^{-m(A)a2(A)}} \times \frac{m(A)}{\omega^x + \lambda_{i,A}^x + m(A)} + \frac{\omega^x}{\omega^x + \lambda_{i,A}^x} \quad (11)$$

where $Q_i(A)$ is the probability of being negative for an individual entering the age-class A and given by the recursive formula:

$$\begin{cases} Q_i^x(1) = 1 \\ Q_i^x(A+1) = \left[Q_i^x(A) - \dfrac{\omega^x}{\omega^x + \lambda_{i,A}^x} \right] e^{-\left[\omega^x + \lambda_{i,A}^x\right](a2(A) - a1(A))} + \dfrac{\omega^x}{\omega^x + \lambda_{i,A}^x} \end{cases} \quad (12)$$

The second step of the method consists in inferring the coefficients $a_k{}^x$ from the data. Since we now know the theoretical probability that each individual is observed positive, we can easily calculate the likelihood as a function of the coefficients $a_k{}^x$:

$$L^x = \prod_{i \in \Omega-} P_i^x(A_{obs}(i)) \prod_{i \in \Omega+} (1 - P_i^x(A_{obs}(i))) \quad (13)$$

where $\Omega-$ and $\Omega+$ stand for the set of negative and positive individuals, respectively, and $A_{obs}(i)$ is the observed age-class of individual i.

Using likelihood maximization we can calculate the coefficients $\widehat{a_k^x}$ for each parasite x, which estimate the coefficients $a_k{}^x$ for the parasite x.

The third step consists of calculating the theoretical table, giving the expected number of individuals positive to none, one, or two parasites, assuming no interaction between the parasites. Four states are possible according to the status (St) to both parasites: neg/neg (St1 = neg, St2 = neg), neg/pos (St1 = neg, St2 = pos), pos/neg and pos/pos. The probability that individual i is in each of the four states is obtained from the formula:

$$P_i(St1, St2|A) = \int_{a2(A)}^{a1(A)} P_i(St1|a) P_i(St2|a)\, da \quad (14)$$

The solution of this equation is given in Appendix A2 (see Eqs. A13–A15).

By summing the probabilities over all individuals when they are in their observed age-class (A_{obs}), we obtain the corrected expected number of individuals ($C_{k,l}$) in each of the four states:

$$C_{k,l} = \sum_i P_i(St1 = k, St2 = l|A_{\text{obs}}(i)) \quad (15)$$

The corrected chi^2 is simply the Pearson chi^2 that compares the expected and the observed tables:

$$\chi^2 = \sum_{k=0}^{1} \sum_{l=0}^{1} \frac{(O_{k,l} - C_{k,l})^2}{C_{k,l}} \quad (16)$$

In the last step, *K in silico* data sets are generated under H0: 'the two parasites are independent' in order to estimate the distribution of the corrected chi^2 under H0. First, for each individual, a serological status to both parasites is sampled assuming the probability of infection given in Eq. (9) by replacing for each parasite the coefficients a_k^x by their estimates ($\widehat{a_k^x}$). Then the *in silico* data set is reanalyzed and the associated corrected chi^2 is calculated.

For each of the *K in silico* data sets, we obtain one estimate of the corrected chi^2 leading to *K* independent realizations of the corrected chi^2 statistic under H0.

The p-value is calculated as follows. Since the chi^2 quantifies the deviation from theoretical quantities, we consider an unilateral *p*-value. If H0 is true, the observed chi^2 combined with the *K* simulated values provide $K+1$ independent realizations of the same distribution. The probability of observing a value as extreme as the one observed is simply given by the number of *in silico* chi^2 that are above the observed corrected chi^2 plus one, divided by $K+1$.

All simulations were run in the R software (R Core Team, 2013) and were performed using the computing facilities of the CC LBBE/PRABI. The method has been included in a new R package that we called *Interatrix*.

2.5. Application to the real dataset

The method described above was then applied to the same dataset of feline viruses than before. Four age-classes were considered (0–2, 2–4, 4–6 years old and older than 6 years), as classically in the study of domestic cats, as well as all risk factors identified previously in our populations (Hellard et al., 2012; Hellard et al., 2011).

3. Results

3.1. The cumulative effect of age and false parasites interactions

3.1.1. Parasite characteristics favoring false interactions: a theoretical approach (deterministic modeling)

Using the dynamic model presented in Section 2.1, we investigated the effect of the force of infection (λ^x) by fixing the antibodies' disappearance rate of both parasites and by varying λ^x between 0.1 and 20 (Table 1). The force of infection of the two parasites were first considered to be similar ($\lambda^1 = \lambda^2 = \lambda$; Fig. 2(B), Fig. 3(B), then dissimilar ($\lambda^1 \neq \lambda^2$; Fig. A3.2 in Appendix A3). The effect of the antibodies' disappearance rate (ω^x) was tested by fixing the force of infection and by varying ω^x between 0 and 2 (Table 1). The antibodies' disappearance rate of the two parasites were first considered similar ($\omega^1 = \omega^2 = \omega$; Fig. 2(A), Fig. 3(A) then dissimilar ($\omega^1 \neq \omega^2$, Fig. A3.2 in Appendix A3).

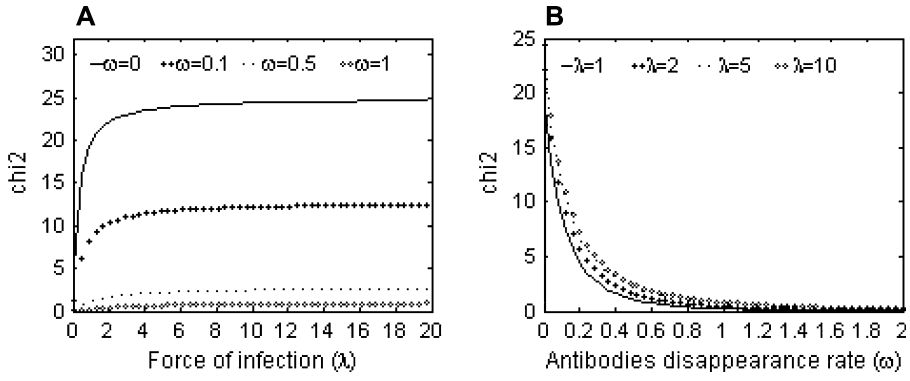

Fig. 2. Pearson's chi^2 statistic in function of parasite x force of infection, λ^x (A) and antibodies' disappearance rate, ω^x (B). The same rates are taken for both parasites (i.e., $\lambda^1 = \lambda^2 = \lambda$; $\omega^1 = \omega^2 = \omega$). In (A), both parasites have an antibodies' disappearance rate, ω, of 0.1; in (B), both parasites have a force of infection, λ, of 10.

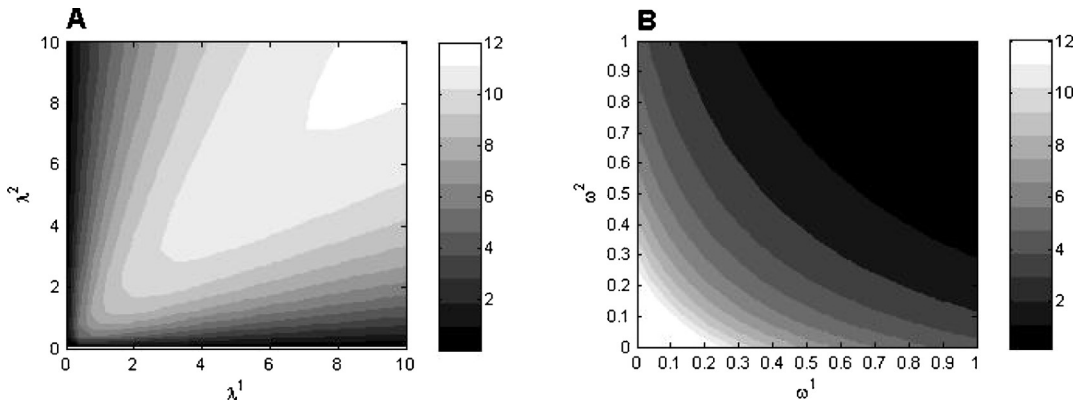

Fig. 3. Pearson's chi^2 statistic in function of the force of infection, λ^x (A) and of the antibodies' disappearance rate, ω^x (B) of each parasite x. In (A), both parasites have a similar antibodies' disappearance rate, $\omega^1 = \omega^2 = \omega$, of 0.1; in (B), both parasites have a similar force of infection, $\lambda^1 = \lambda^2 = \lambda$, of 10.

Using the deterministic version of the compartmental model, we found that the value of the chi^2 increased with the force of infection and decreased with the antibodies' disappearance rate (Fig. 2). In addition, the chi^2 was maximal when both parasites had similar forces of infection and persistent antibodies (Fig. 3; see Appendix A3 for a wider range of ω^x and λ^x values).

3.1.2. *Bias introduced in a real dataset: an empirical approach (stochastic modeling)*

If the cumulative effect of age was not taken into account (i.e., using a classical Pearson's chi^2), five out of the six pairs of feline viruses appeared significantly associated with a type I error of 5% ((Hellard et al., 2012), Table A1 in Appendix A1). Using the dynamic stochastic model, we found that four of these associations could in fact be explained by the cumulative effect of age when both viruses induced long-lasting antibodies (scenario 1; Table A1 in Appendix A1). It was not the case anymore however if at least one virus induced short immunity (scenarios 2 and 3; Table A1 in Appendix A1).

The stochastic modeling confirmed that antibodies persistence was of premium importance. The simulated chi^2 distribution was very strongly shifted compared to the distribution of a Pearson's

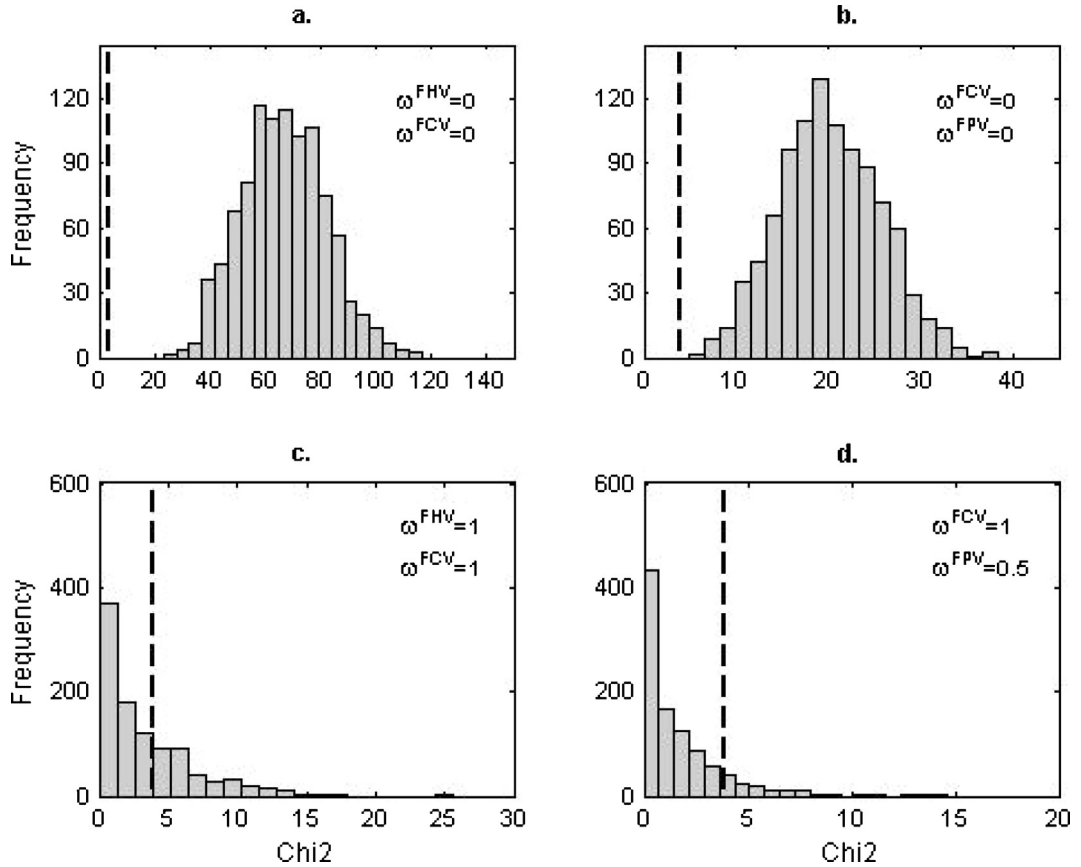

Fig. 4. Examples of simulated chi^2 distribution under H0: "viruses are independent" for two pairs of feline viruses under scenario 1: antibodies' persistence (ω^x) is long for both viruses (a, b) and scenario 3: antibodies' persistence is short for both viruses (c, d). The dashed vertical line indicates the value 3.84, i.e., the theoretical threshold for a Pearson's chi^2 with one degree of freedom and an alpha risk of 5%. The forces of infection of the viruses are as follows: (a) ($\lambda^{FHV} = 0.40$, $\lambda^{FCV} = 1.19$); (b) ($\lambda^{FCV} = 1.19$, $\lambda^{FPV} = 0.08$); (c) ($\lambda^{FHV} = 2.02$, $\lambda^{FCV} = 5.95$); (d) ($\lambda^{FCV} = 5.95$, $\lambda^{FPV} = 0.24$).

Table 2
Chi² tests on pairs of feline viruses after (a) correction for known risk factors (including age as a factor with four age-classes, from (Hellard et al., 2012)) and (b) correction for risk factors (including age as a factor with four age-classes) and for the cumulative effect of age (CEA). Note that there are two scenarios 2 for the three last pairs of viruses, one scenario where the first virus has a long antibodies persistence and the second virus a short antibodies persistence and one scenario with the contrary.

Virus1	Virus2	Scenario	ω^1	ω^2	Chi² | risk factors		Chi² | risk factors and CEA	
					chi²	*P*	chi²	*P*
FIV	FHV	1	0	0	2.24	0.05	1.91	0.10
		2 and 3	0	1			3.25	0.03
FIV	FCV	1	0	0	1.46	0.14	1.85	0.12
		2 and 3	0	1			4.12	0.03
FIV	FPV	1	0	0	0.68	0.23	0.08	0.74
		2 and 3	0	0.5			0.14	0.62
FHV	FCV	1	0	0	20.81	1.9×10^{-8}	7.58	1×10^{-3}
		2	0	1			14.29	1×10^{-3}
		2	1	0			16.49	1×10^{-3}
		3	1	1			19.84	1×10^{-3}
FHV	FPV	1	0	0	54.26	0.00	34.67	1×10^{-3}
		2	0	0.5			40.44	1×10^{-3}
		2	1	0			40.69	1×10^{-3}
		3	1	0.5			45.43	1×10^{-3}
FCV	FPV	1	0	0	26.39	1.7×10^{-11}	14.25	1×10^{-3}
		2	0	0.5			16.02	1×10^{-3}
		2	1	0			16.19	1×10^{-3}
		3	1	0.5			19.31	1×10^{-3}

chi² when antibodies were long-lasting for both viruses (Fig. 4(a). for the FHV-FCV pair and b. for the FCV-FPV pair), especially if viruses had similar forces of infection (Fig. 4(a); Table A1 in Appendix A1). When forces of infection were sufficiently dissimilar and when antibodies were short-lasting for at least one virus, the cumulative effect of age was on the contrary negligible (Fig. 4(d); Table A1 in Appendix A1).

3.2. Application of the correction for the cumulative effect of age to pairs of feline viruses

When the correction presented in Section 2.2 was applied to the feline viruses, three associations remained significant if both viruses induced long-lasting antibodies (scenario 1), and five when at least one virus induced short-lasting antibodies (scenario 2 and 3) (Table 2). These results are coherent with those found in (Hellard et al., 2012), where no correction for the cumulative effect of age was applied but where age was a correcting factor. This suggests that the three associations remaining significant, even when antibodies are persistent, are strong and that the cumulative effect of age is not enough to explain these associations.

4. Discussion

The bias introduced by the cumulative effect of age in the search for parasite interactions in the field has long been suspected (Courchamp et al., 2000; Cross et al., 2009), without providing a theoretical background explaining how this cumulative effect works nor a proper way to take this phenomenon into account into a statistical framework. Intuitive solutions have been proposed, such as the restriction of the dataset to a narrow age-class, most often the young, in which the cumulative effect of age is less important (Kuris and Lafferty, 1994). Age-stratification of the data has also been used as an alternative strategy (e.g., (Farrington et al., 2001)).

Using deterministic and stochastic versions of an epidemiological model, this paper provides a theoretical background to study the degree of false interactions (i.e., statistical associations) that can be generated by the cumulative effect of age. Its effect was indeed important, i.e., found to generate apparent interactions between independent parasites, especially when forces of infection were high (i.e., strong transmission, high prevalence) and when antibodies persisted for a long time within hosts (i.e., long immunity). The maximal risk was when parasites were both transmitted with a similar force of infection. When stochastic modeling was applied to a real dataset of feline viruses, we showed that this mechanical effect of age could be strong in realistic conditions.

The method we propose here is an extension of the corrected chi² method proposed in (Hellard et al., 2012), which enabled the correction for other potential confounding risk factors such as sex, way of life or the population of origin. Age could also be integrated into the analysis, but only using the available information. The method could in particular account for differences in disease exposition between two identified age-classes. But no correction could be done at a finer age-scale, i.e., within the same age-class. When the precise age of individuals was not known, it was impossible to account for the fact that within a given age-class older individuals had a longer history of exposition to all parasites than younger individuals of the same class. This later phenomenon (i.e., the cumulative effect of age) is important since in many species age-classes can be very large (e.g., in long lived species where the precise age of adults is unknown). The method presented in this paper accounts for the cumulative effect of age in the search for parasite interactions, and is rendered available in a new R package, *Interatrix*, which also includes the corrected chi² proposed in (Hellard et al., 2012).

Applied to feline viruses, the mechanical effect of age was not sufficient to explain the observed associations between FHV and FPV, FHV and FCV and FPV and FCV. Results are consistent with what was observed in (Hellard et al., 2012) with the simple corrected chi² that did not correct for the cumulative effect of age. It is not surprising in that case since the three associations were highly significant. Even though the observed associations were explained a little bit better when the cumulative effect of age was accounted for, unexplained links between the three sets of viruses remained high. At this stage it is however too early to make any distinction between real interactions and the existence of one or more other unidentified confounding factors. In either case, these interactions would be interesting to further examine in experimental conditions.

As most statistical tools, our method presents some limitations. To understand these limitations, it is important to consider that the

search for parasite associations in the field can, under no circumstance, provide any proof of a real biological interaction. A high (or low) level of association between two parasites can always result from an existing and unmeasured shared risk factor. Among all, the individual personality (e.g., (Natoli et al., 2005)), in particular, can be a confounding risk factor in all host-parasite systems, and is extremely difficult to capture in the field.

For that reason we believe that the analysis of parasite associations in the field should be viewed as a tool helping to choose which parasite association should be more extensively studied in the lab. In the field one can only investigate which parasite associations can be explained by confounding risk factors. Parasite pairs for which the observed association cannot be explained can motivate further experiments and are more prone to lead to interesting studies in the lab.

As decision tools, statistical methods are imperfect by nature, and can only be designed to take the maximum benefit of the collected data. Introducing the cumulative effect of age into the analysis of parasite associations is clearly a step in this direction. One should remember however that it was done here under restrictive assumptions, making the method more relevant in some situations than in others. In particular, we assumed a constant exposition rate both in time and within age-classes. This simplification is reasonable for endemic pathogens but not for epidemic ones. Studying the association between two epidemic parasites would require including the timing of the two epidemics in the model, a fundamental property that can affect the level of expected mechanical associations. Our method also relies on survival parameters. In particular, survival within the last age-class affects the host life-expectancy and needs to be informed to correctly evaluate the level of potential false interactions due to the cumulative effect of age.

When trying to explain associations between pairs of parasites, we recommend the following procedure. First, one should start by using the method that does not account for the cumulative effect of age presented in (Hellard et al., 2012) (also included in the *Interatrix* package). If no significant association is found there is no need to go further. If a negative association is found there is no need to go further neither, since the cumulative effect of age only generates positive associations. If a positive association is found and if antibodies are suspected to be potentially long lasting, then the full method should be used. The maximum rate of antibody disappearance should be fixed at its smaller realistic value (i.e. larger realistic value for the duration of antibody). If the association remains significant, then nothing in the measured risk factors can explain the observed association and one can suspect a true interaction.

Finally, it is important to remind that the nature of the data analyzed in this framework (transversal data) only allows highlighting potential interactions that could lead to an over (or under) representation of co-positive individuals (positive either to specific antibodies or to antigens, depending on the positivity criterion used). It cannot provide information on the impact of co-infections on the severity of the disease.

5. Conclusion

Parasite interactions have been shown to have serious consequences on parasite circulation, disease severity and management (e.g., (Abu-Raddad et al., 2006; Lello et al., 2004; Corbett et al., 2002; Graham et al., 2005)) but their number still tends to be overestimated. Recent advances in omic technologies now allow the identification of past and present infections by several parasitic agents in wild populations. It offers the growing opportunity to detect associations between numerous pairs of potentially interacting agents. However, deep and robust statistical methods are needed to decide which of these suspected associations should be brought into the lab for further analysis. The method proposed in this paper makes one additional step into that direction by including the cumulative effect of age in the set of potential confounding factors generating false interactions (i.e., statistical associations) between independent parasites.

Authors contributions

EH, DP, DF wrote the manuscript. DF and EH designed the dynamical models and DF designed the statistical method. AS, FS, DF wrote the R code. AS developed the R package. DP is a specialist of the cat-feline viruses system. DF supervised the work. All authors approved the final version of the article.

Acknowledgments

We would like to thank the anonymous reviewers for their valuable inputs on the manuscript; Lyon-Biopôle (FIV-VAX program) for financial support; Guillaume Leblanc for data collection in the field and Hervé Poulet, Béatrice Tarin, Vincent Badol and Christine Coupier for serological analyses. This work was performed within the framework of the LABEX ECOFECT (ANR-11-LABX-0048) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:10.1016/j.epidem.2015.02.004.

References

Anderson, R.M., May, R., Ng, T.W., Rowley, J.T., 1992. Age-dependent choice of sexual partners and the transmission dynamics of HIV in Sub-Saharan Africa. Phil. Trans. R. Soc. B 336, 135–155.

Abu-Raddad, L.J., Patnaik, P., Kublin, J.G., 2006. Dual infection with HIV and malaria fuels the spread of both diseases in sub-Saharan Africa. Science 314, 1603–1606.

Abu-Madi, M., Behnke, J., Lewis, J., Gilbert, F., 1998. Descriptive epidemiology of *Heligmosomoides polygyrus* in *Apodemus sylvaticus* from three contrasting habitats in south-east England. J. Helminthol. 72, 93–100.

Behnke, J.M., Wakelin, D., Wilson, M., 1978. *Trichinella spiralis*: delayed rejection in mice concurrently infected with *Nematospiroides dubius*. Exp. Parasitol. 46, 121–130.

Bentwich, Z., Kalinkovich, A., Weisman, Z., Borkow, G., Beyers, N., Beyers, A.D., 1999. Can eradication of helminthic infections change the face of AIDS and tuberculosis? Immunol. Today 20, 485–487.

Bogaards, J.A., Xiridou, M., Coupe, V.M.H., Meijer, C.J.L.M., Wallinga, J., Berkhof, J., 2010. Model-based estimation of viral transmissibility and infection-induced resistance from the age-dependent prevalence of infection for 14 high-risk types of Human Papillomavirus. Am. J. Epidemiol. 171, 817–825.

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.

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. Z. Parasitenk. 73, 387–410.

Corbett, E.L., Steketee, R.W., ter Kuile, F.O., Latif, A.S., Kamali, A., Hayes, R.J., 2002. HIV-1/AIDS and the control of other infectious diseases in Africa. Lancet 359, 2177–2218.

Conn, P.B., Diefenbach, D.R., 2007. Adjusting age and stage distributions for misclassification errors. Ecology 88, 1977–1983.

Celum, C.L., 2004. The interaction between herpes simplex virus and human immunodeficiency virus. Herpes 11 (Suppl 1), 36A–45A.

Cox, F., 2001. Concomitant infections, parasites and immune responses. Parasitology 122, S23–S38.

Courchamp, F., Say, L., Pontier, D., 2000. Detection, identification, and correction of a bias in an epidemiological study. J. Wildl. Dis. 36, 71–78.

Cross, P., Drewe, J., Patrek, V., Pearce, G., Samuel, M., Delahay, R., 2009. Wildlife population structure and parasite transmission: implications for disease management. In: Delahay, R.J., et al. (Eds.), Management of Disease in Wild Mammals. Springer, pp. 9–29.

Druilhe, P., Tall, A., Sokhna, C., 2005. Worms can worsen malaria: towards a new means to roll back malaria? Trends Parasitol. 21, 359–362.

Dezfuli, B.S., Giari, L., De Biaggi, S., Poulin, R., 2001. Associations and interactions among intestinal helminths of the brown trout, *Salmo trutta*, in northern Italy. J. Helminthol. 75, 331–336.

Fouchet, D., Leblanc, G., Sauvage, F., Guiserix, M., Poulet, H., Pontier, D., 2009. Using dynamic stochastic modeling to estimate population risk factors in infectious disease: the example of FIV in 15 cat populations. PLoS One 4, e7377.

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.

Farrington, C., Kanaan, M., Gay, N., 2001. Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data. J. Roy. Stat. Soc. C-App. 50, 251–292.

Gasparoni, A., Ciardelli, L., Avanzini, A., Castellazzi, A., Carini, R., Rondini, G., Chirico, G., 2003. Age-related changes in intracellular Th1/Th2 cytokine production, immunoproliferative T lymphocyte response and natural killer cell activity in newborns, children and adults. Biol. Neonate 84, 297–303.

Graham, A.L., Lamb, T.J., Read, A.F., Allen, J.E., 2005. Malaria-filaria coinfection in mice makes malarial disease more severe unless filarial infection achieves patency. J. Infect. Dis. 191, 410–421.

Harris, J.B., Podolsky, M.J., Bhuiyan, T.R., Chowdhury, F., Khan, A.I., LaRocque, R.C., Logvinenko, T., Kendall, J., Faruque, A.S.G., Nagler, C.R., Ryan, E.T., Qadri, F., Calderwood, S.B., 2009. Immunologic responses to *Vibrio cholerae* in patients co-infected with intestinal parasites in Bangladesh. PLoS Neglect. Trop. D. 3, e403.

Hellard, E., Pontier, D., Sauvage, F., Poulet, H., Fouchet, D., 2012. True *versus* false parasite interactions: a robust method to take risk factors into account and its application to feline viruses. PLoS One 7, e29618, http://dx.doi.org/10.1371/journal.pone.0029618.

Heisey, D.M., Joly, D.O., Messier, F., 2006. The fitting of general force-of-infection models to wildlife disease prevalence data. Ecology 87, 2356–2365.

Howard, S.C., Donnell, C., Chan, M.S., 2001. Methods for estimation of associations between multiple species parasite infections. Parasitology 122, 233–251.

Hellard, E., Fouchet, D., Santin-Janin, H., Tarin, B., Badol, V., Coupier, C., Leblanc, G., Poulet, H., Pontier, D., 2011. When cats' ways of life interact with their viruses: A study in 15 natural populations of owned and unowned cats (*Felis silvestris catus*). Prev. Vet. Med. 101, 250–264.

Ives, A., Ronet, C., Prevel, F., Ruzzante, G., Fuertes-Marraco, S., Schutz, F., Zangger, H., Revaz-Breton, M., Lye, L.F., Hickerson, S.M., Beverley, S.M., Acha-Orbea, H., Launois, P., Fasel, N., Masina, S., 2011. Leishmania RNA Virus Controls the Severity of Mucocutaneous Leishmaniasis. Science 331, 775–778.

Jolles, A.E., Ezenwa, V.O., Etienne, R.S., Turner, W.C., Olff, H., 2008. Interactions between macroparasites and microparasites drive infection patterns in free-ranging African buffalo. Ecology 89, 2239–2250.

Jones, J., Kruszon-Moran, D., Wilson, M., McQuillan, G., Navin, T., McAuley, J., 2001. *Toxoplasma gondii* infection in the United States: seroprevalence and risk factors. Am. J. Epidemiol. 154, 357–365.

Koch, H., Schmid-Hempel, P., 2011. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. Proc. Natl. Acad. Sci. USA 108, 19288–19292.

Kuris, A.M., Lafferty, K.D., 1994. Analysis of larval trematode communities. Ecology 75, 2275–2285.

Keesing, F., Belden, L.K., Daszak, P., Dobson, A., Harvell, C.D., Holt, R.D., Hudson, P., Jolles, A., Jones, K.E., Mitchell, C.E., Myers, S.S., Bogich, T., Ostfeld, R.S., 2010. Impacts of biodiversity on the emergence and transmission of infectious diseases. Nature 468, 647–652.

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.

Levy, O., 2007. Innate immunity of the newborn: basic mechanisms and clinical correlates. Nat. Rev. Immunol. 7, 379–390.

Montgomery, S., Montgomery, W., 1989. Spatial and temporal variation in the infracommunity structure of helminths of *Apodemus sylvaticus* (Rodentia: Muridae). Parasitology 98, 145–150.

Murrill, C., Weeks, H., Castrucci, B., Weinstock, H., Bell, B., Spruill, C., Gwinn, M., 2002. Age-specific seroprevalence of HIV, hepatitis B virus, and hepatitis C virus infection among injection drug users admitted to drug treatment in 6 US cities, Am. J. Public Health 92, 385–387.

Natoli, E., Say, L., Cafazzo, S., Bonanni, R., Schmid, M., Pontier, D., 2005. Bold attitude makes male urban feral domestic cats more vulnerable to Feline Immunodeficiency Virus. Neurosci. Biobehav. Rev. 29, 151–157.

Pedersen, A.B., Fenton, A., 2007. Emphasizing the ecology in parasite community ecology. Trends Ecol. Evol. 22, 133–139.

Petney, T., Andrews, R., 1998. Multiparasite communities in animals and humans: frequency, structure and pathogenic significance. Int. J. Parasitol. 28, 377–393.

Pontier, D., Guiserix, M., Fouchet, D., Sauvage, F., Gonzalez, J.P., 2009. Emergence of infectious diseases: when hidden pathogens break out. C.R. Biol. 332, 539–547.

Read, A.F., Taylor, L.H., 2001. The ecology of genetically diverse infections. Science 292, 1099–1102.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ⟨http://www.R-project.org/⟩.

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.

Weiss, R.A., McMichael, A.J., 2004. Social and environmental risk factors in the emergence of infectious diseases. Nat. Med. 10, S70–S76.