Introduction

Infectious agents can have major impacts on animal populations through changing population dynamics and stability (Dobson 1992), altering predator prey interactions (Joly 2004), and even causing species decline and extinction (Viana 2015, Daszak 1999). Accurate estimates of parasite-induced host mortality (PIHM) in wild animals are important for understanding what regulates both host and parasite populations and to make predictions about disease transmission in natural systems. Although a negative impact on host fitness is a fundamental component of parasitism (Lafferty and Kuris 2002, Poulin book, other definition sources), it is notoriously difficult to quantify PIHM in wild animal populations (cite, McCallum 2008?, lester 1984).

To conclusively identify PIHM in wild animal populations, it is necessary to experimentally infect and track host populations in ways that are rarely possible in most host-parasite systems (citations). Instead, parasitologists are often only able to collect a certain number of hosts and determine each host's parasite intensity. This snapshot, distributional data is far from the ideal type of data for addressing questions regarding PIHM, but the reality is that this is the type of data on which most questions regarding PIHM are asked (citations). The two primary questions that one may wish to ask given a snapshot host- parasite dataset are: 1) Is PIHM occurring in this system? and 2) How does host survival change as parasite intensity increases? MORE

The first of these two questions was addressed by Crofton (1971) developed a method to test for PIHM using the truncation of the negative binomial distribution. In short, the Crofton Method assumes that the distribution of parasites across hosts before mortality occurs follows a negative binomial distribution (Anderson & May 1978; Shaw *et al.* 1998). As heavily infected hosts begin to die, the negative binomial distribution gets truncated because heavily infected hosts dies and are no

longer observed in a sample. In other words, the observed host-parasite distribution and the pre-mortality host-parasite distribution will predict substantially different numbers of of highly infected hosts (because those have died due to infection) but similar number of hosts with low infection loads (because those have survived). (Crofton 1971) noted that by starting with all the observed data and iteratively fitting a negative binomial distribution to hosts with lower and lower parasite loads, one could determine whether or not PIHM was occurring in the system (Figure XA). This was done graphically by determining whether parameters of the truncated negative binomial distributions showed a substantial change as hosts with smaller and smaller parasite intensities were fitted (Figure XB). The Crofton Method and this graphical technique for determining whether or not PIHM is occurring are both still used (Ferguson *et al.* 2011). A more thorough description and implementation of the Crofton Method is given in Appendix X.

The second question regarding PIHM is how host survival changes as parasite intensity increases. The Crofton Method on its own cannot answer this question and Adjei *et al.* (1986) proposed a method to determine how host survival probability changes with increasing parasite-load. The Adjei Method proceeds by first using the Crofton Method to estimate the pre-mortality parameters for a host-parasite distribution (describe what these are) and then, given these parameters, estimates a host-survival function that describes how the probability of host-survival changes with increasing parasite load (see Appendix X for a full description). From this host-survival function, the Adjei Method can estimate important host-parasite quantities such as the parasite intensity at which 50% of hosts succumb to PIHM (LD_{50}), as well as the percent of hosts in a population succumbing to PIHM (Adjei *et al.* 1986).

Asking questions regardin

Methods

The likelihood method for estimating PIHM

Given the potential deficiencies of the Adjei Method, we provide an alternative approach for estimating parasite-induced host mortality (PIHM) that makes less assumption than the previously proposed Adjei Method. The likelihood method does not require any binning or alteration of the data, potentially reduces the number of parameters that need to be estimated, and allows for standard statistical techniques to be used to assess the significance of PIHM in a system.

The likelihood method makes the following assumption about the host-parasite system. First, as with all previously proposed methods for estimating PIHM, the likelihood method assumes that the pre-mortality distribution follows a negative binomial distribution ($g(x)$) with parameter μ_p and k_p. The validity of the assumption is an inherent problem with all the PIHM methods proposed to date and we address this thoroughly in the Discussion. The second assumption that the likelihood method makes is that the host survival function takes the form of a logistic curve given by

$$h(x|a,b) = h_x = \frac{e^{a+b\log(x)}}{1+e^{a+b\log(x)}} \tag{1}$$

where x is the parasite intensity in a given host and a and b are the two parameters of the function. Generally, a larger a allows for hosts to tolerate larger parasite loads before experiencing parasite-induced mortality and a more negative b leads to a more rapid decline in the probability of host survival as parasite intensity increases. The value $exp(a/|b|)$ (notation) is typically referred to as the LD_50, which gives the parasite intensity at which 50% of host experience mortality. With these two explicit assumptions, the likelihood method tries to estimate the 4 parameters μ_p, k_p, a, and b.

To estimate these parameters, we need to define a probability distribution that gives the probability of having a parasite load of x parasites conditional on host survival. Using the standard rules of conditional probability This distribution can be written as

$$P(x|\text{survival}) = \frac{P(\text{survival}|x) * P(x)}{P(\text{survival})} \tag{2}$$

One can now see that $P(\text{survival}|x)$ is the survival function $h(x; a, b)$, $P(x)$ is the pre-mortality parasite distribution $g(x; \mu_p, k_p)$ and $P(survival) = \sum_0^\infty P(\text{survival}|x) * P(x) = \sum_{x=0}^\infty h(x; a, b) * g(x; \mu_p, k_p)$. Therefore equation 2 can be written as

$$P(x|survival) = \frac{h(x; a, b) * g(x; \mu_p, k_p)}{\sum_{x=0}^\infty h(x; a, b) * g(x; \mu_p, k_p)} \tag{3}$$

Using this probability distribution, one can then find the parameters μ_p, k_p, a, b that maximize the likelihood of an observed host-parasite dataset $\mathbf{x}$.

Alternatively, one could apply the Crofton Method to estimate μ_p and k_p and then find the maximum likelihood estimates of a and b and the corresponding LD_{50}. A final option would be to follow the example of (Ferguson *et al.* 2011) and assume $k_p = 1$ and only estimate a, b and μ_p.

0.1 Testing the ability of approaches to identify PIHM

To test the ability of the Adjei Method and the Likelihood Method to identify whether or not PIHM was occurring in a system, we randomly generated data using the following procedure. First, we drew N_p randomly infected hosts from a negative binomial distribution with parameters μ_p and k_p. This represented the dataset observed before mortality. Second, we chose values of a and b and calculated the probability of survival for all N_p hosts using equation 1. Third, we

drew N_p random numbers from a uniform distribution between 0 and 1 and if host survival probability was less than this random number, the host experienced parasite-induced mortality. The surviving hosts comprised the dataset ($\mathbf{x}$) that would be obtained in the field, after PIHM.

Using both the pre-mortality and post-mortality simulated datasets, we assumed that the values of N_p, μ_p, and k_p were known and tested the ability of both methods to correctly determine whether or not PIHM was occurring. While this scenario is unrealistic because the parameters N_p, μ_p, and k_p are always unknown, we implemented this scenario as a baseline to establish the efficacy of the methods independent of the estimates of N_p, μ_p and k_p. For the Adjei Method, N_p, μ_p, and k_p are estimated using the Crofton Method, while the μ_p and k_p in the likelihood method can be estimated jointly with a and b or via the Crofton Method. If a method could not correctly predict whether or not PIHM was occurring under these idealistic conditions, we considered this strong evidence of the unreliability of this method.

We used three different values of μ_p (10, 100, 500) and for each μ_p we examined three different survival functions that had graduate, moderate, and sharp decreases in host survival with increasing parasite intensity. For a given μ_p, each survival function had the same LD_{50}, but different values of a and b. We examined each μ_p-survival function pair at three levels of parasite aggregation, $k_p = 0.1, 0.5, 1$, which are realistic values of parasite aggregation in natural populations (Shaw *et al.* 1998). For each of these parameter combinations we simulated 150 datasets and tested the probability of each method incorrectly identifying PIHM in the pre-mortality dataset (Type I error) and incorrectly failing to identify PIHM in the post-mortality dataset (Type II error). For each method, we we used likelihood ratio test to determine whether the full model with PIHM provided a significantly better fit than the reduced model without PIHM at significance level 0.05 (Appendix X). We tested all each parameter combinations

for pre-mortality population sizes of $N_p = [300, 500, 1000, 2000, 5000, 7500, 10000]$.

Testing ability of PIHM approaches to recover survival function

To compare the ability of the Adjei Method and the likelihood method to recover LD_{50} and the parameters a and b or the survival function, we used the same simulation procedure and parameter combinations described above.For each parameter combination we simulated 150 datasets, estimated a, b, and LD_{50} and calculated the standardized bias and precision (Walther & Moore 2005) for these estimates over varying pre-mortality host population sizes $N_p = [300, 500, 1000, 2000, 5000, 7500, 10000]$. N_p is not technically the sample size on which the methods are being tested because parasite-induced mortality reduces N_p for each simulated dataset. We therefore used the average number of surviving hosts over all 150 simulations for a given parameter combination as our measure of sample size.

Application to data

We tested the likelihood method on the datasets given in (Crofton 1971) and (Adjei *et al.* 1986) which both papers reported seeing PIHM in the respective populations.

Results

Detecting PIHM

Recovery of the survival function

Application to data

Discussion

Parasite-induced host mortality is of substantial interest in many systems, but determining whether it is occurring given only observational data is notoriously difficult. Many of the previous methods derived to determine the effect of PIHM on a host-parasite system limit their inference to answering the yes or no question of whether or not PIHM is occurring. While a relevant question, it is often of interest to know something about the host survival function which can provide information regarding important properties of the host-parasite system, such as LD_{50} and percent of the population suffering PIHM.

We show that the Adjei Method, the only currently proposed method to estimate the host survival function and the LD_{50} from observational PIHM, has some serious methodological problems that result in biased estimates of the host survival function even under the most idealistic conditions. Interestingly, despite these flaws, the Adjei Method can still produce unbiased and precise estimates of the LD_{50} when host-parasite systems show aggregation close to $k = 1$

To attempt to ameliorate the flaws in the Adjei Method, we proposed a more general method to determine both whether or not PIHM is occurring in a system and the shape of the survival function. We show that this method is asymptotically unbiased when estimating the host-survival function for all of the parameter space that we explored, but can produce seriously biased estimates of the host survival function for sample sizes typically observed in many host-parasite

studies. However, we found that the likelihood method produces unbiased and precise estimates of the LD_{50} for small, realistic sample sizes.

Moreover, the likelihood method...

What do these findings tell us about our ability to go beyond saying whether or not PIHM is occurring in a system? While we have generalized and improved upon the previously existing methods for estimating PIHM, we cannot belie the fact that estimating the host survival function from observational data alone is ladened with assumptions and difficulties.

The most fundamental assumption of all methods for estimating PIHM is that the shape of the pre- mortality host-parasite distribution is known. In the discrete case, this distribution is negative binomial, while in the continuous case the distribution can be gamma of exponential ((Ferguson *et al.* 2011)). While there is substantial empirical and theoretical evidence to justify the use of the negative binomial distribution as the pre-mortality distribution for macroparasites across hosts (Crofton, Shaw, Anderson and Gordon, etc), it is widely recognized that different processes can lead to a variety of distributions of parasites across hosts (Wilber, Duerr etc). However, the critical assumption of the pre-mortality distribution is not that the processes leading the pre-mortality distribution generate a negative binomial distribution, but rather that the pre-mortality distribution is well-fit by a negative binomial. The extreme flexibility of the negative binomial distribution in the discrete case of the gamma distribution in the continuous case make them reasonable candidate distributions for the pre-mortality distributions. Therefore, we do not see this assumption as central problem in any of the proposed methods.

However, to use the pre-mortality distribution to infer whether or not the PIHM is occurring in a system requires an explicit assumption about the host survival function and the shape of the post-mortality distribution. Regarding the host-survival function, all methods of PIHM assume that the host-survival function

is such that uninfected individuals and individuals with low parasite intensity experience essentially no PIHM. Lanciani & Boyett (1989) illustrated this by showing that when hosts experienced a linear decrease in survival probability the Crofton Method could not detect PIHM. This result, relates to the shape of the post-mortality distribution. Given that the pre-mortality distribution is well-fit by a negative binomial, a linear host-survival function will result in a post-mortality distribution that is also well-fit by a negative binomial distribution (why?). In this case, one would be unable to identify PIHM because the pre-mortality distribution

As parasitologist, we often want to get the most out of our

References

1.
Adjei, E. L., Barnes, A. & Lester, R. J. G. (1986). A method for estimating possible parasite-related host mortality, illustrated using data from Caliitetrarhynchus gracilis (Cestoda: Trypanorhyncha) in lizardfish (Saurida spp.). *Parasitology*, 92, 227–243.

2.
Anderson, R. M. & May, R. M. (1978). Regulation and stability of host-parasite interactions: I. Regulatory processes. *Journal of Animal Ecology*, 47, 219–247.

3.
Crofton, H. D. (1971). A quantitative approach to parasitism. *Parasitology*, 62, 179–193.

4.
Ferguson, J. a., Koketsu, W., Ninomiya, I., Rossignol, P. a., Jacobson, K. C. & Kent, M. L. (2011). Mortality of coho salmon (Oncorhynchus kisutch) associated with burdens of multiple parasite species. *International journal for parasitology*, 41, 1197–205. URL `http://www.ncbi.nlm.nih.gov/pubmed/21855547`.

5.
Lanciani, C. A. & Boyett, J. M. (1989). Demonstrating parasitic water mite-induced mortality in natural host populations. *Parasitology*, 81, 465–475.

6.
Shaw, D. J., Grenfell, B. T. & Dobson, a. P. (1998). Patterns of macroparasite aggregation in wildlife host populations. *Parasitology*, 117 (Pt 6, 597–610.

7.
Walther, B. a. & Moore, J. L. (2005). The concepts of bias, precision and accuracy,

231 and their use in testing the performance of species richness estimators, with a
232 literature review of estimator performance. *Ecography*, 28, 815–829.