Comparing methods for estimating parasite-induced host mortality from distributional data: implementation and limitations

Mark Wilber, Sara Weinstein, and Cherie Briggs

May 5, 2015

Abstract

TODO

Introduction

Infectious agents can have major impacts on animal populations through changing population dynamics and stability (Dobson & Hudson 1992), altering predator-prey interactions (Joly & Messier 2004), and even causing species' decline and extinction (De Castro & Bolker 2005; McCallum 2012). Accurately estimating the impact of these infectious agents in wildlife is critical to understanding what regulates host and parasite populations, making predictions about disease transmission, and managing disease outbreaks (Langwig *et al.* 2015). The impact of pathogens, such as rabies(), bovine TB(), and rinderpest(), is typically quantified based on the presence or absence of disease, and does not account for the number of infectious agents present. Although sufficient for many bacterial and viral agents that reproduce within a host, for macroparasites, hosts cannot be simply categorized as infected and uninfected because pathology is linked to the intensity of infection (Anderson & May 1979). Helminths exhibiting this intensity dependent pathology have significant impacts on human health (), domestic livestock economics (), wildlife survival (). While it is generally assumed that some fraction of

wild host populations must succumb to parasitic infections, it is notoriously difficult to actually quantify parasite-induced host mortality (PIHM) in wild animal populations (Lafferty & Kuris 2002; McCallum 2000).

Ideally, parasite-induced host mortality would be quantified by experimentally infecting and tracking individual hosts in the wild population; however, for logistical and ethical reasons this method is rarely feasible (McCallum 2000). Data on parasite intensity is much easier to collect and has often been used to identify the presence PIHM (Crofton 1971; Lester 1977, 1984; Lanciani & Boyett 1989; Royce & Rossignol 1990; Ferguson *et al.* 2011) and quantify the relationship between infection intensity and host mortality (Adjei *et al.* 1986) [others?].

Crofton (1971) first proposed that PIHM could be identified by comparing the observed parasite distribution in the host population to the distribution predicted in the absence of parasite-induced host mortality. This method ("Crofton Method") assumes that, prior to host mortality, parasites are distributed in the host population following a negative binomial distribution; however, as intensity dependent pathology removes heavily infected hosts from the population, the tail of the distribution is truncated. Mortality is assumed to not occur in hosts with low intensity infections, thus by iteratively fitting a negative binomial distribution to hosts with lower and lower parasite loads, and comparing the tail end of this predicted distribution to the observed parasite data, one could determine both whether PIHM was occurring and the parasite distribution in the host population prior to parasite induced mortality. We give a thorough description and implementation of the Crofton Method in *Supplementary Information (SI)* 1 and discuss the validity of its assumptions in the *Discussion*.

The Crofton Method may be able to detect the presence of PIHM however, quantifying the relationship between infection intensity and host survival probability is more complicated. Adjei *et al.* (1986) suggested that these values could be quantified by using the Crofton Method to first estimate the pre- mortality parasite distribution and then, using those parameters, calculate the probability of host survival with increasing

parasite load. To do this, Adjei *et al.* (1986) modeled host survival as a logistic function and then used a generalized linear model (GLM) to estimate the logistic parameters (see *SI 2* for a technical description of the Adjei Method). These methods appeared to provide an estimate the parasite intensity at which 50% of hosts exhibit PIHM (LD_{50}), as well as the unmeasurable fraction of the population that was lost (SI 2). However, to implement this method the observed data must be modified to fit the GLM framework and subjectively binned when mean infection intensity is high or sample sizes are small.

After 30 years, and despite clear limitations (McCallum 2000), these methods (particularly the Crofton Method) are still discussed among parasitologists and are the primary techniques for examining population level impacts of parasitism using parasite intensity data. In these methods, PIHM can only be identified by visually examining plots and, with no clear decision rule; it can be difficult to determine the significance of PIHM across different host-parasite systems. The survival function produced by the Adjei Method offers one solution; however, this method requires manipulating the original data and has never been thoroughly tested.

Estimating parasite impacts on host population using intensity data should only be done if methods exist to permit unbiased and accurate estimates of PIHM. In this study, we first propose a novel method for calculating PIHM. We next use simulations to compare our method with the previous Adjei Method to test the ability of both methods to (1) detect occurrence of PIHM and (2) estimate the lethal parasite load (LD_{50}) and the associated survival function. We then apply both methods to real datasets previously used in PIHM analyses and compare results. Finally, we discuss the limitations of inferring PIHM from distributional data and whether any method for inferring PIHM has a place in quantitative parasitology.

Methods

A novel, likelihood-based method for estimating PIHM

Here we propose a novel, likelihood-based method (henceforth Likelihood Method) that does not require binning or data alteration, reduces the number of parameters to be estimated, and allows the significance of PIHM to be determined using standard statistical techniques.

As with all previously proposed methods for estimating PIHM, the Likelihood Method first assumes that the pre-mortality distribution of parasites across hosts follows a negative binomial distribution $g(x; \mu_p, k_p)$, where μ_p is the mean parasite intensity in hosts before mortality and k_p is the parasite aggregation before mortality (smaller k_p leads to more aggregation). Previous methods have also required a parameter N_p specifying the total number of hosts before mortality (Crofton 1971; Adjei *et al.* 1986), but this is not a necessary parameter in the Likelihood Method.

The second assumption of the Likelihood Method is that the host survival function, the function specifying the probability of a host surviving with x parasites, takes the form of a logistic curve given by

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

For this equation, the parasite intensity at which 50% of host experience PIHM (LD_{50}) can be calculated by $\exp a/b$. With these two explicit assumptions, the Likelihood Method estimates four parameters: μ_p, k_p, a, and b.

To estimate these parameters, we first define a probability distribution that gives the probability of having a parasite load of x parasites conditional on host survival. Using 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 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 $P(\text{survival}) = \sum_{x=0}^{\infty} P(\text{survival}|x) * P(x) = \sum_{x=0}^{\infty} h(x; a, b) * g(x; \mu_p, k_p)$. Therefore equation 2 can be written as

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

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

To estimate the significance of PIHM in a host-parasite system, a likelihood ratio test can be used in which the full model is given by equation 3 and the reduced model is given by a negative binomial distribution. If PIHM is not significant in the system, the resulting likelihood ratio statistic should approximately follow a χ^2 distribution with degrees of freedom equal to 2. We provide the code for implementing this Likelihood Method in SI 3.

Evaluating the Adjei and Likelihood Methods

Question 1: Can we detect PIHM?

To test the ability of the Adjei and the Likelihood Methods 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 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 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, 50, 100) and for each μ_p we examined three different survival functions that had graduate, moderate, and steep decreases in the host survival with increasing parasite intensity (Figure **??**). For a given μ_p, each survival function had the same LD_{50} ([$\mu_p = 10$, $LD_{50} = 7.39$], [$\mu_p = 50$, $LD_{50} = 35.57$], [$\mu_p = 100$, $LD_{50} = 77.3$]), 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, and 1 — 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 correctly identifying PIHM in the post-mortality dataset (power). For each method, we used a 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 $\alpha = 0.05$. We tested each parameter combinations for pre-mortality population sizes of $N_p = [50, 100, 200, 300, 400, 500]$. N_p is not technically the sample size on which the methods are being tested on the post-mortality data because PIHM 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 in the power simulations.

Question 2: Can we estimate fatal parasite intensity and the host survival function?

To compare the ability of the Adjei Method and the Likelihood Method to recover the LD_{50} and the parameters a and b of 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]. We used the average number of surviving hosts over all 150 simulations for a given parameter combination as our measure of sample size. Because parameters a and b showed similar patterns of bias and precision, we only show the results for a.

Application to real data

We tested the ability of the Adjei Method and the Likelihood Method to identify PIHM in 6 host-parasite datasets given in Crofton (1971) and 4 datasets given in Adjei *et al.* (1986) (Table 2). In the Crofton (1971) datasets, the host was the snail *Gammarus pulex* which acts as the intermediate host for the acanthocephalan *Polmorphus minutus*. In the Adjei *et al.* (1986) datasets, the hosts were two species of lizard fish *Saurida tumbil* and *Saurida undosquamis* that were infected by the cestode *Callitetrarhynchus gracilis*. Males and females of both fish species were considered separately.

In both studies, the authors reported PIHM in some of the datasets and we test whether the Adjei Method and/or the Likelihood Method also predicted PIHM. For the 6 datasets from Crofton (1971), we truncated the data at 4 parasites, applied the Crofton Method to estimate the pre-mortality distribution, and then ran the Likelihood Method and Adjei Method using these pre-mortality parameters. For the Adjei *et al.* (1986) datasets, we followed the same procedure as the authors and first truncated the data at 2 parasites and then fit the Crofton Method for the female fish of both

species. Then, following the original authors' methods, we parameterized the male pre-mortality distributions for each species with the results from the females. Finally, we applied the Adjei Method and the Likelihood Method to determine whether or not PIHM was significant for these species and compared our results to those given by the authors. All fitting to data was done with the code provided in SI 3.

Results

Question 1: Detecting presence of PIHM

The Adjei Method showed highly inflated Type I error rates (i.e. falsely detected PIHM) for all parameter combinations that we considered (Figure **??**; SI2 Figs 1, 2). This method also showed the unintuitive pattern of Type I error rate decreasing as sample size decreased. This pattern was due to the issue of binning discussed in the *Introduction* and *SI X*. For small samples sizes, the applicability of the Adjei Method is compromised without binning the observed data in some way. In contrast, the Likelihood Method showed a Type I error rate at or near the pre-set level of 0.05 for all parameter combinations and sample sizes considered.

The ability of the Adjei Method to detect PIHM given that it was occurring in a system (i.e. the power) was close to 1 for larger sample sizes and tended to decrease as sample size decreased (Figure **??**; SI2 Figs 1, 2). The Likelihood Method had a power close to one for all parameter combinations and sample sizes considered. With gradual survival functions, the power of the Likelihood Method decreased slightly for small samples sizes (solid black lines; Figure **??**, SI2 Figs 1, 2).

Question 2: Estimating the LD_{50} and survival function

The Likelihood Method gave asymptotically unbiased estimates of the LD_{50} for all combinations of parameters examined in this study (Figure **??**, SI2 Fig 3, 4). Even for

small sample sizes (< 500 hosts), the Likelihood Method's estimate of LD_{50} was largely unbiased, with small biases occurring for host survival functions that were gradual. The precision of the Likelihood Method's LD_{50} estimates decreased (increasing coefficient of variation) as sample size decreased for all parameter combinations we examined (Figure **??**, SI2 Fig 3, 4).

The Adjei Method always produced biased estimates of the LD_{50} across all parameter combinations (Figure **??**, SI2 Fig 3, 4). For $\mu_p = 10$, the LD_{50} estimates from the Adjei Method were largely unbiased for large samples sizes, but showed increasing bias as sample size decreased, particularly for steep survival functions. As μ_p increased, the Adjei method produced biased estimates of LD_{50} across all sample sizes, with bias increasing as sample size decreased (SI2 Fig 3, 4). The LD_{50} estimates from the Adjei Method showed large decreases in precision occurring for the steepest survival function across all values of μ_p (Figure **??**, SI2 Fig 3, 4).

In terms of the host survival function, the Likelihood Method gave asymptotically unbiased estimates of a as sample size increased for all parameter combinations considered (Figure **??**, SI2 Fig 5, 6). However, as sample size decreased, the Likelihood Method tended to produce severely biased estimates of a. This was generally more pronounced for steeper survival functions and more aggregated pre-mortality distributions (e.g. Figure **??**). The Adjei Method produced biased estimates of a across all sample sizes, with the bias consistently being larger when the survival function was steeper. The bias of the Adjei Method's estimate of a also increased as μ_p increased (Figure **??**, SI2 Fig 5, 6).

Application to real data

Of the 10 datasets we considered, the previous authors visually detected PIHM in 7 of them (Table 2). The Likelihood Method parameterized from the pre-mortality parameters of the Crofton Method detected significant PIHM in 6 of these 7 datasets at a significance level of 0.05. The only dataset in which the Likelihood Method did not detect a significant

effect of PIHM was the Adjei dataset for female *S. tumbil*. For this dataset there was a marginally significant effect of PIHM ($\chi^2_{df=2} = 5.34; p = 0.069$).

The Adjei Method detected PIHM in 9 of the 10 datasets (Table 2), consistent with our simulation results that the Adjei Method has a high Type I error rate [additional data? to emphasize this point. Raccoon data?].

Discussion

Quantifying the impact of parasitism on wild host populations is critical in both disease modeling and wildlife management. Ideally the relationship between infection intensity and host survival would be measured experimentally, but for logistical and ethical reasons, this is often impossible (McCallum 2000). Looking for evidence of mortality in parasite distribution data requires the least amount of information, but is notoriously difficult to implement. The methodological flaws in the Adjei Method limit its utility, so here we propose an alternative, likelihood-based, method to estimate host survival and the LD_{50} from observed parasite intensity data. At a theoretical level this method is a significant improvement over the previous methods because it requires fewer parameters, provides a statistical decision rule for identifying PIHM and does not require any data manipulation.

Using simulated data, we found that the Likelihood Method always out performs the Adjei Method. For simply detecting the presence of PIHM, the Likelihood Method is both more powerful and has fewer false detection events (Type I errors) [tense?]. When both methods were applied to published datasets previously used in PIHM analyses, the Adjei Method tended to detect PIHM where it had not previously been reported, consistent with the high Type I error rate observed in our simulations. The Likelihood Method is also more precise and less biased in calculations of both the parasite LD_{50} and host survival curve over the parameter values we considered. However, while only the Likelihood Method produces precise and unbiased LD_{50} estimates, neither method can provide unbiased estimates of the host survival function at realistic sample sizes. These

simulations demonstrate that the Likelihood Method is more powerful and precise than the previously propose Adjei Method.

Although superior to the Adjei Method, the Likelihood Method may still not always be applicable to real data. The Likelihood Method requires relatively large sample sizes (n>50-100) [Are these relatively large? They seem small to me], that although reasonable to obtain for invertebrates or small fish may be completely unfeasible for many vertebrates, particularly the species of conservation concern where addressing the impact of parasitism would be most important [citation?]. An even larger sample size is required to capture the full parasite distribution when parasites are highly aggregated, mean infection intensity is high, or parasite prevalence is low [same as aggregation], all of which are common in many parasitic helminths. Low parasite-induced host mortality, as might be predicted in many definitive hosts, may also be very difficult to detect and require impossibly large sample sizes. And, even when sample size is sufficient, these methods can only detect PIHM is the host survival curve is non-linear (Lanciani & Boyett 1989). Most host-parasite models assume a linear relationship between survival and infection intensity (Anderson & May 1978; McCallum 2000), however nonlinear survival functions are not uncommon in empirical host-parasite systems (Benesh 2011). And, while linear functions make PIHM undetectable, at the other extreme, steep, non-linear survival curves produce severely biased estimates of the survival function. Even the Likelihood Method is probably limited to detecting PIHM and estimating LD_{50} in systems where greater than 50 hosts can be collected, parasites are common and only moderately aggregated, and substantial host mortality occurs at relatively low parasite intensity.

While we have improved on the existing methods for quantifying PIHM from parasite intensity data, all such methods require several fundamental, and potentially problematic assumptions. Nearly all current methods derive from Crofton (1971) (but see Ferguson *et al.* 2011) and assume that, prior to any PIHM, parasites are distributed in the host population following a negative binomial distribution. But, it is fundamentally impossible to know what the pre-mortality parasite distribution was in a wild host

population and it is widely recognized that different processes can lead to a variety of parasite distributions in hosts (Anderson & Gordon 1982; Duerr *et al.* 2003). However, the negative binomial is extremely flexible and there is substantial empirical and theoretical evidence to support the assumption that, prior to any PIHM, parasite distributions can be fit by a negative binomial distribution (Shaw & Dobson 1995; Shaw *et al.* 1998; Wilson *et al.* 2002). Unfortunately, this flexibility in the distribution may also reduce our ability to detect PIHM. If a negative binomial can be fit to the observed post-mortality parasite distribution then, regardless of how lethal the parasite was, it will be impossible to detect PIHM because there is no need for a more complex model. Most observed parasite distributions are well fit by the negative binomial distribution (Shaw *et al.* 1998), suggesting that systems where these methods are applicable may be more the exception than the rule. Finally, even when truncation of the negative binomial distribution is detected, it may be caused by other processes such as within host density dependence, age dependent variation in host resistance and/or heterogeneous infection rates (McCallum 2000; Anderson & Gordon 1982; Rousset *et al.* 1996). This means that in the event that PIHM is detected, it may actually not be the result of PIHM.

Given these numerous caveats, is the is there a place in parasitology for methods that estimate PIHM from distributional data? We are in agreement with Lester (1984) that, at the very least, methods for estimating PIHM can provide preliminary insight into whether or not PIHM is worth further exploration. However, we stress that these methods should only be used as an exploratory tool when assessing the role of PIHM in a system and potential users should critically evaluate whether they think they have a large enough sample sizes and an appropriate host survival function/post-mortality distribution for the methods developed in this paper to be applicable. Even if they are applicable, inferring PIHM from distributional data is no substitute for field or laboratory experiments and/or in depth understanding of the natural history of the host-parasite system under consideration.

Acknowledgments

We thank X for their useful comments. This work was supported by X.

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 Callitetrarhynchus gracilis (Cestoda: Trypanorhyncha) in lizardfish (Saurida spp.). *Parasitology*, 92, 227–243.

2.
Anderson, R. M. & Gordon, D. M. (1982). Processes influencing the distribution of parasite numbers within host populations with special emphasis on parasite-induced host mortalities. *Parasitology*, 85, 373–398.

3.
Anderson, R. M. & May, R. M. (1978). Regulation and stability of host-parasite interactions: I. Regulatory processes. *Journal of Animal Ecology*, 47, 219–247.

4.
Anderson, R. M. & May, R. M. (1979). Population biology of infectious diseases: Part I. *Nature*, 280, 361 – 367.

5.
Benesh, D. P. (2011). Intensity-dependent host mortality: what can it tell us about larval growth strategies in complex life cycle helminths? *Parasitology*, 138, 913–25. URL `http://www.ncbi.nlm.nih.gov/pubmed/21554844`.

6.
Crofton, H. D. (1971). A quantitative approach to parasitism. *Parasitology*, 62, 179–193.

7.
De Castro, F. & Bolker, B. (2005). Mechanisms of disease-induced extinction. *Ecology Letters*, 8, 117–126. URL `http://doi.wiley.com/10.1111/j.1461-0248.2004.00693.x`.

8.
Dobson, A. P. & Hudson, P. J. (1992). Regulation and stability of a free-living host-parasite system: Trichostrongylus tenuis in red grouse. II. Population models. *Journal of Animal Ecology*, 61, 487–498.

9.
Duerr, H. P., Dietz, K. & Eichner, M. (2003). On the interpretation of age–intensity profiles and dispersion patterns in parasitological surveys. *Parasitology*, 126, 87–101.

10.
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.

11.
Joly, D. O. & Messier, F. (2004). The distribution of Echinococcus granulosus in moose: Evidence for parasite-induced vulnerability to predation by wolves? *Oecologia*, 140, 586–590.

12.
Lafferty, K. D. & Kuris, A. M. (2002). Trophic strategies, animal diversity and body size. *Trends in Ecology and Evolution*, 17, 507–513.

13.
Lanciani, C. A. & Boyett, J. M. (1989). Demonstrating parasitic water mite-induced mortality in natural host populations. *Parasitology*, 81, 465–475.

14.
Langwig, K. E., Voyles, J., Wilber, M. Q., Frick, W. F., Murray, K. a., Bolker, B. M., Collins, J. P., Cheng, T. L., Fisher, M. C., Hoyt, J. R., Lindner, D. L., McCallum, H. I., Puschendorf, R., Rosenblum, E. B., Toothman, M., Willis, C. K., Briggs, C. J. & Kilpatrick, a. M. (2015). Context-dependent conservation responses to emerging wildlife diseases. *Frontiers in Ecology and the Environment*, 13, 195–202. URL http://www.esajournals.org/doi/10.1890/140241.

15.
Lester, R. J. G. (1977). An estimate of mortality in a population of Perca flavescens owing to the trematode Diplostomum adamsi. *Canadian Journal of Zoology*, 55, 288–292.

16.
Lester, R. J. G. (1984). A review of methods for estimating mortality due to parasites in wild fish populations. *Helgoländer Meeresuntersuchungen*, 37, 53–64. URL http://link.springer.com/10.1007/BF01989295.

17.
McCallum, H. (2012). Disease and the dynamics of extinction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 367, 2828–39. URL http://www.ncbi.nlm.nih.gov/pubmed/22966138.

18.
McCallum, H. I. (2000). Host-pathogen and host-parasite models. In: *Population Parameters: Estimation for Ecological Models* (eds. Lawton, J. H. & Likens, G. E.), chap. Chapter 10. Blackwell Science Ltd., pp. 284–312.

19.
Rousset, F., Thomas, F., Meeûs, T. D. & Renaud, F. (1996). Inference of parasite-induced host mortality from distributions of parasite loads. *Ecology*, 77, 2203–2211.

20.
Royce, L. A. & Rossignol, P. (1990). Epidemiology of honey bee parasites. *Parasitology Today*, 6, 348–353.

21.
Shaw, D. J. & Dobson, A. P. (1995). Patterns of macroparasite abundance and aggregation in wildlife populations: a quantitative review. *Parasitology*, 111 Suppl, S111–27.

22.
Shaw, D. J., Grenfell, B. T. & Dobson, a. P. (1998). Patterns of macroparasite aggregation in wildlife host populations. *Parasitology*, 117 (Pt 6, 597–610.

23.
Walther, B. a. & Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, 28, 815–829.

24.
Wilson, K., Bjø rnstad, O. N., Dobson, A. P., Merler, S., Poglayen, G., Read, A. F. & Skorping, A. (2002). Heterogeneities in macroparasite infections: patterns and processes. In: *The Ecology of Wildlife Diseases* (eds. Hudson, P. J., Rizzoli, A., Grenfell, B., Heesterbeek, H. & Dobson, A.), chap. 2. Oxford University Press, Oxford, pp. 6–44.

Table 1: Definition of parameters and functions used in the main text

Parameter	Definition
μ_p	Pre-mortality mean parasite intensity
k_p	Pre-mortality parasite aggregation
N_p	Pre-mortality host population size
x	Number of parasites in a given host
$g(x; \mu_p, k_p)$	Pre-mortality negative binomial distribution parasites distribution
a	Parameter of the logistic host survival function
b	Parameter of the logistic host survival function
$h(x; a, b)$	The host survival function
LD_{50}	$\exp(a/b)$, Parasite intensity at which 50% of hosts die

Table 2: Comparison of the PIHM predictions of previously used host-parasite datasets to those given by the Adjei Method and the Likelihood Method. The first column specifies the identity of the dataset, the second column specifies whether or not the author indicated that PIHM was occurring in the system, the third column indicates whether or not the Likelihood Method with pre-mortality parameters estimated from the Crofton Method detects significant PIHM, and the final column indicates whether the Adjei Method with pre-mortality parameters estimated from the Crofton Method detects PIHM. If a method detected significant PIHM the predicted LD_{50} is given in parentheses

Data Set (sample size)	Author detected PIHM?	Likelihood Method?	Adjei Method?
Crofton, Station 1 ($n = 538$)	Yes	Yes (7.27)	Yes (9.33)
Crofton, Station 2 ($n = 507$)	Yes	Yes (6.92)	Yes (14.95)
Crofton, Station 3 ($n = 633$)	Yes	Yes (5.93)	Yes (5.98)
Crofton, Station 4 ($n = 486$)	No	No	Yes (7.99)
Crofton, Station 5 ($n = 276$)	No	No	Yes (10.58)
Crofton, Station 6 ($n = 191$)	No	No	No
Adjei, *S. tumbil* female ($n = 446$)	Yes (5.7)	No	Yes (6.37)
Adjei, *S. tumbil* male ($n = 452$)	Yes (3.4)	Yes (3.42)	Yes (3.66)
Adjei, *S. undosquamis* female ($n = 2573$)	Yes (3.2)	Yes (3.04)	Yes (3.11)
Adjei, *S. undosquamis* male ($n = 2440$)	Yes (1.8)	Yes (1.83)	Yes (1.78)

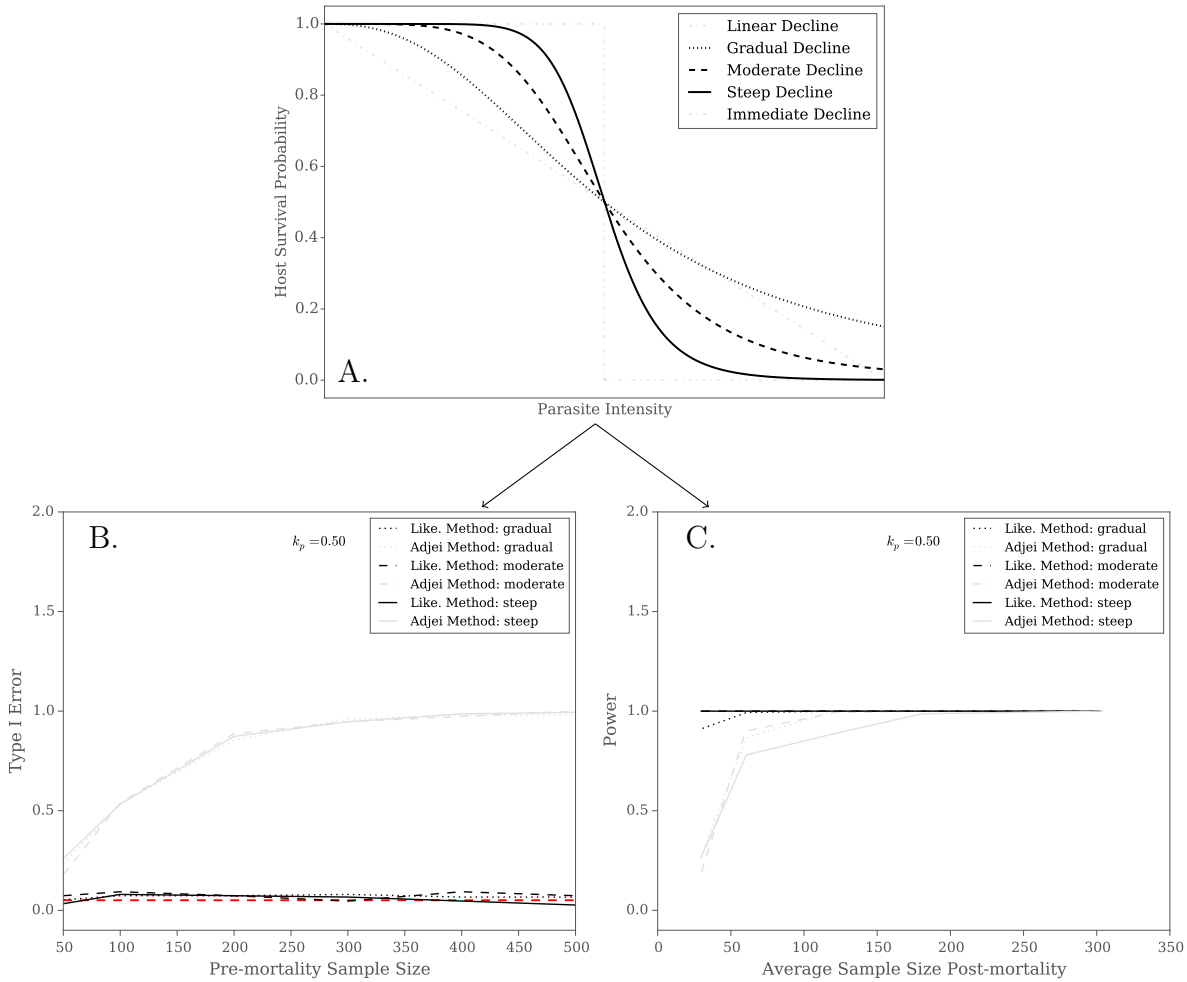

Figure 1: A) Five potential shapes for a host-survival functions. In the simulations we used a gradual survival function (dotted line), and moderate survival function (dashed line), and a steep survival function (solid line). The linear and immediate survival functions represent two potential extremes. For each of these survival functions and the parameter combinations described in the main text, we tested the Type I Error and Power of the Likelihood Method and Adjei Method. B) gives the Type I Error of each method over a range of pre-mortality sample sizes with $\mu_p = 50$ and $k_p = 0.5$. The red line shows the pre-set significance level of 0.05. C) gives the Power of each method for detecting PIHM over a range of post-mortality sample sizes for $\mu_p = 50$ and $k_p = 0.5$. In general, the Likelihood Method has higher power and lower Type I Error than the Adjei Method. See the SI X Fig X - X for Type I Error and Power results for all parameter combinations.

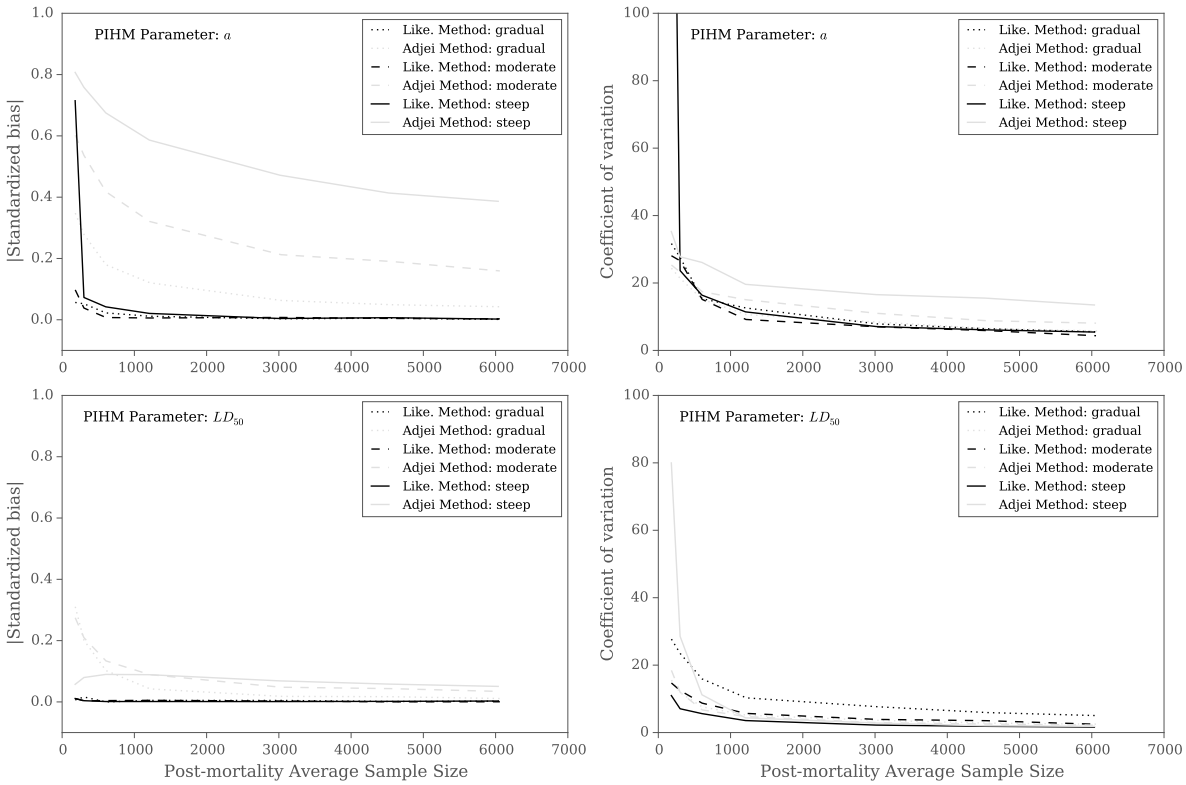

Figure 2: Bias and precision (coefficient of variation) for the Likelihood Method and Adjei Method estimates of the a parameter and the LD_{50} of the host survival function based on simulated PIHM data over a range of post-mortality sample sizes. The pre-mortality parameters for this simulation were $\mu_p = 50$ and $k_p = 0.5$. The figure shows the simulations for three different host survival functions (gradual, moderate, and steep) with the same LD_{50}. Bias and precision results for all other parameter combinations can be found in Fig X - X in SI X.