

# 1 Supporting Information

## 2 Appendix 1: Implementation of the Crofton Method

3 The algorithm for fitting the Crofton Method (Crofton 1971) proceeds as follows.  
4 First, obtain a dataset with  $n$  hosts where each host has some parasite intensity 0  
5 to  $p_{max}$ . Starting with the full dataset, guess a vector of pre-mortality parameters  
6  $(N_p, \mu_p, k_p)$  and given these parameters calculate the predicted number of hosts  
7 with  $0, 1, 2, \dots p_{max}$ . Compare the expected number of hosts with  $0, 1, 2, \dots p_{max}$   
8 parasites to the observed number hosts with  $0, 1, 2, \dots p_{max}$  parasites and calculate  
9 the  $\chi^2$ -squared statistic associated with your observed and predicted vectors.  
10 In reality, one often has to bin the parasite intensity data because all parasite  
11 intensities are not represented in the dataset. Continue to guess  $(N_p, \mu_p, k_p)$   
12 vectors until a set of parameters is found that minimizes the  $\chi^2$ -squared statistic.

13 Second, choose a truncation value  $(t_1)$  such that  $t_1 < p_{max}$ . Truncate the  
14 data such that  $\text{data}_{\text{truncated}} \leq t$  and repeat the above iterative procedure to  
15 calculate another set of parameters  $(N_{p2}, \mu_{p2}, k_{p2})$  that minimizes the  $\chi^2$ -squared  
16 statistic on the truncated data. Choose a new truncated value  $t_2 < t_1$  and repeat  
17 the first two steps. Continue to truncated the dataset until it only contains hosts  
18 with 0, 1, and 2 parasites (or 3 bins). As the method attempts to estimate three  
19 parameters, at least 3 classes are needed if all 3 parameters are to be identifiable  
20 (Royce & Rossignol 1990).

21 We provide an implementation and unit tests of the Crofton Method  
22 Supplementary Information 4. Figure 10 visually shows that our implementation  
23 of the Crofton Method agrees with results previously published by Crofton (1971).

## 24 Appendix 2: Implementation of the Adjei Method

25 The Adjei Method for estimating PIHM has two steps (Adjei *et al.* 1986). The first  
26 step is to estimate the parameters of the pre-mortality host-parasite distribution  
27 using the Crofton Method. The three parameters estimated are the total number of  
28 hosts before mortality  $N_p$ , the mean number of parasites per host before mortality  
29  $\mu_p$ , and the aggregation of parasites before mortality given by the parameter  $k_p$   
30 from a negative binomial distribution. When  $k_p$  is small, parasites are highly  
31 aggregated among hosts and when  $k_p$  is large parasites are more evenly spread  
32 out (Wilson *et al.* 2002). The implementation of the Crofton Method has been  
33 discussed at length elsewhere (e.g. Royce & Rossignol 1990; Lester 1984, and  
34 above) and we provide a tested implementation of the method in SI 4.

35 The second step of the Adjei Method is to estimate the shape of the host  
36 survival function. Adjei *et al.* (1986) assume that the host survival function follows  
37 the logistic form

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

38 where  $x$  is the parasite intensity in a given host and  $a$  and  $b$  are the two parameters  
39 of the logistic function. Generally, a larger  $a$  allows for hosts to tolerate larger  
40 parasite intensities before experiencing parasite-induced mortality and a larger  $b$   
41 leads to a more rapid decline in the probability of host survival as parasite intensity  
42 increases. The value  $\exp(a/b)$  is referred to as the  $LD_{50}$ , which gives the parasite  
43 intensity at which 50% of hosts experience mortality.

44 To estimate this function, the Adjei Method first calculates the expected  
45 number of hosts with a given parasite load  $x$  by using the equation  $g(x; \mu_p, k_p) * N_p$ ,  
46 where  $g(x; \cdot)$  is the negative binomial pre-mortality distribution. Second, the  
47 observed and predicted number of hosts with  $x$  parasites are paired as a single  
48 data point and the method then assumes that this data point follows a binomial

49 distribution with the total number of “trials” equal to the predicted number of  
50 hosts and the total number of “successes” equal to the observed number of hosts.  
51 In some cases, the observed number of hosts is greater than the expected number  
52 of hosts and the Adjei Method alters the data so that the observed is equal to the  
53 predicted (Adjei *et al.* 1986). After this questionable manipulation, the (observed,  
54 predicted) pairs are fit to a standard Generalized Linear Model (McCullagh &  
55 Nelder 1989) with a binomial response variable and a logistic link function given  
56 by equation 1. This model provides estimates for parameters  $a$ ,  $b$  and  $LD_{50}$ .

57 While not included in the original implementation of the Adjei Method,  
58 a  $\chi^2$  test with a degrees of freedom of 1 can be used to assess whether a GLM  
59 model that includes parasite intensity as a predictor of host survival probability is  
60 a “better” model than a GLM without this predictor. This allows the Adjei Method  
61 to determine whether PIHM is a significant factor in a host-parasite system.

62 The Adjei Method’s most glaring deficiency is the need to alter the observed  
63 data in order to fit the model into the binomial GLM framework. A second more  
64 subtle problem with the Adjei Method is the potential need to bin data in order  
65 to predict greater than one host in a given parasite intensity class. For example,  
66 if the total number of hosts pre-mortality was 50, the mean number of parasites  
67 per host pre-mortality was 100 and the aggregation parameter was 1, applying  
68 the equation  $g(x; \mu_p = 100, k_p = 1) * 50$  would result in less than 1 individual in  
69 all parasite intensities  $x$ . In other words, the Adjei Method cannot be applied to  
70 samples with either very high mean parasite loads, small sample sizes, or both  
71 without some sort of binning of the data. While this is not a flaw *per se*, it does  
72 add a certain level of subjectivity (i.e. which bins should you use?) to a method  
73 that already has serious potential issues. In this analysis, we always assume the  
74 Adjei Method is not binning the data, though we provide code for applying the  
75 binning method in SI 4.

## 76 Appendix 3: Additional Figures

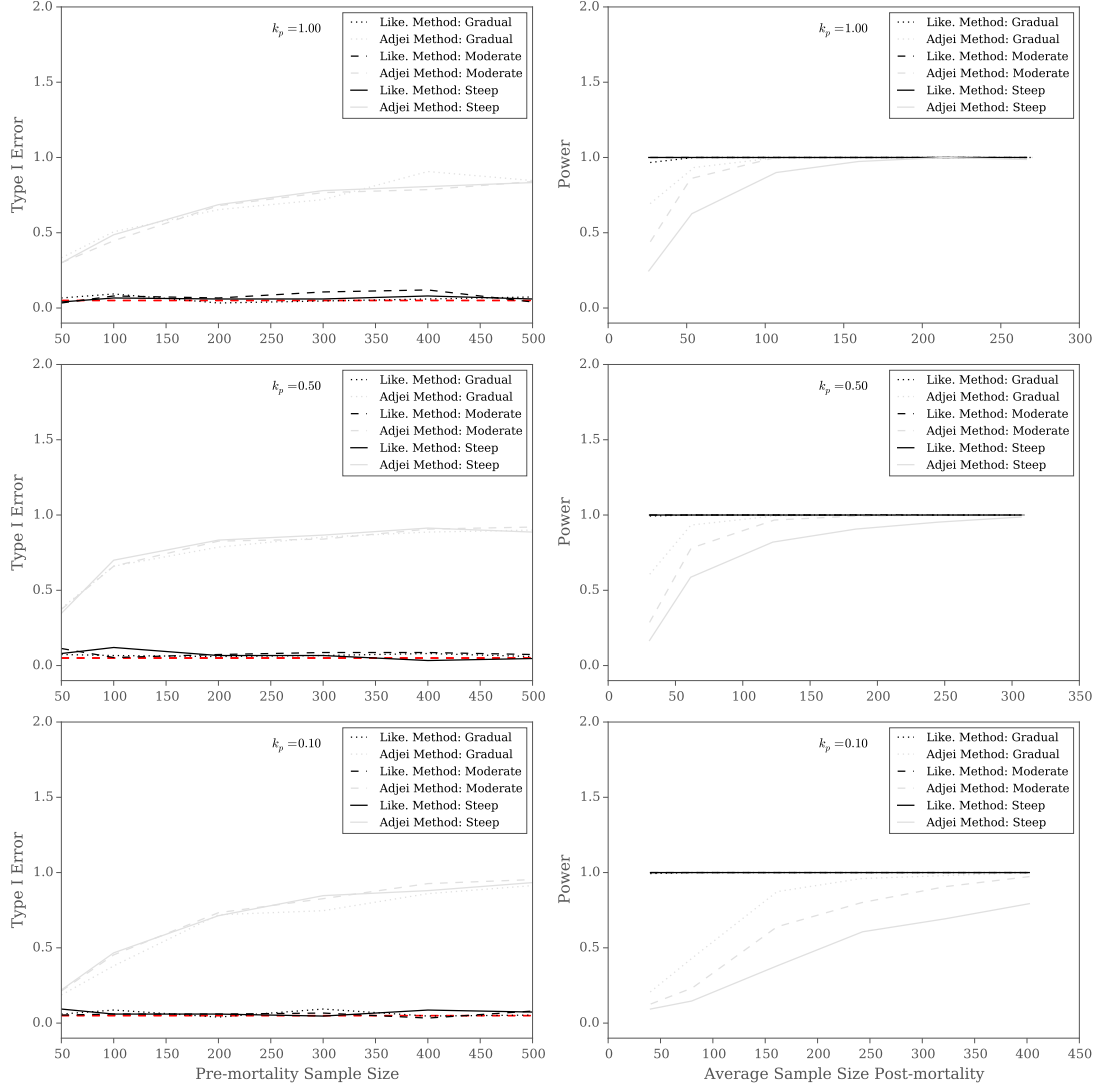
77 See Figures 1 - 9.

## 78 Appendix 4: Code and unit tests for estimating parasite- 79 induced host mortality

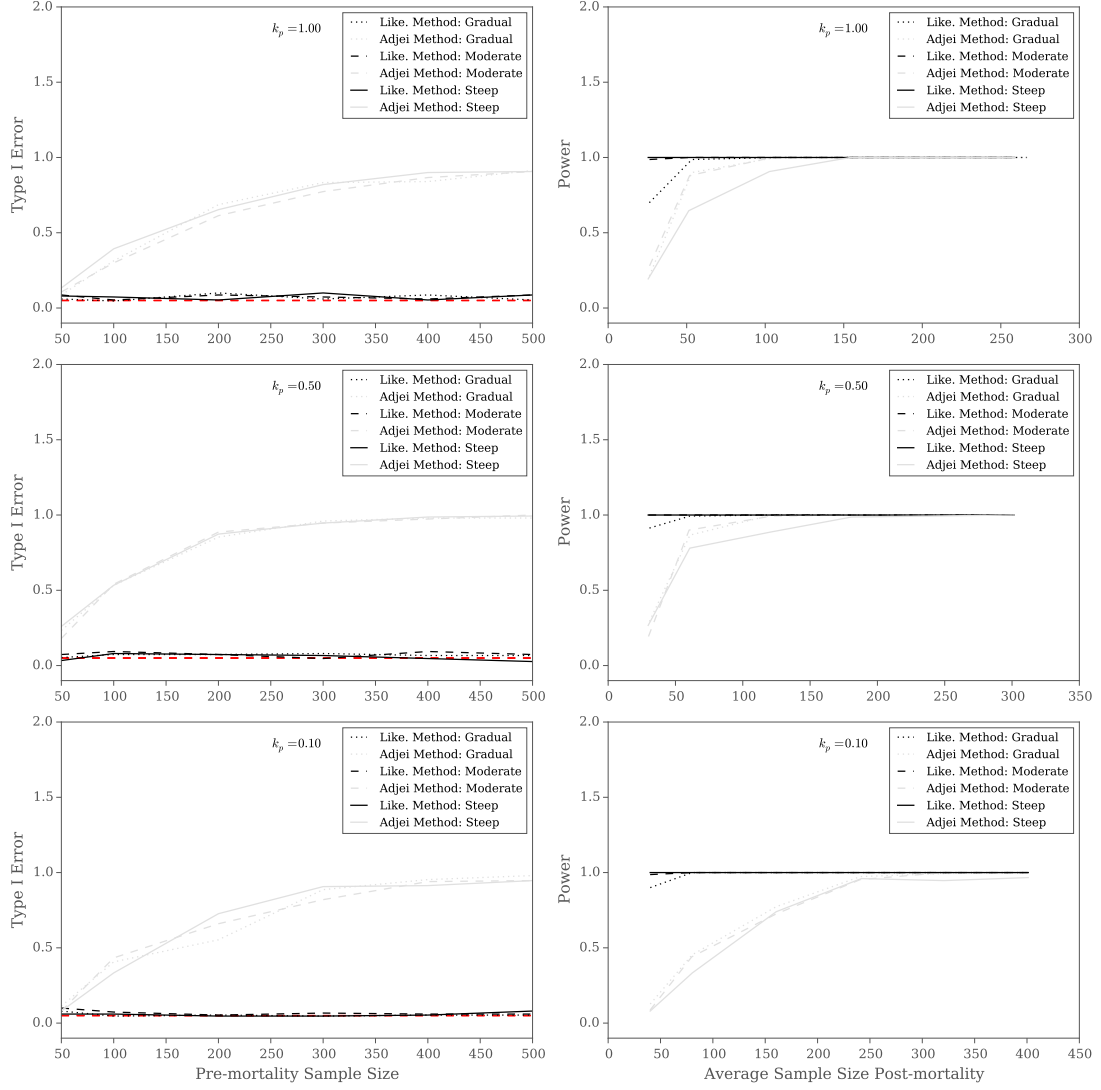
80 Python code, unit tests, and a help file for the Crofton Method, the Ad-  
81 jei Method and the Likelihood Method can be found in `pihm_methods.py`,  
82 `test_pihm_methods.py`, and `help_file.txt`, respectively.

## 83 References

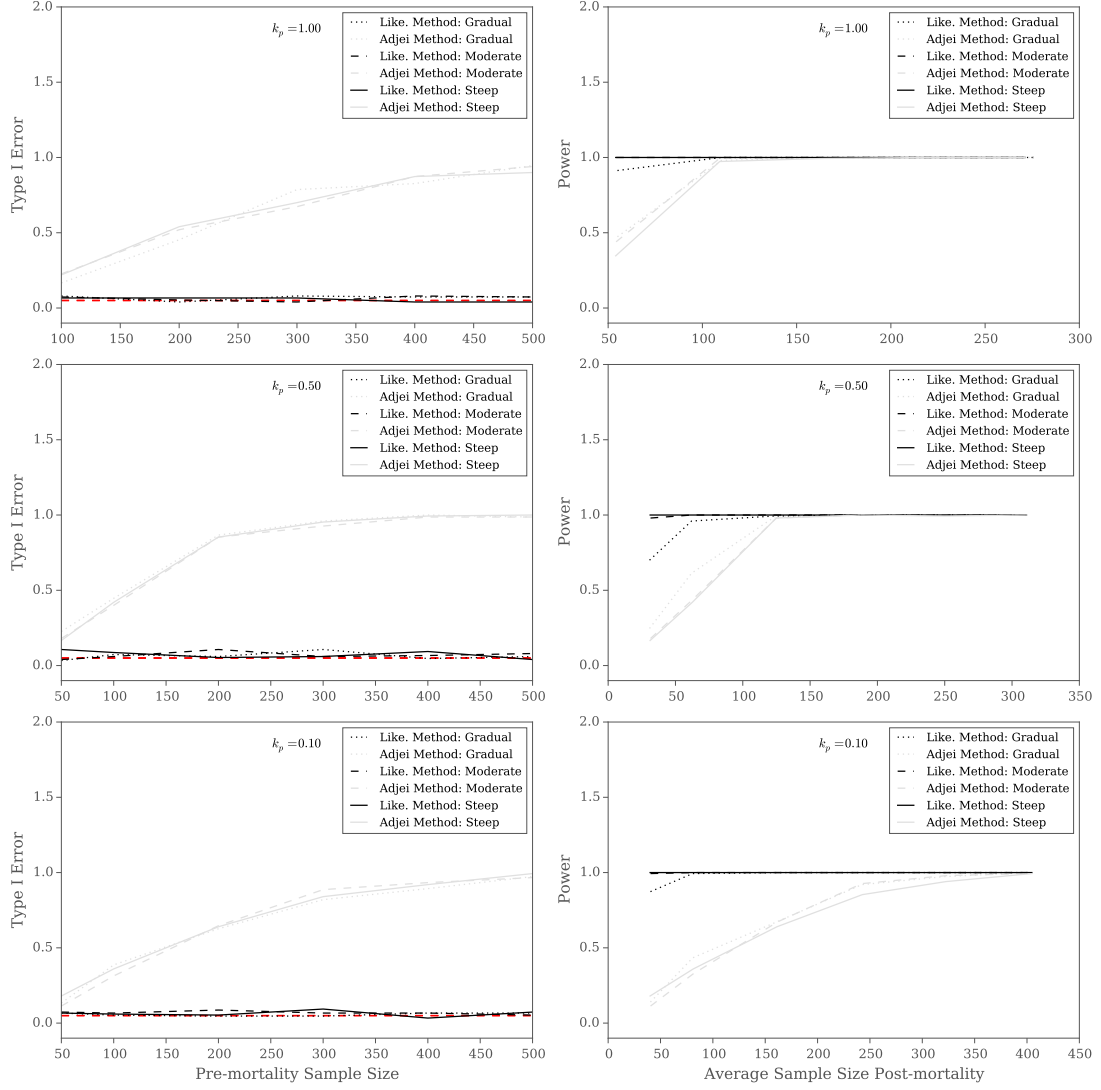
- 84 1.  
85 Adjei, E. L., Barnes, A. & Lester, R. J. G. (1986). A method for estimating possible  
86 parasite-related host mortality, illustrated using data from *Callitetrarhynchus*  
87 *gracilis* (Cestoda: Trypanorhyncha) in lizardfish (Saurida spp.). *Parasitology*, 92,  
88 227–243.
- 89 2.  
90 Crofton, H. D. (1971). A quantitative approach to parasitism. *Parasitology*, 62,  
91 179–193.
- 92 3.  
93 Lester, R. J. G. (1984). A review of methods for estimating mortality due to  
94 parasites in wild fish populations. *Helgoländer Meeresuntersuchungen*, 37, 53–64.  
95 URL <http://link.springer.com/10.1007/BF01989295>.
- 96 4.  
97 McCullagh, P. & Nelder, J. A. (1989). *Generalized Linear Models*. 2nd edn.  
98 Chapman & Hall, New York.
- 99 5.  
100 Royce, L. A. & Rossignol, P. (1990). Epidemiology of honey bee parasites.  
101 *Parasitology Today*, 6, 348–353.
- 102 6.  
103 Wilson, K., Bjørnstad, O. N., Dobson, A. P., Merler, S., Pogliayen, G., Read, A. F.  
104 & Skorpington, A. (2002). Heterogeneities in macroparasite infections: patterns and  
105 processes. In: *The Ecology of Wildlife Diseases* (eds. Hudson, P. J., Rizzoli, A.,  
106 Grenfell, B., Heesterbeek, H. & Dobson, A.), chap. 2. Oxford University Press,  
107 Oxford, pp. 6–44.



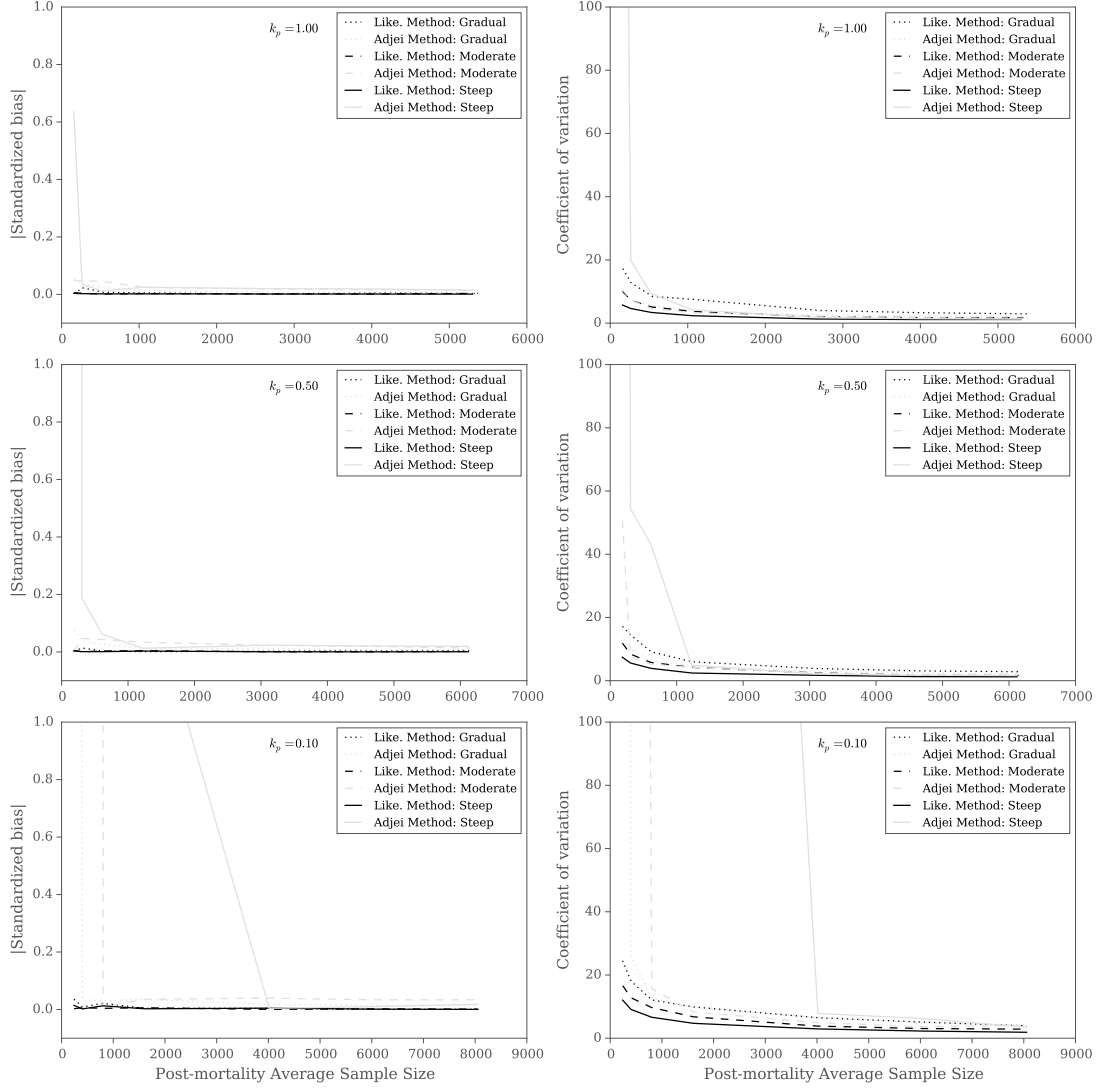
**Figure 1:** The type I error rate and the power of the Likelihood Method (solid line) and the Adjei Method (dashed lines) when  $\mu_p = 10$  for various shapes of the host survival function and levels of aggregation  $k_p$ . The first column gives the type I error rate of each method for falsely detecting PIHM when none is present. The red line gives the the pre-set type I error rate of  $\alpha = 0.05$ . The second column gives the power of a given method to detect PIHM when it is actually occurring.



**Figure 2:** The type I error rate and the power of the Likelihood Method (solid line) and the Adjei Method (dashed lines) when  $\mu_p = 50$  for various shapes of the host survival function and levels of aggregation  $k_p$ . The first column gives the type I error rate of each method for falsely detecting PIHM when none is present. The red line gives the the pre-set type I error rate of  $\alpha = 0.05$ . The second column gives the power of a given method to detect PIHM when it is actually occurring.

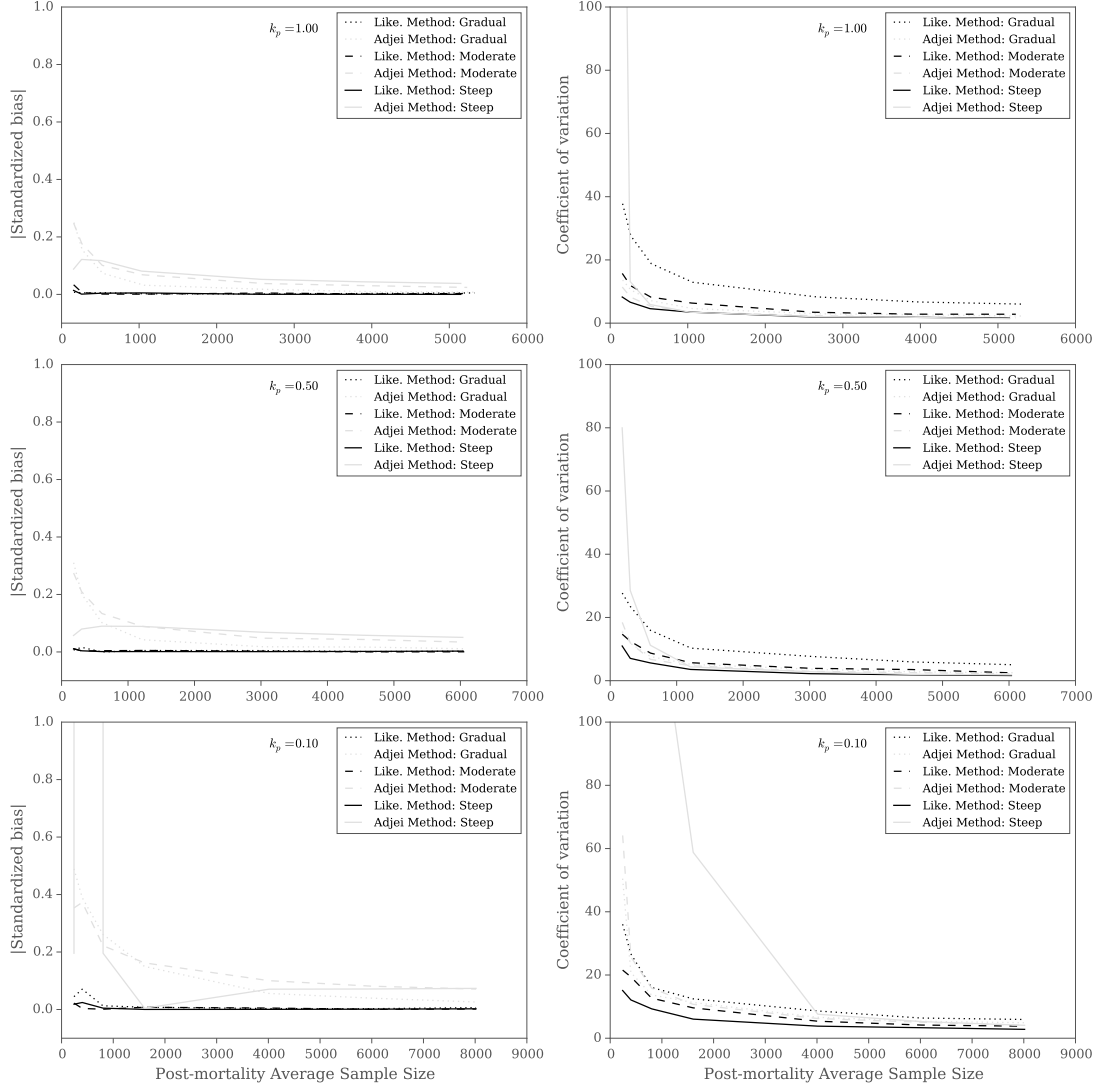


**Figure 3:** The type I error rate and the power of the Likelihood Method (solid line) and the Adjei Method (dashed lines) when  $\mu_p = 100$  for various shapes of the host survival function and levels of aggregation  $k_p$ . The first column gives the type I error rate of each method for falsely detecting PIHM when none is present. The red line gives the the pre-set type I error rate of  $\alpha = 0.05$ . The second column gives the power of a given method to detect PIHM when it is actually occurring.

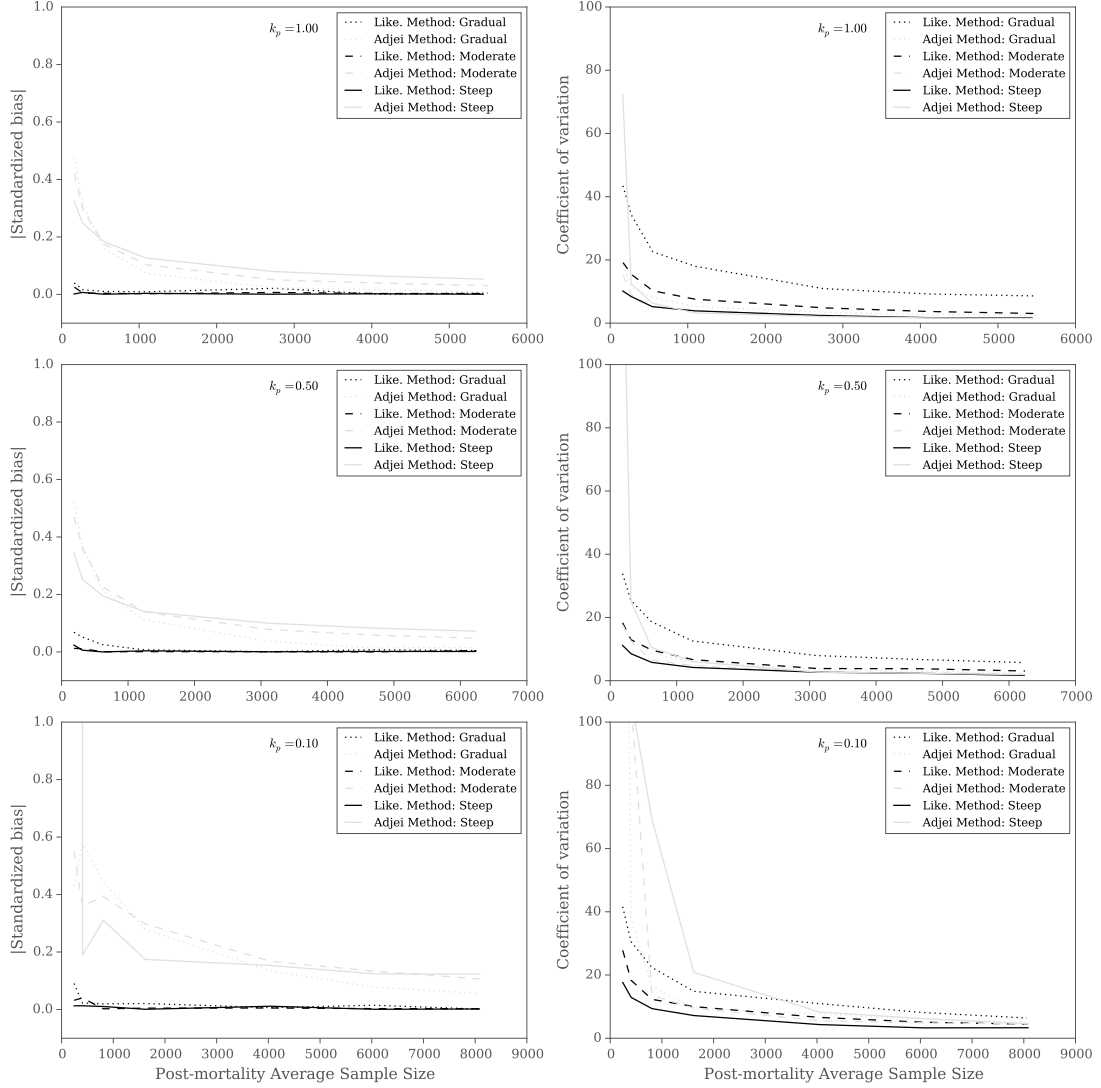


**Figure 4:** The bias and the precision of the Likelihood Method (solid line) and the Adjei Method (dashed lines) when  $\mu_p = 10$  for various shapes of the host survival function and levels of aggregation  $k_p$  when estimating  $LD_{50}$ . The first column gives the bias of each method's  $LD_{50}$  estimate over 150 simulations. The second column gives the precision of each method's  $LD_{50}$  estimate over 150 simulations.

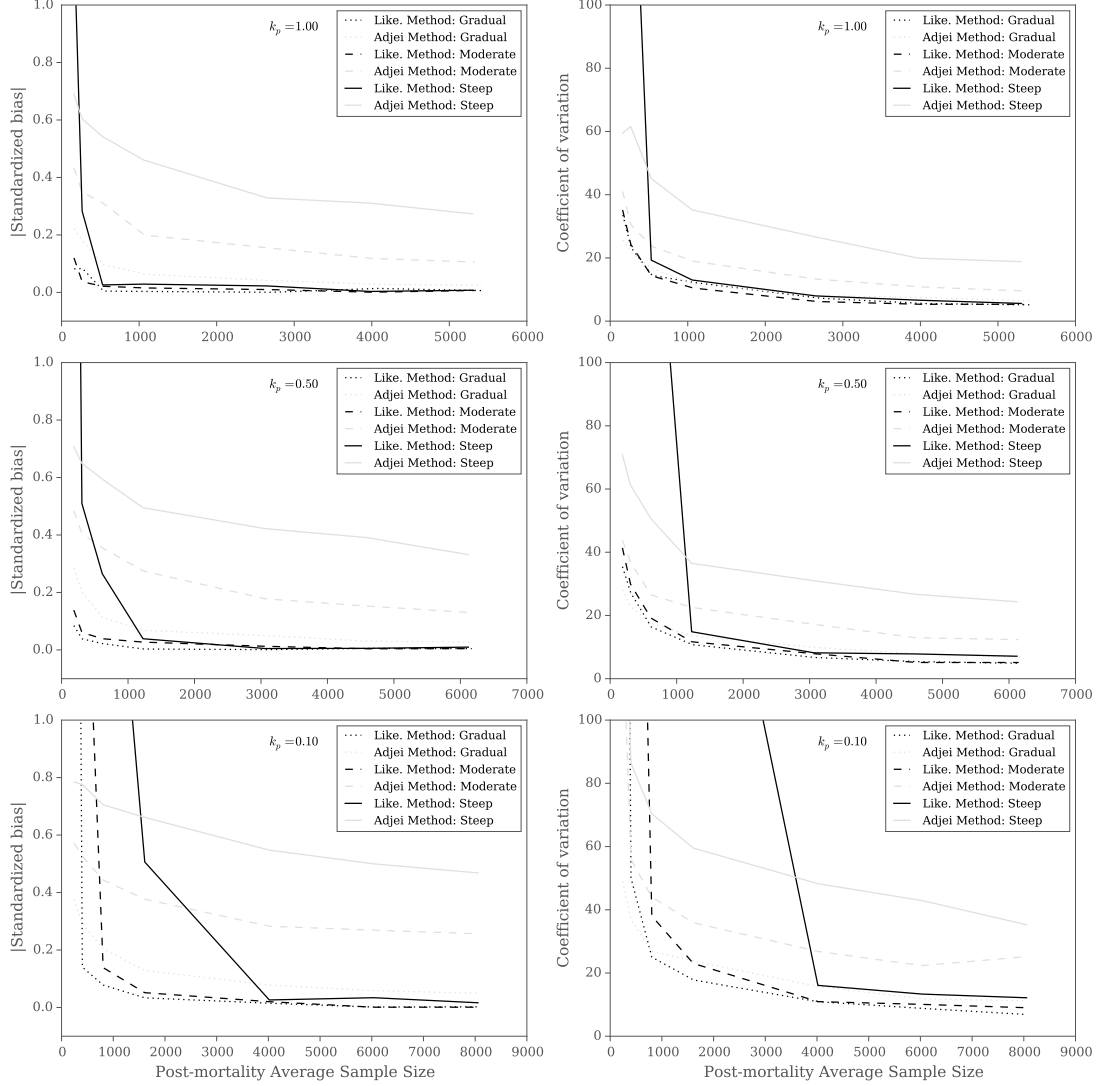




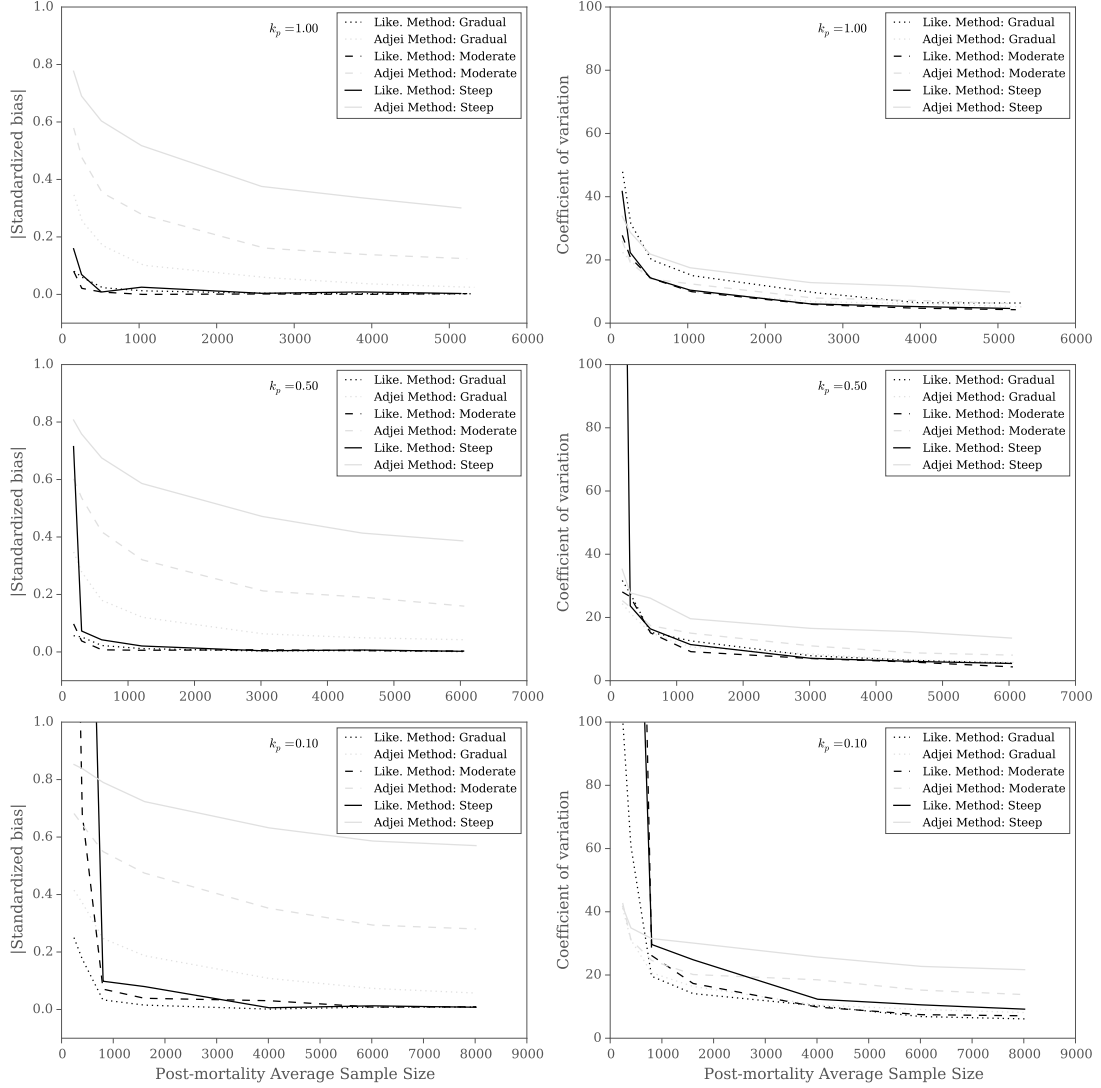
**Figure 5:** The bias and the precision of the Likelihood Method (solid line) and the Adjei Method (dashed lines) when  $\mu_p = 50$  for various shapes of the host survival function and levels of aggregation  $k_p$  when estimating  $LD_{50}$ . The first column gives the bias of each method's  $LD_{50}$  estimate over 150 simulations. The second column gives the precision of each method's  $LD_{50}$  estimate over 150 simulations.



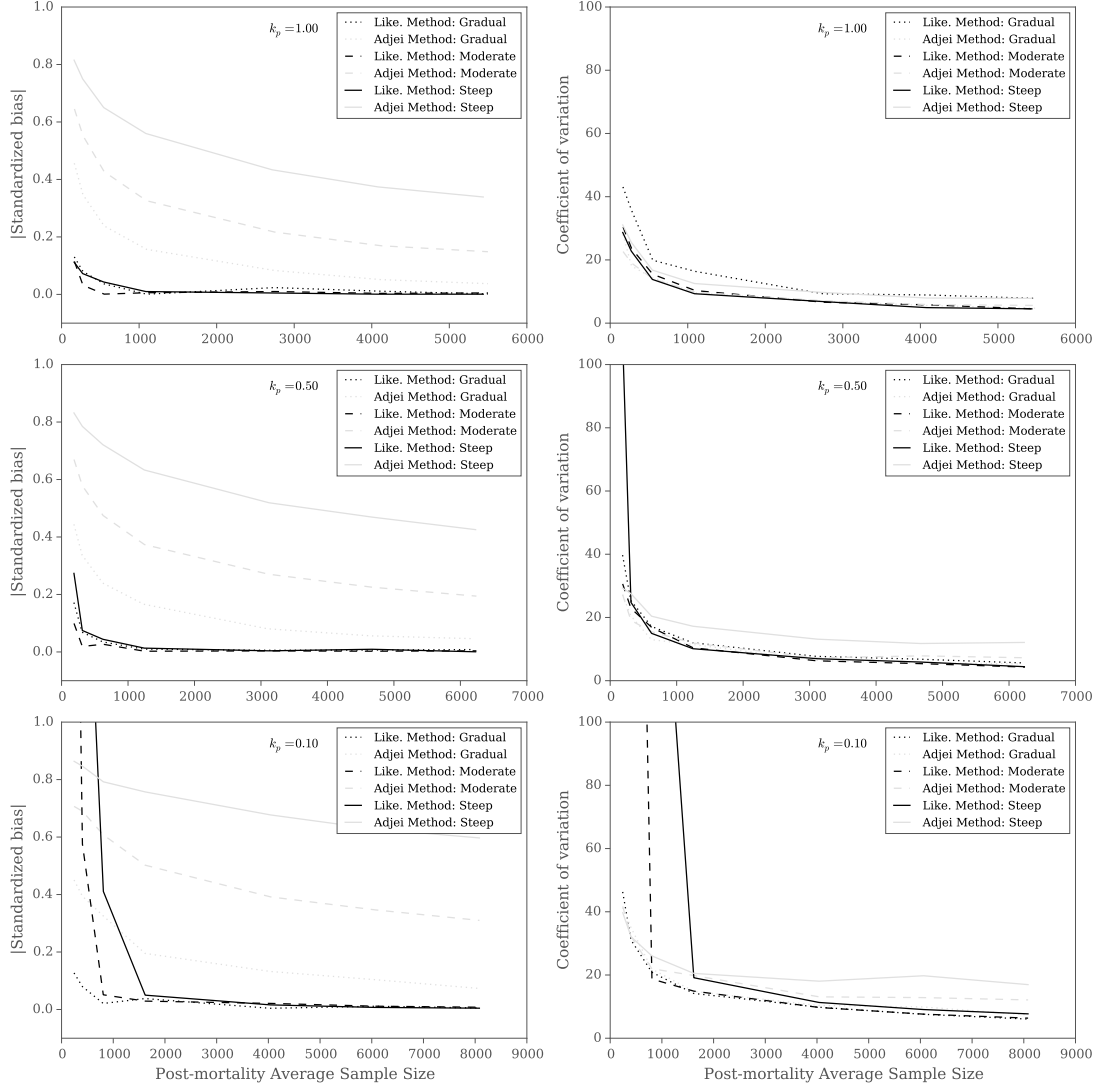
**Figure 6:** The bias and the precision of the Likelihood Method (solid line) and the Adjei Method (dashed lines) when  $\mu_p = 100$  for various shapes of the host survival function and levels of aggregation  $k_p$  when estimating  $LD_{50}$ . The first column gives the bias of each method's  $LD_{50}$  estimate over 150 simulations. The second column gives the precision of each method's  $LD_{50}$  estimate over 150 simulations.



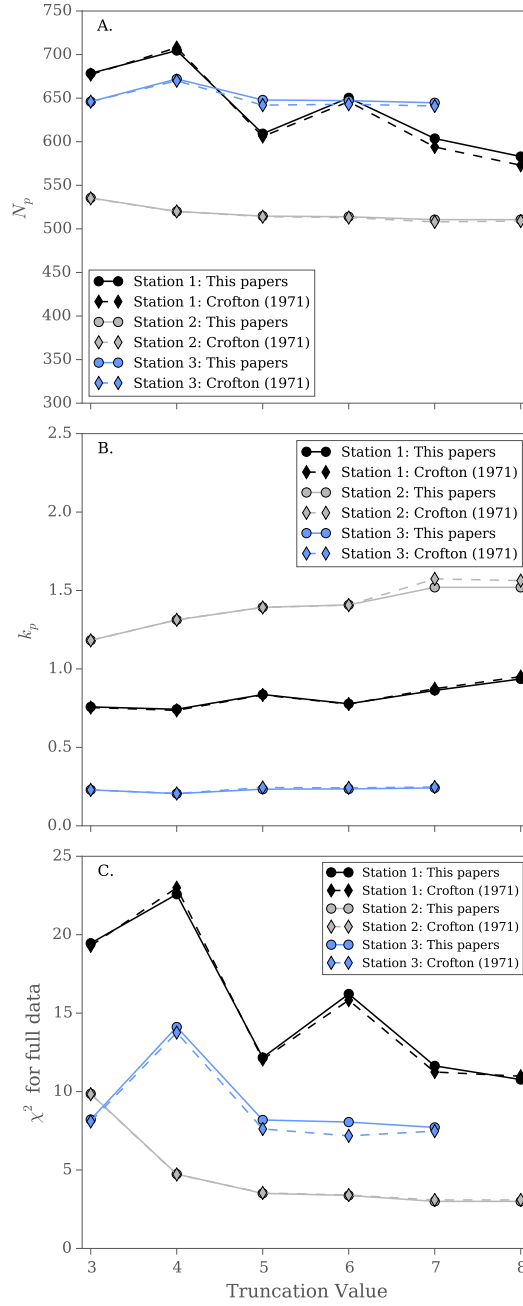
**Figure 7:** The bias and the precision of the Likelihood Method (solid line) and the Adjei Method (dashed lines) when  $\mu_p = 10$  for various shapes of the host survival function and levels of aggregation  $k_p$  when estimating the  $a$  parameter of the host survival function. The first column gives the bias of each method's  $a$  estimate over 150 simulations. The second column gives the precision of each method's  $a$  estimate over 150 simulations.



**Figure 8:** The bias and the precision of the Likelihood Method (solid line) and the Adjei Method (dashed lines) when  $\mu_p = 50$  for various shapes of the host survival function and levels of aggregation  $k_p$  when estimating the  $a$  parameter of the host survival function. The first column gives the bias of each method's  $a$  estimate over 150 simulations. The second column gives the precision of each method's  $a$  estimate over 150 simulations.



**Figure 9:** The bias and the precision of the Likelihood Method (solid line) and the Adjei Method (dashed lines) when  $\mu_p = 100$  for various shapes of the host survival function and levels of aggregation  $k_p$  when estimating the  $a$  parameter of the host survival function. The first column gives the bias of each method's  $a$  estimate over 150 simulations. The second column gives the precision of each method's  $a$  estimate over 150 simulations.



**Figure 10:** A comparison of this papers implementation (solid line, circles) of the Crofton Method with the results given in Crofton (1971) (dashed line, diamonds). Figure A compares the predicted number of hosts in a population pre-mortality ( $N_p$ ). Figure B compares the predicted parasite aggregation pre-mortality ( $k_p$ ). Figure C compares the  $\chi^2$  statistic for each implementation. Three of the 6 stations fit by Crofton are shown here and all show that our implementation gives very similar results to those given by Crofton.