Research

Density-dependent effects on the weight of female *Ascaris lumbricoides* infections of humans and its impact on patterns of egg production

Martin Walker*[1], Andrew Hall[2], Roy M Anderson[1] and María-Gloria Basáñez[1]

Address: [1]Department of Infectious Disease Epidemiology, Faculty of Medicine, St Mary's Campus, Imperial College London, Norfolk Place, London, W2 1PG, UK and [2]Centre for Public Health Nutrition, School of Integrated Health, University of Westminster, 115 New Cavendish Street, London, W1W 6UW, UK

Email: Martin Walker* - m.walker06@imperial.ac.uk; Andrew Hall - a.hall04@westminster.ac.uk; Roy M Anderson - roy.anderson@imperial.ac.uk; María-Gloria Basáñez - m.basanez@imperial.ac.uk

* Corresponding author

Abstract

Background: *Ascaris lumbricoides* exhibits density-dependent egg production, a process which has a marked impact on both the transmission dynamics and the stability of the parasite population. Evidence suggests that the egg production of female *Ascaris* is also associated with the size of the worm. If worm size is mediated by density-dependent processes then the size of female worms may have a causal impact upon patterns of *Ascaris* egg production.

Results: We analyse data collected from a cohort of human hosts, and demonstrate that the per host mean weight (a proxy for size) of female *Ascaris* is dependent on the number of infecting females (worm burden) following a pattern of initial facilitation followed by limitation. Applying a negative binomial (NB) generalized linear model (GLM) and a zero-inflated negative binomial (ZINB) model we confirm that the per host female mean weight is significantly associated with per host egg production. Despite these associations, the mean weight of female *Ascaris* has little causal impact on patterns of density-dependent egg output. The ZINB model is able to account for the disproportionately large number of zero egg counts within the data and is shown to be a consistently better fit than the NB model. The probability of observing a zero egg count is demonstrated as being negatively associated with both female worm burden and female mean weight.

Conclusion: The mean weight of female *Ascaris* is statistically significantly associated with egg output, and follows a consistent pattern of facilitation preceding limitation with increasing female worm burden. Despite these relationships, incorporation of female *Ascaris* mean weight into models of egg output has little effect on patterns of density dependence. The ZINB model is a superior fit to the data than the NB model and provides additional information regarding the mechanisms that result in a zero egg count. The ZINB model is shown to be a useful tool for the analysis of individual-based egg output data.

Background

Density-dependent population processes can occur at each stage of a parasite's lifecycle [1]. For the gastrointestinal (GI) nematodes these include establishment within the host, development and maturation time, adult survival, and female fecundity [2,3]. Density dependence has important implications for both the stability [2] and transmission dynamics [1,3,4] of helminth populations. Incorporation of these processes into mathematical models as accurately as possible is vital for furthering our understanding of important dynamical behaviour, such as the rate of re-infection following chemotherapeutic intervention and the spread of anthelmintic resistance [3-5].

In *Ascaris lumbricoides* infections of humans, density-dependent egg production has been reported; the *per capita* egg output decreasing with increasing number of worms per host (worm burden) [1,6]. Both the severity of density dependence and the level of egg production exhibit marked geographic variability [7]. This variability has implications for the use of egg counts to estimate the intensity of infection [7], and the applicability of transmission models across geographical locations for decision support in view of recent efforts to integrate the control of neglected tropical diseases.

Density-dependent reductions in worm size may be an important factor in *Ascaris* egg output. A positive correlation between worm size and egg production is commonly described in GI nematodes of ruminants (e.g. [8-13]) as well as in *Ascaris* infections of humans [14,15]. A constraint in worm size at high worm burdens may have a causal impact on patterns of density-dependent egg production. Reductions in worm size at high burdens have been described in both natural [12,16] and experimental [17] systems of directly-transmitted helminths in non-human mammals. There is conflicting evidence on the relationship between size and worm burden in *Ascaris* infections of humans. A number of studies have reported no evidence for density-dependent constraints [18-21] whereas the opposite has been described elsewhere [22].

Worm size and egg production may also interact with the host's immune response. In lambs, acquired immune responses to *Teladorsagia* (= *Ostertagia*) *circumcincta* infections are known to control the size of worms and reduce their egg output [8,23-25]. Similar correlations between the host immune response, worm size and egg output have been described in human hookworm infections [26]. Experimental infections of rats with *Strongyloides ratti* have shown that worms are larger and sometimes more fecund in immune-suppressed rats and smaller and sometimes less fecund in immunized animals when compared to controls [27,28]. Furthermore, density-dependent

fecundity effects in this nematode species are known to depend on the host immune response [29,30].

Despite a number of factors potentially influencing the egg production of female *Ascaris*, this important demographic and fitness parameter is ubiquitously described in terms of a single variable; the worm burden. This is largely due to the inherent difficulties in applying suitable statistical models to parasitological data [31,32]. Statistical analyses tend to be complicated by the high degree of variability in the egg output from a single host [33-36], highly overdispersed distributions of worms and egg output across a population of hosts [1,37], and the sensitivity and quantitative reliability of the diagnostic technique [35,36,38]. The estimated concentration of eggs may also be biased by host factors such as the volume of faeces produced (e.g., estimates in children tend to be inflated relative to those in adults) [35]. Typically, density-dependent reductions in egg output are presented in terms of female worm fecundity, a composite parameter describing the *per capita* egg production per unit time (eggs per gram of faeces divided by the number of (female) worms per host). Detection of density dependence has frequently been performed by fitting a functional form to the relationship between fecundity and worm burden. This method contravenes assumptions of statistical independence and may introduce bias via inaccuracies in the estimation of worm burden [39].

A number of studies have characterised density-dependent patterns of *Ascaris* egg output by fitting statistical models to grouped mean egg output data (e.g. [7,19,34]). The advantage of this method is that, assuming a large enough sample size per group, the distribution of means can be assumed to be normal, evoking the central limit theorem. However, this complicates the investigation of other variables which may also be important determinants of egg production. In addition, density-dependent *Ascaris* egg output has been exclusively described using data collected from populations at temporal equilibrium; consequently, whether this phenomenon is static or temporally dynamic is not known.

The analyses in this study are split into two parts. In the first part we define and fit a statistical model to evaluate the evidence for different forms of density dependence in the per host mean weight (a proxy for size) of female *Ascaris*. In the second, we explore the relationship between the mean weight of female worms and per host egg output using a multivariate modelling approach (controlling for both female worm burden and host age). We explore the suitability of two types of statistical model for modelling these individual-based egg output data: a negative binomial (NB) generalized linear model (GLM) [40] and a zero-inflated negative binomial (ZINB) model

[41,42]. The latter is useful in modelling data with a high proportion of zero counts [42-44] and has been applied to GI nematode egg count data in two previous studies [45,46]. Throughout this paper we define the net egg output as the estimated concentration of eggs per gram of faeces per host (regardless of whether they are or not fertilised). Thus we distinguish between egg production (fecundity) and fertility, whereby the latter measures the number of fertilised and embryonated eggs that a female worm produces (i.e. live offspring).

Methods
Study area and data collection
Data were collected from a poor urban suburb of Dhaka, Bangladesh between 1988 and 1989 by Hall and colleagues [7,47,48]. Briefly, households were visited and all their occupants invited to take part in the study with the aim of recruiting as many individuals as possible. All participants were asked to provide a faecal sample from which the number of *Ascaris lumbricoides* eggs were counted using a quantitative ether sedimentation technique [36] and the concentration of eggs per gram of faeces (EPG) estimated. Pyrantel pamoate was administered to each subject and their stools were collected for a period of 48 hours post-treatment. The worms recovered (*A. lumbricoides*) from the faeces of each individual were sexed, counted and weighed. Egg counts, treatment and worm counts were repeated on two further occasions at six-monthly intervals. Pyrantel pamoate paralyses adult *Ascaris* allowing them to be expelled intact from the gut [49] with a cure rate of approximately 88% [50]. Hence, these data provide a reliable and accurate measure of the number and weight of worms per host. The population of worms recovered after the first round of chemotherapy is termed the baseline population, after the 2nd round of chemotherapy, the 1st re-infection population and after the 3rd and final round, the 2nd re-infection population. The pre-treatment egg counts are similarly referred to.

Sample size
To evaluate the evidence for different forms of density dependence affecting the per host mean weight of female *Ascaris*, analyses were performed on the data collected from all individuals who were found to be infected with at least one female worm. To explore the relationship

between the per host mean weight of female *Ascaris* and the per host egg output, data were analysed from those individuals who were found to be infected with at least one female worm and from whom an estimate of egg output had been made. Table 1 summarises the data (available upon request to authors) used in these analyses. Definitions and descriptions of all parameters and variables used throughout this paper are given in Table 2.

Per host female mean weight and worm burden
To explore the relationship between the mean weight of female *Ascaris* in each host infra-population and the female worm burden we define the following statistical model. Let n be the number of female worms in a single host. Given a host harbours n female worms, the weights of the individual female worms are assumed to be independent with true mean M and variance σ_1^2. We further assume that M is a random variable with mean μ_W and variance σ_2^2. The observed per host average weight of a female worm, W, has mean μ_W, variance $\sigma_1^2 / n + \sigma_2^2$ and, for reasonably large n, will approximately follow a normal distribution,

$$W \sim N(\mu_W, \sigma_1^2 / n + \sigma_2^2). \qquad (1)$$

To model different types of density dependence we allow μ_W to be a function of the female worm burden, n, using the following generalised equation,

$$\mu_W = \frac{\alpha_1 + n^{(\alpha_3 - 1)}}{(1 + \alpha_2 n^{\alpha_3})}. \qquad (2)$$

This function can model three different possible relationships between μ_W and n. Model A: μ_W can be a positive constant for all $n (\alpha_1 > 0, \alpha_2 = 0, \alpha_3 = 1)$, indicative of an absence of density-dependent effects. Model B: μ_W can decline asymptotically with increasing $n (\alpha_1 > 0, \alpha_2 > 0, \alpha_3 = 1)$, representing negative or con-

Table 1: Summary of data used in analyses

Population	Hosts sampled		Hosts lost to follow-up		Hosts with a positive female worm count		Hosts with a positive female worm count and from whom egg output was estimated	
	age ≤ 12	age > 12	age ≤ 12	age > 12	age ≤ 12	age > 12	age ≤ 12	age > 12
Baseline	990	775	-	-	834	639	834	639
1st re-infection	699	558	291	217	555	383	528	364
2nd re-infection	592	425	107	133	496	257	463	228

Table 2: Definitions of variables and parameters

Type	Symbol	Definition/description	Units
Random variables, observed values	Λ, λ	Per host net egg output	eggs gram^{-1}
	W, w	Per host mean weight of female *Ascaris*	grams
	n	Per host female worm burden	worms host^{-1}
	a	Host age group ($a = 0$ for age ≤ 12, $a = 1$ for age > 12,)	
Unobserved random variables	M	Per host unobserved true mean weight of female *Ascaris*	grams
Unobserved values	μ_W	Expected value of the per host mean weight of female *Ascaris*	grams
	μ_Λ	Expected value of the per host net egg output	eggs gram^{-1}
	p	The probability of observing a zero count from the Bernoulli process	-
Estimated parameters	$\alpha_a = \{\alpha_{1a}, \alpha_{2a}, \alpha_{3a}\}$	Describe the form of density-dependence of the expected value of the per host mean weight of female *Ascaris* in age group a	-
	$\boldsymbol{\sigma} = \{\sigma_1^2, \sigma_2^2\}$	Describe the relationship between the variance of the per host mean weight of female *Ascaris* and the female worm burden	-
	λ_1	Exponential of the intercept term of statistical models fitting the expected value of the per host net egg output to data	-
	β	Describes the effect of host age category on the expected value of the per host net egg output	-
	c	An inverse measure of the severity of density dependence ($0 < c \leq 1$) on the expected value of the per host net egg output	-
	$\gamma = \{\gamma_1, \gamma_2, \gamma_3\}$	Describe the form of the relationship between the expected value of the per host net egg output and female *Ascaris* mean weight	-
	$\delta = \{\delta_0, \delta_1, \delta_2, \delta_3\}$	Describe the age-dependent relationship between the logit of the probability of observing a zero egg count, the natural logarithm of female worm burden and the natural logarithm of the mean weight of female *Ascaris*	-
	k	Inverse measure of the degree of overdispersion in egg output data	-

straining density dependence. Model C: μ_W can initially increase with increasing n followed by an asymptotic decline $(\alpha_1 > 0, \alpha_2 > 0, \alpha_3 > 1)$ describing a pattern of initial facilitation (positive density dependence) followed by limitation (negative density dependence); for examples and further discussion of using forms of equation (2) to describe density dependence in other host-parasite systems see [51,52]. Each of Models A-C is nested within the following one allowing their respective fits to be compared using the likelihood-ratio statistic (LRS) [53]. Under the null hypothesis the LRS follows a chi-square distribution with degrees of freedom (d.f.) equal to the difference in the number of parameters being estimated [54]. Akaike's information criterion (AIC) [55] was also calculated as an additional measure of goodness-of-fit. Host age, a, was incorporated into the model as a two level factor ($a = 0$ for children ≤12 years, $a = 1$ for teenagers and adults > 12 years) to allow the parameters pertaining to density dependence, $\alpha_a = \{\alpha_{1a}, \alpha_{2a}, \alpha_{3a}\}$, to vary between age groups. The variance parameters, $\boldsymbol{\sigma} = \{\sigma_1^2, \sigma_2^2\}$, were considered independent of host age. Assuming the observed per host mean female weight data to follow a normal distribution of the form given in equation (1), a log-likelihood function was derived and maximised using the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) [56-59] optimisation algorithm to obtain maximum likelihood estimates (MLEs) of the unknown parameters (α_a and σ). The BFGS algorithm was implemented using the optim function in the R statistical program (v.2.8.0) [60,61]. The best-fit form of equation (2) was determined separately in each of the baseline, 1st and 2nd re-infection populations.

Per host female mean weight and net egg output
Model derivation
For *A. lumbricoides*, the relationship between the total net egg output per host (denoted by the random variable Λ) and the per host female worm burden, n, has been empirically well described by a power function [7],

$$E(\Lambda) = \mu_\Lambda = \lambda_1 n^c. \qquad (3)$$

Here E represents the expected value, λ_1 the number of eggs per gram of faeces produced by a sole infecting female and and c inverse measure of the severity of negative density dependence (for $0 < c \leq 1$), with $c = 1$ indicating proportionality or density independence. This model is conveniently linearised by taking natural logarithms,

$$\mathrm{Ln}(\mu_\Lambda) = \mathrm{Ln}(\lambda_1) + c\mathrm{Ln}(n) \qquad (4)$$

A null model was defined by extending this relationship to include host age, a (where c is once again a two-level factor), by adding a multiplicative factor, e^β [in equation (3)], for teenagers and adults ≤12 years. Host age is a potentially important confounding factor associated with both the number of worms per host (the age-intensity profile of *Ascaris* infection is typically convex, for examples see [6,48]), and the concentration of egg counts (egg counts tend to be negatively related with the volume of faeces produced resulting in overestimation in children compared to adults [35]). Thus, the null model (denoted Model 1) describing the relationship between net egg output and female worm burden adjusting for host age was defined as,

$$\mathrm{Ln}(\mu_\Lambda) = \mathrm{Ln}(\lambda_1) + c\mathrm{Ln}(n) + \beta a \qquad (5)$$

In order to extend Model 1 [equation (5)] to reflect the potential dependence of per host net egg output on per host female mean weight the following preliminary analyses were performed to determine appropriate functional forms to describe the relationship. Per host egg output data were stratified by per host mean weight, taking the arithmetic mean per stratum, and regressing these values against polynomial functions (up to 3rd order) of the mean of the per host mean weight of each stratum. Stratum means were centred around their overall arithmetic mean value in order to minimise multicollinearity [62]. A problem of collinearity for non-centred polynomial terms was indicated by high values (consistently greater than 10 [63]) of their variance inflation factors (VIFs) [64] and high standard errors of their estimated coefficients. (We centre the per host female mean weight in all subsequently described polynomial regression models to ensure robust parameter estimation. We do not continue to explicitly indicate this to maintain the clarity of the mathematical notation.) Models were fitted using standard GLM procedures assuming the mean egg output per stratum to be normally distributed with constant variance [40] and implemented using the glm function in R [60,61]. Models were compared using the LRS and AIC. In the baseline, 1st and 2nd re-infection populations, the 3rd order (cubic), 2nd order (quadratic) and 3rd order functions were, respectively, indicated by both test statistics as being the best fits (Figure 1, Table 3).

Using the 3rd order relationship, Model 1 [equation (5)] was extended to define a full model describing the relationship between net egg output, female worm burden, and worm weight adjusting for host age,

Table 3: Models describing the egg output of *Ascaris lumbricoides* as polynomials of the mean weight of female worms

Population	Model order polynomial	Maximum log-likelihood	LRS	*p*-value	AIC
Baseline	1st	-10.68	-	-	27.37
	2nd	4.71	30.80	<0.0001	-1.43
	3rd†	11.22	13.02	0.00031	-12.45
1st re-infection	1st	-12.98	-	-	31.97
	2nd†	2.96	31.88	<0.0001	2.09
	3rd	3.88	1.84	0.17	2.25
2nd re-infection	1st	-12.14	-	-	30.29
	2nd	0.0043	24.30	<0.0001	7.99
	3rd†	3.68	7.34	0.0067	2.65

Comparison of models describing the per host net egg output of *Ascaris* as polynomial functions of the per host mean weight of female worms. Analyses were performed on grouped mean data (see text) using standard GLM procedures. † denotes the best-fit model in each population.
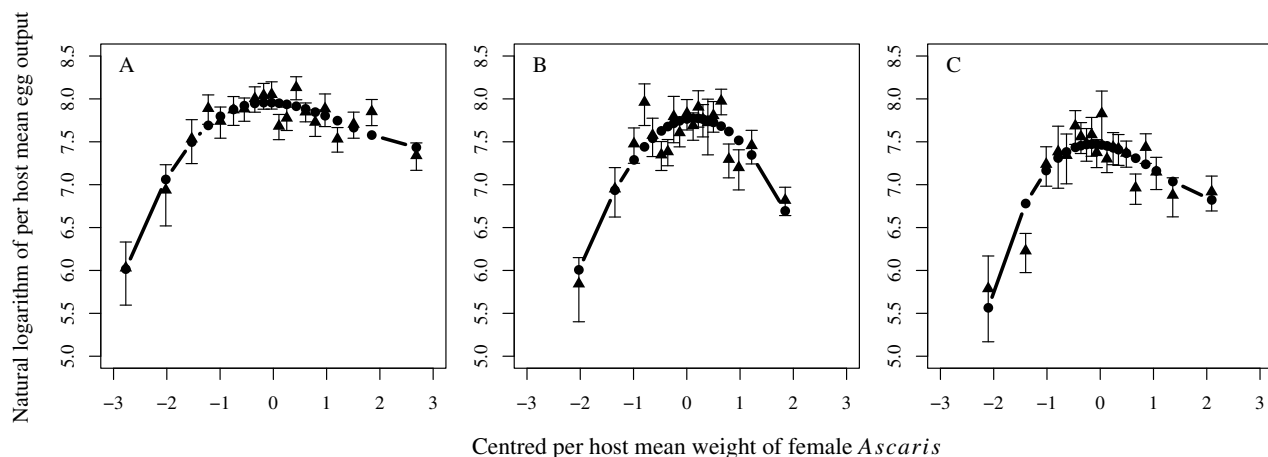
$$Ln(\mu_\Lambda) = Ln(\lambda_1) + cLn(n) + \beta a + \gamma_1 w + \gamma_2 w^2 + \gamma_3 w^3$$

$$(6)$$

Equation (6) is denoted Model 4. Models 2 and 3 are defined as special cases of Model 4, modelling female *Ascaris* mean weight as, respectively, 1st order $(\gamma_1 > 0, \gamma_2 = 0, \gamma_3 = 0)$ and 2nd order $(\gamma_1 > 0, \gamma_2 > 0, \gamma_3 = 0)$ polynomial functions (Table 4). In these models the parameters pertaining to the mean weight of female *Ascaris* per host $(\gamma = \{\gamma_1, \gamma_2, \gamma_3\})$ do not have a direct biological interpretation although from the

unadjusted models (i.e. unadjusted for the effects of host age and female worm burden) in each population (Figure 1), it is clear that the mean egg output tends to initially rise with increasing female mean weight followed by a decline. This functional relationship may be thought of as empirically modelling the antagonistic effects of growth and ageing on worm egg production.

Statistical modelling approach

In order to fit these linear models to the data it is necessary to assume an appropriate probability distribution of the per host net egg output, Λ. Count data are typically modelled assuming either a Poisson or negative binomial distribution (NBD) [40,42]. Given the high level of

Figure 1
The relationship between the per host net egg output and mean weight of female *Ascaris lumbricoides*. The relationship between the per host net egg output and the (centred, see main text) mean weight of female *Ascaris* in the baseline (A), 1st (B) and 2nd (C) re-infection populations. Triangles represent grouped mean egg outputs stratified by female *Ascaris* mean weight. Solid lines and circles represent the fitted values of the best fit polynomial functions (as determined by the likelihood-ratio statistic (LRS), see Table 3).

Table 4: Equations for Models 1–4 and 1I-4I

Models	Negative binomial count component	Bernoulli component (zero-inflated "I" models only)
1,1I	$\mathrm{Ln}(\mu_\Lambda) = \mathrm{Ln}(\lambda_1) + c\mathrm{Ln}(n) + \beta a$	
2,2I	$\mathrm{Ln}(\mu_\Lambda) = \mathrm{Ln}(\lambda_1) + c\mathrm{Ln}(n) + \beta a + \gamma_1 w$	
3,3I	$\mathrm{Ln}(\mu_\Lambda) = \mathrm{Ln}(\lambda_1) + c\mathrm{Ln}(n) + \beta a + \gamma_1 w + \gamma_2 w^2$	$\mathrm{Logit}(p) = \delta_0 + \delta_1 a + \delta_2\mathrm{Ln}(n) + \delta_3\mathrm{Ln}(w)$
4,4I	$\mathrm{Ln}(\mu_\Lambda) = \mathrm{Ln}(\lambda_1) + c\mathrm{Ln}(n) + \beta a + \gamma_1 w + \gamma_2 w^2 + \gamma_3 w^3$	

overdispersion in the egg output data (variance-to-mean ratio (VMR) = 4855, 4081 and 2726 in the baseline, 1st re-infection and 2nd re-infection populations respectively; see also Figure 2), the NBD is more appropriate than the Poisson [42]. However, these data also comprise a high proportion of zero counts which may not be adequately captured by the NBD (zero inflation, Figure 2). For *A. lumbricoides*, a zero egg count represents either an infra-population containing no sexually mature females or a false negative [38], since even unfertilised females can produced (unfertilised) eggs. (In this study, fertilised eggs were not distinguished from those unfertilised.) The distribution of *Ascaris* eggs in faecal samples from infected individuals has been shown to be highly aggregated [65] making false negatives more likely. Furthermore, the probability of a false negative may be dependent on properties of the worm infra-population and the infected host.

Data that are zero inflated relative to the NBD may be better described by a two-component mixture model which defines the response variable as a mixture of a Bernoulli and NBD (zero-inflated negative binomial, ZINB) [41-44,66]. Such a distribution allows zero counts to arise from two distinct mechanisms: a Bernoulli (binary) process generating either a positive or zero count and a count process (including the possibility of a zero count) [42]. Covariates of each process may or may not be the same [44] affording flexibility to construct models with the potential to explain a much higher degree of variability than assuming a single distribution. In these analyses, we fit the linear models derived in the previous section using both a negative binomial and mixture model approach and compare their respective fits.
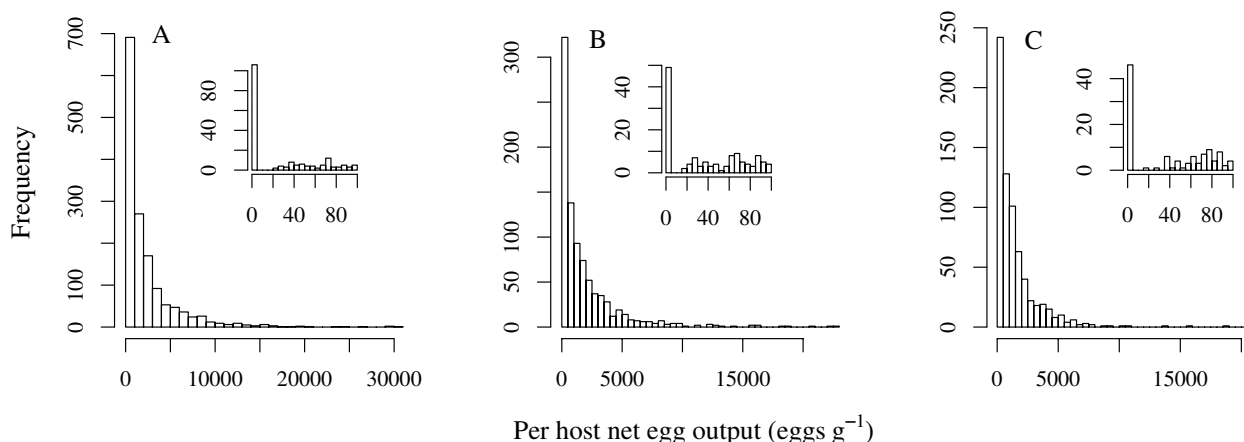

Figure 2
The distribution of per host egg output. Histograms depicting the distribution of the per host egg output in the baseline (A), 1st (B) and 2nd (C) re-infection populations. The insets are histograms of the distribution between 0–100 eggs gram[-1] highlighting the high proportion of zero counts.

Negative binomial (NB) model

For the NBD the probability of observing an egg count λ is,

$$P(\Lambda = \lambda) = \frac{\Gamma(\lambda+k)}{\Gamma(k)\lambda!} \frac{\mu_\Lambda k^k}{(\mu_\Lambda+k)^{\lambda+k}} \qquad (7)$$

where k is an inverse measure of the degree of overdispersion [61] and Γ is the gamma function $[\Gamma(x) = (x - 1)!]$. For a known value of k, equation (7) is of the form of the exponential family of probability distributions [40,67]. The natural logarithm (the link function) of μ_Λ is linearly related to the covariates described in the previous section [equations (5) and (6)] and so, for a given value of k, the parameters can be estimated within the GLM framework [40]. Here, since k is unknown, we employ a frequently used extension of the GLM methodology which allows maximum likelihood estimates of both k and the unknown linear parameters to be obtained [61,68]. The glm.nb function from the MASS package in R [61] was used to implement this technique and fit Models 1–4 (Table 4) to the data in each population. Models were compared using the LRS and AIC to determine the best-fit. We refer to these models as negative binomial or NB models.

Zero-inflated negative binomial (ZINB) model

For the zero-inflated negative binomial distribution the probability of observing an egg count λ is,

$$P(\Lambda = 0) = p + (1 - p)\left[\frac{k}{(\mu_\Lambda+k)}\right]^k$$

$$p(\Lambda = \lambda) = (1 - p)\frac{\Gamma(\lambda+k)}{\Gamma(k)\lambda!}\frac{\mu_\Lambda^\lambda k^k}{(\mu_\Lambda+k)^{\lambda+k}} \quad \lambda = 1, 2, ...$$

$$(8)$$

Here p is the probability of observing a zero count originating from the Bernoulli process and $[k/(\mu_\Lambda + k)]^k$ is the probability of observing a zero count from the NBD. Just as μ_Λ is linearly related to covariates via the logarithmic link function, the logit function $(\ln[p/(1 - p)])$ can be used to linearise the relationship between p and potential covariates [40]. Univariate exploration of the data indicated a negative linear relationship between logit(p) and the natural logarithm of stratified groups of the per host mean weight of female *Ascaris* and the per host female worm burden (Figures 3 and 4). Equation (9) is a model which includes host age (a), the natural logarithm of the mean weight of female *Ascaris* [ln(w)], and the natural logarithm of the female worm burden [ln(n)] as covariates of the probability of observing a zero count. This model was fitted to the data in each population using standard GLM

procedures implemented in R using the glm function [40,61].

$$\log \mathrm{it}(p) = \delta_0 + \delta_1 a + \delta_2 \mathrm{Ln}(n) + \delta_3 \mathrm{Ln}(w) \qquad (9)$$

These preliminary multivariate analyses confirmed ln(n) and ln(w) to be statistically significantly and negatively related to the probability of observing a zero count in each population (Table 5). In the 2nd re-infection population age was found to be positively associated with p (i.e. the probability of observing a zero count is greater in teenagers and adults than in children) (Table 5). The relationship given in equation (9) was used to define the Bernoulli component of the mixture model and extend Models 1–4 into zero-inflated models (denoted by the letter I, Models 1I-4I, Table 4). These models were fitted to the data in each population using maximum likelihood implemented using the zeroinfl function from the pscl package (v.1.02) [68] in R (for further information on fitting zero-inflated mixture models see [41,43,69]). The fitted models were compared to one another using the LRS to determine the best-fit and also to their corresponding NB model (Models 1–4) using AIC. We refer to zero-inflated models as ZINB.

Results
Per host female mean weight and worm burden
Comparisons of nested forms of equation (2) indicated a pattern of facilitation followed by limitation in all populations (Table 6, Figure 5). The LRS and corresponding *p*-values are unambiguous in the baseline and 1st re-infection populations (*p*-value < 0.0001), whereas the facilitation preceding limitation pattern was only marginally preferred over limitation alone in the 2nd re-infection population (*p*-value = 0.046). AIC supported facilitation preceding limitation in all populations. The pattern of density dependence was similar in both age groups, whereas teenagers and adults tended to harbour slightly heavier worms (Figure 5).

Per host female mean weight and net egg output
The LRS and AIC indicated that incorporating the mean weight of female *Ascaris* as a covariate improved the fit of both the NB and ZINB models in all populations (Table 7). The best-fit functional form of the relationship between the per host mean weight of female *Ascaris* and the per host egg output (order of polynomial) varied across populations and with the assumed probability model (Table 8).

The ZINB model provided a consistently better fit to the data than its non zero-inflated counterpart (Table 7, comparing AIC values) and was able to account for the high proportion of zero counts within the data (Table 9). The vast majority of these zero counts were described in the
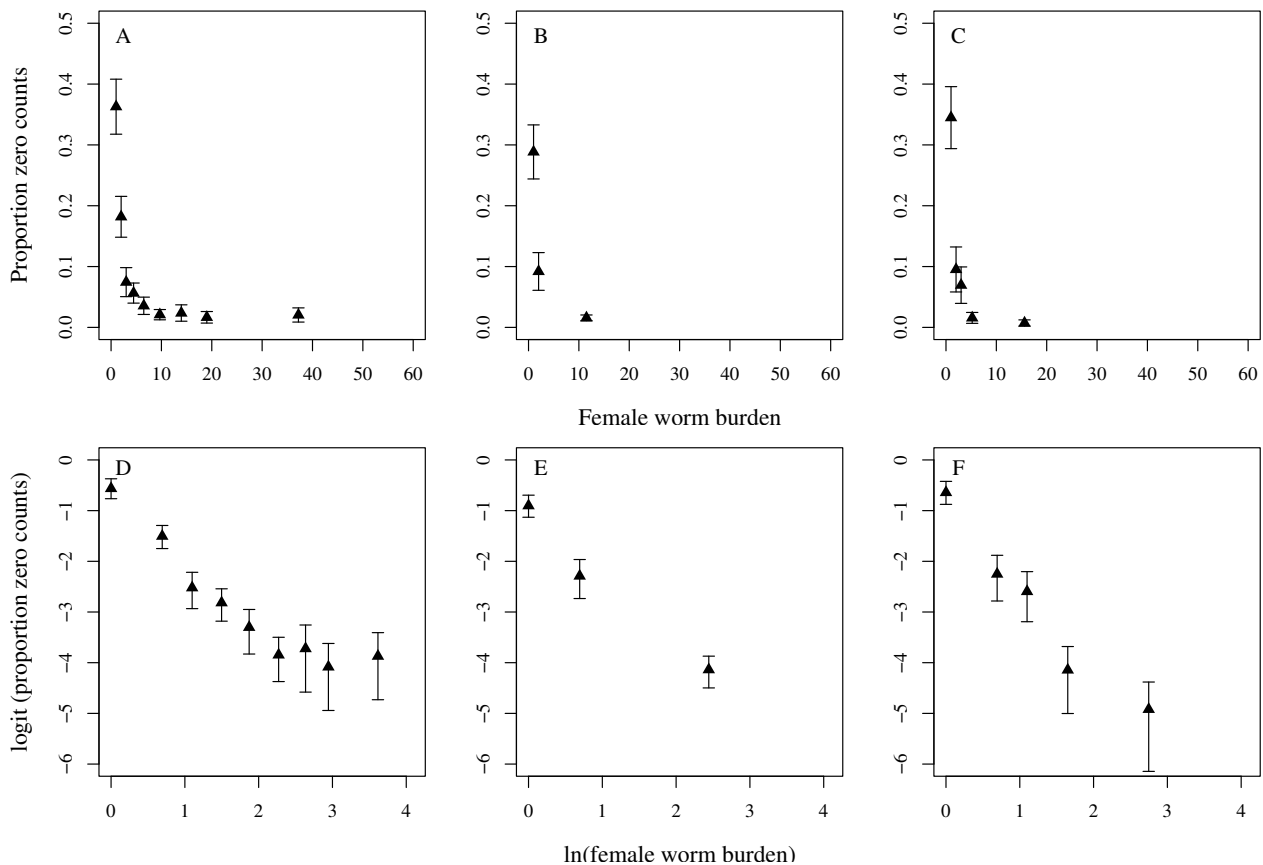
Figure 3
Relationship between the proportion of zero egg counts and female worm burden. Top row: a scatter plot of the proportion of zero egg counts per stratum of female *Ascaris* worm burden in the baseline (A), 1st (B) and 2nd (C) re-infection populations. Bottom row: logit of the proportion of zero egg counts per stratum of the natural logarithm of female worm burden in the baseline (D), 1st (E) and 2nd (F) re-infection populations showing approximately linear relationships.

Bernoulli component of the ZINB model (Table 9); in all populations the per host female worm burden and the per host mean weight of female *Ascaris* were negatively associated with the probability of a zero egg count (Table 8, *p*-value < 0.0001 for both covariates). In the 2nd re-infection population there was also evidence that host age was positively associated with the probability of a zero count (*p*-value = 0.0092 in the best-fit model, Model 2I, Table 8).

The parameter values estimated from the ZINB and NB models are broadly similar (Table 8). The estimated value of the overdispersion parameter k, tends to be higher in the ZINB models (indicating reduced overdispersion) since many of the zeros are accounted for in the Bernoulli

component of the model (Table 9). It is noteworthy that the estimated values of parameter c (the inverse measure of density dependence) tend to be lower in the ZINB models (indicative of more severe density dependence, Table 9).

Discussion
The major objectives of this study were twofold: Firstly, to determine whether there is any evidence for density-dependent processes affecting the per host mean weight of female *Ascaris lumbricoides*. Secondly, to determine whether per host female mean weight is associated with per host egg output and what, if any, causal impact this has on density-dependent egg production. We have
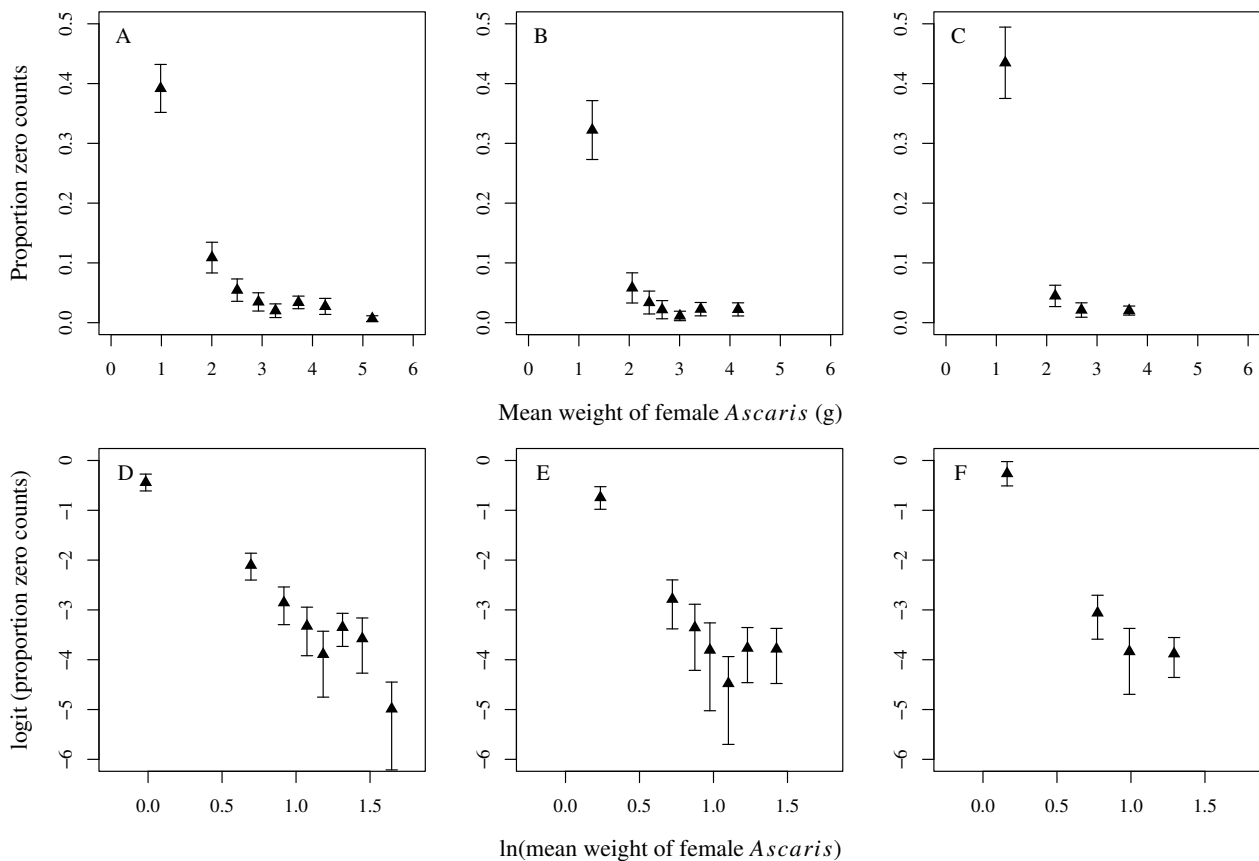
Figure 4
Relationship between the proportion of zero egg counts and the mean weight of female *Ascaris lumbricoides*.
Top row: a scatter plot of the proportion of zero egg counts per stratum of female *Ascaris* mean weight in the baseline (A), 1st
(B) and 2nd (C) re-infection populations. Bottom row: logit of the proportion of zero egg counts per stratum of the natural log-
arithm of female mean weight baseline (D), 1st (E) and 2nd (F) re-infection populations showing approximately linear relation-
ships.

shown that the per host mean weight of female *Ascaris* fol-
lows a pattern of initial facilitation followed by limitation
with worm burden both at endemic equilibrium (baseline
population) and after 6 months re-infection (re-infection
populations). An association between the per host mean
weight of female *Ascaris* and the per host egg output is
demonstrated in the three analysed populations. The
functional form of this relationship is different across
populations and dependent on the assumed probability
model used to estimate the unknown parameters. How-
ever, comparing the zero-inflated negative binomial
(ZINB) models, which provide a better description of the
observed data, we see that at baseline egg output initially
rises with increasing per host female mean weight before
falling at very high weights, whilst in the re-infection pop-
ulations, egg output rises monotonically with increasing

weight. Despite these findings, per host female mean
weight has little discernable causal impact on the well-
characterised patterns of density-dependent egg produc-
tion in *A. lumbricoides* [6,7,70].

The convex pattern of facilitation preceding limitation has
been documented in one previous study of the GI-nema-
tode *Heterakis gallinarum* infecting the ring-necked pheas-
ant (*Phasianus colchicus*) [16]. Constraints on female
weight may be caused by intra-specific (exploitation)
competition for either nutrients or space or by host-medi-
ated effects; such as a non-protective immune response.
Limitation of size due to competition for nutrients is
unlikely since the total energy requirements of even a
heavy *Ascaris* infection is small relative to that of a human
host [71] although it may be possible in the severely

Table 5: Parameter values estimated from the logistic model describing the probability of a zero egg count

Population	δ_0 (SE)	δ_1 (SE)	δ_2 (SE)	δ_3 (SE)
Baseline	0.49* (0.24)	-0.23 (0.27)	-1.77*** (0.18)	-1.01*** (0.14)
1st re-infection	0.91* (0.43)	0.51 (0.40)	-2.63*** (0.36)	-2.00*** (0.33)
2nd re-infection	0.63 (0.41)	1.07** (0.41)	-2.56*** (0.38)	-1.62*** (0.28)

Parameter values were estimated by fitting equation (9) to the egg count data encoded in a binary fashion (positive or zero) using standard GLM procedures. Parameters refer to the following covariates: δ_1: host age, δ_2: the natural logarithm of female worm burden and δ_3: the natural logarithm of female mean weight. Parameters significantly different from 0: * *p*-value < 0.05, ** *p*-value < 0.01, ***p*-value < 0.001.

undernourished. Constraints mediated by the host immune response are also dubitable since teenagers and adults in all populations, who would have a greater immunological experience of *Ascaris* antigens, on average harboured heavier worms. Additionally, negative associations have been reported between various immunological markers and the intensity of *Ascaris* infections [72] and re-infections following chemotherapy [73], suggesting that individuals with heavy worm burdens mount a weaker rather than a stronger immune response. If worm burden relates to the immune response in this manner and the response affects the size of female worms, then the per host mean weight of females would increase with per host worm burden in a facilitative pattern. Immune responses are known to limit worm size in experimental GI infections of rats with *S. ratti* [27,74] and sheep with *T. circumcincta* [24,25]. The data presented in this study are not sufficient to distinguish between the various potential causative mechanisms behind the observed density dependence, however, we speculate that the facilitation is immune mediated whereas the limitation is the result of competition for space.

Two previous studies have described a positive relationship between the size of female *A. lumbricoides* and egg production. Sinniah and Subramaniam [14] dissected the uteruses of females expelled from 50 schoolchildren and showed a moderately positive linear relationship. Seo and Chai [15] took a different approach, relating egg output with female length from hosts harbouring a single female or a male and female pair. Their results point to a more parabolic shape to the relationship, with egg output declining in very large (and presumably old) worms. This is in accordance with the results of the present study in the baseline population. Allometric relationships between body size and egg output are a characteristic feature of parasitic nematode infections [75,76], so it is not surprising that similar mechanisms operate in *Ascaris* infections of

humans. More interesting, however, is how this association influences the host-parasite interaction and ensuing population dynamics; do host responses limit the size of worms? Are some hosts more efficient than others at doing so? Such processes and heterogeneities are known to occur in model non-human nematode systems [8,27,28].

The degree of density-dependent egg output (described by parameter *c*) remains approximately equal in the null and best-fit models in each of the three populations (regardless of the probability model). This consistency shows that the severity of density dependence is not greatly altered by the effects of female weight. Thus, egg production is limited directly by increasing female worm burden and is not simply an artefact resulting from the density dependence of female mean weight, i.e. the association between egg output and female weight does not cause density-dependent fecundity. It is noteworthy that no statistically significant density-dependent fecundity was detected in the 1st re-infection population (*c* = 1.04, 95% C.I. 0.97–1.11, Model 2I, Table 8). The severity of density-dependent *Ascaris* fecundity is known to be weak in Bangladesh relative to other geographical locations [7], and so its detection is likely to be prone to type II statistical errors.

The relationships between the per host net egg output and the female mean weight varied between the baseline and re-infection populations, with a significant decrease for heavier worms present only at baseline (as indicated by the cubic polynomial providing the best-fit functional relationship, Model 4I). This is congruent with the biological interpretation of this functional form representing a decline in egg production in heavier (inferred older) worms. This would be expected to be unimportant in the populations after six months of re-infection since the life-expectancy of *Ascaris* is estimated to range between 1 and 2 years [1,77].

An important result from this work is the evidence that the per host net egg output tends to be higher in children than adults in the baseline population (β = -0.28, *p*-value < 0.0001, Table 8). There is also marginal evidence for this effect in the 2nd re-infection population (β = -0.19, *p*-value = 0.015, Table 8). Egg concentration can be negatively associated with the volume of faeces produced resulting in overestimation of egg output in children compared to adults [35]. However, given the unambiguous result in the baseline population it is surprising that the effect is absent and not more statistically significant in the 1st and 2nd re-infection populations respectively. An alternative explanation is that the decreased egg output in adults is due to an acquired immune response. However, to reconcile this with the results from the re-infection populations, the

Table 6: Comparison of models describing density dependence in the mean weight of female *Ascaris lumbricoides*

Population	Model	Model equation	Model description	Test	LRS	*p*-value	AIC
Baseline	A	$\mu_W = \alpha_1$	no density dependence	-	-	-	4671.33
	B	$\mu_W = \dfrac{\alpha_1 n}{(1+\alpha_2 n)}$	limitation	A vs. B	43.36	<0.0001	4631.97
	C†	$\mu_W = \dfrac{\alpha_1 n^{(\alpha_3-1)}}{(1+\alpha_2 n^{a3})}$	facilitation preceding limitation	B vs. C	38.05	<0.0001	4597.91
1st re-infection	A	$\mu_W = \alpha_1$	no density dependence	-	-	-	2314.29
	B	$\mu_W = \dfrac{\alpha_1 n}{(1+\alpha_2 n)}$	limitation	A vs. B	18.98	<0.0001	2299.31
	C†	$\mu_W = \dfrac{\alpha_1 n^{(\alpha_3-1)}}{(1+\alpha_2 n^{a3})}$	facilitation preceding limitation	B vs. C	17.99	<0.0001	2285.32
2nd re-infection	A	$\mu_W = \alpha_1$	no density dependence	-	-	-	1976.29
	B	$\mu_W = \dfrac{\alpha_1 n}{(1+\alpha_2 n)}$	limitation	A vs. B	12.36	0.0021	1967.93
	C†	$\mu_W = \dfrac{\alpha_1 n^{(\alpha_3-1)}}{(1+\alpha_2 n^{a3})}$	facilitation preceding limitation	B vs. C	6.18	0.046	1965.75

Comparison of the goodness-of-fit of special cases of the generalised equation [equation (2)] describing the relationship between the expected value of the per host mean weight of female *Ascaris*, μ_W, and the per host female worm burden, *n*, using a likelihood ratio test and AIC; *p*-values were calculated assuming that the LRS follows, under the null hypothesis, a chi-square distribution with 1 d.f. † denotes the best-fit of each model type in each population.
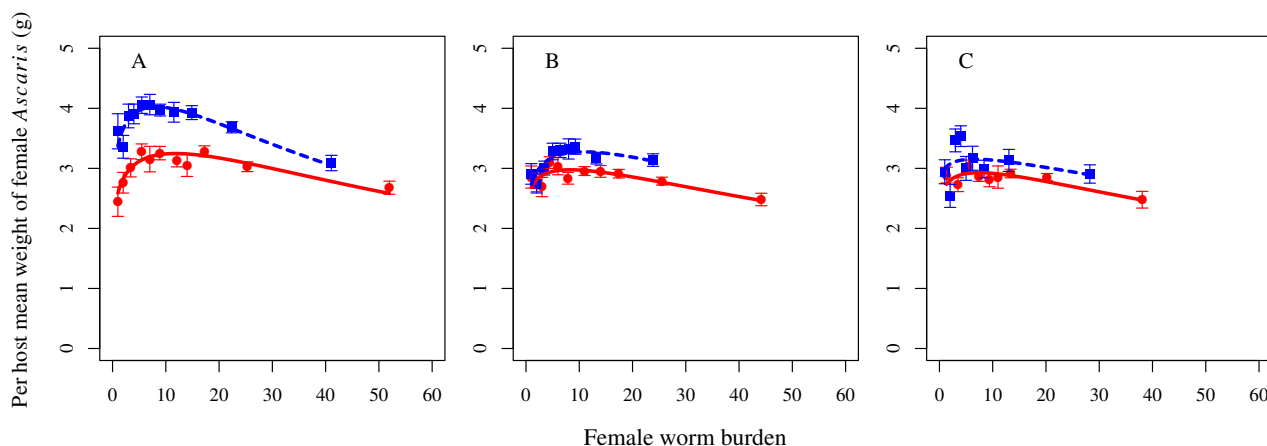
Figure 5
The best-fit relationship between the per host mean weight of female *Ascaris lumbricoides* and female worm burden. The best-fit functional relationships (as determined by the LRS, Table 6) between the per host mean weight of female *Ascaris* and the female worm burden in the baseline (A), 1st (B), and 2nd (C) re-infection populations. The solid red line is the best-fit to children (age ≤ 12 years) and the broken blue line to teenagers and adults (age > 12 years). The best-fit function is given by equation (2) and represents a pattern of initial facilitation followed by limitation. Circular and square data points are grouped means for children and teenagers and adults respectively. Error bars represent the standard error of the mean.

duration of immunological memory should be extremely short and rely on constant exposure to established worms (for transmission models incorporating the effect of immunological memory in helminth infections see [78-80]).

We have shown that a mixture of the negative binomial and Bernoulli distributions (ZINB model) provides a superior description of the distribution of egg output data than a negative binomial (NB model) distribution alone. Similar zero-inflated models have been used frequently in the ecological literature where datasets with many zeros are commonplace (for a review see [81]). In parasitological research, we are aware of only two previous studies that have used zero-inflated models to describe egg output data [45,46]. An added advantage of using a zero-inflated model is the insight which can be gained into the source of zeros (egg counts). Here we show that the probability of a zero count is negatively associated with both the per host female worm burden and the per host female mean weight. These associations suggest that the zero egg counts are false negatives within the data, i.e. failing to detect eggs in truly egg-contaminated faeces. We hypothesise that the greater the total (net) egg production the lower the probability that a sample is taken from a non-contaminated part of the collected faeces. Thus, since total egg production is positively associated both with female worm burden and female mean weight, the probability of sampling from a non-contaminated part of the faeces decreases with increasing female worm burdens and mean weight. This effect will be exacerbated by the highly over-

dispersed distribution of *A. lumbricoides* eggs in faecal samples [65].

The results presented in Table 9 suggest that a very small fraction of the zeros in the data were generated from the negative binomial count process. If we accept the explanation that the vast majority of zeros are false negatives, it is tempting to remove zero counts *a priori* in order to simplify analyses aimed at detecting epidemiologically significant covariates (i.e. covariates that directly impact upon the release of transmission stages). In taking such an approach one must again choose an appropriate distribution with which to model the now zero truncated data. Two suitable contenders are the log-normal and zero-truncated negative binomial distributions. The advantage of the former is that, via a logarithmic transformation, ordinary least squares estimation procedures can be used. For the latter, numerical maximisation of the appropriate log-likelihood function is required [42,44]. Figure 6 compares the results of fitting equation (5) (in which the mean per host egg output is modelled as being dependent on female worm burden and host age only) to the zero-truncated baseline data using the two approaches. Clearly the log-normal assumption provides an inadequate description of the data due largely to the poor approximation of the variance-to-mean relationship, a key aspect in accurate parameter estimation [40,82] (for details of the variance-to-mean relationship for the log-normal and zero-truncated negative binomial distributions see additional file 1 and [83]). Therefore, although removing zeros from the data may be a reasonable approach, more

Table 7: Comparison of models describing the per host net egg output

Population	Model type	Model	Test	LRS	d.f.	p-value	AIC
Baseline	Negative binomial	1	-	-	-	-	24679.09
		2	1 vs. 2	50.87	1	<0.0001	24630.21
		3†	2 vs. 3	8.35	1	0.0039	24623.86
		4	3 vs. 4	0.02	1	0.89	24625.84
	Zero-inflated negative binomial	1I	-	-	-	-	23990.51
		2I	1 vs. 2	26.89	1	<0.0001	23965.62
		3I	2 vs. 3	0.41	1	0.52	23967.20
		4I†	3 vs. 4	6.04	1	0.014	23963.17
			2 vs. 4	6.45	2	0.049	
1st re-infection	Negative binomial	1	-	-	-	-	14404.20
		2	1 vs. 2	44.42	1	<0.0001	14361.78
		3	2 vs. 3	4.06	1	0.044	14359.72
		4†	3 vs. 4	4.50	1	0.025	14356.72
	Zero-inflated negative binomial	1I	-	-	-	-	14043.52
		2I†	1 vs. 2	24.52	1	<0.0001	14020.97
		3I	2 vs. 3	0.25	1	0.62	14022.73
		4I	3 vs. 4	0.68	1	0.41	14024.04
2nd re-infection	Negative binomial	1	-	-	-	-	11050.10
		2	1 vs. 2	24.11	1	<0.0001	11027.99
		3†	2 vs. 3	5.93	1	0.015	11024.06
		4	3 vs. 4	1.72	1	0.19	11024.33
	Zero-inflated negative binomial	1I	-	-	-	-	10715.26
		2I†	1 vs. 2	7.24	1	0.0071	10710.02
		3I	2 vs. 3	0.22	1	0.65	10711.81
		4I	3 vs. 4	0.56	1	0.46	10713.26

The goodness-of-fit of all NB and ZINB models fitted to the data assessed by the LRS and AIC. † denotes the best-fit of each model type in each population. Model equations are given in Table 4.

Table 8: Parameter estimates from the null and best-fit statistical models describing per host net egg output

Population	Model	Negative binomial count component parameters							Bernoulli component parameters			
		$\ln(\lambda_1)$ (SE)	β (SE)	c (SE)	γ_1 (SE)	γ_2 (SE)	γ_3 (SE)	k (SE)	δ_0 (SE)	δ_1 (SE)	δ_2 (SE)	δ_3 (SE)
Baseline	1 (null)	5.98*** (0.081)	-0.14* (0.068)	0.79*** (0.033)	-	-	-	0.61 (0.020)	-	-	-	-
	3 (best-fit)	6.18*** (0.096)	-0.29*** (0.069)	00.74*** (0.036)	0.20*** (0.026)	-0.046** (0.015)	-	0.63 (0.021)	-	-	-	-
	11 (null)	6.25*** (0.066)	-0.20*** (0.053)	0.70*** (0.027)	-	-	-	1.08 (0.034)	0.49* (0.24)	-0.23 (0.28)	-1.02*** (0.14)	-1.78*** (0.18)
	41 (best-fit)	6.26*** (0.076)	-0.28*** (0.055)	0.70*** (0.027)	0.18*** (0.034)	-0.00062 (0.014)	-0.015* (0.0060)	1.10 (0.034)	0.49* (0.24)	-0.23 (0.28)	-1.01*** (0.14)	-1.78*** (0.18)
1st re-infection	1 (null)	5.08*** (0.089)	0.023 (0.078)	1.12*** (0.038)	-	-	-	0.82 (0.036)	-	-	-	-
	4 (best-fit)	5.17*** (0.099)	-0.051 (0.076)	1.11*** (0.039)	0.19** (0.063)	-0.082** (0.031)	0.039* (0.017)	0.86 (0.038)	-	-	-	-
	11 (null)	5.31*** (0.081)	0.024 (0.065)	1.02*** (0.035)	-	-	-	1.28 (0.044)	0.91* (0.44)	0.52 (0.40)	-2.00*** (0.33)	-2.65*** (0.37)
	21 (best-fit)	5.28*** (0.080)	-0.027 (0.065)	1.04*** (0.034)	0.18*** (0.037)	-	-	1.31 (0.04)	0.91* (0.44)	0.51 (0.40)	-2.00*** (0.33)	-2.64*** (0.37)
2nd re-infection	1 (null)	5.71*** (0.10)	-0.17 (0.10)	0.81*** (0.047)	-	-	-	0.71 (0.035)	-	-	-	-
	3 (best-fit)	5.77*** (0.11)	-0.23* (0.10)	0.82*** (0.049)	0.27*** (0.047)	-0.078* (0.033)	-	0.74 (0.037)	-	-	-	-
	11 (null)	5.95*** (0.087)	-0.14 (0.077)	0.71*** (0.040)	-	-	-	1.23 (0.05)	0.63 (0.41)	1.07** (0.41)	-1.62*** (0.29)	-2.56*** (0.39)
	21 (best-fit)	5.92*** (0.086)	-0.19* (0.079)	0.72*** (0.040)	0.11** (0.042)	-	-	1.24 (0.050)	0.63 (0.41)	1.07** (0.41)	-1.63*** (0.29)	-2.56*** (0.39)

Estimated parameter values from the best-fit NB and ZINB models in each population. Parameters significantly different from 0: * *p*-value < 0.05, ** *p*-value < 0.01, ***p*-value < 0.001. Model equations are given in Table 4.

Table 9: Observed and predicted percentage of zero counts from the NB and ZINB models

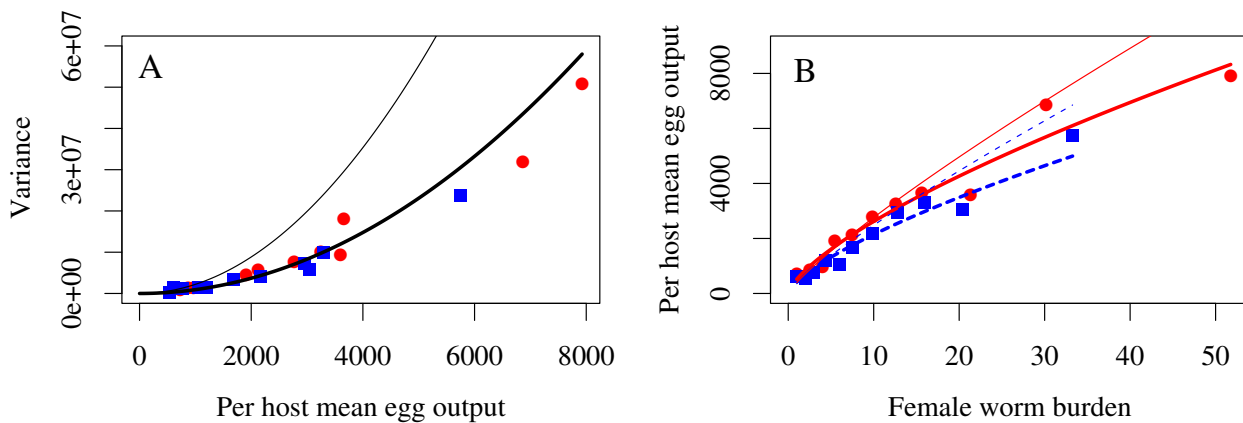| Population	Observed percentage zero counts	Predicted percentage zero counts		
		NB Model	ZINB Model	
			Count component	Bernoulli component
Baseline	7.12%	0.80%	0.039%	7.16%
1st re-infection	5.49%	0.36%	0.033%	5.47%
2nd re-infection	6.66%	0.58%	0.030%	6.64%

Figure 6
Comparison of the fit of a log-normal and zero-truncated negative binomial model. A: The estimated variance-to-mean relationship from the zero-truncated negative binomial model (black thick line) and the log-normal model (black thin line). B: The fitted zero-truncated (thick lines) and log-normal (thin lines) models to data from children (red solid line) and teenagers and adults (blue broken line) in the baseline population. In both figures red circles represent grouped mean data from children and blue squares from teenagers and adults (as defined in Figure 5). Details of the variance-to-mean relationship for the log-normal and zero-truncated negative binomial models are given in additional file 1.

complex and non-standard statistical models are still required for adequate parameter estimation [42].

Conclusion
In this study we have demonstrated that the mean weight of female *A. lumbricoides* infecting a cohort of human hosts follows a pattern of facilitation preceding limitation with increasing worm burden. We verify that weight is associated with net egg output but demonstrate that this has little causal impact on patterns of density-dependent egg production. We show that a zero-inflated negative binomial (ZINB) probability distribution is superior to a negative binomial distribution in modelling individual egg output data.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
MW outlined the paper, performed the analyses and wrote the manuscript. AH collected the data and helped in the interpretation of the analyses. M-GB edited the manuscript for intellectual content. RMA, M-GB and AH jointly supervised the work and helped conceive the paper.

Additional material

Additional file 1
Probability distribution, variance and expected value for the zero-truncated negative binomial and log-normal distributions. A file containing the probability distribution, variance and expected value for the zero-truncated negative binomial and log-normal distributions.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1756-3305-2-11-S1.pdf]

References
1. Anderson RM, May RM: *Infectious Diseases of Humans* Oxford: Oxford University Press; 1992.
2. Anderson RM, May RM: **Regulation and stability of host-parasite population interactions I. Regulatory processes.** *Journal of Animal Ecology* 1978, **47**:219-247.
3. Churcher TS: **Modelling the Spread of Anthelmintic Resistance. PhD thesis.** Imperial College London, Department of Infectious Disease Epidemiology; 2006.

4. Churcher TS, Basáñez M-G: **Density dependence and the spread of anthelmintic resistance.** *Evolution* 2008, **62:**528-537.
5. Churcher TS, Filipe JAN, Basáñez M-G: **Density dependence and the control of helminth parasites.** *Journal of Animal Ecology* 2006, **75:**1313-1320.
6. Anderson RM, May RM: **Helminth infections of humans: mathematical models, population dynamics, and control.** *Advances in Parasitology* 1985, **24:**1-101.
7. Hall A, Holland C: **Geographical variation in *Ascaris lumbricoides* fecundity and its implications for helminth control.** *Parasitology Today* 2000, **16:**540-544.
8. Stear M, Bairden K, Bishop SC: **How hosts control worms.** *Nature* 1997, **389:**27.
9. Stear MJ, Bishop SC: **The curvilinear relationship between worm length and fecundity of *Teladorsagia circumcincta*.** *International Journal for Parasitology* 1999, **29:**777-780.
10. Richards DT, Lewis JW: **Fecundity and egg output by *Toxocara canis* in the red fox, *Vulpes vulpes*.** *Journal of Helminthology* 2001, **75:**157-164.
11. Irvine RJ, Stien A, Dallas JF, Halvorsen O, Langvatn R, Albon SD: **Contrasting regulation of fecundity in two abomasal nematodes of Svalbard reindeer (*Rangifer tarandus platyrhynchus*).** *Parasitology* 2001, **122:**673-681.
12. Dezfuli BS, Volponi S, Beltrami I, Poulin R: **Intra- and interspecific density-dependent effects on growth in helminth parasites of the cormorant, *Phalacrocorax*.** *Parasitology* 2002, **124:**537-544.
13. Skorping A, Read AF, Keymer AE: **Life-history covariation in intestinal nematodes of mammals.** *Oikos* 1991, **60:**365-372.
14. Sinniah B, Subramaniam K: **Factors influencing the egg production of *Ascaris lumbricoides*: relationship to weight, length and diameter of worms.** *Journal of Helminthology* 1991, **65:**141-147.
15. Chai JY, Hong ST, Lee SH, Seo BS: **Fluctuation of the egg production amounts according to worm burden and length of *Ascaris lumbricoides*.** *Korean Journal of Parasitology* 1981, **19:**38-44.
16. Tompkins DM, Hudson PJ: **Regulation of nematode fecundity in the ring-necked pheasant (*Phasianus colchicus*): not just density dependence.** *Parasitology* 1999, **118:**417-423.
17. Michael E, Bundy DAP: **Density dependence in establishment, growth and worm fecundity in intestinal helminthiasis: the population biology of *Trichuris muris* (Nematoda) infection in CBA/Ca mice.** *Parasitology* 1989, **98:**451-458.
18. Mello DA: **A note on egg production of *Ascaris lumbricoides*.** *Journal of Parasitology* 1974, **60:**380-381.
19. Martin J, Keymer A, Isherwood RJ, Wainwright SM: **The prevalence and intensity of *Ascaris lumbricoides* infections in Moslem children from northern Bangladesh.** *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1983, **77:**702-706.
20. Elkins DB, Haswell-Elkins M: **The weight/length profiles of *Ascaris lumbricoides* within a human community before mass treatment and following reinfection.** *Parasitology* 1989, **99:**293-299.
21. Holland CV, Crompton DW, Taren DL, Nesheim MC, Sanjur D, Barbeau I, Tucker K: ***Ascaris lumbricoides* infection in pre-school children from Chiriqui Province, Panama.** *Parasitology* 1987, **95:**615-622.
22. Monzon RB, Cabrera BD, Cruz AC, Baltazar JC: **The "crowding effect" phenomenon in *Ascaris lumbricoides*.** *Southeast Asian Journal of Tropical Medicine and Public Health* 1990, **21:**580-585.
23. Stear MJ, Strain S, Bishop SC: **Mechanisms underlying resistance to nematode infection.** *International Journal for Parasitology* 1999, **29:**51-56.
24. Stear MJ, Bishop SC, Doligalska M, Duncan JL, Holmes PH, Irvine J, McCririe L, McKellar QA, Sinski E, Murray M: **Regulation of egg production, worm burden, worm length and worm fecundity by host responses in sheep infected with *Ostertagia circumcincta*.** *Parasite Immunology* 1995, **17:**643-652.
25. Stear MJ, Park M, Bishop SC: **The key components of resistance to *Ostertagia circumcincta* in lambs.** *Parasitology Today* 1996, **12:**438-441.
26. Pritchard DI, Quinnell RJ, Walsh EA: **Immunity in humans to *Necator americanus*: IgE, parasite weight and fecundity.** *Parasite Immunology* 1995, **17:**71-75.
27. Wilkes CP, Thompson FJ, Gardner MP, Paterson S, Viney ME: **The effect of the host immune response on the parasitic nematode *Strongyloides ratti*.** *Parasitology* 2004, **128:**661-669.
28. Viney ME, Steer MD, Wilkes CP: **The reversibility of constraints on size and fecundity in the parasitic nematode *Strongyloides ratti*.** *Parasitology* 2006, **133:**477-483.
29. Bleay C, Wilkes CP, Paterson S, Viney ME: **Density-dependent immune responses against the gastrointestinal nematode *Strongyloides ratti*.** *International Journal for Parasitology* 2007, **37:**1501-1509.
30. Paterson S, Viney ME: **Host immune responses are necessary for density dependence in nematode infections.** *Parasitology* 2002, **125:**283-292.
31. Wilson K, Grenfell BT: **Generalized linear modelling for parasitologists.** *Parasitology Today* 1997, **13:**33-38.
32. Paterson S, Lello J: **Mixed models: getting the best use of parasitological data.** *Trends in Parasitology* 2003, **19:**370-375.
33. Anderson RM, Schad GA: **Hookworm burdens and faecal egg counts: an analysis of the biological basis of variation.** *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1985, **79:**812-825.
34. Croll NA, Anderson RM, Gyorkos TW, Ghadirian E: **The population biology and control of *Ascaris lumbricoides* in a rural community in Iran.** *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1982, **76:**187-197.
35. Sinniah B: **Daily egg production of *Ascaris lumbricoides*: the distribution of eggs in the faeces and the variability of egg counts.** *Parasitology* 1982, **84:**167-175.
36. Hall A: **Quantitative variability of nematode egg counts in faeces: a study among rural Kenyans.** *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1981, **75:**682-687.
37. Guyatt HL, Bundy DAP, Medley GF, Grenfell BT: **The relationship between the frequency distribution of *Ascaris lumbricoides* and the prevalence and intensity of infection in human communities.** *Parasitology* 1990, **101:**139-143.
38. Seo B-S, Cho S-Y, Chai J-Y: **Egg discharging patterns of *Ascaris lumbricoides* in low worm burden cases.** *Korean Journal of Parasitology* 1979, **17:**98-104.
39. Quinnell RJ, Medley GF, Keymer AE: **The regulation of gastrointestinal helminth populations.** *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 1990, **330:**191-201.
40. McCullagh P, Nelder JA: *Generalized Linear Models* 2nd edition. London: Chapman & Hall; 1989.
41. Lambert D: **Zero-inflated Poisson regression, with an application to defects in manufacturing.** *Technometrics* 1992, **34:**1-14.
42. Hilbe JM: *Negative Binomial Regression* Cambridge: Cambridge University Press; 2007.
43. Ridout M, Demetrio CGB, Hinde J: **Models for count data with many zeros.** *Proceedings of the XIXth International Biometric Conference, Cape Town* 1998:179-192.
44. Welsh AH, Cunningham RB, Donnelly CF, Lindenmayer DB: **Modelling the abundance of rare species: statistical models for counts with extra zeros.** *Ecological Modelling* 1996, **88:**297-308.
45. Denwood MJ, Stear MJ, Matthews L, Reid SWJ, Toft N, Innocent GT: **The distribution of the pathogenic nematode *Nematodirus battus* in lambs is zero-inflated.** *Parasitology* 2008, **135:**1225-1235.
46. Nødtvedt A, Dohoo I, Sanchez J, Conboy G, DesCoteaux L, Keefe G, Leslie K, Campbell J: **The use of negative binomial modelling in a longitudinal study of gastrointestinal parasite burdens in Canadian dairy cows.** *Canadian Journal of Veterinary Research* 2002, **66:**249-257.
47. Hall A, Anwar KS, Tomkins A: **The distribution of *Ascaris lumbricoides* in human hosts: a study of 1765 people in Bangladesh.** *Trans R Soc Trop Med Hyg* 1999, **93(5):**503-510.
48. Hall A, Anwar KS, Tomkins AM: **Intensity of reinfection with *Ascaris lumbricoides* and its implications for parasite control.** *Lancet* 1992, **339:**1253-1257.
49. Abdi YA, Gustaffson LL, Ericson O, Helgren U: *Handbook of Drugs for Tropical Parasitic Infections* London: Taylor & Francis; 1995.
50. Keiser J, Utzinger J: **Efficacy of current drugs against soil-transmitted helminth infections.** *Journal of the American Medical Association* 2008, **299:**1937-1948.
51. Basáñez M-G, Remme JH, Alley ES, Bain O, Shelley AJ, Medley GF, Anderson RM: **Density-dependent processes in the transmission of human onchocerciasis: relationship between the numbers of microfilariae ingested and successful larval development in the simuliid vector.** *Parasitology* 1995, **110:**409-427.

52. Sinden RE, Dawes EJ, Alavi Y, Waldock J, Finney O, Mendoza J, Butcher GA, Andrews L, Hill AV, Gilbert SC, Basáñez M-G: **Progression of *Plasmodium berghei* through *Anopheles stephensi* is density-dependent.** *PLoS Pathogens* 2007, **3:**2005-2016.
53. Kirkwood BR, Sterne JAC: *Essential Medical Statistics* 2nd edition. Oxford: Blackwell Science; 2003.
54. Clayton D, Hills M: *Statistical Models in Epidemiology* Oxford: Oxford University Press; 1993.
55. Akaike H: **New look at statistical model identification.** *IEEE Transactions on Automatic Control* 1974, **AC-19:**716-723.
56. Goldfarb D: **A family of variable metric updates derived by variational means.** *Mathematics of Computation* 1970, **24:**23-26.
57. Broyden CG: **The convergence of a class of double-rank minimisation algorithms.** *Journal of the Institute of Mathematics and its Applications* 1970, **6:**76-90.
58. Shanno DF: **Conditioning of quasi-Newton methods for function minimization.** *Mathematics of Computation* 1970, **24:**647-656.
59. Fletcher R: **A new approach to variable metric algorithms.** *Computer Journal* 1970, **13:**317-322.
60. Ihaka R, Gentleman R: **R: A language for data analysis and graphics.** *Journal of Computational and Graphical Statistics* 1996, **5:**299-314.
61. Venables WN, Ripley BD: *Modern Applied Statistics with S* 4th edition. New York: Springer; 2002.
62. Kleinbaum DG, Kupper LL, Muller KE: *Applied Regression Analysis and Other Multivariate Methods* 2nd edition. Boston: PWS-Kent; 1988.
63. Kutner MH, Nachtsheim CJ, Neter J: *Applied Linear Regression Models* 4th edition. McGraw-Hill/Irwin; 2004.
64. Fox J, Monette G: **Generalized collinearity diagnostics.** *Journal of the American Statistical Association* 1992, **87:**178-183.
65. Ye X-P, Donnelly CA, Fu Y-L, Wu Z-X: **The non-randomness of the distribution of *Trichuris trichiura* and *Ascaris lumbricoides* eggs in the faeces and the effect of stirring faecal specimens.** *Tropical Medicine and International Health* 1997, **2:**261-264.
66. Mullahy J: **Specification and testing of some modified count data models.** *Journal of Econometrics* 1986, **33:**341-365.
67. Dobson AP: *An Introduction to Generalized Linear Models* 2nd edition. London: Chapman & Hall; 2001.
68. Zeileis A, Kleiber C, Jackmon S: **Regression models for count data in R.** *Journal of Statistical Software* 2008, **27:**.
69. Greene WH: **Accounting for excess zeros and sample selection in Poisson and negative binomial regression models.** In *Economics Working Papers* New York: New York University; 1994.
70. Churcher TS, Ferguson NM, Basáñez M-G: **Density dependence and overdispersion in the transmission of helminth parasites.** *Parasitology* 2005, **131:**121-132.
71. Hall A, Hewitt G, Tuffrey V, de Silva N: **A review and meta-analysis of the impact of intestinal worms on child growth and nutrition.** *Maternal and Child Nutrition* 2008, **4(Suppl 1):**118-236.
72. Turner JD, Faulkner H, Kamgno J, Cormont F, Van Snick J, Else KJ, Grencis RK, Behnke JM, Boussinesq M, Bradley JE: **Th2 cytokines are associated with reduced worm burdens in a human intestinal helminth infection.** *Journal of Infectious Diseases* 2003, **188:**1768-1775.
73. Jackson JA, Turner JD, Rentoul L, Faulkner H, Behnke JM, Hoyle M, Grencis RK, Else KJ, Kamgno J, Boussinesq M, Bradley JE: **T helper cell type 2 responsiveness predicts future susceptibility to gastrointestinal nematodes in humans.** *Journal of Infectious Diseases* 2004, **190:**1804-1811.
74. Moqbel R, McLaren DJ: ***Strongyloides ratti*: structural and functional characteristics of normal and immune-damaged worms.** *Experimental Parasitology* 1980, **49:**139-152.
75. Charnov EL: *Life History Invariants. Some Explorations of Symmetry in Evolutionary Ecology* Oxford: Oxford University Press; 1993.
76. Morand S: **Life-history traits in parasitic nematodes: A comparative approach for the search of invariants.** *Functional Ecology* 1996, **10:**210-218.
77. Elkins DB, Haswell-Elkins M, Anderson RM: **The epidemiology and control of intestinal helminths in the Pulicat Lake region of Southern India. I. Study design and pre- and post-treatment observations on *Ascaris lumbricoides* infection.** *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1986, **80:**774-792.
78. Woolhouse MEJ: **A theoretical framework for the immunoepidemiology of helminth infection.** *Parasite Immunology* 1992, **14:**563-578.
79. Anderson RM, May RM: **Herd immunity to helminth infection and implications for parasite control.** *Nature* 1985, **315:**493-496.
80. Berding C, Keymer AE, Murray JD, Slater AF: **The population dynamics of acquired immunity to helminth infection.** *Journal of Theoretical Biology* 1986, **122:**459-471.
81. Martin TG, Wintle BA, Rhodes JR, Kuhnert PM, Field SA, Low-Choy SJ, Tyre AJ, Possingham HP: **Zero tolerance ecology: improving ecological inference by modelling the source of zero observations.** *Ecology Letters* 2005, **8:**1235-1246.
82. Warton DI: **Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data.** *Environmetrics* 2005, **16:**275-289.
83. Lee AH, Wang K, Yau KK, Somerford PJ: **Truncated negative binomial mixed regression modelling of ischaemic stroke hospitalizations.** *Statistics in Medicine* 2003, **22:**1129-1139.