

Analysing risk factors of co-occurrence of schistosomiasis haematobium and hookworm using bivariate regression models: Case study of Chikwawa, Malawi

Bruce B.W. Phiri^a, Bagrey Ngwira^b, Lawrence N. Kazembe^{c,*}

^a Mathematical Sciences Department, Chancellor College, University of Malawi, PO Box 280, Zomba, Malawi

^b Department of Environmental Health, The Polytechnic, University of Malawi, P/Bag 333 Chichiri Blantyre 3, Malawi

^c Department of Statistics and Population Studies, University of Namibia, P/Bag 13301, Pionerspark, Windhoek, Namibia

ARTICLE INFO

Article history:

Received 29 November 2015

Received in revised form 20 February 2016

Accepted 20 February 2016

Available online 2 March 2016

Keywords:

Co-infection

Polyparasitism

Bivariate count models

Hookworm

S. haematobium

Malawi

ABSTRACT

Schistosomiasis and soil-transmitted helminth (STH) infections constitute a major public health problem in many parts of sub-Saharan Africa. In areas where prevalence of geo-helminths and schistosomes is high, co-infection with multiple parasite species is common, resulting in disproportionately elevated burden compared with single infections. Determining risk factors of co-infection intensity is important for better design of targeted interventions. In this paper, we examined risk factors of hookworm and *S. haematobium* co-infection intensity, in Chikwawa district, southern Malawi in 2005, using bivariate count models. Results show that hookworm and *S. haematobium* infections were much localised with small proportion of individuals harbouring more parasites especially among school-aged children. The risk of co-intensity with both hookworm and *S. haematobium* was high for all ages, although this diminished with increasing age, increased with fishing (hookworm: coefficient = 12.29; 95% CI = 11.50–13.09; *S. haematobium*: 0.040; 95% CI = 0.0037, 3.832). Both infections were abundant in those with primary education (hookworm: coef. = 0.072; 95% CI = 0.056, 0.401 and *S. haematobium*: coef. = 0.286; 95% CI = 0.034, 0.538). However, much lower risk was observed for those who were farmers (hookworm: coef. = -0.349, 95% CI = -0.547, -0.150; *S. haematobium*: coef. -0.239, 95% CI = -0.406, -0.072). In conclusion, our findings suggest that efforts to control helminths infection should be co-integrated and health promotion campaigns should be aimed at school-going children and adults who are in constant contact with water.

© 2016 The Authors. Published by Elsevier Ltd on behalf of World Federation of Parasitologists.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Hookworm and schistosomiasis are some of the helminth infections that are prevalent in most tropical and sub-Saharan countries. However, of the two infections, hookworm is more prevalent than schistosomiasis. Estimates indicate that 1.2 billion people worldwide are infected by hookworm (WHO, 1996) and the infection is widely distributed throughout tropical and subtropical areas with prevalence in some communities as high as 90% (Hotez et al., 2003). With regards to Schistosomiasis, an estimated 207 million people worldwide are infected and that 85% of all cases are now in sub-Saharan Africa (Mbabazi et al., 2011). Overall,

* Corresponding author.

E-mail addresses: bruce.williams31@gmail.com (B.B.W. Phiri), bagreyngwira@gmail.com (B. Ngwira), lkazembe@yahoo.com (L.N. Kazembe).

helminth infections affect between 20% and 30% of the general population with prevalence as high as 60–80% in endemic areas (Pan American Health Organization (PAHO), 1997).

Polyparasitism is common in regions where different parasites co-exist (Pan American Health Organization (PAHO), 1997; WHO, 2002; Bethony et al., 2006; Raso, 2004; Utzinger and Keiser, 2004; Pullan and Brooker, 2008), and predominantly in rural areas or where there is poor sanitation found in sub-Saharan Africa, Southeast Asia and tropical regions of the Americas (Pan American Health Organization (PAHO), 1997; WHO, 2002; Bethony et al., 2006). The large overlap in the geographic distribution of geo-helminths implies that co-infection is a norm than an exception (Raso, 2004). Globally, several million children could be concurrently infected with multiple helminth species even at low intensity (Utzinger and Keiser, 2004). Polyparasitism is most prevalent in school-aged children, while adults remain at considerable risk of harbouring multiple helminth species, although at reduced intensity, and are often neglected in the helminthological literature (Pullan and Brooker, 2008). The multiple infections due to helminths are believed to persist throughout the life course, resulting in reduced physical capacity for work (Gilgen et al., 2001). Moreover, co-infection exacerbates common illnesses such as anaemia in infected population. New research provides evidence that the risk of anaemia is amplified in children simultaneously infected with hookworm and *Schistosoma* or hookworm and *Trichuris*, when compared to the sum of risks for children with singular infections (Hotez et al., 2003).

Although co-infections may simply arise by chance; shared risk factors have been a major contribution (Utzinger and Keiser, 2004; Pullan and Brooker, 2008; Gilgen et al., 2001; Mwangi et al., 2006). Fewer studies, however, have examined such (Bethony et al., 2006; Raso, 2004; Utzinger and Keiser, 2004; Pullan and Brooker, 2008; Gilgen et al., 2001; Mwangi et al., 2006). Interest in understanding polyparasitisms or multiple helminths infections has been rekindled with the goal for integrated resource deployment, but also that if risk factors are shared, prevention efforts will be much easier than if separate (Utzinger and Keiser, 2004; Pullan and Brooker, 2008; Gilgen et al., 2001; Mwangi et al., 2006; Centre for Food Security and Public Health, 2005).

Several methods have been employed in the modelling of helminths co-infection. However, most scholars have used univariate models, see for example (Lwambo et al., 1992). Recently, Magalhães et al. (Magalhães et al., 2011) reported use of multinomial geostatistical regression models in predicting *S. haematobium*-hookworm co-infections. Further applications of multinomial spatial models can be found in (Raso et al., 2006; Brooker and Clements, 2009), who used these models for predicting the risk of co-infection with multiple helminth infections. The multinomial approach involves stratifying egg counts, leading to a loss of information whereas the Poisson or the negative binomial approach make full use of infection intensity data on a continuous scale as measured by number of eggs found in both slides per individual. A similar multi-categorical approach is reported in Botelho et al. (Botelho et al., 2008), who used proportional odds models to investigate the relationship between hookworm and *Ascaris lumbricoides* infection. In another study, Sturrock et al. (Sturrock et al., 2013) used bivariate logistic regression for joint spatial analysis of questionnaire and parasitological data in order to predict the prevalence of *S. haematobium* infection for schools with missing questionnaire data. Schur et al. (Schur et al., 2011), again using the multinomial construct, employed a Bayesian geostatistical shared component models (which allows for covariates, disease-specific and shared spatial and non-spatial random effects) to model the geographical distribution and burden of co-infection risk from single-disease surveys in Côte d'Ivoire.

According to our knowledge, little, if any, literature exist that considered co-infection that employs the use of bivariate count models, despite several applications elsewhere (Karlis, 2003; Karlis and Ntzoufras, 2005; Gurmu and Elder, 2011; Efron and Bradley, 1986; Zou et al., 2011; Lao et al., 2011; Jung and Winklemann, 1993; Bermudez and Karlis, 2011). This study, therefore, used bivariate count models to model co-infection of hookworm and *S. haematobium*, to better understand the epidemiology of the two and their co-intensity and account for outcome dependency.

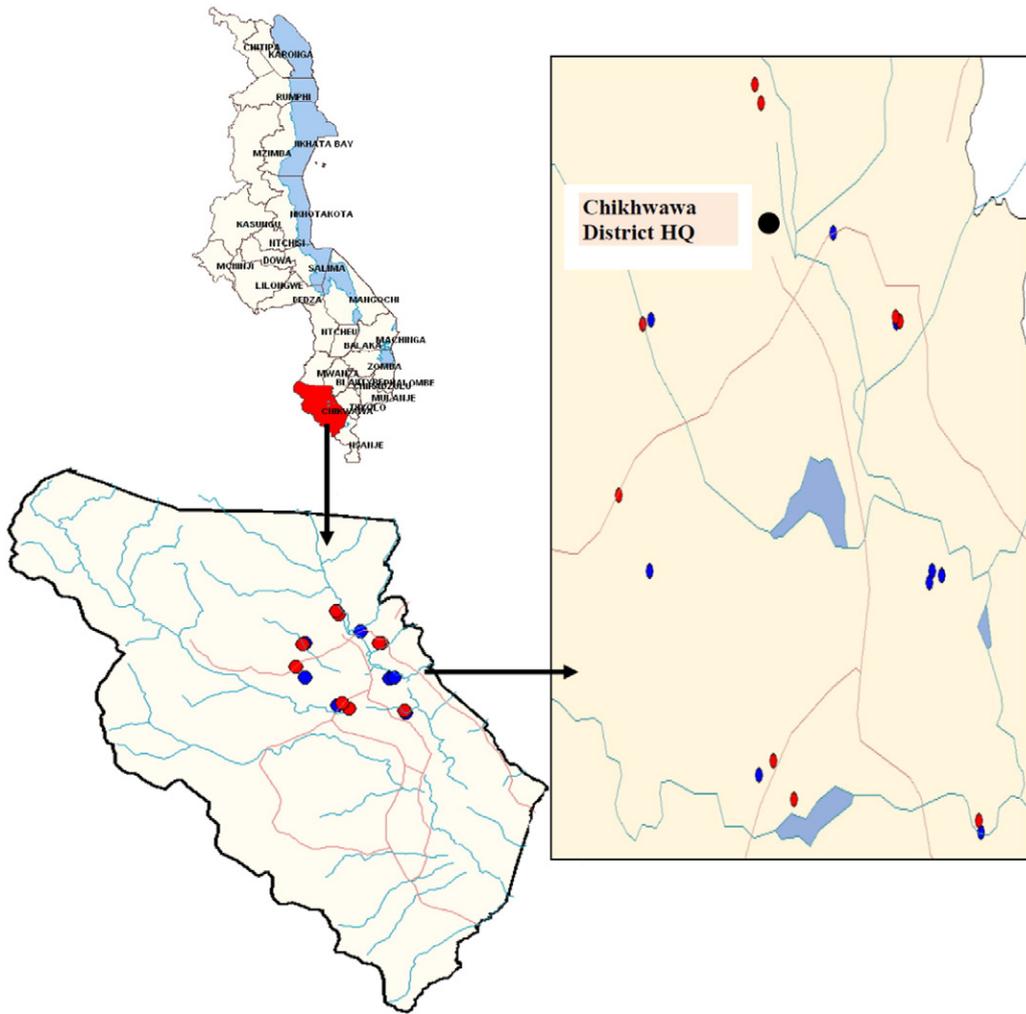
2. Materials and methods

2.1. Dataset

The dataset used was collected in 2005 from a cluster randomised study conducted in Chikwawa district in the lower Shire Valley, southern Malawi (Fig. 1). Full details of the study are found elsewhere (Ngwira, 2005). In brief, the study was designed as follows. Subjects aged 1 year and above were drawn from each selected household to participate in the study, after informed consent was obtained. All consenting participants were given a full body clinical examination for chronic manifestation of human helminth. In addition, each participant was requested to provide a fresh stool and urine sample transported in cooler boxes and examined in the laboratory at Montfort hospital, Nchalo, Chikwawa. Stool samples were examined by a single thick smear technique using *Kato-Katz* template then examined under a light microscope to observe parasite eggs. Ova for each parasite observed were counted and expressed as eggs per gram of stool (EPG). Urine samples were centrifuged at 300 rpm for 5 min and sediments were then examined under a light microscope to detect number of parasite eggs. Helminth infection intensity is measured by counting microscopic eggs that host excretes through urine and faecal. The Malawi College of Medicine Research Ethics Committee (COMREC) and the Ethics Committee of London School of Hygiene and Tropical Medicine (LSHTM) approved the study.

2.2. Statistical methods

Various statistical models have been developed to model helminths infection. The egg counts, obtained from a human urine sample, are a measure of infection intensity and can be modelled as a count variable. Given certain individuals will harbour two or more parasite infections, bivariate models are developed to jointly model co-infection. Four types of models were



Scale 1 cm: 2 km

Fig. 1. Map of Chikhwawa District showing trial villages. Insert showing the geographical location of trial villages at a smaller scale. Intervention villages are in red and control villages are in blue.

investigated and include bivariate Poisson (BP), double Poisson (DP), diagonal inflated bivariate Poisson (DIBP), and bivariate zero inflated Poisson (BIZIP) models. The BP is a basic model for analysing two correlated count data, in our case, Schistosomiasis and hookworms egg counts. Where the two are uncorrelated, the double Poisson is assumed. In some cases, there are many zeros in the data, and in such situations, the DIBP or BIZIP can be used. Detailed treatise of the models can be found in reference (Karlis, 2003; Karlis and Ntzoufras, 2005; Gurmu and Elder, 2011; Efron and Bradley, 1986; Zou et al., 2011; Lao et al., 2011; Jung and Winklemann, 1993; Bermudez and Karlis, 2011).

In brief, the models are defined as follows. The BP regression model can be generated by convolutions of Poisson random variables. Let Y_1 and Y_2 be the egg counts in hookworm and *S. haematobium*, then the observed counts are jointly modelled using the bivariate Poisson distribution, which is given by

$$f_{BP}(x, y | \lambda_1, \lambda_2, \lambda_3) = e^{-\lambda_1 - \lambda_2 - \lambda_3} \sum_{l=0}^{\min(y_1, y_2)} \frac{\lambda_3^l}{l!} \frac{\lambda_1^{y_1-l}}{(y_1-l)!} \frac{\lambda_2^{y_2-l}}{(y_2-l)!} \quad (1)$$

where λ_1, λ_2 and λ_3 denote three Poisson parameters, such that the means intensity of hookworm and *S. haematobium* are given by $E(Y_1) = \lambda_1 + \lambda_3; E(Y_2) = \lambda_2 + \lambda_3$, and the measure of co-infection intensity is captured by λ_3 i.e. $Cov(Y_1, Y_2) = \lambda_3$, respectively. The bivariate Poisson regression model takes the following form:

$$\begin{aligned} (Y_{1i}, Y_{2i}) &\sim BP(\lambda_{1i}, \lambda_{2i}, \lambda_{3i}) \\ \ln(\lambda_{1i}) &= x_{ki}^T \beta_k \end{aligned} \quad (2)$$

for a given set of explanatory variables (x_{ki}) , $k = 1, 2, 3; i = 1, 2, \dots, n$. The term β_k , denotes the corresponding vector of regression coefficients. This model assumes greater flexibility where each parameter is related to some covariates.

If the data is over- or under-dispersed, the BP model may not be an appropriate model. To handle over-dispersion or under-dispersion, one can use several alternatives. A possible alternative is to use the double Poisson (DP). The double Poisson is obtained if $\lambda_3 = 0$. Then the two variables (Y_1, Y_2) are independent and the bivariate Poisson distribution reduces to the product of two independent Poisson distributions, often referred to as the double Poisson.

Another model for correlated count data is the diagonal inflated bivariate Poisson (DIBP) model. The model allows inflation in the diagonal elements, and is an extension of the simple zero inflated model that allows only for an excess in $(0, 0)$ counts (Zou et al., 2011; Lao et al., 2011). The DIBP model can be defined on the basis of the bivariate Poisson regression model as follows:

$$f_{DIBP}(Y_1, Y_2) = \begin{cases} (1-p)f_{BP}(Y_1, Y_2|\lambda_1, \lambda_2, \lambda_3) & Y_1 \neq Y_2 \\ (1-p)f_{BP}(Y_1, Y_2|\lambda_1, \lambda_2, \lambda_3) + pf_D(Y_1; \theta) & Y_1 = Y_2 \end{cases} \quad (3)$$

such that $f_D(Y_1; \theta)$ defines a discrete count distribution on the set $\{0, 1, 2, \dots\}$ with parameter vector θ ; and p is a mixing proportion. If for $p=0$, we have the simple bivariate Poisson model given in Eq. (1).

Related to the diagonal inflated models are the zero inflated bivariate Poisson models, also referred to as zero modified count models. These are used when the observed data displays a high frequency of the zero-zero state. The bivariate zero inflated Poisson (BIZIP) is developed as a mixture of a bivariate Poisson, two univariate Poisson, and a mass point at $(0, 0)$. As such

$$(Y_1, Y_2) \sim \begin{cases} (0, 0) & \text{with probability } p_0 \\ (Po(\lambda_1 + \lambda_3), 0) & \text{with probability } p_1 \\ (0, Po(\lambda_2 + \lambda_3)) & \text{with probability } p_2 \\ BP(\lambda_1, \lambda_2, \lambda_3) & \text{with probability } p_3 \end{cases} \quad (4)$$

Table 1

Characteristics for individuals who had *S. haematobium* and hookworm ($n = 1642$).

Variable	Mean	Std. Dev.	Number (%)	<i>S. haematobium</i> ($n = 233, 14\%$)	Hookworm ($n = 324, 20\%$)
Age (years)	21.7	17.77	-	-	-
Sex					
Male			733 (44.6)	134 (8.0)	137 (8.34)
Female			909 (55.4)	99 (6.0)	187 (11.4)
Education					
None			745 (45.4)	146 (9.0)	95 (5.8)
Primary			850 (51.8)	174 (10.6)	132 (8)
Secondary			47 (2.9)	4 (0.24)	6 (0.4)
Fishing					
Yes			221 (13.5)	51 (3.0)	71 (4.3)
No			1421 (86.5)	182 (11.0)	253 (15.4)
Bathing					
Yes			1280 (78)	205 (12.0)	273 (16.6)
No			362 (22)	28 (1.7)	51 (3.1)
Dimba					
Yes			960 (58.5)	172 (10.5)	218 (13.3)
No			682 (41.5)	61 (3.7)	106 (6.5)
Occupation					
Farmer			733 (44.6)	101 (6.2)	220 (13.4)
Other			909 (55.4)	132 (8.0)	104 (6.3)
Income					
Yes			634 (38.6)	73 (4.44)	43 (2.6)
No			386 (23.5)	78 (4.8)	130 (7.9)
missing			622 (37.9)	82 (5.0)	151 (9.2)
Toilet					
Yes			526 (32.0)	78 (4.8)	81 (4.9)
No			482 (29.4)	72 (4.40)	89 (5.4)
missing			634 (38.6)	83 (5.1)	154 (9.4)
Polyparasitism					
none			807 (49.2)	0	0
One			594 (36.2)	117 (7)	167 (10.2)
Two			200 (12.2)	81 (5)	129 (7.9)
Three			38 (2.3)	33 (2)	25 (1.5)
four			3 (0.2)	2 (0.12)	3 (0.18)
Deworming					
No			805 (74.7)	165 (10)	259 (15.8)
Yes			273 (25.3)	68 (4.14)	65 (4)

where $p_3 = 1 - p_0 - p_1 - p_2$. Hence, we can write $(Y_1, Y_2)' \sim \text{BIZIP}(\lambda_1, \lambda_2, \lambda_3, p_1, p_2, p_3)$ if the two variables $(Y_1, Y_2)'$ follow the distribution given in Eq. (4).

The four models, BP, DP, DIBP, and BIZIP, were explored and estimated in R statistical software and computing system using *BIVPOIS* package (Karlis and Ntzoufras, 2005). The Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) were used for model selection. Models with a small AIC or BIC, were considered better than others. Standard errors for the models were generated through bootstrapping procedure simply because the algorithms for the models employed do not take into account standard errors. The bootstrap standard errors were used for assessing significance of regression coefficients in fitted models. Descriptive analyses were performed in *STATA* (version 12).

3. Results

Table 1 gives summaries of characteristics for study participants. In total, the study recruited 1642 individuals of which 733 (44.6%) were males and 909 (55.4%) were females. The mean age was 21.7 years (standard deviation [SD] = 17.7), of which 903 (55%) were aged less than 20 years. Most people completed primary education ($n = 850$, 51.8%) and only 47 (2.8%) had secondary education ($p < 0.001$).

3.1. Prevalence of infection

The overall prevalence of hookworm and *S. haematobium* was 324 (20%) and 233 (14%) respectively while 48 (4.5%) individuals had both hookworm and *S. haematobium*. Both response variables were positively skewed with mean egg count of 1.89 (SD = 6.86) for hookworm, and 2.57 (SD = 8.11) for *S. haematobium*. Fig. 2 shows the distribution of both outcomes, while Fig. 3 presents the distribution of prevalence of *S. haematobium* and hookworm across different age categories. *S. haematobium* infection was found to increase sharply for the young children (≤ 5 years) to those aged 6–15 years and decreased in female adults aged greater than 15 years and male adults aged greater than 30 years. Hookworm infection increased sharply as age increased.

Multiple parasite infections were recorded in 20.8% of the study population. The majority of those with multiple parasite infections (20.5%) had two to three parasite species per individual and only 0.3% harboured four parasites ($p < 0.001$).

Table 2 presents bivariate frequency table for parasite counts for hookworm and *S. haematobium* after removing all missing observations. The data have two interesting features: over-dispersion and a very high proportion of non-egg execrators of which 77.1% for hookworm and 80.7% for *S. haematobium*, with the sum of the diagonal value approximately at 63% of the total data.

3.2. Risk factors of co-infection of hookworm and *S. haematobium*

Table 3 shows parameters for three models fitted with same covariates on each response variable (hookworm and *S. haematobium*). Overall, the zero inflated bivariate Poisson model was considered the best fitted model because it had the lowest AIC (= 7447.40) or BIC (= 7625.85), and captured about 62% of zeros in the data.

Table 4 provides estimates based on the zero inflated bivariate Poisson models. Estimates from univariate and bivariate Poisson model are also presented to check the efficiency of bivariate Poisson model to determine the risk factors for hookworm and

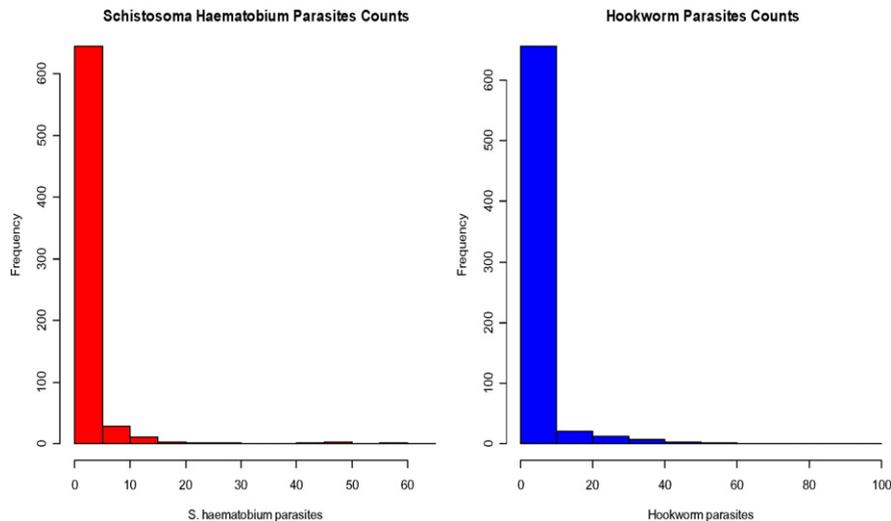


Fig. 2. Egg counts (per gram/ml) for *S. haematobium* and hookworm parasites.

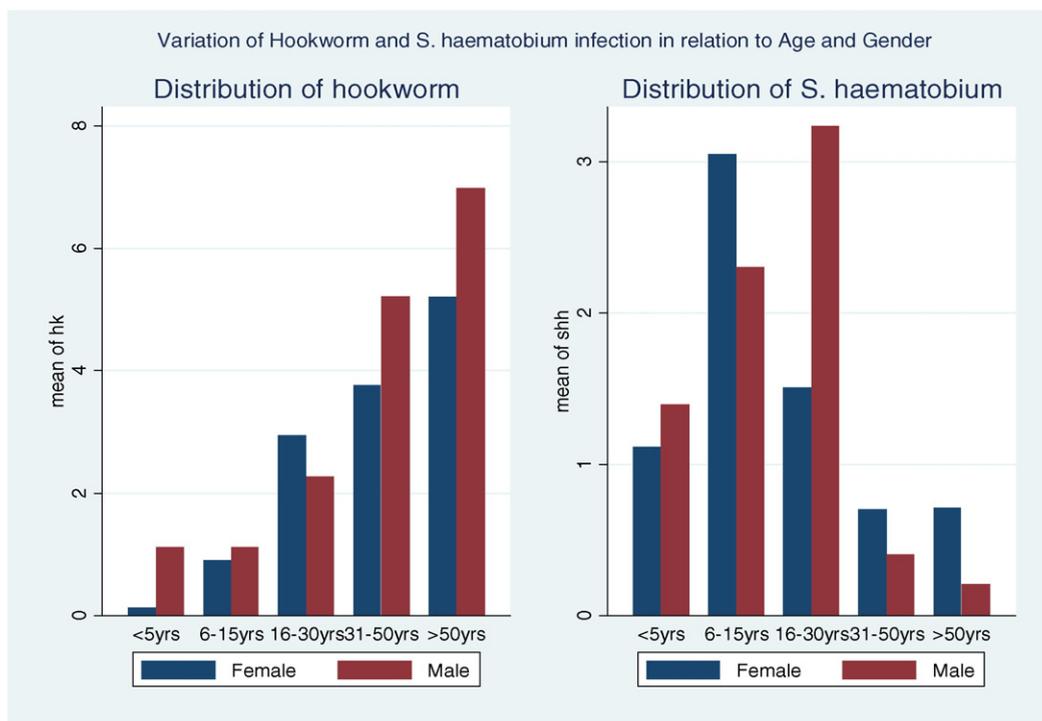


Fig. 3. Variation of hookworm and *S. haematobium* intensity in relation to age and sex of respondent.

S. haematobium infections. Positive (negative) values of the coefficients indicate an increased (reduced) risk of helminth occurrences and intensity.

From the table, the probability of hookworm infection was higher among those aged between 6 and 15 years (λ_1 : coef. = 1.27; 95% CI = 0.46, 2.09), however, generally infection decreased as age increased. There was a positive association between hookworm infection and gender, with male showing high risk than females although the difference was statistically insignificant (λ_1 : coef. = 1.03; 95% CI = -2.72, 4.80). We also observed strong positive association between hookworm and fishing in the Shire river (λ_1 : coef. = 12.29; 95% CI = 11.50, 13.08), whereas working in dimba gardens showed a negative association (λ_1 : coef. = -1.13; 95% CI = -1.375, -0.89). On the contrary, occupation (farmer/other) showed a negative association with infection probability though with marginal significance (λ_1 : coef. = -0.34; 95% CI = -0.54, -0.15). Having a toilet facility had a positive association with hookworm (λ_1 : coef. = 0.29; 95% CI = 0.186, 0.40), contrary to our expectation.

In the same understanding, the probability of *S. haematobium* infection was found to decrease with the increase in age (λ_2 : coef. = -0.06; 95% CI = -0.08, -0.04). The infection was less prevalent among the population aged more than 50 years (λ_2 : coef. = -1.15; 95% CI = -1.56, -0.73). Pre-school children showed high infection intensity relative to those that are in primary school level (λ_2 : coef. = 1.74; 95% CI = 1.22, 2.26). There was a strong association between the risk of *S. haematobium* with bathing in the Shire river, (λ_2 : coef. = 1.74; 95% CI = 1.22, 2.26) and working in the dimba along the Shire valley, (λ_2 : coef. = 0.28; 95% CI = 0.03, 0.53). However, working as a farmer had negative association with *S. haematobium* (λ_2 : coef. = -0.23; 95% CI = -0.40, -0.07) than other occupations.

Table 2

Bivariate frequency for hookworm and *S. haematobium* parasite counts.

Number of <i>S. haematobium</i> parasites											
Number of hookworm parasite	0	1	2	3	4	5	6	7	≥8	Cumulated records	
0	677	23	30	19	16	9	9	7	75	865	
1	12	0	1	1	0	0	0	0	3	17	
2	37	2	1	1	1	1	2	3	5	53	
3	12	0	1	0	2	0	1	0	3	19	
4	21	0	2	0	0	0	0	0	1	24	
5	10	0	0	0	1	0	0	0	1	12	
6	9	0	1	0	0	0	0	0	1	11	
7	6	0	0	0	0	0	0	0	0	6	
≥8	58	2	5	0	0	1	0	0	5	71	
Cumulated records	842	27	41	21	20	11	12	10	94	1078	

Table 3

Summary of models fitted with corresponding model selection criteria.

Model		Diagonal Distribution	Parameter	2LL	AIC	BIC
1	BP	NO	33	−5819.10	11,624.76	11,797.96
2	DP	NO	32	−5779.38	11,622.76	11,790.71
3	DIBP	Discrete(0)	34	−4818.00	7447.40	7625.85
4	DIBP	Discrete(1)	35	−5054.14	7449.40	7633.10
5	DIBP	Discrete(2)	36	−5178.57	7451.40	7640.34

Models compared are Double Poisson (DP); Bivariate Poisson (BP); Diagonal Inflated Bivariate Poisson (DIBP); 2LL: Log-likelihood; AIC: Akaike information criterion; BIC: Bayesian information criterion.

4. Discussion

This study revealed that polyparasitism was common in Chikhwawa, Malawi. The presence of four helminth parasite species was confirmed with 14.5% of the sample harbouring two or more parasite species, and points to common risk factors (Pan American Health Organization (PAHO), 1997; WHO, 2002; Bethony et al., 2006; Raso, 2004; Utzinger and Keiser, 2004; Pullan and Brooker, 2008; Gilgen et al., 2001). Our findings reveal co-occurrence of helminths in individuals below 30 years of age, particularly in children aged 6–15 years. Low socio-economic status, including having no education or obtaining up to primary education increased the risk of co-occurrence of helminth infection (Nguhiu et al., 2009; Xie et al., 2013). Having secondary or higher education may correspond to increased awareness and access to treatment hence reduced infection intensity (Xie et al., 2013).

Contact with contaminated freshwater is the major risk factor of *S. haematobium* infection. In this study, bathing and fishing were found to positively affect *S. haematobium* infection since they involve extensive contact with infected water. Most people in Chikwawa work in cane estates to earn a living (Spear et al., 2004) as such they constantly waded in contaminated water and soil. Few studies that have reported on specific occupations as a risk factor for infection have observed high prevalences in occupational groups with more intensive water contact such as fishing and dimba gardening (Nguhiu et al., 2009). However, being a farmer (Xie et al., 2013) had a positive association with infection intensity. Among the socio-economic determinants, the estimated coefficients showed that working as a farmer, bathing in Shire river, and cultivating in dimba gardens propel the occurrence of hookworm and *S. haematobium* infections in Chikhwawa district.

Bearing in mind that both hookworm and Schistosomiasis are water-related diseases, this finding confirms the dependence of their modes of transmission in the area. Indeed, the positive correlation along the diagonal as depicted in Table 2, implies co-occurrence and transmission patterns of the *S. haematobium* and hookworm are highly correlated in Chikwawa, Malawi. In fact, considering the various ways of transmission for the two infections, the findings strongly suggest common or shared risk factors, be it soil, water and food contamination, that inform the joint occurrence of parasites. Such phenomenon can be exploited to design a single approach for deploying interventions.

Nevertheless, differential patterns of associations were also established. *S. haematobium* was found to be more prevalent among children than adults. This result agrees with what Ngwira (2005); Spear et al. (2004); Nguhiu et al. (2009); Xie et al. (2013) found that prevalence and intensities of hookworm infection generally show a typical convex-shaped curve with a peak at the ages of 6–15 years, and a decrease in adults. Adults have the possibility to develop immunity to schistosome infections than children.

Although common risk factors may lead to increased burden of polyparasitism, several studies have shown that infection with one parasite pose a risk exposure to another infection (Mwangi et al., 2006; Centre for Food Security and Public Health, 2005; Lwambo et al., 1992; Magalhães et al., 2011; Raso et al., 2006). However, there is complex transmission pattern because of synergism and antagonism in multiple species infections. For epidemiological profiling, describing shared risk factors is a critical step towards understanding the complexity of transmission. Be as it may, our understanding of the interaction of parasites is limited in the present study, because it was an observational study.

The study, further, demonstrated the capability of bivariate count models in quantifying the risk of hookworm and *S. haematobium* infections and check their efficiency over univariate count models. Indeed, in contrast to regular univariate count models, the bivariate count models, pertaining to this study, had two dependent predictors which quantify the effects on the reported hookworm and the *S. haematobium* infections, respectively which is not the case with univariate count models which deals with one dependent variable. Although the estimates from the fitted models seem to predict same risk factors for both hookworm and *S. haematobium*, however, the bivariate models seem to differ in the significance of some of the risk factors. For instance basing on the bivariate Poisson model, shows a positive association with hookworm and *S. haematobium* infection. This is not true with univariate Poisson model where it shows a negative association with *S. haematobium*. Generally, the standard errors reported in bivariate model are smaller than those in the univariate model, thus producing inefficient model. This underscores the fact that since the two infections are strongly correlated, ignoring dependence in the two infections may lead to wrong identification of risk factors, hence wrong epidemiological profiling of polyparasitism. This finding is consistent with what Ngwira (2005) found. Indeed, the main advantage in using bivariate count models over univariate count models is clearly an increase in efficiency of parameter estimation.

Our study is not without limitations. First, in our analysis, we assumed that the bivariate Poisson models performed well, despite some possible over-dispersion due to many zeros. We assumed that over-dispersion was negligible especially if zero-adjusted and when covariates are included (Karlis, 2003; Karlis and Ntzoufras, 2005). However, in practice, overdispersed

Table 4Regression estimates obtained from fitting bivariate Poisson models on hookworm and *S. haematobium* co-infection. Also given are estimates from a univariate Poisson model.

Hookworm	Univariate model		Bivariate Poisson		Diagonal inflated bivariate Poisson				
	Coef	95% CI	Coef	95% CI	Coef	95% CI			
Age									
0–5 years	2.51	0.0024	5.02	1.36	0.0004	2.72	0.583	–0.10	1.27
6–15 years	2.15	0.0031	4.30	9.68	–0.0015	19.37	1.275	0.46	2.09
16–30 years	0.99	0.0014	1.98	8.22	0.0016	16.45	0.321	–0.33	0.96
31–50 years	1.07	0.0028	2.14	7.61	–0.0001	15.22	0.424	0.08	0.76
>50 years	0			0			0	0	
Education									
No	–0.25	0.0005	–0.49	–6.34	–6.3362	–6.35	–0.132	–0.56	0.30
Primary	–0.32	0.0005	–0.63	2.25	0.0013	4.49	0.172	–0.06	0.40
Secondary	0			0			0	0	
Sex									
Male	0.26	–0.0022	0.52	1.74	0.0006	3.49	1.039	–2.72	4.80
Female	0			0			0	0	
Bathing in Shire river									
Yes	–0.44	0.0012	–0.88	1.37	–0.0008	2.75	1.032	–2.64	4.70
No	0			0			0	0	
Fishing in Shire									
Yes	13.31	0.0816	26.54	1.39	1.2489	1.53	12.29	11.50	13.09
No	0			0			0	0	
Working in dimba									
Yes	0.59	–0.0003	1.18	2.58	–0.0002	5.17	–1.136	–1.37	–0.90
No	0			0			0	0	
Regular Income									
Yes	–1.02	–2.0446	0.00	–1.12	10.1061	–12.35	0.216	–0.11	0.54
No	0			0			0	0	
Toilet facility									
Yes	–0.41	0.0009	–0.81	–2.84	–0.0005	–5.68	0.294	0.19	0.42
No	0			0			0		
Occupation									
Farmer	0.50	–0.0017	1.01	7.52	–0.0032	15.04	–0.349	–0.55	–0.15
Others	0			0			0		
S. HAEMATOBIIUM									
Age									
0–5 years	0.29	0.0008	0.57	0.31	0.0001	0.63	–0.819	–1.69	0.05
6–15 years	1.16	–0.0004	2.32	1.04	–0.0005	2.07	0.03	–0.80	0.86
16–30 years	0.24	–0.0002	0.47	0.29	–0.0003	0.58	–0.347	–1.07	0.38
31–50 years	–0.09	–0.0002	–0.17	–0.26	0.0002	–0.53	–0.862	–1.32	–0.33
>50 years	0			0			0		
Education									
No	0.99	0.0007	1.98	1.11	0.0004	2.22	1.744	1.22	2.27
Primary	1.01	–0.0011	2.02	1.24	0.0007	2.48	0.286	0.03	0.56
Secondary	0			0			0		
Sex									
Male	0.07	–0.0003	0.14	0.12	0.0014	0.23	2.456	–0.75	5.66
Female	0			0			0		
Bathing in Shire									
Yes	1.49	0.0020	2.97	0.75	0.3356	1.17	2.581	–0.48	5.56
No	0			0			0		
Fishing in Shire									
Yes	–1.30	–0.0003	–2.61	–1.07	–0.0002	–2.15	0.04	0.0032	0.82
No	0			0			0		
Working in dimba									
Yes	0.85	–0.0002	1.70	0.80	0.0002	1.60	0.217	–0.0770	0.551
No	0			0			0		
Regular Income									
Yes	0.61	–0.0003	1.22	0.69	–0.0010	1.38	0.365	–0.0130	0.744
No	0			0			0		

Table 4 (continued)

Hookworm	Univariate model		Bivariate Poisson			Diagonal inflated bivariate Poisson			
	Coef	95% CI	Coef	95% CI	Coef	95% CI	Coef	95% CI	
<i>Toilet facility</i>									
Yes	−0.28	0.0014	−0.55	−0.04	0.0006	−0.08	−0.002	−0.1260	0.122
No	0			0			0		
<i>Occupation</i>									
Farmer	0.46	0.0015	0.92	0.56	0.0013	1.12	−0.239	−0.4100	−0.072
Others	0			0			0		

Models compared are univariate Poisson, bivariate Poisson and Diagonal inflated bivariate model. Standard errors were generated from 1000 bootstrap of estimate coefficients.

count data are well fitted using negative binomial, which is a natural candidate model (Xie et al., 2013). A review of models for skewed parasite counts is provided in Alexander (Alexander, 2012). Although the results are only presented for two correlated parasites, the models can be extended to multivariate Poisson regression of which polyparasitism can be analysed (Alexander, 2012). Second, in our study we used egg counts as indirect measure of worm burden. While faecal egg counts may give measurement error, as they fluctuate daily, many studies, with exception of a few (Bradley et al., 1992), egg counts are seen to be feasible to collect and measure intensity from a large population as opposed to using actual worm burden.

5. Conclusion

Bivariate count models were used to analyse the risk factors for hookworm and *S. haematobium* infections. They were employed to examine the relation of co-occurrence of infection to factors such as age, gender, education level, fishing in Shire river, bathing in Shire river, working in dimba gardens along Shire river, income and nature of occupation. This led to identification of some risk factors that could be used in the control programme planning for hookworm and *S. haematobium*.

This analysis has also demonstrated that the multiple parasite species infections were higher among 6–15 year age group. Presence of multiple parasite species in an individual may be associated with severe morbidity and complication. Therefore, any control initiative that targets only a single parasite species may not be effective in reducing morbidity due to helminth parasitic infections. An integrated approach targeting the major parasite species present would be recommended. Therefore, vaccines for hookworm and intestinal *S. haematobium* could be combined in a multivalent anthelmintic vaccine (Hotez et al., 2003), which may increase vaccine efficiency and reduce the timeframe for widespread distribution in affected areas of Africa and Latin America.

Author contributions

Conceived and designed the experiments: BN. Performed the experiments: BN. Analysed the data: BBW LNK. Contributed reagents/materials/analysis tools: BN LNK BBW. Wrote the paper: BBW BN LNK.

References

- Alexander, N., 2012. Review: analysis of parasite and other skewed counts. *Tropical Med. Int. Health* 17 (6), 684–693.
- Bermudez, L., Karlis, D., 2011. Bayesian multivariate Poisson models for insurance ratemaking. *Insurance* 48, 226–236.
- Bethony, J., Brooker, S., Albonico, M., Hotez, P.J., 2006. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. *Lancet* 367, 1521–1532.
- Botelho, A.J., Raff, S., Rodrigues, R., Hoffman, H.J., Diemert, D.J., Oliveira, R.C., Bethony, J.M., Gazzinelli, M.F., 2008. Hookworm, *Ascaris lumbricoides* infection and polyparasitism associated with poor cognitive performance in Brazilian schoolchildren. *Tropical Med. Int. Health* 13 (8), 994–1004.
- Bradley, M., Chandiwana, S.K., Bundy, D.A., Medley, G.F., 1992. The epidemiology and population biology of *Necator americanus* infection in a rural community in Zimbabwe. *Trans. R. Soc. Trop. Med. Hyg.* 86, 73–76.
- Brooker, S., Clements, A.C.A., 2009. Spatial heterogeneity of parasite co-infection: determinants and geo-statistical prediction at regional scales. *Int. J. Parasitol.* 39 (5), 591–597.
- Centre for Food Security and Public Health, 2005. Hookworms. Retrieved November 13, 2015 from <http://www.cfsph.iastate.edu>.
- Efron, B., Bradley, P.C., 1986. Double exponential families and their use in generalized linear regression. *J. Am. Stat. Assoc.* 81, 709–721.
- Gilgen, D.D., Mascie-Taylor, C.G., Rosetta, L.L., 2001. Intestinal helminth infections, anaemia and labour productivity of female tea pluckers in Bangladesh. *Tropical Med. Int. Health* 6, 449–457.
- Gurmu, S., Elder, J., 2011. Flexible bivariate count data regression models. Andrew Young School of Policy Studies Research Paper Series. Working Paper, pp. 11–33.
- Hotez, P.J., Zhan, B., Bethony, J.M., Loukas, A., Williamson, A., Goud, G., Hawdon, J.M., Dobardzic, A., Dobardzic, R., Ghosh, K., Bottazzi, M.E., Mendez, S., Zook, B., Wang, Y., Liu, S., Essiet-Gibson, I., Chung-Debose, S., Xiao, S.H., Knox, D., Meagher, M., Inan, M., Correa-Oliveira, R., Vilk, P., Shepherd, H.R., Brandt, W., Russell, P.K., 2003. Progress in the development of a recombinant vaccine for human hookworm disease: the human hookworm vaccine initiative. *Int. J. Parasitol.* 33, 1245–1258.
- Jung, R.C., Winklemann, R., 1993. Two aspects of labour morbidity: A bivariate regression approach. *Empir. Econ.* 18, 543–556.
- Karlis, D., 2003. An EM algorithm for multivariate Poisson distribution and related models. *J. Appl. Stat.* 30 (1), 63–77.
- Karlis, D., Ntzoufras, I., 2005. Bivariate Poisson and diagonal inflated bivariate Poisson regression models in R. *J. Stat. Softw.* 14 (10), 1–36.
- Lao, Y., Wu, Y.-J., Corey, J., Wang, Y., 2011. Modelling animal-vehicle collisions using diagonal inflated bivariate Poisson regression. *Accid. Anal. Prev.* 43 (1), 220–227.
- Lwambo, N.J.S., Bundy, D.A.P., Medley, G.F.H., 1992. A new approach to morbidity risk assessment in hookworm endemic communities. *Epidemiol. Infect.* 108, 469–481.

- Magalhães, R.J.S., Biritwum, N.K., Gyaopong, J.O., Brooker, S., Zhang, Y., Blair, L., Fenwick, A., Clements, A.C.A., 2011. Mapping helminth co-infection and co-intensity: geostatistical prediction in Ghana. *PLoS Negl. Trop. Dis.* 5 (6), e1200.
- Mbabazi, P.S., Andan, O., Fitzgerald, D.W., Chitsulo, L., Engels, D., Downs, J.A., 2011. Examining the relationship between urogenital schistosomiasis and HIV Infection. *PLoS Negl. Trop. Dis.* 5 (12), e1396.
- Mwangi, T.W., Bethony, J.M., Brooker, S., 2006. Malaria and helminth interactions in humans: an epidemiological viewpoint. *Ann. Trop. Med. Parasitol.* 100, 551–570.
- Nguhiu, P. N, Kariuki, H. C. Magambo, J. K., Kimani, G., Mwatha, J. K., Muchiri, E., Dunne, D.W., Vennervald, B. J., and Mkoji, G.M.(2009). Intestinal polyparasitism in a rural Kenyan community. *East Afr. Med. J.*, 86, 6.
- Ngwira, B. (2005). The Epidemiology and Control of Lymphatic Filariasis and Intestinal Helminths in the Lower Shire Valley- Chikhwawa District Southern Malawi. Unpublished doctoral dissertation, University of Liverpool, Liverpool.
- Pan American Health Organization (PAHO), 1997. Communicable diseases: control of intestinal parasitosis. Available at: <http://165.158.1.110/english/hcp/hctpep01.html>.
- Pullan, R., Brooker, S., 2008. The health impact of polyparasitism in humans: are we under-estimating the burden of parasitic diseases? *Parasitology* 135, 783–794.
- Raso, G., 2004. Multiple parasite infections and their relationship to self-reported morbidity in a community of rural Cote d'Ivoire. *Int. J. Epidemiol.* 33, 1092–1102.
- Raso, G., Vounatsou, P., Singer, B.H., N' Goran, E.K., Tanner, M., Utzinger, J., 2006. An integrated approach for risk profiling and spatial prediction of *Schistosoma mansoni*-hookworm co-infection. *Natl. Acad. Sci. U. S. A.* 18, 6934–6939.
- Schur, N., Gosoni, L., Raso, G., Utzinger, J., Vounatsou, P., 2011. Modelling the geographical distribution of co-infection risk from single-disease surveys. *Stat. Med.* 30, 1761–1776.
- Spear, R.C.C., Edmund, S., LIANG, S., Birkner, M., Hubbard, A., Dongchuan, Q., Changhong, Y., Zhong, B., Fashen, X., Xueguang, G., George, M.D., 2004. Factors influencing the transmission of *Schistosoma japonicum* in the mountains of Sichuan province of China. *Am. Soc. Trop. Med. Hyg.* 70 (1), 48–56.
- Sturrock, H.J.W., Pullan, R.L., Kihara, J.H., Mwandawiro, C., Brooker, S.J., 2013. The use of bivariate spatial modelling of questionnaire and parasitology data to predict the distribution of *Schistosoma haematobium* in coastal Kenya. *PLoS Negl. Trop. Dis.* 7 (1), e2016. <http://dx.doi.org/10.1371/journal.pntd.0002016>.
- Utzinger, J., Keiser, J., 2004. Schistosomiasis and soil-transmitted helminthiasis: common drugs for treatment and control. *Expert Opin. Pharmacother.* 5, 263–285.
- WHO, 1996. Report of the WHO Informal Consultation on Hookworm Infection and Anaemia in Girls and Women. World Health Organisation, Geneva.
- WHO, 2002. Prevention and control of schistosomiasis and soil-transmitted helminthiasis: report of a WHO expert committee. *WHO Tech. Rep. Ser.* 912, 1–57.
- Xie, H., Tao, J., McHugo, G.J., Drake, R.E., 2013. Comparing statistical methods for analysing skewed longitudinal count data with many zeros: an example of smoking cessation. *J. Subst. Abuse. Treat.* 45 (1), 99–108.
- Zou, Y., Lord, D., Geedipally, S.R., 2011. Over- and Under-Dispersed Count Data: Comparing the Conway-Maxwell-Poisson and Double-Poisson Distributions. Paper No. 12–2801.