

# Severity of Infestation Levels of *Tunga Penetrans* in Central, Kenya: A Bayesian Cumulative Logit Model

Stephen M. Mbunzi<sup>1,2</sup>, Joseph K. Mung'atu<sup>1</sup>, Anthony G. Waititu<sup>1</sup>, Samuel M. Mwalili<sup>1</sup>, Kenneth O. Ogila<sup>3</sup>, Thomas N. O. Achia<sup>4</sup> & Daniel Nthiwa<sup>5</sup>

<sup>1</sup>Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

<sup>2</sup>Department of Mathematics, Computing and Information Technology, University of Embu, Embu, Kenya

<sup>3</sup>Department of Zoology, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

<sup>4</sup>Center for Disease Control, Kemri, Nairobi, Kenya

<sup>5</sup> Department of Biological, University of Embu, Embu, Kenya

Correspondence: Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Received: April 13, 2021 Accepted: August 14, 2021 Online Published: August 20, 2021

doi:10.5539/ijsp.v10n5p70

URL: <https://doi.org/10.5539/ijsp.v10n5p70>

## Abstract

Tungiasis is a neglected parasitic disease that significantly affects communities, especially in developing countries. This study developed a Bayesian severity of the jigger infestation model and its spatial counterpart. Putative determinants leading to different levels of infestation and the most affected areas were to be identified through the model. We collected data through a cross-sectional study with a multi-stage sampling design. A structured questionnaire was administered in each household to capture variables used for modelling jigger infestations. The severity of jigger infestation categorized for each individual was modelled against all the other predictor variables. It was also integrated with spatial data to determine the spatial distribution pattern of jigger infestation. A Bayesian multinomial logistic regression model was used to assess the association between various predictors and different infestation levels. Specifically, an ordered Bayesian Severity Hierarchical (OBSH) categorical model was obtained. This model was categorical based on the Counties (1 – Nyeri, 2 – Murang'a and 3 – Kiambu). Results from this model showed that for a one-unit decrease in the poverty index at level 1 (individuals categorized as poor) there was about a 69% increase in the severity of jigger infestation. A one-unit increase in the percentage of clay in the soil increased the odds ratio of the severity of jigger infestation by a factor of 11.21 while a high percentage of nitrogen in the soil lowered the severity of infestation. Severity of jigger infestation reduced from the baseline, Nyeri County to Kiambu County. It also increased with increasing altitude due to a decrease in nitrogen levels.

**Keywords:** ordered Bayesian severity hierarchical, spatial Bayesian severity hierarchical

## 1. Introduction

*Tunga Penetrans*, is the etiological agent of Tungiasis, a seriously debilitating disease that is prevalent in Central America, South America, India, and tropical Africa as established by Darvin et al., 2018. This disease affects multiple hosts such as humans, cattle, dogs, pigs, cats, horses, mules, sheep, rats and mice (Fein et al., 2001; Escamilla-Martinez et al., 2008; Bouree et al., 2009; Sanusi, 1989). The key drivers of Tungiasis transmission include social neglect, inadequate health care and poor housing conditions, especially in poor-resource communities (Mwangi et al., 2015). The main ecological habitats of *Tunga Penetrans* are warm, dry soil and sand of beaches, stables, and stock farms (Gibbs, 2009). According to Gibbs, 2009, humans are infected by female *Tunga Penetrans* that anchor themselves to the host skin using their biting mouthparts and burrows into the epidermis. The primary typical areas affected include the plantar surface of the foot, the intertriginous regions of the toes, and the periungual regions as Muehlen et al. (2003) showed. However, ectopic sites such as the hands, elbows, thighs, and gluteal region may also be infected (Heukelbach et al., 2002). Whereas earlier studies conducted in the area were more on the social detail (Mwangi et al., 2015; Nyangacha et al., 2019; Nsanzimana et al., 2019; Keiyoro et al., 2016), a mathematical perspective was therefore required to infuse some modelling facet. This would enable finding a better solution to the menace by first establishing the key factors that affect the transmission of *Tunga Penetrans*. The spatial distribution patterns of jigger infestation with their severity distribution paradigm could help deal better with this peril. The intersecting areas of spatial statistical methods, geographical information systems (GIS)

and remote sensing (RS) has witnessed the development of software tools. These tools could enhance the determination of the severity of jigger infestation leading to better large scale implementation of jigger infestation control programmes (Clements et al., 2006 and Lawson et al., 2003).

WinBUGS, a major software tool well versed with modelling spatially referenced small area health data was used. Uncertainty could be incorporated into the modelling process for the observed data and any structured or unstructured random variables through WinBUGS (Lawson et al., 2003). Thus modelling of data was done using the hierarchical Bayesian ordered logistic regression models. Such modelling was through exploiting WinBUGS modern computational advances such as Gibbs sampling.

*intrinsic Conditional autoregressive (iCAR)* models were incorporated to model spatial variation. This was through the incorporation of the standard areal model such as the *Besag* and *Besag – York – Mollie (BYM)*, models by Besag et al., 1991, in form of the structured random effects as expressed by Bakka, 2018. Disease severity refers to the presence and extensiveness of disease in the body as defined by Finlayson et al., 2004. Jigger infestation severity in this work was taken to be the percentage of relevant host tissues or organs covered by a lesion or damaged through such infestation. Thus severity was measured as a categorical variable under four levels. This research, therefore, incorporated a hierarchical Bayesian cumulative logit model to establish the extensiveness of jigger infestation with or without the spatial structure. Through incorporating the spatial structure of the data in a hierarchical Bayesian cumulative logit model, maps together with posterior distribution estimates from the outputs obtained led to informed and objective decision making about targeted areas for disease control through uncertainty assessments. Though documented studies on the comparison on the severity of jigger infestation in these 3 Counties is not found, this study found such severity of infestation to be higher in Nyeri County than Kiambu County while increasing with increase in altitude.

## 2. Materials and Methods

### 2.1 Data Collection

Data collection was through a cross-sectional survey with a multi-stage cluster sampling design. Twelve Counties with the endemic transmission of jiggers were purposefully selected. Among these 12 Counties, 3 (Nyeri, Murang'a and Kiambu) were randomly selected. Two sub-counties were then randomly selected from each county. In Nyeri, we selected Mukurweini and Nyeri South Sub-counties, in Murang'a, Kandara and Gatanga Sub-counties while in Kiambu, Gatundu South and Githunguri Sub-counties were selected (see Figure 1). The population was spread over a large geographical area. Using ArcGIS version 10.4, each of the six sub-counties was sub-divided into approximately 4 square kilometre quadrats or sites (points or patches). One homestead was chosen randomly (see Figure 2) from each quadrat. The number of homesteads selected was 99. Data on jigger infestation was obtained from each member (those found at the time of sampling) of the selected homesteads. The total number of human individuals sampled were 199.

### 2.1 Variables for Modeling Jigger Infestation

Multidimensional Poverty Index (MPI) is designed to measure acute poverty globally, regionally, nationally and across geographical regions (Alkire et al., 2011). A very versatile methodology where different sets of deprivations were defined measuring poverty level vis-a-vis three dimensions using ten indicators as shown by Alkire & Santos, 2010 was described. Thus MPI could be adjusted easily to incorporate alternative indicators, cutoffs and weights as appropriate in regional, national, or sub-national contexts where such indicators are quite effective as key points as established by Setboonsarng, 2005. Representing the number of persons as  $n$ , number of dimensions under consideration as  $d$ , then letting  $y = [y_{ij}]$  be  $n \times d$  matrix of achievements gives the entry  $y_{ij} \geq 0$  as the achievement of individual  $i = 1, \dots, n$  in dimension  $j = 1, \dots, d$ . Each  $y_{i*}$ , a row vector, indicates individual  $i$ 's achievements, while  $y_{*j}$ , a column vector, establishes the distribution of dimension  $j$  achievements across the set of individuals as Alkire & Foster, 2009 indicated. Assuming  $d$  to be given and fixed, and allowing  $n$  to range across all positive integers, then the domain of matrices under consideration becomes  $\mathbf{Y} = \{\mathbf{y} \in \mathbf{R}_+^{nd} : \mathbf{n} \geq \mathbf{1}\}$ .

Let  $z_j > 0$  denote the cutoff below which a person is considered to be deprived in dimension  $j$ , and  $\mathbf{z}$  be the row vector of dimension - specific cutoffs.

An identification method by Alkire & Foster, 2009 and Bourguignon & Chakravarty, 2003 using an **identification function**

$$\rho : \mathbf{R}_+^d \times \mathbf{R}_+^d \longrightarrow \{0, 1\}, \quad (1)$$

mapping from person  $i$ 's achievement, vector  $\mathbf{y}_i \in \mathbf{R}$  and cutoff vector  $\mathbf{Z} \in \mathbf{R}^2$  to an indicator variable in such a way that

$$\rho_p(\mathbf{y}_i, \mathbf{z}) = \begin{cases} 1, & \text{if person } i \text{ is poor} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

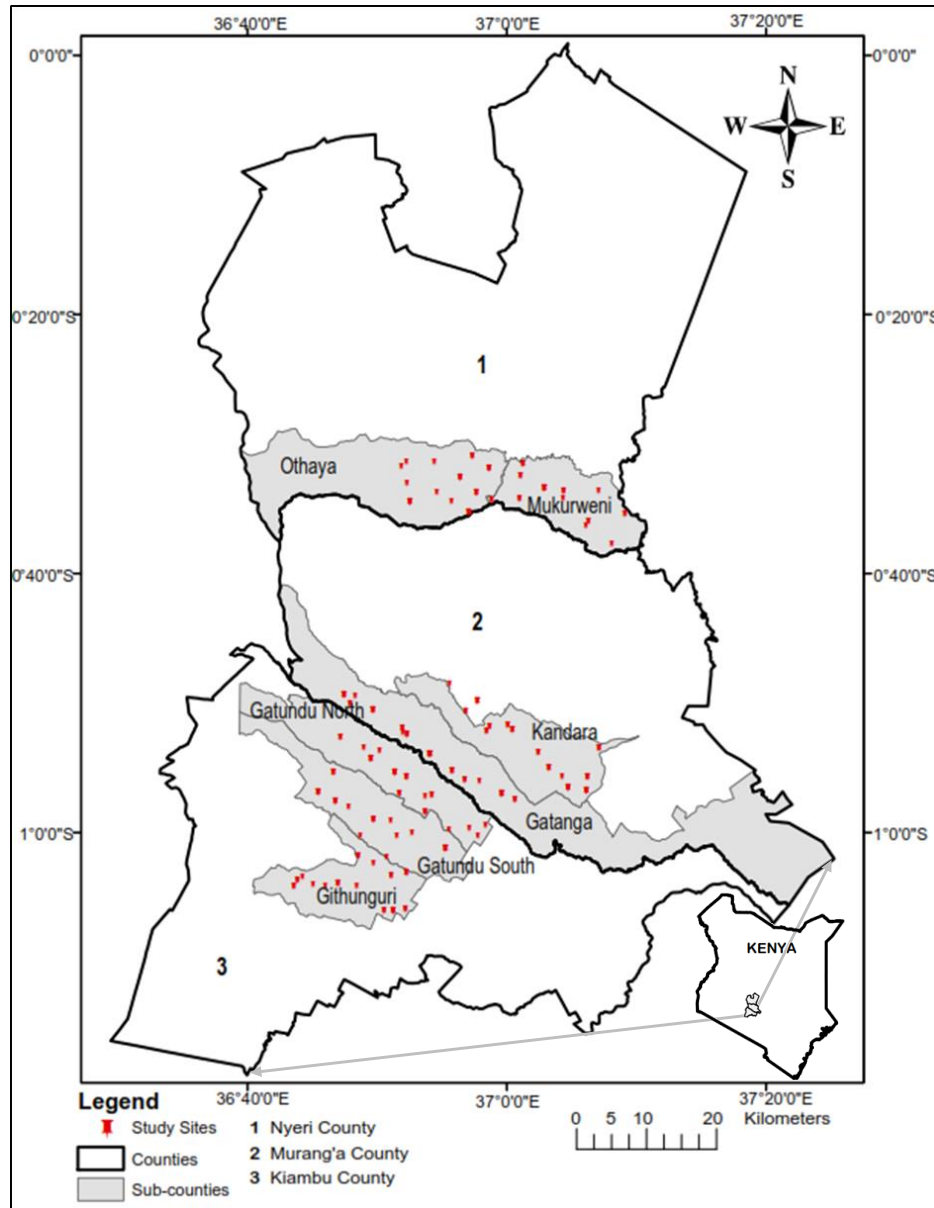


Figure 1. Study Area Map: This is a map of the six sub-counties in the three Counties, Nyeri, Murang'a and Kiambu of Central, Kenya where this study was carried out

Focusing on the dimensional shortfalls, the **union method** of identification describes person  $i$  as multidimensionally poor if deprived in at least one dimension and alternatively, the **intersection approach**, identifies person  $i$ , as being poor only if deprived in all dimensions as established by Alkire & Foster, 2009. Due to shortfall in each, a trade-off of both the two methodologies would be to use an intermediate cutoff level for  $c_i$  that lies somewhere between the two extremes, 1 and  $d$  such that for  $k = 1, \dots, d$  the identification method becomes

$$\rho_k(\mathbf{y}_i, \mathbf{z}) = \begin{cases} 1, & \text{whenever } c_i \geq k \\ 0, & \text{whenever } c_i < k \end{cases} \quad (3)$$

Thus,

### Maps of 6 Sub-counties in the 3 Counties

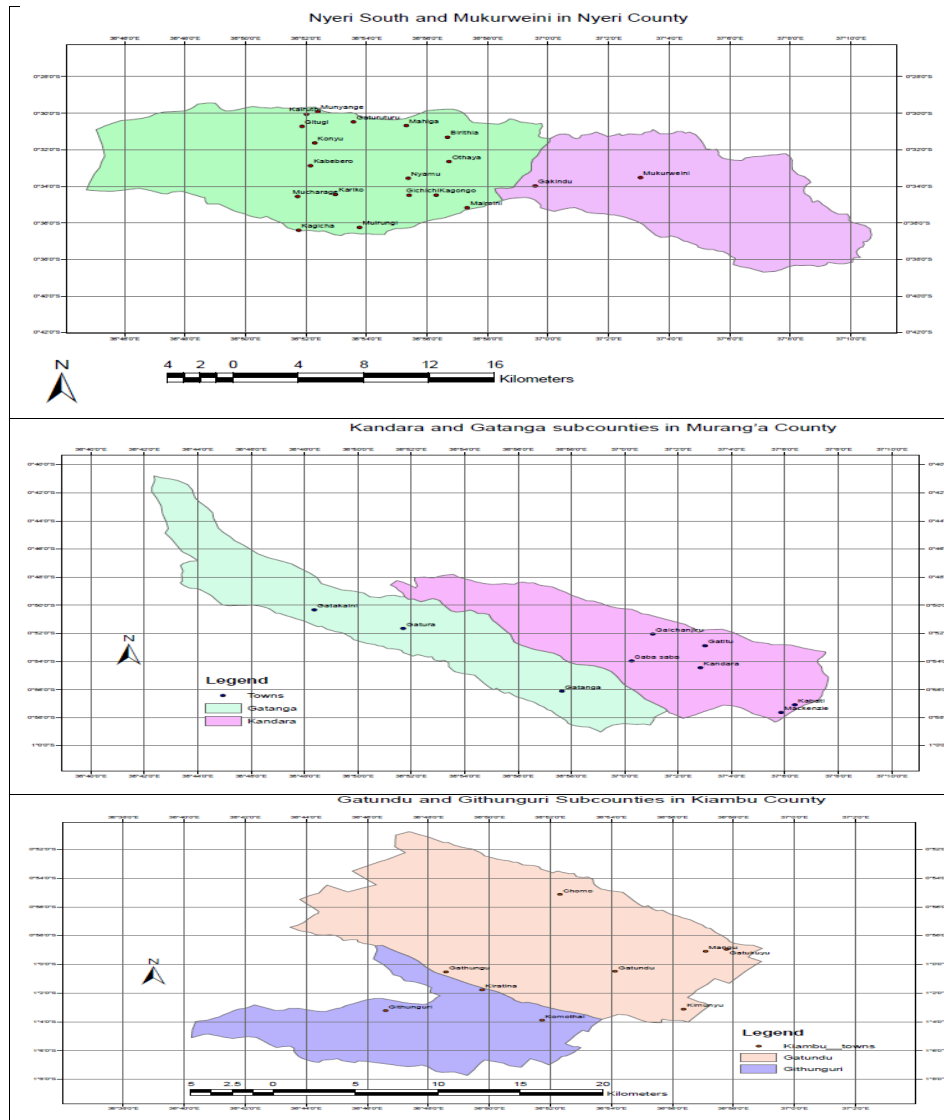


Figure 2. Sampling Map: Map of the sampled six sub-counties showing the sites from which one homestead was sampled

$$\rho_k(\mathbf{y}_i, \mathbf{z}) = \begin{cases} \text{poor,} & \text{if } d_i \geq k \\ \text{not poor,} & \text{if } d_i < k \end{cases} \quad (4)$$

where  $d_i$  is number of dimensions in which  $i$  is deprived.

$\rho_k$  is dependent on both **within dimension cutoffs**  $\mathbf{z}_j$  and the **across dimension cutoff**  $k$ , hence referring it as the **dual cutoff** method of identification.

A questionnaire was used to capture variables for computing multidimensional poverty index, (*MPI*). Using 3, poverty index was determined through excel. Reliability of the index was determined by computing Cronbach's alpha ( $\alpha = 0.710$ ) coefficient using SPSS to confirm consistency. *POVHH*, for poverty index calculated was a categorical variable whereby, if the household calculated score  $c$ , was less than  $\frac{1}{3}$ , then the household was not deprived, hence not poor.

Other variables captured during sampling include Geographical coordinates (altitude, latitude and longitude) of each sampled household. Soil samples were taken from five different locations within each homestead and mixed to make

a pooled sample. Approximately 100 grams of soil were collected from each sampling location. These soil samples were packed and taken to Kenya Agricultural and Livestock Research Organization (KALRO), Nairobi for analysis. Soil types, pH, texture grades, and macronutrients (e.g., total nitrogen, phosphorus and potassium) were each determined. Additional data on precipitation, temperature and humidity for the three main centres of the three-targeted counties were obtained from the central collection point, Nairobi Meteorological (Kenya's headquarters) Department. These County-level variables were taken for the last eight years from the date of the last data collection and averaged values computed.

For each family member, images were taken upon requesting the head of the household, irrespective of whether the person had been infested with jiggers or not, with the main parts of interest being feet, toes, fingers, ankles, hands and heels. A scale of measuring levels of infestation was computed from these images by counting black spots of the infested part with the codes 1, 2, 3 and 4 given on an increasing rate of severity (see Figure 3). The first level of infestation had "0" black spots, followed by "1 - 10" black spots, then "11 - 20" black spots with the fourth level, "over 20" black spots. Severity of Infestation (*SEVTY*) for each person was established.



Figure 3. Images of persons infested with jiggers: These represent images of parts of bodies of persons at different levels of infestation

## 2.2 Ethics Approval and Consent to Participate

Ethical approval for this study was provided by the National Commission for Science, Technology and Innovation (NACOSTI) in Kenya (Reference number: NACOSTI/P/15/6946/6847). Persons aged  $\geq 18$  years provided oral consents for participation in the study. For minors aged  $< 18$  years, oral consents for participation in the study were obtained from their guardians/parents.

### 2.3 Severity of Jigger Infestation Model

Let  $i, (i = 1, \dots, n)$  be an index to represent the  $i^{th}$  individual in this study. Also, let  $s, (s = 1, \dots, S)$  be an index to represent the levels of Severity of Infestation. The index  $s$  took the values 'Not-infested', ( $s = 1$ ), 'Moderately Infested', ( $s = 2$ ), 'Highly Infested', ( $s = 3$ ), and 'Severely Infested', ( $s = 4$ ). This led to the standard ordered response logit model

$$y_i^* = \beta' X_i + \varepsilon_i \quad \text{for } y_i = s \quad \text{if } \tau_{s-1} < y_i^* < \tau_s \quad (5)$$

as shown by Eluru et al., 2008 and Kockelman & Young-Jun, 2002 scholarly works.

In the model,  $y_i^*$  was the latent severely infestation propensity for non - infested individual  $i$  in the event of infestation. By letting  $y_{ijk}$  be the observation variables and  $y_{ijk}^*$  be the latent variables with the threshold/ cut-point parameters obeying  $-\infty = \tau_0 < \tau_1 < \dots < \tau_{S-1} < \tau_S = \infty$  as the scholarly works of Grilli & Rampichini, 2012, Agresti, 2010, Gelman & Hill, 2012, Eluru et al., 2008, Qiu et al., 2002 show, then

$$y_{ijk} = \begin{cases} 1, & \text{if } -\infty < y_{ijk}^* < \tau_1 \text{ (no infestation)} \\ 2, & \text{if } \tau_1 < y_{ijk}^* < \tau_2 \text{ (moderate infestation)} \\ 3, & \text{if } \tau_2 < y_{ijk}^* < \tau_3 \text{ (high infestation)} \\ 4, & \text{if } \tau_3 < y_{ijk}^* < \infty \text{ (severe infestation)} \end{cases} \quad (6)$$

$$y_{ijk}^* = \beta' X_{ijk} + \varepsilon_{ijk}$$

$$\varepsilon_{ijk} \sim \text{logistic}(\beta' X_{ijk}, \sigma^2)$$

represents the severity of jigger infestation for the  $i^{th}$  person, ( $i = 1, \dots, N = 199$ ), in the  $j^{th}$  Ward, ( $j = 1, \dots, n = 30$ ) and in the  $k^{th}$  County, ( $k = 1, 2, 3$ ).

6 could equivalently be expressed as a bayesian cumulative logit model (BCLM), as given by Gelman & Hill (2012), with both models following ordered probit model specification like presented by Kockelman & Young-Jun, 2002.

From 6 assumption that the latent variable  $y_{ijk}^*$  follows a standardized logistic distribution as expressed by Qiu, Z., Song, P. X. K., & Tan, M. (2002), then we have

$$\begin{aligned} \Pr(y_{ijk} = 1 | \beta' x_{ijk}) &= \text{logit}^{-1}(\tau_1 - \beta' x_{ijk}) \\ \Pr(y_{ijk} = s | \beta' x_{ijk}) &= \text{logit}^{-1}(\tau_s - \beta' x_{ijk}) - \text{logit}^{-1}(\tau_{s-1} - \beta' x_{ijk}) \\ \Pr(y_{ijk} = 4 | \beta' x_{ijk}) &= 1 - \text{logit}^{-1}(\tau_3 - \beta' x_{ijk}), \end{aligned} \quad (7)$$

a very important model as it pertains to the sampling scheme implemented and coded in WinBUGS for both severity of jigger infestation with and without the spatial counterpart.

Mapping from any of the above Ordinal Logit Models (OLM), the latent regression to the observed ordinal responses  $y_{ijk}$  introduces a censoring function,

$$\begin{aligned} y_{ijk} &= 0 \quad \text{if } y_{ijk}^* \leq \tau_0 \\ y_{ijk} &= s \quad \text{if } \tau_{s-1} < y_{ijk}^* \leq \tau_s, \\ &\quad s = 1, \dots, S - 1. \quad \text{"Interior Categories"} \\ y_{ijk} &= S \quad \text{if } y_{ijk}^* > \tau_{S-1} \quad \text{"Last Category"}, \end{aligned} \quad (8)$$

to 6 censoring scheme as expressed in Jackman, S. (2009) work.

Having categorized severity of jigger infestation, we ended up with 6 and used the censoring function 8 to map the latent regression to the observed ordinal responses. This led to the probabilities of the outcome on the model for severity of infestation of jiggers as per the bayesian cumulative logit model (BCLM), becoming

$$\begin{aligned}
\Pr(y_{ijk} = 1|\beta'X_{ijk}) &= \Pr(y_{ijk}^* \leq \tau_1) \\
&= \Pr(\beta'X_{ijk} + \epsilon_{ijk} \leq \tau_1) \\
&= \Pr(\epsilon_{ijk} \leq \tau_1 - \beta'X_{ijk}) \\
&= \text{logit}^{-1}(\tau_1 - \beta'X_{ijk}) \\
&= \frac{\exp(\tau_1 - \beta'X_{ijk})}{1 + \exp(\tau_1 - \beta'X_{ijk})} \\
&= \frac{1}{1 + \exp(-\tau_1 + \beta'X_{ijk})} \\
&= F(\tau_1 - \beta'X_{ijk}),
\end{aligned} \tag{9}$$

$$\begin{aligned}
\Pr(y_{ijk} = s|\beta'X_{ijk}) &= \Pr(\tau_{s-1} < y_{ijk}^* \leq \tau_s) \quad \forall \quad s = 2, 3 \\
&= \Pr(\tau_{s-1} < \beta'X_{ijk} + \epsilon_{ijk} \leq \tau_s) \\
&= \Pr(\tau_{s-1} - \beta'X_{ijk} < \epsilon_{ijk} \leq \tau_s - \beta'X_{ijk}) \\
&= \Pr(\epsilon_{ijk} \leq \tau_s - \beta'X_{ijk}) - \Pr(\epsilon_{ijk} < \tau_{s-1} - \beta'X_{ijk}) \\
&= \text{logit}^{-1}(\tau_s - \beta'X_{ijk}) - \text{logit}^{-1}(\tau_{s-1} - \beta'X_{ijk}) \\
&= \frac{\exp(\tau_s - \beta'X_{ijk})}{1 + \exp(\tau_s - \beta'X_{ijk})} - \frac{\exp(\tau_{s-1} - \beta'X_{ijk})}{1 + \exp(\tau_{s-1} - \beta'X_{ijk})} \\
&= \frac{1}{1 + \exp(-\tau_s + \beta'X_{ijk})} - \frac{1}{1 + \exp(-\tau_{s-1} + \beta'X_{ijk})} \\
&= F(\tau_s - \beta'X_{ijk}) - F(\tau_{s-1} - \beta'X_{ijk})
\end{aligned} \tag{10}$$

and

$$\begin{aligned}
\Pr(y_{ijk} = 4|\beta'X_{ijk}) &= \Pr(y_{ijk}^* > \tau_3) \text{ for the last category} \\
&= \Pr(\beta'X_{ijk} + \epsilon_{ijk} > \tau_3) \\
&= \Pr(\epsilon_{ijk} > \tau_3 - \beta'X_{ijk}) \\
&= 1 - \Pr(\epsilon_{ijk} \leq \tau_3 - \beta'X_{ijk}) \\
&= 1 - \text{logit}^{-1}(\tau_3 - \beta'X_{ijk}) \\
&= 1 - \frac{\exp(\tau_3 - \beta'X_{ijk})}{1 + \exp(\tau_3 - \beta'X_{ijk})} \\
&= 1 - \frac{1}{1 + \exp(-\tau_3 + \beta'X_{ijk})} \\
&= 1 - F(\tau_3 - \beta'X_{ijk})
\end{aligned} \tag{11}$$

Using

$$h(\delta_{ijk,s}) = \tau_s + \beta'x_{ijk} + u_i \tag{12}$$

a Bayesian hierarchical severity model having the  $i^{th}$  person covariates  $x_i$ , the  $j^{th}$  Ward covariates  $x_j$  and the  $k^{th}$  County covariates  $x_k$  were fitted through the MCMC simulation approach in WinBUGS package.

Thus the cumulative probability

$$\delta_{ijk,s} = \Pr(y_{ijk} \leq s)$$

represents the probability that the severity of jigger infestation was not as severe than category  $s$ ,  $s = 1, 2, 3, 4$  where the  $X_{ijk}$  was clearly established from our data (see Appendix A). This could be modeled as 12 with the  $\zeta_i$ , being the spatially structured random effect term added to produce a Bayesian hierarchical severity model as presented by Qiu et al., 2002; where

- $\beta' x_{ijk}$  as given in (Appendix A) with  $x_{ijk}$  represented as  $x_i$  incorporated all the variables.
- $u_i$  was the unstructured random effect term specified as an independent normal distribution with an expected mean of zero and variance of  $\sigma_u^2$
- $\zeta_i$  was the spatially structured random effect term whose conditional distribution,  $\zeta_i|\zeta_{(i)}$ , was given as:

$$\zeta_i|\zeta_{(i)} \sim N\left(\bar{\zeta}_i + \phi \frac{1}{n_i} \sum_{j=1}^n a_{ij}(\zeta_j - \bar{\zeta}_j), s_i^2\right), \quad i = 1, \dots, n \quad (13)$$

with

$$\begin{aligned} c_{ij} &= \phi \times w_{ij} \\ \text{where } w_{ij} &= \frac{a_{ij}}{n_i} \\ \text{with } a_{ij} &= \begin{cases} 1, & \text{areas } i \text{ and } j \text{ are neighbors} \\ 0, & \text{otherwise} \end{cases} \\ \text{Thus } a_{ii} &= 0 \Rightarrow w_{ii} = 0; \quad c_{ii} = 0 \end{aligned} \quad (14)$$

with  $c_{ij}$  indicating the spatial proximity and parameter  $\phi$  controlling the properness of the distribution as the scholarly work of Blangiardo & Cameletti, 2015 indicate. Considering  $W$  to be a matrix of generic elements  $w_{ij}$ , and  $S = \text{diag}(s_1, \dots, s_n)$ , we ensured that the distribution for the spatially structured random effect  $\zeta$  was proper, by having the covariance matrix  $(I_n - \phi W)^{-1} S^2$  positive definite.  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$  being the ordered eigenvalues of  $W$ , then the values of  $\phi$  must lie between  $\lambda_n^{-1}$  and  $\lambda_1^{-1}$ . This led to the proper conditional autoregressive specification (CAR) of  $\zeta$  being a multivariate Normal random variable given as:

$$\zeta \sim MVN(\bar{\zeta}, (I_n - \phi W)^{-1} S^2)$$

where the mean vector becomes  $\bar{\zeta} = (\bar{\zeta}_1 \dots \bar{\zeta}_n)' I_n$ , is the  $n \times n$  identity matrix, and  $S^2 = \text{diag}(s_1^2, \dots, s_n^2)$ . Correlation between areas  $i$  and  $j$  depends only on  $\phi$  and  $W$  as demonstrated by Assuncao & Krainski, 2009.  $\phi$  proves to be a bit hard, hence estimation of such proper conditional autoregressive specification is not greatly used in disease mapping. Fixing  $\phi = 1$ , simplifies the above formulation, leading to the covariance matrix to be non positive definite with the conditional distribution for  $\zeta_i$  having the specification known as *intrinsic conditional autoregressive (iCAR)*. Such, coupled with the exchangeable random effect,  $u_i$ , (i.e. the area-specific effect modelled as exchangeable) forms the so-called *Besag – York – Mollie (BYM)* model, by Besag et al., 1991. Due to the non-positive definition of the covariance matrix, there is no proper joint distribution for  $\zeta$  as Blangiardo & Cameletti, 2015 reiterates, as it would be possible to add any constant to each  $\zeta_i$  without changing the distribution, which could be rectified by fixing a constraint such as  $\sum_{i=1}^n \zeta_i = 0$ . Thus assuming  $\bar{\zeta}_i = 0$  for each  $i$ , the conditional distribution now becomes

$$\zeta_i|\zeta_{(i)} \sim N\left(\frac{1}{n_i} \sum_{j=1}^n a_{ij}\zeta_j, s_i^2\right), \quad i = 1, \dots, n \quad (15)$$

which is determined through an *intrinsic conditional autoregressive (iCAR)* distribution that is widely used in spatial statistics.

The vectors in 9, 10 and 11 contain the appropriate conditional probabilities of each of the four ordinal outcomes, for each respondent,  $i$ . Indicating to Bayesian use of Gibbs sampler (BUGS) that there existed a categorical response with the vector of conditional probabilities for each possible response,  $y_i$ , then a link was made between the model and the observed responses through the categorical density.

Bayesian inference for the cumulative link ordered response model has two sets of parameters treated as independent *a priori* such that

$$\begin{aligned} \theta &= (\beta, \tau)'; \\ \text{for } \tau &= (\tau_1, \dots, \tau_s). \\ \Pr(\theta) &= \Pr(\beta) \Pr(\tau) \end{aligned} \quad (16)$$

A vague (diffuse) normal priors for  $\beta$  was assumed in the data as shown by Jackman, 2009 and Qiu et al., 2002, where  $\beta \sim N(\mathbf{b}_0, \mathbf{B}_0)$  resulting from a probit link. The prior for  $\tau$  was found to obey the ordering constraint implied by the model



as per the censoring scheme in 8 which mimics Jackman, 2009 scholarly work. This was accomplished through a simple arrangement of the priors for  $\tau$  so that

$$\begin{aligned}\tau_1 &\sim N(\nu, D) \\ g_j &\sim \text{Exponential}(d), \quad j = 2, \dots, J \\ \tau_j &= \tau_{j-1} + g_j, \quad j = 2, \dots, J\end{aligned}\tag{17}$$

noting that Bayesian hierarchical model would require prior distributions for each of the model parameter. Prior density was arbitrarily chosen. For  $\beta$ , a normal or a multivariate normal prior would be sufficient as indicated by Jackman, 2009. The density for  $\tau$ ,  $Pr(\tau)$ , was given as normal. Other densities like uniform on some specific interval would also suffice. On the quantities  $g_j$ , Jackman, 2009 presented the Exponential distribution as the best to keep an ordering constraint on  $\tau$ . A truncated normal density with  $\mu = 0$  and a large variance of the form  $\sigma^2 = 1000^2$  priors on the interval  $(0, \infty)$  for  $g_j$ 's could still be used. It's arbitrarily chosen since it is supported only on the non-negative half of the real line. Any distribution restricted to the non-negative half of the real line will suffice (like a non-overlapping zero interval on a uniform density, the gamma density or even a log-normal density).

By choosing  $\nu = 0$ , a reasonably large scalar for  $D$  and a small positive quantity for  $d$ , then we obtain a vague prior over the thresholds. Incorporating severity of jigger infestation model, 6, into 12 introduced both the unstructured and spatially structured random effect terms. Such resulted in the Spatial Bayesian Severity Hierarchical (*SBSH*) Models for the severity of jigger infestation we run in WinBUGS.

### 3. Results and Discussion

Basing summaries on frequencies calculated from  $N = 100,000$  simulations followed by an equal number of iterations, then  $N = 200,000$  samples were obtained with acceptable MC errors of less than 5% of the posterior standard deviations. Convergence was monitored through visual examination of trace plots of the samples for each chain as Liu & Zhu, 2017 demonstrated. The ratio of between and within chain variances of the Gelman-Rubin convergence statistic converged to 1.0 as indicated by Spiegelhalter et al., 2003.

Different Bayesian logistic models were subjected to the deviance information criterion (DIC) statistic, where the ordered Bayesian severity hierarchical (OBSH) model 1 captured all personal, ward and County covariates with all the predictor variables well standardized. OBSH 2 had all the County predictors, age and gender removed since they were found nonsignificant in OBSH 1. OBSH 3 had the altitude removed from OBSH 2. Further removal of the soil pH (SOPH) predictor variable led to OBSH Random (with CNTY as a random variable) and OBSH Categorical (with CNTY as a categorical variable). The DIC values for different OBSH models were as given in Table 1.

Table 1. DIC values for different Ordered Bayesian Severity Hierarchical Models

Model	$\bar{D}$	$\hat{D}$	pD	DIC
<i>OBSH1</i>	376.5	358.2	18.27	394.8
<i>OBSH2</i>	376.2	358.7	17.45	393.6
<i>OBSH3</i>	375.3	358.8	16.50	391.8
<i>OBSHRandom</i>	375.1	358.7	16.42	391.6
<i>OBSHCategorical</i>	374.0	357.8	16.13	390.1

OBSH Categorical model was run in WinBUGS (see Appendix B and C for both model and the output, respectively). A negative coefficient in the poverty index, *POVHH*, meant an increase in severity of jigger infestation. *POVHH*, a categorical variable was calculated such that its value was 0, if the household was not deprived, hence not poor and 1, otherwise. An odds ratio of 0.3064 for *POVHH* reported, meant for a one-unit decrease in the poverty index at level 1, approximately 69% increase of severity to jigger infestation was registered as shown in Table 2.

This implied a positive association between the poverty index and the severity of jigger infestation. Its 95% Bayesian CI being a shorter interval (Clements et al., 2006), implying a high level of certainty in predicting that poverty level affects the severity of jigger infestation.

Percentage of clay in the soil contributed to 11.21 times increase in severity of jigger infestation with quite a wide 95% Bayesian CI. A one-unit increase in the Potassium percentage in the soil increased the odds ratio of jigger infestation severity by a factor of 1.725. Thus, Potassium percentage as one of the primary nutrients in the soil increased jigger infestation severity by about 72.5%. From Nyeri, the baseline County to Kiambu County, the severity of jigger infestation

Table 2. Posterior Means and Credible Intervals of Variables of the OBSH Categorical

Variable	Mean (95% Bayesian CI)	Odds Ratio
POVHH,1 - Poverty Index	-1.183 (-1.938,-0.4487)	0.3064 (0.1440,0.6385)
DA,2 - Sheep or Cow or both	-1.873 (-2.871,-0.9096)	0.1537 (0.0566,0.4027)
DA,3 - Cat or Dog or Both	0.3328 (-0.8571,1.507)	1.395 (0.4244,4.513)
DA,4-Shp/Cow and C/D or both	-0.8685 (-1.728,-0.0191)	0.4196 (0.1776,0.9811)
DA,5 - Pig	-2.334 (-4.618,-0.09117)	0.0969 (0.0099,0.9129)
SAND - % of sand in soil	0.9003 (-0.8441,2.668)	2.460 (0.4299,14.41)
CLAY - % of clay in soil	2.417 (0.1963,4.785)	11.21 (1.217,119.7)
Kme - Potassium me%	0.545 (0.1787,0.9182)	1.725 (1.196,2.505)
TOTN - % Total Nitrogen	-0.4105 (-0.808,-0.0300)	0.6633 (0.4457,0.9704)
SILT - % of silt in soil	0.04477 (-0.9111,1.074)	1.0456 (0.4021,2.927)
Ppm - Phosphorus ppm	-0.2257 (-0.6148,0.1543)	0.79780 (0.5407,1.167)
CNTY,2 - Murang'a	-0.3948 (-1.207,0.4083)	0.6738 (0.2991,1.504)
CNTY,3 - Kiambu	-1.458 (-2.369,-0.5701)	0.2327 (0.0936,0.5655)
$g_1$	-2.142 (-3.192,-1.131)	0.1174 (0.04109,0.3227)
$g_2$	2.51 (2.01,3.063)	12.30 (7.463,21.39)
$g_3$	0.387 (0.1702,0.6847)	1.473 (1.186,1.983)

reduced by approximately 77%. The percentage of the severity of infestation was found to be 72.0% and 46.2% comparing County 1 and County 3 respectively as seen in Table 3.

Table 3. Severity of Infestation for the 3 Counties

Label	y	Nyeri	Murang'a	Kiambu
Not Infested	1	14	33	42
Moderately Infested	2	24	26	24
Highly Infested	3	5	0	3
Severely Infested	4	7	12	9

The threshold parameters,  $g$ , representing simple cut-points, though notorious for being highly correlated with one another, displayed some good mixing as displayed by the correlation matrix and the autocorrelation function.

### 3.1 Spatial Analysis

Four Spatial Bayesian Severity Hierarchical (*SBSH*) Models were run in WinBUGS. Their *DIC* and  $\psi$  values were tabulated for assessment. *SBSH Model 1* incorporated all the personal, Ward and County predictor variables in a multilevel Bayesian logistic model. With the highest *DIC* value amongst the models run, none of the County predictor variables was significant together with age and gender. All the non-significant variables were left out in forming *SBSH Model 2*, a better fit than the first. The soil pH, SOPH variable was left out in the last two models. CNTY was used both as a random effect and categorical effect in the standardized *SBSH Random Model* and *SBSH Categorical Model* models respectively. Their fit was almost the same as shown from their *DIC* values in Table 4.

Table 4. *DIC* and  $\psi$  values for different Spatial Bayesian Severity Hierarchical (*SBSH*) Models

Model	$\bar{D}$	$\hat{D}$	pD	DIC	$\psi$
<i>SBSH Model 1</i>	302.155	269.942	32.213	334.367	0.3249
<i>SBSH Model 2</i>	301.409	271.446	29.963	331.372	0.3612
<i>SBSH Random Model</i>	300.646	271.497	29.149	329.794	0.3584
<i>SBSH Categorical Model</i>	300.199	270.783	29.416	329.614	0.2964

*SBSH Random Model* was used in reporting for the spatial severity of jigger infestation since it was the best fit. This model had the lowest number of effective parameters (parsimonious) as shown by its pD value overall. Comparing the

$\psi$  values of *SBSH Random Model* and *SBSH Categorical Model*, it was clear that the CNTY effect affected spatial variation significantly. Using CNTY as a random effect ended up having a higher spatial variation than when using it as a categorical effect, giving another reason for using the standardized *SBSH Random Model*. The value,  $\psi = 0.3584$  implied that slightly more than a third of the total variation was as a result of moving from one Ward to the next.  $\psi$  value near 1 showed that the spatial variation dominated, and was negligible when close to 0 as shown in Lawson et al., 2003 work. The WinBUGS code for the standardized *SBSH Random Model* is as shown in Appendix D. A unit increase in the poverty index increased the odds ratio level of 1 versus 0 by a factor of 0.2712. This led to the severity of jigger infestation increase of 73% with a short 95% Bayesian CI, meaning that predictions could sufficiently be made with some high degree of certainty moving from one ward to the next as reiterated by Clement, et al., 2006. Clay percentage in the soil very highly contributed to the severity of jigger infestation. For every one per cent increase in the amount of clay in the soil, the severity of jigger infestation increased 33 times with a very low degree of certainty in prediction within the area under study. The results of the posterior means and credible intervals of variables of the *SBSH Random Model* were as tabulated in Table 5.

Table 5. Posterior Means and Credible Intervals of Variables of the *SBSH Random Model*

Variable	Mean (95% Bayesian CI)	Odds Ratio
POVHH,1 - Poverty Index	-1.305 (-2.223, -0.4165)	0.2712 (0.1083, 0.6594)
DA,2 - Sheep or Cow or both	-3.370 (-4.908, -1.930)	0.03439 (0.007387, 0.1451)
DA,3 - Cat or Dog or Both	0.3399 (-1.177, 1.861)	1.405 (0.3082, 6.430)
DA,4-Shp/Cow and C/D or both	-1.723 (-2.994, -0.5101)	0.1785 (0.05009, 0.6004)
DA,5 - Pig	-4.443 (-7.442, -1.584)	0.01176 (0.0005861, 0.2052)
SAND - % of sand in soil	1.663 (-0.6564, 4.125)	5.275 (0.5187, 61.87)
CLAY - % of clay in soil	3.508 (0.6754, 6.724)	33.38 (1.965, 832.1)
Kme - Potassium me%	1.128 (0.4646, 1.831)	3.089 (1.591, 6.240)
TotN - % Total Nitrogen	-0.9266 (-1.645, -0.2567)	0.3959 (0.1930, 0.7736)
SILT - % of silt in soil	-0.01054 (-1.236, 1.367)	0.9895 (0.2905, 3.924)
Ppm - Phosphorus ppm	-0.1635 (-0.6415, 0.3045)	0.8492 (0.5265, 1.356)
ALT - Altitude	1.075 (0.4006, 1.792)	2.930 (1.493, 6.001)
CNTY,1 - Nyeri	0.2072 (-8.93, 8.647)	1.230 (0.0001324, 5693)
CNTY,2 - Murang'a	1.246 (-7.723, 9.853)	3.476 (0.0004425, 19010)
CNTY,3 - Kiambu	-3.064 (-12.42, 5.23)	0.04670 (0, 186.8)
$g_1$	-3.37(-12.93, 5.051)	0.03439(0, 156.2)
$g_2$	3.555(2.79, 4.437)	34.99(16.28, 84.52)
$g_3$	0.5216(0.2275, 0.9249)	1.685(1.255, 2.522)

An increase in the total percentage levels of potassium and nitrogen as macronutrients in the soil meant both an increase and decrease in the severity of jigger infestation. From both, a high level of certainty in predicting the exact areas affected by each moving from one Ward to the next was possible. For every one-unit increase in altitude, the severity of jigger infestation increased by 193%. Though very little or none has been done on the area linking soil macronutrients to the severity of Tunga penetrans, Xianjin et al., 2016 and Sundqvist, 2013 showed that foliar and litter nitrogen decrease in montane tropical forests or even wet tropical regions. This explains the fact that as altitude increase, the amount of nitrogen in the soils decrease, ultimately increasing the severity of Tunga Penetrans.

#### 4. Conclusion

From both the OBSH Categorical model and the standardized *SBSH Random Model*, an increase in poverty level and a high percentage of clay and potassium in the soil led to more severe infestation. Increasing the nitrogen percentage in the soil lowered the severity of infestation. Moving from one Ward to the next, predictions could be made with some high degree of certainty considering poverty levels.

#### Acknowledgements

We wish to acknowledge the participants from the sampled six sub-counties, without whose permission, this research would not have taken place. Much appreciation goes to Kanyi, N. C., who helped translate the English written questionnaires into the local dialect, Kikuyu. This work would not have been complete without sampling the cross-section area together with the study area map done by Okeyo, J. M. and Okal, H. A., respectively.

## References

- Agresti, A. (2010). *Analysis of Ordinal Categorical Data*. (2nd Ed.) John Wiley and Sons, Inc., Hoboken, NJ: USA.
- Alkire, S., & Foster, J. (2009). Counting and Multidimensional Poverty In: Von Braun, J., Ed., *The Poorest and Hungry: Assessment, Analysis and Actions, Oxford Poverty and Human Development Initiative*, 32.
- Alkire, S., Roche, J. M., Santos, M. E., & Seth, S. (2011). Multidimensional poverty index: Brief methodological note. *Oxford: Oxford University Press*, 4, 1-14.
- Alkire, S., & Santos, M. E. (2010). Acute multidimensional poverty: A new index for developing countries. *Oxford Poverty and Human Development Initiative*, 38, University of Oxford.
- Assuncao, R., & Krainski, E. (2009). Neighborhood Dependence in Bayesian Spatial Models. *Biometrical Journal*, 51(5), 851-869. <https://doi.org/10.1002/bimj.200900056>
- Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Krainski, E., Simpson, D., & Lindgren, F. (2018). *Spatial modelling with r-inla: A review*, v2., stat.ME, arXiv:1802.06350v1.
- Besag, J., York, J., & Mollie, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Physica*, 43, 1-59
- Blangiardo, M., & Cameletti, M. (2015). Spatial and Spatio-temporal Bayesian Models with R-INLA. *John Wiley and Sons, Ltd*. <https://doi.org/10.1002/9781118950203.ch4>
- Bouree, P., Oss, L., & Rabenandrasana, F. (2009). *Tungiasis, an uncommon ectoparasitic disease*, 59(2), 163-165.
- Bourguignon, F., & Chakravarty, S. R. (2003). The measurement of multidimensional poverty. *Journal of Economic Inequality*, 1(1), 25-49. <https://doi.org/10.1023/A:1023913831342>
- Clements, A. C. A., Lwambo, N. J. S., Blair, L., Nyandindi, U., Kaatano, G., Kinunghi, S., Webster, J. P., Fenwick, A., & Brooker, S. (2006). Bayesian spatial analysis and disease mapping: Tools to enhance planning and implementation of a schistosomiasis control programme in Tanzania. *Tropical Medicine and International Health*, 11(4), 490-503. <https://doi.org/10.1111/j.1365-3156.2006.01594.x>
- Darvin, S. S., Lisa, M., Mark, R. W., Zachary S. W., & Stephanie A. N. (2018). Tungiasis Differential Diagnoses. *American Journal of Clinical Dermatology*. Medscape.
- Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis and Prevention*, 40(3), 1033-1054. <http://doi.org/10.1016/j.aap.2007.11.010>
- Escamilla-Martinez, E., Gómez-Martín, B., Sánchez-Rodríguez, R., Martínez-Nova, A., Martínez-Granada, L. J., & Altube-Arabiurruia, E. (2008). Tungiasis-traveler's ectoparasitosis of the foot: A case report. *Foot Ankle Int.*, 29(3), 354-357.
- Fein, H., Naseem, S., & Witte, D. P. (2001). Tungiasis in North America: A Report of 2 cases in internationally adopted children. *J Pediatr*, 139(5), 744-746.
- Finlayson, T. L., Moyer, C. A., & Sonnad, S. S. (2004). Assessing Symptoms, Disease Severity, and Quality of Life in the Clinical Context: A Theoretical Framework. *The American Journal of Managed Care*, 10(5), 336-344.
- Gelman, A., & Hill, J. (2007). Data Analysis Using Regression and Multilevel/ Hierarchical Models. *Cambridge University Press*. <https://doi.org/10.1017/CBO9780511790942>
- Gibbs, N. F. (2009, February 22). "Tungiasis." . eMedicine.
- Grilli, L., & Rampichini, C. (2012). Multilevel models for ordinal data. *John Wiley and Sons, Ltd*. <https://doi.org/10.1002/9781119961154.ch19>
- Heukelbach, J., Wilcke, T., & Eisele, M. (2002). Ectopic localization of Tungiasis. *Am J Trop Med Hyg*, 67(2), 214-216.
- Jackman, S. (2009). Bayesian Analysis for the Social Sciences. *John Wiley and Sons, Ltd*. <https://doi.org/10.1002/9780470686621.ch8>
- Keiyoro, P., Ngunjiri, J., Mwanda, W., & Bowa, O. (2016). Influence of Disabilities caused by Tungiasis on School Attendance among School Age Children in Murang'a County, Kenya. *International Journal of Current Research*, 8(3), 28826-28830.
- Kockelman, K. M., & Young-Jun, K. (2002). Driver injury severity: An application of ordered probit models. *Elsevier*,

34(3), 313-321. [https://doi.org/10.1016/S0001-4575\(01\)00028-8](https://doi.org/10.1016/S0001-4575(01)00028-8)

- Lawson, A. B., Browne, W. J., & Vidal Rodeiro, C. L. (2003). Disease Mapping with WinBUGS and MLwiN. John Wiley & Sons Ltd.
- Liu, H., & Zhu, X. (2017). Joint modeling of multiple crimes: A bayesian spatial approach. *SPRS Int. J. Geo-Inf.*, 6(1), 16. <https://doi.org/10.3390/ijgi6010016>
- Muehlen, M., Heukelbach, J., & Wilcke, T. (2003). Investigations on the Biology, epidemiology, pathology and control of Tunga penetrans in Brazil. *Parasitol Res*, 90(6), 449-455.
- Mwangi, J. N., Ozwara, H. S., Motiso, J. M., & Gicheru, M. M. (2015). Characterization of Tunga penetrans Antigens in Selected Epidemic Areas in Murang'a County in Kenya. *PLoS Negl Trop Dis.*, 9(3), 1-10. <https://doi.org/10.1371/journal.pntd.0003517>
- Nsanzimana, J., Karanja, S., Kayongo, M., Umuhiza, H., Murangwa, A., Nyiramanzi, N., ... Musafili, A. (2019). Factors associated with tungiasis among primary school children: a cross-sectional study in a rural district in Rwanda. *BMC Public Health*, 19(1192). <https://doi.org/10.1186/s12889-019-7481-y>
- Nyangacha, R., Odongo, D., Oyieke F., Bii, C., Muniu, E., Chasia, S., & Ochwoto, M. (2019). Spatial distribution, prevalence and potential risk factors of Tungiasis in Vihiga County, Kenya. *PLOS Neglected Tropical Diseases*, 13. <https://doi.org/10.1371/journal.pntd.0007244>
- Qiu, Z., Song, P. X. K., & Tan, M. (2002). Bayesian hierarchical models for multi-level repeated data using WINBUGS. *Journal of Biopharmaceutical Statistics*, 12(2), 121-135. <https://doi.org/10.1081/BIP-120014415>
- Sanusi, I. D., Brown, E. B., Shepard, T. G., & Grafton, W. D. (1989). Tungiasis: Report of one case and review of the 14 reported cases in the united states. *J Am Acad Dermatol*, 20(5 Pt 2), 941-944
- Setboonsarng, S. (2005). Child malnutrition as a poverty indicator: An evaluation in the context of different development interventions in Indonesia. *Asian Development Bank*(21). Retrieved from <http://hdl.handle.net/11540/3609>
- Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. (2003). Winbugs user manual version 1.4. *Cambridge: MRC Biostatistics Unit*.
- Sundqvist, M. K., Sanders, N. J., & Wardle, D. A. (2013). Community and Ecosystem Responses to Elevational Gradients: Processes, Mechanisms, and Insights for Global Change. *Annual Review of Ecology, Evolution and Systematics*, 44, 261-280.
- Xianjin, H., Enqing, H., Yang, L., & Dazhi, W. (2016, April 7). *Altitudinal patterns and controls of plant and soil nutrient concentrations and stoichiometry in subtropical China*. PMID: 27052367. <https://doi.org/10.1038/srep24261>

## Appendix A

### The Value of $\beta'X_i$ with $X_{ijk}$ as $X_i$ for Severity cumulative logit model

$$\begin{aligned}\beta'X_i &= b_0 + b_1 * GEN + b_2 * ALTI + b_3 * SOpH + \\ &\quad b_4 * TotN. + b_5 * Pppm + b_6 * Kme. + b_7 * SAND. + \\ &\quad b_8 * CLAY. + b_9 * SILT. + b_{10} * POVHH + \\ &\quad b_{11} * AGE + \alpha_{j(i)} * CNTY + \alpha_{l(i)} * DA \\ \alpha_k * CNTY &\sim N(\gamma\alpha_{m(k)}, \sigma_{\alpha_k}^2) \\ \text{where } \gamma\alpha_{m(k)} &= \alpha_1 * PRECPTTION + \\ &\quad \alpha_2 * MXTMP + \alpha_3 * MNTMP + \alpha_4 * REHUM06Z + \\ &\quad \alpha_5 * REHUM12Z, \\ \text{with } \alpha_l * DA &\sim N(0, \sigma_{\alpha_l}^2) \text{ for } l = 1, \dots, 5\end{aligned}\tag{18}$$

where Gender – GEN, Altitude – ALT, Soil pH – SOpH, Total Nitrogen % – TotN., Phosphorus ppm – Pppm, Potassium me% – Kme., Sand % – SAND., Clay % – CLAY., Silt % – SILT., Poverty Measure HH – POVHH, Age – AGE, County ? CNTY, Precipitation – PRECPTTION, Max Temp. – MXTMP, Min Temp. – MNTMP, Relative Humidity at 06Z – REHUM06Z, Relative Humidity at 12Z – REHUM12Z, with Domestic Animal – DA Categorized as: No DA – 1, Sheep or Cow or Both – 2, Cat or Dog or Both – 3, Sheep or Cow & Cat or Dog or Both – 4 and Pig – 5.

## Appendix B

### The ordered Bayesian severity hierarchical (OBSH) categorical model

```

⇒

model
{
  for( i in 1:N)
  {
    TotN[i]<-(TotN.[i]-mean(TotN.[]))/sd(TotN.[])
    Kme[i]<-(Kme.[i]-mean(Kme.[]))/sd(Kme.[])
    SAND[i]<-(SAND.[i]-mean(SAND.[]))/sd(SAND.[])
    CLAY[i]<-(CLAY.[i]-mean(CLAY.[]))/sd(CLAY.[])
    SILT[i]<-(SILT.[i]-mean(SILT.[]))/sd(SILT.[])
    Ppm[i]<-(Pppm[i]-mean(Pppm[]))/sd(Pppm[])

    SEVTY[i] ~ dcat( p[i,] )
    p[i, 1] <- 1/( 1 + exp( -c[1] + mu[i] ) )
    p[i, 2] <- 1/( 1 + exp( -c[2] + mu[i] ) ) - 1/( 1 + exp( -c[1] + mu[i] ) )
    p[i, 3] <- 1/( 1 + exp( -c[3] + mu[i] ) ) - 1/( 1 + exp( -c[2] + mu[i] ) )
    p[i, 4] <- 1- 1/( 1 + exp( -c[3] + mu[i] ) )

    mu[i] <- b[1]*equals(POVHH[i],1) +b[2]*equals(DA[i],2)+ b[3]*equals(DA[i],3)+ b[4]*equals(DA[i],4)+
      b[5]*equals(DA[i],5) + b[6]*SAND[i] + b[7]*CLAY[i] + b[8]*Kme[i] + b[9]*TotN[i] +
      b[10]*SILT[i] + b[11]*Ppm[i] + b[12]*equals(CNTY[i],2) +b[13]*equals(CNTY[i],3)
  }
  c[1] <- g[1]
  c[2] <- g[1] + g[2]
  c[3] <- g[1] + g[2] + g[3]

  g[1] ~ dnorm(0, 1.0E-6)
  g[2] ~ dgamma(0.001, 0.001)
  g[3] ~ dgamma(0.001, 0.001)

  for(k in 1:13)
  {
    b[k]~ dnorm(0, 1.0E-6)
    exp.b[k]<-exp(b[k])
  }
}

```

⇐

→ Click here to open [data](#) ←

Initial values

list( b = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), g = c( 1, 1, 1))

## Appendix C

### Node statistics from the OBSH categorical model

#### Output of the Node Statistics from the OBSH Categorical

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
b[1]	-1.183	0.3795	0.0031	-1.937	-1.18	-0.4498	100000	100001
b[2]	-1.873	0.5027	0.004763	-2.872	-1.867	-0.9077	100000	100001
b[3]	0.3328	0.6022	0.003515	-0.8576	0.3352	1.509	100000	100001
b[4]	-0.8685	0.4363	0.004229	-1.728	-0.8669	-0.01549	100000	100001
b[5]	-2.334	1.159	0.005597	-4.619	-2.327	-0.08941	100000	100001
b[6]	0.9003	0.9077	0.02083	-0.8494	0.8917	2.683	100000	100001
b[7]	2.417	1.184	0.02769	0.1883	2.387	4.815	100000	100001
b[8]	0.545	0.1885	9.83E-4	0.1786	0.5438	0.9182	100000	100001
b[9]	-0.4105	0.1985	0.001047	-0.8081	-0.4071	-0.0294	100000	100001
b[10]	0.04477	0.5177	0.01123	-0.9187	0.02566	1.107	100000	100001
b[11]	-0.2257	0.1964	0.001168	-0.6154	-0.2243	0.1545	100000	100001
b[12]	-0.3948	0.4142	0.00338	-1.208	-0.3959	0.4126	100000	100001
b[13]	-1.458	0.4591	0.004888	-2.371	-1.452	-0.5691	100000	100001
g[1]	-2.142	0.528	0.006836	-3.194	-2.137	-1.128	100000	100001
g[2]	2.51	0.2698	0.001777	2.01	2.501	3.064	100000	100001
g[3]	0.387	0.1331	4.502E-4	0.1702	0.3727	0.6858	100000	100001

## Appendix D

### Spatial Bayesian Severity Hierarchical (*SBSH*) Models

```

⇒
model
{
  for( i in 1:N)
  {
    TotN[i]<-(TotN.[i]-mean(TotN.[ ]))/sd(TotN.[ ])
    Kme[i]<-(Kme.[i]-mean(Kme.[ ]))/sd(Kme.[ ])
    SAND[i]<-(SAND.[i]-mean(SAND.[ ]))/sd(SAND.[ ])
    CLAY[i]<-(CLAY.[i]-mean(CLAY.[ ]))/sd(CLAY.[ ])
    SILT[i]<-(SILT.[i]-mean(SILT.[ ]))/sd(SILT.[ ])
    Ppm[i]<-(Pppm[i]-mean(Pppm.[ ]))/sd(Pppm.[ ])
    ALT[i]<-(ALT[i]-mean(ALT[i]))/ sd(ALT[i])
    SEVTY[i] ~ dcat( p[i,] )
    p[i, 1] <- 1/( 1 + exp( -c[1] + mu[i] ) )
    p[i, 2] <- 1/( 1 + exp( -c[2] + mu[i] ) ) - 1/( 1 + exp( -c[1] + mu[i] ) )
    p[i, 3] <- 1/( 1 + exp( -c[3] + mu[i] ) ) - 1/( 1 + exp( -c[2] + mu[i] ) )
    p[i, 4] <- 1- 1/( 1 + exp( -c[3] + mu[i] ) )
    mu[i] <- b[1]*equals(POVHH[i],1) +b[2]*equals(DA[i],2)+ b[3]*equals(DA[i],3)+ b[4]*equals(DA[i],4)+
      b[5]*equals(DA[i],5) + b[6]*SAND[i] + b[7]*CLAY[i] + b[8]*Kme[i] + b[9]*TotN[i] +
      b[10]*SILT[i] + b[11]*Ppm[i] + b[12]*ALT[i] +b.CNTY[CNTY[i]] + U[NOWDS[i]] + S[NOWDS[i]]
  }
  c[1] <- g[1]
  c[2] <- g[1] + g[2]
  c[3] <- g[1] + g[2] + g[3]
  g[1] ~ dnorm(0, 1.0E-6)
  g[2] ~ dgamma(0.001, 0.001)
  g[3] ~ dgamma(0.001, 0.001)
  for( j in 1:n.CNTY) {b.CNTY[j] ~ dnorm( 0, tau.CNTY)}
  tau.CNTY ~ dgamma(0.1, 0.001)
  sigma.CNTY <- pow(tau.CNTY, -1)
  # CAR prior distribution for random effects:
  S[1:n] ~ car.normal(adj[], weights[], num[], tau.S)
  for(k in 1:sumNumNeigh) {weights[k] <- 1 }
  for(l in 1:n) {U[l] ~ dnorm(0, tau.U)}

  for(m in 1:12){b[m]~ dnorm(0, 1.0E-6)
    exp.b[m]<-exp(b[m])
  }
  tau.U ~ dgamma(0.1, 0.01) #prior unstructured precision
  tau.S ~ dgamma(0.1, 0.01) #prior spatial precision
  sigma.U <- sqrt(1/tau.U) # unstructured standard deviation
  sigma.S <- sqrt(1/tau.S) # spatial standard deviation
  #check psi to see whether spatial or unstructured variation dominates
  psi <- sigma.S/(sigma.U + sigma.S)
}
⇒

```

## Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).