January 2013

# Modeling The Epidemiological Impact Of Human Hookworm Infection

Laura Skrip
*Yale University*, laura.skrip@yale.edu

# MODELING THE EPIDEMIOLOGICAL IMPACT OF HUMAN HOOKWORM INFECTION

**By Laura Skrip**

A Thesis Presented to

The Faculty of the Department of Epidemiology and Public Health

Yale University

In Candidacy for the Degree of

Master of Public Health

2013

## PERMISSION TO COPY

Permission for photocopying, microfilming, or computer electronic scanning of "Modeling the Epidemiological Impact of Human Hookworm Infection" for the purpose of individual scholarly consultation of reference is hereby granted by the author. This permission is not to be interpreted as affecting publication of this work or otherwise placing it in the public domain, and the author reserves all rights of ownership guaranteed under common law protection of unpublished manuscripts.

Laura Skrip

Signature of Author

May 1, 2013

Date

## ABSTRACT

**Background:** Chronic hookworm infection affects an estimated 576 to 740 million people worldwide. Despite mass drug administration efforts, the morbidity associated with this soil-transmitted helminth remains a significant public health issue. Due to inter- and intra-community heterogeneity in susceptibility to infection, mathematical modeling can serve as an effective and efficient tool for investigating hookworm transmission in different settings and for guiding policymakers to consider new treatment and prevention strategies.

**Methods:** Here we developed an age-structured, compartmental S-I model to identify epidemiological parameters for hookworm infection and to assess rates of attributable anemia within the population of Zanzibar. The model was first used to address the relative contributions of age group (adults versus children) and infection intensity status (high versus low) in transmission. The Markov chain Monte Carlo (MCMC) method was implemented to generate negative binomial distribution parameters for describing population-level parasite aggregation. Gibbs sampling and data on prevalence of hookworm infection were used in a subsequent MCMC to parameterize the overall model and account for uncertainty. Maximum likelihood point estimates for force of infection and natural recovery were derived from the 10,000 posterior distributions generated. Empirical data on average hemoglobin levels for given fecal egg count categories were included in the analysis to give distributions for the relative risk of moderate-to-severe anemia (hemoglobin $< 90$ g/L) among infected versus susceptible individuals.

**Results:** Using maximum likelihood estimates for the negative binomial distribution parameters, it was determined that children prone to high intensity hookworm infections were 4.9 times as infectious as children prone to lower intensity infection; high intensity infection adults were

considered to be 4.7 times as infectious as low intensity infection adults. Model predictions with and without uncertainty were consistent in estimating that approximately 10% of the infected child population experienced moderate-to-severe anemia attributable to hookworm infection.

**Conclusion:** The present model provides a population-level compartmental model framework for assessing hookworm-attributable morbidity independent of assumptions about worm burden thresholds. To evaluate the effectiveness of different intervention strategies at reducing the incidence of infection and rates of hookworm-attributable anemia, the current epidemiological model can be extended structurally to include treatment and vaccine parameters.

## ACKNOWLEDGMENTS

I would like to extend my appreciation to Jennifer Gilbert, MPH for her mentorship and patience in helping me further develop the knowledge, skills, and enthusiasm for considering public health questions through mathematical modeling. Additionally, I would like to thank Professors Alison Galvani and Jeffrey Townsend for their thoughtful guidance and advice in planning and implementing my model. I am incredibly grateful for the time they spent sharing their knowledge and expertise with me.

## TABLE OF CONTENTS

# INTRODUCTION

## *Background*

Affecting an estimated 576-740 million people worldwide, human hookworm infection contributes significantly to the global incidence of iron-deficiency anemia, protein malnutrition, and other comorbid conditions [1-3]. In fact, chronic hookworm infection accounts for nearly 22.1 million of the 57 million DALYS associated with all neglected tropical diseases [1]. Most human infections are attributed to two hookworm species, *Necator americanus* and *Ancylostoma duodenale*, which are endemic in tropical and subtropical regions of sub-Saharan Africa, south China, Southeast Asia, and the Americas [4]. The climate in these regions is conducive to larval survival, and in less developed regions, poor nutritional status combined with lacking sanitation infrastructure provides for significant morbidity due to high transmission rates and chronic infection [5].

Hookworm infection results as third-stage larvae (L3) penetrate exposed skin upon contact with contaminated soil. The larvae enter the circulatory system and migrate to the lungs. After coughing expels them from the respiratory tract, the parasites are swallowed into the digestive system and form a buccal capsule equipped with cutting plates [6, 7]. Larvae ultimately mature into adult hookworms in the small intestine, where their feeding on the mucosa and submucosa leads to a daily loss of up to 9 ml of blood [8]. While in the intestine, adult hookworms reproduce and their eggs are excreted with fecal matter [9]. Low iron bioavailability in hookworm-endemic regions increases susceptibility to iron deficiency anemia due to diets heavily dependent on starch-based foods, such as cassava [10].

## *Current and Projected Control Measures*

Over the past several decades, widespread deworming efforts targeting high-risk

individuals have been undertaken to reduce the prevalence and intensity of STH infection [11]. Such programs have involved the administration of single-doses of benzimidazoles, such as mebendazole (500 mg dose) and albendazole (400 mg dose), in schools and other community settings [12]. In 2001, the World Health Assembly established a goal of treating 75% of all at-risk school-aged children with anthelminthics by 2010 [13]. However, it is projected that long-term mass treatment using benzimidazoles will be increasingly less effective due to rapid rates of reinfection and the likely emergence of drug resistance [14-16]. In addition, individual- and community-level differences in benzimidazole efficacy have been demonstrated [11], with a recent meta-analysis of randomized placebo-controlled studies reporting cure rates for single-dose oral albendazole (400 mg) as low as 40% in Kenya (n = 34) during 1999 to as high as 100% in Haiti (n = 12) during 1990 [17].

In response to limited treatment options and variable efficacy, efforts are currently underway to identify antigens for development of a human hookworm vaccine. The current approach involves investigation into a bivalent vaccine that will affect both the larval and adult stages of the parasite's life cycle [4]. One potential class of antigens includes the aspartic proteases (APRs), which are proteolytic enzymes that interfere with the parasite's digestion of skin macromolecules and thus inhibit larval movement across the epithelial barrier and into the circulatory system [7, 8]. Another class, the *Ancyclostoma* secreted proteins (ASPs), are involved in maturation of L3 larvae to adult hookworms upon human serum stimulation [8]. Table 1 outlines recently investigated vaccine targets and the current phase of development for each.

### *Study Target Population*

The island region of Zanzibar is located off the coast of Tanzania. In 1994, a school-based deworming program was launched by the local Ministry of Health and Social Welfare and involved administration of single does of mebendazole two to three times per year on Pemba, the smaller of the two islands making up Zanzibar [18]. The program was eventually extended to all of Zanzibar with reported coverage rates as high as 90% of all primary school children [19]. The school-based program was disrupted in 2000 due to difficulties in securing sufficient anthelmintic drugs. It was restarted in 2003 with the administration of albendazole and praziquantel [20]. In 2001, Zanzibar also launched a community-wide campaign, The Global Elimination of Lymphatic Filariasis (GELF), which involved annual administration of albendazole and ivermectin to all eligible adults and children. The combined effect of GELF and school-based deworming programs on helminth infection has been investigated [21]. The safety of combining community-based programs to address STH and other endemic parasitic infections has been considered as well [22]. Despite decades of large-scale deworming programs, infection prevalence remains as high as 99% in regions of the islands and some degree of iron deficiency anemia has been found to afflict about 60% of children [18, 19]. Due to extensive evaluation of the impact of mass drug administration in Zanzibar, trends in hookworm prevalence and disease intensity have been well documented. The region would thus be a feasible and logical site for future effectiveness trials on vaccination and other intervention strategies.

### *Existing Models*

Mathematical modeling offers an effective and efficient way of evaluating hookworm transmission dynamics and considering how new intervention strategies could impact rates of

morbidity. Since the 1980's, models on helminth infection have evolved to include new approaches to treatment, prevention, and control. Anderson and May proposed an early worm-based model for evaluating the basic reproductive rate ($R_0$) for macroparasites and considering the impact of treatment in terms of changing the mean worm burden and reducing $R_0$ to below unity [23]. Medley et al expanded upon this model to specifically investigate the effect of anthelminthic treatment on symptoms attributable to roundworm (*Ascaris lumbridoides*) infection [24]. The use of a worm burden threshold to model levels of morbidity was further extended by Chan and his colleagues who incorporated age structuring into the Anderson and May model and compared predictions when different mechanisms of population-level host mixing were considered [25]. Most recently, Sabatelli et al developed an individual-based model to consider the impact of human hookworm vaccination in the context of community-level parasite aggregation and variation in host susceptibility [5].

Despite significant progress in the development of mathematical models for helminth infections over the last several decades, the Disease Reference Group on Helminth Infections (DRG4), established by the Special Programme for Research and Training in Tropical Diseases (TDR), recently issued an agenda highlighting research priorities and gaps to address when modeling helminth infections [26]. Notably, it was recommended that models should be refined to more accurately capture the relationship between morbidity and the evolving infection status of individuals in regions with intermittent treatment and control campaigns [26].

***Study Objectives***

Here we construct and parameterize an age-structured, risk-stratified deterministic model to identify epidemiological parameters for hookworm infection and attributable morbidity within

the population of Zanzibar. For the present model, morbidity was defined as anemia with a hemoglobin level less than 90 g/L. An expanded model representing treatment and vaccination is also presented but not parameterized. As far as we know, this is the first population-based deterministic model for evaluating the epidemiological impact of human hookworm infection.

## METHODS

### *Description of Compartmental Model*

A deterministic model was constructed to evaluate the population-level transmission dynamics of human hookworm infection and estimate moderate-to-severe anemia among infected children in Zanzibar. This age-structured model (Figure 1) includes child compartments (subscript C) representing individuals younger than 15 years old and adult compartments (subscript A) representing individuals 15 years and older. The age cutoff was chosen since treatment programs in Zanzibar target school-age children and the model was developed for ultimate use in evaluating similarly targeted intervention strategies. To reflect community-level overdispersion of hookworm burden, both age classes were subdivided with individuals prone to high intensity infection (subscript 20) distinguished from those prone to low intensity infection (subscript 80). Individuals in the high intensity compartments are expected to be more susceptible to infection due to nutritional, behavioral, and genetic factors as well as past infection experiences [23, 27]. The negative binomial distribution is frequently used when modeling parasite aggregation, and within high aggregation communities, it is often found that 20% of a given population harbors over 80% of the burden [28] such that for the present model, both age classes were risk-stratified to distinguish between those prone to heavy (20% of population) versus light (80% of population) intensity infections. The overall S-I model structure assumed underlying levels of anemia due to nutritional and other factors. Moderate-to-severe anemia was defined in terms of hemoglobin levels less than 90 g/L. This cut off is consistent with that previously used during trials to investigate the effectiveness of hookworm control measures in Zanzibar [10]. Children and adults transition from the S compartments to I compartments at a rate defined by the force of infection $\gamma$, given by the following equation:

$$\gamma = \beta_{C80} \times \frac{I_{C80}}{N} + \beta_{A80} \times \frac{I_{A80}}{N} + \beta_{C20} \times \frac{I_{C20}}{N} + \beta_{A20} \times \frac{I_{A20}}{N} \quad (1)$$

All infected individuals are assumed to acquire hookworms through contact with contaminated soil and have a chance of natural recovery back to the S compartment at a rate, ρ. The movement of individuals between compartments is given by eight (8) ordinary differential equations.

$$\frac{dS_{C80}}{dt} = b_{80}N - d_C S_{C80} - \gamma S_{C80} - \alpha S_{C80} + \rho I_{C80} \quad (2)$$

$$\frac{dS_{C20}}{dt} = b_{20}N - d_C S_{C20} - \gamma S_{C20} - \alpha S_{C20} + \rho I_{C20} \quad (3)$$

$$\frac{dI_{C80}}{dt} = \gamma S_{C80} - d_C I_{C80} - \rho I_{C80} - \alpha I_{C80} \quad (4)$$

$$\frac{dI_{C20}}{dt} = \gamma S_{C20} - d_C I_{C20} - \rho I_{C20} - \alpha I_{C20} \quad (5)$$

$$\frac{dS_{A80}}{dt} = \alpha S_{C80} + \rho I_{A80} - d_A S_{A80} - \gamma S_{A80} \quad (6)$$

$$\frac{dS_{A20}}{dt} = \alpha S_{C20} + \rho I_{A20} - d_A S_{A20} - \gamma S_{A20} \quad (7)$$

$$\frac{dI_{A80}}{dt} = \gamma S_{A80} + \alpha I_{C80} - d_A I_{A80} - \rho I_{A80} \quad (8)$$

$$\frac{dI_{A20}}{dt} = \gamma S_{A20} + \alpha I_{C20} - d_A I_{A20} - \rho I_{A20} \quad (9)$$

***Parameterization and Uncertainty Analysis***

Demographic and disease parameters were used to define the transitions into and out of compartments. All parameters for the transmission-only model are defined in Table 4.

***(a) Demographic Parameters***

Census data were used to identify and calibrate point estimates for birth rate, child- and adult-specific death rates, and the rate of aging from child to adult compartments. Calibration involved setting the birth rate equal to the estimate derived from census data and adjusting the

death and aging rates to achieve both empirical population-level growth rates and age structuring, respectively [29].

***(b) Force of Infection Equation***

According to our model, the force of infection addresses two between-compartment relationships for the infectiousness of different groups of individuals.

The first relationship involves the relative infectiousness of individuals in high versus low infection intensity compartments. This relationship was quantified as a multiplicative factor relating $\beta_{80}$ and $\beta_{20}$ and was calculated as a ratio of the expectation of the lower 80 percent of the density function for population-level fecal egg counts and the expectation of the upper 20 percent of the same density function. Egg counts were used as an indicator of infectiousness since excreted hookworm eggs contaminate soil in areas lacking developed sanitation systems and subsequently result in the development of infectious L3 hookworm larvae [6]. A binned negative binomial distribution and a flat prior were used to model the probability of a given egg count in eggs per gram (epg) of fecal sample. Posterior distributions for mean and variance parameters of a negative binomial distribution were generated through the Markov chain Monte Carlo (MCMC) method. Although the likelihoods of $\mu$ and $\sigma^2$ estimates were considered in the MCMC, equations $R = \left({\mu^2}/{\sigma^2}\right)/\left(1 - {\mu}/{\sigma^2}\right)$ and $P = {\mu}/{\sigma^2}$, where $\sigma^2 = \mu + {\mu^2}/{k}$, were utilized to produce distributions for the number of successes (R) and probability of success (P) parameters specific to the calculation of the negative binomial cumulative distribution function in Matlab. A single MCMC was conducted for children and adults since independence of the study populations could not be assumed; however, separate $\mu$ and $\sigma^2$ parameters of the negative binomial distribution were evaluated for each of the two age classes. The likelihood function (Equation 10) for this analysis was based on population-level data that classified individuals into

four categories of hookworm infection intensity based on fecal egg counts [10]. The function depends on $p_i = F(z) - F(y)$, or an individual's probability of being in a given fecal egg count category (range: z epg – y epg), as determined by the negative binomial cumulative distribution function; $n_{ci}$ and $n_{ai}$ represent the number of children and adults, respectively, from the study sample in the i$^{th}$ non-zero fecal egg count category.

$$\ell(\mu_C, \sigma_C^2, \mu_A, \sigma_A^2) = \sum_{i=1}^{4} n_{ci} * \log(p_{ci}) + \sum_{i=1}^{4} n_{ai} * \log(p_{ai}) \quad (10)$$

Best-fit estimates for the four parameters were considered as those that maximized the log likelihood equation in 200,150 iterations, with an expected burn-in period of 150 iterations. The negative binomial distributions generated by the best-fit parameters resulting from this MCMC were used to determine the expected egg count of an individual in the lower 80% of the density function and the expected egg count of an individual in the upper 20% of the density function, and a ratio was taken. To address the uncertainty in the estimates, the entire parameter distributions from the MCMC were used to produce vectors of ratios to be applied as a prior distribution for the multiplicative factor relating $\beta_{C80}$ and $\beta_{C20}$ or that relating $\beta_{A80}$ and $\beta_{A20}$ in the uncertainty analysis.

The second relationship within the force of infection equation considered the relative infectiousness of adults versus children. Unlike parasite burden of other soil-transmitted helminth infections which peaks during adolescence and subsequently declines, the mean number of hookworms per person increases gradually until plateauing during adulthood [6]. However, considerable between-community heterogeneity has been observed in the relative infectiousness of adults versus children [25, 30]. In accord with the findings of Chan et al [25] for a comparable sub-Saharan African community, it was assumed for the present analysis that Zanzibari children are twice as infectious as adults due to less hygienic practices and a higher

basic reproductive rate of worms; however, other relationships were considered in the uncertainty analysis.

***(c) Disease Parameters***

Point estimates for baseline prevalence of hookworm infection for children and adults in Zanzibar were extracted from the available literature (Table 1). Prevalence estimates were assumed to follow the binomial distribution and be independent of one another. The MCMC method was used to determine the values of $\beta_{C80}$ and ρ that maximized a log likelihood function based on the experimental prevalence data (Equation 11). The beta-binomial distribution was assumed for the prevalence estimates. In the equation, $p_c$ represents the prevalence estimate for children derived from a cross-sectional study by Stoltzfus et al [31] on Pemba Island, Zanzibar prior to mass drug administration and $p_a$ represents a prevalence estimate for Zanzabari adults [32], while $p_c'$ and $p_a'$ represent the prevalence levels predicted with a given set of parameter estimates.

$$\ell(\beta_{C80}, \rho) = p_c n_c * \log(p_c') + (n_c - p_c n_c) * \log(1 - p_c') + p_a n_c * \log(p_a') + (n_a - p_a n_a) * \log(1 - p_a') \quad (11)$$

Uninformative flat prior distributions were assumed for the two disease parameters. For the uncertainty analysis, iterative Gibbs sampling was used to draw from the posterior distributions for the multiplicative factors relating $\beta_{C80}$ and $\beta_{C20}$ or that relating $\beta_{A80}$ and $\beta_{A20}$ to parameterize the force of infection equation.

***(d) Age Class Multiplier***

Significant inter-community heterogeneity in the relative infectiousness of adults and children has been documented by Chan et al [25, 30]. For the present paper, it was assumed that children in Zanzibar are more infectious due to behavioral factors. However, to address the

possibility of other relationships, the age-class multiplier was added as an additional parameter in the MCMC and allowed to vary randomly to consider the possible scenarios of increased infectiousness among adults and equal infectiousness between adults and children; the trend of the resulting MCMC output was visually inspected for convergence.

***(e) MCMC Convergence Diagnostic***

The MCMC process was repeated to generate a series of two chains for each parameter; a burn-in period of 150 iterations of overdispersed values was used. Convergence of posterior distributions was tested using the Gelman-Rubin diagnostic [33, 34]. This diagnostic uses the potential scale reduction factor ($\hat{R}$), involving the estimated variance of a given parameter based on both the within- and between-chain variances, to determine whether the parameter distributions generated through the MCMC process converge to a stationary distribution (Criterion: $\hat{R} < 1.2$). $\hat{R}$ was calculated and the plot reflecting change in Gelman and Rubin's shrink factor was evaluated for each parameter [33].

***Morbidity Measures***

Using published data on hemoglobin levels for a sample of Zanzibari children categorized according to five fecal egg count bins [31], the proportions of the susceptible (0 eggs per gram feces) and infected (>0 eggs per gram feces) populations with anemia were calculated. For individuals within a given fecal egg count range, hemoglobin (Hb) measurements were assumed to follow a normal distribution. Underlying levels of anemia due to nutritional status, malaria, and other factors were also assumed. The proportion of individuals with anemia not attributable to hookworm infection was found as the density of the cumulative distribution function below 90 g/L Hb for uninfected individuals. A point estimate for the proportion of

susceptible children with anemia was calculated using the mean and standard deviation for hemoglobin levels among individuals in the 0 eggs per gram feces category. For each fecal egg count category (i) of infected individuals, the density of the normal cumulative distribution function (A) under 90 g/L was multiplied by the proportion of the population with egg counts in that range (z epg – y epg), as determined using the best-fit parameters of the negative binomial distribution (F) found in the first MCMC (Equation 12).

$$Proportion\ with\ Anemia = \sum_{i=1}^{5}([A_i(90) - A_i(0)] * [F(z_i) - F(y_i)]) \quad (12)$$

For both susceptible and infected individuals, the uncertainty in the empirical data was addressed by drawing 10,000 random numbers from a normal distribution with the published mean and standard deviation for each fecal egg count range and randomly selecting a new mean for the distribution of anemia. Additionally, Gibbs sampling was used to select mean and variance estimates generated in the first MCMC for parameterizing the negative binomial distribution in the uncertainty analysis.

## RESULTS

### *Data*

Data on prevalence of hookworm infection in adult and child populations from Zanzibar or representative areas of eastern Tanzania were extracted from the available literature. Pre-treatment estimates for prevalence of hookworm infection were found for children and adults (Table 2). The number of individuals in the high intensity compartments was calculated as the 20% of the population expected to harbor the majority of the hookworm burden [6, 23].

### *MCMC Convergence*

The Markov chain Monte Carlo (MCMC) method was used to generate posterior distributions for mean and variance parameters of negative binomial distributions describing fecal egg counts in child and adult populations. The Gelman-Rubin diagnostic ($\hat{R}$) was found to be between 1.0 and 1.04 upon comparing two chains of parameter values for each of the four parameters considered in the MCMC. A similar analysis of convergence was conducted to evaluate the chains of posterior distributions of the two disease transmission parameters ($\beta_{C80}$ and $\rho$) generated through the second MCMC. Figure 2 reflects the tendency toward stationary distributions for the disease parameters based on the evolution of Gelman and Rubin's shrink factor.

When the age-class multiplier was allowed to vary to explore other possible relationships in the relative infectiousness of adults versus children, no convergence for this parameter was observed. As a result, the relationship was fixed with children being twice as infectious as adults ($\beta_{A80} = \frac{1}{2} * \beta_{C80}$) for the remaining analyses.

***Model Predictions Using Point Estimates***

Best-fit point estimates were determined as values maximizing the log likelihood functions for parameterizing the negative binomial distribution and subsequently the transmission model. All parameter estimates are presented in Tables 3 and 4. It was found that children in the model had, on average, 1303 hookworm eggs per gram feces and that, on average, adults had 1622 hookworm eggs per gram feces. Children prone to high intensity infections were found to be 4.9 times more infectious as children prone to low intensity infections in terms of their relative fecal egg counts. Adults prone to high intensity infections were similarly found to be 4.7 times more infectious as adults prone to low intensity infections. Using this information to parameterize the force of infection equation and run the transmission model, it was determined that moderate-to-severe anemia affected 6.2% of susceptible children. Among infected children, 16.4% were found to have hemoglobin levels less than 90 g/L. Accordingly, the model predicted the prevalence of hookworm-attributable anemia among children to be 10.2%.

***Model Predictions with Uncertainty***

The Markov chain Monte Carlo process was implemented twice to account for uncertainty in the data used. The first MCMC considered the distributions of fecal egg counts for adults and children. A jump size of 250 (acceptance ratio: 0.5296) and a jump size of 0.32 (acceptance ratio: 0.5834) were used for $\mu_C$ and $\sigma_C^2$, respectively. A jump size of 445 (acceptance ratio: 0.5122) and a jump size of 0.40 (acceptance ratio: 0.6334) were used for $\mu_A$ and $\sigma_A^2$, respectively. The resulting distributions of likely means and variances for the negative binomial distribution were used to find either the risk-class multiplier relating $\beta_{C80}$ and $\beta_{C20}$ or that relating $\beta_{A80}$ and $\beta_{A20}$ (Figures 3 and 4). When taking uncertainty into account, it was found

that the relative infectiousness of children in the high intensity infection compartment compared to that of children in the low intensity infection compartment ranged between 3.4 and 7.2. The infectiousness of the adults in the high intensity infection compartment was found to be between 2.8 and 12.2 times that of adults in the low intensity infection compartment. Ranges of parameter estimates from the uncertainty analysis are presented in Table 3. The second MCMC used Gibbs sampling to evaluate the force of infection equation and prevalence data to parameterize the overall transmission model. A jump size of 0.0025 (acceptance ratio: 0.2699) and a jump size of 0.0001 (acceptance ratio: 0.3115) were used for $\beta_{C80}$ and $\rho$, respectively. It was determined that underlying levels of anemia affect between 0% and 99.7% of the modeled child population ($\mu = 14.2\%$). Among the infected children compartments, it was found that between 0% and 93.1% had hemoglobin levels less than 90 g/L ($\mu = 24.1\%$). Figures 5 and 6 are histograms representing the frequencies of proportions for moderate-to-severe anemia in susceptible and infected child populations, respectively. Accordingly, when accounting for uncertainty, the model predicted rates of childhood anemia attributable to hookworm infection is, on average, 9.9%.

## DISCUSSION

### *Results and Implications*

Chronic hookworm infection is a leading but preventable cause of anemia and other morbidity throughout communities that lack access to sufficient hygiene and sanitation resources. Despite efforts at mass drug administration, the global prevalence of hookworm infection remains high and post-treatment rebounding suggests that current measures are not sustainable [16]. Mathematical modeling offers an effective tool for considering population-level transmission dynamics and provides the framework for evaluating the impact of future interventions against hookworm infection. The present model is unique in that it is the first population-level compartmental model for human hookworm infection and that it does not depend on a worm burden threshold for determining levels of morbidity. Worm- and individual-based models have considered a wide range (40-160 worms) as the threshold for morbidity and extrapolating worm burden from fecal egg counts is variable and often unreliable [35, 36]. Using empirical distributional data for hemoglobin levels and well-established trends in parasite aggregation, we were able to determine the proportion of children with underlying anemia versus anemia attributable to hookworm infection both with and without model uncertainty.

### *Limitations*

The present findings are limited by assumptions about the population-level distribution of hookworm infection. In particular, risk stratification in the current model depends on the "80-20 Rule" which accounts for both parasite aggregation and the observed tendency of a small proportion of the population having increased susceptibility to higher intensity infections. Specifically, according to this rule, it is expected that 20% of the population in a given

community harbors over 80% of the total worm burden [23]. The degree of parasite aggregation that corresponded with the data used to generate negative binomial distributions in the present study was considered in an *a posteriori* analysis. Gibbs sampling was performed to randomly select a mean/variance pair for children or adults and parameterize a negative binomial distribution. Using the resulting distributions, 3213 and 1077 observations were randomly generated to represent the worm burdens for samples comparable in size to the original experimental samples of children and adults, respectively [31, 32]. For the distribution representing each age class, all observations were sorted and normalized by their sum. Upon sorting, the observation representing the cutoff between the lower 80% of observed egg counts and the upper 20% of observed egg counts was determined. This process was repeated for 5000 iterations. The mean density above the 80% cutoff (the 2571$^{st}$ observation for children and the 862$^{nd}$ observation for adults) was found to be 70.09% for children and 71.11% for adults.

The current model uses the 80-20 cutoff as a way of subgrouping the population according to risk level, so the fact that the empirical distribution does not follow the 80-20 expectation does not immediately impact the findings. However, despite wide use of this rule for risk stratification, an empirically supported approach to categorizing by host susceptibility and parasite aggregation should be investigated and the impact of this particular cutoff on our transmission model should be considered in a sensitivity analysis.

***Future Directions***

The transmission model was expanded structurally to account for treatment and vaccination against hookworm infection (Figure 7). The intervention parameters and variables are described in Table 6. Before implementation of the full intervention model, it will be

important to consider how treatment and vaccination efficacy will be measured. Efficacy measured as the transition from infected to susceptible is expected to be an overly conservative approach since reduction in infection intensity without complete cure is sufficient to reduce morbidity. That is, anemia and other comorbidities are more prevalent among individuals with high intensity infections [6]. To capture this in the S-I compartmental framework, we will evaluate shifts in the negative binomial distribution—widely accepted to reflect parasite aggregation and individual-level susceptibility to hookworm infection [23]. It is expected that a compartmental model which stochastically models individual predisposition to infection, parasite aggregation in the host population, and parasite density dependence will sufficiently capture this distributional change at the population level without the computational requirements of an individual-based model. Treatment and vaccine efficacies will be modeled as binomial probabilities of an individual hookworm's death due to intervention. This approach was used by Sabatelli et al in modeling the impact of anthelminthic treatment [5]. Reinfection studies are expected to provide insight into the distributional nature of levels of susceptibility and the sensitivity of the model to risk stratification, as individuals with high intensity infection are more likely to return to the same infection intensity status after treatment effects wane.

***Conclusion***

In summary, despite the limited availability of empirical data, an uncertainty analysis of our hookworm model provides insight into population-level transmission dynamics and allows for estimation of directly unmeasurable quantities such as relative infectiousness and parasite aggregation. Underlying trends in infection were evaluated in terms of age-class and risk-class multipliers and rates of moderate-to-severe anemia among both susceptible and infected child

compartments were determined using distributional data on hemoglobin levels across fecal egg count categories instead of relying on a worm burden threshold. In addition, the structural framework for a compartmental model accounting for treatment and prophylactic measures was provided and recommendations were offered for parameterization and interpretation.

## FIGURES AND TABLES

FIGURE 1: Model of Human Hookworm Transmission

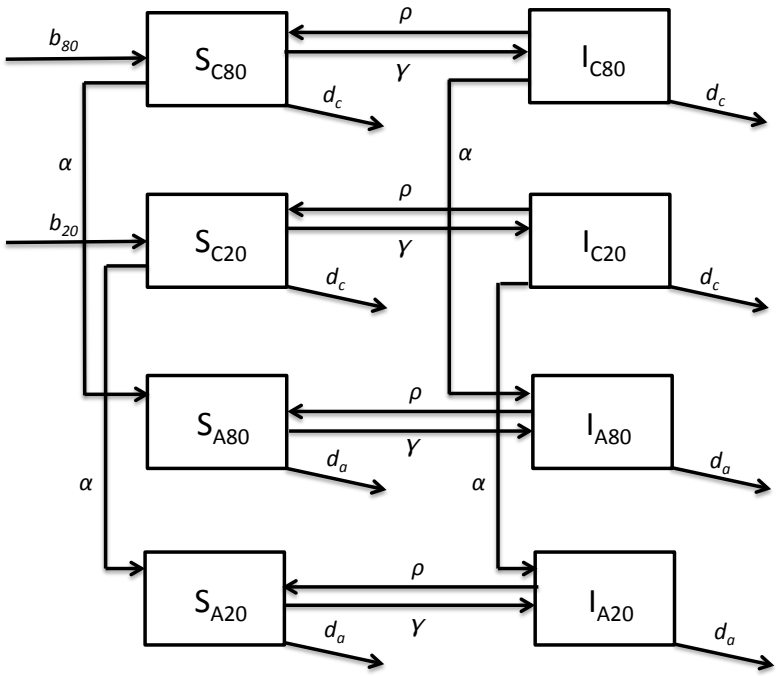


FIGURE 2: Demonstration of Markov chain Monte Carlo (MCMC) Convergence

Figure 2a. Gelman and Rubin shrink factor for $\beta_{C80}$

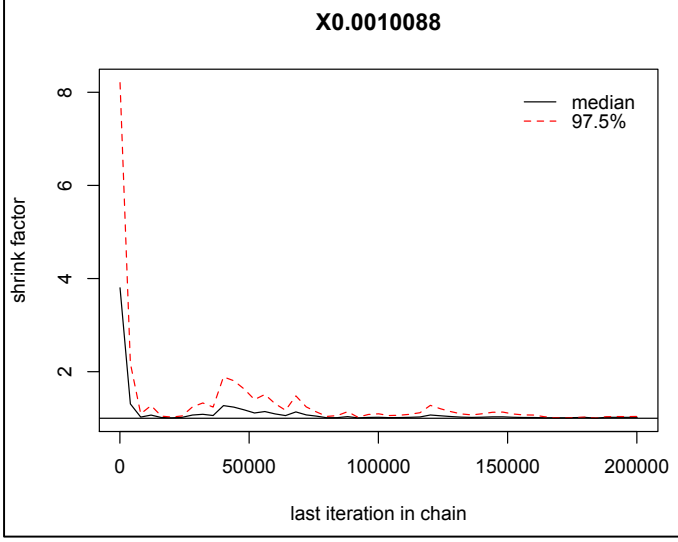


Figure 2b. Gelman and Rubin shrink factor for $\rho$

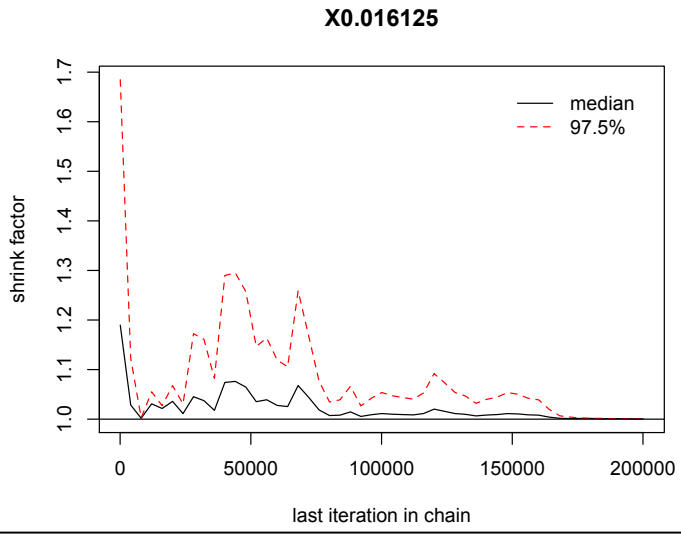

FIGURE 3: Histogram of Multiplier Relating $\beta_{A20}$ and $\beta_{A80}$ for Mean Egg Count of High Intensity Versus Low Intensity Infection Adults

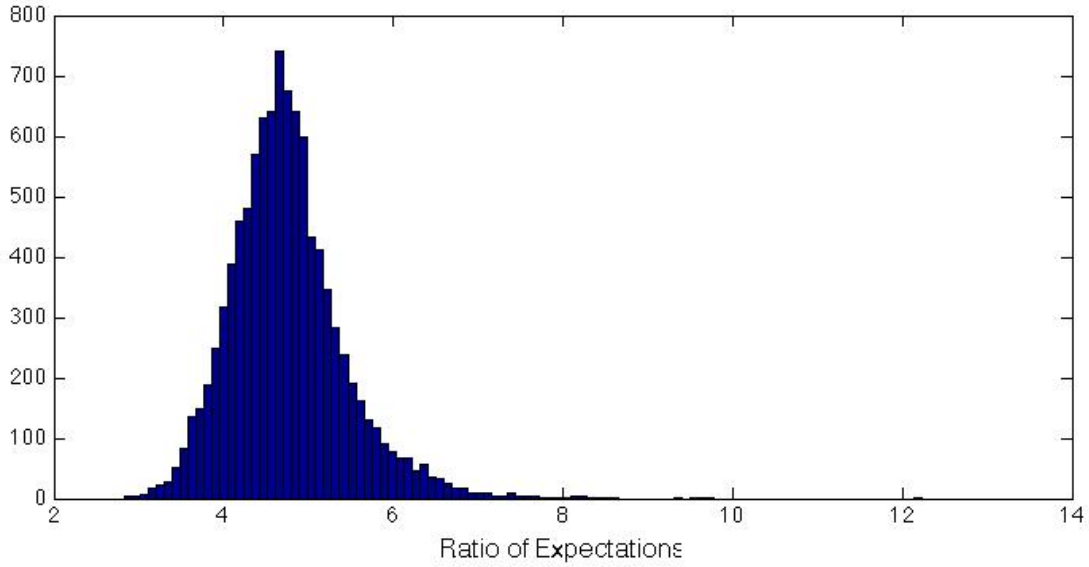


FIGURE 4: Histogram of Multiplier Relating $\beta_{C20}$ and $\beta_{C80}$ for Mean Egg Count of High Intensity Versus Low Intensity Infection Children

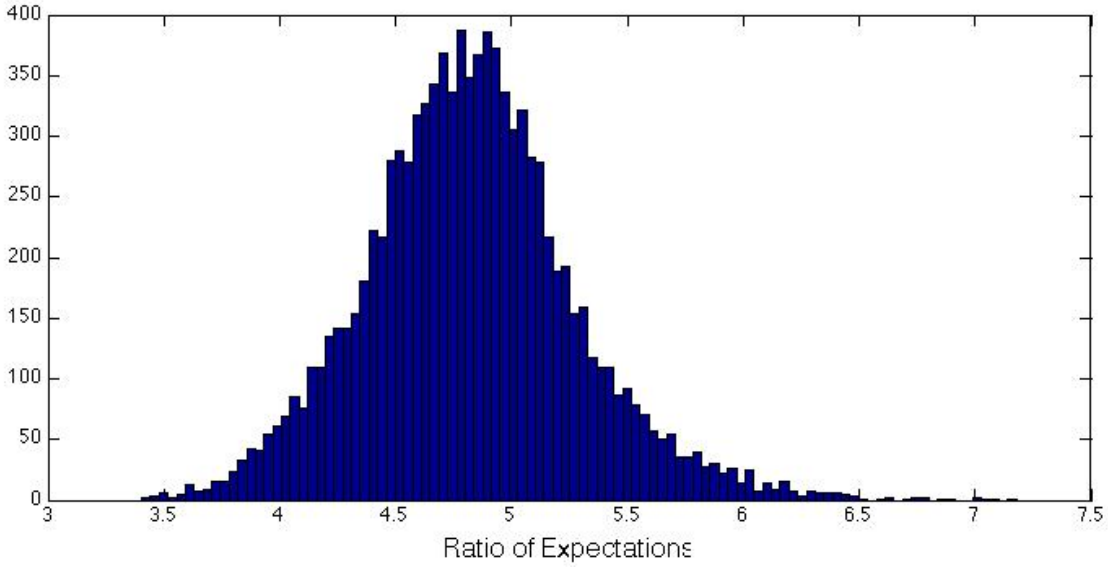

FIGURE 5. Histogram for the Proportion of the Susceptible Child Population with Moderate-to-Severe Anemia

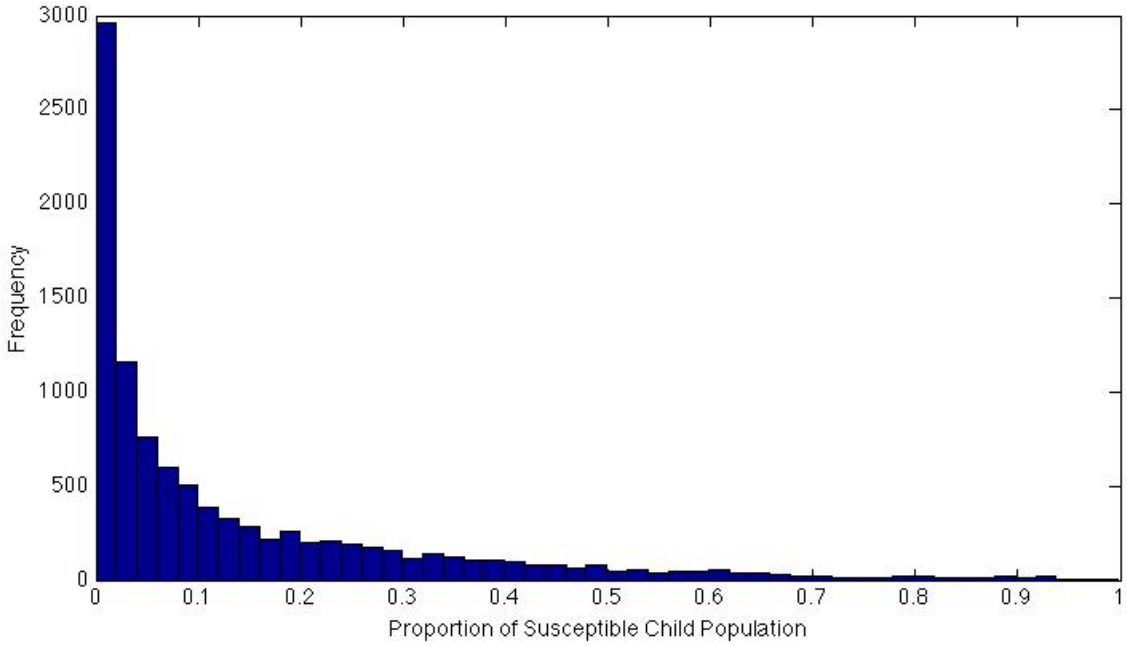


FIGURE 6. Histogram for the Proportion of the Infected Child Population with Moderate-to-Severe Anemia

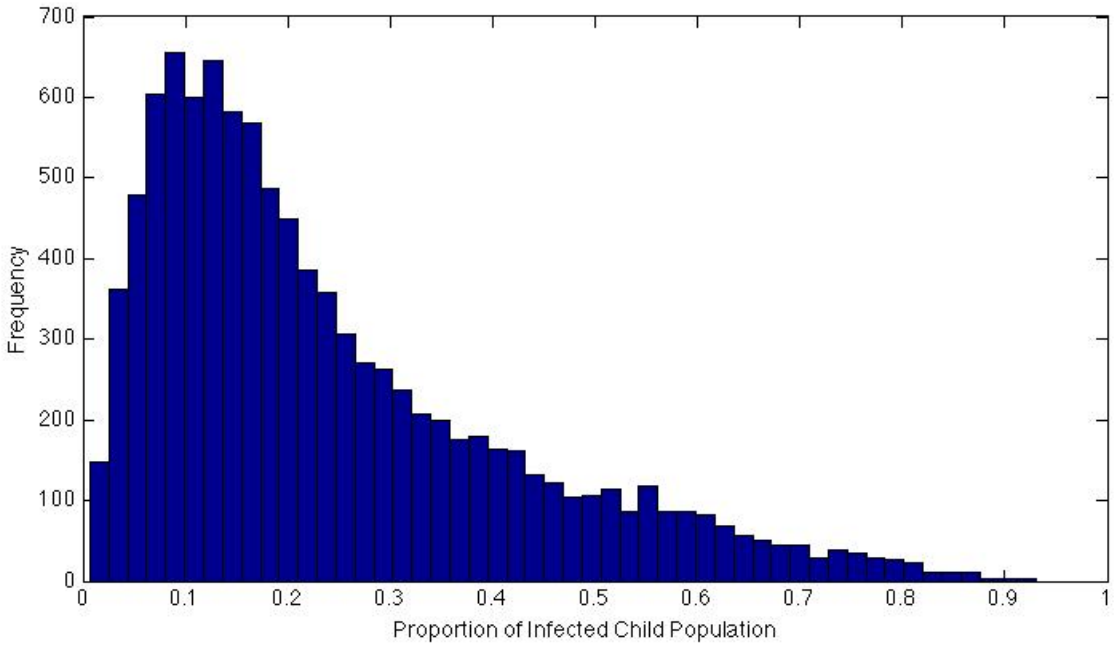

FIGURE 7. Full Epidemiological Model with Treatment and Vaccination Parameters (without risk stratification)

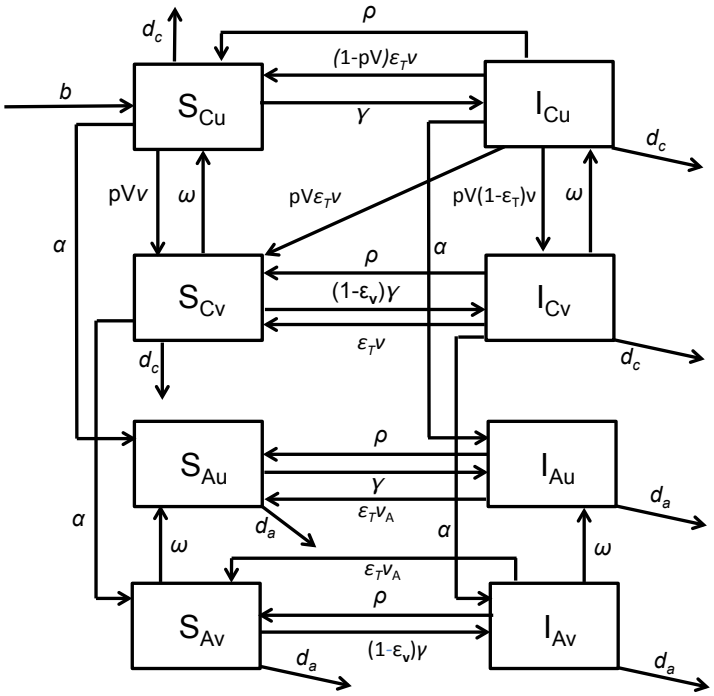

TABLE 1. Human Hookworm Vaccine Targets

| Target | Vaccine Mechanism | Stage in Vaccine Development |
|---|---|---|
| *Necator americanus Ancyclostoma*-secreted protein 2 (*Na*-ASP-2) | Attenuate L3 larval migration to the gastrointestinal tract; generate antibodies to decrease risk of hookworm disease due to heavy infections [37] | Phase I clinical trial suspended in 2007 due to safety concerns regarding reports of allergic reactions [38] |
| Glutathione S-transferase (*Na*-GST-1) | Interfere with the digestive proteolytic cascade that is responsible for heme detoxification [4] | Part 2 of a Phase I clinical trial currently underway among adults in Brazil [39] |
| *Necator americanus* aspartic protease (*Na*-APR-1) | Target macromolecules required by the mature parasite to metabolize erythrocytes; particularly, interfere with hemoglobin proteolysis [7] | Phase 1 study of safety and immunogenicity scheduled to begin in Washington, DC in June 2013 [40] |

TABLE 2: Pre-Treatment Prevalence of Hookworm Infection

| Compartment | Symbol | Measure | Prevalence Estimate | Reference |
|---|---|---|---|---|
| | | | | |
| Infected, Asymptomatic Children | $I_c$ | Prevalence of Hookworm Infection among Children (epg > 0) | 72.5% | [31] |
| Infected Adults | $I_a$ | Prevalence of Hookworm Infection among Adults (epg > 0) | 92.8% | [32] |

TABLE 3: Descriptions and Estimates for Parameterization of the Negative Binomial Distribution

| Model parameter definition | Symbol | Best Fit Point Estimate | Range |
|---|---|---|---|
| | | | |
| Mean fecal egg count for children (eggs per gram feces) | $\mu_C$ | 1303.3 | 1044.6-1691.0 |
| Aggregation parameter for children | $k_C$ | 0.8485 | 0.5196-1.5175 |
| Mean fecal egg count for adults (eggs per gram feces) | $\mu_A$ | 1621.6 | 657.7229-2536.7 |
| Aggregation Parameter (Adults) | $k_A$ | 0.8921 | 0.3054-2.1821 |
| Egg count cutoff between risk stratification categories for children (eggs per gram feces) | | 2123 | 1694-2706 |
| Egg count cutoff between risk stratification categories for adults (eggs per gram feces) | | 2633 | 1071-4151 |
| Ratio of relative infectiousness between children prone to high versus low intensity infection | $\theta_C$ | 4.8842 | 3.4039-7.1777 |
| Ratio of relative infectiousness between adults prone to high versus low intensity infection | $\theta_A$ | 4.7157 | 2.8324-12.2289 |

TABLE 4: Descriptions and Estimates of Transmission Model Parameters

| Model parameter definition | Symbol | Point Estimate (Value or Proportion)/ Point Estimate (Rate in months, if applicable) | Reference |
|---|---|---|---|
| **Demographic Parameters:** | | | |
| Population size | $N$ | 981,750 people | [29] |
| Total Number of Children (15 years and younger) | $N_c$ | 434,886 people | [29] |
| Total Number of Adults (15 years and older) | $N_a$ | 546,864 people | [29] |
| Birth rate | $b$ | 45.5 births per 1000 people annually / 0.00325 per month; $b_{80}$=0.8 $b$ and $b_{20}$=0.2 $b$ | [29] |
| Death rate (children under 15) | $d_c$ | 10.7 per 1000 population per year / 0.00140 per month | [29] |
| Death rate (adults 15 and over) | $d_a$ | 6.4 per 1000 population per year / 0.00127 per month | [29] |
| Aging rate | $\alpha$ | 6.1% of children per year / 0.00335 per month | [29] |
| **Disease Parameters:** | | | |
| Force of infection | $\gamma$ | Equation 1 | Present paper |
| Effective transmission rate in children (low intensity infection) | $\beta_{C80}$ | 0.0164 | Present paper |
| Effective transmission rate in adults (low intensity infection) | $\beta_{A80}$ | $\frac{1}{2} * \beta_{C80}$ = 0.0082 | Present paper; [25] |
| Effective transmission rate in children (high intensity infection) | $\beta_{C20}$ | $\theta_C * \beta_{C80}$ = 0.0801 | Present paper |
| Effective transmission rate in adults (high intensity infection) | $\beta_{A20}$ | $\theta_A * \beta_{A80}$ = 0.0387 | Present paper; [25] |
| Rate of recovery from infected state | $\rho$ | 0.8338 x $10^{-3}$ | Present paper |

TABLE 5: Time Series Prevalence Estimate Data for Hookworm Infection

| Source Population | Year | Prevalence Estimate | N | Reference |
|---|---|---|---|---|
| | | | | |
| Total Infected Children | 1998 | 73.3% | 182 | [41] |
| Total Infected Children | 2004 | 45.6% | 228 | [21] |
| Total Infected Adults | 2005 | 32.3% | 929 | [42] |
| Total Infected Children | 2007 | 18.3% | 367 | [43] |
| Total Infected Population | 2009 | 32.2% | 184 | [20] |

TABLE 6: Description of Parameters and Variables for Full Intervention Model

| Model parameter definition | Symbol | Estimate | Predicted Distribution | Estimation Method/ Reference |
|---|---|---|---|---|
| **Treatment Parameters:** | | | | |
| Chemotherapy treatment coverage | $v$ | Six coverage terms sampled through MCMC and based on different treatment campaigns | n/a | MCMC |
| Treatment efficacy rate | $\varepsilon_T$ | 15% (mebendazole)<br>72% (albendazole) | Beta | [17] |
| **Treatment Variables:** | | | | |
| Vaccine efficacy rate in terms of reducing force of infection | $\varepsilon_{V1}$ | Varied – 25%, 50%, 75%, 100% | Beta | MCMC; [5] |
| Vaccine efficacy rate in terms of reducing progression to clinical disease | $\varepsilon_{V2}$ | Varied – 25%, 50%, 75%, 100% | Beta | MCMC; [5] |
| Vaccine waning rate | $\omega$ | Varied – 1, 2, 3.5, 5 years | Exponential | MCMC; [44] |
| Vaccination coverage | $pV$ | Varied – 25%, 50%, 75%, 100% per month | n/a | MCMC |

## REFERENCES


1. Hotez, P., et al., *Incorporating a Rapid-Impact Package for Neglected Tropical Diseases with Programs for HIV/AIDS, Tuberculosis, and Malaria.* PLoS Medicine, 2006. **3**(5): p. e0102.
2. Bethony, J., et al., *Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm.* The Lancet, 2006. **367**(9521): p. 1521-1532.
3. de Silva, N., et al., *Soil-transmitted helminth infections: updating the global picture.* TRENDS in Parasitology, 2003. **19**(12): p. 547-551.
4. Diemert, D., J. Bethony, and P. Hotez, *Hookworm vaccines.* Clinical Infectious Diseases, 2008. **46**: p. 282-288.
5. Sabatelli, L., et al., *Modelling heterogeneity and the impact of chemotherapy and vaccination against human hookworm.* Journal of the Royal Society Interface, 2008. **5**: p. 1329-1341.
6. Hotez, P., et al., *Hookworm Infection.* New England Journal of Medicine, 2004. **351**(8): p. 799-807.
7. Pearson, M., et al., *Neutralizing Antibodies to the Hookworm Hemoglobinase Na-APR-1: Implications for a Multivalent Vaccine against Hookworm Infection and Schistosomiasis.* Journal of Infectious Diseases, 2010. **201**(10): p. 1561-1569.
8. Pearson, M., et al., *Molecular mechanisms of hookworm disease: Stealth, virulence, and vaccines.* Journal of Allergy and Clinical Immunology, 2012. **130**(1): p. 13-21.
9. Miller, T., *Hookworm infection in man*, in *Advances in Parasitology*. 1979. p. 315-384.
10. Stoltzfus, R., et al., *Hookworm control as a strategy to prevent iron deficiency.* Nutrition Reviews, 1997. **55**(6): p. 223-232.

11. Vercruysse, J., et al., *Assessment of the Anthelmintic Efficacy of Albendazole in School Children in Seven Countries Where Soil-Transmitted Helminths Are Endemic.* PloS Neglected Tropical Diseases, 2011. **5**(3): p. 1-10.
12. Hotez, P., *The development impact of the neglected tropical diseases (NTDs)*, 2011: New York.
13. Organization, W.H., *Elimination soil-transmitted helminthiases as a public health problem in children: Progress report 2001-2010 and strategic plan 2011-2020*, 2012.
14. Albonico, M., D. Engels, and L. Savioli, *Monitoring drug efficacy and early detection of drug resistance in human soil-transmitted nematodes: a pressing public health agenda for helminth control.* International Journal of Parasitology, 2004. **34**: p. 1205-1210.
15. Flohr, C., et al., *Low efficacy of mebendazole against hookworm in Vietnam: Two randomized controlled trials.* American Journal of Tropical Medicine and Hygiene, 2007. **76**: p. 732-736.
16. Anderson, R., *The population dynamics and epidemiology of intestinal nematode infections.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 1986. **80**: p. 686–696.
17. Keiser, J. and J. Utzinger, *Efficacy of Current Drugs Against Soil-Transmitted Helminth Infections: Systematic Review and Meta-analysis.* Journal of the American Medical Association, 2008. **299**(16): p. 1937-1948.
18. Stoltzfus, R., et al., *Effects of the Zanzibar school-based deworming program on iron status of children.* American Journal of Clinical Nutrition, 1998. **68**: p. 179-186.
19. Albonico, M., et al., *Efficacy of mebendazole and levamisole alone or in combination against intestinal nematode infections after repeated targeted mebendazole treatment in Zanzibar.* Bulletin of the World Health Organization, 2003. **81**(5): p. 343-352.

20. Knopp, S., et al., *From morbidity control to transmission control: time to change tactics against helminths on Unguja Island, Zanzibar.* Acta Tropica, 2011.
21. Massa, K., et al., *The combined effect of the Lymphatic Filariasis Elimination Programme and the Schistosomiasis and Soil-transmitted Helminthiasis Control Programme on soil-transmitted helminthiasis in schoolchildren in Tanzania.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 2009. **103**: p. 25-30.
22. Mohammed, K., et al., *Triple Co-Administration of Ivermectin, Albendazole and Praziquantel in Zanzibar: A Safety Study.* PloS Neglected Tropical Diseases, 2008. **2**(1): p. 1-7.
23. Anderson, R. and R. May, *Infectious Diseases of Humans*. 1991, Great Britain: Oxford Science Publications.
24. Medley, G., H. Guyatt, and D. Bundy, *A quantitative framework for evaluating the effect of community treatment on the morbidity due to ascariasis.* Parasitology, 1992. **106**: p. 211-221.
25. Chan, M., et al., *The evaluation of potential global morbidity attributable to intestinal nematode infections.* Parasitology, 1994. **109**: p. 373-387.
26. Basanez, M., et al., *A Research Agenda for Helminth Diseases of Humans: Modelling for Control and Elimination.* PLoS Neglected Tropical Diseases, 2012. **6**(4): p. 1-11.
27. Bundy, D., *Population ecology of intestinal helminth infections in human communities.* Philosophical Transactions of the Royal Society, 1988. **321**: p. 405-420.
28. Anderson, R., *Epidemiology*, in *Modern Parasitology*, F.E.G. Cox, Editor. 1982, Blackwell: Oxford. p. 204-251.
29. *Analytical Report, Volume X*, 2006, National Bureau of Statistics: Dar es Salaam.

30. Chan, M., M. Bradley, and D. Bundy, *Transmission patterns and the epidemiology of hookworm infection.* International Journal of Epidemiology, 1997. **26**(6): p. 1392-1400.
31. Stoltzfus, R., et al., *Epidemiology of iron deficiency anemia in Zanzibari schoolchildren: the importance of hookworms.* American Journal of Clinical Nutrition, 1997. **65**: p. 153-159.
32. Gillespie, S. and J. Johnston. *Expert consultation on anemia determinants and interventions: proceedings of a conference held September 16-17*. 1997. Ottawa, Canada.
33. Brooks, S. and A. Gelman, *General Methods for Monitoring Convergence of Interative Simulations.* Journal of Computational and Graphical Statistics, 1998. **7**(4): p. 434-455.
34. Gelman, A. and D. Rubin, *Inference from iterative simulation using multiple sequences.* Statistical Science, 1992. **7**: p. 457-511.
35. Horton, J., *Albendazole: a review of anthelmintic efficacy and safety in humans.* Parasitology, 2000. **121**: p. S113-S132.
36. Lwambo, N., D. Bundy, and G. Medley, *A new approach to morbidity risk assessment in hookworm endemic communities.* Epidemiology and Infection, 1992. **108**: p. 469-481.
37. Bethony, J., et al., *Antibodies against a secreted protein from hookworm larvae reduce the intensity of hookworm infection in humans and vaccinated laboratory animals.* The FASEB Journal, 2005. **19**(12): p. 1743-1745.
38. Diemert, D., et al., *Generalized urticaria induced by the Na-ASP-2 hookworm vaccine: implications for the development of vaccines against helminths.* Journal of Allergy and Clinical Immunology, 2012. **130**(1): p. 169-176.
39. Harvey, J., *Clinical Trials for First-Ever Human Hookworm Vaccine Advance: Phase 1 Clinical Trial in Brazil begins Testing at Brazilian Field Site*, in *The Sabin Vaccine Institute*2012: Washington, DC.

40. *Phase 1 Study of the Safety and Immunogenicity*, ClinicalTrials.gov, Editor 2013.

41. Montresor, A., et al., *Extending anthelminthic coverage to non-enrolled school-aged children using a simple and low-cost method.* Tropical Medicine and International Health, 2001. **6**(7): p. 535-537.

42. Young, S., et al., *Association of geophagia with Ascaris, Trichuris and hookworm transmission in Zanzibar, Tanzania.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 2007. **101**(8): p. 766-772.

43. Knopp, S., et al., *Changing Patterns of Soil-Transmitted Helminthiases in Zanzibar in the Context of National Helminth Control Programs.* American Journal of Tropical Medicine and Hygiene, 2009. **81**(6): p. 1071-1078.

44. Bethony, J., et al., *Vaccines against blood-feeding nematodes of humans and livestock.* Parasitology, 2006. **133**: p. S63-S79.