Distribution Agreement

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Steph Wraith Date

Approval Sheet

Impact of Schistosomiasis MDA on Antibody Responses among Pre-School Aged Children in Western Kenya

by
Steph Wraith
MSPH

Global Epidemiology

_____

Patrick Sullivan
Faculty Thesis Advisor

_____

Ryan Wiegand
Thesis Field Advisor

Abstract Cover Page

Impact of Schistosomiasis MDA on Antibody Responses among Pre-School Aged Children in Western Kenya

by
Steph Wraith

Bachelor of Science
College of William & Mary
2015

Faculty Thesis Advisor: Patrick Sullivan, PhD, DVM

An abstract of
A thesis submitted to the Faculty of the Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Global Epidemiology
2017

Abstract

Impact of Schistosomiasis MDA on Antibody Responses among Pre-School Aged Children in Western Kenya

by Steph Wraith

Background: Human schistosomiasis is an infectious disease caused by trematode parasites; in a given year at least 230 million people worldwide are infected, although some estimates range as high as 440 million. Pre-school aged children (PSACs; 1-5 years old) are not commonly screened for schistosomiasis. Compared to school-aged children (SACs), PSACs have been thought to be at reduced risk for schistosomiasis. There is mounting evidence, however, demonstrating that PSACs have a non-trivial burden of schistosomiasis.

Purpose: This paper utilized data from a community-randomized study to evaluate whether disparate mass drug administration strategies had a differential effect on schistosomiasis and soil-transmitted helminthic treatment programs. By modeling factors associated with the antigenic responses of the PSACs, we attempted to measure differences in transmission over the course of the study and determine if these differences are moderated by other covariates.

Methods: Thirty villages in the Mbita district bordering Lake Victoria in Western Kenya were enrolled in two separate treatment delivery programs (school-based treatment and community-wide treatment), and the enrolled children were followed over the course of a three-year period from 2012-2014 to conduct both traditional and serological tests for *S. mansoni* infection. Multiplex-bead assays of blood samples drawn over the course of the study period were used to generate serologic data by testing for specific antibody markers.

Results: Generalized linear regression models were constructed for two different antigenic indicators of *S. mansoni* infection. A range of indicators and interactions were assessed with proximity to the coast of Lake Victoria, year of entry into the study, age of the child, levels of tetanus antibody due to vaccination, and co-infections with malaria and strongyloides were all found to be significantly associated with schistosomiasis prevalence.

Conclusion: The overall aim of this study was to evaluate the impact of various mass drug administration strategies on schistosomiasis and soil-based helminthic disease burdens. This study established that factors including coastal proximity and parasitic co-infection are significantly associated with an increased disease burden among PSAC, and highlighted the value of adopting serology-based testing approaches as well as incorporating PSAC into future MDA interventions.

Cover Page

Impact of Schistosomiasis MDA on Antibody Responses among Pre-School Aged Children in Western Kenya

by

Steph Wraith

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of Master of Science in Public Health in
Global Epidemiology
2017

Acknowledgements

I want to express my sincere gratitude to Ryan Wiegand for his time, assistance, revisions, and incredible patience in guiding me through the process of developing and conducting this project – none of this would have been possible without his involvement. I'm also incredibly grateful to Dr. Kimberly Won for her insight and time in providing feedback and guidance on this project. I'd also like to offer my thanks to Dr. Patrick Sullivan for providing initial guidance in identifying resources for this project – including connecting me with the amazing team at DPDM – and for his feedback and advice throughout this process.

Additional acknowledgements go to

Centers for Disease Control and Prevention,
Division of Parasitic Diseases and Malaria,
Parasitic Branch

For allowing me to access and use their data for this thesis project.

As always, I'm so grateful to my friends and family, without whose support I would never have been able to conduct this research.

Table of Contents

Contents **Page**

Background:

What is schistosomiasis, why is it a problem, how big is the problem?

Human schistosomiasis is an infectious disease caused by trematode parasites; the trematodes can reside within the mesenteric veins or venous plexus of the bladder of their human host for many years [8]. Eggs from the parasites can be released into circulation and are shed through the urine or feces; eggs that fail to be excreted remain caught in various organs and can cause granulomata or fibrosis. Symptoms of schistosomiasis are dependent upon the number and location of the eggs caught in the body and the host immunologic response, as well as the species of schistosome. The particular species of interest for this study is *Schistosoma mansoni*, which is distributed throughout Africa, the Middle East, the Caribbean, and parts of South America [30]. Common signs and symptoms of *S. mansoni* infection include diarrhea, abdominal pain, and hepatosplenomegaly, although long duration or high intensity infections can also result in liver fibrosis. If early detection and treatment do not take place, severe morbidities can result from the development of intestinal or urogenital schistosomiasis [31].

The life cycle of the schistosomes involves an intermediate snail host found in fresh water. These snail hosts serve as the site for conversion from the miracidial larvae produced by adult female worms to the cercarial larvae that are infectious to humans. There are several *Biomphalaria* spp. snails native to Africa that are effective hosts for *S. mansoni* [22]. The form endemic to Lake Victoria is *Biomphalaria sudanica*. Once a snail is infected, it can produce two generations of sporocysts in a month that can ultimately generate hundreds to thousands of infectious cercariae over the course of a single year.

Human infection occurs when the cercariae exit the water by penetrating the skin, at which point they transform into schistosomulae and migrate throughout the body before settling in the blood supply of the liver, intestines, and bladder as they mature into fully adult worms [23]. The primary resting points for these schistosomulae are partially correlated with species, with *S. mansoni* tending to settle in the large intestinal tract and develop into intestinal schistosomiasis [32]. Once the worms have matured to adulthood and formed permanently embraced couples – with the females residing within the gynaecophoric channel of the males – the females can begin to produce the eggs that result in the morbidities associated with schistosomiasis [33]. For a more detailed visual representation of the infectious life cycle of *S. mansoni*, refer to Figure 1 [24].

The main treatment is praziquantel (PZQ) an effective, safe, single-dose treatment for schistosomiasis with few adverse events in uninfected individuals, which is the drug of choice against all species including *S. mansoni* [8, 21]. Given the relative absence of adverse side effects, PZQ delivered via a mass drug administration (MDA) intervention is popular in regions with endemic schistosomiasis [38]. Mild side effects associated with PZQ have been reported, but generally resolve within 24 hours of treatment initiation [43]. One of the other reasons for the popularity of PZQ for MDA is its relatively low cost as a result of donations from Merck KGaA, a German pharmaceutical company, beginning in 2007 and expanded in 2010 [38]. When effectively administered, PZQ can result in a cure rate of 60-90%, and a reduction in egg production of 80-95% among those who are not fully cured; one of the drawbacks of PZQ, which is a form of chemotherapy treatment, is that it cannot prevent reinfection [39].

Three species of schistosomes that are capable of infecting humans – *Schistosoma haematobium, S. japonicum,* and *S. mansoni* – are endemic throughout Africa, the Middle East, the Americas, and Asia. For a more detailed depiction of the distribution of schistosomiasis throughout Africa, refer to Figure 2. Much of the distribution of schistosomes is dependent upon the habitat range of the snails that function as the intermediate hosts in their life cycles, which is partially influenced by temperature. *Biomphalaria* are purely aquatic, and cannot survive outside the temperature range of 14-32 degrees Celsius [34].

In a given year at least 230 million people worldwide are infected with schistosomiasis, although some estimates range as high as 440 million [1]. In addition there are at least 800 million people estimated to be at risk of infection [6]. The majority of those both infected and at risk of infection are in sub-Saharan Africa. In Kenya, there are estimated to be approximately 6 million infected, with the highest burdens of disease falling on adolescents and populations concentrated around the shores of Lake Victoria [25]. In regions such as these where schistosomiasis is endemic, the initial infection of a child can occur very early in life, with the burden of disease rising until adolescence, the period with the highest prevalence of infection [7]. The WHO has been making calls since 2001 for countries with endemic schistosomiasis to implement control programs, and in 2012 the World Health Assembly (WHA) designated schistosomiasis elimination as a public health problem [38]. While 2020 has generally been regarded as the elimination target for most neglected tropical diseases (NTDs) the scientific consensus, given the varied distribution and transmission of schistosomiasis, is that 2030 is likely too ambitious a target date for total elimination of this and other STHs. Some countries have

experienced success with schistosomiasis control programs – Japan has succeeded in total elimination, and China, Egypt, and regions of South America have managed to significantly reduce the overall prevalence of disease. These programs have combined broad treatment programs and initiatives targeting socio-economic conditions in endemic regions; following this lead, by 2015 all African countries with endemic schistosomiasis had developed a plan for schistosomiasis control and begin to implement strategies targeting total disease elimination.

As Colley et al. note, "Schistosomiasis does not occur in isolation. It is a disease of poverty that often occurs where other parasites are prevalent and food insecurity is common". In particular, it is noted as being a major cause of global disability concentrated in sub-Saharan Africa [2]. The morbidities associated with schistosomiasis, both acute and chronic, can have a significant effect on the economic production, quality of life, and life expectancy in high prevalence regions. As noted by King et al., the clear linkages that exist between schistosomiasis and the development of long-term disabilities results in a reduction of the chances of combating rural poverty, and highlights the necessity of treating all infected individuals and working to develop new strategies for detection and intervention [7].

Why pre-school aged children?

In general, the prevalence and intensity of schistosomiasis caused by *Schistosoma mansoni* peak between 10-15 years of age and then gradually decline over time. Hence, most of the focus on schistosomiasis treatment efforts has been directed towards school-aged children [19]

Pre-school aged children (PSACs; 1-5 years old) are not commonly screened for schistosomiasis. Compared to school-aged children (SACs), PSACs have been thought to be at reduced risk for schistosomiasis, largely because the population has been understudied as a result of the heavy influence placed on the role of SAC transmission. There is mounting evidence, however, demonstrating that PSACs have a non-trivial burden of schistosomiasis [3]. Reported prevalence of infection in infants and PSACs across sub-Saharan Africa ranges from 14 to 86%, and a range of studies involving PZQ treatment regimens for these age groups have concluded that there is significant evidence to support the inclusion of PSAC in MDA interventions moving forward [10, 11, 12, 13, 14]. There has been some evidence indicating that while the treatment does result in satisfactory cure rates among PSAC, the cure rate is currently significantly lower for younger children included in MDA strategies, indicating that further research into how to optimize treatment among these younger age groups is critical [44].

As noted above, a child's first infection of schistosomiasis is frequently during the PSAC age range. Disease morbidity has primarily been documented in older teenagers and adults but is not yet well understood in young children. However preliminary studies have documented morbidities including fecal occult bleeding, anemia, and ultrasound abnormalities [3, 15, 16, 17, 18]. PSAC are still not generally included in screening or mass treatment programs, largely due to the limitations in diagnostic tools and the absence of a pediatric formulation for PZQ [19].

Why this project?

Currently, impact of public health programs for *Schistosoma mansoni* is largely measured with microscopic outcomes, especially the Kato-Katz stool examination

method [20]. The focus for most impact assessment programs has been placed on diagnostic tools that utilize stool and urine specimens for their testing. These measurement methods have known limitations, including the relatively short time window for specimen collection and examination, the necessity of trained microscopists to generate results, and the known decrease in sensitivity of the testing methods as prevalence decreases [35]. Other approaches for measuring impact therefore may present opportunities for greater facility and usefulness, particularly in assessing antibody presence in contexts with decreased overall prevalence of infection as diseases, such as schistosomiasis, move closer to elimination.

One such approach is the use of multiplex bead assays (MBAs) to analyze serologic outcomes. MBAs simultaneously detect antibodies against an array of antigens, allowing for a broad range of testing using a single serum sample and allowing for the analysis of the influence of co-infections that could potentially play a role in driving schistosomiasis prevalence [45]. This study, in addition to exploring the levels of *S. mansoni* infection among PSAC, also allows for an examination of the utility of MBA testing and serological outcome analysis in gaining an enhanced understanding of the risk factors for schistosomiasis in Western Kenya.

Methods:

Ethics

This study was assessed and approved by the Scientific Steering and Ethics Review Committees of the Kenya Medical Research Institute (KEMRI, SSC No. 2185) and by the Institutional Review Board of the U.S. Centers for Disease Control and Prevention (protocol #6249) through a reliance agreement with KEMRI [40]. The study was clearly explained to all potential participants using an established dialogue, and written informed consent was obtained from all persons who agreed to participate. Parents or guardians provided the consent for all children <18 years of age. In addition, all children between the ages of 7-17 years were asked to provide verbal participation assent. All identifiable information was kept confidential and maintained through the use of a secure database that had all access restricted to solely essential study personnel. This thesis was exempt from Emory IRB consideration as it was a secondary analysis of entirely de-identified data.

Structure of the study

The study was conducted from 2012-2014 in the Mbita district of Western Kenya, which borders Lake Victoria. Relatively high rates of both malaria and schistosomiasis (due to *S. mansoni* infection) have been documented in this region [3, 5]. One of the main commercial activities in Mbita district is fishing in Lake Victoria, which can lead to heightened levels of water contact. In addition, contact with the lake is involved in a wide range of other common daily activities including laundry and washing [4].

This study was part of a larger multi-country project designed to evaluate the impact of various integrated neglected tropical disease (NTD) control programs. All

schools within 5km of the coast of Lake Victoria underwent screening in order to identify communities with a prevalence of *S. mansoni* infection greater than 25%. From the villages that met these selection criteria, thirty were randomly selected and were subsequently randomized into two study arms. The two arms of the study, community-wide treatment (CWT) and school-based treatment (SBT), were based on the most common approaches used among treatment programs for school-aged children [21]. PSAC and their mothers or guardians were enrolled. In both study arms, parasitological and serologic indicators were monitored at baseline (year 1) and annually for two years following treatment. All monitoring was done in cross-sectional surveys in the selected villages. A total of 4,818 PSAC were enrolled in the study, with 3,612 providing serum samples.

The overall structure of data collection and diagnostic testing has already been described in a previous paper on this study [40]. In summary, all visits to villages for the study occurred between May and July of each study year. On the day of a study visit, all participants were asked to meet at a central location where all pertinent data was collected on Android-platform smartphones using the LINKS application [46]. Participants were randomly selected from a previously conducted population census. For each participant, stool and urine samples were collected and tested. Three hundred microliters (300μl) of whole blood was also collected annually via finger stick and used to conduct serology as well as tests for anemia and malaria. The following antigens were included in the MBA panel for this study: SEA (*S. mansoni* soluble egg antigen) and Sm25 (an internal glycoprotein found in adult *S. mansoni* worms) for schistosomiasis;

MSP-1 for malaria, NIE for *Strongyloides stercoralis*; Ascaris hemoglobin for *Ascaris lumbricoides*; VSP 3 and VSP 5 for *Giardia* spp.; tetanus toxoid and diphtheria toxoid.

For treatment, in the CWT arm all eligible individuals were offered annual treatment with PZQ (40mg/kg) and albendazole (400mg) while in the SBT the WHO-recommended protocol of only treating SAC was followed [9]. For both arms of the study, given the absence of any established protocols for PSAC inclusion in schistosomiasis control programs, treatment with PZQ (under the supervision of a physician) was only made available to children who had tested positive for *S. mansoni* using the Kato-Katz method of stool specimen examination.

Analyses conducted/new variables generated

All statistical tests conducted and models were developed in this study using SAS® 9.4 software (SAS Institute, Inc., Cary, NC), with the significance level set at 5%. Data from the study were collected and cleaned within Microsoft Excel prior to being read into SAS software. Frequencies, means, and proportions were compared using F tests and Chi-squared tests of significance to assess whether there were any associations between study arm or study year and the various exposures and co-infections of interest.

There were two main outcomes of interest: SEA and Sm25 median fluorescence intensity (MFI) results. To determine what constituted positive and negative results from the serology, cutoff values of 713.5 MFI – background (bg) units for SEA (sensitivity = 97.5%, specificity = 100%) and 52.5 MFI - bg units for Sm25 (sensitivity = 93.5%, specificity = 97.3%) were calculated. These values were developed at CDC using receiver operator characteristic curves pulling from the sera of 46 stool-positive *S.*

mansoni patients, 65 presumed-negative adult US citizens with no foreign travel history, and 45 presumed-negative US children.

Several new variables were developed and defined using study data: the serology antigen results for malaria, strongyloides, tetanus, and ascaris were all log-transformed, and a new variable defining proximity to the coast of Lake Victoria was defined, with a cut-off for coastal proximity of 1.2 km. This coastal variable was derived from an initial distance variable that was developed from mapping the GPS coordinates, projecting them to UTM zone 36S, and then calculating distance from the shore using ArcGIS version 10.3 (ESRI, Inc., Redlands, CA).

Univariable and multivariable models were fit using a generalized linear model with generalized estimating equations to account for multiple observations per community [36]. Univariable models include only the predictor of interest and are reported as unadjusted analyses. Multivariable modeling was used to explore the significant indicators of antigen variation as well as to check for significant interaction between variables. Differences across both the two treatment administration structures (SBT and CWT) and the three years of the study were examined using model selection strategies, bivariate analyses, as well as biological and historical considerations drawn from the literature. To determine and assess the variables for selection, a range of strategies were used including biological and historical considerations from the literature, as well as comparing AIC/BIC and examining the p-values generated for variables both in unadjusted and partially adjusted models [41,42]. Ultimately two multivariable models were developed, one for each outcome explored (SEA and Sm25). All results are reported with prevalence ratios, 95% confidence intervals, and p-values.

Results:

Descriptive statistics

During the three-year interval of study duration (2012-2014), a total of 4,611 PSAC were enrolled; of those enrolled, there were a total of 3,612 (78.3%) serum samples available to be tested via MBA. Mean age at enrollment was 3.01 years at year 1 of the study (SD = 1.20), decreasing slightly during year 2 (2.87 years, SD = 1.13) and year 3 (2.81 years, SD = 1.15). Mean age at enrollment did not differ significantly by study arm (Table 1). Of the participants in the SBT program, 52.7% were female; in the CWT program, 51.6% were female. Gender distribution did not alter significantly over the course of the study period. The majority of study participants resided in close proximity to the coast (1200m or less) – this difference was more pronounced in the SBT program (74.4%) than the CWT program (69.6%), with a p-value of <0.0001 from the chi-square test.

The percentage SEA positive differed significantly by study arm and across study years. In the SBT program, 53% of PSAC were SEA positive, while only 43.2% were positive in the CWT program (p-value <0.0001). Percentage SEA positive decreased over the course of the study, starting at 50% in year 1 and ending at 44.5% by year 3 (p-value 0.0004). The percentage Sm25 positive also varied significantly by study year (p-value 0.03), but did not differ between the two arms of the study.

Among the covariates describing other medical outcomes, only percentage positive for anemia varied significantly by study arm, with 41.1% positive in the SBT arm and 35.0% positive in the CWT arm. Across years of the study, however, the prevalence of ascaris, malaria, tetanus, anemia, and hematuria all varied significantly.

The overall antibody measures for ascaris increased from 230.55 in year 1 to 313.62 in year 3 (p-value 0.01); the antibody measures for malaria followed a similar pattern, increasing from 7236.87 to 13125.54 over the course of the study (p-value <0.0001), as did the measures for tetanus (year 1 5128.02, year 3 8171.39, p-value <0.0001). The prevalence of anemia increased from 33.85% in year 1 to 39.08% in year 3 (p-value = 0.002), but the prevalence of hematuria decreased over the course of the study from 33.76% in year 1 to 3.91% in year 3 (p-value <0.0001).

Results from regression analyses

For the SEA outcome, study year, sex of participant, coastal proximity, study arm, age, malaria, strongyloides, ascaris, tetanus, and interactions between age and coast as well as age and study year were all incorporated in the final model. For the Sm25 outcome, study year, sex of participant, coastal proximity, study arm, age, malaria, strongyloides, ascaris, tetanus, and interactions between age and coast as well as study arm and coast were all incorporated in the final model.

In the final adjusted model for SEA positive outcomes, neither sex nor study arm had a significant effect on the prevalence of infection, as study arm was randomly assigned and the distribution of sex was equivalent at baseline. There was a significant reduction in seropositivity as a result of MDA across years of the study (year 3 PR = 0.45, CI 0.26, 0.78, p-value = 0.004), demonstrating efficacy of the treatment regimen in decreasing schistosomiasis prevalence as two years of MDA showed an overall reduction of schistosomiasis prevalence. Proximity to the coast was also highly influential, with those residing within 1.2 km of the coastline having a prevalence of SEA positive schistosomiasis infection 4.3 times that of those who lived farther from the lake (PR =

4.32, CI 1.61, 11.59, p-value 0.004). Age of the PSAC also had a significant effect, with prevalence ratios almost doubling when comparing between a one-year difference in age (PR = 1.62, CI 0.75, 3.50, p-value 0.22) and a four-year difference in age (PR = 3.13, CI 1.52, 6.42, p-value 0.002). The log-transformed values of malarial and tetanus MFI values both had significant effects on SEA prevalence, with increases in malarial MFI associated with increased SEA prevalence (PR = 1.02, CI 1.01, 1.03, p-value 0.002) and increases in tetanus MFI, likely as a result of vaccination, having a slightly protective effect (PR = 0.98, CI 0.96, 0.99, p-value 0.007). This is a fairly narrow range of significance, but all calculations and log-transformations of variable values have been verified to confirm accuracy and validity of analysis.

In the final adjusted model for Sm25 positive outcomes, year, sex, and age had no statistically significant effects on prevalence. Proximity to coast once again had a highly significant impact (PR = 2.54, CI 1.16, 3.92, p-value 0.004). Study arm had a borderline significant effect on PSAC schistosomiasis prevalence, with the CWT arm having lower prevalence of Sm25 positive infections (PR = 0.67, CI 0.42, 1.07, p-value 0.09). Among the alternative helminthic infection outcomes for the Sm25 model, only strongyloides had a significant effect on prevalence in the fully adjusted model (PR = 1.38, CI 1.27, 1.50, p-value <0.0001). No significant associations were found in the fully adjusted model between ascaris and malaria MFI values, anemia, and Sm25 positive outcomes, although all three were significantly associated in bivariate analyses (Table 2).

Results from interaction terms

The significant interactions found and incorporated in the multivariable models were between year and age and between coast and age for the SEA model, and between

coast and age and between coast and study arm for the Sm25 model. The results from the interaction term between study year and age for SEA underlined the increased likelihood of infection as children increased in age, but also showed an increase in prevalence by study year, in conjunction with increased PSAC age, with a PR of 2.45 for a child aged 5 in year 1 of the study (CI 1.52, 3.95, p-value = 0.0001) and a PR of 5.22 for a child aged 5 in year 3 of the study (CI 3.24, 8.38, p-value <0.0001). This association was similar, though somewhat less pronounced, among younger children in the study. In the SEA model, proximity to the coast resulted in an increased prevalence of infection across all age groups, but the effect of proximity to coast was greater among the younger PSAC in the study, with a child 1 year of age living in the coastal region having a PR of 4.32 relative to a child the same age living farther from the coast (CI 1.61, 11.59, p-value = 0.004) and a similar association for a child of 2 (PR = 3.12, 1.93, 5.12, p-value <0.0001).

Interestingly, the same associations were not present in the Sm25 data, with coastal proximity still having the same effect of increasing prevalence, but no interaction between coast and age to show a gradient of impact with higher effect among younger children in the study. There was a significant interaction effect between study arm and coastal proximity in the Sm25 data, however, with study populations in near the coast having a significantly higher prevalence of schistosomiasis if they were in the CWT study arm than the SBT study arm (PR = 1.52, CI 1.04, 2.15, p-value = 0.03); in populations farther from the coast, the CWT had a non-significant protective effect relative to the SBT study arm.

Discussion:

Overall summary

In our two models, we explored how a range of factors including age, coastal proximity, study arm, and various co-infections affected the prevalence of two *S. mansoni*-associated MFI levels in populations of PSAC in western Kenya. We found that several of these factors had a significant impact on the risk of schistosomiasis infection in these populations, most notably coastal proximity and intercurrent malarial, strongyloides, and tetanus responses.

Points of interest

The most consistently significant indicator of schistosomiasis infection for both SEA and Sm25 MFI was coastal proximity among PSAC; the strength of association also interacted with age of the child. For the purposes of this study, coastal location is serving as a proxy variable for overall lake water contact among participants. This finding that water contact is a significant driver of infection and that as children age they have increased water contact and thus increased prevalence of schistosomiasis infection supports previous findings from other studies conducted in the area [4]; the clear presence of this association between water contact and infection even among PSAC provides support for the argument that these younger children should be included in treatment strategies; they are vulnerable to infection along the same routes of transmission as SAC, and have the same potential to further transmit disease, so proactive treatment action is warranted [11].

Interestingly, study year had a significant impact on the prevalence ratio for SEA positive infections, but did not have the same level of effect from Sm25 positive infections. One potential reason for the disparities in SEA and Sm25 MFI outcomes is that Sm25 is a recombinant antigen whereas SEA is not, so there are many more epitopes available for SEA. In addition, Sm25 is a microsomal antigen of adult worms. As a general rule, live worms do not shed a lot of this antigen and it is only when they are dying or dead that the worms begin to shed the antigen in greater abundance. Therefore, the lower values of Sm25 in this study could potentially be a factor of availability of Sm25 antigen in infected individuals.

For the modeled SEA outcomes, the third year of the study had a significantly lower prevalence ratio for schistosomiasis than either the first or second year of the study; this effect did not carry over to the Sm25 outcomes, however, where the prevalence ratios are virtually identical for all three years of the study. This effect was also somewhat mediated by the age of the study participants, with older children having an increasing prevalence of infection over the course of the three study years, as noted in the discussion of the interaction term results for each multivariable model.

Additionally, the strong association between anemia and schistosomiasis found in the bivariate regressions disappeared once incorporated into the full model. The prevalence and potential association between anemia and *S. mansoni* infections is particularly relevant for malnutrition and longer-term growth outcomes among PSAC and would benefit from further exploration [7]. It is important to note, however, that while many treatment and control programs tend to use anemia as an indicator for program impact, there are a wide range of confounders for anemia including malaria. Once anemia

and malaria are incorporated into the same full model the association between anemia and schistosomiasis vanishes, indicating that anemia may not be a useful indicator for the overall success of a schistosomiasis treatment program, particularly in areas where malaria is also endemic.

The apparent protective effect of tetanus response on SEA MFI, which was not seen in the Sm25 results, may be associated with an ongoing concern regarding the efficacy of the current tetanus vaccine among individuals with schistosomiasis; studies have found that schistosome infection at the time of immunization has the potential to negatively affect the induction of typical protective responses brought on by vaccination, or may result in a more rapid decrease in antibody response over time [26]. The association between malarial infection and SEA MFI level also supports previous findings in the literature that there may be an association between *Plasmodium falciparum* infection and schistosomiasis among children in sub-Saharan Africa, although the mechanism for this association is still not fully understood [27].

Limitations

One potential limitation of this study lies in the difficulty of applying serology to distinguish between present and past infections with *S. mansoni.* While this may be applicable particularly for older populations, in the context of very young children that are experiencing their first exposure to and infection with *S. mansoni*, using serological outcomes to track decreased prevalence can be a very powerful tool for assessing the impact of treatment programs. Another potential limitation in most studies utilizing serology lies in the challenge of determining appropriate cut-off points for defining a positive test result. The capacity to develop robust cut-offs is largely determined by the extent to which well-characterized sample panels are available for analysis to determine appropriate cut-offs, so if there isn't robust data available and the cut-offs aren't entirely accurate, the reported prevalence values may also be somewhat inaccurate.

An additional limitation is the structuring of water contact as an exposure within this model; here we have simply used proximity to the coastline as an indicator of contact with the water of Lake Victoria, but this is essentially a crude measure. It would be helpful to have more detailed information on the forms of water contact and the levels of risk that they pose for schistosomiasis infection, as this would help us to develop a more nuanced understanding of the key drivers of prevalence and how best to avoid repeated infections among PSAC and the community in general. However the use of coastal proximity as a proxy for water contact is a fairly common approach in studies of these populations, and is unlikely to invalidate the overall findings of this study [37].

Conclusions/Public Health Impact:

The overall goal of this study was to conduct an evaluation of the impact of various mass drug administration strategies on schistosomiasis and other soil-transmitted helminthic disease burden. The results of this thesis indicate a clear association between several factors including coastal proximity and co-infection with malaria and strongyloides on the prevalence of schistosomiasis among PSAC. This study also highlights the value of applying serological testing mechanisms such as MBA for analyzing a range of infections simultaneously. While there are limitations to serology applications, as noted above, there is also evidence that longitudinal monitoring of these serology results can provide information on fluctuation in antibody levels and potential levels of exposure, data that wouldn't be available from simply applying traditional parasitological testing methods. While further research is necessary, this study demonstrates the potential for using serological tools to evaluate program impact in regions with high endemicity and transmission.

This study demonstrates that there is indeed a significant burden of schistosomiasis among PSAC, and while there is still further work to be done to determine the most appropriate strategy for MDA and the adequate safe dosage of PZQ, there is a clear need to address the prevalence of infection among children in order to avoid disabilities and other health complications later on in life. Utilizing serological outcomes provides one route forward to rapidly assess the efficacy of current treatment programs and identify high-risk groups within the population.

References:

1. Colley, D. G., Bustinduy, A. L., Secor, W. E., & King, C. H., 2014. Human schistosomiasis. *The Lancet, 383*(9936), 2253-2264. doi:10.1016/s0140-6736(13)61949-2

2. Davis, S. M., Wiegand, R. E., Mulama, F., Kareko, E. I., Harris, R., Ochola, E., . . . Won, K. Y., 2015. Morbidity Associated with Schistosomiasis Before and After Treatment in Young Children in Rusinga Island, Western Kenya. *American Journal of Tropical Medicine and Hygiene, 92*(5), 952-958. doi:10.4269/ajtmh.14-0346

3. Odiere, M. R., Rawago, F. O., Ombok, M., Secor, W. E., Karanja, D. M., Mwinzi, P. N., . . . Won, K., 2012. High prevalence of schistosomiasis in Mbita and its adjacent islands of Lake Victoria, western Kenya. *Parasites & Vectors, 5*(1), 278. doi:10.1186/1756-3305-5-278

4. Mwinzi, P. N., Montgomery, S. P., Owaga, C. O., Mwanje, M., Muok, E. M., Ayisi, J. G., . . . Karanja, D. M., 2012. Integrated community-directed intervention for schistosomiasis and soil transmitted helminths in western Kenya – a pilot study. *Parasites & Vectors, 5*(1), 182. doi:10.1186/1756-3305-5-182

5. Minakawa N, Kongere JO, Dida GO, Ikeda E, Hu J, Minagawa K, Futami K, Kawada H, Njenga SM, Larson PS, 2015. Sleeping on the floor decreases insecticide treated bed net use and increases risk of malaria in children under 5 years of age in Mbita District, Kenya. *Parasitology 142*: 1516-22.

6. WHO, 2015. Schistosomiasis: number of people treated worldwide in 2013. *Wkly Epidemiol Rec 90*: 25-32.

7. King CH, Dangerfield-Cha M, 2008. The unacknowledged impact of chronic schistosomiasis. *Chronic Illn 4*: 65-79.

8. Heymann, D. L., 2015. *Control of communicable diseases manual* (20th ed.). Washington, DC: American Public Health Association.

9. WHO, 2011. Helminth Control in School-Age Children: A Guide for Managers of Control Programmes.

10. Ekpo UF, Oluwole AS, Abe EM, Etta HE, Olamiju F, Mafiana CF, 2012. Schistosomiasis in infants and pre-school-aged children in sub-Saharan Africa: implication for control. *Parasitology 139*: 835-41.

11. Sousa-Figueiredo JC, Pleasant J, Day M, Betson M, Rollinson D, Montresor A, Kazibwe F, Kabatereine NB, Stothard JR, 2010. Treatment of intestinal schistosomiasis in Ugandan preschool children: best diagnosis, treatment efficacy and side-effects, and an extended praziquantel dosing pole. *Int Health 2*: 103-13.

12. Stothard JR, Sousa-Figueiredo JC, Betson M, Adriko M, Arinaitwe M, Rowell C, Besiyge F, Kabatereine NB, 2011. Schistosoma mansoni Infections in young children: when are schistosome

antigens in urine, eggs in stool and antibodies to eggs first detectable? *PLoS Negl Trop Dis 5*: e938.

13. Stothard JR, Sousa-Figueiredo JC, Betson M, Bustinduy A, Reinhard-Rupp J, 2013. Schistosomiasis in African infants and preschool children: let them now be treated! *Trends Parasitol 29*: 197-205.

14. Verani JR, Abudho B, Montgomery SP, Mwinzi PN, Shane HL, Butler SE, Karanja DM, Secor WE, 2011. Schistosomiasis among young children in Usoma, Kenya. *Am J Trop Med Hyg 84*: 787-91.

15. Betson M, Sousa-Figueiredo JC, Kabatereine NB, Stothard JR, 2012. Use of fecal occult blood tests as epidemiologic indicators of morbidity associated with intestinal schistosomiasis during preventive chemotherapy in young children. *Am J Trop Med Hyg 87*: 694-700.

16. Betson M, Sousa-Figueiredo JC, Rowell C, Kabatereine NB, Stothard JR, 2010. Intestinal schistosomiasis in mothers and young children in Uganda: investigation of field-applicable markers of bowel morbidity. *Am J Trop Med Hyg 83*: 1048-55.

17. Green HK, Sousa-Figueiredo JC, Basanez MG, Betson M, Kabatereine NB, Fenwick A, Stothard JR, 2011. Anaemia in Ugandan preschool-aged children: the relative contribution of intestinal parasites and malaria. *Parasitology 138*: 1534-45.

18. Magalhaes RJ, Clements AC, 2011. Mapping the risk of anaemia in preschool-age children: the contribution of malnutrition, malaria, and helminth infections in West Africa. *PLoS Med 8*: e1000438.

19. Davis SM, Wiegand RE, Mulama F, Kareko EI, Harris R, Ochola E, Samuels AM, Rawago F, Mwinzi PM, Fox LM, Odiere MR, Won KY, 2015. Morbidity Associated with Schistosomiasis Before and After Treatment in Young Children in Rusinga Island, Western Kenya. *Am J Trop Med Hyg 92*: 952-58.

20. WHO, 1991. Basic laboratory methods in medical parasitology. Geneva: World Health Organization.

21. Onkanga, I. O., Mwinzi, P. N., Muchiri, G., Andiego, K., Omedo, M., Karanja, D. M., . . . Montgomery, S. P., 2016. Impact of two rounds of praziquantel mass drug administration on Schistosoma mansoni infection prevalence and intensity: a comparison between community wide treatment and school based treatment in western Kenya. *International Journal for Parasitology,46*(7), 439-445. doi:10.1016/j.ijpara.2016.01.006

22. Lu, L., Zhang, S., Mutuku, M. W., Mkoji, G. M., & Loker, E. S., 2016. Relative compatibility of Schistosoma mansoni with Biomphalaria sudanica and B. pfeifferi from Kenya as assessed by PCR amplification of the S. mansoni ND5 gene in conjunction with traditional methods. *Parasites & Vectors,9*(1), 166. doi:10.1186/s13071-016-1457-x

23. Carter Center, 2017. The Life Cycle of Schistosomiasis. https://www.cartercenter.org/resources/pdfs/health/schistosomiasis/schisto-disease-cycle.pdf

24. CDC, 2012. Parasites – Schistosomiasis. Biology. https://www.cdc.gov/parasites/schistosomiasis/biology.html

25. Standley C., Lwambo N., Lange C., Kariuki H., Adriko M., Stothard J., 2010. Performance of circulating cathodic antigen (CCA) urine-dipsticks for rapid detection of intestinal schistosomiasis in schoolchildren from shoreline communities of Lake Victoria. *Parasite Vector*. 3 (7).

26. Riner, D. K., Ndombi, E. M., Carter, J. M., Omondi, A., Kittur, N., Kavere, E., . . . Colley, D. G., 2016. Schistosoma mansoni Infection Can Jeopardize the Duration of Protective Levels of Antibody Responses to Immunizations against Hepatitis B and Tetanus Toxoid. *PLOS Neglected Tropical Diseases,10*(12). doi:10.1371/journal.pntd.0005180

27. Degarege, A., Degarege, D., Veledar, E., Erko, B., Nacher, M., Beck-Sague, C. M., & Madhivanan, P., 2016. Plasmodium falciparum Infection Status among Children with Schistosoma in Sub-Saharan Africa: A Systematic Review and Meta-analysis. *PLOS Neglected Tropical Diseases,10*(12). doi:10.1371/journal.pntd.0005193

28. Andrade, G., Bertsch, D. J., Gazzinelli, A., & King, C. H., 2017. Decline in infection-related morbidities following drug-mediated reductions in the intensity of Schistosoma infection: A systematic review and meta-analysis. *PLOS Neglected Tropical Diseases,11*(2). doi:10.1371/journal.pntd.0005372

29. Ng'Etich, A. I., Rawago, F. O., Jura, W. G., Mwinzi, P. N., Won, K. Y., & Odiere, M. R., 2016. A cross-sectional study on schistosomiasis and soil-transmitted helminths in Mbita district, western Kenya using different copromicroscopic techniques. *Parasites & Vectors,9*(1). doi:10.1186/s13071-016-1368-x

30. WHO, 2017. Schistosomiasis fact sheet - Parasite species and geographical distribution of schistosomiasis.

31. Kalenda, Y. D., Kato, K., Goto, Y., Fujii, Y., & Hamano, S. (2015). Tandem repeat recombinant proteins as potential antigens for the sero-diagnosis of Schistosoma mansoni infection. *Parasitology International,64*(6), 503-512. doi:10.1016/j.parint.2015.06.012

32. Gryseels, B., Polman, K., Clerinx, J., & Kestens, L. (2006). Human schistosomiasis. *The Lancet,368*(9541), 1106-1118. doi:10.1016/s0140-6736(06)69440-3

33. Gryseels, B. (2012). Schistosomiasis. *Infectious Disease Clinics of North America,26*(2), 383-397. doi:10.1016/j.idc.2012.03.004

34. Mccreesh, N., & Booth, M. (2014). The Effect of Increasing Water Temperatures on Schistosoma mansoni Transmission and Biomphalaria pfeifferi Population Dynamics: An Agent-Based Modelling Study. *PLoS ONE,9*(7). doi:10.1371/journal.pone.0101462

35. Nikolay B, Brooker SJ, Pullan RL. (2014). Sensitivity of diagnostic tests for human soil-transmitted helminth infections: a meta-analysis in the absence of a true gold standard. *Int J Parasitol 44*: 765-74.

36. Liang KY, Zeger SL (1986).Longitudinal data analysis using generalized linear models. *Biometrika*, 13-22.

37. Woodhall, D. M., Wiegand, R. E., Wellman, M., Matey, E., Abudho, B., Karanja, D. M., . . . Secor, W. E. (2013). Use of Geospatial Modeling to Predict Schistosoma mansoni Prevalence in Nyanza Province, Kenya. *PLoS ONE,8*(8). doi:10.1371/journal.pone.0071635

38. Fenwick, A., & Jourdan, P. (2016). Schistosomiasis elimination by 2020 or 2030? *International Journal for Parasitology,46*(7), 385-388. doi:10.1016/j.ijpara.2016.01.004

39. Cioli, D., Pica-Mattoccia, L., Basso, A., & Guidi, A. (2014). Schistosomiasis control: praziquantel forever? *Molecular and Biochemical Parasitology,195*(1), 23-29. doi:10.1016/j.molbiopara.2014.06.002

40. Won, K. Y., Priest, J. W., Mwende, F. M., Lee, Y., Wiegand, R. E., Secor, W. E., . . . Odiere, M. R. (2017). Multiplex Serologic Assessment of Schistosomiasis in Western Kenya: Antibody Responses in Preschool Age Children as a Measure of Reduced Transmission. *The American Journal of Tropical Medicine and Hygiene*. doi:10.4269/ajtmh.16-0665

41. Kuha, J. (2004). AIC and BIC. *Sociological Methods & Research,33*(2), 188-229. doi:10.1177/0049124103262065

42. Bursac, Z., Gauss, C. H., Williams, D. K., & Hosmer, D. W. (2008). Purposeful selection of variables in logistic regression. *Source Code Biol Med.,3*(17). doi:10.1186/1751-0473-3-17

43. Reta, B., & Erko, B. (2013). Efficacy and side effects of praziquantel in the treatment forSchistosoma mansoniinfection in school children in Senbete Town, northeastern Ethiopia. *Tropical Medicine & International Health,18*(11), 1338-1343. doi:10.1111/tmi.12187

44. Sousa-Figueiredo, J. C., Betson, M., Atuhaire, A., Arinaitwe, M., Navaratnam, A. M., Kabatereine, N. B., . . . Stothard, J. R. (2012). Performance and Safety of Praziquantel for Treatment of Intestinal Schistosomiasis in Infants and Preschool Children. *PLoS Neglected Tropical Diseases,6*(10). doi:10.1371/journal.pntd.0001864

45. Moss DM, Priest JW, Boyd A, Weinkopff T, Kucerova Z, Beach MJ, Lammie PJ, (2011). Multiplex bead assay for serum samples from children in Haiti enrolled in a drug study for the treatment of lymphatic filariasis. *Am J Trop Med Hyg 85*: 229-37.

46. Pavluck A, Chu B, Mann Flueckiger R, Ottesen E, (2014). Electronic data capture tools for global health programs: evolution of LINKS, an Android-, web-based system. *PLoS Negl Trop Dis 8*: e2654.

Tables and Figures:

Table 1: Demographics & Outcomes by Study Arm

Characteristic	SBT (n=1805)*	CWT (n=1807)*	p-value**
Demographics			
Age (yr)	2.87 (1.16)	2.91 (1.16)	0.3389
Sex (% female)	953 (52.74)	931 (51.58)	0.5053
Zone***			
A (% le 1200 m)	1345 (74.43)	1257 (69.64)	<0.0001
B (% gt 1200 m)	462 (25.57)	489 (27.09)	
Outcomes			
SEA (MFI value)	13323.72 (13654.8)	10849.8 (13307.2)	<0.0001
Sm 25 (MFI value)	101.09 (12.1)	160.52 (24.13)	0.0277
SEA (% positive)	957 (52.96)	780 (43.21)	<0.0001
Sm 25 (% positive)	324 (17.93)	309 (17.12)	0.5214
Alternative outcomes			
Ascaris (MFI value)	254.07 (21.77)	255.63 (22.66)	0.9605
Malaria (MFI value)	10691.8 (12831.6)	11220.5 (13069)	0.2200
Strongyloides (MFI value)	403.4 (1899.3)3	342.63 (1568.6)	0.2943
Tetanus (MFI value)	6554.88 (8671)	6889.39 (8854.9)	0.2514
Anemia (% positive)	737 (41.1)	629 (35.02)	0.0002
Hematuria (% positive)	214 (12.78)	197 (11.78)	0.3995
Hookworm (% positive)	14 (0.82)	13 (0.77)	0.8650

* either mean(std dev) or N(%)

** via pooled t-test (means), chi-square (proportions), and f-tests (multi-level means)

***If doesn't add to zero, remaining percentage in C (gt 3800)

Table 2: Demographics & Outcomes by Year of Study

Characteristic	year 1 (n=1103)*	year 2 (n=1174)*	year 3 (n=1335)*	p-value**
Demographics				
Age (yr)	3.01 (1.20)	2.87 (1.13)	2.81 (1.15)	<.0001
Sex (% female)	570 (51.68)	627 (53.41)	687 (51.46)	0.5780
Zone***				
A (% le 1200 m)	785 (71.17)	853 (72.66)	964 (72.21)	0.4033
B (% gt 1200 m)	305 (27.65)	297 (25.3)	349 (26.14)	
Outcomes				
SEA (MFI value)	12291.96 (13411.16)	12751.33 (13642.33)	11334.65 (13521.1)	0.0272
Sm 25 (MFI value)	105.53 (624.75)	123.94 (638.99)	157.68 (1048.31)	0.2699
SEA (% positive)	551 (49.95)	592 (50.43)	594 (44.49)	0.0041
Sm 25 (% positive)	155 (14.05)	233 (19.85)	245 (18.35)	0.0008
Alternative outcomes				
Ascaris (MFI value)	230.55 (1004.27)	210.85 (542.31)	313.62 (1146.99)	0.0146
Malaria (MFI value)	7236.87 (11007.02)	11983.16 (13217.16)	13125.54 (13538.87)	<0.0001
Strongyloides (MFI value)	326.76 (1737.11)	378.36 (1705.16)	406.61 (1778.02)	0.5260
Tetanus (MFI value)	5128.02 (7922.1)	6571.55 (8952.1)	8171.39 (9322.98)	<0.0001
Anemia (% positive)	371 (33.85)	478 (40.85)	517 (39.08)	0.0017
Hematuria (% positive)	346 (33.76)	17 (1.55)	48 (3.91)	<0.0001
Hookworm (% positive)	8 (0.76)	13 (1.12)	6 (0.52)	0.2554

* either mean(std dev) or N(%)

** via pooled t-test (means), chi-square (proportions), and f-tests (multi-level means)

***If doesn't add to zero, remaining percentage in C (gt 3800)

Table 3: Regression Results - SEA

Variable		Unadjusted PR	CI	p-value	Adjusted PR	CI	p-value
Year	1	ref cat					
	2	0.9591	(0.8567, 1.0738)	0.4685	0.8552	(0.5062, 1.4450)	0.5590
	3	0.8607	(0.7605, 0.9740)	0.0175	0.4454	(0.2559, 0.7751)	0.0042
Sex	1	ref cat					
	2	0.9495	(0.8841, 1.0108)	0.0999	0.9522	(0.8946, 1.0135)	0.1239
Coast	0	ref cat					
	1	3.1266	(1.9541, 5.0027)	<0.0001	4.3193	(1.6099, 11.5883)	0.0037
Study arm	SBT	ref cat					
	CWT	0.8122	(0.5347, 1.2335)	0.3292	0.8592	(0.6759, 1.0922)	0.2152
Age	1	ref cat					
	2	1.6332	(1.3307, 2.0052)	<0.0001	1.6214	(0.7520, 3.4959)	0.2176
	3	2.3368	(1.9274, 2.8332)	<0.0001	2.7036	(1.5639, 4.6735)	0.0004
	4	2.5902	(2.1519, 3.1177)	<0.0001	2.6578	(1.2923, 5.4668)	0.0079
	5	2.7330	(2.1076, 3.5440)	<0.0001	3.1255	(1.5220, 6.4192)	0.0019
Log msp		1.0479	(1.0191, 1.0464)	0.0015	1.0211	(1.0079, 1.0345)	0.0016
Log nie		1.0327	(0.9967, 1.0511)	<0.0001	1.0202	(0.9918, 1.0494)	0.1653
Log tetanus		0.9320	(0.9161, 0.9482)	<0.0001	0.9789	(0.9640, 0.9942)	0.0069
Anemia		0.8978	(0.8196, 0.9834)	0.0203	1.0057	(0.9192, 1.1004)	0.9014

Table 4: Regression Results - Sm25

Variable		Unadjusted PR	CI	p-value	Adjusted PR	CI	p-value
Year	1	ref cat					
	2	1.3070	(0.9797, 1.7437)	0.0687	1.1958	(0.9202, 1.5539)	0.1810
	3	1.2637	(0.9847, 1.6216)	0.0659	1.1168	(0.8876, 1.4052)	0.3457
Sex	1	ref cat					
	2	1.0273	(0.8556, 1.2334)	0.7731	1.0367	(0.9001, 1.1938)	0.6174
Coast	0	ref cat					
	1	2.1561	(1.5098, 3.0790)	<0.0001	2.5439	(1.1621, 3.9257)	0.0039
Study arm	SBT	ref cat					
	CWT	0.9348	(0.5896, 1.4820)	0.7742	0.6706	(0.4221, 1.0655)	0.0907
Age	1	ref cat					
	2	1.2423	(0.8760, 1.7617)	0.2235	1.0421	(0.7687, 1.4126)	0.7906
	3	2.0307	(1.4775, 2.7910)	<0.0001	0.9685	(0.6474, 1.4490)	0.8764
	4	2.8383	(2.1044, 3.8281)	<0.0001	1.1849	(0.8439, 1.6638)	0.3270
	5	2.6696	(1.7453, 4.0836)	<0.0001	1.1142	(0.5768, 2.1522)	0.7475
Log asc		1.1305	(1.0806, 1.1826)	<0.0001	0.9973	(0.9550, 1.0414)	0.9016
Log msp		1.0590	(1.0288, 1.0900)	0.0001	1.0082	(0.9763, 1.0412)	0.6148
Log nie		1.4151	(1.3075, 1.5316)	<0.0001	1.3829	(1.2735, 1.5018)	<0.0001
Anemia		0.8690	(0.7629, 0.9879)	0.0346	1.0205	(0.8490, 1.1311)	0.7817

Table 5: Interaction term results - SEA

Interaction term		Adjusted PR	CI	p-value
Year*age				
Year 1	Age			
	1	ref		
	2	1.3831	(-0.8177, 2.3077)	0.2143
	3	2.2023	(1.4785, 3.2804)	0.0001
	4	2.1845	(1.3563, 3.5186)	0.0013
	5	2.4508	(1.5215, 3.9475)	0.0001
Year 2				
	1	ref		
	2	1.7286	(1.1724, 2.5487)	0.0057
	3	2.2111	(1.5778, 3.0072)	<0.0001
	4	2.5439	(1.7206, 3.7611)	<0.0001
	5	2.7486	(1.7149, 4.4056)	<0.0001
Year 3				
	1	ref		
	2	2.7674	(1.8002, 4.2543)	<0.0001
	3	4.1795	(2.8648, 6.0976)	<0.0001
	4	4.9744	(3.2485, 7.6172)	<0.0001
	5	5.2164	(3.2410, 8.3788)	<0.0001
Coast*age				
Coast = 1	Age			
	1	4.3193	(1.6100, 11.5879)	0.0037
	2	3.1428	(1.9337, 5.1169)	<0.0001
	3	2.866	(1.4643, 5.6091)	<0.0001
	4	2.9177	(1.7924, 4.7496)	<0.0001
	5	2.6557	(1.3596, 5.1874)	0.0042

Table 6: Interaction term results - Sm25

Interaction term		Adjusted PR	CI	p-value
Coast*age				
Coast = 1	Age			
	1	0.5873	(0.3133, 1.1006)	0.0968
	2	0.7760	(0.4890, 1.2314)	0.2817
	3	1.5987	(1.1174, 2.2878)	0.0102
	4	1.8664	(1.3796, 2.5250)	<0.0001
	5	2.0222	(1.0098, 4.0496)	0.0469
Arm*coast				
Study Arm (CWT/SBT)	Coast			
	0	0.6706	(0.4221, 1.0654)	0.0907
	1	1.5224	(1.0420, 2.1495)	0.029

Figure 1 – Schistosomiasis Life Cycle [24]

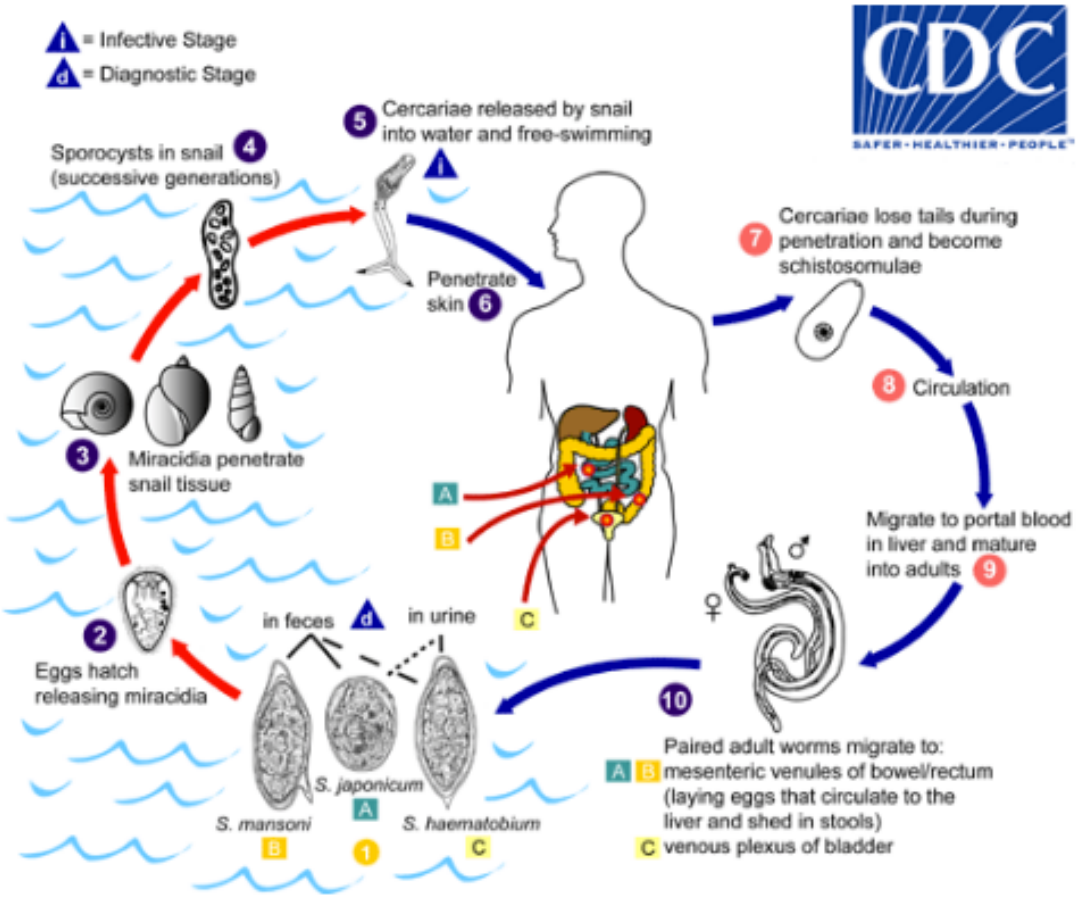


Figure 2 – Global Distribution of Schistosomiasis [32]

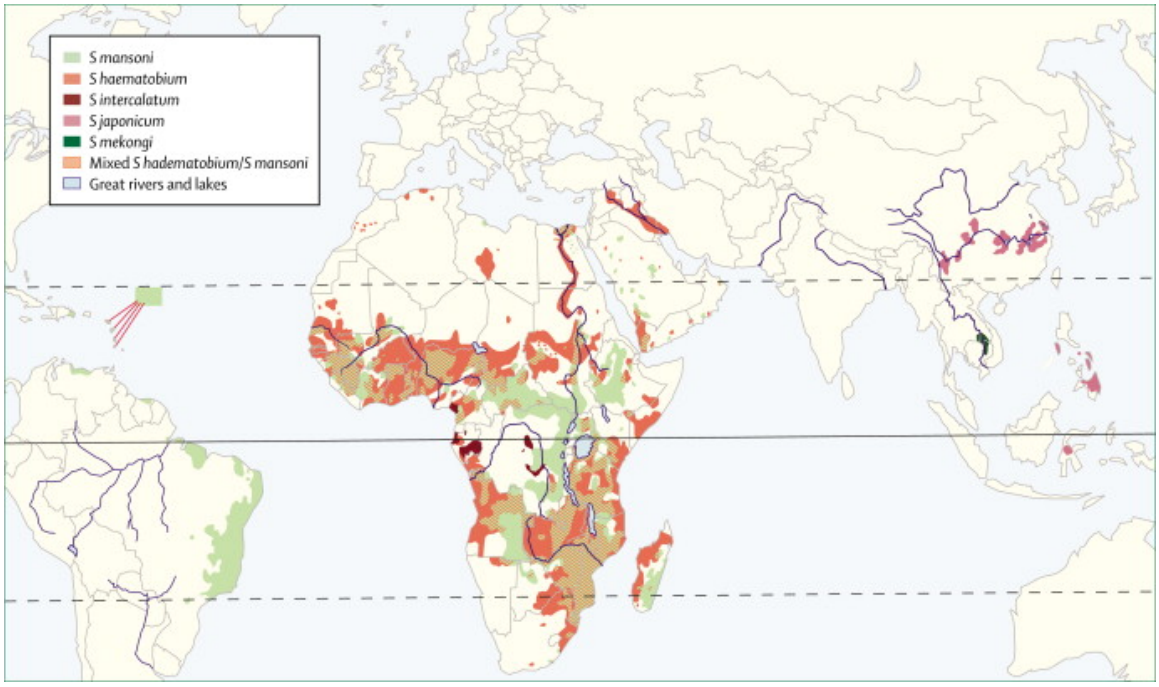
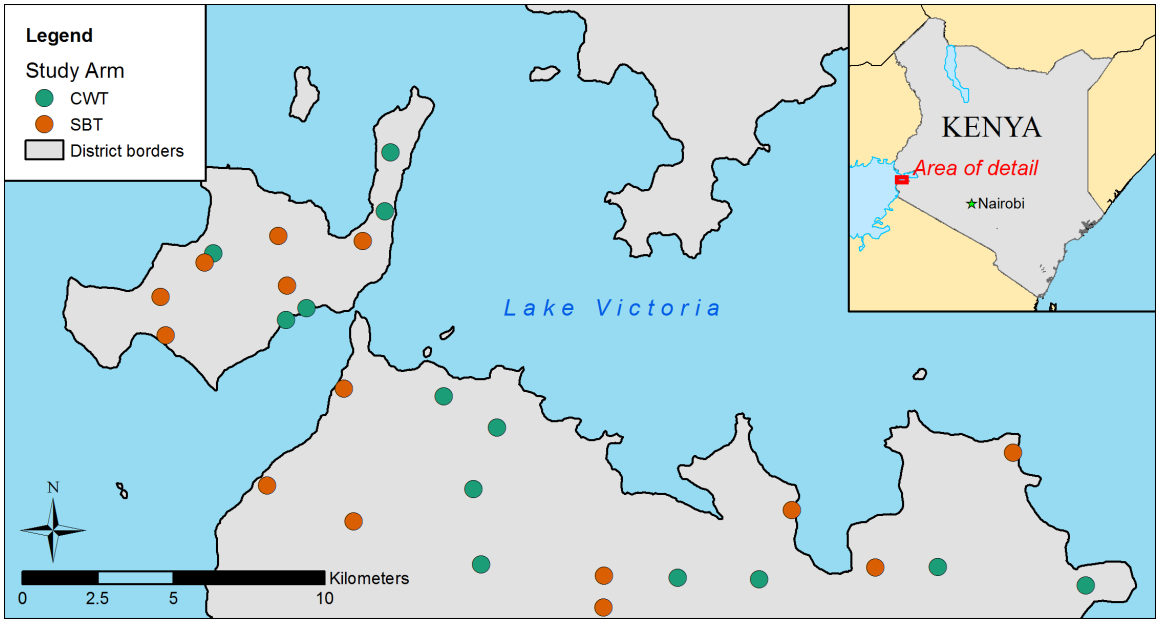


Figure 3 – Area Map



Appendix 1 – SAS Code:

```
*****************************************************************;
* Program: h:\thesis\programs\Wraith_thesis.sas                  *;
* Date: 4/10/17                                                  *;
* Programmer: Steph Wraith                                       *;
*                                                                *;
* Purpose: This program carries out data analysis                *;
* and modeling for an MSPH thesis on schistosomiasis in Kenya    *;
*****************************************************************;

*Importing data for analysis;
libname schisto "h:\thesis\data";

PROC IMPORT OUT= SCHISTO.MBITA
            DATAFILE= "H:\Thesis\Data\Mbita.csv"
            DBMS=CSV REPLACE;
     GETNAMES=YES;
     DATAROW=2;
RUN;

*Creating a working dataset to perform all analyses/calculations in;
data schisto;
       set SCHISTO.MBITA;

       if year=2 then year2=1;
              else year2=0;
       if year=3 then year3=1;
              else year3=0;

       if age=2 then age2=1;
              else age2=0;
       if age=3 then age3=1;
              else age3=0;
       if age=4 then age4=1;
              else age4=0;
       if age=5 then age5=1;
              else age5=0;

       if dist_vic <= 1200 then zone="A";
              else if dist_vic <= 3800 then zone="B";
              else zone="C";
       if dist_vic <= 1200 then coast=1;
              else coast=0;

       if arm = "SBT" then bi_arm = 0;
              else if arm = "CWT" then bi_arm = 1;

       if sea_2 ge 985 then SEA_new = 1;
              else if sea_2 lt 985 then SEA_new = 0;

       if sm_25_4 ge 38 then Sm25_new = 1;
              else if sm_25_4 lt 38 then Sm25_new = 0;

       log_asc = log(ascaris_hb_93);
       log_msp = log(msp_1_23);
       log_nie = log(nie_70);
       log_tetanus = log(tetanus_11);

run;

*Formats;
proc format;
       value coastf
              0 = "z - no"
              1 = "yes";
       run;
```

```
proc format;
        value yearf
                1 = "z - year one"
                2 = "year two"
                3 = "year three";
        run;

proc format;
        value sexf
                1 = "male"
                2 = "female";
        run;

proc format;
        value bi armf
                0 = "SBT"
                1 = "CWT";
        run;

proc format;
        value agef
                1 = "z - 1"
                2 = "2"
                3 = "3"
                4 = "4"
                5 = "5";
        run;

*Exploratory data analysis;
proc contents data=work.schisto;
run;

proc means data=work.schisto;
        var year;
run;

*Initial proq freq run to assess variables;
proc freq data=work.schisto;
run;

*Narrowed proc freq;
proc freq data=work.schisto;
        tables age anemia hematuria hw sex sm village_id year sea_mm sea_lab / list
missing;
run;

proc freq data = work.schisto;
        tables arm;
run;

proc freq data = work.schisto;
        tables year;
run;

*TABLE 1 CALCULATIONS

Variable by variable univariate analysis
Demographics;

*Age;
proc univariate data=work.schisto;
        var age;
        where arm = "SBT";
run;

proc univariate data=work.schisto;
        var age;
        where arm = "CWT";
run;

proc ttest data=work.schisto;
```

```
        class arm;
        var age;
run;

proc univariate data=work.schisto;
        var age;
        where year = 1;
run;

proc univariate data=work.schisto;
        var age;
        where year2 = 1;
run;

proc univariate data=work.schisto;
        var age;
        where year3 = 1;
run;

proc glm data=work.schisto;
        class year;
        model age = year;
        means age;
run;

*Sex;

proc freq data=work.schisto;
        tables sex;
        where arm = "SBT";
run;

proc freq data=work.schisto;
        tables sex;
        where arm = "CWT";
run;

proc freq data=work.schisto;
        tables sex*arm / chisq;
run;


proc freq data=work.schisto;
        tables sex;
        where year = 1;
run;

proc freq data=work.schisto;
        tables sex;
        where year2 = 1;
run;

proc freq data=work.schisto;
        tables sex;
        where year3 = 1;
run;

proc freq data=work.schisto;
        tables sex*year / chisq;
run;

*Zone;
proc freq data=work.schisto;
        tables zone;
        where arm = "SBT";
run;

proc freq data=work.schisto;
        tables zone;
        where arm = "CWT";
run;
```

```
proc freq data=work.schisto;
        tables zone*arm / chisq;
run;

proc freq data=work.schisto;
        tables zone;
        where year = 1;
run;

proc freq data=work.schisto;
        tables zone;
        where year2 = 1;
run;

proc freq data=work.schisto;
        tables zone;
        where year3 = 1;
run;

proc freq data=work.schisto;
        tables zone*year / chisq;
run;

*Outcomes;

*SEA;
proc univariate data=work.schisto;
        var sea_2;
        where arm = "SBT";
run;

proc univariate data=work.schisto;
        var sea_2;
        where arm = "CWT";
run;

proc ttest data=work.schisto;
        class arm;
        var sea_2;
run;

proc univariate data=work.schisto;
        var sea 2;
        where year = 1;
run;

proc univariate data=work.schisto;
        var sea 2;
        where year2 = 1;
run;

proc univariate data=work.schisto;
        var sea_2;
        where year3 = 1;
run;

proc glm data=work.schisto;
        class year;
        model sea_2 = year;
        means sea_2;
run;

*SEA_new;
proc freq data=work.schisto;
        tables SEA_new;
        where arm = "SBT";
run;

proc freq data=work.schisto;
        tables SEA_new;
```

```
        where arm = "CWT";
run;

proc freq data=work.schisto;
        tables SEA_new*arm / chisq;
run;

proc freq data=work.schisto;
        tables SEA_new;
        where year = 1;
run;

proc freq data=work.schisto;
        tables SEA_new;
        where year2 = 1;
run;

proc freq data=work.schisto;
        tables SEA_new;
        where year3 = 1;
run;

proc freq data=work.schisto;
        tables SEA_new*year / chisq;
run;

*Sm 25;
proc univariate data=work.schisto;
        var sm_25_4;
        where arm = "SBT";
run;

proc univariate data=work.schisto;
        var sm_25_4;
        where arm = "CWT";
run;

proc ttest data=work.schisto;
        class arm;
        var sm_25_4;
run;

proc univariate data=work.schisto;
        var sm 25 4;
        where year = 1;
run;

proc univariate data=work.schisto;
        var sm 25 4;
        where year2 = 1;
run;

proc univariate data=work.schisto;
        var sm_25_4;
        where year3 = 1;
run;

proc glm data=work.schisto;
        class year;
        model sm_25_4 = year;
        means sm_25_4;
run;

*Sm25_new;
proc freq data=work.schisto;
        tables Sm25_new;
        where arm = "SBT";
run;

proc freq data=work.schisto;
        tables Sm25_new;
```

```
        where arm = "CWT";
run;

proc freq data=work.schisto;
        tables Sm25_new*arm / chisq;
run;

proc freq data=work.schisto;
        tables Sm25_new;
        where year = 1;
run;

proc freq data=work.schisto;
        tables Sm25_new;
        where year2 = 1;
run;

proc freq data=work.schisto;
        tables Sm25_new;
        where year3 = 1;
run;

proc freq data=work.schisto;
        tables Sm25_new*year / chisq;
run;

*Alternative outcomes;
*Other antibody multiplex results;

*Ascaris;
proc univariate data=work.schisto;
        var ascaris hb 93;
        where arm = "SBT";
run;

proc univariate data=work.schisto;
        var ascaris_hb_93;
        where arm = "CWT";
run;

proc ttest data=work.schisto;
        class arm;
        var ascaris_hb_93;
run;

proc univariate data=work.schisto;
        var ascaris_hb_93;
        where year = 1;
run;

proc univariate data=work.schisto;
        var ascaris_hb_93;
        where year2 = 1;
run;

proc univariate data=work.schisto;
        var ascaris hb 93;
        where year3 = 1;
run;

proc glm data=work.schisto;
        class year;
        model ascaris_hb_93 = year;
        means ascaris_hb_93;
run;

*Malaria;
proc univariate data=work.schisto;
        var msp 1 23;
        where arm = "SBT";
run;
```

```
proc univariate data=work.schisto;
        var msp 1 23;
        where arm = "CWT";
run;

proc ttest data=work.schisto;
        class arm;
        var msp_1_23;
run;

proc univariate data=work.schisto;
        var msp_1_23;
        where year = 1;
run;

proc univariate data=work.schisto;
        var msp_1_23;
        where year2 = 1;
run;

proc univariate data=work.schisto;
        var msp_1_23;
        where year3 = 1;
run;

proc glm data=work.schisto;
        class year;
        model msp_1_23 = year;
        means msp_1_23;
run;

*Strongyloides;
proc univariate data=work.schisto;
        var nie_70;
        where arm = "SBT";
run;

proc univariate data=work.schisto;
        var nie 70;
        where arm = "CWT";
run;

proc ttest data=work.schisto;
        class arm;
        var nie_70;
run;

proc univariate data=work.schisto;
        var nie_70;
        where year = 1;
run;

proc univariate data=work.schisto;
        var nie_70;
        where year2 = 1;
run;

proc univariate data=work.schisto;
        var nie_70;
        where year3 = 1;
run;

proc glm data=work.schisto;
        class year;
        model nie_70 = year;
        means nie_70;
run;

*Tetanus;
proc univariate data=work.schisto;
```

```
        var tetanus_11;
        where arm = "SBT";
run;

proc univariate data=work.schisto;
        var tetanus_11;
        where arm = "CWT";
run;

proc ttest data=work.schisto;
        class arm;
        var tetanus_11;
run;

proc univariate data=work.schisto;
        var tetanus_11;
        where year = 1;
run;

proc univariate data=work.schisto;
        var tetanus_11;
        where year2 = 1;
run;

proc univariate data=work.schisto;
        var tetanus_11;
        where year3 = 1;
run;

proc glm data=work.schisto;
        class year;
        model tetanus 11 = year;
        means tetanus_11;
run;

*Other clinical;

*Anemia;
proc freq data=work.schisto;
        tables anemia;
        where arm = "SBT";
run;

proc freq data=work.schisto;
        tables anemia;
        where arm = "CWT";
run;

proc freq data=work.schisto;
        tables anemia*arm / chisq;
run;

proc freq data=work.schisto;
        tables anemia;
        where year = 1;
run;

proc freq data=work.schisto;
        tables anemia;
        where year2 = 1;
run;

proc freq data=work.schisto;
        tables anemia;
        where year3 = 1;
run;

proc freq data=work.schisto;
        tables anemia*year / chisq;
run;
```

```
*Hematuria;
proc freq data=work.schisto;
        tables hematuria;
        where arm = "SBT";
run;

proc freq data=work.schisto;
        tables hematuria;
        where arm = "CWT";
run;

proc freq data=work.schisto;
        tables hematuria*arm / chisq;
run;

proc freq data=work.schisto;
        tables hematuria;
        where year = 1;
run;

proc freq data=work.schisto;
        tables hematuria;
        where year2 = 1;
run;

proc freq data=work.schisto;
        tables hematuria;
        where year3 = 1;
run;

proc freq data=work.schisto;
        tables hematuria*year / chisq;
run;

*Hookworm;
proc freq data=work.schisto;
        tables hw;
        where arm = "SBT";
run;

proc freq data=work.schisto;
        tables hw;
        where arm = "CWT";
run;

proc freq data=work.schisto;
        tables hw*arm / chisq;
run;

proc freq data=work.schisto;
        tables hw;
        where year = 1;
run;

proc freq data=work.schisto;
        tables hw;
        where year2 = 1;
run;

proc freq data=work.schisto;
        tables hw;
        where year3 = 1;
run;

proc freq data=work.schisto;
        tables hw*year / chisq;
run;

*TABLE 2 CALCULATIONS -
 MODEL DEVELOPMENT;
```

```
*SM25;

*basic sm25 model;
proc reg data=work.schisto;
	model sm_25_4 = year age bi_arm;
run;

*adding all variables to sm 25 model;
proc reg data=work.schisto;
	model sm_25_4 = year age bi_arm sex coast ascaris_hb_93 msp_1_23 nie_70 tetanus_11
anemia hematuria hw;
run;

*running a forward selection on the full model;
proc reg data=work.schisto;
	model sm_25_4 = year age bi_arm sex coast ascaris_hb_93 msp_1_23 nie_70 tetanus_11
anemia hematuria hw
	/ selection = forward;
run;

	*Forward selection model dropped tetanus and hematuria;

*running a backwards selection on the full model;

proc reg data=work.schisto;
	model sm_25_4 = year age bi_arm sex coast ascaris_hb_93 msp_1_23 nie_70 tetanus_11
anemia hematuria hw
	/ selection = backward;
run;

	*Backward selection dropped all but age, study arm, coast, and nie_70;

*running a stepwise selection on the full model;
proc reg data=work.schisto;
	model sm_25_4 = year age bi_arm sex coast ascaris_hb_93 msp_1_23 nie_70 tetanus_11
anemia hematuria hw
	/ selection = stepwise;
run;

	*Stepwise selection dropped all but age, coast, study arm, nie_70, and year;

*new model based on stepwise selection variables;
proc reg data=work.schisto;
	model sm_25_4 = age coast bi_arm nie_70 year;
run;

*SEA2;

*basic sea2 model;
proc reg data=work.schisto;
	model sea_2 = year age bi_arm;
run;

*adding variables for full sea2 model;
proc reg data=work.schisto;
	model sea_2 = year age bi_arm sex coast ascaris_hb_93 msp_1_23 nie_70 tetanus_11
anemia hematuria hw;
run;

*running a forward selection on the full model;
proc reg data=work.schisto;
	model sea_2 = year age bi_arm sex coast ascaris_hb_93 msp_1_23 nie_70 tetanus_11
anemia hematuria hw
	/ selection = forward;
run;

	*Forward selection model dropped hw, hematuria, and anemia;

*running a backwards selection on the full model;

proc reg data=work.schisto;
```

```
        model sea_2 = year age bi_arm sex coast ascaris_hb_93 msp_1_23 nie_70 tetanus_11
anemia hematuria hw
        / selection = backward;
run;

        *Backward selection dropped all but age, study arm, coast, and nie_70;

*running a stepwise selection on the full model;
proc reg data=work.schisto;
        model sea_2 = year age bi_arm sex coast ascaris_hb_93 msp_1_23 nie_70 tetanus_11
anemia hematuria hw
        / selection = stepwise;
run;
quit;

        *Stepwise selection dropped all but age, study arm, coast, nie_70, and sex;

*new model based on stepwise selection variables plus year;
proc reg data=work.schisto;
        model sea_2 = age coast bi_arm nie_70 year sex;
run;
quit;

*Revised modeling using dichotomous outcome variables;

*Model for SEA;

*just year;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = year2 year3 / dist=p link=log;
        repeated sub=village id / type=cs;
        estimate 'SEA Positive' year2 1/exp;
        estimate 'SEA Positive' year3 1/exp;
run;
quit;

*just sex;
proc genmod data=work.schisto;
        class village id;
        model SEA_new = sex / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' sex 1/exp;
run;
quit;

*just coast;
proc genmod data=work.schisto;
        class village id;
        model SEA_new = coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' coast 1/exp;
run;
quit;

*just study arm;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = bi_arm / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' bi_arm 1/exp;
run;
quit;

*Un-adjusted values for age;
proc genmod data=work.schisto;
        format coast coastf. year yearf. bi_arm bi_armf. sex sexf. age agef.;
        class village_id age;
        model SEA new = age / dist=p link=log;
        estimate 'Age' age 1/exp;
run;
```

```
quit;

proc genmod data=work.schisto;
        class village_id;
        model SEA_new = age / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' age 1/exp;
run;
quit;

*combining above variables;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = year2 year3 sex coast bi_arm age / dist=p link=log;
        repeated sub=village id / type=cs;
        estimate 'SEA Positive' year2 1/exp;
        estimate 'SEA Positive' year3 1/exp;
        estimate 'SEA Positive' sex 1/exp;
        estimate 'SEA Positive' coast 1/exp;
        estimate 'SEA Positive' bi_arm 1/exp;
run;
quit;

*looking at interaction terms;

*sex*coast;
proc genmod data=work.schisto;
        class village id;
        model SEA_new = sex coast sex*coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' sex 1/exp;
        estimate 'SEA Positive' coast 1/exp;
        estimate 'SEA Positive' sex*coast 1/exp;
run;
quit;

*age*coast;
proc genmod data=work.schisto;
        class village_id;
        model SEA new = age coast age*coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' age 1/exp;
        estimate 'SEA Positive' coast 1/exp;
        estimate 'SEA Positive' age*coast 1/exp;
run;
quit;

*bi_arm*coast;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = bi_arm coast bi_arm*coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' bi_arm 1/exp;
        estimate 'SEA Positive' coast 1/exp;
        estimate 'SEA Positive' bi_arm*coast 1/exp;
run;
quit;

*year*coast;
proc genmod data=work.schisto;
        class village id;
        model SEA_new = year coast year*coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' year 1/exp;
        estimate 'SEA Positive' coast 1/exp;
        estimate 'SEA Positive' year*coast 1/exp;
run;
quit;

*year*bi_arm;
proc genmod data=work.schisto;
```

```
        class village_id;
        model SEA_new = year bi_arm year*bi_arm / dist=p link=log;
        repeated sub=village id / type=cs;
        estimate 'SEA Positive' year 1/exp;
        estimate 'SEA Positive' bi_arm 1/exp;
        estimate 'SEA Positive' year*bi_arm 1/exp;
run;
quit;

*sex*bi_arm;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = sex bi_arm sex*bi_arm / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' sex 1/exp;
        estimate 'SEA Positive' bi_arm 1/exp;
        estimate 'SEA Positive' sex*bi_arm 1/exp;
run;
quit;

*age*bi_arm;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = age bi_arm age*bi_arm / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' age 1/exp;
        estimate 'SEA Positive' bi_arm 1/exp;
        estimate 'SEA Positive' age*bi_arm 1/exp;
run;
quit;

*sex*age;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = age sex age*sex / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' age 1/exp;
        estimate 'SEA Positive' sex 1/exp;
        estimate 'SEA Positive' age*sex 1/exp;
run;
quit;

*sex*year;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = year sex year*sex / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' year 1/exp;
        estimate 'SEA Positive' sex 1/exp;
        estimate 'SEA Positive' year*sex 1/exp;
run;
quit;

*age*year;
proc genmod data=work.schisto;
        class village_id;
        model SEA new = year age year*age / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' year 1/exp;
        estimate 'SEA Positive' age 1/exp;
        estimate 'SEA Positive' year*age 1/exp;
run;
quit;


*Model with other outcomes;

*Just asc;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = log_asc / dist=p link=log;
```

```
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' log_asc 1/exp;
run;
quit;

*Just msp;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = log_msp / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' log_msp 1/exp;
run;
quit;

*Just nie;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = log_nie / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' log_nie 1/exp;
run;
quit;

*Just tetanus;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = log_tetanus / dist=p link=log;
        repeated sub=village id / type=cs;
        estimate 'SEA Positive' log_tetanus 1/exp;
run;
quit;

*Just anemia;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = anemia / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' anemia 1/exp;
run;
quit;

*Just hematuria;
proc genmod data=work.schisto;
        class village id;
        model SEA_new = hematuria / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' hematuria 1/exp;
run;
quit;

*Just hw;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = hw / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' hw 1/exp;
run;
quit;

*Full antigen model;
proc genmod data=work.schisto;
        class village_id;
        model SEA_new = log_asc log_msp log_nie log_tetanus anemia hematuria hw / dist=p
link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' log_asc 1/exp;
        estimate 'SEA Positive' log_msp 1/exp;
        estimate 'SEA Positive' log_nie 1/exp;
        estimate 'SEA Positive' log tetanus 1/exp;
        estimate 'SEA Positive' anemia 1/exp;
        estimate 'SEA Positive' hematuria 1/exp;
```

```
        estimate 'SEA Positive' hw 1/exp;
run;
quit;

*POTENTIAL FULL MODEL FOR SEA;
proc genmod data=work.schisto;
        format coast coastf.;
        class village_id age year coast;
        model SEA_new = year sex coast bi_arm age age*coast age*year log_msp log_nie
log_tetanus anemia / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SEA Positive' year 1/exp;
        estimate 'SEA Positive' sex 1/exp;
        estimate 'SEA Positive' coast 1/exp;
        estimate 'SEA Positive' bi arm 1/exp;
        estimate 'SEA Positive' age 1/exp;
        estimate 'SEA Positive' age*coast 1/exp;
        estimate 'SEA Positive' age*year 1/exp;
        estimate 'SEA Positive' log_msp 1/exp;
        estimate 'SEA Positive' log_nie 1/exp;
        estimate 'SEA Positive' log_tetanus 1/exp;
        estimate 'SEA Positive' anemia 1/exp;
run;
quit;



*REVISED FULL MODEL FOR SEA;
proc genmod data=work.schisto;
        format coast coastf. year yearf. bi_arm bi_armf. sex sexf. age agef.;
        class village_id year sex coast bi_arm age;
        model SEA new = year sex coast bi_arm age age*coast age*year log_msp log_nie
log_tetanus anemia / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'Year' year 1/exp;
        estimate 'Sex' sex 1/exp;
        estimate 'Coast' coast 1/exp;
        estimate 'Study Arm' bi_arm 1/exp;
        estimate 'Age' age 1/exp;
        estimate 'Log MSP' log msp 1/exp;
        estimate 'Log NIE' log_nie 1/exp;
        estimate 'Log tetanus' log_tetanus 1/exp;
        estimate 'Anemia' anemia 1/exp;
        slice age*coast / sliceby age diff;
        slice age*year / sliceby year diff;
run;
quit;



*Model for SM25;

*just year;
proc genmod data=work.schisto;
        class village_id;
        model Sm25 new = year2 year3 / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' year2 1/exp;
        estimate 'SM25 Positive' year3 1/exp;
run;
quit;

*just sex;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = sex / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' sex 1/exp;
run;
quit;
```

```
*just coast;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' coast 1/exp;
run;
quit;

*just study arm;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = bi_arm / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' bi_arm 1/exp;
run;
quit;

*Un-adjusted values for age;
proc genmod data=work.schisto;
        format coast coastf. year yearf. bi_arm bi_armf. sex sexf. age agef.;
        class village_id age;
        model Sm25_new = age / dist=p link=log;
        estimate 'Age' age 1/exp;
run;
quit;

proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        ods output GEEEmpPEst=coast_Sm25_new;
run;
quit;

*combining above variables;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = year2 year3 sex coast bi_arm age / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' year2 1/exp;
        estimate 'SM25 Positive' year3 1/exp;
        estimate 'SM25 Positive' sex 1/exp;
        estimate 'SM25 Positive' coast 1/exp;
        estimate 'SM25 Positive' bi_arm 1/exp;
run;
quit;

*looking at interaction terms;

*sex*coast;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = sex coast sex*coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' sex 1/exp;
        estimate 'SM25 Positive' coast 1/exp;
        estimate 'SM25 Positive' sex*coast 1/exp;
run;
quit;

*age*coast;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = age coast age*coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' age 1/exp;
        estimate 'SM25 Positive' coast 1/exp;
        estimate 'SM25 Positive' age*coast 1/exp;
run;
```

```
quit;

*bi arm*coast;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = bi_arm coast bi_arm*coast / dist=p link=log;
        repeated sub=village id / type=cs;
        estimate 'SM25 Positive' bi_arm 1/exp;
        estimate 'SM25 Positive' coast 1/exp;
        estimate 'SM25 Positive' bi_arm*coast 1/exp;
run;
quit;

*year*coast;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = year coast year*coast / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' year 1/exp;
        estimate 'SM25 Positive' coast 1/exp;
        estimate 'SM25 Positive' year*coast 1/exp;
run;
quit;

*year*bi_arm;
proc genmod data=work.schisto;
        class village_id;
        model Sm25 new = year bi arm year*bi_arm / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' year 1/exp;
        estimate 'SM25 Positive' bi_arm 1/exp;
        estimate 'SM25 Positive' year*bi_arm 1/exp;
run;
quit;

*sex*bi_arm;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = sex bi_arm sex*bi_arm / dist=p link=log;
        repeated sub=village id / type=cs;
        estimate 'SM25 Positive' sex 1/exp;
        estimate 'SM25 Positive' bi_arm 1/exp;
        estimate 'SM25 Positive' sex*bi_arm 1/exp;
run;
quit;

*age*bi_arm;
proc genmod data=work.schisto;
        class village id;
        model Sm25_new = age bi_arm age*bi_arm / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' age 1/exp;
        estimate 'SM25 Positive' bi_arm 1/exp;
        estimate 'SM25 Positive' age*bi_arm 1/exp;
run;
quit;

*sex*age;
proc genmod data=work.schisto;
        class village_id;
        model Sm25 new = age sex age*sex / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' age 1/exp;
        estimate 'SM25 Positive' sex 1/exp;
        estimate 'SM25 Positive' age*sex 1/exp;
run;
quit;

*sex*year;
proc genmod data=work.schisto;
        class village_id;
```

```
        model Sm25_new = year sex year*sex / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' year 1/exp;
        estimate 'SM25 Positive' sex 1/exp;
        estimate 'SM25 Positive' year*sex 1/exp;
run;
quit;

*age*year;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = year age year*age / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' year 1/exp;
        estimate 'SM25 Positive' age 1/exp;
        estimate 'SM25 Positive' year*age 1/exp;
run;
quit;


*Model with other outcomes;

*Just asc;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = log_asc / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' log_asc 1/exp;
run;
quit;

*Just msp;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = log_msp / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' log_msp 1/exp;
run;
quit;

*Just nie;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = log nie / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' log_nie 1/exp;
run;
quit;

*Just tetanus;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = log_tetanus / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' log_tetanus 1/exp;
run;
quit;

*Just anemia;
proc genmod data=work.schisto;
        class village id;
        model Sm25_new = anemia / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' anemia 1/exp;
run;
quit;

*Just hematuria;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = hematuria / dist=p link=log;
```

```
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' hematuria 1/exp;
run;
quit;

*Just hw;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = hw / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' hw 1/exp;
run;
quit;

*Full antigen model;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = log_asc log_msp log_nie log_tetanus anemia hematuria hw / dist=p
link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' log_asc 1/exp;
        estimate 'SM25 Positive' log_msp 1/exp;
        estimate 'SM25 Positive' log_nie 1/exp;
        estimate 'SM25 Positive' log_tetanus 1/exp;
        estimate 'SM25 Positive' anemia 1/exp;
        estimate 'SM25 Positive' hematuria 1/exp;
        estimate 'SM25 Positive' hw 1/exp;
run;
quit;

*Having issues with this model for some reason - error message "error in computing the
variance function";
*Error possibly due to correlation - maybe issue with negative variance;

*Cut down to significant antigens, trying model again;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = log_asc log_msp log_nie anemia / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' log asc 1/exp;
        estimate 'SM25 Positive' log_msp 1/exp;
        estimate 'SM25 Positive' log_nie 1/exp;
        estimate 'SM25 Positive' anemia 1/exp;
run;
quit;
*Got results for this;

*POTENTIAL FULL MODEL FOR SM25;
proc genmod data=work.schisto;
        class village_id;
        model Sm25_new = year sex coast bi_arm age age*coast bi_arm*coast log_asc log_msp
log_nie anemia / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'SM25 Positive' year 1/exp;
        estimate 'SM25 Positive' sex 1/exp;
        estimate 'SM25 Positive' coast 1/exp;
        estimate 'SM25 Positive' bi arm 1/exp;
        estimate 'SM25 Positive' age 1/exp;
        estimate 'SM25 Positive' age*coast 1/exp;
        estimate 'SM25 Positive' bi_arm*coast 1/exp;
        estimate 'SM25 Positive' log asc 1/exp;
        estimate 'SM25 Positive' log_msp 1/exp;
        estimate 'SM25 Positive' log_nie 1/exp;
        estimate 'SM25 Positive' anemia 1/exp;
run;
quit;


*Can switch reference groups by putting in a format;
proc format;
        value coastf
```

```
                0 = "z - no"
                1 = "yes";
        run;


*NEW FULL MODEL - with slice for interaction;
proc genmod data=work.schisto;
        format coast coastf. year yearf. bi_arm bi_armf. sex sexf. age agef.;
        class village_id year sex coast bi_arm age;
        model Sm25_new = year sex coast bi_arm age age*coast bi_arm*coast log_asc log_msp
log_nie anemia / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'Year' year 1/exp;
        estimate 'Sex' sex 1/exp;
        estimate 'Coast' coast 1/exp;
        estimate 'Study Arm' bi_arm 1/exp;
        estimate 'Age' age 1/exp;
        estimate 'Log ASC' log_asc 1/exp;
        estimate 'Log MSP' log_msp 1/exp;
        estimate 'Log NIE' log_nie 1/exp;
        estimate 'Log Anemia' anemia 1/exp;
        slice age*coast / sliceby age diff;
        slice bi_arm*coast / sliceby bi_arm diff;
run;
quit;

*New model with swapped final interaction term;
proc genmod data=work.schisto;
        format coast coastf. year yearf. bi_arm bi_armf. sex sexf. age agef.;
        class village_id year sex coast bi_arm age;
        model Sm25_new = year sex coast bi_arm age age*coast bi_arm*coast log_asc log_msp
log nie anemia / dist=p link=log;
        repeated sub=village_id / type=cs;
        estimate 'Year' year 1/exp;
        estimate 'Sex' sex 1/exp;
        estimate 'Coast' coast 1/exp;
        estimate 'Study Arm' bi_arm 1/exp;
        estimate 'Age' age 1/exp;
        estimate 'Log ASC' log_asc 1/exp;
        estimate 'Log MSP' log msp 1/exp;
        estimate 'Log NIE' log_nie 1/exp;
        estimate 'Log Anemia' anemia 1/exp;
        slice age*coast / sliceby age diff;
        slice bi_arm*coast / sliceby coast diff;
run;
quit;
```