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:

_____ _____

YuYen Chan Date

Approval Sheet

Geographic, water usage, and sanitation influences on effectiveness of mass drug administration for schistosomiasis control around Lake Victoria

By
YuYen Lorenzo Chan
MPH

Global Health

_____ [Chair's signature]
[Patrick Sullivan]
Committee Chair

_____ [Member's signature]
[Ryan Wiegand]
Committee Member

_____ [Member's signature]
[Evan Secor]
Committee Member

Abstract Cover Page

Geographic, water usage, and sanitation influences on effectiveness of mass drug administration for schistosomiasis control around Lake Victoria

By

YuYen Chan

Bachelor of Science
University of Washington School of Public Health
2012

Thesis Committee Chair: Patrick Sullivan, DVM, PhD, Professor

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 Public Health in Global Health
2016

Abstract

Geographic, water usage, and sanitation influences on effectiveness of mass drug administration for schistosomiasis control around Lake Victoria

By YuYen Chan

Background: Schistosomiasis is an important neglected tropical disease that requires public health intervention. According to the WHO, there are almost 800 million individuals at risk of infection globally. Lake Victoria in western Kenya is the largest tropical lake in the world and also serves as a habitat for *Biomphalaria* snails, which are the intermediate host of *S. mansoni*. Because the lake is used for various activities within the daily lives of the residents around the lake, there is high transmission of *Schistosoma mansoni* in the region, resulting in a large public health burden.

Purpose: An aim of the Schistosomiasis Consortium for Operation Research and Evaluation (SCORE) is finding the most effective interventions to gain control of schistosomiasis. The focus of this paper evaluated what geographical and village-level water and sanitation factors may influence how well the interventions worked.

Methods: This project compared six treatment plans of school based treatment and community-wide treatment over a five year period. One hundred and fifty villages with a primary school participated. Prevalence and intensity of *S. mansoni* infections was monitored in 9-12 year old children using the Kato-Katz thick smear method on stool specimens. Key informants from each village provided survey date of village level behavior.

Results: We found certain geographic areas to be statistically different when looking at the odds of infection for *S. mansoni*. Closer examination identified a primary cluster of high prevalence in Bondo and Rarieda areas with immunization interventions and drinking from bore holes associated with lower impact of the intervention.

Conclusion: Geographic location could have a potential effect on the success of intervention. Village-level event such as experiencing droughts and the presence of immunization interventions or individual habits in these areas may also be associated with treatment results. Further examination of other factors that affect the impact of mass drug administration for *S. mansoni* control is warranted.

Cover Page

Geographic, water usage, and sanitation influences on effectiveness of mass drug administration for schistosomiasis control around Lake Victoria

By

YuYen Chan

Bachelor of Science in Public Health
University of Washington School of Public Health
2012

Thesis Committee Chair: Patrick Sullivan, DVM, PhD, Professor

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 Public Health in Global Health, Infectious Diseases
2016

Acknowledgements

Sincerest appreciations to Mr. Ryan Wiegand for his time, patience, statistical expertise, mentorship, and overall guidance on this project. Much gratitude to Dr. William E. Secor for his time in accommodating my interests, his advice on this thesis, and the wonderful opportunity to participate in the SCORE project at the CDC. Many thanks to Dr. Patrick Sullivan for his time and advice during the thesis process.

Additional acknowledgements to

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

for allowing me to use this data for my analysis of my thesis.

Thanks to friends and family for all their support.

Table of Contents

Contents Page

Introduction

According to WHO estimates, there are more than 237 million individuals who required treatment for schistosomiasis in 2010, along with 779 million individuals globally who are at risk of infection [1]. In areas that are endemic for schistosomiasis, initial infection can be acquired at a young age. One study showed that 14% of l-year-old children along the Kenyan shores of Lake Victoria were positive for *Schistosoma mansoni*, and another study found a prevalence of 47.4% in children less than three years of age [2]. In these areas, overall prevalence of schistosomiasis can be as high as 86-90% [2, 3], with a high risk of reinfection among people who receive treatment. In children, the primary consequence of infection is anemia and individuals with high intensity infections can experience growth stunting, impaired cognitive development, decreased quality of life, and exercise intolerance. Other consequences of schistosomiasis can include increased susceptibility to co-infections, infertility, liver fibrosis, portal hypertension, and death [4].

According to the Global Burden of Disease, it is estimated that the total number of daily adjusted life years (DALYs) lost to schistosomiasis is 1.532 million per year, 77% of which are in sub Saharan Africa, and account for 0.1% of the world global burden of disease [5]. Although the prevalence of schistosomiasis is high, its morbidity is low. Other authors believe these DALYs grossly underestimate the public health burden of schistosomiasis. With recent studies, it is estimated that mortality for schistosomiasis may be up to 280,000 deaths annually in sub-Saharan Africa, where true prevalence is some areas is potentially underestimated by 50% [7]. Because of this, the public health significance and prioritization of schistosomiasis control and treatment has been in constant discussion [5, 6].

Lake Victoria, located in western Kenya, is the largest tropical lake in the world and is also endemic for *Biomphalaria* snails, which are the intermediate host of *S. mansoni* [8]. The lake serves as the main source of water for the people residing in Nyanza Province and is used frequently for bathing, washing, and fishing [9]. Both environmental factors supporting mollusk survival and human behaviors contribute to the high prevalence of schistosomiasis in this region [10].

An aim of the Schistosomiasis Consortium for Operation Research and Evaluation (SCORE) is finding the most effective intervention for achieving schistosomiasis control [11]. The project involves a five-year longitudinal study that consists of six different treatment schedules (arms) [11] in the districts surrounding Lake Victoria. The focus of this paper is to examine how geographic areas around Lake Victoria, health interventions, and behaviors could influence the impact of different treatment schedules on change in prevalence of S *mansoni*.

Literature Review

Biology and background

Schistosomiasis in humans is caused by a parasitic fluke of the genus *Schistosoma,* which dwells in the blood of its host. Schistosomes are whitish-grey in appearance with separate male and female sexes. Adult worms typically are cylindrical, 7-20 mm in length, and, depending on the species, live in either the mesenteric or venous plexus where they mate and feed on blood through anaerobic glycolysis [5]. Thousands of eggs are produced each day by the female worms, and each ovum contains a larval form called a miracidium. The larvae produces several enzymes that stimulate the host immune response, which in turn is thought to aid the egg in traveling to the lumen of the bladder or intestine. The eggs are then excreted in the urine or feces, where they can survive up to 7 days. When the eggs contact a source of water, they release the ciliated miracidia that are within, which are guided to freshwater snails through chemical and light stimuli [5]. The snail serves as the intermediate host in the schistosome life cycle, where the miracidia convert into sporocytes that multiply asexually and turn into cercarial larvae, which are the infectious form for humans (see Figure 1).

About 4-6 weeks after infection of a snail with miracidia snail, the cercariae departs into the surrounding aquatic environment where it can live for up to 72 hours without a definitive host. A snail infected by one miracidium can shed thousands of cercariae every day for several months. This shedding is invoked by light and heat, and occurs mainly during the day [5]. Once the cercariae have located a suitable host, they penetrate the skin and travel through the blood via lungs and into the liver, where they transform into young worms known as schistosomulae. They eventually mature in the portal vein after 4-6 weeks, and then will migrate to the bladder or

mesenteric vasculature to complete the life cycle. The average lifespan of an adult worm is about 3-5 years, although there are many documented cases of persons who maintain infections for more than 30 years after moving out of an endemic area. One schistosome worm pair has the theoretical reproduction potential of up to 600 billion schistosomes [5]

S. mansoni, transmitted by *Biomphalaria* snails, is the main schistosome that infects human beings. It causes intestinal and hepatic schistosomiasis, and is found geographically in South America, Africa, and the Arabian Peninsula [12]. One of the other types of schistosomes, *S. haematobium* causes urinary schistosomiasis in Africa and the Arabian Peninsula; *S. japonicum* causes intestinal and hepatosplenic schistosomiasis in China, the Philippines, and Indonesia [6]. *S. intercalatum* and *S. mekongi* infect a wide range of animals including cattle, dogs, pigs, and rodents but can also infect humans. Although humans are the main host, *S. mansoni* is also found in rodents and primates [6].

In order for individuals to become infected with *S. mansoni*, they either have to live in or have visited an area that fosters disease transmission. Freshwater habitats where snails flourish and people come into contact with that water correlate with high human schistosomiasis prevalence [13, 14].

Studies investigating geospatial and environmental variables have addressed factors such as soil pH, annual rainfall, as well as land surface temperature on schistosomiasis transmission. Schistosomiasis is dependent on the survival of the snails, and environmental factors that predict a snail habitat can be used to help predict prevalence of schistosomiasis [6].

Ecology of snail hosts is the main factor governing geographical distributions of different species, since each schistosome species depends on a different genus of snail intermediate host (see Figure 2). Lakes, ponds, and streams are the primary sources for infection; however, man-

made reservoirs and irrigation systems have also contributed to the spread of schistosomiasis [6]. Rates and intensities of infection occur more frequently at an early stage in humans, typically ages 8-15. The prevalence and intensity of infection decrease in adults, where water contact patterns, along with innate and acquired immunity of the host result in reduced parasite infection [6].

Pathology

The different species of schistosomes have symptoms that vary depending on where in the body the parasite lives and releases its eggs. Symptoms range from severe acute manifestations to chronic disease and from minimal morbidity to death.

Upon entering the host by penetration of the skin, cercariae can cause a temporary urticarial rash that can persist for days as paulopruriginous lesions, seen frequently after primary infection in tourists and migrants [6]. Acute schistosomiasis, also known as Katayama fever, is a systemic acute reaction to infection that occurs several weeks to months after primary infection. Disease symptoms are characterized by sudden fever, followed by fatigue, myalgia, malaise, cough, eosinophilia, and chest infiltrates that can be seen through radiography [6, 15]. Most patients will recover from symptoms after 2-10 weeks, however more severe disease includes weight loss, diarrhea, dyspnea, abdominal pain, toxemia, hepatosplenomegaly, and rash throughout the body. In chronically exposed populations, it is rare to see Katayama fever due *S. mansoni* or *S. haematobium* [6, 16]. While there may be a certain degree of underdiagnoses in endemic populations, it is seen most commonly in tourists, travelers, and individuals who are born to mothers from non-endemic areas [16]. Many people who are diagnosed with schistosomiasis in clinics in Europe and the United States have documented water exposure in Lake Malawi, Lake Victoria, and Lake Volta, along with Zambesi and Niger deltas, as well as several lake resorts in

South Africa. Risky recreational activities including freshwater swimming, scuba diving, water skiing, rafting, and bathing [16].

Different types of Schistosomiasis

Intestinal schistosomiasis is associated with *S. mansoni* and *S. japonicum* infections and is caused by eggs that migrate through the intestinal wall, which then provoke mucosal and granulomatous inflammation, along with pseudopoplyposis, microulcerations, and bleeding [20]. Lesions are mostly situated in the large bowel and rectum and the most common signs are chronic or intermittent abdominal pain and discomfort, loss of appetite, and diarrhea [20].

Hepatic schistosomiasis is the more serious form of intestinal schistosome infections and is caused by a reaction to ova that are trapped in the presinuoidal periportal spaces of the liver, which can cause hepatomegaly in children and adolescents [17]. This inflammation is often difficult to differentiate from similar pathologies caused by malaria. Fibrotic schistosomiasis develops following intense chronic infection, and is generally found in young and middle aged adults. Large deposition of diffused collagen deposits in the periportal spaces can lead to pathognomonic periportal or Symmer's pipestem fibrosis, which results in progressive occlusion of the portal veins, portal hypertension, splenomegaly, collateral venous circulation, portocaval shunting, and gastrointestinal varices [6]. Although the liver is generally hard upon palpitation, unlike cirrhosis, the hepatocellular function remains generally unaffected [6]. The most serious form of hepatic schistosomiasis in characterized by bleeding from the gastro-esophageal varices leading to commonly fatal complications. Blood loss from varices can lead to hypoalbuminemia, cachexia, anemia, and growth retardation [18].

Urinary schistosomiasis is associated with *S. haematobium* infections and occurs due to the inflammation and ulceration of bladder walls in response to parasite eggs [12]. In infected

individuals, blood is seen near the end of urination, and in severe cases, the entire urine sample is a dark color. Chronic lesions can evolve to fibrosis or calcification of the bladder and lower ureters, which then causes chronic compression that can lead to parenchymal damage and kidney failure [19].

Ectopic schistosomiasis is due to portalcaval shunting, which can cause ova to be leaked into the perialveolar capillary beds [6] resulting in granulomas that cause bronchial symptoms and fibrosis, which can be complicated by pulmonary hypertension [21]. Genital schistosomiasis is caused by eggs of *S. mansoni* and *S. haematobium* being deposited in the reproductive organs. This is common, with symptoms in female patients including hypertrophic and ulcerative lesions of the vulva, vagina, and cervix that could potential lead to increased sexual transmission of HIV infection. Lesions within ovaries and fallopian tubes can also lead to infertility, and in men the epididymis, testicles, spermatic chord, and prostate can be affected [21]. Neuroschistosomiasis is a result of worms or eggs in the cerebral or spinal venous plexus that can lead to paralysis [22].

Diagnosis

The standard for the diagnosis of schistosomiasis is the microscopic examination of excreta [23]. Analysis of urine samples for *S. haematobium* eggs includes sedimentation, centrifugation, or filtration. Eggs are then examined and counted directly under microscope. Intensity of is expressed as eggs per 10 mL of urine. The most common method for diagnosis of intestinal schistosomiasis is through fecal exams [24]. In the field, the Kato-Katz method is commonly used because it allows quantification of egg counts, expressed as eggs per gram of feces, which is a relative measure of infection intensity [24]. Rectal snips can also be used [25]. While antibody assays are quite sensitive, they cannot distinguish between exposure history and active infection [20]. Antibody assays can also cross react with other helminths [24, 25]. Immunoassays are more

important for tourists or travelers [15]. They also can be useful for incidence studies among children, and also in low-transmission areas or after the implementation of a control program. Cystoscopy and endoscopy are used to visualize bladder lesions and esophageal varices. Granulomatous inflammation and periportal fibrosis are identified through ultrasonography, laparoscopy and wedge biopsy. Radiography is used in the examination renal, ureteral, and bladder pathology [26].

Treatment

The most widely used treatment today is the drug praziquantel [27]. Praziquantel is a safe drug and is used against all schistosome species [28]. It is sold as 600 mg tablets, with a recommended treatment of 40 mg/kg bodyweight in a single dose [29]. Within 1 hour of ingestion, worms are paralyzed [6]. Side effect are generally mild and include nausea, vomiting, malaise, and abdominal pain. In individuals with heavy infection, acute colic and bloody diarrhea can occur after treatment, which are most likely caused by massive worm shifts and antigen release [30]. Praziquantel has low toxicity, which is important in terms of long term treatment using the drug, especially for control and elimination programs [31].

Eggs and immature worms are generally unaffected by praziquantel, and eggs can still be excreted for several weeks after treatment. 70-100% of patients stop excreting eggs after a single dose of 40mg/kg, and for those who are not cured, urine or fecal eggs concentrations are reduced by more than 95% [32]. Clinical studies with radiographs and ultrasonography have shown regression of intestinal and vesical lesions, reactive hepatomegaly and severe lesions of the upper urinary tract over time [33]. Thus, treatment is ideal for population level control strategies [6]. For populations with rapid reinfection and high levels of initial egg counts, treatment dosage can be increased to 60mg/kg, with a split dosage to avoid general side effects. After 6-12 weeks, a

repeated dose is generally recommended for complete cure of potential prepatent infections and to ensure that adult worms are completely cleared from the system.

Immunology

Individuals who live in endemic areas eventually develop some form of immune resistance to schistosomiasis after years of exposure [34]. Recent epidemiology and mathematical modelling also support the existence of acquired immunity [35]. Studies comparing children and adults show that children are more likely to be susceptible to disease. In populations with recent exposure, age-related immunity also plays an important role in the epidemiology of schistosomiasis [36].

The pathology of schistosomiasis is due to the host immune response to parasite eggs. Reactions around the eggs by CD4-positive T cells that trigger eosinophils, monocytes, and lymphocytes. There have been multiple efforts to develop a vaccine against schistosomiasis, with several antigens having potential efficacy [37]. However, test in animals have shown varied and inconclusive results. Recombinant rShGST-28 (Bilhvax; Eurogentec, Herstal, Belgium) has already undergone phase I and phase II clinical trials, however the question of how feasible and applicable schistosomiasis vaccines are is still being addressed [38].

Control options

Modern schistosomicides such as praziquantel given at the community level have recently demonstrated promising results for disease control [7]. This type of intervention requires long term planning to ensure sustainability towards more effective control and elimination campaigns in the future [6]. Molluscicides that kill snails are generally expensive and logistically complex, with much human and material resources needed, along with detailed epidemiological and malacological surveillance [39]. Also, snail populations are generally reduced but cannot be

eliminated, and the toxicity of molluscicides for other aquatic animals negatively impacts both ecological and economic outcomes [39].

Ideally, schistosomiasis can be eliminated through behavioral changes, sanitation, and provision of a safe water supply, as seen in Japan [40]. Implementing educational programs can improve knowledge about the disease; however, behavior change is difficult without other options for water in many endemic areas [41].

Population based treatment, either community wide or school based with praziquantel, is recommended by the WHO, and is currently the approach being used by most national control programs [29]. Various strategies include indiscriminate mass treatment, active case finding, and treatment of particular risk groups, including school-aged children [29].

Methods: Score project

Background of project

To address practical issues of schistosomiasis control, the Schistosomiasis Consortium for Operation Research and Evaluation (SCORE) was created with goals to identify the most cost effective preventative intervention using praziquantel. SCORE was established through a grant that was provided to the University of Georgia Research Foundation from the Bill & Melinda Gates Foundation in December 2008 [11]. The project targets improving current schistosomiasis control programs and is composed of several components, including: evaluation of screening tests for *S. mansoni*, development of a gold standard diagnostic test for *S. mansoni*, assessment of the impact of drug pressure on parasite genetics, research related to snail control, studies on elimination of *S. mansoni* in areas of very low prevalence, and large field studies to compare multi-year strategies for mass drug administration (MDA) in areas with moderate and high prevalence of schistosomiasis [42,43]. In addition to supporting original research, another aspect of SCORE was the "rapid answers project" (RAP) [11], which synthesizes existing data to answer important programmatic questions.

In accordance with World Health Organization treatment guidelines, MDA should be primarily focused on school aged children living in endemic areas with praziquantel treatment through school based treatment (SBT) [29]. In areas with a higher prevalence of infection, community-wide treatment (CWT), which includes providing praziquantel to adults and pre-school children taller than 94cm may be indicated [45]. One interest of SCORE's study designs focuses on different schedules of MDA for schistosomiasis control in communities with high (>25%) prevalence among school-aged children [45]. The outcome measures for these studies

sampled 9-12 year old children, which was determined to be the best age range to study and correlates with data from adults through the analysis of SCORE baseline data [11,46].

Large scale cluster randomized trials were designed for areas with prevalence higher 25% in school-aged children. Interventions occur once a year, excluding villages that were on "praziquantel holiday" years, in which there would be no drug treatment that year [11]. Data were collected annually, excluding villages that were on drug holiday. Data was collected before the first MDA and a final evaluation was performed a year after the fourth round of treatment [11].

Many studies on schistosomiasis have been conducted in Nyanza Province of western Kenya, all which have found extremely high prevalence of *S. mansoni*, the prevalence of which decreases with distance with increasing distance from to the lake [47, 48]. The prevalence of Schistosomiasis in children as young as 1 year old in this area was approximately 14%. Another study found prevalence of 47.4% in children less than 3 years of age [49].

Study Design

For the gaining control study, there were 6 arms to compare different treatment regimens. The most intensive intervention (arm 1) involves four years of community wide treatment (CWT), while the least intensive intervention (arm 6) involves two year of school based treatment (SBT) alternating with praziquantel (PZQ) "holiday" years.

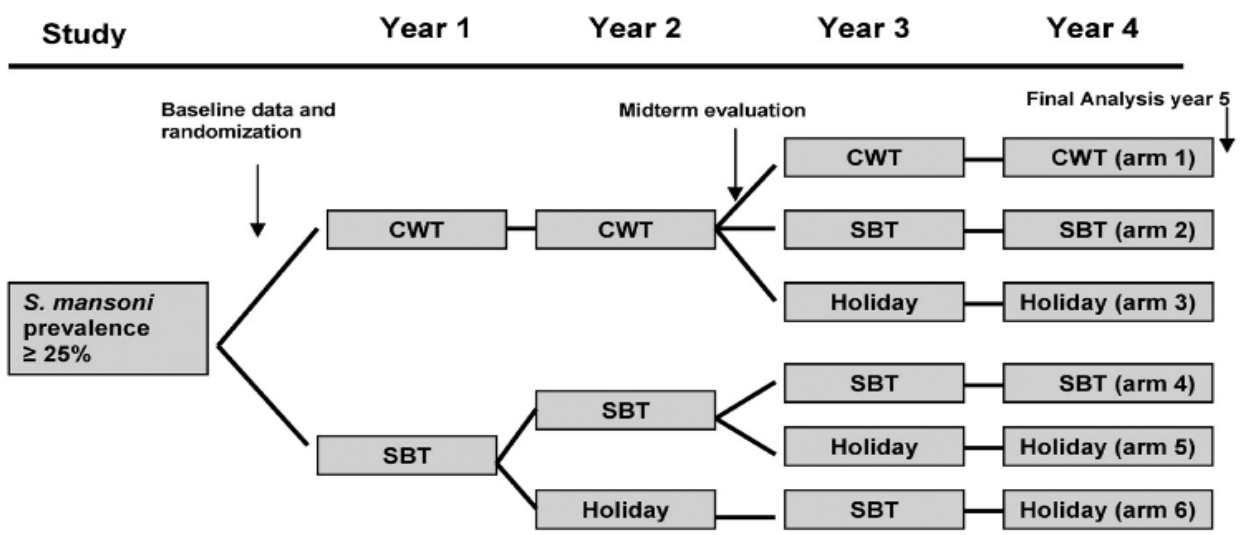


Figure 3. Study design for the Schistosomiasis Consortium for Operational Research and Evaluation (SCORE). Participating communities were randomized into six intervention arms, with community-wide treatment (CWT), and school-based treatment (SBT) [11].

Sample size calculations along with logistical consideration determined the number of villages per arm, and the number of 9-12 year old children tested per village. Infection prevalence was assumed to be 50% at year 1, and was predicted to reduce to 15% by the end of the study in the most intense treatment (arm 1). Minimum effect sizes were computed with 90% power for a two sided $\alpha = 0.05$ level test. Negligible correlation was assumed between measurements at the beginning and end of the study and 25 villages per treatment with 100 children age 9-12 years per village estimated as sufficient to detect treatment differences [11].

Eligibility to participate required that each village had a primary school, at least 100 children aged 9-12 and a starting prevalence of schistosome infection. In each village, fifty children aged 13-14 were tested for *S. mansoni* to estimate likely prevalence among 9-12 year olds.. *S. mansoni* infections in these children were evaluated using the Kato-Katz method [50]. Of those villages, 150 villages with >25% prevalence were identified. All eligible villages were assigned to study arms through simple random allocation [11]. For villages with more than 100 9-12 year old

children, random selection was used to select study participants. A new random selection was performed each year; thus, which children were sampled each year did not necessarily include the same children sampled in other years. However, as praziquantel delivery was through MDA, both sampled and non-sampled children should have been treated each year.

Data collection and management

A list of behaviors that may impact *S. mansoni* infection were collected from a key informant from each village. This village-level data included factors such as sanitation, water contact, and health intervention information that might affect schistosomiasis transmission [11]. A copy of the survey is included in the appendix. Various response choices that were used for analysis were transferred into Microsoft Excel.

Data were collected through a mobile-based system (EpiCollect®) that was developed at Imperial College London. Data entry was performed on-site, then synced to a central server where data were then downloaded for cleaning and analysis [11].

Ethical statement for SCORE project

Written informed consent was obtained from adults (including parents/legal guardians of children in the study) and assent was obtained from children less than 18 years old. This research proposal was reviewed and approved by the National Ethics Review Committee of the Kenya medical research institute (Approval number SCC 1820) and by the Institutional Review Board of the CDC (Approval #: 6016).

Analysis

Statistical tests and models were performed using SAS® 9.4 software (SAS Institute, Inc., Cary, NC), at the 5% significance level. Raw data collected from villages were cleaned in Microsoft Excel and read into SAS software.

Data exploration and cleaning

We summarized continuous data with means, standard deviations, and 95% confidence limits, while categorical data were summarized with frequencies and percentages. The change in village-level *S. mansoni* prevalence was the outcome of interest, and differences in mean prevalence in all five years of data were examined (table 1).

Categories with low numbers (n<10) were collapsed to allow for robust estimation in the model. We used PROC SURVEYREG [51] to test for significance of variables to the change in prevalence of *S. mansoni*, while accounting for clustered data via Taylor series linearization.

Modeling and Variable Selection

A binomial model using generalized estimating equations [52] in PROC GENMOD was chosen to explore significant indicators of the changes in village level *S. mansoni* prevalence. Year, village survey data, geographic area (Rarieda, Bondo, Homa Bay/Rachuonyo, and Kisumu/Nyakach), and interactions were assessed using QIC and QICu [53,54] to determine if the variable improved the model's fit to the data. Because we were interested in the association between geographical areas and village level *S. mansoni* prevalence during each year, we forced in geographical areas, the distance to Lake Victoria, and year variables. The model with the smallest QICu in Table 2 serves as our final model. We also assessed interaction with year and distance to the lake. Results will be presented odds ratios with 95% confidence intervals.

Cluster Analysis

We explored whether villages exhibited any spatial or geographic clustering at the end of the trial. The spatial scan statistic in SaTScan [55] was used to evaluate whether schistosomiasis prevalence at year 5 was evenly distributed over space or presented in elliptically-shaped disease clusters [56].

After identifying a spatial cluster of villages with high prevalence, we fit a logistic regression model to evaluate whether any village-level covariates (found in Table 3) significantly influenced the odds of falling within the high prevalence cluster. The GENMOD procedure in SAS software was used and results are reported as odds ratios.

Results

No significant associations were found between *S. mansoni* prevalence and reported flooding or drought and concurrent HIV, malaria, immunization, or NTD health campaigns nor with drinking, washing, or bathing in rain, river, or tap water. The year of study (χ^2=32.83 $_{(df=149)}$, p<.0001), geographic area ($\chi^2_{=}$11.85$_{(df=149)}$, p<.0001) and distance from Lake Victoria (χ^2=10.15 $_{(df=149)}$, p=0.0018) were significantly associated with *S. mansoni* prevalence (see Table 2, Figure 5).

The final model included survey data, year, and geographic area as well as an interaction terms of geographic area to year to see whether treatment was statistically different between the years in these geographic regions. We saw a drop in the overall prevalence of *S. mansoni* between years 1 and 5, as expected due to praziquantel treatment [26], (see Figure 3).

In analyzing the model that includes geographic areas per year, we used Bondo as a referent group and found that Homa Bay and Kisumu had consistently lower odds by comparison. In year 1, compared to Bondo, the odds ratio (OR) of *S. mansoni* infection and confidence interval (CI) of Homa Bay were 0.38 and 0.23-10.63 (p=0.0002), respectively; for Kisumu the OR was 0.45 and the CI was 0.26-0.77 (p=0.0037). Similar results when comparing to Bondo were found in year 2: Homa Bay (0.31 (CI: 0.19-0.51), p<0.0001), Kisumu (0.40 (CI: 0.23-0.69), p=0.0009), year 3: Homa Bay (0.21(CI: 0.13-0.32), p<0.0001), Kisumu (0.40 (CI: 0.23-0.70), p=0.0014), year 4: Homa Bay (0.18 (CI: 0.11-0.31), p<0.0001), Kisumu (0.31 (CI: 0.18-0.53), p<0.0001), and year 5: Homa Bay (0.14 (CI: 0.09-0.22), p<0.0001), Kisumu (0.38 (CI: 0.22-0.63), p=0.0002). Rarieda showed no significant differences from Bondo in any of the

years. No significant interaction association was found between year and distance to Lake Victoria.

Given these differences between geographic areas, we performed an analysis to determine if there was spatial clustering of high village-level prevalence. Year 5 data revealed an elliptically-shaped cluster along the far western edge of the study area outside of the Winam Gulf area of Lake Victoria, along the shores of both Rarieda and Bondo districts. Villages within the cluster had an increased prevalence of 3.48 times compared to villages outside of the cluster. Villages in the high prevalence cluster had a decreased odds of having had a drought (OR=0.13, (CI: 0.024-0.68, p=0.016) or individuals practicing subsistence farming (OR=0.212, (CI=0.075-0.6001), p=0.0035). There were increased odds of immunization interventions (OR=1.33 (CI: 1.02, 1.74), p=0.038) and wells or bore holes as a water source (OR=4.27 (1.021, 17.87), p=0.047) in the high-prevalence cluster (see Table 3).

Discussion

In our model, we were interested in how village behaviors, intervention years, and geographic location around Lake Victoria affected the village-level prevalence. Individuals who live closer to bodies of water with known *Biomphalaria* snail populations are generally at higher risk for infection [47] (see Figure 5). Village survey data based on human behaviors aggregated and analyzed bivariately across all years of the study showed statistically significant associations between *S. mansoni* prevalence and the presence of malaria, immunization campaigns, and NTD interventions, along with the presence of flood and drought, as well as river drinking, washing, and bathing, rain washing and bathing, and tap water drinking, washing and bathing (See Table 1, Table 2).

Results from several studies have found inconclusive evidence for associations with *S. mansoni* for some of the covariates that we found to not be associated with infection. Thus, the presence of interventions for HIV, malaria, immunizations, or neglected tropical diseases (NTD) may not have a strong impact on *S. mansoni* prevalence. Human behaviors and occupations that involve direct contact with lake water containing schistosome-infected snails [6] are known to affect prevalence, but our results are somewhat surprising since we found no significant associations with *S. mansoni* in our model given the community habits found in our village-level data. The distance to Lake Victoria may serve as a confounder for identifying specific behaviors because proximity to bodies of water with known *Biomphalaria* populations consistently affect odds of schistosomiasis prevalence [10, 58, 59].

We noticed a statistical difference for the odds of infection in the geographic areas of Homa Bay/ Rachuonyo and Kisumu/Nyakach during all years when comparing to our referent

area of Bondo. Many of the villages in Bondo and Rarieda fell into the high prevalence cluster shown in Figure 4. There were several covariates associated with the high prevalence cluster. Drought and subsistence farming were less common in the high prevalence cluster than in other areas. The frequency of immunization campaigns or using a bore hole for drinking water were more common in the high prevalence cluster (see Table 3). A separate study supporting our results found that bore holes did not reduce population level prevalence, due to several reasons including: lack of knowledge of bore hole, contamination, as well as limited numbers of these water sites [60]. Our findings could also have been coincidental, where villages in the high prevalence cluster also happened to have more immunization campaigns and bore holes but they do not affect schistosomiasis transmission. Further exploration of these immunization campaigns and behaviors around the number of well or bore holes per village are needed to examine their effect on risk of *S. mansoni* infection, and there may yet be unmeasured factors that also affect the risk of *S. mansoni*.

Limitations

Survey data were collected at the village level by designated key informants for entire villages. Hence, these data may not be representative of actual individual behavior, leading to possible non-differential misclassification bias. This may have been a reason as to why the village survey data did not improve the fit of the model. Another potential limitation is that prevalence data from children ages 9-12 were used to represent the prevalence of entire villages [10]. Moreover, coverage was not fully monitored, therefore we are not certain of the proportion of the population of each village that actually received treatment. SBT only includes children who attend school so any impact of transmission by children not in school or adults was not affected. Lastly, the population in which prevalence was assessed did not remain completely

constant, due to children aging out of the study, or moving into a different village. As a result, the internal validity of this study is weakened because some of the children were not measured consistently throughout the duration of the study.

Public Health Implications

The results from our studies indicated a notable difference in geographic areas that responded differently to treatment, potentially due to geographic causes that will require investigating other factors that may influence *S. mansoni* acquisition. For example, behaviors collected at the individual level instead of village level and geographic characteristics that may affect snail populations. Overall, praziquantel treatment did decrease *S. mansoni* prevalence and should be administered to treat the disease.

Additionally, re-infection often occurs after without adequate control measures [59], thus long term treatment is often recommended for controlling and reducing the overall morbidity of *S. mansoni* [61, 62]. The study of the treatment arms in currently being analyzed and may present interesting results in terms of treatment efficacy and cost benefits.

It may also be worth exploring potential control combinations, such as praziquantel with molluscicides to further reduce the force of transmission [61, 63]. Chemical molluscicides are generally not recommended because they are visually unappealing and they detrimentally affect other aquatic organisms [63]. The introduction of predatory prawns has been shown to be effective in lab settings, and is potentially a viable option for future snail control [64, 65]. Our analysis of high prevalence clusters noted that environmental circumstances and occupational behavior can reduce risk of *S. mansoni*. Behavior change and improved infrastructure as seen in other countries that have eliminated schistosomiasis may ultimately be the solution [6], though

asking individuals to change behavior that is directly linked to their livelihood is complicated, difficult and poses a continual challenge to the control of this disease.

References

1. WHO, 2012. *Schistosomiasis: population requiring preventive chemotherapy and number of people treated in 2010*. Wkly Epidemiol Rec 87:37–44.
2. 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 84:787–791
3. Stothard JR, Kabatereine NB, Tukahebwa EM, Kazibwe F, Mathieson W, Webster JP, Fenwick A, 2005. *Field evaluation of the Meade Readiview handheld microscope for diagnosis of intestinal schistosomiasis in Ugandan school children*. Am J Trop Med Hyg 73:949–95
4. Colley, DG et al. "Human Schistosomiasis." Lancet (London, England) 383.9936 (2014): 2253–2264. PMC. Web. 19 Apr. 2016.
5. Gryseels B, Polman K, Clerinx J, Kestens L. *Human schistosomiasis*. Lancet. 2006 Sep 23; 368(9541):1106-18. Review. PubMed PMID: 16997665.
6. Utroska JA, Chen MG, Dixon H, et al. An estimate of global needs for praziquantel within schistosomiasis control programmes (WHO/SCHISTO/89.102). Geneva: World Health Organization, 1989.
7. King CH, 2010. Parasites and poverty: the case of schistosomiasis. Acta Trop 113:95–104.
8. Woodhall DM, Wiegand RE, Wellman M, Matey E, Abudho B, Karanja DMS, et al. Use of Geospatial Modeling to Predict Schistosoma mansoni Prevalence in Nyanza Province, Kenya. PLoS Oe 2013; 8.
9. Grimes, JET. et al. "The Relationship between Water, Sanitation and Schistosomiasis: A Systematic Review and Meta-Analysis." Ed. Stephen T. McGarvey. PLoS Neglected Tropical Diseases 8.12 (2014): e3296. PMC. Web. 19 Apr. 2016.
10. Magendantz, Margaret. "The Biology of Biomphalaria Choanomphala and B. Sudanica in Relation to Their Role in the Transmission of Schistosoma Mansoni in Lake Victoria at Mwanza, Tanzania." *Bulletin of the World Health Organization* 47.3 (1972): 331–341. Print.
11. Samuels, AM. et al. "*Schistosoma Mansoni* Morbidity among School-Aged Children: A SCORE Project in Kenya." *The American Journal of Tropical Medicine and Hygiene* 87.5 (2012): 874–882. *PMC*. Web. 13 Apr. 2016.
12. Doumenge JP, Mott KE. Global distribution of schistosomiasis: CEGET/WHO atlas. *World Health Stat Q* 1984; **37:** 186–99.

13. Smith JH, Christie JD, 1986. The pathobiology of Schistosoma haematobium infection in humans. Hum Pathol 17: 333–345

14. Cheever AW, Kamel IA, Elwi AM, Mosimann JE, Danner R, Sippel JE. *Schistosoma mansoni* and *S haematobium* infections in Egypt, III: extrahepatic pathology. *Am J Trop Med Hyg* 1978; 27: 55–75

15. Bottieau E, Clerinx J, De Vega MR, et al. Imported Katayama fever: clinical and biological features at presentation and during treatment. *J Infect* 2006; 52: 339–45

16. Jelinek T, Nothdurft HD, Loscher T. Schistosomiasis in travelers and expatriates. *J Travel Med* 1996; 3: 160–64

17. Gryseels B, Polderman AM. Morbidity, due to schistosomiasis *mansoni*, and its control in Subsaharan Africa. *Parasitol Today* 1991; 7: 244–48

18. Lambertucci RL. *Schistosoma mansoni*: pathological and clinical aspects. In: Jordan P, Webbe G, Sturrock FS. Human schistosomiasis. Wallingford: CAB International, 1993: 195–235

19. Gryseels B. The relevance of schistosomiasis for public health. *Trop Med Parasitol* 1989; 40: 134–42

20. Cheever AW, Kamel IA, Elwi AM, Mosimann JE, Danner R, Sippel JE. *Schistosoma mansoni* and *S haematobium* infections in Egypt, III: extrahepatic pathology. *Am J Trop Med Hyg* 1978; 27: 55–75

21. Lambertucci R L. *Schistosoma mansoni*: pathological and clinical aspects. In: Jordan P, Webbe G, Sturrock FS. Human schistosomiasis. Wallingford: CAB International, 1993: 195–235

22. Ferrari TC. Involvement of central nervous system in the schistosomiasis. *Mem Inst Oswaldo Cruz* 2004; 99 (suppl 1): 59–62

23. Vennervald BJ, Kahama AI, Reimert CM. Assessment of morbidity in *Schistosoma haematobium* infection: current methods and future tools. *Acta Tropica* 2000; 77: 81–89.

24. Richter J, Correia Dacal AR, Vergetti Siqueira JG, et al Sonographic prediction of variceal bleeding in patients with liver fibrosis due to *Schistosoma mansoni*. *Trop Med Int Health* 1998; 3: 728–35

25. Engels D, Nahimana S, Gryseels B. Comparison of the direct faecal smear and two thick smear techniques for the diagnosis of intestinal parasitic infections. *Trans R Soc Trop Med Hyg* 1996; 90: 523–25

26. Feldmeier H. Diagnosis. In: Jordan P, Webbe G, Sturrock FS. Human schistosomiasis. Wallingford: CAB International, 1993: 271–303

27. Jordan P, Webbe G. Epidemiology. In: Jordan P, Webbe G, Sturrock FS. Human schistosomiasis. Wallingford: CAB International, 1993: 87–158

28. Fenwick A, Savioli L, Engels D, Robert BN, Todd MH. Drugs for the control of parasitic diseases: current status and development in schistosomiasis. *Trends Parasitol* 2003; 19: 509–15

29. Renganathan E, Cioli D. An international initiative on praziquantel use *Parasitol Today* 1998; 14: 390–91

30. WHO Expert Committee. Prevention and control of schistosomiasis and soil-transmitted helminthiasis. Technical report series. Geneva: World Health Organisation, 2002.

31. Jordan P, Webbe G. Epidemiology. In: Jordan P, Webbe G, Sturrock FS. Human schistosomiasis. Wallingford: CAB International, 1993: 87–158

32. Dayan AD. Albendazole, mebendazole and praziquantel: review of non-clinical toxicity and pharmacokinetics. *Acta Trop* 2003; 86: 141–59

33. Renganathan E, Cioli D. An international initiative on praziquantel use *Parasitol Today* 1998; 14: 390–91

34. Davis A. Antischistosomal drugs and clinical practice. In: Jordan P, Webbe G, Sturrock FS. Human schistosomiasis. Wallingford: CAB International, 1993: 367–404.

35. Botros S, Sayed H, Amer N, El Ghannam M, Bennett JL, Day TA. Current status of sensitivity to praziquantel in a focus of potential drug resistance in Egypt. *Int J Parasitol* 2005; 35: 787–91

36. Mountford AP. Immunological aspects of schistosomiasis. *Parasite Immunol* 2005; 27: 243–46

37. Gryseels B. Human resistance to schistosoma Infections: age or experience? *Parasitol Today* 1994; 10: 380–84

38. Butterworth AE. Immunology of schistosomiasis. In: Jordan P, Webbe G, Sturrock FS. Human schistosomiasis. Wallingford: CAB International, 1993: 331–66

39. Capron A, Riveau G, Capron M, Trottein F. Schistosomes: the road from host-parasite interactions to vaccines in clinical trials. Trends Parasitol 2005; 21:143–49

40. Cowie RH. Can snails ever be effective and safe biocontrol agents? Int J Pest Manag 2001; 47:23–40

41. Minai M, Hosaka Y, Ohta N. Historical view of schistosomiasis japonica in Japan: implementation and evaluation of disease-control strategies in Yamanashi Prefecture. Paras Int 2003; 52: 321–26

42. Engels D, Ndoricimpa J, Gryseels B. Schistosomiasis *mansoni* in Burundi: progress in its control since 1985. Bull World Health Organ 1993; 71: 207–14

43. King CH, Bertsch D: Historical perspective: snail control to prevent schistosomiasis. *PLoS Negl Trop Dis* 2015, 9(4):e0003657

44. Knopp S, Mohammed KA, Ali SM, Khamis IS, Ame SM, Albonico M *et al*: Study and implementation of urogenital schistosomiasis elimination in Zanzibar (Unguja and Pemba islands) using an integrated multidisciplinary approach. *BMC Public Health* 2012, 12:930

45. Salam RA, Maredia H, Das JK, Lassi ZS, Bhutta ZA: Community-based interventions for the prevention and control of helmintic neglected tropical diseases. *Infect Dis Poverty* 2014, 3:23

46. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C *et al*: Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012, 380(9859):2197-2223

47. Mwinzi PN, Muchiri G, Wiegand RE, et al.: Predictive Value of School-Aged Children's Schistosomiasis Prevalence and Egg Intensity for Other Age Groups in Western Kenya. *Am J Trop Med Hyg.* 2015; pii: 15–0467. 10.4269/ajtmh.15-0467

48. Odiere M, Rawago F, Ombok M, Secor W, Karanja D, Mwinzi P, Lammie P, Won K: High prevalence of schistosomiasis in Mbita and its adjacent islands of Lake Victoria, western Kenya. Parasit Vectors. 2012, 5: 278-10.1186/1756-3305-5-278

49. Sang CS, Muchiri G, Ombok M, Odiere MR, Mwinzi P NM, 2014. *Schistosoma haematobium* hotspots in south Nyanza, western Kenya: prevalence, distribution and co-endemicity with *Schistosoma mansoni* and soil-transmitted helminths. Parasites and vectors 2014 7:125

50. Katz N, Chaves A, Pellegrino J: A simple device for quantitative stool thick-smear technique in schistosomiasis *mansoni*. *Rev Inst Med Trop Sao Paulo* 1972, 14(6):397-400

51. On the Variances of Asymptotically Normal Estimators from Complex Surveys, David A. Binder, International Statistical Review Vol. 51, No. 3 (Dec., 1983), pp. 279-292

52. Kung-Yee Liang and Scott L. Zeger *Longitudinal Data Analysis Using Generalized Linear Models*, Biometrika Vol. 73, No. 1 (Apr., 1986), pp. 13-22

53. Pan, W. (2001), Akaike's Information Criterion in Generalized Estimating Equations. Biometrics, 57: 120–125

54. Searle SR, Speed FM, Milliken GA (1980). "Population marginal means in the linear model: An alternative to least squares means." The American Statistician, 34(4), 216–221

55. Kulldorff M. and Information Management Services, Inc. SaTScanTM v8.0: Software for the spatial and space-time scan statistics. http://www.satscan.org/, 2009

56. Kulldorff M, Huang L, Pickle L, Duczmal L. An elliptic spatial scan statistic. Statistics in Medicine, 2006, 25:3929-3943

57. French MD, Churcher TS, Gambhir M, et al.: Observed reductions in *Schistosoma mansoni* transmission from large-scale administration of praziquantel in Uganda: a

mathematical modelling study. *PLoS Negl Trop Dis.* 2010;4(11):e897. 10.1371/journal.pntd.000089

58. Kloos H (1995) Human behavior, health education and schistosomiasis control: a review. Soc Sci Med 40: 1497–1511

59. Odongo-Aginya EI, Kironde F, Lyazi M, Sempewo H, Correa Oliveira R. Evidence of Long Term Benefit of Morbidity Reduction Due to Praziquantel Treatment Against *Schistosoma Mansoni* in Kigungu Fishing Village in Entebbe, Uganda. *African Journal of Infectious Diseases*. 2011;5(2):33-39

60. El Kholy H, Arap Siongok TK, Koech D, Sturrock RF, Houser H, King CH, et al. Effects of bore hole wells on water utilization in *Schistosoma haematobium* endemic communities in Coast province, Kenya. Am J Trop Med Hyg. 1989;41:212–9

61. Secor, WE. "Early Lessons from Schistosomiasis Mass Drug Administration Programs." *F1000Research* 4 (2015): F1000 Faculty Rev–1157. *PMC*. Web. 11 Apr. 2016

62. Liang S., YC. Zhong B., and Qiu D. 2006. Re‐emerging schistosomiasis in hilly and mountainous areas of Sichuan, China. Bulletin of the World Health Organization 84:139–144

63. Cioli D, Pica-Mattoccia L, Basso A, et al.: Schistosomiasis control: praziquantel forever? *Mol Biochem Parasitol.* 2014; 195(1):23–9. 10.1016/j.molbiopara.2014.06.002

64. Savaya Alkalay, Amit et al. "The Prawn *Macrobrachium Vollenhovenii* in the Senegal River Basin: Towards Sustainable Restocking of All-Male Populations for Biological Control of Schistosomiasis." Ed. Matty Knight. *PLoS Neglected Tropical Diseases* 8.8 (2014): e3060. *PMC*. Web. 13 Apr. 2016.

65. Sokolow, SH., Lafferty KD, and Kuris AM. "Regulation of Laboratory Populations of Snails (*Biomphalaria* and *Bulinus* Spp.) by River Prawns, *Macrobrachium* Spp. (Decapoda, Palaemonidae): Implications for Control of Schistosomiasis." *Acta tropica* 132 (2014): 64–74. *PMC*. Web. 13 Apr. 2016.

66. Life Cycle of Schistosomiais, visited 3/3/16. http://www.cdc.gov/dpdx/schistosomiasis/

67. Geographic distribution of Schistosomiasis visited 3/3/16. http://wwwnc.cdc.gov/travel/yellowbook/2016/infectious-diseases-related-to-travel/schistosomiasis

Tables and Figures

Table 1. The mean change in prevalence of *S. mansoni* in individuals based on village survey data (all years aggregated), years, and arms around Lake Victoria 2011-2015.

Village Data	Response	N	Mean	Std Dev	P-value
Malaria	No	367	64.19	84.95	0.04
Intervention	Yes	228	50.02	78.95	
Immunization	No	202	78.604	88.35	<.0001
Intervention	Yes	393	48.56	78.16	
NTD	No	447	64.55	86.81	0.0010
Intervention	Yes	148	41.27	67.15	
HIV	No	391	61.04	83.65	0.39
Intervention	Yes	204	54.39	81.58	
Flood	No	502	62.65	85.32	0.003
	Yes	93	37.74	64.98	
Drought	No	260	67.79	86.71	0.014
	Yes	335	51.74	79.28	
River	Always/Often	460	66.42	89.16	0.0004
Drinking	Sometimes	111	34.89	51.13	
	Never	23	22.203	34.04	
River	Always	308	79.801	97.51	<.0001
Washing	Often	211	38.89	61.79	
	Sometimes/Never	75	28.74	31.72	
River	Always/Often	512	63.64	87.45	0.003
Bathing	Sometimes/Never	82	28.71	33.13	
Rain	Always/Often	41	50.99	74.24	0.78
Drinking	Sometimes	534	59.09	84.28	
	Never	20	65.71	62.83	
Rain	Always/Often	36	36.81	67.79	0.0022
Washing	Sometimes	460	55.16	77.25	
	Never	99	83.46	105.98	
Rain	Always/Often	33	37.69	70.03	0.012
Bathing	Sometimes	424	54.86	74.86	
	Never	138	75.76	104.49	
Tap Water	Always/Often	119	34.48	44.57	0.0007
Drinking	Sometimes	142	57.48	75.97	
	Never	334	67.95	93.92	
Tap Water	Always/Often	80	34.05	48.03	0.0010
Washing	Sometimes	132	47.97	63.54	
	Never	383	67.64	92.65	
Tap Water	Always/Often	67	35.16	49.09	0.005
Bathing	Sometimes	133	48.74	62.62	
	Never	395	66.14	92.01	
Sand Pit	Always	59	55.68	69.86	0.001
	Often	266	46.19	72.71	
	Sometimes	257	73.74	94.30	
	Never	13	33.73	35.65	
Sand Toilet	Often	10	92.48	100.25	0.23
	Sometimes	51	45.52	61.43	

	Never	529	59.32	84.43	
Year	1	150	09.05	95.82	<0.0001
	2	125	66.02	90.33	
	3	75	68.15	69.91	
	4	75	36.12	50.61	
	5	150	36.03	69.24	
Arm	1	125	73.86	97.15	0.4889
	2	125	48.82	86.03	
	3	75	68.15	69.91	
	4	125	43.69	57.28	
	5	75	60.71	91.82	
	6	75	63.74	85.11	

Table 2. Odds ratios comparing village level events, geographic area, and distance to Lake Victoria, Kenya, 2011-2105.

Parameters	OR	Lower CL	Upper CL	p
Distance to Lake Victoria (km)	0.70	0.59	0.83	<.0001
Geographic areas at specific years Compared to Bondo				
Year 1				
Rarieda	0.80	0.46	1.407	0.43
Homa Bay/Rachuonyo	0.38	0.23	0.63	0.0002
Kisumu/Nyakach	0.45	0.26	0.77	0.0037
Year 2				
Rarieda	0.75	0.43	1.30	0.3036
Homa Bay/Rachuonyo	0.31	0.19	0.51	<.0001
Kisumu/Nyakach	0.40	0.23	0.69	0.0009
Year 3				
Rarieda	0.84	0.49	1.45	0.5357
Homa Bay/Rachuonyo	0.21	0.13	0.32	<.0001
Kisumu/Nyakach	0.40	0.23	0.70	0.0014
Year 4				
Rarieda	0.72	0.42	1.25	0.2459
Homa Bay/ Rachuonyo	0.18	0.11	0.31	<.0001
Kisumu/Nyakach	0.31	0.18	0.53	<.0001
Year 5				
Rarieda	0.89	0.54	1.48	0.0002
Homa Bay/ Rachuonyo	0.14	0.09	0.22	0.5887
Kisumu/Nyakach	0.38	0.22	0.63	<.0001

Table 3. Odds ratios comparing village level events, occupations, and individual behaviors between villages[1] within the high prevalence cluster to villages not within.

Risk Factors	Odds Ratio	Lower	Upper	P Value
Village-level event				
Flood	1.32	0.46	3.77	0.61
Drought	0.15	0.03	0.74	0.012
HIV	0.79	0.27	2.31	0.67
Immunization	1.35	1.03	1.77	0.03
Malaria	1.64	0.55	4.89	0.38
NTD	1.02	0.37	2.79	0.98
Occupations				
Irrigation	1.73	0.64	4.71	0.28
Subsistence Farming	0.21	0.08	0.61	0.004
Fishing	1.33	0.47	3.75	0.59
Car Washing	0.53	0.13	2.11	0.37
Individual Behaviors				
Using Well or Bore hole for:				
Drinking	4.28	1.022	17.901	0.047
Using Tap water for:				
Tap Bathing	2.84	0.23	34.52	0.41
Tap Washing	1.30	0.10	16.67	0.84
Tap Drinking	0.56	0.15	2.13	0.39
Using Rain Water for:				
Rain Bathing	0.16	0.01	2.12	0.17
Rain Washing	0.81	0.07	9.5	0.87
Rain Drinking	2.36	0.21	25.82	0.48
Using River Water for:				
River Bathing	0.95	0.17	5.37	0.95
River Washing	1.61	0.25	10.35	0.62
River Drinking	1.63	0.44	5.99	0.46

[1] SCORE study villages, Lake Victoria 2011-2015

Figure 1. Life stages of Schistosomiasis [66].

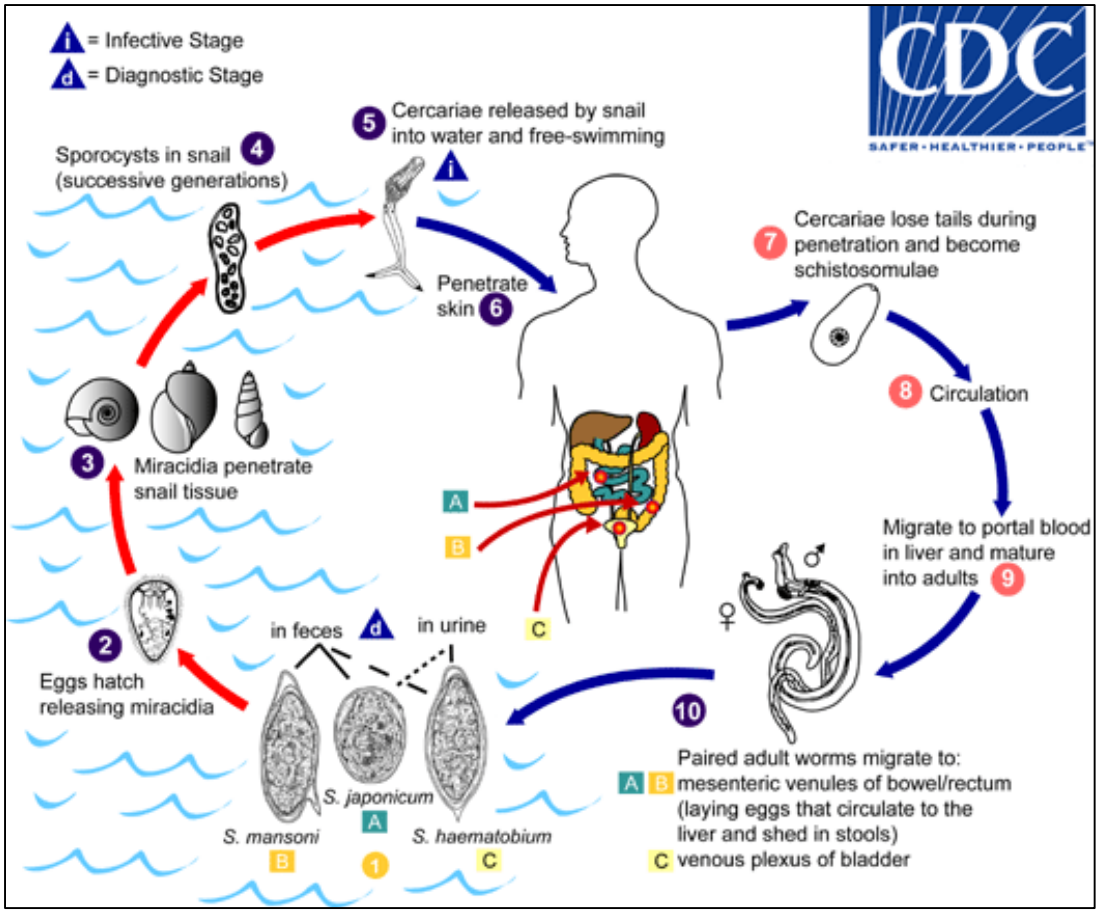


Figure 2. Geographical distribution of Schistosomiasis, CDC [67].

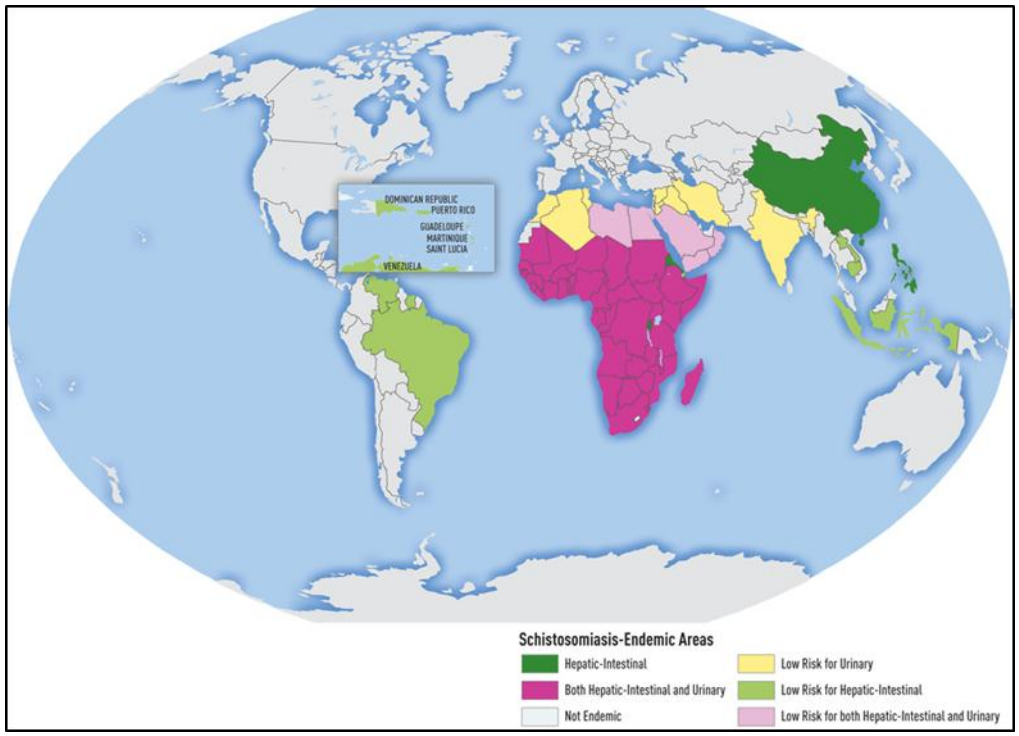


Figure 3. Mean eggs per gram per year over the study period for SCORE.

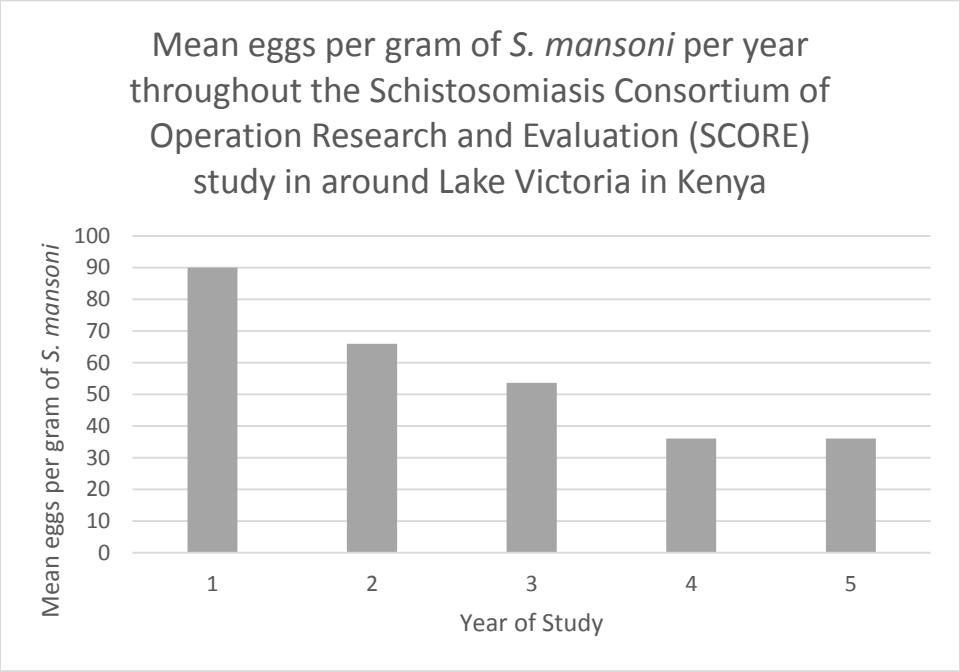


Figure 4. Point distribution of *S. mansoni* prevalence per village. High prevalence clusters are encircled by red. SM2 refers to villages with the SCORE study.

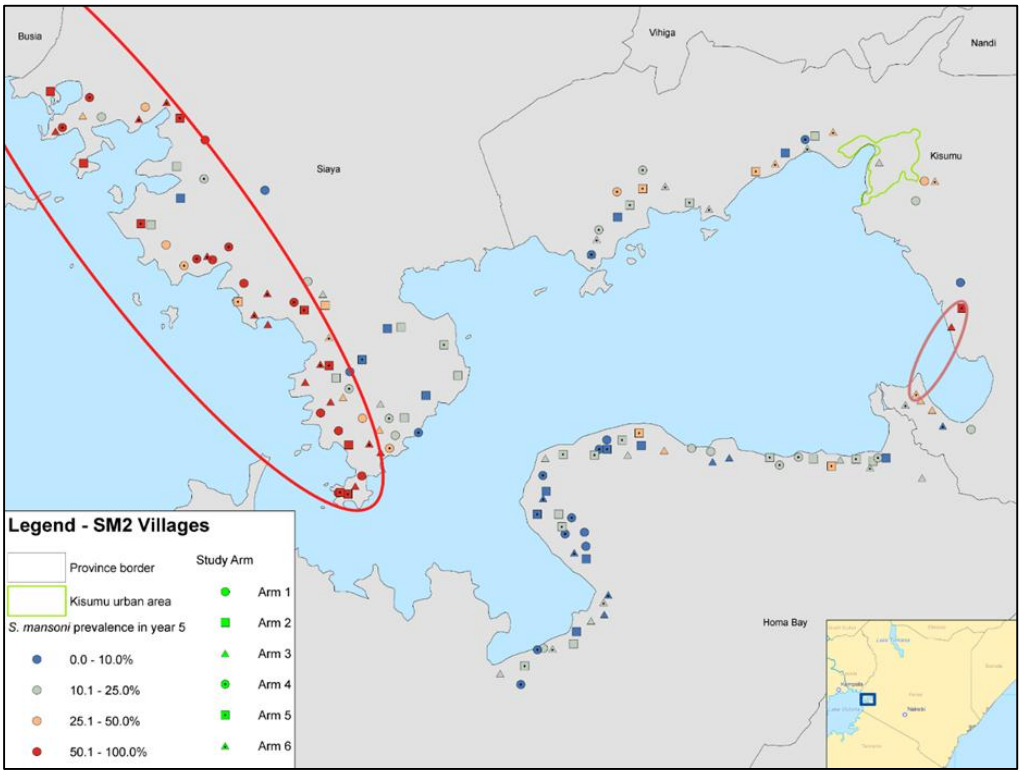
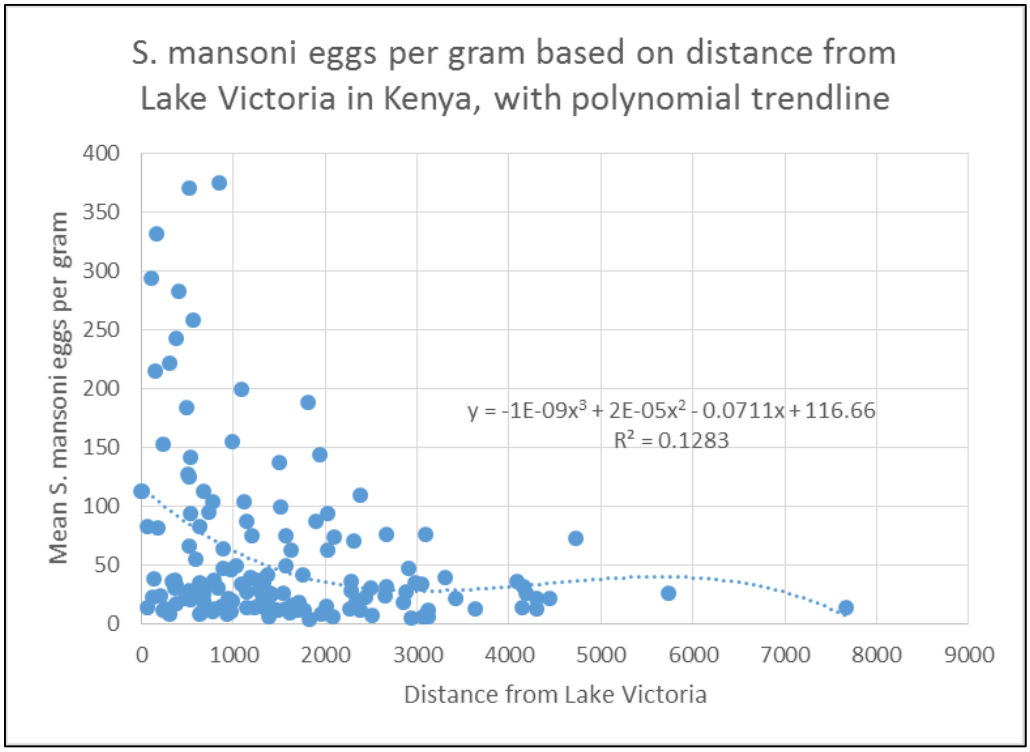


Figure 5 Mean *S. mansoni* prevalence based on distance to Lake Victoria, per village of the study.



Appendix 1: SCORE Project Year 2 survey

A. General Information

1. Village name: _____

2. Village ID (from phone):_____

3. Names of people providing information:

a._____
b._____
c._____

4. Has there been a major change in the population of the village: (Major increase, major decrease, no major change)

If yes, please provide the reasons for this: _____

5. Number of households: _____

What is the basis for the estimates of village population and households (e.g., national census, demographic surveillance system, estimate)?

_____

6. Major climate–related events (Check all that apply)

- □ Drought
- □ Flood
- □ Other (Please specify): _____

7. Health Campaigns in past year (Check all that apply, and if any treatment regimens provided, please specify)

- □ HIV/AIDS and/or TB _____
- □ Malaria _____
- □ Immunization program/campaign _____
- □ NTDs (Please specify - - LF, STH, etc.):_____
- □ Other (Please specify): _____

8. Other changes (e.g., political changes, total community-led sanctions or other sanction efforts), etc., that might affect the study outcomes?

B. Occupations

1. What are the main occupations of the people in this village that are closely related to water contact? (Some examples might be: fishing, car washing, rice or other irrigation-based farming.)

a._____
b._____
c._____
d._____

C. Local Health Facility

1. Longitude _ _._ _ _ _ _ _ degrees (dd.dddddd) and Latitude _ _._ _ _ _ _ _ degrees (dd.dddddd)

2. Is the health facility open regularly? (Yes / No)

3. Do they dispense praziquantel? (Yes / No)
   If yes, please ask to see if there is unexpired praziquantel in the facility and record whether there is any unexpired drug. (Yes / No)

4. Do they dispense ACT? (Yes / No)
   If yes, please ask to see if there is unexpired ACT in the facility and record if there is any unexpired drug. (Yes / No)

E. Water Sources

1. Indicate the use of the following water sources by households in this village for drinking, washing dishes or clothes, or bathing. Do they use them?
   - All the time (code as "1")
   - Often (code as "2")
   - Sometimes (code as "3')
   - Never (code as "4")

	River, pond, lake, or dam	Rainwater Tank	Wells or bore hole	Tap water	Other (please specify)
Drinking					
Washing dishes or clothes					
Bathing					

F. Sanitation

1. Indicate the use of the following options for defecation in this villager. Do they use them?
   - All the time (code as "1")
   - Often (code as "2")
   - Sometimes (code as "3')
   - Never (code as "4")

a. Bush or field _____
b. Pit latrine _____
c. VIP latrine _____
d. Toilet _____
e. Other _____

Recorder's Name_____Date (DD/MM/YY) _/___/___

Appendix 2: SAS CODE

DATA EXPLORATION/ TABLE 1 FORMATION:

```
LIBNAME h 'H:\Practicum';
RUN;
***Collapse inventory data***;

data h.village_level_data_sm2_20160201a;
        set h.village_level_data_sm2_20160201;

        If riverdrk     = '1'   then riverdrk    = '2';
        If riverbath    = '1'   then riverbath   = '2';
        If rainbath     = '1'   then rainbath    = '2';
        If raindrk              = '1'   then raindrk     = '2';
        If rainwash     = '1'   then rainwash    = '2';
        If tapdrk               = '1'   then tapdrk      = '2';
        If tapwash              = '1'   then tapwash     = '2';
        If tapbath              = '1'  then tapbath      = '2';
        If SanVip               = '1'   then SanVip      = '2';
        If tapwash              = '1'   then tapwash   = '2';

        If riverwash    = '4'   then riverwash =  '3';
        If riverbath    = '4'   then riverbath =  '3';


run;


***PROC MEANS***;

proc means data= h.village_level_data_sm2_20160201a mean std;
class year;
var sm_epg;
run;

proc means data= h.village_level_data_sm2_20160201a mean std;
class d_victoria;
var sm_epg;
run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class malaria;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class imm;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class ntd;
var sm_epg;
Run;
```

```
proc means data=h.village_level_data_sm2_20160201a mean std;
class hiv;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class riverdrk;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class riverwash;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class riverbath;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class raindrk;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class rainwash;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class rainbath;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class tapdrk;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class tapwash;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class tapbath;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class sanpit;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
```

```
class sanvipt;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class santoilet;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class drought;
var sm_epg;
Run;

proc means data=h.village_level_data_sm2_20160201a mean std;
class flood;
var sm_epg;
Run;


***PROC TTEST, independence***;

proc ttest data=h.village_level_data_sm2_20160201a;
class hiv imm malaria ntd riverDrk riverwash riverbath raindrk rainwash rainbath tapdrk tapwash tapbath otherdrk
otherwash otherbath sanbush sanpit sanvip santoilets;
var sm_epg;
Run;

proc ttest data=h.village_level_data_sm2_20160201a alpha=.05;
class malaria;
var sm_epg;
Run;

proc ttest data=h.village_level_data_sm2_20160201a alpha=.05;
class imm;
var sm_epg;
Run;

proc ttest data=h.village_level_data_sm2_20160201a alpha=.05;
class ntd;
var sm_epg;
Run;

proc ttest data=h.village_level_data_sm2_20160201a alpha=.05;
class hiv;
var sm_epg;
Run;

proc ttest data=h.village_level_data_sm2_20160201a alpha=.05;
class flood;
var sm_epg;
Run;

proc ttest data=h.village_level_data_sm2_20160201a alpha=.05;
class drought;
var sm_epg;
```

```
Run;

***PROC SURVEYREG, repeated values***;

proc surveyreg data=h.village_level_data_sm2_20160201a alpha=.05;
class malaria ntd imm hiv flood drought;
cluster village_ID;
model sm_epg= malaria ntd imm hiv flood drought;
Run;

proc surveyreg data=h.village_level_data_sm2_20160201a alpha=.05;
class year;
cluster village_ID;
model sm_epg= year;
Run;

proc surveyreg data=h.village_level_data_sm2_20160201a alpha=.05;
class arm;
cluster village_ID;
model sm_epg= arm;
Run;

proc surveyreg data=h.village_level_data_sm2_20160201a alpha=.05;
class raindrk;
cluster village_ID;
model sm_epg= raindrk;
Run;
***PROC GLM, multiple levels of variable***;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class riverdrk;
model sm_epg =riverdrk;
run;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class riverbath;
model sm_epg =riverbath;
run;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class riverwash;
model sm_epg =riverwash;
run;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class raindrk;
model sm_epg =raindrk;
run;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class rainwash;
model sm_epg =rainwash;
run;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class rainbath;
```

```
model sm_epg =rainbath;
run;


PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class tapdrk;
model sm_epg =tapdrk;
run;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class tapwash;
model sm_epg =tapwash;
run;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class tapbath;
model sm_epg =tapbath;
run;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class sanpit;
model sm_epg =sanpit;
run;

PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class santoilet;
model sm_epg =santoilet;
run;


PROC GLM data=h.village_level_data_sm2_20160201a alpha=.05;
class riverdrk riverwash riverbath raindrk rainwash rainbath tapdrk tapwash tapbath sanpit santoilet;
model sm_epg = riverdrk riverwash riverbath raindrk rainwash rainbath tapdrk tapwash tapbath sanpit santoilet;
run;
```

MODEL TESTING

```
LIBNAME h 'H:\Practicum';
RUN;

data dat;
        set h.Dat;
        sm_pos = round(sm * sm_N);
run;

***MODEL TESTING USING RYAN'S QIC/QICu MACRO IN PROC GENMOD***;

proc genmod data=work.dat;***model 1***;
        class Village_ID year(ref="1") arm drought(ref="1") HIV(ref="1") IMM(ref="1") malaria(ref="1") geo_gp
/ param=ref;
        model sm_pos / sm_N = year arm d_victoria_km drought HIV IMM malaria Santoilet sanpit sanbush
geo_gp tapbath tapwash tapdrk rainbath rainwash raindrk riverbath riverwash riverdrk/ dist=b link=logit;
        repeated subject=Village_ID / type=cs;
run;
```

```
proc genmod data=work.dat;***model 2 drop: Santoilet riverdrk riverbath raindrk***;
	class Village_ID year(ref="1") arm drought(ref="1") HIV(ref="1") IMM(ref="1") malaria(ref="1") geo_gp
/ param=ref;
	model sm_pos / sm_N = year arm d_victoria_km drought HIV IMM malaria sanpit sanbush  tapbath
tapwash tapdrk rainbath rainwash   riverwash / dist=b link=logit;
	repeated subject=Village_ID / type=cs;
run;

proc genmod data=work.dat;***model 3 drop: Santoilet riverdrk riverbath raindrk sanpit tapdrk rainbath***;
	class Village_ID year(ref="1") arm drought(ref="1") HIV(ref="1") IMM(ref="1") malaria(ref="1") geo_gp
/ param=ref;
	model sm_pos / sm_N = year arm d_victoria_km drought HIV IMM malaria  sanbush geo_gp tapbath
tapwash  rainwash   riverwash / dist=b link=logit;
	repeated subject=Village_ID / type=cs;
run;

proc genmod data=work.dat;***model 4 drop: Santoilet riverdrk riverbath raindrk sanpit tapdrk rainbath
drought***;
	class Village_ID year(ref="1") arm HIV(ref="1") IMM(ref="1") malaria(ref="1") geo_gp tapbath/
param=ref;
	model sm_pos / sm_N = year arm d_victoria_km  HIV IMM malaria  sanbush geo_gp  tapwash  rainwash
riverwash / dist=b link=logit;
	repeated subject=Village_ID / type=cs;
run;

proc genmod data=work.dat;***model 5 drop: Santoilet riverdrk riverbath raindrk sanpit tapdrk rainbath drought
malaria  tapwash rainwash***;
	class Village_ID year(ref="1") arm HIV(ref="1") IMM(ref="1") geo_gp tapbath/ param=ref;
	model sm_pos / sm_N = year arm d_victoria_km  HIV IMM sanbush geo_gp riverwash / dist=b link=logit;
	repeated subject=Village_ID / type=cs;
run;

proc genmod data=work.dat;***model 6 drop: Santoilet riverdrk riverbath raindrk sanpit tapdrk rainbath drought
malaria  tapwash rainwash HIV IMM riverwash***;
	class Village_ID year(ref="1") arm geo_gp tapbath  / param=ref;
	model sm_pos / sm_N = year arm d_victoria_km sanbush geo_gp  / dist=b link=logit;
	repeated subject=Village_ID / type=cs;
run;

proc genmod data=work.dat;***model 7 Interaction ***;
	class Village_ID year(ref="1") arm HIV(ref="1") malaria(ref="1") NTD(ref="1")  / param=ref;
	model sm_pos / sm_N = year arm d_victoria_km HIV malaria NTD year*arm year*d_victoria_km
arm*d_victoria_km d_victoria_km*HIV d_victoria_km*malaria d_victoria_km*NTD / dist=b link=logit;
	repeated subject=Village_ID / type=cs;
run;

***MACRO FOR VARIABLE TESTING***;
%let vldfldr= H:\Practicum\;
libname h "&vldfldr.";


data dat;
	set h.village_level_data_20151215;
	sm_pos = round(sm * sm_N);
run;
```

```
proc means data=dat;
        class year;
        var sm;
run;

proc genmod data=dat;
        class Village_ID year(ref="1") arm / param=ref;
        model sm_pos / sm_N = year arm / dist=b link=logit;
        repeated subject=Village_ID / type=cs;
        estimate "year 2 vs. year 1" year 1 0 0 / exp;
        estimate "year 3 vs. year 1" year 0 1 0 / exp;
        estimate "year 4 vs. year 1" year 0 0 1 / exp;
run;


/****variable selection****/
%inc"H:\Practicum\variable.selection.macros.sas";

/*put all class variables here*/
/*%let classvars = year arm geo_gp hiv malaria imm ntd drought ;*/

/*put all non class variables here*/
/*%let othervars = d_victoria_km well_bore sanpit sanbush tapwash tapdrk tapbath rainbath rainwash riverwash ;*/

/*%let classvars = year arm geo_gp hiv malaria imm ntd ;*/

/*tested: geo_gp sanbush sanpit*/
%let classvars = year arm hiv malaria imm ntd ;

/*put all non class variables here*/
%let othervars = d_victoria_km well_bore  ;

***;
/*tested: geo_gp sanbush sanpit*/
%let classvars = year arm hiv malaria imm ntd geo_gp;

/*put all non class variables here*/
%let othervars = d_victoria_km well_bore tapbath;


/*creates list of all possible models*/
%modlist;


/*use this step if limiting models*/
/*for example, i have limited to only those models with distance to lake, year, and arm*/
data modlist;
        set modlist;
/*      if index(mod, "x1") > 0 and index(VariablesInModel, "x1a") > 0 then delete;*/
        if index(mod, "d_victoria_km") = 0 then delete;
        if index(mod, "year") = 0 then delete;
        if index(mod, "arm") = 0 then delete;
run;
```

```
/*run macro which computes QIC and QICu*/

/*first give each model a unique id from 1 to the number of models*/
data modlist;
	set modlist;
	id=_N_;
run;

%bsmod(modlist);


data h.modlist;
	set modlist;
run;


/*find difference from minimums*/
proc means data=modlist min;
	var qic qicu;
	ods output summary=icmin(keep = qic_min qicu_min);
run;

data icmin;
	set icmin;
	dummy=1;
	label qic_min= qicu_min=;
run;

data modlist;
	set modlist;
	dummy=1;
run;

data modlist(drop=dummy);
	merge modlist icmin;
	by dummy;
	_qic=qic-qic_min;
	_qicu=qicu-qicu_min;
run;


/*sort by lowest qic*/
proc sort data=modlist;
	by _qic;
run;

/*sort by lowest qicu*/
proc sort data=modlist;
	by _qicu;
run;


***residuals and outliers, deletion diagnostics***;
ods graphics on;
proc genmod data=dat plots=all;
	class Village_ID &classvars. geo_gp / param=ref;
```

```
        model sm_pos / sm_N = year arm hiv ntd d_victoria_km geo_gp / dist=b link=logit;
        repeated subject=Village_ID / type=cs;
/*      ods select GEEFitCriteria;*/
run;
ods graphics off;

proc genmod data=dat;
        class Village_ID &classvars. geo_gp / param=ref;
        model sm_pos / sm_N = year arm ntd d_victoria_km geo_gp / dist=b link=logit;
        repeated subject=Village_ID / type=cs;
        ods select GEEFitCriteria;
run;
```

CONTRAST/ESTIMATE STATEMENTS

```
***CONTRAST STATEMENTS COMPARING YEAR and ARM INTEREACTION***;

LIBNAME h 'H:\Practicum';
RUN;


data data;
        set h.village_level_data_sm2_20160201;
        sm_pos = round(sm * sm_N);
run;

ods trace off;
ods rtf file="H:\Practicum\model1.rtf";
title1 "model OUTPUT TO with contrast";
proc genmod data=data;
        class Village_ID year(ref="1") arm(ref="1");
        model  sm_pos/Sm_N = year|arm d_victoria_km year*d_victoria_km / dist=b link=logit;
        repeated subject=Village_ID / type=cs;
        ods output GEEEmpPEst=pe estimates=e;

*prevalence estimates*;
        estimate "year 1, arm 2"    intercept 1 year 1 0 0 0 0 arm 1 0 0 0 0 0 year*arm 0 1 0 0 0 0  0 0 0 0 0  0 0 0 0
0 0 0  0 0 0 0 0 0;
estimate "year 1, arm 2"    intercept 1 year 1 0 0 0 0 arm 0 1 0 0 0 0 year*arm 0 1 0 0 0 0  0 0 0 0 0  0 0 0 0  0 0 0  0
0 0 0 0 0;
        estimate "year 1, arm 3"    intercept 1 year 1 0 0 0 0 arm 0 0 1 0 0 0 year*arm 0 0 1 0 0 0  0 0 0 0 0  0 0 0 0
0 0 0  0 0 0 0 0 0;
        estimate "year 1, arm 4"    intercept 1 year 1 0 0 0 0 arm 0 0 0 1 0 0 year*arm 0 0 0 1 0 0  0 0 0 0 0  0 0 0 0
0 0 0  0 0 0 0 0 0;
        estimate "year 1, arm 5"    intercept 1 year 1 0 0 0 0 arm 0 0 0 0 1 0 year*arm 0 0 0 0 1 0  0 0 0 0 0  0 0 0 0
0 0 0  0 0 0 0 0 0;
        estimate "year 1, arm 6"    intercept 1 year 1 0 0 0 0 arm 0 0 0 0 0 1 year*arm 0 0 0 0 0 1  0 0 0 0 0  0 0 0 0
0 0 0  0 0 0 0 0 0;

        estimate "year 2, arm 1"    intercept 1 year 0 1 0 0 0 arm 1 0 0 0 0 0 year*arm 0 0 0 0 0 0  1 0 0 0 0  0 0 0 0
0 0 0  0 0 0 0 0 0;
        estimate "year 2, arm 2"    intercept 1 year 0 1 0 0 0 arm 0 1 0 0 0 0 year*arm 0 0 0 0 0 0  0 1 0 0 0  0 0 0 0
0 0 0  0 0 0 0 0 0;
        estimate "year 2, arm 3"    intercept 1 year 0 1 0 0 0 arm 0 0 1 0 0 0 year*arm 0 0 0 0 0 0  0 0 1 0 0  0 0 0 0
0 0 0  0 0 0 0 0 0;
```

```
        estimate "year 2, arm 4"    intercept 1 year 0 1 0 0 0 arm 0 0 0 1 0 0 year*arm 0 0 0 0 0 0  0 0 0 1 0  0 0 0 0
0 0 0  0 0 0 0 0 0;
        estimate "year 2, arm 5"    intercept 1 year 0 1 0 0 0 arm 0 0 0 0 1 0 year*arm 0 0 0 0 0 0  0 0 0 0 1  0 0 0 0
0 0 0  0 0 0 0 0 0;

        estimate "year 3, arm 1"    intercept 1 year 0 0 1 0 0 arm 1 0 0 0 0 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  1 0 0 0
0 0 0  0 0 0 0 0 0;
        estimate "year 3, arm 2"    intercept 1 year 0 0 1 0 0 arm 0 1 0 0 0 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 1 0 0
0 0 0  0 0 0 0 0 0;
        estimate "year 3, arm 4"    intercept 1 year 0 0 1 0 0 arm 0 0 0 1 0 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 1 0
0 0 0  0 0 0 0 0 0;
        estimate "year 3, arm 6"    intercept 1 year 0 0 1 0 0 arm 0 0 0 0 0 1 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 1
0 0 0  0 0 0 0 0 0;

   estimate "year 4  arm 2"                 intercept 1 year 0 0 0 1 0 arm 0 1 0 0 0 0 year*arm 0 0 0 0 0 0  0 0 0 0 0
0 0 0 0  0 1 0  0 0 0 0 0 0;
        estimate "year 4, arm 4"    intercept 1 year 0 0 0 1 0 arm 0 0 0 1 0 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 0
0 0 1  0 0 0 0 0 0;

   estimate "year 5, arm 1"                 intercept 1 year 0 0 0 0 1 arm 1 0 0 0 0 0 year*arm 0 0 0 0 0 0  0 0 0 0 0
0 0 0 0  0 0 0  1 0 0 0 0 0;
        estimate "year 5, arm 2"    intercept 1 year 0 0 0 0 1 arm 0 1 0 0 0 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 0
0 0 0  0 1 0 0 0 0;
        estimate "year 5, arm 3"    intercept 1 year 0 0 0 0 1 arm 0 0 1 0 0 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 0
0 0 0  0 0 1 0 0 0;
        estimate "year 5, arm 4"    intercept 1 year 0 0 0 0 1 arm 0 0 0 1 0 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 0
0 0 0  0 0 0 1 0 0;
        estimate "year 5, arm 5"    intercept 1 year 0 0 0 0 1 arm 0 0 0 0 1 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 0
0 0 0  0 0 0 0 1 0;
        estimate "year 5, arm 6"    intercept 1 year 0 0 0 0 1 arm 0 0 0 0 0 1 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 0
0 0 0  0 0 0 0 0 1;

        run;

ods rtf close;

/*year to year comparisons, each arm*/
        estimate "year 2 vs. year 1 | arm 1" year -1 1 0 0 0 year*arm -1 0 0 0 0 0   1 0 0 0 0  0 0 0 0  0 0 0  0 0 0 0 0
0;
        estimate "year 2 vs. year 1 | arm 2" year -1 1 0 0 0 year*arm 0 -1 0 0 0 0   0 1 0 0 0  0 0 0 0  0 0 0  0 0 0 0 0
0;
        estimate "year 2 vs. year 1 | arm 3" year -1 1 0 0 0 year*arm 0 0 -1 0 0 0   0 0 1 0 0  0 0 0 0  0 0 0  0 0 0 0 0
0;
        estimate "year 2 vs. year 1 | arm 4" year -1 1 0 0 0 year*arm 0 0 0 -1 0 0   0 0 0 1 0  0 0 0 0  0 0 0  0 0 0 0 0
0;
        estimate "year 2 vs. year 1 | arm 5" year -1 1 0 0 0 year*arm 0 0 0 0 -1 0   0 0 0 0 1  0 0 0 0  0 0 0  0 0 0 0 0
0;

        estimate "year 2 vs. year 4 | arm 2" year 0 1 0 -1 0 year*arm 0 0 0 0 0 0   0 1 0 0 0  0 0 0 0  0 -1 0  0 0 0 0 0
0;
        estimate "year 2 vs. year 4 | arm 4" year 0 1 0 -1 0 year*arm 0 0 0 0 0 0   0 0 0 1 0  0 0 0 0  0 0 -1  0 0 0 0 0
0;

        estimate "year 2 vs. year 5 | arm 1" year 0 1 0 0 -1 year*arm 0 0 0 0 0 0   1 0 0 0 0  0 0 0 0  0 0 0  -1 0 0 0 0
0;
```

```
        estimate "year 2 vs. year 5 | arm 2" year 0 1 0 0 -1 year*arm 0 0 0 0 0 0  0 1 0 0 0  0 0 0 0  0 0 0  0 -1 0 0 0
0;
        estimate "year 2 vs. year 5 | arm 3" year 0 1 0 0 -1 year*arm 0 0 0 0 0 0  0 0 1 0 0  0 0 0 0  0 0 0  0 0 -1 0 0
0;
        estimate "year 2 vs. year 5 | arm 4" year 0 1 0 0 -1 year*arm 0 0 0 0 0 0  0 0 0 1 0  0 0 0 0  0 0 0  0 0 0 -1 0
0;
        estimate "year 2 vs. year 5 | arm 5" year 0 1 0 0 -1 year*arm 0 0 0 0 0 0  0 0 0 0 1  0 0 0 0  0 0 0  0 0 0 0 -1
0;

        estimate "year 3 vs. year 1 | arm 1" year -1 0 1 0 0 year*arm -1 0 0 0 0 0  0 0 0 0 0  1 0 0 0  0 0 0  0 0 0 0 0
0;
        estimate "year 3 vs. year 1 | arm 2" year -1 0 1 0 0 year*arm 0 -1 0 0 0 0  0 0 0 0 0  0 1 0 0  0 0 0  0 0 0 0 0
0;
        estimate "year 3 vs. year 1 | arm 4" year -1 0 1 0 0 year*arm 0 0 0 -1 0 0  0 0 0 0 0  0 0 1 0  0 0 0  0 0 0 0 0
0;
        estimate "year 3 vs. year 1 | arm 6" year -1 0 1 0 0 year*arm 0 0 0 0 0 -1  0 0 0 0 0  0 0 0 1  0 0 0  0 0 0 0 0
0;

        estimate "year 3 vs. year 2 | arm 1" year 0 -1 1 0 0 year*arm 0 0 0 0 0 0  -1 0 0 0 0  1 0 0 0  0 0 0  0 0 0 0 0
0;
        estimate "year 3 vs. year 2 | arm 2" year 0 -1 1 0 0 year*arm 0 0 0 0 0 0  0 -1 0 0 0  0 1 0 0  0 0 0  0 0 0 0 0
0;
        estimate "year 3 vs. year 2 | arm 4" year 0 -1 1 0 0 year*arm 0 0 0 0 0 0  0 0 0 -1 0  0 0 1 0  0 0 0  0 0 0 0 0
0;

        estimate "year 3 vs. year 4 | arm 2" year 0 0 1 -1 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 1 0 0  0 -1 0  0 0 0 0 0
0;
        estimate "year 3 vs. year 4 | arm 4" year 0 0 1 -1 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 1 0  0 0 -1  0 0 0 0 0
0;

        estimate "year 3 vs. year 5 | arm 1" year 0 0 1 0 -1 year*arm 0 0 0 0 0 0  0 0 0 0 0  1 0 0 0  0 0 0  -1 0 0 0 0
0;
        estimate "year 3 vs. year 5 | arm 2" year 0 0 1 0 -1 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 1 0 0  0 0 0  0 -1 0 0 0
0;
        estimate "year 3 vs. year 5 | arm 4" year 0 0 1 0 -1 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 1 0  0 0 0  0 0 0 -1 0
0;
        estimate "year 3 vs. year 5 | arm 6" year 0 0 1 0 -1 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 1  0 0 0  0 0 0 0 0 -
1;

        estimate "year 4 vs. year 1 | arm 2" year -1 0 0 1 0 year*arm 0 -1 0 0 0 0  0 0 0 0 0  0 0 0 0  0 1 0  0 0 0 0 0
0;
        estimate "year 4 vs. year 1 | arm 4" year -1 0 0 1 0 year*arm 0 0 0 -1 0 0  0 0 0 0 0  0 0 0 0  0 0 1  0 0 0 0 0
0;

        estimate "year 4 vs. year 5 | arm 2" year -1 0 0 1 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 0  0 1 0  0 -1 0 0 0
0;
        estimate "year 4 vs. year 5 | arm 4" year -1 0 0 1 0 year*arm 0 0 0 0 0 0  0 0 0 0 0  0 0 0 0  0 0 1  0 0 0 -1 0
0;

        estimate "year 5 vs. year 1 | arm 1" year -1 0 0 0 1 year*arm -1 0 0 0 0 0  0 0 0 0 0  0 0 0 0  0 0 0  1 0 0 0 0
0;
        estimate "year 5 vs. year 1 | arm 2" year -1 0 0 0 1 year*arm 0 -1 0 0 0 0  0 0 0 0 0  0 0 0 0  0 0 0  0 1 0 0 0
0;
        estimate "year 5 vs. year 1 | arm 3" year -1 0 0 0 1 year*arm 0 0 -1 0 0 0  0 0 0 0 0  0 0 0 0  0 0 0  0 0 1 0 0
0;
```

```
        estimate "year 5 vs. year 1 | arm 4" year -1 0 0 0 1 year*arm 0 0 0 -1 0 0  0 0 0 0 0  0 0 0 0  0 0 0  0 0 0 1 0
0;
        estimate "year 5 vs. year 1 | arm 5" year -1 0 0 0 1 year*arm 0 0 0 0 -1 0  0 0 0 0 0  0 0 0 0  0 0 0  0 0 0 0 1
0;
        estimate "year 5 vs. year 1 | arm 6" year -1 0 0 0 1 year*arm 0 0 0 0 0 -1  0 0 0 0 0  0 0 0 0  0 0 0  0 0 0 0 0
1;

        run;

***RYAN'S CODE TO USE LEAST MEANS SQUARE COMPARISONS***;

ods trace off;

proc genmod data=data;

    class Village_ID  year arm / param=glm;

    model sm = year|arm / dist=p link=log type3;

    repeated subject=Village_ID / type=cs;

    ods output LSMeans=gee_p_s_mi SliceTests=gee_p_s_sti SliceDiffs=gee_p_s_sdi LSMEstimates=gee_p_s_ei
GEEEmpPEst=gee_p_s_pei;



/*prevalence estimates*/

    lsmeans year*arm / cl;


/*year to year comparisons, each arm*/

    slice year*arm / sliceBy=arm diff cl;


/*arm comparisons at each time point*/

    slice year*arm / sliceBy=year diff cl;


/*arm difference of differences at each time point*/

    lsmestimate year*arm "arm 2 vs. arm 1 | change from year 1 at year 2" 1 -1 0 0 0 0  -1 1 0 0 0  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 2 vs. arm 1 | change from year 1 at year 3" 1 -1 0 0 0 0  0 0 0 0 0  -1 1 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 2 vs. arm 1 | change from year 1 at year 4" 1 -1 0 0 0 0  0 0 0 0 0  0 0 0 0  -1 1 0  0 0
0 0 0 0 / cl;
```

```
    lsmestimate year*arm "arm 2 vs. arm 1 | change from year 1 at year 5" 1 -1 0 0 0 0  0 0 0 0 0  0 0 0 0  0 0 0  -1 1
0 0 0 0 / cl;


    lsmestimate year*arm "arm 3 vs. arm 1 | change from year 1 at year 2" 1 0 -1 0 0 0  -1 0 1 0 0  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 3 vs. arm 1 | change from year 1 at year 5" 1 0 -1 0 0 0  0 0 0 0 0  0 0 0 0  0 0 0  -1 0
1 0 0 0 / cl;


    lsmestimate year*arm "arm 4 vs. arm 1 | change from year 1 at year 2" 1 0 0 -1 0 0  -1 0 0 1 0  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 4 vs. arm 1 | change from year 1 at year 3" 1 0 0 -1 0 0  0 0 0 0 0  -1 0 1 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 4 vs. arm 1 | change from year 1 at year 4" 1 0 0 -1 0 0  0 0 0 0 0  0 0 0 0  -1 0 1  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 4 vs. arm 1 | change from year 1 at year 5" 1 0 0 -1 0 0  0 0 0 0 0  0 0 0 0  0 0 0  -1 0
0 1 0 0 / cl;


    lsmestimate year*arm "arm 5 vs. arm 1 | change from year 1 at year 2" 1 0 0 0 -1 0  -1 0 0 0 1  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 5 vs. arm 1 | change from year 1 at year 5" 1 0 0 0 -1 0  0 0 0 0 0  0 0 0 0  0 0 0  -1 0
0 0 1 0 / cl;


    lsmestimate year*arm "arm 6 vs. arm 1 | change from year 1 at year 3" 1 0 0 0 0 -1  0 0 0 0 0  -1 0 0 1  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 6 vs. arm 1 | change from year 1 at year 5" 1 0 0 0 0 -1  0 0 0 0 0  0 0 0 0  0 0 0  -1 0
0 0 0 1 / cl;



    lsmestimate year*arm "arm 3 vs. arm 2 | change from year 1 at year 2" 0 1 -1 0 0 0  0 -1 1 0 0  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 3 vs. arm 2 | change from year 1 at year 5" 0 1 -1 0 0 0  0 0 0 0 0  0 0 0 0  0 0 0  0 -1
1 0 0 0 / cl;


    lsmestimate year*arm "arm 4 vs. arm 2 | change from year 1 at year 2" 0 1 0 -1 0 0  0 -1 0 1 0  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;
```

```
    lsmestimate year*arm "arm 4 vs. arm 2 | change from year 1 at year 3" 0 1 0 -1 0 0  0 0 0 0 0  0 -1 1 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 4 vs. arm 2 | change from year 1 at year 4" 0 1 0 -1 0 0  0 0 0 0 0  0 0 0 0  0 -1 1  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 4 vs. arm 2 | change from year 1 at year 5" 0 1 0 -1 0 0  0 0 0 0 0  0 0 0 0  0 0 0  0 -1
0 1 0 0 / cl;


    lsmestimate year*arm "arm 5 vs. arm 2 | change from year 1 at year 2" 0 1 0 0 -1 0  0 -1 0 0 1  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 5 vs. arm 2 | change from year 1 at year 5" 0 1 0 0 -1 0  0 0 0 0 0  0 0 0 0  0 0 0  0 -1
0 0 1 0 / cl;


    lsmestimate year*arm "arm 6 vs. arm 2 | change from year 1 at year 3" 0 1 0 0 0 -1  0 0 0 0 0  0 -1 0 1  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 6 vs. arm 2 | change from year 1 at year 5" 0 1 0 0 0 -1  0 0 0 0 0  0 0 0 0  0 0 0  0 -1
0 0 0 1 / cl;



    lsmestimate year*arm "arm 4 vs. arm 3 | change from year 1 at year 2" 0 0 1 -1 0 0  0 0 -1 1 0  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 4 vs. arm 3 | change from year 1 at year 5" 0 0 1 -1 0 0  0 0 0 0 0  0 0 0 0  0 0 0  0 0 -
1 1 0 0 / cl;


    lsmestimate year*arm "arm 5 vs. arm 3 | change from year 1 at year 2" 0 0 1 0 -1 0  0 -1 0 0 1  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 5 vs. arm 3 | change from year 1 at year 5" 0 0 1 0 -1 0  0 0 0 0 0  0 0 0 0  0 0 0  0 0 -
1 0 1 0 / cl;


    lsmestimate year*arm "arm 6 vs. arm 3 | change from year 1 at year 5" 0 0 1 0 0 -1  0 0 0 0 0  0 0 0 0  0 0 0  0 0 -
1 0 0 1 / cl;


    lsmestimate year*arm "arm 5 vs. arm 4 | change from year 1 at year 2" 0 0 0 1 -1 0  0 0 0 -1 1  0 0 0 0  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 5 vs. arm 4 | change from year 1 at year 5" 0 0 0 1 -1 0  0 0 0 0 0  0 0 0 0  0 0 0  0 0
0 -1 1 0 / cl;
```

```
    lsmestimate year*arm "arm 6 vs. arm 4 | change from year 1 at year 3" 0 0 0 1 0 -1  0 0 0 0 0  0 0 -1 1  0 0 0  0 0
0 0 0 0 / cl;

    lsmestimate year*arm "arm 6 vs. arm 4 | change from year 1 at year 5" 0 0 0 1 0 -1  0 0 0 0 0  0 0 0 0  0 0 0  0 0
0 -1 0 1 / cl;



    lsmestimate year*arm "arm 6 vs. arm 5 | change from year 1 at year 5" 0 0 0 0 1 -1  0 0 0 0 0  0 0 0 0  0 0 0  0 0
0 0 -1 1 / cl;

run;
```