To my family

献给我的家人
# Table of Contents

Table of Contents ........................................................................................................... i
List of Tables ....................................................................................................................... iv
List of Figures ...................................................................................................................... vi
List of Abbreviations .......................................................................................................... xii
Acknowledgements ............................................................................................................ xiii
Abstract ............................................................................................................................ 1

Chapter 1 Introduction ....................................................................................................... 3
  1.1 Schistosomiasis ......................................................................................................... 3
  1.2 Epidemiology of schistosomiasis .............................................................................. 7
  1.3 Mathematical modeling of schistosomiasis ............................................................. 10
  1.4 Outline of the dissertation ....................................................................................... 17

Chapter 2 Two-stage snail population dynamics .............................................................. 21
  2.1 Introduction .............................................................................................................. 21
  2.2 Method .................................................................................................................... 24
    2.2.1 Snail Environment of Msambweni region ....................................................... 25
    2.2.2 Resource dependent snail population dynamics ........................................... 26
    2.2.3 Resource-driven growth ................................................................................. 29
  2.3 Calibration procedure .............................................................................................. 31
    2.3.1 Monte-Carlo simulation .................................................................................. 34
2.4 Snail control and analysis of optimal strategy .................................................. 38

Chapter 3 Dynamics of *Schistosoma* infection and transmission for human hosts .... 43

3.1 Review of traditional approaches ........................................................................ 43

3.1.1 Mean Worm Burden (MWB) approach .......................................................... 43

3.1.2 Distributed worm burden: Negative Binomial approach ............................... 46

3.1.3 Heterogeneous human populations .................................................................. 48

3.2 Stratified worm burden (SWB) approach ............................................................. 50

3.2.1 Method ............................................................................................................ 52

3.2.2 Model calibration ............................................................................................ 63

3.3 Parameter estimation and model validation .......................................................... 76

3.3.1 Description of field data and estimation procedure ........................................ 76

3.3.2 Model validation ............................................................................................. 78

3.3.3 Sensitivity analysis .......................................................................................... 79

Chapter 4 Analysis of control and interventions ......................................................... 81

4.1 Overview of schistosomiasis control strategies ..................................................... 81

4.2 Drug treatment based on SWB ............................................................................. 84

4.2.1 Treatment setting ............................................................................................ 84

4.2.2 Model validation with new data ...................................................................... 85

4.2.3 Two guidelines for chemotherapy .................................................................... 87

4.3 Results .................................................................................................................. 91

4.3.1 Stationary snail populations ............................................................................. 91
4.4 Combination of human chemotherapy and snail control ............................................. 96

Chapter 5 Chronic infection and childhood development ............................................. 99

5.1 Introduction ............................................................................................................. 99

5.2 Method .................................................................................................................. 102

5.2.1 Coupled infection, morbidity and retarded growth dynamics ...................... 103

5.2.2 Input data for calibration .................................................................................. 109

5.2.3 Developmental heterogeneity ........................................................................... 111

5.3 Parameter estimation ............................................................................................ 112

5.3.1 Main steps ........................................................................................................ 112

5.3.2 Monte Carlo simulation ..................................................................................... 114

5.4 Drug treatment simulation and analysis of optimal strategy .............................. 120

Chapter 6 Conclusions ............................................................................................... 125

Appendix A Stratified Worm Burden systems .............................................................. 133

A.1 Basic MacDonald system for multiple human risk groups ................................ 133

A.2 Child-adult SWB system ...................................................................................... 136

A.3 Distributed system ............................................................................................... 137

Bibliography ................................................................................................................ 140
List of Tables

Table 2.1  Known biological parameters about snails .............................................. 33

Table 2.2  Proposed ranges of parameters of two-stage model (2.6) for Monte Carlo Simulation. Low sensitivity of one parameter means the change of value of it alone has little effect on the outcome based on simulation in Mathematica 7. ............................. 33

Table 2.3 Mean and standard deviation of parameters based on the 516 ensemble from the 2nd stage of MC simulation. ......................................................................................... 35

Table 3.1  Parameters of basic MacDonald system (3.1) ............................................. 44

Table 3.2 List of notations, parameters and their description in MWB system(Appendix A.1) for a single human and snail population. ........................................................................... 65

Table 3.3 Field data from Msambweni region of eastern Kenya. (a) Demographic and infection data for 10 villages; ‘egg-count’ refers to arithmetic mean/host; ‘prevalence’ = infected fraction of population. (b) Snail density and infection data for 5 water sites. (c) Relative contact rate for each village to the 5 snail sites (each row sums up to 1). .......... 75

Table 3.4 Calibrated transmission coefficients (A’s and B’s) and egg production rates (ρ’s) based on the individual level field data for 2 choices of discretization step (Δw), and infection threshold \( w_f = 10 \). ......................................................................................................................... 78
Table 4.1 minimum number of treatments needed for all prevalence to go below 10% and duration of low level infection for each village for different children coverage under WHO 2002 guidelines, with coverage for high risk adult group is 60%.......................... 91

Table 4.2 Comparison of WHO 2002 strategies without (No) and with (Yes) yearly pretreatment screening........................................................................................................................................... 93

Table 4.3 Number of safe years after 4-year-treatment for each village under strategy 1, 2 and 4 with coverage being 0.7 for children and 0.5 for adults................................................................. 94

Table 5.1 Simulation parameter descriptions and their best-fit values based on infection and anthropometric data from Kajiwe Village, Kenya [129]. ......................................................... 119

Table 5.2 Predicted community height/weight relative to US median at age 20yr, following different treatment regimens. The first element in the parenthesis is the mean, and the second standard deviation ........................................................................................................ 122
List of Figures

Figure 1.1 Global distribution of schistosomiasis[13] ................................................................. 5

Figure 1.2 Diagram of life cycle of schistosome. This diagram is provided by Center for Disease Control and Prevention (CDC) .............................................................. 7

Figure 2.1 Rainfall data (solid bar) and snail population (dotted line); grey bar indicates missing data .................................................................................................................. 26

Figure 2.2 Accumulated rainfall $V(t)$ (shaded area) and resource functions $F(t)$ (thick curves) at 3 different values of $\alpha$: $\alpha = 5 \text{ /year}$ (dotted), $\alpha = 10 \text{ /year}$ (dashed), and $\alpha = 20 \text{ /year}$ (solid black). ............................................................................................................. 28

Figure 2.3 Growth-replication pattern $b(f)$ and snail mortality $\mu(f)$, rescaled relative to threshold ($f_0 = 1$). ...................................................................................................................... 30

Figure 2.4 Marginal parameter distributions based on 516 best choices of the 2nd Monte-Carlo run .................................................................................................................. 36

Figure 2.5 Solution envelope corresponding 516 best parameter choices of the 2nd MC run. Time $t=0$ corresponds to the beginning of year 1984. Black dots represent observed adult snail density, and light gray curves a few selected solutions ........................................ 37

Figure 2.6 Adult snail population rescaled by dividing 1000 (red solid curve) produced by the two-stage snail population model with rainfall data (black) in the same area as the
original data over 5-year period starting from February 2005. The dashed purple curve represents the corresponding habitat \( V(t) \) ............................................................ 38

Figure 2.7 Snail control with different timings. The envelopes are created with the 100 best parameter choices. .......................................................... 40

Figure 2.8 Relative annual mean snail density reduction for different control timing. The envelop is made from mean and standard deviation (std) of the solution ensemble based on 100 best choices of parameters: (mean-std, mean + std) with the solid curve in the middle being the mean. .......................................................... 41

Figure 2.9 Box plot of annual average density of adult snails (scaled by 1000) based on the 100 sets of ‘best-fit’ parameters under different treatment frequencies. ................. 42

Figure 3.1 Examples for the negative binomial fit of frequency distribution of parasite numbers per person. (a) Ascaris Lumbricoides in Korea (m = 2.18 worms per person, k = 0.319). (b) Anecylostoma duodenale in India (m = 24.54 worms per person, k = 0.618) [58].......................................................... 48

Figure 3.2 Infective strata and transitions of SWB model: population strata (numbers or fractions of total) \( \{h_i\} \) move ‘up’ or ‘down’ the scale at certain rates, due to worm acquisition (\( \lambda \)) and resolution (\( \gamma_i \)). Each stratum has its source-term \( S_i \), and loss rate \( \mu \). .......................................................... 55

Figure 3.3 (A) Equilibrium worm burden distribution for different values of the worm acquisition rate \( \lambda \): .2 (black), .6 (medium gray) and 1 (light gray), with identical human vii
loss rate \( \mu = 0.02 \) and stationary source at the lowest strata. (B) Time relaxation of SWB-system coupled to infected snail prevalence. Left panel shows prevalence distributions of 15 human strata \( \{h_i\} \) at \( t = 0 \) (black), \( t = 20 \) (medium gray) and \( t = 40 \) (light gray). For snail prevalence we show the dynamic relaxation pattern (right panel).

Figure 3.4 Mean worm burden and snail prevalence from SWB and standard MacDonald system with the same parameter values. The step size of worm burden \( \Delta w \) is set to be 5 and number of strata equals 30.

Figure 3.5 (A) Plot of uninfected prevalence function \( P = \phi(\lambda, \mu, S) \) on worm acquisition rate \( \lambda \) for 3 values of \( \mu = \{1\gamma, 2\gamma, 3\gamma\} \) (top to bottom) and identical stationary source \( S \) at the lowest strata; (B) Plot of MWB-function \( w = \psi(\lambda, \mu, S) \) on \( \lambda \) for the same values \( \mu \) of as (A) (top to bottom) with dashed lines corresponding to a would-be ‘linear relation’ (slope) \( \frac{\partial w}{\partial \lambda} \approx \frac{1}{\gamma + \mu} \) at \( \lambda = 0 \) (appropriate for Macdonald MWB).

Figure 3.6 Partial area map of the Msambweni region (left), and its schematic view (right) with human sites (black dots), snail sites (gray dots) and established ‘human-to-snail’ links based on observed human contact activity.

Figure 3.7 Comparison between individual level field data (dark disk) and equilibrium solutions of the calibrated model (gray square); the x-axis labels the ‘village’ or snail site and y-axis – ‘mean egg count’ or ‘snail density’.
Figure 3.8 Sensitivity analysis of model parameters (snail-to-human transmission rates) \( \{A_i\} \) for children in 10 villages: panel (A) shows heavily infected villages no. 6 and 7, and panel (B) for all others. Two sets (light,dark) of box-and-whisker plots (‘box’ = .25 - .75 quantiles; ‘whiskers’ = 10 - 90) correspond to infectivity thresholds \( w_T = 10 \) and \( w_T = 20 \). Black dots mark calibrated values of the ‘individual level field data’ in each case.

Figure 4.1 Box-plot of prevalence of each village from model simulation based on SWB model with 100 sets of parameters randomly perturbed around the best-fit one. Gray dots for village 6 and 7 stand for the simulation result from the model with the best-fit parameter values, while black dots represent new data.

Figure 4.2 Bar charts of prevalence for 10 villages after 4-year-session of different treatment strategies under SCORE program, with coverage being 0.7.

Figure 4.3 Prevalences of village no. 2(dashed) and no.6(Solid) as functions of time with treatment strategy 3 in SCORE (CDT-CDT-Holiday-Holiday) and then followed by 4 (gray) or 8(black) yearly school-based treatments.

Figure 4.4 Seasonal snail patterns (same for each year) under snail control of different frequency. blue –baseline (no control); purple-once at best timing (0.37); yellow-twice (at 0.37 and 0.74); green-quarterly (at the beginning of each quarter).

Figure 4.5 Predication of Schistosoma infection prevalences for children in the 10 villages in eastern Kenya starting in December 2009, using Stratified Worm Burden...
model with the same 4-year-session drug treatment scheme under SCORE program (CDT-CDT-CDT-CDT) and snail control of different frequency. upper left - no control; upper right – once/year; lower left – twice/year; lower right – quarterly.

Figure 5.1 Example of normal growth pattern for children: Left-hand panels indicate CDC curves for normal weight and height at different ages during childhood, indicating growth-related gains up to age 20. Right-hand panels indicate their corresponding normal age-specific rates of change (‘growth velocity’), based on NCHS data available at http://www.cdc.gov/growthcharts/percentile_data_files.htm [128]. Data for boys are shown in the upper panels, and for girls in the lower panels. Graphed lines, from bottom to top, represent the expected 3rd, 10th, 50th, 90th, and 97th population-based percentiles of childhood growth, respectively.

Figure 5.2 Illustration of interactive manipulation in Mathematica 7. The left panel is the control panel to choose the values of parameter. The right panel shows the corresponding result.

Figure 5.3 Envelopes of schisto-infection intensity (worm burden represented by egg count) and growth patterns in terms of height (in cm) and weight (in kg) for the hypothetical heterogeneous Kenyan communities (one for girls and the other for boys), each of which is made up of 5 quantile groups i.e., the 5th, 25th, 50th, 75th, and 95th percentiles, containing 20, 60, 40, 60, 20 individuals, respectively. Upper panels for boys, lower for girls, left for egg count, middle for height and right for weight. In each plot, the
solid thick curve (if applicable) represents the US median; the dots represent data and the solid gray line is the best-fit curve with parameters shown in Table 4.1. .......................... 117

Figure 5.4 Histograms of growth achievement in terms of height and weight at age 20 for the communities without treatment. Upper panels are for boys, lower for girls, left for height, and right for weight. Note that they are approximately of normal distribution.. 118

Figure 5.5 Projected effects of treatment at ages 6, 9, and 12 yr (with the fraction of worms killed in each session being 90%) on worm burden, accumulated morbidity, and growth of an individual randomly selected from the communities. Solid curves are for baseline (untreated) and dashed for with the treatment. ......................................................... 118

Figure 5.6 Projected impact of different schistosomiasis treatment regimens on accrued growth (relative to normal) at age 20 for infected girls and boys exposed to possible reinfection in endemic areas, based on model simulations of growth, reinfection and potential for catch-up growth after treatment. In each panel, clustered bars for girls and boys compare the projected outcomes of 5 proposed strategies: i) no treatment, ii) treatment at the time of school entry and completion (ages 5 and 15 yr), iii) treatment at ages 6, 9, and 12 yr old, iv) treatment every 2 years and v) treatment every year during school-ages from 4 up to 15 yr. Upper panels show substantially reduced impact when community participation is low (20%), as compared to where projected uptake and adherence are high (80%, lower panels). ................................................................. 124
List of Abbreviations

BRN - Basic Reproductive Number
CDT - Community Directed Treatment
MWB - Mean Worm Burden
MDT - Mass Drug Treatment
NB - Negative Binomial Distribution
NCHS - National Center for Health Statistics
NGR - Normal Growth Rate
PZQ - Praziquantel
RGR - Remedial Growth Rate
S. - Schistosoma
SWB - Stratified Worm Burden
SBT - School-Based Treatment
WHO - World Health Organization
Acknowledgements

I would like to express my deepest gratitude to Dr. Gurarie and Dr. King for their guidance and support in my PhD study. I really appreciate the opportunities to work with these exciting projects and face many scientific challenges, from which I learned a lot and will be beneficial for my life time.

I also acknowledge my committee members, Dr. Calvetti and Dr. Thomas, for their time and help.

I want to thank all graduate students, professors, and stuff in the department of mathematics for the friendly working environment throughout these years. Special thanks to my friends, who made my life more wonderful in Cleveland.

I would like to thank my husband, Youjiang and my daughter, Tianyu, for giving me the continuous encouragements and energy to make progress. I also thank my parents and all my family members for their selfless support and trust.

I also acknowledge the funding from Fogarty International Center through Center for Global Health and Diseases at Case Western Reserve University.
Mathematical Models of Schistosomiasis Transmission, Morbidity and Control with Applications to Endemic Communities in Coastal Kenya

Abstract

by

XIAOXIA WANG

Schistosomiasis is a tropical parasitic disease caused by blood-dwelling fluke worms of the genus *Schistosoma*. More than 200 million people worldwide, especially in sub-Saharan Africa, the Middle East and Southeast Asia, suffer from this disease. The disease is characterized by focal epidemiology and over-dispersed distribution pattern, with higher infection rates in children than in adults. Beside acute symptoms, chronic schisto-infection results in indirect morbidity such as cognitive and physical impairment in children. While some advances have been made in the control of the disease through population-based chemotherapy more efforts are required to achieve the goal of disease elimination with limited resources. World Health Organization has proposed several schistosomiasis control guidelines and programs by, and their outcomes need to be assessed. Mathematical models can be used for prediction and control analysis. Here we develop several of them that focus on specific parts and processes of schisto-infection and disease. They are i) dynamics model of snail population with rainfall input; ii)
heterogeneous transmission for distributed human/snail population systems, and over-dispersed worm burden in host populations; iii) the effect of chronic schisto-infection on childhood growth and development. The models were calibrated using epidemiological and demographic data from a coastal region in eastern Kenya. We used these models to simulate and predict the effect of different control strategies, to do their cost-benefit analysis and find the optimal regimens. Our models are based on nonlinear differential equations, and exploit diverse mathematical tools and techniques for analysis. For numeric simulations, calibration and symbolic algebra we used Wolfram Mathematica 7.
Chapter 1  Introduction

1.1 Schistosomiasis

Schistosomiasis, which is also known as bilharzia, is a tropical parasitic disease caused by blood-dwelling fluke worms of the genus *Schistosoma*. The main schistosomes infecting human beings are: *S. mansoni*, causing intestinal and hepatic schistosomiasis; *S. haematobium*, causing urinary schistosomiasis; and *S. japonicum*, causing intestinal and hepatosplenic schistosomiasis. In 2005, there were an estimated 207 million people infected, with 89% of these living in the less-developed areas of rural sub-Saharan Africa and South America [1, 2](see Figure 1.1). The disease usually results from contacting with infected water via bathing and swimming, scuba diving, waterskiing and rafting[3]. Symptoms of acute schistosomiasis (called Katayama fever), occurring a few weeks to month after a primary infection in tourists and travelers and people accidentally exposed to transmission[4], include fever, fatigue, myalgia, malaise, non-productive cough, eosinophilia, and abdominal pains. Most patients recover spontaneously after 2-10 weeks, but some develop persistent and more serious disease with weight loss, dyspnea, diarrhoea, diffuse abdominal pains, hepatosplenomegaly and widespread rash. Chronic schistosomiasis develops gradually in the course of infection in people with recurrent infections, commonly in endemic areas. It is mainly caused by the eggs of schistosome that provoke inflammatory and granulomatous reactions, which are progressively replaced by fibrotic deposits [5]. Symptoms of disease vary with types of schistosomiasis
(urinary, intestinal, hepatic) and the severity depends on the intensity of infection and individual immune responses. For example, urinary schistosomiasis has early signs include dysuria, pollakisuria, haematuria [6, 7] and eventually leads to parenchymal damage and kidney failure. Long-term infection causes not only these direct morbidity, but also indirect morbidity such as fatigue and physical or cognitive impairment, with the later receiving more and more attention. Recent studies found significant associations between schistosome infection and anemia, nutritional status and cognitive and physiological capacities [8]. Since 1970s, effective, safe and simple drugs against schistosomiasis become available [9]. Praziquantel, which is active against all schistosome species, and has very low toxicity in animals, is the most widely used [10, 11]. The drug acts within one hour of ingestion by paralyzing the worms and damaging the tegument, but has little or no effect on eggs and immature worms. The drug is very effective. After a single dose of 40 mg/kg, 70-100% of patients cease to excrete eggs [12].
The life cycle of schistosome parasites is complicated and involves two different hosts: human beings and snails. The different species of schistosome parasites have very similar patterns of human-snail life cycle. Figure 1.2, a diagram provided by Center for Disease Control and Prevention (CDC), shows the life cycle of a schistosome. Parasite eggs produced by female schistosome worms are excreted from infected human hosts in the urine or faeces and released into the environment. They can stay viable for up to 7 days. On contact with water, the eggs release the free-swimming miracidium, which searches for the intermediate host, fresh-water snails. After penetrating the snail, the miracidia multiply asexually into multi-cellular sporocysts and later into cercarial larvae. The cercariae start leaving the snail 4-6 weeks after infection and swim around in the water for up to 72 hrs seeking the skin of a suitable definitive host. Snails that shed
cercariae are named shedding snails. One snail, infected by one miracidium, can shed thousands of cercariae every day for 1-2 months. On finding a host (human or domestic livestock in the case of *S. japonicum*), the cercariae penetrate the skin, migrate in the blood via the lungs to the liver, and transform into young worms or schistosomula. These mature in 4-6 weeks in the portal vein. Unlike other trematodes, schistosomes have separate sexes. Once mated, they live as permanently embraced couples and migrate to their perivascular or mesenteric destination where they produce eggs. The lifespan of an adult schistosome averages 3-5 years but can be as long as 30 years. Some of the eggs pass through the walls of the blood vessels, then through the intestinal wall or bladder wall, and are then excreted out of the body in faeces or urine. These eggs are released into the environment and restart the life cycle. However, up to half the eggs produced by the mated worm pairs are trapped in the mesenteric veins or are washed into the liver or kidneys and bladder wall, where they will become permanently lodged. Those trapped eggs release antigens that lead to a vigorous immune response. The eggs are minimally harmful to the body. However the cellular infiltration that results from the immune response to eggs is what causes the pathology associated with schistosomiasis. As a result, duration and intensity of infection play a significant role in disease causation in human schistosomiasis.

From the life cycle of schistosome, we see that water sites like natural streams, ponds, and lakes are typical sources of infection. Transmission determinants include human/snail populations, cercarial density, and patterns of human water contact, all of
which show strong temporal and spatial variations and result in a focal distribution of the infection within countries, regions, and villages.

Figure 1.2 Diagram of life cycle of schistosome. This diagram is provided by Center for Disease Control and Prevention (CDC).

1.2 Epidemiology of schistosomiasis

Before turning to the development and analysis of mathematical models of schistosomiasis (or any other helminthes) transmission, I would briefly talk about schistosomiasis from epidemiological point of view.

General parasites are divided into two categories: microparasites and macroparasites. Microparasites may be thought of as those parasites which have direct reproduction-usually at very high rates within the host[14]. They tend to be characterized by small size and a short generation time. Common diseases caused by microparasites
including measles, influenza, HIV, malaria, etc. Macroparasites may be thought of those having no direct reproduction within the definitive host [14]. They are typically larger and have much longer generation times than microparasites. This category embraces most parasitic helminthes and arthropods, and includes flukes (e.g. schistosome), tapeworms, and nematodes (e.g. filaria). Macroparasitic infections are generally chronic in form and they are more a cause of morbidity than mortality. Another remarkable feature of macroparasitic diseases is that age–dependent reinfection often follows after chemotherapeutic treatment in endemic area [15], which makes the elimination of macroparasitic diseases much more difficult.

Now, let’s consider briefly epidemiological data which are available to calibrate models and test predictions for schistosomiasis transmission (those for other helminthic disease are similar). Options to measure the distribution and abundance of endoparasites within human (or snail) communities are limited by practical and ethical considerations. The most widely used one involves the assessment of parasite burden within individual patients via measures of schistosome eggs abundance in the faeces or urine, depending on the type of schistosome. The association between the rate of production of transmission stages and the adult worm burden, however, is made complicated by density-dependent fecundity [16]. Nevertheless, a simplification to a linear relationship between egg counts and worm burden has been a widely accepted practice [17-20], especially when one is interested at a population level.
The two most widely used statistics in the epidemiological study of macroparasites are the prevalence of infection (the proportion of percentage of hosts infected in the sample) and the intensity of infection (a record of worm abundance measured via direct or indirect methods). With respect to latter quantity, average values and variance around the mean can be calculated for different strata of the population, according to age, sex, or social groups. Changes in prevalence or average intensity of infection are commonly recorded by cross-sectional sampling schemes either horizontally (at one point in time across different sections of the same population) or longitudinally (through time and across the different sections). Horizontal or longitudinal changes in prevalence and intensity of infection with host age are very valuable in the study of transmission dynamics, because age often reflects duration of exposure to infection in areas of endemic parasitism [21].

The microscope has been the traditional tool to count eggs in faeces or urine sample in laboratories. Since 1990’s, more sophisticated techniques, such as DNA probes or the polymerase chain reaction (PCR) have been used to provide more accurate data by improving sensitivity and specificity in the identification of eggs in samples.

No matter how accurate the measurement techniques are, there is always a lot of uncertainty (or errors) when it comes the real infection status of humans or snail hosts. For instance, random errors arise when data are collected by different people; systematic errors may happen when the data variable is dependent on its value (e.g., young snails are hard to collect and count); biases occur when the samples used for record may not be
representative for the population we are interested in. There is another case of uncertainty when we have to look into the subgroups of a population for more information while we are only given summary statistics for the population (like mean, median). So in modeling practice, we should take these into account and make our results as reliable as possible.

1.3 Mathematical modeling of schistosomiasis

Field studies alone would not allow quantitative estimates of various components of transmission or prediction of interventions, like human chemotherapy, molluscicides or other control measures. Control experiments in real communities are expensive in terms of both time and resource, and are limited by ethics and often inconclusive because of other confounders. All these problems can be addressed by mathematical modeling. In general, mathematical models can help us to interpret observed epidemiological trends, to guide the collection of data towards further understanding, and to design programs for the control of infection and disease. With the help of powerful computational tools available today, it is becoming more and more efficient.

A useful mathematical model should properly account for biological and environmental determinants of transmission. From previous description of schistosomiasis, we know that the parasite life cycle requires two hosts-definitive (human/mammal) and intermediate (snail). These two host populations infect each other in a typical “vector borne” transmission pattern through mediating larval stages. Transmission occurs at suitable “contact sites” - water snail habitats. Its intensity depends
on various environmental/weather factors that regulate snail abundance, as well as human biology and behavior (water/land use). The resulting risk factors and infection levels depend on combination of all of those factors. Taking these factors and other critical mechanisms affecting the intensity of infection is important in modeling schistosomiasis transmission.

Many mathematical models of schistosomiasis transmission have been published since 1965. Most of those have had little impact on field studies or on the design of disease control because of the limited interaction between mathematicians and field workers [22]. Some models were developed based on oversimplified assumptions and not easily adapted to field data. However, those previous works laid a foundation for future work and shed light on direction of our research in the modeling the disease transmission. One of the earliest and well-known models was developed by George MacDonald in 1965 [23] (This will be presented in detail in Chapter 3). Although it was based on the oversimplified biological assumptions, the model showed great potential for understanding the disease transmission and control through mathematical analysis.

For a long time after that, modelers used homogeneity assumptions about human hosts and used mean worm burden to describe the infection status of a population. And snail host population is often divided into three compartments: susceptible, infected and shedding (infectious). The infected snails are the ones infected by miracidia and not yet shedding cercariae (infectious). The time interval between the infection and the onset of shedding is called prepatent or latent period. So a typical schistosomiasis transmission
model consists of two parts. The first one deals with intensity of human infection i.e. mean worm burden in the human community. The second one accounts for the dynamics of susceptible, infected and shedding snails. More detailed models also include larval stages: cercariae and miracidia [24]. However, these forms having short life expectancies (less than 1 day) compared to snails and schistosomes, one can assume the cercariae and miracidia populations to be at equilibrium. They would enter the model as part of the transmission parameters. In this model, the transmission parameters such as contact rates, worm establishment probability, and contamination rates are the key elements in the connection between intensity of human infection and different compartments of snail population.

This kind of models suffer from two main unrealistic and over-simplified assumptions on snails: one is that the latent period of infection in snails is ignored and the other that total number of snails keeps constant over time. This latent period is epidemiologically significant to snails since the life expectancy of snail is comparable to the length of latent period. The models that ignore latent period would predict an unrealistically high prevalence of shedding snails [25]. Latent period has been incorporated into transmission models since 1976 [25-27]. These earlier works suggest that only a small fraction of infected snails become shedding after a fixed time period. The assumption that snail population remains steady over time and independent of the intensity of the disease fails to account for seasonal environmental changes that can cause dramatic fluctuation in local snail populations. Parasitic infection can affect mortality and
reduce the life-expectancy of snails as well [28, 29]. Some attempts have been made to
develop more realistic models by including seasonal environmental factors such as
temperature and rainfall in snail dynamics [30, 31]. Although numerical results of these
improved models can mimic some patterns of the snail dynamics observed in the field,
they are not satisfactory because of their limited predictive power. For instance, models
of S. Liang [30] used temperature and rainfall to formulate the effective growth rate of
snail population based on earlier work of Woolhouse and Chandiwana [31]. This model
can be fit to the field data. However, the functional form of environmental factors in this
model fails to account for the mechanisms that trigger snail growth. It explains field data
used for calibration, but it may not predict population response under different
environmental conditions. Moreover, without understanding the mechanisms that drive
snail population growth, one cannot predict the efficacy of snail control interventions. In
brief, a realistic and predictable snail population model should include the critical
environmental factors and account for the driving mechanisms of snail population growth
and establishment.

The homogenous assumptions about human populations used in these models are
also overly simplified and unrealistic. Indeed, field observation showed significant
differences in the mean egg count among different villages. Within one village, one often
observes big differences in infection intensity among different age groups. Even within
an age group, the distribution is invariably highly aggregated such that most individuals
harbor few parasites and a few individuals harbor the majority of the parasite population.
Parasite aggregation may arise as a consequence of a wide variety of factors either acting alone or concomitantly (see [32, 33]). There include heterogeneity in exposure to infection (due to social, environmental, or behavioral factors, or to aggregation in the spatial distribution of infective stages or infected intermediate hosts), differences in susceptibility to infection (due to genetic or nutritional factors, or to varying past experiences of infection), or to variability in parasite survival within different individuals (due to genetic, immunological, or nutritional factors). For instance, contact rates observed in field studies suggest variable usage among various water sites. People from different age groups would have different behavioral patterns in water and varying biological responses to the parasite invasion. Moreover, parasite infection for a population is naturally a random process, in that some individuals may be infected and others may not when exposed to the same infectious source. This may result in an aggregated distribution of infection intensity in a homogenous population. One needs to account all these essential factors to address the heterogeneities in schistosomiasis transmission.

Earlier works of Woolhouse and others [17, 22, 30, 34-39], addressed heterogeneities in patterns of exposure and contamination through the so called contact rate matrix. The results of these works suggested how contact heterogeneities could have significant epidemiological consequences in transmission. Heterogeneities can be studied with spatial-distributed, age-structured, and occupation-specific human populations. For instance, Barbour [34, 35] and Woolhouse [22] accounted for the heterogeneities in
distributed populations without age structure whereas Chan et al. [36] have studied age-structured transmission effect for a single human population. To make their model more realistic, some workers even combined different types of heterogeneity within the human population. For instance, Gurarie [17] extended previous works further by accounting for spatial distribution and age structure of human populations and Liang [20, 30] combined spatial and occupation-specific heterogeneities in their model. The resulting mathematical models involve a large number of equations and variables, which makes them hard to study analytically. However, the development of computing techniques provides many powerful tools to analyze those complicated models and generate numerical solutions. Furthermore, the enhanced computing power enables us to develop and analyze more realistic models accounting for a larger number of relevant factors. Many earlier works shed light on the current modeling efforts to account for geographic and human behavior features, effects of acquired immune memory and control of chronic morbidity [23, 36, 40-44].

All of the above heterogeneous models can be categorized into two groups in terms of how to describe the infection status of human populations. In the first group, the observed pattern of worm distribution is ignored entirely and only mean worm burden is used; the second group empirically uses the negative binomial distribution. This distribution is characterized by two parameters: the mean, $m$, and a parameter, $k$, which varies inversely with the degree of parasite aggregation ($k \to 0$ as the total population is concentrated on fewer and fewer people, while $k \to \infty$ as the distribution become more
random in character) (see [45]). Most observed distributions of *Schistosoma* within human communities are remarkable for their high degrees of aggregation. Values of the aggregation parameter $k$ typically lie in the range 0.1-1. For example, for *S. mansoni* in Brazil, it is between 0.03 and 0.5 [46]. It is not uncommon, in these circumstances, for more than 80 percent of the total parasite population to be harbored by less than 20 percent of the human community. As such, severe morbidity due to schistosomiasis is often restricted to a small fraction of ‘wormy people’.

The observation that the negative binomial model is a good description of recorded patterns in human communities should therefore simply be regarded as a convenience, in the sense of giving a useful two-parameter distribution model to describe observed trends. Its goodness of fit to data can never, by itself, suffice to ‘explain’ the pattern of parasite abundance. We shall develop a dynamic model of schistosomiasis transmission that accounts for over-dispersion of infection intensity in host populations by studying the mechanism.

Another topic of interest for us is the *Schistosoma* infection related morbidity in endemic areas. As is mentioned above, schistosomiasis is a very widespread disease in the developing world and is one of the most important helminthic diseases in public health terms. The intensity of current infection is important for some acute forms of diseases: for example, the presence of hematuria is closely related to the intensity of infection in *S. haematobium* [47]. Other forms of morbidity may develop over many years and relate to previous experience of infection: for instance, *S. mansoni* infections
are generally most prevalent in younger people but the liver disease due to this infection is slow to develop and may mostly show up decades later [48]. Medley and Bundy ([44]) proposed a general framework that modeled schistosome-induced damage as a function of accumulated experience of infection. They suggested different models of morbidity with different assumptions of how the morbidity resolves after treatment, but did not attempt to relate model predications to empirically observed measures of morbidity. Chan et al (1996) [49] developed a more detailed model of schistosomal morbidity. They categorized morbidity into that due to current heavy infection, and early and late stages of chronic disease. The model was parameterized using clinical data of observed diseases such as hepatomegaly, splenomegaly, fibrosis of the liver, and diarrhea. As did Medley and Bundy, they aimed to quantify the different resolution rates of morbidity for early or late forms of disease through treatment. Childhood growth stunting due to chronic schistosome infection and catch up growth after treatment have never been qualitatively investigated by mathematical modeling. We will do this in our study.

1.4 Outline of the dissertation

In order to better understand the schistosomiasis transmission and childhood growth retardation due to chronic schistosome infection in a schistosomiasis endemic area, Msambweni of the Kwale District in eastern coastal Kenya, and to estimate potential outcomes of various intervention strategies proposed by WHO, we develop several dynamic models based on nonlinear ordinary differential equations. Based on
field data, we calibrate our models and run control experiments by numerical simulation in Mathematica 7 (Wolfram Research, Champaign, IL).

In Chapter 2, we develop a population model for snail dynamics coupled to seasonal factors. We utilize longitudinal biweekly data about snail density and rainfall over a 3 year period in a coastal region of eastern Kenya for model calibration. The underlying logic is that rainfall plays critical role in snail population dynamics. As it fills water pools and creeks during rainy season, the accumulated rainfall creates the habitat for snail breeding and provides for their food supplies. In the model, young and old snails are distinguished, since only the latter contribute to schistosomal transmission. Cohabiting the same environment, snails of the two stages compete for resources reflected by resource-dependent growth rate and death rate. Resource is quantified by volume of water sites (pools) through seasonal accumulation of rainfall. The model developed in this way has good predictive power. We expect our model to generate the proper snail response under variable conditions such as seasonal rainfall and mollusciside. After calibrating the model, we identify the optimal timing for snail control under fixed rainfall pattern, and we also investigate the outcome of control programs with different frequencies of snail control treatments. The result is then coupled with the transmission model developed in Chapter 3 to compare the effect of various integrated disease control strategies.

In Chapter 3, we propose a new modeling approach, called *stratified worm burden* (SWB), to schistosome transmission by stratifying human populations according to worm
burden, and replacing *mean worm burden* (MWB) dynamics with that of 'population strata'. We then developed proper calibration procedures for such multi-component systems in different level of populations: a single homogeneous population, multiple risk groups, child-adult system, and a distributed system in which there are multiple child-adult systems connected through sharing multiple water sites. We apply the methodology to a special environment- Msambweni region of eastern Kenya. The cross sectional data we use for calibration include community level (village) human populations, snail densities, human and snail infection prevalence, and contact rates for villages to different water sites. We do sensitivity analysis as well as model validation. This modeling methodology is potentially applicable to all macroparasitic transmission.

In Chapter 4, we continue to explore the advantage of SWB approach developed in Chapter 3 in terms of prediction of the effects of cyclic chemotherapy. Compared with MWB methodology, drug treatment for SWB can be formulated in a different way. This allows us to access the impact of different drug treatment strategies. We may also accommodate human/snail population change in this model to make the predication more realistic. We then do treatment experiments according to the current WHO guidelines, as well as the newly proposed protocols of the SCORE project, the newly launched anti-schisto program by WHO targeting endemic communities with high infection level in Africa, and predict their possible outcomes. For example, we investigate how many years of treatments are needed to bring the regional infection prevalence to below 10% and how long control areas will remain below this safe level after the program finishes. We
also compare results for two settings of snail populations: one in which snail population is stationary; the other is based on the snail model in chapter 2, which is a dynamic with a seasonal pattern and is subjected to control with different frequencies of interventions.

In Chapter 5, we study the infection-induced morbidity for children and linked dynamically the childhood growth retardation and chronic schistosomal infection. We use cross sectional data sets including age-specific infection intensity and community median height/weight from 4 to 20 yr old children. We take corresponding US anthropometric data, available at CDC website, as growth standards. Because gender is an important factor in body development, we distinguish boys and girls for analysis. People in different quantile groups make up a heterogeneous community, and we should treat them differently. So in order to interpret the data as realistically as possible, we construct a hypothetical community for our prediction.

In Chapter 6 we give an overall summary of our models, discuss the results, and present some ideas for their potential improvement in the future.
Chapter 2  
Two - stage snail population dynamics

2.1 Introduction

Understanding seasonal dynamics of snail populations is important for modeling vector- borne diseases and control interventions where snails serve as intermediate host, as occurs with schistosome infection. Indeed, many transmission areas exhibit strong seasonal variations that can affect transmission and control [42, 50]. The classical models of schistosomiasis transmission [21, 23, 28, 34, 46, 51] typically consider only stationary environments and snail populations. That leaves open a significant question on how seasonality can affect such predictions.

Many ecological and environmental factors drive snail population growth, among them temperature, precipitation, and differences in habitat due to differences in steady or moving water flows, vegetation, etc. Several earlier works [30, 31, 50, 52] focus on the role temperature in snail reproduction. The key concept here is temperature-day [30, 31], whereby snails need to accumulate a certain (threshold) level of “days” within suitable “temperature window” to start reproduction.

The effect of “temperature-day” accumulation is a time-lag term in the dynamic snail equation. The mathematical formulation of such model for population (density) $N(t)$ with per capita reproduction rate $b(t)$, introduces a time-lag $\tau$ determined by accumulated temperature-days, as well as snail mortality and other intervening factors.
(e.g. rainfall). Different patterns of temperature and rainfall dependence have been proposed by Liang et al. [30, 50].

Here we take another approach to snail population dynamics for regions where seasonal precipitation plays crucial role. In such environments (e.g. eastern coast of Kenya) the rainfall accumulates in water pools (static) during rainy season, which creates snail habitat and food resources required for reproduction, a process followed with some delay by an upsurge in snail population (Figure 2.1). We attribute this time-lag to resource development and snail maturation processes. In the present work we develop a resource-driven snail population model, with growth and mortality determined by availability of resource, which is a combination of food, habitat, water and other necessities. The environment has a carrying capacity and the total snail population cannot exceed a maximal load. We study snail population dynamics mainly for understanding schistosomiasis transmission. Since only mature snails serve as intermediate hosts for the parasite, we divide snail population into two stages, adult and young, and let them compete for resources.

We apply this approach to a specific environment and data set collected in the coastal region of eastern Kenya. The local snail species include *Bulinus globosus* and *Bulinus nasutus* - key intermediate hosts for schistosomiasis transmission. Both species are hermaphroditic, possessing male and female organs and being capable of self- or cross-fertilization. Thus, a single snail can invade and populate a new habitat. The eggs are laid at intervals in batches of 5 - 40, each batch being enclosed in a mass of jelly-like
material. The young snails hatch after 6 - 8 days and reach maturity in 4 - 7 weeks, depending on the species and environmental conditions. A snail can lay up to 1000 eggs during its life, which may last more than a year. Temperature and food availability are among the most important limiting factors. In our current model we shall focus primary on rainfall as a proxy for the “source of food”.

Having formulated such model we calibrate and validate it with the field data collected over a 3-year period [42, 53]. We used the 1st year data for model calibration, and then validated it over an extended 3-year period. The model exhibited good predictive skills.

Based on this model, we can examine various snail-control strategies, and identify the optimal seasonal-based control. Indeed, our formulation directly accommodates rainfall input in snail population dynamics. Thus we can analyze and predict population response to any environmental change (precipitation) or molluscicides (pesticides against snails). We can also study the effect of future (climate) change of weather patterns on snail populations.

In the forthcoming work, we apply our snail model to schistosomiasis transmission to a distributed human-snail meta-population environment of Msambweni region in Eastern Kenya.
2.2 Method

Our goal is to construct a predictive snail population model based on the environmental constraints and the basic reproductive biology. We take the standard birth-mortality balance

$$ \frac{dN}{dt} = \left[ b(F) - \mu(F) \right] N $$

(2.1)

where both rates depend on the available resource $F$ in a threshold-type pattern. There is a transition level $F_0$, above which the population starts growing (replicating) at increasing rates approaching maximal value $b_0$, while its mortality will remain close to its natural relatively low level $\mu_0$. Below $F_0$ snail replication rate will decrease approaching zero, while its mortality rate will grow higher with diminished $F_0$. Under adverse environmental conditions, the population will experience negative growth (provided $F$ falls well below the threshold), but the growth will resume once it rises above it, based on empiric observations [42]. Resource $F$ in turn will depend on accumulated seasonal rainfall. We shall extend the basic system (2.1) to a two-stage population model made of young and adult stages. Two groups will differ in their mortality rates, besides both will compete for the available resource resulting from accumulated precipitation. There's several reasons to distinguish two developmental stages: (i) only the numbers of the adult stage can be measured in the field; (ii) snail maturation plays role in the observed time-lags between rainfall-peak and snail-peak; (iii) only adult snails contribute to schistosomiasis transmission (after a proper pre-patency period).
Before proceeding to model description we briefly outline the specific environment and field data collected in the Msambweni region of Eastern Kenya used for parameter estimation and model validation.

### 2.2.1 Snail Environment of Msambweni region

The field data was collected between March 1984 and July 1986. The water bodies here are mainly the rain-filled ponds and the snail sampling method was elaborated in previous works [42, 53]. Sampling is conducted by one man searching each snail site for five minutes with a standard flat wire-mesh scoop with a mesh size of 2 mm. Therefore, only adult snails can be found at sampled sites. According to Kariuki et al.[42], the water levels of ponds fluctuated considerably leading to large seasonal variation of snail populations. The ponds fill during the rainy season, and then the snail population peaks with roughly 2-month delay. During the dry season water levels drop, some ponds drying out completely, and snail population declines significantly. Figure 2.1 shows these patterns over a 3-year period.
Accumulated water level in the ponds creates the vital resource for snail abundance. While other environmental and weather factors (temperature, water turbidity, pH value, aquatic vegetation, hydrology of the habitats) can also play the role, their effects in the Msambweni area appear limited. Indeed Kariuki et al.[42] shows rainfall to be the only reliable predictor of snail populations.

### 2.2.2 Resource dependent snail population dynamics

In our model, accumulated rainfall creates snail habitats and provides the food resource that sustains snail population. The resource levels directly impact snail growth rate and mortality. The field study found that long drought of 2001-2003 significantly reduced snail numbers, but they rebounded dramatically after short period of heavy rains. Typically after the dry season, the first rains saturate the soil, then ponds start filling.
When accumulated rainfall exceeds a certain level, habitat and resources are created to trigger population growth. We divide this process into 2 steps:

(i) rainfall accumulation in pools and creeks (habitat)

(ii) growth of vegetation - food resource for snails

We represent habitat by variable \( V(t) \), referring to “volume” (or area) of water pools, and designate food resource (vegetation) by \( F(t) \). We describe the creation of habitat by a simple accumulation-decay process for “volume variable” \( V(t) \) relative to a fixed (background) level \( V_0 \) maintained throughout the season,

\[
\frac{dV}{dt} = P(t) - \theta(V - V_0)
\]  

(2.2)

with known precipitation source \( P(t) \) and decay (loss) rate \( \theta \). The loss-term accounts for all possible hydrologic and other “losses” (evaporation, drainage, seepage etc.). The solution \( V(t) \) of (2.2) is given by a convolution integral

\[
V(t) = V_0 + \int_0^t e^{\theta(t-s)} P(s) ds
\]  

(2.3)

We assume the accumulated habitat creates a carrying capacity (proportional to \( V(t) \)) for vegetation growth, therefore the resource (food) function \( F(t) \) obeys a logistic equation

\[
\frac{dF}{dt} = \alpha \left(1 - \frac{F(t)}{V(t)}\right) F(t)
\]  

(2.4)
Here $\alpha$ denotes the maximal growth rate of $F(t)$, or its maximal relaxation rate to $V(t)$. Let us note that the exact explicit solution of equation (2.4) is

$$F(t) = \frac{1}{(1/F_0)e^{-\alpha t} + \alpha \int_0^t e^{-\alpha(t-\tau)} \int V(\tau) d\tau}$$

This solution formally approaches $V(t)$ as $\alpha \to \infty$, so increased $\alpha$ would drive $F(t)$ faster to the current level of $V(t)$. However, any finite value $\alpha$ creates a time lag between $V(t)$ and $F(t)$ (see Figure 2.2).

The two rates $\theta$ and $\alpha$, ‘rainfall loss’ and ‘resource relaxation’, along with ‘background’ levels $V_0$ and $F_0$ will be among the parameters of the system that will be estimated.

![Figure 2.2 Accumulated rainfall $V(t)$ (shaded area) and resource functions $F(t)$ (thick curves) at 3 different values of $\alpha$: $\alpha = 5$/year (dotted), $\alpha = 10$/year (dashed), and $\alpha = 20$/year (solid black).](image)
2.2.3 Resource-driven growth

Figure 2.1 shows the peak of rainfall and the peak of snail population separated by 2-4 month delay. The rainfall peak typically falls in May while peak of snail abundance (including shedding snails) occurs in August, September, or even October. Two factors can contribute to this delay: resource development from rainfall, represented by logistic equation (2.4), and snail maturation from juvenile to adult stage. Accordingly we divide the snail population into young and old stages competing for the same resource $F$. They allocate resource in proportion to their populations (or population densities per unit habitat), and adult snails get a competitive advantage expressed through the factor $q > 1$, which is yet to be determined. It gives an average amount of resource available to adult relative to that of a young snail, independent of the overall resource level $F$. More precisely, the two stages, $Y$ - young and $X$ - adult, allocate resource $F$ according to

$$ F_Y = \frac{F}{qX + Y}, \quad F_X = \frac{qF}{qX + Y}. \quad (2.5) $$

Now we can formulate a two-stage snail population model with resource-dependent birth (replication) rate $b(f)$, stage-specific mortality rates $\nu(f), \mu(f)$ (for young and adults), and maturation period (young - to - adult) $\tau = 2$ months,

$$ \begin{cases} \frac{d}{dt} Y = b(F_X)X - \left( \nu(F_Y) + \frac{1}{\tau} \right)Y \\ \frac{d}{dt} X = \frac{1}{\tau}Y - \mu(F_X)X \end{cases} \quad (2.6) $$
All rates (birth and mortality) depend on allocated resources $F_x, F_y$ (2.5). And we represent them by sigmoid functions

$$b(f) = \frac{b_0}{1 + \left(\frac{f_0}{f}\right)^{p_y}}; \quad \nu(f) = \nu_0 \left[1 + \left(\frac{f_1}{f}\right)^{p_y}\right]; \quad \mu(f) = \mu_0 \left[1 + \left(\frac{f_0}{f}\right)^{p_x}\right]$$

(2.7)

Here $b_0$ designates maximal reproduction rate, $\nu_0, \mu_0$ - natural mortality rates (under abundant resource), $f_0, f_1$ - threshold levels (for adult/young) snails, and $p_B, p_Y, p_X$ are the Hill exponents. Typical patterns of $b(f)$ and $\mu(f)$ are shown in Figure 2.3. When the available resource $f$ exceeds threshold $f_0$, birth rate increases to its maximal value $b_0$, while mortality drops to its minimal value $\nu_0$. Below $f_0$, growth rate falls to near zero level while mortality keeps increasing.

Additional factors, like schistosome infection can also affect snail population. Indeed, as infected snails become shedding (patent) they stop reproducing and die faster [68]. However, the number of infectious snails is usually small compared to the total
population, typical not exceeding 5-10% of the infected snails. So in the current analysis, we shall ignore the effect of schistosomiasis on snail population.

2.3 Calibration procedure

The combined resource-population system (2.4)-(2.6) can be brought to a dimensionless form via rescaling:

\[ V \mapsto V / V^*; \quad F \mapsto F / F^*; \]
\[ X, Y \mapsto (X, Y) / N^* \]

with scales \( V^* = F^* = 1000 \text{ mm (mean annual rainfall)} \), and \( N^* = 1000 \) maximal snail number. Some biological parameters of the model are known (Table 2.1), and we shall fix them. The remaining uncertain and unknown ones will be estimated within suitable ranges (Table 2.2). Ranges of some parameters, e.g. maximal snail replication \( b_0 \) come from the known biological data, others (model specific ones) were obtained by numeric simulations of the coupled differential system (2.4)-(2.6), by varying these parameters and exploring their ranges with manipulation function in Mathematica 7. The goal is to identify parameter values (or their combinations) “best able” to reproduce the dynamic (rainfall-snail) pattern of Figure 2.1.

The problem of parameter identification can be approached in several ways. Given a solution, or more general function, \( Y = Y(a_1, \ldots, a_n) \) depending on parameters \( \{a_1, a_2, \ldots, a_n\} \), and a proper choice of error, \( \varepsilon(Y) \) (often taken as a square mean distance
from \( Y \) to a “reference” data- point(s) \( Y_0 \), \( \mathcal{E} = \| Y - Y_0 \|^2 \), we seek to minimize an objective function

\[
\mathcal{E}(a_1, \ldots, a_n) = \| Y(a_1, \ldots, a_n) - Y_0 \|^2
\] (2.8)

In our case \( Y = \{ F(t;\ldots); y(t;\ldots); x(t;\ldots) \} \) is a solution of coupled differential system (2.4)-(2.6), which depends on the function of parameters of Table 2.2. We define the objective function to be the square-mean distance for “adult-population” variable \( x(t) \) sampled at the data points \( \{ t_k \} \) of Figure 1- collected snail numbers \( \{ X_k \} \) over the 1\textsuperscript{st} year of study

\[
\mathcal{E} = \sum_k \left[ x(t_k;\ldots) - X_k \right]^2
\] (2.9)

One difficulty solving the minimization problem is that the system (2.6) has no explicit solutions\(^1\). We solve it numerically and produce large ensembles of solutions by sampling different parameter values. In principle, having a properly sampled set of solutions (over a regular grid in the parameter space) one can produce an approximate solution \( Y(t; a_1, \ldots, a_n) \), hence error \( \mathcal{E}(a_1, \ldots, a_n) \), and use it for minimization. Such direct approach however, can work only for problems with small number of unknown parameters (the grid size grows exponentially with the dimension of the parameter space).

\(^1\) While “resource equation” (2.4) by itself is decoupled from snail system (2.6), and allows analytic solution, such procedure is untenable for (2.6) with any realistic pattern \( F(t) \). So we solve the coupled system (2.4)-(2.6) numerically.
An alternative widely used approach for multi-parameter problems, like ours is Monte-Carlo (MC) simulations, described below.

**Table 2.1 Known biological parameters about snails**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturation time-lag</td>
<td>$\tau = 2$ month = $1/6$ year</td>
</tr>
<tr>
<td>Natural mortalities</td>
<td>$\nu_0 = 10$/year - young</td>
</tr>
<tr>
<td></td>
<td>$\mu_0 = 1$/year - adult</td>
</tr>
</tbody>
</table>

**Table 2.2 Proposed ranges of parameters of two-stage model (2.6) for Monte Carlo Simulation.** Low sensitivity of one parameter means the change of value of it alone has little effect on the outcome based on simulation in Mathematica 7.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Name</th>
<th>Proposed Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss rate of water habitats</td>
<td>$\theta$</td>
<td>3-10/year</td>
</tr>
<tr>
<td>Relaxation rate of resource</td>
<td>$\alpha$</td>
<td>3-15/year</td>
</tr>
<tr>
<td>Minimal sustained water level</td>
<td>$V_0$</td>
<td>15-30mm</td>
</tr>
<tr>
<td>Initial resource level</td>
<td>$F_0$</td>
<td>.02-.15 mm</td>
</tr>
<tr>
<td>Competition factor</td>
<td>$q$</td>
<td>5-15</td>
</tr>
<tr>
<td>Hill exponents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>growth</td>
<td>$p_B$</td>
<td>4 (low sensitivity)</td>
</tr>
<tr>
<td>young mortality</td>
<td>$p_Y$</td>
<td>1-10</td>
</tr>
<tr>
<td>adult mortality</td>
<td>$p_X$</td>
<td>3-15</td>
</tr>
<tr>
<td>Resource threshold for young</td>
<td>$f_i$</td>
<td>.1-.6</td>
</tr>
<tr>
<td>Resource threshold for adult</td>
<td>$f_0$</td>
<td>.02-.2</td>
</tr>
<tr>
<td>Maximal replication rate</td>
<td>$b_0$</td>
<td>50-500/year</td>
</tr>
</tbody>
</table>
2.3.1 Monte-Carlo simulation

Our methods combine mathematical analysis with numeric simulations performed on Mathematica 7. Computer limitations and data noise make “best-fit optimization” methods untenable. Instead, we adopt a Monte Carlo (MC) simulation technique, based on some proper prescribed ranges of parameters (Table 2.2 last column). By a MC ‘choice’ of parameters, we mean that we select randomly and independently a value for each parameter as it is uniformly distributed in its own range. For each choice of parameters we run numeric DE solver, compare its output with the data, compute error given by (2.9), and select the choices which yield relative small errors (less than an error threshold).

We use the three-year data shown in Figure 2.1 to calibrate and validate our model. Note that data shows only the population adult snails (young snails are too tiny to be counted in data collection). The first year is used for calibration, while the remaining two serve to validate model predictions.

Our calibration procedure involves 2 steps. The first consists of choosing 10,000 random values of parameters uniformly filling a hypercube as described in Table 2.2. Of those we select the 289 ‘good choices’ with error $\varepsilon \leq .2$. On the second MC step, each of 289 first choices had its range extended from a single value $a_i$ to an interval $(.8a_i,1.2a_i)$, from which another 100 random values are tested. So the 2$^{nd}$ step consisted of 29189 choices of parameters. Once again we compute errors (2.9) and select those values of parameters for which the error $\varepsilon \leq .15$, thus arriving at 516 choices. Finally we compute
the marginal conditional distribution of each parameter and generate their histograms (Figure 2.4). Most of the parameters exhibit localized pattern around certain (most likely) values, which can also been seen by examining the mean and standard deviation of each parameters (Table 2.3). In the following part of the analysis, we use the set of 516 good choices of parameters.

Table 2.3 Mean and standard deviation of parameters based on the 516 ensemble from the 2nd stage of MC simulation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_Y$</td>
<td>3.36641</td>
<td>1.85615</td>
</tr>
<tr>
<td>$p_X$</td>
<td>9.67214</td>
<td>3.57578</td>
</tr>
<tr>
<td>$f_0$</td>
<td>0.367057</td>
<td>0.0742765</td>
</tr>
<tr>
<td>$f_1$</td>
<td>0.110497</td>
<td>0.0526611</td>
</tr>
<tr>
<td>$q$</td>
<td>9.14996</td>
<td>3.13432</td>
</tr>
<tr>
<td>$b_0$</td>
<td>231.902</td>
<td>128.184</td>
</tr>
<tr>
<td>$V_0$</td>
<td>19.3521</td>
<td>4.16016</td>
</tr>
<tr>
<td>$\theta$</td>
<td>6.83202</td>
<td>1.02326</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>8.79472</td>
<td>1.84833</td>
</tr>
<tr>
<td>$F_0$</td>
<td>0.0817582</td>
<td>0.0316966</td>
</tr>
</tbody>
</table>

For the validation part, we run simulations of each good choice over a 3-year period and compare the prediction with the last two years of data. The ensemble of predicted histories generates a band that contains most of data points and follows closely the seasonal pattern (Figure 2.5). We conclude that our model with estimated parameters
captures the essential (resource-dependent) snail population dynamics and can be adopted for the follow up studies of snail control strategies.

Figure 2.4 Marginal parameter distributions based on 516 best choices of the 2nd Monte-Carlo run
Given any rainfall data as input, we are able to predict the snail population using our calibrated model. For example, we took precipitation of the period February 2005 to January 2010 (available in Mathematica data base), and produced the adult snail population shown in Figure 2.6.
Figure 2.6 Adult snail population rescaled by dividing 1000 (red solid curve) produced by the two-stage snail population model with rainfall data (black) in the same area as the original data over 5-year period starting from February 2005. The dashed purple curve represents the corresponding habitat $V(t)$.

2.4 Snail control and analysis of optimal strategy

Snail control is currently considered a viable option to supplement human chemotherapy. Such control can be easily accommodated in our model through an additional loss term. Clearly, snail control will reduce the total number of snails and their shedding fraction proportionally.

Here we adapt our two-stage snail model to examine the effect of snail control on snail population. Such analysis can give some crude estimate of the effect of snail control
on transmission. Indeed, the per capita force of human infection in simple Macdonald-type transmission models is proportional to snail population.

The additional loss rate due to snail control takes the form

$$\frac{1}{T} \ln(1 - \phi)$$  \hspace{1cm} (2.10)

where $T$ is the duration of the program and $\phi$ is the efficiency of the program, namely the fraction of snail eliminated during the program. This loss term is added to the natural loss terms (mortality) over the duration of the program. Given large seasonal variations, the timing and duration of control can have significant effect on its outcome. Here we search for an optimal seasonal control setting in terms of the overall population reduction.

To this end we take the best 516 of the selected parameter values obtained in the second MC run. For each choice, we solve the ODE system (2.4)-(2.6) under different choices of control timing (the time when a snail control is started) within the first year. The other control parameters are set constant: $T = 2$ weeks and $\phi = 0.9$ for all the simulations. Different timing of control has varying impact on snail population. In Figure 2.7, we showed the results of four choices of timing, from the beginning to the end of the year. In all cases snail population is reduced significantly during control program, but very typically, it quickly rebounds. The resurgence rates depending on the choices of parameter. Indeed, the envelope widths widen during snail control runs, but they revert to normal patterns after longer period.

To compare the effect of various control scenarios, we use the ratio of seasonal mean snail densities with and without control. We do this by checking the solution
ensemble based on the first 100 best choices of parameters. Figure 2.7 shows the envelope with different timings.

![Graphs showing snail control with different timings.](image)

**Figure 2.7** Snail control with different timings. The envelopes are created with the 100 best parameter choices.

Based on our 100 ensemble of parameters, we find that a single optimized snail control can reduce the one-year average of snail density by about 32% (Figure 2.8). The efficacy of snail control in terms of seasonal mean reduction is not significant since the snail growth recovery rate of the snail population is overall very high. But we can still get some useful information for setting up control strategy. By studying the minimum of the mean of normalized one-year average, we find the optimal timing is 0.37 in 1 year scale (between middle-May to end-June during a year), which suggests that the optimal timing
of snail control should be at the end of first rainy season. This matches the result in Jordan book [54] for static habitats in St Lucia region in eastern Africa.

![Graph of Relative annual mean snail density reduction for different control timing. The envelop is made from mean and standard deviation (std) of the solution ensemble based on 100 best choices of parameters: (mean-std, mean + std) with the solid curve in the middle being the mean.](image)

Compared to drug treatment on humans, snail control can prove a more convenient and economical means to reduce overall transmission efficiency in a given environment[55]. It is generally observed that the frequency of snail treatments need be much higher than once a year to be effective. We would like to the compare the outcomes of control with different frequencies: biweekly, monthly, quarterly based on the same setting. Figure 2.9 shows different effects of such control strategies. We see that biweekly control is so intensive as to bring snail density down to nearly zero, and monthly control is good enough to reduce the snail density by more than 75%.
In practice, the ultimate goal of snail control is to bring down the schistosomiasis transmission. In the next step, we should take our snail population model and couple it with schistosomiasis transmission. In reality, the human-snail environment is distributed and interconnected with multiple snail habitats (multiple lakes or ponds) and human residential clusters. In this case, each snail habitat has its own snail dynamics driven by the seasonal environmental factors. The analysis of snail control based on this extended model would be more sophisticated and suitable for transmission intervention.
Chapter 3  Dynamics of *Schistosoma* infection and transmission for human hosts

3.1 Review of traditional approaches

A typical schistosomiasis transmission model is a coupled system of two parts: one describes the infection dynamics for human hosts and the other describes the dynamics of for intermediate snail hosts; these are linked through the infection transmission forces from each other.

3.1.1 Mean Worm Burden (MWB) approach

The simplest and most influential homogeneous model was proposed by Macdonald in 1965 [23]. Assuming that all individuals in the host population have the same infection-related characteristics (susceptibility, contact frequency, and rate of worm establishment etc), it exploits a mean worm burden (MWB) formulation for human hosts and a suitable infection prevalence estimate for intermediate snail hosts. This can be described by a coupled system of differential equations

\[
\frac{dw}{dt} = (\alpha \eta N) y - \gamma w = Ay - \gamma w
\]

\[
\frac{dy}{dt} = (\beta \eta H) w(1 - y) - \mu y = Bw(1 - y) - \mu y
\]

(3.1)
where the variables $w(t)$ represents mean worm burden of the human population and $y(t)$ prevalence of infectious snails. It assumes homogeneous stationary host populations $H$ (for human) and $N$ (for snail density) and fixed (time independent) transmission parameters.

**Table 3.1 Parameters of basic MacDonald system (3.1)**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>Worm establishment probability per contact per shedding snail</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Contamination rate per susceptible snail per contact per adult worm</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Human-snail contact rate</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Mortality of worm in human host</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Mortality of shedding snail</td>
</tr>
</tbody>
</table>

The equilibria of system (3.1) under stationary conditions depends on the basic reproduction number (BRN), $R_0$, which, in the case of macroparasite, is defined as the average number of offspring produced throughout the reproductive life span of a mature parasite that themselves survive to reproductive maturity in the absence of density-dependent constraints on population growth (see [21, 28, 46, 51]). Here, it can be described through the following expression:

$$R_0 = \frac{\alpha N}{\gamma} \frac{\beta H}{\mu} \eta^2 = \frac{AB}{\gamma \mu}, \text{ with } A = \alpha N \eta \text{ and } B = \beta H \eta;$$

If $R_0 > 1$, the system has a stable endemic equilibrium:

$$\begin{align*}
    w^* &= \frac{A}{\gamma} (1 - \frac{1}{R_0}) \\
    y^* &= 1 - \frac{1}{R_0}
\end{align*}$$

44
and a unstable disease free equilibrium \((w^*, y^*) = (0,0)\). This means that the transmission, once initiated, will go to the endemic equilibrium level and stay there. If \(R_0 < 1\), the system has only one stable disease free equilibrium, indicating that transmission will stop. This result shows that the disease can be eradicated when the relevant controls reduce \(R_0\) to less than 1. From this analysis, we see that possible ways to control the disease are mollusciciding (reducing snail density \(N\)), chemotherapy (increasing worm mortality \(\gamma\)) and health and hygienic education (reducing contact rate \(\eta\)).

Since not all worms can be successfully mated and produce eggs in reality. It is natural to improve the MWB model (3.1) by including a mating probability \(\phi(w)\). For illustration, let’s assume that

\[
\phi(w) = \frac{w}{w_0 + w},
\]

with \(w_0\) being a threshold under which worm mating is rare to happen. The effect of this probability is modeled by replacing \(w\) in the second equation of (3.1) with

\[
f(w) = \frac{w^2}{w_0 + w}.
\]

If \(R_0 > 1 + 2w_0\sqrt{\frac{B}{\mu}}\), then there are three equilibria for the modified system:
The first and the third equilibria are stable, and the second one is unstable. This has significant implications in disease control practice. If we can bring the mean worm burden below \( w_1 \) (not necessary total clearance), then the system will approach the disease free equilibrium eventually.

While simple and amenable to direct mathematical analysis, the biggest drawbacks of the Macdonald model are oversimplified and its unrealistic assumptions of host populations and transmission patterns. An over-dispersed pattern of infection intensity in human populations should be taken into account to make a model descriptive of the real world problems to provide reliable disease control guidelines.

3.1.2 Distributed worm burden: Negative Binomial approach

As an improvement of the MWB methodology, empirical probability distributions have been used to fit field data to describe the distribution of infection intensity (or worm burden) in a human population [32, 45, 56]. Among them, the negative binomial (NB) distribution has been the most widely used [32, 45, 57]. Figure 3.1 serves as an illustration of Negative Binomial fit of frequency of macroparasitic worms in a population. A negative binomial distribution can be characterized by two parameters: the mean, \( m \), and clumping parameter, \( k \). Its variance is \( m + m^2 / k \). The parameter \( k \) takes
positive values. The smaller $k$ is, the more overdispersed is the distribution. The negative binomial distribution approaches the Poisson distribution as $k \to \infty$. Under the assumption that the infection intensity in a host population follows a negative binomial distribution with mean $m$ and overdispersion parameter $k$, a series of formulas may be deduced, including the prevalence of infection:

$$p(m,k) = 1 - (1 + m/k)^{-k};$$

and mating probability of worms[24] or worm acquisition rate[20]:

$$\phi(w,k) = 1 - (1 + w/2k)^{-(k+1)}.$$  

These formulas have been proved handy for parameter estimation and for linking different stages of parasites. For example, the value of $k$ can be estimated by fitting data.

The introduction of a probability element makes it possible to account for the overdispersion of infection in human hosts and to perform systematic mathematical/statistical analysis. However, statistically fitting data cannot explain the underlying causal mechanism. It leaves many uncertainties about its application to general epidemiological data and its predictions about the impact of control programs. Clearly, a new modeling approach to explain the dynamic relationship between the essential factors in transmission and the resulting over-dispersed worm patterns should be developed to provide policymaker with more reliable control strategies.
3.1.3 Heterogeneous human populations

Instead of individual human hosts, we can divide population into different groups by sex, age or occupation. People from the same group of the population tend to have similar behavioral pattern and biological characteristics. Based on either MWB or NB
methodologies, one can develop heterogeneous models. Some earlier works (e.g. [24, 59]) used age-structured MWB to assess infection-associated morbidity and the effects of drug treatment for idealized host populations. Other works [17, 22, 30, 34-38] addressed heterogeneities in patterns of exposure and contamination through the so called contact rate matrix. Barbour (1978)[34] first developed model with multiple water contact sites

\[
\begin{align*}
\dot{w}_i &= \alpha \sum_{j=1}^{Q} \eta_{ij} N_j y_j - \gamma w_i \\
\dot{y}_j &= \beta \left( \sum_{i=1}^{H} \eta_{ij} w_i \right) (1 - y_j) - \mu y_j
\end{align*}
\]

(3.3)

Here, \( w_i \) is the mean worm burden of the \( i \)-th host population (\( i=1,2,\ldots,M \)), \( y_j \) is prevalence of shedding snails at the \( j \)-th site (\( j=1,2,\ldots,Q \)), and \( \eta_{ij} \) is the contact rate of \( i \)-th host site to \( j \)-th water site. Other parameters are exactly as described in Table 3.1.

The model by D. Gurarie et al. [17] accounts for the dynamics of transmission in a distributed environment consisting of multiple human population sites (villages) and snail population sites. In this environment, each site of aggregated human residences can be thought of as a village and population of each site is stratified by age with time-step \( \Delta t \). The whole human population can be represented by a matrix \( \bar{H} = (H_{r,a})_{M \times P} \) where \( M \) is the number of human residence sites and \( P \) is the number of age strata in each site. The snail populations were distributed among multiple water sites and densities were assumed constant labeled by \( \bar{N} = (N_j)_Q \) where \( Q \) is the number of water contact sites. In their model, they assumed that contact rate is a product of age-behavioral and spatial factors. Contamination rates, worm establishment probability and worm mortality rate are
age-dependent because of the difference in age-behavioral pattern, acquired immunity or other age-protective mechanics. Among different water sites, snail densities and mortalities are site-dependent due to variance of environmental factors. The system of differential equations in the model takes the form

\[
\dot{W}_{i,a} = \alpha_a \sum_{j=1}^{N_{i,a,j}} N_{j,y_j} y_j - \gamma_{a,i} w_{i,a} + \frac{1}{\Delta a} \left( \frac{H_{i,a-1}}{H_{i,a}} w_{i,a-1} - w_{i,a-1} \right) \\
\dot{y}_j = \left( \sum_{i,a} \beta_a \eta_{i,a,j} H_{i,a} w_{i,a} \right) (1 - y_j) - \mu_j y_j
\]  

(3.4)

The parameter symbols are similar to those that we used before, only now furnished with indices [38, 39]. The state variable \(w_{i,a}\) represents mean worm burdens of the specific age group of human population at village \(i\) and \(y_j\) represents shedding snail prevalence in water site \(j\). With this model, we can analyze the age-specific or site-specific chemotherapy strategies and identify the high-risk villages and age-groups. One of the drawbacks of this model is that it ignores the over-dispersed worm distribution among human hosts.

### 3.2 Stratified worm burden (SWB) approach

*Stratified-Worm-Burden* (SWB) system is an alternative to the classical MWB approach. Several papers have studied such or similar systems for macro-parasites in the past [39-41], but most of this work aimed at deriving reduced (MWB-type) equations, or their extended versions to couple MWB to higher moments (variance and/or aggregation)
of SWB distribution. Among several examples are papers on: (i) regulation and stability of host–parasite populations [60, 61]; (ii) aggregation for single parasite species and multi-strain competition, coexistence, invasion [62-64]; and (iii) parasite strain diversity and evolution of drug resistance [65]. In contrast, for the present study, we develop and implement an extended SWB approach to model, as necessary, the complexity of schistosomiasis transmission for structured host populations in a geographically distributed environment with varying water contact patterns. We take advantage of the extended model systems to consider age-structured communities made of children and adult groups having suitable rates of birth, mortality, and aging. Finally, we combine age-behavioral and spatial heterogeneities of human population with two-stage snail model, and apply it for the specific environment of Msambweni area of the Kwale District in Kenya.

We start with an outline of general SWB-scheme, its implications for burden/prevalence distribution, and the analysis of endemic equilibriums. Then we proceed to develop a model calibration scheme that exploits typical demographic and epidemiological infection data to estimate the baseline human-to-snail and snail-to-human transmission rates. These models are built upon each other, starting with a model describing a single homogeneous population, adding an age-structured (adult-child) population, and finally a geographically distributed (human-snail) transmission environment.
Our choice of calibration scheme is partly motivated by the available data from the Msambweni region of eastern Kenya, which were collected in several multi-year studies [66-70]. In a sequel chapter we will apply our methodology (SWB model and calibration scheme) to this specific environment, comprised of 10 villages, each one subdivided into adult and child subgroups with 5 shared water contact (snail) sites. When calibrated and validated, this model then allows us to make projections on the transmission of vector-borne disease, and to simulate the effects of hypothetical control interventions based on single or combined drug therapy and vector control schemes.

3.2.1 Method

The host population $H$ is divided into burden strata $\{h_i\}$ based on the number of parasites they carry.

<table>
<thead>
<tr>
<th>Infection level</th>
<th>0</th>
<th>1</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worm burden</td>
<td>$0 &lt; w &lt; w_i$</td>
<td>$w_i &lt; w &lt; w_{i+1}$</td>
<td>...</td>
</tr>
<tr>
<td>Population strata</td>
<td>$h_0$</td>
<td>$h_i$</td>
<td>...</td>
</tr>
</tbody>
</table>

So each group $h_i$ viewed as a local subpopulation carries a parasite range $w_i < w < w_{i+1}$.

The dividers $\{w_i\}$ can be chosen arbitrarily, the simplest (extreme) choice being ‘zero-nonzero’ (‘infected’ – ‘non infected’) strata with prevalences $\{h_0, h_i\}$. Such partition would result in the standard SIR-type prevalence model. The SIR methodology, however, is too crude for macro-parasite transmission, where force of infection depends primarily on the worm burden of the host population rather than the infection frequency. On the
opposite extreme is continuously distributed worm burden \( h(w) : 0 \leq w < \infty \), which has limited practical applications outside mathematical analyses. Our stratified worm burden (SWB) models extends the SIR (prevalence-based) modeling to accommodate both the worm burden and infection frequency relevant to population dynamics of macroparasite infections using the Macdonald-type “mean burden” formulation, as explained below.

In what follows, we shall use uniform strata \( w_i = i\Delta w (i = 0, 1, 2, \ldots) \) counted in steps \( \Delta w \) worms. The choice of \( \Delta w \) is related to infectivity threshold \( w_r \) - the ‘effective’ number below which population is considered free of infection. It is convenient to take \( w_r \) as a multiple of \( \Delta w \; ; \) thus \( w_r = \Delta w \) means \( h_0 = \) ‘uninfected strata’. For example, if \( w_r = 5 \) then the first stratum \( h_0 = \) ‘uninfected’. Indeed, host infectivity is proportional to the number of mated worm pairs, which in turn depends on their mating probability \( \phi(w) \) ([71], see also [24]). Derivation of such a probability function is based on a ‘negative-binomial’ assumption, suggesting a somewhat low threshold value \( w_r = 5 \) (independent of aggregation). On the other hand, experimental/field data for primate infection predicts a substantially higher threshold for egg detection (\( w_r \approx 40 \)).

We assume the infectivity threshold to lie somewhere in between \( 10 \leq w_r \leq 20 \), and numerically explore 2 values: \( w_r = 10 \) and \( w_r = 20 \). Accordingly, we choose to convert the continuous values to discrete values at 10 worms per step (discretization step, \( \Delta w \)). While the choice of step \( \Delta w \) is not essential for calibration and analysis (as explained below), the choice of threshold \( w_r \) has a significant effect (will talk about it later).
For higher burden levels \((w > 10)\) we assume host infectivity as measured by egg counts, is roughly proportional to \(w_i\). The choice of uniform strata is not essential from the computational standpoint, but has some mathematical advantages. We used 15 strata (i.e. burden levels up to 150 worms/host) in our analysis, which, according to our data set, is sufficient to cover a broad range of human infection, even heavily infected populations.

Figure 3.2 shows the diagram of transitions among strata with ‘up-arrows’ \(\rightarrow\) designating the *force of human infection* \(\hat{\lambda}\) (determined by infected snail/ larval densities, water contact patterns, probability of worm establishment, etc.), and ‘down arrows’, \(\leftarrow\) indicating *burden resolution* \(\gamma_i\) (due to worm mortality or other causes, including drug treatment). Different strata \(i\) have specific resolution rates \(\gamma_i\) that increase with burden level. Additional source terms, \(S_i\), represent new recruits to the infection pool (newborns, migrants into the area, and transition from younger to older age groups). Loss rates, \(\mu\), combine host mortality as well as out-migration, aging, etc. In some papers \(\mu\)’s are considered to depend on level \(i\) to account for ‘burden-specific’ mortality. We ignore disease-specific mortality as quite small, *i.e.*, not having impact on population levels or community levels of schistosomiasis (even its more pathogenic *S. japonicum* species [72]). So our \(\mu\) is purely demographic and considered identical for homogeneous populations.
Figure 3.2 Infective strata and transitions of SWB model: population strata (numbers or fractions of total) \( \{ h_i \} \) move ‘up’ or ‘down’ the scale at certain rates, due to worm acquisition (\( \lambda \)) and resolution (\( \gamma_i \)). Each stratum has its source-term \( S_i \), and loss rate \( \mu \).

For age-structured populations (e.g., children and adult subgroupings) each group has its own infection strata, but their source terms differ. The children’s source terms come primarily in the lowest (uninfected) strata \( S_0 \) (as newborns are free of infection). But the adult source, \( \{ S_i \} \), will be distributed over all strata and depend on the burden distribution among children by the time they age into adulthood.

For a single (homogeneous) population \( H = \sum_{i=0}^{n} h_i \), variables \( \{ h_i(t) \} \) obey a system of differential equations (SWB)

\[
\begin{align*}
\frac{dh_0}{dt} &= - (\lambda + \mu) h_0 + \gamma_1 h_1 + S_0 \\
\vdots \\
\frac{dh_i}{dt} &= \lambda h_{i-1} - (\lambda + \gamma_i + \mu) h_i + \gamma_{i+1} h_{i+1} + S_i \\
\vdots \\
\frac{dh_n}{dt} &= \lambda h_{n-1} - (\gamma_n + \mu) h_n + S_n
\end{align*}
\]

(3.5)
We used the following notations: (i) transition rate $\lambda$ is proportional to (per capita) force of infection: $\lambda = \frac{A_z}{\Delta w}$, made of ‘transmission coefficient’/’infected snail’ - $A$, infected snail density - $z$, and $\Delta w$ - ‘burden increment’ in the definition of strata $\{h_i\}$; (ii) resolution rates $\{\gamma_i = \gamma t\}$ are proportional to burden levels $\{w_i\}$ of each strata ($\gamma$ being natural worm mortality); (iii) host loss rate(s) = $\mu$, and new recruit sources = $\{S_i\}$.

Here we assume identical force of infection and loss for all strata, i.e., a homogeneous community in terms of behavior and environmental factors. This, however, does not preclude uneven (aggregated) burden distribution within system$(3.5)$. Figure 3.3 (A) shows a few equilibrium distributions of human prevalence $\{h_i(t)\}$ for several values $\lambda$ contributed by a stationary source $S_0$ at the lowest (uninfected) stratum (a typical pattern for child group). Figure 3.3(B) illustrates dynamic relaxation (to equilibrium) of the coupled SWB-system $(1)$ and snail prevalence $y(t)$ (see Appendix A). Here, human force of infection $\lambda = Ay$, and snail force $A$ is proportional to the cumulative worm burden $(2)$ of the human population. For dynamic simulation, we initialized the coupled system at a low-level human infection (light bars) – a hypothetical effect of ‘drug clearing’, and let it relax to its natural equilibrium aggregated at higher burden levels (black bars).

System $(3.5)$ can be interpreted as cumulative outcome of a stochastic process of ‘worm acquisition-death’ on individual level. Indeed, the RHS of system $(1)$ has a matrix
made of ‘subdiagonal’ $\lambda$ and ‘diagonal’ $\{r_{jj}\}$ that can be viewed as discretized Markov generator of such stochastic process. In this sense the standard MWB [23] gives the ‘statistical mean dynamics’ for such process (as function of time), whereas our SWB model gives its complete probability distribution function.

Figure 3.3 (A) Equilibrium worm burden distribution for different values of the worm acquisition rate $\lambda$: .2 (black), .6 (medium gray) and 1 (light gray), with identical human loss rate $\mu=.02$ and stationary source at the lowest strata, (B) Time relaxation of SWB-system coupled to infected snail prevalence. Left panel shows prevalence distributions of 15 human strata $\{h_i\}$ at $t=0$ (black), $t=20$ (medium gray) and $t=40$ (light gray). For snail prevalence we show the dynamic relaxation pattern (right panel).
Human system (3.5) can be coupled in the usual way to snail infection (pre-patent and patent groups), through the force of human infection $\lambda$.

\[
\begin{align*}
\frac{dh_0}{dt} &= -\left(\lambda + \mu\right)h_0 + \gamma_i h_i + S_0 \\
\vdots
\frac{dh_i}{dt} &= \lambda h_{i-1} - \left(\lambda + \gamma_i + \mu\right) h_i + \gamma_{i+1} h_{i+1} + S_i \\
\vdots
\frac{dh_n}{dt} &= \lambda h_{n-1} - (\gamma_n + \mu) h_n + S_n
\end{align*}
\]  

(3.6)

\[
\text{(Snail)} \quad \frac{dy}{dt} = \Lambda(1-y) - \mu y
\]

Here the force of snail infection $\lambda$ will depend on human MWB, similarly as in the standard Macdonald approach[23]. Unlike the Macdonald system, our MWB is not an independent variable but computed from prevalence variables $\{h_i\}$

\[
w = (\Delta w) \frac{1}{H} \sum_{i=1}^{n} i h_i
\]  

(3.7)

In the follow up analysis and calibration we rescale the SWB-system (3.5) and MWB (3.7) using non-dimensional (primed) variables and parameters, along with time and source terms. Specifically,

\[
w \leftrightarrow w' = \frac{w}{\Delta w} = \{0,1,2,...\} \quad \text{-dimensionless burden}
\]

\[
h_i \leftrightarrow h'_i = \frac{h_i}{H} \quad \text{-prevalence strata}
\]

\[
S_i \leftrightarrow S'_i = \frac{S_i}{(\Delta w)H} \quad \text{-source terms in prevalence strata}
\]
Then the \(i^{\text{th}}\) equation for system (3.5) becomes

\[
\frac{d h'_i}{dt} = \lambda h'_{i-1} - (\lambda + \gamma i + \mu) h'_i + \gamma(i+1) h'_{i+1} + S'_i,
\]

for the sake of brevity we shall drop ‘prime’ from all rescaled variables.

The SWB system yields several reduced equations based on statistical moments of distribution \(\{h_i\}\). Among them are (i) ‘total population dynamics’ for \(H\) (0-moment (sum)),

\[
\frac{dH}{dt} = S - \mu H \tag{3.8}
\]

with \(S = H \sum_{i=0}^{n} S_i\); (ii) Macdonald-type MWB equation for \(w\) (1\textsuperscript{st} moment (mean))

\[
\frac{dw}{dt} = \lambda(1-h_n) - (\gamma + \mu)w + \bar{S} \tag{3.9}
\]

Here \(\bar{S} = \sum_{i=0}^{n} iS_i\) designates the cumulative source. Let us note that \(\bar{S} = 0\), if new recruits enter system (3.9) only through its non-infected (zero) compartment (like newborn children). Derivation of (3.9) exploits a special form of resolution rates \(\gamma_i = \gamma i\) for \(i\)-th strata). So worms in all hosts are assumed to have the same (mean) life-span, independent of the host state (age, burden level etc.). We also ignore all other confounding factors, like density-dependent mortality (due to crowding), or immune regulation of parasites. Such factors can, in principle, be included in the SWB scheme, but to calibrate such a model would require a more detailed demographic/infection data of host population (e.g. age, behavioral risk, susceptibility etc.), than provided in most community studies of schistosomiasis. Our formulation tries to keep what we believe are
the most essential components of the system, suited for the most frequently available data.

Equation (3.9) looks like the standard Macdonald MWB (3.1) for dimensionless $w(t)$ with an additional term $h_n$ that disappears in the infinite limit systems ($n \to \infty$), used in many theoretical studies. In our case (finite size SWB) this term remains positive, and sufficiently small over a wide range of parameter values, to be safely dropped in the follow-up analysis and calibration. Figure 3.4 shows negligible difference (two curves are overlapped by each other) between the two systems when the values of parameters are set to be same.

![Figure 3.4](image)

Figure 3.4 Mean worm burden and snail prevalence from SWB and standard MacDonald system with the same parameter values. The step size of worm burden $\Delta w$ is set to be 5 and number of strata equals 30.

Equations (3.8) (total population growth) and (3.9) (MWB dynamics) represent the first two statistical moments of the distribution $\{h_i\}$ driven by the SWB system (3.5). It can be further extended to higher moments: variance, aggregation etc. Such schemes were adopted in some theoretical studies (e.g. [39]) as a way to reduce a large (infinite)
system to its low-dimensional (mean-field) approximation. In this chapter, our goal and approach is different. We utilize reduced (moment) equations (3.8) and (3.9) for model calibration. But all dynamic simulations (for prediction-analysis-control) are run with full SWB system (3.5), using its calibrated parameters. Thus a system of 10 villages in a particular study area, with each population subdivided into child and adult groups, will contain 2x10x15 ‘human’ variables plus a suitable number of ‘snail’ variables proportional to the number of contact sites.

It should be stressed that increased order (dimension) of SWB system (compared to its MWB cousin) posed no substantial mathematical or computational obstacle by itself. Indeed, we can generate such a large, coupled dynamical system using Mathematica, which also allows efficient ways to solve and manipulate them in various ways for analysis, calibration, prediction and control experiments.

The equilibrium solution of system(3.5), with stationary source S and prescribed \( \lambda \), depends on two parameters \( \mu \) and \( \lambda \). The former \( \mu \) is determined by known demographic or biological factors; the latter \( \lambda \) (proportional to the force of human infection) can vary depending on the transmission environment, and will be calibrated along with other parameters.

The complete SWB system (3.6) has no exact explicit solutions, but we can solve it numerically for a broad range of parameters \( \{ \lambda, \mu \} \). Figure 3.3(A) shows equilibrium burden distributions for several \( \lambda \). Qualitatively they resemble negative binomial (NB)
distributions of decreased aggregation (smaller $\lambda$ having higher aggregation, hence a larger fraction of low-level infections).

MWB equation (3.5) can be used either in dimensional form for variable $w(t)$, or non-dimensional $w = \sum_{i}^{n} i h_i$. Our calibration scheme exploits two epidemiological data values, uninfected human prevalence $P = h_0$, and MWB, $w$ (which is related to measured egg-counts in the field data). Both are functions of parameters $\lambda, \mu$, and source $S$

$$P = \phi(\lambda, \mu, S)$$
$$w = \sum_{i}^{n} i h_i = \psi(\lambda, \mu, S).$$

Figure 3.5 (A) Plot of uninfected prevalence function $P = \phi(\lambda, \mu, S)$ on worm acquisition rate $\lambda$ for 3 values of $\mu = \{.1\gamma, .2\gamma, .3\gamma\}$ (top to bottom) and identical stationary source $S$ at the lowest strata; (B) Plot of MWB-function $w = \psi(\lambda, \mu, S)$ on $\lambda$ for the same values $\mu$ of as (A) (top to bottom) with dashed lines corresponding to a would-be ‘linear relation’ (slope $\frac{\partial w}{\partial \lambda} \approx \frac{1}{\gamma + \mu}$ at $\lambda = 0$ (appropriate for Macdonald MWB).
The functions $\phi$ and $\psi$ are plotted for different values of $\mu$ in Figure 3.5. While there are no explicit formulae, they can be easily computed numerically by interpolation (sampling a range of values $\lambda$ and $\mu$). These procedures are used extensively in our calibration schemes below.

Let us note that the range of uninfected human prevalence from our data (Table 2), $0.2 < P < 1$, constrains $\lambda / \gamma$ within relatively narrow interval $[0, 2.5]$ (Figure 3.5(A)). In that range, both $\lambda$ and $w$ obey a nearly linear (Macdonald type) relationship (Figure 3.4 (B)),

$$\lambda \approx (\gamma + \mu)w$$

which is important for our calibration scheme. At large $\lambda > 8$ (a strong force of infection) we observe nonlinear departure from simple Macdonald model outcomes (Figure 3.5 (B)), due to the extra-term of our MWB equation. But this range (and level of transmission) is well above the likely range of values in our study environment.

### 3.2.2 Model calibration

Data that are typically available for calibrating models of macro-parasite transmission include: (i) Demographics: human population $H$, and snail density $N$; (ii) Biological factors: human loss rates (mortality, turnover) $-\mu$, snail mortality $-\nu$, worm mortality $\gamma$; (iii) Epidemiological infection data: infected and patent snail prevalences ($y$, $p$).
z), human egg-counts (arithmetic and geometric means) $E, G$. Depending on the choice of model, all or part of these data can be used for estimating unknown transmission parameters (see, e.g. [73]).

In our calibration scheme below, we use both egg-counts (arithmetic mean) and infection prevalence as two independent parameters. For the sake of exposition we precede the discussion of larger, distributed systems, with a few special cases.

The general scheme involves 3 steps:

i) Estimate human force of infection $\lambda$ (for suitable populations) from demographic and human prevalence data. This part utilizes prevalence function $\phi(\lambda, \mu, S)$

ii) Next estimate snail-to-human transmission coefficient $\{A\}$ for appropriate population groups, as well as biological parameter $\rho$ (egg-production/worm).

iii) Finally, estimate human-to-snail transmission coefficients $\{B\}$ (Table 3.2).
### Table 3.2 List of notations, parameters and their description in MWB system(Appendix A.1) for a single human and snail population.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>Human population density</td>
</tr>
<tr>
<td>$E$</td>
<td>Egg release/ host (arithmetic mean egg count /10ml urine)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Human mortality (turnover) rate/year</td>
</tr>
<tr>
<td>$N$</td>
<td>Snail population density per unit habitat</td>
</tr>
<tr>
<td>$y; Y$</td>
<td>Infected snail prevalence and snail density (per unit habitat)</td>
</tr>
<tr>
<td>$z; Z$</td>
<td>Infectious(patent) snail prevalence and snail density / habitat</td>
</tr>
<tr>
<td>$T$</td>
<td>Total contact per person to all water sites for villagers</td>
</tr>
<tr>
<td>$\eta_{ij}$</td>
<td>Fraction of total contacts from village $i$ to snail habitat $j$</td>
</tr>
<tr>
<td>$a$</td>
<td>Probability of worm establishment per cercariae per contact</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Human growth rate (estimated .05 /year)</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Transition rate from youth to adulthood (estimated .05/year)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Worm mortality (life span: 5 years)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Snail mortality( life span: 9 weeks)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Released egg production rate per worm</td>
</tr>
<tr>
<td>$\psi$</td>
<td>‘Patency conversion’ fraction of infected snails</td>
</tr>
<tr>
<td>$\pi_c; V_c$</td>
<td>Rate of cercariae output/shedding snail; cercariae mortality</td>
</tr>
<tr>
<td>$\pi_m; V_m$</td>
<td>Miracidia production rate/worm and miracidium mortality</td>
</tr>
<tr>
<td>$b$</td>
<td>Probability of snail infection per unit miracidium density</td>
</tr>
<tr>
<td>$A = aT(\frac{\pi_c}{V_c})\psi$</td>
<td>Snail-to-human transmission coefficient (can be site specific)</td>
</tr>
<tr>
<td>$B = bT(\frac{\pi_m}{V_m})\rho$</td>
<td>Human-to-snail transmission coefficient (can be site specific)</td>
</tr>
</tbody>
</table>
3.2.2.1 The single population case

In this case our calibration scheme will utilize two reduced equations resulting from SWB (3.5): (i) MWB (3.9) coupled to a snail prevalence equation (as in the standard Macdonald setup), and (ii) an equilibrium zero-prevalence function(3.10), computed numerically. The former gives a coupled system of 2 differential equations for $w(t)$ and $y(t)$

\[
\frac{dw}{dt} = \lambda \Delta w - (\gamma + \mu)w; \\
\frac{dy}{dt} = \Lambda (1 - y) - \nu y;
\]

with per capita forces of infection

\[
\lambda \Delta w = ANy \\
\Lambda = BHw
\]

The transmission coefficients $A$ (snail to human) and $B$ (human to snail) depend on several biological and environmental/behavioral factors (Table 3.2). Let us note that having stratified human population is irrelevant for snail infection, as long as all strata have similar behavioral (water contact) patterns. So the force of snail infection $\Lambda$ is the same as in the classical Macdonald setup (i.e., proportional to total human burden, $H_w$)

System (3.11) has stable endemic equilibrium provided its Basic Reproduction Number (BRN)

\[
R_0 = \frac{ANBH}{(\gamma + \mu)\nu} > 1.
\]
Our goal is to estimate transmission coefficients $A, B$ from equilibrium solutions and the available data. The data include egg-count $E$ (related linearly to $w$) and snail prevalence $(y)$. The relation of $E$ to $w$ is given by another (unknown) parameter egg production rate

$$
\rho = \frac{\text{Egg release}}{\text{worm} \times \text{day}}
$$

Our goal is to calibrate parameters $A, B, \rho$. We proceed in two steps:

1) solve equation (3.10) for a given (known) human prevalence $P$, relative loss rate $\mu$, (demographics), and source $S$, to find rescaled force of infection $\lambda$.

2) compute $A$ from the first expression of (3.12), $\rho$ and $B$ equilibrium of equations (3.11)

The resulting explicit formulae give calibrated parameters in terms of the data

$$
A = \frac{\lambda(\Delta w)}{Ny}; \quad \rho = \frac{(\gamma + \mu)E}{\lambda(\Delta w)}; \quad B = \frac{\rho}{HE} \frac{vy}{(1 - y)}
$$

(3.13)

In reality, infection field data, either human (egg-counts $E$, prevalence $P$) or snail data (prepatent and patent prevalences $y, z$), can be highly inaccurate [68, 74] due to low sensitivity of the standard tests and observational errors. One way to account for data errors (underestimation) is through a suitable statistical ensemble of data values and the resulting parameters ($A, B, \rho$). Here explicit formulae (3.13) prove helpful, as any *a priori* information of data errors (statistics or distribution) can be translated via it into suitable estimates for calibrated parameters. Indeed, we shall explore sensitivity of our calibration scheme to data errors for a specific data set.


3.2.2.2 Multiple risk groups

Next we subdivide human population within a single village (or human - snail site pairing) into several risk groups labeled by index $s$, that differ in their behavior (contact rate) hence the resulting force of infection, $\lambda_s$. In the following, those will be children and adult groups. The extended system couples human transmission (either MWB equations for dimensional $\{w_s\}$ or the corresponding SWBs for each risk-group) and a single snail equation (contributed by all groups)

$$
\frac{dw_s}{dt} = A_s Ny - (\gamma + \mu_s) w_s; \\
\frac{dy}{dt} = \sum_s B_s H_s w_s (1 - y) - v y.
$$

As above, we use equilibrium solutions (3.14) along with the prevalence equation $\phi(\lambda_s, \mu_s, S_s) = P_s$, for each group. The calibration problem now is to estimate transmission coefficients $\{A_s, B_s\}$. The above two steps yield (i) the rescaled force of infection for each group $\{\lambda_s\}$

(ii) the corresponding A-coefficients

$$A_s = \frac{\lambda_s (\Delta w)}{Ny} \quad (3.15)$$

To get B-coefficients we observe that both sets (A’s and B’s) are proportional to contact rates $\{\eta_s\}$, assumed to be the only risk factor,
Here $a$ is worm establishment rate /per single infected snail, assumed to be identical for all groups, $b$ -snail infection rate /per single worm x contact. Then we get $B$-coefficients

$$B_s = \frac{A_s}{\sum_q A_q H_q E_q \rho_q} \nu_y (1 - y)$$  \hspace{1cm} \text{(3.17)}$$

expressed through group specific egg-production/worm rates

$$\rho_s = \frac{(\gamma + \mu_s) E_s}{\lambda_s (\Delta w)}$$  \hspace{1cm} \text{(3.18)}$$

Equations (3.15)-(3.17)-(3.18) extend the above calibration scheme (3.13) to heterogeneous populations with multiple risk groups.

### 3.2.2.3 Child - adult system

Next we further specify the above ‘multi-risk’ setup for separate child- adult groups, still within a single human-snail pairing context. These two human age-related subgroups differ in their risk factors (contact behavior) and infection levels. This case however, differs from the previous one, in that the two groups are also coupled by a demographic birth- aging process, whereby infection is carried from the younger to the older group. The population dynamics for variables \{ $H_c$, $H_a$ \} (total child/adult populations) is given by a Leslie-type differential system
\[
\frac{dH_c}{dt} = -(\mu_c + \tau)H_c + \beta H_a \\
\frac{dH_a}{dt} = \tau H_c - \mu_a H_a
\]

The demographic parameters include (i) mean per capita birth rate $\beta$, (ii) specific child and adult mortalities $\mu_c, \mu_a$, (iii) transition (aging) rate $\tau$.

Note that population dynamics (3.19) follow from the extended child-adult SWB system, the same way that the previous simple population model (3.8) follows from (3.5). The source term for child SWB, $S_c = \beta H_a$, enters through the uninfected (zero) stratum (newborns), whereas the adult source, $S_a$, is equal to children’s SWB distribution.

The real population and transmission dynamics are more involved. Clearly, different ages have different SWB – distributions, gradually changing from early childhood (aggregated at low burden) to older age-group patterns shifted towards higher $w$ (Figure 3.3). The adult source would then come from the older children’s distribution (15-20 yr), rather than the mean distribution over all ages (5-20yr). In principle, both systems (demographic and SWB) can be extended to accommodate refined population structure (e.g. 5-year bins). But here we confine ourselves to the coarse child-adult partition, as the available epidemiological data for our specific environment has only those two groups.

The corresponding child-adult Macdonald-type MWB equations account for demographics.

\textbf{70}
\[
\frac{dw_c}{dt} = A_c N y - (\gamma + \mu_c + \tau) w_c \\
\frac{dw_a}{dt} = A_a N y - (\gamma + \mu_a) w_c + \tau \frac{H_c}{H_a} w_c \\
\frac{dy}{dt} = (B_c H_c w_c + B_a H_a w_a)(1 - y) - \nu y
\]  \hspace{1cm} (3.20)

As in the previous (multiple risk) case, we utilize zero-prevalence to compute forces of infection,

\[
\phi(\hat{\lambda}_c, \mu_c + \tau, S^c) = P_c \Rightarrow \hat{\lambda}_c \\
\phi(\hat{\lambda}_a, \mu_a, S^a) = P_a \Rightarrow \hat{\lambda}_a.
\]  \hspace{1cm} (3.21)

Specifically, we first find children’s force of infection $\hat{\lambda}_c$ with known prevalence $P_c$ and source $S^c$ (in terms of birth rate), then compute the resulting equilibrium-prevalence distribution $\hat{h}_c = \{h_i^c\}_{i=0}^a$, substitute it to get adult source $S^a = \{\tau h_i^c H_c / H_a\}_{i=0}^a$ and get $\hat{\lambda}_a$ from the second equation in (3.21). Having estimated both $\hat{\lambda}$’s we proceed as in the previous section to get transmission coefficients $\{A_c : A_a ; B_c : B_a ; \rho_c ; \rho_a\}$ for two groups (see Appendix A for details).

### 3.2.2.4 Distributed system

Our next goal is to extend the above models and calibration scheme for geographically distributed systems made up of several different human sites (villages), labeled by $i = 1, 2, \ldots, n$, and connected snail (water), sites labeled by $j = 1, 2, \ldots, m$. Mathematically, calibration of such dynamical system is ill-posed (heavily under-
determined) problem as far as the available human and snail infection data. Indeed, the \( n+m \) equations of a typical ‘MWB + snail prevalence’ system ((Eq.A.8)) have altogether \((2n \times m)\) unknown transmission coefficients \(\{A_{ij}, B_{ij}\}\), if all connections are allowed [73, 75-77]. Different reduction schemes can be used to constrain the resulting parameter set. Thus Gurarie and King (2005) and Gurarie and Seto (2009) proposed a version of a landscape ‘gravity’ model, whereby geographic distances (or other hurdles) control human water contact rates. A suitable parameterization of this hurdle-function yields a drastic reduction of the parameter set.

![Figure 3.6 Partial area map of the Msambweni region (left), and its schematic view (right) with human sites (black dots), snail sites (gray dots) and established ‘human-to-snail’ links based on observed human contact activity.](image)

Other procedures are possible depending on the available data. For example, data collected from our specific environment, Msambweni region of eastern Kenya(Figure
3.6), [66-70]) has additional information in terms of relative contact rates from various human sites to water (snail) sites. Here we propose a distributed calibration scheme designed specifically for such type of data. Each contact pair \((i, j)\) contributes to human infection in village \(i\) and snail infection at site \(j\) through transmission coefficients \(A_j; B_j\) both proportional to absolute contacts per person from village \(i\) to water site \(j\), \(T_i \eta_{ij}\) with \(T_i\) being the group-specific total contacts to all water sites and \(\eta_{ij}\) relative contact \((\sum_j \eta_{ij} \leq 1)\). Our basic hypothesis is that each of them is a product of a suitable local (biological/environmental) factor \((A_j\) times ‘absolute contact rate’,

\[
A_j = A_0 T_i \eta_{ij} = A_0 \eta_{ij}; (i = 1, 2, ..., n)
\]

\[
B_j = B_0 T_i \eta_{ij} = B_0 \eta_{ij}; (j = 1, 2, ..., m)
\]  

(3.22)

It means that humans at site \(i\) will pick up infection from snail site \(j\) in proportion to their ‘absolute contact’ \(\times\) ‘the site’s snail infection level’. By the same pattern, snails at \(j\) will get infected by humans from different sites \(i\) in proportion to their absolute contact times human infection level (egg release). Since the establishment of miracidia in snails involves two stages- hatching of miracidia and the miracidium’s successful invasion into a snail body- which depends heavily on the local environmental conditions, we set the local environmental factor \(B_{0j}\) as site-specific. This hypothesis clearly reduces the number of unknown parameters to a manageable set \(\{A_i; B_j\}\), consistent with the number of MWB equations.
Accordingly, for age-structured populations, the MWB system (Eq.A.3 in Appendix A) has the following unknown parameters to be estimated: (i) local transmission coefficients for child - adult groups \( \{A_i^c; A_i^a\} \); (ii) local transmission coefficients for snail sites \( \{B_j\} \); (iii) egg production / release rate for children and adults \( \{\rho_c; \rho_a\} \), the latter two are biological factors independent of location.

The calibration scheme proceeds similarly to the previous case. Namely, we first use local prevalence and loss rate for children and adults to get their local rescaled forces of infection, \( \{\lambda_i^c; \lambda_i^a\} \). Then we write the proper distributed versions of the calibration formula:

\[
A_i^s = \frac{\lambda_i^s(\Delta w)}{\sum_j \eta_j Z_j} \approx \frac{(\gamma + \mu_i^c + \tau)E_i^c}{\lambda_i^c(\Delta w)}; \quad \rho_c^s \approx \frac{(\gamma + \mu_i^a)E_i^a}{\tau(\frac{H_i^c}{H_i^a})(E_i^c + \lambda_i^a(\Delta w))} \quad (3.23)
\]

Symbol \( s = \{c, a\} \) indicates child/adult groups, and denominator of \( A \) gives a weighted snail infection density (with relative ‘contact weights’). While all \( A \)-coefficients are local (i.e., depend on \( i \)), the egg production term, \( \rho_c \), should be biological and site independent. So, based on (3.23) we expect the numerator of the right hand site of \( \rho_c \) expression to be proportional to the denominator at all sites, with slope \( \rho_c \). The real data shows departures from simple linear relations, so we approximate parameters \( \rho_c^c; \rho_a \) with their ‘best-fit’ slopes. Having estimated \( \{A_i^s\} \) and \( \{\rho_c\} \), we compute coefficients \( \{B_j\} \) from the snail equilibrium equations ((Eq.A.11)).
Table 3.3 Field data from Msambweni region of eastern Kenya. (a) Demographic and infection data for 10 villages; ‘egg-count’ refers to arithmetic mean/host; ‘prevalence’ = infected fraction of population. (b) Snail density and infection data for 5 water sites. (c) Relative contact rate for each village to the 5 snail sites (each row sums up to 1).

(a)

<table>
<thead>
<tr>
<th>Village no.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>children</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>population</td>
<td>1134</td>
<td>791</td>
<td>867</td>
<td>1584</td>
<td>1572</td>
<td>357</td>
<td>753</td>
<td>403</td>
<td>294</td>
<td>492</td>
</tr>
<tr>
<td>egg count</td>
<td>58</td>
<td>91</td>
<td>134</td>
<td>62</td>
<td>115</td>
<td>319</td>
<td>257</td>
<td>139</td>
<td>148</td>
<td>170</td>
</tr>
<tr>
<td>prevalence</td>
<td>0.23</td>
<td>0.42</td>
<td>0.49</td>
<td>0.28</td>
<td>0.32</td>
<td>0.67</td>
<td>0.72</td>
<td>0.51</td>
<td>0.53</td>
<td>0.5</td>
</tr>
<tr>
<td>adults</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>population</td>
<td>932</td>
<td>632</td>
<td>701</td>
<td>1328</td>
<td>1287</td>
<td>358</td>
<td>714</td>
<td>367</td>
<td>268</td>
<td>399</td>
</tr>
<tr>
<td>egg count</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>4</td>
<td>13</td>
<td>22</td>
<td>42</td>
<td>11</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>prevalence</td>
<td>0.12</td>
<td>0.14</td>
<td>0.18</td>
<td>0.12</td>
<td>0.12</td>
<td>0.26</td>
<td>0.34</td>
<td>0.23</td>
<td>0.3</td>
<td>0.29</td>
</tr>
</tbody>
</table>

(b)

<table>
<thead>
<tr>
<th>Snail water sites</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative area</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Total density</td>
<td>1609</td>
<td>568</td>
<td>721</td>
<td>507</td>
<td>1005</td>
</tr>
<tr>
<td>Infected density</td>
<td>483</td>
<td>204</td>
<td>336</td>
<td>213</td>
<td>387</td>
</tr>
<tr>
<td>Patent (infectious) density</td>
<td>18</td>
<td>1</td>
<td>2</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

(c)

<table>
<thead>
<tr>
<th>Village\snail site</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.44</td>
<td>.33</td>
<td>0</td>
<td>.26</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.93</td>
<td>.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>.93</td>
<td>.03</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>.76</td>
<td>.11</td>
<td>.03</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>.07</td>
<td>.43</td>
<td>.31</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>.43</td>
<td>.25</td>
<td>.04</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>.01</td>
<td>0</td>
<td>.46</td>
<td>.01</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>.45</td>
<td>.03</td>
<td>.45</td>
<td>.03</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>.09</td>
<td>0</td>
<td>.83</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
3.3 Parameter estimation and model validation

3.3.1 Description of field data and estimation procedure

In this section we apply SWB methodology to a specific environment of the Msambweni region of Eastern Kenya sampled from 10 villages and 5 contact snail sites observed in year 2000, shown in Figure 3.6 [66-70]. Infection prevalence and egg-counts (arithmetic means) were observed using standard filtration of urine samples (10ml) from the resident populations. Basic biological parameters were included from well known estimates of snail and worm mortality, as well as basic life table estimates (among other parameters in Table 3.2). Field data (Table 3.3) for the models included (i) demographic data: human populations (divided into child/adult groups) and snail densities; (ii) infection data: mean egg count for children in each village and density of susceptible, infected and shedding adult snails in all water sites; and (iii) behavioral data: relative contact distribution for each village to adjacent snail sites.

We first demonstrate calibration results of the individual level field data, utilizing formulae and procedures of the previous methods sections (Model Calibration). As above we run it in 3 steps:

i) Estimate ‘local forces’ of human infection in 10 villages (first children, then adults) from the human data (prevalence and demographics), by solving equations at each site. Here we utilize the known (‘universal’) childhood prevalence
function $\phi(\lambda, \mu, S^c)$ (Figure 3.5) and the derived adult prevalence function $\phi(\lambda, \mu, S^a)$ with source $S^a$ from the children’s group.

ii) Next we estimate village-specific transmission coefficients $A_i (i = 1, 2, ..., 10)$ for child/ adult groups via formulae (3.23), and forces of infection $\lambda_i$ (output of step (i)). We also compute the ‘best-fit’ estimates for biological parameters $\rho_c, \rho_a$ (egg-production/worm).

iii) Finally, we estimate snail transmission coefficients $B_j (j = 1, 2, ..., 5)$ from equation (Eq.A.11).

Table 3.4 shows estimated snail to human transmission coefficients $\{A_i, A_w\}$ for child and adult groups in 10 villages, along with the best fit egg-production rate estimates $\{\rho_c, \rho_a\}$ for the two groups, and human to snail transmission rates $\{B_j : j = 1, ..., 5\}$. We also looked at the effect of worm discretization step $\Delta w$ on estimated parameters. To this end we run two calibrations of the individual level field data, with steps $\Delta w = 10$ and $\Delta w = 5$, and fixed infectivity threshold $w_T = 10$. The 2nd case gave overall higher estimates of transmission rates $\{A_i\}$, but the children group was less sensitive than adults. Since children carry the bulk of infection and transmission, our scheme seems overall insensitive to $\Delta w$.  

77
Table 3.4 Calibrated transmission coefficients (A’s and B’s) and egg production rates (ρ’s) based on the individual level field data for 2 choices of discretization step (Δw), and infection threshold \( w_T = 10 \).

<table>
<thead>
<tr>
<th>Village no.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_c )</td>
<td>( \Delta w = 10 )</td>
<td>.51</td>
<td>.72</td>
<td>.92</td>
<td>.52</td>
<td>2.92</td>
<td>23.71</td>
<td>6.95</td>
<td>2.87</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>( \Delta w = 5 )</td>
<td>.82</td>
<td>.90</td>
<td>1.09</td>
<td>.76</td>
<td>4.03</td>
<td>25.02</td>
<td>7.05</td>
<td>3.34</td>
<td>2.58</td>
</tr>
<tr>
<td>( A_a )</td>
<td>( \Delta w = 10 )</td>
<td>.08</td>
<td>.02</td>
<td>.035</td>
<td>.05</td>
<td>.18</td>
<td>.76</td>
<td>.34</td>
<td>.22</td>
<td>.29</td>
</tr>
<tr>
<td></td>
<td>( \Delta w = 5 )</td>
<td>.23</td>
<td>.11</td>
<td>.13</td>
<td>.16</td>
<td>.68</td>
<td>2.01</td>
<td>.68</td>
<td>.52</td>
<td>.52</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Snail site no.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B )</td>
<td>( \Delta w = 10 )</td>
<td>1.04e-5</td>
<td>6.50e-6</td>
<td>1.73e-5</td>
<td>2.26d-5</td>
</tr>
<tr>
<td></td>
<td>( \Delta w = 5 )</td>
<td>7.98e-6</td>
<td>5.79e-6</td>
<td>1.53e-5</td>
<td>2.13e-5</td>
</tr>
</tbody>
</table>

| \( \rho_c \)  | \( \Delta w = 10 \) | 17.62  | 6.48  |
| \( \rho_a \)  | \( \Delta w = 5 \)  | 16.09  | 4.28  |

### 3.3.2 Model validation

We ran a simple self-consistency (validation) test for our calibrated system by comparing the predicted equilibria (computed from the model) with the individual level field data. An exact match is not expected as parameters of the distributed system are no longer exact solutions of equilibrium equations. Indeed, the complete set of equations (MWB + human prevalence + snail prevalence) is over-determined now, so we use best-fit approximation to obtain the estimates of Table 3.4. The comparison test shows good agreement (Figure 3.7).
Figure 3.7 Comparison between individual level field data (dark disk) and equilibrium solutions of the calibrated model (gray square); the x-axis labels the ‘village’ or snail site and y-axis – ‘mean egg count’ or ‘snail density’.

### 3.3.3 Sensitivity analysis

The infection data based on egg and snail counts is notoriously inaccurate partly due to imperfect diagnostic tools. Another major uncertainty in our model is the infectivity threshold $w_T$. To examine such uncertainties on calibrated parameters we took two threshold values $w_T = 10$ and 20, and in each case run an ensemble of 200 data choices, randomly selected within suitable ranges: $\pm 10\%$ of the individual level field data, for egg counts; $\pm 5\%$ - for human prevalences, and $\pm 20\%$ - for snail infection. The error margins are somewhat arbitrary, but they reflect overall level of uncertainty. For instance, individual egg-counts can be much ‘noisier’ than 10%, but the community mean would reduce their variance by factor $1/n$ – community size. Figure 3.8 shows the results of sensitivity analysis for some calibrated parameters. The overall estimates remain fairly stable (error $< 5\%$). Therefore, we plan to use the calibrated ensemble of parameters to
provide us error bars for predicted control outcomes in the future work on control analysis with SWB.

Figure 3.8 Sensitivity analysis of model parameters (snail-to-human transmission rates) \( \{A_i\} \) for children in 10 villages: panel (A) shows heavily infected villages no. 6 and 7, and panel (B) for all others. Two sets (light, dark) of box-and-whisker plots (‘box’ = .25 - .75 quantiles; ‘whiskers’ = 10 - 90) correspond to infectivity thresholds \( w_T = 10 \) and \( w_T = 20 \). Black dots mark calibrated values of the ‘individual level field data’ in each case.
Chapter 4 Analysis of control and interventions

4.1 Overview of schistosomiasis control strategies

Schistosomiasis is an environmentally mediated parasitic disease like malaria, hookworm, onchocerciasis, and Chagas disease which result in high morbidity and increased mortality and infect millions of people living in tropical and subtropical regions [78-82]. Different control initiatives have been applied to reduce worm infections and the disease caused by schistosomiasis parasitic infections. However, the effects of those initiatives are quite uneven because of the regional and country-specific differences in available resources and political commitment. With combined approaches of snail control, chemotherapy, health education and hygienic movement, China has made substantial progress against schistosomiasis and it has been successfully eliminated several provinces over the past several decades [83, 84]. In Brazil, chemotherapy programs achieved 50% to 70% decrease in prevalence [85]. However, schistosomiasis still remains a major public health problem in Africa.

Measures for controlling schistosomiasis include chemotherapy using praziquantel, snail control (molluscicide or changing environmental habitats), cercariae and miracidia control, provisional safe water supply, and health education, etc. Among those measures, chemotherapy and molluscicide are the most widely used in Africa.

Traditionally, when using mean worm burden as the metric for human infection intensity, chemotherapy treatment (praziquantel) can be modeled as an additional loss in
the mean worm burden over a time period. This additional loss rate can be derived from the basic pharmacodynamics. If we define the efficacy of a drug \( e_c \) to be the fraction of worm killed by the drug within its effective duration and denote \( T_c \) be the effective duration of the chemotherapy, then the additional loss due to chemotherapy takes the following form (see [86]):

\[
-\frac{1}{T_c} \ln(1 - e_c). 
\]  

(4.1)

This loss rate is added to the equations (for mean worm burden) only for the period of \( T_c \) since the additional loss rate is not valid for other times. In reality, we cannot cover the whole population or target subpopulation with chemotherapy because of limitations of drug delivery. So the target group is divided into two subgroups: covered and uncovered. For instance, if we choose a group of children for a drug treatment of efficacy \( e_c \) and effective duration \( T_c \), we use \( W_c \) to represent the mean worm burden of the covered children and \( W_U \) for the uncovered subgroup. The differential equations for \( W_c \) is same as the uncovered group except the additional loss rate (4.1) of worms. The mean worm burden of whole children group is \( f_c W_c + (1 - f_c) W_U \), where \( f_c \) is the fraction of children treated by the drug. I would like to mention here that the drug treatment on human can be modeled in a more realistic and flexible way with our stratified worm burden (SWB) modeling approach. We will talk about it in detail shortly.

In Chapter 3, we described a new approach to predicting schistosomiasis transmission and community-level infectious burden using models based on a SWB
formulation, combined with calibration using data from an existing multi-year schistosomiasis control program [42, 87, 88]. In this approach, a host population is subdivided into burden strata in terms of individual parasite load, with the \(i^{th}\) stratum carrying \((i - 1)\Delta w\) to \(i\Delta w\) worms, where the size of the stratum, \(\Delta w\), is prescribed e.g., at increments of an additional 10 worm pairs per person. The transitions among strata depend on (i) force of infection (determined by the transmission environment and human risk factors); (ii) worm attrition (natural or drug-induced); and (iii) human demographics (birth, aging, mortality, growth sources).

Taking burden strata as model variables instead of the population mean worm burden, SWB carries detailed information of the infection status of a host population and readily explains its highly aggregated distribution (typical of *Schistosoma* and other helminth infections [89]), without imposing *ad hoc* distribution patterns. Statistical moments (variance and/or aggregation) of such a distribution can include human population dynamics and, in simplified format reflect the classical MacDonald [90] MWB equations. With such information, we are able to develop refined calibration procedures that take better advantage of available epidemiological data. The thrust of our calibration methodology is to first estimate local force of infection through fitting of SWB equilibrium equation parameters to prevalence data, then using reduced MWB and snail prevalence systems to estimate transmission parameters and other unknown biotic and abiotic parameters associated with transmission.
In this chapter, we shall use our age-structured heterogeneous model based on the calibrated SWB formulation and estimate the potential effects of various chemotherapies on the distributed environment made of multiple interconnected human villages and snail habitats in eastern Kenya (discussed in Chapter 3) in two cases: first, with stationary snail populations; second, with seasonal dynamic snail populations and snail control based on the snail model discussed in Chapter 2.

### 4.2 Drug treatment based on SWB

#### 4.2.1 Treatment setting

Due to various limitations (resources, delivery system etc.), drug treatment does not cover every infected human host in the area and the frequency of the treatment cannot be too high. To achieve the best result of chemotherapy, one generally needs to identify the high-risk groups and then concentrate effort on those. In this study, we estimate the effect of age-based treatment and focal treatment using our SWB model. Non-stationary and nonlinear models of this type are difficult to analyze analytically. However, we can simulate different chemotherapy scenarios numerically and observe their effects.

The effect of drug treatment on the prevalence of *Schistosoma* infection was implemented in the model by instantaneously shifting humans at higher burden levels (stratum) to lower levels according to the established efficacy of praziquantel [91, 92]. As an example, within the model, a treatment session with an efficacy of 90% (*i.e.*, killing 90% of worms) would bring a person in the 10th stratum, who is carrying between 90 and
100 worm pairs before therapy, down to the 0th strata (carrying 0 to 10 worms), as the remaining number of worm pairs will most likely be between 9 and 10 of worms after the treatment. This is a more realistic effect than modeled in the traditional approaches used in the past, in which a treatment term is added to the natural mortality of worms to indicate an extra loss. In fact, the drug-killing effect on worms occurs more quickly (in less than one month) than natural death (~ 4-5 years [93]), and this more flexible scheme allows predictions for various timing and coverage among different control strategies, such as treating only the most heavily-infected groups. In our model’s simulations we assumed at-random treatment coverage, in the sense that a random subset of each population was treated each time [94, 95]. In this sense, the ‘coverage’ of a control program means the fraction of all people treated in each treatment session, regardless of individual year-to-year participation.

4.2.2 Model validation with new data

Before we apply our SWB model and the corresponding treatment setting to do any predications, we need to convince ourselves that ours is a valid model with good predictive power. To this end, we compare the result of model simulation with field data in a specific scenario. Fortunately, we have new data about infection prevalence from the same region of Kenya. Two interventions were implemented in that region in 2000 and 2003: in 2000, 79% of infected for all of the 10 villages were treated; in 2003, only the most heavily infected villages-village 6 and 7-were treated, with coverage of 41% for the
infected. New data 2009 shows that the infection prevalence of children is 61.7% for village 6 and 62% for village 7.

Two additional assumptions were made for better realism of the model:

I. The population of each of the 10 villages grows at an annual rate of 1.5%, based on population census of village 6.

II. All infected were able to be diagnosed and treated with the same likelihood. Treatments covered only the infected and coverage for children and adults was the same.

Figure 4.1 shows the comparison of children’s prevalences of village 6 and 7 from simulation based on estimated parameters (grays dot) and from new data (black dots). We can see fairly good agreement. If we take into account with the noisiness of field data by allowing a broader range of parameters instead of point estimate, we can even get stronger support that our model for this specific treatment setting is good enough to provide reliable prediction. Keeping everything else the same, we allow parameters to vary randomly in a range of ±20% of their exact value estimated from raw data and run simulation to get the child-group prevalence at the end of 2009. The boxes in Figure 4.1 were derived from 100 runs of this procedure. The two sets of dots are enclosed by the boxes, showing that the parameters are not extremely sensitive and our model for this setting has good predictive power.
Figure 4.1 Box-plot of prevalence of each village from model simulation based on SWB model with 100 sets of parameters randomly perturbed around the best-fit one. Gray dots for village 6 and 7 stand for the simulation result from the model with the best-fit parameter values, while black dots represent new data.

4.2.3 Two guidelines for chemotherapy

Now we are ready to apply our model to predict potential outcomes of various chemotherapy strategies. Currently, the ‘standard-of-care’ for population-based drug treatment of *S. haematobium* is based on the 2002 WHO guidelines [96]. Under these guidelines, populations are divided into 3 classes according to local prevalence of infection among children. These are: high prevalence (> 50% children infected), moderate prevalence (10 – 49% infected), and low prevalence (<10% prevalence), respectively. For highly infected population, it is recommended that children between 5 and 15 year old, as well as high risk adults, are treated annually; for moderately infected populations, only children between 5 and 15 yr of age are treated every 2 years; for populations with low prevalence of infections, it is recommended that school children be treated twice, upon their entry into school and at school completion.
In late 2010, new large-scale operational research programs will begin in seven schistosomiasis-endemic locations in sub-Saharan Africa, situated in Cote d’Ivoire, Niger (2 sites), Kenya (2 sites), Tanzania and Mozambique [97]. These studies will seek to identify the more cost-effective approaches to regional and national schistosomiasis control, within the usual present-day framework of mass treatment campaigns for helminth control. These SCORE trials will involve implementation in 75 to 150 villages per location, and will compare the relative costs and effectiveness of different drug delivery strategies for population-based control of either Schistosoma haematobium or S. mansoni infections. The various SCORE project studies will compare relative effectiveness of program implementation for communities having lower vs. higher prevalence of infection, measured at rates of 10-24% among school-age children (low) and ≥ 25% (high), respectively. Among the selected SCORE villages, there will be randomized assignment to one of several different treatment options and schedules: Some communities will receive community directed treatment (CDT) which will include non-standard programmatic treatments for all consenting adults; other communities will receive the more standard school-based targeted treatments (SBT) along the lines of current W.H.O. recommendations [96]. With each type of delivery, some communities will receive yearly therapy, some will transition from CDT to SBT, while others will transition to every other year treatments or to ‘drug holidays’. CDT will involve the most extensive treatment coverage, aiming to include as many people as possible, whereas SBT will target schools (with non-attenders of school age also encouraged to participate),
while drug holiday means there will be no treatment for the year. For communities with high infection prevalence (≥ 25%), there are 6 proposed strategies over the 4-year study period:

1. CDT – CDT – CDT – CDT.

For populations with lower infection prevalence (10 to 24%), no CDT is recommended and treatment will be either school – based every year, two years of SBT followed by two year holiday, or SBT every two years.

An important consideration in choosing to implement a more intensive drug treatment program will be its incremental cost-effectiveness relative to standard-of-care [98]. For purposes of this paper’s analysis, we focus on two major costs: the cost for community screening (in determining the prevalence of active *Schistosoma* infection in the population) and that for drug treatment (drug costs + delivery costs). In the SCORE program screening and assignment will be performed only at the beginning of the 4 year treatment regimen, so that screening costs will be the same in each arm. However, for comparison, we have also modeled the possible impact of rescreening and reassigning treatment regimens on a yearly basis, and for these approaches (not actually researched
by SCORE) we will include the costs of repeated rapid screening among school age children [99-101]. Although it requires extra expense, screening before treatment has the potential to reduce the drug cost by adapting to a less expensive strategy when prevalence of infection becomes less intense. For our analysis, we took the cost of initial community sensitization and screening to be approximately $1 U.S. dollar (USD) per person, based on in-country costs documented by the Partnership for Child Development and other national control programs [102]. The cost of annual drug dose was estimated at $ 0.25 per child and $ 0.50 per adult. Rapid program screening is carried out by sampling 50-100 persons for each village [99], so $100 dollar is the maximum estimated screening cost per each village, no matter how large the population is. The cost for drug depends on the number of people treated, so information about total population, coverage of each treatment, and number of villages treated at each time is needed for estimates tailored to specific locations.

Based on the same assumption of population growth (at annual rate of 1.5%) over time, we will take the end of 2009 as the baseline and investigate the following metrics of a drug treatment strategy: i) the number of years to bring down the infection prevalence of the whole community to a low-risk level; ii) years to remain a safe level after the termination of a treatment; iii) cost of the program with or without pretreatment screening.
4.3 Results

4.3.1 Stationary snail populations

4.3.1.1 WHO 2002 guidelines

As indicated in the guideline, only the children entering or completing school are treated for populations with low infection. I made a simplification by treating a fraction of school children at the first year of one session. Coverage for adults for highly infected populations is set to be 60% and restricted to high risk individuals. Since our model is able to target different burden strata, we interpret here “high risk groups” as those in the 5 highest strata. Table 4.1 shows the result for different coverage for children.

Table 4.1 minimum number of treatments needed for all prevalence to go below 10% and duration of low level infection for each village for different children coverage under WHO 2002 guidelines, with coverage for high risk adult group is 60%.

<table>
<thead>
<tr>
<th>Village index</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence in 2000</td>
<td>0.23</td>
<td>0.4</td>
<td>0.47</td>
<td>0.27</td>
<td>0.3</td>
<td>0.65</td>
<td>0.74</td>
<td>0.53</td>
<td>0.52</td>
<td>0.53</td>
</tr>
<tr>
<td>Prevalence in 2009</td>
<td>0.16</td>
<td>0.29</td>
<td>0.35</td>
<td>0.18</td>
<td>0.22</td>
<td>0.57</td>
<td>0.69</td>
<td>0.46</td>
<td>0.44</td>
<td>0.45</td>
</tr>
<tr>
<td>Coverage</td>
<td>no. of treatments</td>
<td>no. of safe years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>8</td>
<td>6 3 2 6 5 1 0 2 0 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>6</td>
<td>7 4 3 7 6 2 1 2 1 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>4</td>
<td>8 4 3 7 6 2 1 2 1 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>4</td>
<td>10 7 6 10 8 4 3 5 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The above hypothetical treatments were set to be carried out according to the WHO 2002 guidelines based on the baseline prevalence for children in 2009. If screening is carried out before each treatment, then the strategy for each village should be subsequently adjusted in late years because the prevalence is reduced after treatment. Since villages are not to be treated if found to have low infection in screening, it is unlikely that all villages have prevalence to be far below 10%, and further treatment will be needed if low infection level is to be maintained. In Table 4.2, we show a comparison of cost of treatments with and without pretreatment screening. Minimum years are calculated based on those without pretreatment screening except the first year. We can see that although fewer treatments are implemented for the same period of time, strategy with screening before each treatment is more expensive than no screening and less efficient, given low treatment coverage. For treatment programs with coverage as high as 80%, then the cost–effectiveness of prescreening become more visible, due to dramatic reduction of drug cost. Despite of this, the minimum prevalence of the 10 villages is still above 10%, which means that we need to do more years of treatment (then more cost) in order to achieve this goal.
Table 4.2 Comparison of WHO 2002 strategies without (No) and with (Yes) yearly pretreatment screening

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Minimum years of treatment needed for all villages to be at low-risk</th>
<th>Total treatment units (sum over no. of village treated each year)</th>
<th>Maximum prevalence(children) after treatment among villages</th>
<th>Total cost (in $)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No/Yes</td>
<td>No/yes</td>
<td>No/Yes</td>
</tr>
<tr>
<td>0.6</td>
<td>8</td>
<td>40/32</td>
<td>0.09/0.35</td>
<td>6723/9899</td>
</tr>
<tr>
<td>0.7</td>
<td>6</td>
<td>30/24</td>
<td>0.07/0.28</td>
<td>6162/7374</td>
</tr>
<tr>
<td>0.8</td>
<td>4</td>
<td>20/17</td>
<td>0.08/0.25</td>
<td>5178/5098</td>
</tr>
<tr>
<td>0.9</td>
<td>4</td>
<td>20/17</td>
<td>0.04/0.15</td>
<td>5825/4664</td>
</tr>
</tbody>
</table>

4.3.1.2 SCORE program

Since the prevalence of infection for children for all of the villages at the year 2009 are above 25%, all villages are treated with strategies for high infection. Figure 4.2 shows the prevalence of each village after treatment with one of the 6 different strategies. Strategies 1, 2 and 4 bring the prevalence of all villages below 10% after 4 years’ treatment. Among the three, strategy 4 costs the least (Figure 4.2). The results for prevalence at the end of treatment of the three strategies are similar, which implies that treating adults has very limited effect for the control of transmission. The number of subsequent safe years (minimum prevalence<10%) is shown in Table 4.3. Strategies 3, 5, 6 are not able to bring the region to a safe level of infection, so further treatment is necessary for the villages with prevalence larger than 10% after the 4-year session. Based on the previous experience (Figure 4.2 and Table 4.3), we know that treating adults will...
contribute little to the control of transmission in the region and that school-based treatment for every two years is better than two consecutive treatments followed by two holiday years in terms of reduction of prevalence. So we think it is wise to follow up community treatment either with SBT for every two years or every year.

Figure 4.2 Bar charts of prevalence for 10 villages after 4-year-session of different treatment strategies under SCORE program, with coverage being 0.7

Table 4.3 Number of safe years after 4-year-treatment for each village under strategy 1, 2 and 4 with coverage being 0.7 for children and 0.5 for adults

<table>
<thead>
<tr>
<th>Village index</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence 2009</td>
<td>0.16</td>
<td>0.29</td>
<td>0.36</td>
<td>0.19</td>
<td>0.22</td>
<td>0.57</td>
<td>0.69</td>
<td>0.46</td>
<td>0.44</td>
<td>0.45</td>
</tr>
<tr>
<td>Strategy index</td>
<td>Cost (in $)</td>
<td>no. of safe years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>8455</td>
<td>9 8 7 10 8 2 0 3 4 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6272</td>
<td>9 8 7 10 8 2 0 3 4 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4185</td>
<td>9 8 7 10 7 2 0 2 3 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.3 shows the prevalence profiles of village 2 and village 6 over time under first 4-year CDT-CDT-Holiday-Holiday strategy followed by 4 or 8 yearly SBT. Village 2 and 6 are representatives for moderately-infected and highly-infected groups, respectively. After the first 4 treatment session, village 2 become low-risk, and will not be treated in the subsequent years. However, village 6 still has prevalence above 10% and needs to be further treated. As we can see from this figure, after intensive yearly treatment (4 years or 8 years), village 6 goes to a safe level but rebounds back within certain years. And the discrepancy of the two dashed curves after certain time shows subtle influence of treatment for village 6 on village 2.

Figure 4.3 prevalences of village no. 2(dashed) and no.6(Solid) as functions of time with treatment strategy 3 in SCORE (CDT-CDT-Holiday-Holiday) and then followed by 4 (gray) or 8(black) yearly school-based treatments
4.4 Combination of human chemotherapy and snail control

For all simulations up to now, the total snail populations of the 5 water sites are assumed stationary at their respective equilibrium levels, in the sense that there is no fluctuation by environmental factors such as seasonal rainfall or sudden change by snail control. By interrupting cercarial production, snail control in water sites is an effective and cheap way to reduce schistosome transmission and is supposed to help bring down the infection prevalence of nearby villages. In the following analysis, we would like to utilize the seasonal snail population model we developed in Chapter 2 and predict quantitatively the integrated outcome of human chemotherapy and snail control.

For simplicity, we assume that all the 5 snail sites in this region have the same population pattern resulted from the same rainfall—the 1984 rainfall data we used for model calibration in the snail model— and snail control. We also assume that this one-year seasonal pattern is periodically repeated each year. Figure 4.4 shows the normalized three-year snail population profile from the snail model with different snail control strategies: no control; once a year at the best timing (0.37); twice a year at 0.37 and 0.74; quarterly at the beginning of each quarter. By normalization, we mean dividing the snail population profile without any control by its integration over one year so that the resulting profile has its annual average being 1. Then the snail population dynamic for each water site is this normalized profile multiplied by their respective stationary constant which is given by the snail data used in Chapter 3. In this way, we replace the stationary snail population of each water site by a seasonal dynamic with the same average.
Figure 4.4 Seasonal snail patterns (same for each year) under snail control of different frequency. Blue—baseline (no control); purple—once at best timing (0.37); yellow—twice (at 0.37 and 0.74); green—quarterly (at the beginning of each quarter).

With this setting for snail population, we do simulation of human chemotherapy based on SCORE strategies as in the previous section. Figure 4.5 is an example of the resulting prevalence profiles for child-group for the 10 villages with the same CDT-CDT-CDT-CDT scheme and different snail control strategies. The snail control is continued even after termination of 4-year human chemotherapy. We can see that for 4-year CDT session, a continued quarterly snail control program is necessary and good enough to maintain all of the 10 villages at a low-risk level for a long time.
Figure 4.5 Predication of *Schistosoma* infection prevalences for children in the 10 villages in eastern Kenya starting in December 2009, using Stratified Worm Burden model with the same 4-year-session drug treatment scheme under SCORE program (CDT-CDT-CDT-CDT) and snail control of different frequency. upper left - no control; upper right – once/year; lower left – twice /year; lower right – quarterly
Chapter 5  Chronic infection and childhood development

5.1 Introduction

Schistosomiasis remains one of the most serious and prevalent diseases worldwide. In 2005, there were an estimated 207 million people infected, with 89% of these living in the less-developed areas of rural sub-Saharan Africa and South America [1, 2]. Although highly effective anti-schistosomal drugs have been marketed for over 25 years [103], there remain significant challenges to providing treatment (or preventive therapy) to those who are at highest risk for disease [104, 105]. Because, until recently, implementation of large-scale of anti-schistosomal treatment has been very limited, there remain significant gaps in our knowledge about the expected benefits of repeated treatment in areas that continue to have high risk for Schistosoma reinfection following therapy [106]. These communities, which often have the highest prevalence and intensity of infection, pose a particular challenge to program development for schistosomiasis morbidity control [107, 108].

Clinical and epidemiological studies indicate that 10-15 year old children typically carry the highest rates of schistosome infection and the highest risk of

---

3 First published in American Journal of Tropical Medicine and Hygiene [148]
inflammation-related disease associated with infection [109, 110]. The more lethal, late outcomes of infection are more common among adult age groups, and result from progressive infection-associated fibrosis of vital organs [111, 112]. However, of recent note, researchers and policymakers have also come to appreciate that schistosomiasis can also be a significant risk factor for chronic anemia, childhood growth stunting, protein calorie malnutrition, cognitive disability, and poor school performance [113-119]. These sub-clinical morbidities are physiologically important but more subtle than the easily-recognized advanced forms of schistosomiasis [112]. Nevertheless, these ‘subtle morbidities’ can have serious day-to-day consequences in the setting of rural poverty [117, 120], and may, in fact, given the substantial numbers of persons who are affected by these pathologies, represent the bulk of schistosomiasis-associated disability and health burden among endemic populations [121].

How available resources are best allocated to prevent both the prevalent sub-clinical morbidity associated with schistosomiasis and the more severe forms of advanced disease (including developmental stunting)? In the present analysis, we use a calibrated computer simulation to estimate the relative benefits of different treatment strategies for school-age schistosomiasis control programs. WHO presently advocates schistosomiasis control by a strategy of periodic drug treatment of affected populations, focusing on school-age children as the highest risk group for infection (and consequent disease formation) [122]. Large-scale control programs have already begun in many countries [104], but important operational questions regarding the optimal timing and distribution
of treatment efforts still remain. To address these questions in better detail, our present modeling approach builds partly on earlier modeling efforts of Medley & Bundy [123] Chan et al. [95], and of Gurarie-King [76, 124]. These papers focused on late-term Schistosoma-related morbidity outcomes and the analysis of the optimal timing for targeted or population-based therapy for control or prevention of those ‘classical’ forms of schistosomiasis. The current study takes a different approach, focusing instead on predicting the best means to use available therapy for prevention of the detrimental impact of schistosomiasis on childhood growth and development [115, 116, 125-127]. Because available field data are presently quite limited, we calibrated a growth-development model for ages 0-20 yr using the CDC’s NCHS data [128], and evaluated the age-dependent impact of 'schistosomiasis-like' chronic infection that causes growth retardation. The model was benchmarked using detailed anthropometric and infection data collected for Kenyan villagers in a S. haematoobium endemic area [129]. In addition, the approach was nuanced to include innate age- and gender-related differences in individual risk for growth related disease formation [130, 131], and the projected impact of preventing reversible nutritional morbidities of childhood [113, 115, 132, 133].

As a basis for discussion in current program planning, we addressed the unanswered questions about long-term treatment impacts by developing predicted outcomes based on a bounded modeling system that accounted for: a) age-related parasite exposure; b) the kinetics for development of inflammation-related disease; and c) the impact of infection on early and mid-childhood growth retardation [117, 134]; and d) a
child’s potential for catch-up growth at different stages of childhood and adolescence. Our analysis indicates certain optimal drug-based control strategies, and suggests that these have the potential for a substantial reduction of developmental morbidities found among schistosomiasis-affected populations.

5.2 Method

This model of physiological growth/development over young ages 0-20 years accounts for different growth strata (“developmental quantiles”) with a population, and the possible growth retardation due to chronic infection (such as schistosomiasis). It is meant to represent a dynamic process, i.e., the life history of an individual or of a given population growth stratum, driven by natural (physiological) growth-factors and growth-inhibitors (e.g., chronic infections).

Model Assumptions:

1. For the purposes of this analysis, human populations and transmission environment are stationary, i.e., unchanging.

2. The normal childhood growth/development patterns are based on CDC-WHO standard curves (see http://www.cdc.gov/growthcharts/percentile_data_files.htm, and [128], See Figure 5.1) reflecting the expected pattern of growth, where growth rates (velocity) can be modified downward (decelerated) or upward (accelerated) to model, respectively, growth inhibition by infection, and its potential remediation by catch-up growth [135] after anti-parasite treatment.
3. Persisting parasite burden due to lack of treatment or reinfection after treatment, and its resulting inflammation [136] can impair growth-related developmental processes over the first 20 years of life. However, we assume that the retardation process is partly reversible (remedial growth), once the sources of infection and chronic inflammation, are removed.

4. The basic biological growth parameters and behavioral patterns (e.g. water contact rates and reinfection risk) are age- and gender-specific, which results in different patterns of worm burden and morbidity between girls and boys.

The basic dynamic variables of the model are: (i) \( w(a) \) - worm burden of a host (or mean worm burden (MWB) of a homogeneous population stratum); (ii) development index (height/weight) \( h(a) \) of a host (or stratum); both of these being functions of age \( 0 < a < L \) (L, life span). In general, both infection and developmental variables should be considered functions of ‘age’ and ‘time’ \( \{ w(a,t); h(a,t) \} \) [93], but here we ignore temporal changes, assuming population structure and transmission to remain stationary in time.

5.2.1 Coupled infection, morbidity and retarded growth dynamics

Our analysis of drug intervention is based on age-patterned strategies and their (cohort/community) outcomes:

A. Infection dynamics Age dynamics of burden \( w(a) \) is described by differential equation with source term \( S = S(a) \) (force of infection that accounts for transmission and
parasite establishment), and resolution rate \( \gamma = 0.2 \text{/year} \) (assuming worm life-span = 5 years),

\[
\frac{dw(a)}{da} = S(a) - \gamma w(a)
\]

\[ (5.1) \]

We assume source \( S(a) = S_0 \eta(a) \) to be a product of environmental ‘risk factor’ \( S_0 \) (maximal ‘worm accumulation rate’ per ‘unit transmission’), and relative age-dependent transmission pattern \( 0 \leq \eta(a) \leq 1 \). The latter \( \eta(a) \) should account for age-dependent behavioral (risk) changes and changing ‘probability of worm establishment’ due to acquired immunity. Chan et al [21], proposed to take

\[
\eta(a) = \eta_p \left( \frac{a}{a_0} \right); \text{ with } \eta_p(z) = z^p e^{-p(z-1)^2}
\]

\[ (5.2) \]

It highlights peak-transmission age \( a_0 \) (\( \eta(a_0) = 1 \)), and exponent \( p > 0 \) that measures steepness of ‘low to high’ transitions in early or late ages (away from \( a_0 \)), and the follow up 'high to low' drop. The parameters \( \{ a_0, p, S_0 \} \) were calibrated using mean worm burden (reflected by egg count) data from community-based field studies of \( S. \ haematobium \) infection and reinfection in Kenya [106, 129]. The Kenyan data represents a cross sectional community survey, and in our calibration scheme below we apply it to the appropriate age cohorts.

**B. Accumulated morbidity** Chronic worm infection results in morbidity caused by persisting inflammatory immune response to parasite eggs deposited into host tissues [110]. Quantitatively, we describe this process by the following differential equation:
Here variable \( z(a) \) serves as a cumulative proxy for the damage caused by chronic infection that is relevant to impaired host development ([137]). Variable \( z(a) \), is acquired at a rate \( \rho \) proportional to worm burden, and is resolved at an age-dependant rate \( \nu(a) \), given by a sigmoid function,

\[
\nu(a) = \nu_0 \frac{\alpha + (a / a_i)^q}{1 + (a / a_i)^q}
\]  

(5.4)

with maximum - \( \nu_0 \), minimum - \( \alpha \nu_0 (0 < \alpha < 1) \), threshold age \( a_i \) and hill-exponent \( q_i \).

Such form of burden resolution differs from earlier works ([79], [80]).

**C. Drug treatment**  Anti-schistosomal drugs such as praziquantel remove about 90% of each child’s mature worm burden for each treatment, as estimated by reductions in egg output [103]. For the purposes of our analysis, we assumed that this happens very quickly once treatment is implemented. To do this, chemotherapy was modeled in equation (1) by augmenting natural worm death rate \( \gamma \) with an additional clearing term, made of step-function of short duration \( T \) (or its Dirac delta approximation) whose strength \( r = -\log(1 - e) \) depends on the efficacy of drug \( e \) = fraction of worms killed per session). Thus a sequence of several sessions administered at ages \( \{a_j\} \), gives the combined clearing term

\[
\gamma + \frac{r}{T} \sum \delta(a - a_j)
\]  

(5.5)
Our goal in this study is to construct a predictive model based on the environmental constraints and basic biology of snails. We start with a simple model of snail population, and then proceed to more general one. The basic “single population” balances the difference between its birth rate and mortality rates.

**D. Impaired growth** For the present analysis, we examine the cumulative impact of infection at age 20 (adulthood), when normal development would typically terminate. *Normal growth* was assumed to be driven by an age-dependent *natural growth rate* (NGR) $g(a)$ (per unit weight or height, etc). We estimated $g(a)$ using CDC (NCHS) data on juvenile development in the USA (Figure 5.1). Specifically, NCHS data gives median quantile development curves (indices) $\{H(a) = H_q(a)\}$ from which we compute NGR function

$$g(a) = g_q(a) = \frac{H_q'(a)}{H_q(a)}$$  \hspace{1cm} (5.6)

for each quantile $0 < q < 1$. Here $H'(a) = \frac{dH(a)}{da}$ denotes the derivative of H(a). We further assumed that NCHS growth patterns are inborn characteristics, so that an individual born into any $q$-quantile would, with some oscillation, tend to maintain through later normal development. *Chronic infection* (along with malnutrition) interferes with normal development at younger ages, and can ultimately result in lifelong developmental retardation [135, 138]. To model the dynamics of disease-related impairment, we chose a suitable physiological characteristic (e.g., one’s weight or
height), \( h(a) \), and relate it to the normal developmental index \( H(a) \), for each quantile.

Furthermore, we included the effect of remediation via ‘catch-up’ growth potential at
different ages [138]: The catch-up growth is believed to reflect a feedback response to
an impaired developmental state that deviates from an expected set-point. This departure
then stimulates an additional component of growth [135, 138, 139]. For children with
developmental deficiencies, we estimated a remedial growth rate (RGR) that, like normal
growth, was assumed to depend on age \( \{a\} \), but also on some form of ‘departure from
the norm’ i.e., deficit, (either an ‘absolute’, or a ‘relative’ one),

\[
\begin{align*}
  d(a) &= H(a) - h(a) \\
  d(a) &= 1 - \frac{h(a)}{H(a)}
\end{align*}
\]  

such that \( \text{RGR} = R(a, d) \). One can think of ‘catch-up’ growth as a built-in ability by the
body to ‘read’ its current state \( h(a) \), to ‘compare’ it with the ‘programmed’ normal value
\( H(a) \), and provide an additional growth term in the \( h \)-equation (through enhanced
metabolic-hormone regulation). The deleterious effect of a chronic infection such as
schistosomiasis on growth-development could be linked either to accumulated worm
burden \( w(a) \), or a suitable morbidity variable \( z(a) \) (e.g. inflammation) derived from \( w \).

Using \( z \) as a the main source of inhibition, we represented its diminishing effect on both
growth rates, NGR and RGR by factor \( 0 < \phi(z) < 1 \), its value decreasing with greater \( z \).

The combined inhibited growth equation for index \( h(a) \)
\[
\frac{h'(a)}{h(a)} = \phi(z) g(a) + \phi(z) R(a, d)
\]

Let us note that in the absence of infection \((\phi = 1, \text{or } d=0)\), equation (4) should turn into the normal growth equation (3) for \(h = H(a)\). So we expect \(R(a, d) = 0\), if \(d=0\).

The simplest such choice should be linear function in \(d = H - h\), or \(1 - \frac{h}{H}\). So we assumed

\[
R(a, 1 - \frac{h}{H}) = r(a) \left(1 - \frac{h}{H}\right)
\]

where \(r(a)\) is the (hypothetical) maximal catch-up rate (under 'extreme deficit' conditions). We made two additional assumptions:

(i) Growth inhibition factor \(\phi\) is a sigmoid function between 2 levels: \(\phi_0 < \phi(z) < \phi_1\), with a 50\% threshold \(-z_0\) and a Hill exponent \(m\)

\[
\phi(z) = \phi_0 + \frac{\phi_1 - \phi_0}{1 + (z/z_0)^m}
\]

Its maximal value \(\phi_1 < 1\) (at \(z=0\)), rather difference \(1 - \phi_1\), should account for other (non-schistosomiasis) inhibiting factors (e.g. co-infections or malnutrition). Here, the schistosomiasis-specific contribution corresponds to the gradual transition from high value \(\phi_1\) (at low \(z\)) to low \(\phi_0\) (at large \(z\)).
(ii) The observed height/weight retardation for a single human host or a homogeneous growth stratum results from the same accumulated damage (morbidity) of this group.

(iii) In the absence of more detailed information on ‘natural’ and ‘remedial’ growth mechanisms we assumed a simple linear relation between two rates

\[ r(a) = Kg(a) \]  

(5.12)

with parameter K equal to ‘relative efficiency’ of remediation.

The net result is a coupled system of equations (5.1), (5.3) and (5.9) for variables \( \{w(a), z(a), h(a)\} \) that described the life progression of a chronic infection and its deleterious effect on growth.

5.2.2 Input data for calibration

Normal growth data for weight and height of boys and girls aged up to 20 years were obtained from CDC tables available at http://www.cdc.gov/growthcharts/percentile_data_files.htm. Figure 5.1 shows the appropriate range of height and weight values and their typical deceleration (rate of change in growth rate) as a function of age. While these CDC standards were developed from USA resident, following recent WHO guidance [128], we took several quantile-specific weight-for-age and height-for-age as the normal growth-development curves (designated by \( H(a) \)) and applied them for children in schistosomiasis-endemic areas. Estimation of age-specific, and gender specific schistosomiasis-related impairment of
childhood development \( (h(a)) \) was taken from anthropometric data from community surveys performed in Coast Province, Kenya, as part of a treatment study [129]. The infection data included age-specific mean egg-counts typical for \( S. \ haematobium \)-endemic areas in eastern Kenya ([129]). In this single village, cross-sectional study of 356 children, co-infection with hookworm was 93\%, \( Trichuris \) 84\%, and \( Ascaris \) 36\%. Details of nutritional intake were not studied.

![Graph of normal growth pattern for children](http://www.cdc.gov/growthcharts/percentile_data_files.htm)

Figure 5.1 Example of normal growth pattern for children: Left-hand panels indicate CDC curves for normal weight and height at different ages during childhood, indicating growth-related gains up to age 20. Right-hand panels indicate their corresponding normal age-specific rates of change (‘growth velocity’), based on NCHS data available at http://www.cdc.gov/growthcharts/percentile_data_files.htm [128]. Data for boys are shown in the upper panels, and for girls in the lower panels. Graphed lines, from bottom to top, represent the
expected 3rd, 10th, 50th, 90th, and 97th population-based percentiles of childhood growth, respectively.

5.2.3 Developmental heterogeneity

We apply the above 'infection-morbidity-development' system to create heterogeneous communities both in terms of worm acquisition (contact patterns) and in terms of their normal and retarded growth/development. The latter requires to stratify such community into developmental *quantiles* (labeled by index \(q\)), each stratum having a specific normal growth curve \(H_q(a)\) along with NGR and RGR \(\{g_q(a); r_q(a)\}\) (equations (5.9)-(5.12)).

We assumed all hosts have statistically similar levels of exposure described by 3 parameters \(\{S_0, a_0, p\}\) of contact pattern \(\eta(a)\) (equation (5.1)), and their growth-developmental parameters: 4 for chronic damage (5.3), and 5 - for anthropometrics height and weight (each) (see Table 4.1). All 17 parameters will vary randomly (within 20% relative range) about their best-fit values estimated from the Kenyan data (below). The resulting virtual community of 200 Kenyan-like hosts consists of 5 developmental quantiles, each one having its specific normal- growth curve and randomly varying `exposure’ + ‘development’ parameters. The underlying assumption here is that developmental quantile of each host is independent of his/her behavioral pattern.
5.3 Parameter estimation

5.3.1 Main steps

Let us note that the available data sets from Kenya are of community level: the egg-count (worm burden) represents a community geometric mean rather than individual life history of infection; and these for height and weight are community median. In order to interpret it precisely in our model, we need to proceed in three steps. First we calibrated the transmission part of the system (worm burden accumulation), then extended it to calibrate its ‘accumulated inflammation + growth impairment’ part, and finally an ensemble of parameters values represent a heterogeneous community. The first two steps are to get a reasonable representative of the population.

Step 1. There are only 3 unknown parameters in the worm equation (5.3). And the solution to it with initial value $w(0) = 0$ (newborn are not infected) can be analytical expressed as

$$w(a; S_0, p, a_0) = S_0 \int_0^a e^{r(t-a)} \eta_p(t/a_0) dt \quad (5.13)$$

where $\eta_p(t/a_0)$ is the infection pattern as in (5.2). Using this express to fit egg count data (representing worm burden), we are able to get the least square estimate $\{ \hat{S}_0, \hat{p}, \hat{a}_0 \}$.

Step 2. Having estimated transmission parameters of (5.1), we use the resulting life-history $w(a; \hat{S}_0, \hat{p}, \hat{a}_0)$ as an input for morbidity equation (5.3) along with developmental equation (5.9). This allowed us to calibrate chronic morbidity parameters.
and two sets of developmental parameters \( z_0, \phi_0, \phi_1, q, K \) (one for height, the other for weight). Note that the parameter \( \rho \) in (5.3) is excluded from estimation after parameter reduction. Once again our goal is to find most reasonable estimates of the 14 in-host parameters (4 for chronic infection + 10 for developmental retardation) by fitting the height and weight data. We take US median height/weight as the norm (\( H \) in equation (5.10)) and use Monte Carlo simulation to get the most reasonable estimate (will be explained in detail in the following subsection).

**Step 3.** The point estimates of the unknown parameters we get from the above 2 steps can only represent the development history of a typical host in the community. Assuming that infection history and developmental patterns are similar for all individuals in the community, we construct the virtual community in the following way: the community is made up with 5 quantile groups: 5%, 25%, 50%, 75%, and 95%; the ratio of population of the groups is 1:3:2:3:1. That is, if the community has population size 200, then there are 20 people in the 5% group, 60 in 25% group, and so on. When building their developmental histories, each group has different standard norm (height/weight for the corresponding US percentile) but the parameters are random perturbation (up to 20% ) from the best estimate from step 1 and step 2. For example, one individual in the 25% quantile group has a height/weight development history given by the solution to the system (5.1), (5.3), (5.9) with \( H(a) \) being 25% percentage height/weight of US counterparts and model parameter values randomly chosen from an interval \((0.8, 1.2)\times \) point estimate.
A virtual community built in such a way has several advantages: first, it takes into account possible data errors (egg-counts are notoriously inaccurate) and behavioral heterogeneity; second, it allows us to predict the control outcomes more accurate and reliable. Figure 5.2 illustrates such heterogeneous group of 200 hosts over 20-year history (its envelope, mean curve and a few selected cases against real data 'dots' and the 'best-fit' curve).

5.3.2 Monte Carlo Simulation

First of all, we create the following error function on the 14-dimensional parameter vector $\theta$:

$$\text{error}(\theta) = \sum_{i \in \text{ages}_{\text{data}}} (\text{weight}(_i \theta) - \text{weight}_{\text{data}})_i^2 + \sum_{i \in \text{ages}_{\text{data}}} (\text{height}(_i \theta) - \text{height}_{\text{data}})_i^2 (5.14)$$

Here $\{\text{weight}(_i \theta), \text{height}(_i \theta)\}$ is part of the numerical solution to the coupled differential system with equations (5.1), (5.3), (5.9) for weight and (5.9) for height with prescribed value of $\theta$. In other words, weight and height of a human host result from the same schisto infection history and the resulting accumulated damage. The subscript index $i$ stands for a specific age at which height and weight were included in the data sets.

Then we are to decide a proper range for each of the unknown parameters. Noting that all the parameters should be positive and some (the relative rates) should be between 0 and 1, we need to further narrow the ranges so that the simulation will be more efficient. To do this, we take advantage of the manipulation function in Mathamatica 7. We choose a proper value for the parameter vector by trying a wide range of flexible
values and checking both the resulting error value and the visual distance of the solution curve and data points. Figure 5.2 serves as an illustration.

![Interactive manipulation in Mathematica 7](image)

**Figure 5.2** Illustration of interactive manipulation in Mathematica 7. The left panel is the control panel to choose the values of parameter. The right panel shows the corresponding result.

In this way, we decide the proper ranges of parameters. Without further information of parameters, we prescribe uniform as the prior distribution on the
determined range for each parameter, that is, any value in the range has the equal probability to be chosen in the simulation. A choice is said to be ‘good’ when this choice of parameter values (one for each of the 14 parameters) produce the small value error based on (5.14).

We tried a random ensemble of 10,000 parameter choices and selected the best one with least error (shown in Table 5.1). Based on this best choice, we built a virtual community as explained in Step 3. Figure 5.3 shows community envelopes of the worm infection, height and weight development histories. Note that boys and girls are different quantitatively both in infection and growth development, so they are treated separately. The resulting community nicely represents a real community in the sense that anthropometrics height and weight at a certain age (e.g. 20yr ) at a community level follow closely normal distributions.(see Figure 5.4)
Figure 5.3  Envelopes of schisto-infection intensity (worm burden represented by egg count) and growth patterns in terms of height (in cm) and weight (in kg) for the hypothetical heterogeneous Kenyan communities (one for girls and the other for boys), each of which is made up of 5 quantile groups i.e., the 5th, 25th, 50th, 75th, and 95th percentiles, containing 20, 60, 40, 60, 20 individuals, respectively. Upper panels for boys, lower for girls, left for egg count, middle for height and right for weight. In each plot, the solid thick curve (if applicable) represents the US median; the dots represent data and the solid gray line is the best-fit curve with parameters shown in Table 4.1.
Figure 5.4 Histograms of growth achievement in terms of height and weight at age 20 for the communities without treatment. Upper panels are for boys, lower for girls, left for height, and right for weight. Note that they are approximately of normal distribution.

Figure 5.5 Projected effects of treatment at ages 6, 9, and 12 yr (with the fraction of worms killed in each session being 90%) on worm burden, accumulated morbidity, and growth of an individual randomly selected from the communities. Solid curves are for baseline (untreated) and dashed for with the treatment.
Table 5.1 Simulation parameter descriptions and their best-fit values based on infection and anthropometric data from Kajiwe Village, Kenya [129].

<table>
<thead>
<tr>
<th>Type</th>
<th>Symbol</th>
<th>Description</th>
<th>Girls height/weight</th>
<th>Boys height/weight</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Worm burden</strong></td>
<td>$S_0$</td>
<td>Maximum force of infection among ages</td>
<td>31.2</td>
<td>52.64</td>
</tr>
<tr>
<td></td>
<td>$a_0$</td>
<td>Age when rate of new infections begins to decline</td>
<td>8.68</td>
<td>7.08</td>
</tr>
<tr>
<td></td>
<td>$p$</td>
<td>Exponent in contact pattern</td>
<td>15.13</td>
<td>2.96</td>
</tr>
<tr>
<td><strong>Morbidity</strong></td>
<td>$v_0$</td>
<td>Maximum morbidity resolution rate</td>
<td>10.18</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
<td>Rate of min/max of resolution rate</td>
<td>.2</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>$a_i$</td>
<td>Threshold age for resolution</td>
<td>11.26</td>
<td>8.33</td>
</tr>
<tr>
<td></td>
<td>$q_i$</td>
<td>Hill exponent for morbidity resolution</td>
<td>5.47</td>
<td>1.</td>
</tr>
<tr>
<td><strong>Remedial growth</strong></td>
<td>$z_0$</td>
<td>Threshold of stunting factor $\phi$</td>
<td>7.67/9.85</td>
<td>122.96/103.61</td>
</tr>
<tr>
<td></td>
<td>$\phi_0$</td>
<td>Maximum growth reduction rate due to morbidity</td>
<td>.5/.13</td>
<td>.52/.18</td>
</tr>
<tr>
<td></td>
<td>$\phi_1$</td>
<td>Baseline stunting rate</td>
<td>.99/.82</td>
<td>.96/.8</td>
</tr>
<tr>
<td></td>
<td>$m$</td>
<td>Hill exponent in $\phi$</td>
<td>7.1/5.15</td>
<td>4.93/1.7</td>
</tr>
<tr>
<td></td>
<td>$K$</td>
<td>Ratio of remedial to normal growth rate</td>
<td>11.13/8.</td>
<td>9.37/7.84</td>
</tr>
</tbody>
</table>
5.4 Drug treatment simulation and analysis of optimal strategy

Anti-schistosoma drug treatment was modeled into the system as adding a loss rate to the natural death rate, as has been introduced in the method section. Efficacy of the commonly used drug, praziquantel, can be as high as 90% in killing adult worms. So we set efficacy \( e \) in expression (5.5) to be .9. Coverage of a treatment in a community is interpreted as the fraction of people treated (randomly selected for treatment) in the whole population of the community. The virtual heterogeneous community we constructed above enables us to do treatment experiments with varying coverage.

Based on modeling of the possible remediation of schistosomiasis-associated growth retardation observed among untreated children (shown in Figure 5.3), Figure 5.5 demonstrates the modeled growth improvements with a hypothetical, three-session praziquantel (PZQ) treatment regimen (given at ages 6, 9, and 12 yrs of age) for a typical individual. The analysis indicates that, despite continuing transmission and risk of reinfection, improvements can occur gradually at the end of in the childhood (e.g. at 20 yr old) in terms of height and weight for treated as compared to untreated children. The distributions of heights or weights for the untreated communities are close to normal distributions, as expected (shown in Figure 5.4). Therefore, we reflect the effect of drug treatment through shift of the population mean and standard deviation of height/weight at age 20 yr. Results of the model predict the potential for recovery from infection-associated weight deficits, and reductions in stunting (impaired height gain) for both girls and boys, especially in the treatment campaigns with greater adherence (i.e., 50% and
80%). Numerical details of the projected outcomes at the community level are presented in Table 5.1.

We next explored the relative impact of three different regimens currently recommended by WHO [96] for school-aged treatment of schistosomiasis in high, medium, and low prevalence areas. These are: i) treat children every year from age 5 to 15 yrs of age in high (≥ 50%) prevalence areas, ii) treat children every 2 years from age 5 to 15 yrs of age in medium (10% to 30%) prevalence areas, and iii) treat children on school entry and at primary school completion in low (< 10%) prevalence areas. We also examined differences in outcomes among programs that commenced treatment either at age 4 or 6 yr of age. Table 5.2 shows the relative height and weight (to US median) achieved at age 20 for these different strategies at different levels of adherence. Figure 5.6 shows the same results from a different point of view, i.e., the resulting deficits (as a percentage of normal growth) in height and weight for boys and girls after either no therapy, or participation in differently timed treatment programs during the childhood years from 5 to 15 yr. From Table 5.2, we see that starting treatment at an earlier age appears to make little difference in outcomes, both for the every-1 year and every-2 year strategies; and given the high frequency, every-1 year strategies are not as efficient as their every-2 year counterparts. Also, we see that outcomes of treatment strategies at ages 6, 9 and 12 yr are comparable to those of every-2 year’. In contrast, the 2-session regimen (i.e., at only ages 5 and 15 yr) appears not to be so helpful in improving the catch-up growth, no matter how high the adherence of treatment is. Overall, boys, who
are more wasted or stunted than girls when untreated, were predicted to have better improvement after treatment than girls, for whatever strategy used.

As shown in the lower panels of Figure 5.6, in communities with 80% or better adherence to treatment, cumulative deficits (estimated for age =20 yr in the face of continuing risk for reinfection with S. haematobium) were reduced in magnitude for height outcomes by almost 89% in boys and by about half in girls. Corresponding weight deficits at age 20 were reduced by 90% for boys and by 38% in girls. Where treatment was less frequent (only 2-3 treatments in childhood) or where adherence was less good (e.g., 20% see Figure 5.6, upper panels) the impact of drug-based treatment campaigns on the childhood population’s growth and development profile was projected to be much more modest, on the order of 16-20% reductions in cumulative height and weight deficits overall.

Table 5.2 Predicted community height/weight relative to US median at age 20yr, following different treatment regimens. The first element in the parenthesis is the mean, and the second standard deviation

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Untreated</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Height</td>
<td>Weight</td>
</tr>
<tr>
<td>A. Untreated</td>
<td>0.962 (0.036)</td>
<td>0.939 (0.138)</td>
</tr>
<tr>
<td>B. Treat at school entry and completion (ages 5 and 15 yr) with:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20% adherence</td>
<td>0.962 (0.036)</td>
<td>0.939 (0.138)</td>
</tr>
<tr>
<td>50% adherence</td>
<td>0.962 (0.036)</td>
<td>0.940 (0.138)</td>
</tr>
<tr>
<td>80% adherence</td>
<td>0.963 (0.036)</td>
<td>0.941 (0.138)</td>
</tr>
</tbody>
</table>
C. Treat at ages 6, 9 and 12 yr, with:

<table>
<thead>
<tr>
<th>Adherence</th>
<th>20% adherence</th>
<th>50% adherence</th>
<th>80% adherence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.968 (0.037)</td>
<td>0.976 (0.038)</td>
<td>0.985 (0.034)</td>
</tr>
<tr>
<td></td>
<td>0.948 (0.142)</td>
<td>0.963 (0.146)</td>
<td>0.978 (0.144)</td>
</tr>
<tr>
<td></td>
<td>0.924 (0.040)</td>
<td>0.950 (0.043)</td>
<td>0.967 (0.041)</td>
</tr>
<tr>
<td></td>
<td>0.793 (0.135)</td>
<td>0.846 (0.146)</td>
<td>0.896 (0.145)</td>
</tr>
</tbody>
</table>

D. Treat every other year beginning age 6 yr, with:

<table>
<thead>
<tr>
<th>Adherence</th>
<th>20% adherence</th>
<th>50% adherence</th>
<th>80% adherence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.968 (0.038)</td>
<td>0.979 (0.039)</td>
<td>0.989 (0.035)</td>
</tr>
<tr>
<td></td>
<td>0.949 (0.143)</td>
<td>0.965 (0.147)</td>
<td>0.980 (0.145)</td>
</tr>
<tr>
<td></td>
<td>0.935 (0.042)</td>
<td>0.954 (0.046)</td>
<td>0.973 (0.043)</td>
</tr>
<tr>
<td></td>
<td>0.799 (0.142)</td>
<td>0.861 (0.154)</td>
<td>0.919 (0.150)</td>
</tr>
</tbody>
</table>

E. Treat every other year beginning age 4 yr, with:

<table>
<thead>
<tr>
<th>Adherence</th>
<th>20% adherence</th>
<th>50% adherence</th>
<th>80% adherence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.968 (0.038)</td>
<td>0.979 (0.035)</td>
<td>0.989 (0.035)</td>
</tr>
<tr>
<td></td>
<td>0.949 (0.143)</td>
<td>0.965 (0.147)</td>
<td>0.980 (0.145)</td>
</tr>
<tr>
<td></td>
<td>0.935 (0.042)</td>
<td>0.955 (0.046)</td>
<td>0.974 (0.043)</td>
</tr>
<tr>
<td></td>
<td>0.799 (0.142)</td>
<td>0.861 (0.155)</td>
<td>0.920 (0.151)</td>
</tr>
</tbody>
</table>

F. Treat every year beginning age 6 yr, with:

<table>
<thead>
<tr>
<th>Adherence</th>
<th>20% adherence</th>
<th>50% adherence</th>
<th>80% adherence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.969 (0.039)</td>
<td>0.980 (0.040)</td>
<td>0.991 (0.035)</td>
</tr>
<tr>
<td></td>
<td>0.949 (0.143)</td>
<td>0.965 (0.148)</td>
<td>0.982 (0.145)</td>
</tr>
<tr>
<td></td>
<td>0.936 (0.043)</td>
<td>0.956 (0.048)</td>
<td>0.976 (0.044)</td>
</tr>
<tr>
<td></td>
<td>0.802 (0.146)</td>
<td>0.868 (0.160)</td>
<td>0.930 (0.155)</td>
</tr>
</tbody>
</table>

G. Treat every year beginning age 4 yr, with:

<table>
<thead>
<tr>
<th>Adherence</th>
<th>20% adherence</th>
<th>50% adherence</th>
<th>80% adherence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.969 (0.039)</td>
<td>0.980 (0.040)</td>
<td>0.991 (0.035)</td>
</tr>
<tr>
<td></td>
<td>0.949 (0.143)</td>
<td>0.965 (0.148)</td>
<td>0.982 (0.145)</td>
</tr>
<tr>
<td></td>
<td>0.936 (0.043)</td>
<td>0.957 (0.048)</td>
<td>0.977 (0.044)</td>
</tr>
<tr>
<td></td>
<td>0.802 (0.146)</td>
<td>0.869 (0.160)</td>
<td>0.931 (0.155)</td>
</tr>
</tbody>
</table>
Figure 5.6 Projected impact of different schistosomiasis treatment regimens on accrued growth (relative to normal) at age 20 for infected girls and boys exposed to possible reinfection in endemic areas, based on model simulations of growth, reinfection and potential for catch-up growth after treatment. In each panel, clustered bars for girls and boys compare the projected outcomes of 5 proposed strategies: i) no treatment, ii) treatment at the time of school entry and completion (ages 5 and 15 yr), iii) treatment at ages 6, 9, and 12 yr old, iv) treatment every 2 years and v) treatment every year during school-ages from 4 up to 15 yr. Upper panels show substantially reduced impact when community participation is low (20%), as compared to where projected uptake and adherence are high (80%, lower panels).
Chapter 6    Conclusions

In the previous chapters, we’ve talked in detail a snail model with seasonal rainfall as the main driving environmental factor; a new modeling approach to account for over-dispersed pattern of parasite distribution in a host population; a dynamic relationship between the chronic Schistosoma infection and childhood growth retardation. Using field data from Msambweni region in Kenya, the models were calibrated and applied to estimate outcomes of different intervention strategies. We shall discuss the main findings and shortcomings of these models in this chapter.

In Chapter 2, we developed a population model for snail dynamics. With rainfall as the main input, the model outputs a seasonal snail population dynamic. We also extended the basic snail population model into a more sophisticated two-stage model to accommodate the time-lag for snail maturation. Comparing with the previous works of Liang [30] and Woolhouse [31], this two-stage snail population model is more concise and more realistic because it includes the most critical environmental driver in that area, rainfall, and accounts for the essential mechanisms of snail population establishment and growth in statics habitats (pools). With our model and computational tools, we can explore all choices of timing of the snail control. In particular, we found that the most effective intervention should be implemented at the end of first raining season. This result agrees with that from an analysis of static habitats in St Lucia in east Africa by Jordan et al [54]. We also estimated the effect of snail controls with different frequencies, and found that monthly application of mollusicides is intensive enough to reduce annual snail
density by more than 80%. Although our model was validated in a particular study area in Msambweni (Kenya) where seasonal rainfall is the principal driving force of snail growth and snails live in static habitats, some of its basic principles may be applied to other environments. One just needs to identify specific mechanisms and environmental factors that most strongly affect snail dynamics. For instance, for flowing snail habitats such as rivers as described in [54], modeling would involve different effects such as snail immigration and migration.

In Chapter 3, we proposed a new approach to modeling schistosomiasis transmission (and other vector-borne macroparasites), based on a Stratified Worm Burden (SWB) formulation. In this approach, a community of hosts is subdivided into burden strata in terms of their infection levels. The transitions among strata depend on (i) force of infection (determined by the transmission environment and human risk factors); (ii) worm attrition (natural or drug-induced); and (iii) human demographics (aging, mortality, growth sources). Such stratified systems offer many advantages compared to the conventional MWB approach [23], or Ross SIR –type[140] prevalence systems. Indeed, both MWB and SIR scheme can be directly derived from SWB system. By naturally encoding both types of information, these models allow refined calibration procedures with the available data. Furthermore, they can also include demographics and/or other measured (\textit{e.g.}, behavioral) risk factors for structured human populations (age, occupation, etc.). Finally, SWB naturally explains highly aggregated burden distribution in host populations without imposing ad-hoc distributions.
We also demonstrated model calibration procedures for the SWB approach in several settings: the single homogeneous population, a population with multiple risk/age groups (child-adult systems), and a more complex area with geographically distributed human-snail populations and contact patterns. In each case, the modeling combines demographics (population levels, births, deaths, and aging), human-snail infection levels (egg-counts and prevalences), and for distributed systems, a minimum of environmental data (snail habitats, human contact patterns, etc.).

In simple cases such dynamical (differential) systems can be analyzed mathematically to determine their endemic equilibrium levels in terms of transmission parameters such as the Basic Reproduction Number, $R_0$. In all cases they are amenable to efficient numeric implementation using desktop software, despite large number of variables and parameters. For instance, a complete list of variables for a distributed human-snail system would typically include number of population sites x population age/risk groups x burden strata + all snail variables. Here, we were able to implement it using accessible commercial desktop software.

The SWB methodology has many potential applications. In calibration, it allows using both infection intensity and population prevalence data; in application, it enables one to use more a reasonable formulation for the treatment effect. Knowing the distribution of infection intensity, we are able to do more flexible treatment simulations such as treating only the highly infected subgroups. We also conducted a simple validation test and sensitivity analysis of our calibration scheme to uncertainties in the
The results show good agreement of model simulation and new data and the system is not so sensitive to small perturbation of parameters. This means our model is applicable and reliable to draw predicative conclusions.

In Chapter 4, we provided further details and applications of our SWB methodology in exploring the predicted impact of different control strategies such as the currently advocated targeted treatment of school-aged children [122] as compared to community-based delivery to children and adults [141]. The effect of drug treatment in SWB model is formulated through immediate transition from higher strata to lower strata. It is more realistic compared to traditional setting where treatment is modeled as an additional term to the natural mortality of schistosome worms, since the latter is a far slower process. Based on our model simulation, we found that prescreening each year before drug treatment is less cost-effective when treatment coverage is relative small (less than 80%), but will show its benefit when coverage is high. We also found that treating adults makes little difference in bringing down the overall infection prevalence of the region and that 4-year school-based treatment (SBT) is the most efficient one among the six SCORE strategies. The three 4-year strategies with holidays in any of the years are not intensive enough to lower the village prevalences to a safe level of 10%. And in these cases, more intensive SBT is recommended for the subsequent years.

As we see in Figure 4.3, infection prevalences of the villages, no matter high-risk or low-risk at the beginning, can be reduced to a safe level with mass treatments, but will eventually go back to their respective endemic levels. These endemic levels are the same
as in the data (year 2000) we used for model calibration. For a dynamic system with a stable equilibrium, this is inevitable. However, if we change the parameters, for example, the contact rates for villagers to water sites, then theoretically we are able to change the situation. If possible, a theoretical analysis of the dynamic system to study the condition of parameters for the solution to approach a stable disease-free equilibrium is desirable. Unfortunately, it is a huge system with 10 villages and 5 water sites, a child-group and an adult-group for each village, and 15 strata (subgroup) for each group, and is difficult to analyze directly. So we have to resort to numerical simulations. On the other hand, the setting of the system, such as snail habitat and human behavior, is changing in reality, so our model gives a crude estimate of the effect of interventions.

If stationary snail populations are replaced by more realistic population dynamics driven by rainfall and snail control, the SWB model with human chemotherapy shows a more hopeful picture of reduced infection prevalence of the region. The tendency to rebound after the cessation of drug treatment is much reduced with regular snail control.

We concentrated on the specific region of eastern Kenya in our study. Other endemic districts with different geographical features may have different outcomes with the same mass-treatment strategies. However, as long as children play a more important role than adults in the disease transmission, which is the case for most *Schistosoma* species, it’s very likely that the conclusions are qualitatively similar to ours. In this sense, our predication is able to be generalized to provide strategic recommendations for disease elimination.
We’ve talked about the unstable non-disease-free equilibrium for modified Macdonald MWB model after introducing a mating probability in Chapter 3. In SWB setting, this idea can be more realistically adopted by applying the mating probability to each stratum, each with different level of worm burden. Unfortunately, it is not easy to analyze to see if such an equilibrium exists or not for SWB formulation. Our numerical simulation of a single population doesn’t imply this.

In Chapter 5, we developed a nonlinear dynamic of chronic *Schistosoma* infection and growth stunting during childhood. Our modeling analysis, benchmarked to available individual-level field data, was based on the well-recognized potential of children for ‘catch-up growth’ following recovery from chronic diseases [135]. Individuals can acutely increase their post-insult growth velocity up to four-fold after the growth restriction ceases. This is known as Type A catch up growth. In another form of catch up growth, known as Type B, puberty can be delayed by several years, allowing linear bone growth to continue for an extended compensatory period [135]. However, for schistosomiasis, the net impact of any Type B recovery is likely to be complex, in that pubertal hormonal changes *per se* have been shown to down-modulate anti-parasite inflammation and reduce nutritional deficits associated with *S. japonicum* infection [142]. We recognize that one of the limitations of our simulation model is that we are not able to distinguish the relative contributions of these two types of growth recovery. Future longitudinal studies, using careful Tanner staging for sexual maturity [142] will be needed to clarify the relative effects of Type A and Type B recovery following treatment.
Other factors, such as diet quality and co-infection with other chronic diseases, including soil-transmitted helminths, may serve to limit actual catch-up growth in treatment campaigns. Our study is limited in that it focused primarily on schistosomiasis and data on the impact of *S. haematobium* infection.

We should stress that, as constructed, our model gives a lower bound estimate of growth remediation, and the real benefits could be higher. Indeed, mass drug administration (MDA) may have a double effect in some communities— it can lower human infection levels and may also reduce transmission in some locales, particularly if high-risk adults are included in the treatment campaign. Our present simulation does not account for the coupled process of ‘human-to-snail transmission’ (only its ‘snail-to-human’ part), so that part of the model system could underestimate the effect of drug treatments on the process of contamination and snail infection. However, human-to-snail transmission remains a patchy, non-linear phenomenon in which a single infected individual (alone) can continue to contaminate one or more snail contact sites and maintain transmission for several months within any given community. This is the likely reason that mass drug treatment (MDT) programs have not reliably reduced transmission in many high-prevalence areas [75, 108]. Nevertheless, our projections do suggest an increasing benefit from repeated treatments during childhood, even in the face of continuing reinfection.

Our results suggest that repeated treatment during childhood has the potential to reverse most, but not all, growth impairment associated with schistosomiasis. In
particular, early treatment of *S. haematobium* with PZQ beginning at or before the age 6, with repeated treatments into the adolescent years, appears likely to be most effective in facilitating catch-up growth among repeatedly infected children. The gender-specific differences in growth observed among our benchmark Kenyan population were consistent with those found in treatment studies of *S. mansoni* infection in Brazil [127, 143]. In those studies, infected males were found to suffer more undernutrition, but they also had more dramatic improvements after anti-schistosome therapy [143]. Among our calibration sample of children, boys had higher average egg burdens than girls (geometric mean = 96 eggs/10 ml urine vs. 36 eggs/10 ml urine), which might explain a higher risk of inflammation with corresponding worsening of growth-related morbidity outcomes.

Data are scarce on the true prevalence of schistosomiasis among pre-school children [144]. Our analysis was calibrated on detailed information from school age children (5 to 20 yr) in one affected village. Based on this setting, our model did not indicate a benefit for early age therapy. In the future, in order to assess the potential benefit of anti-schistosomal treatment during pre-school years, it will be important to include these younger subjects in community-based studies, including more sensitive diagnostics for *Schistosoma* infection than the standard, relatively insensitive assays based on egg-detection in stool and urine [145-147]. Provided with this kind of data from areas where earlier growth deficits can be tied to schistosomiasis, our model is very likely to show that initiation of treatment in preschool years will be beneficial.
Appendix A  Stratified Worm Burden systems

A.1 Basic MacDonald system for multiple human risk groups

All snail and larvae populations \( \{N; c, m\} \) are measured by [density]/[unit snail habitat]. Our basic assumptions include

i) Snails are minimally mobile and relatively attached to their local habitat, so we ignore hydrologic (or other) transport of snails / larvae.

ii) There is an approximately uniform snail density throughout each habitat (their local numbers are part of the collected field data)

iii) All human-snail contacts and transmission (via larval densities) take place locally. So the force of human-to-snail infection (miracidial production term) is proportional to human population density/unit habitat, rather than total human population.

The basic Macdonald setup [23] for single habitat modified to include multiple risk groups (labeled by \( s \)), consists of (i) mean worm burden (MWB) and egg counts \( \{w_s, E_s\} \); (ii) susceptible-infected-patent snail densities (or prevalences) \( \{x, y, z\} \); (iii) cercarial and miracidial densities \( \{c, m\} \).
\[
\begin{align*}
\text{(Human)} & : \quad \frac{dw}{dt} = a \eta_c (\gamma + \ldots) w_c; \\
& \quad \quad E_s = \rho w_s \\
\text{(Snail)} & : \quad \frac{dy}{dt} = bm(1 - y) - \nu y; \\
& \quad \quad z = \psi y \\
\text{(Larvae)} & : \quad \frac{dm}{dt} = \sum_s Q_s - \mu_m m; \\
& \quad \quad \frac{dc}{dt} = \pi_c N_z - \mu_c c.
\end{align*}
\]

(Eq.A.1)

The miracidial production term is given by \( Q_s = \pi_s \eta_s H_s E_s \) - proportional to the population density (per unit habitat).

When specified for child-adult groups (the human part of the MWB system) takes the form

\[
\begin{align*}
\frac{dw_c}{dt} &= a \eta_c (\gamma + \mu_c + \tau) w_c \\
\frac{dw_a}{dt} &= a \eta_a c + \tau \frac{H_c}{H_a} w_c - (\gamma + \mu_a) w_a.
\end{align*}
\]

(Eq.A.2)

Here we allow for aging, i.e. transition of worm burden from ‘child’ to ‘adult’ compartment at the age of 20. Hence we estimated loss-rate \( \tau = .05 \) in the child-equation, balanced by a gain term, \( \tau \frac{H_c}{H_a} \), in the adult-equation, weighted in proportion to population fraction \( \frac{H_c}{H_a} \). In this formulation we assumed stationary populations, but the
equations can be easily adjusted for variable (time dependent) case. The notations are listed in Table 2.1.

The complete system can be reduced by eliminating larval equations, due to their ‘fast’ death rates \((\mu_c, \mu_m)\) compared to other ‘slow’ processes. So we replace \((c, m)\) with their quasi-equilibrium values

\[
c^* = \frac{\pi_c N_z}{\mu_c}
\]

\[
m^* = \sum_s \frac{\pi_M}{\mu_M} \eta_s H_s E_s
\]

The resulting human-snail system has MWB-type human equations:

\[
\frac{dw_s}{dt} = A_s N_y - (\gamma + \ldots) w_s \quad \text{(for s-group)}, \quad \text{or specifying it for a ‘child-adult’ setup we get}
\]

\[
\begin{align*}
\frac{dw_c}{dt} &= A_c N_y - (\gamma + \mu_c + \tau) w_c \\
\frac{dw_a}{dt} &= A_a N_y + \tau \frac{H}{H_a} w_c - (\gamma + \mu_a) w_a.
\end{align*}
\]

(Eq.A.3)

The snail part assuming stationary (seasonal mean) total population \(N\), consist of 3 prevalence variables: \(x\) – susceptible, \(y\) – infected (prepatent), \(z\) - patent, that obey a system of equations

\[
\frac{dy}{dt} = \Lambda(1 - y) - \nu y; \\
x = (1 - y - z); \\
z = \psi y;
\]

(Eq.A.4)

The snail force of infection \(\Lambda\) (contributed by all contact groups) is given by
\[ \Lambda = \sum_i B_j H_j w_j \]  
(Eq.A.5)

The transmission coefficients \( \{ A_i, B_i \} \) (‘snail-to-human’ and ‘human-to-snail’) are shown in Table 2.1.

**A.2 Child-adult SWB system**

Differential equations for SWB system for children and adults

\[
\begin{align*}
\frac{dh_0^c}{dt} &= -\left( \lambda_c + \mu_c + \tau \right)h_0^c + \gamma_i h_i^c + \beta \frac{H_a}{H_c}; \\
\vdots \\
\frac{dh_i^c}{dt} &= \lambda_c h_{i-1}^c - \left( \lambda_c + \gamma_i + \mu_c + \tau \right)h_i^c + \gamma_{i+1} h_{i+1}^c; \\
\vdots \\
\frac{dh_n^c}{dt} &= \lambda_c h_{n-1}^c - \left( \gamma_n + \mu_c + \tau \right)h_n^c; \\
\end{align*}
\]

(Child)

\[
\begin{align*}
\frac{dh_0^a}{dt} &= -\left( \lambda_a + \mu_a \right)h_0^a + \gamma_i h_i^a + \tau \frac{H_c}{H_a} h_0^c; \\
\vdots \\
\frac{dh_i^a}{dt} &= \lambda_a h_{i-1}^a - \left( \lambda_a + \gamma_i + \mu_a \right)h_i^a + \gamma_{i+1} h_{i+1}^a + \tau \frac{H_c}{H_a} h_i^c; \\
\vdots \\
\frac{dh_n^a}{dt} &= \lambda_a h_{n-1}^a - \left( \gamma_n + \mu_a \right)h_n^a + \tau \frac{H_c}{H_a} h_n^c. \\
\end{align*}
\]

(About)

(Eq.A.6)

The condition for stationary population in demographic equations (3.19) is

\[
\frac{H_c}{H_a} = \frac{\beta}{\mu_c + \tau} = \frac{H_a}{\tau}. 
\]
We estimated $\tau$ and $\beta$ from the demographic data of the Msambweni region. The child-adult death rates were computed from

$$
\mu_c = \beta \frac{H_a}{H_c} - \tau; \mu_a = \tau \frac{H_c}{H_a}.
$$

They depend on local populations, hence we make them site specific.

Having estimated rescaled forces of infection $\lambda_c$ and $\lambda_a$, we get the transmission coefficients and egg release rate from the equilibrium solution of system (3.20):

$$
A_c = \frac{\gamma \lambda_c \Delta w}{Ny}; A_a = \frac{\gamma \lambda_a \Delta w}{Ny};
$$

$$
\rho_c = \frac{(\gamma + \mu_c + \tau)E_c}{\gamma \lambda_c (\Delta w)}; \rho_a = \frac{(\gamma + \mu_a)E_a}{\gamma \lambda_a (\Delta w) + \tau \frac{H_c E_c}{H_a \rho_c}};
$$

(Eq.A.7)

$$
B_c = \frac{\nu \gamma A_y}{(1-y)\sum_s A_s H_s E_s}; B_a = \frac{\nu \gamma A_y}{(1-y)\sum_s A_s H_s E_s}.
$$

**A.3 Distributed system**

MWB equations for distributed human-snail meta-population system:

$$
\begin{align*}
\frac{dw_i}{dt} &= \sum_{j=1}^{m} A_{ij} N_j y_j - (\gamma + \mu_c)w_i; (i = 1,\ldots,n) \\
\frac{dy_j}{dt} &= \sum_{i=1}^{m} B_{ij} H_i w_i (1-y_j) - \nu y_j; (j = 1,\ldots,m)
\end{align*}
$$

(Eq.A.8)

MWB equations for an age-structured (child-adult) human population with simplified transmission coefficients:
Here we use local patent snail densities \( \{ Z_j \} \) instead of infected snail prevalences of earlier special cases (formulae(Eq.A.1)). Since patency conversion rates of infected snails \( \{ \psi_j = \frac{Z_j}{Y_j} \} \) are local and vary across snail sites (Table 3.3), we need to account of their contribution to human infection separately. Another modification is the force of infection for snail in (Eq.A.5)

\[
\Lambda_j = \frac{B_j}{U_j} \sum_i \eta_{ij} (A_i^c H_i^c E_i^c + A_i^a H_i^a E_i^a) \tag{Eq.A.10}
\]

Here \( \{ U_j \} \) are relative sizes of snail habitats (based on the number of contact sites of Table 2). We divide ‘human population’ by ’contact size’ \( \frac{H}{U_j} \), since humans distribute their infection in inverse proportions to contact size of each habitat. Also the mean egg count \( E_i \) is used instead of the mean worm burden \( w_i \) in the snail force of infection \( \Lambda \), as different age groups may have different egg-release rates. The transmission
coefficients $\{A_i; B_j\}$ in formula are somewhat different from earlier cases (Table 2.1).

Namely,

$$A_i = a \frac{\pi_c}{v_c} T_i^t;$$

$$B_j = \frac{a}{v_a} \frac{\pi_c}{v_c}.$$

where $T_i^t$ represents the total contact rate for the risk group $s$ in village $i$ to all water sites. Having estimated $A_i^c, A_i^a, \rho_c$ and $\rho_a$, we compute $B_j$ from the snail equation (Eq.A.7) at site $j$, using the relation between $A$’s and $B$’s equation(18):

$$B_j = \frac{\nu Y_j U_j}{(N_j - Y_j) \sum_i \eta_i (A_i^c H_i^c E_i^c + A_i^a H_i^a E_i^a)}.$$  \hspace{1cm} (Eq.A.11)
Bibliography


Validation of a chart to estimate true Schistosoma mansoni prevalences from simple egg counts. Parasitology. **114 (Pt 2)**: p. 113-121.


86. Li, Z., Schistosomiasis Transmission and Control in A Distributed Heterogeneous Human-Snail Environment in Coastal Kenya, in Mathematics. 2008, Case Western Reserve University: Cleveland. p. 69-70.


157


