Title

Exploration of Surveillance and Control Strategies for Re-emerging Schistosomiasis Environments in Sichuan Province, China - The Development and Application of an Individually-based Model

Permalink

Author

Wang, Shuo

Publication Date

2013

Exploration of Surveillance and Control Strategies for Re-emerging Schistosomiasis
Environments in Sichuan Province, China
– The Development and Application of an Individually-based Model

by

Shuo Wang

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Environmental Health Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Robert C. Spear, Chair
Professor Alan Hubbard
Professor Mark Nicas
Professor Travis C. Porco

Fall 2013

Exploration of Surveillance and Control Strategies for Re-emerging Schistosomiasis Environments in Sichuan Province, China – The Development and Application of an Individually-based Model

Abstract

Exploration of Surveillance and Control Strategies for Re-emerging Schistosomiasis
Environments in Sichuan Province, China –
The Development and Application of an Individually-based Model

by

Shuo Wang

Doctor of Philosophy in Environmental Health Sciences

University of California, Berkeley

Professor Robert C. Spear, Chair

Schistosomiasis is still prevalent in seven provinces in China despite several decades of continuous effort. Suppression of transmission in low-infection areas proves to be extremely difficult, and re-emergence of the disease in areas it was previously controlled or eliminated further complicates the situation. New surveillance and control strategies based on a better understanding of current transmission patterns are therefore crucial to the achievement of disease elimination. This dissertation explores these questions with the focus on the hilly and mountainous regions in Sichuan Province.

The main contents of this dissertation are centered on the development and simulation of a stochastic individually-based model. Focused on two cohorts representing the post-equilibrium and re-emerging environments, respectively, survey protocols and data description on human demographics, infection status, snail density, and cercarial mouse bioassay are presented and compared. A health education intervention was also introduced as an example of how individual data could be used for hypothesis testing related to model component determination. Upon development of the model, Monte-Carlo simulations were conducted to account for randomness from multiple sources including parasite acquisition, worm development, egg shedding and infection testing.

Surveillance results in the re-emerging villages indicated low sensitivity of current techniques for low infection-intensity environment. The distribution of individual susceptibility to *S. japonicum* infection appeared to be log-normally distributed with a high skewness. Given the cercarial density and water contact magnitude distributions of the surveyed populations, a considerable fraction of people had susceptibility levels that would not lead to infection. For infection surveillance purposes, sampling individuals with high susceptibility has markedly greater efficiency than other indices focused on either water or cercarial exposure. Based on these findings, priorities should be given to the development of more sensitive environmental monitoring and infection testing methods and exploration of an appropriate and quantifiable index for susceptibility. Environmental and economic modifications in addition to the traditional control methods need to be conceptualized and implemented to further conquer the disease and achieve comprehensive elimination.

1

Table of Contents

Acknowledgement

Chapter 1   Introduction

"Heading for a destination one hundred-Li away, one who has marched fifty is merely halfway accomplished."

--- Chronicles of the Warring States: Strategies of Qin, Chapter 5

1.1   Prelude

Formulated in the first decade of the twenty-first century, the concept of "neglected tropical diseases (NTDs)", a group of diseases particularly endemic in tropical and subtropical regions, has been increasingly recognized by public health professionals [1]. Although multiple health organizations have suggested NTD lists which are partially different from one another, commonly shared by all diseases in this category is their close relationship with extreme poverty, poor hygiene, lack of clean water and food, and non-mechanized agriculture in remote and inaccessible rural areas [2]. Because of these characteristics, along with the high priorities set for tuberculosis, malaria, and HIV/AIDS, often referred to as the "big three" diseases, funding for scientific research and disease control regarding NTDs has been far less than necessary [3]. As a consequence, while the availability of inexpensive drugs for many NTDs does shed light on diminishing the morbidity and mortality, the coverage of massive treatment programs is still considerably inadequate [4, 5].

On the other hand, although "neglected", the impacts of NTDs are simply too large to be neglected. There is an estimate of 1.4 billion people worldwide who are infected with at least one type of NTD [6]. While causing less deaths than the big three diseases, most NTDs can cause a series of chronic conditions from limited working capacity to severe permanent disability, resulting in approximately 57 million years of lost life (YLL) [7]. Furthermore, some NTDs have shown their ability to compromise the immune system, which greatly facilitates co-infection with other diseases and, in some circumstances, weakens the effectiveness of disease treatment and prevention [7-9]. It is due to their public health significance and tremendous social burdens that, while faced with financial and implementation challenges, the World Health Organization still set an ambitious blueprint of NTD control for the next several years, including the eradication, global elimination and regional elimination of 10 diseases [10]. Amongst these, the goal regarding schistosomiasis elimination in China by the year of 2015, a proposal in line with that of China's national government, is the main topic and motivation of this dissertation.

1.2 The disease

1.2.1 Global epidemiology

Schistosomiasis is one of the most of the prevalent parasitic diseases of humans. Referring to the statistics in 2011, at least 240 million people were infected with the disease, and more than 700 million people resided in endemic areas of 78 countries with various risks of infection [11]. Accounting for approximately 280 thousand deaths per year [12], schistosomiasis is also the neglected tropical disease that causes the most mortality [2]. There are three important species of the schistosome blood fluke that cause human schistosomiasis: *Schistosoma mansoni* which occurs mainly in Africa and South America, *Schistosoma haematobium* which is mainly endemic in Africa and the Middle East, and *Schistosoma japonicum*, the Asian species that occurs in the Far East including China [13]. Similar with other NTDs, the prevalence of schistosomiasis is highly correlated with underdevelopment and low income: it is estimated that 90% of infected people live in Africa, mainly the sub-Saharan regions where more than half of the population live below the extreme poverty level [1, 14]. Lacking sufficient financial support from both inside and outside these undeveloped countries, infection surveys have not been comprehensively conducted in many endemic regions of Africa [15]. Therefore, it has been suggested that the current prevalence and morbidity of schistosomiasis are considerably underestimated. Some people further suspect that the real burden of the disease is actually much closer to that of malaria and tuberculosis [16].

1.2.2 Etiology

Schistosomiasis is transmitted between intermediate host snails and various mammals including humans which are the definitive hosts. While differentiated by the size of parasites, the corresponding species of the snail hosts and the target organs, the three main species of schistosome present similar life cycles including a sexual reproduction phase in the adult worm stage and the asexual reproduction phase in the larval stages (Figure 1.1). Taking *S. japonicum* for example, the asexual phase initiates when eggs are excreted into the water with mammalian feces. Under optimal conditions, the eggs hatch and release free-living miracidia, which aims to find and infect the intermediate snail host, *Oncomelania hupensis*. Within several weeks to months, schistosomes inside the snail hosts develop into two generations of sporocysts and ultimately, the infective cercariae. Upon release from the snail, the tailed cercariae are able to penetrate human skin, which takes as short as 10 seconds, to cause the infection [17]. During this process, cercariae shed their tails and become schistosomulae, which then move through the circulatory system and to reside in the portal veins where they will mature to adult worms in four to eight weeks [13]. Migrating to the mesenteric venules in various locations, the paired worms produce the pathogenetic eggs which will eventually circulate to liver or be shed in stools. Finally, accompanying the emission of these eggs into water is the reactivation of the asexual phase.

It is thus evident that transmission relies greatly on interactions between humans and the environment. Human water contact is the most direct cause of infections, with the odds varying by frequency, duration and location of such activities. On the other hand, the way by which people dispose of their excretion is a key factor in the transmission cycle. In many rural areas of developing countries, residents, either due to absence of

private or public toilets or for convenience purposes, tend to urinate and defecate in the fields which are potentially connected with water environments through geographical characteristics and/or the runoff from precipitations [14]. Also, night soils (human feces) are broadly used as fertilizers in these less developed regions instead of chemical compounds. All of these behaviors can lead to the entry of urinary and fecal materials, together with the schistosome eggs they may contain, into surface waters [14, 18].
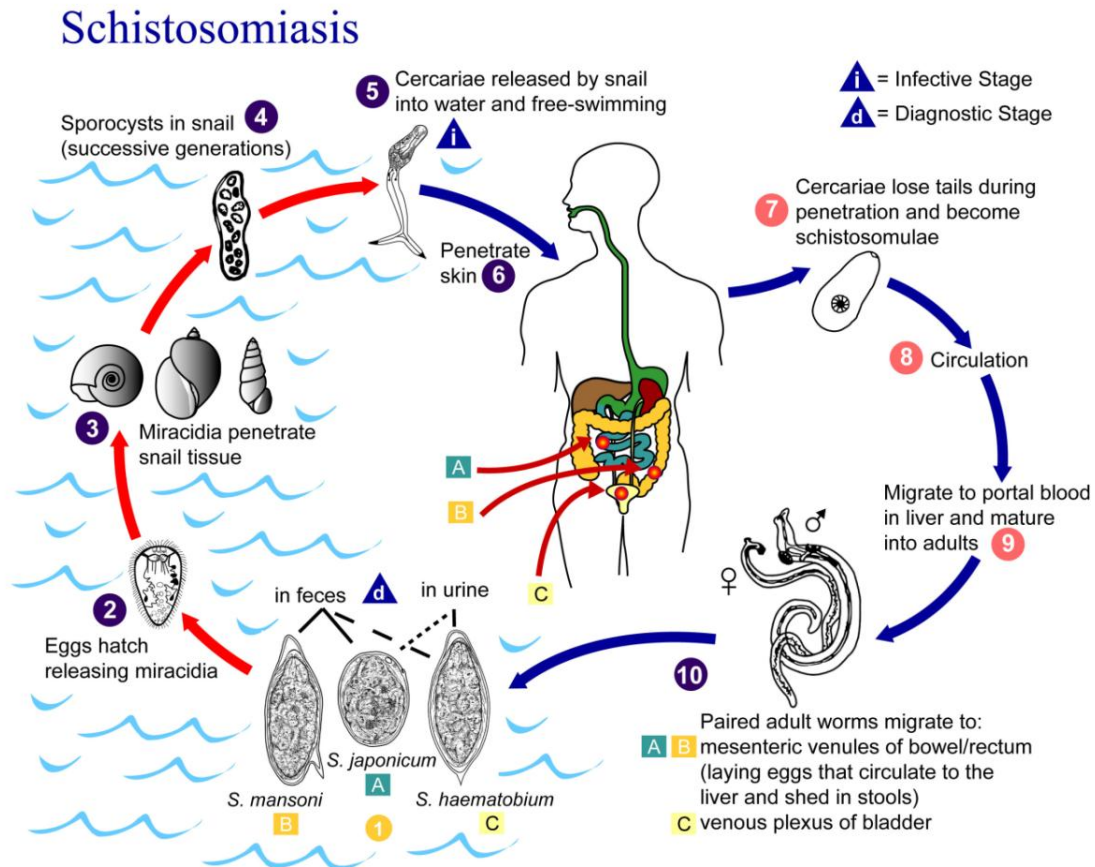
Schistosomiasis

Figure 1.1 Transmission cycle of schistosomes [19]

1.2.3   Clinical manifestations

Schistosomiasis can cause a variety of clinical manifestations which are associated with acute or chronic infection. Acute infection typically occurs among people who are exposed to cercariae for the first time. However, *S. japonicum* may also trigger acute infection in individuals with an infection history and repeated treatments [17]. The first manifestation of acute infection is usually dermatitis, which occurs within 24 hours after exposure and develops a pruitic rash which lasts for a few days. Later on, a much more severe consequence, namely the acute schistosomiasis or the "Katayama fever", can develop approximately one month after exposure and lead to symptoms such as fever, cough, increased eosinophils, and enlarged liver, spleen or lymph nodes. Although neither of the consequences is species-specific, acute infection of *S. japonicum* often causes more damage compared to the other two human species.

The severity of chronic infection is associated to the schistosome type, intensity and duration of the infection [13]. For *S. japonicum*, common symptoms include weakness, abdominal pain, irregular bowel movement and bloody stool, although more than half of chronic cases are asymptomatic. Corresponding to high infections, clinical diseases of chronic infection include liver fibrosis, hepatosplenomegaly, ascites, and ectopic lesions in the brain. Especially, severe infection in children may hinder normal growth and result in dwarfism. Co-infection with hepatitis B has also been suspected, which is of particular concern to China because of the high prevalence of the latter.

1.2.4 Interventions

The life cycle of schistosome reveals that transmission is conditioned on three key components: an optimal water environment in which intermediate host snails can inhabit and reproduce, access of parasite-containing excretions into the environment, and interactions between definitive hosts and cercaria-contaminated areas through water contact. Corresponding to each component, intervention methods are designed to target either the intermediate host or the definitive host, mainly humans.

Target the intermediate host:

Snail control with chemical molluscicides is a direct method to eliminate the intermediate hosts [20]. As one of the widely adopted methods, snail control has played a significant part in eradicating the disease in Japan and lowering the transmission level tremendously in China [21, 22]. The most commonly used chemical, niclosamide, has also proved to be effective in killing all stages of snails and the larval stage of the schistosome [12, 23]. However, complete eradication of snails through molluscicide only is extremely difficult, if possible at all, to achieve. Often times, residual snails can still sustain the transmission and, without necessary follow-up controls, reproduce rapidly to reach the pre-control level again. Furthermore, the high cost of molluscicide makes it financially impractical in many poor but highly endemic countries; the toxicity of niclosamide to the environment and the ecological system also adds to the disadvantages of chemical molluscicide [12].

Environmental modification aims at permanently reducing the habitat of snails. Compared with molluscicide, it ensures not only short-time but also sustainable effects of eliminating snails and preventing further reproductions. Primarily a one-time project, it can also be more cost-effective than repetitive snail controls [24, 25]. That having been said, environmental modification can still require large investments for extensive construction work. Barriers may also appear when such projects, which may very likely correlate with economic development and urbanization planning, need approvals and funding from multiple governmental agencies.

Targeting the definitive host (humans):

Chemotherapy with praziquantel is the most important means of reducing the infection intensity and morbidity of infected hosts [9]. The drug is able to kill the adult worms of all schistosome species with rather high effectiveness, which is estimated to be between 70% and 95% [5, 26]. Also rigorously tested for its safety, praziquantel can be applied to people of a broad range of ages, from school-aged children to the elderly

population. Despite the promising and inexpensive drug, unfortunately, treatment coverage is extremely low – of the 240 million people infected in 2011, only 11% were indeed treated [11]. Problems associated with massive applications of praziquantel have also emerged: First, because praziquantel eliminates only adult worms but not eggs or schistosomulae older than two days, timing is often a crucial determinant of the treatment effects [27]. Second, residual infections are common due to imperfect treatment effectiveness, particularly among people who are heavily infected. Third, noncompliance with treatment is of specific concern in endemic regions where chemotherapy is conducted repeatedly. Finally, resistance to praziquantel of certain schistosome species has been reported since 1994. Although convincing evidence is still lacking, this problem certainly requires scrutiny when mass chemotherapy is extended to the majority of infected people [28].

Sanitation improvement attempts to lower disease transmission by protecting humans from snails or protecting snails from humans. The former goal can be achieved through the development and application of municipal water and irrigation systems, so that residents can reduce farming and living related exposure to potentially contaminated water sources. The notion of achieving the latter goal is to prevent schistosome eggs from entering the water environment. This can be fulfilled through building household toilets, especially the biogas toilet, where human wastes are collected, stored and fermented for a week or more until schistosome eggs, if any, are destroyed [29]. In addition, the use of toilets can also lower the chance of people excreting in the fields. In terms of the downside, like environmental modification, large funding and labor inputs are necessary for such projects. On the other hand, more important than building up the facilities is to make sure that people comply and make maximum use of them, which cannot be achieved without education.

Health education to individuals living in endemic regions is an indispensable part of schistosomiasis control. Cheap and feasible to conduct, it can be easily applied and tailored to all people regardless of their age, occupation and education level. Most education programs focus on three aspects: teaching knowledge about schistosomiasis, enhancing individuals' attitudes towards disease diagnosis, treatment and prevention, as well as changing individuals' behaviors on water contact, excretion, and self-protection . A more detailed discussion on health education will be presented in Chapter 2. Yet, it is worth mentioning at this point that however seemingly modifiable through health education, human behavior is, as a matter of fact, highly resistant to change, sometimes even despite increased knowledge and awareness regarding the disease.

1.2.5 Diagnosis

Without exaggeration, diagnosis is the most important prerequisite of disease control. It depends thoroughly on the results of infection testing to inform the design of the control strategy, such as defining target populations, selecting intervention methods, setting priorities and allocating resources, etc. Many techniques have been developed for schistosomiasis diagnosis and most fall into either one of the two categories: direct parasitological methods or immunological methods [12]. The former category includes the miracidial hatching test and the Kato-Katz smear test, which are able to directly detect schistosome eggs in urine or feces and yield quantitative and/or qualitative results. Due to the feasibility and relatively low cost, they are also the most widely adopted methods of field tests in endemic areas. However, as specific as these methods are,

insensitivity to low infection levels is a major drawback, especially for the Kato-Katz test which infers the infection status and intensity from only a tiny portion of the sample [30, 31]. On the other hand, antibody and antigen based methods, as well as the PCR test, have been researched extensively in recent years to pursue higher sensitivity. Although there have been some reports suggesting the superiority of these immunology-based methods, more validation is needed to endorse these findings and reveal under what circumstances could this conclusion stand [32]. Effectiveness for low infections and affordability of these methods are key factors in their applicability in massive-scale field diagnosis.

1.2.6 Monitoring

Other than human infection testing, another important means of disease surveillance is to monitor parasites and snails in the environment. Ideally, a monitoring system should be designed such that the techniques are sensitive enough to adapt to low-transmission environments, the locations selected can represent and imply the overall distributions, and the results are indicative of human infection patterns. Furthermore, because the distributions of snails, cercariae and miracidia, all free-living in the water, are highly subject to local climatic characteristics, data collection regarding these targets needs to be not only spatial but also longitudinal.

There are a variety of ways to make use of the surveillance data. A direct application, for instance, is to identify the areas with large density of snails, particular infected snails, and deploy molluscicide accordingly for snail control [20]. The monitoring data is also important for the development of mathematical models which mimic the transmission cycle of schistosomiasis. As the intermediate host, snail is an indispensable state variable which requires valid data input from field surveys. Information on parasites can also contribute to the calibration of such dynamics models [24, 33-36]. However, like human infection testing, current techniques of snail and parasite monitoring are also flawed due to their insensitivity in low-transmission environments [37]. Despite these important applications, much less effort have been devoted to advancing surveillance techniques in contrast to other control methods.

1.3 Schistosomiasis in China

1.3.1 History

Having eggs identified in a female corpse dating back to the Western Han Dynasty, *Schistosoma japonicum* is believed to have been prevalent in China for more than 2100 years [38]. Although symptoms indicating the Katayama fever were recorded in ancient Chinese medicine books thousands of years ago, the earliest case confirmed by modern diagnostic methods occurred much later in 1905, when this disease was, for the first time, officially documented in China [38, 39]. Unfortunately, decades of civil and international war only facilitated the spread of the disease, and it was not until the advent of People's Republic of China that government-funded control programs were initiated. Conducted during the mid-1950s, the first national infection survey revealed that there were then about 11.6 million infected people and some 100 million people at risk. Geographically, 380 counties in 12 provinces were defined as endemic regions [18]. All distributed along the Yangtze River watersheds, the 12 provinces could be further divided into two categories due to their territorial characteristics: the marshland environment

represented by Hunan and Jiangxi, and the hilly and mountainous environment represented by Sichuan and Yunnan [40].

In response, priorities of disease control in earliest years were focused on snail elimination and snail habitat destruction through labor-intensive physical methods, incorporated with health education to mobilize such activities [41]. The later application of chemical molluscicide greatly improved the efficiency of snail control. However, it also became clear from accumulated experience that additional measures had to be taken for more sustainable outcomes. With the availability of the effective drug, praziquantel, mass chemotherapy started being applied to people living in endemic regions, which significantly reduced the prevalence and morbidity of infected people [42]. Meanwhile, health education was extended to cover much broader contents, such as behavior changing and improving the compliance of infection testing and treatment [43, 44]. Environmental modification projects have also been being launched since late last century, and the notion of integrated control, which tends to incorporate a combination of methods, was finally recognized as important. With more than five decades of continuing efforts, the number of infected people have dropped by more than 93% to approximately 726,000, with five provinces having wiped out the disease, according to the 2004 nationwide survey [18].

To achieve long–term comprehensive controls, a target to reduce the prevalence to less than 1% by 2015 was set by the National Schistosomiasis Control Program in 2004, and led to the selection of Sichuan province as a pilot region to achieve transmission elimination through the use of the integrated control strategy [45]. However, pursuit of this target has been considerably challenged by the phenomenon occurring in Sichuan, namely the re-emergence of schistosomiasis.

1.3.2 Re-emergence in Sichuan

Re-emergence of schistosomiasis is defined as the confirmed diagnosis of new human acute infections or local presence of infected snails in an area where transmission was formerly terminated or suppressed below detectable levels [36]. According to the results of surveillance in late 2004, re-emergence had occurred in eight counties in Sichuan, with seven located together in the hilly regions of northern Sichuan and one in the mountainous areas of the Anning river valley in the southwest. The average "return time" for counties of transmission control (e.g. from the declaration of disease control to report of first human infection with confirmed local transmission) was 8.1 years, with the shortest being two years and longest fifteen years. Human prevalence of infection ranged from 0.9% to 12.7%; acute cases were reported in seven of the eight counties [36, 46]. While successively detected, little was known about the cause of the renewed transmission in the eight counties where effective control had been attained. What was clear, however, was that the traditional control strategy had proven unable to prevent re-emergence or, to date, elimination of the disease. It was in this specific context that this research was conducted, with the aims of exploring the potential factors underlying re-emergence in the low-transmission setting, and making practical suggestions on surveillance and control programs to terminate future re-emergence and ultimately, the transmission of schistosomiasis in Sichuan [47].

7

1.4 Dissertation structure

1.4.1 Preliminary results

As a major part of this dissertation, the development of the individually-based model, including the primary assumptions and the model structure, was based upon some preliminary findings resulting from the same survey data as will be described in Chapter 3. These findings suggested, through hypothesis testing using statistical models, that: (1) no village-level characteristics were associated with the likelihood of re-emergence, (2) social connectivity, or individual mobility, did not appear to be a factor in individual infection or disease spread, and (3) individual water exposure measured by the total amount of time was only a weak indicator of infection. While, to some extent, these findings failed to support some of our previous assumptions regarding village and individual risk factors, they did suggest the plausible existence of stochasticity from multiple sources, and led us to suspect individual susceptibility to *S. japonicum*, an issue that has rarely been explored particularly through epidemiological studies [48].

1.4.2 Dissertation chapters

In Chapter 2, I present a randomized-controlled trial educational intervention and evaluate its impacts on villagers' knowledge, attitude and behavior regarding disease transmission. Although the results, due to lack of statistically significant findings, were not eventually incorporated in the individually-based model, the intervention itself, along with the surveys included, provides an example of individual data collection and use of the data to test for hypotheses that are potentially applicable to the model development. This example also attributed to the preliminary results noted above.

In Chapter 3, I present the protocols of field surveys from which the data to be incorporated in the individually-based model was collected. Two separate studies are included to represent quite different transmission scenarios. The first study, conducted earlier from 2000 to 2002 and having been reported extensively elsewhere, is associated with the endemic equilibrium environments where transmission is more intense and stable. The second study, designed specifically for this research and conducted in 2007 and 2008, reflects the low transmission level associate with the re-emerging situations. In addition, descriptive statistics of the survey results of both studies will also be shown.

Development of the individually-based model will be presented next in Chapter 4. In detail, I will discuss the process of parameterization, particularly for individual susceptibility and cercarial density, two pieces of information that have not been explored in our previous research. Model simulations will be specified in Chapter 5. Based on the characteristics for the two study populations, "ghost populations" will be generated randomly to explore the transmission patterns of more generic scenarios. I will first show from Monte-Carlo simulations that the distribution and magnitude of individual infections of the ghost populations are broadly consistent with epidemiological surveys. The results will then be used to explain, phenomenologically, the clustering of re-infection which we detected in our study populations, and to explore the implications regarding disease surveillance.

Finally in Chapter 6, suggestions based on the findings will be made towards disease elimination and prevention of disease re-mergence in Sichuan and other endemic regions in China.

8

References

1.Hotez, P.J. and Kamath, A., *Neglected tropical diseases in sub-saharan Africa: review of their prevalence, distribution, and disease burden.* PLoS Negl Trop Dis, 2009. **3**(8): p. e412.

2.Hotez, P.J., Molyneux, D.H., Fenwick, A., Kumaresan, J., Sachs, S.E., Sachs, J.D., and Savioli, L., *Control of neglected tropical diseases.* N Engl J Med, 2007. **357**(10): p. 1018-27.

3.Gyapong, J.O., Gyapong, M., Yellu, N., Anakwah, K., Amofah, G., Bockarie, M., and Adjei, S., *Integration of control of neglected tropical diseases into health-care systems: challenges and opportunities.* Lancet, 2010. **375**(9709): p. 160-5.

4.Reddy, M., Gill, S.S., Kalkar, S.R., Wu, W., Anderson, P.J., and Rochon, P.A., *Oral drug therapy for multiple neglected tropical diseases: a systematic review.* JAMA, 2007. **298**(16): p. 1911-24.

5.Utzinger, J. and de Savigny, D., *Control of neglected tropical diseases: integrated chemotherapy and beyond.* PLoS Med, 2006. **3**(5): p. e112.

6.Bethony, J.M., Cole, R.N., Guo, X., Kamhawi, S., Lightowlers, M.W., Loukas, A., Petri, W., Reed, S., Valenzuela, J.G., and Hotez, P.J., *Vaccines to combat the neglected tropical diseases.* Immunol Rev, 2011. **239**(1): p. 237-70.

7.Hotez, P.J., Molyneux, D.H., Fenwick, A., Ottesen, E., Ehrlich Sachs, S., and Sachs, J.D., *Incorporating a rapid-impact package for neglected tropical diseases with programs for HIV/AIDS, tuberculosis, and malaria.* PLoS Med, 2006. **3**(5): p. e102.

8.Borkow, G. and Bentwich, Z., *HIV and helminth co-infection: is deworming necessary?* Parasite Immunol, 2006. **28**(11): p. 605-12.

9.McManus, D.P. and Loukas, A., *Current status of vaccines for schistosomiasis.* Clinical microbiology reviews, 2008. **21**(1): p. 225-242.

10.WHO, *Accelerating work to overcome the global impact of neglected tropical diseases – A roadmap for implementation.* Geneva: WHO, 2012.

11.WHO, *http://www.who.int/schistosomiasis/en/.*

12.Gryseels, B., Polman, K., Clerinx, J., and Kestens, L., *Human schistosomiasis.* Lancet, 2006. **368**(9541): p. 1106-18.

13.Jordan, P., Jordan, P., Webbe, G., and Sturrock, R.F., *Human schistosomiasis.* 1993, Wallingford, Oxon, UK: CAB International.

14.Rollinson, D., Knopp, S., Levitz, S., Stothard, J.R., Tchuente, L.A., Garba, A., Mohammed, K.A., Schur, N., Person, B., Colley, D.G., and Utzinger, J., *Time to set the agenda for schistosomiasis elimination.* Acta Trop, 2012.

15.King, C.H. and Bertino, A.M., *Asymmetries of poverty: why global burden of disease valuations underestimate the burden of neglected tropical diseases.* PLoS Negl Trop Dis, 2008. **2**(3): p. e209.

16.Steinmann, P., Keiser, J., Bos, R., Tanner, M., and Utzinger, J., *Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk.* Lancet Infect Dis, 2006. **6**(7): p. 411-25.

17.Secor, W.E., Colley, D.G., and SpringerLink, *Schistosomiasis*, 2005, Springer: New York, NY.

18. Zhou, X.N., Guo, J.G., Wu, X.H., Jiang, Q.W., Zheng, J., Dang, H., Wang, X.H., Xu, J., Zhu, H.Q., Wu, G.L., Li, Y.S., Xu, X.J., Chen, H.G., Wang, T.P., Zhu, Y.C., Qiu, D.C., Dong, X.Q., Zhao, G.M., Zhang, S.J., Zhao, N.Q., Xia, G., Wang, L.Y., Zhang, S.Q., Lin, D.D., Chen, M.G., and Hao, Y., *Epidemiology of schistosomiasis in the People's Republic of China, 2004.* Emerging infectious diseases, 2007. **13**(10): p. 1470-1476.

19. CDC, *http://www.cdc.gov/parasites/images/shistosomiasis/Schistomes_LifeCycle.gif.*

20. Seto, E., Liang, S., Qiu, D., Gu, X., and Spear, R.C., *A protocol for geographically randomized snail surveys in schistosomiasis fieldwork using the global positioning system.* The American Journal of Tropical Medicine and Hygiene, 2001. **64**(1-2): p. 98-99.

21. Tanaka, H. and Tsuji, M., *From discovery to eradication of schistosomiasis in Japan: 1847-1996.* Int J Parasitol, 1997. **27**(12): p. 1465-80.

22. Yang, G.J., Zhou, X.N., Sun, L.P., Wu, F., Zhong, B., Qiu, D.C., Utzinger, J., and Bradshaw, C.J., *Compensatory density feedback of Oncomelania hupensis populations in two different environmental settings in China.* Parasit Vectors, 2011. **4**: p. 133.

23. Pan, J.X., Ding, K., and Wang, C.Y., *Niclosamide, an old antihelminthic agent, demonstrates antitumor activity by blocking multiple signaling pathways of cancer stem cells.* Chin J Cancer, 2012. **31**(4): p. 178-84.

24. Liang, S., Seto, E.Y., Remais, J.V., Zhong, B., Yang, C., Hubbard, A., Davis, G.M., Gu, X., Qiu, D., and Spear, R.C., *Environmental effects on parasitic disease transmission exemplified by schistosomiasis in western China.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(17): p. 7110-7115.

25. Spear, R.C., Seto, E., Liang, S., Birkner, M., Hubbard, A., Qiu, D., Yang, C., Zhong, B., Xu, F., Gu, X., and Davis, G.M., *Factors influencing the transmission of Schistosoma japonicum in the mountains of Sichuan Province of China.* The American Journal of Tropical Medicine and Hygiene, 2004. **70**(1): p. 48-56.

26. Liu, R., Dong, H.F., Guo, Y., Zhao, Q.P., and Jiang, M.S., *Efficacy of praziquantel and artemisinin derivatives for the treatment and prevention of human schistosomiasis: a systematic review and meta-analysis.* Parasit Vectors, 2011. **4**: p. 201.

27. Wang, W., Wang, L., and Liang, Y.S., *Susceptibility or resistance of praziquantel in human schistosomiasis: a review.* Parasitol Res, 2012. **111**(5): p. 1871-7.

28. Seto, E.Y., Wong, B.K., Lu, D., and Zhong, B., *Human schistosomiasis resistance to praziquantel in China: should we be worried?* Am J Trop Med Hyg, 2011. **85**(1): p. 74-82.

29. Sow, S., Polman, K., Vereecken, K., Vercruysse, J., Gryseels, B., and de Vlas, S.J., *The role of hygienic bathing after defecation in the transmission of Schistosoma mansoni.* Trans R Soc Trop Med Hyg, 2008. **102**(6): p. 542-7.

30. Lin, D.D., Liu, J.X., Liu, Y.M., Hu, F., Zhang, Y.Y., Xu, J.M., Li, J.Y., Ji, M.J., Bergquist, R., Wu, G.L., and Wu, H.W., *Routine Kato-Katz technique underestimates the prevalence of Schistosoma japonicum: a case study in an endemic area of the People's Republic of China.* Parasitol Int, 2008. **57**(3): p. 281-6.

31. Yu, J.M., de Vlas, S.J., Jiang, Q.W., and Gryseels, B., *Comparison of the Kato-Katz technique, hatching test and indirect hemagglutination assay (IHA) for the diagnosis of Schistosoma japonicum infection in China.* Parasitol Int, 2007. **56**(1): p. 45-9.

32. Fung, M.S., Xiao, N., Wang, S., and Carlton, E.J., *Field evaluation of a PCR test for Schistosoma japonicum egg detection in low-prevalence regions of China.* Am J Trop Med Hyg, 2012. **87**(6): p. 1053-8.

33. Liang, S., *A spatio-temporal modeling of schistosomiasis transmission dynamics and control in Sichuan, China*, 2003.

34. Liang, S. and Spear, R.C., *Model-based insights into multi-host transmission and control of schistosomiasis.* PLoS medicine, 2008. **5**(1): p. e23.

35. Liang, S., Spear, R.C., Seto, E., Hubbard, A., and Qiu, D., *A multi-group model of Schistosoma japonicum transmission dynamics and control: model calibration and control prediction.* Tropical medicine & international health : TM & IH, 2005. **10**(3): p. 263-278.

36. Liang, S., Yang, C., Zhong, B., and Qiu, D., *Re-emerging schistosomiasis in hilly and mountainous areas of Sichuan, China.* Bulletin of the World Health Organization, 2006. **84**(2): p. 139-144.

37. Carlton, E.J., Bates, M.N., Zhong, B., Seto, E.Y., and Spear, R.C., *Evaluation of mammalian and intermediate host surveillance methods for detecting schistosomiasis reemergence in southwest China.* PLoS neglected tropical diseases, 2011. **5**(3): p. e987.

38. McManus, D.P., Gray, D.J., Li, Y., Feng, Z., Williams, G.M., Stewart, D., Rey-Ladino, J., and Ross, A.G., *Schistosomiasis in the People's Republic of China: the era of the Three Gorges Dam.* Clinical microbiology reviews, 2010. **23**(2): p. 442-466.

39. Faust, E.C., *Schistosomiasis in China: Biological and Practical Aspects.* Proceedings of the Royal Society of Medicine, 1924. **17**(Sect Trop Dis Parasitol): p. 31-43.

40. Utzinger, J., Zhou, X.N., Chen, M.G., and Bergquist, R., *Conquering schistosomiasis in China: the long march.* Acta Trop, 2005. **96**(2-3): p. 69-96.

41. Mao, S.P. and Shao, B.R., *Schistosomiasis control in the people's Republic of China.* Am J Trop Med Hyg, 1982. **31**(1): p. 92-9.

42. Zhou, X.N., Wang, L.Y., Chen, M.G., Wu, X.H., Jiang, Q.W., Chen, X.Y., Zheng, J., and Utzinger, J., *The public health significance and control of schistosomiasis in China--then and now.* Acta Trop, 2005. **96**(2-3): p. 97-105.

43. Hu, G.H., Lin, D.D., Zhang, S.J., Zen, X.J., Kang, Z.H., and Cheng, G.J., *The role of health education for schistosomiasis control in heavy endemic area of Poyang Lake region, People's Republic of China.* The Southeast Asian journal of tropical medicine and public health, 2000. **31**(3): p. 467-472.

44. Yin, S., Hu, G.H., and Chen, H.Y., *Progress of research on health education for schistosomiasis control.* Chinese Journal of Schistosomiasis Control, 2009. **21**(4): p. 337-339.

45. Seto, E.Y., Remais, J.V., Carlton, E.J., Wang, S., Liang, S., Brindley, P.J., Qiu, D., Spear, R.C., Wang, L.D., Wang, T.P., Chen, H.G., Dong, X.Q., Wang, L.Y., Hao, Y., Bergquist, R., and Zhou, X.N., *Toward sustainable and comprehensive control of schistosomiasis in China: lessons from Sichuan.* PLoS Negl Trop Dis, 2011. **5**(10): p. e1372.

46. Wang, L.D., Guo, J.G., Wu, X.H., Chen, H.G., Wang, T.P., Zhu, S.P., Zhang, Z.H., Steinmann, P., Yang, G.J., Wang, S.P., Wu, Z.D., Wang, L.Y., Hao, Y., Bergquist, R., Utzinger, J., and Zhou, X.N., *China's new strategy to block Schistosoma japonicum transmission: experiences and impact beyond schistosomiasis.* Trop Med Int Health, 2009. **14**(12): p. 1475-83.

47. Spear, R.C., Seto, E.Y., Carlton, E.J., Liang, S., Remais, J.V., Zhong, B., and Qiu, D., *The challenge of effective surveillance in moving from low transmission to elimination of schistosomiasis in China.* Int J Parasitol, 2011. **41**(12): p. 1243-7.

48. Carlton, E.J., Hubbard, A., Wang, S., and Spear, R.C., *Repeated Schistosoma japonicum infection following treatment in two cohorts: evidence for host susceptibility to helminthiasis?* PLoS Negl Trop Dis, 2013. **7**(3): p. e2098.

Chapter 2   Evaluation of an educational intervention on villagers' knowledge, attitude and behavior: an example of hypothesis testing for mathematical model development

2.1   Background

Health education has been an important means of suppressing transmission of many infectious diseases [1-5]. Despite the various mechanisms of disease transmission, the common goal is to enhance people's awareness of the infection risks, emphasize the significance and methods of disease prevention, increase their adoption of self-protective behaviors and improve compliance with infection testing and treatment. For schistosomiasis, massive education programs have been carried out in endemic areas [6-8]. Not only have these programs improved people's knowledge about the disease, they have also potentially reduced infection risks and enhanced the effectiveness of other control methods, including chemotherapy and environmental modification [9].

In China, health education has been a core element of schistosomiasis control programs for several decades [10-13], although its effects have seldom been evaluated through epidemiological studies until recently. Most papers in the Chinese literature suggest that participants' knowledge of schistosomiasis and attitudes towards disease control were significantly improved as measured by survey results before and after the educational intervention [14-16]. However, most studies failed to adopt randomized controlled trials to account for effects that may not have been due to the intervention, and detailed information on participant selection and enrolment, intervention protocols, and survey designs are rarely reported. Also, few studies considered potential confounders or conducted statistical analyses beyond pre-post mean comparisons. English articles on China's schistosomiasis educational interventions have also been scarce, and, to my knowledge, no original research papers specifically focused on this topic have been published in English journals since 2005. Moreover, most previous studies were conducted in the marshland environment along the lower Yangtze River [17-20], which is quite different from the focus of our group in the irrigated agricultural environment of the hilly and mountainous regions. These differences include the landscape, economy, population demographics, and occupational and recreational activities that bring people in contact with potentially contaminated water sources [21-23].

Here I present the results of a randomized controlled educational intervention trial

targeting adults living in rural areas of two counties in Sichuan province. Surveys were conducted over a 6-month period on multiple endpoints that allowed for comprehensive analyses regarding the effects of the intervention on health-related knowledge, attitudes and behaviors. Accordingly, the hypotheses proposed for testing in this study were that: community-based health interventions improve people's knowledge of schistosomiasis, improve positive attitudes towards infection testing and treatment, reduce transmission-facilitating behaviors including defecating in the field and water exposure, and motivate self-protective behaviors including wearing personal protective equipment (PPE) for reducing dermal contact with potentially contaminated water sources.

2.2 Methods

2.2.1 Study population

The intervention study is part of ongoing efforts to characterize determinants of schistosomiasis re-emergence in Sichuan [24, 25], which thus included surveys to be specified in more details in Chapter 3. However, to ensure the integrity of this chapter, I will still introduce the relevant surveys briefly in this chapter.

This study was conducted in conjunction with a case-control study whose objective was to identify the impact of individual behavior on infection risk. The initial cohort included residents in 53 villages in three counties. A village is defined as the smallest organizational unit in a region, which is also referred to as a production group or natural village. As described in detail elsewhere [24], in 2007, all residents of the 53 study villages, aged 6 years and older, were invited to be tested for *Schistosoma japonicum* infection (Table 2.1). Due to the severe damage caused by the May 12th, 2008 Sichuan earthquake, 17 villages from one county had to be dropped from the study. This resulted in limiting eligible villages to the 28 villages in the two remaining counties where one or more human *S. japonicum* infections were detected 2007. Due to the sensitive nature of conducting infection surveys in regions where schistosomiasis transmission control criteria have officially been met, and to promote candid reporting, the names and exact locations of the counties and study villages have been withheld.

In each study village, all individuals who were infected with *S. japonicum* in 2007 were selected. For each infected individual, four individuals randomly selected from those who tested negative for *S. japonicum* infection in 2007 (or, if more than one in four individuals were infected in a village, all uninfected individuals) were included in this study as well. As health education programs focused on schistosomiasis are commonly implemented at schools in areas where the disease is endemic, people less than 18 years old were excluded. As a result, 706 people in 28 villages were eligible for participation in the intervention study. The 28 selected villages were then stratified by county and randomized by administrative village. An administrative village typically contains several natural villages that are geographically adjacent to each other, so residents of the same administrative village are much more socially connected than with outside residents. Thus, randomization was applied on the administrative village level to reduce contamination bias. Using a random number generator, half of the administrative villages in each county (five in county 1 and three in county 2) were selected as intervention groups (13 total villages and 358 total eligible individuals). Correspondingly, the control group included the same numbers of administrative villages in each county as the intervention group, with a total of 348 eligible individuals in 15 villages.

14

2.2.2　Educational intervention

The educational intervention consisted of two sessions which were administered to the intervention group in April and late June of 2009. All interventions were class-based and led by trained staff from Sichuan Center for Disease Control and Prevention (Sichuan CDC). While eligible participants in the study were notified specifically by village leaders beforehand, all of the residents in the "intervention villages" were welcomed to participate shortly before the class began. The protocols of both intervention sessions were similar. Posters and display boards designed by the Chinese Ministry of Health and Sichuan CDC were put up 15 minutes before the class, and informal tutoring was made available to interested participants. The formal tutoring began with a brief outline of the format and contents of the class, followed by a verbal presentation that elaborated on the transmission, prevention, protection and treatment of schistosomiasis. To improve the effectiveness and attract more attention and participation, an educational video produced by China CDC was played in the first class, and quizzes were conducted in the second class regarding some of the key points and prizes were given to winners (Figure 2.1). In addition, a set of educational materials, including pamphlets, towels, schoolbags and other small items that had schistosomiasis-relevant knowledge printed thereon, were given to each household that had members eligible for the intervention study regardless of their participation (Figure 2.2). Each intervention class lasted for 1 to 1.5 hours based on the number of participants. The control group received the same intervention at the end of the evaluation in 2010.

2.2.3　Assessment of knowledge, attitudes and behaviors

Two series of surveys were used for data collection. The first survey contained sections regarding people's knowledge, attitudes, use of personal protective equipment (PPE) and defecation behavior, which was termed the KAE survey. KAE surveys were conducted three times: once before the intervention in 2008 (further referred to as "baseline"), the first follow-up in early June 2009 (after the first intervention but before the second intervention), and the second follow-up in October, 2009. Survey questions were selected from the "Question Base of Knowledge and Behavior of Schistosomiasis" published by China CDC in 2007 [for the full survey, see appendix] [26]. The knowledge questions tested participants' familiarity with the cause, symptoms, prevention and treatment of the disease. The attitude section inquired about the individual's willingness to be tested for *S. japonicum* and to take the anti-schistosomal drug, praziquantel, with the options ranging from "seeking the service actively" to "refusing the service" and ordered by ascending degree of compliance and initiative. The PPE section asked how frequently in the past 2 weeks the individuals used PPE (i.e. rubber boots or gloves) when contacting surface water during agricultural work. If responders reported no use of any equipment, participants were asked why, and responders claiming that they "did not farm at all" were not included in the analyses of this section. Each section was scored using grading criteria described in detail in *Appendix I-III*. It should be noted that, although these scores were all formalized to the percentage scale, the significance of "one point" varied across the three sections. Due to the different number of questions and types of answers (ordinal in the attitude and behavior sections, binary in the knowledge section), the scores of different sections should not be compared with one another. Finally, participants were also asked whether they had defecated in the fields in the past two

15

weeks. An index to be used for regressions was created such that "1" stands for "frequently or occasionally defecated" and "0" for "never defecated".

The second survey was designed to assess the water contact (WC) magnitude of individuals, which was therefore termed as the WC survey. In 2008, WC surveys were conducted monthly between June and October of 2008 (WC data of May was collected during the June survey as, due to earthquake relief efforts, no interviews were conducted in May). While the surveys were originally designed as a means of investigating the determinants of schistosomiasis re-emergence, a portion of the results reflected people's dermal exposure levels prior to the intervention and thus were used as the baseline data. For comparison, information on WC behaviors was collected in May, July, September and October of 2009. Unlike the KAE surveys which required all participants to respond in person, family members were allowed to report WC patterns of study participants that were absent at the time of the WC survey. A total of 10 categories of activities were used in the surveys: washing clothes, washing agricultural tools, washing hands or feet, playing and swimming, ditch cleaning or diverting, rice planting, rice harvest, fishing, washing vegetables, and drinking or cooking. For each activity, noted as A, participants were asked to report as many as three locations L where this activity occurred, as well as the frequency F and duration D, measured by minutes/event, at each location. Accordingly, the total amount of WC for the ith individual in each month m, measured by the total number of minutes, is estimated by:

$$WC_{i,m} = \sum_A \sum_L F_{i,m,A,L} \cdot D_{i,m,A,L} \qquad \text{(Equation 2.1)}$$

To compare reported WC before and after the intervention, the total WC reported in May, June, September and October each year was calculated. Individuals whose total WC declined from 2008 to 2009 or remained zero both years were classified as improved; individuals whose WC increased from 2008 to 2009 or remained constant and above zero were classified as unimproved.

2.2.4 Statistical analyses

The individuals included in subsequent analyses were defined based on their degree of participation in the various surveys (see timeline Table 2.1). For KAE surveys, only people who participated in all three surveys were included for analyses to yield more meaningful comparisons. On the other hand, WC survey results were utilized to compare each individual's WC magnitude between the years of 2008 and 2009, which thus required that the study population must also have provided complete WC profiles for all four peak months of the infection season (May, July, September and October) in both years.

An intention-to-treat approach was applied to analysis, including all individuals in their assigned treatment group regardless of whether they attended the intervention sessions. Prior to analysis, the distribution of all continuous outcomes was examined to confirm that parametric analysis methods were appropriate.

Although the intervention group assignment was randomized and therefore theoretically independent of any variables that might impact knowledge, attitudes and behaviors, the study population included in the analyses was a subset of the eligible population due to incomplete participation. As age, gender, education, occupation and county were possibly associated with people's participation of the surveys, these

variables were considered as potential confounders in the analyses.

Linear regression was used to evaluate the impact of the intervention on knowledge scores, attitude scores and the use of PPE, comparing the intervention to control group at each follow-up period. Logistic regression was used to evaluate the impact of the intervention on defecation practices and WC. Generalized estimating equations (GEE) were used to account for within-village correlation [27]. Regarding model options, robust regression was applied with individuals clustered by village and exchangeable correlation assumed. For each outcome, both unadjusted and adjusted effect estimates were calculated.

Finally, as an educational intervention study, it is meaningful to identify the factors that were associated with participation in the intervention, i.e. which subset of people were most likely to attend. Using the records of the first intervention (June, 2009), associations between attendance and each demographic variable are listed separately below.

All survey data were input using EpiData© version 3.1 (The EpiData Association, Odense, Denmark) and analyzed with STATA© version 12.1 (StataCorp LP., College Station, TX, USA).

2.2.5 Ethics

The protocols of this study, including all of its relevant sections, were reviewed and approved by the Sichuan Institutional Review Board and the University of California, Berkeley, Committee for the Protection of Human Subjects. All participants provided written informed consent before participating in this study. All of the surveys and interventions were conducted by provincial and/or county CDC staff using the local dialect. Each person who tested positive for *S. japonicum* was provided treatment with 40 mg per kg of praziquantel tablets by the county Anti-Schistosomiasis Control Station.

2.3 Results

2.3.1 Study population summarization

Table 2.2 describes the eligible population and the people who completed the KAE and WC surveys. Farmer was the primary occupation for the majority due to the agriculture-based economy in these rural areas. Nearly two-third of the eligible individuals in both groups were people aged 40 years and above. This rather disproportionate fraction was associated with the nationwide phenomenon in China that younger people from rural areas tend to work in cities and leave children and the elderly population home. Only 29.4% of the eligible individuals completed the KAE surveys; the intervention group had relatively better participation than the control group (37% vs. 22%).There was also a difference in participation by county: the proportion of people who completed the KAE surveys in county 1 was approximately one third of county 2 (16.8% vs. 58.4%). For the WC surveys, because family members were allowed to report WC patterns of study participants that were absent at the time of the survey, the proportion of people enrolled in the analyses was much greater (62.5%).

2.3.2 Baseline result summarization

Table 2.3 presents the results of the baseline KAE and WC surveys conducted in

2008. For the knowledge section, the percent of participants answering each question correctly was similar in intervention and control groups, but the responses reveal gaps in people's knowledge of schistosomiasis: 76% of the people could not name the anti-schistosomal drug, and 53% did not know any symptom of infection. Only half (51%) were aware of the possibility of re-infection after treatment, and the proportion that could list any self-protective method was 58%.

The results of the attitude section were relatively positive: 94% of the people would be willing to either "proactively seek infection testing and treatment" whenever they suspect being infected, or "comply actively when these services are available in the village"; only a few people in each groups would completely refuse the services.

In comparison, the compliance of PPE using was rather poor. More than 85% of all people did not use hand-protection equipment or protective ointments during agricultural work. Even the most-frequently used equipment (rubber shoes or boots), only 40% of people used them on a regular basis. Regarding other behaviors associated with disease transmission, 19% of people had defecated in the fields during the 2 weeks before the survey, and 87% of people reported some level of WC in 2008.

2.3.3 Knowledge

The unadjusted averaged scores from the knowledge section show that the intervention group outscored the control group by only a small margin at baseline but by greater margins in the two follow-ups. This trend was consistent with the regression results, which showed that the point estimates of within-group score difference in the follow-ups were all much higher than those of the baseline. However, none of the estimates of between-group differences were statistically significant, with the only exception being the adjusted result of the first follow-up [95% confidence internal (CI): 1.03, 13.12]. On the other hand, both groups had marked longitudinal improvements. Comparing the unadjusted results between the baseline and the second follow-up, the score increased by 13.3 points in the control group and 17.7 points in the intervention group.

2.3.4 Attitudes

For the attitude section, the averaged unadjusted scores of the two groups were very close in the baseline survey (Table 2.4) and statistically insignificant. Although the score difference by which the intervention group surpassed the control group increased slightly in the follow-up surveys, this pattern was not supported by the regression results where the minimum estimates of both unadjusted and adjusted models were found at the first follow-up. Further, all estimates of the between-group differences were relatively small and likely to be due to chance, so there was little evidence provided by intergroup comparisons that indicate effects of intervention on attitudes. However, similar to the knowledge section, both groups' scores improved from the baseline to a noticeable extent, and as expected, greater improvement was found in the intervention group.

2.3.5 PPE wearing behavior

Little difference was found between the two groups in the baseline survey due to both the unadjusted score comparisons and model regressions. For the first follow-up, the between-group difference was even smaller. While the intervention group had 5.5 more averaged points, or 4.4 points after adjusted for confounders, than the control group in the

second follow-up, this finding was again insignificant. Unlike the previous two endpoints, there was no sign that people from either group had any improvement in self-reported PPE wearing behavior.

2.3.6 Defecation behavior

In the first two surveys, the proportion of people who defecated in the fields was slightly smaller in the intervention group. While this finding was consistent with what were suggested by the odds ratios (ORs) from model regressions (OR>1), none of these estimates was statistically significant. In the last survey, the proportion in the intervention group was less than half of the control group, and much higher ORs were detected with lower p-values (OR=3.28 and p=0.185 for the adjusted estimate). Also, consistent "improvement" was found for this behavior within groups as well, meaning that the proportions that defecated decreased with time in both groups. Compared to the control group that had a 63% drop from baseline to the final survey, this proportion was reduced by 81% in the intervention group.

2.3.7 WC behavior

Table 2.5 shows the results of the logistic regressions regarding WC change from 2008 to 2009. Both ORs, with and without confounders adjusted, were marginally larger than unity, meaning that those in the intervention group were slightly more likely to reduce their level of WC compared to the control group. Again however, the p-values indicated that none of the findings were statistically significant. Thus, no evidence was found in the dataset which suggested an effect of education in reducing people's WC behavior.

2.3.8 Longitudinal effects

As mentioned above, while little difference was detected between groups, some clear longitudinal effects were found in both groups in three endpoints: knowledge, attitudes and defecation behavior. Regressions on the individual level confirmed these trends (Table 2.6). For the control group, statistically significant improvements were found on the knowledge endpoint in the first follow-up survey, and on all three endpoints in the second follow-up with greater margins. Longitudinal improvements in the intervention group were even more striking – stronger effects were found for all three endpoints in both follow-ups, and most results were highly unlikely to be due to chance.

2.3.9 Intervention participation

Table 2.7 shows the characteristics of the individuals who participated in the intervention in April 2009. As expected and noted earlier, participants were not evenly distributed by the various demographic variables. Specifically, elderly people, females, farmers, people without post-elementary education and people from county 2 were more likely to participate than their counterparts in other subgroups. However, the infection status of people in the previous two years (2007, 2008) was not related to participation. As all individuals diagnosed as infected were notified, knowledge of one's infection apparently did not prompt the individual to participate in the education classes.

2.4 Discussion

The results of this intervention trial showed no evidence of an improvement

between the intervention and control groups in four endpoints – attitude and three types of behavior – defecation, PPE wearing and WC. Although, in the first follow-up survey, the intervention group had a significantly higher score in the knowledge section, this pattern was not sustained in the second follow-up. In sum, no clear evidence was found suggesting any effect of the intervention through intergroup comparisons.

On the other hand, both groups showed quite significant longitudinal improvement in knowledge, attitude and defecation behavior, although somewhat greater improvements were seen in the intervention group as might be expected. First, it is possible that some people in the control group may have become more familiar with the subjects covered in the questionnaire by participating in repeated surveys. However, it is speculated that this result is a more general example of the theory of diffusion of innovations (DOI), first introduced by Rodgers over 40 years ago [28]. DOI theory concerns the spread of ideas through social networks. In addition to introducing new ideas into the social network through the intervention itself, many other activities, such as snail surveys, molluscicide treatment and the development of rural infrastructure were carried out during the same period, and in the same general region, as part of disease control programs. While contamination bias was specifically considered during group formation, there were no restrictions preventing people in different groups from communicating or discussing the contents of the educational intervention during or subsequent to any element of the trial. Hence, there was opportunity for the diffusion of the information among farmers and other village residents in the region.

While the self-reported defecation behavior appears to have been affected, WC and PPE wearing behaviors were not, perhaps because they are difficult to alter in any sustainable way in this environment. In the hilly and mountainous rural areas of Sichuan, agricultural machinery is less available and useful, and farming is mostly accomplished by humans and bovines. As only adults are focused and, as earlier stated, >80% of the study population were farmers, their work-associated WC is very difficult to avoid. In this context, self-protection during agricultural work plays an important role of reducing the risk of infection, at least in principal. However, PPE wearing behaviors were not improved by the intervention, and both groups had similarly low scores across all three surveys. The lowest scores were found in the first follow-up (June 2009), suggesting that the least protection was used during the spring planting season when most WC occurred. When asked about the reason for not using any protection, most people responded that "however useful in interrupting transmission, it was uncomfortable and inconvenient to wear rubber gloves or boots while working in the fields". While this illustrates the difference between understanding the benefit and modifying behavior, as suggested in other studies of occupationally-related behavior change [19], it also raises the issue of the practicalities of sustained personal protection in an inherently risky environment. As having been argued elsewhere, a focus on environmental improvements and monitoring systems to signal early risks of infection are better long-term solutions than personal protection in rural China as they are in occupational settings worldwide [25].

An important objective of this study is to facilitate improved design of education and control programs. For instance, it was shown in all three KAE surveys that a large proportion of people were not aware of the re-infection possibility or familiar with disease symptoms, both of which may lower their chance of receiving timely diagnostic testing and treatment. These are examples of the points that should be emphasized in

future educational programs. Also, the analysis suggested there to be "vulnerable" members of the community that should be specifically targeted in the future. For example, it was found that the elderly people and women appeared to be less knowledgeable about schistosomiasis compared to those in other subgroups both before and after the intervention. However, these same groups were well-represented in the intervention trial suggesting the need for alternative educational strategies.

The advantages of this study over earlier work included a combination of various educational formats and multiple media to engage participants' attention. Second, longitudinal surveys were carried out on multiple endpoints that could all be impacted by the intervention. Third, the finding that improvements over time were also observed in the control group clearly demonstrated the value of a randomized controlled trial in understanding the effects of intervention.

There are some limitations of this study. First, participation in the surveys and interventions was markedly poorer in county 1 than in county 2.This might be due to the fact that a greater fraction of the study population from county 1 worked in urban areas and because these emigrant workers come back only occasionally, their participation was inherently more difficult. Second, as mentioned above, contamination bias was not completely preventable, as social connections were fairly common and frequent in the rural areas. Finally, considering the inherent challenges of knowledge assimilation and behavior change, it is possible that a longer study period would have been more effective.

Although comprehensive control strategies tailored to the current re-emerging conditions remain under study, health education is still widely considered as a cost-effective method of lowering the individual risk of infection [9, 29]. However, unlike what is suggested in many other studies that also enrolled a control group [16], the effects of the education intervention, evaluated through knowledge, attitudes and behaviors, were not strong enough to yield statistically significant differences between the intervention and control groups. Also, occupation-related behavior change was shown to be extremely difficult. These findings, while not denying the potentially important role of health education in future schistosomiasis control, does suggest the necessity of exploring new ways of conducting education programs that work more effectively in the current low transmission environments.

Other than its implication for disease control, this study, as stated earlier, also presents an example of individual data collection and its potential of being used for hypothesis testing which can be incorporated into mathematical models. First of all, individual water exposure data is a core component of almost all models which mimic the dynamics of schistosomiasis transmission, and I will specify its application in the individually-based model in Chapter 4. Second, the hypotheses, while not supported by the data in this specific study, can be used for scenario setting in model simulations, particularly for exploring the effects of control strategies. For instance, if health education were shown to be effective in reducing water contact, this finding could be easily reflected in simulations by multiplying a coefficient less than unity to the water contact magnitude, which would otherwise be set constant or fluctuate within a small range, in the simulation years post intervention. Since this reduction is likely to be activity-specific, i.e., washing vegetables in the ditch is potentially more "adjustable" than fishing given that the latter is occupation-related exposure, another coefficient, the (mean) body-surface area averaged across all activities, can be also modified in response

to the statistical findings. Similarly, the use of PPE can be recognized as an equivalent way of water contact reduction and handled using the same method. On the other hand, individuals' attitudes towards infection testing are directly related with the efficiency of diagnosis, which is an important simulation endpoint to suggest on disease surveillance. Finally, in multi-year simulations, residual infections following year-end chemotherapy depend on people's attitudes of complying, which is another factor subject to modification should the intervention have a positive impact. To sum up, the design and results of hypothesis testing through statistical methods can substantially impact the structure and simulation results of mathematical models. But most crucially, no quantitative analyses have been possible without the collection of high-quality data, which will be presented in more detail in the next chapter.

Table 2.1. Timeline of the health education intervention and evaluation surveys.

	2007		2008					2009				
	Jun	Nov – Dec	Jun	Jul	Sep	Oct	Nov	Apr	Jun	Jul	Sep	Oct
Demographic survey *	X											
Infection survey †		X					X					
Education intervention Ŧ								X	X			
KAE survey ‡							X		X			X
Water contact survey **			X	X	X	X			X	X	X	X

* Collected demographic information including age, sex, occupation, education, and village of residency;
† Tested for *S. japonicum* infection status; see detailed methods of infection testing in (Carlton et al., 2011);
Ŧ The intervention survey consisted of class-based education interventions to the intervention group;
‡ Included three sections: knowledge of schistosomiasis, attitude towards infection testing and treatment, defecation practices and self-protective behaviors of wearing personal protective equipment;
**Asked about WC behaviors in the past 2 weeks. The June, 2008 survey asked about WC behaviors in both May and June; the October, 2009 survey asked about WC behaviors in both September and October.

Figure 2.1 Educator from Sichuan CDC conducting the second intervention class in uses of educational posters (on the back) and in-class quizzes (prizes on the table) in two villages

Figure 2.2 Educational materials distributed to families with members eligible for the educational intervention

Table 2.2. Characteristics of the eligible population and study participants in an evaluation of a health education intervention in Sichuan, China.

	Eligible population		Complete participation in the KAE surveys *		Complete participation in the WC surveys **	
	Control	Intervention	Control†	Intervention†	Control†	Intervention†
	No.	No.	%	%	%	%
Total	348	358	22	37	61	64
Age (years)						
18-29	23	25	0	8	43	52
30-39	75	61	15	23	57	44
40-49	100	83	22	34	63	69
≥50	150	189	29	46	65	70
Sex						
Female	177	187	28	45	60	64
Male	171	171	16	28	63	64
Education						
Elementary school or no schooling	211	230	23	41	63	73
At least some middle school or higher	111	89	18	26	73	66
Occupation						
Not Farmer	44	65	5	20	61	65
Farmer	274	253	24	42	67	74
County						
1	284	203	13	22	59	66
2	64	155	64	56	73	62

*: Included three sections: knowledge of schistosomiasis, attitude towards infection testing and treatment, and self-protective behaviors of wearing personal protective equipment

†: Percentages are calculated by dividing the number of enrolled people in each cell by the number of eligible population in the corresponding cell.

**: Asked about WC behaviors in the past 2 month, including the location, frequency and duration associated with each of the 10 WC categories

Table 2.3. Schistosomiasis knowledge, attitudes towards testing and treatment and behaviors reported at baseline.

		Control group (N = 77)	Intervention group (N = 131)
		%	%
Knowledge			
Know that the cause of infection is contacting contaminated water		68.8	67.2
Know that schistosomiasis will lead to enlarged abdomen		67.5	67.9
Know the shape of the intermediate snail		79.2	77.1
Know the months during which most of the transmission occurs		64.9	77.9
Know the drug for schistosomiasis treatment		23.4	23.7
Know that the faeces of infected animals can transmit the disease		70.1	60.3
Know that people can be re-infected after treated		55.8	48.9
Know how many symptoms of schistosomiasis infection	0	61.0	48.1
	1	20.8	37.4
	2	15.6	8.4
	≥3	2.6	6.1
Know how many self-protective methods	0	42.9	41.9
	1	45.5	52.7
	≥2	11.7	5.4
Attitudes			
Infection testing			
I will take the test proactively whenever I suspect that I am infected		19.5	14.5
I will comply actively when infection testing is provided in the village		71.4	82.5
I will take the test only when I am asked to		3.9	1.5
I refuse to take the test		5.2	1.5
Infection treatment			
I will take the treatment proactively if I am infected		22.1	13.7
I will comply actively when treatment is provided in the village		70.1	85.5
I will take the treatment only when I am asked to		5.2	0
I refuse to take the treatment		2.6	0.8
Behaviors			
Personal protective equipment (PPE) used while working in the fields			
Wear rubber shoes or boots	Frequently	39.0	42.8
	Sometimes	54.5	45.0
	Never	6.5	12.2
Wear rubber gloves or use handled tools	Frequently	2.6	0.7
	Sometimes	9.1	15.3
	Never	88.3	84.0
Use protective ointment	Frequently	0	0
	Sometimes	7.8	13.0
	Never	92.2	87.0
Defecated in the fields in the past 2 weeks		20.8	18.3
WC*			
Reporting any WC (sum of June, July, September and October of 2008)		85.5	88.6

*: 214 individuals in the control group and 229 individuals in the intervention group

Table 2.4. Comparisons of schistosomiasis knowledge, attitudes towards treatment, PPE wearing and defecation practices in the intervention and control groups.

	Averaged unadjusted results*		Unadjusted group difference (95% CI)†	Adjusted group difference (95% CI)†‡
	Control (N = 77)	Intervention (N =131)		
Knowledge score (out of 100)				
Baseline	37.3	37.5	1.91 (-6.64, 10.46)	2.31 (-5.17, 9.79)
First follow-up (06/2009)	42.1	46.8	6.09 (-1.11, 12.55)	7.07 (1.03, 13.12)
Second follow-up (12/2009)	50.6	55.2	5.23 (-2.60, 13.06)	4.91 (-4.06, 13.87)
Attitude score (out of 100)				
Baseline	81.2	82.6	3.21 (-1.53, 9.74)	2.26 (-2.68, 7.19)
First follow-up (06/2009)	82.5	84.5	1.87 (-3.32, 7.06)	1.55 (-4.20, 7.29)
Second follow-up (12/2009)	87.3	91.0	2.56 (-2.28, 7.40)	3.65 (-1.75, 9.05)
PPE wearing behavior score (out of 100)				
Baseline	28.8	30.2	2.20 (-6.12, 10.53)	1.74 (-6.12, 9.60)
First follow-up (06/2009)	14.1	14.1	1.62 (-4.24, 7.48)	1.78 (-4.18, 7.74)
Second follow-up (12/2009)	24.9	30.4	5.05 (-2.26, 12.38)	4.42 (-3.38, 12.22)
Defecation behavior (binary)				
Baseline	18.7	17.0	1.09 (0.27, 4.43)	1.54 (0.45, 5.24)
First follow-up (06/2009)	10.3	8.1	1.19 (0.27, 5.21)	1.21 (0.32, 4.53)
Second follow-up (12/2009)	7.0	2.4	2.53 (0.44, 14.39)	3.28 (0.57, 18.98)

*: For the knowledge, attitude and PPE wearing sections, cells represent the averaged unadjusted scores; for the "defecation" question, cells represent the proportions of people who **DID** defecate in the fields.
†: For the knowledge, attitude and PPE wearing sections, cells represent score differences based on regression results; positive numbers correspond to higher scores in the intervention group. For the "defecation" question, cells represent the odds ratios (intervention group/control group) of **NOT** defecating in the fields.
‡: Adjusted for age, gender, education, occupation and county.

Table 2.5. Unadjusted and adjusted odds ratios between the intervention and control groups regarding WC change from 2008 to 2009.

Outcome variable	Unadjusted odds ratio* (95% CI)	Adjusted odds ratio (95% CI) *†
WCC_i	1.29 (0.59, 2.83)	1.15 (0.55, 2.43)

*: The odds ratio describing the odds of an improvement in WC in the intervention group vs. the control group. Individuals whose total WC (the sum of May, June, September and October) declined from 2008 to 2009 or remained zero both years were classified as improved. Individuals whose WC increased from 2008 to 2009 or remained constant and above zero were classified as unimproved.
†: Adjusted for age, gender, education, occupation and county.

Table 2.6. Longitudinal changes on knowledge and attitudes scores and the defecation behavior in each study group.

	Control group (95% CI)		Intervention group (95% CI)	
	First follow-up (06/2009)	Second follow-up (12/2009)	First follow-up (06/2009)	Second follow-up (12/2009)
Knowledge*	4.76 (0.19, 9.33)	13.25 (8.15, 18.34)	9.36 (5.47, 13.25)	17.71 (13.99, 21.43)
Attitudes*	1.24 (-3.04, 5.51)	6.09 (2.12, 10.05)	1.95 (-0.10, 3.99)	8.45 (6.09, 10.81)
Defecation Behavior †	2.00 (0.75, 5.31)	3.02 (1.04, 8.77)	2.33 (1.13, 4.79)	8.38 (2.47, 28.45)

GEE models with robust estimations were applied. Options were made to cluster individuals and assume exchangeable correlations. All cells are relative to the baseline KAE survey (2008).
*: Cells represent score differences
‡: Cells represent odds ratios

Table 2.7. Associations between demographic characteristics and participation of the first intervention.

	Eligible (No.)	Participated (No.)	Participation (%)	p-value
Total	407	276	(68)	
Age (years)				
<18	50	28	(58)	
18-29	24	11	(42)	
30-39	61	35	(57)	0.001*
40-49	83	57	(69)	
≥50	189	145	(77)	
Sex				
Female	212	155	(73)	0.017
Male	195	121	(62)	
Education				
Elementary school or no schooling	257	183	(71)	0.019
At least some middle school or higher	98	57	(58)	
Occupation				
Not Farmer	73	38	(52)	<0.001
Farmer	254	189	(74)	
County				
County 1	220	118	(54)	<0.001
County 2	187	158	(84)	
Previous infection†				
Yes	240	170	(71)	0.809
No	104	75	(72)	

*: Test for trend; age categories are treated as ordinal
†: Based on the results of the 2007 and 2008 infection tests; "Yes" means infection in either test

References

1. Aiello, A.E., Coulborn, R.M., Perez, V., and Larson, E.L., *Effect of hand hygiene on infectious disease risk in the community setting: a meta-analysis.* American Journal of Public Health, 2008. **98**(8): p. 1372-1381.
2. Amoran, O.E., Fatugase, K.O., Fatugase, O.M., and Alausa, K.O., *Impact of health education intervention on insecticide treated nets uptake among nursing mothers in rural communities in Nigeria.* BMC research notes, 2012. **5**: p. 444-0500-5-444.
3. Gao, X., Wu, Y., Zhang, Y., Zhang, N., Tang, J., Qiu, J., Lin, X., and Du, Y., *Effectiveness of school-based education on HIV/AIDS knowledge, attitude, and behavior among secondary school students in Wuhan, China.* PloS one, 2012. **7**(9): p. e44881.
4. Ghosh, S.K., Patil, R.R., Tiwari, S., and Dash, A.P., *A community-based health education programme for bio-environmental control of malaria through folk theatre (Kalajatha) in rural India.* Malaria journal, 2006. **5**: p. 123.
5. Manabe, T., Pham, T.P., Vu, V.C., Takasaki, J., Dinh, T.T., Nguyen, T.M., Shimbo, T., Bui, T.T., Izumi, S., Tran, T.H., Ngo, Q.C., and Kudo, K., *Impact of educational intervention concerning awareness and behaviors relating to avian influenza (H5N1) in a high-risk population in Vietnam.* PloS one, 2011. **6**(8): p. e23711.
6. Lansdown, R., Ledward, A., Hall, A., Issae, W., Yona, E., Matulu, J., Mweta, M., Kihamia, C., Nyandindi, U., and Bundy, D., *Schistosomiasis, helminth infection and health education in Tanzania: achieving behaviour change in primary schools.* Health education research, 2002. **17**(4): p. 425-433.
7. Asaolu, S.O. and Ofoezie, I.E., *The role of health education and sanitation in the control of helminth infections.* Acta Tropica, 2003. **86**(2-3): p. 283-294.
8. Kloos, H., *Human behavior, health education and schistosomiasis control: a review.* Soc Sci Med, 1995. **40**(11): p. 1497-1511.
9. Rollinson, D., Knopp, S., Levitz, S., Stothard, J.R., Tchuente, L.A., Garba, A., Mohammed, K.A., Schur, N., Person, B., Colley, D.G., and Utzinger, J., *Time to set the agenda for schistosomiasis elimination.* Acta Trop, 2012.
10. Sleigh, A., Jackson, S., Li, X., and Huang, K., *Eradication of schistosomiasis in Guangxi, China. Part 2: Political economy, management strategy and costs, 1953-92.* Bull World Health Organ, 1998. **76**(5): p. 497-508.
11. Mao, S.P. and Shao, B.R., *Schistosomiasis control in the people's Republic of China.* Am J Trop Med Hyg, 1982. **31**(1): p. 92-9.
12. Wang, L.D., Chen, H.G., Guo, J.G., Zeng, X.J., Hong, X.L., Xiong, J.J., Wu, X.H., Wang, X.H., Wang, L.Y., Xia, G., Hao, Y., Chin, D.P., and Zhou, X.N., *A strategy to control transmission of Schistosoma japonicum in China.* The New England journal of medicine, 2009. **360**(2): p. 121-128.
13. Hong, Q.B., Yang, K., Huang, Y.X., Sun, L.P., Yang, G.J., Gao, Y., Gao, Y., Zhang, L.H., Zhou, M., Steinmann, P., and Liang, Y.S., *Effectiveness of a comprehensive schistosomiasis japonica control program in Jiangsu province, China, from 2005 to 2008.* Acta Trop, 2011. **120 Suppl 1**: p. S151-7.
14. Yin, S., Hu, G.H., and Chen, H.Y., *Progress of research on health education for schistosomiasis control.* Chinese Journal of Schistosomiasis Control, 2009. **21**(4): p. 337-339.

15. Chen, H.Y., Hu, G.H., Song, K.Y., Xiong, Z.W., Hu, J., Yang, P.Y., Peng, G.H., Hu, W.C., Yu, S.S., Fu, G.L., Liu, Z.H., Qi, J.C., Ge, J., and Wan, B.P., *[Establishment and application of school-based health promotion and intervention model of schistosomiasis in lake-type endemic area].* Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi, 2009. **27**(2): p. 125-9, 134.

16. Zhou, L.Y., Deng, Y., Steinmann, P., and Yang, K., *The effects of health education on schistosomiasis japonica prevalence and relevant knowledge in the People's Republic of China: A systematic review and meta-analysis.* Parasitol Int, 2012. **62**(2): p. 150-156.

17. Hu, G.H., Dandan, L., Shaoji, Z., Xiaojun, Z., Zenghua, K., and Guojun, C., *The role of health education for schistosomiasis control in heavy endemic area of Poyang Lake region, People's Republic of China.* The Southeast Asian journal of tropical medicine and public health, 2000. **31**(3): p. 467-472.

18. Yuan, L.P., Manderson, L., Ren, M.Y., Li, G.P., Yu, D.B., and Fang, J.C., *School-based interventions to enhance knowledge and improve case management of schistosomiasis: a case study from Hunan, China.* Acta Tropica, 2005. **96**(2-3): p. 248-254.

19. Hu, G.H., Hu, J., Song, K.Y., Lin, D.D., Zhang, J., Cao, C.L., Xu, J., Li, D., and Jiang, W.S., *The role of health education and health promotion in the control of schistosomiasis: experiences from a 12-year intervention study in the Poyang Lake area.* Acta Trop, 2005. **96**(2-3): p. 232-41.

20. Yuan, L., Manderson, L., Tempongko, M.S., Wei, W., and Aiguo, P., *The impact of educational videotapes on water contact behaviour of primary school students in the Dongting Lakes region, China.* Trop Med Int Health, 2000. **5**(8): p. 538-44.

21. Zhou, X.N., Guo, J.G., Wu, X.H., Jiang, Q.W., Zheng, J., Dang, H., Wang, X.H., Xu, J., Zhu, H.Q., Wu, G.L., Li, Y.S., Xu, X.J., Chen, H.G., Wang, T.P., Zhu, Y.C., Qiu, D.C., Dong, X.Q., Zhao, G.M., Zhang, S.J., Zhao, N.Q., Xia, G., Wang, L.Y., Zhang, S.Q., Lin, D.D., Chen, M.G., and Hao, Y., *Epidemiology of schistosomiasis in the People's Republic of China, 2004.* Emerging infectious diseases, 2007. **13**(10): p. 1470-1476.

22. Spear, R.C., Seto, E., Liang, S., Birkner, M., Hubbard, A., Qiu, D., Yang, C., Zhong, B., Xu, F., Gu, X., and Davis, G.M., *Factors influencing the transmission of Schistosoma japonicum in the mountains of Sichuan Province of China.* The American Journal of Tropical Medicine and Hygiene, 2004. **70**(1): p. 48-56.

23. Liang, S., Seto, E.Y., Remais, J.V., Zhong, B., Yang, C., Hubbard, A., Davis, G.M., Gu, X., Qiu, D., and Spear, R.C., *Environmental effects on parasitic disease transmission exemplified by schistosomiasis in western China.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(17): p. 7110-7115.

24. Carlton, E.J., Bates, M.N., Zhong, B., Seto, E.Y., and Spear, R.C., *Evaluation of mammalian and intermediate host surveillance methods for detecting schistosomiasis reemergence in southwest China.* PLoS neglected tropical diseases, 2011. **5**(3): p. e987.

25. Spear, R.C., Seto, E.Y., Carlton, E.J., Liang, S., Remais, J.V., Zhong, B., and Qiu, D., *The challenge of effective surveillance in moving from low transmission to elimination of schistosomiasis in China.* Int J Parasitol, 2011. **41**(12): p. 1243-7.

26. Chen, L., Cao, C.L., Wan, X.X., Liu, Q., Cai, X.P., Qiu, D.Q., and Guo, J.G., *Reliability and Validity of Question Base of Knowledge and Behavior of Schistosoiasis Control.* Parasitoses and Infectious Diseases, 2008. **6**(3): p. 3.

27. Liang, K.Y. and Zeger, S.L., *Longitudinal data analysis using generalized linear models.* Biometrika, 1986. **73**(1): p. 13-22.

28. Rogers, E.M., *Diffusion of Innovations*. 5th ed. 2003: Free Press.

29. Liu, R., Dong, H.F., and Jiang, M.S., *What is the role of health education in the integrated strategy to control transmission of Schistosoma japonicum in China?* Parasitol Res, 2012. **110**(5): p. 2081-2.

Appendix to Chapter 2

The knowledge section questions

There are three types of questions in the knowledge section. For the multiple choice and true/false questions shown in Table A2-1, only one answer is correct for each question, full grade is given only when the right option is selected.

Table A2-1. Multiple-choice and true/false questions in the knowledge section

Multiple-choice questions (options omitted)	*What is the cause of schistosomiasis?*
	Untreated schistosomiasis will cause what type of body change?
	What is the shape of the intermediate snail (Oncomelania hupensis)?
	Most of the transmission occurs in which months during a year?
	What is the drug for schistosomiasis treatment used in the local areas?
True/False questions	*Will the faeces of infected animals be able to transmit the disease?*
	After treatment, will people be infected again if they contact contaminated water?

Further, responders are required to list the symptoms and protective methods of schistosomiasis. The pre-set options, shown in Table A2-2, will be checked if the corresponding symptoms are mentioned by the responder. Extra grades will be assigned if the responder mentions other symptoms that are correct but not listed in the options.

Table A2-2. Self-report questions in the knowledge section

Self-report questions	**Pre-set options to be checked if mentioned**
What types of symptoms might be associated with schistosomiasis infection?	Fever
	Diarrhoea
	Bloody faeces
	Fatigue
	Cough
Do you know any methods that are preventive of getting infected while working in the fields?	Wear boots or rubber shoes
	Wear rubber gloves or tools with handles
	Use protective ointment

All questions in the knowledge section are weighted equally with a total of 100 points. Since there are 15 questions/options in total, each question is worth 6.6 points.

The attitude section questions

There are two questions for grading and two follow-up questions in the attitude section. The questions as well as the grading criteria are shown in Table A2-3:

Table A2-3. Questions and grading criteria of the attitude section

Question for grading	Grading criteria	Follow-up question (not for grading)

1. Regarding infection tests in the future, you will:
  1) Take the test proactively whenever suspecting being infected
  2) Comply actively whenever the test is conducted in the village by schistosomiasis control organizations
  3) Take the test only when being asked to
  4) Refuse to take the test
	1) => 5 points
	
	2) => 4 points
	
	3) => 2 points
	
	4) => 0 point
		If (4), then ask:
		Why will you refuse to take the test?
		  (1) The blood test hurts
		  (2) It is troublesome to collect the stools
		  (3) It is unnecessary as I don't feel uncomfortable
		  (4) It is unnecessary as I must be infected
		  (5) Others, please specify_____

2. Regarding treatment in the future, you will:
  1) Take the treatment proactively if infected
  2) Comply actively when treatment or mass chemotherapy is provided in the village by schistosomiasis control organizations
  3) Take the treatment or mass chemotherapy only when being asked to
  4) Not take the treatment due to other diseases
  5) Refuse to take the treatment
	1) => 5 points
	
	2) => 4 points
	
	3) => 2 points
	
	4) => "N/A"
	5) => 0 point
		If (5), then ask:
		Why will you refuse to take the test?
		  (1) The side effects are uncomfortable
		  (2) Unwilling to comply with mass chemotherapy if found uninfected
		  (3) Unwilling to take the drugs until feeling sick
		  (4) Others, please specify_____

Not applicable is assigned to option (4) of the second question because the special situation that it indicates does not represent the responder's attitude towards treatment. However, this option was not selected in any of the three surveys.

The PPE section questions

The five questions and grading criteria of the behavior section are shown in Table A2-4.

Table A2-4. Questions and grading criteria of the PPE wearing behavior section

Question for grading		Grading criteria	Follow-up question (not for grading)
In the past month, when working in the snail-residing water environments, did you…	(1) wear rubber shoes?	1) Frequently 2) Occasionally 3) Never	1) => 5 points 2) => 3 points 3) => 0 point	*If answering "never" to all questions, then ask:* Why do you refuse to use protection? (1) Don't think these methods are effective (2) However useful, these methods are inconvenient for working in the fields (3) Wearing boots and rubber shoes cannot protect hands so I will get infected anyway (4) Others, please specify_____
	(2) wear long rain boots?			
	(3) wear rubber gloves?			
	(4) use protective ointment?			

As methods (1) and (2) can practically be substitutes with one another, the higher point from these two questions was taken when calculating the total score.

Chapter 3 Data collection for the individually-based model

"Data! Data! Data!" he cried impatiently. "I can't make bricks without clay!"
— Sherlock Holmes, the Adventure in the Copper Beeches

In this chapter, I present the protocols and descriptive analyses of data collected from two cohorts residing in an endemic equilibrium (Cohort 1) and the re-emerging environments (Cohort 2), respectively. Some of the data used in this chapter can also be found in previous publications, although much more is available regarding Cohort 1 than Cohort 2 [1-9]. However, unlike the previous applications, the surveyed data will be used for yielding point estimates and/or distributions for key parameters in the individually-based model. Consequently, it is helpful to summarize the relevant surveys before moving forward to the model development and simulation sections.

3.1 Surveys in the endemic equilibrium environment (Cohort 1)

The study population of Cohort 1 was from 20 villages of four townships in Xichang County. To provide a brief timeline of the surveys, the study was initiated in 2000, when demographic and infection surveys were conducted and covered all residents in the twenty villages. Meanwhile, a subset of the residents was invited to participate in a retrospective survey on monthly water exposure. From May to October in 2001, five villages were selected for monthly bioassay surveys to estimate cercarial distributions in surface waters. Finally, a second round of infection survey was conducted in October, 2002 among people from 10 selected villages. I will give details of each survey in the next sections.

3.1.1 The environment

Xichang County is located in the Anning River Valley in southwestern Sichuan. Having a subtropical climate in the hilly and mountainous environments, its temperature varies only modestly within a year, from 9.6℃ in January and 22℃ in the summer. The annual precipitation in Xichang is around 1000mm, with more than 90% occurring between May and October, which is therefore referred to as the "wet season" with relatively less sunshine time as opposed to the dry season representing the remaining months [10].

Residents in this region are mainly farmers. Unlike Cohort 2, however, farmers in Xichang rarely raise livestock such as water buffalos or cows for agricultural purposes, so the main definitive host of transmission is humans. Due to the climatic conditions, there are two growing seasons. The spring season starts from March and ends in September. The spring crops include rice, corn, tobacco and cabbage. Other crops, mainly garlic, wheat and beans, are planted in the fall planting season between October and the following February [11]. Corresponding to the specific requirements for water and fertilization of the various crops, human water contact behaviors can be quite different during a year.

3.1.2 Demographic survey

Demographic surveys were conducted in November, 2000 in four townships of Xichang County: Chuanxing, Daxing, Gaojian and Hainan. The survey was targeted to all residents aged between 4 and 60, and a total of 3897 people from 20 villages actually participated. Information was collected regarding each individual's age, gender, occupation and socioeconomic status.

3.1.3 Infection survey

Nested with the demographic survey, an infection survey was conducted among the same 3897 people in November and December of 2000. Stool samples from as many as three different days were collected for each individual. The miracidial hatch test and Kato-Katz thick smear test were used to identify infection status and intensity. The detailed protocols of the two methods are as follows.

Miracidial hatch test
The miracidial hatch test was conducted according to the Chinese Ministry of Health Protocols, which includes the following steps [12]:
    (1) Pre-treat the water used for rinsing stool samples to ensure that the free chlorine level is less than $1ppm$ and the PH level is between 6.5 and 7.5;
    (2) Suspend 30 grams of stool in water. Strain the sample with copper mesh to remove large particles first, then with nylon mesh and retain the portion of the sample that does not pass through the mesh;
    (3) Suspend the sample in an Erlenmeyer flask or hatch test container ;
    (4) Incubate the samples in a room where the temperatures can be maintained between 28 and 30 degrees C;
    (5) Examine the sample for the presence of miracidia 2, 4, 6 and 8 hours after step (4) is completed. Each time, examine the sample for at least 2 minutes.
If miracidia were observed, a "positive" result was recorded for the sample.

Kato-Katz test
One stool sample per person should be examined using the Kato Katz method. Three slides should be prepared per sample, and the test was conducted according to the WHO protocols, which include the following steps:
    (1) Soak cellophane strips in 50% (or greater) glycerol solution with malachite-green or methylene blue stain;
    (2) Place a small amount of stool on a piece of scrap paper. Place mesh (60-105 mesh) over the sample and force the stool through the screen;

(3) Place the 41.7 mg Kato-Katz template on a microscope slide. Transfer the screen stool to the template. Fill the template hole, level with an applicator stick and carefully remove the template;

(4) Cover the sample with a cellophane strip that has soaked in the glycerol solution for at least 24 hours. Wipe off any excess glycerol;

(5) Invert the slide and press against a smooth surface to spread the sample and gently slide the microscope sideways. Store the slide out of direct sunlight for at least 24 hours before reading;

(6) Examine each slide using a dissecting microscope for the presence of *S. japonicum* eggs. If present, another researcher should also examine the slide to confirm the presence and number of *S. japonicum* eggs.

The number of *S. japonicum* eggs on each slide was recorded. The sum of the three slides, after multiplied by 8.032, yielded the EPG (eggs per gram of stool) of each individual. For quality control, 5% of the slides should be re-examined each day by an experienced laboratory.

In addition to the 2000 survey, a second infection survey was conducted in 2002 to the same population, except that one instead of three stool samples was collected for each individual. Each person who tested positive for *S. japonicum* in the 2000 or 2002 infection surveys was provided treatment with 40 mg per kg of praziquantel tablets by the county Anti-Schistosomiasis Control Station.

3.1.4 Water contact survey

Human water contact (WC) surveys were conducted in October, 2000 to a subset of the study population. Stratified by occupation, 25% of the people in each village were invited to participate. A retrospective questionnaire was used to collect individual WC information from April to October. Eight categories of activities were used: washing clothes and vegetables, washing agricultural tools, washing hands and feet, playing and swimming, ditch diverting and cleaning, plough and rice planting, rice cutting, and fishing. For each activity, individuals were asked to report the locations where this activity occurred, as well as the frequency (times/month) and duration (minutes/event), at each location.

3.1.5 Mouse bioassay

A mouse bioassay was used to yield cercariae distributions in the surface water. Also accordingly to MOH protocol, compared to alternative methods such as infected snail survey, the results from bioassays were less sensitive to water velocity and turbidity and associated with human risks in a more direct way. In 2001, mouse bioassays were conducted in five highly prevalent villages according to the 2000 infection surveys, namely Xinmin7, Shian5, Minhe3, Xinlong7, and Tuanjie2 [9]. In each village, four to six sites were selected as the locations where bioassay was deployed. Most selected sites met one or more of the following three criteria, i.e., locations which (1) were close to residential areas; (2) had infected snails detected in previous surveys; and (3) were near village borders. Some other considerations, including whether sites were close to economic crops or near school, were also applied to the selection of a few sites. Altogether, twenty five sites were selected, and in each site, a bioassay was conducted for two consecutive days, five hours per day, at the end of each month from May to September. For the experiment, a metal screened cage with five laboratories mice was

suspended such that the bottom of the cage was set just below the surface water, in which case the mice's tails, paws and lower abdomens were dipped in the water. To allow for the development of adult worms, these mice were kept in the laboratories for six month before dissected. The number of worms in each mouse was counted to quantify the infection intensity.

3.2 Surveys in the re-emerging environment (Cohort 2)

The second series of data collection occurred in the re-emerging environment which, as opposed to endemic equilibrium, corresponded to much lower transmission levels. Of the eight counties where re-emergence had been found by 2006, three were selected for this study [4]. Their selection was due to the availability of historical surveillance data and the willingness of local CDC staff to collaborate on the study. In March 2007, historical surveillance records were examined for all three counties to detect re-emergence. Villages were classified as reemerging if schistosomiasis was endemic prior to disease control activity, and either one of the three kinds of infection was detected after control: an acute human infection, an infected child under 12, or an infected snail. Based on this classification, 25 villages were selected from the re-emergent (R) villages and 28 villages were selected from the non-re-emergent (NR) villages. Accordingly, Cohort 2 included all residents in the 53 villages. As a subset of Cohort 2 was enrolled in the educational intervention introduced in Chapter 2, some surveys to be discussed next have already been discussed. However, to provide an integrated introduction on all data incorporated in the individually-based model, detailed study populations, protocols and survey contents will be specified in this chapter as necessary. The names and locations of the counties and villages are withheld due to the sensitivity of reporting infection survey data in regions where disease control had been officially announced.

3.2.1 Demographic survey

Demographic surveys were conducted to the complete study population, which was comprised of all residents aged between 6 and 65 from the 53 villages, in June 2007. Information was collected on each individual's age, gender, primary occupation, education, and anti-schistosomiasis treatment history. For each family, the household head was further asked to finish a more detailed questionnaire regarding household contents (toilet and well), agricultural practices, cattle ownership, and socioeconomic status which was represented by the type of house (concrete, brick or adobe) and ownership of the following items: car, tractor, motorcycle, computer, television, washing machine, air conditioner, and refrigerator.

3.2.2 Baseline infection survey

The first round of infection survey was conducted in November and December of 2007. Stool samples from three consecutive days were requested from each individual. Using the same protocols as earlier introduced for the Xichang study, the miracidial hatch test was applied to all samples and for one of them, the Kato-Katz method was used to quantify the infection intensity. Of the 3009 people who provided any sample, three samples were collected from 2504 people, two samples from 207 people and one sample from 298 people. As a result, 195 people were found infected with *S. japonicum*, and no infection was detected for the remaining 2814 people (Figure 3.1). More detailed results

will be presented later.
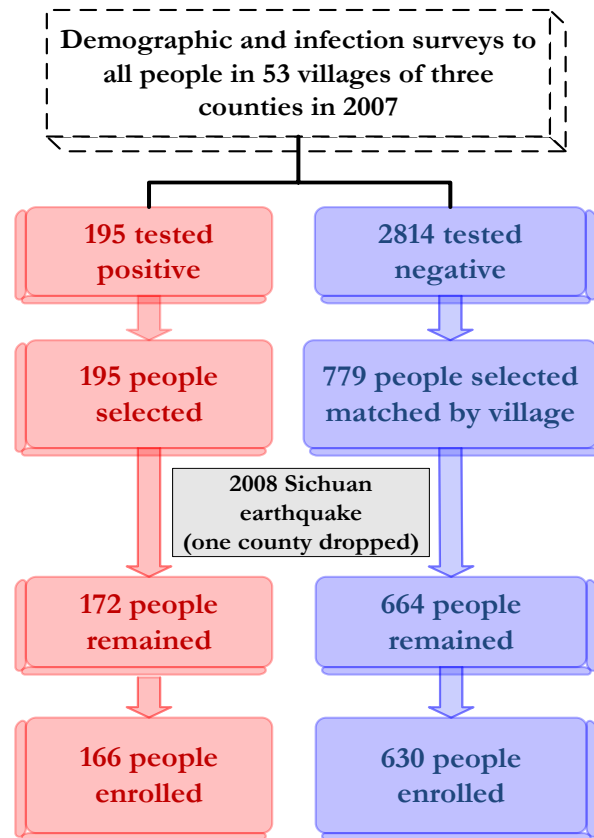
Figure 3.1 Study population of the baseline surveys and the case-control study

3.2.3 Water contact survey and the second infection survey

Based on the result of the baseline infection survey and considering efficiency and logistic feasibility, a subset of the whole population was invited to participate in a series of longitudinal surveys regarding individual water exposure. In reference to Figure 3.1, all of the 195 people who tested positive in 2007 were selected. From the same village in which each infected individual resided, four individuals were randomly selected from those who tested negative. If more than one in four individuals were infected in a village, all uninfected individuals were selected as well. Unfortunately, due to the 7.9 magnitude Sichuan earthquake which occurred on May 12, 2008, one of the three counties was catastrophically damaged and thus excluded from the study. Finally, 796 people from the remaining 31 villages consented to enroll, including 166 infected and 630 uninfected people according to the baseline infection survey.

The water contact surveys were conducted monthly between June (including data collection regarding both May and June) and October of 2008. The surveys were questionnaire-based and conducted by trained Sichuan CDC and county Anti-Schistosomiasis Control Station personnel using the local dialect. Eligible individuals were notified beforehand and encouraged to participate in person. If absent, family members of the eligible individuals were allowed to respond on their behalf.

Each time, people were asked to report on 10 categories of water contact activity in the past half month: washing clothes, washing agricultural tools, washing hands or

feet, playing and swimming, ditch cleaning or diverting, rice planting, rice harvest, fishing, washing vegetables, and drinking or cooking. For each activity, i.e., "washing clothes" as the example in Table 3.1, participants were asked to report as many as three locations where this activity occurred. Shown in

Figure 3.2 is the simplified map of a specific village, numbered 210302, the area of which is divided into 13×13 cells. Water bodies, including ditches, rivers and ponds, are represented by the black dots. Thus, each location at which individuals had water contact could be recorded by its coordinates in this map, i.e., F5 or J9, etc. Finally, the frequency (times/month), f, and duration (minutes/time), d, associated with each activity/location combination were recorded, as well as a qualitative estimate of the water flow velocity, which, at least conceptually, is related to the likelihood of infection if cercarial exposure is involved.

Table 3.1 Sample of the water contact survey questionnaire with hypothetical data

Activity	Location		Frequency	Duration	Flow velocity
Washing clothes	1	F5	2	30	Fast
	2	J9	3	20	Slow
	3	D10	5	15	Moderate

Also for these 796 people, a second infection survey was conducted in October and November of 2008 using the same protocols and methods as the 2007 infection survey. Each person who tested positive for *S. japonicum* in the 2007 or 2008 infection surveys was provided treatment with 40 mg per kg of praziquantel tablets by the county Anti-Schistosomiasis Control Station.
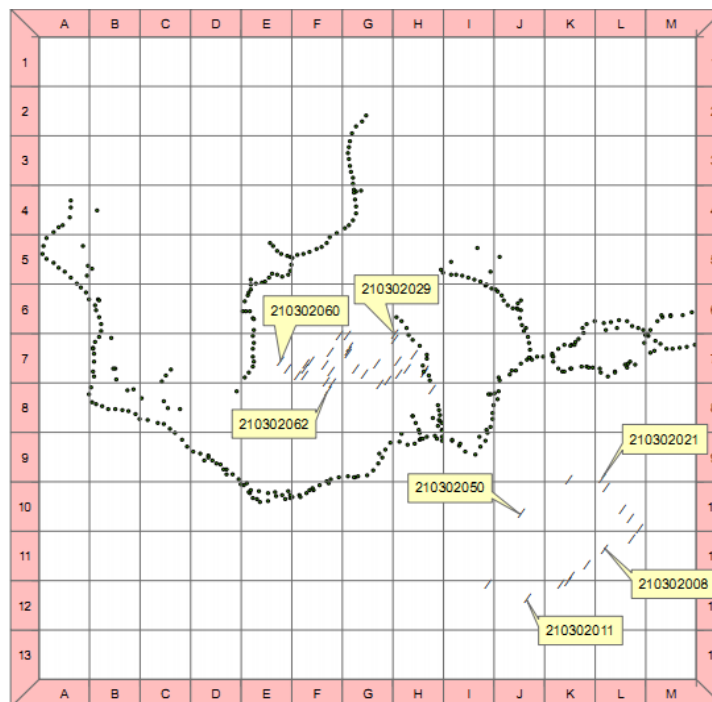
Figure 3.2 A conceptual map of one village (210302) with dots representing the water bodies (rivers, streams or ponds) and several households labeled

39

3.2.4 Infected snail survey

Snail surveys were conducted in all irrigation ditches, including mud, block and concrete ditches of all sizes, of the study villages in April, 2007. In each village, *O. hupensis* snails were collected using the standard kuang method ($0.11m^2$ square frames collected at 10-meter intervals) for all ditches. A GPS coordinate was collected corresponding to each kuang site. Furthermore, as moisture-abundant terraces were also potential habitat of snails, in villages where terraces were present, 10 terraces were randomly chosen for snail sample collection along the base of the terrace at three locations along the terrace wall (the two ends and in the middle of the terrace wall). Collected snails were deposited in envelopes with location ID and coordinates labeled, which were subsequently transferred to a lab for snail crushing to test for snail infection. The presence of schistosome cercariae detected by microscopy indicated infection.

3.2.5 Mouse bioassay

Mouse bioassay were originally scheduled for 2008 but postponed to 2009 due to the earthquake. They were conducted three times: during the primary planting season between late April and early May, June, and the secondary planting season during September. Experiments were deployed in the eight villages where at least 9 people were found infected in 2007. A total of 50 sites were selected from the locations where water contact behaviors were reported in the 2008 surveys. In each time, 5 mice were exposed in the water for 5 hours a day in the two-day interval and dissected in the laboratory six weeks post exposure to test for infection.

3.3 Water exposure estimate

3.3.1 Water contact duration

Using data collected from the water contact surveys in both cohorts, the total water contact duration, measured by the number of minutes, can be calculated by summarizing the duration multiplied by frequency across all locations, activities and months. It can be mathematically expressed by the following equation:

$$WC_i = \sum_m \sum_a \sum_l f_{i,m,a,l} \cdot D_{i,m,a,l} \qquad \text{(Equation 3.1)}$$

Where *i* stands for individual, *m* for month, *a* for activity, *l* for location, *f* for frequency and *D* for duration.

3.3.2 Body surface area

In addition to duration, another important variable used to estimate water exposure degree is the body surface area involved. It is defined as the wetted area of the body part while having water contact activities, which is measured by m^2 [10]. The detailed classification of body surface area used in this dissertation can be found in Table 3.2, which followed the estimations of Mosteller [13] and Ross [14]. The proportion of total body area related with each activity varies considerably from 5% to 32%. Ignoring the difference between genders, body surface area is also determined by the age of the individual, and the total area of adults (aged 15 or older) is set to be approximately 1.4 times more than children (less than 15 years old).

Table 3.2 Body surface area involved in multiple water contact activities for children (less than 15 years old) and adults (at least 15 years old)

Activity	Fraction of total body area* (%)	Body surface area (m^2)	
		Children (<15 yo)	Adults (≥ 15 yo)
Washing clothes and vegetables	5	0.06	0.08
Washing agricultural tools	3	0.03	0.05
Washing hands and feet	12	0.14	0.20
Playing and swimming	20	0.23	0.33
Ditch diverting and cleaning	5	0.06	0.08
Plough and rice planting	5	0.06	0.08
Rice cutting	5	0.06	0.08
Fishing	32	0.36	0.52

*: Total body area is 1.13 m^2 for children and 1.63 m^2 for adults.

3.3.3 Water exposure magnitude

Synthesizing both the body surface area and duration of water contact activities yields a comprehensive estimate of the water exposure magnitude (WE) measured by $\min \cdot m^2$. Adjusting Equation 3.1 to incorporate body surface area ($S_{i,a}$) associated with each activity, the total WE of each individual can be calculated by:

$$WE_{i,m,a,l} = S_{i,a} \cdot D_{i,m,a,l}$$
$$WE_i = \sum_m \sum_a \sum_l \sum_f WE_{i,m,a,l}$$

(Equation 3.2)

Further, the mean body surface area for each individual is the total water exposure magnitude divided by the total water contact duration:

$$\bar{S}_i = \frac{WE_i}{WC_i}$$

(Equation 3.3)

3.4 Ethics

The research protocols and informed consent procedures of all surveys for both cohorts were approved by the Sichuan Institutional Review Board and the University of California, Berkeley, Committee for the Protection of Human Subjects. Oral informed consent was obtained from participants of Cohort 1 and recorded by Sichuan provincial CDC staff. Written, documented informed consent was obtained from Cohort 2. Assent was provided by minors and their parents or guardians provided written, informed permission for them to participate in this study.

3.5 Survey results regarding Cohort 1

3.5.1 Demographic distribution

The demographic characteristics of the participants of the two infection surveys and the water contact survey in Xichang are summarized in Table 3.3.

Table 3.3 Demographic characteristics of participants of the 2000 infection survey, the 2000 water contact survey, and the 2002 infection survey among Cohort 1 in Xichang

	Demographic / Infection survey (2000)		Water contact survey (2000)		Infection survey (2002)	
Total	3897		1067		1973	
	No.	%	No.	%	No.	%
Age (years)						
5-15	1046	27	270	25	457	23
16-30	1125	29	301	28	515	26
31-45	1031	26	295	28	579	29
46+	695	18	197	19	422	21
Sex						
Female	1912	49	557	52	960	49
Male	1985	51	506	48	1010	51
Education						
≤ Elementary	2503	64	686	65	1289	65
≥ Middle school	1390	36	376	35	681	35
Occupation						
Farmer	2259	58	599	56	1150	58
Student	1042	27	261	25	306	26
Other	596	15	203	19	517	16

Table 3.4 Prevalence and mean EPG vs. demographic characteristics from the 2000 and 2002 infections surveys among Cohort 1 in Xichang

	Infection 2000 (N=3987)		Infection 2002 (N=1973)	
	Prevalence (%)	Mean EPG	Prevalence (%)	Mean EPG
Total	28.6	25.4	30.9	14.1
Age (years)				
5-15	24.3	18.2	22.1	8.9
16-30	30.5	36.4	31.0	18.7
31-45	32.5	24.3	36.7	14.8
46+	26.7	22.0	32.1	14.0
Sex				
Female	28.9	27.7	28.7	14.1
Male	28.4	20.7	33.0	14.1
Education				
Elementary	31.2	28.3	32.6	16.7
Middle school	23.6	22.5	28.3	9.4
Occupation				
Farmer	32.8	28.7	36.7	16.5
Students	22.8	18.1	22.1	14.7
Other	22.5	25.3	24.5	8.8

The distribution of each demographic variable was similar in the three surveys. For age, the number of participants was close in the three younger categories but smaller in the oldest group. The sex ratio was approximately unity, and in terms of the education level, only a little over 1/3 of all participants had some middle school or higher education. Finally, regarding the primary occupation, more than half were farmers and about 1/4 were students (including pre-school children); the remained, labeled as "other", had miscellaneous occupations such as governmental official, fisherman, businessman, and housework.

3.5.2   Infection surveys

Shown in Table 3.4 are the results of the infection surveys. Recall that the 2000 survey has consistently been assumed to correspond to the endemic equilibrium level. The overall prevalence was 28.6% and mean EPG was 25.4 among the 3893 people. Prevalence was between 24.3% and 34.5% in the four age groups, with the youngest and oldest groups having lower prevalences. For infection intensity, young adults aged between 16 and 30 had the highest mean EPG, which was twice that of the lowest group and about 50% more than the two older groups. Males had higher mean EPG than females, although the prevalence was close between genders. People with more education had lower prevalence and infection intensity than less educated people. As for occupation, farmers were more infected than students and people with "other" occupations as a whole, although the latter category had much internal variability. In addition, differences in infection by village were significant (Figure 3.3): the prevalence varied from 68% to 2%, and the mean EPG was between 109 and 0.14, which was a 778-fold difference.
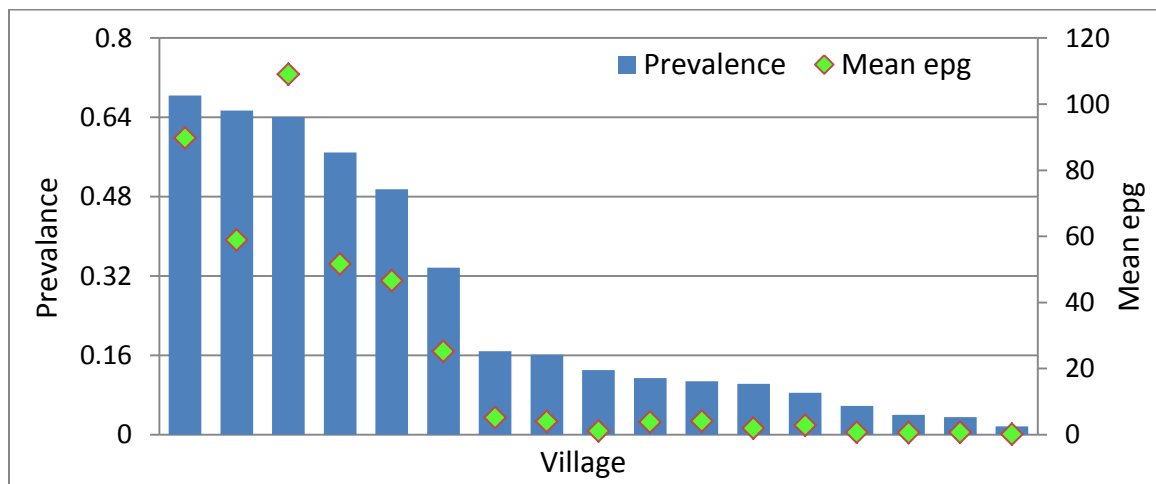
Figure 3.3 Prevalence and mean EPG by village in the 2000 infection surveys in Xichang

In comparison, the 2002 survey results reflected the two-year incidence following the massive chemotherapy in the end of 2000. While the mean EPG decreased largely as expected, the prevalence, surprisingly, increased by a small margin. Infections patterns within age, education and occupation were consistent with the previous results, although the mean EPG in the "other" occupation group had a relatively greater reduction. But regarding the gender variable, somewhat different prevalences were observed between males and females but not infection intensity, which was to the opposite of the corresponding pattern in the 2000 survey result.

43

3.5.3 Water contact surveys

Results of the water contact surveys are presented in Figure 3.4 and Figure 3.5. The top histogram of Figure 3.4 shows the distribution of individual water contact duration of the 529 people whose data will be used for parameterization of the individually-based model (due to the availability of cercarial data in their corresponding villages; see more details in Chapter 4). A clear long-tail pattern can be observed, with the mean being 3657 minutes and varying between 0 and 31940 minutes. The individual averaged body surface area, shown in the bottom histogram, also appears to be right-skewed despite the bump at 0.23. The 89 individuals having exactly the same averaged body surface area were all children aged between 4 and 14, whose annual water contact behaviors were exclusively playing in the water mostly during the summer.
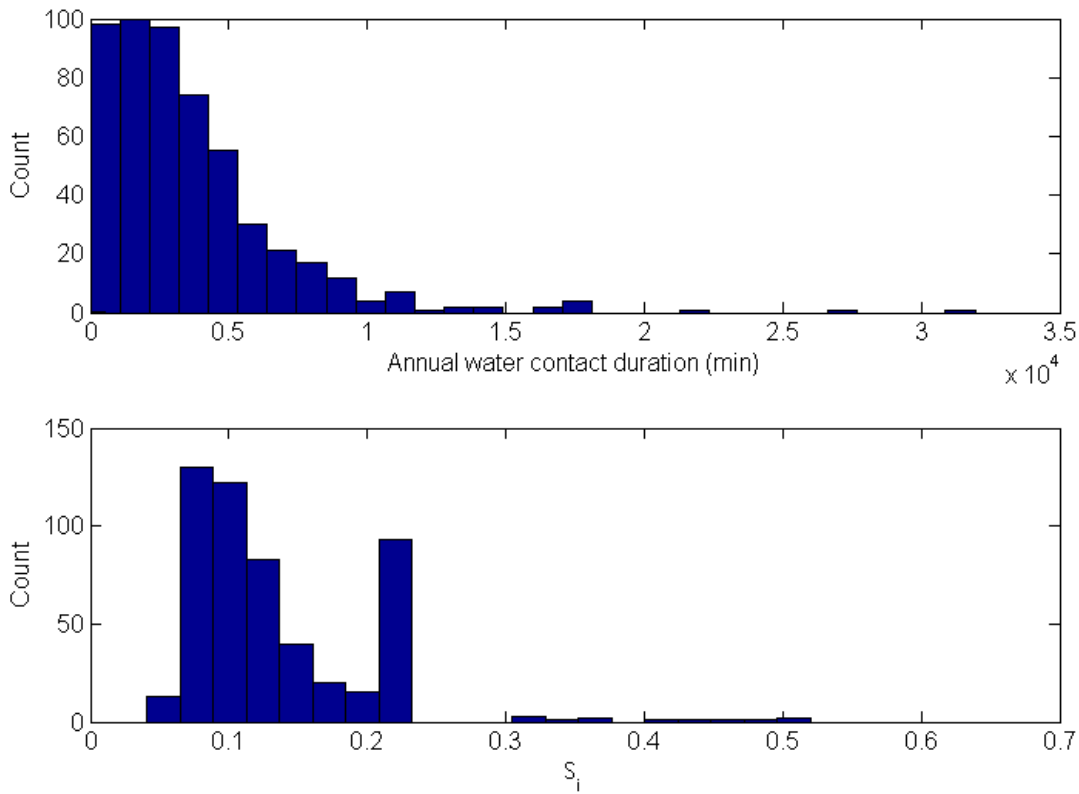
Figure 3.4 Histograms of annual water contact (WC) duration and average body surface area (S_i) of 529 people from five villages of Xichang

Accounting for both duration and body surface area, Figure 3.5 shows the mean water exposure magnitude versus demographic variables. A consistently increasing trend was found within age groups, indicating older people had more exposure than younger people. Males and people with less education were more exposed than their counterparts for potentially involving in more farming and/or fishing related work. Finally, farmers had more annual exposure than students despite the latter group's intensive recreational exposure in the summer. While fishermen alone had more exposure than farmers, when averaged with people having less-exposed occupations, the mean exposure of all "others" was still lower than farmers and slightly higher than students.
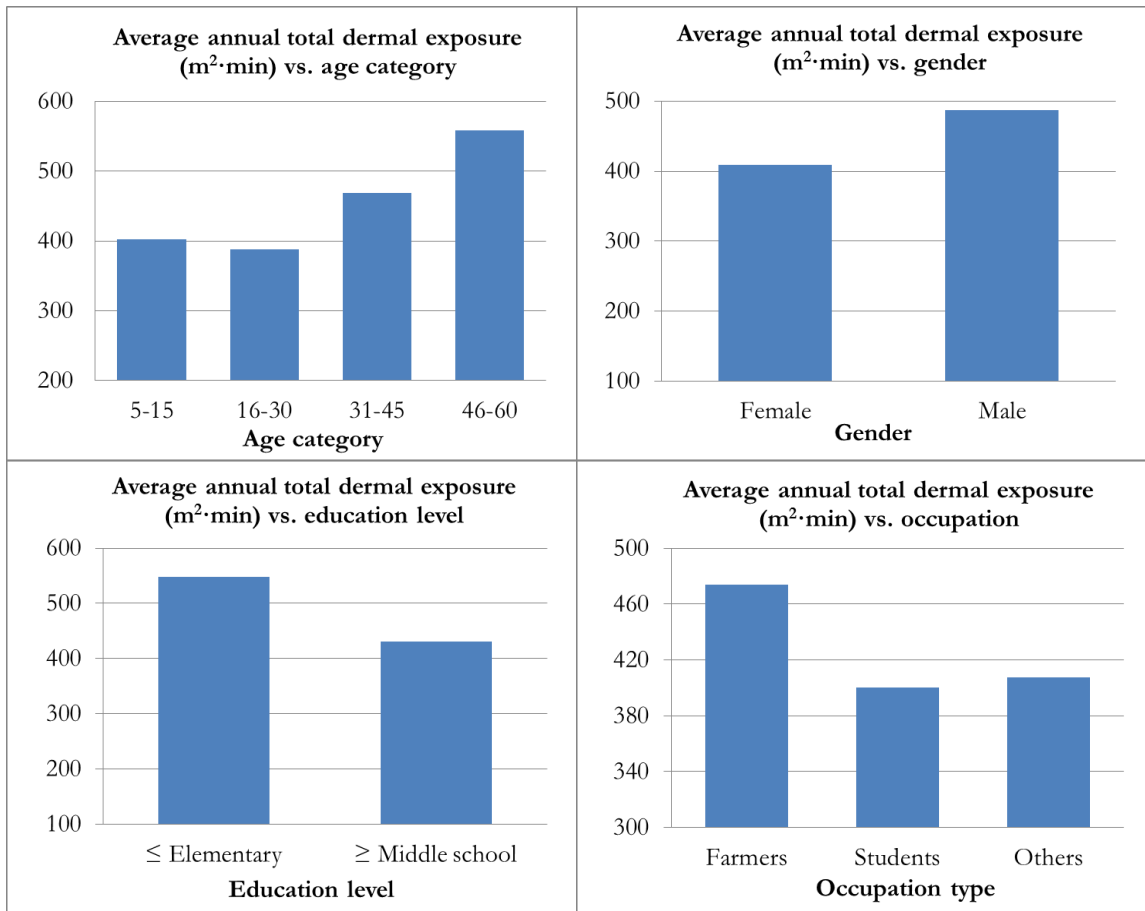
Figure 3.5 Averaged annual water exposure ($\min \cdot m^2$) versus demographic variables among Cohort 1 in Xichang

3.5.4 Mouse bioassay result

The results of mouse bioassay are summarized in Table 3.5. The number corresponding to each site and month represents the mean number of worms found in the dissected mice, which is assumed to be proportional to the cercarial density in the water. A huge variance could was found on three dimensions: inter-village, intra-village and seasonal. First, among the five villages, the overall mean worm count of all months varied from 0.54 (Tuanjie2) to 30.56 (Xinmin7), which suggested a 56-fold difference. Second, within each village, a wide range of mean worm counts was found at different sites. For instance, the ratio between the largest and the smallest non-zero annual means of each village was between 3.5 (for Shian5) and 72.8 (for Minhe3), reflecting significant geospatial variety among the multiple locations. Finally, the worm counts of each month show considerable time-variation but inconsistent trends. Figure 3.6 shows such patterns in the three villages where mean worms were relatively large. In Xinmin7 (note the logarithmic scale on the y-axis), the highest count at site 5 (259.6) occurred in August; but for three of the other four sites, the count in August was below the annual average. A clear "W" shape was found at site 4 of Shian5, which suggested lower counts in June and August and higher counts in May, July and September. However, an "M" shaped pattern is seen at site 5 of Minhe3 which suggests just the opposite trend. Consequently, there

45

was no consistency among the sites from which an explicitly temporal pattern could be determined. Rather, the cercarial density in water will be assumed as a stationary random process in model development.

Table 3.5 Mouse bioassay results by month in 5 villages from Xichang in 2002

| Village	Site	May	June	July	Aug	Sept	Mean	Variance
Xinmin7	1	32.5	96	2.8	8	7.6	29.38	1217.19
	2	20	3	5.6	0.8	1.2	6.12	51.03
	3	73	4.8	6.8	7.2	16.3	21.62	675.74
	4	2	5.8	5.25	8	11	6.41	8.95
	5	79.8	41.5	7.8	259.6	1.2	77.98	9028.07
	Mean	39.5	31.7	5.6	68.9	7.1	30.56	544.91
Shian5	1	0.4	6	8.6	11	2.4	5.68	15.11
	2	0	1	2.8	2.4	2.8	1.8	1.25
	3	1.8	0	1	2.8	16.6	4.44	37.81
	4	11	0.7	7	0.8	12.4	6.38	24.27
	Mean	2.5	1.5	4.4	4.3	7.6	4.06	4.34
Minhe3	1	8.6	0	0.8	0	0.4	1.96	11.11
	2	4.2	26	0.8	2.5	5.6	7.82	85.23
	3	16.5	30.5	4.8	4.8	6.3	12.58	99.40
	4	36	8	2.8	25	12.5	16.86	145.56
	5	3	16	0	22.2	4.4	9.12	72.27
	6	0.4	0	0	2.8	0	0.64	1.19
	Mean	7.5	17.8	1.5	10.5	4.5	8.36	31.28
Xinlong7	1	9.2	0	3.4	16.5	0	5.82	39.82
	2	4.4	0	1.6	0	1	1.4	2.62
	3	0	0	0	0.4	0	0.08	0.03
	4	0	0	0	0	0	0	0.00
	5	0	0	0	0	0	0	0.00
	6	0	0	0	0	0	0	0.00
	mean	2.7	0	1	1.8	0.1	1.12	1.05
Tuanjie2	1	7	0	1.2	0	2.6	2.16	6.77
	2	0	0	0	0	0	0	0.00
	3	0	0	0	0.4	1	0.28	0.15
	4	1.2	0	0	0	0.4	0.32	0.22
	mean	1.5	0	0.3	0.1	0.8	0.54	0.31
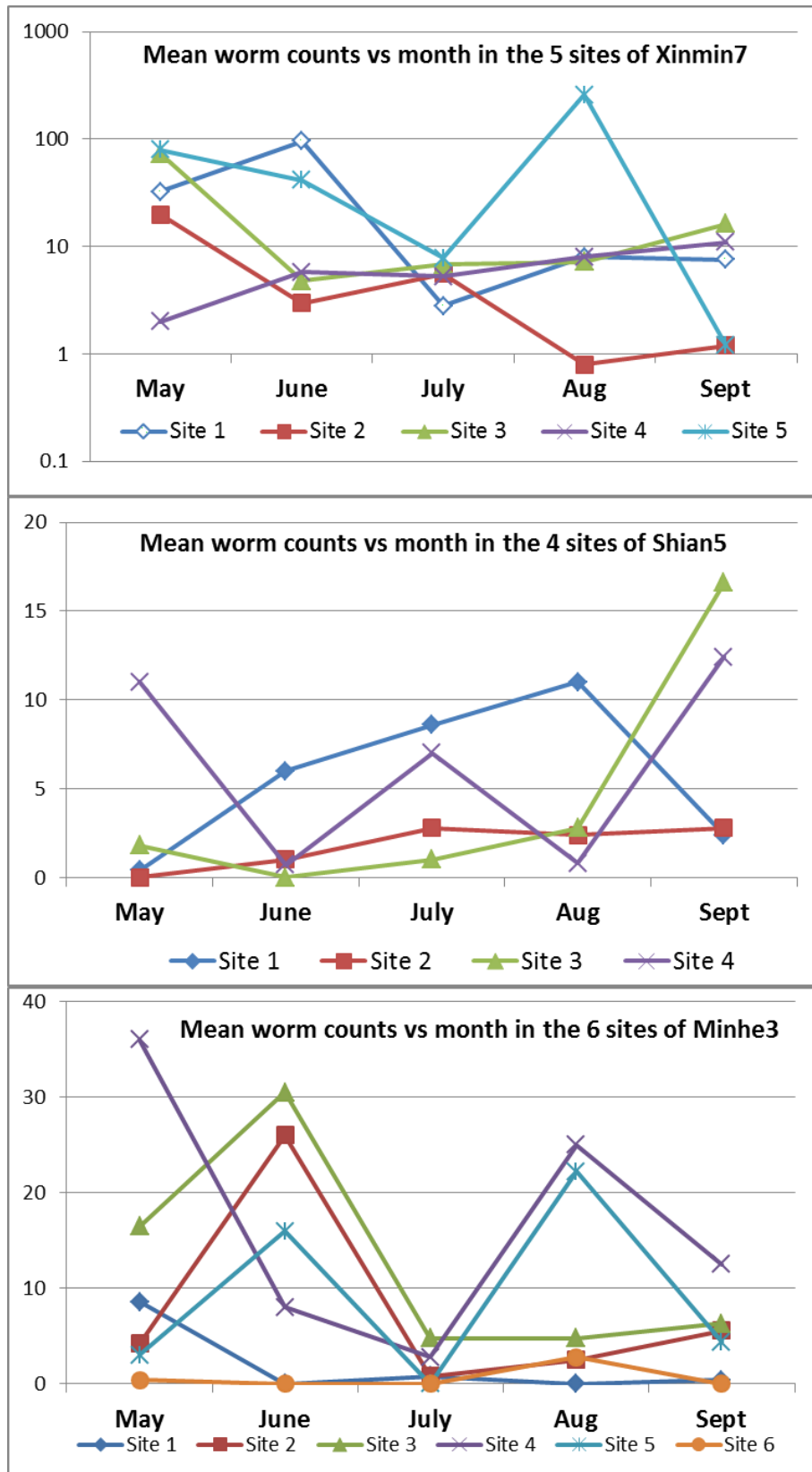
Figure 3.6 Mean worm counts vs. month in each study site of three villages in Xichang

3.6 Survey results for Cohort 2

3.6.1 Demographic distribution

Shown in Table 3.6, the demographic distributions of Cohort 2 were somewhat different than Cohort 1. While the sex ratio was again unity and about one third of the people had post-elementary education as in Cohort 1, there was a higher percentage of people in the upper two age categories, and farmers were a more dominate proportion among occupational groups. In the following surveys, the participation of children and young adults aged between 16 and 30 was particularly low. People having less education were more likely to be tracked, and the participation rates regarding gender and occupation groups were relatively stable across all surveys.

Table 3.6 Demographic characteristics of participants of the surveys among Cohort 2

	Demographic survey (2007)		Infection survey (2007)		Completion of water contact survey (2008)		Infection survey (2008)	
Total	2719		2019		634		611	
	No.	%	No.	%	No.	%	No.	%
Age (years)								
5-15	316	12	196	10	43	7	55	9
16-30	368	13	183	10	47	7	31	5
31-45	968	36	709	36	225	36	199	33
46+	1067	39	875	44	319	50	326	53
Sex								
Female	1366	50	1013	51	306	50	328	54
Male	1356	50	964	49	328	50	282	46
Education								
≤ Elementary	1706	63	1313	67	435	69	407	73
≥ Middle school	1005	37	635	33	197	31	151	27
Occupation								
Farmer	2299	84	1752	89	482	81	440	85
Students	371	14	192	10	103	17	26	10
Others	46	2	18	1	8	2	54	5

3.6.2 Infection surveys

Summarized in Table 3.7 are the infection survey results of the whole population in 2007 and the subset in 2008. Compared with Cohort 1, the overall infection prevalence and intensity were all much lower, reflecting the substantial difference between endemic and re-emerging environments. In the baseline survey, people aged over 30 had higher prevalence and mean EPG than the other two age groups, and so did people with lower education. The difference between genders was not large: males were about 11% more prevalent than females but the mean infection intensity was about 13% lower. Farmers had a much higher prevalence and mean EPG than any other occupational groups. Specifically, no infections were found in people who were neither farmers nor students.

Figure 3.7 presents the prevalence and mean EPG of each village having people who were infected. Compared with Cohort 1, less variance on prevalence was found between villages, which ranged from 0.01 to 0.41. Also, the correlation between prevalence and EPG was not as evident as the previous pattern, particularly in higher prevalent villages.

Table 3.7 Prevalence and mean EPG vs. demographic characteristics from the 2000 and 2002 infections surveys of Cohort 2.

|	Infection 2007 (N = 2019)		Infection 2008 (N = 611)	
	Prevalence (%)	Mean EPG	Prevalence (%)	Mean EPG
Total	8.3	2.1	11.7	4.78
Age (years)				
5-15	4.1	0.4	7.2	0.30
16-30	6.3	1.3	9.1	0.78
31-45	9.4	3.4	10.0	0.44
46-60	8.9	1.6	12.6	8.56
Sex				
Female	7.9	2.2	9.5	2.24
Male	8.8	1.9	13.1	7.45
Education				
≤ Elementary	9.3	2.3	12.5	6.85
≥ Middle school	6.6	1.4	7.3	0.64
Occupation				
Farmer	8.9	2.3	11.8	6.20
Students	4.7	0.05	19.2	4.63
Others	0	0	7.4	0.59
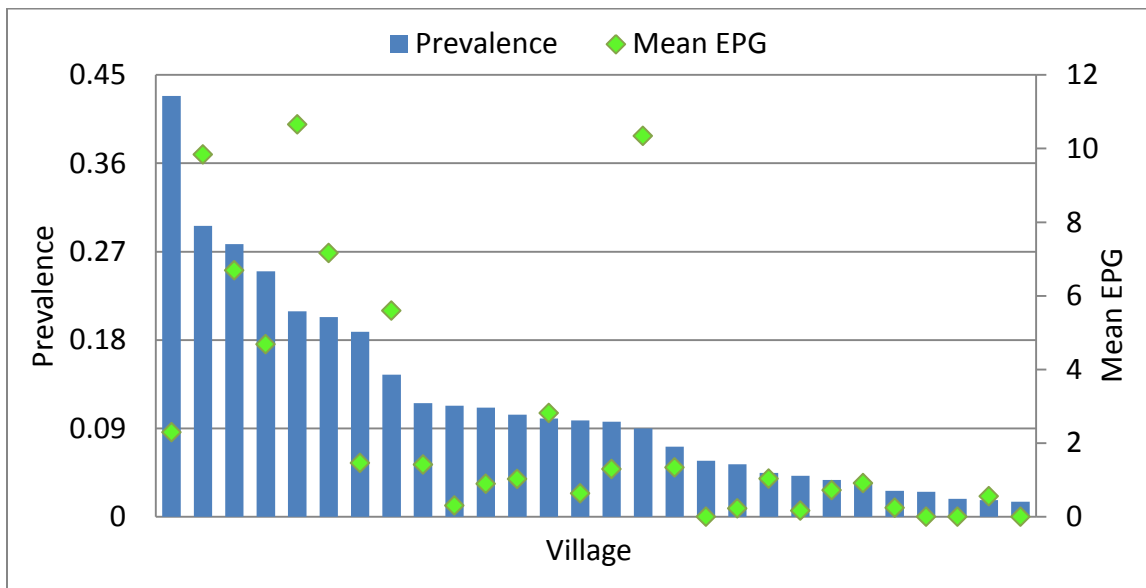
Figure 3.7 Prevalence and mean EPG by village in the 2007 infection surveys in 2 counties of Cohort 2

49

The prevalence in the 2008 survey was higher although enrolling fewer people but including all people who were infected in the previous year, potentially those who had risks higher than average. As in the 2007 results, older, male and less educated people had more infections. However, students were 63% more prevalent than farmers, and the infection incidence of "others" was 7.4% during 2008. Despite the 68 people tested positive, egg counts were detected for only 27 people. Two individuals, having 578 and 1606 EPG respectively, accounted for 75% of the total infection intensity. Thus, the mean EPG of either of the subgroups or all participants was largely due to chance and was not able to reliably reflect the population infection levels.

3.6.3 Water contact surveys

In the longitudinal water contact surveys between May and October of 2008, a total of 611 people provided complete profiles for all months. The demographic information from this dataset, including gender, age, and village of residence were used to develop full profiles for another 116 people who had partial survey data. The distributions of water contact duration and S_i for the 727 total people are presented in Figure 3.8. Compared with Cohort 1, the distribution of total duration was more right-skewed, with the population-mean being 18% greater and the variance more than doubled.
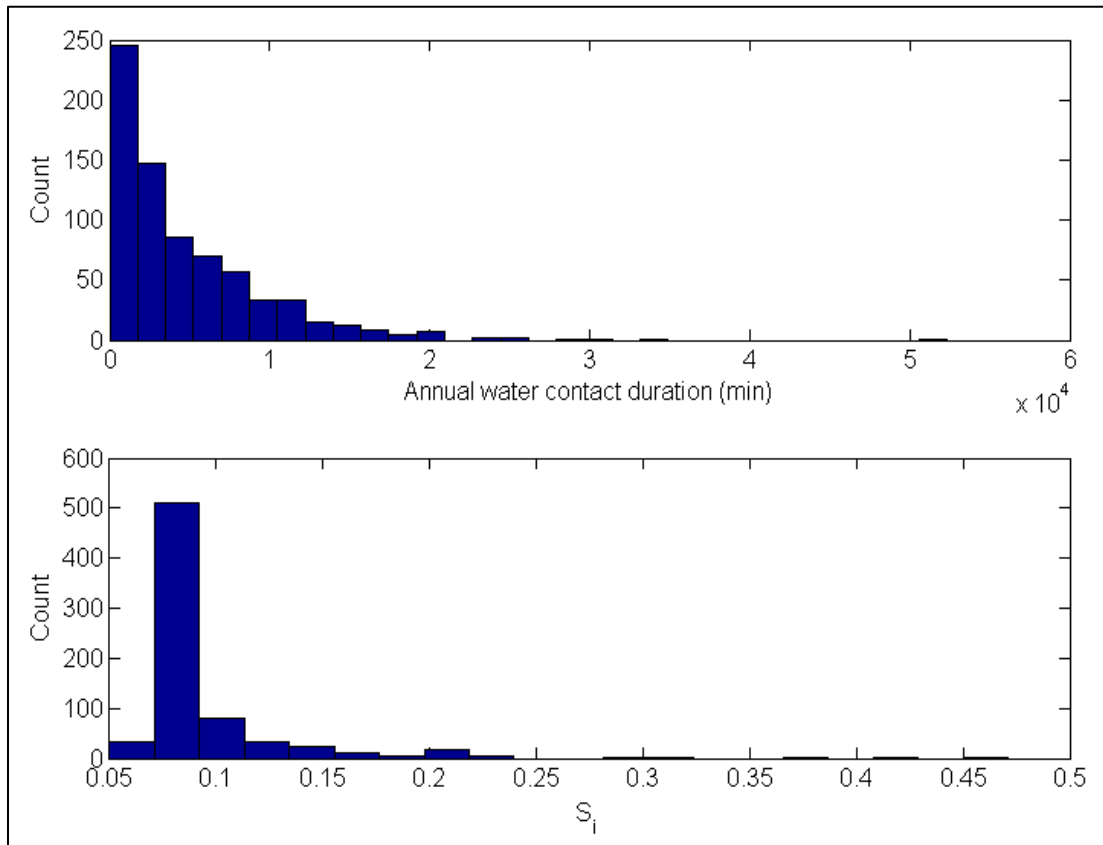
Figure 3.8 Histograms of annual water contact (WC) duration and average body surface area (S_i) of 727 people from Cohort 2

50

On the other hand, the distribution of individual average body surface area was more centralized than Cohort 1. Dividing the observed numbers into 20 categories, 70% of people concentrated in the second group, which ranged between 0.062 and 0.082. This was consistent with the fact that the dominant majority of people were farmers, and the body surface area associated with farming-related activities was all 0.08 for adults except for cleaning tools, which was 0.06.

Because of the smaller average S_i, the mean water exposure magnitude of Cohort 2 was lower than Cohort 1 despite the longer exposure duration (379 vs 447 min$\cdot m^2$). However, the patterns within the demographic variables were consistent (Figure 3.9): it is still the older, male and lower-educated people who had more exposure, and the ratios between subgroups were close to those of Cohort 1, except that the exposure of the youngest group was relatively lower. Farmers were still more exposed than students, although the ratio was larger as the latter group did not appear to have intensive recreational water contact in the summer. Two fishermen accounted for 95% of the total water exposure in the "other" category, which included only 11 people altogether. Unsurprisingly, the overly dispersed distribution yielded a mean which was higher than the farmer group.
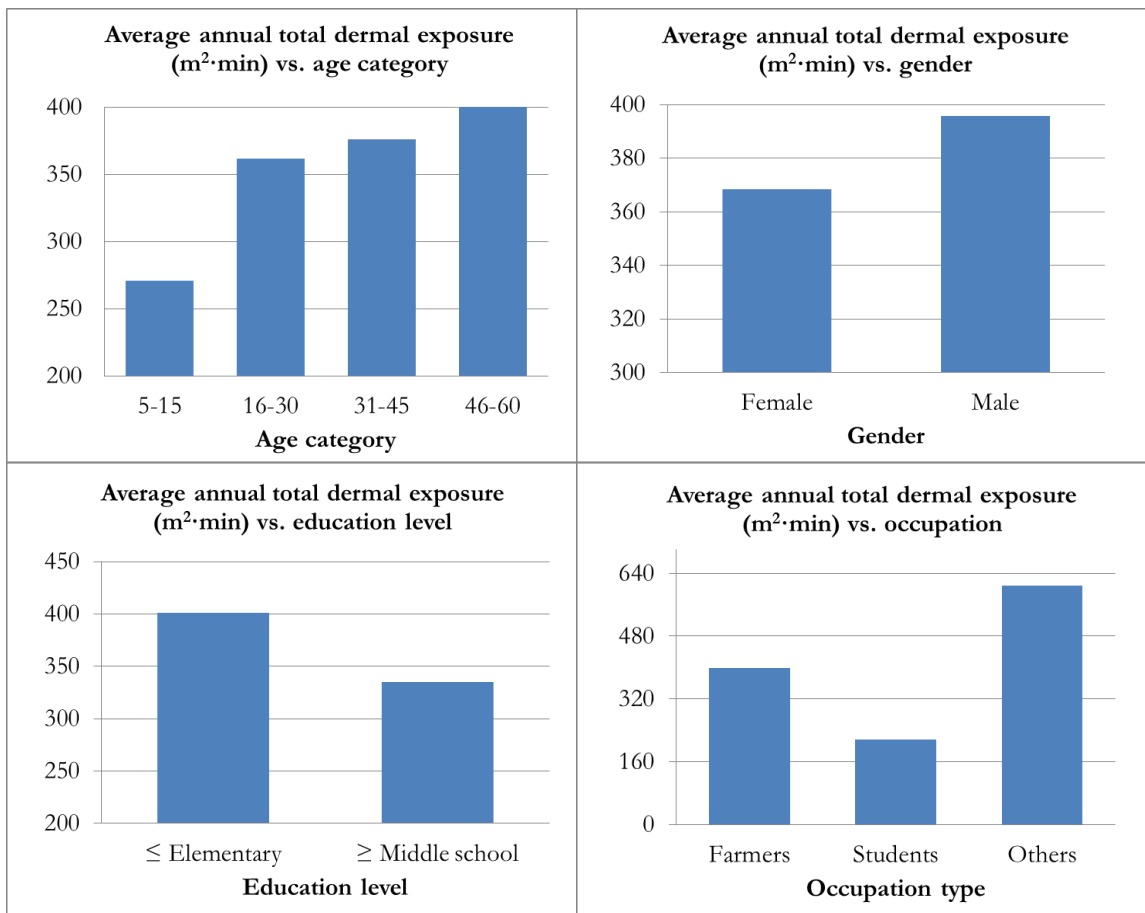

Figure 3.9 Averaged annual water exposure (min$\cdot m^2$) versus demographic variables among Cohort 2 in the re-emerging environment

3.6.4 Mouse bioassay and snail survey results

Table 3.8 summarizes the results of mouse bioassay in Cohort 2 villages, which, unfortunately, were not informative. Of the 657 mice deployed at a total of 150 sites in three months, only 1 tested positive. Alternatively, the distribution of infected snails can also be used to estimate a cercarial density distribution. However, while snail surveys were conducted in all villages and snail density by village was estimated, only 1 out of 7325 snails was found infected (Figure 3.10) [1]. Thus, data from neither survey was useful in providing information on cercarial density. Also apparently, the ongoing mollusciciding was effective in most villages.

Table 3.8 Mouse bioassay results by month in 5 villages from Cohort 2 in 2009

Month	Sites	Mean exposure time in hours (range)	Mean mice/site (range)	Total mice	Infected mice
May	51	10.2 (8.0 – 11.9)	4.2 (1-5)	216	1
June	48	9.6 (5.0 – 11.0)	4.2 (2-5)	203	0
September	51	10.1 (9.9 – 11.1)	4.7 (1-5)	238	0
Total	150	10.0 (5.0 – 11.9)	4.4 (1-5)	657	1
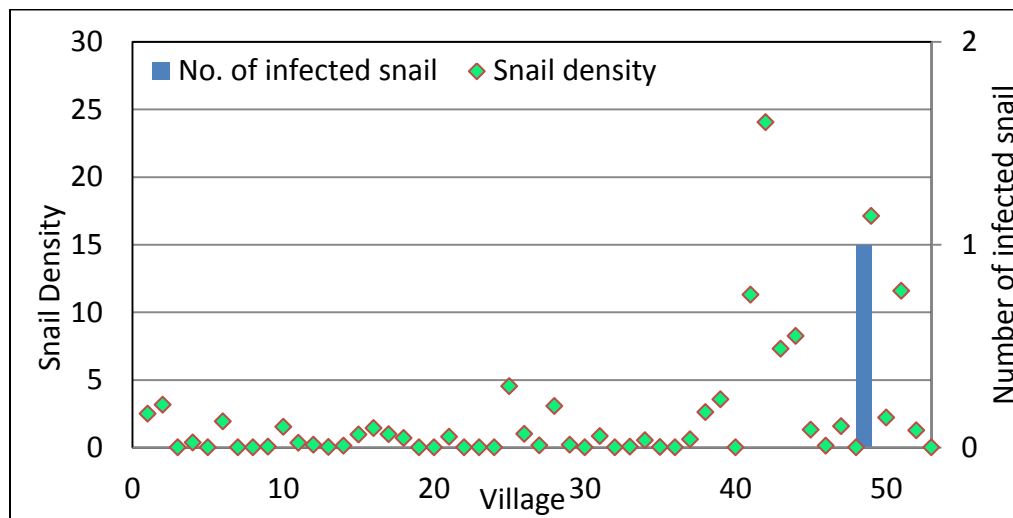
Figure 3.10 Snail density and number of infected snails by village

3.7 Discussion

This chapter presented the study protocols and descriptive analyses of various survey data to be used in the development of the individually based model. Surveys were conducted in two cohorts who both resided in hilly and mountainous regions of Sichuan but corresponded to different transmission environments. There were some consistencies between the two cohorts: (1) individual water contact duration was highly right-skewed with a large variance; (2) intergroup differences regarding water exposure were found according to age, gender, education and occupation; (3) infection prevalence and intensity varied significantly by village. The different capability and hydrological condition of each village to sustain the transmission, along with individual factors such as water exposure and susceptibility, no doubt contributed to this pattern [3, 7, 15].

52

On the other hand, the results clearly reflected the substantial differences between the endemic equilibrium and re-emerging transmission environments: (1) mouse bioassay and snail infection surveys were not able to indicate the cercarial distribution in the low-transmission levels; (2) quantitative estimates of infection intensity using the Kato-Katz method proved much less effective in the re-emerging environment. For instance, of all people who were tested positive, egg counts were available for 85% and 70% in the 2000 and 2002 surveys for Cohort 1, respectively, but only for 45% and 40% in the 2007 and 2008 surveys for Cohort 2. These findings strongly suggest the inadaptability of current parasite monitoring and infection testing techniques in the low-transmission setting and the urgent needs for improvement and the development of more advanced alternative methods [2, 15].

The results also showed the importance of conducting surveys regionally and making control strategies tailored to local conditions. For example, while many studies in other regions of the world all generally suggest that younger people had much more exposure than the elderlies , this pattern was not observed in either of our study populations. This was mainly due to the fact that younger people in rural areas of China have been increasingly tending to work in cities as immigrant workers, so it is the older people who have most burdens of agricultural work. But even between the two cohorts which shared much similarity regarding geographical characteristics, living environments and economic structures, somewhat different patterns were found for the distribution of body surface area, an index directly associated with the specific occupation and water contact behavior patterns in the study populations. In addition, although not presented in this dissertation, the possession of definitive host animals, mainly cows and water buffalo, was quite different in the two cohorts. Opposed to cohort 1 where almost no bovines were raised, they were commonly used for agricultural purposes in cohort 2 and played an important role in disease transmission [1]. Similarly, comparing transmission characteristics between Sichuan and the marshland environments in lower Yangtze River only led to more diverse patterns [16]. Thus, disease control methods should be designed in accordance with the various types and levels of risk factors in different regions.

Finally, it is worth mentioning that the results presented in this chapter were mainly at the population level to provide a snapshot of the two cohorts and compare the overall exposure and transmission levels. However, it is individual data that will be incorporated in the individually-based model. Thus, in the next chapter, I will introduce the process of model development, specify how these data are used for parameterization, and how the model was set up for simulations.

References

1.Carlton, E.J., Bates, M.N., Zhong, B., Seto, E.Y., and Spear, R.C., *Evaluation of mammalian and intermediate host surveillance methods for detecting schistosomiasis reemergence in southwest China.* PLoS neglected tropical diseases, 2011. **5**(3): p. e987.

2.Carlton, E.J., Hubbard, A., Wang, S., and Spear, R.C., *Repeated Schistosoma japonicum infection following treatment in two cohorts: evidence for host susceptibility to helminthiasis?* PLoS Negl Trop Dis, 2013. **7**(3): p. e2098.

3.Liang, S., Seto, E.Y., Remais, J.V., Zhong, B., Yang, C., Hubbard, A., Davis, G.M., Gu, X., Qiu, D., and Spear, R.C., *Environmental effects on parasitic disease transmission exemplified by schistosomiasis in western China.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(17): p. 7110-7115.

4.Liang, S., Yang, C., Zhong, B., and Qiu, D., *Re-emerging schistosomiasis in hilly and mountainous areas of Sichuan, China.* Bulletin of the World Health Organization, 2006. **84**(2): p. 139-144.

5.Seto, E., Liang, S., Qiu, D., Gu, X., and Spear, R.C., *A protocol for geographically randomized snail surveys in schistosomiasis fieldwork using the global positioning system.* The American Journal of Tropical Medicine and Hygiene, 2001. **64**(1-2): p. 98-99.

6.Seto, E.Y., Lee, Y.J., Liang, S., and Zhong, B., *Individual and village-level study of water contact patterns and Schistosoma japonicum infection in mountainous rural China.* Tropical medicine & international health : TM & IH, 2007. **12**(10): p. 1199-1209.

7.Spear, R.C., Seto, E., Liang, S., Birkner, M., Hubbard, A., Qiu, D., Yang, C., Zhong, B., Xu, F., Gu, X., and Davis, G.M., *Factors influencing the transmission of Schistosoma japonicum in the mountains of Sichuan Province of China.* The American Journal of Tropical Medicine and Hygiene, 2004. **70**(1): p. 48-56.

8.Spear, R.C., Seto, E.Y., Carlton, E.J., Liang, S., Remais, J.V., Zhong, B., and Qiu, D., *The challenge of effective surveillance in moving from low transmission to elimination of schistosomiasis in China.* Int J Parasitol, 2011. **41**(12): p. 1243-7.

9.Spear, R.C., Zhong, B., Mao, Y., Hubbard, A., Birkner, M., Remais, J., and Qiu, D., *Spatial and temporal variability in schistosome cercarial density detected by mouse bioassays in village irrigation ditches in Sichuan, China.* The American Journal of Tropical Medicine and Hygiene, 2004. **71**(5): p. 554-557.

10.Liang, S., *A spatio-temporal modeling of schistosomiasis transmission dynamics and control in Sichuan, China*, 2003.

11.Seto, E.Y., Zhong, B., Kouch, J., Hubbard, A., and Spear, R.C., *Genetic and household risk factors for Schistosoma japonicum infection in the presence of larger scale environmental differences in the mountainous transmission areas of China.* The American Journal of Tropical Medicine and Hygiene, 2005. **73**(6): p. 1145-1150.

12.Carlton, E.J., Hsiang, M., Zhang, Y., Johnson, S., Hubbard, A., and Spear, R.C., *The impact of Schistosoma japonicum infection and treatment on ultrasound-detectable morbidity: a five-year cohort study in Southwest China.* PLoS

neglected tropical diseases, 2010. **4**(5): p. e685.

13. Mosteller, R.D., *Simplified calculation of body-surface area.* N Engl J Med, 1987. **317**(17): p. 1098.

14. Ross, A.G., Yuesheng, L., Sleigh, A.C., Williams, G.M., Hartel, G.F., Forsyth, S.J., Yi, L., and McManus, D.P., *Measuring exposure to S. japonicum in China. I. Activity diaries to assess water contact and comparison to other measures.* Acta Trop, 1998. **71**(3): p. 213-28.

15. Spear, R.C., *The Role of Environmental Determinants of R0 in Suppressing Re-emergent Transmission of S.japonicum in China.* Journal of the Royal Society, Interface, 2011.

16. Wang, L.D., Guo, J.G., Wu, X.H., Chen, H.G., Wang, T.P., Zhu, S.P., Zhang, Z.H., Steinmann, P., Yang, G.J., Wang, S.P., Wu, Z.D., Wang, L.Y., Hao, Y., Bergquist, R., Utzinger, J., and Zhou, X.N., *China's new strategy to block Schistosoma japonicum transmission: experiences and impact beyond schistosomiasis.* Trop Med Int Health, 2009. **14**(12): p. 1475-83.

Chapter 4 Development of the Individually-based Model

"All that is valuable in human society depends upon the opportunity for development accorded the individual."

--Albert Einstein

4.1 Introduction

Mathematical modeling has been widely used and proven to be a powerful tool for exploring the transmission patterns of infectious diseases [1-4]. For schistosomiasis, considerable effort has been devoted to applying and modifying the ordinary differential equation (ODE) models initially developed by MacDonald and later revised by Anderson and May [5-14]. These models typically define human worm burden and infection prevalence in snails as state variables and yield outcomes on the group level, although some variants have introduced more specific sub-groups due to, for instance, occupation, age, or infection intensity of individuals. A main application of such models, after calibrated to field data from one or multiple surveys, is to make predictions on future infections under different control strategies and thus facilitate decision making.

It is this paradigm that our group followed in previous years to understanding the transmission dynamics of *S. japonicum* in Sichuan. Our community-based model incorporating village-specific internal transmission potentials proved to have fitted well with the endemic equilibrium scenarios corresponding to moderate to high infection levels in Cohort 1 [15-25]. However, in the low-transmission setting in Cohort 2, the community-based model is no longer suitable for several reasons. First, the detection limits of snail surveys and cercarial bioassays, as introduced in Chapter 3, are no longer capable of identifying the spatial and temporal distributions of the intermediate snail host and the free-swimming forms of the schistosome, making it impossible to calibrate the relevant modules. Second, the stochastic nature of interactions between human water contact and cercariae, which is less important in the endemic environment, might now play a significant role determining the likelihood, timing and magnitude of infection. Therefore, the deterministic structure applied in the community-based model needs to be revised to allow for stochastic factors. Finally, as the overall infection level is quite low in the re-emerging environment, only a small portion of people have quantifiable
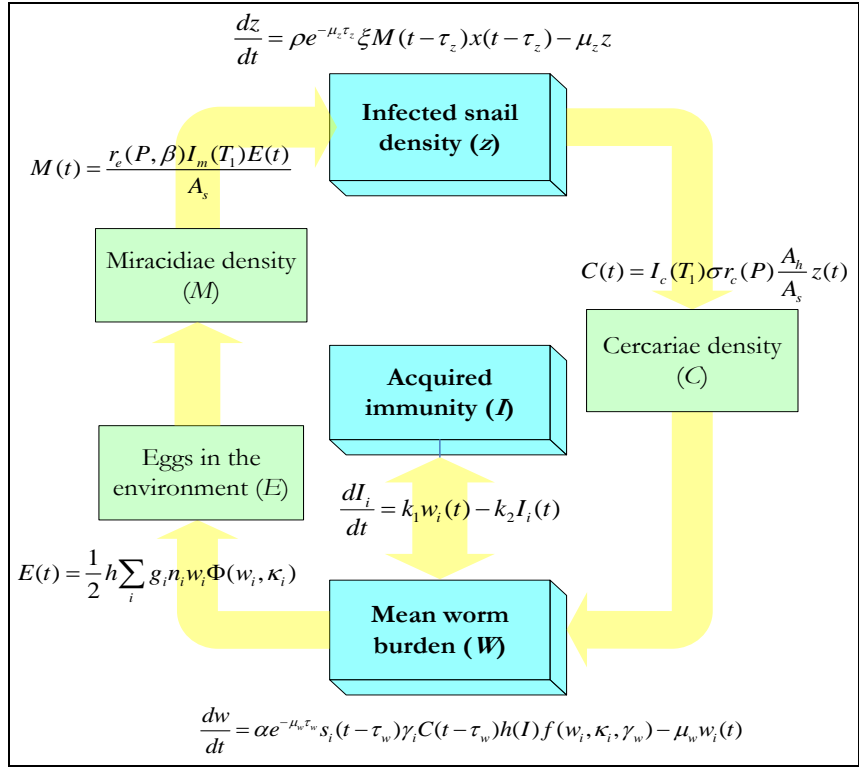
infections under the current infection testing techniques. Thus, it is the individual-level infection, rather than the group-mean, that becomes the more appropriate outcome to identify individuals with higher risks of infection and help prioritize resources accordingly. In consideration of these changes, an individually-based model is necessary to address to the current goal of *S. japonicum* elimination in Sichuan.

Individual-based models are computational simulations of the actions of autonomous individuals and their interactions with the environment [26]. They are motivated by the inability to relate compartmental models to real-world events and the need to consider individual-level variability. This modeling paradigm has two distinguishing features [27]. First, instead of aggregating individuals into a few groups due to specific characteristics like age or occupation, the individually-based model is developed for every individual, thus ensuring individual heterogeneity to be considered to the greatest possible extent. Second, an individually-based model allows for simulations of each individual with specific temporal and spatial information embedded, the latter being particularly crucial when interactions between human and environment are associated with the study question. Thus, in many circumstances an individually-based model is a more "natural description" of a system [26].

4.2   Development of the individually-based model

4.2.1   The community based model

Before introducing the development of the individually-based model, I will start with a brief description of the community-based model, which provides an explicit example of modeling the transmission dynamics with mathematical equations, and then present how it is modified to form the individually-based model. Conceptually, the community-based model is designed to mimic each step of the life cycle of schistosome [16, 18] and, by integrating all transmission steps b, yields a complete simulation of disease transmission between the definitive and intermediate hosts. Shown in Figure 4.1 is the latest version of the community-based model that our group has developed, which includes three state variables: the mean worm burden of the population (*W*), the mean infected snail density in the environment (*Z*), and the mean level of acquired immunity (*I*), the latter module being added most recently. Without going through the details of the equations which can be found elsewhere [16, 18, 23-25], it is worth noting that the model, as mentioned above, has been being applied to simulate the endemic transmission patterns in Cohort 1. In Cohort 2, however, due to the lack of data on a key state variable, infected snail density, the loop cannot be completed and therefore, creating an individually-based model which has the same components as the community-based model becomes impossible. On the other hand, the magnitude of acquired immunity to *S. japonicum* in our study populations was shown to be associated with individuals' infection intensity rather than demographic variables such as age or kinship [23]. According to the logistic association between immunity and worm burden (Figure 4.2) [25], which was suggested by statistical analysis of the epidemiological data of Cohort 1, it requires a high worm burden to trigger the immunity effect, which is rarely seen in the re-emerging environment after years of repetitive treatments. Thus, it is solely the infection intensity indices, including both the worm burden and EPG levels, that will be the focus of the individually-based model.

$$\frac{dz}{dt} = \rho e^{-\mu_z \tau_z} \xi M(t-\tau_z) x(t-\tau_z) - \mu_z z$$

Infected snail density (z)

$$M(t) = \frac{r_e(P,\beta) I_m(T_1) E(t)}{A_s}$$

Miracidiae density (M)

$$C(t) = I_c(T_1)\sigma r_c(P)\frac{A_h}{A_s} z(t)$$

Cercariae density (C)

Acquired immunity (I)

Eggs in the environment (E)

$$\frac{dI_i}{dt} = k_1 w_i(t) - k_2 I_i(t)$$

$$E(t) = \frac{1}{2} h \sum_i g_i n_i w_i \Phi(w_i, \kappa_i)$$

Mean worm burden (W)

$$\frac{dw}{dt} = \alpha e^{-\mu_w \tau_w} s_i(t-\tau_w)\gamma_i C(t-\tau_w) h(I) f(w_i,\kappa_i,\gamma_w) - \mu_w w_i(t)$$

Figure 4.1 The community-based model structure with equations describing all steps of *S. japonicum* transmission.

$$h(I) = 1 - \frac{1}{1 + e^{-\lambda(I - I_{50})}}$$
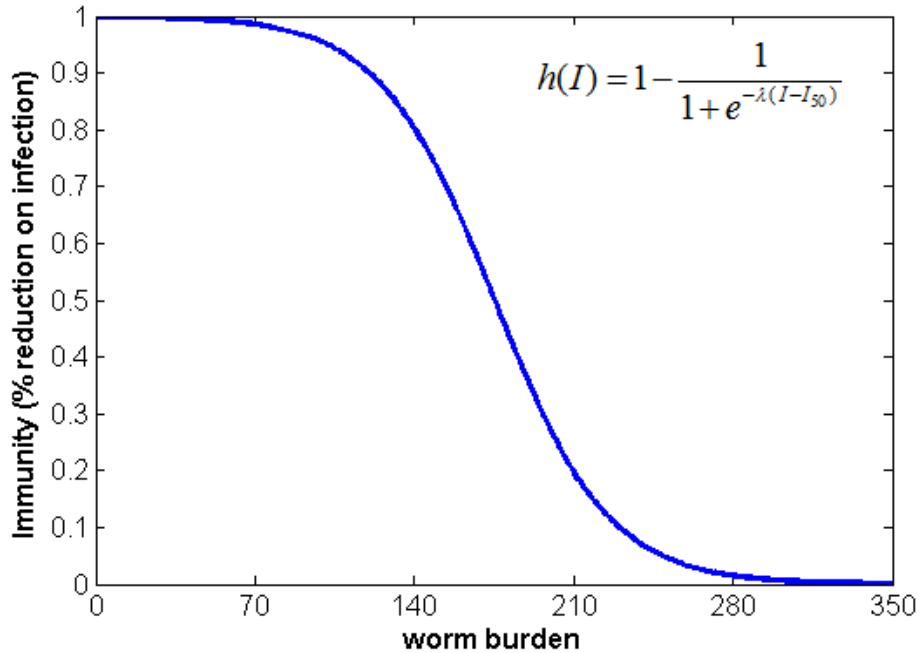
Figure 4.2 The logistic curve between individual acquired immunity *S. japonicum* [*h(I)*] and worm burden (*W*) at equilibrium. The immunity index, *I*, is calculated by 10 times the equilibrium worm burden level. I_{50} is the immunity index that leads to a 50% of infection reduction, which corresponds to the worm burden of 175 [25].

4.2.2 Linking community-based model to individually-based model

The modifications mentioned above lead to the focus on the worm development step in the community-based model, which is shown in Equation 4.1 [25]:

$$\frac{dw}{dt} = \alpha e^{-\mu_w \tau_w} s(t - \tau_w) \gamma C(t - \tau_w) h(I) f(w, \kappa, \gamma_w) - \mu_w w(t) \qquad \text{(Equation 4.1)}$$

The qualitative implication of this equation is straightforward: the rate of worm burden change with respect to time is determined by the rate of new worms developed less the rate of worm death. I will next interpret each parameter in the equation and explain whether or not it will be kept or modified in the individually-based model.

α is the number of parasites acquired per cercaria per m^2 skin water contact, which is used to represent the susceptibility to *S. japonicum* infection. In the community-based model, this parameter was randomly selected from a uniform distribution for each simulation, and was assumed to be invariant in the population. However, little was known regarding the likely magnitude of this parameter, hence it was set to have a 5000-fold range between 10^{-4} to 0.5 [16]. Further, using data from the two cohorts, some of our group's recent findings suggested that this parameter actually had a considerable variation between among different people. Thus, the magnitude and distribution of individual susceptibility, α_i, will be specifically explored and then utilized in the individually-based model.

γ is a spatial parameter that reflects the interaction between human water contact and the cercarial distribution, which was set to be spatially uniform within a village in the community-based model. In the individually-based model, however, cercarial density will be specifically associated with each water contact time interval (to be specified in Section 4.2.3.1). Therefore, the interaction term, γ, will be omitted.

$C(t - \tau_w)$ is the cercarial density at time $(t - \tau_w)$; τ_w is the period (days) required for cercariae to develop into adult worms in human hosts. The time delay corresponds to the fact that the number of worms developed at day t depends on the cercarial exposure at day $(t - \tau_w)$. For the individually-based model, given that neither infected snail data nor mouse bioassay data from cohort 2 was applicable, alternative methods have to be developed to estimate the distribution of cercariae density.

$s(t - \tau_w)$ is the averaged water contact index of the population at day $(t - \tau_w)$. Water contact will certainly be kept but reformed at the individual level and segmented into small intervals in the individually-based model, as will be introduced in Section 4.2.3.1.

$e^{-\mu_w \tau_w}$ is the fraction of worms surviving development in human. This parameter will be modified in the individually-based model to fit the individual-level simulations.

$h(I)$ is the immunity effect function which will be omitted in the individually-based model for the reasons addressed in section 4.2.1.

$f(w, \kappa, \gamma_w)$ is the density-dependent worm development function which describes the establishment of new worms in human hosts as restrained by current worm burden. However, given that most infected people would carry only a few – often times even

59

undetectable – worms in the low-transmission environment, this effect is unlikely to play an important role and thus omitted in the individually-based model.

Finally, μ_w is the worm mortality rate. In the individually-based model simulations, annual treatment of humans with the drug praziquantel is conducted at the end of each year assuming perfect compliance and efficacy. As the mean life span of schistosome worms is 5 years, worm death will not be specifically considered in each one-year simulation. However, note that μ_w is also a parameter in the survival fraction function, $e^{-\mu_w \tau_w}$.

In sum, derived from the community-based model, the worm development equation of the individually-based model needs to include the following components: water contact, cercarial density, individual susceptibility, and the survival proportion during development. A detailed structure of the individually-based model will be specified next.

4.2.3 Development of the individually-based model

Development of the individually-based model is broken into four steps corresponding to the actual processes of schistosome infection and detection: cercarial exposure, schistosome acquisition, worm development *in vivo*, and infection testing.

4.2.3.1 Cercarial exposure

The total number of <u>cercariae</u> which individual i acquires in one year, noted as C_i, is estimated by Equation 4.2 and 4.3:

$$C_i = \sum_{j=1}^{n_i} r_s s_i c_j \Delta t = r_s s_i \Delta t \sum_{j=1}^{n_i} c_j \qquad \text{(Equation 4.2)}$$

$$n_i = \frac{WC_i}{\Delta t} \qquad \text{(Equation 4.3)}$$

Where: s_i is the time weighted averaged body surface area (m^2) of all water contact activities of the ith individual; Δt is the number of minutes in each water contact interval, which is to be defined below; c_j is the time-weighted average cercarial density ($No./m^2$) associated with the ith individual's jth water contact interval; n_i is the total number of water contact intervals of the ith individual; r_s is the rate of cercarial acquisition (fraction/minute), i.e., of the total available cercariae that an individual could contact with in one minute given cercarial density and the individual's body surface area, what proportion actually hits the person's skin; finally, WC_i, as before, is the total number of water contact minutes of the ith individual in a year.

The water contact interval Δt is associated with the temporal distribution of cercariae. Referring to the mouse bioassay results shown in Section 3.5.4, cercarial density in water is assumed to be approximated by a stationary random process. Thus, the average density for data collected at any fixed sampling interval, Δt, is invariant, but the variance of this distribution varies and depends on the sample duration and the autocorrelation between sampling periods. Thus, human water contact is also segmented into Δt-minute intervals to accord with such variance. The concrete form and quantification of cercarial distribution will be discussed in Section 4.2.4.1 but note that

the cercariae density term in Equation 4.2, c_j, is a stochastic factor associated with each water contact interval, while Δt, r_s and s_i are all constants to be assigned prior to simulations.

4.2.3.2 Schistosome acquisition and worm development

Having the total number of cercarial hits estimated, the number of acquired schistosome is calculated by Equation 4.4:

$$J_i = \lfloor \alpha_i C_i \rfloor \qquad \text{(Equation 4.4)}$$

It is worth noting that individual susceptibility, α_i, is here defined as the proportion of the *annual total* exposed cercariae which actually penetrates people's skin and enter the circulation, rather than the cercariae associated with each water contact interval. Note also that only an integer number of schistosome acquisitions is meaningful for each individual.

The survival fraction $e^{-\mu_w \tau_w}$ is applied next to determine the number of schistosomes developing into adult worms. This term is treated as the probability of "succeeding" in a Bernoulli trail, each trail corresponding to the successful development of one schistosome. The rationale, again, is related to conducting simulations on the individual level with low-infections: if multiplied by this fraction and having the decimal part rounded towards zero again, infection with one worm will then have no infection by definition, and infection with a few worms will have none or much lower chance of pairing and producing eggs. Thus, each of the J_i schistosomes is simulated separately with stochasticity on its survival incorporated. Finally, the number of worms that individual *i* acquires *at the end of each year* is calculated by:

$$w_i = \sum_{J_i} B(1, e^{-\mu_w \tau_w}) \qquad \text{(Equation 4.5)}$$

4.2.3.3 Infection testing

The last element of the individually-based model relates to infection testing. As both the miracidial hatch test and the Kato-Katz test depend on the existence of schistosome eggs, the first step is to determine the number of single and paired worms. In order to do so, two assumptions are made: (1) each worm could be either female or male with equal probability; (2) any two unpaired worms of the opposite sex will pair. In the simulation, accordingly, the sex of each worm is randomly assigned upon development. The total number of worms, w_i, can therefore be expressed by the sum of single worms, $w_{i,s}$, and the double of the worm pairs noted as P_i:

$$w_i = w_{i,s} + 2P_i \qquad \text{(Equation 4.6)}$$

Clearly, egg production relies on the number of worm pairs, P_i. Previous studies of our group have shown that the number of eggs produced by one pair of worms per day, over time, follows a negative binomial distribution such that:

$$E_p \sim NB(\bar{E}_p, r_E) \qquad \text{(Equation 4.7)}$$

where $\bar{E}_p$ is the mean egg output in EPG and r_E is the aggregation parameter, a measure of the variability in egg output by a given worm pair [28, 29]. Assuming that egg output of each worm pair is independent (a reasonable assumption at low infection density) and letting E_{p1}, E_{p2},..., E_{pN} denote the EPG level corresponding to the 1st, 2nd, ... , nth worm pair, which are all randomly selected from the distribution shown in Equation 4.7, for an individual with P_i pairs of worms, the total EPG of individual i at any given day can be calculated by summing up the EPG which each pair "contributes":

$$E_i = \sum_{k=1}^{P_i} E_{pk} \qquad \text{(Equation 4.8)}$$

Finally, with individual EPG defined, the result of infection testing can be simulated. Data from previous infection tests allowed for summarizing the detection probability of the two methods at different EPG levels, i.e., of all people tested positive in either of the two methods, what proportion was found positive in the hatch and Kato-Katz method, respectively. It is worth noting that the detection probability defined here is not equivalent to the sensitivity of the method. Rather than comparing the result of either test to a "true" result from a gold standard, it is simply the combined result of the two tests that is used as the benchmark.
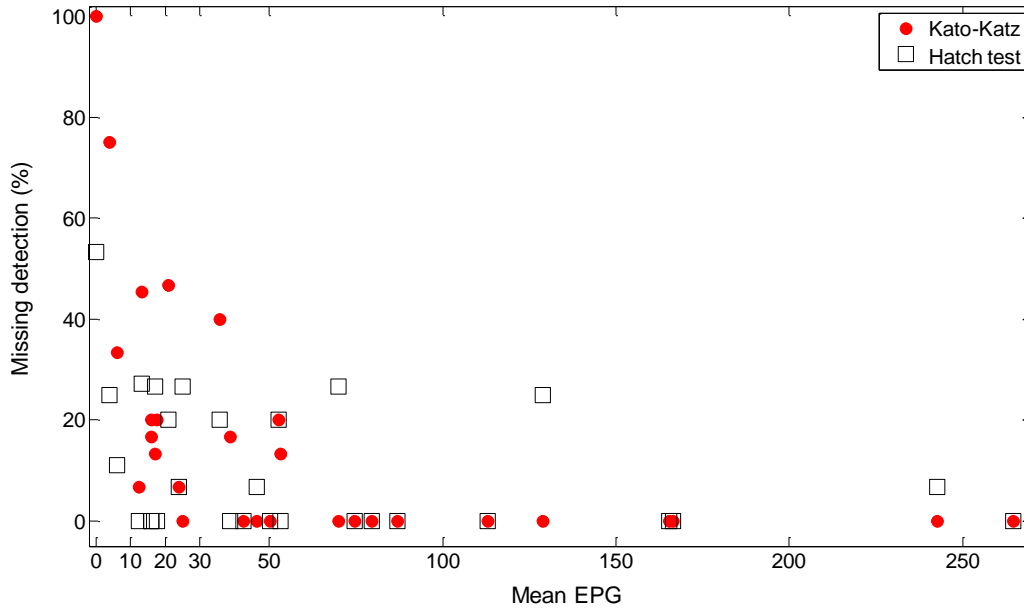
Figure 4.3 Percentage of individuals classified as uninfected by the Kato-Katz and miracidial hatch tests as a function of true mean EPG level

Shown in Figure 4.3 is the percentage of individuals classified as uninfected in both methods corresponding to different infection levels. At EPG levels greater than 150, the Kato-Katz test was able to detect all infection. While also having high detection

probability, the hatch test missed a few cases at a quite large infection level around 240. Recall, however, that the hatch test only detects "hatchable eggs" whereas the Kato-Katz test can detect non-hatchable eggs as well. At the lower end, the missing percentage of the hatch test does not seem to be related to EPG levels either positively or negatively, but rather fluctuates between 0% and 20% with only one exception being around 55% where the mean EPG is less than 1. On the other hand, the negative association between missing percentage of Kato-Katz and mean EPG levels was clear: this test method worked rather poorly when the mean EPG was very low but improved rapidly as the EPG got larger. Particularly, when the EPG was greater than 70 (still a quite large infection intensity level in the re-emerging environment), the Kato-Katz appeared to be sufficiently sensitive to detect all the infections. The patterns in these data are consistent with the nature of the Kato-Katz method, which examines only a tiny proportion of the sample and tends to miss an infection when the intensity is low. For the purposes of the individually-based model, these patterns are summarized by an exponential function to describe the detection probability of Kato-Katz vs. EPG, and for the hatch test, the probability is set to be constant at 0.8:

$$S_K(E_P) = 1 - \exp(-\frac{E_P}{12})$$

$$S_H = 0.8$$

(Equation 4.9)

Again, in the simulation, each infection test will be treated as a Bernoulli trail with S_K and S_H being the probabilities of observing a positive result using the respective methods. Note that there will be two types of infection status outcome in the simulation for each individual: (1) the "true" status defined as the existence/absence of worms, which can be determined directly by counting the number of worms, and (2) the "observed" status as the result of infection testing, which encloses a series of stochastic factors including the number of worm pairs, the EPG level at the day(s) when infection testing is conducted (simulated), and the detection probabilities of both methods.

4.2.4 Parameterization

Having developed the structure of the individually-based model, parameterization will be introduced next. Table 4.1 summarizes the values of parameters derived from either previous studies or survey data. For the four biological parameters, $\bar{E}_p$, r_E, μ_w and τ_w, relatively narrow ranges have been suggested and used in the community-based model. While their variances still add to the uncertainties of the simulation results, it is not a primary goal of this dissertation to explore the impact of these factors. Thus, the mean value or a point estimate of each parameter will be a constant in all simulations. For the parameter, r_s, the rate of cercariae acquisition, it is assumed to be 100% per minute considering the rapid penetration rate of cercariae, which can take as short as 10 seconds. Finally, the values of individual s_i and WC_i are directly estimated from survey data in the two cohorts and implicitly define their respective distributions, which will be specified in the next chapter detailing model simulations.

Other than the parameters listed in Table 4.1, there are two parameters whose values and distributions need to be specifically estimated for the individually-based

model: the cercarial density, c_j, and the individual susceptibility, α_i.

Table 4.1 Values of parameters derived from previous studies or survey data

Parameter	Interpretation and unit	Value	References
r_s	Rate of cercarial acquisition (percentage/minute)	100%	Assumption
$\bar{E}_p$	The average number of EPG produced by one pair of worm (EPG)	1.44	[28]
r_E	The aggregation parameter associated with the negative binomial distribution of egg production	0.5	[28]
μ_w	Worm mortality rate (percentage/day)	7.28×10^{-4}	[16, 23]
τ_w	Period required for cercariae to develop into worms in human hosts (days)	30	[16, 23]
s_i	Averaged body surface area of all water contact activities of the ith individual (m^2)	Varies	Survey data
WC_i	Total water contact minutes of the ith individual (min)	Varies	Survey data

4.2.4.1 Cercarial distribution

The contents of this section mainly refer to relevant notes from Prof. Robert Spear in a personal communication, except for Figure 4.4 which is adopted from [24].

Following the discussion in Section 4.2.3.1, the average cercarial density in each village appears to be relatively invariant for data collected at any fixed sampling interval (Δt) across the infection season. I will assume this average density can be described by a stationary random process. Hence, the variance of this distribution of a random sample collected over interval depends on Δt. Specifically, if a long-term sampling period ($N\Delta t$) is comprised of N short-term samples, the variance relationship is then:

$$\sigma_y^2 = \frac{\sigma_x^2}{N}[1 + \frac{2}{N}\sum_{T=1}^{N-1}(N-T)r_T] \qquad \text{(Equation 4.10)}$$

where σ_y^2 is the variance of samples collected in the total $N\Delta t$ period, which is 10 hours for a mouse bioassay as in our study in Xichang, σ_x^2 is the variance of samples sampled over time interval Δt, and r_T is the autocorrelation coefficient between different intervals. There is no data available on the likely autocorrelation structure; Spear argued that independence, even over short terms sampling intervals, is a reasonable assumption insofar as the density is the result of a set of independent infected snail inputs. Therefore, Equation 4.10 can be simplified to the following form:

$$\sigma_y^2 = \frac{\sigma_x^2}{N} \qquad \text{(Equation 4.11)}$$

Refer back to Section 4.2.3.1, and considering the probability of cercarial density changes in relatively short periods, Δt is set to be 30 minutes, which yields $N = 20$ to cover over the 10 hours. On the other hand, with the variance change given by Equation 4.11 for short-term exposures, it is clear that the degree of over-dispersion is extreme. To pursue the issue further, the path of least resistance is to make a distributional assumption, in this case that the cercarial density of any location in a village, over time, follows the negative binomial distribution:

$$c_j \sim NB(\overline{c}_v, k_c)$$ (Equation 4.12)

where $\overline{c}_v$ is the spatiotemporally-averaged cercarial density of each village, and k_c is the aggregation parameter corresponding to the skewness of the distribution. Table 3.5 in Chapter 3 allows for estimates of the mean and variance of the 10 hour data from various sites during the infection season, which yields the 30-minute k_c of 0.06. Note that while this parameter is estimated from all data in Table 3.5 and assumed to be invariant in all villages, $\overline{c}_v$ is obviously village-specific and needs to be estimated separately.

For Cohort 1 villages, the most straightforward means of estimating the village average annual cercarial density is to back calculate its values using data from the 2002 infection survey and the 2000 water contact survey. The method is based on the analysis presented in [24] and utilizes what is there termed the "averaged model", mainly the equation below which is also derived from the community-based model:

$$\frac{dw}{dt} = \alpha S_m \overline{\alpha}_{12} \gamma c(t)$$ (Equation 4.13)

where $\overline{\alpha}_{12}$ is the averaged time-varying effect and S_m is max water contact index. The result is presented in Table 4.2 for the 10 villages where water contact data was available; $\overline{c}_v$ of each village for both the years of 2001 and 2002 are estimated.

Table 4.2 Village-specific average annual cercarial density in 10 villages from Cohort 1 for the years of 2001 and 2002.

Village	Average annual cercarial density ($No./m^2$)	
	$\overline{c}_{v,01}$	$\overline{c}_{v,02}$
1	8	3
2	20	10
3	7	4
4	22	6
5	2	2
6	2	2
7	12	3
8	28	15
9	11	10
10	2	2

For Cohort 2 villages, due to the lack of information about infected snail density, the values estimated for Cohort 1 are used as ballpark estimates. Shown in Figure 3 of reference [24], which is shown here as Figure 4.4, mean worm burdens in humans less than 10 are generally associated with infected snail densities less than $0.06/m^2$ in the field observations and $0.15/m^2$ in community-based model simulations. Hence, where the mean worm burden is below 5 and no infected snails are observed (in a decent sample size), an upper bound on infected snail density is about $0.06/m^2$ simulated [24]. This density produces cercarial densities of about $8/m^2$, which is set to be the upper limit of $\overline{c}_v$ in Cohort 2 villages.
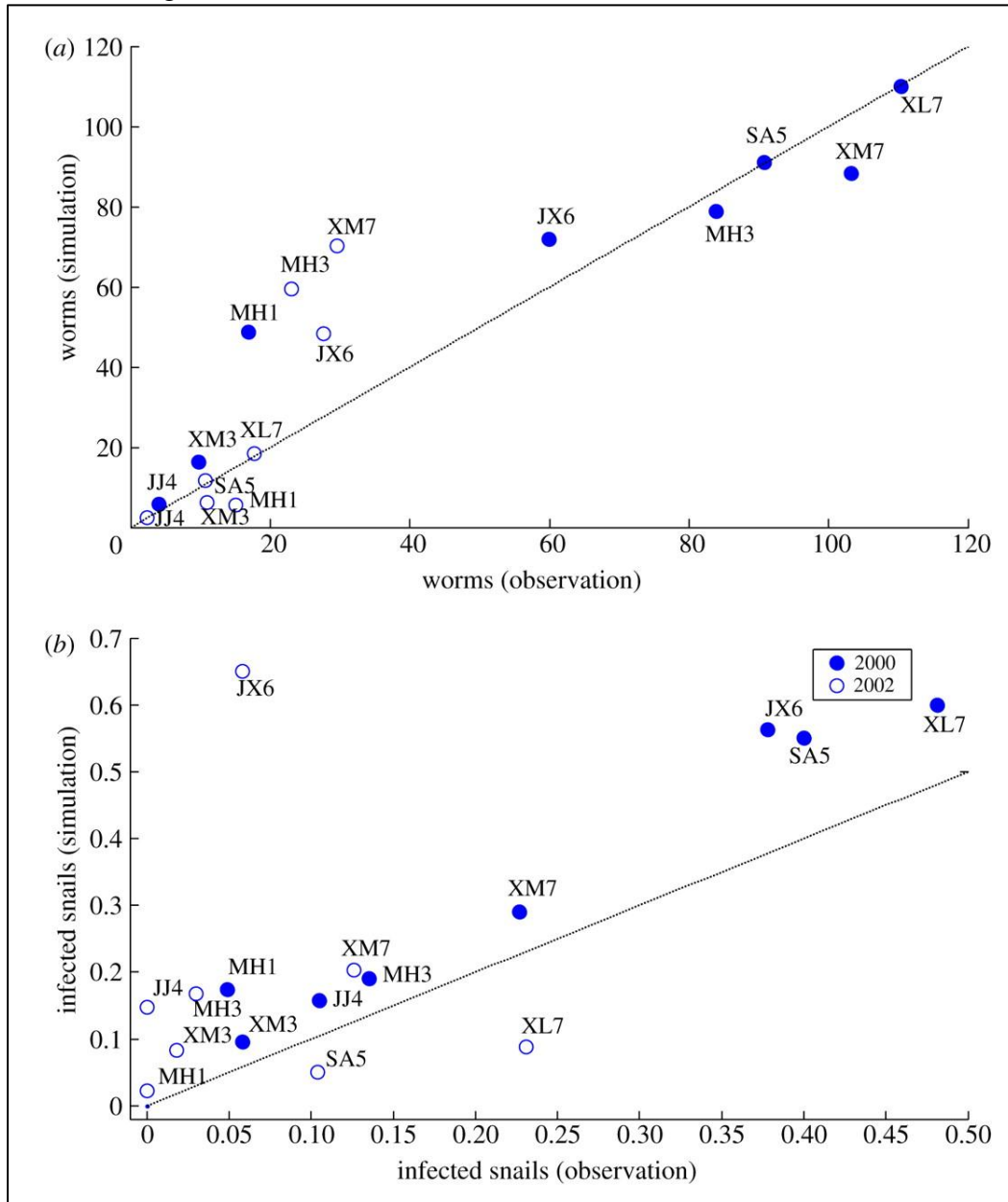

Figure 4.4 Year 2000 equilibrium levels and 2002 re-infection levels of worm burden and infected snail density predicted by the averaged model versus those simulated by the time-variable model (Figure was originally published in [24])

The next step is to decide what value to use for the mean cercarial density for each Cohort 2 village in the re-emerging environment. The most direct approach is to classify the villages by a combination of human and bovine prevalence. Setting the highest $\bar{c}_v$ to 8/m^2 and lowest $\bar{c}_v$ to 2/m^2, the two by two classification of high or low human and bovine prevalence is set to correspond to the annual average cercarial density of 2, 4, 6, 8 per m^2, respectively (Table 4.3).

Table 4.3 Village average cercariae/m^2 as a function of infection risk where no infected snails were detected.

		Bovine Prevalence	
		Low	High
Human Prevalence	Low	2	4
	High	6	8

The infection testing results of 2007 were used to generate the classification. Of the 28 villages where human infection was found (and also where water contact surveys were conducted), human prevalence varied between 1.5% and 42.9%. Accordingly, the cut point is set to be 10%, yielding 13 "high prevalence" and 15 "low prevalence" villages. Infected bovines were found in 15 of the 28 villages, with prevalence varying between 4% and 65.4%. Thus, villages having more than 15% of all bovines infected are categorized as "high prevalence", which includes 7 villages; the remained 21 villages, therefore, are "low bovine prevalence" villages. Finally, $\bar{c}_v$ of each village in Cohort 2 was determined using Table 4.3. The number of villages with each annual average snail density is shown in Figure 4.5.
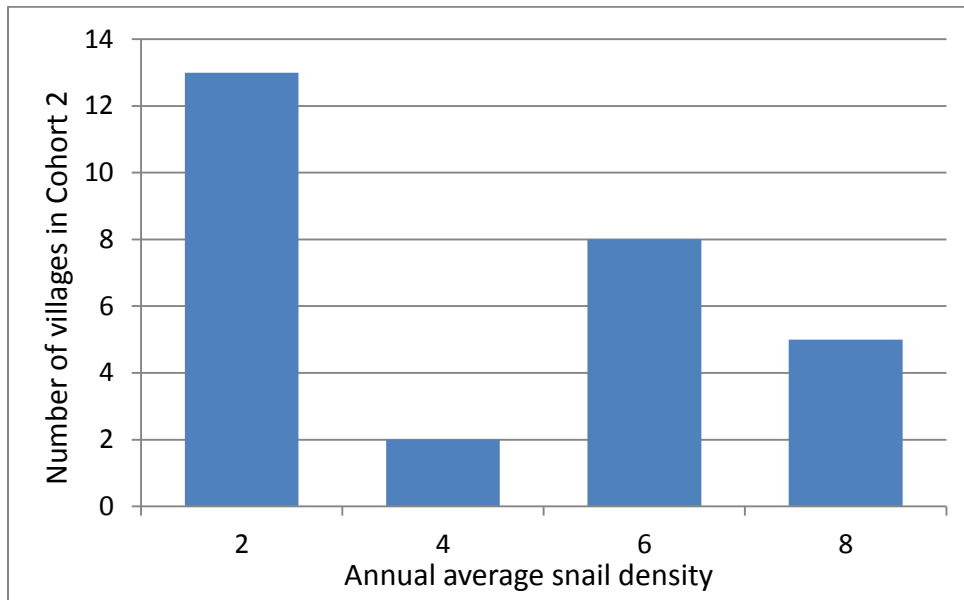
Figure 4.5 Number of villages with the annual average snail density of 2, 4, 6, 8 per m^2 in the 28 villages of Cohort 2

4.2.4.2   Individual susceptibility

The last part of parameterization regards individual susceptibility, α_i. Since little is known about this biological parameter from our previous work, the goal herein, instead of finding the "correct" values, is rather to explore its likely distribution among the population, and make the best estimates of its magnitude using currently available data. With all other parameters specified already, the individually-based model can be used for calculating the values of α_i. For this purpose, the model is temporarily modified to the following form:

$$E_i = \alpha_i r_s s_i \Delta t \sum_{j=1}^{n_i} c_j$$

(Equation 4.14)

Compared with the original individually-based model, the main change in Equation 4.14 is to use EPG as a substitute of worm burden, which is used to estimate α_i for two reasons. First, although individual worm burden can be explicitly calculated in model simulations, it cannot be directly estimated in infection tests. Thus, there is no alternative infection intensity outcome other than EPG that can be used for the calculation. Second, recall that on average, one pair of worms can produce 1.44 EPG per day. Given this association, the number of worms and the corresponding EPG level are expected to be of the same magnitude and numerically close. Then, re-writing Equation 4.14 gives the equation to be used to calculate α_i:

$$\alpha_i = \frac{E_i}{r_s s_i \Delta t} \cdot \frac{1}{\sum_{j=1}^{n_i} c_j}$$

(Equation 4.15)

In Equation 4.15, the first term $E_i / r_s s_i \Delta t$ is completely deterministic, so the core component of simulation is to randomly select the cercarial density associated with each water contact interval. Monte-Carlo simulation was used to take the stochasticity into account. Individual data collected from the two cohorts was used separately for the estimation, and only individuals with both E_i and n_i greater than zero were included for obvious reasons. The algorithm for each simulation was as follows:
(1)   For each individual, year-to-year variation of his/her total water contact was assumed to vary between 0.8 and 1.2 times his/her original annual n_i to yield the number of water contact intervals.
(2)   A random sample was drawn from the cercarial density associated with each interval from the negative binomial distribution in Equation 4.12. In cohort 1, each simulation is conducted on a two-year basis, with the first and the second year's $\bar{c}_v$ of each village (where the individual resides) selected from Table 4.2. In cohort 2, the simulation duration was one year, and $\bar{c}_v$ of each village is pre-determined based on the criteria in Table 4.3.
(3)   Calculate the individual susceptibility from each simulation, which is noted as

68

$\alpha_{i,k}$ for the *k*th simulation.

The model is coded in MATLAB$^©$ (The MathWorks Inc., Natick, Massachusetts, USA), and a total of 5000 simulations were run for each cohort. In Cohort 1, α_i was calculated for 108 out of 470 people; in cohort 2, only 24 out of 608 people had EPG values greater than zero and were therefore available to be used for the simulation. Shown in Figure 4.6 are the histograms of the 5000 results for four individuals in Cohort 1. It can be seen that, mainly depending on the number n_i, the distribution of $\alpha_{i,k}$ for each individual has quite diverse patterns, varying from highly right-skewed (top left) to nearly normal (bottom right). Furthermore, when n_i is relatively small, the total number of cercarial hits will be zero in many simulations, which leads to invalid estimate of $\alpha_{i,k}$. To account for all simulation results, for each individual, the medium of the 5000 $\alpha_{i,k}$ was adopted for the point estimate of individual α_i.
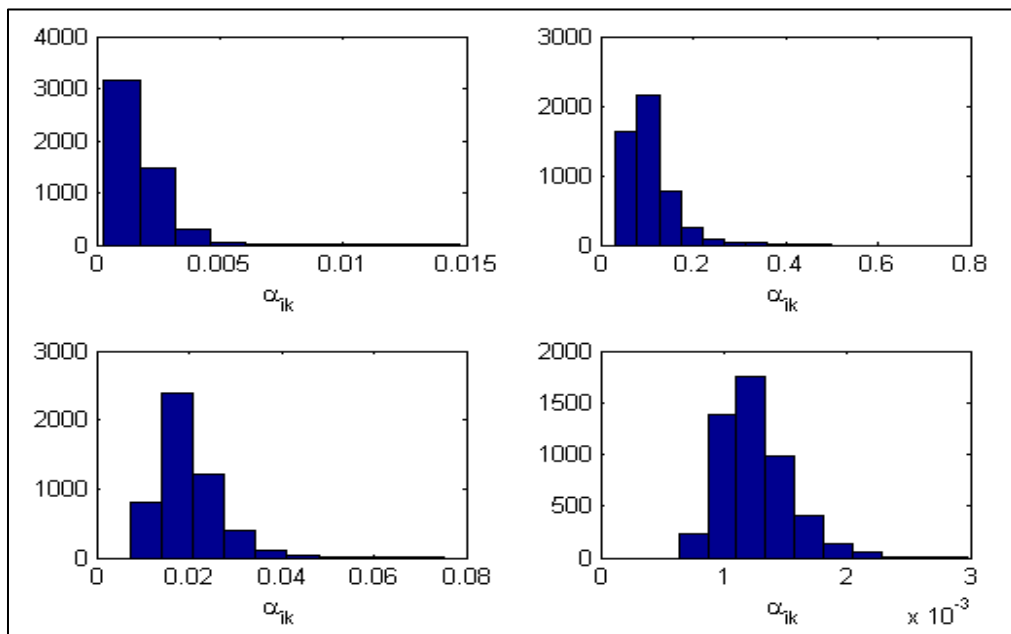
Figure 4.6 Distribution of 5000 simulated $\alpha_{i,k}$ for four individuals from Cohort 1

The result of the 132 estimated α_i is shown in Figure 4.7 (note the logarithmic scale of the y-axis). Highly consistent patterns in terms of both the magnitude and distribution can be observed between the two cohorts: the values of both cohorts ranged between 10^{-4} and 10^{-2} with similar lower and upper limits, and while many fewer points were available from Cohort 2, they were distributed along the combined curve fairly evenly. Thus, it is reasonable to conclude that the estimates from the two cohorts are likely to be derived from the same distribution. This turns out to be also a biologically meaningful observation regarding individual susceptibility among people living in the similar environments.
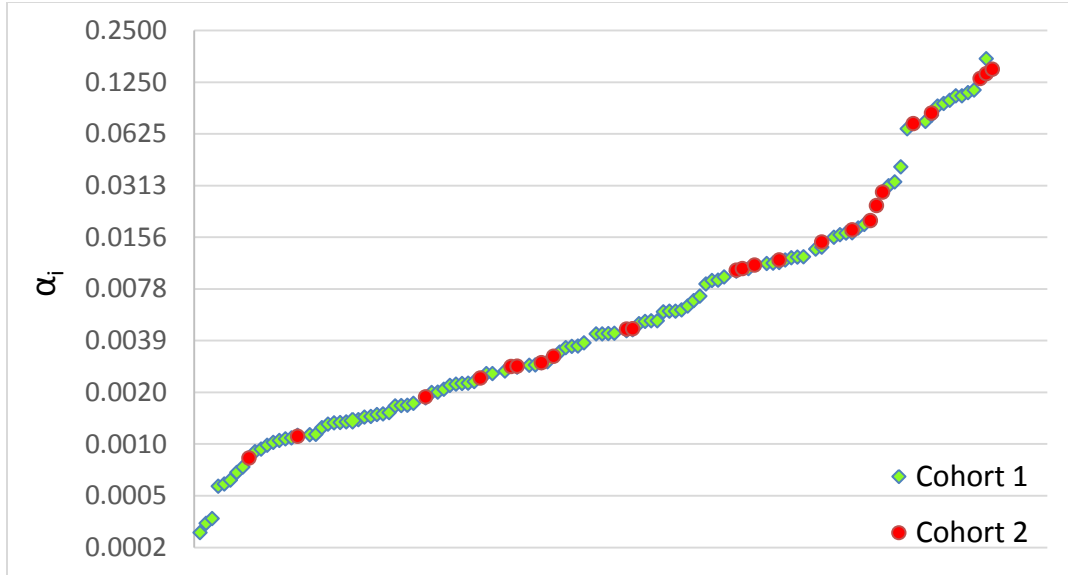
**Figure 4.7** Simulated α_i for 108 people in Cohort 1 (green) and 24 people in Cohort 2 (red)

The final step is to generate the overall distribution which will be assumed to apply to the whole population. Compared with the rest of the people in Cohort 1, the mean water exposure minutes of the 108 people was only 6% greater, and the difference was insignificant (p=0.54). Moreover, the 24 infected people in Cohort 2, on average, had only less than half of the exposure minutes than the other 584 people. Therefore, the 132 people included in the α_i calculation are convincingly the much more susceptible group, and, considering the majority who had much exposure but no detectable EPG, the overall distribution of individual susceptibility appears to be highly right-skewed. Thus, a log-normal distribution is assumed to describe the pattern of the population:

$$\alpha_i \sim \ln(GM_\alpha, GSD_\alpha) \qquad \text{(Equation 4.16)}$$

Combining the two cohorts together, the smallest estimated α_i, 2.98×10^{-4}, corresponds to the percentile of 88%. Thus, a series of log-normal curves with different combinations of GM (geometric mean) and GSD (geometric standard deviation) can be fit to the data. As GSD is the parameter to determine the variance of the distribution, three levels of GSD, respectively 1.5, 3.0 and 4.5, are used to represent the low, medium and high variance. The GM associated with each GSD can be calculated by:

$$\log(GM_\alpha) = \log(2.98 \times 10^{-4}) - Z_{88\%} \cdot \log(GSD_\alpha) \qquad \text{(Equation 4.17)}$$

Finally, the cumulative density curves of the three log-normal distributions are shown in Figure 4.8.
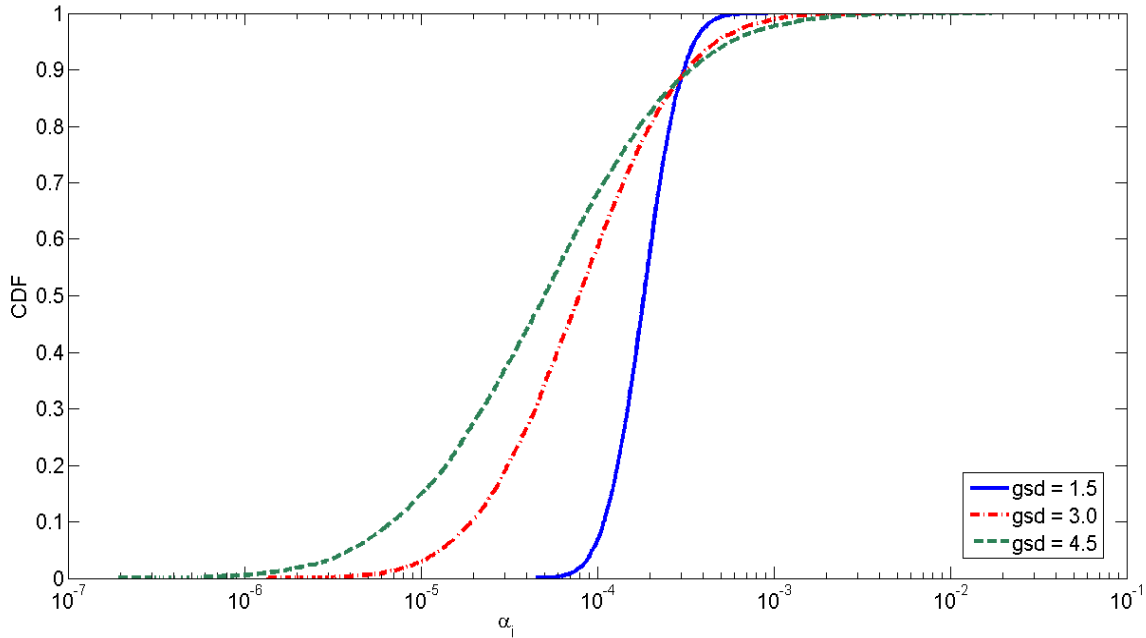
Figure 4.8 Distribution of α_i corresponding to GSD of 1.5, 3.0 and 4.5

4.3   Discussion

Following the introduction on survey protocols and data collection in the two cohorts, development of the individually-based model was specified in this chapter. The model was designed to simulate three components on the individual level: cercarial acquisition, worm development and infection testing. Parameterization was conducted particularly to two variables: the village-specific distribution of cercarial density and the individual susceptibility to schistosome, with the latter problem, which had not been explored in our former researches, being the first application of the individually-based model. The result suggested a highly right-skewed, log-normal shaped distribution of susceptibility, which confirmed our previous inference from a machine-learning model that, instead of invariant, this biological parameter was more likely to be varied among the population. Having the model set up, I will focus on the simulations in the next chapter and present how the results are used for interpreting epidemiological findings and making implications on disease control.

While the individually-based model was developed tailored to low-transmission environments and made maximum use of the available data, it was limited by the incompetency of current infection testing and environment monitoring techniques. Regarding parameterization, for instance, it has been shown in Figure 4.3 that Kato-Katz was not an effective method when the infection intensity was low. Therefore, the number of people whose data was actually incorporated to estimate individual susceptibility in cohort 2 was much less than the total number of people tested positive, which significantly reduced the sample size. But compared with susceptibility, the biological parameter that was not location or cohort specific, there were even more uncertainties regarding the estimation of cercarial density in cohort 2 villages due to unavailability of either snail infection or mouse bioassay data. However acceptable in this context, the method proposed and applied in this dissertation, which was to arbitrarily assign four

71

different levels to all villages based on their human and bovine infection prevalence, was far from being ideal. Thus, as also discussed elsewhere [24], a much more sensitive, robust and convenient method of detecting parasites in the surface water is of substantial need. Presumably, not only will the new technique be able to provide more informational data which can better facilitate model parameterization, more importantly, it can also identify the "hot spots" of parasite residence, which is crucial to the prevention and control of disease transmission.

On the other hand, further modification on the model structure is also subject to improvement on environmental monitoring. If new methods are developed such that the cercariae can be directly quantified, depending on the frequency of such measurements, a series of much more site-specific, time-varying patterns of cercarial density can be developed. Human water contact behaviors can be then segmented into intervals or events as appropriate to match the cercariae monitoring results. As an example, assuming that cercarial density data is available on a daily basis, the current individually-based model can be modified to an event-driven structure to be specified as follows.

Adopted from the previous chapter, Table 4.4 shows again part of a hypothetical water contact survey result. With this individual profile, an "event list" can be generated by reorganizing all reported water contact behaviors. Table 4.5 gives the example of the event list derived from the first row of Table 4.4 (the Italic, bold row corresponding to the first location). For instance, this individual washes clothes at Location F7 for 30 minutes twice in May, so two "events" should be generated. As the specific dates of each activity were not asked in the questionnaire due to recall difficulties, they will be selected using non-repeated random number generation methods. In this hypothetical example, the two events occur on the 5th and 25th day of May, respectively. These dates are further transferred to the "day of year" for simulation conveniences. Repeating this process for all types of activities at all locations and months, the full event list of an individual can be generated, and the total number of events, V_i should therefore be $\sum_m \sum_a \sum_l Freq$.

Table 4.4 Hypothetical water contact survey result (partial)

ID	Activity	Location		Frequency	Duration	Flow velocity
110103001001	Washing clothes	*1*	*F5*	*2*	*30*	*Fast*
		2	J9	3	20	Slow
		3	D10	5	15	Moderate

Table 4.5 Example of the event list using partial data from Table 4.4

ID	Activity	Month	**Day of month**	Day of year	Location	Duration (min)
110103001001	W/ clothes	5	**5**	125	F5	30
110103001001	W/ clothes	5	**25**	145	F5	30

Finally, the events of all people are sorted by date. In the simulation, one event will be taken out each time and, based on the date and location where this event occurs, a cercarial density can be determined from monitoring data and used for calculating the cercarial exposure. Summarizing the results from all events of one individual yields

his/her total cercarial hits, and the following steps will be the same as the current individually-based model. If proved necessary, a distribution can still be assumed for cercarial density change within one day, and each event can again be segmented into smaller intervals. However, an improved cercariae detection method will certainly allow for better and more specific estimations on the "averaged" density level, which is why it is the prerequisite of any forms of model modification in the future.

References

1. Grassly, N.C. and Fraser, C., *Mathematical models of infectious disease transmission.* Nat Rev Microbiol, 2008. **6**(6): p. 477-87.
2. Anderson, R.M. and May, R.M., *Infectious diseases of humans : dynamics and control.* Oxford science publications. 1991, Oxford; New York: Oxford University Press.
3. Gambhir, M. and Michael, E., *Complex ecological dynamics and eradicability of the vector borne macroparasitic disease, lymphatic filariasis.* PloS one, 2008. **3**(8): p. e2874.
4. Michael, E., Simonsen, P.E., Malecela, M., Jaoko, W.G., Pedersen, E.M., Mukoko, D., Rwegoshora, R.T., and Meyrowitsch, D.W., *Transmission intensity and the immunoepidemiology of bancroftian filariasis in East Africa.* Parasite immunology, 2001. **23**(7): p. 373-388.
5. Anderson, R.M. and May, R.M., *Helminth infections of humans: mathematical models, population dynamics, and control.* Advances in Parasitology, 1985. **24**: p. 1-101.
6. Castillo-Chavez, C., Feng, Z., and Xu, D., *A schistosomiasis model with mating structure and time delay.* Mathematical biosciences, 2008. **211**(2): p. 333-341.
7. Chan, M.S. and Bundy, D.A., *Modelling the dynamic effects of community chemotherapy on patterns of morbidity due to Schistosoma mansoni.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 1997. **91**(2): p. 216-220.
8. Feng, Z., Li, C.C., and Milner, F.A., *Schistosomiasis models with density dependence and age of infection in snail dynamics.* Mathematical biosciences, 2002. **177-178**: p. 271-286.
9. Finkelstein, J.L., Schleinitz, M.D., Carabin, H., and McGarvey, S.T., *Decision-model estimation of the age-specific disability weight for schistosomiasis japonica: a systematic review of the literature.* PLoS neglected tropical diseases, 2008. **2**(3): p. e158.
10. Ishikawa, H. and Ohmae, H., *Modeling the dynamics and control of transmission of Schistosoma japonicum and S. mekongi in Southeast Asia.* The Korean journal of parasitology, 2009. **47**(1): p. 1-5.
11. Medley, G.F. and Bundy, D.A., *Dynamic modeling of epidemiologic patterns of schistosomiasis morbidity.* The American Journal of Tropical Medicine and Hygiene, 1996. **55**(5 Suppl): p. 149-158.
12. Milner, F.A. and Zhao, R., *A deterministic model of schistosomiasis with spatial structure.* Mathematical biosciences and engineering : MBE, 2008. **5**(3): p. 505-522.
13. Woolhouse, M.E., *Mathematical models of transmission dynamics and control of schistosomiasis.* The American Journal of Tropical Medicine and Hygiene, 1996. **55**(5 Suppl): p. 144-148.
14. Macdonald, G., *The dynamics of helminth infections, with special reference to schistosomes.* Trans R Soc Trop Med Hyg, 1965. **59**(5): p. 489-506.
15. Liang, S., *A spatio-temporal modeling of schistosomiasis transmission dynamics and control in Sichuan, China*, 2003.
16. Liang, S., Seto, E.Y., Remais, J.V., Zhong, B., Yang, C., Hubbard, A., Davis, G.M., Gu, X., Qiu, D., and Spear, R.C., *Environmental effects on parasitic disease transmission exemplified by schistosomiasis in western China.* Proceedings of the

National Academy of Sciences of the United States of America, 2007. **104**(17): p. 7110-7115.

17. Liang, S. and Spear, R.C., *Model-based insights into multi-host transmission and control of schistosomiasis.* PLoS medicine, 2008. **5**(1): p. e23.

18. Liang, S., Spear, R.C., Seto, E., Hubbard, A., and Qiu, D., *A multi-group model of Schistosoma japonicum transmission dynamics and control: model calibration and control prediction.* Tropical medicine & international health : TM & IH, 2005. **10**(3): p. 263-278.

19. Maszle, D.R., *Dynamic modeling for the control of schistosomiasis in China in light of parametric uncertainty*, 1998.

20. Seto, E.Y. and Carlton, E.J., *Disease transmission models for public health decision-making: designing intervention strategies for Schistosoma japonicum.* Advances in Experimental Medicine and Biology, 2010. **673**: p. 172-183.

21. Spear, R.C., Hubbard, A., Liang, S., and Seto, E., *Disease transmission models for public health decision making: toward an approach for designing intervention strategies for Schistosomiasis japonica.* Environmental health perspectives, 2002. **110**(9): p. 907-915.

22. Spear, R.C., Seto, E., Liang, S., Birkner, M., Hubbard, A., Qiu, D., Yang, C., Zhong, B., Xu, F., Gu, X., and Davis, G.M., *Factors influencing the transmission of Schistosoma japonicum in the mountains of Sichuan Province of China.* The American Journal of Tropical Medicine and Hygiene, 2004. **70**(1): p. 48-56.

23. Wang, S., *Exploring the Impact of Infection-induced Immunity on Endemic Levels of Schistosoma japonicum in Hilly and Mountainous Environments in China*, 2009.

24. Spear, R.C., *The Role of Environmental Determinants of R0 in Suppressing Re-emergent Transmission of S.japonicum in China.* Journal of the Royal Society, Interface, 2011.

25. Wang, S. and Spear, R., *Exploring the impact of infection-induced immunity on the transmission of Schistosoma japonicum in hilly and mountainous environments in China.* Acta Tropica (under review), 2013.

26. Bonabeau, E., *Agent-based modeling: methods and techniques for simulating human systems.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99 Suppl 3**: p. 7280-7287.

27. Burke, D.S., Epstein, J.M., Cummings, D.A., Parker, J.I., Cline, K.C., Singa, R.M., and Chakravarty, S., *Individual-based computational modeling of smallpox epidemic control strategies.* Academic emergency medicine: official journal of the Society for Academic Emergency Medicine, 2006. **13**(11): p. 1142-1149.

28. Hubbard, A., Liang, S., Maszle, D., Qiu, D., Gu, X., and Spear, R.C., *Estimating the distribution of worm burden and egg excretion of Schistosoma japonicum by risk group in Sichuan Province, China.* Parasitology, 2002. **125**(Pt 3): p. 221-231.

29. Fung, M.S., Xiao, N., Wang, S., and Carlton, E.J., *Field evaluation of a PCR test for Schistosoma japonicum egg detection in low-prevalence regions of China.* Am J Trop Med Hyg, 2012. **87**(6): p. 1053-8.

Appendix to Chapter 4 – Core Matlab codes for the model development

```
load ('result_sum_combined_alpha.mat');

loggm_merged = result_sum(1,:);
loggsd = result_sum(2,:);
gsd_merged = result_sum(3,:);

load ('REC_we_ori_dbled.mat');  % "real" water contact measured by total min

meanwe = mean(REC_we_ori_dbled);
varwe = var(REC_we_ori_dbled);
kwe = meanwe^2 / (varwe - meanwe);  % get the parameters of the NB distribution for
water exposure
pwe = 1 / (1 + meanwe/kwe);

n = length(REC_we_ori_dbled);   %number of people (real data and the ghost data)

testcut = 1;  % cut point due to current infection testing methods

mu_w = 0.00072839;    %No death assumed, mainly for the exp(-mu_w*tau_w) term

tau_w = 30.0;
surv = exp(-mu_w*tau_w);    %survival rate during development

simurep = 2000; % simulation time
meanepg = 1.44; % epg associated with one worm pair per day
kepg = 0.5; %aggregation paramter
pepg = 1/(1+meanepg/kepg);


for i = 1:simurep   % time of simulation

   % to assign the individual s_i for the ghost population

   load ('si_pool_rec.mat');

      for people = 1:n;
      randnum(people) = rand();
      roww = max(find(randnum(people)>si_pool(:,2)));
      si_ghost(people,1) = si_pool(roww+1,1)-[si_pool(roww+1,1)-
si_pool(roww,1)]*[si_pool(roww+1,2)-randnum(people)]/[si_pool(roww+1,2)-
si_pool(roww,2)];
      end;

   load ('village_REC');
```

```matlab
    load ('cbar_REC.mat');  % the list of individual(village-level) cbar

% to create the ghost population

    interval = 30; % 30 minutes per interval

    ghostprofile(:,1) = nbinrnd(kwe, pwe, n, 1);    % assign the water contact in total
minutes (again, doubled "real" minutes)

    ghostprofile(:,2) = ceil(ghostprofile(:,1)./interval);    % calculate the n_i

    ghostprofile(:,3) = si_ghost;                   % assign the s_i
    ghostprofile(:,4) = cbar_REC;
    ghostprofile(:,5) = village;

    gsdtesttotal = 3;
    for gsdtest = 1:gsdtesttotal
        gfile{gsdtest}(:,1:5) = ghostprofile;
        posipre = abs(loggsd-log(gsdtest*1.5));
        position = find(posipre == min(posipre));
        gfile{gsdtest}(:,6) = lognrnd(loggm_merged(position), loggsd(position), n, 1);   %
assign the alpha_i
        alpha_i{gsdtest} = gfile{gsdtest}(:,6);
        gfile{gsdtest} = sortrows(gfile{gsdtest},1);
    end
    n_i = gfile{gsdtest}(:,2);
    s_i = gfile{gsdtest}(:,3);      % after sorting by WATER exposure
    cbar_i = gfile{gsdtest}(:,4);

    n_iall{gsdtest}(:,i) = n_i;
    s_iall{gsdtest}(:,i) = s_i;
    cbar_iall{gsdtest}(:,i) = cbar_i;
    alpha_iall{gsdtest}(:,i) = alpha_i{gsdtest};
    village_iall{gsdtest}(:,i) = gfile{gsdtest}(:,5);

    kc = 0.04;
    pc = 1./(1+cbar_i.*.5./kc);

        % the simulation

    for gsdtest = 1:3   % three alpha_i distributions
    for year = 1:2  % to get re-inf probability
        for people = 1:n;
            cdins = nbinrnd(kc, pc(people), n_i(people), 1);    % simulate the number of
cercarial hits
```

sch_raw = alpha_i{gsdtest}(people) .* sum(cdins) .* s_i(people) .* interval;   %number of schistosoma acquired from cercarial hits

sch_int = fix(sch_raw); % only integer is allowed

sch_adj = random('binomial',sch_int,surv); % number of survived schistosome; for each schistosome, its survival probability follows the Bernoulli traial with success probability: exp(-mu_w*tau_w)

sch_indtotal{gsdtest}(people,2*(i-1)+year) = sum(sch_adj); % year-end worm burden

nmale = binornd(sch_indtotal{gsdtest}(people,2*(i-1)+year),0.5); % number of male worms

nfemale = sch_indtotal{gsdtest}(people,2*(i-1)+year) - nmale; % number of female worms

npair{gsdtest}(people,2*(i-1)+year) = min(nmale, nfemale); % number of worm pairs

npairtem = npair{gsdtest}(people,2*(i-1)+year);

%for infection testing:
if npairtem > 0
    epg{gsdtest}(people,2*(i-1)+year) = mean(sum(nbinrnd(kepg, pepg, npairtem, 3)));
else
    epg{gsdtest}(people,2*(i-1)+year) = 0;
end
kktest{gsdtest}(people,2*(i-1)+year) = (epg{gsdtest}(people,2*(i-1)+year)>0) * (rand() > exp(-1/12*epg{gsdtest}(people,2*(i-1)+year))); % result of kato-katz

hatch {gsdtest}(people,2*(i-1)+year) = (epg{gsdtest}(people,2*(i-1)+year)>0) * (rand() > 0.2); % result of hatch

test = max(kktest{gsdtest}(people,2*(i-1)+year), hatch{gsdtest}(people,2*(i-1)+year)); % test result: max of (kato-katz, hatch)

testres{gsdtest}(people,2*(i-1)+year) = test; % record the test result

ctotal{gsdtest}(people,2*(i-1)+year) = sum(cdins); % record total number of cercarial hits
        end

        % extra result summarization
        Etemp{gsdtest}(:,2*(i-1)+year) = s_i.* ctotal{gsdtest}(:,2*(i-1)+year);   % total E
        hnotkk{gsdtest}(:,2*(i-1)+year) = hatch{gsdtest}(:,2*(i-1)+year) > kktest{gsdtest}(:,2*(i-1)+year);   % hatch positive but kato-katz negative
        kknoth{gsdtest}(:,2*(i-1)+year) = hatch{gsdtest}(:,2*(i-1)+year) < kktest{gsdtest}(:,2*(i-1)+year);   % kato-katz positive but hatch negative
     end
  end
end

Chapter 5 Simulations of the Individually-based Model

"We can only see a short distance ahead, but we can see plenty there that needs to be done."

--Alan Turing

With the development and parameterization of the individually-based model, I present the simulation protocols and results in this chapter. The simulations described here are designed for two purposes: to phenomenologically explain the epidemiological findings, and to explore the significance of individual susceptibility on risk identification and disease control. Thus, instead of simulating the two original cohorts, the survey data will be used to generate more generic scenarios to be incorporated in the simulations.

5.1 The hypothetical populations

Previous chapters have shown that the two cohorts showed considerable variability in water contact as well as in village-specific cercarial density. These factors, together with individual susceptibility which will be assumed site invariant and, as a result, to be described by the same distribution for both cohorts, yield two transmission scenarios that we expect to result in medium and low mean infection intensities. But, is that a reasonable expectation? That is, given the residual uncertainty and variability in the individual parameters, is the low/medium transmission level an emergent property at the community level or might different or mixed patterns of community response be observed. To investigate this issue, two hypothetical populations (noted as HP1 and HP2) corresponding to the two original cohorts are created such that:

(1)   The number of people in HP1 and HP2 follows the original data. There were 529 people from 10 villages in Cohort 1 and 727 people from 28 villages in Cohort 2, from which the respective water contact distributions were estimated. In the simulation, the number of villages in each HP and the number of people in each village will be exactly the same as in the epidemiological data (Table 5.1 and Table 5.2).

(2)   The annual mean cercarial density of each village is as estimated in Chapter 4 (Table 5.1 and Table 5.2). For each hypothetical individual, his residential village is randomly assigned and so is the corresponding $\bar{c}_v$. Note again that for HP1, the annual cercarial density was estimated separately for two consecutive years. Accordingly, each simulation will be conducted on a two-year basis. For HP2, on the other hand, the simulation interval is one year.

Table 5.1 The number of people and the average annual cercarial density of both years of the 10 hypothetical villages in HP1

| Village	No. of people for simulation	Average annual cercarial density ($No./m^2$)	
		Y_1	Y_2
1	62	8	3
2	40	20	10
3	44	7	4
4	59	22	6
5	58	2	2
6	55	2	2
7	61	12	3
8	55	28	15
9	42	11	10
10	53	2	2

Table 5.2 The number of people and the average annual cercarial density of the 28 hypothetical villages in HP2

Village	No. of people in simulations	Average annual cercarial density	Village	No. of people for simulation	Average annual cercarial density
1	53	8	15	46	2
2	45	6	16	26	2
3	11	6	17	17	2
4	18	6	18	12	4
5	5	8	19	6	2
6	18	8	20	17	2
7	6	6	21	29	2
8	10	6	22	69	2
9	11	6	23	33	2
10	52	8	24	17	2
11	53	6	25	3	2
12	16	8	26	42	2
13	34	6	27	23	2
14	43	4	28	12	2

(3)　　Prior to running the simulations, the values of three other parameters also need to be specified for each individual: individual susceptibility (α_i), the water contact minutes (or equivalently, n_i), and the mean body surface area (s_i). The α_i of each individual is randomly selected from the corresponding distribution shown in Figure 4.8. Derived from the survey data (also shown in the top plots of Figures 3.4 and 3.8) is a negative binomial distribution that is parameterized for each cohort to describe the total number of water contact minutes. The value of this variable for each hypothetical individual is therefore randomly chosen from the respective distribution. Finally, due to the unique patterns of s_i shown in the bottom plots of Figures 3.4 and 3.8, no

distributional types are arbitrarily assumed. Instead, an empirical cumulative density curve is generated based on the originally calculated body surface area of each cohort (Figure 5.1), from which a random number will be selected and assigned to each hypothetical individual.
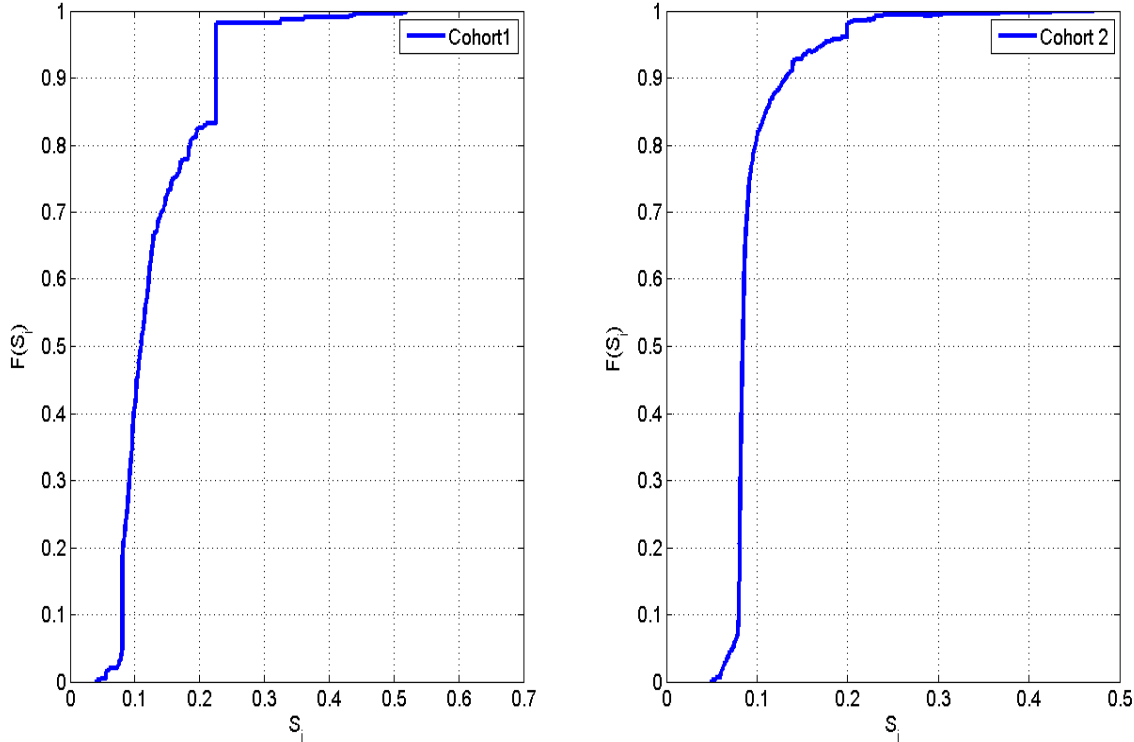
Figure 5.1 The empirical cumulative density curve of s_i of Cohort 1 and Cohort 2

5.2   Simulation protocols

Two series of simulations were conducted to comprehensively account for stochasticity from various sources. The first series (Series I) mimics transmission among the same group of individuals over multiple years. So the simulation results reflect the fluctuations of water contact and random variability in cercarial exposure from year to year. Hence, each individual's α_i, s_i and $\bar{c}_v$ (or equivalently, the village in which this individual resides) are fixed in all simulations once randomly assigned. For water exposure, it is assumed that the total number of 30-minute intervals randomly varies between 0.8 and 1.2 of the originally selected n_i (the same assumption as that earlier used in simulations to estimate α_i). For the second series (Series II), on the other hand, each individual's exposure profile is randomly re-assigned in every simulation, so a new hypothetical population is created each time. The main goal of conducting Series II simulations is to confirm that the patterns observed in the Series I simulations are not due to chance, i.e., they do not apply to only one specific population at one particular time, but can represent more generic scenarios. Both series of simulations were conducted 2000 times for each cohort. The results are summarized around three bullet points derived from the epidemiological findings, plus one bullet point suggesting a new approach to facilitating disease monitoring in low-transmission environments.

81

5.3   The bullet points

5.3.1   Infection intensity distribution

The first bullet point regards the distribution of infection intensity among the hypothetical populations. Our epidemiological data from previous infection surveys as well as that of many other investigators has shown that EPG is negative binomially distributed for essentially all the transmission levels and exposure durations [1-9]. To verify this pattern in the two HPs, each simulation was run for three time units (TU) – that is, three 2-year intervals for HP1 and three years for HP2 – and the distributions of EPG at the end of each time category will be summarized. Furthermore, based on the results corresponding to each of the three α_i distributions, one set of GM/GSD will be selected for all further simulations.

5.3.2   Correlation between exposure and infection

The second point concerns the correlation between exposure and infection. One notable finding in our previous statistical analyses was that the water exposure magnitude measured by $min \cdot m^2$ was not a statistically significant predictor of EPG [10, 11]. This finding will also be tested by simulation of HP1.

5.3.3   Re-infection clustering

Re-infection clustering refers to the epidemiological finding that repeated infections were found to be non-randomly distributed in the population even when controlled for cercarial exposure: people who were infected in previous surveys were more likely to be re-infected after treatment than those who were not previously infected [12]. Table 5.3 summarizes the distributions of re-infection in the follow-up surveys by individuals' infection status in the baseline survey. In Cohort 1, among previously infected people 41.1% were re-infected. That is 1.63 times more than the re-infected among the non-infected group (25.2%). In comparison, this ratio was larger in Cohort 2 (4.05) where the transmission level was, on the contrary, much lower.

Table 5.3 Re-infection distributions by infection status in the baseline survey

	No. of people	No. infected at follow-up (percentage)
Cohort 1		
Infected at baseline (2000)		
No	222	56 (25.2)
Yes	202	83 (41.1)
Cohort 2		
Infected at baseline (2007)		
No	315	21 (6.7)
Yes	85	23 (27.1)

The results suggested that, as cercarial exposure decreased, so did the fraction of the population vulnerable to *S. japonicum* infection. Therefore, re-infection was more likely to occur in a smaller portion of people. It was also this finding that initially

suggested the plausibility of variable individual susceptibility across the population and led to its further exploration using the individually-based model. However, the statistical model presented in [12] focused strictly on the two original cohorts but not more generalized scenarios. Furthermore, instead of having a cercarial concentration parameter, the statistical analysis used proxies such as village infection prevalence, county of residence, and individual demographic variables as substitutes, which may have compromised the accuracy of individuals' cercarial exposure assessment.

In comparison, the individually-based model not only offers an alternative, mechanism-oriented model structure, but also addresses these two aspects by incorporating the hypothetical populations and explicitly estimated cercarial density. In each simulation, the model is run for two time units (two 2-year intervals for HP1 and two years for HP2). At the end of each TU, a mass chemotherapy treatment is simulated assuming perfect compliance and efficacy. That is, every individual's worm burden will be reset to zero. Any infection occurring at TU1 will be regarded as baseline infection and that of TU2 is the re-infection. If re-infection is random among the population, the proportion of people who are re-infected in each simulation is calculated by:

$$P_R = IR_1 \times IR_2 \qquad \text{(Equation 5.1)}$$

Where IR_1 and IR_2 are fraction of the population infected at TU1 and TU2. On the other hand, the *actual* proportion of re-infections in each simulation is:

$$P_A = \frac{N_R}{N_{HP}} \qquad \text{(Equation 5.2)}$$

Where N_R is the number of people who have infections (EPG>0) at both time points, which can be directly counted from each simulation, and N_{HP} is the number of people in the corresponding hypothetical population. The ratio of the two proportions gives the index to be compared between the two HPs which represent two different levels of cercarial exposure and transmission intensity:

$$R_{AR} = \frac{P_A}{P_R} \qquad \text{(Equation 5.3)}$$

5.3.4   Infection monitoring with susceptibility

The last bullet point concerns disease surveillance in the low transmission environment. Specifically, the question is that if individual susceptibility could be measured, what percentage of infections would be detected by only testing the top 5% to 50% of most susceptible people. The rationale for exploring this new surveillance method arises from two advantages it can potentially offer. First, water contact magnitude alone has proven to be a poor predictor of infection [11]. While cercarial exposure is an incomplete measure of individual risks, a sensitive technique for assessing cercarial concentration in the environment is still unavailable. Thus, there exists a substantial need to seek an alternative, effective indicator that can identify individuals with higher risks of infection. Second, compared with cercarial exposure which can vary significantly from year to year, individual susceptibility as a biological parameter is likely to be a constant or varies slowly over time. If this can be confirmed to be the case, once estimated for an individual, it does not require repeated efforts to monitor this index on a regular basis. Consequently, surveillance of susceptibles, if proven efficient as an index of individual risk, may serve as a new and very cost-effective strategy of disease control.

5.4　Simulation results

　　The simulation results of 2000 Series I simulations of the exposure and infection of HP1 are shown in Figure 5.3. For all subplots, the X-axis is EPG and the Y-axis is the count of people. Qualitatively, all of the plots present clear negative binomial distributions which are phenomenologically consistent with the survey data of cohort 1(Figure 5.2 [1]). Thus, the selection of the GSD of the distribution of individual susceptibility relies on quantitative comparisons between the simulated and survey results regarding the same duration of transmission – one TU, or two years.
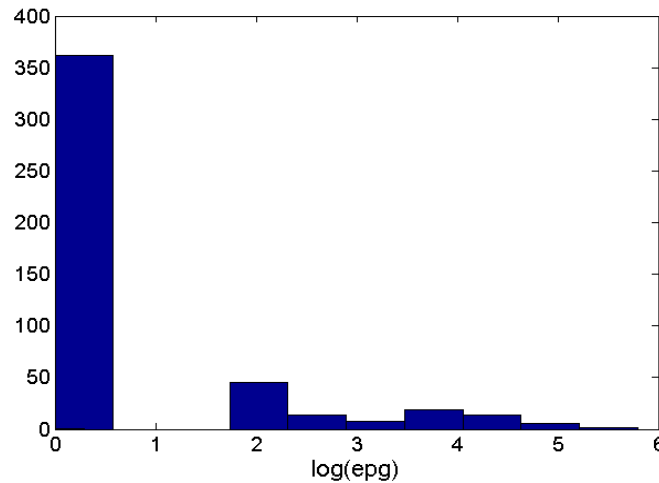

Figure 5.2 Histogram of EPG2002 from survey results at cohort 1

　　Shown in Table 5.4 are the infection prevalence of the population and the mean, variance and aggregation parameter of EPG, the last calculated with the assumption of the NB distribution. Having a similar mean EPG (9.0) to the survey result (8.5), the distribution with GSD of 1.5 yielded more than twice as high a prevalence in HP1 than cohort 1 (51.8% vs. 23.0). In addition, the variance and k_{EPG} were not nearly close to the survey results. Comparing the other two groups, the results corresponding to GSD of 4.5 were in general more consistent with the survey results without sacrificing the accuracy of k_{EPG} estimation. Consequently, a GSD of 4.5 will be adopted in all further simulations. The histograms of simulated prevalence and mean EPG in HP1 are shown in Figure 5.5. The normal-alike distributions of simulated means of prevalence and EPG as shown in Figures 5.4 and 5.5, as the central limit theorem predicts, suggest the adequacy of 2000 Monte-Carlo simulations to yield representative outcomes.

Table 5.4 Mean, variance and aggregation parameter of EPG regarding cohort 1 (survey results) and the corresponding HP1 (simulation results) for several GSDs

	Prevalence (%)	Mean(EPG)	Var(EPG)	k_{EPG}
Cohort 1 survey data	23.0	8.5	824.0	0.08
GSD = 1.5	51.8	9.0	278.7	0.30
GSD = 3	34.9	6.3	426.9	0.09
GSD = 4.5	27.8	6.9	692.1	0.07

---

1 For formatting expedience, Figure 5.2 and Figure 5.3 are presented in the reverse order of quotation.
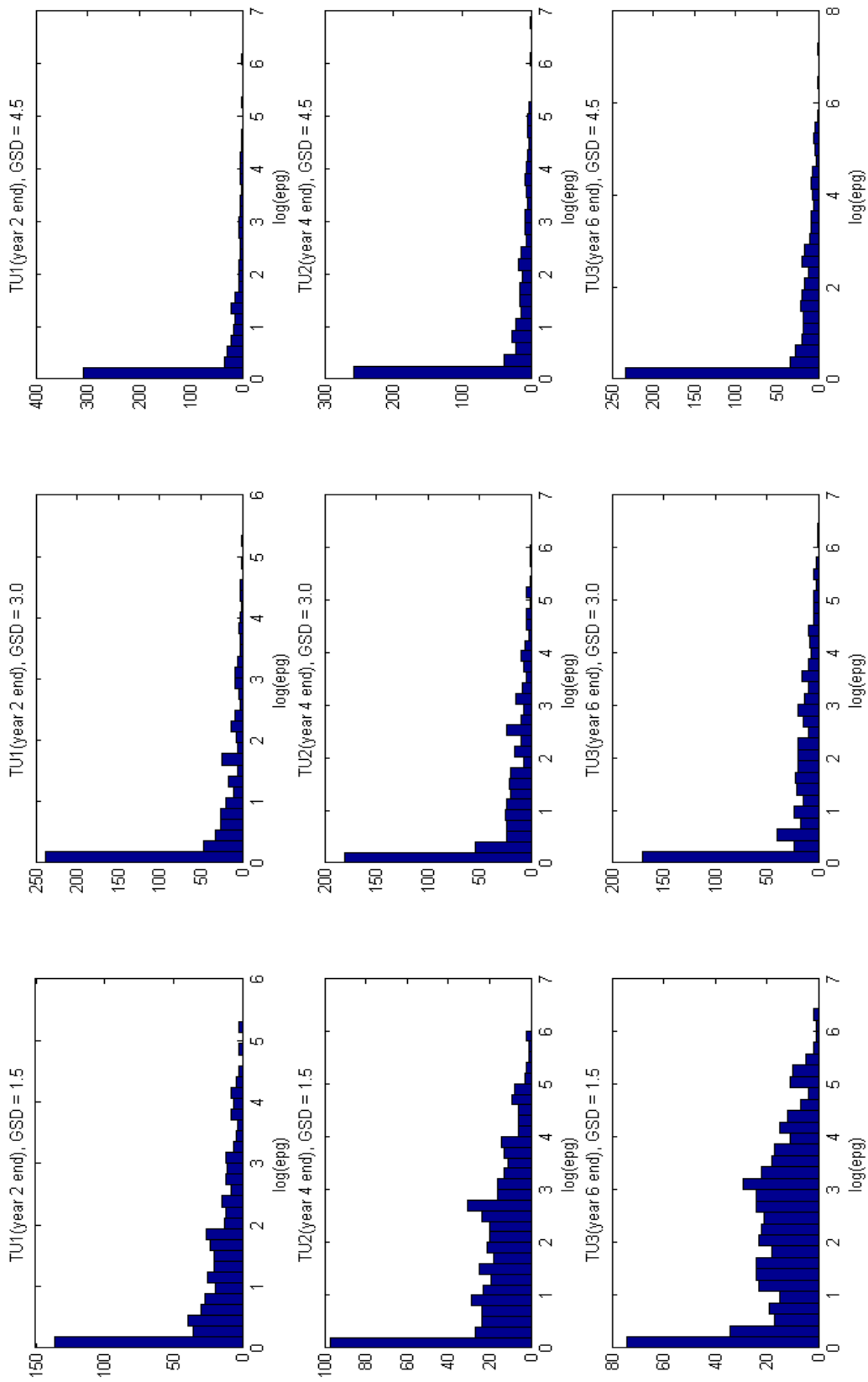
Figure 5.3 Simulated EPG histogram corresponding to three GSDs of the individual susceptibility distribution for HP1 (2000 Series I simulations)
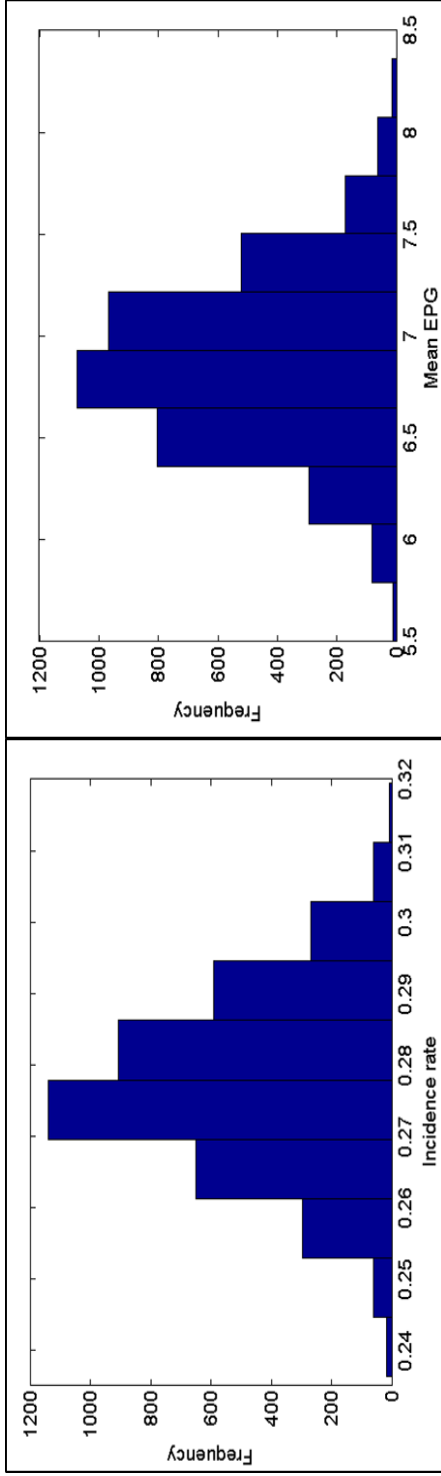
Figure 5.5 Distribution of 1-TU (2-year) incidence rate (left) and mean(EPG) (right) from 2000 Series I simulations for HP1
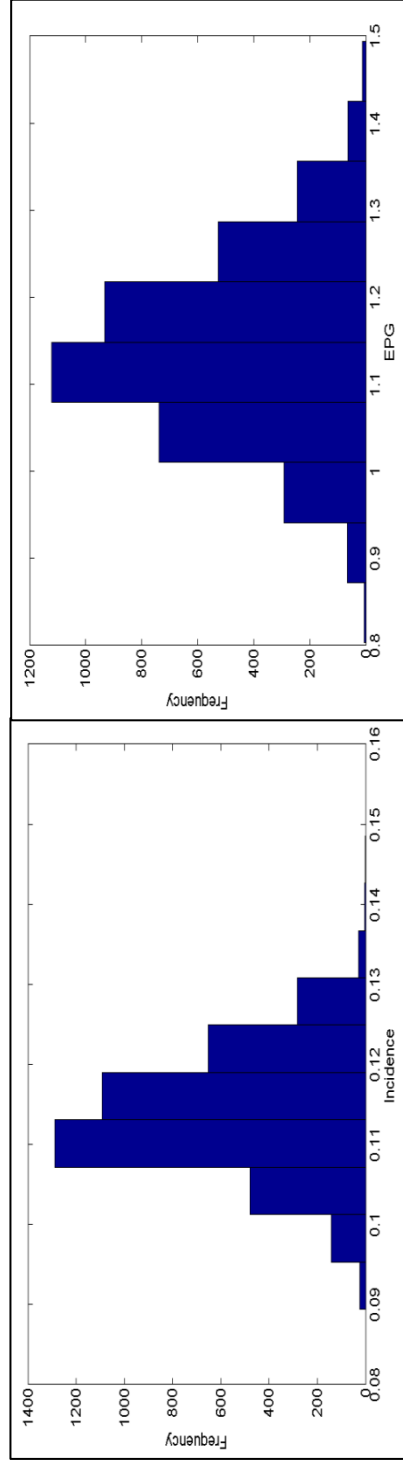
Figure 5.4 Distribution of 1-TU (1-year) incidence rate (left) and mean(EPG) (right) from 2000 Series I simulations for HP2

Using a susceptibility GSD of 4.5, the results of Series I simulations for HP2 are shown in . The mean incidence is quite close to the survey result of cohort 2 (11.4% vs. 11.7%), but the mean EPG is considerably lower than the survey result (1.14 vs. 4.78). However, recall that 75% of the total EPG observed in cohort 2 was attributed to two individuals; upon exclusion of these two outliers, the survey would have yielded a mean EPG of 1.19. Thus, while the individually-based model cannot account for extreme cases potentially due to current assumptions and/or parameterizations of susceptibility and cercarial density, the simulation results in the two hypothetical populations have proved reasonably capable of capturing the incidence rate and infection intensity for the two cohorts.

Turning to the second bullet point, the scatter plot between EPG and body surface area-adjusted water exposure ($n_i \times s_i$) is shown for cohort 1 (top) and HP1 (bottom, Series I simulations) in Figure 5.6, respectively. While the pattern for HP1 is largely subject to the results of random assignment of all individual parameters, it is, on the phenomenological level, quite consistent with that observed from the survey. Qualitatively, there is clearly no positive, if not somewhat negative, relationship between exposure and EPG, according to both the survey and simulation results. Although the correlation is found relatively larger for HP1 than cohort 1 (0.25 vs. 0.03), it is in fact still quite low, and water exposure, therefore, is a poor indicator of infection in both the simulation and epidemiological results.
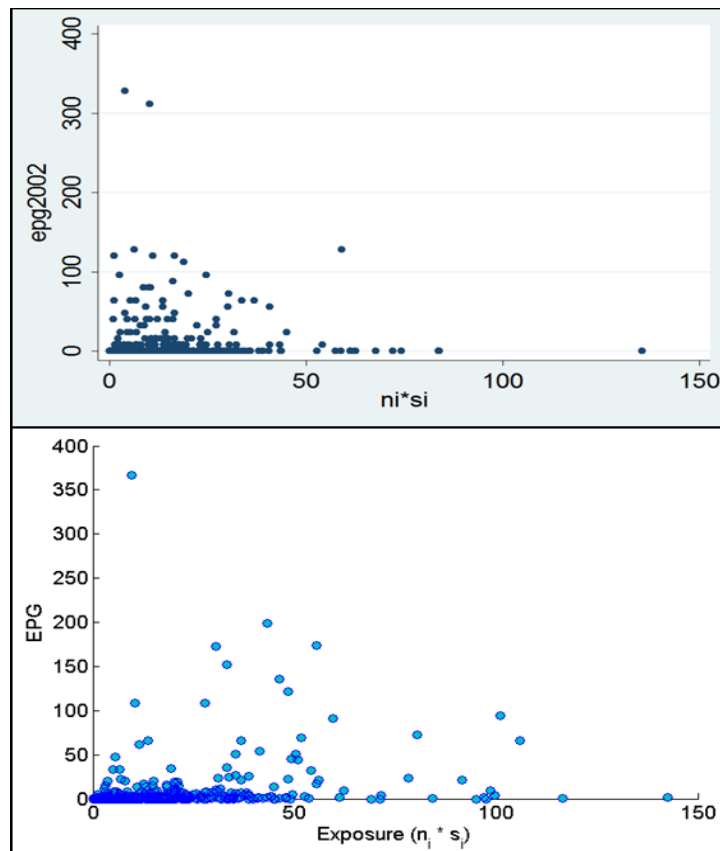
Figure 5.6 Scatter plot of individual EPG vs. water exposure (measured by $n_i s_i$) in Cohort 1 (top) and HP1(bottom)

For both cohorts and the corresponding hypothetical populations, the surveyed and simulated R_{AR} are presented in Table 5.5. The simulation results reveal two trends which are consistent with survey results. First, re-infection is not random among the population – individuals who are infected in the first TU are more likely to be re-infected in the second TU. Second, the ratio of R_{AR} is greater in lower-transmission environment than in higher-transmission environment, as the estimates for HP2 and HP1 reflect.

Table 5.5 Surveyed and simulated results of re-infection percentage ratio (R_{AR})

	Cohort 1/HP1	Cohort 2/HP2
Survey	1.63	4.04
Simulation	2.56	5.72

Both phenomena have been heuristically explained through a conceptual model [12] Specifically, it was suggested that, in a specified exposure environment, there exists a benchmark value of susceptibility, α^*, such that for any individual with α_i below α^*, the probability of an exposure leading a diagnosis of infection is essentially zero. In the individually-based model, this value can be quantified through simulations by taking the minimum α_i that corresponded to a positive infection test result. Corresponding to the distribution of α, the z-scores of α^* estimated in both hypothetical populations are summarized in Table 5.6.

Table 5.6 Z-scores of α^* estimated from model simulations in both hypothetical cohorts

|	Series I simulations	Series II simulations
HP1	-0.0500	-0.2730
HP2	-0.1678	-0.4480

Note that the α^* in each cell is calculated based on more than 2 million samples from each hypothetical population. Taking the most conservative estimate, which is observed in Series II simulations for HP2, the z-score yields a cumulative percentage of 31.3%. In other words, given the assumed distributions for susceptibility and cercarial exposure accounting for water exposure, cercarial density and body surface area, at least 31.3% of all people will essentially never test positive.

The significance of Hatch test (H) versus Kato-Katz (KK) to identify infections in the two hypothetical populations is compared next. Shown in Figure 5.7 are the mean frequency and percentage of infections detected by only one infection testing method but not the other, corresponding to each incidence rate simulated in HP1. It can be observed that the number of H+/KK- (i.e., hatch positive and KK negative) increases by the incidence rate while H-/KK+ remains more constant. However, in terms of the proportion of respective counts out of all positive cases, both patterns appear to be steady although that of H+/KK- has a slight increase in the right-end (the right plot of Figure 5.7). In comparison, the counts of both testing scenarios for HP2 are smaller due to lower transmission levels. However, the percentage of H+ only tests is even greater than that of HP1, and the percentage of KK+ only is clearly lower. This result is consistent with the well documented difficulty of applying the KK method to detect low-intensity infections.
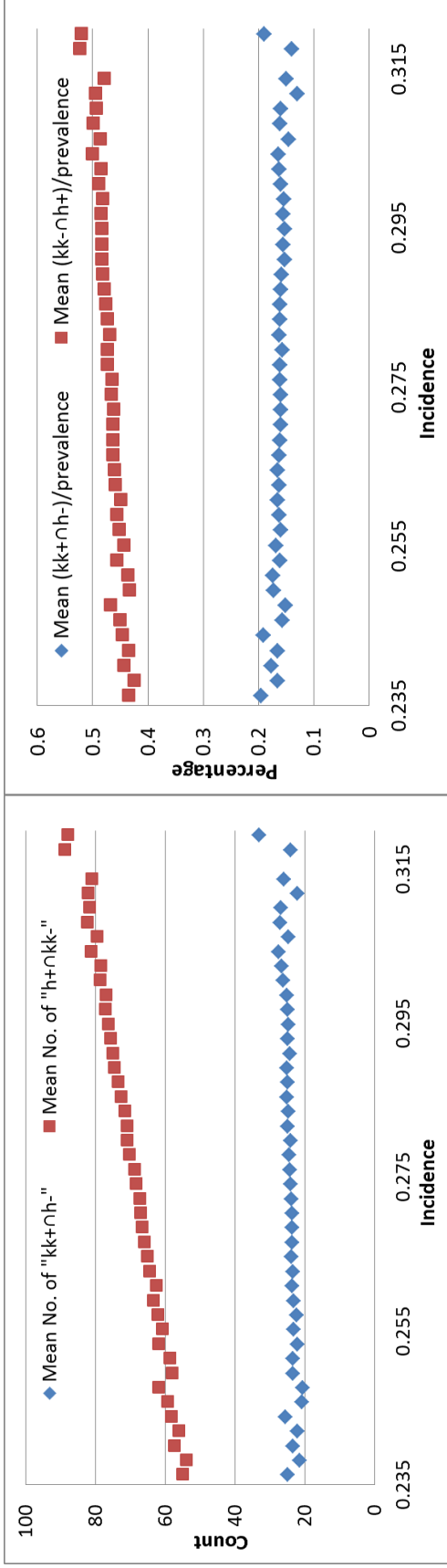
Figure 5.7 The mean count (left) and percentage (right) of infections tested by one method only versus incidence rate in HP1
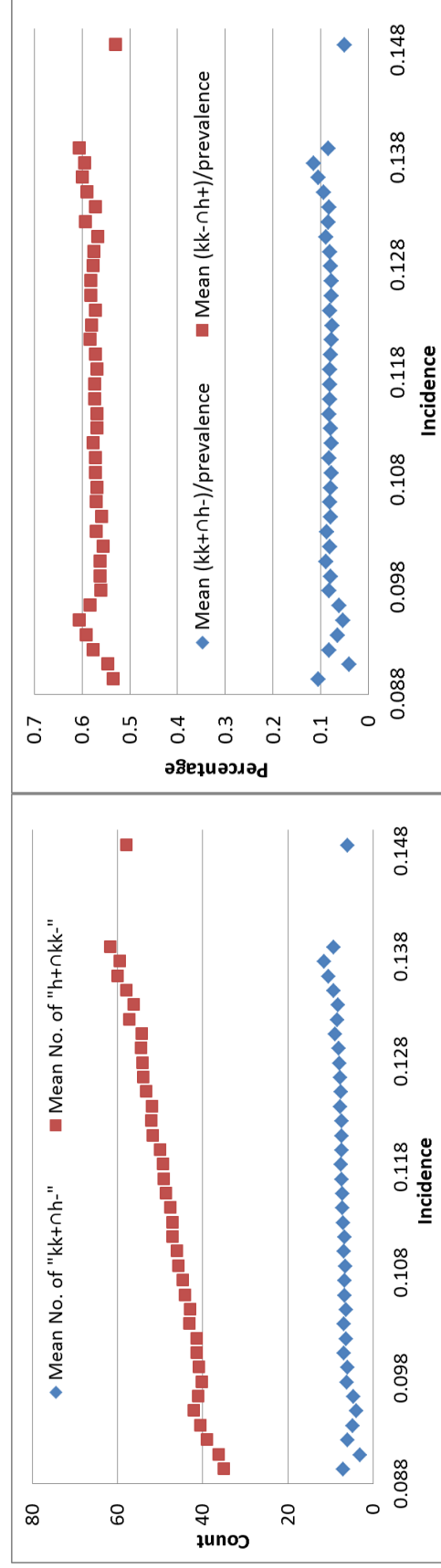
Figure 5.8 The mean count (left) and percentage (right) of infections tested by one method only versus incidence rate in HP2

89

The final issue regards the efficiency of using susceptibility as an index for disease surveillance under the assumption that an individual's level of susceptibility could be determined. Table 5.7 summarizes the results from 2000 Series I simulations in both cohorts; the numbers represent the percentage of all infections that can be detected by only testing people who rank 5% through 50% of each of the four indices, respectively susceptibility (α_i), water contact intervals or equivalently, water contact minutes (n_i), water exposure level measured by ($n_i s_i$), and mean cercarial exposure ($n_i s_i \bar{c}_v$).

For HP1, the efficiency of focusing on susceptibility is consistently better than either the total number of exposure intervals or the body surface area accounted exposure for all sample percentages. Compared with cercarial exposure, using susceptibility as the index offers higher or equivalent surveillance efficiency on both ends of the sampling schemes (top 10% or lower, or top 40% or higher), and is only slightly less efficient in others. For HP2, on the other hand, testing most susceptible individuals proves to have quite compelling efficiency. Not only are the percentages associated with susceptibility all much higher than any of the other indices, they are also markedly higher than the corresponding results found for HP1. For instance, conducting infection tests on the top 25% of most susceptible people can detect 82.3% of all infections in HP2, in contrast to 53.2% for HP1 and 60.4% for HP2 if cercarial exposure, the second most efficient index, is adopted.

Table 5.7 Percentages of all infections that can be detected by testing people ranking 5% thorough 50% of each of the four indices: α_i, n_i, $n_i s_i$, and $n_i s_i \bar{c}_v$ from Series I simulations

Top % of people tracked	Percentage of total infected people							
	HP1				HP2			
	α_i	n_i	$n_i s_i$	$n_i s_i \bar{c}_v$	α_i	n_i	$n_i s_i$	$n_i s_i \bar{c}_v$
5%	15.2	11.0	12.4	15.4	28.8	16.2	17.1	20.0
10%	27.5	21.3	21.6	26.4	47.9	27.3	32.2	32.5
15%	36.7	30.8	31.2	38.3	60.9	43.0	42.6	44.5
20%	44.1	38.5	39.8	48.8	71.3	48.0	54.1	55.4
25%	53.2	45.0	45.7	56.5	82.3	56.0	59.3	60.4
30%	59.5	52.8	53.6	61.2	87.4	61.2	64.6	66.1
35%	65.3	58.6	60.1	67.0	91.1	66.9	69.6	71.9
40%	71.5	64.7	67.4	71.5	94.0	70.7	75.2	79.3
45%	77.2	70.1	74.1	77.0	97.0	76.9	77.9	83.8
50%	81.9	75.9	77.3	80.8	98.1	81.0	81.2	88.9

To further confirm the generality of this finding, Series II simulations were also conducted in both hypothetical populations. Referring to the results in Table 5.8, although the susceptibility-focused percentages are slightly less for HP2 compared to the corresponding numbers in Table 5.7, they are still convincingly more efficient than all alternatives, and this finding also applies to HP1 regardless of the sampling scheme. Thus, the results again suggest the benefits of using susceptibility for disease surveillance over any other indices that have been commonly estimated previously.

Table 5.8 Percentages of all infections that can be detected by testing people ranking 5% thorough 50% of each of the four indices: α_i, n_i, $n_i s_i$, and $n_i s_i \overline{c}_v$ from Series II simulations

Top % of people tracked	Percentage of total infected people							
	HP1				HP2			
	α_i	n_i	$n_i s_i$	$n_i s_i \overline{c}_v$	α_i	n_i	$n_i s_i$	$n_i s_i \overline{c}_v$
5%	14.2	10.9	11.8	13.8	28.8	18.1	19.4	21.6
10%	25.4	19.9	21.2	24.8	47.0	31.1	32.7	36.1
15%	35.7	28.4	30.1	34.9	61.2	42.3	44.0	48.1
20%	44.8	36.1	38.0	43.7	71.7	51.7	53.3	57.8
25%	53.2	43.6	45.6	51.9	79.7	59.8	61.4	65.8
30%	60.4	50.3	52.4	59.1	86.0	67.0	68.6	72.6
35%	67.1	56.8	58.9	65.7	90.4	73.2	74.5	78.3
40%	72.8	62.7	64.7	71.5	93.6	78.4	79.6	83.0
45%	78.1	68.3	70.2	76.7	96.0	83.0	84.0	87.0
50%	82.6	73.4	75.2	81.3	97.6	86.8	87.7	90.3

5.5   Discussion

A fundamental advantage of the individually-based model, implied by the name itself, is its structure which mimics the characteristics and behavior of each individual rather than treating the population as a whole. Regarding the transmission of *S. japonicum* in Sichuan, the individually-based model is particularly well suited to the current situation in which disease transmission is sustained by a much smaller proportion of individuals compared to the endemic equilibrium environment. In the individually-based model, stochastic factors describing water exposure and cercarial density were incorporated for each individual, and Monte-Carlo simulations were applied to capture variability from multiple sources. Furthermore, based on the distributions of key parameters estimated from the epidemiological data, hypothetical populations derived from the same distributions but with different individual profiles were able to be created and simulated to explore more generic scenarios. This, again, can only be accomplished with a model structure focused on individuals. Regarding the simulation results, group-level outcomes such as incidence rates and mean EPG can still be calculated from individual-level outcomes as primary criteria of calibration, but it relies on the latter to identify individuals with higher priorities for disease surveillance. Specifically, the potential of using individual susceptibility as the indicator is explored in this chapter, with the comparisons made with three other indices traditionally estimated in the fields: water contact duration, water exposure incorporated with contact surface area, and cercarial exposure accounting for cercarial density. These results provide motivation for seeking a practical means of identifying susceptible individuals.

The first important conclusion arising from the simulation results derives from the model calibration process using the three GSDs for the distribution of the susceptibility parameter α. It is shown that the simulated mean incidence rate and infection intensity which are most consistent with epidemiological data correspond to the largest GSD, ergo yielding a highly right-skewed distribution of susceptibility. This finding further suggests a highly variable degree of susceptibility and shows the large degree to which this

parameter can vary among the population. In addition, this calibration process also completes the parameterization of this biological variable following the estimation in Chapter 4.

Another benefit of the individual-level simulations, the "benchmark level" of susceptibility conceptualized in our previous paper and, noted as α^*, was quantified. Note that this value is highly dependent on the distributions of other variables, including water exposure, cercarial density and individual body surface area, and thus likely to be spatially and temporally specific. Using the estimates of these variables from our two cohorts, it was found that theoretically almost 1/3 of all people were not at risk of infection even following the most restrictive criteria. This finding raises realistic concerns regarding the efficiency of massive infection testing. Provided that infections are not to be found among a considerable proportion of people, targeting the appropriate individuals in disease surveillance can be efficient, cost-effective, and even necessary in the lower-transmission environments.

Comparing the efficiency of the four indices suggests the merit of focusing on the most susceptible people if an appropriate biomarker of susceptibility could be identified. By testing the same fraction of top-ranked individuals, the highest percentage of infections was found when susceptibility is used as the indicator. More importantly, the advantage is more compelling in the re-emergent. Because susceptibility, as earlier stated, is likely to be much more stable than behavior-related or field-specific variables which require repeated measurements, much less labor and monetary inputs are potentially needed to acquire an accurate portrait of disease prevalence in the population. Of course, none of these can be achieved without finding the appropriate marker of susceptibility and a feasible and accurate test method in the first place. However, the efficiency of this new surveillance strategy, if made feasible and reliable, promises the potential benefits of such exploration.

As the final argument in this chapter, it is worth noting as well that infection testing methods are also in urgent needs of improvement. The simulation results, in addition to the survey results in Cohort 2, showed again how poorly the current K-K method performs in the low-transmission situation. Not only is an alternative test needed to better quantify infection intensity, disease prevalence would also be more accurately estimated by new methods to supplement to the currently used hatch test. An example of newly developed tests is PCR detection for *S. japonicum* eggs, which was shown able to find a couple of cases not otherwise identified by the two traditional methods in our Cohort 2 in the 2010 survey[13]. However, the PCR test, like Kato-Katz, suffers from the same problem in the sense that the result relies on the presence of eggs. It was previously shown in that, in individuals infected with a single worm pair, the probability a one gram stool sample contains zero eggs is 45% to 62%, depending on the aggregation parameter of the particular negative binomial distribution. If samples are collected on three different days, there is a 9.0% to 23.9% chance that no eggs will be present in any sample, for an individual infected with a single worm pair. Keep in mind that these percentages represent the possibility of no eggs produced, rather than no eggs contained in the selected piece of stool sample, the possibility of which should of course be even higher. Even for the hatch test which uses the whole sample, the chance of having no viable eggs hatched increases due to quite limited numbers of total eggs. Thus, dealing with stochasticity of egg production with low worm burdens is a primary consideration of any

newly invented or improved tests.

References

1. Carlton, E.J., Bates, M.N., Zhong, B., Seto, E.Y., and Spear, R.C., *Evaluation of mammalian and intermediate host surveillance methods for detecting schistosomiasis reemergence in southwest China.* PLoS neglected tropical diseases, 2011. **5**(3): p. e987.
2. Gurarie, D., King, C.H., and Wang, X., *A new approach to modelling schistosomiasis transmission based on stratified worm burden.* Parasitology, 2010. **137**(13): p. 1951-1965.
3. Hubbard, A., Liang, S., Maszle, D., Qiu, D., Gu, X., and Spear, R.C., *Estimating the distribution of worm burden and egg excretion of Schistosoma japonicum by risk group in Sichuan Province, China.* Parasitology, 2002. **125**(Pt 3): p. 221-231.
4. Liang, S., Yang, C., Zhong, B., and Qiu, D., *Re-emerging schistosomiasis in hilly and mountainous areas of Sichuan, China.* Bulletin of the World Health Organization, 2006. **84**(2): p. 139-144.
5. Rollinson, D., Knopp, S., Levitz, S., Stothard, J.R., Tchuente, L.A., Garba, A., Mohammed, K.A., Schur, N., Person, B., Colley, D.G., and Utzinger, J., *Time to set the agenda for schistosomiasis elimination.* Acta Trop, 2012.
6. Spear, R.C., Hubbard, A., Liang, S., and Seto, E., *Disease transmission models for public health decision making: toward an approach for designing intervention strategies for Schistosomiasis japonica.* Environmental health perspectives, 2002. **110**(9): p. 907-915.
7. Spear, R.C., Seto, E., Liang, S., Birkner, M., Hubbard, A., Qiu, D., Yang, C., Zhong, B., Xu, F., Gu, X., and Davis, G.M., *Factors influencing the transmission of Schistosoma japonicum in the mountains of Sichuan Province of China.* The American Journal of Tropical Medicine and Hygiene, 2004. **70**(1): p. 48-56.
8. Wang, L.D., Guo, J.G., Wu, X.H., Chen, H.G., Wang, T.P., Zhu, S.P., Zhang, Z.H., Steinmann, P., Yang, G.J., Wang, S.P., Wu, Z.D., Wang, L.Y., Hao, Y., Bergquist, R., Utzinger, J., and Zhou, X.N., *China's new strategy to block Schistosoma japonicum transmission: experiences and impact beyond schistosomiasis.* Trop Med Int Health, 2009. **14**(12): p. 1475-83.
9. Zhou, X.N., Guo, J.G., Wu, X.H., Jiang, Q.W., Zheng, J., Dang, H., Wang, X.H., Xu, J., Zhu, H.Q., Wu, G.L., Li, Y.S., Xu, X.J., Chen, H.G., Wang, T.P., Zhu, Y.C., Qiu, D.C., Dong, X.Q., Zhao, G.M., Zhang, S.J., Zhao, N.Q., Xia, G., Wang, L.Y., Zhang, S.Q., Lin, D.D., Chen, M.G., and Hao, Y., *Epidemiology of schistosomiasis in the People's Republic of China, 2004.* Emerging infectious diseases, 2007. **13**(10): p. 1470-1476.
10. Seto, E.Y. and Carlton, E.J., *Disease transmission models for public health decision-making: designing intervention strategies for Schistosoma japonicum.* Advances in Experimental Medicine and Biology, 2010. **673**: p. 172-183.
11. Seto, E.Y., Lee, Y.J., Liang, S., and Zhong, B., *Individual and village-level study of water contact patterns and Schistosoma japonicum infection in mountainous rural*

China. Tropical medicine & international health : TM & IH, 2007. **12**(10): p. 1199-1209.

12. Carlton, E.J., Hubbard, A., Wang, S., and Spear, R.C., *Repeated Schistosoma japonicum infection following treatment in two cohorts: evidence for host susceptibility to helminthiasis?* PLoS Negl Trop Dis, 2013. **7**(3): p. e2098.

13. Fung, M.S., Xiao, N., Wang, S., and Carlton, E.J., *Field evaluation of a PCR test for Schistosoma japonicum egg detection in low-prevalence regions of China.* Am J Trop Med Hyg, 2012. **87**(6): p. 1053-8.

Chapter 6 Conclusion

"DON'T PANIC."

— *The Hitchhiker's Guide to the Galaxy*

After several decades of continuous effort, the battle against schistosomiasis in China has come to a stage where a glimmer of ultimate control has finally come into sight. Disease elimination, defined as prevalence of less than 1%, has been made an official goal to be achieved by the end of 2015 as the first of final steps towards eradication [1]. When originally proposed, this aim was considered quite feasible and practical, but progress in disease control has slowed down significantly in the past decade [2-6]. Further suppression of transmission in regions where infection levels are already low has proven to be extremely difficult; re-emergence of the disease in areas where it was controlled or even eliminated has further complicated the situation [2]. In this context, and focusing on the hilly and mountainous regions in Sichuan Province, this dissertation aimed to explore and evaluate new possibilities of prevention, control and surveillance that can potentially overcome the current barriers.

6.1   The contents

The main contents of this dissertation are centered on the development and exploration of an individually-based model. Prior to conceptualizing the model structure, I used a health education intervention as an example of showing how individual data could be incorporated in the model, and how it could be used to test for hypotheses that were applicable to model development through statistical analyses. In the following chapter, I provided details on the protocols of each field survey, and presented descriptive statistics of the results on demographics, infection testing, snail survey, and cercarial mouse bioassay. Specifically, results from two cohorts that represented the middle and low level transmission environments were presented and compared. Correspondingly, the hypothetical populations to be studied in individually-based model simulations were derived from either one of the cohorts.

The design of model structure was mainly driven by two factors: data availability and the necessity of focusing on individuals in the low-emergence scenario. Low sensitivity of current parasite monitoring techniques compromised the usefulness of our community-based model whose performance relied largely on calibration of the snail and cercariae modules. Similar problems also applied to the human infection module – only a small proportion of people now have quantifiable infections under the current infection testing techniques. Thus, it was the individual-level infection intensity, rather than the group-mean, that became the appropriate outcome to identify individuals with higher risks and higher priorities of surveillance.

The structure of the individually-based model was therefore designed to accommodate both aspects: it should include the best estimates of parasite distribution in the environment that can be made with the fewest assumptions, and should be able to yield individual-level infection outcomes through comprehensive simulation of parasite development *in vivo* given the sensitivity of the infection testing process. As human infection is directly caused by contact with cercariae, the density of this parasite stage, rather than miracidia or schistosome eggs, was quantified using previous data from higher transmission environments as the reference and benchmark. To capture the stochasticity of worm development, egg production and the infection testing process which was particularly important in the low-infection scenario, stochastic factors were incorporated in each of these steps. Accordingly, Monte-Carlo simulations of the individually-based model were applied to account for randomness from multiple sources.

Another unique feature of the individually-based model, as stated in the Chapter 5, was its capability of exploring more generic scenarios. In addition to interpreting the epidemiological findings in our study regions, the individually-based model was used to verify the universality of these findings by simulating hypothetical populations which are derived from the same distributions as the surveyed populations. Thus, the simulation results were robust since they were consistently observed across different hypothetical populations.

6.2 The conclusions

The health education intervention was not able to differentiate the intervention group from the controls based on any of the five indices on knowledge, attitudes and behaviors. Improvement of knowledge, attitudes and the defecation behavior found in both groups indicated the "diffusion of innovation" effect, but the intervention failed to have impact in reducing people's water contact or increasing their frequency of using personal protection equipment.

This study is the first randomized-controlled trial education intervention regarding schistosomiasis in China since 2005. Among a handful of similarly well-designed studies in the past decade, this study is the first to suggest that intervention had no statistically significant effects based on intergroup comparisons. For the three indices on which the control group also had improvement, lacking of significant intergroup differences may result from the diffusion of innovation phenomenon. In the rural area environment where kinship can be commonly found in villages within relatively close distance with each other, although contamination bias was specifically considered during group stratification, it was almost impossible to completely prevent inter-group communication.

From the perspective of disease control, the improvement in both groups does indicate the importance of health education. Despite the mechanism, the "radiation effects" of education to ambient non-intervention groups clearly benefits the whole community in the endemic regions. However, it is of particular concern that the two most important behaviors related to infection risk, water contact and the use of self-protective equipment, could not be changed by education. Future education programs should be designed to target these two aspects specifically.

Surveillance results in the re-emerging villages indicated low sensitivity of current techniques for low infection-intensity environment.

Using the same protocols that have been applied in past years, the results of mouse bioassays and infected snail surveys in the low-transmission environment were not able to characterize the distribution of schistosome parasites in the water. The extreme low sensitivity of both methods suggested the urgent needs of more advanced techniques of parasite monitoring. Regarding infection testing, of the 72 people tested positive in 2008, only 25 had measureable EPG levels as estimated by the Kato-Katz method. Clearly, more sensitive methods that can quantify infection intensity are required to counter the stochasticity of egg production and detection among low-infected individuals.

The distribution of individual susceptibility to S. japonicum infection appears to be log-normally distributed with a high skewness. Given the cercarial density and water contact magnitude distributions of our study populations, approximately 1/3 of all people have the susceptibility levels that would never lead to a positive infection testing result.

Upon development of the individually-based model, Monte-Carlo simulations were conducted to explore the distribution and benchmark level of susceptibility. Combining the results of point estimates from both cohorts, susceptibility was suggested to be right-skewed with a high degree of dispersion. It was further shown through simulations on hypothetical populations that the benchmark level was relatively high even following the most rigorous criteria. Thus, a relatively large fraction of people, due to their susceptibility level, would not test positive regardless of their water contact magnitude (or, to be more precise, as long as their magnitude does not exceed the maximum of the 6 million plus samples in the simulation). This result implies the possibility of improving the efficiency of disease surveillance by targeting a subset of individuals with high risk. So the next logical question is which index can best indicate individual risk of infection.

For the purpose of human infection surveillance, sampling individuals with high susceptibility has more efficiency than sampling individuals with high levels of any of the following indices: water contact duration, water contact magnitude accounting for averaged body surface area, and cercarial exposure.

Again based on the results of Monte-Carlo simulations, the efficiency of focusing on the most susceptible individuals for disease surveillance was shown to be consistently

greater than focusing on any other indices that have been conventionally estimated. Moreover, the advantage is more striking in the low-infection environment than the post-equilibrium moderate-infection environment. This finding indicates a promising strategy of disease surveillance particularly well suited for the current re-emerging scenario. If susceptibility could be identified and quantified, the labor-intensive and costly mass infection surveys could be modified to target the most appropriate individuals and still achieve a satisfactory picture of disease prevalence and infection intensity.

6.3   The finale

It is the best of times – with the resolution and support of the Chinese government, never has there been a better timing for wiping out schistosomiasis in China. It is the worst of times – the beauty and clearness of the integrated control strategy is at present obscured by re-emergence[2]. However, it is exactly these obstacles that prove that more needs to be done, and more can be done.

Although solutions that can lead to schistosomiasis elimination or even eradiation are still to be determined, there is little doubt that traditional control strategies combining snail control, chemotherapy and health education are inadequate for achieving these goals. Future directions of technique improvement are clearly suggested from field survey results and model simulations. Priorities should be made to develop more sensitive environmental monitoring and infection testing methods, and exploration of an index for susceptibility will greatly benefit the efficiency of disease surveillance [3, 7, 8].

However, more important than enhancement on the technical side are changes on the strategy side. From disease control to elimination, high-level strategic methods have to be conceptualized, specified and well implemented. In addition to environmental modification, a concept recommended in our previous studies aiming to diminish the habitat of parasites and snails, economic modification is another new strategy managing to reduce individual risk by changing people's economy-associated demographics [9-11]. One possibility, for example, is to replace bovines with tractors for agricultural uses as the rural economy develops [12]. In endemic regions where bovines are a major definitive host, this method can potentially be effective of lowering the infection level. However, whether it is applicable also depends on geographic characteristics of the specific area. In our hilly and mountainous study regions, for instance, the utilization of tractors is apparently quite infeasible. Consequently, just as environmental modification, the implementation of economic modification or any other new strategy has also to be highly site and population specific.

---

2 Modified from William Thompson's famous words:" The beauty and clearness of the dynamical theory, which asserts heat and light to be modes of motion, is at present obscured by two clouds".

References

1. Wang, L.D., Chen, H.G., Guo, J.G., Zeng, X.J., Hong, X.L., Xiong, J.J., Wu, X.H., Wang, X.H., Wang, L.Y., Xia, G., Hao, Y., Chin, D.P., and Zhou, X.N., *A strategy to control transmission of Schistosoma japonicum in China.* The New England journal of medicine, 2009. **360**(2): p. 121-128.
2. Liang, S., Yang, C., Zhong, B., and Qiu, D., *Re-emerging schistosomiasis in hilly and mountainous areas of Sichuan, China.* Bulletin of the World Health Organization, 2006. **84**(2): p. 139-144.
3. Spear, R.C., Seto, E.Y., Carlton, E.J., Liang, S., Remais, J.V., Zhong, B., and Qiu, D., *The challenge of effective surveillance in moving from low transmission to elimination of schistosomiasis in China.* Int J Parasitol, 2011. **41**(12): p. 1243-7.
4. Carlton, E.J., Bates, M.N., Zhong, B., Seto, E.Y., and Spear, R.C., *Evaluation of mammalian and intermediate host surveillance methods for detecting schistosomiasis reemergence in southwest China.* PLoS neglected tropical diseases, 2011. **5**(3): p. e987.
5. Zhou, X.N., Bergquist, R., Leonardo, L., Yang, G.J., Yang, K., Sudomo, M., and Olveda, R., *Schistosomiasis japonica control and research needs.* Adv Parasitol, 2010. **72**: p. 145-78.
6. Zhou, Y.B., Liang, S., and Jiang, Q.W., *Factors impacting on progress towards elimination of transmission of schistosomiasis japonica in China.* Parasit Vectors, 2012. **5**: p. 275.
7. Wang, S. and Spear, R., *Exploring the impact of infection-induced immunity on the transmission of Schistosoma japonicum in hilly and mountainous environments in China.* Acta Tropica (under review), 2013.
8. Spear, R.C., *The Role of Environmental Determinants of R0 in Suppressing Re-emergent Transmission of S.japonicum in China.* Journal of the Royal Society, Interface, 2011.
9. Gurarie, D. and Seto, E.Y., *Connectivity sustains disease transmission in environments with low potential for endemicity: modelling schistosomiasis with hydrologic and social connectivities.* Journal of the Royal Society, Interface / the Royal Society, 2009. **6**(35): p. 495-508.
10. Liang, S., Seto, E.Y., Remais, J.V., Zhong, B., Yang, C., Hubbard, A., Davis, G.M., Gu, X., Qiu, D., and Spear, R.C., *Environmental effects on parasitic disease transmission exemplified by schistosomiasis in western China.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(17): p. 7110-7115.
11. Liang, S., Spear, R.C., Seto, E., Hubbard, A., and Qiu, D., *A multi-group model of Schistosoma japonicum transmission dynamics and control: model calibration and control prediction.* Tropical medicine & international health : TM & IH, 2005. **10**(3): p. 263-278.
12. Wang, L.D., Guo, J.G., Wu, X.H., Chen, H.G., Wang, T.P., Zhu, S.P., Zhang, Z.H., Steinmann, P., Yang, G.J., Wang, S.P., Wu, Z.D., Wang, L.Y., Hao, Y., Bergquist, R., Utzinger, J., and Zhou, X.N., *China's new strategy to block Schistosoma japonicum transmission: experiences and impact beyond schistosomiasis.* Trop Med Int Health, 2009. **14**(12): p. 1475-83.