

ESTIMATING AND MODELING THE DYNAMICS OF THE INTENSITY OF INFECTION WITH *SCHISTOSOMA JAPONICUM* IN VILLAGERS OF LEYTE, PHILIPPINES. PART II: INTENSITY-SPECIFIC TRANSMISSION OF *S. JAPONICUM*. THE SCHISTOSOMIASIS TRANSMISSION AND ECOLOGY PROJECT

STEVEN RILEY, HÉLÈNE CARABIN, CLAIRE MARSHALL, REMIGIO OLVEDA, A. LEE WILLINGHAM,
AND STEPHEN T. MCGARVEY

Department of Community Medicine, Faculty of Medicine, The University of Hong Kong, Hong Kong; Department of Biostatistics and Epidemiology, College of Public Health, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma; Division of Primary Health Care and Public Health, Faculty of Medicine, Department of Epidemiology and Public Health, Imperial College, London, United Kingdom; Department of Health, Research Institute for Tropical Medicine, Manila, The Philippines; Department of Experimental Parasitology, Royal Veterinary and Agricultural University, Copenhagen, Denmark; International Health Institute, Brown University, Providence, Rhode Island

Abstract. A dynamic model of *Schistosoma japonicum* transmission is presented that incorporates effects of infection intensity, age, and sex. We use four infection intensity classes to investigate the impact of ecologic changes and public health interventions on the burden of infection within communities. Age- and sex-specific infection data from three disease-endemic villages in the Philippines are used to estimate the parameters of the model. The model gives good qualitative agreement with observed fecal egg counts adjusted for the accuracy of the Kato-Katz examination. Our results suggest that differences in infection burden between villages are caused by differences in both the infection process and the recovery process in humans. We describe the potential impact of mass treatment of all humans on the numbers with high infection. Furthermore, we show that a sudden reduction in snail population size would affect high prevalence and low prevalence communities in different ways.

INTRODUCTION

Schistosoma japonicum causes schistosomiasis and imposes a substantial health burden on the people of the Philippines. The comparison of possible intervention strategies and the evaluation of the impact of ecologic change on the transmission of the infection and burden of disease are of considerable interest. There are an estimated 6.7 million people at risk with 1.8 million directly exposed to the infection.^{1,2} Symptoms of infection include fever and muscular pain with acute schistosomiasis, malnutrition and cognitive deficits in youth, and hepatomegaly and hepatic fibrosis of the portal vein with chronic schistosomiasis.^{3–5} All schistosomes require either aquatic or amphibious molluscan intermediate hosts to complete the life cycle of the parasite with infected intermediate hosts living for up to half a year. Therefore, although inexpensive effective treatment with praziquantel is available, people who remain in disease-endemic areas usually reacquire infection after treatment. Any substantial changes to local water courses are likely to have an impact on the snail populations and people's water contact patterns. These changes may, in turn, have a significant effect on the local prevalence and severity of schistosomiasis in humans.⁶ One further complication of *S. japonicum*, compared with other schistosomes, is its potential ability to infect other mammalian hosts⁷: cattle and water buffaloes in China,¹ as well as dogs, cats, goats, pigs, and rats in the Philippines.⁸ These hosts may act as a reservoir for the adult phase of the parasite.

Deterministic dynamic models of infectious disease have been used extensively to study the transmission of the African schistosomes.^{9–11} However, with the exception of the report by Hairston,¹² application of these models to *S. japonicum* has been less frequent until recently.^{1,13} These two latter models adopt a similar approach to that used here. Specifically, the free-living stages of the parasite are not included explicitly. Effectively, it is assumed that infected snails are capable of directly infecting mammalian hosts. Given that the

free-living stages have much shorter life spans than the parasitic stages, this is a reasonable approximation.¹⁴

The model described here extends those previously reported^{1,13} by allowing for different intensities of infection and age classes in the human host. This allows the model parameters to be estimated using detailed, corrected estimates of infection prevalence in different age and sex categories from a study conducted in three villages in Leyte Province, the Philippines, before any control program had been implemented.¹⁵ The three villages were separated from each other by a minimum of 10 km and at the time of the study, there was no significant movement of residents between villages. Almost all residents in the study villages were farmers and water was supplied mainly by open wells, although some houses were equipped with hand pumps.

MATERIALS AND METHODS

Data. The population and methodology used to obtain the estimates of the proportion of the population in each age, sex, and intensity of infection category and its 95% confidence intervals (CIs) are described in a companion paper.¹⁶ Briefly, these cross-sectional data are from the baseline of a eight-year longitudinal study conducted in three villages in Leyte, the Philippines.⁴ A Bayesian cumulative-logit model with non-proportional odds, within which we adjusted for the sensitivity and specificity of the single stool Kato-Katz test, was used to obtain the village-specific and average prevalence of each age, sex, and intensity of infection category and its 95% CIs.¹⁶ These data are presented in Figure 1.

Dynamic model. The intensity of infection and the way in which it changes with age is an important feature of the data described, and, more generally, of human schistosome infections. Morbidity is associated with duration of the different intensities of infection. Therefore, in this model, the human population of each village is stratified using both intensity of infection and age. Three age categories are used: 0–6, 7–12,

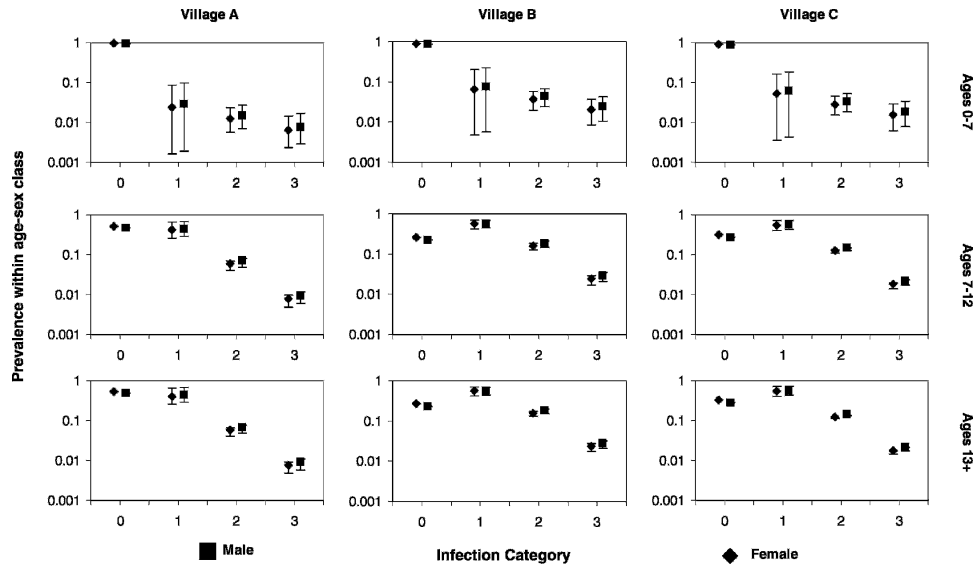


FIGURE 1. Observed infection profiles for men and women in each of the three age categories for each village in the study. A log scale is used for the y-axis to distinguish between differences in the higher infection classes. The infection classes are defined by the concentration of eggs per gram (epg) of stool: 0 = no infection or 0 epg; 1 = light or 1–100 epg; 2 = moderate or 101–800 epg; 3 = heavy or ≥ 800 epg. Error bars show the 95% confidence intervals for prevalence. Since these are based on contrasts, they are not shown for the uninfected class (see Appendix 2 for a description of a contrast).

and ≥ 13 years. Four classes of infection are used: none (0 eggs per gram [epg] of stool), light (1–100 epg), moderate (101–800 epg), and heavy (≥ 801 epg). Sex is also modeled as a further sub-categorization because water contact patterns differ greatly between men and women due to divisions of labor. The key features of the dynamic model are given in the remainder of this section and a full model definition is presented in Appendix 1. All parameters used and their values (either assumed or estimated) are given in Tables 1 and 2.

The principle feature of the dynamic model is the mechanism through which humans change their infection status: how quickly they acquire infection and how the intensity of infection increases and decreases over time. We assume that

these processes may differ between villages, that they vary between age-sex groups and that they are dependent on the number of infectious snails within each habitat. The “per snail” rate at which humans in infection class i , age category j and sex group k progress to the next infection class is defined by the indexed parameter $\beta_{i,j,k}^{SH}$. For example, the per snail rate with which young male children acquire light infection is given by $\beta_{0,0,0}^{SH}$ and the rate at which those same children may progress from mild to moderate infection is given by $\beta_{1,0,0}^{SH}$. It should be noted that one can only progress to moderate infection from mild infection: it is impossible to become moderately infected if one is not already infected. Therefore, the progression described by $\beta_{1,0,0}^{SH}$ is conditional on individuals

TABLE 1
Key to parameters used in differential equations to define the dynamic model*

Parameter	Description	Value	Notes
v_i^H	Human birth rate	0 for $i \neq 0$ 0.0235 for $i = 0$	Average lifetime of approximately 50 years in slightly growing population
μ_i^H	Human death rate	0.0001	As above
γ_j^H	Human aging rates	1/7 for $j = 0$ 1/6 for $j = 1$ 1/48 for $j = 3$	Reflects age-class boundaries at years 7 and 13 with a maximum of 60 years
v^S	Snail birth rate	2	Average lifetime of 6 months and constant population size
μ^S	Snail death rate	2	As above
ϕ^S	Snail recovery rate	0	Snails usually die before they recover. We do not include parasite-associated mortality in this model.
$\beta_{i,j,k}^{HS}$	Rate of transmission from humans to snails	1	Set to unity because no information was available on snail population size or parasite prevalence
β^{RS}	Rate of transmission from reservoir mammals to snails	Set equal to β_{base}^{HS} , which is estimated	No data was available on possible reservoir mammal populations. Therefore, it is assumed to behave as humans.
ν^R	Reservoir mammal birth rate	0.33	Average lifetime of 3 years and constant population size
μ^R	Reservoir mammal death rate	0.33	As above
β^{SR}	Rate of transmission from snail to reservoir mammal	1	Set to unity because no information was available on snail population size or parasite prevalence
ϕ^R	Reservoir mammal recovery rate	Set equal to ϕ_1^H , which is estimated	No data was available on possible reservoir mammal populations. Therefore, it is assumed to behave as humans.

* These parameters are assumed to be the same for each of the three villages.

TABLE 2
Estimated parameters values and 95% confidence intervals for dynamic model*

Parameter	Description [range explored]	Village	Estimate	95% CI
β_{base}^{SH}	Scaling parameter for snail to human transmission [1×10^{-8} , 10]	A	1.5×10^{-5}	9.3×10^{-6} 2.1×10^{-5}
		B	3.8×10^{-5}	1.8×10^{-5} 6.2×10^{-5}
		C	2.3×10^{-5}	1.8×10^{-5} 2.8×10^{-5}
α_1	Snail to human transmission rate from light infection to moderate infection, relative to rate from no infection to light infection [0.001, 100]	A	1.6	0.64, 7.5
		B	0.69	0.64, 16
		C	13	4.5, 20
α_2	Snail to human transmission rate from moderate infection to heavy infection, relative to rate from no infection to light infection [0.001, 100]	A	20	2.6, 35
		B	17	4.5, 29
		C	9.4	5.3, 19
ϵ_1	Snail to human transmission for men relative to women [-1, 1]	A	0.15	-0.0063, 0.23
		B	0.11	-0.020, 0.27
		C	0.11	-0.0043, 0.28
λ	First parameter used to define how transmission varies with age category [0.001, 10]	A	0.54	0.37, 1.8
		B	0.074	0.0011, 0.83
		C	0.46	0.11, 0.67
δ	Second parameter used to define how transmission varies with age category [0.001, 10]	A	0.52	0.020, 0.72
		B	1.1	0.53, 1.8
		C	0.60	0.28, 1.1
ϕ^H	Basic rate of recovery for humans from light infection to no infection [1×10^{-5} , 10]	A	0.060	0.023, 0.12
		B	0.058	0.013, 0.14
		C	0.034	0.015, 0.074
ϕ_2^H	Human recovery rate from moderate infection to light infection, relative to that from light to no infection [0.001, 10]	A	0.73	0.30, 2.8
		B	0.60	0.0023, 6.1
		C	8.3	3.75, 10
ϕ_3^H	Human recovery rate from heavy infection to moderate infection relative to that from light to no infection [0.001, 10]	A	6.3	0.64, 10
		B	9.1	2.3, 10
		C	3.4	2.1, 7.0

* Other parameter values (not estimated) are given in Table 1. With these parameter values, the model is consistent with the data for each village: $P = 0.92, 0.85,$ and 0.90 for villages A, B, and C, respectively. The ranges of parameter values investigated are given in brackets in the Description column. CI = confidence interval.

already having mild infection. A functional form is used for $\beta_{i,j,k}^{SH}$, so as to reduce the total number of parameters used, details of which are given in the Appendix 1. The scalar parameter β_{base}^{SH} determines the base per snail rate of infection in each village.

At any given time it must also be possible for humans to reduce the intensity of their infection. We assume that this process is dependent on the age-sex group of the individual, but not on the size and spatial distribution of the local snail population. It is important to note that the class of infection of an individual reflects the concentration of excreted eggs and that this measure is unlikely to be directly proportional to the number of established adult parasites.¹⁷ Therefore, the process of recovery attempts to parameterize how the levels of infection would decrease in a population if all snails were removed from the local habitat. It should not be interpreted as being a linear process directly proportional to the death rate of the adult worms. Recovery takes a more simple functional form than transmission, details of which are also given in Appendix 1. The parameter ϕ^H determines the baseline rate for recovery between intensity of infection classes and is allowed to vary between villages but not between age-sex groups.

The parameters of the dynamic model can be divided into three types. The first type contains those parameters that are well understood biologically and can have values assigned with a reasonable degree of accuracy, e.g., the human birth rate. The values used for these parameters and the corresponding description are given in Table 1. The second type consists of those that are not constrained by these data, or by

data from other studies. Some parameters of this type are not constrained because no data are available, e.g., rates of transmission from snails to mammalian reservoirs. Others are not constrained due to the cross-sectional nature of the data, e.g., rates of transmission from humans to snails. Values used in the model and the rationale for their choice are also given in Table 1. The third type is those that are not well understood biologically but are constrained by the data, e.g., the base rate of transmission from snails to humans. These values are presented in Table 2. By definition, estimated values for constrained parameters are not sensitive to the values chosen for unconstrained parameters.

To investigate the potential impact of non-human mammalian reservoir, its average lifetime was assumed to be three years and its population was assumed to be of constant size with births balancing deaths. Transmission to and from the reservoir, infection rates, and recovery rates were all assumed to be the same as that for humans with light infection (i.e., $i = 1$). A significant difference was not obtained for either the goodness-of-fit (GOF) P value or parameter estimates when the model was fitted with and without a reservoir. The values given in Table 2 are for parameters estimated without a reservoir.

Since there is no information available for the prevalence of *S. japonicum* infection in snails, the rates with which humans in different intensity of infection classes infect snails ($\beta_{i,j,k}^{HS}$) were not constrained by these data. Parameters were estimated using different constant values. Again, no significant difference was found in the GOF P value or the parameter estimates for the estimates of $\beta_{i,j,k}^{HS}$. However, the steady-

state prevalence of infection in snails and the time needed by the system to reach an equilibrium were sensitive to this parameter. Therefore, for simplicity, unity was used for the rates with which humans infect snails. The speed of recovery of the system from interventions should nonetheless be thought of as conditional on these rates.

The model was formulated as a set of ordinary differential equations and is described in detail in Appendix 1. Some form of inferential framework is required to be able to make precise statements relating the mathematical model to the data. Specifically, such a framework is required to identify which values of the transmission parameters provide the best fit to the data and to check for correlation among those parameters near the best-fit values. Here, we use a maximum GOF approach that permits both point and interval estimates for the parameters of our model. As far as we are aware, the inferential approach is novel in its application to macroparasite transmission models. Details of this approach are given in Appendix 2.

Comparisons of interventions under different scenarios were made using the maximum GOF parameter estimates. The two interventions modeled were a single mass treatment of humans and a sudden change in snail numbers. The latter may result from a large-scale ecologic change, e.g., introduction of cemented irrigation canals. For mass treatment of humans, it was assumed that a 95% coverage was achieved with a drug with 100% efficacy, i.e., 95% of infected people were moved to the uninfected class. The step change in numbers of snails was assumed to be a decrease in the numbers of susceptible and infected snails in equal proportions. It should be noted that substantial ecologic change may increase the size of the snail population rather than decrease it.

RESULTS

The maximum goodness-of-fit parameter estimates (MGOFE) obtained are shown in Table 2. A comparison of model results

with the data is presented in Figure 2. The model is able to reproduce the main features of the observed infection classes. In particular, the characteristic differences between the infection classes for the older age categories for village A, compared with those of villages B or C, are accurately reproduced. The model fails to reproduce the peculiar difference between the no-infection and light-infection classes for all three villages in the youngest age category. We suggest that this is due to the difference in behavior within the lowest age category. Infants and toddlers have very little opportunity to acquire infection. However, in the model they are treated the same way as 4–6-year-old children. It is likely that this discrepancy explains why the model is not consistent with the data (GOF P value = 0.0087).

The base parameters that govern the per snail rate of human infection, β_{base}^{SH} , and recovery, ϕ^H , are well constrained by these data (Table 2). The MGOFE values for transmission, β_{base}^{SH} , differ between the three villages, with, as would be expected from the observed infection classes, village A having lower transmission than villages B or C. The differences in β_{base}^{SH} are not statistically significant, but the point estimate for village A (1.5×10^{-5} ; 95% CI = 9.3×10^{-6} to 2.1×10^{-5}) is considerably lower than that for village B (3.8×10^{-5} ; 95% CI = 1.8×10^{-5} to 6.2×10^{-5}) or for village C (2.3×10^{-5} ; 95% CI = 1.8×10^{-5} to 2.8×10^{-5}). In contrast, relative to the base per snail rate of transmission, the conditional per snail rate of progression to a higher class of infection is impossible to differentiate between the three villages due to very large 95% CIs. However, it must be noted that individuals have a far greater probability of moving to higher infection intensities class when already infected (α_1 and α_2). The basic rates of recovery, ϕ^H , are similar for all three villages, except for the recovery rate from moderate to light infection in village C compared with village A (ϕ_2^H).

Parameter samples close to (not significantly different from) the MGOFE for β_{base}^{SH} and ϕ^H are plotted for each

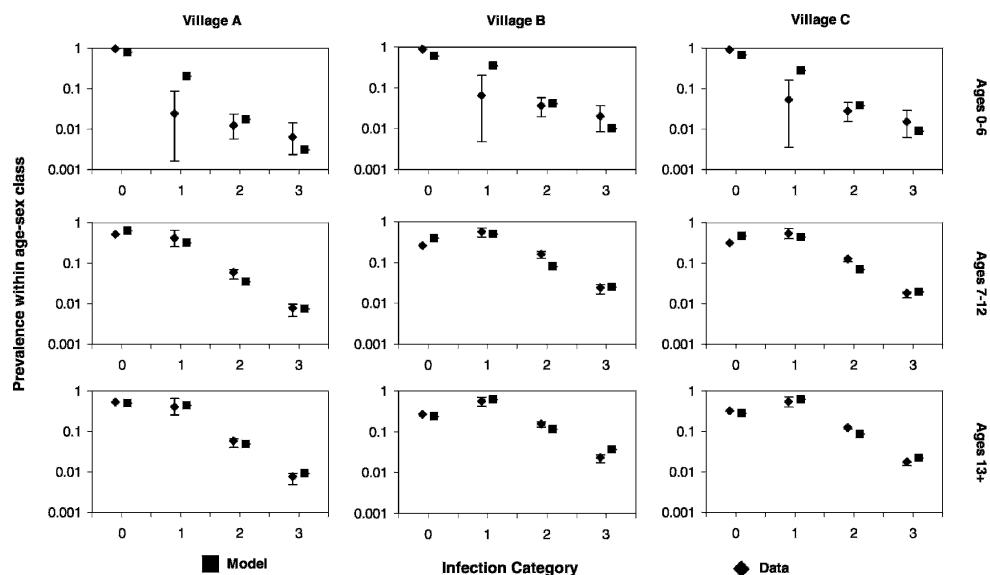


FIGURE 2. Comparison of data and model infection classes for men in each of the three villages in the study. Similar results are obtained for women. However, to aid visual clarity, only the results for men are presented. A log scale is used for the y-axis to distinguish between differences in the higher infection classes. The infection classes are defined by the concentration of eggs per gram (epg) of stool: 0 = no infection or 0 epg; 1 = light or 1–100 epg; 2 = moderate or 101–800 epg; 3 = heavy or ≥ 800 epg. Error bars show the 95% confidence intervals for prevalence. Since these are based on contrasts, they are not shown for the uninfected class (see Appendix 2 for a description of a contrast).

village in Figure 3. If one remembers that the MGOFE is the parameter value best able to reproduce the data, this chart allows one to see how these two parameters are correlated near to their optimal values. Thus, in the two dimensional space of $\beta_{\text{base}}^{SH} - \phi^H$, village A is significantly different from villages B and C, even though in each of the single dimensions the interval estimates overlap. Note also that the shape of the areas covered by the samples for each of the villages suggest a positive correlation between β_{base}^{SH} and ϕ^H . This is as one might have expected because individuals who, on average, recover more quickly would be able to absorb more infections for a given overall burden.

Villages B and C have similar transmission characteristics, whereas village A is quite different. Therefore, to aid clarity, further comparisons are restricted to be between villages A and B.

A single parameter, ϵ_1 , was used to allow for different rates of infection for males and females. The MGOFE for this parameter is similar for all three villages and suggests that men are more likely to be infected than women, although the effect is not statistically significant in this model. Furthermore, the similarity in point and interval estimates for this parameter for all three villages suggests that there is no major difference in the force of infection resulting from the different water exposure patterns of men and women from village to village.

The impact of mass chemotherapy on villages A and B is shown in Figure 4A. The relatively detailed structure of the human population within the model allows us to show explicitly the proportion of the population with heavy infections. The steady-state prevalence proportion ratio of heavy infection in village B compared with village A is 3.5. After chemotherapy, as indicated by the steepness of the curve post-treatment, the rate of re-acquiring heavy infection in village B is higher than that seen in village A. This leads to a similar period of time being needed for both villages to return to

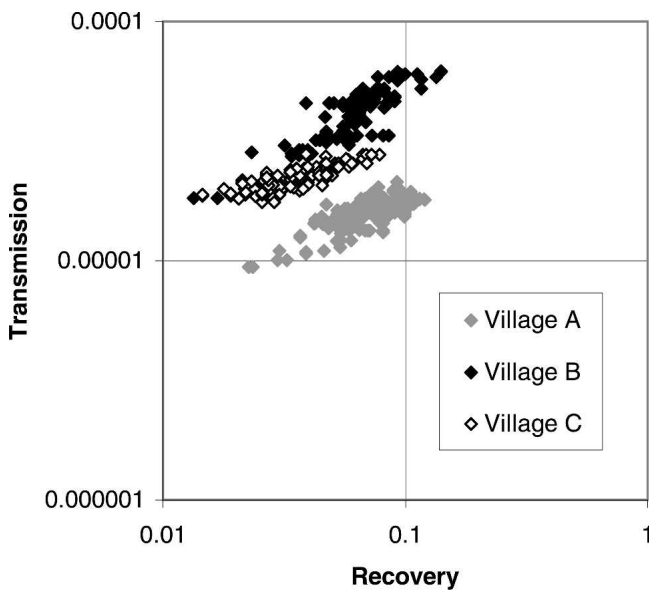


FIGURE 3. Maximum goodness-of-fit parameter estimates (MGOFEs) for each of the three villages. The figure shows the parameter samples that are not significantly different from the MGOFE for each village for the base rates of transmission (β_{base}^{SH}) and recovery (ϕ^H).

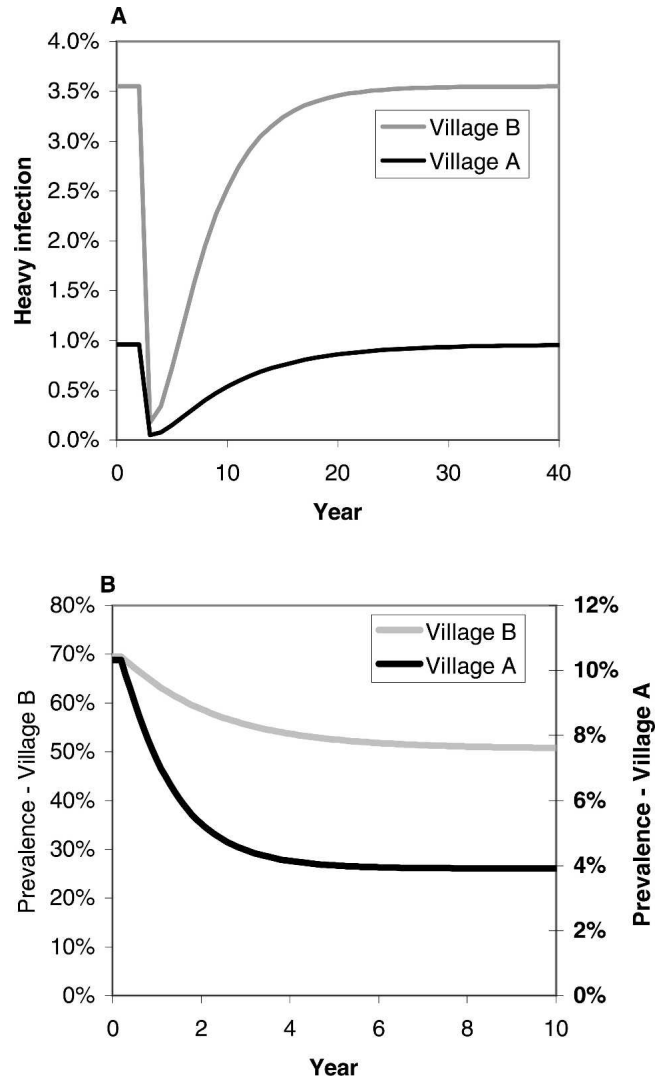


FIGURE 4. Impact of interventions. **A**, Prevalence of heavy infection for villages A and B before and after a community-wide treatment of chemotherapy. It is assumed that the drug treatment is 100% effective with a coverage of 95% in each community. **B**, Possible impact of a sudden change in snail numbers on villages A and B. Separate y-axes are used so as to allow comparison of the relative rates at which over prevalence could be affected in communities with different transmission profiles.

their pre-treatment prevalence proportion of heavy infection. Therefore, the current model does not suggest that communities with a larger proportion of heavily infected individuals would take longer to rebound from treatment. Given the cross-sectional nature of the data used to estimate parameters, the units of time on the x-axis should be used only as a guide.

The possible impact of a sudden reduction in snail numbers on overall snail prevalence is shown in Figure 4B. This can be viewed as a simple representation of a large ecologic change. Separate y-axes are used to show the difference in relative impact of that change. As opposed to the similar, proportional, response predicted for mass treatment with chemotherapy, the transmission characteristics of villages A and B respond differently to a sudden change in the numbers of infective snails interacting with the human population. The

higher prevalence in village B is driven by slower recovery as well as greater transmission. Therefore, a sudden decrease in transmission from snails to humans in village B has a proportionately smaller impact than might be expected in the lower prevalence village A because more people are infected before and after the ecologic change in village B. As with the results described above, the timescale of the dynamic changes should be viewed as indicative only as the parameters are estimated using cross-sectional data.

DISCUSSION

The age-sex specific infection profiles (as defined above) produced by the dynamic model faithfully reproduce the main features of those observed in the three villages. Significant differences have been found in some key underlying parameters for those villages that have allowed us to quantify local variation in the transmission characteristics of *S. japonicum* in disease-endemic areas of The Philippines. These differences may manifest themselves in the response of these communities to some interventions (such as a sudden change in snail numbers), whereas they may not be important for other interventions (such as mass chemotherapy). Although we have not drawn on detailed ecologic data, these results are complementary to other recent studies seeking to characterize local variation of *S. japonicum* transmission in China.^{18,19}

We have shown that the prevalences of higher levels of infection are associated with both higher transmission and longer recovery times (Figure 3). The recovery process is only dependent on the human population already infected, whereas the transmission process is dependent on both snails and humans. This suggests that there are significant differences in both human behavior and snail ecology between village A and villages B and C. Therefore, more detailed data on both snail and human populations, when available, may suggest different optimal intervention strategies for certain snail or human population characteristics.

The model presented here describes medium and high levels of infection for small spatially distinct populations. Therefore, it should be of considerable use in building detailed morbidity models for *S. japonicum*, which could be used to characterize morbidity and the impact of interventions on morbidity within a small community at a specific location. Use of systematic parameter estimation techniques, i.e., MGOFE with prevalence contrasts, should allow for more robust choices to be made between different strategies. Although this was facilitated by the relatively simple structure of the model, in future work, even as the model is allowed to become more complex to capture increasing levels of biologic detail, it should be possible to extend these inferential techniques. This approach has the advantage of allowing traditional inferential statements and contrasts well with more categorical approaches to model-fit used elsewhere for highly structured models of *S. japonicum*.⁶

A potential animal reservoir was investigated for each of the three villages. With the MGOFE parameter values, the presence of an animal reservoir leads to a small difference in the period of time required for the infection classes to return to their initial levels in the villages after mass-chemotherapy. However, the magnitude of the effect was small, even when taking into account the cross-sectional nature of the data.

Without any information on the potential size and cross-infectivity of any animal reservoir, we choose not to show these results here. However, in the absence of any demographic data on potential reservoir species or even concrete evidence for the presence of strain specific cross-infectivity, the possibility that animal reservoirs may impact public health measures against human schistosome infections in The Philippines cannot be ruled out. Field studies currently underway in The Philippines will measure infection in other mammals and compare strains of parasite obtained from other mammals with those obtained from humans.

Transmission of schistosomiasis japonicum is spatially heterogeneous; this is a necessary consequence of the fact that transmission occurs primarily when humans are in direct contact with infected water. Here, we have captured some of that spatial heterogeneity by looking at data from three different spatially distinct villages. It should be possible to build a more informative spatially explicit model. Geographic information systems can be used to characterize local water courses, generating statistical robust inputs for such a model. The very local level of transmission, coupled with the small human population size of high prevalence communities, make the transmission of schistosomiasis particularly suitable to be used as an example for an explicitly spatial infectious disease model.

It is unlikely that *S. japonicum* will ever be eliminated from The Philippines. However, it is not yet clear what impact large-scale ecologic changes, such as a widespread change in the use of modern irrigation techniques or change in plowing techniques from water buffaloes to tractors, will have on the overall infection profile of the disease. Therefore, the optimization of control strategies, and the ability to predict the impact of major ecologic change will remain important issues for the foreseeable future.

Received April 13, 2004. Accepted for publication December 4, 2004.

Financial support: This project was supported by the National Institutes of Health (NIH)/National Science Foundation Ecology of Infectious Diseases program (NIH grant R01 TW01582).

Authors' addresses: Steven Riley, Department of Community Medicine, Faculty of Medicine Building, 5/F, William M. K. Mong Block, 21 Sassoon Road, Hong Kong. Telephone: +852-2819-9283, Fax: +852-2855-9528, E-mail: steven.riley@hku.hk. Hélène Carabin, Department of Biostatistics and Epidemiology, College of Public Health, University of Oklahoma Health Sciences Center, Room 303, 801 NE 13th Street, Oklahoma City, OK 73116. Telephone: 405-271-2229 extension 48083, Fax: 405-271-2068, E-mail: helene-carabin@ouhsc.edu. Clare Marshall, Department of Epidemiology and Public Health, Division of Primary Health Care and Public Health, Faculty of Medicine, Imperial College, St. Mary's Campus, Norfolk Place, London, W2 1PG, United Kingdom. E-mail: clare.marshall@imperial.ac.uk. Remigio Olveda, Research Institute for Tropical Medicine, Department of Health Compound, FILINVEST Corporate City, Alabang, Muntinlupa City 1781, The Philippines. Telephone: +632-809-7599, Fax: +632-842-2245, E-mail: Email: r.olveda@ritm.gov.ph. A. Lee Willingham, International Livestock Research Institute, PO Box 30709, 00100 Nairobi, Kenya. Telephone: 254-20-422-3069 or 650-833-6660 extension 4955, Fax: 254-20-422-3001 or 650-833-6661, E-mail: a.willingham@cgiar.org. Stephen T. McGarvey, Institute of International Health, Brown University, 171 Meeting Street, Box G-B495, Providence, RI 02912. Telephone: 401-863-1354, Fax: 401-863-1243, E-mail: Stephen_McGarvey@Brown.edu.

Reprint requests: Hélène Carabin, Department of Biostatistics and Epidemiology, College of Public Health, University of Oklahoma

Health Sciences Center, Room 303, 801 NE 13th Street, Oklahoma City, OK 73116.

REFERENCES

- Williams GM, Sleight AC, Li Y, Feng Z, Davis GM, Chen H, Ross AG, Bergquist R, McManus DP, 2002. Mathematical modelling of schistosomiasis japonica: comparison of control strategies in the People's Republic of China. *Acta Trop* 82: 253–262.
- Mott KE, Nuttall I, Desjeux P, Cattand P, 1995. New geographical approaches to control of some parasitic zoonoses. *Bull World Health Organ* 73: 247–257.
- Ross AG, Bartley PB, Sleight AC, Olds GR, Li Y, Williams GM, McManus DP, 2002. Schistosomiasis. *N Engl J Med* 346: 1212–1220.
- Olveda RM, Daniel BL, Ramirez BD, Aligui GD, Acosta LP, Fevidal P, Tiu E, de Veyra F, Peters PA, Romulo R, Domingo E, Wiest PM, Olds GR, 1996. Schistosomiasis japonica in the Philippines: the long-term impact of population-based chemotherapy on infection, transmission, and morbidity. *J Infect Dis* 174: 163–172.
- Wiest PM, Wu G, Zhang S, Yuan J, Peters PA, McGarvey ST, Tso M, Olveda R, Olds GR, 1992. Morbidity due to schistosomiasis japonica in the People's Republic of China. *Trans R Soc Trop Med Hyg* 86: 47–50.
- Zhou XN, Malone JB, Kristensen TK, Bergquist NR, 2001. Application of geographic information systems and remote sensing to schistosomiasis control in China. *Acta Trop* 79: 97–106.
- He YX, Salafsky B, Ramaswamy K, 2001. Host-parasite relationships of *Schistosoma japonicum* in mammalian hosts. *Trends Parasitol* 17: 320–324.
- Fernandez TJ Jr, Petilla T, Banez B, 1982. An epidemiological study on *Schistosoma japonicum* in domestic animals in Leyte, Philippines. *Southeast Asian J Trop Med Public Health* 13: 575–579.
- Chan MS, Anderson RM, Medley GF, Bundy DA, 1996. Dynamic aspects of morbidity and acquired immunity in schistosomiasis control. *Acta Trop* 62: 105–117.
- Chan MS, Bundy DA, 1997. Modelling the dynamic effects of community chemotherapy on patterns of morbidity due to *Schistosoma mansoni*. *Trans R Soc Trop Med Hyg* 91: 216–220.
- Chan MS, Guyatt HL, Bundy DA, Medley GF, 1996. Dynamic models of schistosomiasis morbidity. *Am J Trop Med Hyg* 55: 52–62.
- Hairston NG, 1965. An analysis of age-prevalence data by catalytic models. A contribution to the study of bilharziasis. *Bull World Health Organ* 33: 163–165.
- Barbour AD, 1996. Modeling the transmission of schistosomiasis: an introductory view. *Am J Trop Med Hyg* 55: 135–143.
- Anderson RM, May RM, 1991. *Infectious Diseases of Humans: Dynamics and Control*. Oxford, United Kingdom: Oxford Science Publications.
- Olveda RM, Tiu E, Fevidal P, Deveyra F, Icatlo FC, Domingo EO, 1983. Relationship of prevalence and intensity of infection to morbidity in schistosomiasis japonica - a study of three communities in Leyte, The Philippines. *Am J Trop Med Hyg* 32: 1312–1321.
- Carabin H, Marshall CM, Joseph L, Riley S, Olveda R, McGarvey ST, 2005. Estimating the intensity of infection with *Schistosoma japonicum* in villagers of Leyte, Philippines. Part I: A Bayesian cumulative logit model. The schistosomiasis transmission & ecology project (STEP). *Am J Trop Med Hyg* 72: 745–753.
- de Bont J, Shaw DJ, Vercruyse J, 2002. The relationship between faecal egg counts, worm burden and tissue egg counts in early *Schistosoma mattheei* infections in cattle. *Acta Trop* 81: 63–76.
- Liang S, Maszle D, Spear RC, 2002. A quantitative framework for a multi-group model of schistosomiasis japonicum transmission dynamics and control in Sichuan, China. *Acta Trop* 82: 263–277.
- Spear RC, Hubbard A, Liang S, Seto E, 2002. Disease transmission models for public health decision making: toward an approach for designing intervention strategies for schistosomiasis japonica. *Environ Health Perspect* 110: 907–915.
- Vanamail P, Subramanian S, Das PK, Pani SP, Rajagopalan PK, Bundy DA, Grenfell BT, 1989. Estimation of age-specific rates of acquisition and loss of *Wuchereria bancrofti* infection. *Trans R Soc Trop Med Hyg* 83: 689–693.
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP, 2002. *Numerical Recipes in C++: The Art of Scientific Computing*. Cambridge, United Kingdom: Cambridge University Press.
- Aitchison J, 1982. The statistical-analysis of compositional data. *J R Stat Soc B* 44: 139–177.
- Robert CP, Casella G, 1999. *Monte Carlo Statistical Methods*. New York: Springer-Verlag.

APPENDIX 1 MODEL DEFINITION

Humans $H_{i,j,k}$ are indexed by intensity of infection class i , age category j and sex group k . Infection class $i = 0$ represents susceptible humans with each $0 \leq i \leq i_{\max}$ corresponding to higher infection classes. Similarly, age categories are bounded such that $0 \leq j \leq j_{\max}$. Sex group $k = 0$ represents females and $k = 1$ males. The rates at which humans move between age and intensity classes are determined by the numbers in the different classes and the per capita rates (i.e., for age classes, these rates would be the per capita aging and birth and death rates). For simplicity, it is assumed that death only occurs in the final age category and rates are chosen so as to generate an average lifetime of 60 years with a population growth rate of 4%. For example, births are proportional to the birth rate v_j^H and the total numbers of females at that time. The units of the rates are years⁻¹ throughout. In general, all parameters are defined for all classes. However, for example, $v_{i,j,k}^H = 0$ for all $i \neq 0$ and $j \neq 0$, so that births only occur into the uninfected classes of the first age groups. Therefore, with additional parameters defined below,

$$\begin{aligned} \frac{dH_{i,j,k}}{dt} = & v_j^H \sum_{\text{all } i,j} H_{i,j,0} - \mu_j^H H_{i,j,k} + \gamma_{j-1}^H H_{i,j-1,k} - \gamma_j^H H_{i,j,k} \\ & + \phi_{i+1}^H H_{i+1,j,k} - \phi_i^H H_{i,j,k} + \beta_{i-1,j,k}^{SH} H_{i-1,j,k} S_1 \\ & - \beta_{i,j,k}^{SH} H_{i,j,k} S_1. \end{aligned}$$

Human hosts die at rate μ_j^H and age at rate γ_j^H . The constant rate of recovery of humans infected with intensity i ($H_{i,j,k}$) to humans with a lower level of infection $i - 1$ ($H_{i-1,j,k}$) is ϕ_i^H . The rates at which human host's infection levels are increased are defined by the age-sex-intensity specific per capita rate of infection between snail and human ($\beta_{i,j,k}^{SH}$) and the number of infected snails (S_1). Therefore, the matrix $\beta_{i,j,k}^{SH}$ contains the key transmission parameters for this system. However, the values of $\beta_{i,j,k}^{SH}$ summarize diverse human behavioral and immunologic heterogeneities and must be estimated. More detail is given below on the functional form and parameters used to define $\beta_{i,j,k}^{SH}$. Where the indices of the human class are outside the ranges defined for i , j and k , it is assumed that the class takes a value of 0.

Snail dynamics are defined by the following two equations, with parameters defined below,

$$\begin{aligned} \frac{dS_0}{dt} = & v^S (S_0 + S_1) - S_0 \sum_{\text{all } i,j,k} \beta_{i,j,k}^{HS} H_{i,j,k} - S_0 \beta^{RS} R_1 - \mu^S S_0 + \phi^S S_1 \\ \frac{dS_1}{dt} = & S_0 \sum_{\text{all } i,j,k} \beta_{i,j,k}^{HS} H_{i,j,k} + S_0 \beta^{RS} R_1 - \mu^S S_1 - \phi^S S_1. \end{aligned}$$

Snails are born at rate v_s and die at rate μ^S . Snails recover from infection at rate ϕ^S . It is assumed that the rate at which snails are infected is proportional to the number of infectious humans. In general, the rate at which humans infects snails is dependent on the number of humans in each infection intensity-age-sex class. The parameter matrix $\beta_{i,j,k}^{SH}$ defines the specific rates at which this occurs. Similarly, β^{RS} defines the rate at which infected non-human reservoir mammals infect snails.

The class R_1 is infected reservoir mammals. Reservoir mammals are born at rate v^R and die at rate μ^R . They are infected by infectious snails at rate β^{SR} and they recover from infection at a rate ϕ^R . Therefore, the rates of change of the two reservoir mammal classes can be defined as

$$\begin{aligned}\frac{dR_0}{dt} &= v^R (R_0 + R_1) - \beta^{SR} S_1 R_0 - \mu^R R_0 + \phi^R R_1 \\ \frac{dR_1}{dt} &= \beta^{SR} S_1 R_0 - \mu^R R_1 - \phi^R R_1.\end{aligned}$$

These equations can be reduced to the three equations given in both references 1 and 13. If it is assumed that births balance deaths, then the population is of constant size. Each absolute number representing the size of each infected group in the snail (S_1), reservoir (R_1), and human ($H_{1,j,k}$ for $i > 0$) populations can be divided by the total population size of their respective species-age-sex group to obtain prevalence estimates. In the case of humans, the proportion of each age-sex group in each intensity of infection class can also be obtained. With only one infection class ($i_{\max} = 1$) one age class ($j_{\max} = 0$) and one sex class ($k_{\max} = 0$), we are left with only infection and recovery terms. A single infection class and constant population size allows the human system to be specified completely by just one equation.

The parameter matrix $\beta_{i,j,k}^{SH}$ is of particular interest as it characterizes the transmission dynamics from snails to humans. The behavioral and immunologic heterogeneities of the human population are summarized in these parameters. For example, if it is more likely that males rather than females acquire additional infection, then we would expect $\beta_{0,j,1}^{SH} > \beta_{0,j,0}^{SH}$.

To reduce the numbers of parameters used in the model, we do not estimate each element of $\beta_{i,j,k}^{SH}$ separately. The following functional form is used to populate the matrix:

$$\beta_{i,j,k}^{SH} = \beta_{\text{base}}^{SH} \alpha_i \varepsilon_k C \lambda \left\{ \frac{e^{-\delta(j+1)} - e^{-\lambda(j+1)}}{\lambda - \delta} \right\}.$$

Infection and sex effects are scaled by the parameter vectors α_i and ε_k respectively. Both $\alpha_0 = 1$ and $\varepsilon_0 = 1$. We choose not to use a parameter vector for all values of age-dependent relative transmission. Instead, a scaled difference

of exponents is used as a two-parameter function capable of different peak ages of infection. This functional form has been used to describe the age profile of filarial infection in an Indian town.²⁰ Parameters λ and δ define the shape of this function and C is the scaling parameter such that, effectively, the peak age-dependent transmission factor is always 1. More precisely,

$$C = \left(\left[\frac{\lambda \{ e^{-\delta(j+1)} - e^{-\lambda(j+1)} \}}{\lambda - \delta} \right]_{\max} \right)^{-1},$$

where the function inside the square brackets is maximized over $0 \leq j \leq j_{\max}$. Note that independence between the effects of infection status, age, and sex is assumed implicitly in the above formulation.

APPENDIX 2 INFERENCE

The prevalence of each age-sex infection class are defined in Appendix 1 with steady-state solutions for the differential equations. The equations are solved numerically using the adaptive step size Runge-Kutta algorithm.²¹ Our interest is in fitting a single longitudinal data set using this model. Therefore, only a steady-state solution for the model was required. This need not be steady state in the strict mathematical sense. Our criteria were that the system could be assumed to be stationary if no single infection-age-sex prevalence had changed by a factor of more than 0.01 in the previous 10 years. It is very unlikely that changes smaller than this would ever be observed through community surveillance systems.

Let the steady state prevalence for infection class i , age category j , and sex group k be $y_{i,j,k}$. Each prevalence, where $i \neq 0$, was converted into a contrast.²² The expected contrast $E_{i,j,k}$ is defined to be

$$E_{i,j,k} = \ln \left\{ \frac{y_{i,j,k}}{y_{0,j,k}} \right\},$$

with $i \neq 0$. A χ^2 statistic was then obtained from a likelihood ratio goodness-of-fit (GOP) test using the posterior distributions for each observed contrast.¹⁶ Parameters were estimated and GOF P values were obtained. Simulated annealing was used for optimisation and interval estimation.²³ As well as being robust against local extremes, this algorithm can be used to generate confidence bounds and maximum goodness-of-fit estimates in a single process. Typically, 10^5 samples were taken during which the scaling constant (analogous to temperature) was allowed to drop from 50 to 0.1. Repeated runs were performed for all parameter estimates from different starting points to ensure results were not sensitive to the initial state of the algorithm.