Published in final edited form as:

Parasitology. 2010 November ; 137(13): 1951–1965. doi:10.1017/S0031182010000867.

A new approach to modelling schistosomiasis transmission based on stratified worm burden

D. GURARIE^{1,*}, C. H. KING², and X. WANG¹

¹Department of Mathematics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106 USA

²Center for Global Health and Diseases, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106 USA

SUMMARY

Background/Objective—Multiple factors affect schistosomiasis transmission in distributed meta-population systems including age, behaviour, and environment. The traditional approach to modelling macroparasite transmission often exploits the ‘mean worm burden’ (MWB) formulation for human hosts. However, typical worm distribution in humans is overdispersed, and classic models either ignore this characteristic or make *ad hoc* assumptions about its pattern (e.g., by assuming a negative binomial distribution). Such over simplifications can give wrong predictions for the impact of control interventions.

Methods—We propose a new modelling approach to macro-parasite transmission by stratifying human populations according to worm burden, and replacing MWB dynamics with that of ‘population strata’. We developed proper calibration procedures for such multi-component systems, based on typical epidemiological and demographic field data, and implemented them using Wolfram Mathematica.

Results—Model programming and calibration proved to be straightforward. Our calibrated system provided good agreement with the individual level field data from the Msambweni region of eastern Kenya.

Conclusion—The Stratified Worm Burden (SWB) approach offers many advantages, in that it accounts naturally for overdispersion and accommodates other important factors and measures of human infection and demographics. Future work will apply this model and methodology to evaluate innovative control intervention strategies, including expanded drug treatment programmes proposed by the World Health Organization and its partners.

Keywords

models; theoretical; schistosomiasis; disease transmission; infections; environmental and behavioural risk

INTRODUCTION

Helminth infection typically involves non-linear systems that link, human populations intermediate hosts and multiple environmental factors contributing to efficient transmission in endemic areas. For policymakers, capturing the dynamics of this process has proven

challenging in terms of developing predictive models that have quantitative accuracy. A traditional mathematical approach to modelling macro-parasite systems (Macdonald, 1965) exploits a mean worm burden (MWB) formulation for human hosts and a suitable infection prevalence estimate for intermediate invertebrate hosts. However, worm distributions among host populations often exhibit an overdispersed pattern. In the standard modelling approach, overdispersion is either ignored or over-simplified using *ad hoc* assumptions about its distribution (e.g., the negative binomial distribution). Furthermore, the pattern of infectious burden is assumed to remain static (unchanged) over time. This approach cannot explain overdispersion or justify its 'stationarity', and thus leaves many uncertainties about its application to observed epidemiological data from the field and its predictions about the impact of control programmes.

We propose a new approach to modelling schistosomiasis transmission (or any vector-borne macro-parasite), based on a *Stratified-Worm-Burden* (SWB) system as an alternative to the classical MWB approach. Several papers have studied such or similar systems for macro-parasites in the past (Pugliese, 2000; Feng *et al.* 2001; Rosa and Pugliese, 2002), but most of this work aimed at deriving reduced (MWB-type) equations, or their extended versions that couple MWB to higher moments (variance and/or aggregation) of SWB distribution. Among several examples are papers on (i) regulation and stability of host-parasite populations (Anderson and May, 1978a, b); (ii) aggregation for single parasite species and multi-strain competition, coexistence, invasion (Dobson and Roberts, 1994; Pugliese, 2000; Rosa and Pugliese, 2002) and (iii) parasite strain diversity and evolution of drug resistance (Feng *et al.* 2001).

In contrast, for the present study, we developed and implemented an extended SWB approach to model, as necessary, the complexity of schistosomiasis transmission for structured host populations in a geographically distributed environment with varying water contact patterns. We take advantage of the extended model systems to consider age-structured communities made of children and adult groups having suitable rates of birth, mortality, and ageing.

Our approach extends the standard Macdonald-type MWB and other prevalence-based models (e.g., Riley *et al.* 2008). We gain several advantages by combining both infection intensity and prevalence in the model. Primarily, we can naturally account for overdispersed worm burden in host populations (even for homogeneous behavioural risk subgroups) without *ad hoc* assumptions on distribution patterns commonly used in MWB-type approach such as the negative binomial (Anderson and May, 1991; Alexander *et al.* 2000). Inclusion of additional features reflecting the distribution of human infection (the prevalence of infected/non-infected, or lightly infected/heavily infected, and mean egg counts (as a measure of infection intensity)), proves particularly helpful for comparison of model outputs with the observed field data for calibration/validation purposes. In using the model, these variables can be derived from data or entered as dynamic variables. In addition, our model accommodates host demographics, so that one can study the effect of multiple factors (both stationary and variable) on transmission and long-term outcomes of infection. These changing variables can include population changes, environmental/behavioural modifications, snail variability, and intermittent control interventions.

Some earlier works (Chan *et al.* 1996; Medley and Bundy, 1996; Chan and Bundy, 1997) used age-structured MWB to assess infection-associated morbidity and the effects of drug treatment for idealized host populations. The MWB approach, however, severely limited our ability to project outcomes for promising control interventions that target both age and geographical risk factors together (see e.g., Gurarie and King, 2005, 2008; Gurarie and Seto, 2009). The new SWB approach opens new possibilities for modelling the outcomes of

control programmes that focus interventions by targeting high burden strata by age or location.

MATERIALS AND METHODS

We start with an outline of general SWB-scheme, its implications for burden/prevalence distribution, and the analysis of endemic equilibria. Then we proceed to develop a model calibration scheme that exploits typical demographical and epidemiological infection data to estimate the baseline human-to-snail and snail-to-human transmission rates. These models are built upon each other, starting with a model describing a single homogeneous population, adding an age-structured (adult-child) population, and finally a geographically distributed (human-snail) transmission environment.

Our choice of calibration scheme is partly motivated by the available data from the Msambweni region of eastern Kenya, which were collected in several multi-year studies (Sturrock *et al.* 1990; Muchiri *et al.* 1996; Hamburger *et al.* 2004; Kariuki *et al.* 2004; Clennon *et al.* 2006). In a sequel paper we will apply our methodology (SWB model and calibration scheme) to this specific environment, comprised of 10 villages, each one subdivided into adult and child subgroups with 5 shared snail sites. When calibrated and validated, this model will then allow us to make projections on the transmission of vector-borne disease, and to simulate the effects of hypothetical control interventions based on variously targeted, single or combined drug therapy- and vector control-based schemes.

Infection strata

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

Infection level	0	1	...
Worm burden	$0 < w < w_1$	$w_1 < w < w_2$	...
Population strata	h_0	h_1	...

So each group h_i viewed as a local subpopulation carries a parasite range $w_i < w < w_{i+1}$. The dividers $\{w_i\}$ can be chosen arbitrarily, the simplest (extreme) choice being ‘zero-nonzero’ (‘infected’ – ‘non in-fected’) strata with prevalences $\{h_0, h_1\}$. Such partition would result in the standard SIR-type prevalence model. The SIR methodology ($h_0; h_1$) however, is too crude for macro-parasite transmission, where force of infection depends primarily on the worm burden of the host population, rather than the infection frequency. On the opposite extreme is continuously distributed worm burden $\{h(w): 0 \leq w < \infty\}$, which has limited practical applications outside mathematical analyses. Our stratified worm burden (SWB) models extends the SIR (prevalence-based) modelling to accommodate both the worm burden and infection frequency relevant to population dynamics of macroparasite infections using the Macdonald-type ‘mean burden’ formulation, as explained below.

In what follows, we shall use uniform strata $w_i = i\Delta w$; ($i = 0, 1, 2, \dots$) counted in steps Δw worms. The choice of Δw is related to infectivity threshold w_T – the ‘effective’ number below which a population is considered free of infection. It is convenient to take w_T a multiple of Δw ; thus $w_T = \Delta w$ means h_0 = ‘uninfected strata’, $w_T = 2\Delta w$ the first two strata ($h_0 + h_1$) = ‘uninfected’, etc. Indeed, host infectivity is proportional to the number of mated worm pairs, which in turn depends on their *mating probability* $\phi(w)$ (Nasell, 1978; see also Anderson-May, 1991). Derivation of such a probability function is based on ‘negative-binomial’ assumption, suggesting a some-what low threshold value $w_T = 5$ (independent of

aggregation). On the other hand, experimental/field data for primate infection (Wilson, 2006) predict a substantially higher threshold for egg detection ($w_T \approx 40$). We assume infectivity threshold to lie somewhere in between $10 \leq w_T \leq 20$, and numerically explore 2 values: $w_T = 10$ and $w_T = 20$. Accordingly, we choose to convert the continuous values to discrete values at 10 worms per step (discretization step, Δw). While the choice of step Δw is not essential for calibration and analysis (as explained below), the choice of threshold w_T has significant effect (see Fig. 6).

For higher burden levels ($w > 10$) we assume that host infectivity, as measured by egg counts, is roughly proportional to w_i . The choice of uniform strata is not essential from computational standpoint, but has some mathematical advantages. We used 15 strata (i.e. burden levels up to 150 worms/host) in our analysis, which, according to our data set, is sufficient to cover a broad range of human infection, even heavily infected populations.

Fig. 1 shows the diagram of transitions among strata with ‘up-arrows’ $\rightarrow$ designating the *force of human infection* λ (determined by infected snail/larval densities, water contact patterns, probability of worm establishment, etc.), and ‘down arrows’, $\leftarrow$ indicating *burden resolution* λ_i (due to worm mortality or other causes, including drug treatment). Different strata i have specific resolution rates γ_i that increase with burden level. Additional source terms, S_i , represent new recruits to the infection pool (newborns, migrants into the area, and transition from younger to older age groups). Loss rates, μ , combine host mortality as well as out-migration, aging, etc. In some papers μ ’s are considered to depend on level i to account for ‘burden-specific’ mortality. We ignore such disease-specific mortality as quite small, i.e., not having impact on population levels or community levels of schistosomiasis (even its more pathogenic *S. japonicum* species (van der Werf and de Vlas, 2001)). So our μ are purely demographic and considered identical for homogeneous populations.

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

For a single (homogeneous) population $H = \sum_0^n h_i$, variables $\{h_i(t)\}$ obey a system of differential equations (SWB)

$$\begin{aligned} \frac{dh_0}{dt} &= -(\lambda + \mu) h_0 + \gamma_1 h_1 + S_0; \\ &\dots \\ \frac{dh_i}{dt} &= \lambda h_{i-1} - (\lambda + \gamma_i + \mu) h_i + \gamma_{i+1} h_{i+1} + S_i; \\ &\dots \\ \frac{dh_n}{dt} &= \lambda h_{n-1} - (\gamma_n + \mu) h_n + S_n; \end{aligned} \quad (1)$$

We used the following notations (i) transition rate λ is proportional to (per capita) force of

infection: $\lambda = \frac{A_z}{\Delta w}$, made of ‘transmission coefficient’/‘infected snail’ – A, infected snail density – z, and Δw – ‘burden increment’ in the definition of strata $\{h_i\}$ (see Appendix A and Table 1); (ii) resolution rates $\{\gamma_i = \gamma_i\}$ are proportional to burden levels $\{w_i\}$ of each strata (γ being natural worm mortality); (iii) host loss rate(s) = μ and new recruit sources = $\{S_i\}$. Here we assume identical force of infection and loss for all strata, i.e., a homogeneous community in terms of behaviour and environmental factors. This, however, does not preclude uneven (aggregated) burden distribution within system (1). Fig. 2A shows a few

equilibrium distributions of human prevalence $\{h_i(t)\}$ (1) for several values λ contributed by a stationary source S_0 at the lowest (uninfected) stratum (a typical pattern for child group). Fig. 2B illustrates dynamic relaxation (to equilibrium) of the coupled SWB-system (1) and snail prevalence $y(t)$ (see Appendix A). Here human force of infection $\lambda = Ay$ and snail force Λ is proportional to the cumulative worm burden (2) of human population. For dynamic simulation Fig. 2B we initialized the coupled system at a low-level human infection (light bars) – a hypothetical effect of ‘drug clearing’, and let it relax to its natural equilibrium aggregated at higher burden levels (black bars).

System (1) can be interpreted as cumulative out-come of a stochastic process of ‘worm acquisition-death’ on individual level. Indeed, the RHS of system (1) has a matrix made of ‘subdiagonal’ λ and ‘diagonal’ $\{\lambda_j\}$ that can be viewed as discretized Markov generator of such stochastic process. In this sense the standard Macdonald (1965) MWB gives the ‘statistical mean dynamics’ for such process (as function of time), whereas our SWB – its complete probability distribution function (PDF).

Human system (1) can be coupled in the usual way to snail infection (pre-patent and patent groups), through the force of human infection λ and the corresponding, extra-human, larval densities (Appendix A). Here the *force of snail infection* Λ will depend on human MWB, similar to the standard Macdonald (1965). But, unlike the Macdonald system, our MWB is not an independent variable but computed from prevalence variables $\{h_i\}$

$$w = (\Delta w) \sum_1^n (ih_i/H). \quad (2)$$

Rescaling

In the follow-up analysis and calibration we rescale SWB-system (1) and MWB (2) using non-dimensional (primed) variables and parameters, along with time and source terms. Specifically,

$w \rightarrow w' = \frac{w}{\Delta w} = \{0, 1, 2, \dots\}$	– dimensionless burden
$h_i \rightarrow h'_i = h_i / H$	– prevalence strata
$S_i \rightarrow S'_i = S_i / (\Delta w H)$	– source terms in prevalence strata

Then the i^{th} equation for system (1) becomes

$$\frac{dh'_i}{dt} = \lambda h'_{i-1} - (\lambda + \gamma i + \mu) h'_i + \gamma (i+1) h'_{i+1} + S'_i;$$

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

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

$$\frac{dH}{dt} = S - \mu H \quad (3)$$

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

$$\frac{dw}{dt} = \lambda (1 - h_n) - (\gamma + \mu) w + \bar{S} \quad (4)$$

Here $\bar{S} = \sum_{i=0}^n i S_i$ designates the cumulative source. Let us note that $\bar{S} = 0$, if new recruits enter system (1) only through its non-infected (zero) compartment (like newborn children). Derivation of (4) exploits a special form of resolution rates $\gamma_i = \gamma$ for i -th strata. So worms in all hosts are assumed to have the same (mean) life-span, independent of the host state (age, burden level etc.). We also ignore all other confounding factors, like density-dependent mortality (due to crowding), or immune regulation of parasites. Such factors can in principle be included in the SWB scheme, but to calibrate such a model would require a more detailed demographic/infection data of host population (e.g. age, behavioural risk, susceptibility etc.), than provided in most community studies of schistosomiasis. Our formulation tries to keep what we believe are the most essential components of the system, suited for the most frequently available data.

Equation (4) looks like the standard Macdonald MWB for dimensionless $w(t)$ with an additional term h_n that disappears in the infinite limit systems ($n \rightarrow \infty$), used in many theoretical studies. In our case (finite size SWB) this term remains positive, but sufficiently small over a wide range of parameter values, to be safely dropped in the follow-up analysis and calibration.

Equations (3) (total population growth) and (4) (MWB dynamics) represent the first two statistical moments of distribution $\{h_i\}$ driven by SWB system (1). It can be further extended to higher moments: variance, aggregation etc. Such schemes were adopted in some theoretical studies (e.g. Pugliese 2000) as a way to reduce a large (infinite) system (1) to its low-dimensional (mean-field) approximation. In this paper, our goal and approach is different. We utilize reduced (moment) equations (3)–(4) for model calibration. But all dynamic simulations (for prediction-analysis-control) are run with full SWB system (1), using its calibrated parameters. Thus a system of 10 villages in a particular study area (Msambweni), with each population subdivided into child and adult groups will contain $2 \times 10 \times 15$ ‘human’ variables plus a suitable number of ‘snail’ variables proportional to the number of contact sites.

It should be stressed that increased order (dimension) of SWB system (compared to its MWB cousin) posed no substantial mathematical or computational obstacle by itself. Indeed, we can generate such a large, coupled dynamical system using Wolfram Mathematica 7 (Champaign, IL). It also allows efficient ways to solve and manipulate them in various ways for analysis, calibration, prediction and control experiments. Details of these programs are available from the corresponding author.

Analysis of equilibria

The equilibrium solution of system (1), with stationary source S and prescribed λ , depends on 2 parameters μ and λ . The former μ is determined by known demographical or biological

factors, the latter λ (proportional to the force of human infection) can vary depending on the transmission environment, and will be calibrated along with other parameters.

The complete SWB system (1) has no exact analytic solutions (even for stationary equilibria), but we can explore a broad range of parameters $\{\lambda, \mu\}$ numerically. Fig. 2A shows equilibrium burden distributions at several λ . Qualitatively they resemble negative binomial (NB) distributions of decreased aggregation (smaller λ having higher aggregation, hence a larger fraction of low-level infections).

MWB equation (4) can be used either in dimensional form for variable $w(t)$, or non-dimensional $w = \sum_{i=1}^n ih_i$. Our calibration scheme exploits 2 epidemiological data values, uninfected human prevalence $P = h_0$, and MWB, w (which is related to measured egg-counts in the field). Both are functions of parameters λ, μ , and source S

$$\begin{aligned} P &= \phi(\lambda, \mu, S) \\ w &= \sum_{i=1}^n ih_i = \psi(\lambda, \mu, S) \end{aligned} \quad (5)$$

Both functions ϕ and ψ are illustrated in Fig. 3. While there are no analytic formulae, they can be easily computed numerically by interpolation (sampling a range of values λ and μ). These procedures are used extensively in our calibration schemes below.

Let us note that the range of uninfected human prevalence from our data (Table 2), $0.2 < P < 1$, constrains λ/γ within relatively narrow interval $[0, 2.5]$ (Fig. 3A). In that range, both λ and w obey a nearly linear (Macdonald type) relationship (Fig. 3B),

$$\lambda \approx (\gamma + \mu) w \quad (6)$$

which is important for our calibration scheme. At large $\lambda > 8$ (a strong force of infection) we observe nonlinear departure from simple Macdonald model outcomes (Fig. 3B), due to the extra-term of our MWB-equation (4). But this range (and level of transmission) is well above the likely range of values in our study environment.

MODEL CALIBRATION

Data that are typically available for calibrating models of macro-parasite transmission include: (i) Demographics: human population H , and snail density N ; (ii) Biological factors: human loss rates (mortality, turnover) $-\mu$, snail mortality $-\nu$, worm mortality λ ; (iii) Epidemiological infection data: infected and patent snail prevalences (y, z), human egg-counts (arithmetic and geometric means) E, G . Depending on the choice of model, all or part of these data can be used for estimating unknown transmission parameters (see, e.g. Riley *et al.* 2008).

In our calibration scheme below, we use both egg-counts (arithmetic mean) and infection prevalence as 2 independent parameters. For the sake of exposition we precede the discussion of larger, distributed systems, with a few special cases.

The general scheme involves 3 steps:

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

- ii. Next estimate snail-to-human transmission coefficient $\{A\}$ for appropriate population groups, as well as biological parameter ρ (egg-production/worm).
- iii. Finally, estimate human-to-snail transmission coefficients $\{B\}$ (see Table 1 for notations).

The single population case

In this case our calibration scheme will utilize 2 reduced equations resulting from SWB (1): (i) MWB (4) coupled to a snail prevalence equation (as in the standard Macdonald setup), and (ii) an equilibrium zero-prevalence function (5), computed numerically. The former gives a coupled system of 2 differential equations for $w(t)$ and $y(t)$

$$\begin{aligned}\frac{dw}{dt} &= \lambda \Delta w - (\gamma + \mu) w; \\ \frac{dy}{dt} &= \Lambda (1 - y) - \nu y;\end{aligned}\tag{7}$$

with per capita forces of infection

$$\begin{aligned}\lambda \Delta w &= A N y \\ \Lambda &= B H w.\end{aligned}\tag{8}$$

The transmission coefficients A (snail to human) and B (human to snail) depend on several biological and environmental/behavioural factors (Table 1). Let us note that having stratified human population is irrelevant for snail infection, as long as all strata have similar behavioural (water contact) patterns. So the force of snail infection Λ is the same as in the classical Macdonald setup (i.e., proportional to total human burden, Hw)

System (7) has stable endemic equilibrium provided its Basic Reproduction Number (BRN)

$$R_0 = \frac{A N B H}{(\gamma + \mu) \nu} > 1.$$

Our goal is to estimate transmission coefficients A , B from equilibrium solutions (7) and the available data. The data include egg-count E (related linearly to w) and snail prevalence (y). The relation of E to w is given by another (unknown) parameter egg production rate

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

Our goal is to calibrate parameters A , B , ρ . We proceed in 2 steps:

1. solve equation (5) for a given (known) human prevalence P , relative loss rate μ , (demographics), and source S , to find rescaled force of infection λ .
2. compute A from the first expression of (8), ρ and B equilibrium of equations (7)

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

$$A = \frac{\lambda (\Delta w)}{N y}; \quad \rho = \frac{(\gamma + \mu) E}{\lambda (\Delta w)}; \quad B = \frac{\rho}{H E} \frac{\nu y}{(1 - y)}.\tag{9}$$

In reality, infection field data, either human (egg-counts E , prevalence P) or snail data (pre-patent and patent prevalences y , z), can be highly inaccurate (de Vlas *et al.* 1997; Hamburger *et al.* 2004) due to low sensitivity of the standard tests and observational errors. One way to account for data errors (underestimation) is through a suitable statistical ensemble of data values and the resulting parameters (A , B , ρ). Here explicit formulae (9) prove helpful, as any *a priori* information of data errors (statistics or distribution) can be translated via (9) into suitable estimates for calibrated parameters. Indeed, we shall explore sensitivity of our calibration scheme to data errors for a specific data set (below).

Multiple risk groups

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

$$\begin{aligned}\frac{dw_s}{dt} &= A_s N_y - (\gamma + \mu_s) w_s; \\ \frac{dy}{dt} &= \left(\sum_s B_s H_s w_s \right) (1 - y) - \nu y.\end{aligned}\quad (10)$$

As above, we use equilibrium solutions (10) along with the prevalence equation $\phi(\lambda_s, \mu_s, S_s) = P_s$, for each group. The calibration problem now asks to estimate transmission coefficients $\{A_s, B_s\}$. The above 2 steps yield (i) the rescaled force of infection for each group $\{\lambda_s\}$, (ii) the corresponding A-coefficients

$$A_s = \frac{\lambda_s (\Delta w)}{N_y} \quad (11)$$

To get B-coefficients we observe that both sets (A's and B's) are proportional to contact rates $\{\eta_s\}$, assumed to be the only risk factor,

$$\begin{aligned}A_s &= a \eta_s \\ B_s &= b \eta_s\end{aligned}\quad (12)$$

Here a is worm establishment rate/per single infected snail, assumed to be identical for all groups, b – snail infection rate/per single worm $\times$ contact. Then we get B-coefficients

$$B_s = \frac{A_s}{\sum_q (A_q H_q E_q / \rho_q)} \frac{\nu y}{(1 - y)} \quad (13)$$

expressed through group specific egg-production/worm rates

$$\rho_s = \frac{(\gamma + \mu_s) E_s}{\lambda_s (\Delta w)} \quad (14)$$

Equations (11)–(13)–(14) extend the above calibration scheme (9) to heterogeneous populations with multiple risk groups.

Child–adult system

Next we further specify the above ‘multi-risk’ setup for separate child–adult groups, still within a single human–snail pairing context. These two human age-related subgroups differ in their risk factors (contact behaviour) and infection levels. This case, however, differs from the previous one, in that the two groups are also coupled by a demographic birth–aging process, whereby infection is carried from the younger to the older group. The population dynamics for variables H_c ; H_a (total child/adult populations) is given by a Leslie-type differential system

$$\begin{aligned}\frac{dH_c}{dt} &= -(\mu_c + \tau) H_c + \beta H_a \\ \frac{dH_a}{dt} &= \tau H_c - \mu_a H_a.\end{aligned}\quad (15)$$

The demographic parameters include (i) mean per capita birth rate β , (ii) specific child and adult mortalities μ_c , μ_a , (iii) transition (aging) rate τ .

Note that population dynamics (15) follow from the extended child–adult SWB system, the same way that the previous simple population model (3) follows from (1). The source term for child SWB, $S_c = \beta H_a$, enters through the uninfected (zero) stratum (new-borns), whereas the adult source, S_a , is equal to children’s SWB distribution ((Eq.B.1) in Appendix B).

The real population and transmission dynamics are more involved. Clearly, different ages have different SWB-distributions, gradually changing from early childhood (aggregated at low burden) to older age-group patterns shifted towards higher w (Fig. 2). The adult source would then come from the older children’s distribution (15–20 years), rather than the mean distribution over all ages (5–20). In principle, both systems (demographic and SWB) can be extended to accommodate refined population structure (e.g. 5-year bins). But here we confine ourselves to the coarse child–adult partition, as the only available epidemiological data for our specific environment has only those two groups.

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

$$\begin{aligned}\frac{dw_c}{dt} &= A_c N y - (\gamma + \mu_c + \tau) w_c \\ \frac{dw_a}{dt} &= A_a N y - (\gamma + \mu_a) w_a + \tau \frac{H_c}{H_a} w_c \\ \frac{dy}{dt} &= (B_c H_c w_c + B_a H_a w_a) (1 - y) - \nu y.\end{aligned}\quad (16)$$

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

$$\begin{aligned}\phi(\lambda_c, \mu_c + \tau, S^c) &= P_c \Rightarrow \lambda_c \\ \phi(\lambda_a, \mu_a, S^a) &= P_a \Rightarrow \lambda_a.\end{aligned}\quad (17)$$

Specifically, we first find children’s force of infection λ_c with known prevalence P_c and source S^c (in terms of birth rate), then compute the resulting equilibrium-prevalence

distribution $\widehat{h}_c = \{h_i^c\}_{i=0}^n$, substitute it to get adult source $S^a = \{\tau h_i^c H_c / H_a\}_{i=0}^n$ and get λ_a from the second equation in (17). Having estimated both λ ’s we proceed as in the previous section

to get transmission coefficients $\{A_c; A_a; B_c; B_a; \rho_c; \rho_a\}$ for two groups (see Appendix B for details).

Distributed system

Our next goal is to extend the above models and calibration scheme for geographically distributed systems made up of several different human sites (villages), labelled by $i = 1, 2, \dots, n$, and connected snail (water), sites labelled by $j = 1, 2, \dots, m$. Mathematically, calibration of such dynamical system is an ill-posed (heavily *under-determined*) problem as far as the available human and snail infection data. Indeed, the $n + m$ equations of a typical ‘MWB + snail prevalence’ system (Eq.C.1) have altogether $(2n \times m)$ unknown transmission coefficients $\{A_{ij}; B_{ij}\}$, if all connections are allowed (Woolhouse, 1991; Woolhouse *et al.* 1991; Gurarie and King, 2005; Riley *et al.* 2008). Different reduction schemes can be used to constrain the resulting parameter set. Thus Gurarie and King (2005) and Gurarie and Seto (2009) proposed a version of a landscape ‘gravity’ model, whereby geographical distances (or other hurdles) control human water contact rates. A suitable parameterization of this hurdle-function yields a drastic reduction of the parameter set.

Other procedures are possible depending on the available data. For example, data collected from our specific environment (Msambweni region of eastern Kenya (Sturrock *et al.* 1990; Muchiri *et al.* 1996; Hamburger *et al.* 2004; Kariuki *et al.* 2004; Clennon *et al.* 2006)) has additional information in terms of relative contact rates from various human sites to water (snail) sites. Here we propose a distributed calibration scheme designed specifically for such type of data. Each contact pair (i, j) contributes to human infection in village i and snail infection at site j through transmission coefficients $A_{ij}; B_{ij}$ both proportional to absolute contacts per person from village i to water site j , $T_i \eta_{ij}$ with T_i being the group-specific total contacts to all water sites and η_{ij} relative contact ($\sum_j \eta_{ij} = 1$). Our basic hypothesis is that each of them is a product of a suitable local (biological/environmental) factor (A_0 times ‘absolute contact rate’),

$$\begin{aligned} A_{ij} &= A_0 T_i \eta_{ij} = A_i \eta_{ij}; & (i=1, 2, \dots, n) \\ B_{ij} &= B_0 T_j \eta_{ij} = B_j A_{ij}; & (j=1, 2, \dots, m). \end{aligned} \quad (18)$$

It means that humans at site i will pick up infection from snail site j in proportion to their ‘absolute contact’ $\times$ ‘the site’s snail infection level’. By the same pattern, snails at j will get infected by humans from different sites i in proportion to their absolute contact times human infection level (egg release). Since the establishment of miracidia in snails involves 2 stages – hatching of miracidia and the miracidium’s successful invasion into a snail body – which depends heavily on the local environmental conditions, we set the local environmental factor B_{0j} as site specific. This hypothesis clearly reduces the number of unknown parameters to a manageable set $\{A_i; B_j\}$, consistent with the number of MWB equations.

Accordingly, for age-structured populations, the MWB system ((Eq.C.2) in Appendix C) has the following unknown parameters to be estimated: (i) local transmission coefficients for child-adult groups $\{A_i^c; A_i^a\}$; (ii) local transmission coefficients for snail sites $\{B_j\}$; (iii) egg production/release rate for children and adults $\{\rho_c; \rho_a\}$, the latter two are biological factors independent of location.

The calibration scheme proceeds similarly to the previous case. Namely, we first use local prevalence and loss rate for children and adults to get their local rescaled forces of infection,

$\{\lambda_i^c; \lambda_i^a\}$, as in Table 3. Then we write the proper distributed versions of the calibration formula (11), (14)

$$A_i^s = \frac{\lambda_i^s (\Delta w)}{\sum_j \eta_{ij} Z_j}; \quad \rho_c \approx \frac{(\gamma + \mu_i^c + \tau) E_i^c}{\lambda_i^c (\Delta w)}; \quad \rho_a \approx \frac{(\gamma + \mu_i^a) E_i^a}{\tau (H_i^c / H_i^a) (E_i^c / \rho_c) + \lambda_i^a (\Delta w)}. \quad (19)$$

Symbol $s = \{c, a\}$ indicates child/adult groups, and denominator of A gives a weighted snail infection density (with relative ‘contact weights’). While all A -coefficients are local (i.e., depend on i), the egg production term, ρ_s , should be biological and site independent. So, based on (19) we expect the numerator of the right hand side of ρ_s expression to be proportional to the denominator at all sites, with slope ρ_s . The real data show departures from simple linear relations, so we approximate parameters $\rho_c; \rho_a$ with their ‘best-fit’ slopes. Having estimated $\{A_i^s\}$ and $\{\rho_s\}$, we compute coefficients $\{B_j\}$ from the snail equilibrium equations ((Eq.C.4) in Appendix C).

RESULTS

Calibration results for a specific environment

In this section we apply SWB methodology to a specific environment of the Msambweni region of Eastern Kenya sampled from 10 villages and 5 contact snail sites observed in year 2000, shown in Fig. 4 (Sturrock *et al.* 1990; Muchiri *et al.* 1996; Hamburger *et al.* 2004; Kariuki *et al.* 2004; Clennon *et al.* 2006). Infection prevalence and egg-counts (arithmetic means) were observed using standard filtration of urine samples (10 ml) from the resident populations (Table 2). Basic biological parameters were included from well-known estimates of snail and worm mortality, as well as basic life table estimates (among other parameters in Table 1). Field data (Table 2) for the models included (i) demographic data: human populations (divided into child/adult groups) and snail densities; (ii) infection data: mean egg count for children in each village and density of susceptible, infected and shedding adult snails in all water sites; and (iii) behavioural data: relative contact distribution for each village to adjacent snail sites.

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

- i. Estimate ‘local forces’ of human infection in 10 villages (first children, then adults) from the human data (prevalence and demographics), by solving equations at each site. Here we utilize the known (‘universal’) childhood prevalence function $\varphi(\lambda, \mu, S^c)$ (Fig. 3) and the derived adult prevalence function $\varphi(\lambda, \mu, S^a)$ with source S^a from the children’s group.
- ii. Next we estimate village-specific transmission coefficients $A_i (i = 1, 2, \dots, 10)$ for child/adult groups via formulae (19), and forces of infection λ_i (output of step (i)). We also compute the ‘bestfit’ estimates for biological parameters $\rho_c; \rho_a$ (egg-production/worm).
- iii. Finally, we estimate snail transmission coefficients $B_j (j = 1, 2, \dots, 5)$ from equation (Eq.C.4) (Appendix C).

Table 3 shows estimated snail to human transmission coefficients $\{A_c, A_a\}$ for child and adult groups in 10 villages, along with the best fit egg-production rate estimates $\{\rho_c, \rho_a\}$ for the two groups, and human to snail transmission rates $\{B_j; j = 1, \dots, 5\}$. We also looked at

the effect of worm discretization step Δw on estimated parameters. To this end we run 2 calibrations of the individual level field data, with steps $\Delta w = 10$ and $\Delta w = 5$, and fixed infectivity threshold $w_T = 10$ (Table 3). The 2nd case gave overall higher estimates of transmission rates $\{A_i\}$, but the children group was less sensitive ($< 15\%$) than adults. Since children carry the bulk of infection and transmission, our scheme seems overall insensitive to Δw .

Model validation

We ran a simple self-consistency (validation) test for our calibrated system by comparing the predicted equilibria (computed from the model) with the individual level field data. An exact match is not expected as parameters of the distributed system are no longer exact solutions of equilibrium equations. Indeed, the complete set of equations (MWB + human prevalence + snail prevalence) is over-determined now, so we use best-fit approximation to obtain the estimates of Table 3. The comparison test shows good agreement (Fig. 5).

Sensitivity analysis

The infection data based on egg and snail counts is notoriously inaccurate partly due to imperfect diagnostic tools. Another major uncertainty in our model is the infectivity threshold w_T . To examine such uncertainties on calibrated parameters we took 2 threshold values $w_T = 10$ and 20, and in each case run an ensemble of 200 data choices, randomly selected within suitable ranges: $\pm 10\%$ of the individual level field data, for egg counts; $\pm 5\%$ – for human prevalences, and $\pm 20\%$ – for snail infection. The error margins are somewhat arbitrary, but they reflect the overall level of uncertainty. For instance, individual egg-counts can be much ‘noisier’ than 10%, but the community mean would reduce their variance by factor $1/n$ – community size. Fig. 6 shows the results of sensitivity analysis for some calibrated parameters. The overall estimates remain fairly stable (error $< 5\%$). We plan to use the calibrated ensemble of parameters to provide error bars for predicted control outcomes in the future work on control analysis with SWB.

SUMMARY AND CONCLUSIONS

We propose a new approach to modelling schistosomiasis transmission (and other vector-borne macroparasites), based on a *Stratified Worm Burden* (SWB) formulation. In this approach, a community of hosts is subdivided into burden strata in terms of their infection levels. The transitions among strata depend on (i) force of infection (determined by the transmission environment and human risk factors); (ii) worm attrition (natural or drug-induced); and (iii) human demographics (ageing, mortality, growth sources). Such stratified systems offer many advantages compared to the conventional MWB approach (Macdonald, 1965), or SIR-type prevalence systems. By naturally encoding both types of information, these models allow new (refined) calibration procedures with the available data. Furthermore, they can also include demographics and/or other measured (e.g., behavioural) risk factors for structured human populations (age, occupation, etc.). Finally, SWB naturally explain highly aggregated burden distributions in host populations (typical of *Schistosoma* and other helminth infections (Anderson and May, 1991)) without imposing *ad hoc* distribution patterns.

We have demonstrated the SWB approach and model calibration procedures in several settings: the single homogeneous population, a population with multiple risk/age groups (child–adult systems), and a more complex area with geographically distributed human–snail populations and contact patterns. In each case, the modelling combines demographics (population levels, births, deaths, and ageing), human–snail infection levels (egg-counts and

prevalences), and for distributed systems, a minimum of environmental data (snail habitats, human contact patterns, etc.).

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

The SWB methodology has many potential applications. In the present paper, we outlined calibration results for a specific environment (the Msambweni region of eastern Kenya). We also conducted a simple validation test and sensitivity analysis of our calibration scheme to uncertainties in the data (random errors). In subsequent work we shall provide further details and applications of our methodology in exploring the predicted impact of different control strategies such as the currently advocated targeted treatment of school-aged children (WHO, 2006) as compared to community-based delivery to children and adults (Gabielli *et al.* 2006).

One might wonder how our extended model and results compare to the standard (Macdonald) MWB approach. There is no direct way to compare the two systems, as MWB by itself has no human prevalence information. Other significant differences can arise in implementation of drug treatment in the two systems. The drug-based clearing of infection can enter an MWB-system only through increased worm mortality, whereas in SWB we implement it by shifting burden strata from ‘high’ to ‘low’ to ‘lowest’ levels as hosts get cleared. The significant difference has to do with the slow process of natural worm mortality versus fast worm clearing by drug. We shall discuss this aspect in more detail in the forthcoming work on control interventions.

While our approach allows more detailed parasite distribution in host populations, it leaves out many other important within-host factors, such as density-dependent survival and egg-production and immune regulation of parasite levels and establishment. Some of these factors (immunity) are age-dependent, and would require more detailed age-structured (doubly stratified) SWB. In the current analysis, we had limited parasitological data (child–adult), and structured our model accordingly. Hence age-dependent factors (immunological etc) enter our system through different transmission and egg-production rates. Table 3 shows marked differences between two groups, and we attribute this fact partly to a protective role of immunity that develops with age.

It is, however, somewhat easier to accommodate density-dependent factors (worm survival and host infectivity) within SWB approach. Resolution rates $\{\gamma_m\}$ (for m -th strata) related to worm mortality, can be made to increase faster (than linearly) in ‘ m ’ to account for crowding mortality. One can also make infectivity (egg-production) of m -th strata a nonlinear function of m in terms of reduced fertility. While there is no simple way to accommodate such features in MWB, we have demonstrated how SWB provides a natural representation of such macroparasite dynamics.

Acknowledgments

The authors thank the referees for their pointed questions and stimulating comments that helped us to improve the results and exposition of the work.

FINANCIAL SUPPORT

This project was supported in part by National Institutes of Health Research Grant R01 TW008067 funded by the Fogarty International Center. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

APPENDICES

A. Basic Macdonald system for multiple human risk groups

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

- i. Snails are minimally mobile and relatively attached to their local habitat, so we ignore hydrologic (or other) transport of snails/larvae.
- ii. There is an approximately uniform snail density throughout each habitat (their local numbers are part of the collected field data).
- iii. All human–snail contacts and transmission (via larval densities) take place locally. So the force of human-to-snail infection (miracidial production term) is proportional to human population density/unit habitat, rather than total human population.

The basic Macdonald setup (Macdonald, 1965) for single habitat modified to include multiple risk groups (labelled by s), consists of (i) mean worm burden (MWB) and egg counts $\{w_s, E_s\}$; (ii) susceptible-infected-patent snail densities (or prevalences) $\{x, y, z\}$; (iii) cercarial and miracidial densities $\{c, m\}$:

$$\begin{aligned} \text{(Human)} & \begin{cases} \frac{dw_s}{dt} = a\eta_s c - (\gamma + \dots) w_s; \\ E_s = \rho w_s \end{cases} \\ \text{(Snail)} & \begin{cases} \frac{dy}{dt} = bm(1-y) - \nu y; \\ z = \psi y \end{cases} \\ \text{(Larvae)} & \begin{cases} \frac{dm}{dt} = \sum_s Q_s - \mu_M m; \\ \frac{dc}{dt} = \pi_c N z - \mu_c c. \end{cases} \end{aligned} \quad (\text{Eq.A.1})$$

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

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

$$\begin{cases} \frac{dw_c}{dt} = a\eta_c c - (\gamma + \mu_c + \tau) w_c \\ \frac{dw_a}{dt} = a\eta_a c + \tau \frac{H_c}{H_a} w_c - (\gamma + \mu_a) w_a. \end{cases} \quad (\text{Eq.A.2})$$

Here we allow for aging, i.e. transition of worm burden from ‘child’ to ‘adult’ compartment at the age of 20. Hence we estimated loss-rate ($\tau = 0.05$) in the child-equation, balanced by a

gain term, $\tau \frac{H_c}{H_a}$, in the adult-equation, weighted in proportion to population afraction H_c/H_a . In this formulation we assumed stationary populations, but the equations can be easily adjusted for variable (time dependent) case. The notations are listed in Table 1.

The complete system (Eq.A.1) can be reduced by eliminating larval equations, due to their ‘fast’ death rates ($\mu_C; \mu_M$) compared to other ‘slow’ processes. So we replace (c, m) with their quasi-equilibrium values

$$c^* = \frac{\pi_C}{\mu_C} N z$$

$$m^* = \sum_s \frac{\pi_M}{\mu_M} \eta_s H_s E_s.$$

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

$$\frac{dw_s}{dt} = A_s N y - (\gamma + \dots) w_s \text{ (for s-group), or specifying it for a ‘child-adult’ setup we get}$$

$$\begin{cases} \frac{dw_c}{dt} = A_c N y - (\gamma + \mu_c + \tau) w_c \\ \frac{dw_a}{dt} = A_a N y + \tau \frac{H_c}{H_a} w_c - (\gamma + \mu_a) w_a. \end{cases} \quad (\text{Eq.A.3})$$

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

$$\begin{aligned} \frac{dy}{dt} &= \Lambda (1 - y) - \nu y; \\ x &= (1 - y - z); \\ z &= \psi y; \end{aligned} \quad (\text{Eq.A.4})$$

The snail force of infection Λ (contributed by all contact groups $\{s\}$) is given by

$$\Lambda = \sum_s B_s H_s w_s. \quad (\text{Eq.A.5.})$$

The transmission coefficients $\{A_s; B_s\}$ (‘snail-to-human’ and ‘human-to-snail’) are shown in Table 1.

B. Child-adult SWB system

Differential equations for SWB system for children and adults

$$\begin{aligned}
 \text{Child} \left\{ \begin{aligned} \frac{dh_0^c}{dt} &= -(\lambda_c + \mu_c + \tau) h_0^c + \gamma_1 h_1^c + \beta \frac{H_a}{H_c}; \\ &\dots \\ \frac{dh_{i-1}^c}{dt} &= \lambda_c h_{i-1}^c - (\lambda_c + \gamma_i + \mu_c + \tau) h_i^c + \gamma_{i+1} h_{i+1}^c; \\ &\dots \\ \frac{dh_n^c}{dt} &= \lambda_c h_{n-1}^c - (\gamma_n + \mu_c + \tau) h_n^c; \end{aligned} \right. \\
 \text{Adult} \left\{ \begin{aligned} \frac{dh_0^a}{dt} &= -(\lambda_a + \mu_a) h_0^a + \gamma_1 h_1^a + \tau \frac{H_c}{H_a} h_0^c; \\ &\dots \\ \frac{dh_{i-1}^a}{dt} &= \lambda_a h_{i-1}^c - (\lambda_a + \gamma_i + \mu_a + \tau) h_i^c + \gamma_{i+1} h_{i+1}^c + \tau \frac{H_c}{H_a} h_i^c; \\ &\dots \\ \frac{dh_n^a}{dt} &= \lambda_a h_{n-1}^a - (\gamma_n + \mu_a + \tau) h_n^a + \tau \frac{H_c}{H_a} h_n^c. \end{aligned} \right.
 \end{aligned} \tag{Eq.B.1}$$

The condition for stationary population in demographic equations (15) is

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

We estimated τ and β from the demographic data of the Msambweni region (Sturrock *et al.*, 1990; Muchiri *et al.*, 1996; Hamburger *et al.*, 2004; Kariuki *et al.*, 2004; Clennon *et al.*, 2006). The child-adult death rates were computed from

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

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

Having estimated rescaled forces of infection λ_c and λ_a , we get the transmission coefficients and egg release rate from the equilibrium solution of system (16)

$$\begin{aligned}
 A_c &= \frac{\gamma \lambda_c \Delta w}{N_y}; \quad A_a = \frac{\gamma \lambda_a \Delta w}{N_y}; \\
 \rho_c &= \frac{(\gamma + \mu_c + \tau) E_c}{\gamma \lambda_c (\Delta w)}; \quad \rho_a = \frac{(\gamma + \mu_a) E_a}{\gamma \lambda_c (\Delta w) + \tau \frac{H_c E_c}{H_a \rho_c}}; \\
 B_c &= \frac{\nu \gamma A_c}{(1 - \gamma) \sum_s A_s H_s E_s / \rho_s}; \\
 B_a &= \frac{\nu \gamma A_a}{(1 - \gamma) \sum_s A_s H_s E_s / \rho_s}
 \end{aligned} \tag{Eq.B.2}$$

C. Distributed system

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

$$\begin{cases} \frac{dw_i}{dt} = \sum_{j=1}^m A_{ij} N_j y_j - (\gamma + \mu_c) w_i; & (i=1, \dots, n) \\ \frac{dy_j}{dt} = \sum_{i=1}^n B_{ij} H_i w_i (1 - y_j) - \nu y_j; & (j=1, \dots, m) \end{cases} \tag{Eq.C.1}$$

MWB equations for an age-structured (child–adult) human population with implied transmission coefficients:

$$\begin{aligned}
 & \text{(Human)} \begin{cases} \frac{dw_i^c}{dt} = A_i^c \sum_{j=1}^m \eta_{ij} z_j - (\gamma + \mu_c + \tau) w_i^c; \\ \frac{dw_i^a}{dt} = A_i^a \sum_{j=1}^m \eta_{ij} z_j + \tau \frac{H_i^c}{H_i^a} w_i^c - (\gamma + \mu_a) w_i^a; (i=1, \dots, n) \end{cases} \\
 & \text{(Worm)} \begin{cases} E_i^c = \rho^c w_i^c; \\ E_i^a = \rho^a w_i^a; (i=1, \dots, n) \end{cases} \\
 & \text{(Snail)} \begin{cases} \frac{dY_j}{dt} = \Lambda_j (N_j - Y_j) - \nu Y_j; \\ Z_j = \psi_j Y_j; (j=1, \dots, m) \end{cases}
 \end{aligned} \tag{Eq.C.2}$$

Here we use local *patent snail densities* $\{Z_j\}$ instead of infected snail prevalences of earlier special cases (formulae (9)–(11)–(13)). Since patency conversion rates of infected snails $\{\psi_j = Z_j/Y_j\}$ are local and vary across snail sites (Table 2) we need to account of their contribution to human infection separately. Another modification is the force of infection for snail in (Eq.C.2).

$$\Lambda_j = B_j \sum_i \eta_{ij} (A_i^c H_i^c E_i^c + A_i^a H_i^a E_i^a) / U_j. \tag{Eq.C.3}$$

Here $\{U_j\}$ are relative sizes of snail habitats (based on the number of contact sites of Table 2). We divide ‘human population’ by ‘contact size’ (H_i/U_j), since humans distribute their infection in inverse proportions to contact size of each habitat. Also the mean egg count E_i is used in (Eq.C.3) instead of the mean worm burden w_i in the snail force of infection Λ , as different age groups may have different egg-release rates. The transmission coefficients $\{A_i, B_j\}$ in formula (Eq.C.3) are somewhat different from earlier cases (Table 1). Namely,

$$A_i^s = a \left(\frac{\pi_c}{\nu_c} \right) T_i^s;$$

$$B_j = b_j \left(\frac{\pi_m}{\nu_m} \right) / a \left(\frac{\pi_c}{\nu_c} \right),$$

where T_i^s represents the total contact rate for the risk group s in village i to all water sites. Having estimated A_i^c , A_i^a , ρ_c and ρ_a , we compute B_j from the snail equation (2) at site j , using the relation between A ’s and B ’s equation (18):

$$B_j = \frac{\nu Y_j}{(N_j - Y_j) \sum_i \eta_{ij} (A_i^c H_i^c E_i^c + A_i^a H_i^a E_i^a) / U_j}. \tag{Eq.C.4}$$

REFERENCES

- Alexander N, Moyeed R, Stander J. Spatial modelling of individual-level parasite counts using the negative binomial distribution. *Biostatistics*. 2000; 1:453–463. [PubMed: 12933567]
 Anderson RM, May RM. Regulation and stability of host-parasite population interactions. I. *Journal of Animal Ecology*. 1978a; 47:219–247.

- Anderson RM, May RM. Regulation and stability of host-parasite population interactions. II. *Journal of Animal Ecology*. 1978b; 47:249–267.
- Anderson, RM.; May, RM. *Dynamics and Control*. Oxford University Press; New York, USA: 1991. *Infectious Diseases of Humans*.
- Chan MS, Bundy DA. 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; 91:216–220. [PubMed: 9196775]
- Chan MS, Guyatt HL, Bundy DA, Medley GF. Dynamic models of schistosomiasis morbidity. *American Journal of Tropical Medicine and Hygiene*. 1996; 55:52–62. [PubMed: 8702023]
- Clennon JA, Mungai PL, Muchiri EM, King CH, Kitron U. Spatial and temporal variations in local transmission of *Schistosoma haematobium* in Msambweni, Kenya. *American Journal of Tropical Medicine and Hygiene*. 2006; 75:1034–1041. [PubMed: 17172362]
- Dobson A, Roberts M. The population dynamics of parasitic helminth communities. *Parasitology*. 1994; 109(Suppl.):S97–S108. [PubMed: 7854855]
- Feng Z, Curtis J, Minchella DJ. The influence of drug treatment on the maintenance of schistosome genetic diversity. *Journal of Mathematical Biology*. 2001; 43:52–68. [PubMed: 12120867]
- Gabrielli AF, Toure S, Sellin B, Sellin E, Ky C, Ouedraogo H, Yaogho M, Wilson MD, Thompson H, Sanou S, Fenwick A. A combined school- and community-based campaign targeting all school-age children of Burkina Faso against schistosomiasis and soil-transmitted helminthiasis: performance, financial costs and implications for sustainability. *Acta Tropica*. 2006; 99:234–242. [PubMed: 16997268]
- Gurarie D, King CH. Heterogeneous model of schistosomiasis transmission and long-term control: the combined influence of spatial variation and age-dependent factors on optimal allocation of drug therapy. *Parasitology*. 2005; 130:49–65. [PubMed: 15700757]
- Gurarie D, King CH. Age- and risk-targeted control of schistosomiasis-associated morbidity among children and adult age groups. *The Open Tropical Medicine Journal*. 2008; 1:21–30.
- Gurarie D, Seto EY. Connectivity sustains disease transmission in environments with low potential for endemicity: modelling schistosomiasis with hydrologic and social connectivities. *Journal of the Royal Society Interface*. 2009; 6:495–508. doi: 10.1098/rsif.2008.0265.
- Hamburger J, Hoffman O, Kariuki HC, Muchiri EM, Ouma JH, Koech DK, Sturrock RF, King CH. Large-scale, polymerase chain reaction-based surveillance of *Schistosoma haematobium* DNA in snails from transmission sites in coastal Kenya: A new tool for studying the dynamics of snail infection. *American Journal of Tropical Medicine and Hygiene*. 2004; 71:765–773. [PubMed: 15642969]
- Kariuki HC, Clennon JA, Brady MS, Kitron U, Sturrock RF, Ouma JH, Ndlovu ST, Mungai P, Hoffman O, Hamburger J, Pellegrini C, Muchiri EM, King CH. Distribution patterns and cercarial shedding of *Bulinus nasutus* and other snails in the Msambweni area, Coast Province, Kenya. *American Journal of Tropical Medicine and Hygiene*. 2004; 70:449–456. [PubMed: 15100463]
- Macdonald G. The dynamics of helminth infections, with special reference to schistosomes. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1965; 59:489–506. [PubMed: 5860312]
- Medley GF, Bundy DA. Dynamic modeling of epidemiologic patterns of schistosomiasis morbidity. *American Journal of Tropical Medicine and Hygiene*. 1996; 55:149–158. [PubMed: 8940969]
- Muchiri EM, Ouma JH, King CH. Dynamics and control of *Schistosoma haematobium* transmission in Kenya: an overview of the Msambweni Project. *American Journal of Tropical Medicine and Hygiene*. 1996; 55:127–134. [PubMed: 8940966]
- Nasell I. Mating for schistosomes. *Journal of Mathematical Biology*. 1978; 6:21–35. [PubMed: 750626]
- Pugliese A. Coexistence of macroparasites without direct interactions. *Theoretical Population Biology*. 2000; 57:145–165. doi: S0040-5809(99)91443-0 [pii]. [PubMed: 10792979]
- Riley S, Carabin H, Belisle P, Joseph L, Tallo V, Balolong E, Willingham AL, Fernandez TJ, Gonzales RO, Olveda R, Mcgarvey ST. Multi-host transmission dynamics of *Schistosoma japonicum* in Samar province, the Philippines. *PLoS Med*. 2008; 5:e18. doi: 10.1371/journal.pmed.0050018. [PubMed: 18215106]

- Rosa R, Pugliese A. Aggregation, stability, and oscillations in different models for host-macroparasite interactions. *Theoretical Population Biology*. 2002; 61:319–334. doi: 10.1006/tpbi.2002.1575. [PubMed: 12027618]
- Sturrock RF, Kinyanjui H, Thiongo FW, Tosha S, Ouma JH, King CH, Koech D, Siongok TK, Mahmoud AA. Chemotherapy-based control of schistosomiasis haematobia. 3. Snail studies monitoring the effect of chemotherapy on transmission in the Msambweni area, Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1990; 84:257–261. [PubMed: 2117788]
- Van Der Werf, MJ.; De Vlas, SJ. Report for WHO Parasitic Diseases and Vector Control. Erasmus University; Rotterdam, The Netherlands: 2001. Morbidity and Infection with Schistosomes or Soil-Transmitted Helminths; p. 1-103.
- de Vlas SJ, Engels D, Rabello AL, Oostburg BF, Van Lieshout L, Polderman AM, Van Oortmarssen GJ, Habbema JD, Gryseels B. Validation of a chart to estimate true *Schistosoma mansoni* prevalences from simple egg counts. *Parasitology*. 1997; 114:113–121. [PubMed: 9051920]
- Wilson RA, Dam GJ, Kariuki TM, Farah IO, Deedler AM, Coulson PS. The detection limits for estimates of infection intensity in schistosomiasis mansoni established by a study in non-human primates. *International Journal for Parasitology*. 2006; 36:1241–1244. [PubMed: 16930605]
- World Health Organization. Preventive Chemotherapy in Human Helminthiasis: Coordinated Use of Anthelmintic Drugs in Control Interventions: a Manual for Health Professionals and Programme Managers. World Health Organization Press; Geneva, Switzerland: 2006.
- Woolhouse ME. On the application of mathematical models of schistosome transmission dynamics. I. Natural transmission. *Acta Tropica*. 1991; 49:241–270. [PubMed: 1684260]
- Woolhouse ME, Watts CH, Chandiwana SK. Heterogeneities in transmission rates and the epidemiology of schistosome infection. *Proceedings of the Royal Society of London B*. 1991; 245:109–114.

Fig. 1.

Infective strata and transitions of SWB model. Population strata (numbers or fractions of total) $\{h_i\}$ move 'up' or 'down' the scale at certain rates, due to worm acquisition (λ) and resolution (γ_i). Each stratum has its source-term S_i , and loss rate μ .

Fig. 2.

(A) Equilibrium worm burden distribution for different values $\lambda = 0.2$ (light grey), 0.6 (medium grey) and 1 (black), with identical human loss rate $\mu = 0.02$ and stationary source at the lowest strata. (B) Time relaxation of SWB-system coupled to infected snail prevalence. Left panel shows prevalence distributions of 15 human strata $\{h_i(t)\}$ at $t = 0$ (light grey), $t = 20$ (medium grey) and $t = 40$ (black). For snail prevalence we show the dynamic relaxation pattern (right panel).

Fig. 3.

(A) Plot of uninfected prevalence function $P = \varphi(\lambda, \mu, S)$ on worm acquisition rate λ for 3 values $\mu = \{0.1\gamma, 0.2\gamma, 0.3\gamma\}$ (top to bottom) and identical stationary source (S) at the lowest strata; (B) Plot of MWB-function $w = \psi(\lambda, \mu, S)$ on λ for the same values of μ as (A) (top to bottom) with dashed lines corresponding to a would-be 'linear relation' (slope) $\frac{\partial w}{\partial \lambda} \approx \frac{1}{\gamma + \mu}$ at $\lambda = 0$ (appropriate for Macdonald MWB).

Fig. 4.

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

Fig. 5. Comparison between individual level field data (dark disk) and equilibrium solutions of the calibrated model (grey square); the x-axis labels the ‘village’ or snail site and y-axis – ‘mean egg count’ or ‘snail density’.

Fig. 6. Sensitivity analysis of model parameters (snail-to-human transmission rates) $\{A_i\}$ for children in 10 villages: panel (A) shows heavily infected villages nos 6 and 7, and panel (B) for all others. Two sets (light/dark) of box-and-whisker plots ('box' = 0.25–0.75 quantiles; 'whiskers' = 10–90) correspond to infectivity thresholds $w_T = 10$ and $w_T = 20$. Black dots mark calibrated values of the 'individual level field data' in each case.

Table 1

List of notations, variables and parameters

Notation	Description
H	Human population density
E	Egg release/host (counted as arithmetic mean egg count per 10 ml of urine)
μ	Human mortality (turnover) rate/year
N	Snail population density per unit habitat
$y; Y$	Infected snail prevalence and snail density (per unit habitat)
$z; Z$	Infectious(patent) snail prevalence and snail density per unit habitat
T_i	Total contact per person to all water sites for people in village i
η_{ij}	Fraction of total contacts from village i to snail habitat j
a	Probability of worm establishment per cercariae per contact
β	Human growth rate (estimated 0.05/year)
τ	Transition rate from youth to adulthood (estimated 0.05/year)
γ	Worm mortality (life span: 5 years)
ν	Snail mortality(life span: 9 weeks)
ρ	Released egg production rate per worm
$\psi_j = z_j/y_j$	'Patency conversion' fraction of infected snails
$\pi_C; \nu_C$	Rate of cercariae output per shedding snail and cercariae mortality
$\pi_M; \nu_M$	Miracidia production rate/worm and miracidium mortality
B	Probability of snail infection per unit miracidium density (can be site specific)
$A = aT \left(\frac{\pi_C}{\nu_C} \right) \psi$	Snail-to-human transmission coefficient (can be site specific)
$B = aT \left(\frac{\pi_M}{\nu_M} \right) \rho$	Human-to-snail transmission coefficient (can be site specific)
w_T	Infectivity threshold ('effective worm burden' below which hosts are considered free of infection).

Table 2

Human and snail demographics and infection data: 'egg-count' refers to arithmetic mean/host (Prevalence, infected fraction of population.)

10 villages		1	2	3	4	5	6	7	8	9	10
Children	population	1134	791	867	1584	1572	357	753	403	294	492
	egg count	58	91	134	62	115	319	257	139	148	170
	prevalence	0.23	0.42	0.49	0.28	0.32	0.67	0.72	0.51	0.53	0.5
Adults	population	932	632	701	1328	1287	358	714	367	268	399
	egg count	7	9	10	4	13	22	42	11	27	27
	prevalence	0.12	0.14	0.18	0.12	0.12	0.26	0.34	0.23	0.3	0.29
Snail water sites											
		1	2	3	4	5					
Relative area		1	3	3	5	2					
Total density		1609	568	721	507	1005					
Infected density		483	204	336	213	387					
Patent (infectious) density		18	1	2	12	10					

Calibrated transmission coefficients and egg production rates based on the individual level field data for 2 choices of discretization step ($\Delta\omega$), and infection threshold $\omega_T = 10$.

Table 3

Village no.	1	2	3	4	5	6	7	8	9	10	
A_c	$\Delta\omega=10$	0.5130	0.7241	0.9245	0.5165	2.882	23.71	6.945	2.874	2.254	1.339
	$\Delta\omega=5$	0.8203	0.9014	1.089	0.7612	4.025	25.02	7.046	3.335	2.578	1.564
A_a	$\Delta\omega=10$	0.0804	0.0196	0.0346	0.0496	0.1752	0.7590	0.3402	0.2238	0.2880	0.1933
	$\Delta\omega=5$	0.2283	0.106	0.1317	0.1608	0.6823	2.016	0.6819	0.5210	0.5176	0.3562

Snail site no.	1	2	3	4	5	
B	$\Delta\omega=10$	1.045×10^{-5}	6.503×10^{-6}	1.730×10^{-5}	2.264×10^{-5}	5.457×10^{-5}
	$\Delta\omega=5$	7.983×10^{-6}	5.791×10^{-6}	1.534×10^{-5}	2.130×10^{-5}	4.533×10^{-5}

ρ_c		ρ_a	
$\Delta\omega=10$	17.62	6.479	
$\Delta\omega=5$	16.09	4.282	