Epidemiol. Infect. (1996), **117**, 537–550

The consequences of uncertainty for the prediction of the effects of schistosomiasis control programmes

M. S. CHAN

Centre for the Epidemiology of Infectious Disease, Department of Zoology, South Parks Road, Oxford OX1 3PS, UK

(*Accepted 31 July 1996*)

SUMMARY

Progress in the development of schistosomiasis models for use in control programmes is limited by the considerable uncertainty in many of the biological parameters. In this paper, this problem is addressed by a comprehensive sensitivity analysis of a schistosomiasis model using the Latin Hypercube method. Fifty simulations with different parameter contributions are run for 50 years with treatment during the first 20 years and reinfection thereafter. The analysis shows only a relatively small divergence between simulations during the chemotherapy treatment programme but considerable divergence in reinfection levels after treatment is stopped. A skewed distribution of outcomes was seen with most simulations showing effective control and a few where control had less impact. The most important uncertainty source was due to the unknown levels of acquired immunity and also uncertainty in the true worm burden. In particular, the strength of the immune response was most important in determining whether control was effective with higher immunity leading to less effective control. Among those simulations in which control was not very effective, those in which the mean worm burden was high showed the least effective control. Since both these are areas of genuine uncertainty, it is proposed that uncertainty analysis should be an integral part of any projection of control programmes.

INTRODUCTION

Schistosomiasis is a disease which causes serious morbidity in many parts of Africa, Asia and Latin America [1]. It is caused by a parasitic worm that lives in the bloodstream of the host. People in endemic countries are typically infected for much of their lives and it is this long term infection that causes serious pathological changes to the internal organs [2]. All species of schistosomes can be effectively treated with the drug praziquantel and mass drug treatment is considered a cost-effective control strategy [3].

Mathematical models of schistosomiasis have a long history [4–9] and have been used to demonstrate important processes determining patterns of infection. However, to date, models are not used routinely in the control of schistosomiasis although they have been for another helminth disease, onchocerciasis [10]. An age-structured transmission and morbidity model is being currently developed with the intention of eventually using it to aid control decisions [9, 11–13].

One reason why it has been difficult to use models of schistosomiasis in a predictive way is due to considerable uncertainty in the magnitude of many parameters in the models and hence potential uncertainty in the simulation results. These reflect fundamental gaps in our knowledge of schistosome biology. The consequences of these uncertainties can be examined with formal sensitivity analysis in which sets of parameter values are sampled in parameter space, either randomly or systematically and the simulations arising from these examined. These types

of sensitivity analysis were developed for engineering applications and are only now starting to be used in epidemiological models [14–17].

In this paper, a full sensitivity analysis of the schistosomiasis transmission and morbidity model is undertaken using parameter values appropriate for *Schistosoma mansoni* to examine both the likely variability in outcomes of control programmes given our level of uncertainty in the parameter values and to identify the particular parameter uncertainties that contribute most to the overall variability. The method of Latin Hypercube Sampling [18, 19] was used to generate a sample of parameter values for the simulations using the software package UNCSAM. A linear multiple regression model was used to determine the specific parameter contributions to the model outcomes.

METHODS

The schistosomiasis transmission model used in this paper has been described in full in earlier publications [9, 11–13] and is therefore only described in summary here. The model consists of an age structured partial differential equation framework that describes the changes in intensity of infection (as mean worm burden and mean egg count) over age and time. The model can also predict the impact of mass and age targeted chemotherapy programmes.

Individuals are assumed to get infected at a rate dependent on their age, the current level of infection (concomitant immunity) and their previous experience of infection (acquired immunity). Worms die at a constant rate (μ) with average lifespan $1/\mu$. Infected individuals are at risk of developing early disease (hepatomegaly in *S. mansoni*) and those with early disease are at risk of developing late disease (liver fibrosis in *S. mansoni*). Both types of disease can resolve but early disease resolves over a shorter time scale (about 1 year) compared with late disease (about 10 years).

The model is specified by state variables describing mean worm burden (M), the level of acquired immunity (I), The prevalence of early disease (D_E) and the prevalence of late disease (D_L) over age (a) and time (t) as follows:

$$\frac{\partial M(a,t)}{\partial t}+\frac{\partial M(a,t)}{\partial a}=\Lambda(a,t)\,e^{-(\delta I(a,t))^x}-\mu M(a,t) \quad (1)$$

$$\frac{\partial I(a,t)}{\partial t}+\frac{\partial I(a,t)}{\partial a}=M(a,t)-sI(a,t) \quad (2)$$

$$\frac{\partial D_E(a,t)}{\partial a}+\frac{\partial D_E(a,t)}{\partial t}=r_E M(a,t)(1-D_E(a,t))-\mu_{DE} D_E(a,t) \quad (3)$$

$$\frac{\partial D_L(a,t)}{\partial a}+\frac{\partial D_L(a,t)}{\partial t}=r_L D_E(a,t)(1-D_L(a,t))-\mu_{DL} D_L(a,t). \quad (4)$$

In these equations $\partial M(a,t)/\partial a+\partial M(a,t)/\partial t$ represent the partial derivatives of mean worm burden over age and time. For example $\partial M(a,t)/\partial a$ is the rate of change of mean worm burden by age at a particular instant in time.

δ is the strength of immunity and $1/s$ is the average duration of immune protection, r_E is the rate of development of early disease and r_L is the rate of development of late disease. The resolution rates of early and late disease are μ_{DE} and μ_{DL} respectively with the inverses of the parameters being the average resolution times. x is a parameter which describes the shape of the immune response function.

$\Lambda(a,t)$ is the rate of infection and is given by:

$$\Lambda(a,t)=\frac{\mu R_0 \rho(a) f(M(a,t))\int_a \pi(a)\kappa(a)M(a,t)\,da}{\int_a \pi(a)\kappa(a)\rho(a)\,da}, \quad (5)$$

where R_0 is the basic reproductive number of the parasite, and is biologically equivalent to the number of offspring produced by a worm which themselves reach maturity in the absence of density dependent contraints, $f(M(a,t))$ is a density dependent concomitant immunity function and $\pi(a)$ is a demographic function defining the proportion of the population in each age class. $f(M(a,t)$ is given by:

$$f(M)=\left(1+\frac{M}{k(M)}(1-e^{-\gamma})\right)^{-(k(M)+1)} \quad (6)$$

γ describes the strength of concomitant immunity. $k(M)$ is the aggregation parameter of the negative binomial distribution and this is assumed to increase linearly with M with intercept k_0 and slope k_{lin}. $\rho(a)$ and $\kappa(a)$ are the age dependent exposure and contamination functions respectively. These are both assumed to take the same form and have the same parameter values and are given by:

$$\rho(a)=\kappa(a)=(ae^{-(\beta a)^2})+c, \quad (7)$$

where β and c are constants.

Mean egg count (arithmetic, including negatives) was assumed to be linearly related to mean worm

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

burden with a constant factor e_i representing the number of eggs per gram faeces (epg) per worm.

Further details of the model and its numerical implementation are given in earlier publications [9, 13].

The purpose of the sensitivity analysis is to look at the variability in predicted consequences of mass treatment for a given observed endemic situation due to uncertainties in values of parameters and the structure of the model. All the simulations therefore had the same peak intensity of 500 eggs per gram of faeces (epg). Furthermore, the age-dependent contact function was kept constant with peak water contact at age 15. The treatment schedule was also constant with children between the ages of 7 and 18 being treated every 4 years in the first 20 years of the simulation at a coverage of 80% and with drug efficacy 95%. Simulations were run for 50 years. Two types of sensitivity analysis were carried out, structural sensitivity analysis, where the structural assumptions of the model were tested, and parameter sensitivity analysis where the values of the parameters were varied within a constant model structure.

Structural sensitivity analysis

The aim of this analysis was to examine whether changes in the mathematical structure of the model had significant consequences for the model output. Two structural variations were used in this analysis. Firstly, the consequences of changing the density dependent establishment function ($f(M(a,t))$ into a density dependent fecundity function was assessed. This involves the replacement of equation (5) by:

$$\Lambda(a,t) = \frac{\mu R_0 \rho(a) \int_a \pi(a)\kappa(a) f(M(a,t))\, M(a,t)\, da}{\int_a \pi(a)\kappa(a)\rho(a)\, da}. \quad (8)$$

The same function for $f(M(a,t))$ was used.

The second structural adjustment was that the value of the immune response parameter x in equation (1) was varied at three values of 0·2, 1 and 5. The purpose of this analysis was to ascertain whether it is possible to exclude variation in this parameter from the parameter sensitivity analysis and use a simple exponential function instead.

These structural changes were examined at two levels of immunity $\delta = 0{\cdot}001$, $l/s = 5$ years giving a 30% reduction of mean worm burden at age 20 compared with no immunity and $\delta = 0{\cdot}002$, $l/s = 10$ years, giving a 60% reduction. Other parameter values for this analysis are those listed under the 'mode' column in Table 1.

Parameter sensitivity analysis

The Latin Hypercube Sampling method requires that the distribution of each variable parameter is specified. The method then samples systematically from the parameter space to generate n parameter sets (where n is the number of simulations to be carried out) as follows. First the range of each parameter is divided into n intervals of equal probability. A random value for the parameter is taken from each of these intervals according to the probability distribution. The values for each parameter are then combined at random with the samples for other parameters [21].

Twelve parameters were varied in the parameter sensitivity analysis, other parameters relating to the determination of the endemic infection intensity were kept constant. Since the distributions of the parameters are generally unknown, triangular distribution was used for all parameters for simplicity. The mode values for these distributions were the values used for this model in previous papers [20, 9, 13] and are appropriate for the species *S. mansoni*. The maximum value was taken to be twice the mode and the minimum value was taken to be half the mode (or zero if appropriate). In the case of the rate of development of early disease (r_E), estimated maximum and minimum values are taken from Chan and colleagues [11]. The parameter x in equation (1) is also included in this analysis. The values of the variable parameters, their distributions and the values of the constant parameters are all shown in Table 1.

The above parameter distributions are input into the software package UNCSAM version 1.1 [21]. This software samples the distributions and outputs the parameter sets for each simulation. Fifty simulations were performed which is more than adequate for 12 variable parameters ($4N/3$ samples is considered to be sufficient, if N is the number of parameters) [21].

Several outcome variables were calculated from the simulations to allow output to be compared. The mean egg count, prevalence of early disease and prevalence of late disease were output for ages at 5-year intervals (i.e. age 5, 10, 15 etc) at years 0 (initial), 20 (treatment) and 50 (reinfection). In addition, the time profile of mean egg count was calculated at 5-year intervals for individuals of age 12 (treatment

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

Table 1. *Values of variable and constant parameters*

Parameter	Description	Mode	Minimum	Maximum
$1/\mu$	Worm lifespan (yr)	4	2	8
γ	Density dependence	0·001	0·0005	0·002
k_0	k intercept	0·132	0·1	0·3
k_{lin}	k slope	0·002	0	0·004
e_i	epg per worm	5·26	1	10
r_E	Early disease	0·0071	0·002	0·016
$1/\mu_{DE}$	Resolution ED (yr)	1	0·5	2
rL	Late disease	0·015	0·007	0·03
$1/\mu_{DL}$	Resolution LD (yr)	13	7	26
δ	Strength of immunity	0·001	0	0·002
x	Exponent	1	0·2	5
$1/s$	Duration immunity (yr)	5	2	10
	Peak egg count		constant: 500 epg	
	Peak water contact		constant: age 15	
c	Contact function constant		constant: 0·4	
	Treatment ages		constant: 7–18	
	Treatment coverage		constant: 80%	
	Treatment interval		constant: 4 years	
	Drug efficacy		constant: 95%	

group). The basic reproductive number (R_0) was calculated for each simulation. The benefit of the treatment programme is calculated as the average egg count reduction per person per year over the 50 years of the treatment compared with the absence of treatment. This was calculated for the whole community (B_A) and for the treated age group only (B_C). These outcome variables are defined as follows:

$$R_0 = \frac{\alpha \int_a \pi(a)\,\kappa(a)\,\rho(a)\,da}{\mu \int_a \pi(a)\,\kappa(a)\,M^*(a)\,da} \tag{9}$$

or, in the fecundity model:

$$R_0 = \frac{\alpha \int_a \pi(a)\,\kappa(a)\,\rho(a)\,da}{\mu \int_a \pi(a)\,\kappa(a)\,M^*(a) f(M^*(a))\,da}, \tag{10}$$

where α is a parameter fitted by iteration to give the pre-treatment peak mean egg count of 500 epg.

$$B_A = \frac{e_i \int_{t=0}^{t=50} \int_{a=0}^{a=80} [M^*(a) - M(a,t)]\,\pi(a)\,da\,dt}{50} \tag{11}$$

$$B_C = \frac{e_i \int_{t=0}^{t=50} \int_{a=7}^{a=18} [M^*(a) - M(a,t)]\,\pi(a)\,da\,dt}{50 \int_{a=7}^{a=18} \pi(a)\,da}, \tag{12}$$

where $M^*(a)$ is the equilibrium mean worm burden.

For the structural sensitivity analysis, the benefit is given as the reduction in mean worm burden due to the complication with reduced egg counts due to reduced fecundity in the density dependent fecundity model. These benefit values are calculated by taking e_i out of equations 11 and 12.

To compare the contribution of each parameter to the overall uncertainty in the results, a multiple linear regression model is used which is also performed by the UNCSAM package. The Standardised Regression Coefficient is compared for the different parameters. This coefficient is measured relative to the standard deviation of the parameters and hence takes into account the differences in absolute values between parameters. The value of the t-statistic is also calculated.

RESULTS

Structural sensitivity analysis

The values of the outcome variables for the simulations in the structural sensitivity analysis are shown in Table 2. The values of the outcome variables are very similar for the establishment and fecundity models of density dependence. This suggests that this particular structural change has little impact on the model results. In contrast, changing the exponent, x, of the immune response function has a profound

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

Table 2. *Values of outcome variables for the structural sensitivity analysis. Benefits are reduction in worm burden*

Density dependence model	Strength of immunity (δ)	Duration of immunity ($1/s$)	Exponent (x)	Basic reproductive number (R_0)	Community benefit (B_A)	Benefit to children (B_C)
Establishment	0·001	5	0·2	3·31	16	42
			1	2·07	14	38
			5	1·57	23	58
	0·002	10	0·2	3·96	14	40
			1	3·69	5	19
			5	2·61	2	11
Fecundity	0·001	5	0·2	3·33	15	41
			1	2·08	13	38
			5	1·59	21	55
	0·002	10	0·2	3·96	14	40
			1	3·58	5	21
			5	2·48	2	15

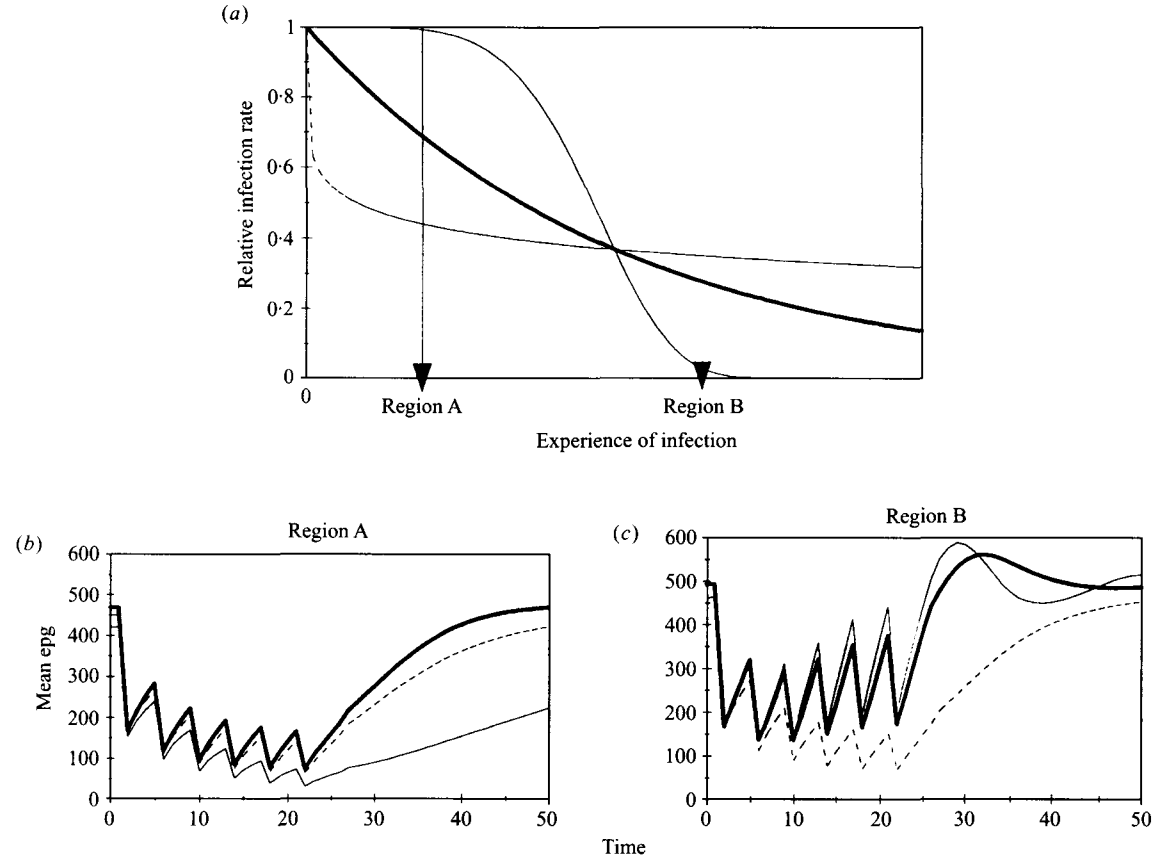

Fig. 1. Diagram to demonstrate the consequences of the shape of the immune function on the transmission dynamics during and after mass chemotherapy. (*a*) shows schematically the shapes of three response functions with the value of the exponent x (equation 1) varied between 0·2 (dotted line), 1 (heavy solid line) and 5 (light solid line). The x axis is an unscaled measure of experience of infection. In the area marked region A, the transmission dynamics are shown in (*b*) (with the same line legend), which shows the time profiles at age 12 for the simulations where the immune parameters are $\delta = 0{\cdot}001$, $1/s = 5$. In the area marked region B, the transmission dynamics are shown in (*c*) which shows the time profiles at age 12 for the simulations where the immune parameters are $\delta = 0{\cdot}002$, $1/s = 10$.

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

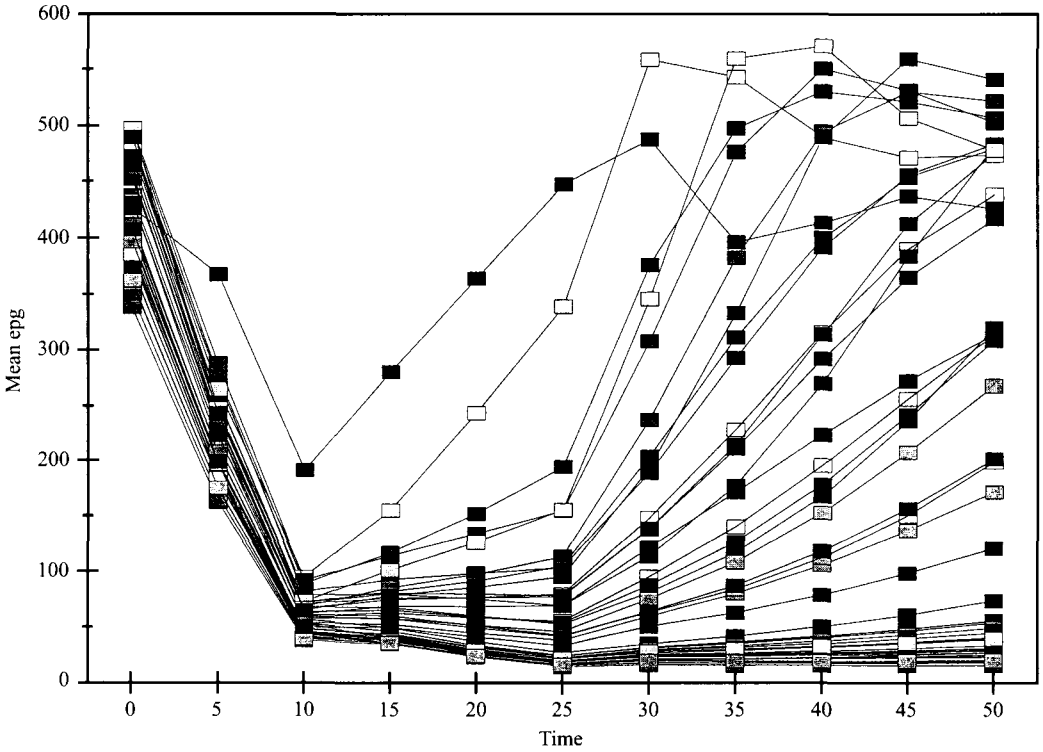

Fig. 2. Time profiles of mean egg count at age 12 plotted separately for all 50 simulations.

effect on the outcome variables. Furthermore, this effect is highly non-linear, with the benefit values being highest for a high exponent at the lower level of immunity and the reverse being true at the higher level of immunity.

The explanation for the above counter-intuitive result is shown in Figure 1 which includes the actual simulation profiles. The top figure shows the relative infection rate ($e^{-\delta I(a,t)}$) against experience of infection ($I(a,t)$). The x axis is on an arbitrary scale. We examine the changes in the infection rate as the population 'experience of infection' first decreases after treatment and increases again during the period of reinfection. If the endemic experience of infection is in region A relative to the immune response function (as it is for the lower immunity parameter set), there is a greater degree of immunity when x is low and hence a greater potential reinfection rate. In region B (higher immunity parameter set), however, as the effective experience of infection decreases due to lack of exposure during the treatment programme, the simulation with the highest exponent switches from being the simulation with most effective immunity to having the least effective immunity and hence gives the highest reinfection rate and a substantial overshoot of the mean egg count above the initial endemic level. Given the substantial influence of the shape of the immune response function on the simulation results, the value of the exponent, x, was also included as a variable in the parameter sensitivity analysis.

Parameter sensitivity analysis

Prediction of mean egg count

The change in mean egg count over time at age 12 years for all the simulation runs is shown in Figure 2. The simulations show similar patterns during the treatment programme (up to year 20) and quite different patterns on reinfection (after year 20). All the simulations show an initial rapid fall in infection intensity up to year 10. Between years 10 and 20, the infection intensity remains at a constant low level for the majority of simulations but starts to increase in a few simulations, with two simulations showing a large increase in infection during this late treatment phase. These increases are due to the breakdown in herd immunity in the population since they are not acquiring immunity so rapidly in the absence of intense infection. During the reinfection phase (years 20–50) the infection intensity remains low for most of the simulations, but for some simulations there is rapid reinfection. Only in very few simulations is there any overshooting to above the initial level, a phenom-

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

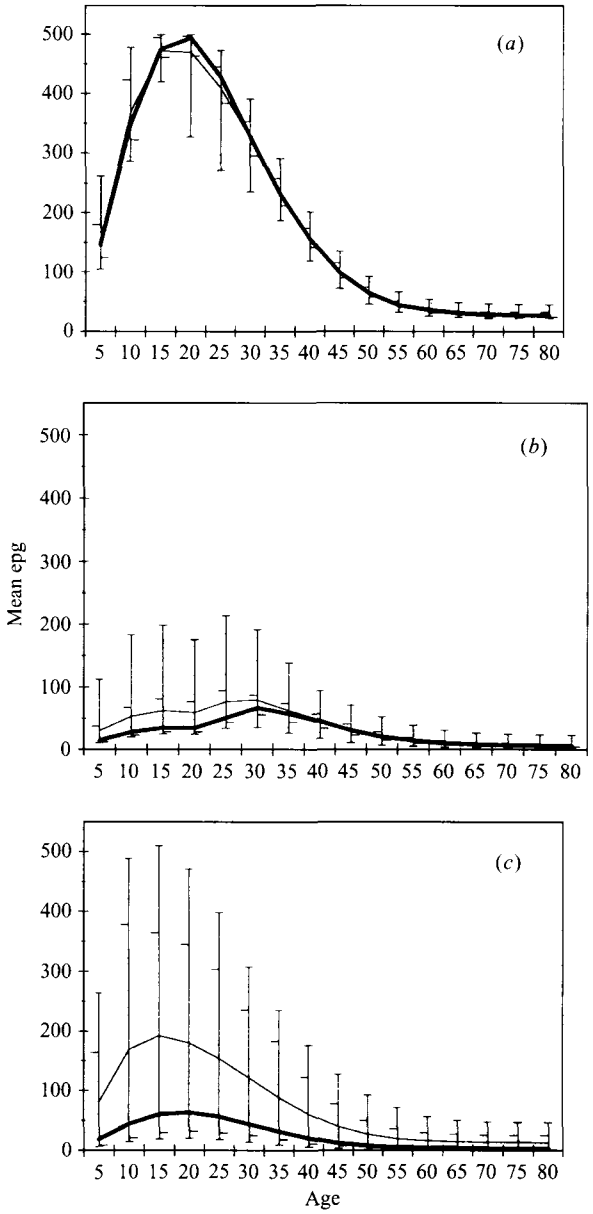

Fig. 3. Age profiles of mean egg count at years 0 (*a*), 20 (*b*) and 50 (*c*). The heavy solid lines show the median values and the light solid lines show the means. The top and bottom bars of the I bar represent the 95th and 5th percentile and the internal ticks the interquartile range.

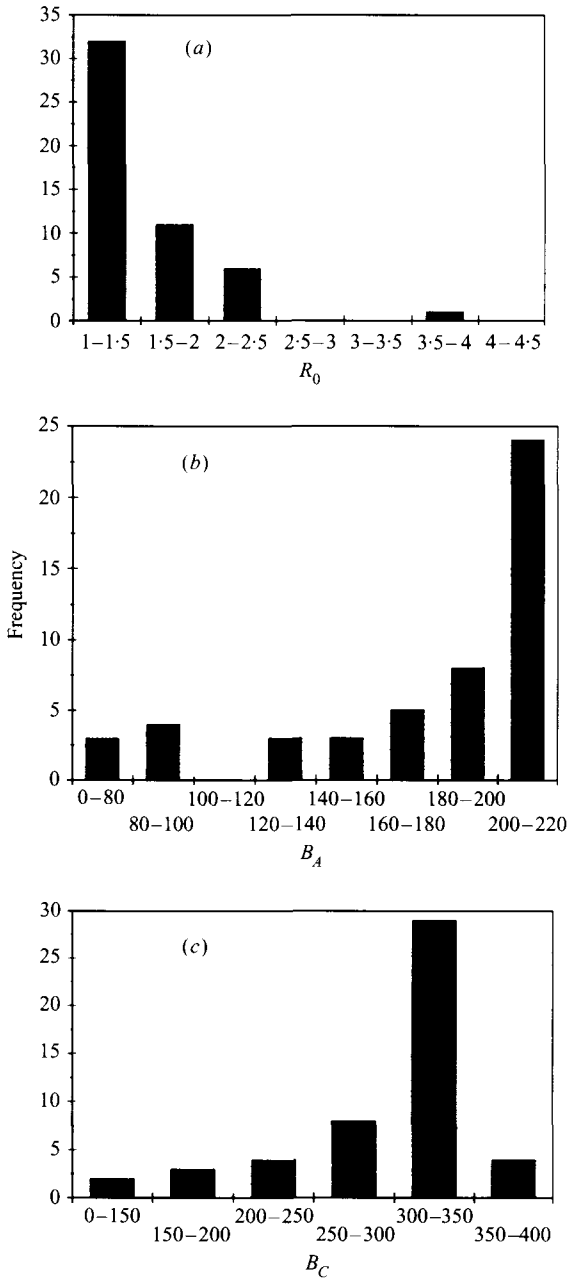

Fig. 4. Histograms showing the frequency (number of simulations) distribution of the outcome variables, R_0, B_A (Fig. 4*b*) and B_C (Fig. 4*c*).

enon often suggested as being important in dynamic helminth models with immunity.

The age profiles at different times in the simulation are shown in Figure 3. The quantities plotted are the mean of all the simulations, the median, and the 5, 25, 75 and 95% percentiles. The initial age profile (Fig. 3*a*) shows little variation, as expected since the contact functions are all the same. In some profiles there is a lower peak age due to strong immunity. At year 20, there is also little variation in the age profiles (Fig. 3*b*), suggesting that the prediction of infection levels during the treatment phase is not that dependent on the initial parameters. However, an interesting pattern is emerging which was not present initially. There is an asymmetry in the distribution with the upper percentiles being much more spread out than the lower percentiles and the mean being higher than the median. This indicates that the distribution of infection intensity is skewed, with many simulations with low intensities and a few with higher intensities.

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

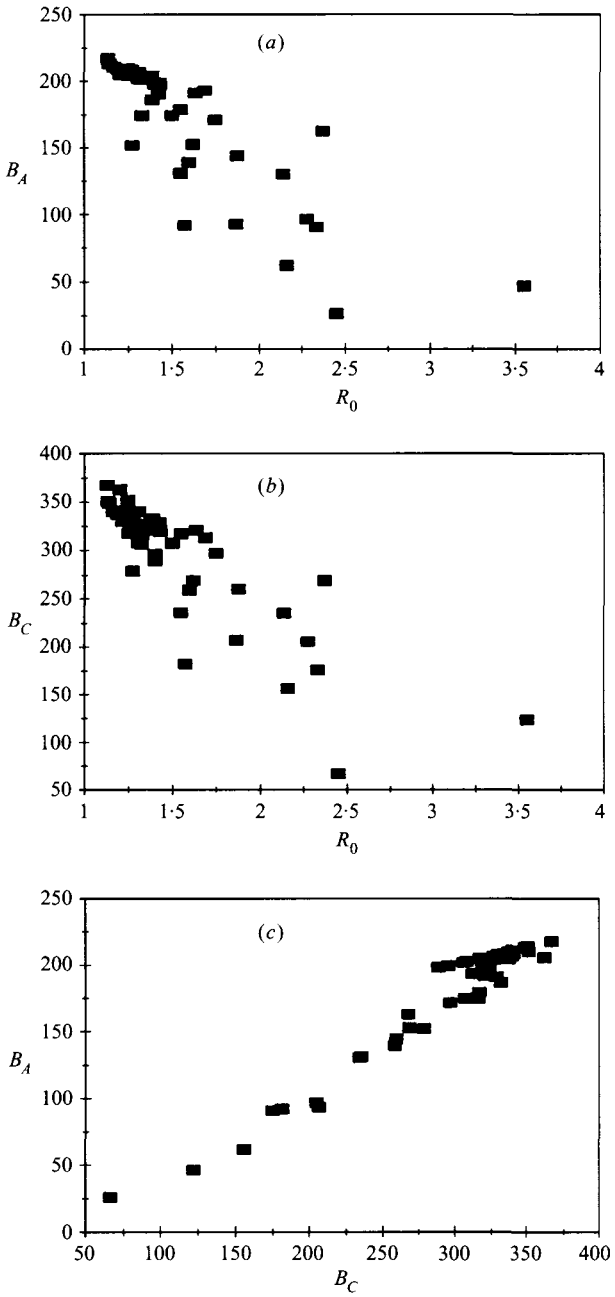

Fig. 5. Scatter graphs (*a–c*) showing the relationship between the different outcome variables.

The skewed distribution becomes much more exaggerated at year 50 (Fig. 3*c*) where the median simulation shows very low reinfection although the 95% percentile shows reinfection to equilibrium levels. This indicates that high rates of reinfection are possible but unlikely within the parameter space examined. Note that the median is a better measure of central tendency compared to the mean in this situation because the median is the expected consequence of an 'average' simulation.

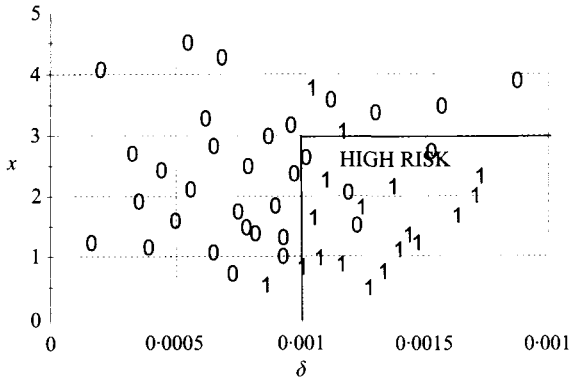

Fig. 6. Figure showing the influence of the immune parameters δ (strength of immunity) and x (immune function exponent) on the benefit to the community (B_A). The benefits are categorized into values above 180 (0) and below 180 (1) with each plotted point representing the position of one simulation in parameter space. The solid line encloses the area in which most of the low benefit values lie and is labelled as the high risk area.

Table 3. *Standardized regression coefficients (SRC) and t-statistics for the analysis of parameter uncertainty contributions to the outcome variables for those simulations for which $B_A < 180$. Only parameters significant at the 5% level are shown. The value of the correlation coefficient (R^2) is also given for the full regression model for each outcome variable*

Outcome variable	Parameter	SRC	*t*-statistic
R_0	e_i	−0·8126	−2·5211
$R^2 = 0{\cdot}89$			
B_A	e_i	1·1319	3·9903
$R^2 = 0{\cdot}91$	$1/s$	−0·5971	−2·7552
	δ	−0·4654	−2·1563
	γ	−0·3588	−2·3461
B_C	e_i	1·1882	3·7601
$R^2 = 0{\cdot}89$	$1/s$	−0·5473	−2·2673
	γ	−0·3799	−2·2300

The distributions of the outcome variables (the basic reproductive number, R_0, and the benefits to the population and the treated ages, B_A and B_C respectively) are plotted as histograms in Figure 4. All the distributions are highly skewed as would be expected from the simulation results. The distribution of R_0 is skewed towards the right with most of the values under 2 and a few higher values. However, the values of the benefits are skewed towards the left with the majority of values being at the higher end of the distribution. These results reflect the observation that

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

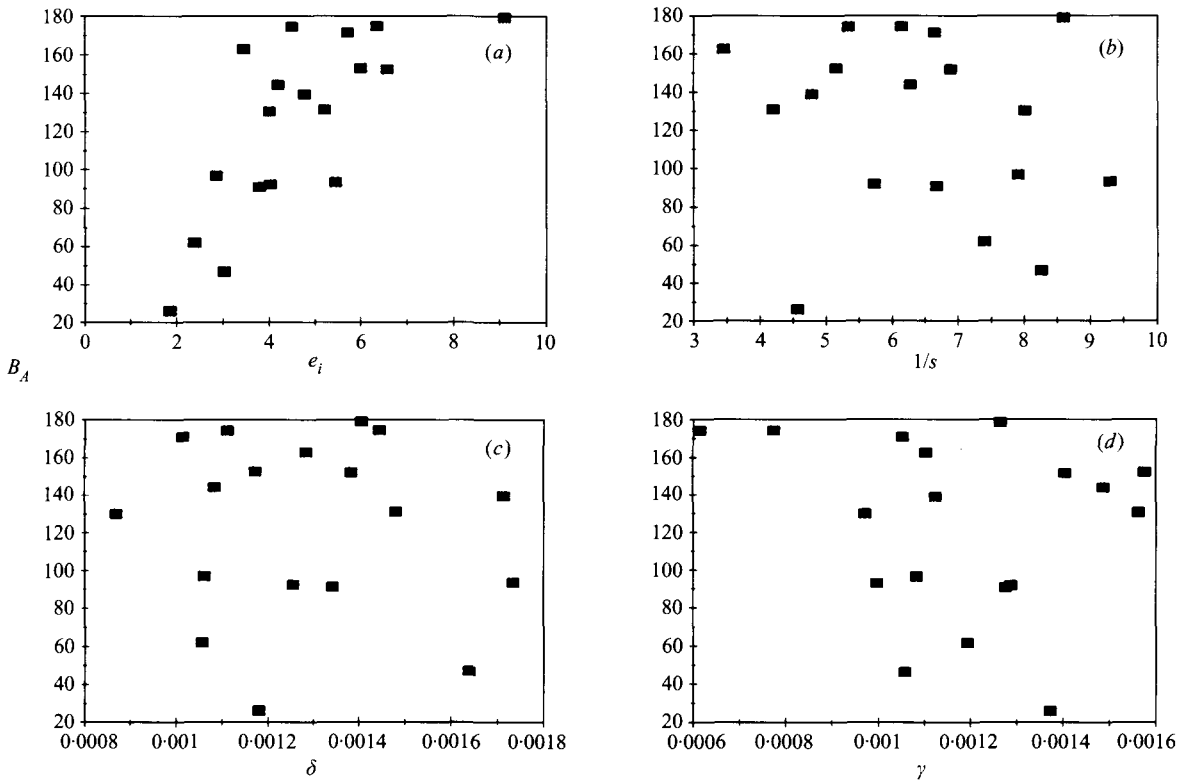

Fig. 7. Scatter graphs (*a–d*) showing the correlations between the benefit of the treatment programme (B_A) and input parameters, 7a. epg per worm (e_i), 7b. duration of immunity ($1/s$), 7c. strength of immunity (δ) and 7d. strength of density dependence (γ).

reinfection levels are very low for the majority of simulations but that a few simulations show considerable reinfection.

The correlation between the outcome variables is shown in Figure 5. It is observed that both benefit measures are negatively correlated with the basic reproductive number (Fig. 5*a*, *b*), as would be expected, but that neither correlation is complete ($R^2 = 0{\cdot}7$) which implies that the knowledge of the value of R_0 does not completely predict the benefit of a control programme. However, the two benefit values are almost completely positively correlated ($R^2 = 0{\cdot}95$) and therefore can be predicted from each other (Fig. 5*c*).

Initial statistical analysis of the simulation results suggested that a multiple linear regression model was not appropriate ($R^2 < 0{\cdot}7$). Further examination of the results revealed that this was due to two of the parameters showing a threshold type effect on the results. As observed previously, the benefit values showed a very skewed distribution with most values being high. The values of B_A can be divided into those above 180 (0) and those below (1). When these transformed data are plotted against the immunity parameters, δ and x, (Fig. 6) the simulations with low benefit values can be shown to almost all fall in the region $\delta > 0{\cdot}001, x < 3$. This means that whether there is significant reinfection can be considered as a categorical variable which depends on the level of immunity and not on other parameters. To determine the level of reinfection in those simulations with significant reinfection, a multiple linear regression model was fitted to these simulations (18 simulations) only.

The analysis for the selected simulations shows a very good fit with the linear model ($R^2 = 0{\cdot}89$–$0{\cdot}91$) (Table 3). The parameter giving the most significant correlation is e_i, the estimate of eggs per gram of faeces per worm which gave positive linear correlation with the benefit measures and negative linear correlation with the basic reproductive number. The other parameters which give significant correlations with the benefit measures are the strength (δ) and duration ($1/s$) of acquired immunity (this is in addition to the effect of acquired immunity determining whether or not there is reinfection). The strength of density dependence (γ) is also significant. The correlations of the significant parameters for B_A are shown

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

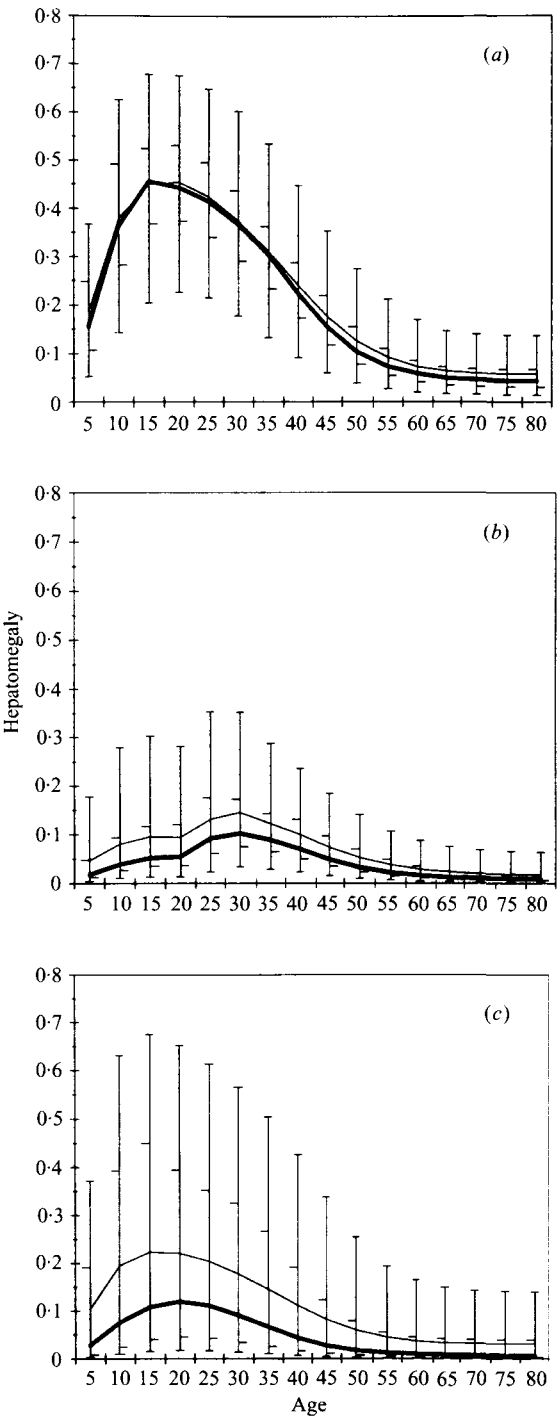

Fig. 8. Age profiles of proportion with hepatomegaly at years 0 (*a*), 20 (*b*), and 50 (*c*). The heavy solid lines show the median values and the light solid lines show the means. The top and bottom bars of the I bar represent the 95th and 5th percentile and the internal ticks the interquartile range.

in Figure 7. This figure shows a strong positive correlation with e_i but only weak correlations with the other parameters.

It is also worth noting that two of the simulations showing 'unusual' behaviour, namely the simulation showing an early increase in intensity during the treatment phase and one of the simulations outside the 'high risk' area on Figure 6 but showing substantial reinfection have a combination of a high value of the immune function exponent x and a low value for e_i (implying a high worm burden). This suggests that they may be in the region of parameter space denoted as region B in Figure 1 (where breakdown of immune protection during treatment and substantial reinfection occurs with high values of x).

Of equal interest to those parameters showing correlation with simulation outcomes are those which, perhaps surprisingly, show no correlation. The lifespan of the worm ($1/\mu$) and both the intercept and slope of the aggregation parameter function (k_0 and k_{lin} respectively) do not show significant correlation with any of the outcome variables (note that the morbidity parameters are not expected to, and do not, correlate with these outcomes). This is perhaps surprising in the case of the worm lifespan since this parameter is of importance in models without im-

Table 4. *Standardized regression coefficients (SRC) and t-statistics for the analysis of parameter uncertainty contributions to the observed levels of morbidity. The early disease is evaluated at age 15 and late disease at age 35 (the respective peaks) and at years 0, 20 and 50. Only the three most significant parameters are shown. The value of the correlation coefficient (R^2) is also given for the full regression model for each outcome variable*

Output	Parameter	SRC	*t*-statistic
Early Disease	r_E	0·6270	22·8477
Year 0	e_i	−0·5868	−21·4135
$R^2 = 0·97$	μ_{DE}	0·4472	16·3310
Early Disease	e_i	−0·5691	−4·9563
Year 20	δ	0·2738	2·3751
$R^2 = 0·51$	r_E	0·2597	2·2584
Early Disease	e_i	−0·5651	−7·4033
Year 50	δ	0·4212	5·4958
$R^2 = 0·79$	x	−0·3218	−4·1980
Late Disease	r_L	0·5736	13·7747
Year 0	r_E	0·4442	10·6355
$R^2 = 0·94$	μ_{DL}	0·3949	9·4269
Late Disease	e_i	−0·4759	−9·4223
Year 20	r_L	0·4696	9·3125
$R^2 = 0·91$	r_E	0·4402	8·7031
Late Disease	e_i	−0·5559	−7·3150
Year 50	δ	0·3920	5·1382
$R^2 = 0·79$	r_L	0·3298	4·3458

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

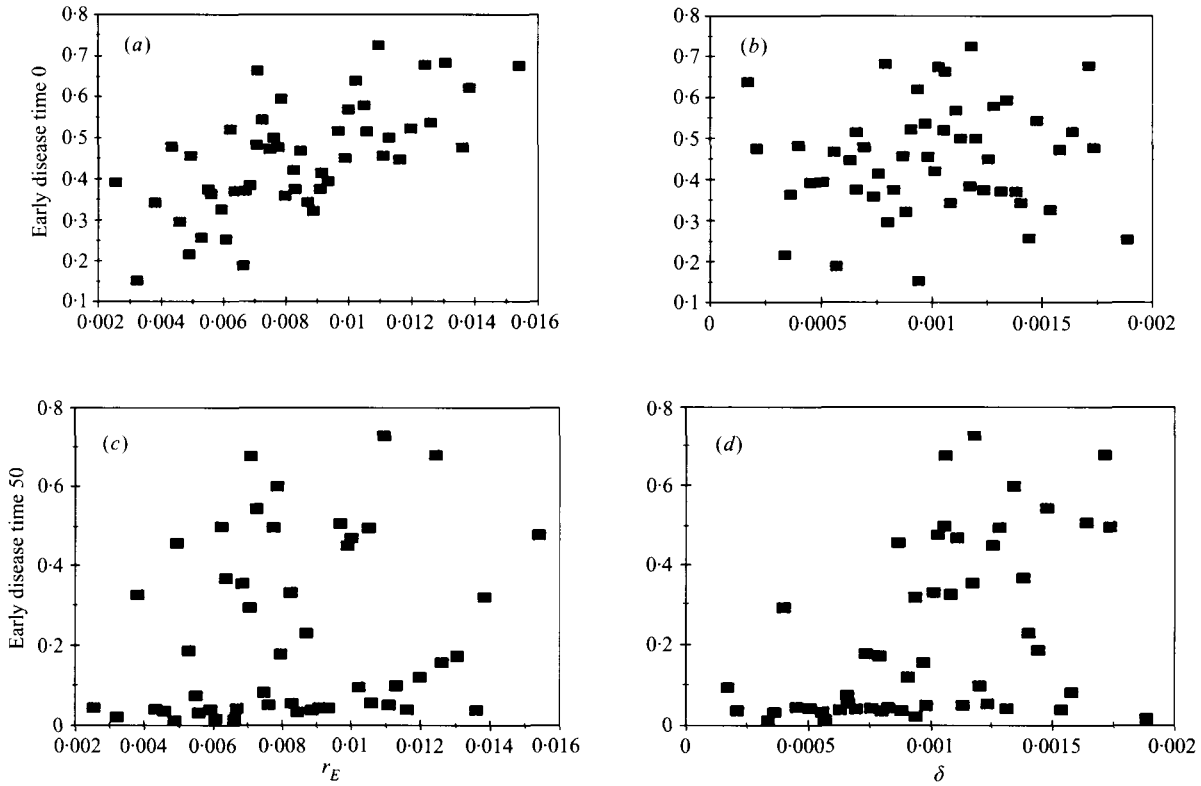

Fig. 9. Graphs comparing the relationship of the percentage of people aged 15 with hepatomegaly before the treatment programme (*a*, *b*) and during the reinfection phase (year 50) (*c*, *d*) with the rate of development of early disease r_E and the strength of immunity, δ.

munity but it appears that when immunity is present, the effects of immunity are much stronger than that of worm lifespan.

Prediction of chronic morbidity levels

Age profiles of hepatomegaly (early disease) are shown in Figure 8. There is a wide distribution of levels of early disease initially and this variation shows a symmetrical distribution (Fig. 8*a*). At year 20 (Fig. 8*b*), the median level of early disease has reached very low levels and the distribution of simulation has become skewed with a minority of simulations retaining high levels of disease, mirroring the pattern observed with the mean egg count. At year 50 (Fig. 8*c*), the skewed distribution has become much more pronounced with a high level of morbidity in a few simulations.

The parameter contributions to the uncertainty in both early disease and late disease (fibrosis) are shown in Table 4. Since there are many significant parameters, the three with the highest t statistics for each outcome variable only are shown. The values of early disease are given at age 15 and those for late disease at age 35 (the respective peaks). For both outcome variables, it is observed that the initial levels are mostly determined by the morbidity parameters and the reinfection levels (year 50) are mostly determined by the immunity parameters. Hence, as with the infection intensity, the immunity parameters are of primary importance in determining the consequences of control. The patterns are further illustrated in Figure 9 which shows that a stronger correlation is seen with the morbidity parameter initially and with the immunity parameter on reinfection. There is also a strong correlation with e_i at all time points. Also of interest is the fact that the parameter for early disease development is significant with respect to late disease (although less so than the development rate of late disease) and that in all cases it is the development of strength parameters (r_E, r_L and δ) which appears more significant than the rates of decay (s, μ_{DE}, μ_{DL}) of either immunity or morbidity.

DISCUSSION

The results of this investigation should be considered with an understanding of the necessary constraints imposed by the considerable uncertainty in the system.

https://doi.org/10.1017/S0950268800059239 Published online by Cambridge University Press

With respect to the structural sensitivity analysis it is not possible to include all types of schistosome models in the analysis. The sensitivity analysis was therefore restricted within the type of model to which the original model belonged. The central assumptions of this class of models are that the dynamics of the system can be represented in terms of differential equations of the mean worm burden, that the dynamics of the parasite in the intermediate host can be set to equilibrium and that immunity is modelled as a function of accumulated experience of infection [6, 7, 22]. Structural adjustments in the mechanism of density dependence and the type of immune response function are made within this framework. Therefore the analysis does not question the general approach to modelling schistosomiasis but only varies the form of some of the functional relationships between different model components.

Similar considerations apply to the parameter sensitivity analysis. There is a genuine considerable uncertainty in the expected values of most of the parameters and very little is known about their distributions. Therefore the same proportional change in the parameter values were included for those parameters where the distribution was unknown. The analysis is therefore actually estimating the relative parameter sensitivity of the model in the region of parameter space that is thought to reflect reality. The analysis is also intended to reflect biological uncertainty (the consequences of uncertainty in the knowledge of biological processes and parameters) and therefore experimental errors in estimating egg counts, for example, are not included. The results of the analysis should therefore be viewed from the perspective of these necessary limitations.

One general observation from the sensitivity analysis is the marked skewed distribution of model outcomes despite initial input parameter distributions which were almost symmetrical. The pattern observed was that in most simulations there was little reinfection but in the minority of cases reinfection levels were very high. Whether or not there was significant reinfection was almost entirely dependent on the parameters determining the strength and shape of the immune response function. Note that this should not be interpreted to mean that in most cases there will be little reinfection if control programmes are terminated, since this is dependent on the initial range of endemic parameters used.

It was also observed that the trajectories of infection intensity for the different simulations were very close during the treatment phase which suggests prediction of consequences of treatment can be carried out with relative confidence during his phase. In 2 out of 50 runs there was a 'breakdown' in control with infection intensity increasing while control was still occurring which suggests that this is a relatively unlikely event if the parameter space explored reflects reality. However, the reinfection phase is much more sensitive to parameter values with a wide range of outcomes being possible.

The most important source of uncertainty is very clearly demonstrated to be those relating to acquired immunity. Of the immune parameters the most important is the strength of immune protection followed by the shape parameter of the immune response function. The duration of immune protection appears to have less of an influence on the outcomes. This contrasts with results from sensitivity analyses where parameters are varied one at a time [13, 22].

The other parameter which strongly influences the results is e_i, the measure of eggs per gram of stool per worm. Since the endemic mean egg count was set to the same value for all simulations, e_i is essentially an inverse measure of worm burden. With the current formulation of the model, the rate of development of immunity and morbidity is scaled with respect to the worm burden, hence the worm burden will have a profound effect on the outcome of simulations.

An equally important result is the fact that all other parameters have very little effect on the model outcomes. Even for the prediction of morbidity, where the initial levels are determined by the morbidity parameters, the reinfection levels are also mainly determined by the immunity parameters and e_i. This suggests that uncertainty in the other parameters may not have significant influence on the accuracy of predictions.

The above results have several implications for the prediction of outcomes of control. The results suggest that the focus of prediction should be on the risk of the breakdown of control (since levels of infection during control where there is no breakdown are very similar) and on the amount of reinfection. This is important since sometimes control programmes may be abandoned abruptly when funding for a project terminates. The results also suggest that uncertainty in the majority of parameters has little effect on the overall results and it is only the immunity parameters and e_i which have a major influence on the results.

There are two approaches to tackling the uncertainty in the parameter e_i. Firstly, it may be

possible to improve the estimates of this parameter by the use of data from post-mortem studies [23], and perfusion studies [24] which are the only direct method currently available to estimate worm burden. Reliable estimates are difficult to obtain since it is not possible to use large sample sizes and there is considerable variability in the results. Alternatively, statistical methods can be used [9, 25, 26] but these present similar problems. Another approach would be to remove this parameter from the model altogether and rescale the model formulation in terms of mean egg count. Since there is a linear relationship between the worm burden and egg count in the model, this reformulation will not involve any change in the model structure but has the added advantage that parameters will be easier to estimate in terms of egg counts.

Despite numerous laboratory studies of schistosomiasis immunology, there are considerable gaps in our current understanding of the role of immunity in schistosomiasis epidemiology and in particular very little quantitative information from which to estimate parameters is available. Again, there are two approaches to dealing with this uncertainty. Greater understanding of both the actual role of immunity in schistosomiasis infection [27, 28] and of the behaviour of different immuno-epidemiological models [22] will eventually reduce the uncertainty in the predictions. Immunoepidemiological studies in which both specific antibody levels and infection intensity are measured can be used to increase our understanding of immunological processes. Model parameterization from such studies is extremely difficult due to the variability in immunological variables and the number of parameters which would need to be estimated. It is recognized that for the foreseeable future, most of the uncertainty will still be unresolved. In this case, sensitivity analysis can be used as an integral part of the prediction process. For example, all projections can be accompanied by a base case (median) and a bad (75 percentile) and worst case (95 percentile) scenarios. In this way, both a prediction based on our understanding of the situation and the risk of adverse outcomes can be taken into account.

ACKNOWLEDGEMENTS

I thank Graham Medley for help with the computing. This study was supported by the Edna McConnell Clark Foundation.

REFERENCES

1. WHO. The control of schistosomiasis. Geneva: World Health Organisation, 1993. (Technical Report Series; vol. 830.)
2. WHO. Progress in the assessment of morbidity due to schistosomiasis. Geneva: World Health Organisation, 1989.
3. Warren KS, Bundy DAP, Anderson RM, et al. Helminth infections. In: Jamison DT, Mosley WH, Measham AR, Bobadilla JL, eds. Disease control priorities in developing countries. Oxford: Oxford University Press, 1993.
4. Hairston NG. On the mathamatical analysis of schistosome populations. Bull WHO 1965; **33**: 45–62.
5. MacDonald G. The dynamics of helminth infections, with special reference to schistosomes. Trans Royal Soc Trop Med Hyg 1965; **59**: 489–506.
6. Anderson RM, May RM. Herd immunity to helminth infection and implications for parasite control. Nature 1985; **315**: 493–6.
7. Anderson RM, May RM. Helminth infections of humans: mathematical models, population dynamics and control. Adv Parasitol 1985; **24**: 1–101.
8. Woolhouse MEJ. On the application of mathematical models of schistosome transmission dynamics. 1. Natural transmission. Acta Tropica 1991; **49**: 241–70.
9. Chan MS, Guyatt HL, Bundy DAP, Booth M, Fulford A, Medley GF. The development of an age structured model for schistosomiasis transmission dynamics and its validation for *Schistosoma mansoni*. Epidemiol Infect 1995; **115**: 325–44.
10. Plaisier AP, van Oortmarsenn GJ, Habbema JDF, Remme J, Alley ES. ONCHOSIM: a model and computer simulation program for the transmission and control of onchocerciasis. Comp Methods Prog Biomed 1990; **31**: 43–56.
11. Chan MS, Guyatt HL, P Bundy DAP, Medley GF. Dynamic models of schistosomiasis morbidity. Am J Trop Med Hyg 1996; **51**: 52–62.
12. Chan MS, Anderson RM, Medley GF, Bundy DAP. Dynamic aspects of morbidity and acquired immunity in schistosomiasis control. Acta Tropica 1996. In press.
13. Chan MS, Bundy DAP. The effects of community chemotherapy on patterns of morbidity due to *Schistosoma mansoni*. Trans Roy Soc Trop Med Hyg 1996. In press.
14. Blower SM, Dowlatabadi H. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. Internat Stat Rev 1994; **62**: 229–43.
15. Blower SM, Hartel D, Dowlatabadi H, Anderson RM, May RM. Drugs, sex and HIV: a mathematical model for New York City. Philosoph Trans Royal Soc London, Series B, Biological Sciences 1991; **331**: 171–87.
16. Rowley JT, Dowlatabadi H, Anderson RM. The potential magnitude and demographic consequences of the HIV epidemic: parameter uncertainties and the reliability of model projections. J AIDS 1996. In press.

17. Lord CC, Woolhouse MEJ, Rawlings P, Mellor PS. Simulation studies of African Horse Sickness and *Culicoides imicola* (Diptera: Ceraopogonidae). J Med Entomol 1996; **33**: 328–38.
18. McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 1979; **21**: 239–45.
19. Iman RL, Conover WJ. Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. Communications Stats 1980; **A9**: 1749–842.
20. Butterworth AE, Sturrock RF, Ouma JH, et al. Comparison of different chemotherapy strategies against *Schistosoma mansoni* in Machakos District, Kenya: effects on human infection and morbidity. Parasitol 1991; **103**: 339–55.
21. RIVM. UNCSAM version 1.1 Uncertainty/ sensitivity analysis using Monte Carlo sampling. 1.1 ed. The Netherlands: National Institute of Public Health and Environmental Protection, 1992.
22. Woolhouse MEJ. A theoretical framework for the immunoepidemiology of human helminth infection. Parasite Immunol 1992; **14**: 563–78.
23. Cheever AW. A quantitative post-mortem study of *Schistosomiasis mansoni* in man. Am J Trop Med Hyg 1968; **17**: 38–64.
24. Domingues L, Silveira M, Lima JF, Carreiro JC, Kelner S. Removal of *S. mansoni* in patients with hepatosplenic schistosomiasis: an estimate of the parasitological load by means of quantitative coproscopy. Rev Inst Med Trop Sao Paulo 1983; **25**: 2–15.
25. Booth M. The epidemiology and population biology of multiple infections with *Ascaris lumbricoides*, *Trichuris trichiura* and hookworms. Doctoral thesis: London: University of London, Imperial College, 1994.
26. de Vlas SJ, Gryseels, B, van Oortmarssen GJ, Polderman AM, Habbema JDF. A model for variations in single and repeated egg counts in *Schistosoma mansoni* infections. Parasitology 1990; **104**: 451–9.
27. Butterworth AE, Dunne DW, Fulford AJ, et al. Human immunity to *Schistosoma mansoni*: observations on mechanisms, and implications for control. Immunol Invest 1992; **21**: 391–407.
28. Maizels RM, Bundy DAP, Selkirk ME, Smith DF, Anderson RM. Immunological modulation and evasion by helminth parasites in human populations. Nature 1993; **365**: 797–805.