Review TRENDS in Parasitology Vol.20 No.11 November 2004

Evaluating the epidemiological impact of national control programmes for helminths

Simon Brooker[1], Sarah Whawell[2], Narcis B. Kabatereine[2,3], Alan Fenwick[2] and Roy M. Anderson[4]

[1]Department of Infectious and Tropical Disease, London School of Hygiene and Tropical Medicine, Keppel St, London, WC1E 7HT, UK
[2]Schistosomiasis Control Initiative, Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College, Norfolk Place, London, W2 1PG, UK
[3]Vector Control Division, Ministry of Health, Kampala, PO Box 1661, Uganda
[4]Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College, Norfolk Place, London, W2 1PG, UK

During the past decade, there has been renewed commitment to programmes for helminth control, and several international initiatives have been launched. Scientific evaluation of large-scale interventions to reduce the incidence of infection and associated morbidity and mortality is vital, not only to demonstrate health benefits, but also to assess cost-effectiveness to show that monies from donor agencies have been used wisely. Using the example of schistosomiasis, this article outlines a framework for the evaluation of the impact of national control programmes, highlighting the important epidemiological and practical issues that must be addressed.

An increasing number of international initiatives have been established that aim either to reduce or to eliminate the disease burden caused by a variety of helminth parasites prevalent in the resource-poor regions of the world. Two examples are the global elimination programmes for lymphatic filariasis [1] and onchocerciasis [2]. For soil-transmitted helminthiasis and schistosomiasis, the 54th World Health Assembly passed a resolution urging member states to provide regular drug treatment for high-risk groups (http://www.who.int/gb/ebwha/pdf_files/WHA54/ea54r19.pdf). In support of this important milestone in helminth control, WHO has set the global target of regularly treating at least 75% of all school-aged children by the year 2010. This has received new sources of funding, notably from the Bill and Melinda Gates Foundation, which has, for example, recently supported the Schistosomiasis Control Initiative (SCI) (http://www.schisto.org). In this five-year initiative, country-wide control in six African countries is to be based on chemotherapy using praziquantel [3]. Target groups include school-aged children and at-risk occupational groups, and also whole communities where infection prevalence exceeds 50%. However, to ensure long-term sustainability, this approach will need to be integrated with preventive measures including health education and improved access to safe water and sanitation [4].

Defining targets

The first task in any control programme is to define programme targets in as precise a manner as possible to facilitate scientific evaluation of IMPACT (see Glossary). The SCI has the objectives to achieve a reduction in intensity of infection and to reduce and prevent morbidity associated with schistosomiasis, including improved haemoglobin and reversal of liver fibrosis. Linked to this is the careful design of evaluation programmes to assess whether the defined aims have been achieved. Too often in the past, helminth control is implemented with little thought given either to targets or to the design of sampling

Glossary

Adequacy: comparison of the performance or impact of the programme against previously established targets.
Coverage: measure of the extent an intervention covers the need for its use, and defined by the percentage of the target population that received the intervention.
Effectiveness: evaluative measure of the impact of the intervention in the context of its routine or programme-based implementation.
Efficacy: evaluative measure of the extent to which a defined intervention produces a beneficial result under ideal conditions, typically in the context of a randomised placebo-controlled trial.
Impact: the health and educational benefits of a defined intervention.
Performance (provision, utilization and coverage): measures of programme activities among intended recipients.
Plausibility: the programme has an effect above and beyond the effect of external, non-programme influences, by the inclusion of control groups.
Power: the probability that the null hypothesis is rejected when it is false. For example, the null hypothesis is that there is no difference between mean intensities pre- and post-intervention. Mathematically, power is $1-\beta$, where β is the type 2 error, or probability of failing to reject the null hypothesis when it is false.
Probability: the programme has an effect, with a small and defined probability that the difference between programme and control areas were due to chance or confounding.
Provision: measure of the quantity of an intervention or service delivered.
Significance level (or type 1 or α error): the probability of rejecting a true null hypothesis; 5% is normally seen as the acceptable level of significance. α and β are required to be as small as possible (and therefore power as large as possible) but there is a trade-off between the two.
Utilization: measure of the percentage of delivered intervention being used.

Corresponding author: Simon Brooker (simon.brooker@lshtm.ac.uk).
Available online 15 September 2004

 1471-4922/$ - see front matter doi:10.1016/j.pt.2004.08.012

frames to permit robust evaluation of what has been achieved. For example, the major problem of the earlier schistosomiasis control efforts implemented through Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) funding was that chemotherapy-based morbidity control was evaluated on the basis of infection indicators rather than morbidity indictors.

Impact and performance

An important conceptual distinction in the context of evaluating the impact of health interventions is EFFICACY versus EFFECTIVENESS (see Glossary for definition). The evaluation of efficacy is concerned with the beneficial result of an intervention under ideal conditions, and is usually based on a randomized, double-blind, placebo-controlled trial design for the evaluation of interventions delivered at the individual level, or large-scale, cluster-randomized, controlled trials for interventions delivered at the community level [5,6]. For national control programmes, evaluation should aim to assess effectiveness: the impact of an intervention in the context of its programme-based implementation.

Recently, Habicht *et al.* [7] reviewed appropriate evaluation designs for large-scale health programmes. Their framework involves two axes, the first of which distinguishes between programme impact (on health and behavioural indicators; i.e. effectiveness) and programme PERFORMANCE (PROVISION, UTILIZATION and COVERAGE), whereas the second axis relates to the type of inference to be made, including ADEQUACY, PLAUSIBILITY or PROBABILITY assessments). Adequacy evaluations assess how well the programme targets have been met; for example, 75% coverage among at-risk school-aged children or a predefined reduction in intensity of infection. The obvious shortcoming of this form of assessment is that observed changes might have been caused by unmeasured secular changes, which could only be tracked by monitoring similar communities in which no control was implemented. Plausibility assessments involve an attempt to control for the influence of confounding factors by choosing either internal or external control groups for comparison over the same time period. However, as Habicht *et al.* [7] highlight, such assessments cannot completely rule out alternative explanations for observed changes. Probability evaluation must aim to ensure that there is only a small and clearly defined probability that the difference between intervention and control groups is due to confounding factors. This type of evaluation requires randomization of treatment and control groups, either from outset or in a carefully designed introduction over the course of the intervention, which could be months or years.

Whereas evaluation of programme impact is usually done at the end of a programme, evaluation of programme performance is undertaken at periodic intervals during implementation to inform planning, and to adjust and improve the intervention [8,9]. An example would be to assess drug efficacy to monitor for the emergence of drug resistance [10]. Several other performance indicators can also be usefully included in monitoring programmes: detection of side-effects of treatment; degree of local financial and political support; evenness (equity) of drug distribution; and training of teachers and community drug distributors. Treatment coverage should be included both in performance evaluation, to assist both in routine planning [9], and in impact evaluation, where the comparison of populations with varying degrees of treatment coverage can help assess impact by establishing a relationship between degree of programme coverage and programme impact [11].

The precise type of evaluation will ultimately depend on the intended audience of the evaluation [7]. For example, a district health officer might be interested in performance evaluation, such as the provision and coverage of treatment, easily assessed by using routine information systems, whereas the Ministry of Health or donors may be more interested in an impact evaluation on morbidity, based on a more robust scientific approach, as outlined here.

What is to be evaluated?

Because the main aim underlying helminth control is to prevent disease rather than to reduce or eradicate transmission, it is important that evaluation focuses on both infection and morbidity measures where possible. One ubiquitous feature of helminth parasites is that the intensity of infection is a key determinant of morbidity with high levels likely to induce greater morbidity. As a consequence, helminth control has typically been based on changes in the intensity of infection. However, increasingly, quantitative morbidity measures are possible such as ultrasound for assessing complications of urinary (due to *Schistosoma haematobium*) and intestinal (due to *Schistosoma mansoni*) schistosomiasis [12], and anthropometric and cognitive assessment for measuring child growth or educational impairment [13,14]. The advantages and disadvantages of the various direct and indirect indicators for the detection of exposure to infection and the occurrence of morbidity are discussed in Box 1.

In selecting appropriate indicators, quantitative measures are the most desirable, with due account taken of variability, sensitivity and specificity in measurement, plus acceptability, ease and cost of measurement [9]. In situations where there are multiple national programmes, as with the SCI, there is a need to establish standardization because this will help make comparison within and between programmes possible. Especially important is the validation of indicators over time because it remains poorly understood how the sensitivity and specificity of indicators, such as haematuria and self-reported morbidity, varies alongside the implementation of chemotherapy-based programmes [15].

As indicated above, it is also important to make every effort to take potentially confounding variables into account. In the case of schistosomiasis, geographical differences in pathology patterns [16], past experience of infection [17], concomitant infections and conditions, and immunogenetic background [18] may confound outcomes (see Box 1). Where possible, these require assessment both pre- and post-intervention.

The ultimate choice of indicators is likely to depend on local relevance and the scale of evaluation, with different

Box 1. Indicators and methods for the evaluation of schistosomiasis control

Prevalence and intensity of infection

Detection by microscopy and counting of eggs in urine or faeces is the most frequently used method to determine prevalence and intensity of infection. Methods for detection of circulating antigen or antibody are increasingly being developed, but as these do not, at present, reflect intensity of infection, they are unlikely to have large-scale application in evaluation [38]. Morbidity questionnaires administrated by health or educational staff have been shown to be useful for approximating prevalence of infection, and can be used for the rapid screening of the need for treatment [15]. They also have a potential use for monitoring the impact of control, although this has yet to be investigated.

However, measurement of prevalence alone is not a useful indicator of impact because of the non-linear relationship between prevalence and mean intensity of infection within a community, which depends on the degree of worm aggregation within the host populations [25] (Figure Ia). It follows that changes in intensity in response to treatment might not be reflected in changes in prevalence, such that intensity of infection should always be used for the parasitological assessment of impact. Owing to high day-to-day variability of egg counts within individuals, there is a need for the collection of multiple samples [39], although this is tedious and might reduce compliance.

Haematuria

Haematuria (blood in urine) is easily assessed by urinary reagent strips and indicates the pathological effects of *Schistosoma haematobium* infection. It has been used successfully for community or individual diagnosis and has been confirmed to be an effective surveillance tool after chemotherapy [40]. However, the potential for evaluating impact of intervention remains poorly understood.

Bloody stool

Although intestinal morbidity might be reflected by the presence or recent history of bloody stool, this indicator is too nonspecific to be a reliable marker of morbidity and the impact of intervention [41].

Clinical complications

The pathology as a result of schistosomiasis mainly arises from the host's cell-mediated inflammatory granulomatous reaction around trapped schistosome eggs, with the frequency and severity, in selected populations, non-linearly related to the prevalence of infection (Figure Ib). Infection levels and disease are not strictly related however because the development of pathology could depend on factors other than infection levels including nutrition, toxins, concurrent infections and intrinsic host characteristics [18]. Ultrasonography is a safe, rapid, non-invasive technique that enables complications of schistosomiasis, such as hydronephrosis, bladder cancer and portal hypertension, to be easily assessed, and the course after therapy to be monitored [12]. Recent WHO workshops have contributed to standardizing the methodology, thereby enabling comparable results between programmes. Although palpating the liver and spleen can assess hepatic and splenic morbidity, there is high inter-examiner variability and a lack of accepted standardization. Furthermore, in malaria-endemic regions, splenomegaly is not a reliable indicator for schistosome morbidity. In general, reliable assessment of clinical pathology due to intestinal schistosomiasis is more difficult because of its generalized form.

Growth

Despite some studies demonstrating an association between schistosomiasis and soil-transmitted helminths and under-nutrition [14], trials on the nutritional impact of treatment have yielded less convincing results [42]. However, these studies have a follow-up period of <1 year, during which time it is not possible to detect any treatment effect [43]; thus, a follow-up period of >1 year is recommended to assess impact.

Anaemia

Although schistosomiasis can result in direct blood loss as the spined eggs produced by female parasites rupture blood vessels surrounding the intestine (*Schistosoma mansoni*) and the bladder (*Schistosoma haematobium*), the extent schistosomiasis might contribute to anaemia is controversial [14,16]. Furthermore, the effect of treatment on anaemia will depend on many factors, including the intensity of infection, the underlying iron status and the presence of concomitant infections, such as malaria and hookworm [44,45]. As a consequence, it is not always easy to assess the impact of treatment on anaemia, which underscores the importance of assessing for other parasitic infections.

Educational achievement, cognition and productivity

For more than 60 years, it has been suggested that schistosomiasis might adversely affect outcomes of educational achievement, cognition and productivity [13,46]. The mechanism by which these outcomes are affected is uncertain, but evidence suggests that the mechanism is indirect, perhaps mediated through iron-deficiency anaemia and undernutrition [47]. Several drawbacks exist in assessing educational achievement and cognition: tests need to be developed over an extended period to ensure face validity and content validity of the tests (correlations among tests, and between school achievement

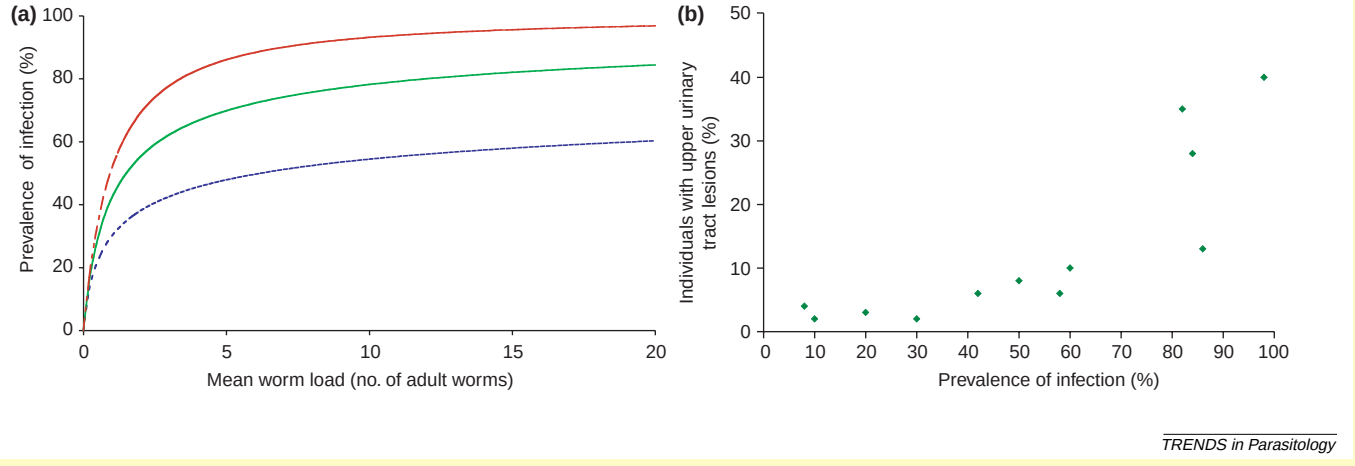

Figure I. The non-linearities of schistosomiasis. **(a)** Theoretical relationship between mean intensity and prevalence of infection based on the assumption that worms are distributed in a negative binomial manner with aggregation parameter k [k=0.2 (blue); k=0.5 (green); k=1.2 (red)]. **(b)** Relationship between prevalence of disease and intensity of infection in communities in Burundi (Ref. [49]). The relationship is strongly non-linear with the percentage of individuals with detectable lesions of the urinary tract increasing sharply at an infection prevalence of $>60\%$.

and tests, follow the predicted pattern), to ensure that children were familiar and comfortable with all testing materials and to ensure adequate test–retest reliability; and it is likely that long-term follow-up is required. Research on the cognitive impacts of parasites in developing countries has therefore been limited by the considerable expertise and extensive training typically required for fieldworkers to conduct cognitive assessments. However, recent methodological innovations have found that class-based tests of cognitive function administered to groups of children by teachers with minimal training can be sensitive to the impact of health interventions (M. Jukes, pers. commun.).

Mortality

Schistosoma haematobium infection can result in mortality as a consequence of kidney failure and bladder cancer, and mortality associated with *S. mansoni* infection can be caused by haematemesis (vomiting blood) or liver failure. However, there is a complete lack of data on the impact of treatment on schistosome-related mortality, and evaluation of this outcome would require prohibitively large sample sizes [48]. Encouragingly, the contribution of schistosomiasis to mortality and the impact of treatment might potentially be assessed by data from demographic surveillance systems (DSS) [22].

indicators appropriate for different levels of evaluation, and will correspond to a different subset of indicators for different schistosome species. At the national level, indicators are necessarily restricted, and therefore we suggest that intensity of infection might be an appropriate indirect indicator for evaluation on a large scale. In a limited number of sentinel schools, ultrasonography, nutritional assessment (growth and anaemia) and education performance assessment is recommended. As well as epidemiological indicators, integrated control strategies will necessitate additional indicators such as the coverage of clean water and sanitation and the impact of health education [4].

Who to evaluate?

Generally, schoolchildren are used as the sentinel group with which to evaluate the impact of schistosomiasis control. There are several reasons for this. First, age-stratified epidemiological surveys consistently show that age–infection profiles for schistosome infection are typically convex, with prevalence and intensity peaking in older children or in young adults [19]. The rapid rise in intensity within the child-age classes, before acquired immunity and decreased exposure reduce the incidence of new infections, provides a good marker against which to measure change induced by control. Second, school-aged children are particularly at risk from the clinical manifestations of disease resulting from frequent infection. Severe morbidity might result from cumulative exposure and be expressed later in life. Third, reversibility of pathology with treatment is quickest among school-aged children; treatment of children, even with one round of chemotherapy, has been shown to reduce morbidity that would have been manifest later in life [16]. Multiple treatment rounds are necessary to achieve sustained reduction of morbidity among adults. Finally, there are also practical considerations for selecting schoolchildren because schools are accessible; experience shows that there is generally good compliance and follow-up is relatively easy within such settings [20].

By contrast, evaluation among the whole community is more problematic. The currently most-promising platform for longitudinal, community-based, evaluation of health-related interventions is demographic surveillance systems (DSS) in sentinel sites. The INDEPTH Network and its partnering sites (http://www.indepth-network.org/) have been collecting longitudinal data on pregnancies, births, deaths and migration for several years. Building on this unique platform, there is potential to complement core data collection systems by the inclusion of disease-specific information into routine activities. Recently, for example, DSS have been used to estimate the impact of social marketing of treated nets on child survival in Tanzania [21] and the impact of malaria control activities on malaria-related mortality [22]. Apart from documenting the impact of intervention on health outcomes, DSS also offer the potential to assess the impact of intervention on the broader aspects of social and economic equity. The establishment of SCI-funded programmes in multiple African countries provides a unique opportunity to explore the suitability of DSS as evaluative platforms for schistosomiasis control.

How many individuals to evaluate?

Determining the sample size required to detect a defined reduction in either intensity of infection or morbidity measures following treatment is vitally important because samples that are too small might have insufficient POWER to estimate precisely a significant intervention effect; excessive power demanding large sample sizes will waste resources [6]. For many infectious diseases, mathematical models of transmission dynamics can provide an insight into the expected impact of a defined intervention such as the treatment of a certain fraction of the population at defined intervals or similar treatment targeted at school-aged children [23–25]. Aside from generic work on the conceptual framework required to address questions of predicted impact based on knowledge of the population biology of the parasite, specific simulation packages have been developed for particular helminth species. For example, ONCHOSIM, a microsimulation model for onchocerciasis transmission, has been used to explore the implications of different control strategies in varying settings with defined intensities of infection [26,27].

Deterministic models of schistosomiasis [23–25] describe the dynamics of worm burdens and immunity, and, by predicting rates of re-infection over time, provide estimates of infections, heavy infections and morbidity cases prevented each year [28,29]. These models indicate why evaluation of control impact should be based on measures of intensity, not prevalence. The reason for this is that large reductions in transmission intensity (as measured by the basic reproductive number, Ro, which defines the average number of eggs produced by one female worm over her reproductive life that themselves survive to reproductive maturity [23]) will induce changes in infection intensity in a proportional manner, whereas identical changes will induce little measurable change in infection prevalence until transmission drops to very low levels.

Box 2. Calculating sample size for detecting a change in intensity of infection in Uganda

Detecting a reduction in mean intensity of infection pre- and post-intervention can be computed using a two-sample *t*-test if the raw data are a sample from an underlying normal distribution. If measurements are repeated on the same individual, then the correlated nature of the data can be accommodated using a paired two-sample *t*-test. The test requires knowledge of: the two means, μ_1 and μ_2; their standard deviations, σ_1 and σ_1; the correlation coefficient *r*; the SIGNIFICANCE LEVEL at which we want to detect a difference, α; and the desired power $(1-\beta)$ of the test (see Glossary). Power calculations simply rearrange the *t*-test formula so that power can be calculated as sample size varies (see Ref. [50] for details).

On the basis of available *S. mansoni* data for schoolchildren from Uganda [Schistosomiasis Control Initiative (SCI), pers. comm.], the necessary sample sizes for cohort groups aged 6, 7, 8 and 11 years can be calculated to satisfy a power requirement of 90% as a benchmark and a significance level of 5%. Because egg counts are highly over-dispersed, a $\log(\chi_i + k/2)$ transformation of the raw egg counts was used [51] and *k*, an inverse measure of worm aggregation in the host population, was estimated by maximum likelihood methods using the negative binomial relationship between prevalence and intensity [25] (Figure I) This transformation serves to stabilize the variance, permitting a quadratic relationship to be established between the mean and standard deviation of the transformed counts using regression techniques.

Sample size theory indicates that, as the correlation between two measurements increases, the necessary sample sizes for a given power decreases; the increased correlation reduces the combined standard deviation using the paired *t*-test. In epidemiological studies of helminths, the degree of correlation between egg counts in faecal samples from individuals before successful treatment and after a period of further exposure and reinfection is referred to as predisposition. The degree of predisposition can be assessed by the non-parametric Kendall's correlation coefficient, *r*, and published studies have estimated a range of 0.19 to 0.24 [30–31]. As shown in Figure IIa, the necessary sample size for a given power decreases as *r* increases (the strength of predisposition rises). The sample size for a given power is also related to the baseline mean intensity of infection (Figure IIb).

An alternative method to estimate sample sizes for over-dispersed data involves calculating asymptotic standard errors of the logarithm of the mean in a negative binomial generalized linear model, extended to a two-sample hypothesis test [52,53].

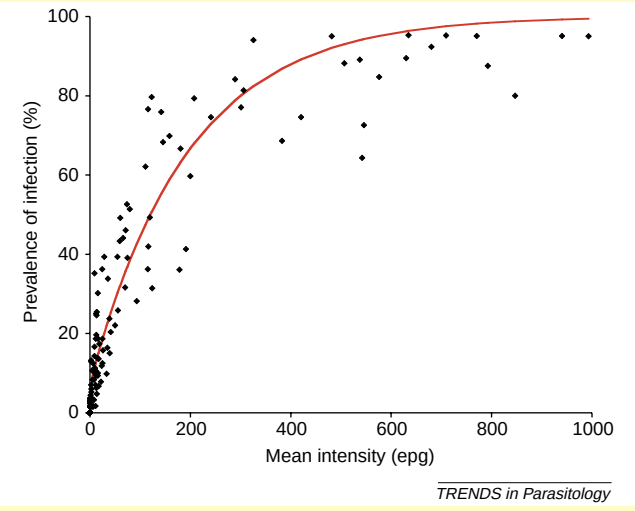

Figure I. The relationship between prevalence and intensity of infection. Best-fit prevalence-intensity curve based on 13 796 children in 201 Ugandan schools [54], where *k* was a linear function of mean intensity ($k=0.001+0.0007$ m). Estimates of *k* should be based on analysis of country-specific epidemiological data because differences in the magnitude of *k* will influence sample sizes: as *k* increases, sample sizes rise.

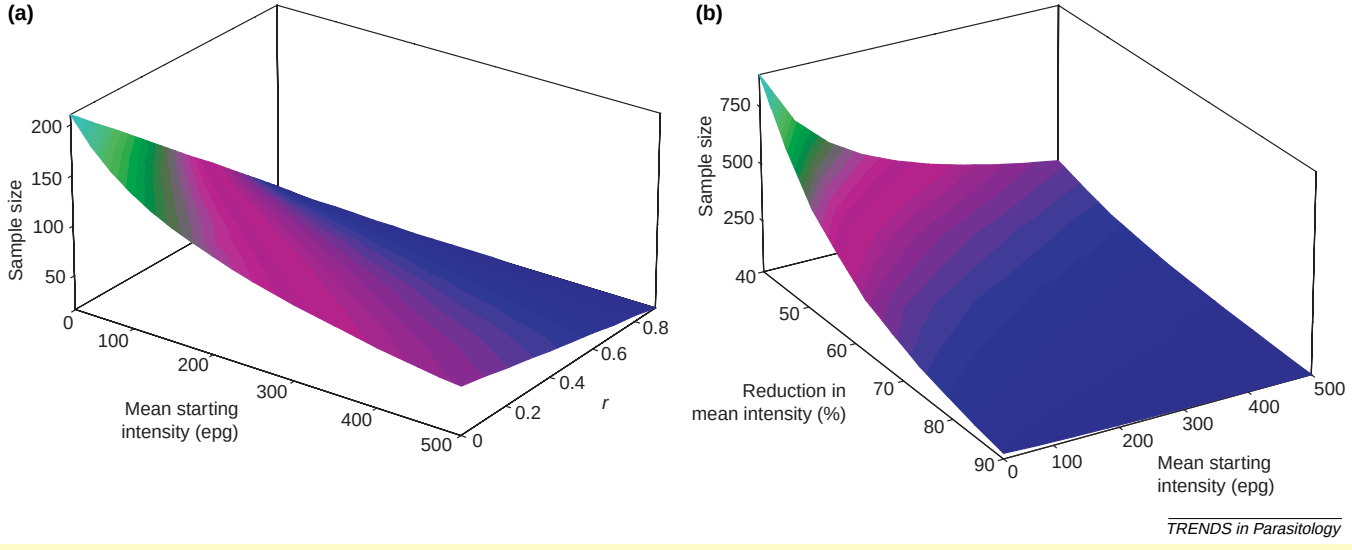

Figure II. Factors determining sample size. The relationships between sample size required to detect a 70% reduction in mean intensity of infection with a power of 90% given **(a)** varying degrees of predisposition (as assessed by Pearson's correlation coefficient, *r*) for an initial mean intensity of infection of 400 eggs per gram faeces (epg), and **(b)** varying initial intensity of infection, where $r=0.2$.

Such models of schistosomiasis have been incorporated into the EPISCHISTO software tool (http://www.schoolsandhealth.org/epidynamics.htm), and can be used, for example, to explore the consequences of a hypothetical five-year control programme for *S. mansoni* in Uganda. Incorporating differing assumptions of treatment schedule and coverage, drug efficacy, and differing baseline intensities of infection [ranging from 20 to 600 eggs per gram faeces (epg), as a population average taking account of population structure and changes in intensity with age], the intervention impact can be estimated and the appropriate sample size required for a defined power

Box 3. Use of geographical information systems and remote sensing in guiding sampling procedures

The spatial dynamics of parasitic transmission has received increased attention in recent years, not least because of advances in the use of geographical information systems (GIS) and remotely sensed satellite data [55]. Figure Ia presents available survey data for Uganda and shows marked variation in the prevalence of infection (data described in Ref. [54] in Box 2). Figure 1b shows ecological zones, based on remotely sensed derived ecological variables (methodology is described in Ref. [56]). In Uganda, these zones include arid areas and highlands where transmission of *Schistosoma mansoni* does not occur, or occurs at very low levels. It also includes three ecological zones suggested to have different levels of transmission. This is corroborated by estimations of the force of infection (FOI) derived from the analysis of age-stratified data for each zone (Figure Ic–e). The maximum likelihood method described by Williams and Dye [57] is used to estimate FOI. The resultant estimates of FOI show different patterns of transmission as indicated by the ecological zones. Such zones can help guide the planning of sampling procedures; thereby, locations of sentinel schools can be identified in each ecological zone, excluding arid and highland areas.

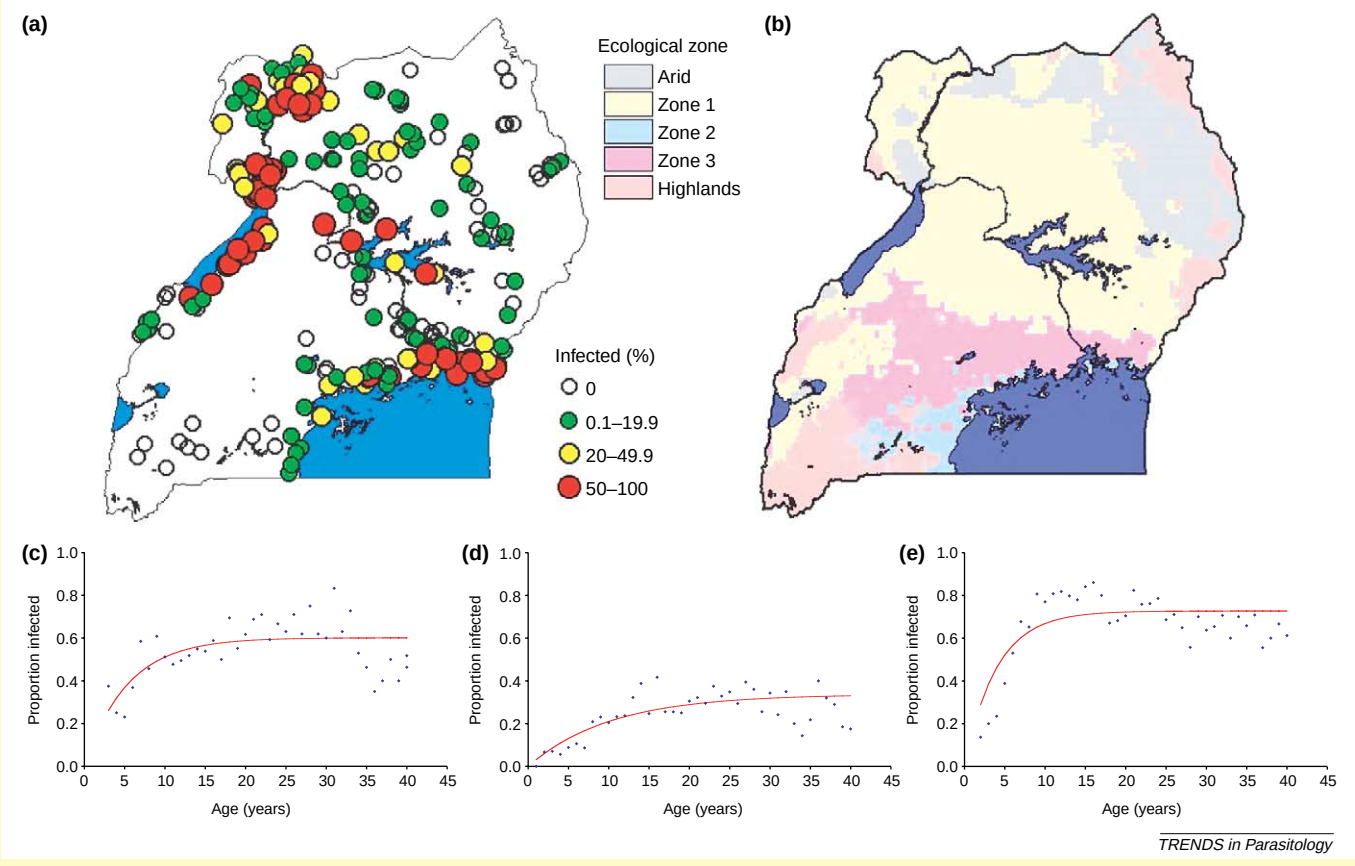

Figure I. Spatial variation in schistosomiasis transmission in Uganda. **(a)** The geographical distribution of *Schistosoma mansoni* and **(b)** ecological zones in Uganda, and differences in the transmission of infection as indicated by estimates of the force of infection (FOI) in **(c)** zone 1 [FOI=0.11 (11 new infections per 1000 people per year) based on 2898 individuals], **(d)** zone 2 (FOI=0.03, based on 2599 individuals) and **(e)** zone 3 (FOI=0.18, based on 2716). Solid circles represent observed data and the solid lines indicate maximum likelihood fitted estimates of FOI [57].

can be derived (Box 2). In Uganda, it is predicted that the percentage reductions in mean intensities following one round of treatment targeted at school-aged children with high coverage (>90%), would range 58–83%. On this basis, the expected reduction in mean intensity can conservatively be estimated to be 60–70%. This estimate can be used to determine sample sizes for detecting significant changes in the intensity of infection.

One important consideration in determining sample size is the occurrence of predisposition to heavy or light infection. Predisposition is indicated by a positive correlation between egg counts in faecal samples from individuals before successful treatment and after a period of further exposure and reinfection, and has been reported for both *S. mansoni* [30,31] and *S. haematobium* infection [32]. It might arise as a consequence of differences either in individual susceptibility or in environmental exposure, or both. This raises the question of genetic factors influencing susceptibility to infection and disease, which has recently been suggested by studies of genetic markers [33]; the importance of heterogeneity in exposure has long been recognized [34]. Analysis presented in Box 2 indicates that, as the degree of predisposition increases, the sample size for a given power decreases (Figure IIa in Box 2). The results also show that the sample size for a given power varies according to the baseline mean intensity of infection (Figure IIb in Box 2).

A clear inference arising from these calculations is that choosing how many people to sample must be based on a detailed analysis of country-specific epidemiological data. In Uganda, assuming a power of 90%, a correlation coefficient of $(r)=0.2$, a reduction in mean intensity of

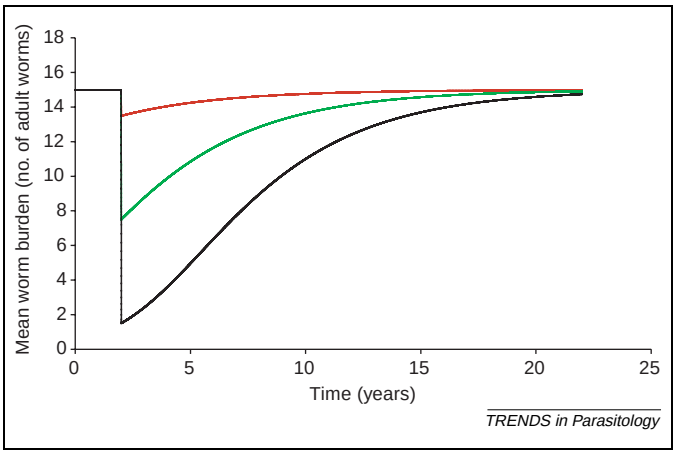

Figure 1. Recovery of mean worm burden after a single round of chemotherapy. A single round of chemotherapy depressed the average burden in the target age group (5–14 years) by 10% (green), 50% (red) and 80% (black). Predictions are based on a mathematical model of transmission (see Ref. [22]; parameter values: negative binomial, $k=0.2$, $R_0=3$; life expectancy of adult worms at five years, and density dependency parameter, $z=0.98$). Adapted from Ref. [25].

60%, a follow-up rate of 70% and an initial mean intensity of *S. mansoni* infection of 400 epg, it is calculated that the necessary sample size in each of the child age classes age class is of the order of 136. With a starting mean intensity of infection of 100 epg, the sample size rises to 208 individuals. In practice, such numbers are difficult to achieve in a single school, and it might be necessary to reduce the desired power of the evaluation to reduce sample size requirements. The trade-off between statistical robustness and logistical constraints must depend on who the evaluator is and what type of evaluation is required [7]. Evaluation of measures other than intensity of infection will require similar sample size calculations and model-based predictions of intervention impact.

Where to evaluate?

Overlaying spatially referenced, epidemiological data with environmental data remotely sensed from satellites may be valuable in selecting study areas for evaluation; for example, to identify areas of environmental homogeneity and where disease transmission is predicted to be similar (Box 3). Geo-statistical analysis of spatial patterns of infection and disease can help identify an appropriate sampling grid [35]. To date, little use has been made of modern geographical tools for guiding the planning of sampling, despite their useful application in targeting helminth control [36].

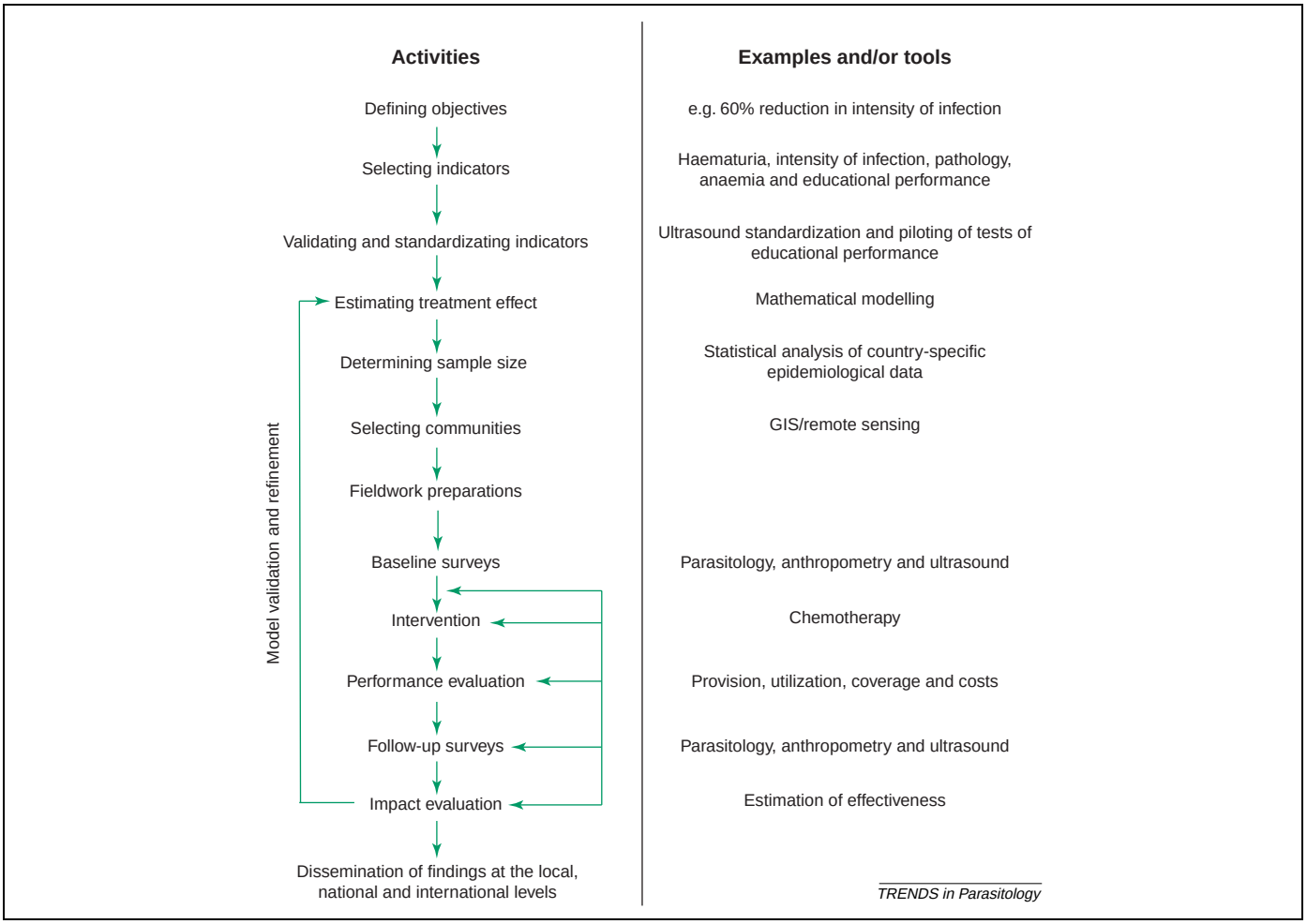

Figure 2. Proposed framework for the evaluation of national schistosomiasis control programmes. Evaluation includes both: (i) impact evaluation that assesses effectiveness of the intervention in the context of its programme-based implementation; and (ii) performance evaluation that monitors the provision, utilization and coverage of treatment [7]. The choice of evaluation indicators should reflect local appropriateness and ease of assessment. The design of the sampling design should be based on detailed analysis of country-specific epidemiological data and sample size calculations. Abbreviation: GIS, geographical information systems.

How often to evaluate?

The frequency at which evaluation should be undertaken can be determined from knowledge of (i) the transmission dynamics of infection and disease and (ii) the nature of schistosome-associated pathology and its reversal following treatment. We currently know a great deal more about transmission dynamics than reversal of pathology.

The central question in chemotherapy programmes concerns how best to achieve long-term suppression of worm burdens in terms of the fraction treated and the interval between treatments, given the constraint of limited resources for drug purchase. Various theoretical and field studies have demonstrated that targeting treatment to school-aged children is highly beneficial in terms of impact and cost [37]. More problematic is the optimal interval between treatments given rapid re-infection and bounce-back to pre-control worm burdens once treatment ceases. The rapidity of bounce-back on cessation of chemotherapy depends on the degree to which mass treatment has reduced the mean worm burden in the target community, the magnitude of R_0 in a given setting, and various other parameters that influence the transmission dynamics of the parasite such as life expectancy, degree of worm aggregation and the severity of density-dependent constraints on reproduction or survival [25]. As illustrated in Figure 1, for schistosome species with longish life expectancies in the human host (3–5 years) and low-to-moderate reproductive numbers (1–4) [23], the bounce-back time can be many years, provided one or a few rounds of treatment suppress the mean epg by 80%. Targeting treatment at a few age groups will reduce the overall impact on the community-based mean worm burdens because all infected individuals in the community contribute to the overall transmission and the magnitude of R_0. However, for schistosomes, the greatest fraction of the community worm burden typically lies within the age range 5–15 years. As such, targeting treatment at school-aged children can be very effective. Thus, from the perspective of disease transmission dynamics, the frequency of evaluation critically depends on the reduction in mean intensity of infection; if an 80% reduction in mean intensity of infection is achieved through treatment, then re-treatment and subsequent evaluation might be required only every 2–3 years.

This conclusion is supported by the limited evidence on the reversal of pathology following treatment (reviewed by Refs [12,16]). Although some forms of intestinal schistosomiasis, such hepatic fibrosis, peak later in life, children with heavy infections suffer considerable intestinal morbidity. Generally, reversal of hepatic abnormalities in children occurs 1–3 years post-treatment [16], such that follow-up evaluation should take place at the very least one year later. In areas where transmission is very intense, multiple rounds of chemotherapy might be required to resolve intestinal pathology, necessitating longer evaluation than the five-year evaluation currently envisaged by the SCI, which underscores the need for continued long-term evaluation. Reversal of urinary tract pathology can occur within 12 months, and resurgence of clinically important pathology can arise 12–18 months post-treatment, thereby highlighting important species-specific differences. Long-term follow-up is also required to detect potential treatment effects on growth and education (see Box 1).

Implications and future prospects

The epidemiological and theoretical evidence outlined above provides a framework for the evaluation of large-scale schistosomiasis control, as is being implemented through the SCI (Figure 2). This is subject to modification and revision according to the epidemiological setting and the predominant schistosome species, as well as country resources, expertise and requirements. A similar framework can usefully be developed for soil-transmitted helminthiasis, taking into account the transmission dynamics of parasites and the clinical sequelae of infection and its reversal following treatment. There is also a need to achieve a balance between implementation and evaluation, but this can be achieved by pilot intervention in one area while baseline data are being collected in another area. In this way, the careful collection of data needed for subsequent evaluation can be achieved without delaying the control plan of the countries involved. There remains a need to use the opportunity afforded by SCI to undertake rigorous evaluation of the impact of large-scale helminth control. Without such evaluation, it is unclear whether current large-scale political and financial support would continue.

Acknowledgements

The Schistosomiasis Control Initiative (SCI) is generously supported by a grant from the Bill and Melinda Gates Foundation. S.B. is supported by a Wellcome Trust Advanced Training Fellowship (73656). We thank J. Webster, J. Utzinger and the referees for their contribution to the development of this paper, as well as the SCI staff, R. Stothard, H. Thompson, C. Kamenka, L. Blair, A. Clements and A. Gabrielli for their work. We thank F. Kazibwe and E. Tukahebwa for their contribution from Uganda.

References

1 Molyneux, D.H. *et al*. (2003) Mass drug administration for lymphatic filariasis and onchocerciasis. *Trends Parasitol.* 19, 516–522
2 Richards, F.O. *et al*. (2001) Control of onchocerciasis today: status and challenges. *Trends Parasitol.* 17, 558–563
3 Fenwick, A. *et al*. (2003) Drugs for the control of parasitic diseases: current status and development in schistosomiasis. *Trends Parasitol.* 19, 509–515
4 Utzinger, J. *et al*. (2003) Sustainable schistosomiasis control – the way forward. *Lancet* 362, 1932–1934
5 Hayes, R.J. *et al*. (2000) Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Stat. Methods Med. Res.* 9, 95–116
6 Smith, P.G. and Morrow, R.H. (1996) *Field Trials of Health Interventions in Developing Countries: A Toolbox*, Macmillan Education Ltd
7 Habicht, J.P. *et al*. (1999) Evaluation designs for adequacy, plausibility and probability of public health programmes. *Int. J. Epidemiol.* 28, 10–18
8 Gyorkos, T.W. (2003) Monitoring and evaluation of large scale helminth control programmes. *Acta Trop.* 86, 275–282
9 Tanner, M. (1989) Evaluation and monitoring of schistosomiasis control. *Trop. Med. Parasitol.* 40, 207–213
10 Albonico, M. (2003) Methods to sustain drug efficacy in helminth control programmes. *Acta Trop.* 86, 233–242
11 Kirkwood, B.R. *et al*. (1997) Issues in the design and interpretation of studies to evaluate the impact of community-based interventions. *Trop. Med. Int. Health* 2, 1022–1029
12 Hatz, C.F. (2001) The use of ultrasound in schistosomiasis. *Adv. Parasitol.* 48, 225–284

13 Jukes, M.C. *et al*. (2000) Heavy schistosomiasis associated with poor short-term memory and slower reaction times in Tanzanian schoolchildren. *Trop. Med. Int. Health* 7, 104–117
14 Stephenson, L.S. (1987) *Impact of Helminth Infection in Human Nutrition*, Taylor and Francis
15 Lengeler, C. *et al*. (2002) Questionnaires for rapid screening of schistosomiasis in sub-Saharan Africa. *Bull. World Health Organ.* 80, 235–242
16 Richter, J. (2003) The impact of chemotherapy on morbidity due to schistosomiasis. *Acta Trop.* 86, 161–183
17 Booth, M. *et al*. (2004) Hepatosplenic morbidity in two neighbouring communities in Uganda with high levels of *Schistosoma mansoni* infection but very different durations of residence. *Trans. R. Soc. Trop. Med. Hyg.* 98, 125–136
18 Hoffmann, K.F. *et al*. (2002) Cytokine-mediated host responses during schistosome infections; walking the fine line between immunological control and immunopathology. *Adv. Parasitol.* 52, 265–307
19 Jordan, P. and Webbe, G. (1993) Epidemiology. In *Human Schistosomiasis* (Jordan, P. *et al*., eds), pp. 87–158, CAB International
20 Montresor, A. *et al.* (1998) *Guidelines for the Evaluation of Soil-Transmitted Helminthiasis and Schistosomiasis at the Community Level*, Schistosomiasis and Intestinal Parasite Unit, Division of Control of Tropical Diseases, World Health Organization, (WHO/CTD/SIP/98.1)
21 Hanson, K. *et al*. (2003) Cost-effectiveness of social marketing of insecticide-treated nets in the United Republic of Tanzania. *Bull. World Health Organ.* 81, 269–276
22 Korenromp, E.L. *et al*. (2003) Measurement of trends in childhood malaria mortality in Africa: an assessment of progress towards targets based on verbal autopsy. *Lancet Infect. Dis.* 3, 349–358
23 Anderson, R.M. and May, R.M. (1991) *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press
24 Anderson, R.M. and Medley, G.F. (1985) Community control of helminth infections of man by mass and selective chemotherapy. *Parasitology* 90, 629–660
25 Anderson, R.M. and May, R.M. (1985) Helminth infections of humans: mathematical models, population dynamics and control. *Adv. Parasitol.* 24, 1–101
26 Habbema, J.D.F. *et al*. (1996) The ONCHOSIM model and its use in decision support for river blindness control. In *Epidemic Models: Their Structure and Relation to Data* (Isham, V. and Medley, G.F. eds), pp. 360–380, Cambridge University Press
27 Winnen, M. *et al*. (2002) Can ivermectin mass treatments eliminate onchocerciasis in Africa? *Bull. World Health Organ.* 80, 384–390
28 Chan, M.S. *et al*. (1995) The development of an age structured model for schistosomiasis transmission dynamics and control and its validation for *Schistosoma mansoni*. *Epidemiol. Infect.* 115, 325–344
29 Chan, M.S. *et al*. (1996) Dynamic models of schistosomiasis morbidity. *Am. J. Trop. Med. Hyg.* 55, 52–62
30 Bensted-Smith, R. *et al*. (1987) Evidence for predisposition of individual patients to reinfection with *Schistosoma mansoni* after treatment. *Trans. R. Soc. Trop. Med. Hyg.* 81, 651–654
31 Tingley, G.A. *et al*. (1988) Predisposition of humans to infection with *Schistosoma mansoni*: evidence from the reinfection of individuals following chemotherapy. *Trans. R. Soc. Trop. Med. Hyg.* 82, 448–452
32 Etard, J.F. *et al*. (1990) *Schistosoma haematobium* infection in Mauritania: two years of follow-up after a targeted chemotherapy – a life-table approach of the risk of reinfection. *Parasitology* 100, 399–406
33 Quinnell, R.J. (2003) Genetics of susceptibility to human helminth infection. *Int. J. Parasitol.* 33, 1219–1231
34 Woolhouse, M.E. *et al*. (1998) Heterogeneities in schistosome transmission dynamics and control. *Parasitology* 117, 475–482
35 Srividya, A. *et al*. (2002) A geostatistical analysis of the geographic distribution of lymphatic filariasis prevalence in southern India. *Am. J. Trop. Med. Hyg.* 67, 480–489
36 Brooker, S. *et al*. (2002) Use of remote sensing and a geographical information system in a national helminth control programme in Chad. *Bull. World Health Organ.* 80, 783–789
37 Warren, K. *et al*. (1993) Helminth infections. In *Disease Control Priorities in Developing Countries* (Jamison, D. *et al*., eds), pp. 131–160, Oxford University Press
38 Doenhoff, M.J. *et al*. (2004) Specific and sensitive diagnosis of schistosome infection: can it be done with antibodies? *Trends Parasitol.* 20, 35–39
39 Engels, D. *et al*. (1996) Day-to-day egg count fluctuation in *Schistosoma mansoni* infection and its operational implications. *Am. J. Trop. Med. Hyg.* 54, 319–324
40 Lwambo, N.J.S. *et al*. (1997) The relationship between prevalence of *Schistosoma haematobium* infection and different morbidity indicators during the course of a control programme on Pemba Island. *Trans. R. Soc. Trop. Med. Hyg.* 91, 643–646
41 Gryseels, B. (1992) Morbidity due to infection with *Schistosoma mansoni*: an update. *Trop. Geogr. Med.* 44, 189–200
42 Dickson, R. *et al*. (2000) Effects of treatment for intestinal helminth infection on growth and cognitive performance in children: systematic review of randomised trials. *BMJ* 320, 1697–1701
43 Bundy, D. and Peto, R. (2000) Treatment for intestinal helminth infection. Studies of short term treatment cannot assess long term benefits of regular treatment. *BMJ* 321, 1225
44 Olsen, A. *et al*. (2000) The impact of iron supplementation on reinfection with intestinal helminths and *Schistosoma mansoni* in western Kenya. *Trans. R. Soc. Trop. Med. Hyg.* 94, 493–499
45 Taylor, M. *et al*. (2001) The effect of different anthelmintic treatment regimens combined with iron supplementation on the nutritional status of schoolchildren in KwaZulu-Natal, South Africa: a randomized controlled trial. *Trans. R. Soc. Trop. Med. Hyg.* 95, 211–216
46 Kvalsvig, J.D. (1988) The effects of parasitic infection on cognitive performance. *Parasitol. Today* 4, 206–208
47 Guyatt, H. (2000) Do intestinal nematodes affect productivity in adulthood? *Parasitol. Today* 16, 153–158
48 Peto, R. (1992) Statistics of chronic disease control. *Nature* 356, 557–558
49 Gryseels, B. and Nkulikyinka, L. (1990) The morbidity of schistosomiasis mansoni in the highland focus of Lake Cohoha (Burundi). *Trans. R. Soc. Trop. Med. Hyg.* 84, 542–547
50 Breslow, N.E. and Day, N.E. (1987) *Statistical Methods in Cancer Research. The Design and Analysis of Cohort Studies*, Vol. 2, International Agency for Research on Cancer
51 Anscombe, F.J. (1948) The transformation of poisson, binomial and negative-binomial data. *Biometrika* 35, 246–254
52 Krebs, C.J. (1999) *Ecological Methodology*, 2nd edn, Benjamin Cummings
53 Lachin, J.M. (1981) Introduction to sample size determination and power analysis for clinical trials. *Control. Clin. Trials* 2, 93–113
54 Kabatereine, N.B. *et al*. (2004) Epidemiology and geography of *Schistosoma mansoni* in Uganda: implications for planning control. *Trop. Med. Int. Health* 9, 372–380
55 Hay, S.I. *et al*. (2000) *Remote Sensing and Geographical Information Systems in Epidemiology*, Academic Press
56 Brooker, S. *et al*. (2002) Tools from ecology: useful for evaluating infection risk models? *Trends Parasitol.* 18, 70–74
57 Williams, B. and Dye, C. (1994) Maximum likelihood methods for parasitologists. *Parasitol. Today* 10, 489–493