# Research

# Estimating the number of helminthic infections in the Republic of Cameroon from data on infection prevalence in schoolchildren

Simon Brooker,[1] Christl A. Donnelly,[2] & Helen L. Guyatt[3]

**Introduction** The prevalence of infection with helminths is markedly dependent on age, yet estimates of the total number of infections are typically based on data only from school-aged children. Such estimates, although useful for advocacy, provide inadequate information for planning control programmes and for quantifying the burden of disease. Using readily available data on the prevalence of infection in schoolchildren, the relation between the prevalence of infection in school-aged children and prevalence in the wider community can be adequately described using species-specific models. This paper explores the reliability of this approach to predict the prevalence infection in the community and provides a model for estimating the total number of people infected in the Republic of Cameroon.

**Methods** Using data on the prevalence of helminthic infection in school-aged children in Cameroon, the prevalence of infection in pre-school children and adults was estimated from species-specific linear and logistic regression models developed previously. The model predictions were then used to estimate the number of people infected in each district in each age group in Cameroon.

**Results** For Cameroon, if only the prevalence of infection in schoolchildren is used, the number of people infected with each helminthic species will be overestimated by up to 32% when compared with the estimates provided by the species-specific models. The calculation of confidence intervals supports the statistical reliability of the model since a narrow range of parameter estimates is evident. Furthermore, this work suggests that estimation of national prevalence of infection and the number infected will be enhanced if data are stratified by age; this model represents a useful planning tool for obtaining more accurate estimates. Estimates based on data aggregated from three geographical levels (district, regional, and national) show that summarizing prevalence data at the national level will result in biases of up to 19%. Such biases reflect differences in the geographical distribution for the prevalence of each species.

**Discussion** Developing more accurate estimates requires a better understanding of the differences in the spatial heterogeneity of each species and also better methods of incorporating this information when making estimates.

**Keywords:** helminthiasis, epidemiology; schistosomiasis, epidemiology; ascariasis, epidemiology; trichuriasis, epidemiology; prevalence; children; linear models; logistic models; regression analysis; Republic of Cameroon.

*Bulletin of the World Health Organization*, 2000, **79**: 1456–1465.

## Introduction

When planning interventions to control a disease, it is essential to understand the burden of the disease. In the absence of reliable data on the total number of helminthic infections in a country, estimates have often been based on prevalence data from a few limited studies which have then been extrapolated to the country as a whole (*1–5*). Although these estimates may provide an indication of the scale of the problem, they do not provide data on the age distribution of the infection. Yet because of the marked association between age and prevalence in many diseases (Fig. 1) information on patterns of distribution by age is essential in planning control measures within countries and for quantifying the burden of disease (*5*). One reason that age distribution has not been incorporated in previous estimates is because most data on prevalence have come primarily from school-aged children (*6*).

[1] Research Fellow, Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3FY, England (email: simon.brooker@ceid.ox.ac.uk). Correspondence should be addressed to this author.

[2] Head of Statistics Unit and Reader in Statistical Epidemiology, Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, England.

[3] Research Fellow, Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, England.

Ref. No. **99-0079**

We investigated the relation between the prevalence of helminthic infection in the community and the prevalence among schoolchildren in studies across Africa (7) and showed that at all levels of infection prevalence in schoolchildren is typically higher than prevalence in the entire community, with the absolute difference depending on the species. These observations suggest that if used in isolation prevalence data from schoolchildren will tend to produce an overestimate of the number of people infected when compared with estimates based on prevalence data for the community. However, prevalence among the whole community can be predicted from models that are species specific.

The geographical distribution of infection is also important in providing reliable estimates of the number of people infected. Current procedures either ignore spatial distribution or assume a normal distribution (4, 8–10). In practice there are few studies that have explored this issue, despite spatial heterogeneity in prevalence having the potential to introduce further bias into any estimate.

This paper builds on recent work (7) and provides a practical means for extrapolating from prevalence data on school-aged children to the prevalence of infection in the total community and thus the total number infected for each of the major species of helminth. This approach uses the example of the Republic of Cameroon because nationwide survey data exist (11–13), and investigates the differences in estimates of the total number of infections based on prevalence data for school-aged children alone or extrapolated to the whole community. Because of the inherent potential for error in the use of any extrapolation procedure, this paper also assesses the reliability of model-based estimates to shed light on whether further investigations into these relations is required. Finally, investigations are carried out to explore how estimates alter in relation to spatial scale by aggregating infection data for three geographical levels (district, regional, and national).

## Methods

### Overview

Using data on the prevalence of helminthic infection in schoolchildren in Cameroon, the prevalence of infection in pre-school children and adults was estimated from species-specific logistic regression models and linear regression models that were developed earlier (7). Estimates and confidence intervals are provided in order to assess the reliability of the predictions. Both the best-fit logistic regression models and the linear regression models are used to estimate the prevalence of infection, to examine the differences arising from the two approaches and to determine whether a single, consistent approach could be used. The predictions from the linear regression model are then used to estimate the number of people infected in each district in other age groups (pre-school and adult populations); these are then summed



Fig. 1. **Prevalence of different helminths in infected people of different ages in the Republic of Cameroon.** Data adapted from *22–25*
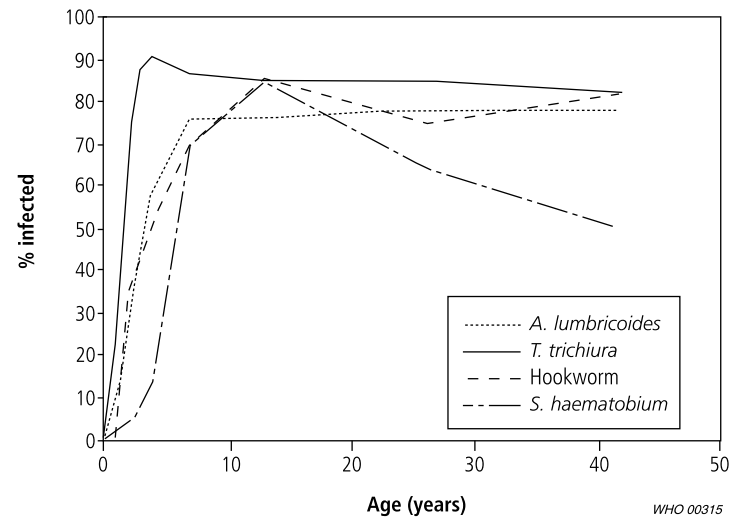
to estimate the total number of people in the country who are infected. The improvement in large-scale estimates of the number infected that is gained by using such an approach is investigated by comparing estimates based only on crude infection, data for schoolchildren with estimates based on modelled infection data for the total population. Finally, the implication of using data aggregated at different geographical scales is investigated.

### Sources of empirical data

The data are taken from prevalence estimates for Cameroon found in a nationwide survey of schistosome infection (with *Schistosoma mansoni* and *S. haematobium*) and intestinal nematode infection (with *Ascaris lumbricoides*, *Trichuris trichiura*, and hookworm) (11–13). The survey was conducted during 1985 and 1987 and covered each of the 52 districts in Cameroon. It also included a representative sample of 22 166 children aged 10–19 years in 512 schools. This survey represents the most comprehensive survey of both schistosomes and intestinal nematodes in sub-Saharan Africa. District-level prevalences were calculated by dividing the number of children infected in the district by the number of children examined.

Data on the population of the district came from a 1990 national population forecast based on the 1990 national census (14); they were projected to 1999 assuming a national inter-census growth rate of 2.8%, which was obtained from the United States Census Bureau (15). The administrative boundaries of Cameroon were obtained from the African population database (14). Geographical data are displayed using ArcView (version 3.0, Environmental Systems Research Institute, Inc., Redlands, CA, 1996).

### The models

The methods used for estimating the prevalence of infection in the general community have already been

described in detail (7). In summary, linear models and three variants on logistic regression models are fitted using a maximum likelihood approach and data on the number who are infected out of the number sampled. These data for pre-school children, school-aged children, and adults are abstracted from studies in sub-Saharan Africa. (Further details of the regression models are given in the Appendix which is available on the web at http://www.who.int/bulletin/tableofcontents/2000/issue12/appendix).

In fitting the different models, data points at which the prevalence of infection in schoolchildren was ≤ 5% were excluded. This modification was not used in the previous study but was incorporated here since in some instances the prevalence among pre-school children and adults was found to be greater than zero while the prevalence in schoolchildren was zero. This is considered to be unrealistic for intuitive reasons, and it was therefore assumed that at low prevalences (<5%) the predicted prevalence of infection in the community is equal to that in schoolchildren. The consequence is that estimates for *S. mansoni* and hookworm differed by 3–5% compared with estimates derived from models without this modification. Prevalences for *S. haematobium*, *A. lumbricoides*, and *T. trichiura* differed by less than 2%.

The resulting age-prevalence relations were then combined to predict the relation between prevalence of infection among schoolchildren and in the total population (Fig. 2). The lines show the regression models for the maximum likelihood fit for each species with the precise functional form varying according to the species of parasite. The model equations and maximum likelihood parameter esti-

mates are given in the Appendix available on the web at http://www.who.int/bulletin/tableofcontents/2000/issue12/appendix. The study sites and references for all data sources are published (7). The predictions are standardized by age (by weighting the prevalence in each age group by the proportion of infections expected in each age class given in a Coale and Demeny "West" model life table (female life expectancy of 50 years and an annual growth rate of 3%). In order to investigate the differences in estimates of prevalence that arise when different regression models are used, estimates are provided using both the best fit logistic regression model and the linear regression model.

Confidence intervals for the model's parameters (Tables A1 and A2 available on the web at http://www.who.int/bulletin/tableofcontents/2000/issue12/appendix) were obtained using maximum likelihood methods. Confidence intervals for national prevalence and the number of individuals infected were obtained from the variance of the parameter estimates and the variance in the prevalence estimate among schoolchildren, allowing for correlations between the estimated prevalence in the different age groups. The estimated variance of the model's parameter estimates was then used to calculate confidence intervals for the estimates of the number infected.

## Estimating the total number infected within the country

The models described above and detailed in the appendix provided estimates of the proportion of pre-school children and adults infected in the

Fig. 2. **Predicted prevalence of helminthic infection among school-aged children and prevalence among the community in the Republic of Cameroon.** a) Linear regression; b) best fit logistic regression model
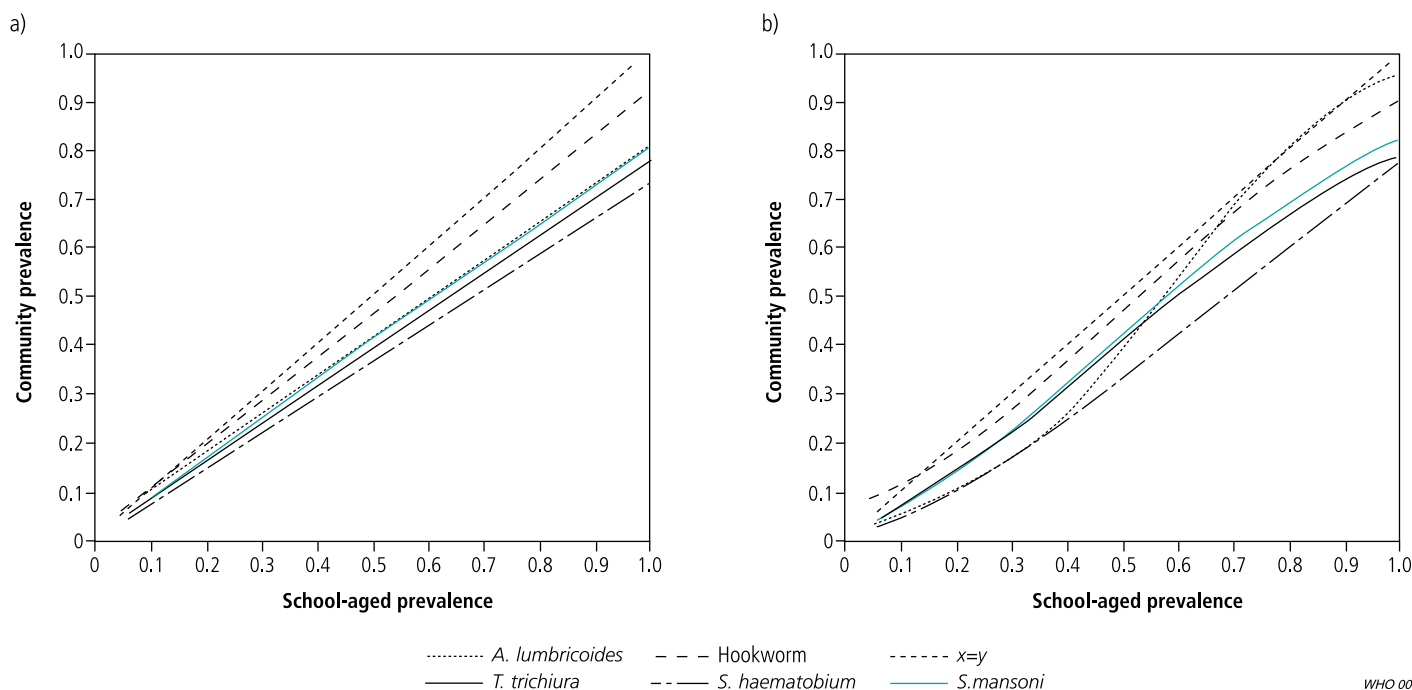
country. The number infected in each age group in each district was then estimated by multiplying the prevalence estimates by the age-specific population size using information on the age structure of the population taken from the 1987 census (*15*). These age-specific estimates were summed to provide total estimates at the district level; these in turn were summed to provide estimates for the country as a whole. Since no large-scale programmes have been undertaken since the time of the survey, we assumed that the prevalence remained stable.
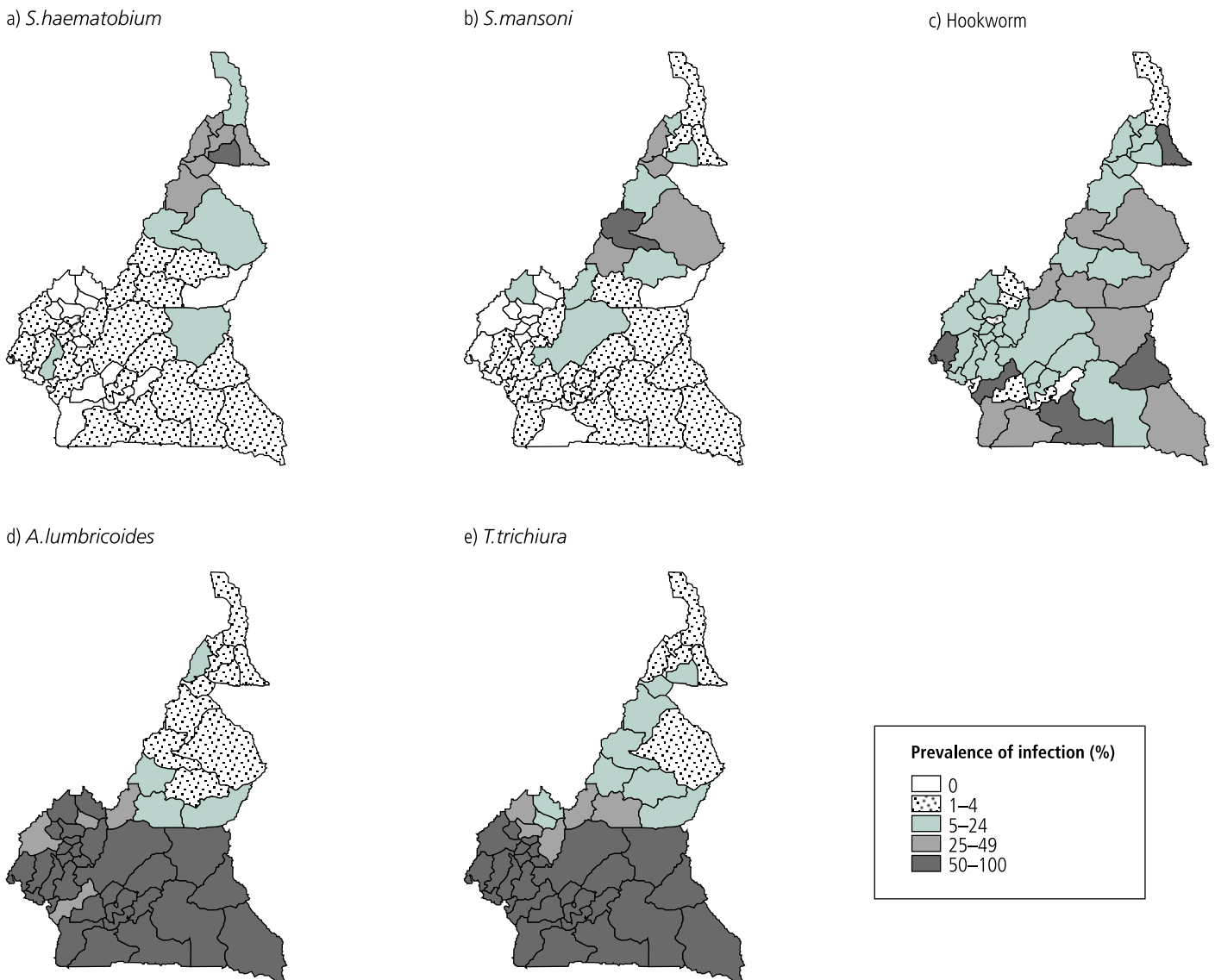
We provide estimates of the number of people infected in Cameroon using data for schoolchildren; additional estimates are based on the modelled, age-stratified data. We also provide estimates based on infection data aggregated at three geographical levels: district, regional and national.

# Results

## Distribution of prevalence in schoolchildren by district

The distribution of the prevalence of helminthic infection in schoolchildren in Cameroon by species and district is shown in Fig. 3. For each species, the prevalence of infection varies considerably across the country. While the highest levels of infection prevalence with *S. haematobium* and *S. mansoni* are found in the extreme north and in the northern regions, the distribution is patchy elsewhere (Fig. 3). There is no clear pattern to the variation in the prevalence of hookworm infection (Fig. 3): in one district there are no infections, two districts have prevalences >50%, but most districts have moderate prevalences (between 1–49%). *A. lumbricoides* and *T. trichiura* infections are present in all districts in the

Fig.3. **Distribution by district of prevalence of infection with helminths in the Republic of Cameroon using data from school-aged children**



a) *S.haematobium*

b) *S.mansoni*

c) Hookworm

d) *A.lumbricoides*

e) *T.trichiura*

Prevalence of infection (%)

- 0
- 1–4
- 5–24
- 25–49
- 50–100

Source: *11–13*.

WHO 00317

country but there is a distinct geographical gradient, ranging from 75–100% in the equatorial south of the country to 1–9% in the north (Fig. 3).

### Estimates of infection in the total population

The estimates of the prevalence of infection across the country are shown in Table 1. For each species these estimates are derived by calculating the number sampled and the number infected in each district. These estimates are then summed for the whole country. Three estimates are presented: these are based on data for schoolchildren and the modelled, age-stratified data for the total population using both the best-fit logistic regression model and the linear regression model. Despite the logistic regression model giving a better statistical fit to the empirical data, the differences in the maximum likelihood estimates were marginal (see Appendix at http://www.who.int/bulletin/tableofcontents/2000/issue12/appendix); the estimates of national prevalence using each approach differed only slightly (Table 1). Therefore, although the linear regression model may result in some variability in the estimate of prevalence, its use can be justified because of the simplicity of the model and the consistency in the approach. For this reason, the following analyses use simple linear regression rather than the complex and varied functional forms of the logistic models.

The estimates of prevalence made using the linear regression model are shown with the observed prevalence in schoolchildren (Table 1). In the case of hookworm, the difference in these estimates is negligible. For the other species of helminth, the model-based estimates are lower than those derived using only data from infections among schoolchildren. The difference in estimates tends to vary for each species, being greatest for *A. lumbricoides* and *T. trichiura*. In part, this reflects observed species-specific and age-dependent prevalences incorporated into the models (Fig. 2).

### Estimates of the total number infected

The modelled population size for Cameroon is 14 849 714, which was distributed unevenly across the country (Fig. 4). The estimates of the total number of people infected are shown in Table 2. For each species, six estimates are given reflecting the different procedures used and the different levels of geographical stratification. The estimated number infected is substantially higher when it is based only on the observed prevalence of infection among schoolchildren, with the exception of hookworm in which case the estimate is 0.3% lower. The degree of bias varies for each species, being greatest for *S. haematobium* (32% higher). In part, this reflects the greater difference between prevalence among schoolchildren and prevalence in the entire community (Fig. 2). Using the prevalence of infection among schoolchildren, the estimates for infection with *S. mansoni*, *A. lumbricoides*, and *T. trichiura* were up to 18%, 20%, and 27% higher respectively than those based on the modelled, age-stratified data.

Table 2 also shows that the estimates of the number of people infected at the regional and national level are either overestimates or underestimates of those at the district level. This difference is as much as 19% for infections with hookworm and 14% for *S. mansoni* infections. This suggests that the use of the region or country as a unit of aggregation is inappropriate for estimating the number infected with hookworm and *S. mansoni* within Cameroon. The total number of infections with other species calculated using district-level prevalences are not very different from those based on data from the regional or national level. Thus, the appropriate geographical scale for estimation differs according to the species of the parasite.

The geographical distribution of the number infected as calculated from the predicted prevalence of infection in the entire community is shown in Fig. 5. The concentration of schistosomiasis is greater in the north of the country, where the prevalence of infection is highest (Fig. 3). The greatest number of people infected with *A. lumbricoides* and *T. trichiura* is concentrated in the centre of the country and in certain districts in the south (Fig. 5), although the prevalence was highest in the south of the country (Fig. 3). This has occurred, at least in part, because of differences in the distribution of the population across the country (Fig. 4), and it reflects the importance of the differential distribution of infection and population.

## Discussion

### Model applications

Obtaining accurate estimates of the number of people infected with helminths within any given country has been difficult due to the paucity of

Table 1. **Estimates of the national prevalence of helminthic infection in the Republic of Cameroon obtained using data from schoolchildren alone and results obtained using modelled, age-stratified data[a, b]**

| Species | Estimated national prevalence (%) | | |
| --- | --- | --- | --- |
| | Data from schoolchildren only | Linear model community prediction | Best fit community prediction |
| *Schistosoma haematobium* | 9.5 (9.1–9.8)[c] | 7.2 (6.9–7.4) | 6.1 (5.8–6.4) |
| *Schistosoma mansoni* | 5.7 (5.4–6.0) | 4.8 (4.5–5.1) | 4.5 (4.2–4.8) |
| *Ascaris lumbricoides* | 42.6 (41.7–43.4) | 35.5 (34.8–36.2) | 37.7 (37.0–38.4) |
| *Trichuris trichiura* | 53.4 (52.7–54.1) | 42.1 (41.3–42.8) | 43.6 (42.9–44.3) |
| Hookworm | 14.7 (14.2–15.2) | 14.7 (14.3–15.2) | 14.4 (13.9–14.9) |

[a] Modelled, age-stratified data for the community were aggregated at three geographical levels.

[b] Best-fit logistic models and linear models were used.

[c] Values in parentheses are 95% confidence intervals.

detailed prevalence data stratified by age. In an earlier study we developed species-specific models to help predict infection prevalence in a community using data on the prevalence of infection in school-aged children (7). This paper explores the reliability of this approach and its application in estimating the number of people infected in Cameroon. The omission of age-stratification introduced bias of up to 32% in the estimates (Table 2). This discovery has consequences for the accuracy of estimates of disease burden caused by helminthic infection that are measured using disability-adjusted life years (DALYs) (5, 10).

In terms of using the model specifically for Cameroon, it is useful to contrast the present estimates of the number infected with previous estimates. Earlier work has been limited to estimating the number of people infected with schistosomes. The present estimates suggest that a total of 1 779 000 (95% confidence intervals (CI) = 1 750 000–1 835 000) people may be infected with either *S. haematobium* or *S. mansoni* (not allowing for infections with more than one species). This is approximately half of previous estimates (3), which suggested that 3 935 000 people are infected with both species (here adjusted to 1999 population levels). This probably reflects differences in the data used. While the present estimates incorporated age stratification and, to some degree, geographical stratification, Utroska et al. (3) used a crude national prevalence of 26.5%; and much of that data came from surveys of schoolchildren typically conducted in restricted foci of the country where prevalence was relatively high (16). The very large difference in these two sets of estimates highlights the need for collecting age-specific prevalence data from a wide range of environments at the district level (17). In the absence of such detailed data, these models may

Fig. 4. **Distribution of the population in the Republic of Cameroon by district**



Population
☐ < 100 000
▦ 100 000 – 200 000
▨ 200 000 – 500 000
▩ > 500 000
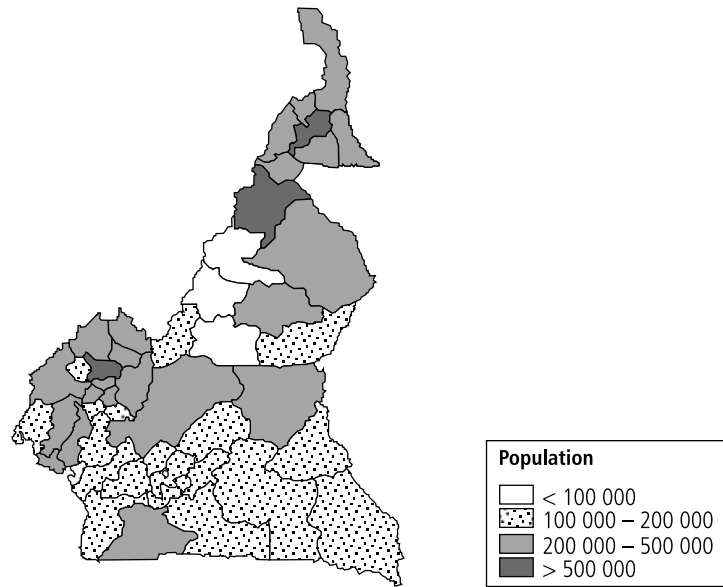
Source: *14.*

WHO 00318

provide a potentially useful means for extrapolating to predict the age-specific prevalence of infection.

Extrapolating from the prevalence of infection among schoolchildren to the prevalence in the wider community has also been done by Ratard et al. (8) who used the same prevalence data as used here and estimated that there were 394 000 cases of infection with *S. haematobium* and 420 000 cases of infection with *S. mansoni*. Adjusting these estimates to 1999 population levels suggests that 579 000 (95% CI = 301 000–671 000) people would be infected with *S. haematobium* and 609 000 (95% CI = 347 000–

Table 2. **Comparison of estimates of the total number of people infected with helminths obtained using data from schoolchildren alone and estimates obtained using modelled, age-stratified data**[a]
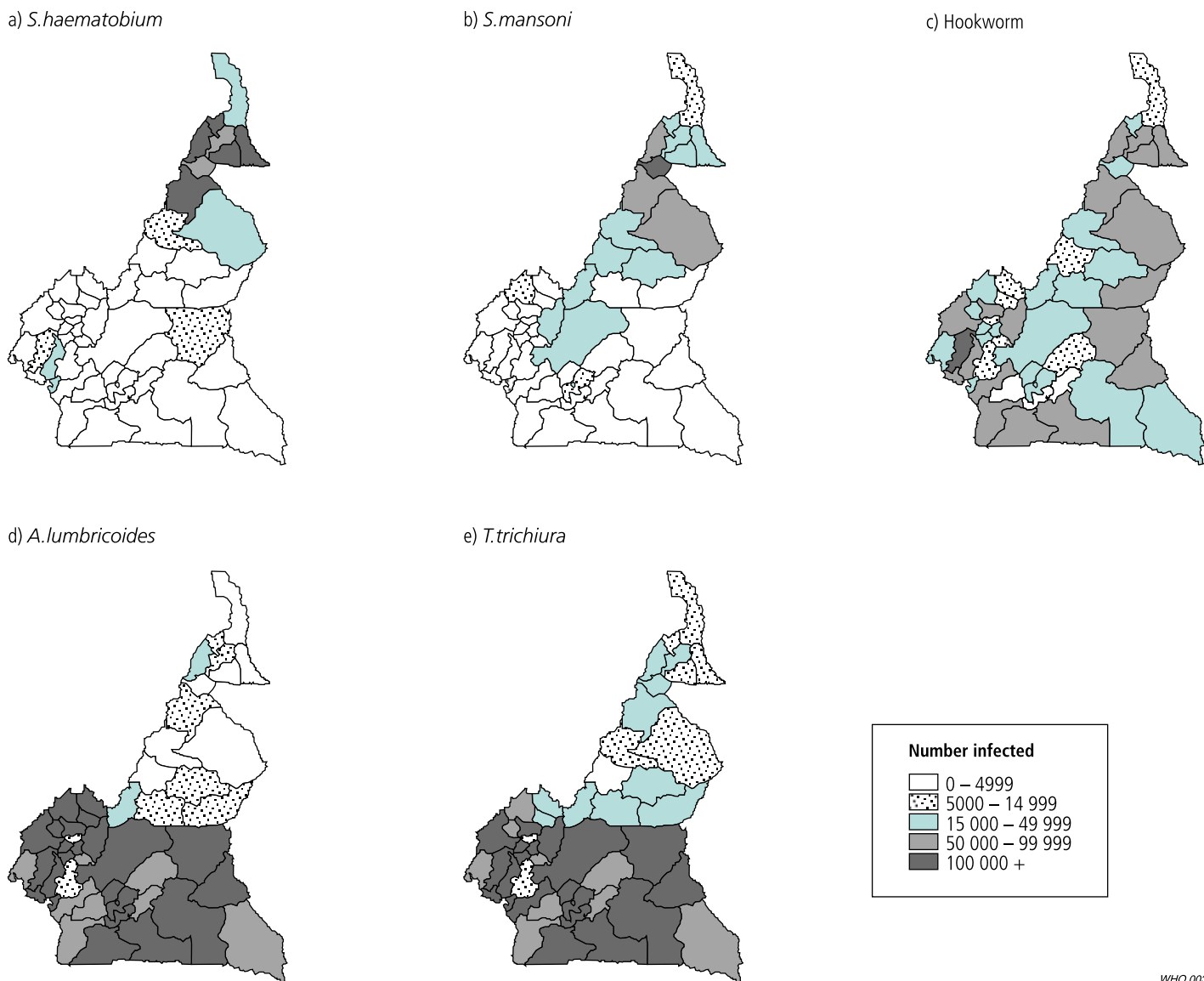
| Species | Level of aggregation | | | | | |
|---|---|---|---|---|---|---|
| | District[b] | | Regional | | National | |
| | School-children | Community prediction | School-children | Community prediction | School-children | Community prediction |
| *Schistosoma haematobium* | 1408 (1357–1460)[c] | 1066 (1026–1106) | 1530 (1466–1593) | 1081 (1036–1126) | 1612 (1547–1676) | 1129 (1082–1176) |
| *Schistosoma mansoni* | 845 (801–890) | 713 (674–753) | 778 (740–816) | 614 (581–647) | 876 (830–922) | 664 (624–704) |
| *Ascaris lumbricoides* | 6321 (6198–6443) | 5273 (5164–5383) | 6677 (6588–6767) | 5459 (5303–5615) | 6769 (6672–6865) | 5639 (5247–6032) |
| *Trichuris trichiura* | 7932 (7825–8040) | 6248 (6138–6359) | 8315 (8239–8390) | 6159 (5985–6332) | 8716 (8620–8811) | 6457 (6032–6882) |
| Hookworm | 2183 (2111–2255) | 2190 (2117–2263) | 2575 (2499–2650) | 2469 (2370–2568) | 2706 (2631–2781) | 2615 (2363–2866) |

[a] Modelled, age-stratified data for the community were aggregated at three geographical levels.

[b] Derived by summing district-level estimates derived from prevalence at district level and calculated using projected 1999 population levels (*14*).

[c] Values in parentheses are 95% confidence intervals.

Fig.5. **Estimates of the number of people infected with different helminths in the Republic of Cameroon by district**

a) *S.haematobium*

b) *S.mansoni*

c) Hookworm



d) *A.lumbricoides*

e) *T.trichiura*



**Number infected**

☐ 0 – 4999
▦ 5000 – 14 999
▨ 15 000 – 49 999
▨ 50 000 – 99 999
■ 100 000 +

*WHO 00319*

879 000) with *S. mansoni*. The estimates in this paper imply that 1 066 000 people (95% CI = 1 026 000– 1 106 000) are infected with *S. haematobium* and 713 000 (95% CI = 674 000–753 000) with *S. mansoni*. Ratard et al. multiplied the prevalence of infection in schoolchildren by a species-specific constant ratio of the prevalence in the total population and the prevalence in school-aged children (derived from selected age-stratified studies in Cameroon) to predict the prevalence in the total population; their results included very wide confidence intervals. More recently, Traore et al. (*9*) have developed an index based on linear regression to estimate the prevalence of infection in other age groups based on the prevalence in children aged 7–14 years; they do not incorporate any estimate of error.

Building on these previous studies of schisto-somiasis, this study explores a more reliable method and extends the approach to the five major species of

helminth. In addition, we use models developed from empirical data collected across a wide range of countries in Africa (*7*), whereas previous extrapola-tions have been based on data for both the school-aged population and the total community within a single country. The consistency of the results suggests that there is a common model for Africa, which can be applied to the many parts of Africa for which detailed, country-specific data on infection in different age groups may not be readily available.

**Differences for statistical models**

The slight differences in estimates obtained from the various statistical models show that the method is not sensitive to the precise regression model used. However, there are several reasons why this method may not truly estimate the prevalence of infection in the community and thus the total number of infected people in the country. Firstly, there are no direct

estimates of the number of people within Cameroon who may be infected so the present estimates cannot be validated against them. Secondly, although the applicability of this approach to estimating prevalence in the community and estimating the number infected is statistically reliable (see Table 2 for confidence intervals) and is based on empirical data, there remains a need for further evaluation of the approach in other countries to determine if this model will be useful. Thirdly, it is unclear whether the relation between age and prevalence and subsequent models would differ for other parts of the world (for example, in Asia) or whether a universal model is possible. Fourthly, the approach adopted here has assumed an equivalence of prevalence at low levels of infection (<5%). Although undertaking a sensitivity analysis in the final estimates of the number infected assessed the implications of such assumptions, and revealed little difference in the estimates, there remains uncertainty until further data are available. Finally, this approach does not control for potential biases in individual studies, such as the diagnostic method or sampling method used.

One area that warrants particular attention is the effect of differences in prevalence across different geographical areas on the method of extrapolation. Here we show that the spatial resolution incorporated in the estimation procedure is critical in determining the number infected. This is an observation that has been made previously for estimates of the potential morbidity attributable to intestinal nematodes (5, 10) and points to the need to obtain more accurate estimates of differences in geographical distribution in helminthic infections. In this study we used the district as the unit of analysis because the data were available and because approaches to controlling infections are increasingly implemented at the district level. In doing so, however, we have made assumptions about the geographical heterogeneity of prevalence within districts. Although the original sampling plan was designed to ensure that the communities sampled within each district were as representative as possible, it remains unclear whether district units are appropriate units for estimation since these units may give a false impression of the spatial heterogeneity of infection within districts. Indeed, the average difference between the minimum prevalence and the maximum prevalence within districts varied 20–33% for different species of parasites (R. Ratard, personal communication, 1999). To resolve this issue more detailed data at a finer spatial resolution are required to incorporate more precisely within the

estimation procedure the effect of geographical heterogeneity in prevalence. Maps of the distribution of the prevalence of helminthic infection are being developed for malaria (18) and lymphatic filariasis (19). Ongoing work, in collaboration with the World Health Organization, is being undertaken to develop an atlas of the geographical distribution of species-specific helminthic infection at the district level (6).

A final interesting observation is that the degree of underestimation or overestimation of the number of people infected using different levels of the geographical scale varied for each species, being greatest for the hookworm and *S. mansoni* species. In part, these differences between the species may reflect variability in the focal nature of each species and indicate the need for further investigation into the inherent differences in the prevalence of each species across geographical areas of each species, and how this might vary with the environment.

## Future work
Within its limitations, this study shows the importance of incorporating age stratification for each major species of helminth when estimating the number of people at risk of infection, and it also explores the reliability of a method to estimate the prevalence of infection in the community from readily available data on schoolchildren. Further refinement of the approach would incorporate a more detailed understanding of the effects of differences in geographical distribution on prevalence and differences between species within the estimation procedure. It might be possible to develop high-resolution modelled maps of the distribution of prevalence across districts using relations with environmental variables (20). This approach is being developed for malaria (21) and is an area of helminthic epidemiology which clearly warrants further work. Such studies, together with models using age stratification, as developed here, will provide the basis for a wider and more detailed approach to predicting the prevalence of helminthic infection, defining the burden of disease, and planning and evaluating control efforts. ∎

## Résumé

### Cameroun : estimation du nombre d'helminthiases à partir de la prévalence de l'infestation chez les enfants d'âge scolaire

Il est essentiel de disposer de données sur le nombre de personnes infestées par une espèce d'helminthe particulière si l'on veut planifier les efforts de lutte. La prévalence des helminthiases est nettement fonction de l'âge ; or les estimations relatives au nombre total d'infestations reposent habituellement sur des données

relatives aux seuls enfants d'âge scolaire. Ces estimations, si elles sont utiles pour promouvoir la lutte, fournissent en revanche moins d'informations pour ce qui est de planifier les programmes de lutte et de quantifier la charge de morbidité. A partir des données relatives à la prévalence de l'infestation chez les enfants d'âge scolaire, on peut très bien décrire la relation qui existe entre la prévalence de l'infestation chez ces enfants et celle qui prévaut dans la communauté en général, au moyen de modèles spécifiques d'espèces. La méthode repose sur les rapports qu'on a observés entre les diverses prévalences de l'infestation chez les enfants d'âge préscolaire, chez les enfants d'âge scolaire et chez les adultes. Les données stratifiées sur l'âge provenant de l'ensemble de l'Afrique ont été ajustées au moyen de modèles linéaires et de trois modèles d'analyse de régression logistique. Dans cet article, on analyse la fiabilité de cette approche et l'on propose un modèle permettant d'estimer le nombre total de personnes infestées au Cameroun. Le modèle de régression logistique a permis d'obtenir un meilleur ajustement statistique des données empiriques que le modèle de régression linéaire mais seules de légères différences ont été observées dans les estimations de la prévalence nationale obtenues par chacune de ces méthodes. C'est pourquoi on a utilisé le modèle de régression linéaire. Les estimations relatives au nombre de personnes infestées fondées uniquement sur les données de la prévalence chez des enfants d'âge scolaire étaient supérieures aux estimations modélisées, stratifiées sur l'âge. Ces dernières ont intégré les données relatives à la variation des degrés d'infestation en fonction de l'âge. Les différences d'estimations selon les modèles n'étaient pas les mêmes pour chaque espèce d'helminthes. Dans le modèle ne faisant appel qu'aux données relatives aux enfants d'âge scolaire, l'estimation de l'infestation par l'ankylostome obtenue est inférieure de 0,3 %, mais elle peut être majorée d'une valeur pouvant atteindre 32 % pour d'autres espèces. Cette estimation est majorée de 18 % pour *S. mansoni*, de 20 % pour *Ascaris lumbricoides,* de 27 % pour *Trichuris trichiura* et de 32 % pour *Schistosoma haematobium*, entraînant donc une surestimation du nombre de personnes infestées par chaque espèce. Le calcul des intervalles de confiance montre que notre méthode statistique est fiable, puisqu'il est évident que l'éventail des estimations des paramètres est étroit. Ces observations laissent à penser qu'on parviendra à une meilleure estimation de la prévalence du nombre de personnes infestées si l'on stratifie les données sur l'âge, et que ce modèle représente un outil de planification utile permettant d'obtenir des estimations plus précises. Les estimations réalisées à partir des données globales recueillies à trois niveaux géographiques (district, région et pays) montrent que le fait de récapituler les données de la prévalence à l'échelle nationale entraîne des erreurs de 19 % pour l'ankylostomiase et de 14 % pour l'infestation à *S. mansoni* ; cette erreur est <7 % pour chacune des espèces de nématodes intestinaux. Cela reflète probablement des différences dans la répartition géographique de chaque espèce. Ce résultat indique que, pour parvenir à des estimations plus précises, il nous faut mieux comprendre l'hétérogénéité géographique de chaque espèce et disposer de méthodes améliorées permettant de tenir compte de cette information lorsqu'on procède aux estimations.

## Resumen

### Estimación del número de infecciones por helmintos en la República del Camerún a partir de la prevalencia de la infección entre escolares

La disponibilidad de información sobre el número de personas infectadas por una determinada especie de helminto es un requisito crucial para planificar las actividades de control. La prevalencia de las helmintiasis depende mucho de la edad, pese a lo cual las estimaciones del número total de infecciones se basan normalmente en datos referentes sólo a niños en edad escolar. Esas estimaciones, aunque útiles a efectos de sensibilización, proporcionan menos información para planificar los programas de lucha y cuantificar la carga de morbilidad. A partir de los datos disponibles sobre la prevalencia de las helmintiasis en los escolares, es posible describir adecuadamente la prevalencia de la infección en los niños en edad escolar y en el conjunto de la comunidad mediante modelos específicos para cada especie. El método se basa en la relación observada entre la prevalencia de la infección entre los niños de edad preescolar, los escolares y los adultos. Se utilizaron modelos lineales y tres variantes de regresión logística para ajustar datos estratificados por edades procedentes de toda África. En este artículo se analiza la fiabilidad de ese enfoque y se facilita un modelo para estimar el número total de personas infectadas en la República del Camerún. El modelo de regresión logística permitió ajustar los datos empíricos mejor que el de regresión lineal, pero sólo se observaron ligeras diferencias entre las estimaciones de la prevalencia nacional obtenidas con cada método. Por ese motivo, se decidió emplear el método de regresión lineal. Las estimaciones del número de personas infectadas basadas únicamente en los datos sobre la prevalencia entre los escolares fueron mayores que las estimaciones obtenidas con el modelo de datos estratificados por edades. Estas últimas incorporaban datos sobre la dependencia de la edad de los niveles de infección. El grado de diferencia entre las estimaciones arrojadas por los modelos era distinto para cada especie de helminto. La estimación de la infección por anquilostoma fue un 0,3% inferior con el modelo basado exclusivamente en los datos sobre escolares, pero fue superior, hasta en un 32%, con otras especies. Concretamente, fue un 18% mayor con *S. mansoni*, un 20% mayor con *Ascaris lumbricoides*, un 27% mayor con *Trichuris trichiura*, y un 32% mayor con *Schistosoma haematobium*, con la consiguiente sobreestimación del número de personas infectadas por cada una de esas especies. Los intervalos de confianza calculados de-

muestran la fiabilidad estadística de nuestro enfoque, pues definen un margen claramente estrecho para las estimaciones de los parámetros. Estas observaciones llevan a pensar que las estimaciones de las prevalencias nacionales y del número de personas infectadas se verán facilitadas si los datos se estratifican por edades, y que este modelo constituye un valioso instrumento de planificación pues aporta estimaciones más precisas. Las estimaciones efectuadas a partir de los datos agregados a tres niveles geográficos (distrital, regional y nacional) muestran que la síntesis de los datos de

prevalencia a nivel nacional adolecen de sesgos del 19% para la infección por anquilostoma y del 14% para *S. mansoni*; el sesgo fue inferior al 7% en todas las especies de nematodo intestinal. Ello se debe probablemente a la existencia de diferencias en la distribución geográfica de cada especie. Este resultado indica que para conseguir estimaciones más precisas es necesario comprender mejor las diferencias en la heterogeneidad geográfica de cada especie, y emplear métodos más idóneos para incorporar esa información al realizar las estimaciones.

## References

1. **Warren KS et al.** Helminth infections. In: Jamison DT et al, eds. *Disease control priorities in developing countries.* Oxford, Oxford University Press, 1993: 131–160.

2. **Crompton DWT, Tulley JJ.** How much ascariasis is there in Africa? *Parasitology Today,* 1987, **3**: 123–127.

3. **Utroska J et al.** *An estimate of global needs for praziquantel within schistosomiasis control programmes.* Geneva, World Health Organization, 1989 (unpublished document WHO/SCHISTO/89.102).

4. **Bundy DAP et al.** The epidemiological implications of a multiple-infection approach to the control of human helminth infections. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 1991, **85**: 274–276.

5. **Bundy DAP et al.** Intestinal nematode infections. In: Murray CJL., Lopez, AD. *The global epidemiology of infectious diseases.* Harvard University Press, Cambridge, MA, 2000.

6. **Brooker S et al.** Towards an atlas of human helminth infection in sub-Saharan Africa: the use of geographical information systems (GIS). *Parasitology Today,* 2000 **16**: 303–307.

7. **Guyatt HL et al.** Can prevalence of infection in school-aged children be used as an index for assessing community prevalence? *Parasitology,* 1999, **118**: 257–268.

8. **Ratard RC et al.** Estimation of the numbers of cases of schistosomiasis in a country: the example of Cameroon. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 1992, **86**: 274–276.

9. **Traore M et al.** *Schistosoma haematobium* in Mali: prevalence rate in school-age children as an index of endemicity in the community. *Tropical Medicine and International Health*, 1998, **3**: 214–221.

10. **Chan M-S et al.** The evaluation of potential global morbidity attributable to intestinal nematode infections. *Parasitology,* 1994, **109**: 373–387.

11. **Ratard RC et al.** Human schistosomiasis in Cameroon. 1. Distribution of schistosomiasis. *American Journal of Tropical Medicine and Hygiene,* 1990, **6**: 561–572.

12. **Ratard RC et al.** Distribution of hookworm infection in Cameroon. *Annals of Tropical Medicine and Parasitology,* 1992, **86**: 413–418.

13. **Ratard RC et al.** Ascariasis and trichuriasis in Cameroon. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 1991, **85**: 84–88.

14. **Deichmann U.** African population database documentation. United Nations Environment Programme Global Resource Information Database website (UNEP-GRID) (available on the Internet at http://grid2.cr.usgs.gov/globalpop/africa/index.php3; accessed 1996).

15. **US Census Bureau.** International Data Base (available on the Internet at http://www.census.gov/ipc/www/idbnew.html).

16. **Doumenge JP et al.** *Atlas of the global distribution of schistosomiasis.* Bordeaux, Presses Universitaires de Bordeaux, 1987.

17. **Montresor A et al.** *Guidelines for the evaluation of soil-transmitted helminthiasis and schistosomiasis at the community level.* Geneva, Schistosomiasis and Intestinal Parasite Unit, Division of Control of Tropical Diseases, World Health Organization, 1998 (document WHO/CTD/SIP/98.1).

18. **Snow RW et al.** The need for maps of the transmission of transmission intensity to guide malaria control in Africa. *Parasitology Today,* 1996, **12**: 455–457.

19. **Michael E, Bundy DAP.** Global mapping of lymphatic filariasis. *Parasitology Today,* 1997, **13**: 472–476.

20. **Brooker S, Michael M.** The potential of geographical information systems and remote sensing in the epidemiology and control of human helminth infections. *Advances in Parasitology,* 2000, **47**: 245–288.

21. **Snow RW et al.** A preliminary continental risk map for malaria mortality among African children. *Parasitology Today,* 1999, **15**: 99–104.

22. **Bradley DJ, McCullough FS.** Egg output stability and the epidemiology of *Schistosoma haematobium.* Part II. An analysis of the epidemiology of endemic *S. haematobium. Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1973, **67**: 491–500.

23. **Bundy DAP et al.** Epidemiology and population dynamics of *Ascaris lumbricoides* and *Trichuris trichiura* infection in the same community. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 1987, **81**: 987–993.

24. **Bradley M et al.** The epidemiology and population biology of *Necator americanus* infection in a rural community in Zimbabwe. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1991, **86**: 73–76.

25. **Bundy DAP et al.** Evaluating measures to control intestinal parasitic infections. *World Health Statistics Quarterly,* 1992, **45**: 168–179.