Validation of a chart to estimate true *Schistosoma mansoni* prevalences from simple egg counts

S. J. DE VLAS[1,2]*, D. ENGELS[3], A. L. T. RABELLO[4], B. F. J. OOSTBURG[5], L. VAN LIESHOUT[2], A. M. POLDERMAN[2], G. J. VAN OORTMARSSEN[1], J. D. F. HABBEMA[1] *and* B. GRYSEELS[6]

[1] *Department of Public Health, Faculty of Medicine, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands*
[2] *Department of Parasitology and Tropical Medicine, University of Leiden, P.O. Box 9605, 2300 RC Leiden, The Netherlands*
[3] *Schistosomiasis Control Programme, B.P. 337, Bujumbura, Burundi*
[4] *Centro de Pesquisas René Rachou, Fundação Oswaldo Cruz, P.O. Box 1743, 30190–131, Belo Horizonte, Brazil*
[5] *Laboratory of Parasitology, University of Surinam, P.O. Box g212, Paramaribo, Surinam*
[6] *Prince Leopold Institute of Tropical Medicine, Nationalestraat 155, B-2000 Antwerp, Belgium*

(*Received 8 May 1996 ; revised 26 July 1996 ; accepted 26 July 1996*)

SUMMARY

Schistosoma mansoni egg counts by faecal examination vary considerably and are not very sensitive, so prevalences are underestimated. The distribution of egg counts can adequately be described by a stochastic model which distinguishes variation in counts between persons and variation in repeated counts within a person. Based on this model a pocket chart has been developed which predicts the proportion of individuals harbouring at least 1 *S. mansoni* worm pair – the 'true prevalence' – from a simple single survey prevalence and geometric mean egg count (using common duplicate 25 mg Kato–Katz smears). The current paper describes the validation of this chart by comparing predicted true prevalences with prevalences observed after 5–7 repeated Kato–Katz faecal examinations (Burundi), by examination of a large quantity of stool using the Visser filter (Brazil) or a selective sedimentation–filtration method (Surinam). Because 5–7 repeated examinations do not suffice to measure all infections, predictions have been made of the cumulative proportion positives over 5–7 surveys – the 'approximate true prevalence' – as well. After dividing the data into age groups, 12 different subsets were considered for validation. In all 12 cases, predicted true prevalences (or approximate true prevalences for the Burundi data) agree well with those observed. The overall agreement depends only slightly on the assumed relationship between worm numbers and mean egg counts, with a good fit for a productivity between 0·8 and 4·4 eggs per gramme faeces (EPG) per worm pair (WP). This interval includes the most plausible value from the literature, i.e. 1·0 EPG/WP, which has been applied in the initial pocket chart. These findings support the validity of the chart to predict true prevalences for a wide range of productivity assumptions, and reinforces the applicability of its underlying stochastic model to describe egg count variation. However, as predictions appear to vary importantly when using only part of the data, it is also concluded that the pocket chart never compensates for limited validity of initial single survey prevalences and geometric means in consequence of small sample sizes.

Key words: *Schistosoma mansoni*, egg counts, prevalence, stochastic model.

INTRODUCTION

Detection and quantification of human *Schistosoma mansoni* infection is mainly based on counting eggs in stools. The faecal thick-smear technique (Kato & Miura, 1954; Katz, Chaves & Pellegrino, 1972) is widely accepted as the best diagnostic tool for use in the field (WHO, 1993). Slides prepared from templates with 20–30 mg of stools can already be screened after 15 min (Peters *et al.* 1980), and if necessary be followed by treatment on the spot. Since no alternatives for such an easily and rapidly applicable quantitative method are available as yet, the Kato–Katz technique is also the method of choice for population-based research and intervention studies on *S. mansoni*. Its lack of sensitivity (especially in detecting light infections), however, hampers interpretation of research results and evaluation of control programmes (De Vlas & Gryseels, 1992). Moreover, common knowledge on human worm burdens and *S. mansoni* egg production is conflicting, causing complications in relating egg counts to the number of worms actually harboured (Gryseels & De Vlas, 1996).

Recently, a stochastic model for egg count variation has been developed which incorporates the distributions of worms and worm pairs in the

* Corresponding author: Department of Public Health, Faculty of Medicine, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. Tel: +31 10 408 7714. Fax: +31 10 436 6831. E-mail: devlas@mgz.fgg.eur.nl.

population, as well as the variability of egg counts in repeated stool samples from an individual with a given worm pair load (De Vlas *et al*. 1992). The model thus explicitly distinguishes inter- and intra-individual variation in egg counts, and relates egg counts to worm burdens. Empirical data based on single and repeated faecal egg counts from several endemic communities could be described with only a few parameters. On the basis of this model a practically applicable pocket chart to predict true *S. mansoni* prevalences has been constructed (De Vlas *et al*. 1993*a*). This chart uses 2 measures from field observations (1) the prevalence from 50 mg single stool surveys (duplicate 25 mg Kato–Katz smears), and (2) the corresponding geometric mean egg count of detected positive individuals. By plotting the observed prevalence against the geometric mean, one can predict the true prevalence from the contour lines (Fig. 1). The pocket chart is based on 2 hypotheses (1) the underlying model for egg count variation can adequately be explained by only these 2 simple field measures, and (2) the true prevalence can properly be predicted from this model. In the original paper, researchers were invited to test the validity of this chart by comparing predicted true prevalences from the chart with observed true prevalences after using more sensitive techniques (De Vlas *et al*. 1993*a*). The current paper describes the experiences so far.

MATERIALS AND METHODS

Epidemiological data

The data come from 4 population surveys in 3 different countries: Burundi, Brazil and Surinam. 'Simple' prevalence and geometric mean EPG (eggs per gramme faeces) for positive individuals were based on examination of approximately 50 mg of faeces from a single stool sample. Duplicate 25 mg Kato–Katz faecal samples were prepared as described by Katz *et al*. (1972) and Polderman *et al*. (1985). More reliable prevalences for validation of the pocket chart (the 'observed true prevalence') were obtained after repeated Kato–Katz surveys, or by the additional examination of much larger amounts of stools using filtration techniques as described below. All individuals excreting eggs received treatment by praziquantel (40 mg/kg). In Brazil, this was 60 mg/kg for patients < 15 years and 50 mg/kg for patients $\geqslant 15$ years.

In Gihungwe (Burundi), repeated surveys were carried out on 7 occasions (days 1, 3, 5, 8, 10, 32, 37) in a study population of 200 individuals (100 adults and 100 children). Follow-up was almost complete, only 17 individuals missed one survey. The observed true prevalence is calculated as the cumulative proportion of positives after all 7 (or 6) measurements. The details of study-design and further outcomes are described by Engels, Sinzinkayo & Gryseels (1996).

In Buhandagaza (Burundi), 5 repeated surveys were performed with intervals of about 3 months (Gryseels & Nkulikyinka, 1988; Gryseels, Nkulikyinka & Engels, 1991). This data-set has previously been used to estimate parameters of the egg count model (De Vlas *et al*. 1992). For the present validation of the pocket chart, we only use those 231 persons (out of 435) with complete follow-up. The observed true prevalence is calculated by the accumulation of all 5 surveys.

In Sabará, Minas Gerais state (Brazil), the Visser method was used for detecting all *S. mansoni* infections in a random sample of 141 (out of 347) school children. The Visser filter enables examination of several grammes of faecal material (Pitchford & Visser, 1975; Schutte *et al*. 1994). For the current study, faecal samples were calibrated at around 1 g. Specimens to be filtered were formalinized, and eggs were stained with acid fuchsin on filter paper (Bell, 1963). Use of this highly sensitive method left only 1 positive individual using the Kato–Katz method (from the same stool sample) undetected. This lightly infected person, who only showed 1 egg, was included for the observed true prevalence.

In Catharina Sophia, district of Saramacca (Surinam), the sedimentation-selective-filtration (SSF) technique was used as a sensitive technique. Approximately 2–3 g of stools from each of 205 subjects were investigated. SSF is comparable to the Visser filter, but includes several steps of washing and rinsing. Probably this had lead to some 'loss of eggs', because 2 Kato–Katz surveys (1 from the same and 1 from an additional stool sample) showed another 16 individuals positive in addition to the 70 detected by SSF. See Polderman *et al*. (1994) for a description of the SSF technique and further details of this study. The observed true prevalence is assumed to be the combination of positives from SSF and both surveys.

The 4 data-sets were divided into age categories of about 50 individuals, with the condition of at least 20 of them being positive after 1 survey in order to guarantee a reliable estimate of the geometric mean EPG among positives. In total, 12 different subsets resulted for validation of the pocket chart (Table 1). In cases with repeated examinations – i.e. Gihungwe (7), Buhandagaza (5), and Catharina Sophia (2) – we made predictions of the true prevalence using the single survey prevalence and geometric mean of each survey separately, and then considered the average result for validation. This was done in order to diminish the effect of variation in initial measurements of single stool prevalence and geometric mean of positives due to the low number of individuals investigated. However, for the Gihungwe data we

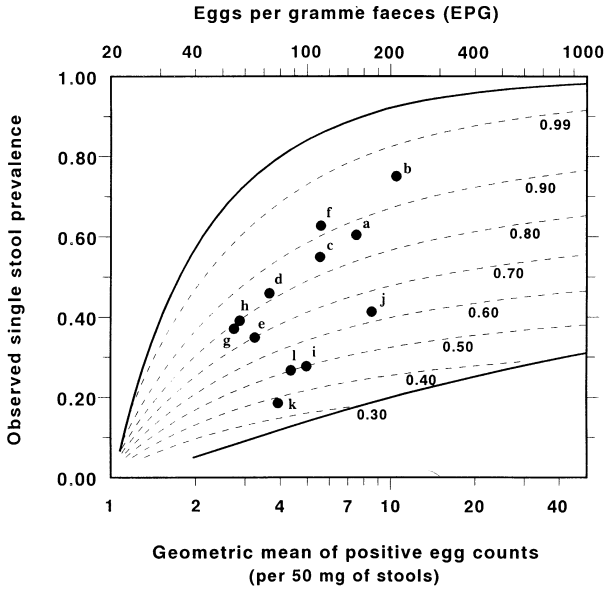

Fig. 1. Pocket chart to estimate true *Schistosoma mansoni* prevalences. By plotting observed single survey prevalence against the geometric mean of positive egg counts, the proportion of individuals with at least 1 worm pair (i.e. the true prevalence) can be predicted from the dashed contour lines. The model which underlies the predictions only applies for pre-control situations and is only defined between the solid lines (De Vlas *et al*. 1993*a*). The dots indicate the 12 different subsets that are used for validation of the chart, with indices referring to Table 1.

will also evaluate what would have happened in case only 1 survey had been available.

Model structure and validation

All 'true' prevalences inferred from the pocket chart (Fig. 1) have been based on an existing stochastic model which distinguishes variation in *S. mansoni* egg counts *between* individuals due to differences in the number of worm pairs harboured, and *within* individuals due to the variability of egg counts (De Vlas *et al*. 1992). The model is essentially based on 4 parameters: the mean M and aggregation parameter k of the underlying negative binomial distribution of individual worm burdens, the relationship between egg counts and worm burdens, and the aggregation parameter r of the negative binomially distributed egg counts within an individual. The smaller the value of k, the more the worms are aggregated in a small, highly infected part of the population. Similarly, a smaller value of r means that eggs are less homogeneously distributed over the subsequent stool samples and corresponds with more variation in repeated measurements. The relationship between egg counts and worm burdens has been assumed 1·0 EPG/WP (see also Gryseels & De Vlas, 1996). The aggregation in repeated individual egg counts r was earlier estimated at $r =$ 0·87 (De Vlas *et al*. 1992), and has also been assumed fixed for all further calculations.

This leaves only M and k as 'free' parameters to describe a particular endemic situation (for a specific age group). The pocket chart assumes that the prevalence and geometric mean EPG among positive individuals can adequately determine M and k, and thereby be used to explain the whole model. The chart was constructed by first calculating the values of M and k that correspond to each combination of prevalence and geometric mean, and then predicting from these parameters the proportion of individuals with at least 1 worm pair (i.e. the true prevalence). Mathematical details of the model and the pocket chart can be found in previous papers (De Vlas *et al*. 1992, 1993*a*, *b*).

Validation of the pocket chart occurs through comparison of predicted true prevalences with the observed true prevalences using the more sensitive approaches described above. Overlap of 90% confidence intervals, a high correlation and no systematic differences between observed and predicted true prevalences are criteria for a good fit. We use 90% intervals rather than the standard 95% in order to diminish the chance of unwarranted acceptance of the chart. The mathematical background of the calculation of confidence intervals and the deviance are given in the Appendix. We employ the deviance *dev* as an overall goodness-of-fit criterion to investigate alternative assumptions of both fixed parameters: the EPG/WP productivity and the aggregation r in repeated individual egg counts.

RESULTS

Table 1 gives an overview of the results for the 12 data sets considered. Single stool prevalences and geometric mean EPGs varied considerably between the communities and different age groups. A substantial part of the pocket chart is thereby covered (Fig. 1). The corresponding true prevalences as predicted from the chart are given in Table 1, and can be compared with the observed prevalences from the sensitive methods.

Although the observed and predicted 'true' prevalences in most cases overlap, the pocket chart systematically overestimates all 8 observed prevalences from Burundi (Table 1). This is due to the fact that the repeated Kato–Katz faecal examinations do not result in detection of all (lowly) infected individuals. The 5 repeated surveys in Buhandagaza and the 7 in Gihungwe still showed new infections in the last survey (De Vlas & Gryseels, 1992; Engels *et al*. 1996), so that it was very likely that more infections would have been detected in further surveys. Thus, the cumulative proportion of positive individuals only approximates and undoubtedly underestimates the true prevalence, and is not adequate for comparison with predictions from

Table 1. Results of the pocket chart for predicting true *Schistosoma mansoni* prevalences from simple egg counts, validated through comparison with results from more sensitive parasitological methods

(Geometric mean EPGs (of positives only) and simple observed prevalences ($\times 100\,\%$) were obtained from community surveys based on examination of about 50 mg Kato–Katz thick smears per individual. Predicted true prevalences represent percentages of individuals with at least 1 worm pair as can be read from the pocket chart using these two measures (Fig. 1). 'True' prevalences were empirically obtained by using more sensitive methods: repeated Kato–Katz surveys and/or filtration methods (see last column). For adequate comparison of model predictions with the observations from repeated surveys, additional predictions have been made of the cumulative proportion positive after the actual number of measurements (viz. 7 for Gihungwe and 5 for Buhandagaza). In the text these are referred to as 'approximate true prevalences'. Validation of the pocket chart is based on the comparison of the predicted (approximate) true prevalences with the observed prevalences from the sensitive method (both in **bold** typeface). Fig. 2 gives a graphical representation of the agreement. All intervals are 90 % CI.)

Focus (country)/ age category		No. of individuals	Geometric mean EPG	Observed single survey prevalence	Predicted true prevalence from the pocket chart	Predicted prevalence after actual number of measurements	Observed prevalence with sensitive method	Sensitive method
Gihungwe (Burundi)								
a	5–8	48	150	60·4	86·4 ± 3·3	**81·0** ± 3·2	**83·3** ± 7·4	7 repeated
b	9–16	52	209	75·0	95·4 ± 1·8	**92·2** ± 2·2	**92·3** ± 5·3	surveys
c	18–33	49	111	54·9	84·7 ± 3·7	**78·4** ± 3·6	**77·6** ± 7·9	
d	34+	51	73	45·9	82·1 ± 5·1	**73·9** ± 4·7	**68·6** ± 8·4	
Buhandagaza (Burundi)								
e	5–9	55	65	34·9	70·0 ± 7·6	**57·7** ± 5·5	**69·1** ± 8·9	5 repeated
f	10–19	45	112	62·7	91·7 ± 3·7	**84·7** ± 4·0	**84·4** ± 7·4	surveys
g	20–39	63	55	37·1	79·5 ± 7·2	**65·4** ± 5·9	**66·7** ± 8·2	
h	40+	68	57	39·1	79·9 ± 6·0	**66·9** ± 5·3	**63·2** ± 7·6	
Sabará (Brazil)								
i	11–13	47	99	27·7	**49·9** ± 13·5	—	**48·9** ± 9·3	Visser filter
j	14–18	80	171	41·3	**63·3** ± 7·4	—	**55·0** ± 6·0	+1 survey
Catharina Sophia (Surinam)								
k	1–20	100	79	18·5	**37·2** ± 6·5	—	**38·0** ± 6·3	SSF method
l	21+	105	87	26·7	**49·9** ± 6·0	—	**45·7** ± 6·0	+2 surveys

the chart. We have therefore added to Table 1 a column with the predicted prevalence for the actual number of measurements: 7 for Gihungwe and 5 for Buhandagaza. These 'approximate true prevalences' can be derived from single survey data on prevalence and geometric mean in much the same way as the 'true prevalences' have been estimated. 'Pocket charts' that predict the cumulative proportion positive after a specific number of surveys can be obtained from the first author on request. Note that, hypothetically, the original pocket chart (Fig. 1) predicts the cumulative proportion positive after an infinite number of repeated measurements. Henceforth, we will apply the predicted approximate true prevalences for comparison with the Burundi data. The validation of the pocket chart is now based on the agreement of the predicted (approximate) true prevalences and the observed prevalences using the sensitive methods (both in **bold** typeface, Table 1).

Fig. 2 shows that for all 12 situations confidence intervals of the observed and predicted (approximate) true prevalences clearly overlap. Although the confidence intervals of observations and predictions are quite large, in 10 cases the agreement between point estimations is striking with differences between observation and prediction $\leqslant 5\,\%$, or just 2–4 persons. Plotting observed and predicted prevalences against the corresponding single survey prevalences shows that there was no relationship with the level of endemicity (Fig. 2). There was furthermore no predominance in over- or underestimation of true prevalences, nor a relationship with age. The agreement, however, depended largely on the variation in initial measurements of single stool prevalence and geometric mean. The minimum and maximum predictions in Gihungwe differed dramatically if only 1 of the 7 surveys would have been used (Table 2). Without detracting from the overall statistical validation, this demonstrates that the chart is less informative for operational situations where examinations on only a few persons are available. Note that the corresponding (much wider) confidence intervals in Table 2 in 7 out of the 8 cases still overlap with the observed prevalences in Table 1.

All predictions were based on a relationship

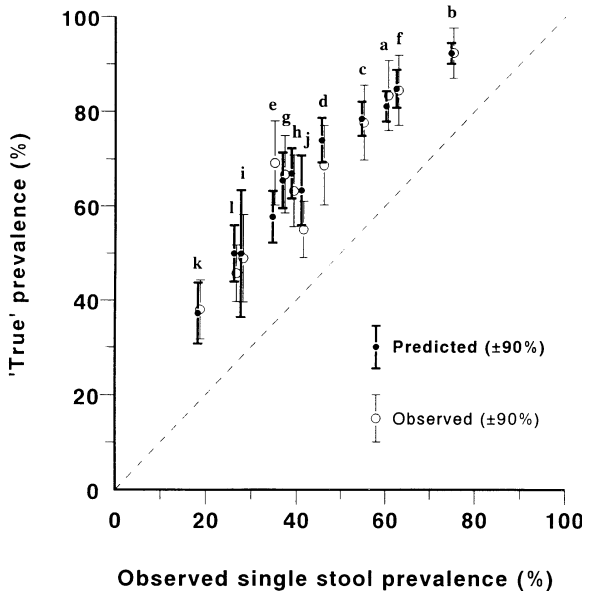

Fig. 2. Graphical representation of the comparison between observed and predicted (approximate) true prevalences plotted against the corresponding single survey prevalences. The indices correspond with the 12 subsets in Table 1. Intervals represent 90% CI.

between worm pair numbers (WP) and faecal egg counts (EPG) as 1·0 EPG/WP. At first sight, the results for alternative productivity assumptions would not differ much, as for most of the 12 data sets there would still be a good agreement between observation and prediction if a 5 times lower or 5 times higher productivity had been assumed. However, a productivity of 0·2 EPG/WP resulted in a systematic overestimation (Fig. 3A) and a productivity of 5·0 EPG/WP in a systematic underestimation (Fig. 3C) of (approximate) true prevalences, respectively, whereas for the initial assumption of 1·0 EPG/WP all 12 dots were evenly distributed around the dashed line (Fig. 3B). This suggests that, based on this agreement of predicted and observed true prevalences, an optimal estimate for productivity can be found between both extremes. Fig. 4 shows the joint goodness-of-fit, expressed by the deviance (see Appendix), as a function of the productivity of *S. mansoni* worms. The best fit was obtained for 2·0 EPG/WP ($dev = 9{\cdot}88$), with a broad 95% confidence interval ranging from 0·8 to 4·4 EPG/WP. This interval includes the value 1·0 EPG/WP which was assumed in the pocket chart. A separate estimate for the relationship between worm burdens and egg counts according to age of the human host, showed a higher estimate for adults (3·0 EPG/WP) than for children (1·6 EPG/WP), but this extension of the model did not lead to a significant improvement of the fit ($dev = 9{\cdot}48$; so $\chi^2_{\text{D.F.}=1} = 0{\cdot}40$, $P \approx 0{\cdot}5$). Fig. 4 also shows that assuming a larger value for r, or less variation in repeated individual egg counts, did not alter the results significantly, although the best fitting productivity appeared to be somewhat lower (1·6 EPG/WP).

DISCUSSION

This study tests whether true *S. mansoni* prevalences can be predicted from a pocket chart which uses only 2 simple field measures: the single survey prevalence and geometric mean among positives. In an earlier study we have already demonstrated that the underlying stochastic model for egg count variation can explain the relationship between prevalences from single and 3 stool examinations using an independent data source (Jordan *et al.* 1975; De Vlas, Van Oortmarssen & Gryseels, 1992). The current re-

Table 2. Ranges (90% CI) in predicted cumulative proportion individuals positive for *Schistosoma mansoni* infection after 7 and an infinite number of repeated stool examinations, i.e. the true prevalence, based on the observed single stool prevalence and geometric mean EPG of positive individuals

(For each of 4 age groups in Gihungwe village (Burundi), those 2 surveys (out of the 7 performed) have been selected that result in the lowest and the highest predictions.)

Age category	No. of individuals	Geometric mean EPG	Observed single survey prevalence	Predicted prevalence after 7 surveys	Predicted true prevalence	
5–8	48	220	50·0	59–73	63–80	Lowest
	48	134	66·7	81–97	86–100	Highest
9–16	52	280	75·0	85–96	89–99	Lowest
	51	181	76·5	88–100	93–100	Highest
18–33	49	170	46·9	57–72	61–80	Lowest
	49	88	59·2	77–96	84–100	Highest
34+	51	96	37·3	48–67	53–77	Lowest
	50	75	54·0	72–96	79–100	Highest

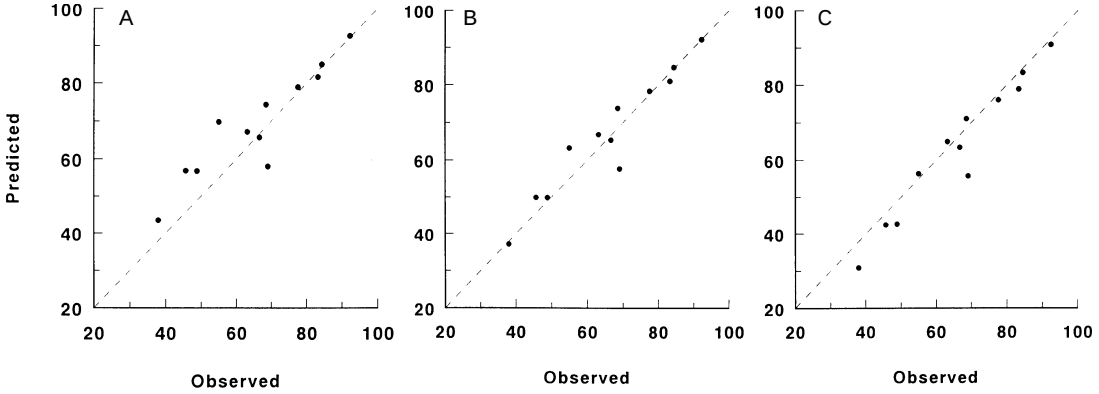

Fig. 3. Relationship between observed and predicted (approximate) true prevalences for 3 alternative values of the productivity of *Schistosoma mansoni* worms: 0·2 (A), 1·0 (B) and 5·0 (C) eggs per gramme faces (EPG) per worm pair (WP). Each dot represents a data set. Ideally, all dots would be positioned on the dashed line, leading to a deviance equal to 0. The value of 1·0 EPG/WP is the most plausible value from the literature and has been used in the original pocket chart.

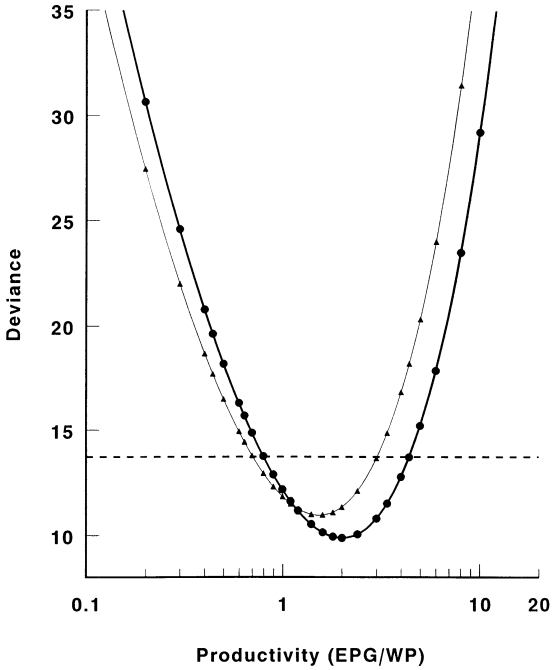

Fig. 4. Deviance for different values of the productivity in eggs per gramme faeces (EPG) per *Schistosoma mansoni* worm pair (WP). The lowest deviance indicates the value for which the predicted (approximate) true prevalences fit the observed prevalences best. The horizontal dashed line is 3·84 (i.e. $\chi^2_{1;5\%}$) above the curve minimum and indicates the 95% confidence interval for the best estimated productivity. (●) Results for the aggregation in repeated individual egg counts as has been assumed in the pocket chart ($r = 0{\cdot}87$); (▲) calculations for a less intense aggregation ($r = 1{\cdot}0$).

search is an extension in as much that now the model's ability to estimate (approximate) true prevalences has been tested.

Justification of the chart and model by measuring cumulative prevalences after repeated stool examinations will necessarily be incomplete because some infections are probably still being missed. Seven consecutive stool examinations is, however, at the limit of what can reasonably be obtained from community studies. Much longer periods between repeated examinations could avoid this problem but might, especially for young children, violate the initial model assumption of individual worm burdens not changing over time (De Vlas *et al*. 1992). The 3-month period between repeated examinations in Buhandagaza might therefore explain the relatively poor fit for 5 to 9-year-olds. Without showing a clear trend, the prevalences for this group fluctuated considerably over time (between 29 and 44%). Furthermore, delaying treatment of positives until the end of a long series of examinations is unacceptable for ethical reasons. In this respect, the data set from Gihungwe, 200 persons examined at 7 occasions over a 1-month period with almost no loss to follow-up, can be considered as one of the most valuable sets ever collected.

Filtration of large amounts of stool also offers no guarantee that all infected individuals will be found, as empirically demonstrated by the detection of some additional infections by the Kato–Katz surveys. Theoretically, one can expect a few very light infections still being missed by examination of 1 g of faecal material (as by the Visser filter), given the assumed one-to-one relationship between individual worm pair burdens and EPGs (Gryseels & De Vlas, 1996). The more qualitative SSF method starts from 2 to 3 g of faecal material, but leads to significant loss of eggs as demonstrated by the mean EPG count which can be 10 times lower than by the Kato–Katz method (Polderman *et al*. 1995).

Taking into account these considerations, we conclude that the high level of agreement between observations and predictions strongly supports the validity of the pocket chart to predict true levels of *S. mansoni* infection, and its underlying model to describe variations in egg counts. Still, the cumulative proportion positive after 5 (Buhandagaza) and 7 (Gihungwe) repeated measurements only approximates the desired true prevalences, and

thus provides only partial evidence. Comparable field experiments (in other endemic areas) would therefore still be welcome, especially if the endemic situation corresponds with a combination of prevalence and geometric mean that is located at parts of the pocket chart not yet covered. Determination of circulating antigens can provide another sensitive technique for testing the pocket chart. In the same Surinam community, the prevalence as predicted from the pocket chart was found in concordance with the results from immunodiagnosis by detection of the circulating antigens CAA and CCA (Van Lieshout *et al*. 1995). However, more research is necessary to find out to what extent the false negatives by the Kato–Katz method were compensated by false positive individuals due to cross-reactivity reactions.

The broad range of adequate values for productivity of *S. mansoni* worms demonstrate that the pocket chart is not very sensitive to the assumed relationship between worm pair burdens and mean egg counts. This means that the current chart can still be considered valid in case future evidence would point out that other, not too different, values are more likely. On the population level, indeed, variability in productivity might exist because of differences in *S. mansoni* strains or immunity levels. The presence of some density dependence in worm fecundity will also leave the current chart unaffected. In general, a lower productivity corresponds with a higher mean worm burden which, in turn, coincides with a higher individual probability of harbouring at least 1 worm pair and hence a higher true prevalence. Given the wide range in productivity, it is nevertheless reassuring that the best estimate of 2·0 EPG/WP is of the same order of magnitude as the initially assumed value of 1·0 EPG/WP. This is in contrast with a productivity of 5·0 EPG/WP, as can be estimated from autopsy data (Cheever, 1968), which is outside the interval. It has earlier been pointed out that a ratio of 1·0 EPG/WP corresponds with individual worm burdens numbering up to thousands or even tens of thousands in areas of moderate to high endemicity (Gryseels & De Vlas, 1996). In contrast to egg productivity, the aggregation parameter r is more important for describing community data with repeated individual measurements and therefore predicting true prevalences. Its value depends on the duration between successive surveys and, for example, the schistosome species involved (De Vlas *et al*. 1992). Re-analysis of the data originally used for fitting the underlying egg count model revealed that missing values could have resulted in an overestimation of the level of aggregation in repeated examinations, and therefore in an underestimation of r (unpublished observations). A value of $r = 1{\cdot}0$, which seems to be more accurate, however, hardly influences the general validation of the pocket chart.

The applicability of the pocket chart to estimate, or at least approximate, true prevalences is hereby proven to be statistically valid, and very robust to the main assumptions on parameter values. This does, however, not certify that the chart is always reliable. In operational conditions field researchers will principally apply the pocket chart based on only 1 survey, and for example in Gihungwe the predictions would diverge considerably if only 1 out of the 7 surveys had been used. In 3 age groups confidence intervals of minimum and maximum predictions would not even overlap! This is clearly due to the fact that prevalence and geometric mean are not reliable if only based on a single examination of 50 individuals. For practical purposes, field researchers should therefore convince themselves that the number of investigated individuals is at least high enough for an accurate measurement of single survey prevalence and geometric mean to be used in the pocket chart. The true prevalence can obviously never be more trustworthy than the measurements used to estimate it.

A fully validated pocket chart for obtaining the real number of infected people is a helpful tool for several purposes, for example, to decide whether indiscriminate mass treatment or selective treatment based on screening should be carried out, or to provide a 'statistical gold standard' for new diagnostic methods (De Vlas *et al*. 1993*a*). Obviously, the chart only acts on the population level and cannot reveal whether particular stool-negative individuals are infected or not. Based on the same underlying model for egg count variation, similar charts can be developed which provide predictions of the number of repeated surveys necessary to leave only a small proportion of infected persons undetected. In combination with careful cost considerations, such charts will provide an even more practical starting point for planning chemotherapy interventions based on screening.

The present study was not possible without the work carried out by Ms K. Deibel in Brazil, Ms H. Noordpool in Surinam, and E. Sinzinkayo and L. Nkulikyinka in Burundi. The Visser filter apparatus and detailed instruction for use was kindly donated by P. S. Visser and C. H. J. Schutte, Nelspruit, South Africa. We further thank Ms K. Polman for critical reading of the manuscript and N. Nagelkerke for invaluable statistical advice. The project was funded by the WHO/UNDP/World Bank special programme Tropical Diseases (TDR) and by the Science and Technology Programme of the European Communities.

APPENDIX

Statistical comparison of observed and predicted 'true' prevalences is based on the individuals negative after 1 stool examination, because only they can show up as positives using the more sensitive technique. A proper confidence interval of the proportion P_w positive after w repeated surveys (with P_w approximating the true preva-

lence if $w \to \infty$) therefore depends on the observed prevalence after one survey P_1. If N represents the total number of individuals in the data set, then from the number of individuals negative after one survey, $n = N(1-P_1)$, a proportion $p_{obs} = (P_{w,obs}-P_1)/(1-P_1)$ is observed to be positive using additional surveys or filtration. Similarly, a proportion $p_{pred} = (P_{w,pred}-P_1)/(1-P_1)$ is predicted to be positive if the pocket chart is used. An approximate 90% confidence interval of p_{obs} simply equals $p_{obs} \pm u_{0\cdot05}\sqrt{\{p_{obs}(1-p_{obs})/n\}}$. Multiplication of the interval with $(1-P_1)$ results in the desired 90% confidence interval of the observed prevalence using the sensitive technique:

$$P_{w,obs} \pm 1\cdot645 \frac{\sqrt{[(P_{w,obs}-P_1)(1-P_{w,obs})]}}{\sqrt{[(1-P_1)N]}}.$$

Calculation of the variation of p_{pred} is more complex, since the egg count model is used with the (correlated) geometric mean EPG as a second statistic. The jackknife resampling technique has been used to provide a confidence interval. By leaving out 1 individual at a time from the complete data set and determining the corresponding p_{pred}, a confidence interval can be obtained (Efron, 1982). Let $p_{(i)}$ be the ith jackknife replication of p_{pred} (i.e. from the data set with individual $(i) = 1 \ldots N$ removed), then from the pseudovalues $\tilde{p}_{(i)} = Np_{pred} - (N-1)p_{(i)}$ the 90% confidence interval of p_{pred} can be estimated as:

$$\tilde{p}(.) \pm t_{0\cdot05,N-1}\sqrt{\{\Sigma(p_{(i)}-\tilde{p}_{(.)})^2/[(N-1)N]\}}$$

with $\tilde{p}_{(.)}$ the mean of all $\tilde{p}_{(i)}$. Multiplication of the interval with $(1-P_1)$ again gives the desired 90% confidence interval of the predicted true prevalence $P_{\infty,pred}$ or approximate true prevalence $P_{w,pred}$.

Thus, the probability distribution of $x = N(P_w - P_1)$ positive individuals in a sample of $n = N(1-P_1)$ individuals negative after 1 survey is binomial with probability p_{pred}. The likelihood L of observing x infected individuals is therefore

$$L = p_{pred}^{x}(1-p_{pred})^{(n-x)} \quad \{\text{terms not in } p_{pred}\}$$

where the 'terms not in p_{pred}' are the combinational factors concerning the order of observations. The deviance, *dev*, between the $-2\log L$ function for all data sets and the best possible model, i.e. with $p_{obs,(j)} = x_{(j)}/n_{(j)}$ as probability for each subset $(j) = 1 \ldots 12$, is a joint indicator of the goodness-of-fit and equals

$$\begin{aligned} dev = -2\sum_{1}^{12}\{ & x_{(j)}\log(p_{pred,(j)}) \\ & + (n_{(j)}-x_{(j)})\log(1-p_{pred,(j)}) - x_{(j)}\log(x_{(j)}) \\ & - (n_{(j)}-x_{(j)})\log(n_{(j)}-x_{(j)}) \\ & + n_{(j)}\log(n_{(j)})\}. \end{aligned}$$

Test of significance and confidence intervals can be based on analysis of *dev*, with differences in *dev* between 2 hierarchical models following a χ^2 distribution.

REFERENCES

BELL, D. R. (1963). A new method for counting *Schistosoma mansoni* eggs in faeces. *Bulletin of the World Health Organization* **29**, 525–530.

CHEEVER, A. W. (1968). A quantitative post-mortem study of schistosomiasis mansoni in man. *American Journal of Tropical Medicine and Hygiene* **17**, 38–64.

DE VLAS, S. J. & GRYSEELS, B. (1992). Underestimation of *Schistosoma mansoni* prevalences. *Parasitology Today* **8**, 274–277.

DE VLAS, S. J., GRYSEELS, B., VAN OORTMARSSEN, G. J., POLDERMAN, A. M. & HABBEMA, J. D. F. (1992). A model for variations in single and repeated egg counts in *Schistosoma mansoni* infections. *Parasitology* **104**, 451–460.

DE VLAS, S. J., GRYSEELS, B., VAN OORTMARSSEN, G. J., POLDERMAN, A. M. & HABBEMA, J. D. F. (1993*a*). A pocket chart to estimate true *Schistosoma mansoni* prevalences. *Parasitology Today* **9**, 305–306.

DE VLAS, S. J., NAGELKERKE, N. J. D., HABBEMA, J. D. F. & VAN OORTMARSSEN, G. J. (1993*b*). Statistical models for estimating prevalence and incidence of parasitic diseases. *Statistical Methods in Medical Research* **2**, 3–21.

DE VLAS, S. J., VAN OORTMARSSEN, G. J. & GRYSEELS, B. (1992). Validation of a model for variations in *Schistosoma mansoni* egg counts. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **86**, 645.

EFRON, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Philadelphia.

ENGELS, D., SINZINKAYO, E. & GRYSEELS, B. (1996). Day-to-day egg count fluctuation in *Schistosoma mansoni* infection and its operational implications. *American Journal of Tropical Medicine and Hygiene* **54**, 319–324.

GRYSEELS, B. & DE VLAS, S. J. (1996). Worm burdens in schistosome infections. *Parasitology Today* **12**, 115–119.

GRYSEELS, B. & NKULIKYINKA, L. (1988). The distribution of *Schistosoma mansoni* in the Rusizi Plain (Burundi). *Annals of Tropical Medicine and Parasitology* **82**, 581–590.

GRYSEELS, B., NKULIKYINKA, L. & ENGELS, D. (1991). Repeated community-based chemotherapy for the control of *Schistosoma mansoni*: effect of screening and selective treatment on prevalences and intensities of infection. *American Journal of Tropical Medicine and Hygiene* **45**, 509–517.

JORDAN, P., WOODSTOCK, L., UNRAU, G. O. & COOK, J. A. (1975). Control of *Schistosoma mansoni* transmission by provision of domestic water supplies: a preliminary report of a study in St Lucia. *Bulletin of the World Health Organization* **52**, 9–20.

KATO, T. & MIURA, M. (1954). On the comparison of some stool examination methods. *Japanese Journal of Parasitology* **3**, 35 (in Japanese).

KATZ, N., CHAVES, A. & PELLEGRINO, J. (1972). A simple device for quantitative stool thick-smear technique in *Schistosoma mansoni*. *Revista do Instituto de Medicina Tropical de São Paulo* **14**, 397–400.

PETERS, P. A., EL ALAMY, M., WARREN, K. S. & MAHMOUD, A. A. F. (1980). Quick Kato smear for field quantification of *Schistosoma mansoni* eggs. *American Journal of Tropical Medicine and Hygiene* **29**, 217–219.

PITCHFORD, R. J. & VISSER, P. S. (1975). A simple and rapid technique for quantitative estimation of helminth eggs in human and animal excreta with special reference to *Schistosoma* spp. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **69**, 318–322.

POLDERMAN, A. M., KAYITESHONGA MPAMILA, MANSHANDE,

J. P. & BOUWHUIS-HOOGERWERF, M. L. (1985). Methodology and interpretation of parasitological surveillance of intestinal schistosomiasis in Maniema, Kivu Province, Zaire. *Annales de la Société Belge de Médecine Tropicale* **65**, 243–249.

POLDERMAN, A. M., GANGARAM PANDAY, U., RAMKISOEN, S., VAN LIESHOUT, L. & OOSTBURG, B. F. J. (1994). A sedimentation-selective filtration method for the diagnosis of light infections with *Schistosoma mansoni*. *Acta Tropica* **58**, 221–227.

SCHUTTE, C. H. J., PIENAAR, R., BECKER, P. J., VILJOEN, E., CROWTHER, G., VAN DEVENTER, J. M. G. & MARKUS, M. B. (1994). Observations on the techniques used in the qualitative and quantitative diagnosis of schistosomiasis. *Annals of Tropical Medicine and Parasitology* **88**, 305–316.

VAN LIESHOUT, L., GANGARAM PANDAY, U., DE JONGE, N., KRIJGER, F. W., OOSTBURG, B. F. J., POLDERMAN, A. M. & DEELDER, A. M. (1995). Immunodiagnosis of schistosomiasis mansoni in a low endemic area in Surinam by determination of the circulating antigens CAA and CCA. *Acta Tropica* **59**, 19–29.

WORLD HEALTH ORGANIZATION (1993). *The Control of Schistosomiasis : Second Report of the WHO Expert Committee*. WHO Technical Report Series No. 830. WHO, Geneva.