VARIATIONS IN FECAL SCHISTOSOMA JAPONICUM EGG COUNTS

J. M. YU, S. J. DE VLAS, H. C. YUAN, AND B. GRYSEELS

Department of Epidemiology, Shanghai Medical University, Shanghai, People’s Republic of China; Department of Public Health, Faculty of Medicine, Erasmus University Rotterdam, Rotterdam, The Netherlands; Prince Leopold Institute of Tropical Medicine, Antwerp, Belgium

Abstract. Variations in fecal Schistosoma japonicum egg counts were studied in ZhuXi administrative village, JiangXi Province, China. Population stool examinations were collected with duplicate, standard, 41.5-mg Kato-Katz thick smears on seven consecutive days for 570 individuals from two natural (individual) villages: village I with high endemicity and village II with low endemicity. The proportion of individuals with at least one positive count increased from 42.4% after a single measurement to 68.3% after seven measurements in village I (n = 356), and from 17.0% to 36.0% in village II (n = 214), respectively. This demonstrates a very high variation in repeated S. japonicum egg counts and a considerable lack of sensitivity of the Kato-Katz technique; light and moderate infections are especially missed with a single or a few measurements. The observed day-to-day variation in individual egg counts is highly aggregated (variance higher than the mean) and suggestive of a negative binomial distribution. For five individuals on three days, repeated sampling from different locations of a stool specimen shows a clear trend with egg counts decreasing from the beginning of the stool to the end and from the outside layer to the center. Ten multiple samples from a particular subsection (10–30 g) of a stool specimen for 44 positive individuals still showed aggregation in egg counts, particularly for high intensities of infection. This means that the aggregation in repeated daily S. japonicum egg counts cannot be explained alone by a specific day-to-day component and variation in the concentration of eggs at different locations in the stool. There also exists clustering of eggs within parts of the stool.

Most control programs and epidemiologic studies on Schistosoma japonicum and S. mansoni infection are based on the detection of parasite eggs using the modified Kato-Katz method. Generally, levels of egg excretion are used as an index of the intensity of infection. Individual diagnosis of infection is commonly based on one stool examination. However, for S. mansoni it has been shown that a single stool examination considerably underestimates the prevalence; many light infections are especially missed. Fecal egg counts as a quantitative measure of infection on S. mansoni in epidemiologic studies are the result of inter- and intra-individual variations: the former reflecting egg counts varying from individual to individual caused by differences in worm loads, and the latter reflecting egg count fluctuation for an individual with a given worm load from time to time and/or within a fecal specimen. Based on a stochastic model that takes into account this inter- and intra-individual variation in egg counts, projections can be made of how many individuals are truly infected with S. mansoni in a population.

We expect similar mechanisms to exist for S. japonicum infection, but thus far detailed studies have not been undertaken. Imperfect diagnosis of S. japonicum by routine stool examination has however been demonstrated earlier. In an epidemiologic study in the Philippines, use of a merthiolute-iodine-formaldehyde concentration method demonstrated several additional S. japonicum infections to those detected by a duplicate Kato-Katz examination.

In the present study, we describe the sensitivity of the Kato-Katz technique for S. japonicum infection by population studies in two Chinese villages of which inhabitants were investigated on seven consecutive days. Subsequently, we explore various factors that can cause the observed egg count variation.

Theoretically, four factors can be considered responsible for the overall egg count distribution in a population: 1) differences between individual counts due to the underlying differences in worm loads or worm pair loads; 2) differences from day-to-day within an individual due to fecal production and/or oviposition patterns of their female worms; 3) concentration of eggs at specific locations of a stool specimen; and 4) variation of repeated egg counts from the same location of a stool specimen due to micro-clustering of eggs. Since factors 2, 3, and 4 contribute together to the observed overall day-to-day variation in repeated individual egg counts, additional research is necessary to know more about the contribution of each separate factor. We have therefore applied multiple sampling of single stool specimens from selected infected individuals with the aim of relating egg counts to the location in the stool and testing the remaining variability of egg counts from a particular subsection of a stool specimen.

The resulting information on the degree of egg count variation and its origin will form the basis for future research in which the existing stochastic model for variation in S. mansoni egg counts will be adapted to S. japonicum and applied to estimate the proportion missed S. japonicum infections by the Kato-Katz technique for any endemic situation.

MATERIALS AND METHODS

Data. XinZi county, JiangXi Province, China is a well-known endemic area for S. japonicum. Control effects have been rather erratic, and mainly remained confined to irregular mollusciciding or individual treatment. The present study is based on data from ZhuXi administrative village (XinZi county), that consists of village I and II. Natural (individual) village I (366 inhabitants) is near the PoYan Lake, one of the largest fresh water lakes in China. The endemicity for S. japonicum infection in village I is relatively high. Village II (267 inhabitants) is located 1–2 km from PoYan Lake and is moderately or lightly endemic.

Written informed consent was obtained from all individuals who participated in the study. Ethical approval for the study was obtained from village (local government), county (anti-schistosomal), and provincial (schistosomiasis head-
For the population study, every inhabitant was asked to produce stool specimens on seven different days. Normally this happened on consecutive days, sometimes with a two- or three-day interval. From each specimen duplicate 41.5-mg thick smear slides (the standard in China) were prepared according to the modified Kato-Katz method.\textsuperscript{1,2,12} All slides were prepared and read by the same member of the project (JMY). To maintain quality control, part of the slides were re-examined by Dr. Q. J. Yang (Department of Epidemiology, Shanghai Medical University). A total of 356 cases in village I had complete data. The 10 individuals with missing values were excluded from the study. In village II, there were 214 persons with complete data and 53 were excluded because of missing data.

Five positive cases were selected for more detailed study of egg count variation. All five persons were males, one 12-year-old child and four adults. They were asked to produce complete stools for three consecutive days, of which 30 stool samples were taken in a very systematic way. Every stool specimen was divided into five equal sections, with the beginning of the stool (the part that comes out first) called section 1 and the last part section 5. To know what is the beginning and end of a stool, the persons were asked to straighten the stool on a board. The consistency of the stool specimens was formed. Duplicate Kato slides were taken from each section at three different depths: 1) the outside layer, 2) more to the middle, and 3) the center of the stool.

For studying random variation of egg counts within parts of a stool specimen, 44 stool-positive cases (23 adults and 21 children up to 18 years old) in village I were asked to deliver another stool specimen of which 10 Kato slides were taken. For each individual, a 10–30 g subsection of the complete stool (\(\approx 250 \text{ g}\)) was used, from which all 10 samples were prepared.

**Statistical analyses.** The sensitivity of Kato-Katz method has been studied by comparing the proportion positive cases after a single examination (i.e., single stool prevalence) with the proportion showing at least one positive count after the seven accumulated surveys. This cumulative prevalence is used as the gold standard. The overall trend in cumulative proportion positives has been calculated on the basis of all possible permutations. This means that the chronologic order of surveys was not explicitly taken into account. For example, the cumulative prevalence after two surveys is the average value of 21 possible combinations (survey 1 and 2, 1 and 3, \ldots, 6 and 7).

To determine the sensitivity of Kato-Katz method for different intensities of infection, individual geometric mean egg outputs over all seven surveys are classified into four egg per gram (epg) categories: negative, light (1–100 epg), moderate (101–400 epg), and heavy (> 400 epg) infections. The geometric mean epg is expressed as \(12 \times \text{antilog } [\text{arithmetic mean log (egg count + 1)]} - 1\). The value 12 is the correction factor because the total egg count in duplicate 41.5-mg examinations from each specimen was used to obtain the mean log egg count.

Variation in repeated measurements on different days, or within stool specimens were studied by plotting the variances against the arithmetic means for each individual and calculating the regression line: log variance = \(a + b \times \log \text{ mean}\). In case of a random or homogenous distribution of egg counts (i.e., Poisson), we expect that the variance equals the mean. This would be reflected by a slope \(b\) equal to 1.0 and a small intercept \(a\). If the variance exceeds the mean, the distribution is called overdispersed or aggregated. A slope of \(b\) near 2.0 indicates that the negative binomial distribution could be an adequate representation of such data.\textsuperscript{13,14}

Deming regression was used instead of ordinary linear regression because measurement variation has to be considered in both the dependent (log variance) and the independent variable (log mean).\textsuperscript{15} We chose to minimize the perpendicular distance of each point to the regression line, so that equal weight was given to measurement error of the \(x\)-value and the \(y\)-value. We applied the likelihood ratio test for the statistical comparison of different regression models. In particular, linear models were compared with stepwise linear models that include a break in the slope.

Two-way analysis of variance on the log (egg count + 1) was used to test the relationship between egg counts and their origin (5 sections and 3 depths) in the stool specimen.

### RESULTS

The data in Table 1 reconﬁrm that the part of ZhuXi administrative village close to PoYan Lake (village I) has for all age groups a much higher endemicity than the more distinct part (village II). Village I has relatively more adults than village II, with sexes evenly distributed among both villages. Individuals in village I reach the maximum level of infection at a younger age than in village II.

The examination of duplicate Kato-Katz slides from a single stool specimen seriously underestimates the final cumulative prevalence after seven repeated surveys. Although point prevalences of infection detected by a single stool examination varied considerably, between 39.6\% and 46.4\% in village I and between 14.5\% and 20.6\% in village II, there was no clear trend over time. The average cumulative percentage of people with at least one positive egg count in all examined duplicate slides increased from 42.4\% (single stool) to 68.3\% (7 measurements) in village I, and from 17.0\% to 36.0\% in village II, respectively (Figure 1). Thus,
in both villages six extra stool examinations increased the detected prevalence of infection about 20%. Since there is no clear leveling-off in the cumulative proportion of positive individuals, probably more infections would have been found after additional surveys.

The data in Table 2 show the sensitivity to detect individuals with different intensities of infection, assuming the geometric mean egg count after seven repeated examinations as the gold standard. As expected, the sensitivity of one stool (or a few) stool examinations to detect *S. japonicum* infection is higher among heavy than among light or moderately infected individuals. Because village II shows a lower level of infection, sensitivity is lower than in village I. Therefore, additional stool examinations are more beneficial in the lightly infected group (1–100 epg) than in the moderately (101–400 epg) and intensely infected groups (>400 epg).

Intra-individual variation in seven consecutive daily stool examinations is high, with almost for every infected person the variance exceeding the mean (Figure 2). The overall slope of the regression line equals 1.68. Inclusion of a breakpoint at 5.0 results in a significantly better fit (*P* < 0.001), with the difference between 2 × log maximum likelihood of both models (with and without a breakpoint) being 13.6. For more intensely infected individuals after the break, the slope (1.85) approximates 2.0 (dashed line) and indicates that the negative binomial distribution seems to be adequate to describe the intra-individual variation in *S. japonicum* egg counts.

This day-to-day variation will partly be due to differences in the location of the samples from each stool specimen. Figure 3 shows a systematic relationship between egg counts and the section and depth of the sample, which clearly indicates that distribution of eggs in the stool is nonrandom. Apparently, most eggs are concentrated in the outside layer of the beginning of a stool specimen. From this location egg counts are easily up to five times higher than from the inner parts of the end of a stool specimen (Figure 3).

Taking 10 repeated samples from a particular subsection of a stool still shows that generally individual variances are higher than the mean egg counts, especially for highly infected individuals (Figure 4). Thus, there even exists a certain clustering of eggs on the spot.

**DISCUSSION**

The modified Kato-Katz method, because of its many practical advantages for examining large numbers of people

---

**Table 2**

<table>
<thead>
<tr>
<th>Repeated times</th>
<th>Village I</th>
<th>Village II</th>
<th>Light infection (n = 202)</th>
<th>Moderate infection (n = 70)</th>
<th>Heavy infection (n = 48)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>62 (58–68)</td>
<td>47 (40–57)</td>
<td>40 (38–46)</td>
<td>87 (83–93)</td>
<td>96 (92–98)</td>
</tr>
<tr>
<td>2</td>
<td>77 (70–82)</td>
<td>65 (53–73)</td>
<td>60 (51–67)</td>
<td>98 (94–100)</td>
<td>99 (96–100)</td>
</tr>
<tr>
<td>3</td>
<td>85 (79–90)</td>
<td>76 (69–83)</td>
<td>73 (65–80)</td>
<td>99 (98–100)</td>
<td>99 (98–100)</td>
</tr>
<tr>
<td>4</td>
<td>91 (85–95)</td>
<td>84 (77–91)</td>
<td>83 (75–91)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>95 (89–98)</td>
<td>90 (83–97)</td>
<td>90 (83–96)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>98 (95–99)</td>
<td>95 (91–100)</td>
<td>95 (92–99)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

* Numbers are the average values for all possible permutations (thus without taking into account chronologic order of examinations). The figures in parentheses are the lowest and highest sensitivities in this respect. Duplicate Kato-Katz stool examinations were taken from 570 individuals on seven consecutive days. The cumulative proportion positive is considered the gold standard.
at low cost within a short time, is the standard method for the field diagnosis of \textit{S. japonicum} infection. However, a single stool examination with the Kato-Katz method will miss a certain number of infections, especially among lightly infected individuals, and misdiagnosis is considerable. In this study, only about half of all infections diagnosed by seven repeated measurements were found by a single stool examination with duplicate slides of 41.5 mg. Also given the still increasing trend in cumulative prevalences after five, six, and seven measurements (Figure 1), the lack of sensitivity in using the Kato-Katz method to diagnose \textit{S. japonicum} infection is of the same order of magnitude as for \textit{S. mansoni}.

Obviously, intense infections are easier to diagnose (Table 2) because higher worm loads will lead to higher densities of eggs in the stool. In operational circumstances, with the main objective of controlling morbidity, a single stool measurement can still be very useful because most of the pathology is connected with heavy infections. However, if control or even eradication of \textit{S. japonicum} infection is the objective, single or a few fecal egg count measurements will never suffice. In that case application of more sensitive method, perhaps including serologic methods are necessary.

As for a comparable study on variations in \textit{S. mansoni} egg counts, modeling can reveal the number of additional infections to be detected after eight or more repeated examinations, with in theory the true prevalence as the predicted proportion of individuals with at least one worm pair. Development of a stochastic model of egg count variation in \textit{S. japonicum} infection to predict true prevalences, will be the subject of a forthcoming report (Yu JM and others, unpublished data). Results of the current study will be applied to select promising mathematical distributions and to quantify parameter values.

The slope of 2.0 between log variance and log mean for Figure 2. Day-to-day variation in seven repeated egg counts (duplicate Kato-Katz smears) for \textit{Schistosoma japonicum} infection from all 320 positive cases in Villages I and II. Each dot indicates one individual. For most cases the variance exceeds the mean, reflecting aggregation in repeated daily egg counts. The straight line represents the results of Deming linear regression between log variance and log mean egg counts. The best fitting line has a breakpoint at 5.0. Before the breakpoint the slope equals 1.55 with an intercept of 0.44; after the breakpoint the slope equals 1.85, which is highly suggestive of a negative binomial (dashed line with a slope of 2.0).

Figure 3. Variation in \textit{Schistosoma japonicum} egg counts due to taking samples from different locations of a stool specimen. Geometric mean egg counts (duplicate Kato-Katz smears) for five different individuals on three days are presented for five different sections (where section 1 indicates the beginning and 5 the end of the stool) and three different depths (the outside layer, more to the middle, and the center of the stool section). Trends against section ($F_{1,223} = 23.10; P < 0.0001$) and depth ($F_{1,223} = 6.50; P = 0.012$) are significant, thereby indicating that eggs are nonrandomly mixed in the stool with most eggs concentrated in the outside layer of the beginning of the stool.

Figure 4. Results of 10 multiple stool samples (single, 41.5-mg Kato-Katz smears) from the same part (10–30 g) of a stool specimen from 44 selected positive cases in Village I. Each dot indicates one individual. The best fitting Deming linear regression line between log variance and log mean egg counts (straight line) contains no break and has a slope of 1.50 with intercept of $-0.15$. For high mean egg counts, the variance is higher than the dashed unity line, thereby indicating aggregation or clustering of eggs even at particular parts of a stool specimen.
high egg counts indicates that the negative binomial distribution is a plausible choice for describing day-to-day variations in individual *S. japonicum* egg counts. Such a distribution for variation in repeated daily egg counts is a common assumption for other helminthic infections, such as hookworm and *S. mansoni*.\(^{13,14,19}\) The observed day-to-day variation is a combination of within-stool variation (both concentration of eggs at particular locations in the feces and/or microclustering of eggs) and a certain day-to-day component. This day-to-day component probably is a common phenomenon for other helminths living in or next to the intestine. It may exist of changes in stool bulk, with more fecal material diluting the eggs, and/or it may be due to diurnal or longer-term fluctuations in the egg production or oviposition of female worms. Intense within-stool variation egg count variation also is common for other helmith infections, and typically shows an aggregated distribution.\(^{3,20,21}\) Remarkably, in the only study we know of on within-stool variation for *S. japonicum* infection, the Poisson distribution (i.e., homogeneous distribution, thus no aggregation) appeared to be adequate to describe the observations.\(^{22}\) However, this study was based on animal hosts.

All previous studies on within-stool variation only concern samples being taken systematically or randomly from complete stools. Our study is the first that distinguishes variation in within-stool variation according to the location of the sample and on the micro level. For *S. mansoni* infection, the possibility of a systematic pattern in egg count according to the origin of the sample has been the subject of debate. Concentration of *S. mansoni* eggs at the outside layer of stools has been reported by e.g., Khalil and Salah El Din,\(^{23}\) but this was not confirmed in studies by Martin and Beaver\(^{20}\) and Ratard and others.\(^{24}\) Also, from experiments on human volunteers that had ingested glass beads, it was suggested that helmith eggs should be randomly scattered in the stools.\(^{25}\) The highly significant trend in *S. japonicum* egg counts with location in the stool (highest egg counts at the outside layer of the beginning of the stool) that was observed in our study could be due to poor mixing of eggs, knowing that eggs are deposited at the outside layer of the fecal mass, and extraction of more water from the fecal material that stayed longest in the large intestine (i.e., beginning of the stool) leading to a higher concentration of eggs per amount of fecal material in those places.

Since the 10 samples from the same location in a stool (Figure 4) show much less aggregation than for daily samples (Figure 2), it can be concluded that the day-to-day component, in combination with the variation according to location of the sample, contribute largely to the observed aggregation of repeated egg counts. Nevertheless, the still existing aggregation in repeated sampling from the same location (Figure 4) demonstrates a certain clumping of eggs at the micro level. This microclustering of eggs is not likely to be observed directly in the Kato-Katz slides because the fecal samples are sieved before screening. In direct smears without sieving, however, sticking of eggs into groups can sometimes be observed (Polderman AM, unpublished data). Clumped deposition of eggs by the female worms can be the reason for this phenomenon.

Comparison with studies for *S. mansoni* shows that the variation in *S. japonicum* daily egg counts, as observed in our study, is relatively more intense.\(^{14,19}\) In situations in which we found a scattering around the line with a slope of 2 after plotting the variation against the mean of repeated *S. japonicum* egg counts (Figure 2), similar plots for *S. mansoni* show that all individual cases are located below such a line.\(^{14}\) Perhaps this difference in variability is caused by the differences in the location of both worms since *S. japonicum* worms normally live in lower parts of the intestinal system, so that eggs deposited in clumps are therefore less prone to mixing.\(^{26}\) Differences in diets, and therefore stool production, between Chinese and Brazilians might further affect this discrepancy. By means of more sophisticated modeling, we will further explore this phenomenon.

Acknowledgments: The field work would not have been possible without the help of Q. J. Yang (Department of Epidemiology, Shanghai Medical University). Caspar Looman of the Department of Public Health, Erasmus University Rotterdam, kindly offered statistical advice.

Financial support: The project was funded by the UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases and by the Science and Technology Program of the European Communities.

Authors’ addresses: J. M. Yu, Department of Epidemiology, Shanghai Medical University, Shanghai 200032, People’s Republic of China and Department of Public Health, Faculty of Medicine, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. S. J. de Vlas, Department of Public Health, Faculty of Medicine, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. H. C. Yuan, Department of Epidemiology, Shanghai Medical University, Shanghai 200032, People’s Republic of China. B. Gryseels, Prince Leopold Institute of Tropical Medicine, Nationalestraat 155, B-2000 Antwerp, Belgium.

REFERENCES


