Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
International Journal for Parasitology 36 (2006) 895–902
INTERNATIONAL Journal for PARASITOLOGY
www.elsevier.com/locate/ijpara

Bayesian estimation of community prevalences of *Schistosoma japonicum* infection in China☆

Xian-Hong Wang, Xiao-Hua Wu, Xiao-Nong Zhou *

National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, 207 Rui Jin Er Road, Shanghai 200025, China

Received 6 February 2006; received in revised form 5 April 2006; accepted 11 April 2006

Abstract

A Bayesian approach to overcome the imperfections of an immunological test (an antibody-based ELISA) and a parasitological test (Kato-Katz) in the detection of *Schistosoma japonicum* infection, was used to estimate community prevalences of *S. japonicum* infection in China. At the same time, the similarity between the prevalence estimates based on data from ELISA alone and those using data from both ELISA and Kato-Katz tests was explored. The database from the third nationwide sampling survey of schistosomiasis in China, 2004, was used for analysis, in which a total of 239 endemic villages were sampled from seven endemic provinces through a stratified cluster sampling technique and 250,987 residents aged from 6 to 65 years, were examined by ELISA followed by a Kato-Katz test applied to the seropositives. Bayesian hierarchical models incorporating random effects to reflect the nested data structure and uncertainty about test properties were employed to analyse the data. Our analysis suggested that using data from ELISA alone or both ELISA and Kato-Katz tests resulted in similar prevalence estimates, probably owing to the lack of sensitivity of Kato-Katz and the fact that Kato-Katz was only applied to the seropositives. We conclude that it is feasible to employ only ELISA, instead of combined ELISA and Kato-Katz tests, to estimate prevalence of *S. japonicum* infection in large-scale epidemiological settings. This study confirmed heterogeneity in the prevalence of *S. japonicum* infection in space by the fact that the estimated prevalences of *S. japonicum* infection in the sampled villages ranged from 0.02% to about 56% (posterior median). It is indicated that the disease remains a threat in some areas along the Yangtze River, although great achievements have been made in the control programme of schistosomiasis in China.

Keywords: Bayesian statistics; Prevalence; ELISA; Kato-Katz; *Schistosoma japonicum*; China

1. Introduction

Schistosomiasis due to *Schistosoma japonicum* infection was once a devastating disease in China (Chen and Feng, 1999; Ross et al., 2001; Zhou et al., 2005b). Approximately 12 million people were infected with *S. japonicum* in the mid-1950s and the disease disabled and killed millions of persons (Chen and Feng, 1999; Ross et al., 2001). The national schistosomiasis control programme initiated in 1955 and the 10-year World Bank Loan Project for schistosomiasis control launched in 1992, led to an interruption or control of transmission in many formerly endemic counties of China (Yuan, 1995; Chen and Feng, 1999; Chen et al., 2005). Data from periodic epidemiological surveys at both regional and national levels revealed that both the prevalence of *S. japonicum* infection and the estimated number of infected people dramatically declined during the last five decades (Zheng, 1993; Office of EDC, 1998; Li et al., 2000, 2005); only seven out of 12 formerly endemic provinces, namely Anhui, Hubei, Hunan, Jiangsu, Jiangxi, Sichuan and Yunnan, are still in the status of transmission of *S. japonicum*. However, an increase in *S. japonicum* infections was observed in some areas of the aforementioned seven endemic provinces along the Yangtze River after 1998, which was attributed to changes in ecological and social factors (Zhou et al., 2004, 2005b).

The diagnostic methods used in epidemiological surveys have been improved significantly during the last few decades (Wu, 2002; Bergquist, 2002; Zhu, 2005). In the early days of the national control programme, only stool examination such as the Kato-Katz thick smear method was applied. With the successful implementation of control activities, which consequently resulted in a decrease in the infection rate and intensity in humans, the problem of low sensitivity of stool examination methods became a crucial issue in epidemiological surveys and

☆ Note: Supplementary material associated with this paper.

* Corresponding author. Tel.: +86 21 64738058; fax: +86 21 64332670.
E-mail address: ipdzhouxn@sh163.net (X.N. Zhou).

0020-7519/$30.00 © 2006 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijpara.2006.04.003

precise estimation of community prevalence (Zhu, 2005). With advances in molecular biology, immunodiagnostic kits were developed and have been used as screening tools in local surveillance systems (Zhu, 2005; Zhu et al., 2005). For example, an indirect hemaglutination test (IHA) was used as a tool to screen and identify targets for mass chemotherapy in the World Bank Loan Project on schistosomiasis control in China (Wang et al., 2000). Dipstick dye immunoassay (DDIA) was used for detecting cases infected with *S. japonicum* (Zhu et al., 2005). Antibody-based ELISA was used in the third nationwide sampling survey of schistosomiasis in China (or the 3rd NSC), organised by the Chinese Ministry of Health in 2004, for screening followed by Kato-Katz stool examination which was applied to seropositives from ELISA tests. The reason for using a combination of ELISA and Kato-katz was to overcome the problem of low specificity of serological tests (Wu, 2002; Doenhoff et al., 2004) and greatly reduce the labour and cost requirements of the Kato-Katz technique (Zhu, 2005). However, there is still a need for a proper statistical method to incorporate the test errors into the community prevalence estimation if a combination approach of screening by a serological test (ELISA) and subsequent examination of the seropositives by a parasitological test (Kato-Katz) is used, particularly in the post-transmission areas where both prevalence and infection intensity are usually quite low (Giboda and Bergquist, 1999; Wu et al., 2005). Meanwhile, given the low sensitivity of the Kato-Katz test and its labour-intensive and time-consuming features (Zhu, 2005), it will be useful to explore whether the Kato-Katz test contributes to prevalence estimation when combined with an ELISA test, especially in post-transmission areas.

Recently, Bayesian analysis has been increasingly used for estimating measures of test accuracy and prevalence inference in epidemiological studies of schistosomiasis (Carabin et al., 2005) and other parasitic diseases (Dendukuri and Joseph, 2001; Black and Craig, 2002; Utzinger et al., 2002; Carabin et al., 2003; Basanez et al., 2004; Dorny et al., 2004; Geurden et al., 2004). A Bayesian approach not only treats sensitivity and specificity as well as prevalence as variables, but can also incorporate prior information about these into the analysis, thus improving the estimation when substantive prior information exists (Joseph et al., 1995; Carabin et al., 2005; Schurink et al., 2005). Therefore, the purpose of our study is: (i) to estimate the community prevalence of *S. japonicum* infection by Bayesian analysis based on data obtained from the 3rd NSC; and (ii) to explore if prevalence estimates based on data from an ELISA test alone are similar to those using data from both ELISA and Kato-Katz tests.

2. Materials and methods

2.1. Prevalence data

The 3rd NSC was carried out among all the endemic villages in the seven provinces with a current epidemic of schistosomiasis (Fig. 1) recorded by organisations of schistosomiasis control at administrative levels in China (Zhou et al., 2005a). According to geographical patterns of the endemic areas and ecological characteristics of the intermediate host snail of *S. japonicum*, these villages have been classified into several ecotypes (Chen and Feng, 1999), i.e. fork-beach, islet without embankment, islet with embankment, inner embankment in the marshland and lake regions, the plain regions with waterway networks, and the plateau, mountain and hill in the hilly and mountainous regions. In order to facilitate the implementation of different control strategies, these villages have also been cataloged into four prevalence levels, i.e. $\geq$10%, 5-10%, 1-5% and less than 1%, which were estimated

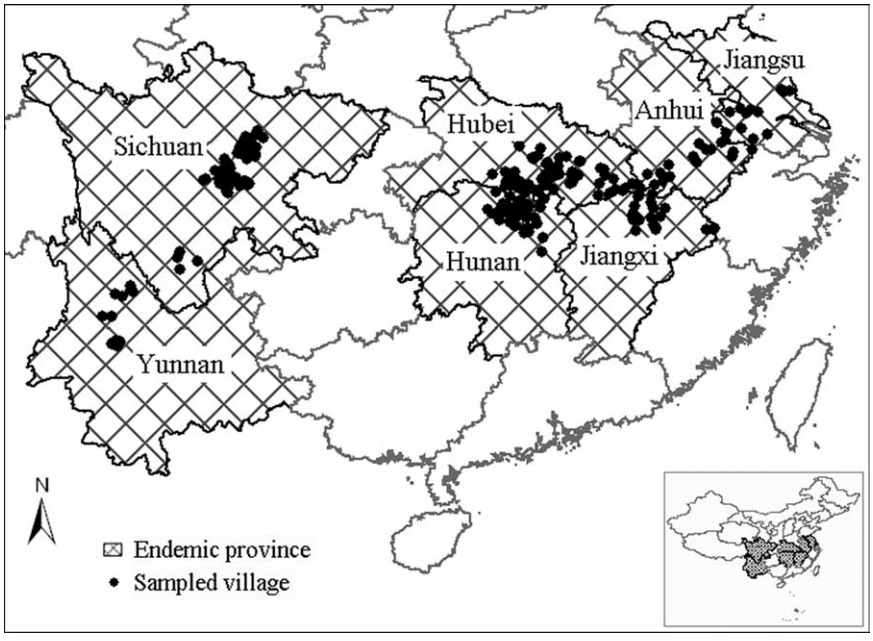

Fig. 1. Map of southern China showing the screening zone (endemic provinces: Anhui, Hubei, Hunan, Jiangsu, Jiangxi, Sichuan and Yunnan) and the location of sampled villages in the third nationwide sampling survey of schistosomiasis in China, 2004.

and adjusted by experts based on the latest regional sampling survey. In order to estimate the prevalence of infection for each province, ecotype and prevalence level, all sampled villages in the 3rd NSC have to be well representative of those villages within different provinces, ecotypes and prevalence levels, therefore a stratified cluster sampling technique with an endemic village as a unit, as with the second nationwide sampling survey on schistosomiasis in 1995 (Office of EDC, 1998), was applied to the endemic villages which were stratified by province, ecotype and prevalence level estimated by experts. It was required that the sampling ratio of villages would not be less than 1%. Residents aged 6–65 years were asked to participate in the survey. At first, an antibody-based ELISA test was used according to the manufacturer's instructions as a serological screening test. A parasitological test was then applied to the seropositives by reading three Kato-Katz thick smears from one stool specimen (Katz et al., 1972; WHO workshop, 1980). The data from whether a resident was examined by ELISA or Kato-Katz and that of the result (positive or negative) were used for analysis in this study.

2.2. Questionnaire survey on sensitivities and specificities of the tests

A two-round interview based on a questionnaire was carried out to collect expert knowledge on sensitivities and specificities of the tests when used in the field. In the first round, an open questionnaire was developed and experts from the National Institute of Parasitic Diseases (IPD), the provincial IPDs or CDCs or professors from the medical universities of the seven endemic provinces were asked about their estimates of the average, minimum and maximum sensitivity and specificity of each test. The medians of the minima and maxima were calculated, respectively, for the sensitivity and specificity of each test. In the second round, these values were sent back to the experts with a semi-structured questionnaire for reference, and he or she was asked to give alternatives if he or she did not agree with the given values. The medians of values on upper and lower limits were calculated, respectively, for the sensitivity and specificity of each test and served as ranges of test properties.

2.3. Statistical analysis

2.3.1. Conventional statistics

All data management of the 3rd NSC was performed under the supervision of IPD, which also kindly provided the database for the present study. The number of subjects examined by serological or parasitological tests and the respective test outcomes were calculated for each sampled village using SAS version 8.0 (SAS Institute, Inc., NC, USA).

2.3.2. Bayesian modeling

A total of three steps in Bayesian analysis were undertaken to establish proper models as follows:

Prior distributions. The first important step in Bayesian analysis is to obtain prior distributions for all model parameters. This was accomplished in the present study by drawing upon expert knowledge on the sensitivity and specificity of each test and we chose beta(α, β) distributions for these (Joseph et al., 1995; Willian, 2004). Let π be the prior mean, matched with the center of the range of the test property, and let σ be the prior SD, matched with one quarter of the range of the test property. Then the formulas defining α and β are:

$$\alpha = \pi\left(\frac{(1-\pi)\pi}{\sigma^2} - 1\right)$$

$$\beta = (1-\pi)\left(\frac{(1-\pi)\pi}{\sigma^2} - 1\right)$$

We chose beta(1,1), a non-informative prior distribution, for the prevalence.

Model establishment. In the second step, in order to see if the estimated prevalences of *S. japonicum* infection using data from both ELISA and Kato-Katz tests were similar to those using data from ELISA alone and thus understand the contributions of the Kato-Katz test to the prevalence estimation, we assumed two different situations: In situation 1, data from both ELISA and Kato-Katz tests in the 3rd NSC were used. A Bayesian hierarchical model was developed to draw inferences about the prevalence of *S. japonicum* infection in each sampled village, that in each ecotype within a province as well as that in each province. Independence between ELISA and Kato-Katz was assumed since they measure different biological phenomena (Georgiadis et al., 2003). We assumed the sensitivity and specificity of ELISA and Kato-Katz, respectively, was the same in all villages within each ecotype. Random village, ecotype and province effects were included in the model to reflect the data structure (Schaik et al., 2003), i.e. villages were nested in ecotype and ecotypes were nested in provinces (see supplementary material for details). In situation 2, it was supposed that all the residents were merely examined by ELISA, so we used the data from ELISA alone in the 3rd NSC to estimate the prevalence of *S. japonicum* infection in each sampled village, that in each ecotype within a province as well as that in each province. The model assumptions and the random effect structure were the same as those in situation 1 except that no relationship between ELISA and Kato-Katz was needed to be defined since there was only an ELISA test in situation 2 (see supplementary material for details). Models were run in WinBUGS 1.4.1 (Available at: http://www.mrc-bsu.cam.ac.uk/bugs/) and Markov chain Monte Carlo (MCMC) (Smith and Roberts, 1993) simulation was used to estimate all parameters of the models. The WinBUGS code is available upon request from the corresponding author. Model convergence was assessed by visually inspecting the time series plot for each parameter and Gelman–Rubin statistic (Gelman and Rubin, 1992).

Sensitivity analysis. In the final step, we considered perturbations of the priors for all the parameters to see if the estimates were strongly dependent on the priors (Congdon, 2003).

Table 1
The descriptive statistics of the population examined by ELISA or Kato-Katz and the test results in 239 sampled villages in the third nationwide sampling survey of schistosomiasis in China, 2004

	Min	Q_1	Median	Q_3	Max	Total in all villages
Number of residents examined by ELISA	490	975	1,003	1,056	3,121	250,987
Number of positive ones detected by ELISA	0	48	87	168	1,766	30,680
Number of residents examined by Kato-Katz	0	40	82	168	1,448	28,908
Number of positive ones detected by Kato-Katz	0	0	1	6	262	2,624
Seroprevalence[a] (%)	0	4.51	8.55	16.60	56.58	12.22[b]

[a] Number of residents examined by ELISA divided by number of positive ones detected by ELISA in each village.
[b] Average seroprevalence (30,680/250,987).

3. Results

3.1. Data descriptions

In total, 239 villages were sampled (Fig. 1) and 250,987 residents were examined by ELISA. More than 94% (28,908/30,680) of the seropositives were examined by Kato-Katz test and among these, 2,624 egg-positives were found (see Table 1). This table also shows the minimum, lower quartile (25% percentile, Q_1), median, upper quartile (75% percentile, Q_3) and the maximum of the population examined by each test and the test results. There were 47.70% (114/239) of the villages with no infection identified by Kato-Katz test. The seroprevalences in the sampled villages differed from each other, ranging from 0 to 56.58% with the median of 8.55% (see Table 1 and Fig. 2).

3.2. Bayesian model establishment and sensitivity analysis

The prior distributions of both the sensitivity and specificity of each test used in the field obtained from expert knowledge were summarised in Table 2. A non-informative prior distribution for the prevalence of *S. japonicum* infection, beta(1,1), was used.

In WinBUGS, two chains were simulated in parallel, each with different initial values. The time series plot for each parameter and Gelman–Rubin statistic showed that convergence occurred within several thousand iterations (data not shown), thus a burn-in phase of 10,000 iterations was used and the models were run another 20,000 iterations for inference.

In order to find out how the priors influence the prevalence estimation, first we used the maximum ranges of expert knowledge on test properties, i.e. 80–100%, 70–95%, 20–80% and 85–100% for the sensitivities and specificities of ELISA and Kato-Katz tests, respectively, to obtain prior distributions. Again, a non-informative prior for the prevalence, beta(1,1), was used. The differences of the posterior medians of the prevalence of infection compared with those estimated with priors shown in Table 2 ranged from 0.05 to 6.07% with the median of 1.83% in situation 1 and from 0.10 to 5.04% with the median of 2.25% in situation 2. We then set the ranges as 70–100%, 65–100%, 15–85% and 75–100% for the sensitivities and specificities of ELISA and Kato-Katz tests, respectively, and found similar changes. Further enlarged ranges of the test properties resulted in much broader 95% credible intervals for the prevalences of *S. japonicum* infection.

3.3. Estimated prevalences of S. japonicum infection and comparison

Fig. 2 shows the estimated prevalence of *S. japonicum* infection (median with 95% credible intervals) for each village in situations 1 and 2, which were summarised in Table 3. The posterior medians of the prevalences in the 239 sampled villages might differ from each other dramatically, ranging from 0.02 to 55.68% with a median of 3.51% in situation 1, and from 0.02 to 55.52% with a median of 4.30% in situation 2. Most of the estimated prevalences were significantly smaller than the seroprevalences in both situations (the 95% credible intervals did not include the corresponding seroprevalences), especially when the seroprevalences were small (such as

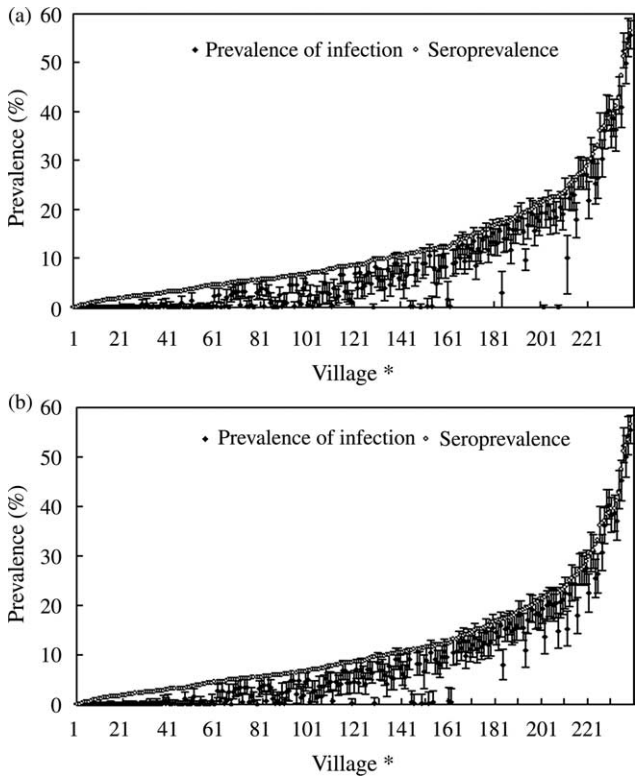

Fig. 2. Seroprevalence and estimated prevalence of infection (posterior median with 95% credible intervals) in each sampled village in the third nationwide sampling survey of schistosomiasis in China, 2004. (a) Prevalences of infection were estimated using data from both ELISA and Kato-Katz tests (situation 1). (b) Prevalences of infection were estimated using data from ELISA alone (situation 2). *Villages were sorted by the seroprevalence (from the minimum to the maximum).

Table 2
Expert knowledge and parameters of the beta(α, β) prior distributions for the sensitivities and specificities of ELISA and Kato-Katz tests used to detect *Schistosoma japonicum* infection in the field in China

Test	Sensitivity			Specificity		
	Range (%)	α	β	Range (%)	α	β
ELISA	90–95	409.775	33.225	85–90	611.625	87.375
Kato-Katz	20–70	6.678	8.162	95–100	151.125	3.875

Table 3
Bayesian estimates (posterior median) of prevalences of *S. japonicum* infection within villages, ecotypes and provinces, respectively, in the third nationwide sampling survey of schistosomiasis in China, 2004

	Min	Q_1	Median	Q_3	Max
Village prevalence (%)					
Situation 1[a]	0.02	0.12	3.51	12.43	55.68
Situation 2[b]	0.02	0.21	4.30	13.49	55.52
Ecotype prevalence (%)					
Situation 1[a]	0.19	0.91	1.25	2.30	5.94
Situation 2[b]	0.32	1.45	1.75	3.01	6.81
Province prevalence (%)					
Situation 1[a]	0.20	0.90	1.24	2.41	5.66
Situation 2[b]	0.37	1.42	1.74	2.98	6.40

[a] Using data from both ELISA and Kato-Katz tests.
[b] Using data from ELISA alone.

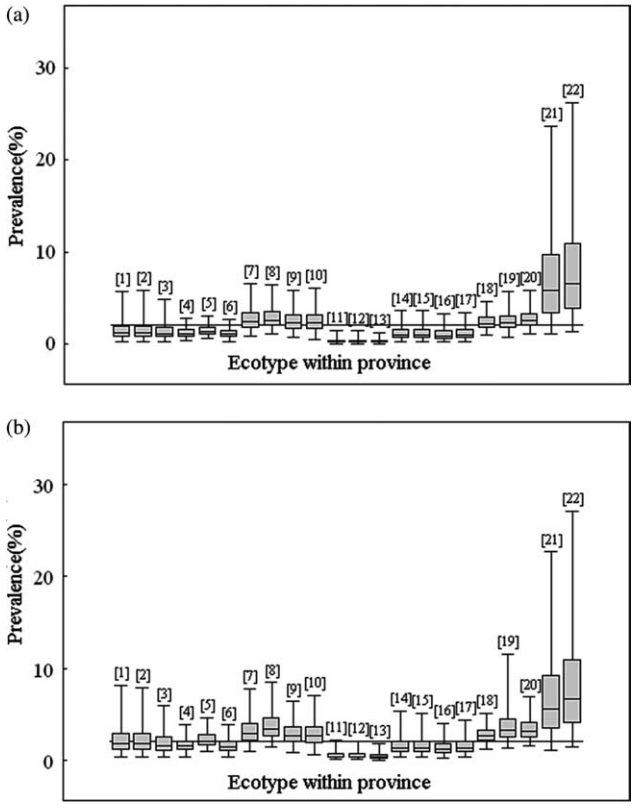

Fig. 3. Estimated prevalence of infection (posterior median, lower quartile (Q_1), upper quartile (Q_3) and 95% credible intervals) in each ecotype within provinces in the third nationwide sampling survey of schistosomiasis in China, 2004. (a) Prevalences of infection were estimated using data from both ELISA and Kato-Katz tests (situation 1). (b) Prevalences of infection were estimated using data from ELISA alone (situation 2). [1–3]: fork-beach, islet without embankment, and the hill within Anhui; [4–6]: islet without embankment, inner embankment and the hill within Hubei; [7–10]: fork-beach, islet with embankment, inner embankment and the hill within Hunan; [11–13]: islet without embankment, the plain regions with waterway networks and the hill within Jiangsu; [14–17]: fork-beach, islet without embankment, islet with embankment and the hill within Jiangxi; [18–20]: the plateau, mountain and hill within Sichuan; [20,21]: the plateau and mountain within Yunnan.

smaller than 10%). The estimated prevalences were similar in situations 1 and 2, while those in several villages in situation 2 are a bit higher than in situation 1.

Fig. 3 shows the estimated prevalence of *S. japonicum* infection (median with 95% credible intervals) for each ecotype within a province in situations 1 and 2 which were also summarised in Table 3. The estimates in ecotypes did not differ from each other significantly since the 95% credible intervals overlapped each other in both situations. However, the variations of the estimates were diverse in different ecotypes within provinces. The estimated prevalence of infection for each ecotype within a province in situation 1 was not significantly different from that in situation 2.

Fig. 4 shows the estimated prevalence of *S. japonicum* infection (median with 95% credible intervals) for each province in situations 1 and 2. Table 3 contains the summary information of the posterior medians. The prevalence estimates in provinces did not differ significantly from each other; while the variation of the prevalence estimates in Jiangsu province was the smallest, those for Jiangxi and Hubei provinces were the second smallest, and Yunnan province the biggest, about 10 times as large as the variation in Jiangsu province. The estimates for each province were not significantly different between situations 1 and 2 and the variations of estimates for each province in the two situations were similar.

4. Discussion

Schistosomiasis can be diagnosed by parasitological methods or immunological tests (Doenhoff et al., 2004). The Kato-Katz thick smear test is the most extensively used method for the detection of *S. japonicum* infection since it can provide information on both infection rate and intensity at the same time (Katz et al., 1972). However, its insensitivity limits its reliability in areas where the intensity of infection is low (Ross et al., 2001; Doenhoff et al., 2004). Its sensitivity can be improved by repeated stool collection and examination but this makes it more labour intensive and costly (Ross et al., 1998). Although promising and superior to many other immunological tests, antibody-based ELISA is an imperfect method for diagnosis of *S. japonicum* infection due to its low specificity (Wu, 2002; Doenhoff et al., 2004). The approach used in the 3rd NSC, namely that only those with positive ELISA tests would be examined by Kato-Katz, greatly reduced the time and manpower required but still could not compensate for the lack of sensitivity because of the inherent drawback of tests in series which are designed to increase specificity (Gardner et al., 2000; Christensen and Gardner, 2000). Thus, a proper statistical method was needed to precisely estimate community

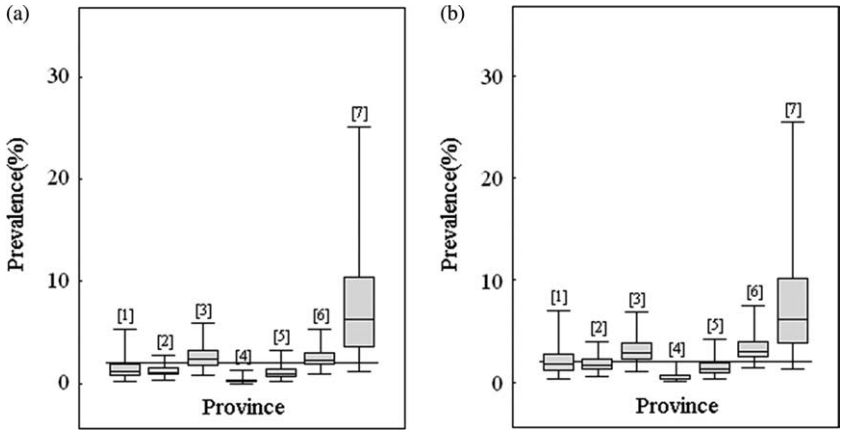

Fig. 4. Estimated prevalence of infection (posterior median, lower quartile (Q_1), upper quartile (Q_3) and 95% credible intervals) in each endemic province, in the third nationwide sampling survey of schistosomiasis in China, 2004. (a) Prevalences of infection were estimated using data from both ELISA and Kato-Katz tests (situation 1). (b) Prevalences of infection were estimated using data from ELISA alone (situation 2). [1–7]: Anhui, Hubei, Hunan, Jiangsu, Jiangxi, Sichuan and Yunnan provinces in sequence.

prevalence based on the data derived from the 3rd NSC while incorporating the test properties into the estimation.

Bayesian statistics have been successfully applied to estimate population prevalence and test properties when the true disease status was unknown (Joseph et al., 1995; Tu et al., 1999; Enoe et al., 2000; Black and Craig, 2002; Carabin et al., 2003, 2005; Orr et al., 2003; Dorny et al., 2004; Evans and Erlandson, 2004; Geurden et al., 2004). The following reasons favor the estimation of prevalence by Bayesian methods in the absence of a gold standard: firstly, traditional methods treat test properties as constants, thus inevitably underestimating the variability of prevalence estimates (Tu et al., 1999; Greiner and Gardner, 2000). Bayesian statistics allow us to incorporate the uncertainty about sensitivity and specificity into prevalence estimation (Tu et al., 1999), by treating them as variable parameters. Second, traditional methods use only the current data to make inferences, while Bayesian statistics incorporate prior information into the estimation, which improves estimation when substantive prior information exists (Carabin et al., 2005; Schurink et al., 2005). Based on the aforementioned reasons, we employed a Bayesian approach to estimate the community prevalences of *S. japonicum* infection based on the database from the 3rd NSC in order to achieve the original objectives of the 3rd NSC to understand the current epidemiological status and transmission trends of schistosomiasis and to provide a scientific basis for the formulation of further control strategies.

We only used the expert knowledge as prior information about test properties although data on test evaluations can also serve as prior information. This is because most published traditional test evaluations were based on the assumption that a certain parasitological test was the gold standard, which is not necessarily the case (Wu, 2002; Doenhoff et al., 2004), and thus might introduce selection bias and lead to misleading conclusions (Toft et al., 2005). Although, Branscum et al. (2004) and Mintiens et al. (2005) suggested the probability of disease freedom should be considered in the analysis, a non-informative beta prior for prevalence was assumed in this study for the following reasons: firstly, many more parameters would need to be estimated if the probability of disease freedom was taken into account, which aggravated the problem of non-identifiability discussed below and made the model unstable based on our data. Second, the probability that all the residents in a village were free from *S. japonicum* infection might be negligible since all the 239 villages were sampled from endemic villages.

Non-identifiability occurs when the number of parameters to be estimated in the model is larger than the degrees of freedom provided by the data, which was discussed by many authors such as Andersen (1997), Johnson et al. (2001) and Toft et al. (2005). This problem can be solved by assuming the test properties are the same in several populations (Johnson et al., 2001; Schaik et al., 2003; Branscum et al., 2004; Borel et al., 2004; Toft et al., 2005) when more than one population is involved, so that the degrees of freedom of data is at least equal to the number of parameters. Another possible way is to provide informative priors on some parameters (Dorny et al., 2004; Mintiens et al., 2005). The data for each village had two degrees of freedom at most in situation 1 and one degree of freedom in each village in situation 2 while many parameters needed to be estimated: prevalence in each village, test properties (sensitivity and specificity) and random effects (discussed below). We assumed the test properties were the same in all villages within each ecotype within each province to reduce the number of parameters and the prior distributions of test properties derived from the knowledge of experts who are veterans in the field, were informative. Hence the prevalence estimates were credible.

We also included random effects in our models to reflect the nested data structure that resulted from the stratified cluster sampling survey. Villages were considered to be directly nested in ecotype and the stratum of estimated prevalence level was discarded in the analysis since it was very common that

only one or two villages were nested in each part of this stratum and there was no between-village variation within this stratum or the between-village variation was very small.

Missing data is an inevitable phenomenon in a large-scale survey and may affect the precision of the estimation of community prevalence. Nearly 6% of the seropositives were not examined by Kato-Katz test for two reasons: firstly, there was an interval of about 1 week between blood and stool sample collection and some residents left during this period. Second, some residents were not willing to provide stool samples. However, the missing data hardly biased our results since its proportion was not big in such a large-scale survey and Bayesian statistics can handle this kind of missing data.

In this study, we did not attempt to estimate the test properties of ELISA or Kato-Katz. The sample size per village was not small, but Kato-Katz test was only applied to the seropositives and the residents examined by Kato-Katz were few in villages with very low prevalences. Meanwhile, we used informative priors for the test properties and allowed the prevalence estimation to be influenced by the priors to a certain degree (as revealed by sensitivity analysis). Therefore, it was not appropriate to estimate the test properties based on the data and the modeling.

We assumed two situations in this study and obtained similar prevalence estimates of infection in both situations (Figs. 2–4), suggesting it is feasible to employ only antibody-based ELISA instead of combined ELISA and Kato-Katz tests to estimate the prevalence of *S. japonicum* infection in large-scale epidemiological settings. The fact that Kato-Katz test was only applied to the seropositives, combined with its low sensitivity, means it made little contribution to the estimation. Meanwhile, that only blood samples are collected and examined may improve residents' compliance, reduce the required manpower and cost and decrease missing data. As a result we will consider using ELISA only in future investigations or surveillance of the prevalence of *S. japonicum* infection.

This study confirmed heterogeneity in the prevalence of *S. japonicum* infection in space (Zhou et al., 2004, 2005b) since the prevalence of *S. japonicum* infection was lower than 1% in many sampled villages whereas schistosomiasis was highly prevalent in some other villages (Fig. 2 and Table 3). The low prevalence status in Jiangsu province was encouraging, while the prevalence status might be very severe in some areas of Yunnan province. The prevalences were similarly higher in some areas of Hubei and Jiangxi provinces than those in Jiangsu province, but lower than those in some areas of Anhui, Hunan and Sichuan provinces (Figs. 3 and 4). Our study suggested that great achievements have been made in the control of schistosomiasis in China. Still, the disease remains a threat and control should be strengthened in some areas along the Yangtze River.

In conclusion, we have explored how a Bayesian approach can be used to obtain prevalence estimates of *S. japonicum* infection and we have shown that it is possible to use an antibody-based ELISA test alone to gain estimates similar to those using data from both ELISA and Kato-Katz tests in large-scale surveys.

Acknowledgements

This work received financial support from the National Natural Science Foundation of China (no. 30590373) and UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) (no. A30298). We thank the staff of the Department of Schistosomiasis Control, National Institute of Parasitic Diseases (IPD), the IPDs or CDCs of the seven endemic provinces for their excellent work on the data collection and management. Our sincere thanks go to the experts from the IPD, the IPDs or CDCs of the seven endemic provinces and from medical universities for their comments on test properties. We also thank the two very thoughtful referees and P. Steinmann from the Swiss Tropical Institute for comments and suggestions that greatly improved the article.

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijpara.2006.04.003.

References

Andersen, S., 1997. Re: 'Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard'. Am. J. Epidemiol. 145, 290–291.

Basanez, M.G., Marshall, C., Carabin, H., Gyorkos, T., Joseph, L., 2004. Bayesian statistics for parasitologists. Trends Parasitol. 20, 85–91.

Bergquist, N.R., 2002. Schistosomiasis: from risk assessment to control. Trends Parasitol. 18, 309–314.

Black, M.A., Craig, B.A., 2002. Estimating disease prevalence in the absence of a gold standard. Stat. Med. 21, 2653–2669.

Borel, N., Doherr, M.G., Vretou, E., Psarrou, E., Thoma, R., Pospischil, A., 2004. Seroprevalences for ovine enzootic abortion in Switzerland. Prev. Vet. Med. 65, 205–216.

Branscum, A.J., Gardner, I.A., Johnson, W.O., 2004. Bayesian modeling of animal- and herd-level prevalences. Prev. Vet. Med. 66, 101–112.

Carabin, H., Escalona, M., Marshall, C., Vivas-Martinez, S., Botto, C., Joseph, L., Basanez, M.G., 2003. Prediction of community prevalence of human onchocerciasis in the Amazonian onchocerciasis focus: Bayesian approach. Bull. World Health Organ. 81, 482–490.

Carabin, H., Balolong, E., Joseph, L., McGarvey, S.T., Johansen, M.V., Fernandez, T., Willingham, A.L., Olveda, R., Schistosomiasis Transmission And Ecology In The Philippines Step Project, 2005. Estimating sensitivity and specificity of a faecal examination method for *Schistosoma japonicum* infection in cats, dogs, water buffaloes, pigs, and rats in Western Samar and Sorsogon Provinces, The Philippines. Int. J. Parasit. 35, 1517–1524.

Chen, M.G., Feng, Z., 1999. Schistosomiasis control in China. Parasitol. Int. 48, 11–19.

Chen, X.Y., Wang, L.Y., Cai, J.M., Zhou, X.N., Zheng, J., Guo, J.G., Wu, X.H., Engels, D., Chen, M.G., 2005. Schistosomiasis control in China: the impact of a 10-year World Bank Loan Project (1992–2001). Bull. World Health Organ. 83, 43–48.

Christensen, J., Gardner, I.A., 2000. Herd-level interpretation of test results for epidemiologic studies of animal diseases. Prev. Vet. Med. 45, 83–106.

Congdon, P., 2003. Applied Bayesian Modelling. Wiley, England pp. 1–30.

Dendukuri, N., Joseph, L., 2001. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. Biometrics 57, 158–167.

Doenhoff, M.J., Chiodini, P.L., Hamilton, J.V., 2004. Specific and sensitive diagnosis of schistosome infection: can it be done with antibodies? Trends Parasitol. 20, 35–39.

Dorny, P., Phiri, I.K., Vercruysse, J., Gabriel, S., Willingham III., A.L., Brandt, J., Victor, B., Speybroeck, N., Berkvens, D., 2004. A Bayesian approach for estimating values for prevalence and diagnostic test characteristics of porcine cysticercosis. Int. J. Parasitol. 34, 569–576.

Enoe, C., Georgiadis, M.P., Johnson, W.O., 2000. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Prev. Vet. Med. 45, 61–81.

Evans, R.B., Erlandson, K., 2004. Robust Bayesian prediction of subject disease status and population prevalence using several similar diagnostic tests. Stat. Med. 23, 2227–2236.

Gardner, I.A., Stryhn, H., Lind, P., Collins, M.T., 2000. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. Prev. Vet. Med. 45, 107–122.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulations using multiple sequences. Statist. Sci. 7, 457–472.

Georgiadis, M.P., Johnson, W.O., Gardner, I.A., Singh, R., 2003. Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. Appl. Stat. 52, 63–76.

Geurden, T., Claerebout, E., Vercruysse, J., Berkvens, D., 2004. Estimation of diagnostic test characteristics and prevalence of *Giardia duodenalis* in dairy calves in Belgium using a Bayesian approach. Int. J. Parasitol. 34, 1121–1127.

Giboda, M., Bergquist, N.R., 1999. Post-transmission schistosomiasis. Parasitol. Today 15, 307–308.

Greiner, M., Gardner, I.A., 2000. Application of diagnostic tests in veterinary epidemiologic studies. Prev. Vet. Med. 45, 43–59.

Johnson, W.O., Gastwirth, J.L., Pearson, L.M., 2001. Screening without a 'gold standard': the Hui-Walter paradigm revisited. Am. J. Epidemiol. 153, 921–924.

Joseph, L., Gyorkos, T.W., Coupal, L., 1995. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am. J. Epidemiol. 141, 263–272.

Katz, N., Chaves, A., Pellegrino, J., 1972. A simple device for quantitative stool thick-smear technique in Schistosomiasis mansoni. Rev. Inst. Med. Trop. Sao Paulo 14, 397–400.

Li, Y.S., Sleigh, A.C., Ross, A.G., Williams, G.M., Tanner, M., McManus, D.P., 2000. Epidemiology of *Schistosoma japonicum* in China: morbidity and strategies for control in the Dongting Lake region. Int. J. Parasitol. 30, 273–281.

Li, Y.S., Zhao, Z.Y., Ellis, M., McManus, D.P., 2005. Applications and outcomes of periodic epidemiological surveys for schistosomiasis and related economic evaluation in the People's Republic of China. Acta Trop. 96, 266–275.

Mintiens, K., Verloo, D., Venot, E., Laevens, H., Dufey, J., Dewulf, J., Boelaert, F., Kerkhofs, P., Koenen, F., 2005. Estimating the probability of freedom of classical swine fever virus of the East-Belgium wild-boar population. Prev. Vet. Med. 70, 211–222.

Office of EDC, M.O.H., Expert Advisory Committee on Schistosomiasis, M.O.H., Institute of Parasitic Diseases, C.A.P.M., 1998. An introduction of a nationwide sampling survey on schistosomiasis in China. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi 16, 84–88.

Orr, K.A., O'Reilly, K.L., Scholl, D.T., 2003. Estimation of sensitivity and specificity of two diagnostics tests for bovine immunodeficiency virus using Bayesian techniques. Prev. Vet. Med. 61, 79–89.

Ross, A.G., Li, Y., Sleigh, A.C., Williams, G.M., McManus, D.P., 1998. Faecal egg aggregation in humans infected with *Schistosoma japonicum* in China. Acta Trop. 70, 205–210.

Ross, A.G., Sleigh, A.C., Li, Y., Davis, G.M., Williams, G.M., Jiang, Z., Feng, Z., McManus, D.P., 2001. Schistosomiasis in the People's Republic of China: prospects and challenges for the 21st century. Clin. Microbiol. Rev. 14, 270–295.

Schaik, G.V., Schukken, Y.H., Crainiceanu, C., Muskens, J., VanLeeuwen, J.A., 2003. Prevalence estimates for paratuberculosis adjusted for test variability using Bayesian analysis. Prev. Vet. Med. 60, 281–295.

Schurink, C.A., Lucas, P.J., Hoepelman, I.M., Bonten, M.J., 2005. Computer-assisted decision support for the diagnosis and treatment of infectious diseases in intensive care units. Lancet Infect. Dis. 5, 305–312.

Smith, A.F.M., Roberts, G.O., 1993. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. J. Roy. Statist. Soc. B. 55, 3–24.

Toft, N., Jorgensen, E., Hojsgaard, S., 2005. Diagnosing diagnostic tests: evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. Prev. Vet. Med. 68, 19–33.

Tu, X.M., Kowalski, J., Jia, G., 1999. Bayesian analysis of prevalence with covariates using simulation-based techniques: applications to HIV screening. Stat. Med. 18, 3059–3073.

Utzinger, J., Vounatsou, P., N'Goran, E.K., Tanner, M., Booth, M., 2002. Reduction in the prevalence and intensity of hookworm infections after praziquantel treatment for schistosomiasis infection. Int. J. Parasitol. 32, 759–765.

Wang, X.Y., Hong, X.L., Yan, J.L., Lu, R.G., Shao, D., Yang, S.G., Zhong, J.H., Li, Z.H., 2000. Study on cost-effectiveness of strategy of examination and therapy with IHA after mass questionnaire in middle-endemic areas of schistosomiasis. Chin. J. Schisto. Control 12, 277–279.

WHO workshop, 1980. Quantitative aspects of the epidemiology of *Schistosoma japonicum* infection in a rural community of Luzon, Philippines. Philippines. Bull. World Health Organ. 58, 629–638.

Willian, M.B., 2004. Introduction to Bayesian Statistics. Wiley, Hoboken, NJ pp. 129–146.

Wu, G.L., 2002. A historical perspective on the immunodiagnosis of schistosomiasis in China. Acta Trop. 82, 193–198.

Wu, W.L., Wang, J.R., Wen, L.Y., Huang, Y.Y., Xu, X.F., Yu, W.M., 2005. Surveillance and control of post-transmission schistosomiasis in Jiaxing prefecture, Zhejiang province, China. Acta Trop. 96, 282–287.

Yuan, H.C., 1995. Schistosomiasis control in China. Mem. Inst. Oswaldo Cruz 90, 297–301.

Zheng, J., 1993. A brief introduction to a nation-wide sampling survey on schistosomiasis. Chin. Med. J. (Engl.) 106, 569–575.

Zhou, X.N., Wang, T.P., Wang, L.Y., Guo, J.G., Yu, Q., Xu, J., Wang, R.B., Chen, Z., Jia, T.W., 2004. The current status of schistosomiasis epidemics in China. Zhonghua Liu Xing Bing Xue Za Zhi 25, 555–558.

Zhou, X.N., Chen, J.X., Chen, M.G., Bergquist, R., 2005a. The National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention: a new administrative structure for schistosomiasis control. Acta Trop. 96, 296–302.

Zhou, X.N., Wang, L.Y., Chen, M.G., Wu, X.H., Jiang, Q.W., Chen, X.Y., Zheng, J., Utzinger, J., 2005b. The public health significance and control of schistosomiasis in China—then and now. Acta Trop. 96, 97–105.

Zhu, Y.C., 2005. Immunodiagnosis and its role in schistosomiasis control in China: a review. Acta Trop. 96, 130–136.

Zhu, Y.C., Socheat, D., Bounlu, K., Liang, Y.S., Sinuon, M., Insisiengmay, S., He, W., Xu, M., Shi, W.Z., Bergquist, R., 2005. Application of dipstick dye immunoassay (DDIA) kit for the diagnosis of schistosomiasis mekongi. Acta Trop. 96, 137–141.