Use of Ultrasonography to Evaluate *Schistosoma japonicum*-Related Morbidity in Children, Sichuan Province, China, 2000–2007

Michelle S. Hsiang,* Elizabeth J. Carlton, Yi Zhang, Bo Zhong, Qiu Dongchuan, Pierre-Alain Cohen, Christopher C. Stewart, and Robert C. Spear

Global Health Sciences, University of California, San Francisco, California; School of Public Health, University of California, Berkeley, California; Institute of Parasitic Disease, Sichuan Center for Disease Control and Prevention, Chengdu, Sichuan, People’s Republic of China; Department of Radiology, University of California, San Francisco, California; Department of Pediatrics, University of California, San Francisco, California

Abstract. Liver ultrasonography is a convenient way to evaluate *Schistosoma japonicum*-related morbidity; however, no consensus standards exist, and data on use in Chinese children are scant. We describe 7 years of ultrasound findings in a prospective cohort of 151 children from an endemic area in Sichuan Province, China and evaluate technical aspects of the ultrasound methodology. Although prevalence of infection decreased over time, prevalence of hepatomegaly increased, which was likely caused by re-infections. The prevalence of late findings such as parenchymal fibrosis and splenomegaly were rare and did not increase over time; however, when present, they were associated with stunting. The use of adult thresholds versus height-adjusted standards underestimated pathology in children. Reliability of all measures except parenchymal grade was poor to fair. Our findings highlight the importance of early intervention and screening. We also suggest methodological refinements to improve reliability of ultrasonography for large-scale assessment of *S. japonicum*-related subclinical morbidity in children.

INTRODUCTION

More than 200 million people are afflicted with schistosomiasis worldwide, and 600 million people are at risk for infection.1 In China, it is estimated that over 730,000 are infected and 14 million people are at risk.2 Historically, schistosomiasis had a devastating effect on chronically infected patients, causing liver fibrosis with portal hypertension and frequently, resulting in life-threatening upper-gastrointestinal hemorrhage. Today, chemotherapy regimens as well as educational and environmental programs to control schistosomiasis are widespread in China. As a result, severe disease is less common, but fatigue, abdominal pain, and diarrhea are common as are anemia3 and growth retardation.4 Also, reductions in economic production5 and learning capacity6 related to schistosomiasis infection remain major problems.

Assessment of morbidity is important when evaluating the effectiveness of control programs. Since the 1970s, ultrasonography in community surveys has been shown to be a safe, rapid, and relatively inexpensive way to identify schistosomiasis-induced hepatic disease as well as monitor reversal after treatment.7 In 1992, the World Health Organization (WHO) proposed standards to detect and quantify pathology resulting from schistosomiasis infections, including *Schistosoma japonicum*.8 In 2000, the WHO published updated guidelines on the standardized use of ultrasonography for the assessment of schistosomiasis-related morbidity. However, morbidity caused by *S. japonicum* was not discussed in detail because of insufficient data. The guidelines for *S. japonicum* have yet to be updated.9

In Sichuan Province, China, there is ongoing transmission of *S. japonicum*.10 The basic epidemiology of schistosomiasis infection in rural Sichuan has been described. In 2000, the overall prevalence of infection in the area was 29%.11 Reemergence of infection has also been identified in previously well-controlled areas, likely because of weakened surveillance and environmental changes. Early re-emergent cases have occurred among young people, 6–25 years of age, presum-ably because young people lack the relative immunity that is acquired from repeated past exposures.12 In Sichuan Province, children represent a significant portion of the affected population,13 and in most parts of the world, children are disproportionately affected by schistosomiasis.13,14

Ultrasoundography is widely accepted as a safe and convenient way to evaluate schistosomiasis-related morbidity in affected communities. However, community-based longitudinal studies of *S. japonicum*-related ultrasound findings in Chinese children are rare.15 Therefore, the relative usefulness of the various hepatosplenic measures to reflect morbidity and infection history in the pediatric population is unclear. The long-term public health impact of chemotherapy programs on subclinical hepatosplenic morbidity in this population is also unknown. The goal of this study was to evaluate the usefulness of ultrasonography for community-based screening of *S. japonicum*-related subclinical morbidity in Chinese children. We performed a longitudinal study describing ultrasound findings from children living in a *S. japonicum*-endemic area of rural Sichuan Province. We also evaluated two technical aspects of the ultrasound methodology that have previously not been well-addressed. We compared the use of adult thresholds with the use of height-adjusted standards to identify morbidity, and we performed an inter-observer variance study to evaluate reliability of ultrasound measures.

MATERIALS AND METHODS

Study design and population. This longitudinal study was conducted from 2000 to 2007 in a population of children and young adults in Xichang County, Sichuan Province, China, a *S. japonicum*-endemic area where there have been ongoing chemotherapy campaigns. We describe the epidemiology of *S. japonicum*-related subclinical morbidity in this population as evidenced in ultrasound findings. We also compared two different standards to identify morbidity in children and assessed the reliability of ultrasound measures in an inter-observer variance study.

In 2000, residents 5 years of age and older were recruited to participate in an ultrasound survey. A random–number gener-
ator was used to select approximately 30% of the population, which was stratified by village and occupation. Children from 10 villages with higher infection prevalence were followed-up in 2002, 2005, and 2007. During the follow-up data collection, local schistosomiasis controllers contacted villagers by telephone and/or direct contact. Attempts were made to recruit all participants from the prior years for the new study. Subjects were included in the study if they were less than 18 years of age in 2000, resided in one of the selected 10 villages, and had at least one ultrasound within the study period. Informed consent was obtained from all participants, and the institutional review boards of Sichuan Province, China; University of California, Berkeley, CA; and University of California, San Francisco, CA all approved the study.

Demographics and other clinical data. Demographic data including age, sex, education, occupation, and ethnicity were obtained in 2000 and again in 2007. Self-reported history of alcohol use (categorized as “none,” “occasionally,” or “daily”) and hepatitis was also obtained in 2007.

Nutritional assessment. In 2002, height and weight measurements were conducted on 99 of 103 children who had ultrasound exams. In 2007, heights and weights were measured in all 83 children who had ultrasound exams. Subjects wore light clothing and no shoes. In 2002, weight was measured to the nearest 0.5 kg (Xiangshan BR9701 scale, Zhongshan Huibao Weighing Apparatus Co., Zhongshan, China). In 2007, weight was measured to the nearest 0.1 kg (Chengdu Henchichang CB335-65 scale, Wuhan Panscale Hardware & Electronics Co., Wuhan, China). Height was measured to the nearest centimeter with a stadiometer affixed to a wall. Height-for-age z-score (HAZ), BMI (calculated as weight in kilograms divided by height squared in meters), and BMI z-score (BMIZ) were calculated. A low BMIZ (< 2 SD below median reference) is the index of choice for the assessment of wasting or recent undernutrition, whereas a low HAZ (< 2 SD below median reference) represents stunting or chronic undernutrition. Currently available international growth standards do not account for ethnic and geographical diversity. Therefore, z-scores for each nutritional index were calculated using reference data from Chinese children. The purpose of including growth data in our study was to correlate it with infection and ultrasound findings. Therefore, to limit confounders associated with rapid, heterogeneous socioeconomic development in China, we used reference data from rural Sichuan Province.

Ultrasound examination. In 2000, 2002, and 2005, examination was performed using a Hitachi EUB 405 ultrasound machine with Hitachi EUP-C314T 3.5-mHz probe (Hitachi Medical Systems Co., Beijing, China). Each subject had one ultrasound study done, and all grading was done by the same experienced ultrasonographer in the field (Y.Z. has 12 years of experience performing ultrasonography for schistosomiasis-related hepatosplenic morbidity). In 2007, examination was done using a Medison Sonoline 5200 ultrasound machine with Hitachi EUP-C314T 3.5-mHz probe (Hitachi Medical Systems Co.). The two ultrasound machines used were of comparable quality, and imaging or interpretation should not have been affected by the use of a different machine in 2007. To evaluate for inter-observer variance, readings in 2007 were performed twice, one time by the ultrasonographer who performed all prior ultrasounds (Y.Z.) and a second time by another experienced ultrasonographer (Shifen Zhou has 16 years of ultrasonography experience, including 6 years in performing ultrasound for schistosomiasis-related hepatosplenic morbidity). Observers were blinded to each other’s results and to the infection status of the individual. In 2007, 10% of subjects also had their ultrasound repeated after at least 1 hour by Y.Z. to evaluate intra-observer variance. Readings by Shifen Zhou were only used for inter-observer variance analysis; all other outcomes analyzed were based on readings by Y.Z.

The ultrasound methodology included measurements described by the WHO 1992 Cairo White Paper and the study by Ca. Briefly, parenchymal fibrosis and perportal fibrosis were graded 0, 1, 2, or 3. Grade 2 and above was considered pathologic. Other measures were evaluated as height-adjusted SD scores (SDS; z-score) if reference data from a healthy Chinese reference population was available. Measurements were classified as abnormal if they were greater than two SDS above the mean measurement for the appropriate 0.2-m height bracket. For two measurements, spleen-vein diameter and portal-vein branch average diameter, reference data were not available, and morbidity was defined with standard cutoffs used widely in China. Standard cutoffs, most of which were recommended by the WHO in 1992, were also used for other measures in a separate analysis comparing morbidity defined by the two methods.

Height data were obtained in 125 children (82%) at least one time in 2002 and 2007, and 59 children have two height measurements. To impute heights for the years when height was not measured and for the children missing height measurements, an age-dependent (a) and sex-dependent (s) height model was generated using xtmixed in Stata 8.0. Comparing observed with predicted heights, the SD of the error was 0.022 m including a random intercept, 0.071 m, when the random intercept is 0. There were no differences between participants with and without height data in terms of sex, participation in the infection surveys, or infection status. Subjects with height data were younger, on average, than those without height data (mean age = 10.1 versus 12.1 years; P = 0.004).

Stool examination. S. japonicum infection was determined by stool examination at the end of the transmission seasons in 2000, 2002, and 2006 using methods described previously. In 2000 and 2006, three consecutive stool samples were collected from each subject, and in 2002, one stool sample was collected. All samples were analyzed by the miracidial hatch test. For each collection year, one sample per subject was analyzed by the Kato-Katz thick-smear procedure. Briefly, the Kato-Katz thick-smear procedure used three slides of homogenized stool (124.5 mg total) to determine presence and intensity of infection by measuring eggs per gram (EPG) of feces. If a sample was positive by miracidial hatching but negative by Kato-Katz, an EPG of 0 was used. Anyone testing positive with either method was classified as infected. Infected individuals were treated with praziquantel.

Data management and statistical analyses. All data were entered twice and analyzed using Stata 8.0 and Microsoft Excel (Microsoft Co., Redmond, WA). The relationships between each hepatosplenic measurement and infection data, nutritional status, and demographic characteristics were analyzed using xtmixed in Stata 8.0 (StataCorp, College Station, TX).
examined. Infection data were also analyzed in relation to demographics and nutritional status. Unadjusted analyses were conducted using a χ² test for dichotomous data, t test for continuous data, and ordinal logistic regression for ordinal data. EPG was transformed to the square root of EPG to normalize the skewed data. The Mann-Whitney U test was used for data that could not be normalized. For analyses of binary outcomes using data from multiple years, we accounted for correlations between repeated measurements using fixed effect logistic regression models with exchangeable correlation and robust inference (xtgee in Stata). For ordinal outcomes (liver grade and periportal fibrosis), ordinal logistic regression with the cluster option was used. P values of less than 0.05 were considered to be statistically significant.

All hepatosplenic measurements from 2007 were tested for inter-observer variance using the kappa statistic. Weighted κ statistic was used for the two ordinal measures: parenchymal fibrosis and periportal fibrosis. Reliability was defined as being almost perfect if κ > 0.81, substantial if between 0.61 and 0.80, moderate if between 0.41 and 0.60, fair if between 0.21 and 0.40, and slight to poor if between 0 and 0.20.²³ The variance of continuous measurements between observers relative to the total variance was also evaluated by calculating the intraclass correlation coefficient (ICC) from a random effects model (xtreg in Stata).²⁴ Ten percent of ultrasounds were read twice by one ultrasonographer (Y.Z.) in 2007 and tested for intra-observer variance using the same methods.

RESULTS

Recruitment and follow-up. Ultrasounds were performed in 145 children in 2000, 103 in 2002, 56 in 2005, and 83 in 2007. The total cohort included 151 children; 145 entered the study in 2000, and 6 entered the study in 2002. Of the 145 subjects who joined in 2000, 68 were lost to ultrasound follow-up: 26 in 2002, 19 in 2005, and 19 in 2007. Of the six subjects who joined in 2002, four were lost to ultrasound follow-up in 2005. At the final collection in 2007, reason for loss of ultrasound follow-up was determined: 4 had died (suicide for one and etiology of death in the other 3 was unknown), 19 moved away, 5 refused, and 40 could not participate because they were working. The attrition rate was lower for stool studies. Of the 151 subjects in the cohort, stool samples were obtained from 143 in 2000, 117 in 2005, and 120 in 2006. Subjects not lost to ultrasound follow-up did not differ significantly from subjects lost to ultrasound follow-up in terms of infection status (62.2% of retained participants were infected at least one time versus 55.1% of those who were lost to ultrasound follow-up; P = 0.167) and sex (55.4% versus 51.5% male; P = 0.628). Retained participants were slightly younger (mean age = 16.2 years; 95% CI = 15.5–16.9 years) compared with subjects lost to ultrasound follow-up (mean age = 17.6 years; 95% CI = 16.8–18.3 years).

Sampling. Table 1 presents descriptive characteristics of participants in the ultrasound cohort compared with children who were part of regular Xichang County schistosomiasis surveillance but did not participate in the ultrasound surveys. Sex, age, education, occupation, and infection status were not significantly different between the two groups.

Infection data. Over 6 years, the prevalence of positive infection status fell 79.5%, and the mean intensity of infection fell 99.0%. Percent of subjects infected decreased from 44.8% in 2000 to 27.4% in 2002 to 9.2% in 2006. Mean intensity of infection decreased from 47.9 EPG (range = 0–1180.7 EPG) in 2000 to 6.0 EPG (range = 0–313.3 EPG) in 2002 to 0.5 EPG (range = 0–16.1 EPG) in 2006. There were no relationships between infection status and age, sex, or farming as an occupation. Higher adult educational attainment in the child’s household, used as a marker of socioeconomic status, was marginally associated with infection status. Children living in households where no adults had completed schooling past elementary school had higher odds of infection compared with children living in households where at least one adult had attended middle school or higher (odds ratio (OR) = 1.61; 95% CI = 1.00–2.59; P = 0.054 adjusting for correlation within household).

Figure 1 presents follow-up percentages of infection, new infection, and re-infection from 2000–2006. Of those infected in 2000, 30.2% were re-infected in 2002; of those previously not infected, 25.4% were newly infected (P = 0.565). Infection in 2000 was not associated with infection in 2002. In 2006, everyone who was infected had been infected at least one time in 2000 or 2002. In looking at 119 subjects from whom infection data were collected all 3 years, the risk of infection in 2006 was significantly associated with the number of past infections (OR = 4.61; 95% CI = 1.57–13.52). By 2006, 50% of children in the cohort were infected at least one time.

Growth parameters. In 2002, the prevalence of wasting was 1.0% (BMI Z < 2 SD below median reference) and 6.6% showed stunting (HAZ < 2 SD below median reference). In 2007, the prevalence of wasting increased slightly to 2.4%, and the prevalence of stunting increased even more to 12.1%. Stunting in 2007 was found to be associated with a history of

![Figure 1](image-url)
2.8% at baseline in 2000 and then increased at each subsequent stage. No participants had grade 3 fibrosis (Table 2).

Consistent with the practice of many Chinese ultrasonographers, we considered grade 1 to be normal or very mild disease. No participants had grade 3 fibrosis (Table 2).

In contrast, prevalence of hepatomegaly and portal-vein fibrosis increased over time. Left liver lobes were enlarged in 2.8% at baseline in 2000 and then increased at each subsequent data collection time to 13.3% in 2007. Right liver lobe enlargement was identified in 10.3% of children in 2000, decreased to 1.9% in 2002 and 1.8% in 2005, and then increased to 21.7% in 2007. As for periporal fibrosis, prevalence of stage 2 disease was higher in 2007: 12.1% versus 0.7% in 2000 and 0% in 2002 and 2005 (Table 2).

Relationship between ultrasound findings and age, sex, reported alcohol use, and reported history of hepatitis. Bivariate analyses were performed to examine the relationship between age, sex, history of hepatitis, and history of alcohol use with each ultrasound measure. Older age was associated with a higher parenchymal grade (OR = 1.15; 95% CI = 1.00–1.31), enlarged left liver lobe (OR = 1.17; 95% CI = 1.08–1.26), and enlarged right liver lobe (OR = 1.13; 95% CI = 1.04–1.23).

Parenchymal fibrosis (defined as grade 2 or higher) was rare and stable over time; it was present in one participant in 2000 (0.75%), two in 2002 (1.9%), and two in 2007 (2.4%). Consistent with the practice of many Chinese ultrasonographers, we considered grade 1 to be normal or very mild disease. No participants had grade 3 fibrosis (Table 2).

Ultrasound findings. For quantitative measures with height-specific standards available, morbidity was defined as > 2 SD above the mean. In comparison, use of standard adult cutoffs severely underestimated morbidity for right liver lobe, left liver lobe, and portal–vein diameter. With spleen size, the difference was minimal (Table 3).

Parenchymal fibrosis (defined as grade 2 or higher) was rare and stable over time; it was present in one participant in 2000 (0.75%), two in 2002 (1.9%), and two in 2007 (2.4%). Consistent with the practice of many Chinese ultrasonographers, we considered grade 1 to be normal or very mild disease. No participants had grade 3 fibrosis (Table 2).

In contrast, prevalence of hepatomegaly and periporal fibrosis increased over time. Left liver lobes were enlarged in 2.8% at baseline in 2000 and then increased at each subsequent data collection time to 13.3% in 2007. Right liver lobe enlargement was identified in 10.3% of children in 2000, decreased to 1.9% in 2002 and 1.8% in 2005, and then increased to 21.7% in 2007. As for periporal fibrosis, prevalence of stage 2 disease was higher in 2007: 12.1% versus 0.7% in 2000 and 0% in 2002 and 2005 (Table 2).

Relationship between ultrasound findings and age, sex, reported alcohol use, and reported history of hepatitis. Bivariate analyses were performed to examine the relationship between age, sex, history of hepatitis, and history of alcohol use with each ultrasound measure. Older age was associated with a higher parenchymal grade (OR = 1.15; 95% CI = 1.00–1.31), enlarged left liver lobe (OR = 1.17; 95% CI = 1.08–1.26), and enlarged right liver lobe (OR = 1.13; 95% CI = 1.04–1.23).

Parenchymal fibrosis (defined as grade 2 or higher) was rare and stable over time; it was present in one participant in 2000 (0.75%), two in 2002 (1.9%), and two in 2007 (2.4%). Consistent with the practice of many Chinese ultrasonographers, we considered grade 1 to be normal or very mild disease. No participants had grade 3 fibrosis (Table 2).

In contrast, prevalence of hepatomegaly and periporal fibrosis increased over time. Left liver lobes were enlarged in 2.8% at baseline in 2000 and then increased at each subsequent data collection time to 13.3% in 2007. Right liver lobe enlargement was identified in 10.3% of children in 2000, decreased to 1.9% in 2002 and 1.8% in 2005, and then increased to 21.7% in 2007. As for periporal fibrosis, prevalence of stage 2 disease was higher in 2007: 12.1% versus 0.7% in 2000 and 0% in 2002 and 2005 (Table 2).

Relationship between ultrasound findings and age, sex, reported alcohol use, and reported history of hepatitis. Bivariate analyses were performed to examine the relationship between age, sex, history of hepatitis, and history of alcohol use with each ultrasound measure. Older age was associated with a higher parenchymal grade (OR = 1.15; 95% CI = 1.00–1.31), enlarged left liver lobe (OR = 1.17; 95% CI = 1.08–1.26), and enlarged right liver lobe (OR = 1.13; 95% CI = 1.04–1.23).
enlarged splenic vein diameter (OR = 1.12; 95% CI = 1.01–1.25), and periportal fibrosis (OR = 1.30; 95% CI = 1.21–1.40). Males were associated with higher parenchymal grade (OR = 8.45; 95% CI = 1.89–37.76). This difference in males compared with females could not be explained by differences in infection history, age, or farming as an occupation. Reported alcohol use in 2007 (“none” in 74.7%, “occasionally” in 21.7%, and “daily” in 3.6%) was associated with a larger portal vein (P = 0.015). Reported history of hepatitis in 2007 was 3.9% and was not associated with any ultrasound findings. All other relationships between ultrasound measures and descriptive characteristics were not significant.

Relationship between infection and ultrasound findings. The relationships between infection data (stool positivity and EPG) and each ultrasound measurement were analyzed. No significant relationships were found between current infection status and ultrasound findings. Given that *S. japonicum*-induced liver changes should follow infection, we also looked at the relationship between infection status and ultrasound findings obtained after the time that infection data were collected. In this analysis, only left liver lobe enlargement was found to be associated with past infection. Adjusting for stool positivity in 2006 and 2002, left liver lobe enlargement in 2007 was associated with stool positivity in 2000 (OR = 11.40; 95% CI = 1.31–99.53). Left liver lobe enlargement in 2007 was not associated with infection intensity in prior years. A trend was noted in the association between left liver–lobe enlargement in 2007 and increased frequency of past infection; however, the relationship was not significant (OR = 2.70; 95% CI = 0.98–7.50). No other associations between prior infection status and ultrasound findings were identified.

Inter- and intra-observer variance. Inter-observer variance analysis was performed on results from 2007 when readings from both ultrasonographers were available. Inter-observer agreement on parenchymal grade was good (weighted $\kappa = 0.79$), but $\kappa$ statistics for all other hepatosplenic measures were poor to fair (Table 4). Similar results were obtained using the ICC to assess inter-observer variance of continuous measurements relative to the total variance. ICC was poor to fair with all values; ICCs for periportal fibrosis and splenic vein diameter were especially poor (Table 4).

Intra-observer agreement was good ($\kappa$ statistic of 1.00 for parenchymal grade, 0.63 for left liver–lobe size, and 0.75 for right liver–lobe size). For second-order portal branch, portal vein diameter, spleen thickness, splenic vein diameter, and spleen length, there were too few rating categories for a $\kappa$ statistic to be calculated.

**DISCUSSION**

To summarize, prevalence and intensity of infection in our cohort decreased dramatically during the study period. Nonetheless, prevalence of infection at the final collection point was still high, and re-infections were common. The most notable ultrasound finding in our study was that, despite a decrease in infections, the prevalence of hepatomegaly increased. Also, left liver–lobe enlargement was significantly associated with prior infections. Late findings such as parenchymal fibrosis and splenomegaly were rare and stable over time, although they were associated with stunting when present. Reliability of all measures except parenchymal grade was poor to fair, and the use of adult thresholds versus height-adjusted standards underestimated pathology in children.

To assess the effectiveness of schistosomiasis-control programs on children, it is important to have reliable and useful measures of subclinical morbidity. Liver ultrasonography is widely accepted as a safe and convenient tool for this purpose. However, updated standards for *S. japonicum* are not available, and the most recent standards from 1992 do not address how the methodology can be applied to the pediatric population. Several longitudinal community-based surveys look at the use of ultrasonography to evaluate *S. japonicum*-related liver morbidity. Our study is unique in its long follow-up period and focus on a cohort of Chinese children living in an area where there have been large reductions in infection prevalence.

The observed reductions in infections reflect the success of widespread chemotherapy campaigns in China. Similar patterns have been seen in other parts of China after the use of chemotherapy. However, despite improving trends, the prevalence of infection in 2006 was still high at 9.2%. Furthermore, 9.2% may be an underestimation given that the sensitivity of Kato-Katz and hatch tests decline with decreased infection intensity. Control measures should be strengthened, and given that those infected in 2006 were more likely to have been previously infected, this population especially should be targeted. Modeling by Liang and others suggests that environmental modifications in concert with chemotherapy could be the best way to interrupt transmission and sustain control.

The most notable ultrasound finding in our study was that, despite a decrease in infection prevalence, the prevalence of hepatomegaly increased. In the natural progression of disease, hepatomegaly is followed by fibrosis and portal abnormalities. In our study, the prevalence of parenchymal fibrosis and portal abnormalities, including portal–vein enlargement, splenomegaly, and splenic vein enlargement, remained low and stable. This finding is consistent with other studies that suggest
that fibrosis and portal abnormalities reflect long standing pathology not yet evident in our population of young children and adults. Weist and others followed Chinese subjects over a 3-year period and found that hepatomegaly decreased in response to treatment, especially in younger individuals. However, a longer study by Olveda and others in the Philippines showed similar increases in hepatomegaly, despite treatment. Hepatomegaly decreased within 3–4 years of instituting treatment, but rebound hepatomegaly developed within 2 years of re-infection, despite continual population-based chemotherapy. At the start of our study, chemotherapy programs had been ongoing for almost a decade. We found an association between left liver-lobe enlargement and past infection. We also noted an association between left liver-lobe enlargement and the number of past infections. Although the finding was of borderline statistical significance, similar findings have been reported by others, and we conclude that re-infections are likely contributing to the rise in hepatomegaly in our population, independent of treatment. Another possibility is that prevalence of infection is being underestimated by the Kato-Katz and hatch tests and that undiagnosed low-intensity infections are contributing to the increasing prevalence of hepatomegaly.

There were no significant relationships found between any of the other ultrasound abnormalities and infection, past or active. In an analysis of data from the larger cohort in this area (578 subjects, including adults), parenchymal fibrosis was significantly associated with infection status. Our study may have been underpowered to detect the association between parenchymal fibrosis and infection status, because only 0.7–2.4% in our pediatric cohort had grade 2 disease (0.7–2.4% versus 18.5–21.8% in the larger cohort) and no subjects had grade 3 disease. Nonetheless, given that parenchymal fibrosis is rare in the pediatric population, left liver-lobe enlargement might be a more useful indicator of infection control in the Chinese pediatric community.

Prevalence of periportal fibrosis also increased from 0.7% in 2000 to 12.1% in 2007. The increased prevalence of periportal fibrosis versus parenchymal fibrosis could suggest that this measure is an early marker of liver involvement in children. However, measurement of periportal fibrosis is time consuming and difficult to perform. Reliability was very poor (ICC = 0.02). Also, in contrast to Schistosoma mansoni eggs, S. japonicum eggs are smaller and therefore, deposit in the periphery of the liver rather than in the periportal area. For these reasons, we do not recommend that this measure be included in the ultrasound methodology for assessment of S. japonicum-related liver morbidity in children. In the context of large-scale programs using ultrasonography to evaluate S. mansoni liver morbidity in Malian children, periportal fibrosis was not included, because the measure is time consuming and unreliable. Whether or not this recommendation applies to adults requires further study; it is possible that measurement of periportal fibrosis is more reliable in adults, because their second-order portal branches are larger and perhaps, easier to measure.

Growth parameters were obtained as another measure of possible S. japonicum-related morbidity. Despite the decrease in infection and a general improvement in economic status of the study participants, stunting increased over time. Compared with 2002, stunting increased from 2.4% to 12.1% in 2007. We considered whether or not cohort aging would have impacted this outcome. For example, there were no children 5–9 years of age in 2007, and delayed puberty could falsely overestimate nutritional deficits in 10–14 year olds. However, the average age of subjects in 2007 was 16.2 years (95% CI = 15.5–17.0) at which point most children will have completed puberty. Stunting is a multi-factorial process, and our study did not account for many other factors that can contribute to poor growth. Nonetheless, others have found that hepatic fibrosis, as a chronic inflammatory condition, may lead to growth stunting and wasting through systemically increased levels of pro-inflammatory cytokines. The likely association between chronic hepatosplenic changes and nutritional morbidity speaks to the importance of prevention, early detection, and early treatment of S. japonicum infection.

Last, we evaluated two technical aspects of the ultrasound methodology: we compared the use of adult thresholds with the use of height-adjusted standards to identify morbidity, and we performed an inter-observer variance study to evaluate reliability of ultrasound measures. To identify abnormal measures, height-adjusted standards were published by Li and others in 2004. These references have been used in research studies, but they are not routinely used in Chinese hospitals and in the field where ultrasonographers use standard adult cutoffs for children. To our knowledge, this is the first documented study that compares both standards using community-based data from Chinese children living in a S. japonicum-endemic area. With spleen thickness and spleen length, classification was similar. However, use of standard adult cutoffs in children was found to underestimate hepatomegaly and portal-vein enlargement (Table 3). For classification of S. mansoni-associated liver morbidity, the use of cutoffs based on a Senegalese reference population overestimated morbidity in Kenyan, Egyptian, and Malian children. To prevent over or underestimations of schistosomiasis-related liver morbidity, measures should be height-adjusted using reference measures from the appropriate population. Heights should be obtained on all subjects when performing ultrasonography.

Another major challenge to the use of ultrasonography is that inter-observer variation may be considerable. Although the WHO guidelines recommend that ultrasonography studies include tests of inter-observer variance, most studies ignore this recommendation. Reliability of all measures except parenchymal grade was poor to fair. Reliability of ultrasound measures also needs to be improved through better training and quality control. Also, the current ultrasound methodology used in Sichuan Province, China includes several redundant measures. A simplified methodology would make wide-scale screening less cumbersome and enable ultrasonographers to master fewer measures. For hepatomegaly, it would be ideal to only have one measure. In most parts of the world, liver size is measured at the right mid-axillary line. However, height-adjusted standards based on a Chinese reference population are not available for this measurement in China, where the standard measurements are the maximum oblique diameter (right liver lobe) or left sternal line (left liver lobe). Our study identified a significant relationship between left liver-lobe enlargement and history of infection. Furthermore, left liver-lobe hypertrophy and right liver-lobe atrophy have been ascribed to higher vascular flow in the left lobe. Therefore, we recommend left liver-lobe size as a single measure of S. japonicum-related hepatomegaly in China. As discussed above, periportal fibrosis is a measure that should not be included. To assess for
portal hypertension, we suggest that only portal–vein size and spleen size need to be measured. Only one measure of spleen size—either length or thickness—is necessary. We recommend that splenic vein size not be included in ultrasound methodology. Splenic vein size increases along with portal–vein size, and splenic vein measurement in our study had poor ICC. See Table 5 for a summary of recommendations on a simplified ultrasound methodology to evaluate *S. japonicum*-related hepatosplenic morbidity in Chinese children.

Although our study included an inter-observer variance study, it was performed only in 2007. Therefore, the longitudinal outcomes and associations reported are based on readings from a single ultrasonographer (Y.Z.) who performed readings all years. Given good inter-observer agreement on 2007 readings of parenchymal grade, we feel that the longitudinal outcomes and associations made for this important measure are valid. Regarding our finding that hepatomegaly increased despite decreased infections, which is likely caused by recurrent infections, this finding was also observed in our analysis of the larger cohort that included adults and has been documented by others. However, the reliability of this measure along with all other hepatosplenic measures was poor to fair (Table 4). To confirm our conclusions on these ultrasound measures, more studies are needed.

There were three other potential weaknesses in our study. First, although the attrition rate for infection data was reasonably low, we lost a significant number on ultrasound follow-up. Given our long follow-up period, participant retention was difficult. We compared subjects not lost to ultrasound follow-up to those who were lost to ultrasound follow-up in terms of age, sex, and infection status. Although retained participants were slightly younger, the difference is small and likely, clinically insignificant. There was no significant difference in terms of sex, and most importantly, there was no significant difference in terms of infection status, suggesting that loss to follow-up is unlikely to bias our findings.

Second, as discussed above, we may have underestimated infection prevalences, especially in later years. For each round of infection surveys, we examined three Kato-Katz slides and conducted hatch tests. Lin and others recommend using six Kato-Katz smears to ensure detection of infections when EPG infection prevalences, especially in later years. For each round of infection surveys, we examined three Kato-Katz slides and conducted hatch tests. Lin and others recommend using six Kato-Katz smears to ensure detection of infections when EPG were slightly younger, the difference is small and likely, clinically insignificant. There was no significant difference in terms of sex, and most importantly, there was no significant difference in terms of infection status, suggesting that loss to follow-up is unlikely to bias our findings.

Second, as discussed above, we may have underestimated infection prevalences, especially in later years. For each round of infection surveys, we examined three Kato-Katz slides and conducted hatch tests. Lin and others recommend using six Kato-Katz smears to ensure detection of infections when EPG were slightly younger, the difference is small and likely, clinically insignificant. There was no significant difference in terms of sex, and most importantly, there was no significant difference in terms of infection status, suggesting that loss to follow-up is unlikely to bias our findings.

Strengths of our study include the 7-year follow-up and relevance to the present day situation where communities have low transmission because of ongoing chemotherapy campaigns. Longitudinal community-based surveys looking at the use of ultrasonography to evaluate *S. japonicum*-related liver morbidity are few. Those that exist are from over a decade ago and do not have as long of a follow-up period as our study. Another strength of our study is its specific focus on Chinese children, which are an at-risk group that needs to be evaluated differently than adults because of their growth and development. Also, our study provides useful recommendations for schistosomiasis control and surveillance. The likely role of re-infection in subclinical liver morbidity points to the need to target this population. Also, we make concrete recommendations on a simplified ultrasound methodology. Most studies looking at *S. japonicum*-related hepatosplenic disease focus only on fibrosis and hepatomegaly. To our knowledge, our study is the only documented study that looks at all of the recommended ultrasound measures.

**Table 5**

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parenchymal grade</td>
<td>Recommended; parenchymal fibrosis is an important late finding, good reliability</td>
</tr>
<tr>
<td>Left liver lobe</td>
<td>Recommended; may be more useful than parenchymal grade as an indicator of infection control, reliability needs improvement</td>
</tr>
<tr>
<td>Right liver lobe (maximum oblique diameter)</td>
<td>Not recommended; poor reliability and unlike with left liver lobe, this study did not find an association with infection</td>
</tr>
<tr>
<td>Liver size (international measure at right mid-axillary line)</td>
<td>Not a standard measure in China and not evaluated in this study; studies may be warranted, Chinese height-adjusted standards not currently available</td>
</tr>
<tr>
<td>Portal vein branch average diameter (periportal fibrosis)</td>
<td>Not recommended; measure is time consuming and technically difficult to perform, poor reliability</td>
</tr>
<tr>
<td>Portal vein diameter</td>
<td>Recommended; portal–vein enlargement is an important late finding, reliability needs improvement</td>
</tr>
<tr>
<td>Spleen thickness</td>
<td>Recommended; select one measure because two measures are redundant, splenomegaly is an important late finding, reliability of both measures needs improvement</td>
</tr>
<tr>
<td>Spleen length</td>
<td>Not recommended; redundant to measures of portal–vein diameter and spleen size, poor reliability</td>
</tr>
<tr>
<td>Splenic vein diameter</td>
<td>Poor reliability</td>
</tr>
</tbody>
</table>
In conclusion, in our cohort of children from rural Sichuan province where there have been ongoing chemotherapy campaigns, we found a decrease in infections but an increase in the prevalence of hepatomegaly, which is likely caused by re-infections. Prevalence of late findings such as parenchymal fibrosis and some portal abnormalities did not increase. However, these abnormalities, when present, were associated with stunting, highlighting the importance of early intervention and screening. We also suggest methodological refinements to improve the reliability of ultrasonography for large-scale assessment of *S. japonicum*-related subclinical morbidity in children. As long as *S. japonicum* infection remains endemic, it will be important to develop reliable and useful measures of morbidity for children. Our study provides insight into the usefulness of ultrasonography as such a tool.

Received January 30, 2009. Accepted for publication September 28, 2009.

Acknowledgments: The authors thank the study participants and field staff in Xichang County for their participation and assistance on the project, Shifen Zhou for her assistance in performing ultrasounds for the inter-observer variance study, and Edmund Seto, John Swartzberg, and Sam C. Wang for their critical review of this manuscript.

Financial Support: This study was funded in part by a grant from the National Institutes of Health (R01 AI-50612-5) and support from Sichuan Centers for Disease Control and Prevention.

Authors’ addresses: Michelle S. Hsiang, Global Health Sciences, University of California, San Francisco, CA, E-mail: hsiangm@globalhealth.ucsf.edu. Elizabeth J. Carlton, School of Public Health, University of California, Berkeley, CA, E-mail: ejcarlton@berkeley.edu. Yi Zhang, Institute of Parasitic Disease, Sichuan Center for Disease Control and Prevention, Chengdu, Sichuan, People’s Republic of China, E-mail: zhangxxdept@163.com. Bo Zhong, Institute of Parasitic Disease, Sichuan Center for Disease Control and Prevention, Chengdu, Sichuan, People’s Republic of China, E-mail: zhongbo1968@163.com. Qiudong Chuan, Institute of Parasitic Disease, Sichuan Center for Disease Control and Prevention, Chengdu, Sichuan, People’s Republic of China, E-mail: qiudongchuan@163.com. Pierre-Alain Cohen, Department of Radiology, University of California, San Francisco General Hospital, San Francisco, CA, E-mail: PCohen@sfgghad.ucsf.edu. Christopher C. Stewart, Department of Pediatrics, University of California, San Francisco General Hospital, San Francisco, CA, E-mail: CStewart@sfgbpedics.ucsf.edu. Robert C. Spear, School of Public Health, University of California, Berkeley, CA, E-mail: spear@berkeley.edu.

Reprint requests: Michelle S. Hsiang, Global Health Sciences, University of California, 50 Beale Street, 12th Floor, San Francisco, CA 94105, Tel: 415-597-8180, Fax: 415-597-8299, E-mail: hsiangm@globalhealth.ucsf.edu.

REFERENCES

26. Li YS, Sleigh AC, Li Y, Tanner M, Dessein A, Williams GM, McManus DP, 2002. Five-year impact of repeated praziquantel...
ULTRASONOGRAPHY TO EVALUATE S. JAPONICUM MORBIDITY


