Unraveling Trends in Schistosomiasis: Deep Learning Insights into National Control Programs in China

Qing Su[1,2,3], Cici Xi Chen Bauer[4], Robert Bergquist[5], Zhiguo Cao[6], Fenghua Gao[6], Zhijie Zhang[1,2,7,*], Yi Hu[1,2,7,*]

[1] Department of Epidemiology and Biostatistics, School of Public Health, Fudan University, Shanghai 200032, China

[2] Key Laboratory of Public Health Safety, Ministry of Education, Shanghai 200032, China

[3] Xuhui District Center for Disease Control and Prevention, Shanghai 200032, China

[4] Department of Biostatistics and Data Science, University of Texas Health Science Center at Houston, 1200 Pressler Street, Houston, TX 77030, USA

[5] Ingerod, Brastad, Sweden

[6] Anhui Institute of Parasitic Diseases, Wuhu, People's Republic of China 230061, China

[7] Laboratory for Spatial Analysis and Modeling, School of Public Health, Fudan University, Shanghai 200032, China

* Corresponding author: Yi Hu (huyi@fudan.edu.cn); Zhijie Zhang (epistat@gmail.com)

Running title: Unraveling Schistosomiasis Trends by Deep Learning

Abstract

Background: To achieve the ambitious goal of eliminating schistosome infections, the Chinese government has implemented diverse control strategies. This study explored the progress of the 2 most recent national schistosomiasis control programs in an endemic area along the Yangtze River in China.

Methods: We obtained village-level parasitological data from cross-sectional surveys combined with environmental data in Anhui Province, China from 1997 to 2015. A convolutional neural network (CNN) based on a hierarchical integro-difference equation (IDE) framework (i.e., CNN-IDE) was used to model spatio-temporal variations in schistosomiasis. Two traditional models were also constructed for comparison with 2 evaluation indicators: the mean-squared prediction error (MSPE) and continuous ranked probability score (CRPS).

Results: The CNN-IDE model was the optimal model, with the lowest overall average MSPE of 0.04 and the CRPS of 0.19. From 1997 to 2011, the prevalence exhibited a notable trend: it increased steadily until peaking at 1.6 per 1000 in 2005, then gradually declined, stabilizing at a lower rate of approximately 0.6 per 1000 in 2006, and approaching zero by 2011. During this period, noticeable geographic disparities in schistosomiasis prevalence were observed; high-risk areas were initially dispersed, followed by contraction. Predictions for the period 2012 to 2015 demonstrated a consistent and uniform decrease.

Conclusions: The proposed CNN-IDE model captured the intricate and evolving dynamics of schistosomiasis prevalence, offering a promising alternative for future risk modeling of the disease. The comprehensive strategy is expected to help diminish schistosomiasis infection, emphasizing the necessity to continue implementing this strategy.

Key words: deep learning; schistosomiasis; spatio-temporal model; national schistosomiasis control programs

Introduction

Schistosomiasis is a chronic parasitic disease associated with poverty, caused by blood flukes belonging to the genus *Schistosoma* [1, 2]. Considered by the World Health Organization as a neglected tropical disease, schistosomiasis is mainly prevalent in low-resource tropical and

subtropical areas, with an estimated 140 million cases worldwide in 2019 [3]. Three species of *Schistosoma* are commonly seen in Asia, with *Schistosoma japonicum* the most prevalent, followed by *S. mekongi* and *S. malayensis* [4]. *S. japonicum* has a history of more than 2200 years in China [5]. Over the past 60 years, China has made significant strides in reducing the prevalence of schistosomiasis; currently, fewer than 29,000 people nationwide are estimated to be living with the disease, and only 5 new cases were reported in 2020 [6]. However, the risk of *S. japonicum* infection still exists in some areas of China, and the goal of eliminating the infection by 2030 remains a challenge [7].

Over the past 6 decades, China has made great strides toward reducing the prevalence of schistosomiasis. The World Bank Loan Project (WBLP) for schistosomiasis control and prevention from 1992 to 2001 made substantial progress; however, it focused on the treatment (e.g., praziquantel chemotherapy) [8] and not the transmission sources (e.g., intermediate snail host), and the epidemic rebounded following the end of the project. The integrated control strategy implemented in 2005, which focused on eliminating the intermediate snail host, again reduced the number of *S. japonicum* infections [9, 10], and the prevalence of schistosomiasis in the country has now stabilized at a historical low.

In order to better understand the transmission patterns and the temporal trend of schistosomiasis, previous studies employed numerous spatio-temporal models to estimate the infection risk and the contributing factors. Most studies have used Kriging [11, 12], regression techniques [13, 14], or linear dynamic models (e.g., integro-difference equations [IDE]) [15]. However, the dynamic transmission patterns of schistosomiasis are not fully captured by descriptive or linear models due to the complexity of the process and the multitude of contributing factors, which include environmental changes, human behavior, and evolving interventions. Therefore, in this study, we have adopted a deep learning (DL) approach with the aim of capturing the intricate and dynamic patterns of the disease. Our analysis provides detailed information on the annual prevalence of schistosomiasis across a grid of 1 × 1 km. This information is valuable for both researchers and local policymakers, enabling them to comprehend the evolution of

schistosomiasis distribution patterns under the 2 national schistosomiasis control programs (NSCPs) and to identify areas in urgent need of disease intervention.

Materials and Methods

Study area

The study was conducted at the village level in Anhui Province, located in eastern China, where the Yangtze River traverses the province (Figure 1). Anhui Province covers a geographic area of approximately 140,100 km² and had a population of nearly 65 million in 2019 (http://tjj.ah.gov.cn/). The province experiences a humid subtropical climate, with average annual temperatures ranging from 15°C to 16°C. Summers are typically hot, with temperatures frequently exceeding 30°C, while winters are cooler, with temperatures often falling below 10°C. The region receives a significant amount of rainfall, averaging between 1,000 and 1,500 mm annually, with the heaviest precipitation occurring in late spring and early summer. Humidity levels are generally high, often surpassing 80% during the rainy season. These climatic conditions—high temperatures, ample rainfall, and elevated humidity—create an ideal environment for the proliferation of *Oncomelania hupensis*, the freshwater snail that serves as the intermediate host for *S. japonicum*. Therefore, snail populations increase during the warmer and wetter months, which in turn affects the transmission dynamics of schistosomiasis in Anhui Province.

Figure 1 about here

Parasitological data

Prevalence data for schistosomiasis from 1997 to 2015 were obtained from cross-sectional surveys conducted by the Anhui Institute of Parasitic Diseases. In China, schistosomiasis is classified as a Class B notifiable infectious disease, ensuring that case detection at the village level is comprehensive, with a coverage rate of 100%. Upon identifying a schistosomiasis case, healthcare providers are required to complete an infectious disease report card within 24 hours of diagnosis for network reporting. Data collection occurred annually at the village level, targeting the population aged 5-65. A 2-tiered diagnostic approach was employed: all

participants were initially screened using the indirect hemagglutination assay [16], and positive readings were confirmed by Kato-Katz stool examinations [17] of all seropositive individuals. Individuals testing positive by both methods were diagnosed with *S. japonicum* infection. The study's annual selection of sample villages was systematically conducted in accordance with the "Control and Elimination of Schistosomiasis (GB15976 - 1995)" guidelines established by the National Health Commission of the People's Republic of China in 1995. Each year, villages categorized as level 1, with a human prevalence exceeding 5% in the previous year, were included in the study. In contrast, villages categorized as level 2, with a human prevalence below 5% in the previous year, were included every 2 years. Furthermore, villages categorized as level 3, indicating a human prevalence of less than 1% in the year before, were selected every 3 years. To ensure the accuracy of our disease prevalence calculations and maintain consistency with the methodologies of our previous studies [12, 18], sample villages with fewer than 100 participants were omitted to mitigate the impact of statistical outliers. Throughout the duration of the study, the number of sample villages varied between 1028 and 1683. Detailed enumerations are provided in the Supplementary Material (Table S1).

Environmental data

Considering the environmental factors that affect snail habitats and the growth and reproduction of *O. hupensis,* the intermedia host snail of *S. japonicum*, we included the following covariates: precipitation, hours of daylight, distance to water bodies, daytime land surface temperature (LST_{day}) and the normalized difference vegetation index (NDVI). We collected raster variable data for the corresponding years (1997-2015) for the areas included in the study, and the data for the distance to water bodies were from 2015. Average daily precipitation and hours of daylight were obtained from the China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn/home.do), from 610 weather stations across China. We performed Kriging interpolation to produce continuous overlays each year during the study period for all of China and then extracted the interpolated meteorological measures for Anhui Province using ArcGIS 10.5 (ESRI, Redlands, CA, USA). Water-body data were downloaded from the World Wildlife Fund's Conservation Science Data Sets (http://worldwildlife.org), and the Euclidean

distances between the geographic centroids of each sampled villages and water bodies were calculated. The NDVI and LST_{day} data were obtained from NASA's Level 1 and Atmosphere Archive and Distribution System (http://ladsweb.nascom.nasa.gov),which included 8-day 1-km^2 images for LST_{day} and the monthly 1-km^2 for NDVI. All the above data were processed using annual data and the raster data with a resolution of 1 km. The detailed sources of all the environmental data, as well as the resolution of time and space, are presented in the Supplementary Material (Table S2).

Statistical analysis

We treated the prevalence of schistosomiasis as a continuous variable and converted it to a Gaussian distribution, and the prevalence was transformed by the Box-Cox method [19] to satisfy the assumptions of a Gaussian model. We used univariate analysis for initial variable screening, and retained any variables with a p-value <0.1 [20]. We then assessed the correlations among the remaining variables, where correlation coefficients >0.6 indicated strong collinearity [21]. The 5 covariates we selected proved not to be colinear.

Convolutional neural network (CNN)

To evaluate the potential nonlinear spatio-temporal trends of schistosomiasis and the influence of environmental factors, we considered a convolutional neural network (CNN) based on an IDE framework (i.e., CNN-IDE) [22] to model the nonlinear trend. We employed a 2-level hierarchical structure [23] in the IDE model, with a data level and a process level. The former modeled the data generating mechanism, conditioned on the underlying spatial-temporal process and parameters, while the latter considered the unobserved process given by the parameters. Details on this CNN-IDE model can be found in the Supplementary Material.

Model comparison and validation

In addition to the CNN-IDE, we also implemented 2 spatio-temporal models (IDE and ST Kriging [7]) to estimate the risk of *S. japonicum* infection. The IDE model was the same as the IDE framework in CNN-IDE. Similarly, we constructed a hierarchical structure including a data

level and a process level. Detailed descriptions of the 2 models can be found in the Supplementary Material.

Cross-validation is used to evaluate model predictions by splitting the data into training and validation samples, then training the model with the training sample, and evaluating the model with the validation sample. In K-fold cross-validation, the available data are randomly divided into K folds. Each fold is excluded, the model is trained on the remaining K−1 folds, and then the model is evaluated on the initially excluded fold. Specifically, for k = 1..., K folds, the model is fit after removing the Kth fold, and the prediction result $\widehat{Z_v}^{(-k)}$ is obtained for i=1…m_k, where m_k is the number of data in the Kth fold. We conducted 10-fold cross-validation [23] in to evaluate the performance of the 3 models and to identify the optimal model. We used 2 evaluation indicators, the mean-squared prediction error (MSPE) and continuous ranked probability score (CRPS), which are defined as follows:

$$MSPE = \frac{1}{Tm}\sum_{j=1}^{T}\sum_{i}^{m}\{Z_v(s_i;t_j) - \widehat{Z_v}(s_i;t_j)\}^2 \quad (4)$$

$\{Z_v(s_i;t_j)\}$ represent observations of a randomly selected 10% of samples, and $\widehat{Z_v}(s_i;t_j)$ are predictions from modeling the rest of the observations.

$$CRPS = \int \left(1\{Z_v \le \widehat{Z_v}\} - F(\widehat{Z_v})\right)^2 d\widehat{Z_v} \quad (5)$$

Where $1\{Z_v \le \widehat{Z_v}\}$ indicates that if Z_v is less than x, the value is 1, otherwise 0, and $F()$ is the cumulative distribution function of the observed 10% of samples. Smaller values of MSPE and CRPS indicate better model performance. Our analyses were all done in R (version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria; http://cran.r-Project.org). A high-resolution spatial prediction (1 km^2) of schistosomiasis prevalence was mapped using the optimal model.

Results

As shown in Figure 2, the median annual prevalence of schistosomiasis began to increase in 1997 and reached a peak in 2005 (1.6 per 1000). There was a sharp decline to 0.6 per 1000 in 2006, and the prevalence continued to decrease rapidly, approaching zero after 2010. The mean

annual prevalence exhibited a similar pattern. The interquartile range widened from 0-0.8 (per 1000) in 1997 to 1.2-2.4 (per 1000) in 2005, then diminished to 0% by 2013.

Figure 2 about here

Figure 3 presents the results of model comparison. As shown in the figure, both the MSPE and CRPS for the CNN-IDE model were lower than those from the other 2 models for most of years. The overall average MSPE values of the CNN-IDE model, the IDE model, and ST Kriging model were 0.04, 0.05 and 0.06, respectively, and the overall CRPS values were 0.19, 0.22 and 0.25, respectively.

Figure 3 about here

Table 1 shows the final environmental covariates and parameters in the CNN-IDE model. The average daily precipitation (p=0.02) and NDVI (p <0.01) showed statistically significant positive associations with schistosomiasis prevalence, while LST_{day} (p <0.01), with longer hours of daylight (p <0.01), and distance to a water body (p <0.01) exhibited statistically significant negative associations. The estimate for the diffusion parameter $\theta_{p,1}$ was 2.31E+02, and those for the shift parameters $\theta_{p,2}$, and $\theta_{p,3}$ were 6.50E-03 and 2.45E-02, respectively.

Table 1. Estimations of parameters for schistosomiasis in the CNN-IDE model.

Parameters	Estimate	Standard error	z	p
Intercept	-2.76	1.23	-7.78	<0.01
Average daily precipitation	1.35E-05	1.63	2.26	0.02
Hours of daylight	-1.13E-03	0.42	-10.21	<0.01
Distance to water body	-8.44E-03	0.19	6.61	<0.01
LST_{day}	-1.17E-05	1.12	-7.24	<0.01
NDVI	4.16E-06	0.10	-6.71	<0.01
$\theta_{p,1}$ [a]	2.31E+02	10.11	1.89	0.21
$\theta_{p,2}$ [b]	6.50E-03	13.01	2.01	0.14

$\theta_{p,3}$ [d]	2.45E-02	10.92	3.92	0.04

CNN, convolutional neural network; IDE, integro-difference equation; LST_{day}, daytime land surface temperature; NDVI, normalized difference vegetation index.

[a] Diffusion parameter.

[b] Shift parameter.

A map of the annual predicted prevalence for schistosomiasis is displayed in Figure 4. Starting in 1997, the prevalence was relatively high and showed a gradual increase, as indicated by the expanding yellow areas and the occasional red spots. The epidemic reached its peak in 2005, characterized by extensive light yellow and small red areas. Following this peak, the prevalence began to decline and remained relatively stable at a low level from 2006 to 2011. The predictions, represented by a dark green shade, showed a consistent uniformity across the study area, with values nearing zero from 2012 to 2015. Figure 5 illustrates the standard error of the corresponding estimates, indicating that the values were higher in areas where the sample villages were less densely distributed. However, the standard errors remained low throughout the study period.

Figures 4 and 5 about here

Discussion

Our study presents a comprehensive application of advanced DL methods for quantifying local trends of schistosomiasis prevalence to assess the effectiveness of 2 NSCPs in the Yangtze River Basin, China. These estimates highlight substantial differences within the study area in levels and trends. The annual predicted prevalence map illustrates the disease's progression, showing fluctuating trends until a relatively stable low level after 2011. These findings contribute to our understanding of schistosomiasis dynamics and control strategies currently implemented in China.

The environmental factors that affect snail habitats and the growth and reproduction of the snails have been confirmed in previous studies [12, 24-26]. A study found that an ideal snail habitat is distributed less than 1 km from the water source [24]. In line with this, our study

found a negative association between the proximity to water bodies and the prevalence of schistosomiasis. Increased rainfall can facilitate the dispersal of snails to new areas, including rivers, lakes, and wetlands [4]. The optimal survival temperature for the eggs of the parasite varies between 16°C and 35°C [27], a fact that supports the transmission of *S. japonicum* in snails that reproduce and grow under conditions of higher LST_{day} and therefore longer hours of daylight. Vegetation could reduce solar radiation and regulate the water temperature for host snails, thus providing comfortable spawning shelters [25]. As a result, a higher NDVI is more conducive to snail survival.

The results of model comparison indicated that our CNN-IDE was the optimal model that best accounted for the spatio-temporal variation in schistosomiasis prevalence. Figure 3 shows its superiority in predicting the schistosomiasis prevalence data, possibly because the complexity of the dynamics presented in a latent process can be captured flexibly if a sufficient number of parameters is available, and the CNN was trained on a massive amount of spatial data to obtain this [22]. Deep neural nets, especially CNNs, contain the necessary structure to harness this complexity. Furthermore, CNNs offer a global prior model for the dynamics that is both realistic and computationally efficient [22].

We found that the distribution patterns of schistosomiasis infection varied over space and time throughout the study period. These shifts in pattern can likely be attributed to changes in control strategies implemented at different stages of schistosomiasis management, potentially leading to a nonlinear dynamic process in the prevalence of the disease [15, 28]. The 10-year WBLP, launched in 1992, has effectively facilitated the praziquantel chemotherapy strategy [10]. Figure 4 shows that the predictions indicated schistosomiasis was maintained at a relatively stable and low level until 2001, after which there was a resurgence following the conclusion of the WBLP. This resurgence may be due to the fact that chemotherapy measures were limited to bovines and humans. Considering that over 40 species of mammals can act as potential zoonotic reservoirs, these measures are insufficient to completely interrupt the life cycle of the parasite [29]. Another possible explanation for the rebound is environmental and social factors,

such as the floods of the Yangtze River in the late 1990s, ecological changes, and population movements [30, 31].

To address this issue, a comprehensive national control strategy was put into place in 2005, encompassing agricultural mechanization, improved sanitation, and health education [32, 33]. This strategy has been effective, leading to positive changes in both environmental conditions and human behaviors. Socioeconomic improvements, including better access to healthcare, enhanced sanitation facilities, and changes in water-related activities, have contributed to the reduction in disease transmission [34]. Furthermore, modifications in agricultural practices and water management may have impacted snail habitats [35]. The success of this comprehensive control strategy is evident based on the reduced number of schistosomiasis foci from 2006 to 2011, as shown in Figure 4. The sustained low prevalence of the disease from 2012 to 2015 further demonstrates the effectiveness and durability of these measures, underscoring the importance of continuing with this strategy.

Our CNN-IDE model has proven its superiority over traditional models in capturing the intricate spatio-temporal variations of schistosomiasis prevalence. This success is rooted in the model's flexible ability to comprehend latent process complexities through an ample parameter set, bolstered by extensive training on spatial data using a CNN [24]. The CNN's inherent ability to process complex multilayered information enables it to accurately represent the dynamics of schistosomiasis. This advantage highlights the potential of our model as a promising tool for future spatio-temporal risk modeling in schistosomiasis, contributing to the advancement of precision public health methodologies.

The limitations of the study need to be discussed. First, the low prevalence of schistosomiasis might result from the suboptimal specificity of serological analysis and sensitivity of stool examinations. It will be necessary to consider diagnostic errors in future modeling research to improve prediction accuracy. Second, we only considered a limited number of environmental factors and socioeconomic factors, such as household financial situations and medical

conditions, were not included because these covariates were not available at the village level. Nevertheless, the random effect, $\eta_t(s)$, as shown in Supplementary Material Equation (2), is the "residual" after discounting what the covariates explain [36]. This suggests that covariates not included in our analysis are left in this "residual." Third, we obtained data on water bodies in 2015, since historical and updated data on tributaries of the Yangtze River were not available.

In summary, the proposed CNN-IDE model effectively captured the complex dynamic process of schistosomiasis prevalence. The high-resolution 1 × 1-km grid-level maps in our study facilitate the quantification of inequalities in prevalence to guide the efficient deployment of resources and interventions to those with the greatest need. As researchers, policymakers, and program implementors need to work together to achieve schistosomiasis elimination, our study provides a precision tool, guiding them where to go next.

Ethics approval

The collection of parasitological data was part of a continuing public health investigation determined by the National Health and Family Planning Commission. Hence, this study was exempt from institutional review board assessment.

Competing Interests

We declare that we have no conflicts of interest.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request and from a website repository (http://cdc.cma.gov.cn/home.do, http://ladsweb.nascom.nasa.gov, http://worldwildlife.org).

Funding

This work is primarily being funded by the National Natural Science Foundation of China (81773487,81973102).

Acknowledgements

We would like to express gratitude to all the people who participated in data collection, as well as to our teachers for guiding and revising the paper.

References

1. King CH, Dickman K, Tisch DJ. Reassessment of the cost of chronic helmintic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis. Lancet 2005; 365:1561-1569.
2. WHO. Control of Neglected Tropical Diseases. World Health Organization; 2021 [cited 2022 Apr 17]. Available from:https://www.who.int/health-topics/schistosomiasis.
3. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018; 392:1789-1858.
4. Gordon CA, Kurscheid J, Williams GM, Clements ACA, Li Y, Zhou XN, et al. Asian Schistosomiasis: Current Status and Prospects for Control Leading to Elimination. Trop Med Infect Dis 2019; 4
5. Shou-Pai M, Bao-Ruo S. Schistosomiasis control in the People's Republic of China. The American journal of tropical medicine and hygiene 1982; 31:92-99.
6. Zhang LJ, Xu ZM, Yang F, Dang H, Li YL, Lü S, et al. Endemic status of schistosomiasis in People's Republic of China in 2020. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi 2021; 33:225-233.
7. Heuvelink GB, Griffith DA. Space–Time Geostatistics for Geography: A Case Study of Radiation Monitoring Across Parts of Germany. . Geographical analysis 2010; 42:161-179.
8. Xianyi C, Liying W, Jiming C, Xiaonong Z, Jiang Z, Jiagang G, et al. Schistosomiasis control in China: the impact of a 10-year World Bank Loan Project (1992-2001). Bull World Health Organ 2005; 83:43-48.
9. Collins C, Xu J, Tang S. Schistosomiasis control and the health system in P.R. China. Infect Dis Poverty 2012; 1:8.
10. Liu R, Dong HF, Jiang MS. The new national integrated strategy emphasizing infection sources control for schistosomiasis control in China has made remarkable achievements. Parasitol Res 2013; 112:1483-1491.
11. Hu Y, Gao J, Chi M, Luo C, Lynn H, Sun L, et al. Spatio-temporal patterns of schistosomiasis japonica in lake and marshland areas in China: the effect of snail habitats. Am J Trop Med Hyg 2014; 91:547-554.
12. Hu Y, Li R, Bergquist R, Lynn H, Gao F, Wang Q, et al. Spatio-temporal transmission and environmental determinants of Schistosomiasis Japonica in Anhui Province, China. PLoS Negl Trop Dis 2015; 9:e0003470.
13. Hu Y, Zhang Z, Chen Y, Wang Z, Gao J, Tao B, et al. Spatial pattern of schistosomiasis in Xingzi, Jiangxi Province, China: the effects of environmental factors. Parasit Vectors 2013; 6:214.
14. Zhang ZJ, Carpenter TE, Lynn HS, Chen Y, Bivand R, Clark AB, et al. Location of active transmission sites of Schistosoma japonicum in lake and marshland regions in China. Parasitology 2009; 136:737-746.

15. Hu Y, Bergquist R, Chen Y, Ke Y, Dai J, He Z, et al. Dynamic evolution of schistosomiasis distribution under different control strategies: Results from surveillance covering 1991-2014 in Guichi, China. PLoS Negl Trop Dis 2021; 15:e0008976.
16. Miyashita K, Itoh H, Nakao K. [Idiopathic hyperaldosteronism (IHA)]. Nihon Rinsho 2006; Suppl 1:624-627.
17. Yu JM, de Vlas SJ, Jiang QW, Gryseels B. Comparison of the Kato-Katz technique, hatching test and indirect hemagglutination assay (IHA) for the diagnosis of Schistosoma japonicum infection in China. Parasitol Int 2007; 56:45-49.
18. Su Q, Bergquist R, Ke Y, Dai J, He Z, Gao F, et al. A comparison of modelling the spatio-temporal pattern of disease: a case study of schistosomiasis japonica in Anhui Province, China. Trans R Soc Trop Med Hyg 2021;
19. Box GE, Cox DR. An analysis of transformations. J R Stat Soc: Series B (Methodological) 1964; 26:211-243.
20. Diggle PJ, Tawn JA, Moyeed RA. Model□based geostatistics. J R Stat Soc: Series C (Applied Statistics) 1998; 47:299-350.
21. Hu Y, Xiong C-L, Zhang Z-J, Bergquist R, Wang Z-L, Gao J, et al. Comparison of data-fitting models for schistosomiasis: a case study in Xingzi, China. Geospatial health 2013:125-132.
22. Mangion AZ, Wikle CK. Deep Integro-Difference Equation Models for Spatio-Temporal Forecasting. CoRR, vol. abs/1910.13524 2019;
23. Wikle CK, Zammit-Mangion A, Cressie N. Spatio-temporal statistics with R: CRC Press; 2019, p. 258-266.
24. Zhu HR, Liu L, Zhou XN, Yang GJ. Ecological Model to Predict Potential Habitats of Oncomelania hupensis, the Intermediate Host of Schistosoma japonicum in the Mountainous Regions, China. PLoS Negl Trop Dis 2015; 9:e0004028.
25. Ajakaye OG, Adedeji OI, Ajayi PO. Modeling the risk of transmission of schistosomiasis in Akure North Local Government Area of Ondo State, Nigeria using satellite derived environmental data. PLoS Negl Trop Dis 2017; 11:e0005733.
26. Gao F, Ward MP, Wang Y, Zhang Z, Hu Y. Implications from assessing environmental effects on spatio-temporal pattern of schistosomiasis in the Yangtze Basin, China. Geospat Health 2018; 13
27. Pflüger W. Experimental epidemiology of schistosomiasis. Zeitschrift für Parasitenkunde 1980; 63:159-169.
28. Hu Y, Xiong C, Zhang Z, Luo C, Ward M, Gao J, et al. Dynamics of spatial clustering of schistosomiasis in the Yangtze River Valley at the end of and following the World Bank Loan Project. Parasitol Int 2014; 63:500-505.
29. Zhou YB, Liang S, Jiang QW. Factors impacting on progress towards elimination of transmission of schistosomiasis japonica in China. Parasit Vectors 2012; 5:275.
30. Zhou XN, Guo JG, Wu XH, Jiang QW, Zheng J, Dang H, et al. Epidemiology of schistosomiasis in the People's Republic of China, 2004. Emerg Infect Dis 2007; 13:1470-1476.
31. Wang L, Utzinger J, Zhou XN. Schistosomiasis control: experiences and lessons from China. Lancet 2008; 372:1793-1795.
32. Wang LD, Guo JG, Wu XH, Chen HG, Wang TP, Zhu SP, et al. China's new strategy to block Schistosoma japonicum transmission: experiences and impact beyond schistosomiasis. Trop Med Int Health 2009; 14:1475-1483.
33. Wang LD, Chen HG, Guo JG, Zeng XJ, Hong XL, Xiong JJ, et al. A strategy to control transmission of Schistosoma japonicum in China. N Engl J Med 2009; 360:121-128.
34. Ben-Jiao H, Hong-Ling X, Sheng-Ming L, Zheng-Yuan Z, Yi-Biao Z, Zhi-Hong L, et al. [Measures and achievements of schistosomiasis control in the Yangtze River Basin]. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi 2018; 30:592-595.
35. Guo S, Dang H, Li Y, Zhang L, Yang F, He J, et al. Sentinel Surveillance of Schistosomiasis - China, 2021.

China CDC Wkly 2023; 5:278-282.
36. Wikle CK, Zammit-Mangion A, Cressie N. Spatio-temporal statistics with R: CRC Press; 2019, p. 210-226.

Figure legend

Figure 1. Endemic areas of schistosomiasis in Anhui Province, China. Anhui Province is located in the lower reaches of the Yangtze River in eastern China.

Figure 2. Box plot for the observed prevalence of schistosomiasis in sample villages in Anhui, China, from 1997 to 2015. The blue line represents the average annual prevalence, and the red points are the median annual prevalence. The boxes denote minimum, median, maximum, and interquartile ranges.

Figure 3. Average mean squared prediction error (MSPE) (top) and continuous ranked probability score (CRPS) (bottom) of the CNN-IDE predictions (red), the IDE predictions (green) and the ST Kriging predictions (blue) as a function of time. CNN, convolutional neural network; IDE, integro-difference equation.

Figure 4. Plot for the annual prevalence of schistosomiasis predicted from 1997 to 2015 in Anhui Province, China.

Figure 5. Plot for the annual standard error of the predicted prevalence of schistosomiasis from 1997 to 2015 in Anhui Province, China.