CHAPTER TEN

Epidemiological big data and analytical tools applied in the control programmes on parasitic diseases in China: NIPD's sustained contributions in 70 years

Shang Xia[a,b,c,d,e], Jin-Xin Zheng[a,b,c,d,e], Xin-Yi Wang[a,b,c,d,e], Jing-Bo Xue[a,b,c,d,e], Jian-Hong Hu[a,b,c,d,e], Xue-Qiang Zhang[a,b,c,d,e], Xiao-Nong Zhou[a,b,c,d,e], Shi-Zhu Li[a,*]

[a]National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, Shanghai, People's Republic of China
[b]Chinese Center for Tropical Diseases Research, Shanghai, People's Republic of China
[c]Key Laboratory of Parasite and Vector Biology, Ministry of Health, Shanghai, People's Republic of China
[d]National Health Commission of China, Shanghai, People's Republic of China
[e]WHO Collaborating Centre for Tropical Diseases, Shanghai, People's Republic of China
*Corresponding author: e-mail address: lisz@chinacdc.cn

Contents

1. Introduction 320
2. Empirical analysis 322
   2.1 Statistical methods for analysing the causative factors and disease incidences 323
   2.2 Descriptive methods for characterizing the temporal and spatial disease distribution 325
   2.3 Experimental methods for test epidemiological hypotheses 328
3. Theoretical and mathematical analysis 333
   3.1 Theoretical modelling for analysing the dynamics of disease transmission 333
   3.2 Mathematical analysis for estimated the effectiveness of diseases intervention measures 334
4. Computational analysis 335
   4.1 Computational modelling for analysing the dynamics of disease transmission 335
   4.2 Data mining and machine learning 337
   4.3 Systems thinking for disease control 339

Advances in Parasitology, Volume 110
ISSN 0065-308X
https://doi.org/10.1016/bs.apar.2020.04.009

© 2020 Elsevier Ltd
All rights reserved.

5. Remaining challenges and future opportunities 339
5.1 Factors influencing diseases are rooted in multiple sources and at multiple scales 342
5.2 Complexity in epidemiological system of parasitic diseases needs integrated more factors in the modelling 342
5.3 Systems thinking plays an emerging important role for the prediction and control of parasitic diseases 343
6. Conclusion 344
References 344

Abstract

The analysis of epidemiological data has played an important role for the academic research carried out by the National Institute of Parasitic Diseases, China CDC, since its foundation in 1950s. Those researches, e.g., the temporal-spatial patterns of disease transmission and the identification of risk factors, have contributed significantly to the national parasitic disease control and elimination programmes in China. With the development and application of epidemiological data analysis in the last decade, all research results improve our understanding of parasitic diseases epidemiology and related health issues through the application platform of epidemiological big data and analytical tools. In particular, implementation research on analytical predictions on disease outbreak or epidemic risks have provided references to the scientific guidance on effective preventions and interventions in the parasitic disease elimination in China, such as fliariasis, malaria and schistosomiasis. This review has reflected the function of data accumulation and application of temporospatial tools in parasitic diseases control, and the ways of the NIPD's sustained contributions to the disease control programmes in China.

1. Introduction

Epidemiology serves as the cornerstone of public health. According to the definition of World Health Organization (WHO), epidemiology is "the study of the distribution and determinants of health-related states or events (including disease), and the application of this study to the control of diseases and other health problems". Since the main tasks for the National Institute of Parasitic Diseases at China CDC (NIPD) is to organize or coordinate national parasitic diseases control efforts with more works on surveillance, scientific research on parasitic diseases, emergency response to outbreaks of parasitic and tropical diseases, services on detection and consultation on parasite infections, and international cooperation, etc. Therefore, the epidemiological investigation and field surveillance are two major components for the NIPD's routine work, with supported by big databases and

analytical tools. In this regards, epidemiological studies carried out in NIPD mainly focus on the transmission patterns of parasitic and tropical diseases, e.g., the temporal-spatial patterns of parasitic disease transmission and the identification of risk factors, etc.

The major epidemiological researches in NIPD are driven by the problem-solving approach, particularly on development of new control strategy, application of modern tools in surveillance and response to outbreaks, and providing the updated references as baseline data to the decision making for formulation of 5-year's working plan or standards in the control programmes. All of these epidemiological studies are only enabled by the confluence of data from various sources, including (i) surveillance data, e.g., records of when, where, and who has been infected, (ii) biological data of parasites and etiological data of disease vectors, e.g., mosquitos biting rate in malaria, data on variables of parasites' lifecycle, (iii) environmental data, e.g., temperature, precipitation, vegetation, land usage, (iv) socio-economic data, e.g., income level or poverty level, and (v) population's behavioural data, e.g., host population mobile and migrant activities. The variety of epidemiological data will break through the exiting scientific research paradigms, enable a new generation of scientific research. For instance, the goal of understanding the impacts of climate change on transmission of schistosomiasis transmission is a driving research for problem-solving, then the questions is what kinds of datasets we need to accumulated, what kind of analytical tools we could use for the study, and how to integrate the datasets and analytical tools to explain the what extents of climate changes' impacts on the schistosomiasis transmission or spreading in the future. So that all datasets, such as plant physiology, soil science, meteorology, oceanography, hydrology, and fluvial geomorphology, are necessary for the analysis, and the tools of theoretical and mathematical analysis, and computational analysis are essential for the study, as well as bridging the diverse length and time scales involved will require a collection of cooperating models. Thus, the purpose of this article is to review the application of various analytical tools in the control programmes on parasitic diseases in China, in particular, the NIPD's sustained contributions in her 70 years' development, through application of epidemiological methodologies on empirical analysis, theoretical and mathematical analysis, and computational analysis.

The search strategy for literature review for this investigation was done through the NCBI PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) and CNKI (http://www.cnki.net/, one of the biggest biomedical literature

databases in China), by searching the related English and Chinese literature for application of epidemiological analysis in parasitic disease studies of China.

Based on the science of epidemiology that is of interdisciplinary nature, e.g., disease aetiology, vector biology, surveillance and monitoring, meteorological and geographical environments, ecological evaluation, molecular and genetic analysis, and pharmaceutical studies on treatment effects including clinical trials, our review has focus on the development of epidemiological approaches, which are typically on three types of research paradigms, as illustrated in Fig. 1 and 2, namely empirical analysis, theoretical and mathematical analysis, and computational analysis.

2. Empirical analysis

In the earlier stage of epidemiological investigations, epidemiologists from NIPD were mainly relying on expertise from multiple scientific disciplines (Lv et al., 2011; Xia et al., 2017). For example, biologists can help provide a better understanding of parasitic disease processes and the attributes of parasites to the disease development. Empirical observations or experiences may be analysed quantitatively or qualitatively. By taking

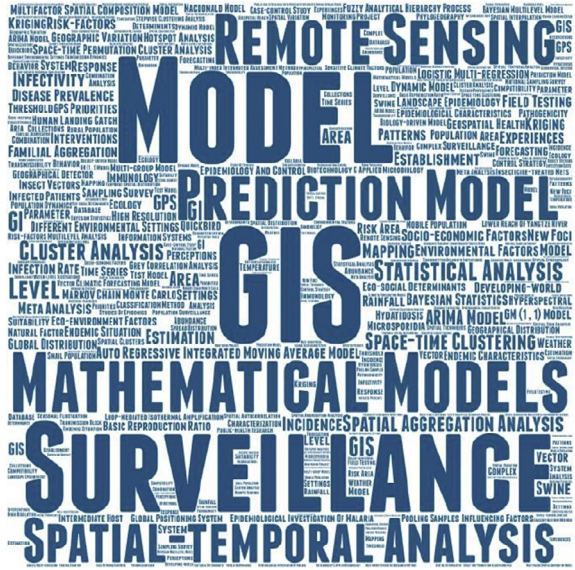
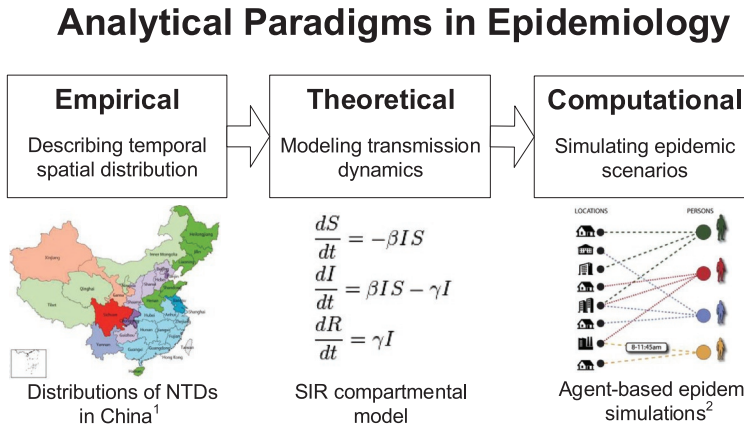


Fig. 1 The cloud mapping from the frequency of key words collected by our literature review.



[1] Yang G J, et al. China's sustained drive to eliminate neglected tropical diseases. The Lancet infectious diseases, 2014, 14(9): 881-892.
[2] Venkatramanan S. et al. Using data-driven agent-based models for forecasting emerging infectious diseases. Epidemics, 2018, 22: 43-49.

Fig. 2 Three types of research paradigms in epidemiological data analysis, namely empirical analysis, theoretical and mathematical analysis, and computational analysis.

different disciplines into consideration, the investigations on empirical epidemiology carried out in NIPD can be divided into three catalogues, such as (1) analytical epidemiology, in which statistical analysis of epidemiological data in an attempt to establish relationships between causative factors and incidences of a disease; (2) descriptive epidemiology, in which information about the distribution of a disease, and some of its mathematical descriptions; (3) experimental epidemiology, in which prospective population experiments are designed to test epidemiological hypotheses, and usually attempt to relate the postulated causes to the observed effects. Trials of new antiviral drugs and vaccines are an example.

2.1 Statistical methods for analysing the causative factors and disease incidences

In the history, scientists from NIPD did a lot of epidemiological studies to link the multi-impact factors with the disease prevalence. For instance, the impacts of the Three Gorges Dam on the schistosomiasis transmission supported by the ecological studies, the results were turned in formulation of a policy to take the surveillance in the Three Gorges region of Chongqian city. Up to now, the surveillance measures are still carried out in the Three Gorges region to monitor the potential risks of schistosomiasis transmission there. In addition, more tools of ecology and geographic information systems have introduced to the schistosomiasis epidemiological studies, which

improve the prediction capacity of NIPD in the risk assessment of schistosomiasis transmission, and now those modern tools have been used for routine application for the field evaluation of the schistosomiasis elimination programmeme every year.

In the field of malaria studies, the impact of climate changes on malaria transmission were explored in Yunnan province by the joint-team from NIPD and Yunnan Provincial Institute on Parasitic Diseases, the logistic regression models employed with climate data and malaria incidence data were developed. The major results of these studies indicated a positive correlation between warming climate or El Nino events and malaria incidence in Yunnan province, therefore, those impact of climate changes can used for the risks prediction to help mitigation strategy formulation. At the same time, a risk prediction model was established by employed the environmental data, malaria incidence data, and mosquito density data in Yunnan province. In addition, the impact of earthquake and social economic status on malaria transmission were evaluated in Yunnan and Hainan provinces by statistic models, which help local government to take consideration in the decision making (Feng et al., 2015b); More recent study aiming to understand the dynamics of malaria vectors in implementing malaria control strategies was undertaken in the Three Gorges Reservoir areas. So a total of six villages were selected from different sections in the Three Gorges Reservoir for exploring the relationship between the climatic factors and its malaria vector density from 1997 to 2007 using the auto-regressive linear model regression method. The result indicated that both temperature and precipitation were better modelled as quadratic rather than linearly related to the density of Anopheles sinensis (Wang et al., 2014).

For other parasitic diseases, several studies have been performed to correlate the social economic and environmental factors with the prevalence of echinococcosis in western and northwestern China, either through the cross-sectional surveys or case-control studies. Most of those results showed that the prevalence was highly correlated to the ecosystems and social behaviours (Xue et al., 2018; Yang et al., 2015a). Thus more interventions to reduce transmissibility through the staydogs, improve sanitation, and promote health education have put forward to control echinococcosis both in human and animals. An epidemiological analysis of the cases of visceral leishmaniasis in the non-endemic province of China, conducted a case investigation in the form of telephone and letter, and made a descriptive analysis of the obtained data in the region, time and population, gender, age and occupation distribution; Another epidemiological study to

understand impact factors of Blastocystis hominis infection among specific population, human immunodeficiency virus (HIV), carried out in Yunnan province by logistic regression model, results illustrated that among 3.7% (12/324) positives in SSU-RNA gene of Blastocystis hominis, the quality of drinking water, livestock raised at home, and immunocompromised status were the major factors related to the blastocystosis (Teng et al., 2018).

2.2 Descriptive methods for characterizing the temporal and spatial disease distribution

More researches from NIPD were focus on establishment of temporal spatial database for various parasitic diseases together with local environmental data, and then applied the GIS tools to integrate the epidemiological features to presented the transmission patterns with both space and time dimensions.

In schistosomiasis studies, scientists from NIPD have identified the cluster patterns and transmission risk areas where schistosomiasis heavily transmitted, by using temporal-spatial models. First, the scale of the study ranged from the national, provincial, county and village levels. Based on the local settings, both human and animal transmission patterns have been explored and compared, it was found the animal transmission more focalized along the lakes or rivers, particularly in the marshland and lake ecozoons. Second, another study on correlation between location of patients and distribution of *Oncomelania* snails found the spatial autocorrelation relationship strongly existed between these two variables, indicating people's infection was highly related to the environmental factors by exposure to the infested water in the lake region. Third, more researches on the detection of snail habitats by using remote sensing technology, i.e., Landcet, QuickBird images (Huang et al., 2007; Yang and Guo, 2006). It also found both *Oncomelania* snails or infected snails presented in cluster distribution in the marshland where snail habitats appeared as under water in summer and being wetland in winter (Gao et al., 2013b). These characters of the snail habitats are able to detect or layout from analysed remote sensing images after performing the temporal-spatial modelling in image analysis. Fourth, the landscape genetics on *Oncomelania* snail was also undertaken to illustrated the correlation of spatial and genetic distances of *Oncomelania hupensis* in mainland China (Guan et al., 2016; Li et al., 2009) using microsatellite loci and mitochondrial DNA techniques.

In malaria investigations, several studies have been done to provide a baseline for the ongoing national malaria elimination programmeme in China by illustrating the correlation between malaria incidences or

prevalence and other factors. First, a study on a 6-year longitudinal surveillance on density of malaria vectors was carried out in the 62 national malaria surveillance sites, to reveal the spatio-temporal distribution of malaria vectors in the national malaria surveillance sites from 2005 to 2010. The spatial and temporal analyses of the four primary vectors, including *Anopheles sinensis*, *An. lesteri*, *An. dirus* and *An. minimus* distribution were conducted by the methods of kernel k-means and the cluster distribution of the most widely distribution vector of *An. sinensis* was identified. It was revealed that *An. sinensis* was the most widely distributed, accounting for 96.25% of all collections, and its distribution was divided into three different clusters with a significant increase of density observed in the second cluster which located mostly in the central parts of China (Huang et al., 2015). Second, one report on spatial aggregation of malaria cases and specific vectors in the areas along the Huang-Huai River of central China where malaria had outbreak during 2006–2008, the data on geographical, meteorological and vectorial factors were analysed to determine the key factors related to malaria re-emergence in the areas (Zhou et al., 2012). It was found the spatial distribution between malaria cases and water-body, the changing of meteorological factors, and increasing vectorial capacity and basic reproductive rate of *An. sinensis* leaded to malaria re-emergency. Third, to identify the effective strategy to control the malaria re-emergency, a study was done to control the sources of malaria infection, based on the identification of risk areas and population. The geographic information of malaria cases and their surrounding water bodies were collected from Suixi, Guoyang, Guzhen, Yingshang, Fengyang and Yongqiao County in Anhui province, Yongcheng and Tongbai County in Henan province. Data on environment and malaria cases were mapped and analysed in 113 villages of the eight counties to identify the spatial correlation between malaria cases and water bodies. A total of 357 malaria cases and their geographic data as well as surrounding water bodies were collected and analysed. It was found that 74% of malaria cases were located within the extent of 60 m proximity to the water bodies, and distribution of water bodies is an important factor influencing the occurrence and distribution of malaria cases in the *An. sinensis* distributed areas. Therefore, the intervention for malaria case management needs to mainly target on the population living within 60 m around water bodies (Yan et al., 2014). Fourth, studies in mosquito vectors, a spatial-temporal variation and primary ecological drivers of *An. sinensis* human biting rates in malaria epidemic-prone regions of China was performed, and the temporal variation in biting rates was much larger than

the spatial variation, with 19.3% of temporal variation attributable to differences in minimum temperature and enhanced vegetation index and 16.9% of spatial variance due to distance to the nearest river and the paddy index. The results indicated that substantial spatial-temporal variation in *An. sinensis* human biting rates exists in malaria epidemic-prone regions of China, with minimum temperature and enhanced vegetation index accounting for the greatest proportion of temporal variation and distance to nearest river and paddy index accounting for the greatest proportion of spatial variation among observed ecological drivers (Huang et al., 2011; Ren et al., 2015).

In epidemiological studies on echinococcosis, two major areas on methodologies of temper-spatial distribution of human echinococcosis as well as well cluster patterns of the endemic areas were explored, which can provide the precision information on high risk areas where interventions need to be strengthened both in human and animals. For instance, the spatial scan analysis and spatial autocorrelation methods were used to identify the spatial clusters on the detection rate of cystic echinococcosis and alveolar echinococcosis patients who were newly detected based on the screening data of echinococcosis among population in 18 counties of Ganzi Tibetan Autonomous Prefecture, Sichuan province of China, in 2014. It was found that both the primary cluster of cystic and alveolar echinococcosis was majorly located in the northwest part of the Ganzi Tibetan Autonomous Prefecture, but the former one more concentrated in the northeastern Dege county and northeastern Shiqu county, while later one more focalized in the northeastern Dege county, northern Ganzi county, and northeastern Shiqu (Zhao et al., 2015). Then, the prediction of echinococcosis prevalence was also tried by use of spatial interpolation methods in Qinghai province, and the Kriging interpolation got the best estimation of echinococcosis prevalence which arranged from 2.04% to 5.28% in two counties (Fang et al., 2014).

In the other parasitic diseases, more studies were to identify the cluster distribution of infections. For example, one study on food-borne helminthic zoonosis, angiostrongyliasis, caused by infections with the third-stage larvae of *Angiostrongylus cantonensis*, was performed in three villages from Nanao Island, Guangdong Province to understand the spatial distribution of hosts of the parasites as a reference to the control strategy formulation. The density and natural infection of *Pomacea canaliculata*, and the snail intermediate host of *A. cantonensis*, and various rat species, the reservoir host of *A. cantonensis*, were surveyed from December 2015 to September 2016, and the survey data was analysed by ArcGIS, scan statistics, ordinary least

squares (OLS) and geographically weighted regression (GWR) models. Results indicated that there was a risk of angiostrongyliasis in this region of China, with 6.1% (72/1190) larval infection rate in snails and 29.1% (32/110) infection rate in the rate definitive hosts. The distribution of positive *P. canaliculata* and rats exhibited spatial correlation, and the GWR model had advantage over the OLS model in the spatial analysis of hosts of *A. cantonensis* (Hu et al., 2017).

Another study was on the spatial analysis of visceral leishmaniasis in the oases of the plains of Kashi Prefecture, Xinjiang Uygur Autonomous Region, results showed that the spatial distribution of visceral leishmaniasis-infected families by Poisson model spatial aggregation of the Scan statistical analysis was clustered, with three clustered zones, including Zone A located in a centre point of 76.153447°E, 39.528477°N within its 1.11 mile radius, where the cumulative life-incidence of leishmaniasis was 1.95 times as high as that in surrounding areas ($P < 0.05$); Zone B located in a centre point of 76.111968°E, 39.531895°N within its 0.54 mile radius, where the cumulative life-incidence of leishmaniasis was 1.82 times as high as that in surrounding areas ($P < 0.01$); and Zone C located in a centre point of 76.195427°E, 39.563835°N within its 0.68 mile radius, where the cumulative life-incidence of leishmaniasis was 1.31 times as high as that in surrounding areas ($P < 0.05$) (Wang et al., 2016).

2.3 Experimental methods for test epidemiological hypotheses

Some prospective studies on transmission risk assessment in high endemic areas of parasitic diseases, including schistosomiasis, malaria, echinococcosis and leishmaniasis, have been carried in China in order to verify current transmission risks and identify the potential new foci of those diseases driven by environmental changes or social features, which may provide assistance for the local decisions on control or elimination strategy (Table 1).

In schistosomiasis, active surveillance in selected risk areas of three provinces, namely Jiangsu, Anhui and Shandong, near the norther border of schistosomiasis endemic areas, was carried out from 2008 to 2010, including screening of potential infections in human and snail detection. Sentinel surveillance showed eight, nine and five confirmed cases of acute schistosomiasis in mobile populations (fishermen, migrant workers) in 2008–2010, respectively. All these cases were detected in Chaohu County, which must therefore be deemed an area at risk. In sentinel sites of Anhui, for example, the local number of acute cases account for 43% (19/44) in 2008, 33%

Table 1 Characteristics of studies using experimental analysis in the control programmes on parasitic diseases by NIPD.

				Methodology		
Disease	**Study aim**	**Study scales**	**Data set**	**Method**	**Algorithm/Model/ Technique/Tools**	**References**
Malaria	To study the impact of earthquake and social economic status on malaria transmission	Province (Yunnan & Hainan)	Earthquake data, social economic status data and malaria data	Analytical study	Statistic models	Feng et al. (2015a)
Malaria	To explore the relationship between the climatic factors and its malaria vector density	Village (the Three Gorges Reservoir areas)	Climate data and malaria vector density data	Analytical study	Auto-regressive linear model	Wang et al. (2014)
Echinococcosis	To study the correlation between the social economic & environmental factors and the prevalence of echinococcosis	Country	Social economic data, environmental factor data and echinococcosis incidence data	Descriptive study	Cross-sectional surveys & case-control studies	Xue et al. (2018) and Yang et al. (2015a)
Visceral leishmaniasis	To study visceral leishmaniasis cases	Country	Leishmaniasis cases incidence data	Descriptive analysis	Case investigation	Chen et al. (2011)
Blastocystis hominis disease	To study the impact factors of Blastocystis hominis infection among specific population	Province (Yunnan)	Gene data and impact factors data	Analytical study	Logistic regression models	Teng et al. (2018)
Schistosomiasis	To explore a new method of Oncomelania snails distribution and density prediction	County and Village	Oncomelania snails distribution GIS database	Descriptive method	QuickBird images	Huang et al. (2007)

Continued

Table 1 Characteristics of studies using experimental analysis in the control programmes on parasitic diseases by NIPD.—cont'd

				Methodology		
Disease	**Study aim**	**Study scales**	**Data set**	**Method**	**Algorithm/Model/ Technique/Tools**	**References**
Schistosomiasis	To study on correlation between location of patients and distribution of Oncomelania snails	Province	Schistosomias epidemiological data, GIS database	Descriptive method	Arc View3.x,S-PLUS and Spatial Statistics software	
Schistosomiasis	To detection of Oncomelania snails habitats	County and Village	Environmental data, malaria incidence data, and mosquito density data	Descriptive method	Temporal-spatial model	Yang ang Guo (2006)
Schistosomiasis	To illustrated the correlation of spatial and genetic distances of Oncomelania hupensis	National	Spatial distribution data, gene data	Descriptive method	Microsatellite loci and mitochondrial DNA techinquies	Guan et al. (2016) and Li et al. (2009)
Malaria	To survey the density of malaria vectors	National	Malaria vectors temporal distribution data	Descriptive method	Kernel k-means and the cluster distribution	Huang et al. (2015)
Malaria	To determine the key factors related to malaria re-emergence in the areas	Area research	Geographical, meteorological and vectorial factors	Descriptive method	Spatial analysis tools	Zhou et al. (2012)
Malaria	To identify the effective strategy to control the malaria re-emergency	County	Malaria cases spatial aggregation data	Descriptive method	Statistic analysis	Yan et al. (2014)
Malaria	To determine the primary ecological drivers of malaria vector human biting rates in malaria epidemic-prone regions of China	Area research	62 Malaria surveillance sentinel sites data	Descriptive method	Linear mixed effects model	Ren et al. (2015) and Huang et al. (2015)

Echinococcosis	To the spatial clusters on the detection rate of echinococcosis	Area research	Screening spatial data	Descriptive method	Spatial scan analysis and spatial autocorrelation methods	Zhao et al. (2015)
Echinococcosis	To estimate the echinococcosis prevalence	Province	Echinococcosis prevalence data	Descriptive method	Spatial interpolation method	Fang et al. (2014)
Schistosomiasis	To carry out active surveillance	Province	Potential infections in human and snail detection Sentinel surveillance data	Experimental method	Sample survey	Qian et al. (2013)
Malaria	To study on the epidemic factors of malaria	County	Malaria epidemiological investigation data	Experimental method	ArcGIS, excel	Feng et al. (2015a)
Echinococcosis	To determine the distribution of the intermediate hosts of Echinococcus multilocularis	County	Density and antigen positive rate of Intermediate hosts	Experimental method	Systematic sampling method, remote sensing (RS), regression analysis	Xue et al. (2018)
Echinococcosis	To study epidemiological characteristics and risk factors of human hydatidosis	Area research	The prevalence of cystic echinococcosis in slaughtered sheep	Experimental method	Logistic regression analysis	Yang et al. (2015a)
Leishmaniasis	To predict the risks and high risk areas for leishmaniasis transmission	National	Data in the region, time and population, gender, age and occupation distribution	Experimental method	Case investigation	Chen et al. (2011)

(25/75) in 2009 and 40% (17/42) in 2010, compare to the total number in the country. While the proportion of newly detected snail breeding areas in Anhui accounted for 42% (5.03/11.98) in 2008, 95% (8.39/8.79) in 2009 and 79% (8.52/10.81) in 2010, compared a total of 31.58 km^2 snail breeding sites that were newly detected nationwide through the year 2008–2010 (Qian et al., 2013).

In echinococcosis, to determine the distribution of the intermediate hosts of *Echinococcus multilocularis* in Shiqu County, Sichuan, China. A systematic sampling method was used to investigate the density of burrows of rodents and lagomorphs at 97 pasture sites in winter and summer pastureland and remote sensing (RS) technology was used to correlate their densities to the distribution of these animals in different landscape types. The regression relationship between the densities of *Ochotona curzoniae* and *Microtus fuscus* and their burrow densities indicated that burrow densities could reflect the animal densities and was greater in summer pastureland than that in winter pastureland. The main distribution areas of the intermediate hosts were in bareland and half-bareland. Xue et al. used a stratified cluster sampling method to select 364 villages in 70 counties (district) from August to October 2016. In each village, 200 residents were randomly selected for B-ultrasound examination, accompanied by a serological test for those suspected. Results showed that the highest risk for the transmission was the family with antigen positive rate in dog feaces, and human ages (Xue et al., 2018); Yang et al. analyse epidemiological characteristics and risk factors of human hydatidosis in Habahe, Xinjiang uygur autonomous region, and provide evidence for local prevention and control of human hydatidosis. Results demonstrated that the prevalence of cystic echinococcosis in slaughtered sheep may be used as an indicator to assess and monitor the transmission status during and after control programme providing information for betterment of performance to sustain control strength (Yang et al., 2015a).

In malaria, Feng et al. analysed the epidemic factors of malaria after earthquake in the recent 3 years in Cangyuan County of Yunnan Province, results indicated the western Cangyuan County had higher malaria transmission risk which related to the natural environment, social factors (Feng et al., 2015a).

In leishmaniasis, Chen et al. made an epidemiological analysis of the cases of visceral leishmaniasis in the non-endemic province of China, conducted a case investigation in the form of telephone and letter, and made a descriptive analysis of the obtained data in the region, time and population,

gender, age and occupation distribution. Results showed that among 13 non-endemic provinces, the top three provinces had higher risks for the transmission were Chongqing, Hunan and Guangxi. Where those infections were occurred mainly in Gangshu, Sichuan and Xingjiang (Chen et al., 2011).

3. Theoretical and mathematical analysis

When applied mathematics introduced into the research on epidemiology of parasitic diseases, the use of mathematical models to explain and reproduce the empirical observations in the national control or elimination programmes on parasitic diseases that will improve the precision effectiveness in the implementation of control strategies or diseases surveillance. Mathematical models can project how parasites infections progress and show the likely outcomes of an epidemic for parasitic diseases, and thus help inform public health decisions on interventions in the national programmes, technically supported by NIPD (Chen et al., 2018; Xia et al., 2017; Zhu et al., 2015). Two approaches developed for specific purposes, such as modelling for the dynamics of diseases transmission, and the intervention evaluation, were performed in NIPD up to now.

3.1 Theoretical modelling for analysing the dynamics of disease transmission

Several studies were carried out to apply autoregressive integrated moving average (ARIMA) model to predict the monthly reported malaria cases in China, for example, results showed that ARIMA is a good model for the short-term prediction, which can provide a reference for prevention and control of malaria (Hou et al., 2017); ARIMA model was established based on the month malaria incidences from 1996 to 2005 in Huaiyuan County of Anhui Province. Parameters of the model were estimated through maximum likelihood method; the structure was determined according to criteria of residual uncorrelation and concision, and the order of model was confirmed through Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (BSC). The constructed model was used to predict the month incidence in 2006 and the result was compared with the actual incidence (Zhou et al., 2007). While other models also used for the imported malaria prediction, such as Xia builds a GM (1, 1) model for forecasting of imported falciparum malaria cases in provinces free of indigenous falciparum malaria (Xia, 2009).

In another field of active surveillance on malaria guided by the modelling results under scenario of global warming, which have been integrated with surveillance system in the national malaria elimination programmes (Wang et al., 2014). For instance, Tian et al. applied the prediction modelling to understand the impact of global warming on malaria transmission, and Yang et al. established the prediction model by using integrated biology-driven and statistical models to understand the risk areas distribution which located in the border of China, under scenario of global warming.

In helminthiasis, biology driven two-host model and grey model were used in helminthiasis transmission prediction (Xu et al., 2015). For instance, barbour's two-host model was applied to predict and evaluate the effect of different control measures by computer simulation. Pilot samples in two villages of Shanghai suburb in 1950s, and three villages in Jiangxi provinces in 2000s were applied for the analysis (Gao et al., 2013a, 2017), all results indicated the basic reproduction (R0) can been changed due to intervention. In addition, construction of grey model for prediction of *Clonorchis sinensis* infection (Zang et al., 2015), and establishment of the mathematical model by the turning point of filariasis transmission were also established, as well as application of transmission dynamic model for echinococcosis and ascariasis.

3.2 Mathematical analysis for estimated the effectiveness of diseases intervention measures

The effectiveness of intervention on interrupting parasitic diseases transmission is also can be estimated or simulated by mathematical analysis (Wu, 2005). In schistosomiasis transmission was monitored after anti-flooding policy implemented in lake region after 1989 when the big floods occurred along whole Yangtze river. In comparison of predicting effect of schistosomiasis prevalence, five statistical models was established in the areas of "breaking dikes or opening sluice for water store" in Dongting Lake, it was found this policy had high impact to the snail habitats re-established (Sai et al., 2006). A drug-based intervention was studied on interrupt the transmission from buffaloes to human Schistosoma japonicum infection around Poyang Lake, which showing the priorities of controlling buffalo transmission in lake region (Guo et al., 2006). Optimizing molluscicide treatment strategies in different control stages of schistosomiasis in China (Yang et al., 2012b).

After the national malaria elimination programmeme launched in 2010, the surveillance is the major intervention to reach the goal of malaria free in

China (Yang et al., 2012a). Therefore, how to understand and to establish the surveillance-response systems with focus on the mobile populations as the basis of the elimination efforts (Chen et al., 2018; Yin et al., 2013; Zhou et al., 2015). One of big surveillance studies was carried out in Tengchong County of Yunnan province, where highest number of malaria cases were observed in the beginning of the malaria elimination programmeme launched (Li et al., 2016). Not only the surveillance of the "1-3-7" approach at the China-Myanmar Border were adopted (Feng et al., 2016; Wang et al., 2017), but also management of imported cases were strengthened (Wang et al., 2015). All these countermeasures produced great outcomes which help Tengchong County to reach the free malaria status in 2015 (Li et al., 2016) (Table 2).

4. Computational analysis

Along with the development of computer science, artificial intelligence, data mining, and geographic science, epidemiology becomes a research domain that applies multidisciplinary techniques to better understand the processes of disease spread, the patterns of disease infection and transmission, as well as the effectiveness of disease control and interventions from public health authorities. In computational epidemiology, the primary focus is the application of computational concepts and resources to disease related issues, while, providing analytical tools to public health authorities and epidemiological scientists. With the tools of modelling, simulation, and visualization, epidemiologists will expand their capabilities in predicting and analysing disease infection and transmission in a given population. Furthermore, the public health authorities can more effectively conduct scenario analysis and, therefore, facilitate their strategic planning of on allocating public health resources.

4.1 Computational modelling for analysing the dynamics of disease transmission

The computational modelling approaches have been used to understand the dynamics of parasitic diseases transmission, such as endemic types of echinococcosis transmission, infection or prevalence rates of various parasitic infections, severity of the disease and mortality (Zhang et al., 2018).

In schistosomiasis epidemiology, Li et al. used the endemic data of schistosomiasis in Xinjian District which were collected by taking the village as a unit from 2009 to 2014 to establish an endemic database, and the

Table 2 Characteristics of studies using theoretical and mathematical analysis in the control programmes on parasitic diseases by NIPD.

Disease	Study aim	Study scales	Data set	Methodology		References
				Method	Algorithm/Model/Technique/Tools	
Malaria	To predict the monthly reported malaria cases	National	Time series of monthly reported cases of malaria	Theoretical method	Autoregressive integrated moving average model (ARIMA)	Hou et al. (2017)
Malaria	To predict the month incidence	County	Month malaria incidences	Theoretical method	Maximum likelihood method	Zhou et al. (2007)
Malaria	To forecast imported falciparum malaria cases	National	Imported falciparum malaria incidence	Theoretical method	GM (1, 1) model	Xia (2009)
Malaria	To understand the impact of global warming on malaria transmission	National, area research	Meteorological data and malaria incidence	Theoretical method	Prediction model, integrated biology-driven and statistical models	Wang et al. (2014)
Clonorchis sinensis	To predict the Clonochis sinensis infection	Province	Epidemiological survey data	Theoretical method	Grey model	Zang et al. (2015)
Schistosomiasis	To compare predicting effect of schistosomiasis prevalence	Area research	Prevalence of schistosomiasis	Mathematical method	Moving average method, exponential smoothing method, autoregression method, ARIMA model and grey prediction method	Sai et al. (2006)
Schistosomiasis	To estimate the effectiveness of drug intervention	Area research	Surveillance data	Mathematical method	Predicted modelling	Guo et al. (2006)
Schistosomiasis	To estimate molluscicide treatment strategies	National	Surveillance data	Mathematical method	Predicted modelling	Yang et al. (2012b)
Malaria	To survey the malaria cases	County	Surveillance data	Mathematical method	Sample survey	Yang et al. (2012a), Zhou et al. (2015), Yin et al. (2013), Chen et al. (2018), Li et al. (2016), Feng et al. (2016) and Wang et al. (2017)

SaTScan software was applied to analyse the spatiotemporal aggregation areas of Schistosoma japonicum infection in people ' Oncomelania hupensis snails and cattle (Li et al., 2018; Xu et al., 2015); Application of climatic forecasting models to rank the various risk areas or patterns for schistosomiasis transmission in Chin Transmission risks of schistosomiasis japonica were catalogued in lake or mountainous regions, by extraction from back-propagation artificial neural network and logistic regression models, and by Barbour's two-host schistosomiasis model with seasonal fluctuations (Gao et al., 2017; Xu et al., 2013; Yang et al., 2009, 2015). Xue et al. were collected and analysed the schistosomiasis epidemiological data in Jiangling County from 2009 to 2013 together with the related geographical information. The tempo-spatial distribution patterns were analysed by the spatial autocorrelation analysis and spatial clustering analysis (Xue et al., 2016).

In malaria transmission, models were set up to predict the monthly distribution of malaria cases or classify the distribution zones based on the transmission risks. For instance, Zhou et al. collected the data of malaria incidence of 156 counties or cities along the Yellow River and Huai River in 2005 to establish the geographical information system data base by Arcgis 9.0 software. Mapping the malaria probability distribution based on the GIS data base by the spatial local interpolation method in the extension function. The predictive incidence probability map and semivariance function was produced by unbiased criterion. Cross-validation technique was used to evaluate the fitness of the distribution maps by mapping the error distribution map (Zhou et al., 2007); Another study by using multifactor spatial composite model to predict transmission tendency of malaria at national level (Yang et al., 2002). Many field investigations were carried out in Hainan provinces where malaria was heavily epidemic in early 1980s, so the transmission features were computerized for understand the epidemiological features, impacts of the social development projects.

4.2 Data mining and machine learning

With the complicated data derived from the large-scaled epidemiological studies, the machine learning approach starts to use for computerizing the problems solving, for instance, what is the optimized intervention for the control or elimination of a parasitic disease in certain settings, and how long or how much fund needed. For a given task, these tasks all seek to learn from data, based on several or huge number of parameters, either by

supervised or unsurprised learners. In schistosomiasis, we have applied these methods in identification of transmission risk patters in schistosomiasis transmission in China by back-propagation artificial neural network model (Xu et al., 2013) This neural network consists of two layers, one is the input layer and the other one is output layer. The former one is the input values supplied by the user in terms of the age of a person, the sex of a person and the region where the person lives. The system automatically generates weights for the input. A transformation function is applied on the weighted sum of the input to produce an output. The later one gives a prediction of if the person stands a high or low risk of being infected. So that we can understand the risk of infections in the study areas and population.

In addition to the application of machine learning in epidemiological studies, the large amount of genetic sequence information generated has provided massive data banks that lift the ability of researchers to examine and process this information with the aid of computational methods. For instance, computational methods can be used for identification of the pathogen origins and co-evaluation with their hosts. An integrated immunoproteomics and bioinformatics approach for the analysis of Schistosoma japonicum tegument proteins has been studied in NIPD (Chen et al., 2014). In this study, an integrated immunoproteomics and bioinformatics approach was used to profile the tegument of the human blood fluke Schistosoma japonicum, with an assistance of machine learning. It was found that 30 highly immunoreactive tegument proteins and 10 antigens with an AUC value greater than 0.90 were identified for the first time. In particularly, we found 17 of tegument immunoproteomes having putative interaction networks with other proteins of *S. japonicum*. The results will provide clues of potential target molecules for vaccine development and biomarkers for diagnostics of schistosomiasis.

In malaria, the impact of heterogeneous factors on geographic variations of *Plasmodium vivax* incidences is systematically investigate in Tengchong, Yunnan province, by data mining of geographic variations of *P. vivax* for active surveillance (Shi et al., 2015). The transmission model is built upon relevant demographic, environmental, and biophysical factors to describe the local infections of *P. vivax*. While the hidden time-dependent process is assessed by several socioeconomic factors to account for the imported cases of *P. vivax*. To quantitatively assess the impact of heterogeneous factors on geographic variations of *P. vivax* infections, a Markov chain Monte Carlo (MCMC) simulation method is developed to estimate the model parameters by fitting the space-time model to the reported spatial-temporal disease incidences. Although this study focuses mainly on the transmission of *P. vivax*,

the proposed space-time model is general and can readily be extended to investigate geographic variations of other diseases. Practically, such a computational model will offer new insights into active surveillance and strategic planning for disease surveillance and control.

4.3 Systems thinking for disease control

Systems thinking is a method for analysing and understanding the dynamic behaviours of complex systems in virtue of computer simulation technology, and has incomparable advantages in the study of complex nonlinear systems, but less in the fields of medicine and public health. Therefore, NIPD together with Yunnan Provincial Institute for Endemic Diseases, explored an innovative system model for schistosomiasis elimination in Erhai Lake in combination with ecological approaches based on the characteristics of schistosomiasis transmission in Eryuan County. The system modelling method was used to evaluate and predict the effectiveness of interventions, based on the outcomes of the multifarious interventions applied in the study. Results found that ecological approaches to control schistosomiasis in Eryuan County consist of three major components: (i) implementing precise interventions to stop schistosomiasis transmission by means of controlling the source of infection, blocking the biological transmission chains and cutting off the route of disease transmission; (ii) employing ecological approaches to improve the co-effectiveness of environmental protection and schistosomiasis prevention in the study area; and (iii) strengthening the professional skills of personnel involving in the schistosomiasis control programme. Therefore, ecological approaches implemented in schistosomiasis endemic areas of the Eryuan region are able to improve the co-effectiveness of environmental protection and schistosomiasis control, providing a new avenue for eliminating schistosomiasis thanks to the application of precise interventions (Table 3).

5. Remaining challenges and future opportunities

It has been nearly 70 years since the beginning of applying epidemiological data analysis techniques for the control and prevention of parasitic diseases in China. These researches carried out by the NIPD have made great developments and contributions. The achievements in using epidemiological data analysis in parasitic diseases' control and prevention are obvious, but novel developments in methodology and filed practices are still limited. Here some key problems and challenges are highlighted as follows:

Table 3 Characteristics of studies using computational analysis in the control programmes by NIPD.

				Methodology		
Disease	**Study aim**	**Study scales**	**Data set**	**method**	**Algorithm/Model/ Technique/Tools**	**References**
Schistosomiasis	To analyse the spatiotemporal aggregation areas of Schistosoma japonicum infection in people	County (Xinjian)	The endemic data of schistosomiasis	Computational modelling	Spatial clustering model	Li et al. (2018)
Schistosomiasis	To rank the various risk areas or patterns for schistosomiasis transmission	Country	Climate data	Computational modelling	Forecasting models	Xu et al. (2013), Yang et al. (2009, 2015b), Gao et al. (2017)
Schistosomiasis	To catalogue transmission risks of schistosomiasis japonica	In lake or mountainous regions	Climate data and he endemic data of schistosomiasis	Computational modelling	Back-propagation artificial neural network and logistic regression models, and Barbour's two-host schistosomiasis model with seasonal fluctuations	Xu et al. (2013), Yang et al. (2015b), Gao et al. (2013a)
Schistosomiasis	To analyse the tempo-spatial distribution patterns of schistosomiasis	County (Jiangling)	The endemic data of schistosomiasis	Computational modelling	Spatial clustering model	Xue et al. (2016)

Malaria	To map the malaria probability distribution	The Yellow River and Huai River	The data of malaria incidence of 156 counties or cities	Computational modelling	The spatial local interpolation method	Zhou et al. (2007)
Malaria	To predict transmission tendency of malaria	Country	The endemic data of malaria	Computational modelling	Multifactor spatial composite model	Yang et al. (2002)
Schistosomiasis	To identify the transmission risk patters in schistosomiasis transmission	Country	The endemic data of schistosomiasis	Data mining and machine learning	Back-propagation artificial neural network model	Xu et al. (2013)
Schistosomiasis	To analysis the Schistosoma japonicum tegument proteins	–	The large amount of genetic sequence information	Data mining and machine learning	An integrated immunoproteomics and bioinformatics approach ' the application of machine learning	Chen et al. (2014)
Malaria	To quantitatively assess the impact of heterogeneous factors on geographic variations of *P. vivax* infections	County (Tengchong)	Demographic, environmental, and biophysical factors	Data mining and machine learning	Markov chain Monte Carlo (MCMC) simulation method	Shi et al. (2015)

5.1 Factors influencing diseases are rooted in multiple sources and at multiple scales

The driven forces root in two emerging observations, in terms of multiple factors, and the coupled interactions among them. At the micro level, e.g., molecules and pathogens, biologists carry out fundamental research to explore and explain the mechanism of disease infection and transmission, as well as develop the antiviral drugs and vaccines. As for middle level, e.g., host individuals and vector species, epidemiologists conduct front filed interviews and collect samples for analyses. Ecologists analyse and study the relationships among diseases, vectors and the host population, by exploring their mutual interactions as well as the correlations with their environments. As for a higher middle level, e.g., population or communities, health authorities are in charge of the implementation of disease control and interventions by evaluating the effectiveness of different intervention strategies, as well as balancing the costs and benefits associated. At the macro level, e.g., nations and continents, the global climate and the landscape usage are changing, and epidemiologists must quantify significant linkages between atmospheric, oceanic, and terrestrial processes to properly study the influence on the emergence and varnish of a certain disease.

5.2 Complexity in epidemiological system of parasitic diseases needs integrated more factors in the modelling

Parasitic diseases, intermedia vector species, as well as their relied environments are both complicatedly interacted and act with complexity. Due to such a complexity, the emerging epidemiological issues requires large synthesis studies being attempted by collaborations beyond the traditional domain-based studies. To address these challenges in epidemiological studies, researchers first need to realize that impact factors are represented at multiple levels, including the micro level biochemical, etiological, genetic, and the macro level environmental, socio-economical, and behavioural, as well as at multiple scales at temporal and spatial aspects, such as from the seasons, years to decades, and from the villages, nations to continental regions. Second, these factors interrelate with each other and the influences are also represented at multiple levels and scales. An epidemiological system consists of all organizations, people and actions whose primary intent is to promote, restore or maintain health.

5.3 Systems thinking plays an emerging important role for the prediction and control of parasitic diseases

In many parts, the problem in the existing research paradigms of epidemiological data analysis, namely, lack of a comprehensive systems perspective. The spread of infectious diseases and its controls are, by and large, studied in separate disciplines. The policies and decisions on disease preventions and interventions, i.e., reducing morbidity and mortality, are developed within each research domain, which encumbering efforts from the public health authorities for an efficiently allocation of scarce resources. Diseases are not stand-alone studies with clear boundaries among multiple displaces. Disease parasites, human population, as well as the applied interventions are coupled with each other with the complexities in their causes and interactions at multiple levels and scales. The interrelationships in terms of their environmental, biological, socio-economical as well as behavioural conditions, which defies a simple representation of epidemiological systems. They call for novel ways of thinking to improve our ability to predict and control infectious diseases in host population. As illustrated in Fig. 3, a holistic systems framework is needed to capture diseases, host population, and interventions, and thus their intricate relationships in-to a unified platform. Such frameworks are developed using complex systems the problem in the existing research paradigms of epidemiological studies, namely, lack of a comprehensive systems perspective.

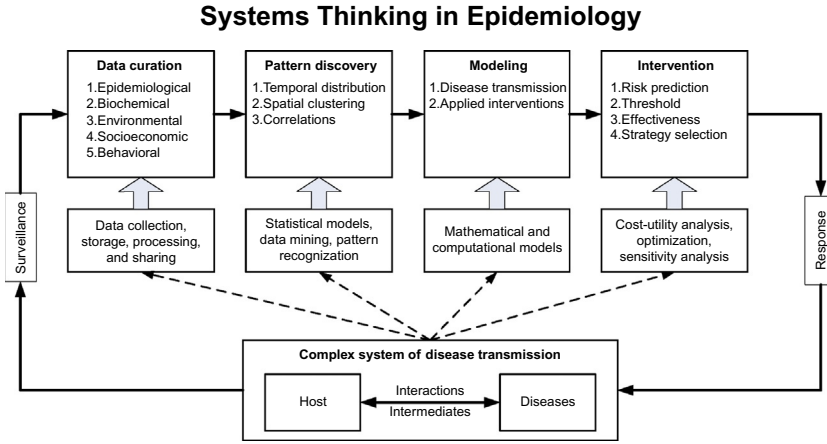


Fig. 3 Systems thinking for the control and prevention of parasitic diseases.

6. Conclusion

Since 1950s, the 70s years' continuous efforts by the NIPD on the analysis of epidemiological data have contributed significantly for the control and elimination of parasitic diseases in China. These mentioned paradigms and methods of epidemiological data analysis have significantly accelerated the in explorations in the prediction and control studies of parasitic diseases. While, prompt adaptations in epidemiological studies are also required as a response to the rapid changes in the understanding of disease related ecological environments and human related social systems. We have provided a complete review on epidemiological data analysis in parasitic diseases control and prevention by the NIPD's 70 years efforts and have identified the possible challenges and opportunities, hoping to shed lights on future research directions.

References

Chen, H.T., Wang, J.Y., Zhu, X.H., 2011. Epidemiological analysis of cases with visceral leishmaniasis reported from non-epidemic provinces in China. Int. J. Med. Parasit. Dis. 38, 177–179.

Chen, J.H., Zhang, T., Ju, C., Xu, B., Lu, Y., Mo, X.J., Chen, S.B., Fan, Y.T., Hu, W., Zhou, X.N., 2014. An integrated immunoproteomics and bioinformatics approach for the analysis of Schistosoma japonicum tegument proteins. J. Proteomics 98, 289–299.

Chen, T.M., Zhang, S.S., Feng, J., Xia, Z.G., Luo, C.H., Zeng, X.C., Guo, X.R., Lin, Z.R., Zhou, H.N., Zhou, S.S., 2018. Mobile population dynamics and malaria vulnerability: a modelling study in the China-Myanmar border region of Yunnan Province, China. Infect. Dis. Poverty 7.

Fang, Q., Wu, W.P., Wang, L.Y., Zeng, X.M., 2014. Use of spatial interpolation to predict the prevalence of hydatidosis. J. Parasit. Biol. 203–206 (in Chinese).

Feng, J., Xia, Z.G., Zhou, S.S., Xia, S., Wang, Q., Xiao, N., Zhou, X.N., 2015a. Preliminary evaluation of potential transmission risk of malaria after earthquake in Cangyuan County, Yunnan Province. Chin. J. Parasitol. Parasit. Dis. 33, 301–305.

Feng, J., Xiao, H., Xia, Z., Zhang, L., Xiao, N., 2015b. Analysis of malaria epidemiological characteristics in the People's Republic of China, 2004–2013. Am. J. Trop. Med. Hyg. 93 (2), 293–299.

Feng, J., Liu, J., Feng, X., Zhang, L., Xiao, H., Xia, Z., 2016. Towards malaria elimination: monitoring and evaluation of the "1-3-7" approach at the China-Myanmar border. Am. J. Trop. Med. Hyg. 95, 806–810.

Gao, S.J., He, Y.Y., Liu, Y.J., Yang, G.J., Zhou, X.N., 2013a. Field transmission intensity of Schistosoma japonicum measured by basic reproduction ratio from modified Barbour's model. Parasit. Vectors 6, 141.

Gao, F.H., Zhang, S.Q., He, J.C., Li, T.T., Wang, T.P., Zhang, G.H., Li, S.Z., 2013b. Spatial-temporal analysis on the distribution of infected snail at the village level in Anhui province, 2006–2012. Chin. J. Epidemiol. 34, 1101–1104 (in Chinese).

Gao, S.J., Cao, H.H., He, Y.Y., Liu, Y.J., Zhang, X.Y., Yang, G.J., Zhou, X.N., 2017. The basic reproductive ratio of Barbour's two-host schistosomiasis model with seasonal fluctuations. Parasit. Vectors 10, 42.

Guan, W., Li, S.Z., Abe, E.M., Webster, B.L., Rollinson, D., Zhou, X.N., 2016. The genetic diversity and geographical separation study of Oncomelania hupensis populations in mainland China using microsatellite loci. Parasit. Vectors 9, 28.

Guo, J.G., Li, Y.S., Gray, D., Ning, A., Hu, G.H., Chen, H.G., Davis, G.N., Sleigh, A.C., Feng, Z., McManus, D.P., Williams, G.M., 2006. A drug-based intervention study on the importance of buffaloes for human Schistosoma japonicum infection around Poyang Lake, People's Republic of China. Am. J. Trop. Med. Hyg. 74, 335–341.

Hou, H.Y., Song, H.Q., Zhang, S.X., Ai, L., Lu, Y., Cai, Y.C., Li, S.Z., Teng, X.J., Yang, C.L., Hu, W., Chen, J.X., 2017. Application of ARIMA model to predict number of malaria cases in China. Chin. J. Schistosom. Control 29, 436–440. 458 (in Chinese).

Hu, Q.A., Zhang, Y., Guo, Y.H., Lv, S., Xia, S., Liu, H.X., 2017. Investigation of Pomacea canaliculata and rodent densities and their infection status with Angiostrongylus cantonensis in Nan'ao island of Guangdong Province. Chin. J. Parasitol. Parasit. Dis. 35, 239–245.

Huang, Q.N., Tang, L.L., Jiang, X.G., Chen, Z., Zhou, X.N., 2007. Retrieving eco-environment factors relevant to Oncomelania snail distribution based on QuickBird image. Chin. J. Parasitol. Parasit. Dis. 304–309 (in Chinese).

Huang, F., Zhou, S.S., Zhang, S.S., Zhang, H.W., Li, W.D., 2011. Meteorological factors–based Spatio-temporal mapping and predicting malaria in Central China. Am. J. Trop. Med. Hyg. 85, 560–567.

Huang, J.X., Xia, Z.G., Zhou, S.S., Pu, X.J., Hu, M.G., Huang, D.C., Ren, Z.P., Zhang, S.S., Yang, M.N., Wang, D.Q., 2015. Spatio-temporal analysis of malaria vectors in national malaria surveillance sites in China. Parasit. Vectors 8, 146.

Li, S.Z., Wang, Y.X., Yang, K., Liu, Q., Wang, Q., Zhang, Y., Wu, X.H., Guo, J.G., Bergquist, R., Zhou, X.N., 2009. Landscape genetics: the correlation of spatial and genetic distances of Oncomelania hupensis, the intermediate host snail of Schistosoma japonicum in mainland China. Geospat. Health 3, 221–231.

Li, S.G., Yin, S.Q., Wang, J.Z., Li, X.S., Feng, J., 2016. Shifting from control to elimination: analysis of malaria epidemiological characteristics in Tengchong County around China-Myanmar border, 2005-2014. Malar. J. 15, 1–7.

Li, Y.L., Xue, J.B., Hu, W.C., Mao, Y.H., Zhang, Z.S., Zhou, L., Zhang, P., Yu, Z.K., Xu, J., 2018. Endemic features and space-time clustering of schistosomiasis in Xinjian District, Nanchang City from 2009 to 2014. Chin. J. Schistosom. Control 30, 260–263.

Lv, S., Zhang, Y., Steinmann, P., Yang, G.J., Yang, K., Zhou, X.N., Utzinger, J., 2011. The emergence of angiostrongyliasis in the People's Republic of China: the interplay between invasive snails, climate change and transmission dynamics. Freshw. Biol. 56, 717–734.

Qian, Y.J., Li, S.Z., Xu, J., Yang, K., Huang, Y.X., Cao, Z.G., Feng, M., Dang, H., Zhang, L.J., Zhang, L., Wang, Q., Bergquist, R., Xiao-Nong, Z., 2013. Potential schistosomiasis foci in China: a prospective study for schistosomiasis surveillance and response. Acta Trop. 141, 342–348.

Ren, Z.P., Wang, D.Q., Hwang, J., Bennett, A., Sturrock, H.J.W., Ma, A., Huang, J.X., Xia, Z.G., Feng, X.Y., Wang, J.F, 2015. Spatial-temporal variation and primary ecological drivers of Anopheles sinensis human biting rates in malaria epidemic-prone regions of China. PLoS One 10, e0116932.

Sai, X.Y., Xing, Q.J., Meng, D.R., Jia, Y.R., Cai, K.P., Li, Y.S., 2006. Comparison of predicting effect of schistosomiasis prevalence by 5 statistical models in the areas of "breaking dikes or opening sluice for water store" in Dongting Lake. J. Fourth Mil. Med. Univ. 1603–1605 (in Chinese).

Shi, B.Y., Tan, Q., Zhou, X.N., Liu, J.M., 2015. Mining geographic variations of plasmodium vivax for active surveillance: a case study in China. Malar. J. 14, 216.
Teng, X.J., Chu, Y.H., Zhai, C.C., Yu, Y.F., Cai, Y.C., Chen, S.H., Ai, L., Tian, L.G., Chen, J.X., 2018. The epidemiological characteristics and influencing factors for Blastocystis hominis infection among human immunodeficiency virus seropositive individuals in Tengchong of Yunnan Province. Chin. J. Parasitol. Parasit. Dis. 36, 129–134 (in Chinese).
Wang, D.Q., Gu, Z.C., Zheng, X., Guo, Y., Tang, L.H., 2014. Application of autoregressive linear model in understanding the effect of climate on malaria vectors dynamics in the three gorges reservoir. Biomed. Environ. Sci. 27, 811–814.
Wang, D.Q., Li, S.G., Cheng, Z.B., Xiao, N., Cotter, C., Hwang, J., Li, X.S., Yin, S.Q., Wang, J.Z., Bai, L., Zheng, Z., Wang, S.B., 2015. Transmission risk from imported plasmodium vivax malaria in the China-Myanmar border region. Emerg. Infect. Dis. 21, 1861.
Wang, L.Y., Wu, W.P., Fu, Q., Guan, Y.Y., Han, S., Niu, Y.L., Tong, S.X., Osman, I., Zhang, S., Kaisar, K., 2016. Spatial analysis of visceral leishmaniasis in the oases of the plains of Kashi prefecture, Xinjiang Uygur autonomous region, China. Parasit. Vectors 9, 1–7.
Wang, D.Q., Cotter, C., Sun, X.D., Bennett, A., Gosling, R.D., Xiao, N., 2017. Adapting the local response for malaria elimination through evaluation of the 1-3-7 system performance in the China-Myanmar border region. Malar. J. 16.
Wu, K.C., 2005. Math ematical model in prediction and evaluation of the effects on control measures for schistosomiasis. Chin. J. Parasitol. Parasit. Dis. 23, 408–414.
Xia, Z.G., 2009. Forecasting of imported falciparum malaria cases in provinces free of indigenous falciparum malaria using GM(1,1)model. Pract. Prev. Med. 16, 1744–1746.
Xia, S., Zhou, X.-N., Liu, J., 2017. Systems thinking in combating infectious diseases. Infect. Dis. Poverty 6 (5), 57–63.
Xu, J.F., Xu, J., Li, S.Z., Jia, T.W., Huang, X.B., Zhang, H.M., Chen, M., Yang, G.J., Gao, S.J., Wang, Q.Y., Zhou, X.N., 2013. Transmission risks of schistosomiasis japonica: extraction from back-propagation artificial neural network and logistic regression model. PLoS Negl. Trop. Dis. 7, e2123.
Xu, J.F., Lv, S., Wang, Q.Y., Qian, M.B., Liu, Q., Bergquist, R., Zhou, X.N., 2015. Schistosomiasis japonica: modelling as a tool to explore transmission patterns. Acta Trop. 141, 213.
Xue, J.B., Xia, S., Zhang, X., Hu, H.H., Wang, Q., Zhou, Y.B., Li, S.Z., 2016. Pattern analysis of tempo-spatial distribution of schistosomiasis in marsh-land epidemic areas in stage of transmission control. Chin. J. Schistosom. Control 28, 624–629 (in Chinese).
Xue, C.Z., Wu, W.P., Han, S., Zheng, C.J., Wang, Y., Wang, L.Y., Yu, Q., Zhang, S.S., Zhou, X.N., Xiao, N., 2018. Prevalence and influencing factors of hydatid disease among children in Tibet Autonomous Region. Chin. J. Parasitol. Parasit. Dis. 36, 20–25 (in Chinese).
Yan, R.Z., Zhou, S.S., Xia, Z.G., Zhang, S.S., Xiao, H.H., 2014. Spatial-temporal characteristics of malaria transmission in China. J. Parasit. Biol. 198–202.
Yang, G.J., Zhou, X.N., Malone, J.B., Mccarroll, J.C., Liu, J.X., 2002. Application of multifactor spatial composite model to predict transmission tendency of malaria at national level. Chin. J. Parasitol. Parasit. Dis. 20, 145–147 (in Chinese).
Yang, D.P., Guo, J.G., 2006. The application of remote sensing in rapid identification of the snails' breeding areas. Int. J. Med. Parasit. Dis. 33 (05), 268–272 (in Chinese).
Yang, G.J., Sun, L.P., Hong, Q.B., Yang, K., Deng, Y., Li, S.Z., Lv, S., 2009. Application and perspective of climatic forecasting models related to schistosomiasis transmission. Chin. J. Schistosom. Control 21, 432–436.

Yang, G.J., Sun, L.P., Hong, Q.B., Zhu, H.R., Yang, K., Gao, Q., Zhou, X.N., 2012a. Optimizing molluscicide treatment strategies in different control stages of schistosomiasis in the People's Republic of China. Parasit. Vectors 5, 260.

Yang, G.J., Tanner, M., Utzinger, J., Malone, J.B., Bergquist, R., Chan, E.Y.Y., Gao, Q., Zhou, X.N., 2012b. Malaria surveillance-response strategies in different transmission zones of the People's Republic of China: preparing for climate change. Malar. J. 11, 426.

Yang, S.J., Li, M., Tian, T., Zhao, J.S., Wu, W.P., 2015a. Epidemiological characteristics of human hydatidosis and risk factors in Habahe County, Xinjiang. Dis. Surveillance 30, 485–488 (in Chinese).

Yang, K., Zhou, X.N., Jia, T.W., Yang, G.J., Wu, X.H., Shi, X.W., Li, H.J., Steinmann, P., Utzinger, J., Bergquist, R., 2015b. Eco-social determinants of Schistosoma japonicum infection supported by multi-level modelling in Eryuan county, People's Republic of China. Acta Trop. 141, 391–398.

Yin, J.H., Yang, M.N., Zhou, S.S., Wang, Y., Feng, J., Xia, Z.G., 2013. Changing malaria transmission and implications in China towards National Malaria Elimination Programmeme between 2010 and 2012. PLoS One 8, e74228.

Zang, W., Lu, G.H., Zhu, T.J., Chen, Y.D., 2015. Construction of grey model for prediction of Cllaborating sinensis infection. Int. J. Med. Parasit. Dis. 42, 170–172.

Zhang, M.Y., Wang, L.Y., Guan, Y.Y., Wu, W.P., 2018. Meta-analysis on severity of hydatidosis. Chin. J. Parasitol. Parasit. Dis. 36, 156–160.

Zhao, Y., Wu, W.P., Li, W., Xu, K.J., Wang, L.Y., Wang, Y., 2015. Spatial clustering analysis of echinococcosis in Ganzi Tibetan autonomous prefecture, Sichuan Province. Int. J. Med. Parasit. Dis. 42, 164–169.

Zhou, S.S., Huang, F., Shen, Y.Z., 2007. Application of ARIMA model on prediction of malaria incidence. J. Pathogen Biol. 2, 284–286 (in Chinese).

Zhou, S.S., Zhang, S.S., Wang, J.J., Zheng, X., Huang, F., Li, W.D., Xu, X., Zhang, H.W., 2012. Spatial correlation between malaria cases and water-bodies in Anopheles sinensis dominated areas of Huang-Huai plain, China. Parasit. Vectors 5, 106.

Zhou, S.S., Zhang, S.S., Li, Z., Rietveld, A.E.C., Ramsay, A.R., Zachariah, R., Bissell, K., Bergh, R.V.D., Xia, Z.G., Zhou, X.N., 2015. China's 1-3-7 surveillance and response strategy for malaria elimination: is case reporting, investigation and foci response happening according to plan? Infect. Dis. Poverty 4, 1–9.

Zhu, H.R., Liu, L., Zhou, X.N., Yang, G.J., 2015. Ecological model to predict potential habitats of Oncomelania hupensis, the intermediate host of Schistosoma japonicum in the mountainous regions, China. PLoS Negl. Trop. Dis. 9, e0004028.