A village-based multidisciplinary study on factors affecting the intensity of cystic echinococcosis in an endemic region of the Tibetan plateau, China

H. H. HU[1,2], W. P. WU[1]*, Y. Y. GUAN[1], L. Y. WANG[1], Q. WANG[3], H. X. CAI[1,4] AND Y. HUANG[3]

[1] *National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention; WHO Collaborating Center for Malaria, Schistosomiasis and Falariasis, Key Laboratory of Parasite and Vector Biology, Shanghai, People's Republic of China*
[2] *Laboratory of Exercise Epidemiology, Graduate School of Sport Sciences, Waseda University, Mikajima, Saitama, Japan*
[3] *Institute of Parasitic Disease Control, Sichuan Province Center for Disease Control and Prevention, Chengdu City, Sichuan, People's Republic of China*
[4] *Qinghai Institute for Endemic Disease Prevention and Control, Xining City, Qinghai, People's Republic of China*

SUMMARY

We investigated and quantified the factors which may affect the prevalence of cystic echinococcosis caused by *Echinococcus granulosus* in Rangtang County using a multidisciplinary approach. From a previously performed field survey, epidemiological data were linked with environmental data. Altitude and land surface temperature were extracted from remote-sensing images. Cumulative logistic regression models were used to identify and quantify the potential risk factors. The multiple regression models confirmed that yaks ($\chi^2 = 4 \cdot 0447$, $P = 0 \cdot 0443$), dogs ($\chi^2 = 8 \cdot 3455$, $P = 0 \cdot 0039$) and altitude ($\chi^2 = 7 \cdot 6223$, $P = 0 \cdot 0058$) were positively correlated with the prevalence of cystic echinococcosis, while land surface temperature may have a negative association. The findings showed that dogs and yaks play the most important role in the transmission of cystic echinococcosis, while altitude and land surface temperature may also be involved in the transmission.

Key words: Altitude, epidemiology, prevalence, temperature, yak.

INTRODUCTION

Echinococcus granulosus, a zoonotic disease, which is primarily transmitted between domestic dogs and livestock, is the cause of human cystic echinococcosis (CE) and has a worldwide geographical distribution [1]. Within the endemic zones (whether large or small scale), the occurrence of the parasite varies from sporadic to high [2–4]. It was reported that CE prevalence depends on various biological, social, and environmental risk factors including hosts, livestock husbandry, dog management, customs, as well as natural environment [2, 5]. To get a better understanding on the factors which are involved in the transmission, a multidisciplinary approach is needed [6]. The main challenge is to find ways to connect various fields of science and to consider the transmission systems as a whole.

Satellite sensor data have been used in epidemiology and health science in the past decades [7, 8]. Some reports on the use of remote sensing for

monitoring, mapping and predicting *E. multilocularis* transmission have been published in recent years [9–14]. However, the use of the technology and its associated data when studying *E. granulosus* is less common. Remotely sensed data and spatial analysis provide an opportunity to further the study of CE transmission and this paper reports one aspect of current research for a study conducted in the Tibetan plateau, Sichuan Province, China.

Human CE continues to be a substantial cause of morbidity and mortality in the Tibetan plateau. Surveys have shown that CE prevalence in some villages may reach as high as 12·1% [15–19]. Risk factors associated with *E. granulosus* infections include the number of owned dogs, frequency of contact with dogs, and sources of drinking water. However, there is still no comprehensive understanding of the transmission of the disease in the plateau and the role of the possible risk factors, which are important for predicting CE prevalence and developing a targeted control measure in this area.

It was reported that the prevalence of human echinococcosis was about 1·4% in Rangtang County [20, 21]. In 2009, we conducted a field survey on CE prevalence in yaks in Rangtang County. About 20·6% of yaks died during March–May in this area. Of those dead yaks, infection rate of *E. granulosus* ranged from 0% to 70·5% by village. The analyses results showed that yak mortality in spring was not significantly associated with the prevalence of yak infection [22]. In order to investigate and quantify the factors which may influence the prevalence of yak infection, the present study used satellite-derived environmental variables combined with epidemiological data obtained from our previous work [22]. In this paper, we report the results of our efforts to further explore the factors affecting the variation in CE endemicity.

METHODS

Study area

As described previously, a field survey regarding CE prevalence in yaks was conducted in 390 households from 35 administrative villages in Rangtang County in July 2009 [22]. Geographically, Rangtang County is located in the southeast part of the Qinghai–Tibetan plateau (longitude 100° 30′–101° 20′ N and latitude 31° 30′–32° 40′ E). It covers an area about 6600 km^2. The average altitude is ~3280 m. Ethnic

Tibetans comprise most of the population; they are primarily engaged in livestock production and herding. The main domestic livestock are yaks. Approval for the survey was given by the academic board of the National Institute of Parasitic Diseases, China CDC and all participants were informed about the objectives and procedures of the study.

Questionnaires and procedures

During the field survey which was performed in July 2009, a questionnaire was administered to the herdsmen after obtaining data about the death of yaks in spring and the infection status of CE in the dead yaks. The questionnaire covered items like the number of dogs and the number of yaks. These variables are important for the current study which aims at identifying and quantifying factors potentially affecting the CE transmission.

Buffer zone definition

Research on *E. multilocularis*, another *Echinococcus* species, has demonstrated that 2000 m is an important distance from a village centre for landscape variables to influence the transmission of *E. multilocularis*. This is most likely due to the maximum range that domestic dogs will generally forage away from the village centre [10–13]. Considering that domestic dogs are also the main definitive host of *E. granulosus*, a 2000 m buffer was used in this study.

Annual land surface temperature

It was reported that the survival of eggs of *Echinococcus* in the environment may be affected by temperature, and lower temperature may prolong the survival of eggs [2]. Considering this, the land surface temperature was considered to be a potential risk factor in this study. To extract land surface temperature of this study area in 2008, images from a moderate-resolution imaging spectroradiometer (MODIS 8-Day L3 1 km; NASA, USA) were used. Since the images for land surface temperature were produced every 8 days, day and night respectively, a total of 92 images (46 images for day, 46 images for night) were obtained for this study. All the images were registered with a Universal Transverse Mercator (UTM) coordinate system with the MODIS reprojection tool. This tool was designed to reproject MODIS images into more standard map projections. Annual average land surface
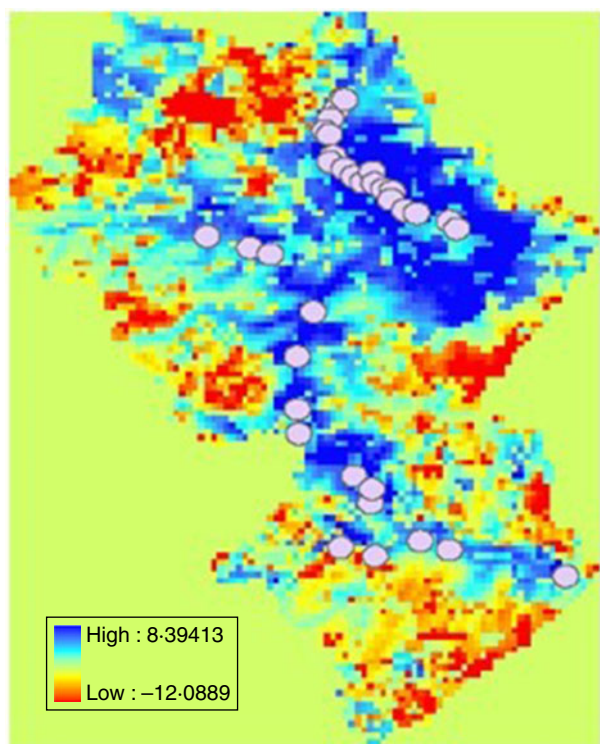
Fig. 1 [*colour online*]. Annual daytime land surface temperature in 2008, Rangtang County; 2000 m buffer (○).

Fig. 2 [*colour online*]. Annual night-time land surface temperature in 2008, Rangtang County; 2000 m buffer (○).

temperature for day and night, respectively, was calculated by the equation:

$$t = (\text{float}(b1) + \text{float}(b2) + \cdots + \text{float}(b92)) \\ * 0.02/92 - 273.15.$$

To quantify the land surface temperature around the target villages, circular buffer zones, centred on each village, were defined and data was extracted from these buffer zones with ArcGis software version 8.3 (ESRI, USA) (Figs 1 and 2). The annual average land surface temperature for each village was $T=(t(\text{day})+t(\text{night}))/2$.

Altitude

Considering that altitude may affect CE transmission by influencing the distribution of vegetation and grazing activity, altitude was treated as a risk factor in this study. In order to obtain data on altitude in the study area, topographic data at 90 m resolution from the Shuttle Radar Topography Mission (SRTM) was obtained. The image was transferred to a UTM coordinate system. To quantify the altitude around the target villages, circular buffer zones with radii of 2000 m, centred on each village, were
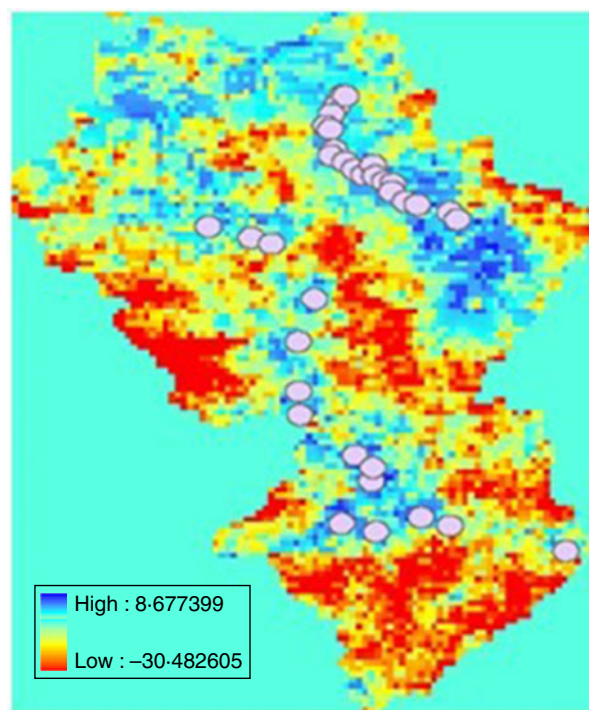
defined and data was extracted from these buffer zones with ArcGis software version 8.3.

Data management and statistical analysis

The 35 administrative villages were classified into four types of areas: class I, highly endemic (prevalence >20%); class II, moderately endemic (9% ≤ prevalence ≤ 20%); class III, endemic (0% < prevalence < 9%); class IV, suspected (no infected yaks were found during the research period; however, herdsmen were aware of CE). The average number of dogs per household in each village was calculated with M1= total number of dogs in the households participating/ number of households participating, the average number of yaks per household in each village was calculated with M2=total number of yaks in the households participating/number of households participating.

Data analysis focused on the examination of the correlation between village CE prevalence and potential risk factors. Cumulative logistic regression was used to investigate correlation between risk factors and endemic intensity of CE (classes I, II, III, IV). The dependent variable (endemic intensity of CE) was given a score 1, 2, 3 and 4 for villages in

classes I, II, III and IV, respectively. The lower the scores, the more endemic the area. The model included the following variables: average number of dogs per household in each village, average number of yaks per household in each village, altitude of each village, and annual average land surface temperature of each village (both altitude and land surface temperature were transferred to be ordinal variables). For the multiple regression models, Pearson's χ^2 test was used to assess goodness-of-fit.

All statistical analysis was also performed in Epi-Info version 6 (CDC, USA). The level of statistical significance was set at $P = 0.05$, unless otherwise stated.

RESULTS

Structure of dogs and yaks in Rangtang County

Dogs

According to the survey, about 64% of households did not keep a dog, 31% kept one dog and a few households (5%) kept more than one dog. The average number of dogs per household for classes I, II, III and IV was 0.66 ± 0.454, 0.46 ± 0.243, 0.39 ± 0.266 and 0.39 ± 0.293, respectively. The difference between classes I, II, III and IV was not significant ($F = 1.284$, $P = 0.297$). The herdsmen indicated that dogs ate the offal and remains of dead yaks, especially during spring, in which season large numbers of yaks die.

Yaks

About 60% of households kept 20–60 yaks. The average number of yaks per household for classes I, II, III and IV was 63.7 ± 32.42, 22.97 ± 11.38, 32.8 ± 15.84 and 29.8 ± 14.53, respectively. The difference between classes I, II, III and IV was significant ($F = 5.834$, $P = 0.003$). There were more yaks in class I compared to other classes.

Annual average land surface temperature

The annual average land surface temperature in the 35 administrative villages ranged between -0.39 and $5.38\,°C$. The annual average land surface temperature for classes I, II, III and IV was about $3.82 \pm 1.502\,°C$, $3.44 \pm 1.369\,°C$, $2.84 \pm 1.466\,°C$ and $3.49 \pm 1.140\,°C$, respectively. The difference between classes I, II, III and IV was not significant ($F = 0.822$, $P = 0.492$).

Altitude

The altitude in the 35 administrative villages ranged between 2955 m and 3784 m, with the average altitude about 3510.6 ± 157.79 m. The average altitude for classes I, II, III and IV were 3572.3 ± 117.49 m, 3449.5 ± 257.79 m, 3518.2 ± 141.81 m and 3489.7 ± 132.76 m, respectively. The difference between classes I, II, III and IV was not significant ($F = 0.547$, $P = 0.854$).

Factors affecting the endemic intensity of CE

A series of univariate cumulative logistic regression models were constructed to test the individual risk factors on CE prevalence (Table 1). For cumulative logistic regressions, the odds ratio represents the effect of a unit increase in the independent variable on the likelihood of a more adverse categorical outcome on the dependent variable. Thus, yaks were significantly associated with CE [odds ratio (OR) 0·964, 95% confidence interval (CI) 0·934–0·995], indicating that, with each unit increase in that variable, the odds of an endemic intensity score of 2, 3 and 4 are 4·6% lower than the odds of a score of 1. This implies that yaks may promote CE prevalence. Due to the proportionality assumption imposed by the logistic regression, the same 4·6% reduction was associated with each other division of endemic intensity score into more and less severe dichotomizations (e.g. scores [3, 4] vs. [1, 2]; [4] vs. [1, 2, 3]). In addition to yaks, altitude also showed a positive association with CE prevalence (OR 0·258, 95% CI 0·080–0·832). However, the positive association between dogs and CE prevalence was not significant in this univariate analysis (OR 0·143, 95% CI 0·019–1·064).

To examine effects in a multiple regression framework, an initial stepwise cumulative logistic regression was conducted, with all four variables as covariates. Results from Table 2 show that altitude, yaks and dogs were all positively correlated with CE prevalence. In this model, dogs were highly significant correlated with the disease with an odds ratio of 0·024 (95% CI 0·002–0·300), indicating that a 1-point increase in the average number of dogs per household in each village was associated with a 97·6% increase in the odds of a lower score. The adjusted R^2 for this model is 0·4426.

When annual land surface temperature was forced into the regression model (Table 3), it was noted that it was negatively correlated with CE prevalence,

Table 1. *Univariate cumulative logistic regression models*

Model	Dependent variable	Independent variable	β	OR (95% CI)	P	Adjusted R^2
1	Score	Altitude	−1·3542	0·258 (0·080–0·832)	0·0233	0·1450
2	Score	Temperature	−0·2147	0·807 (0·359–1·815)	0·6038	0·0082
3	Score	Dogs	−1·9432	0·143 (0·019–1·064)	0·0576	0·0980
4	Score	Yaks	−0·0366	0·964 (0·934–0·995)	0·0235	0·1650

OR, Odds ratio; CI, confidence interval.
Score: the dependent variable (endemic intensity of CE) was given a score of 1, 2, 3 and 4 for villages in classes I, II, III and IV, respectively.
Altitude: altitude of each village.
Temperature: annual land surface temperature of each village.
Dogs: average number of dogs per household in each village.
Yaks: number of yaks per household in each village.

Table 2. *Multiple regression (stepwise) cumulative logistic regression model*

Parameter	D.F.	β	OR (95% CI)	Wald χ^2	$Pr > \chi^2$
Intercept4	1	5·5242		8·7665	0·0031
Intercept3	1	7·3015		12·9750	0·0003
Intercept2	1	8·9765		16·3604	<0·0001
Altitude	1	−2·0786	0·125 (0·029–0·547)	7·6223	0·0058
Dogs	1	−3·7484	0·024 (0·002–0·300)	8·3455	0·0039
Yaks	1	−0·0348	0·996 (0·934–0·999)	4·0447	0·0443

D.F., Degrees of freedom; OR, odds ratio; CI, confidence interval.

Table 3. *Cumulative logistic regression model with full variables*

Parameter	D.F.	β	OR (95% CI)	Wald χ^2	$Pr > \chi^2$
Intercept4	1	4·8287		6·0388	0·0140
Intercept3	1	6·6306		9·8573	0·0017
Intercept2	1	8·3246		13·2926	0·0003
Altitude	1	−2·0866	0·124 (0·028–0·542)	7·6960	0·0055
Dogs	1	−4·3618	0·013 (0·001–0·214)	9·1924	0·0024
Yaks	1	−0·0371	0·964 (0·932–0·996)	4·7423	0·0294
Temperature	1	0·4850	1·624 (0·607–4·344)	0·9338	0·3339

D.F., Degrees of freedom; OR, odds ratio; CI, confidence interval.

although association was not significant. The adjusted R^2 increased to 0·4579.

An additional model was run to examine the interaction effects of altitude and land surface temperature. Results from Table 4 show that yaks and dog were positively correlated with CE prevalence. The interaction of altitude and land surface temperature introduced into this model, made it difficult to understand the role of these variables in the regression model. However, the interaction can improve the prediction of this model by increasing the adjusted R^2 to 0·59.

DISCUSSION

The field survey performed in July 2009, was retrospective. Data were collected up to 3 months after a large number of yaks began to die in spring. Since many survey participants were not able to recall the exact number of yaks infected with CE, we classified the 35 administrative villages into four classes based on the CE prevalence of yaks and other related information that was obtained from participants. It was inevitable that some information was lost. The current

Table 4. *Cumulative logistic regression model with full variables and interactions*

Parameter	D.F.	β	OR (95% CI)	Wald χ^2	$Pr > \chi^2$
Intercept4	1	−3·5500		0·7944	0·3728
Intercept3	1	−1·5196		0·1462	0·7022
Intercept2	1	0·4848		0·0146	0·9038
Altitude	1	2·1080		1·2978	0·2546
Dogs	1	−4·6735	0·009 (0·001–0·196)	9·0448	0·0026
Yaks	1	−0·0286	0·972 (0·932–1·013)	1·7836	0·1817
Temperature	1	5·5078		5·8053	0·0160
Altitude*temperature	1	−2·6363		5·3925	0·0202

D.F., Degrees of freedom; OR, odds ratio; CI, confidence interval.

study followed on from a previous field survey that identified and quantified the role of dogs (0·43 dog per family) and yaks (36 yaks per family) in CE transmission. The current study confirmed that the survey methods used in that field survey were feasible and reliable.

Our results demonstrate that number of dogs and yaks play the most important role in CE transmission, but altitude and land surface temperature may also be involved in CE transmission. Although the association between dogs and CE prevalence was not significant when univariate analysis was used, the multiple regression models confirmed that dogs were positively correlated with CE prevalence. The multiple regression models also demonstrated that there was a positive association between altitude and CE prevalence, and a negative association between land surface temperature and CE prevalence. One explanation for this finding could be that in higher altitudes, there is more livestock production which may promote CE transmission and that lower temperature may prolong the survival of *Echinococcus* eggs [2]. With a higher sample size, this relationship may be better confirmed. When interaction of altitude and land surface temperature was entered into the regression model, the prediction of model was improved. The current results show that CE transmission is a complex process, which needs further study.

In order to indentify and quantify the factors which may influence CE transmission, this study combined landscape ecological data with epidemiological data and tried to consider CE transmission as a whole with spatial techniques. In this study, the SRTM image at 90 m resolution and MODIS images at 1 km resolution used for the extraction of altitude and land surface temperature provided meaningful results. If finer spatial resolutions could be obtained, the results could change. Our study tested the role of altitude and land surface temperature in CE transmission. If more related environmental variables could be studied with advanced spatial techniques, a more comprehensive understanding of CE transmission could be obtained.

Limitations

The findings from this study should be interpreted with caution. Some data in this study were obtained by self-report questionnaire and recall bias is inevitable to some extent. The subjects who agreed to participate in this study could be different from those who did not participate. Last, this study is a local survey and cannot represent all Tibetan areas.

DECLARATION OF INTEREST

None.

REFERENCES

1. **McManus DP, *et al.*** Echinococcosis. *Lancet* 2003; **362**: 1295–1304.
2. **Eckert J, *et al.*** *WHO/OIE Manual on Echinococcosis in Humans and Animals: a Public Health Problem of Global Concern*. World Organization for Animal Health 2001, pp. 101–135.
3. **Torgerson PR, *et al.*** Present situation of cystic echinococcosi in central Asia. *Parasitology International* 2006; **55** (Suppl. 1): 207–212.

4. **Chai JJ,** *et al.* An investigation on the epidemiologic baseline of hydatid disease in Xinjiang, China II. *Echinococcus granulosus* infection in dogs and sheep. *Endemic Diseases Bulletin* 1989; **4**: 9–13.

5. **Craig PS** *et al.* **(eds).** Control of cystic echinococcosis/hydatidosis: 1863–2002. *Advances in Parasitology* 2006; **61**: 443–508.

6. **Giraudous P,** *et al. Echinococcus multilocularis*: why are multidisciplinary and multiscale approaches essential in infectious disease ecology? *Tropical Medicine and Health* 2007; **35**: 293–299.

7. **Graham AJ, Atkinson PM, Danson FM.** Spatial analysis for epidemiology. *Acta Tropica* 2004; **91**: 219–225.

8. **Green RM, Hay SI.** The potential of Pathfinder AVHRR data for providing surrogate climatic variables across Africa and Europe for epidemiological applications. *Remote Sensing of Environment* 2002; **79**: 166–175.

9. **Graham AJ, Danson FM, Pleydell D.** Remote sensing for disease transmission: small mammal and vegetation interactions. *Geoscience and Remote Sensing Symposium* 2003; **7**: 4546–4548.

10. **Danson FM,** *et al.* Landscape dynamics and risk modeling of human alveolar echinococcosis. *Photogrammetric Engineering and Remote Sensing* 2004; **70**: 359–366.

11. **Pleydell DRJ,** *et al.* Landscape composition and spatial prediction of alveolar echinococcosis in Southern Ningxia, China. *PLoS Neglected Tropical Diseases* 2008; **2**: e287.

12. **Graham AJ,** *et al.* Ecological epidemiology: landscape metrics and human alveolar echinococcosis. *Acta Tropica* 2004; **91**: 267–276.

13. **Danson FM, Giraudoux P, Craig PS.** Spatial modeling and ecology of *Echinococcus multilocularis* transmission in China. *Parasitology International* 2006; **55**: s227–s231.

14. **Pleydell DRJ,** *et al.* Mapping HAE disease risk using remotely sensed data. *Geosicence and Remote Sensing Symposium* 2003; **5**: 3362–3364.

15. **Yang YR,** *et al. Echinococcus granulosus* infection and options for control of cystic echinococcosis in Tibetan communities of Western Sichuan Province, China. *PLoS Neglected Tropical Diseases* 2009; **3**: e426.

16. **Budke CM,** *et al.* A canine purgation study and risk factor analysis for echinococcosis in a high endemic region of the Tibetan plateau. *Veterinary Parasitology* 2005; **127**: 43–49.

17. **Li TY,** *et al.* Echinococcosis in Tibetan populations, Western Sichuan Province, China. *Emerging Infectious Disease journal* 2005; **11**: 1866–1872.

18. **Heath D,** *et al.* Control of hydatidosis. *Parasitology International* 2006; **55**: s247–s252.

19. **Bai YN,** *et al.* Survey on cystic echinococcosis in Tibetans, West China. *Acta Tropica* 2002; **82**: 381–385.

20. **Huang Y,** *et al.* Prevalence investigation and evaluation of human echinococcosis in Sichuan Province. *Journal of Preventive Medicine Information* 2012; **28**: 594–596.

21. **He SQ, Zhang FB.** Epidemiology of echinococcosis in Rangtang County. *Guide of China Medicine* 2010; **8**: 144–146.

22. **Hu HH,** *et al.* Study of infection of *Echinococcus granulosus* in yak in spring and its potential role in transmission of cystic echinococcosis in Rangtang County of Sichuan, China. *Biomedical and Environmental Sciences* 2003; **26**: 226–229.