Meteorological conditions, elevation and land cover as predictors for the distribution analysis of visceral leishmaniasis in Sinkiang province, Mainland China

Xiang Gao, Zheng Cao

**HIGHLIGHTS**

- Meteorological factors and land cover may affect the distribution of VL.
- Jackknife method was used to measure the importance of risk variables.
- Temperature and precipitation affect the distribution of VL in Sinkiang province.
- Distribution of VL has been found high in grassland and shrubland area.
- A risk forecast map was built based on the results of MaxEnt model.

**ABSTRACT**

Visceral leishmaniasis (VL) is a fatal disease caused by sandfly-borne protozoa of the *Leishmania* genus. This study explored the influence of environmental factors on the distribution of VL in Sinkiang province, Mainland China, which is a known natural focus of leishmaniasis. Disease identification records were obtained from publicly available data, in which the existence of VL at each geographical location had been recorded as part of the surveillance of leishmaniasis in Sinkiang province. Maximum entropy modelling (MaxEnt) was used to predict the distribution of VL across Sinkiang province, and to match this distribution against environmental variables relating to elevation, climate and land cover, obtained from the WorldClim database, China Meteorological Data Sharing System and the National Geomatic Center of China dataset, respectively. Finally, a regional-scale map was developed to show the potential distribution of VL in the Sinkiang province. Receiver-Operating characteristic (ROC) analysis was used to evaluate the performance of the model. The daily average temperature, maximum temperature of the warmest quarter, daily precipitation and precipitation of the driest month were each found to be predictive of the distribution of VL in Sinkiang. Moreover, we found that presence of VL was significantly influenced by the distribution of grassland and shrubland. The results demonstrate that with proper construction and design, probability surfaces using Maxent can be used as an accurate method by which to predict the distribution of VL in Sinkiang province. The information generated by the model could be used to inform the design of detailed prevention and control strategies for leishmaniasis in this region of Mainland China.

© 2018 Elsevier B.V. All rights reserved.
1. Introduction

Visceral leishmaniasis (VL), which is caused by Leishmania genus, is a fatal global disease. Every year, approximately 200,000 to 400,000 VL cases occur in the world, which result in 20,000 to 40,000 deaths per year (Alvar et al., 2012; WHO, 2010). Due to the expenses incurred in treating the disease, as well as the interruption to development programmes that results from disease outbreaks, VL is also associated with the propagation of poverty (Alvar et al., 2006; Cortes et al., 2012).

Moreover, because there is an overlap in the distribution of leishmaniasis and human immunodeficiency virus (HIV), there has been an emergence of Leishmania/HIV co-infection (Desjeux, 2004). Because this background, leishmaniasis is ranked 9th in the list of global infectious disease and constitutes a serious global public health risk (Dilger, 2013).

VL was a very prevalent infectious disease in Mainland China before the 1950s. Only in 1951, >530,000 VL cases have been reported in this region (Wang et al., 2012). However, as a result of a major national control program, the disease has now effectively been controlled in the east and southeast China (Lun et al., 2015). In the northwest of Mainland China, including Sinkiang province, the disease is still endemic. Previous research has clearly demonstrated that Sinkiang province is a natural focus of VL (Wang et al., 2011).

The female sandfly is the main vector of the Leishmania parasite. Individuals feed on blood from vertebrates as a protein source for egg development and can then infect humans with the blood-borne parasites if they are bitten (Ladopoulos et al., 2015). Phlebotomine sandflies, include Phlebotomus chinesis, Phlebotomus longiductus, Phlebotomus mongolensis, Phlebotomus wuiri and Phlebotomus alexandri are the vector of Leishmania parasite of Sinkiang province. Among them, Phlebotomus wuiri is the main vector of Leishmania parasite in this region (Guan et al., 2000, 2016).

Distribution of Phlebotomine sandflies has been proved to be affected by the landscape (Guernanui et al., 2006). Moreover, the number of endemic regions and prevalence of leishmaniasis have increased in many parts of the world in recent years, primarily due to global warming (Mohebali, 2013). In order to better understand the spread of this disease, it is essential to study the effect of land cover, elevation and meteorological factors on the geographic distribution of VL. Study in this paper aim to explore the impacts of environmental factors on the distribution of VL in Sinkiang province, using published records of disease occurrence between 2010 and 2016. In addition, this study may provide references to neighboring regions because that Sinkiang province sits adjacent to regions that are more traditionally known to be epidemic areas for leishmaniasis (e.g. Afghanistan, Pakistan and India).

2. Material and methods

2.1. Study area and disease occurrence data

Sinkiang province is located in the northwest of Mainland China and borders with Afghanistan, India, Kazakhstan, Kyrgyzstan, Pakistan, Russia and Tajikistan. The region is a natural focus of VL and represents the long-term surveillance of VL across China and is based upon both active and passive monitoring methods. Disease distribution locations for VL were obtained by matching positive test results from both clinical and laboratory examinations. All of the records contained sufficient detailed information on the geographical locality and are able to produce latitude and longitude coordinates by using Google Map.

2.2. Environmental variables

Variables relating to climate and land cover were included in the final model. Meteorological data came from the Chinese Meteorological Data Sharing Service System (http://data.cma.cn/) with a spatial resolution of 1 km² and in a gridded data form. Meteorological Data can be assigned to every county, all VL reported point took climatic information from the grid point closest to their administrative centroids. Meteorological factors that were obtained from Chinese Meteorological Data Sharing Service System can be seen in Table 1.

Spearman’s correlation coefficients were used to exam correlated meteorological variables. Correlated variables have not been included in the same model. Adjusted Akaike’s information criterion (AICc) values were calculated to assess every meteorological variable’s possible contributions to the models. For the correlated variables, the one with the lowest AICc score can be included in the finally model (James et al., 2015).

Elevation data was obtained from the WorldClim website (http://www.worldclim.org) with a spatial resolution of 1 km². Land cover data come from the National Geomatic Center of China (http://www.webmap.cn/mapDataAction.do?method=globalLandCover) with a spatial resolution of 1 km² (Table 2).

<table>
<thead>
<tr>
<th>Land cover type Abbreviation Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water L1 Land covered with rivers, lakes, reservoir and pond (with area &gt;0.1 square meters)</td>
</tr>
<tr>
<td>Wetland L2 Land covered with swamp, flooding wetlands, forest/shrubland, peat bog and salt marsh</td>
</tr>
<tr>
<td>Needleleaf forest L3 Land covered with needle-leaved trees, with vegetation cover over 30%</td>
</tr>
<tr>
<td>Broadleaf forest L4 Land covered with broad-leaved trees, with vegetation cover over 30%</td>
</tr>
<tr>
<td>Theropencedrymion L5 Land covered with needle-wide mixed forest, with vegetation cover over 30%</td>
</tr>
<tr>
<td>Wooded grassland L6 Land covered with trees, with vegetation cover between 10 and 30%</td>
</tr>
<tr>
<td>Grassland L7 Land covered by natural grass with cover 10%</td>
</tr>
<tr>
<td>Shrubland L8 Land covered by natural shrubs with cover 10%</td>
</tr>
<tr>
<td>Cropland L9 Land used for agriculture</td>
</tr>
<tr>
<td>Bare ground L10 Land with vegetation cover lower than 10%</td>
</tr>
<tr>
<td>Permanent ice and snow L11 Lands covered by permanent snow, glacier and icecap</td>
</tr>
<tr>
<td>Urban and Built L12 Lands modified by human activities</td>
</tr>
</tbody>
</table>

Table 1

<table>
<thead>
<tr>
<th>Meteorological factor</th>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily average temperature</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>Daily mean relative humidity</td>
<td>H</td>
<td></td>
</tr>
<tr>
<td>Daily precipitation</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>Daily mean wind speed</td>
<td>W</td>
<td></td>
</tr>
<tr>
<td>Minimum temperature of the coldest quarter</td>
<td>Tmin</td>
<td></td>
</tr>
<tr>
<td>Maximum temperature of the warmest quarter</td>
<td>Tmax</td>
<td></td>
</tr>
<tr>
<td>Precipitation of the wettest month</td>
<td>Rmax</td>
<td></td>
</tr>
<tr>
<td>Precipitation of the driest month</td>
<td>Rmin</td>
<td></td>
</tr>
</tbody>
</table>
2.3. Modelling of VL distribution

Maximum entropy modelling was used to predict the distribution of VL across Sianking province (Maxent, version 3.3.1, http://www.cs.princeton.edu/~schapire/maxent/) (Phillips et al., 2006). In the analysis, Maxent divided the VL presence data into two parts: 75% were used to assess the effect of environmental variables on the presence of VL and the rest 25% was used to examine the predictive capacity of the model. Jackknife testing was used to determine which variables contributed most to the distribution of disease. Variables were considered to be the important variables when they produce high training gains or reduce the training gain when left out of the model (Kantzoura et al., 2011; Slater, 2012).

To assess the performance of Maxent, partial Receiver operating characteristic (partial ROC) analysis was performed. The Area Under the ROC Curve (AUC) is calculated and its value varies from 0.5 to 1, with a value >0.8 indicating high performance of the model (Lobo and Real, 2008; Phillips and Dudík, 2008; Khanum et al., 2013). A bootstrap algorithm was performed to assess the error related to these values, where we compute partial AUC for each sample and sample 200 times with replacement from the testing data.

2.4. Validation of the predictive model

To validate the performance of the predictive model, a follow-up survey was conducted to record the presence of VL in Sinkiang Province of Mainland China during 2016: 8 records of absence of VL and 8 records of distribution of VL were used for this purpose. Then, Cohen’s kappa coefficient was calculated to assess the consistency between the distribution predictions of Maxent and those obtained in reality from the follow-up survey. A Cohen’s kappa coefficient between 0 and 0.20 was taken to indicate very poor consistency between the model and reality; a coefficient between 0.21 and 0.40 was taken to indicate consistency was weak; a coefficient between 0.41 and 0.60 was taken to indicate consistency was moderate; a coefficient between 0.61 and 0.80 was taken to indicate consistency was good, and; a coefficient between 0.81 and 1.00 was taken to indicate consistency was very good (Quintana et al., 2013).

3. Results

3.1. Importance of environmental variables for predicting VL distribution

Based on VL surveys in Sinkiang published between 2010 and 2015, a total of 105 specific data points comprising VL distribution data in Sinkiang province were identified for the present analysis (Fig. 1).

Through Spearman correlation analysis, we found that rainfall and relative humidity are correlated variables (Spearman’s $r > 0.7$). Rainfall had a lower AICc score and thus was included in the further analyses. According to the results of the Maxent modelling and subsequent jackknife testing, the environmental variable that achieved the high training gain when only a single variable was included in the model were the maximum temperature of the warmest quarter, the precipitation of the driest month, daily average temperature and daily precipitation. The results of jackknife test also indicated that the distribution of grassland and shrubland significantly influence the presence of VL (Fig. 2).

3.2. VL prediction map

Fig. 3 shows the predictions generated by the final model regarding the spatial characteristics of the distribution of VL in Sianking. From this, it can be seen that the risk of the distribution of VL is highest in the northwest and east Sinkiang, and are lowest in the central region of Sinkiang.

3.3. Performance of the predictive model

Partial ROC analysis showed that the values of partial AUC of predictive model in this paper were uniformly higher than the random classifier with a ratio of 1 ($P < 0.01$).

3.4. Validation of the predictive model

Fig. 3 shows the predictions of the model regarding the distribution of VL and the observed distribution as obtained by our follow-up survey. Comparison of the two revealed that the model correctly predicted 81.25% of the new records obtained in the follow-up survey: 7 out of 8
of the presence records and 6 out of 8 of the absence records. The Cohen’s kappa coefficient was calculated as 0.625, which proved a good consistency between the model’s predictions and the field observations from the follow-up survey.

4. Discussion

This study attempted to explain and predict distribution of the VL in relation to landscape, climate and elevation. Our models indicate that the maximum temperature of the warmest quarter, the daily average temperature, the average precipitation of the driest month, the daily average precipitation and the distribution of grassland and shrubland may each have a significant influence on the distribution of VL. Maxent was used to investigate the effect of elevation, meteorological factors and land cover on the presence of VL and predict the distribution of VL in Sinkiang province of Mainland China in this study. Maxent is well known as a presence-only ecological niche model that, using limited input data, can make highly-accurate distribution predictions (Matyukhina et al., 2014; Gao et al., 2017). A study by Stockwell and Peterson (2002) demonstrated that the model could make accurate species distribution predictions when only 50 data points were included.

The study of Zhao et al. (2015) indicated that the bite of phlebotomine sandfly is the major risk factor of VL outbreak in Sinkiang. On the one hand, Sinkiang is one of the major pastoral areas of Mainland China. Most of the residents in local are herdsmen. Grassland and shrubland is suitable for grazing. Weather conditions can affect the growth of pasture and then affect the grazing and migration of herdsmen. Thus, grassland, shrubland and climate variables may influence the contact between herdsmen and phlebotomine sandfly.

On the other hand, climate variables have a profound effect on the biology and behavior of phlebotomine sandfly (Teles et al., 2013). A

Fig. 2. Relative importance of risk variables in predicting the probability of VL presence, as determined by Jackknife analysis. Variables producing higher trainings gain are considered to be more important (more predictive). L1-Water; L2-Wetland; L3-Needleleaf Forest; L4-Broadleaf Forest; L5-Therophytes; L6-Grassland; L7-Shrubland; L8-Cropland; L9-Bare Ground; L10-Permanent ice and Snow; L11-Urban and Built; T-Daily average temperature; R-Daily precipitation; W-Daily mean wind speed; Tmin-minimum temperature of coldest quarter; Tmax-maximum temperature of warmest quarter; Rmax-precipitation of wettest month; Rmin-precipitation of driest month; Total-all risk variables.

Fig. 3. Risk map of VL distribution predicted by Maxent model and potential distribution of VL and observation data of VL in the Sinkiang province of Mainland China. The gray rectangles are the records of absence of VL, and the black rectangles are the records of presence.
study by Alten et al. (2003) observed that at 32°C with a relative humidity of 65 to 75%, the adult phlebotomine sandfly emerged within about 28 days, whilst at 18°C it took approximately 245 days for adults to emerge, and no emergence was observed at 15°C. Cross and Hyams (1996) conducted a temperature test on phlebotomine sandfly and found that all individuals died within 2 h at temperatures above 40°C combined with a relative humidity below 33%. They also reported that temperatures below 10°C are unfavorable for the survival of larvae. As well as affecting growth and mortality, environmental temperature also controls the seasonal activity of phlebotomine sandfly. Boussaa et al. (2005) found sandflies to be active when temperatures fluctuate between 11 and 36°C. A further study of Faraj (2011) in the Asir region of Saudi Arabia found that in that environment phlebotomine sandfly were only active from June to September and this activity was significantly influenced by the environmental temperature.

From the above, it is clear that temperature and relative humidity influence phlebotomine sandfly population growth, development and activity. Collectively, the previous study suggests that any increase in temperature between a range of 10°C and 40°C, combined with a relative humidity above 33%, is likely to increase the prevalence of phlebotomine sandfly populations by speeding up growth and development (Cross and Hyams, 1996). In Sinkiang province, the precipitation of the driest month and the maximum temperature of the warmest month is between a range of 10°C and 40°C, combined with a relative humidity between 31% and 71%.

Spending between sand fly and livestock feces in the lab, and other studies have reported strong associations of phlebotomine sandfly with livestock and vegetation. This is due to the fact that the sandfly has an association with livestock and vegetation for survival. This means that they require feeding predominantly on blood meals as their food, shading and protection. A number of studies have reported rodent and livestock feces to be active when temperatures fluctuate between sand fly and livestock feces and troughs in different parts of the world (Dilger, 2013). In Sinkiang province, as in other regions of northwest mainland China, cattle and sheep are the main livestock animals that graze on the grassland habitat, therefore their feces are more likely to be found in grassland habitats.

Another aspect of the zoonotic nature of leishmaniasis is the involvement of rodent habitats as a reservoir for the phlebotomine sandfly. Rodents predominantly dig their burrows next to vegetation as a source of food, shading and protection. A number of studies have reported strong associations between sandfly and livestock feces and troughs in different parts of the world (Dilger, 2013).

Finally, it is also worth noting that whilst female sandfly depend mostly on blood meals as their food, male sandfly depend mainly on soft stemmed edible plants for survival. This means that they require vegetation cover nearby in order to survive (Wasserberg et al., 2003). This occurrence of males in vegetative habitats exerts a further influence on the distribution of females, in terms of enabling mating and breeding behaviors.

From the distribution prediction map generated by the developed model, it was evident that the risk of the distribution of VL disease was highest in the northwest and eastern Sinkiang, and was lowest in the central region of this province. This most likely because the northwest and east Sinkiang are dominated by grassland, whilst the central region of Sinkiang is mostly desert.

5. Conclusion

Maxent and jackknife testing of VL distribution data against defined environmental variables produced a predictive map of the distribution of VL in Sinkiang province and indicated that temperature, precipitation and land cover all independently influenced the distribution of VL. This information could be used to inform the design of more detailed surveillance programmes and more evidence-based planning for the control of VL in the Sinkiang region, making better use of the limited availability of human resources and funding for such programmes. In addition, the results will improve scientific understanding of the risks of the spread of VL. Further research that incorporates socio-economic information into the model would further enhance predictions about likely changes in the distribution of VL in the Sinkiang region.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgment

Special thanks go to researchers who made a great contribution to this study reporting distribution of VL in the Sinkiang province of Mainland China. Research reported in this study was supported by the University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province under grant number UNPYSCT-2017006.

References


Slater, Claire H., 2012. Spatial Epidemiology and the Integrated Control of Malaria and Lymphatic Filariasis in Africa. Imperial College London.


