International Journal of Infectious Diseases 26 (2014) 1–8

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Ecological niche modeling for predicting the potential risk areas of severe fever with thrombocytopenia syndrome

Zhaohui Du [a], Zhiqiang Wang [b], Yunxia Liu [a], Hao Wang [a], Fuzhong Xue [a], Yanxun Liu [a,*]

[a] Department of Epidemiology and Biostatistics, School of Public Health, Shandong University, PO Box 100, 44 West Wenhua Road, Jinan 250012, Shandong, China

[b] Shandong Center for Disease Control and Prevention, Jinan, Shandong, China

ARTICLE INFO

Article history:
Received 29 November 2013
Received in revised form 17 March 2014
Accepted 9 April 2014

Corresponding Editor: Eskild Petersen, Aarhus, Denmark

Keywords:
SFTS
Ecological niche model
MaxEnt
Geographical information system

SUMMARY

Background: Severe fever with thrombocytopenia syndrome (SFTS) is an emerging infectious disease caused by a novel bunyavirus. The spatial distribution has continued to expand, while the areas at potential high risk of SFTS have, to date, remained unclear.

Methods: Using ecological factors as predictors, the MaxEnt model was first trained based on the locations of human SFTS occurrence in Shandong Province. The risk prediction map of China was then created by projecting the training model onto the whole country. The performance of the model was assessed using the known locations of disease occurrence in China.

Results: The key environmental factors affecting SFTS occurrence were temperature, precipitation, land cover, normalized difference vegetation index (NDVI), and duration of sunshine. The risk prediction maps suggested that central, southwestern, northeastern, and the eastern coast of China are potential areas at high risk of SFTS.

Conclusions: The potential high risk SFTS areas are distributed widely in China. The epidemiological surveillance system should be enhanced in these high risk regions.

1. Introduction

Severe fever with thrombocytopenia syndrome (SFTS) is an emerging infectious disease caused by a novel *Phlebovirus* genus in the *Bunyaviridae* family, which was first identified in 2009 in rural areas of Hubei and Henan in China.[1] The clinical presentations of most patients are characterized by symptoms including high fever, thrombocytopenia, leukocytopenia, gastrointestinal symptoms, and lymphadenopathy. Although most patients experience no symptoms during the incubation period, the clinical condition in patients with severe disease may progress quite rapidly and end in multiorgan dysfunction and intravascular coagulation.[2] The disease showed an unusually high fatality rate of 30% when first identified,[1] and subsequent reports from several regions have illustrated a case fatality rate ranging from 0.0% to 23.8%.[3] The clinical symptoms of SFTS are nonspecific and are difficult to distinguish from those of human anaplasmosis, hemorrhagic fever with renal syndrome, and leptospirosis. Thus, SFTS is likely to be a fatal disease if left untreated or treated inappropriately, especially when confused with other diseases.

The limited published data on the epidemiological characteristics show that the majority of SFTS patients have been farmers living in wooded and hilly areas with a history of outdoor work before disease onset. The disease shows an obvious seasonal variation, starting around March, peaking between May and July, and ending around November.[3] SFTS cases first emerged in Henan Province in 2007, and when the virus was first identified in 2009, six provinces in Central and Northeast China had discovered SFTS cases; by the end of 2011, at least 11 provinces in China had reported the occurrence of SFTS according to the surveillance data, showing an expanding tendency of the endemic areas. SFTS patients have recently been reported in South Korea and Japan.[4,5]

Although several studies have reported that there is a risk of SFTS virus (SFTSV) being transmitted from person to person through contact or exposure to patient blood,[6–8] most SFTS cases have shown a sporadic distribution without evidence of human-to-human transmission. The tick is believed to be a candidate vector species of SFTSV, and *Haemaphysalis longicornis*, which has been

* Corresponding author. Tel.: +86 531 8838 0280; fax: +86 531 8838 0280.
E-mail address: liu-yx@sdu.edu.cn (Y. Liu).

http://dx.doi.org/10.1016/j.ijid.2014.04.006

detected as carrying SFTSV RNA, is dispersed widely in mainland China and has a broad host range.[9] Therefore, it has been classified as a tick-borne disease of natural infectious focus that may be affected by ecological factors.[1]

Given the fact that neither specific treatment nor a vaccine is yet available, SFTS may pose a huge threat to public health, especially for farmers and travelers. Consequently, it is very important to identify those areas at potential high risk of the disease in China in order to conduct effective regional prevention and control strategies.

In this study, we sought to identify the potential high risk areas for SFTS in China using an ecological niche model (ENM) under the framework of a geographic information system (GIS). The ENM was developed based on the ecological niche theory.[10] The ecological niche of a species, which was first presented by Joseph Grinnell in 1917, is the set of conditions permitting it to maintain populations without the immigration of individuals from other areas.[11,12] The ecological niche model contains a suite of models that can develop a quantitative picture, characterizing a species' potential distribution, based on the idea that the known occurrence of a species can be related to raster geographic datasets summarizing the factors that likely influence the suitability of the environment for the species across the occurrence landscapes.[13] It has been utilized by researchers as a powerful tool to characterize the ecological distribution of species, to identify the habitats of undocumented diseases, and thus to anticipate the areas at high risk for disease occurrence.[14] By using the maximum entropy ENM[15] with MaxEnt software version 3.3.3k (http://www.cs.princeton.edu/~schapire/maxent/), the initial ENM was first trained based on the locations of human SFTS occurrence and ecological variables in Shandong Province. The predictive potential high risk map of mainland China was then created by projecting the training model onto the whole country. Furthermore, the performance of the model was assessed using the known locations of disease occurrence in China.

2. Materials and method

2.1. Human SFTS occurrence data

Shandong Province, located in eastern China between latitude 34°25′–38°23′ and longitude 114°36′–112°43′, which has been reported to be a main endemic region of SFTS, was selected as the training area to construct the initial training MaxEnt model. Data of known human SFTS cases in Shandong Province during January 2010 to April 2013 were obtained from the China Information System for Disease Control and Prevention (CISDCP). Data included sex, age, residential address, and personal disease characteristics. In this study, a human SFTS case was defined in the presence of both clinical diagnosis criteria and laboratory confirmation criteria, in accordance with the Chinese SFTS diagnosis and treatment guidelines (Ministry of Health, China, 2010). SFTS cases were first detected by a sentinel surveillance system in the local medical institutions and were quickly reported, within 24 h, through a national internet-based surveillance system for communicable diseases (http://1.202.129.170/). All records were georeferenced using the patient's residential village center as the occurrence point by consulting Google Maps. Three hundred and twelve occurrence locations with 416 SFTS cases were used to train the model. Provinces with known SFTS occurrence records in China were identified by referring to several published articles.[1,16,17]

2.2. Environmental data

Environmental data from within China were downloaded as detailed below.

Climate data were derived from the China Meteorological Data Sharing System (http://www.cdc.cma.gov.cn/home.do), which included climate data from 722 ground meteorological stations throughout mainland China during 1990–2011. Of the numerous monthly and annual variables in the datasets, we chose annual mean temperature, annual mean precipitation, annual mean relative humidity, annual mean air pressure, annual mean vapor pressure, and annual mean solar radiation hours to represent annual average climate characteristics. The January mean temperature, January mean precipitation, January mean relative humidity, July mean temperature, July mean precipitation, and July mean relative humidity were recorded to standardize the seasonal variation in meteorological conditions, on the basis of suggestions from researchers and field epidemiologists who have studied SFTS for years, as well as the results from some of our initial model selection. To meet the data format requirement of the MaxEnt software, the kriging interpolation method[18] was used to generate 1-km resolution raster layers for each of the climatic variables based on the weather station point data by ArcGIS 9.3.

Elevation data were obtained from WorldClim (http://www.worldclim.org).[19]

The European Space Agency (ESA) GlobCover data, a global land cover (LC) layer with native resolution of 300 m, was provided by the Database of Global Change Parameters, Chinese Academy of Sciences (http://globalchange.nsdc.cn); this was the only categorical variable in the study. It was then re-sampled to a pixel resolution of 1×1 km and cut using a mask of mainland China in order to keep the same resolution and geographic extent with other layers.

The monthly maximum normalized difference vegetation index values (NDVI) were drawn from the same source as the land cover layer. All the raster layers with the original geographic extent of China were further extracted by the mask of Shandong Province to build the database of Shandong, which was used to train the model. Both the variable sets of Shandong Province and mainland China were converted to the ASCII raster data format required by MaxEnt using ArcGIS 9.3.

2.3. Training model construction

MaxEnt is a machine learning approach that is suitable for presence-only datasets to model the potential distributions of species.[21,22] It utilizes a maximum entropy algorithm to estimate the probability distribution of a species.[15,24] We used the program MaxEnt version 3.3.3k (http://www.cs.princeton.edu/~schapire/maxent/) to build the training model for SFTS in Shandong Province. Within MaxEnt processing, 312 occurrence locations were partitioned as follows: 75% (234 locations) were randomly selected for constructing the model and the remaining 25% (78 locations) were set aside for external validation. To obtain a robust model, we developed 10 replicated models based on independent random partitions, and we used the median output grids as the final predictive potential high risk map of SFTS and imported it into ArcGIS 9.3 to visualize the risk map.

Response curves outputted from MaxEnt 3.3.3k were used to elucidate the range of each environmental condition that was suitable to the occurrence of SFTS. Furthermore, the logistic output format of the predictive map with probability values ranging from 0 (impossible occurrence area) to 1 (optimal occurrence area) was used to visualize the potential risk of SFTS. The 'balanced' threshold value was chosen as the risk cut-off distinguishing potential presence areas from potential absence areas.[15,22,24,25]

2.4. Model evaluation

The threshold-dependent binomial test using the extrinsic omission rate as statistic was performed to detect the statistical