Spatial and Spatio-temporal Epidemiology 19 (2016) 125–135

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

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

Modeling of spatio-temporal variation in plague incidence in Madagascar from 1980 to 2007

Emanuele Giorgi[a,*], Katharina Kreppel[b], Peter J. Diggle[a], Cyril Caminade[c], Maherisoa Ratsitorahina[d], Minoarisoa Rajerison[d], Matthew Baylis[c,e]

[a] *Lancaster Medical School, Lancaster University, Lancaster, UK*
[b] *Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK*
[c] *Institute of Infection and Global Health, Department of Epidemiology and Population Health, University of Liverpool, Leahurst Campus, Neston, Cheshire, UK*
[d] *Unité Peste, Institut Pasteur de Madagascar, Antananarivo, Madagascar*
[e] *Health Protection Research Unit for Emerging and Zoonotic Infections, University of Liverpool, UK*

A R T I C L E   I N F O

Article history:
Received 16 May 2016
Revised 12 September 2016
Accepted 12 October 2016
Available online 18 October 2016

Keywords:
Cox process
Distributed lag-model
Gaussian process
Plague
Spatio-temporal point pattern

A B S T R A C T

Plague is an infectious disease caused by the bacterium *Yersinia pestis*, which, during the fourteenth century, caused the deaths of an estimated 75–200 million people in Europe. Plague epidemics still occur in Africa, Asia and South America. Madagascar is today one of the most endemic countries, reporting nearly one third of the human cases worldwide from 2004 to 2009. The persistence of plague in Madagascar is associated with environmental and climatic conditions. In this paper we present a case study of the spatio-temporal analysis of plague incidence in Madagascar from 1980 to 2007. We study the relationship of plague with temperature and precipitation anomalies, and with elevation. A joint spatio-temporal analysis of the data proves to be computationally intractable. We therefore develop a spatio-temporal log-Gaussian Cox process model, but then carry out marginal temporal and spatial analyses. We also introduce a spatially discrete approximation for Gaussian processes, whose parameters retain a spatially continuous interpretation. We find evidence of a cumulative effect, over time, of temperature anomalies on plague incidence, and of a very high relative risk of plague occurrence for locations above 800 m in elevation. Our approach provides a useful modeling framework to assess the relationship between exposures and plague risk, irrespective of the spatial resolution at which the latter has been recorded.

1. Background

Plague, also known as the “Black Death” during the fourteenth century, is a vector-borne disease caused by the zoonotic bacterium *Yersinia pestis*, usually found in small mammals, especially rodents. It can be transmitted to humans through the bite of an infected flea or, in some cases, through direct contact, inhalation or digestion of infected material. Three different forms of plague can be distinguished: bubonic plague, which is the most common form and characterized by swollen lymph nodes, called “buboes”; septicaemic plague, when the infection spreads from the buboes through the bloodstream; pneumonic plague, the most virulent form, that occurs when the infection, usually from advanced bubonic plague, spreads to the lungs. If diagnosed in time, plague can be effectively treated with antibiotics. If untreated, it can be rapidly fatal.

* Corresponding author.
E-mail address: e.giorgi@lancaster.ac.uk (E. Giorgi).

http://dx.doi.org/10.1016/j.sste.2016.10.001
1877-5845/

Today, plague is mainly present in South America, Asia and Africa. Madagascar is one of the most endemic countries, accounting for almost one third of the human cases worldwide from 2004 to 2009 (WHO, 2009). Plague reached the coast of Madagascar in 1898 with steamboats from India (Brygoo, 1966), following the third pandemic that originated in China in 1855. It then spread to the Central Highlands in 1921 where it became endemic. The main host for plague on the island is the black rat, with its fleas acting as the vector. The persistence of plague foci in Madagascar is associated with climatic, environmental and sociological factors (Ben Ari et al., 2011) all of which affect the distribution of flea and rodent populations, and their contact with humans. Among the climatic factors, temperature, rainfall and relative humidity are considered to affect the disease dynamics, mostly through their effects on the survival and reproduction of host and vector (Stenseth et al., 2006). In Madagascar, plague is mainly found in rural areas that provide the preferred habitat of the main host for *Y. pestis*, the black rat *Rattus rattus*. Traditional burial customs and lack of access to affordable health care also impact on the spread and persistence of the disease, but data on these factors is currently lacking. Understanding how all these factors interact and jointly affect human plague incidence requires the use of appropriate multivariate statistical methods.

In this paper we analyze monthly data on district-level plague incidence in Madagascar from January 1980 to December 2007. Motivated by Diggle et al. (2013), we model monthly district-level plague incidence as an aggregated outcome of a spatio-temporally continuous process. The objective of our analysis is to quantify the effects of temperature and precipitation anomalies (henceforth TAs and PAs), and of elevation, on plague incidence whilst accounting for unexplained spatio-temporally structured variation. One of the main features of the analyzed data is that temporal variation in incidence dominates its spatial variation, making the joint spatio-temporal analysis of the data computationally cumbersome. To address this issue, we carry out marginal temporal and spatial analyses based on a spatio-temporal log-Gaussian Cox process model. For the spatial analysis, we also introduce a spatially discrete approximation to the underlying continuous Gaussian process, for which inference is computationally more feasible, but whose parameters retain a spatially continuous interpretation.

Kreppel et al. (2014) have previously analyzed country-level plague incidence anomalies from 1960 to 2007. They used wavelet analysis to study the relationship between plague and climatic anomalies caused by the El Niño Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD). They found evidence of a temporally varying relationship between plague incidence anomalies and TAs, PAs, ENSO and IOD. However, one of the main limitations of their approach is that only marginal associations between plague and climatic risk factors were investigated. Additionally, in their analysis plague incidence was treated as a continuous outcome, which is questionable in view of the low case-counts that are observed throughout the time series. In the review paper by Andrianaivoarimanana et al. (2013), maps of empirical plague incidence show its restriction mainly to the Central Highlands districts. Based on summary descriptive statistics, the authors also mention a threshold of 800 m in elevation as critical in determining the geographical limits of plague foci. The modeling approach we propose in this paper overcomes the inherent limits of the analysis by Kreppel et al. (2014), and also allows us to validate the statement by Andrianaivoarimanana et al. (2013) on the effect of elevation, in a statistically rigorous fashion.

The structure of the paper is as follows. In Section 2, we briefly describe the different data sources for plague incidence, elevation, TAs, PAs and population density. In Section 3, we first carry out an exploratory analysis to highlight the main features of the data, then develop a spatio-temporal log-Gaussian Cox process for plague incidence and fit the model using Bayesian inference. In Section 4, we report the results of our analysis and check the validity of the modeling assumptions. Section 5 is a concluding discussion.

2. Data

We use the following data sources, including climatic, environmental and demographic datasets, to obtain the relevant explanatory variables and offset quantities.

- *Plague cases.* We use data on all cases of human plague confirmed by a bacteriological test from January 1980 to December 2007. These data were made available by the World Health Organisation Plague Reference Laboratory of the Institut Pasteur de Madagascar.
- *Population density.* Spatial population density estimates for Madagascar for the year 2010 were derived from the GPWv4 dataset, released by the Center for International Earth Science Information Network, Columbia University (Doxsey-Whitfield et al., 2015). We use 5-yearly country-level human population growth estimates from the United Nations to estimate the total population of Madagascar from 1980 to 2007, based on the latest population census conducted in 1993.
- *Temperature and precipitation.* Monthly average values of temperature and precipitation were obtained from the NCEP-NCAR reanalysis (Kalnay et al., 1996). These were available as averaged values over a large domain including Madagascar [42.18° E–49.68°E; 24.76°S – 11.42°S].
- *Elevation.* Spatial estimates of elevation were obtained from the ETOP01 dataset made available by the National Oceanic and Atmospheric Administration (Amante and Eakins, 2009).

3. Model formulation

3.1. Exploratory analysis

In Madagascar, plague is mainly found on the Central Highlands but, as with all infectious diseases, it can also travel to non-endemic areas with its host. Cases of disease may be recorded, therefore, outside of plague's natural (endemic) range. In Fig. 1, the average annual incidence of plague, from 1980 to 2007, is shown for each district,