Human-Altered Landscapes and Climate To Predict Human Infectious Disease Hotspots

Soushieta Jagadesh (✉ soushieta.jagadesh@usys.ethz.ch)  
ETH Zurich: Eidgenossische Technische Hochschule Zurich  https://orcid.org/0000-0003-2985-7430

Marine Combe  
Institut de recherche pour le developpement

Mathieu Nacher  
Andree Rosemon Hospital Centre: Centre Hospitalier Andree Rosemon

Rodolphe Elie Gozlan  
IRD: Institut de recherche pour le developpement

Research Article

Keywords: Emerging Infectious Diseases, human-altered landscapes, Biogeography, Topography, Climate, Land-use

Posted Date: December 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1104482/v1

License: ☒ ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background

Zoonotic diseases account for more than 70% of emerging infectious diseases. Due to their increasing incidence, and impact on global health and economy, anticipating the emergence of zoonoses is a major public health challenge. Here, we use a biogeographic approach to predict future hotspots and determine the factors influencing disease emergence. We have focused on three viral disease groups of concern: Filoviridae, Coronaviridae, and Henipaviruses.

Methods

We modelled presence-absence data in spatially explicit binomial and zero-inflation binomial logistic regression with and without autoregression. Presence data were extracted from published studies for the three EID groups. Various environmental and demographical rasters were used to explain the distribution of EIDs. True Skill Statistic and deviance parameters were used to compare the accuracy of the different models.

Results

For each group of viruses, we were able to identify and map areas at high risk of disease emergence based on the spatial distribution of disease reservoirs and hosts, as well as data on the distribution of each disease. Common influencing drivers are climatic covariates (minimum temperature and rainfall) and human-induced land modifications.

Conclusions

Using topographical, climatic and previous disease outbreaks reports, we show that we can identify and predict future high-risk areas for disease emergence, such as the current COVID-19 pandemic, and their specific underlying human and environmental drivers. We suggest that such a predictive approach to EIDs should be carefully considered in the development of active surveillance systems for pathogen emergence and epidemics at local and global scales.

Introduction

Shifting geographical footprints of pathogens and/or infected hosts due to ecosystem disruption lead to emerging infectious diseases (EIDs) (Ogden, AbdelMalik and Pulliam, 2017), of which COVID-19 is a current example at the centre of international attention. Infectious diseases of animal origin (or zoonoses) account for more than 70% of emerging infectious diseases in recent decades (Taylor, Latham and Woolhouse, 2001; Jones et al., 2008). With the onset of SARS-CoV-2, anticipating the emergence of
new pathogens has become the major public health challenge of our time. The spatial dynamics of zoonotic diseases make it difficult to study and detect hotspots, as they depend on the spatial distribution of mammalian hosts and reservoirs and their interaction with humans (Meentemeyer, Haas and Václavík, 2012). Studies show that disease emergence is closely linked to human-modified landscapes, such as fragmented peri-urban forests, which disrupt the human-animal-environment interface (Myers et al., 2013; Daszak, Olival and Li, 2020; Gibb et al., 2020). Therefore, ecological processes, landscape alterations, especially agricultural development, changes in water ecosystems, deforestation and reforestation, (Patz et al., 2004; Meentemeyer, Haas and Václavík, 2012) and climate change (Patz et al., 2008) are the main drivers of EIDs.

Natural landscape attributes like elevation, and human-modified landscape factors, such as deforestation and agricultural expansions, influence the spatial extent of hosts and reservoirs. For example, high elevation and water bodies could function as geographical barriers by preventing host movement (Smith et al., 2002), while rapid landscape changes can increase the likelihood of contact with a reservoir host and thus promote the emergence of micro-organisms previously unknown to humans (Fenollar and Mediannikov, 2018). Implementing a landscape approach to detect EID risk requires the integration of spatial complexity using geographic information systems (GIS) and remote sensing data with the distribution to EID and its immediate environment (Biek and Real, 2010). Mathematical models help to integrate spatial data to measure the predictive risk of disease emergence. Recent work has shown that using a spatial Bayesian framework on species distribution modelling (SDM) produces more accurate results when dealing with limited and clumped data and takes into account random effects, thus providing better results on landscape factors influencing risk (Redding et al., 2017; Jagadesh et al., 2019). Hierarchical Bayesian SDMs allow observations to be interpreted as the result of ecological processes such as climate change and human-altered landscapes.

Here, we have mapped and compared the predictive risk of three viral epidemics of infectious diseases transmissible to humans that are under surveillance: Filoviridae, including Ebola and Marburg viral diseases (EVD & MVD), Coronaviridae such as SARS, MERS and COVID-19, and Henipaviruses (Nipah & Hendra diseases) of the Paramyoviridae family. We are also identifying potential hotspots and quantifying the importance of environmental factors such as climate and human-induced landscape on the emergence of viral infectious disease.

**Methods**

Using a Bayesian framework, we modeled the presence-absence data using a two stage spatially explicit hierarchical logistic regression (Gelfand et al., 2006). First, we model the potential presence of EID occurrence in each cell grid (local) of bioclimatic variables and population density variables using disease-level coefficients, and a spatial random effect. Then, if the frequency of presence/absence data at each site follows a binomial distribution, the sampling intensity of each site varies accordingly. Once the models were fitted, we compared the different models based on parameter summaries and model deviance.
Presence-absence data

When considering presence-absence data for SDM, the most common biases arise from the assumption of perfect detection and stationary hosts. Disease occurrence depends on the spatial distribution of the disease reservoir and intermediate hosts. We used zero-inflation binomial models (Latimer et al., 2006) to recognize the imperfect detection of occurrence. Autocorrelation and non-stationarity of mammalian hosts were accounted using intrinsic conditional autoregressive models (iCAR) to avoid overestimation of spatial inference and prediction in the models. We extracted distributional data on the global occurrence of Filoviridae, Coronaviridae and Henipavirus human disease outbreak over time from WHO archives and published studies (Supplementary Table.1). In cases where the origin of the outbreaks was unclear, we restricted the analysis down to the region or district of origin. Laboratory outbreaks, outbreaks resulting in asymptomatic disease (Reston Ebola disease in the Philippines) and domestic (Hendra outbreaks in horses) and wildlife (Ebola in gorilla populations) outbreaks were also excluded. We excluded the recent SARS-CoV-2 outbreak as the origin of the infection remains controversial. The analysis and coordinates of the suspected origin of SARS-CoV-2 are included in the Supplementary material. We geo-referenced the sites of origin and constructed spatial buffers of 10km around the geographical coordinates. For each group of viruses, we generated 500 random spatial points in the spatial buffers to constitute presence points. Pseudo-absences were randomly generated in the spatial extents of the reservoirs and intermediate hosts of each virus group in a 1:2 ratio, leading to 1000 absence points.

Bioclimatic and population predictor datasets

We extracted climatic and elevation covariates such as monthly maximum and minimum temperatures, rainfall and elevation from global Bioclim data (Fick and Hijmans, 2017) at a spatial resolution of 2.5 minutes or about 4.5km at the equator. We used Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover Type (MCD12Q1) data (Broxton et al., 2014) and land use changes included in human activities from the Global Human Modification of Terrestrial Systems data set (Kennedy et al., 2020). Finally, we used Gridded Population of the World, Version 4 (GPWv4) for the human population density raster (University, 2018). The raster layers were resampled to a fixed resolution of 4.5km and stacked to a raster brick. We obtained the geographical distribution and spatial extent of the primary hosts and reservoir mammals from the IUCN red list (Isberg, Balaguera-Reina and Ross, 2017) and the list of mammals is included in Supplementary material.

Model fitting and Model prediction

The models were fitted using the r package “hsdm” which uses a hierarchical Bayesian approach incorporating spatial dependency into the analysis by accounting for geographical clumping which can be explained by biological (reservoirs and hosts movement) or bioclimatic variables. In our study, we analyzed the data using hierarchical binomial and ZIB hierarchal SDM models with and without spatial autoregression. ZIB models combine a binomial process for observability and a Bernoulli process for habitat suitability (Plumptre et al., 2016; Wilson and Jetz, 2016). To model spatial autocorrelation, we used SDM with Gaussian intrinsic conditional autoregressive (iCAR) (Besag, 1974). The model is fitted
using a Bayesian framework that allows the use of pre-validated predictors and the generation of parameter uncertainties. A mixture of topographical, climatic, landscape, and human dependent predictors was used. The effect of a predictor was considered significant if it fell within a 95% confidence interval of the posterior distribution parameter. We used non-informative priors with a large variance of $10^6$ (mean = 0), except for spatial random effects, for which a weak informative prior: Uniform (min = 0, max = 10) was used. Two parallel MCMCs were run for each parameter and the convergence of the chains was checked visually using traceplots (Supplementary information) and the Gelman and Ruben’s convergence diagnostic. High-risk areas or hotspots for each viral EID group are predicted using a maximum sensitivity + specificity threshold selection and the accuracy of the model was determined by True Skill Statistic (TSS) (Liu, White and Newell, 2011).

Results

Model comparison

We used hierarchical SDM binomial, ZIB, binomial iCAR and ZIB iCAR models to map the predictive risk of viral EIDs. To compare the models with respect to deviance, we constructed a null model in each case to be compared and calculated the percentage of deviance explained by the null model. The spatial autoregression models performed better than their counterparts in the three groups (Table 1). In the Filoviridae prediction model, we find that 69% of the null deviance could be explained by the bioclimatic and population predictors using the ZIB model, which does not allow for accurate identification of hotspots. In contrast, the inclusion of the random effects through iCAR allowed us to explain 100% of the null deviance resulting in a perfect or saturated model. Similarly, ZIB models with imperfect detection performed slightly better with Coronaviridae, with 74% and 100% of the null deviance explained without and with iCAR respectively. However, with the Henipavirus EID models, the binomial iCAR model was slightly better, but we chose to summarize the ZIB with iCAR as the possibility of imperfect detection of outbreak events remains a concern. In addition, the ZIB model was able to explain 71% of the null deviance using predictors for Henipavirus events, which is superior to the binomial model and facilitates model standardization and comparison. Prediction maps, mcmc traceplots and TSS evolution for each model are available in the Supplementary information.

Detection of EID hotspots

The hotspots for filovirus diseases, EVD and MVD, were found in the forest regions of Uganda, southern Sudan and eastern parts of Democratic Republic of Congo, with smaller areas in West and Central Africa, as far as Angola (Figure 1). The ZIB iCAR model had a high TSS of 0.99 due to the addition of spatial autoregulation. High-risk regions for EIDs caused by Coronaviridae predominate across the Indian subcontinent, with some areas in China and Southeast Asia (Figure 2), detected by a ZIB iCAR model with a high TTS of 1. Henipavirus diseases hotspots are scattered along the west coast of India, in
Bangladesh, along the coast in Malaysia and in smaller areas of the Indonesian archipelago (Figure 3). The model had a high TSS of 0.99 and a probability threshold of 0.5.

**Significant environmental predictors**

We observed that minimum temperature [2.49; 95%CI 1.41 to 3.63] and rainfall [1.49; 95%CI 1.01 to 2.05] were the significant predictors of Filoviridae outbreaks in the ZIB iCAR model. As there were sufficient observations for MVD and EVD outbreak events, we ran separate ZIB models with iCAR for each group of diseases to identify the significant predictors. The MVD model was found to have positive correlations with minimum temperature [3.11; 2.03 to 4.40] and negative maximum temperature [-2.21; -3.28 to -1.34]. While the significant predictors of EVD were an increase in minimum temperature [2.66; 1.14 to 3.91] and human-induced land cover change. [1.50; 1.02 to 1.99]. We observed the significant influence of population density [3.75; 2.15 to 5.43] and land cover changes [2.55; 1.48 to 3.69] on the distribution of hotspots for Coronaviridae events. Nipah and Hendra virus outbreaks were negatively influenced by elevation [-8.84; -16.27 to -4.95] and positively influences by increases in human-induced land change [2.13; 1.41 to 2.88] and mean rainfall [2.43; 1.62 to 3.12]. The estimated means and quantiles of the MCMC chains for each variable are detailed in Table 2.

**Discussion**

Here we show that ecological, climatic and landscape factors could predict future hotspots of human viral diseases emergence on a global scale and could thus serve as a basis for surveillances and early warning systems. For the three groups of viral diseases studied, we were able to map areas at high risk of disease emergence based on the spatial distribution of disease reservoirs and hosts as well as WHO data on the distribution of each disease. We found that human-related factors, and in particular the impact of population growth on human-modified landscape, were a common predictor of disease emergence. Filoviridae and henipavirus outbreaks were also linked to rainfall, while Filoviridae and Coronaviridae emergences were favoured by increased in minimum night-time temperature. In addition, for Filoviridae, we noted the potential involvement of “unknown” variables (variables not used in this study). In Africa, these variables could relate to human behaviour, such as bushmeat consumption, biodiversity loss or even other bioclimatic covariates. Interestingly, coronavirus diseases are the only ones to be positively impacted by human population density. Similarly, hotspots of henipaviruses depended on areas of low elevation and low rainfall.

Recent research has shown that increased surface temperature and unpredictable seasonal rainfall due to climate change have an indirect effect on disease emergence through sudden ecological changes of their reservoir, loss of biodiversity and migration of small mammal hosts (García et al., 2018; El-Sayed and Kamel, 2020). For example, minimum temperature is the limiting factor for parasite development and vector distribution in malaria transmission (Patz and Olson, 2006) and other vector-borne disease epidemics such as Crimean Congo Hemorrhagic Fever and Zika (Myers et al., 2013; Broxton et al., 2014).
Unfortunately, research outside of vector-borne diseases is limited. However, this direct spatial dependence of disease emergence on minimum temperatures is worrying. Indeed, with climate change, increasing night-time minimum temperatures lengthening the frost-free season in most mid- and high latitude regions (Folland, Karl and Salinger, 2002) could potentially increase the latitudinal extent of infectious disease emergence.

We also found that low elevation and high rainfall have a significant influence on the distribution of Henipavirus outbreaks. Consistent with our results, studies have hypothesized that the emergence of Nipah in the lower Gangetic plains and low-lying marshes could be attributed to flooding, which leads to the destruction of mammalian habitats (Ambat et al., 2019). Rapid changes in ecological habitats due to human land-use change lead to starvation and migration of fruit bat species (Family Pteropodidae), reservoirs of Nipah virus, with contamination of fruit trees near human habitations and increased exposure to the pathogen (JH et al., 2006; Ambat et al., 2019; Royce and Fu, 2020). Our results support this hypothesis of the increased risk of Nipah outbreaks associated with lowland plains, flooding, and rapid human-induced habitat changes.

EVD and coronaviral diseases have also been found to be associated with human-modified landscapes. EVD has long been linked to landscape alterations such as deforestation, mining, population growth and land fragmentation (Castillo-Chavez et al., 2015; Olivero et al., 2017; Redding et al., 2019). Our results show that EVD outbreaks are not directly related to population density, contrary to a recent study (Redding et al., 2019), but rather to the effects of population increase on the human-modified landscape, such as urbanization, deforestation, mining and hunting. In contrast, population density was significantly related to coronavirus hotspots. Whether high population density leads to observer bias and thus to increased reporting of outbreaks needs to be examined in detail. The report of a SARS-like pneumonia in 2012 in miners in the Tongguan, Mojiang (Rahalkar and Bahulikar, 2020) raises the issue of potentially unreported sporadic outbreaks in regions with limited populations. Studies show that the emergence of coronaviral disease such as SARS (Shi and Hu, 2008) and MERS (Dromedaris en ‘Middle East respiratory syndrome’ | Nederlands Tijdschrift voor Geneeskunde, no date) is directly related to exposure to body fluids from mammal raised in confined spaces for bushmeat and recreation activities, respectively. “Wild flavour” bushmeat restaurants and markets are often located in densely populated cities where the demand for exotic proteins is high (Wolfe et al., 2005; Lee et al., 2020) and cases are therefore more likely to be reported in densely populated areas. In the case of MERS, there is an increase in reporting in large cities as camel owners seek treatment for respiratory distress in tertiary hospitals located in large cities and are therefore likely to report cases. The effect of population density is, however, crucial in the spread of the epidemic and therefore remains an important factor in the detection of hotspot and active surveillance.

We suggest here the urgent need of alternatives to rapid land-use changes such as deforestation, land fragmentation for agriculture and livestock, and changes in cultural practices for bushmeat consumption. More importantly, the results highlight the major impact of increasing population and human activities on land alteration and ecological changes, as well as the dependence of viral disease emergences on
bioclimatic changes (minimum temperature, rainfall, low elevation and flooded areas) at the global scale. We show the potential of using climatic, topographic and population data to identify and predict areas at high risk of disease emergence. Although our study focused on three viral diseases of concern, we suggest that such biogeographic approach to predicting disease emergence should be considered and tested for other diseases under surveillance in a global active surveillance context.

Conclusions

Landscape (deforestation, urbanization, elevation, flooded-areas, etc.), climatic (rainfall, temperature, etc.) and epidemic data were used in combination to estimate the potential role of human-induced landscape and climate change on the emergence of Filoviridae, Coronaviridae and henipaviruses on a global scale. Such predictive approach identifying regions at high risk of disease emergence should be urgently considered for monitoring local and global pathogen emergence and for the identifying potential future epidemics. By recognizing the influence of predictive environmental factors on EIDs and adopting a predictive approach to disease emergence, unprecedented EID outbreaks can be made predictable.

Declarations

Ethics approval and consent to participate.

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All data and R code for the models used in this manuscript can be accessed at https://github.com/soushie13/ecohealth

Competing interests

The authors declare that they have no competing interests.

Funding

The manuscript was a part of the doctoral dissertation of SJ during her PhD funded by the University of Guyane.

Authors' contributions

SJ extracted, analyzed, and modeled the data. SJ, MC and REG interpreted the results. MC, MN and REG were major contributors in writing the manuscript. All authors read and approved the final manuscript.
Acknowledgements

Not applicable

References


**Tables**

Due to technical limitations, table 1,2 is only available as a download in the Supplemental Files section.
Figure 1

Predictive risk of Filoviridae disease emergence modeled from the ZIB iCAR model with the spatial distribution of mammalian reservoirs in grey.
Figure 2

Predictive risk Coronaviridae disease emergence modeled from the ZIB iCAR model with the spatial distribution of mammalian reservoirs in grey.

Figure 3
Predictive risk of Henipavirus disease emergence modeled by the ZIB iCAR model with spatial distribution of mammalian reservoirs in grey.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TablesLandscape.pdf
- supplementaryinformation.pdf