Spotted Fever Rickettsiosis Dynamics in the United States

Murphy John, Department of Mathematics and Statistics
Yan Lin, Ph.D., Department of Geography and Environmental Studies.

Abstract

Spotted Fever Rickettsiosis (SFR) refers to a group of tick-borne diseases caused by Rickettsia bacteria. Over the past two decades, there has been a significant surge in SFR cases, rising from 495 reported cases in 2000 to a peak of 6,248 in 2017. A significant concentration of SFR is observed in the southeastern United States, with five states contributing to more than 50% of reported instances. The high prevalence of SFR in these states has led to a scarcity of research on cases occurring in other regions of the United States as well as the spatio-temporal pattern of the burden. To bridge this gap, this study leveraged comprehensive social vulnerability index, healthcare shortage, and environmental data from all U.S. counties. To investigate the impact of canine hosts in the prevalence of SFR, this study incorporates veterinary accessibility and dog population data. Utilizing advanced spatial analysis techniques, this project characterized SFR geographical patterns and identified contributing factors. This research also identified and analyzed high-incidence SFR areas, offering crucial insights for public health officials. These findings have potential to inform targeted interventions in high-risk areas, aiding in the effective mitigation of SFR transmission.

Background

Spotted Fever Rickettsiosis (SFR) encompasses a group of tick-borne diseases attributed to Rickettsia bacteria, posing a significant public health concern globally. The incidence of SFR has exhibited a notable increase in the United States over the past two decades, as seen by a surge from 495 documented cases in 2000 to a peak of 6,248 cases reported in 2017 (CDC, 2020). Predominantly observed in the southeastern United States, this rise in cases has prompted substantial attention due to its implications for public health. Five states within this region, namely Alabama, Arkansas, North Carolina, Tennessee, and Texas, collectively account for over 50% of the reported instances, emphasizing the concentrated burden of SFR in this geographical area (CDC, 2020).

Rocky Mountain Spotted Fever (RMSF) is a part of the SFR group caused by the bacterium Rickettsia rickettsii and spread by multiple species of ticks. This intracellular, coccobacillus bacterium harvests nutrients from its host cell to carry out respiration. If left untreated, it can cause nerve damage, hearing loss, incontinence, partial paralysis, gangrene of toes or fingers, and potentially death (Johns Hopkins Medicine, 2024).

A study of Rocky Mountain Spotted Fever (RMSF) in Sonora, Mexico found that better socioeconomic conditions and improved access to healthcare typically result in fewer cases and lower case fatality rates (Álvarez-López et al, 2021). Furthermore, a 2017 study in the state of Illinois found associations between temperature and precipitation variables and increased incidence of SFR (Kerins et al, 2017). The investigation of an RMSF outbreak in Mexicali, Mexico, revealed that in areas with high rates of tick

infestations, dogs serve as hosts and have a key role in the transmission of RMSF (López-Pérez et al, 2020).

Despite the pronounced prevalence of SFR in the southeastern United States, limited research has focused on cases occurring in other regions of the country, thereby limiting understanding of the spatio-temporal distribution of the disease. There exists a critical gap in knowledge regarding the broader epidemiology of SFR and the factors contributing to its occurrence and spread across diverse geographic areas.

To address this gap, this study draws upon various data sources, including comprehensive social vulnerability indices, healthcare accessibility metrics, and environmental parameters, spanning counties in the contiguous United States. Additionally, the study integrates data pertaining to veterinary accessibility and dog population to explore the potential role of canine hosts in influencing SFR prevalence. Leveraging advanced spatial analysis techniques, this research aims to explain the geographical patterns of SFR incidence and the underlying factors contributing to its spatial variability.

Materials

The study area consists of all counties in the contiguous United States, excluding Alaska and Hawaii. After tidying the data and removing all null observations, the resulting dataset includes complete information for 2,743 counties in the United States.

The data utilized in this research encompassed a comprehensive list of social and environmental factors potentially contributing to SFR. Spotted Fever Rickettsiosis Rate from 2016 to 2019 was sourced from the Centers for Disease Control's (CDC) National Center for Emerging and Zoonotic Infectious Diseases Division of Vector-Borne Diseases. Additionally, the Social Vulnerability Index (SVI) data, acquired from the CDC's Agency for Toxic Substances and Disease Registry in 2020, provided essential insights across multiple themes including socioeconomic status (Theme 1), household characteristics (Theme 2), racial and ethnic minority status (Theme 3), housing type/transportation (Theme 4), and an overall percentile ranking (Overall Theme). Temperature parameters, comprising average and maximum temperatures from 2016 to 2019, were extracted from the National Centers for Environmental Information. The study also integrated average annual precipitation data from the same source. To evaluate healthcare infrastructure, the Healthcare Shortage Index from the Rural Health Information Hub in 2023 was incorporated alongside Rural Continuum Codes obtained from the U.S. Department of Agriculture in 2013. Furthermore, veterinary accessibility metrics such as the Veterinary Accessibility Score and Dog to Human Population Ratio were sourced from The Veterinary Care Accessibility Project in 2022.

Methods

Data Cleaning

Variance Inflation Factor and Correlation Matrix reveal strong multicollinearity between SVI Themes 1 through 4 and SVI Overall Theme. There was also strong multicollinearity between Average Temperature and Maximum Temperature. Thus, Overall SVI Theme and Maximum Temperature were dropped.

Regression Analysis

A linear regression model is fit to the resulting variables with SFR Rate as the response. After removing those that are not significant at the 0.05 level, we produce a final regression model.

Table 1: Regression Analysis Results

	Estimate	SE	P value
Intercept	-0.6436	0.115	0.000
Minority Status SVI	-0.3349	0.054	0.000
Average Temperature	0.0084	0.002	0.000
Average Precipitation	0.0048	0.001	0.000
Rural-Urban Continuum Codes	0.0184	0.006	0.001
Veterinary Care Accessibility Score	-0.0015	0.000	0.002
Dog to Human Population Ratio	1.0960	0.168	0.000
R-squared	0.104		
AIC	5002		

Hot Spot Analysis

A Hot Spot Analysis is conducted with SFR Rate as the response. Descriptive statistics for the hot and cold spots are calculated.

Figure 1: Hot Spot Analysis

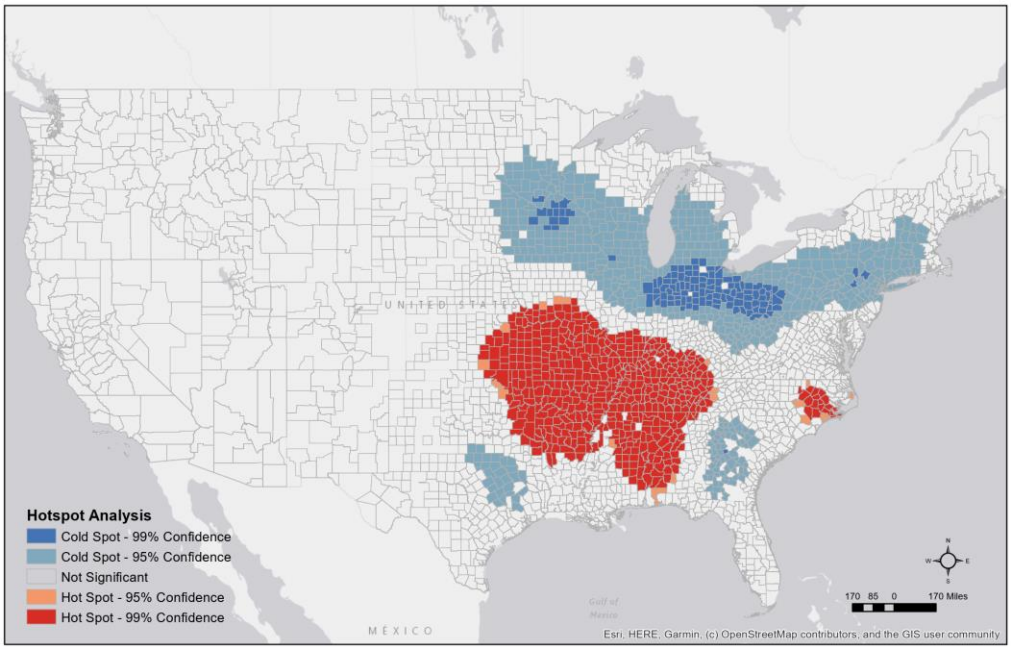

Table 2: Hot Spot Descriptive Statistics

	Mean	Minimum	Median	Maximum	Standard Deviation
SFR Rate (x1000)	0.817	0.0	0.393	13.188	1.261
Minority Status SVI	0.413	0.014	0.399	0.966	0.252
Average Temperature	60.008	52.7	59.6	67.3	3.155
Average Precipitation	53.686	36.15	55.14	67.52	7.276
Health Shortage Value	1.823	0	2	2	0.435
Rural-Urban Continuum Codes	5.181	1	6	9	2.555
Veterinary Care Accessibility Score	38.426	0.798	33.490	99.809	25.511
Dog to Human Population Ratio	0.348	0.170	0.355	0.475	0.0625

Table 3: Cold Spot Descriptive Statistics

	Mean	Minimum	Median	Maximum	Standard Deviation
SFR Rate (x1000)	0.017	0.0	0.0	0.496	0.043
Minority Status SVI	0.442	0.0003	0.408	0.997	0.277
Average Temperature	52.247	41.4	51.2	69.9	6.715
Average Precipitation	44.289	27.19	44.41	65.6	5.685
Health Shortage Value	1.342	0	1	2	0.6243
Rural-Urban Continuum Codes	4.158	1	4	9	2.4056
Veterinary Care Accessibility Score	53.488	0.665	55.512	100	28.191
Dog to Human Population Ratio	0.252	0.126	0.254	0.411	0.060

Spatial Lag Regression

To assess the spatial clustering component of the data, a Spatial Lag Regression is fit.

Table 4: Spatial Lag Regression Results

	Estimate	SE	P value
Weight	0.631	0.018	0.000
Intercept	-0.225	0.094	0.016
Minority Status SVI	-0.189	0.044	0.000
Average Temperature	0.004	0.002	0.012
Average Precipitation	0.002	0.001	0.045
Rural-Urban Continuum Codes	0.010	0.005	0.026
Veterinary Care Accessibility Score	-0.001	0.000	0.016
Dog to Human Population Ratio	0.350	0.137	0.012
R-squared	0.407		
AIC	4086.79		

Results

On the nation-wide scale, Temperature, Precipitation, Rural-Urban Continuum Codes, and Dog to Human Population Ratio have significant positive correlations with SFR Rate. That is, as these variables increase, the SFR Rate increases. Note that an increase in Rural-Urban Continuum Code is interpreted as a more rural area. Conversely, the variables Minority Status SVI and Veterinary Care Accessibility Score

have significant negative correlations with SFR Rate on the nation-wide scale. This means as these variables decrease, the SFR Rate increases. Notably, a decrease in Minority Status SVI implies fewer individuals identifying with minority groups. A decrease in Veterinary Accessibility Score implies less access to veterinary care.

Hot Spots are where features with high values of SFR Rates cluster spatially. The Hot Spots in this analysis are located in the southeastern United States. These areas have increased Temperature, Precipitation, Health Shortage Value, Rural-Urban Continuum Codes, and Dog to Human Population Ratios. Note that increased Health Shortage Value is interpreted as lack of access to healthcare. The Cold Spot areas are where low values of SFR Rate cluster spatially. These areas are identified predominately in the northeastern United States but small Cold Spots are found in the southeast. They are characterized by increased Minority Status SVI and Veterinary Accessibility Score.

Decreased Minority Status (smaller minority population) is consistently associated with increased SFR Rates. This is hypothesized to be a result of the location of ticks being in more rural areas, which have decreased minority populations. This hypothesis is further supported by the inverse association between Rural-Urban Codes and Minority Status. That is, there are greater minority populations in urban settings. We also observe a positive association between Rural-Urban Codes (more rural areas) and SFR Rate. The role of canine hosts is apparent by the association of Veterinary Accessibility Score and Dog Population Ratio variables. As Veterinary Accessibility decreased, SFR rate increases. As Dog population Ratio increased (more dog population) SFR Rate increases. The Spatial Lag Model produces similar associations as the Regression Model but the inclusion of a spatial lag term improves the R-squared and AIC values. This emphasizes the importance of a spatial indicator in a model of SFR Rates.

Future Work

Further investigation is warranted to understand the nationwide correlations between environmental factors, socio-economic indicators, and SFR Rates. Additionally, expanding the hot spot analysis could provide insights into localized patterns and aid in the development of targeted interventions to mitigate SFR Rates in high-risk areas. A closer look at the role of canine hosts in the spatial distribution of SFR Rates is also suggested. Understanding how these factors influence SFR Rates can inform targeted interventions aimed at controlling the spread of tick-borne diseases.

Acknowledgements

Funding for this project is provided by the National Science Foundation, Award Number 2019609, the University of New Mexico Arts & Sciences Support for Undergraduate Research (ASSURE), and the Frank O. and Sadie M. Lane Scholarship.

References

Álvarez-López D.I, Ochoa-Mora E, Nichols Heitman K, Binder A.M, Álvarez-Hernández G, Armstrong P.A. Epidemiology and Clinical Features of Rocky Mountain Spotted Fever from Enhanced Surveillance,

Sonora, Mexico: 2015-2018. Am J Trop Med Hyg. 2021 Jan;104(1):190-197. doi: 10.4269/ajtmh.20-0854. PMID: 33146112; PMCID: PMC7790062.

Centers for Disease Control and Prevention Agency for Toxic Substances and Disease Registry. (2023). CDC/ATSDR Social Vulnerability Index 2020. https://www.atsdr.cdc.gov/placeandhealth/svi/index.html

Centers for Disease Control and Prevention National Center for Emerging and Zoonotic Infectious Diseases Division of Vector-Borne Diseases. (2023). Reported Cases of Selected Tickborne Diseases 2016-2019. https://www.cdc.gov/ticks/data-summary/geographic-distribution.html

Johns Hopkins Medicine. (2024). Rocky Mountain Spotted Fever. https://www.hopkinsmedicine.org/health/conditions-and-diseases/rocky-mountain-spotted-fever

Kerins J.L, Dorevitch S, Dworkin M.S. Spotted Fever Group Rickettsioses (SFGR): weather and incidence in Illinois. Epidemiology and Infection. 2017;145(12):2466-2472. doi:10.1017/S0950268817001492

López-Pérez A.M, Orozco L, Zazueta O.E, Fierro M, Gomez P, Foley J. An exploratory analysis of demography and movement patterns of dogs: New insights in the ecology of endemic Rocky Mountain-Spotted Fever in Mexicali, Mexico. PLoS One. 2020 May 21;15(5):e0233567. doi: 10.1371/journal.pone.0233567. PMID: 32437470; PMCID: PMC7241830.

National Centers for Environmental Information. (2023). Climate at a Glance County Mapping 2016-2019. https://www.ncei.noaa.gov/

Rural Health Information Hub. (2023). Health Professional Shortage Areas: Primary Care, by County, 2023. https://www.ruralhealthinfo.org/charts/5

United States Department of Agriculture. (2013). Rural-Urban Continuum Codes. https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/

Veterinary Care Accessibility Project. (2024). Veterinary Care Accessibility Score. https://www.accesstovetcare.org/

Virginia Department of Health. (2018). Spotted Fever Rickettsiosis (including Rocky Mountain Spotted Fever). https://www.vdh.virginia.gov/epidemiology/epidemiology-fact-sheets/spotted-fever-rickettsiosis-including-rocky-mountain-spotted-fever/