A WEB-BASED GEOGRAPHIC FRAMEWORK TO DETECT AND VISUALIZE SPACE-TIME CLUSTERS OF INFECTIOUS DISEASES

by

Yu Lan

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Geography

Charlotte

2022

Approved by:

Dr. Eric Delmelle

Dr. Deborah Thomas

Dr. Jing Yang

Dr. Rajib Paul

ABSTRACT

YU LAN. A Web-based Geographic Framework to Detect and Visualize Space-time Clusters of Infectious Diseases
(Under the direction of Dr. ERIC DELMELLE)

Infectious diseases pose a significant threat to public health worldwide as evidenced by the recent coronavirus 2019 (COVID-19) pandemic. Despite significant human losses, the advent of web-accessed, map-based “data dashboards” that can monitor disease outbreaks, proved essential in managing public health responses. In many cases, the backend of these dashboards employs basic mapping functionality, displaying counts or rates. As the pandemic advanced, the identification of elevated rates was increasingly important in the geographical allocation of public health resources. However, such maps miss the opportunity to provide accurate information to policy decision makers such as the rate of disease spread, cyclicity, direction, intensity, and the risk of diffusion to new regions. Space-time geoanalytics, when coupled with rich visualizations, can address these shortcomings. Moreover, when implemented over the web, such functionality can be accessed from virtually anywhere.

This dissertation presents a web-based geographic framework for detecting and visualizing explicit space-time clusters of infectious diseases. First, I conduct a systematic review of the literature around the theme of space-time cluster detection for infectious diseases to identify state-of-the-art techniques that should be included in the proposed web-based framework. Second, I develop a tightly coupled, web-based analytical framework for the detection of clusters of infectious diseases using interactive and animated 3D visualizations to aid epidemiologists in readily and adequately uncovering the characteristics of space-time clusters. As a proof of concept, I populate the framework with COVID-19 county-level data for the 48 contiguous states in the US, and demonstrate data retrieval and storage, space-time cluster detection analysis, and 3D visualization within an open source WebGIS environment. Third, I evaluate the prototype in two steps: 1) present this and two existed COVID-19 systems to a group of infectious diseases experts and solicit feedback, 2) and evaluate functionalities on the prototype by conducting a user study with graduate students in a setting of online surveys.

This tightly coupled approach facilitates the detection of space-time clusters of diseases in a computationally acceptable timeframe. The characteristics of this framework (generic, open source, highly accurate, modifiable) will enable low-cost monitoring of the spatial and temporal trends of diseases causing high risks of infection.

ACKNOWLEDGMENTS

Pursuing this Ph.D. started with unhappy work experiences seven years ago when I was in Nanchang. Before that time, I had no idea what my future would be. But I knew I was unhappy and did not want to retain this work as my future. However, I found myself passionate about learning new techniques and knowledge during the long work experience, although studying could have been more interesting to me as an undergraduate. Then, I decided to quit that good-paid job in my home city and prepare to study overseas in the US. I was lucky that the master's coordinator (Dr. Eric Delmelle) recruited me, and then I said goodbye to my family and friends in China and flew to the US alone.

After graduating from the master's program, I continued as a PhD student in the same department, and Eric became my advisor. In the first two years, I enjoyed learning knowledge and technology from class and conducting research with my advisor and colleagues. Then, the pandemic started. At the beginning of 2020, I checked the news of COVID-19 in China and cried about terrible things there. Later in that spring semester, UNCC also changed all classes to virtual. I was isolated in my bedroom, along with the fear of this pandemic. However, the long isolation pushed me to think about what life I wanted to have in the future, and I realized that unexpected things always happen. My future path in the academy became clearer to me.

I would like to first thank my advisor for his mentor and patience in this long journey. My English was not well at the beginning and he reviewed my writing sentence by sentence and word by word. He taught me the way how to do research and generate critical thinking, and also showed me how to deal with issues from students as an instructor. His passion for research also inspires me to devote my future into academy.

I am very grateful to my committee (Drs. Eric Delmelle, Deborah Thomas, Jing Yang, and Rajib Paul) for their support and comments on my dissertation, especially on the evaluation section. Without their timely responses and regular meetings with me in a short time, I would not be able to finish it on time. I also thank Drs. Wenwu Tang and Doug Shoemakers for their support and help as Center for Applied GIScience (CAGIS) members. I also would like to acknowledge other faculty during my studies in Charlotte: Drs. Elizabeth Delmelle (my master advisor), Heather Smith, Wei-Ning Xiang, Harrison Campbell.

I have also made many friends from the department: Dr. Michael Desjardin, Dr. Alexandar Hohl, Dr. Minrui Zheng, Tianyang Chen, Tarini Shukla, Zachery Slocum, Dr. Claudio Owusu, Dr. Wenpeng Feng, Jiyang Shi, Dr. Jing Deng, Yanfang Su, Fazizeh Hatami, Providence Adu, Connor Moore, Jiaxin Liu, Shuli Hao. Many of them are/were CAGIS students, and I appreciate the friendship with all of them.

I am grateful that my parents have supported my studying abroad. My mother always trusts and supports me with no condition, and my father provides all the financial support. Besides my parents, everyone from my family is also very supportive during PhD study, and I want to thank them as well.

I want to thank my favorite person and best friend, Yin Leung, for all the support and love in this journey. Although our physical distance is so long, our heart is so close. Without your support, I have no confidence to complete it alone. I look forward to our future together.

November 2022

Charlotte, NC

DEDICATION

To my family.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Background

On the last day of 2019, the World Health Organization (WHO) made the first response to an infectious disease caused by a newly discovered coronavirus (named COVID-19), which became a pandemic on March 11, 2020 (WHO, 2020a). As a novel and highly contagious infectious disease among humans, COVID-19 spreads from communities to cities, states, countries, and globally at a rapid transmission rate. Two years later, despite the development and rollout of effective vaccines, the number of daily new cases of COVID-19 is still substantial, and the COVID-19 pandemic has had enormous health, economic and societal impacts in every country. Unfortunately, threats of infectious diseases to humans have existed and increased since 1980 and will likely continue (Smith et al., 2014). Thus, the question is not "... if another outbreak will occur, but when and where" (Boulos & Geraghty, 2020).

Infectious diseases, also known as communicable diseases, are defined as "diseases caused by pathogenic microorganisms, such as bacteria, viruses, parasites or fungi; the diseases can be spread, directly or indirectly, from one person to another" (WHO, n.d.). The spread or transmission route could be through direct and/or indirect contact, airborne, waterborne or foodborne, vector-borne, and the environment. For instance, the transmission of COVID-19 among humans includes both direct and indirect contacts (such as contaminated surfaces) and airborne aerosol/droplet routes; the latter is considered the dominant transmission mechanism for this disease (Zhang et al., 2020).

According to the WHO, the number of deaths caused by communicable diseases has dropped since 2000. However, several communicable diseases, including lower respiratory infections (ranked 4th), neonatal conditions (ranked 5th), and diarrheal diseases (ranked 8th), claimed 6.1 million deaths in 2019[1], are still among the top 10 leading causes of death globally (see Figure 1). In a comparison, deaths caused by the COVID-19 pandemic already surpassed 5 million by the first day of November 2021 worldwide (World Health Organization, 2020), and this number is expected to increase given emerging variants of SARS-CoV-2, the virus that causes COVID-19. Therefore, stopping or slowing the transmission of COVID-19 and other infectious diseases is critical.

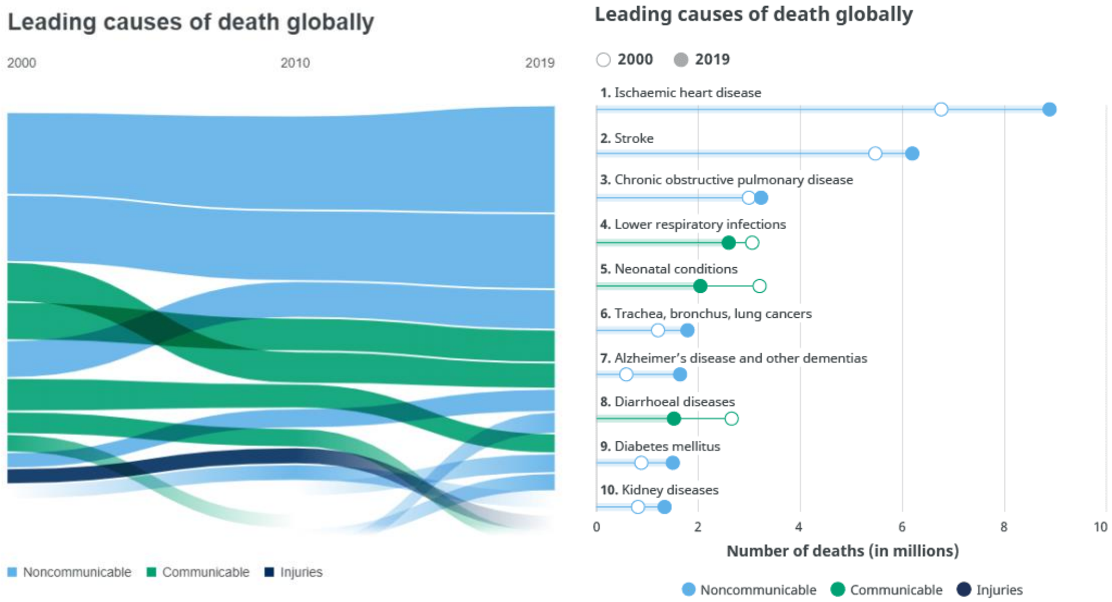


Figure 1. Leading causes of death globally. The left graph shows the trend of leading causes of death globally from 2010 to 2019. Each belt represents one disease in three types of causes, and its height is the death at that time (WHO, 2020b). The right graph shows the numbers of death of the top 10 leading causes in 2000 and 2019 (WHO, 2020c).

In public health, epidemic models are used to measure the dynamics of an infectious disease through different groups of the population. Those groups contain individuals who are *susceptible* to being infected (S), people who are *infected* or *exposed* (I), and individuals who are *recovered* (R). When a new infectious disease is introduced in a community, the number of

---

[1] WHO only updated data on leading causes of death globally until 2019.

susceptible (S) is extremely elevated since no immunity has yet been built. As the number of infected individuals (I) increases, the number of susceptible decreases, and so does the number of recovered people (R). A common approach to model infectious diseases is the Susceptible-Infected-Recovered (SIR) compartmental model. However, recovered individuals could become infected again, especially when new variants are emerging and introduced in a community, leading to an SIR-S model or when immunity vanishes. Another critical measure of the potential spread of infectious diseases in the population is the basic reproduction number R_0, the expected number of cases infected from one case in a population with the assumption that other individuals are not infected or immunized, calculated using different compartmental models (van den Driessche & Watmough, 2008). An R_0 higher than 1 reflects a spreading disease and indicates the contagiousness or transmissibility of infectious agents. In contrast, a value of less than 1 marks the end of an outbreak (Delamater et al., 2019). It is worth noting that R_0 is not a constant value and changes over time (R_t is used in this case). It is dynamic throughout space and time because of different human-environmental interactions (e.g., various levels of lockdowns) and health policies (e.g., vaccine adoption). For example, R_0 would be smaller if an area adopted a lockdown policy than areas with no such policy. As to COVID-19, the range of estimated R_0 by March 2020 was 1.90 to 6.49 (Alimohamadi et al., 2020).

Geography plays a pivotal role in disease spread. Barrows (1923) defined geography as the science of human ecology that focuses on the relationships between the environment and the distribution of human activities. Changes in human ecology, such as long-distance mobility, raise unprecedented challenges to limiting the spread of infectious diseases (Weiss & McMichael, 2004).

As noted by Kirby et al. (2017) and Kwan (2016), the health outcomes of a population are the results of "an interplay of different factors, such as individual characteristics, the physical and social environments an individual interacts with, cultural norms, and both the provision and utilization of health services." In other words, the "activity spaces" are connected with the natural, social, and economic surrounding environments. This 'activity space' can manifest itself at different scales (local, regional, national to global) and should be considered a potential disease determinant (Cromley & McLafferty, 2011; Diez Roux, 2001), because it will improve our understanding of health outcomes. Rogers & Randolph (2003) point out that the key to understanding the diffusion of infectious diseases is to uncover the dynamic process of spreading patterns with the investigation of time, place, and person at the same time.

1.2 GIS and infectious diseases surveillance

When monitoring the burden of infectious diseases, timely surveillance is an essential epidemiological component to describe the ongoing dynamics of a disease, identify trends, and detect outbreaks and new pathogens (Murray & Cohen, 2017). Three key elements are critical when conducting space-time surveillance of infectious diseases; (1) the time when the disease occurs, (2) the geographic location where cases are reported, and (3) the segments of the population who get infected. Documenting each of these factors is complicated and made even more challenging when data is uncertain or incomplete. For instance, individuals may be infected days before a test reveals positivity due to the incubation period. Or individuals may be asymptomatic, increasing uncertainty in the number of daily cases. Reported cases may originate from another region than where the test for the disease was conducted at. In fact, too often in geography, the residential location of an individual is assumed to be the place where the infection occurred, but as mentioned earlier, humans have complex space-time paths making up the so-called ‘activity space.’ Despite those issues and uncertainties, effective prevention methods, monitoring and early detection of outbreaks, and controls and treatments aimed to stop or slow down the spread of infection have improved (Smith et al., 2014). Among the technologies that contribute to infectious diseases surveillance are geospatial technologies, which are particularly suited to capture the geographic complexity of infectious diseases (Kirby et al., 2017).

A geographic information system (GIS) is a computer system for creating, managing, analyzing, and displaying geographic data, a valuable and practical tool for monitoring the infectious disease research (Cromley, 2003; Eisen & Eisen, 2014; Kirby et al., 2017). As early as 1854, John Snow, the father of modern epidemiology, identified the cluster around a public water pump, which became the source of the cholera outbreak in London due to contaminated water found in that pump (Newsom, 2006). Back at that time, cholera was a dangerous infectious disease which would mainly spread through contaminated water. This example has been used in many textbooks to showcase how a simple map of deaths could uncover the distribution of an infectious disease. Obviously, the display of data on the map is not the most important tool in the epidemiology (Kistemann et al., 2002); other geospatial functionality can be used for geographic data collection, management, and spatial analysis, such as geocoding, geographic management, and proximity analysis to name a few. GIS can be used in concert with detection tools to monitor and respond to health issues, ultimately assisting health professionals in identifying cases, spatial trends, disease clusters, and correlation with other spatial data (Carroll et al., 2014; Delmelle et al., 2011; Delmelle et al., 2015). As a vital component of disease surveillance, cluster detection can identify high-risk areas, and it can facilitate the investigation of the spread of infectious diseases (Aamodt et al., 2006).

Space-time cluster detection methods play a pivotal role in monitoring the spread of infectious disease. The key to understanding the diffusion of infectious diseases is to uncover the dynamic process of spreading patterns with the investigation of time, place, and person at the same time (Rogers & Randolph, 2003). The combination of these three components makes the major difference of infectious disease from other diseases. Epidemic models are used to measure the dynamics of an infectious disease through different groups of the population during its spread. However, infectious diseases are not merely related to infected populations and populations at risk, as the risks of infection and transmission are caused by a myriad of covariates, such as demographics, socioeconomic factors, and environmental characteristics. These conditions can act as confounding factors leading to variations in the risk of infection and transmission (Delmelle et al., 2016; McMichael, 2004; Taylor et al., 2001; Weiss & McMichael, 2004; Wichmann et al., 2007). In addition, as evidenced by the COVID-19 pandemic, interactions among individuals can accelerate the spread of disease, adding another layer of complexity to the analysis of infectious diseases.

1.3 Cluster, outbreak, or hotspot?

Before introducing cluster detection techniques, it is fundamental to highlight the conceptual differences between clusters, outbreaks, and hotspots. These terms are often used interchangeably in infectious diseases surveillance. Knox (1989) provided a non-mathematical definition of a *cluster* as “a geographically bounded group of occurrences of sufficient size and concentration to be unlikely to have occurred by chance.” As to *hotspot*, Lessler et al. (2017) summarized three distinct types in epidemiology and suggested alternative terms such as *transmission hotspot* (elevated transmission efficiency), *emergence hotspots* (a high frequency of emergence or re-emergence of diseases), and *burden hotspot* (elevated disease incidence or prevalence or a geographic cluster of cases). Among these three types, the definition of *burden hotspots* is similar to clusters, as it was defined as “a geographic cluster of cases”. Farrington and Beale (1998) defined an *outbreak* as the increment in the number of cases beyond expected levels. According to those definitions above, an outbreak refers to the status of unexpectedly elevating infected cases in areas during a particular time. During an outbreak, hotspots and clusters are areas with elevated incidence or prevalence or sufficient concentration of the disease, and these two terms are interchangeable. For example, the current pandemic is an outbreak of COVID-19, while hotspots and clusters of COVID-19 cases are dynamic in both spatial and temporal dimensions. This dissertation is primarily concerned with “clusters,” particularly space-time clusters, as they also imply statistical significance that hotspots do not provide.

1.4 Techniques for cluster detection

Although forecasting is a necessary tool in public health responses, the focus of my dissertation is to understand the 'nowcasting,' the current situation, and the presence of clusters. In most commercial GISs, methods of spatial modeling are integrated to predict the trend of infectious diseases, such as geographically weighted regression (GWR) and agent-based modeling (ABM), while many spatial statistical methods are used to evaluate and map risk areas (e.g., through kernel density estimation for instance) (Carroll et al., 2014). As the focus here is to detect clusters, only spatial statistical methods able to do so are introduced. An inherent advantage of using spatial statistics to detect clusters is their ability to reveal the current distribution patterns, which is essential to finding high-risk areas in need of a timely response during an outbreak (Aamodt et al., 2006). Several spatial statistics methods have been developed to detect clusters (see Figure 2), and these can broadly be classified into three categories based on their mechanisms: (1) distance-based methods that measure distances among cases, such as the *K* function; (2) area-based methods that analyze cases within subset regions of a study area, such as spatial autocorrelation and scan statistics; and (3) continuous methods that estimate a risk surface, such as kernel density analysis.

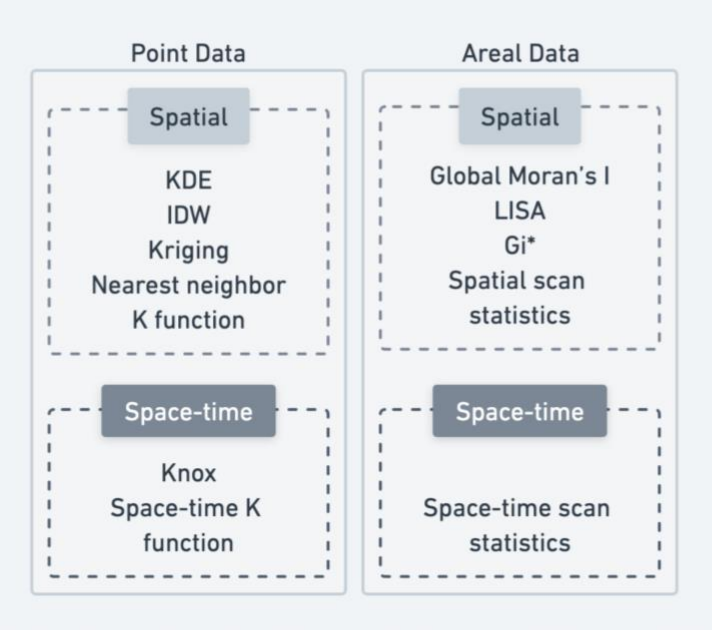


Figure 2. Popular techniques for spatial and spatiotemporal cluster detection among point and areal data.

The two most common data types used in health surveillance are disaggregated (e.g., point) data that contain spatial coordinates for each case and areal or aggregated data that aggregate cases for a specific region (e.g., postal code) during a given time. Aggregated data are typically more available and less sensitive to privacy issues (Olson et al., 2006).

Some cluster detection techniques have been extended to deal with temporal information in two different ways: (1) the spatial statistical method is repeated over different time periods; (2) a more robust approach that explicitly takes space and time into account can be developed. These cluster detection methods are typically implemented as part of specific standalone software such as CrimeStat (Levine, 2013), Geoda (Anselin et al., 2010), SaTScan (Kulldorff, 2010), SpaceStat and ClusterSeer from Biomedware, or codes in opensource programming platforms using R (Bivand et al., 2008; Gómez-Rubio et al., 2005; Moraga, 2017), Python (Rey & Anselin, 2010), etc. However, those methods suffer from poor visualization, especially in space and time.

In the following section, I introduce several methods for space-time cluster detections, categorized by the type of data they require (point or area, see Figure 2). Although the list is not exhaustive, this is representative of the most popular methods found in the literature review (see Chapter 2).

1.4.1 Techniques for point data

Techniques to evaluate clustering among point data can be categorized into methods with first-order and second-order variation or properties. First-order variation defines that the point process varies over space because of spatial variation in the mean (for example, cholera cases may vary depending on contaminated water distribution). In contrast, second-order variation describes that the variation in a spatial process is associated with spatial dependency (Gatrell et al., 1996; Rogerson & Yamada, 2008). For example, areas with high community transmission could result in high rates of COVID-19. Among first-order variation techniques, I describe the Kernel Density Estimation (KDE) and interpolation methods, including Inverse Distance Weighted and Kriging; and among second order techniques, I explain the nearest neighbor statistic and the *K*-function.

The first-order variation techniques can map the variation of certain event, such as the patterns of the disease. One such techniques is KDE, which calculates the intensity of observed points over the study region. The study region is divided into a grid of square cells, and the intensity at each cell is estimated using the Kernel function. This weighted distance function measures the intensity between the cell's centroid to all events within a predefined bandwidth.

The result of the KDE for a given cell is the sum of the intensities within the bandwidth. As the choice of bandwidth is arbitrary, KDE results are subject to a tradeoff between bias due to a considerable bandwidth and uncertainty due to a small bandwidth (Rogerson & Yamada, 2008).

Although less prevalent in infectious diseases, interpolation methods such as Inverse Distance Weighted or Kriging can monitor the spatial variation of infectious diseases, but they require an attribute (Z-value) to be interpolated. Typically, data aggregated at an areal level (e.g., count or rate of events in a zip code), however the data are generally reduced to area point, such as the centroid of that area. Interpolation methods will then use points and attribute as an input to generate a continuous surface. One crucial assumption of interpolation methods is that the spatial correlation structure is spatially constant, suitable for environmental variables such as water or air pollution, but not infectious cases or rates (Pfeiffer et al., 2008).

Unlike those first-order techniques that only map disease patterns, second-order techniques such as the nearest neighbor statistic and K-function can estimate clustering of the disease. The nearest neighbor statistic calculates the distance to the nearest neighbor, which is used to test whether closer incidents are randomly distributed or not (Delmelle, 2009). Two popular nearest neighbor techniques are the average nearest neighbor or *k*-nearest neighbors. The average nearest neighbor averages all nearest neighbor distances between each event and its nearest neighbor (also an event) and then returns the ratio of the observed mean distance to the expected mean distance for total events given in a random pattern. If this ratio is less than one, it indicates clustering. In the *k*-nearest neighbors (*k*NN) test, Cuzick and Edwards (1990) extended the nearest neighbor to *k* nearest neighbors of events, and this test can be used to detect spatial clustering of events with the consideration of the inhomogeneous populations.

The (Ripley's) *K*-function is another popular method to estimate second-order properties but assumes that no first-order effects exist in terms of the spatial pattern (Pfeiffer et al., 2008; Ripley, 1977). In essence, the *K*-function counts the number of events within a circular search window around each case, and the window moves to the next event until all the events are visited (Hohl et al., 2017). The process is repeated for increasing radius values up to a maximum distance, coinciding with the two most distant point events. For evaluation of statistical significance, (random or population-based) simulations are generated by locating the same number of events in the study area, and the *K*-function is estimated for these simulations. Significance envelopes can be generated from these *K*-function falls inside, above, or below them (inside = randomness, below = dispersion, above = clustering).

When point events have a **temporal signature** (e.g., time of occurrence), several techniques described earlier can be extended to handle the temporal dimension, including the Knox test or the space-time *K* function. The Knox test is a pairing method to detect space-time clustering present in data points, based on the assumption that events' spatial and temporal features are independent of one another (Knox & Bartlett, 1964). The Knox test requires setting both a spatial and a temporal threshold distance, and it counts the number of pairs of events separated by critical space and time thresholds. It compares the observed and expected number of pairs of points using a Chi-Square test, where the expected number of pairs of points are calculated from simulated space-time point events. It is then possible to identify the space-time distance at which the Chi-square statistic is the greatest (note that multiple maxima can occur, suggesting clustering at different scales). The *K*-function has been extended by Diggle et al. (1995) to account for its temporal counterpart. The mechanism of the space-time *K* function is, in essence, similar to the Knox test. Thus, the space-time *K* function can be considered a series of

Knox tests with different critical distances, and it further measures the spatial and temporal bandwidth for significant clusters, which would be used for further space-time analysis.

1.4.2 Techniques for areal data

One key feature inherent to spatial data is spatial autocorrelation, which refers to the correlation of a variable of interest between two locations. Based on Tobler's first law of geography, correlation decreases with increasing spatial separation (Schabenberger & Gotway, 2017). Many tests can examine spatial autocorrelation, and these statistics are categorized into global and local statistics based on whether the returned result of clusters is nonspecific (global) or specific (local).

As to global statistics, these methods measure whether the pattern of aggregated events is clustered, dispersed, or random. The global Moran's I is one of the most popular spatial autocorrelation methods for areal data, and the reported Moran's I value indicates whether events are statistically clustered (when $I > 0$). This test reflects the similarity among areas based on the assumption of the even distribution of the population at risk within the study area (Moran, 1950). The algorithm computes the mean and variance for the attribute of the variable being studied. Then, for each areal feature, it subtracts the mean, creating a deviation from the mean. Deviation values for all neighboring features (features within the specified distance band, for example) are multiplied together to create a cross-product. Thus, it provides a statistic (I) for each location with an assessment of significance (p-value). Second, it establishes a proportional relationship between the sum of the local statistics and a corresponding global statistic.

Local spatial autocorrelation methods are different from their global counterparts in that they are aimed at identifying the locations and extent of clusters (Pfeiffer et al., 2008). The local Moran's I test (LISA) is the local level version of the global Moran's I test, estimated by decomposing the Moran's *I* statistic geographically, resulting in local indicators of spatial association (Anselin, 1995). The LISA statistic measures the strength of patterns among nearby geographic units, resulting in four different types of clusters (see Figure 3). Low-Low clusters indicate low values surrounded by low values, High-High clusters indicate high values surrounded by high values, Low-High outliers indicate low values surrounded by high values, and High-Low outliers indicate high values surrounded by low values. These four categories can be used to lay out not only clusters of high values (e.g., High-High: the high infection rate area surrounded with other high infection rate areas) but also clusters with different surrounding situations (e.g., High-Low: the high infection rate area while surrounding areas have low infection rates). From Figure 3, the center of the map is classified as High-High clusters from Figure 3 (b), and the rates for the same area also have the highest number from Figure 3 (a).

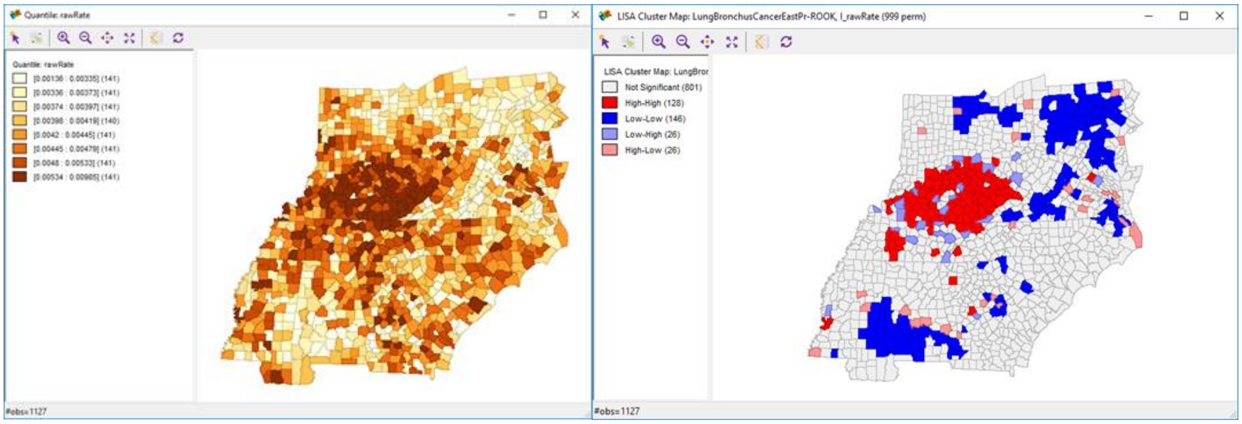


Figure 3. An example of LISA results of lung and tracheal cancer among males from 2011 to 2015 in the eastern US. (a) average annual rates for the lung and tracheal cancer among males using the 65+ male census population as the denominator; (b) LISA results of the same dataset.

Another local clustering detection method for aggregated or areal data is the Getis-Ord Gi* statistic, which compares local estimates of spatial autocorrelation with global averages to detect clusters in spatial data (Ord & Getis, 1995). It has been called the hot spot analysis since the test returns significant clusters of high (hot spot) and low (cold spot) values.

Another important clustering detection at the local level is Kulldorff's spatial scan statistic. The scan statistic detects spatial clusters by scanning the data via a circular or elliptic window with the radius ranging from zero to a maximum value specified by the user (Kulldorff, 1997). A cluster is defined as a circle with a significant maximum likelihood ratio, and only events with centroids located within this circular are affiliated to this cluster. In addition, each cluster contains a relative risk (*RR*) value, which is the ratio of the estimated risk within the cluster to the estimated risk outside the cluster.

The Kulldorff's spatial scan statistic can be extended in time to account for the temporal dimension by replacing the scanning circular or elliptic window with a cylindrical window where the height represents the period of potential clusters (Kulldorff et al., 2005). The window with the maximum likelihood is the most likely cluster, that is the cluster least likely to be due by chance. Thus, this cylindrical window not only moves in space but also in time. In public health surveillance, this statistic can be applied to both retrospective and prospective studies depending on the different focus either on past patterns or on current trends. Retrospective methods carry out analyses -such as hypothesis tests- for a fixed geographical region and a fixed study period to estimate the prevalence of diseases or compare diseases patterns in the different areas in one frame; in contrast, prospective methods use ongoing collected data and repeat analysis to detect a significant change in a timely manner (Sonesson & Bock, 2003).

1.5 Web-based GIS for infectious diseases

Although desktop GIS software generally contains several techniques to analyze spatial patterns, web-based GIS or WebGIS can support methods to any individual through the internet with no system requirements. In particular, lower-income countries can potentially access these web-based tools to monitor the spread of infectious diseases locally, without the need for extra resources. According to Luan and Law (2014) who reviewed web-based GIS public health surveillance systems (WGPHSS) between 2000 and 2013, WGPHSS are not just good platforms to share and display data in (near) real-time, but they can also include advanced analytical techniques, such as cross-correlation and cluster analysis. For instance, EpiScanGIS is an online geographic surveillance system for meningococcal disease in Germany, and this system implemented the detection and visualization of spatiotemporal clusters of disease; it integrated the SaTScan module, which can retrieve weekly reported data, perform cluster detection and store results in a PostgreSQL database (M. Reinhardt et al., 2008). However, the system is now defunct. Other similar web-based infectious diseases surveillance systems include Flumapper (Padmanabhan et al., 2014), VBD-AIR (Huang et al., 2012), Dengue-GIS (Hernandez-Avila et al., 2013), and others. As to the COVID-19, several infectious diseases surveillance systems were implemented with GIS technologies mostly to map patterns and conduct limited analysis. Several organizations and studies[2] display COVID-19 related count data along with brief statistical results via web-based GIS dashboards, but these lack analytical capabilities (Boulos & Geraghty, 2020). Dashboards are usually designed to share and visualize data on a single screen in a dynamically updating manner (Ivanković et al., 2021). Current web-based GIS infectious diseases surveillance systems have several weaknesses. One of them is the system’s lack of timely updates of the currently used dataset because of formatting issues. Second, not many web-based GIS infectious diseases surveillance systems were implemented with spatial statistical methods, and even those with such techniques did not allow users to experience different parameters to gain insight into the sensitivity of the results, and rarely were these systems able to generate dynamic results based on customized parameters. One barrier to improving this issue is

---

[2] 1) COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University; 2) The World Health Organization dashboard; 3) HealthMap.

that advanced spatial statistic methods require complicated parameters calibration and computationally demanding calculations. Another weakness of web-based GIS surveillance systems is the visualization of resulting clusters. Lan et al. (2021) call attention to advanced geovisualization techniques of COVID-19, as most WebGIS platforms merely display and share cases data in a two-dimensional framework that is challenging to identify space-time patterns. Finally, these systems are not adequate to reveal uncertainty; although some researchers have attempted to develop novel visualization to display the stability of clusters (Chen et al., 2008; Preim & Lawonn, 2020), these approaches have not been tested in a web-environment, nor have they been tested in a dynamic (i.e., space-time) environment.

1.6 Dissertation's Objectives

This dissertation is aimed at developing a web-based GIS solution that can be customized to detect space-time clusters in a timely manner, which is critical to help monitor the spread of infectious diseases. Based on this aim, three chief research objectives are formulated.

Objective 1. I conduct a systematic review of the literature around the theme of space-time cluster detection for infectious diseases. I evaluate current research trends, including methods for cluster detection, visuals of space-time clusters, and efforts to integrate cluster detection in web-based applications.

Objective 2. I propose a web-based geographic framework (see Figure 4) to detect and visualize space-time clusters for infectious diseases and demonstrate its usefulness using reported COVID-19 cases in the contiguous US. Based on this framework, I develop a tight-coupled WebGIS system that incorporates a spatial and temporal analysis module and a visualization module.

According to the systematic review, the space-time scan statistics and the LISA algorithm are found to be the most popular techniques for space-time cluster detection, with the caveat that the LISA statistic is not actually using the temporal dimension. As to the visualization module, animated and interactive mapping techniques are chosen as methods of visualization. In this system, the users can obtain visualizations of space-time clusters in various ways.

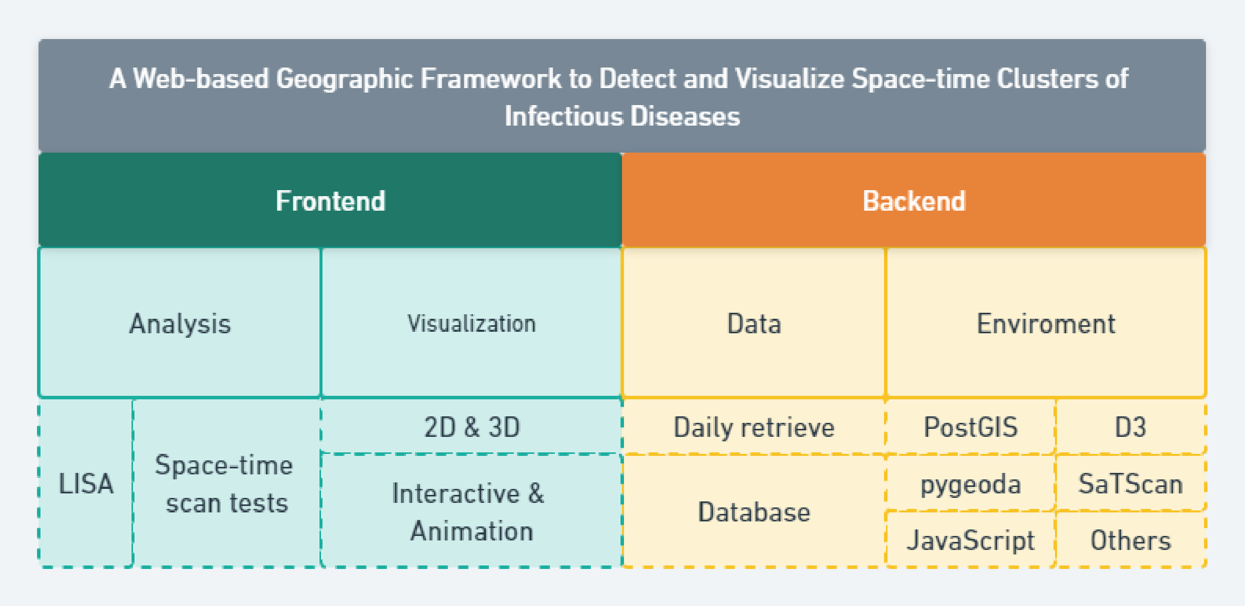


Figure 4. The illustration of the proposed framework to detect and visualize space-time Clusters of Infectious Diseases.

Objective 3. I evaluate the framework using a combination of qualitative methods. For the evaluation, I first conduct a user study with graduate students using an online survey on the web-based prototype, then a focus group interview with four health disease surveillance experts is also conducted. All collected results from the evaluation are analyzed and discussed.

1.7 Contributions

This dissertation addresses important gaps in research around the detection and visualization of space-time clusters of infectious diseases. The overarching question this dissertation attempts to answer is: *in the context of detecting infectious diseases, which methods are sufficiently robust for the timely detection of disease clusters and powerful enough to display the dynamics of these clusters, and how can we integrate data with a fine temporal scale into a robust framework?* We can tackle the first question throughout a systematic literature review, while the second question has no solid answers in the literature. My dissertation is dedicated to addressing that question, and it makes two significant contributions to the literature.

First, it integrates techniques, including 3D, WebGIS, and interactivity, to robust visualize space-time patterns, which is a novelty. While the literature on the combination of 3D, WebGIS, and interactivity is scant, the integration of 2D and the latter two techniques are commonly seen because of their advantages in accessing and exploring spatial data. However, only combining those two techniques in a 2D environment is not conducive to reflect the dynamic characteristics of space-time data. The temporal dimension is a crucial element to understand the dynamics of infectious diseases. A three-dimensional approach has the advantage of using this extra space to represent temporal information. Therefore, patterns revealed from a 3D visualization could be hidden in a 2D visualization. Several studies in computer science introduced visualizations with 3D, WebGIS, and interactivity, but none of them apply those techniques to health data with a spatial and temporal signature. Furthermore, the visualization interaction of space-time clusters enhances the ability to discover patterns from both dimensions at the same time. Thus, a robust geovisualization with the integration of 3D, WebGIS, and interactivity is desirable to reveal space-time patterns of infectious diseases.

Second, the system I propose is tightly coupled; meaning that it integrates data retrieving and preparation, analysis, and visualization into a connected and automatic manner, facilitating timely surveillance. Early cluster detection, such as daily spatiotemporal clusters detection, can facilitate “outbreak recognition and investigation, provider and community outreach, and timely intervention (Greene et al., 2016).” However, it could take a significantly longer time to obtain results when each step is loosely connected and not automated. Therefore, for daily surveillance, an automatic and routine process is needed, and this process should include 1) downloading and preparing data from the data source, 2) running statistical analyses on daily updated data, and 3) generating visualization results for that analysis.

1.8 Road Map

This dissertation is organized as follows (Figure 5). In Chapter 2, a systematic literature review on detecting spatiotemporal patterns of infectious diseases is presented. Based on gaps identified in the review, the problem statement is introduced in Chapter 3. The scope and its limitations are discussed therein as well. Chapter 4 describes the architecture and implementation of a web-based GIS system to detect space-time clusters of infectious diseases. In Chapter 5, I introduce an evaluation scheme for the proposed web-based GIS system. Finally, discussion and conclusion are provided.

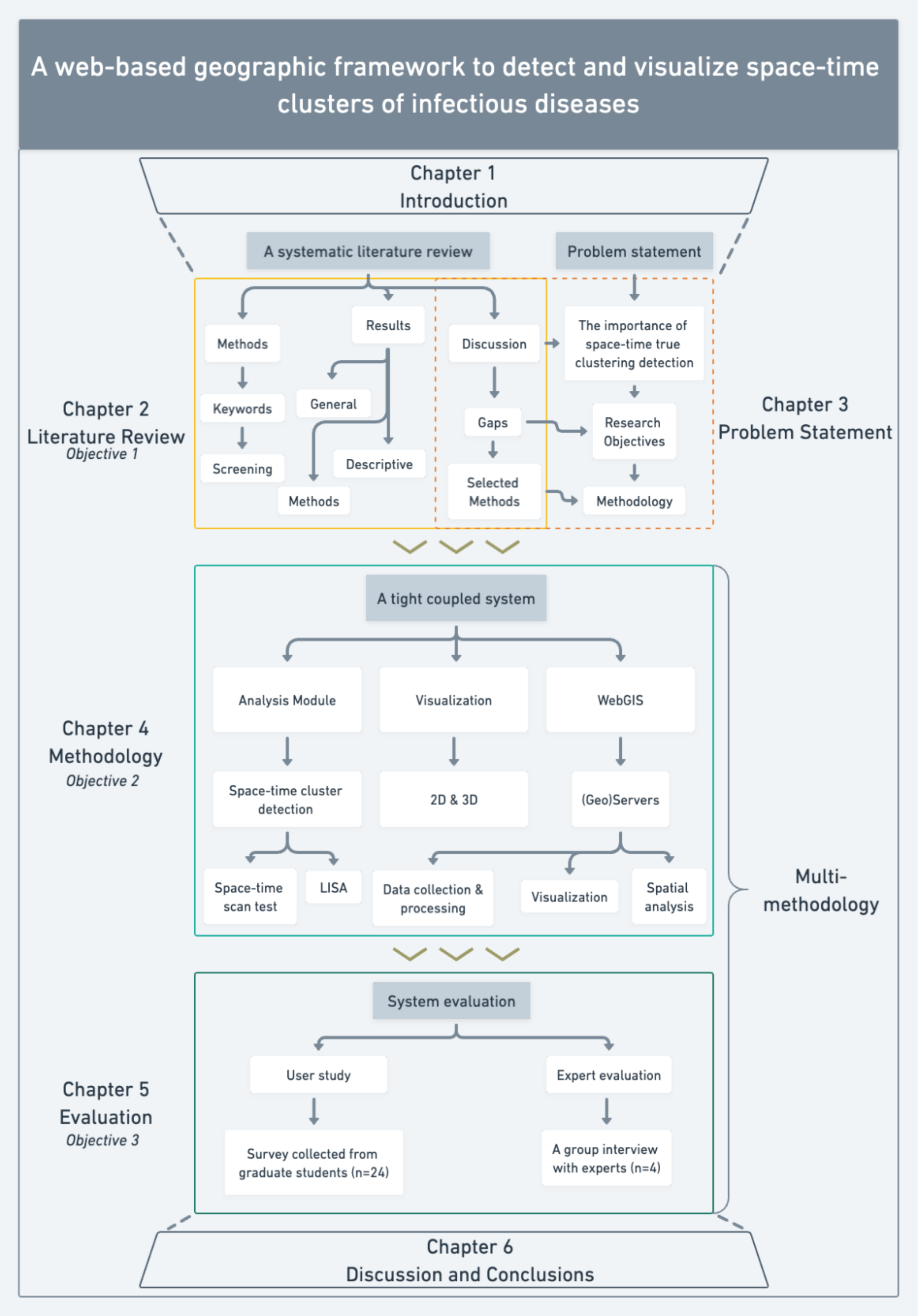


Figure 5. The structure of this dissertation.

CHAPTER 2: LITERATURE REVIEW[3]

This chapter presents a systematic literature review that reflects recent trends in space-time cluster detection for infectious diseases. The first section introduces the search and screening criteria for this review. The second and third sections provide an overview of the topic in both general and descriptive ways. The fourth section presents two very different approaches to space-time cluster detection, namely the temporal repetition of existing spatial methods and "true" space-time cluster detection methods. Meaningful findings and relevant discussion are provided in the fifth section. The last section ends this chapter with a discussion of gaps in the literature.

2.1 Search and screening strategies

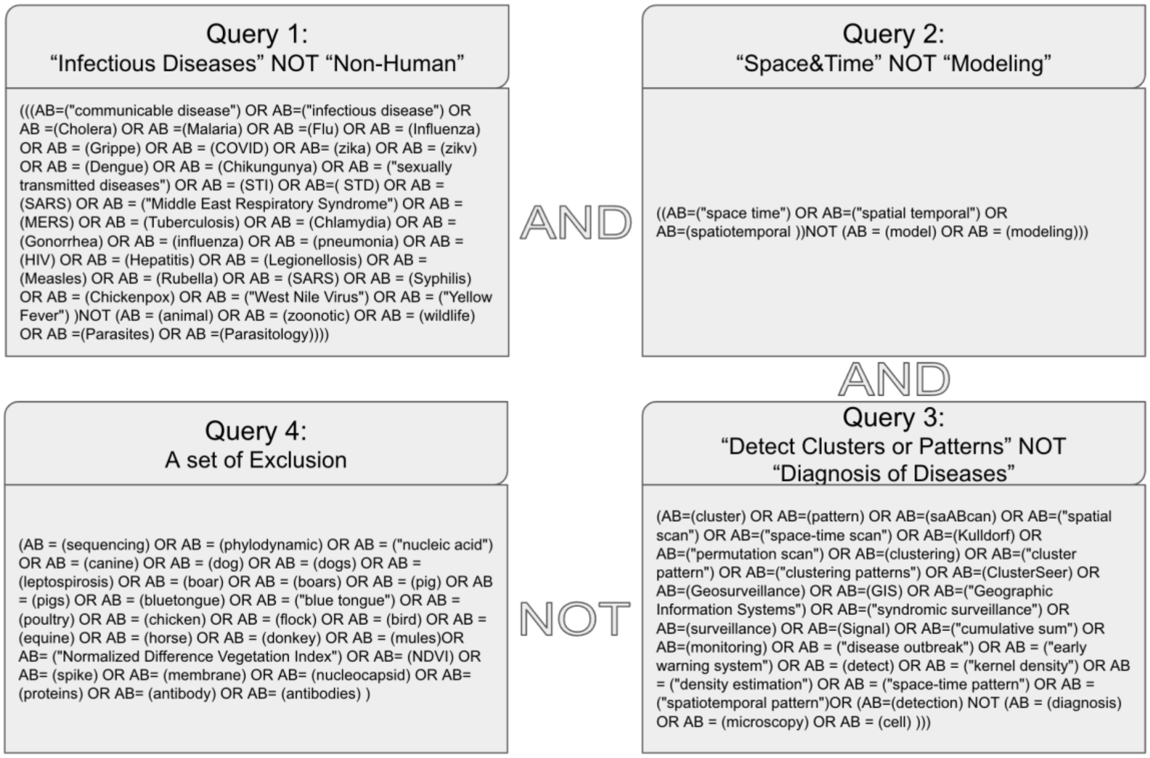
2.1.1 Search strategy including keywords

I conducted an electronic literature search for relevant articles from PubMed, the Web of Science (WoS), and Scopus databases on August 27, 2022, articulated around four main queries (see Figure 6). The first query included different types of infectious diseases but excluded non-human infectious studies. The second query attempted to incorporate articles that dealt with the spatial and temporal nature of contagious diseases (purely predictive studies, such as the ones using regression techniques, and which did not use clustering techniques, were excluded). The third query retained articles that focused on detecting spatial or space-time clusters and excluded genotype clustering papers. Finally, the fourth query further ruled out papers that were not

---

[3] This chapter was submitted to *Spatial and Spatio-temporal Epidemiology* and under review after a first round of revision.

relevant using specific. Therefore, the first three queries are connected using the “AND” operator, while the last query uses “NOT” as a set of exclusion.



Query 1:
“Infectious Diseases” NOT “Non-Human”

```
(((AB=("communicable disease") OR AB=("infectious disease") OR
AB =(Cholera) OR AB =(Malaria) OR AB =(Flu) OR AB = (Influenza)
OR AB = (Grippe) OR AB = (COVID) OR AB= (zika) OR AB = (zikv)
OR AB = (Dengue) OR AB = (Chikungunya) OR AB = ("sexually
transmitted diseases") OR AB = (STI) OR AB=( STD) OR AB =
(SARS) OR AB = ("Middle East Respiratory Syndrome") OR AB =
(MERS) OR AB = (Tuberculosis) OR AB = (Chlamydia) OR AB =
(Gonorrhea) OR AB = (influenza) OR AB = (pneumonia) OR AB =
(HIV) OR AB = (Hepatitis) OR AB = (Legionellosis) OR AB =
(Measles) OR AB = (Rubella) OR AB = (SARS) OR AB = (Syphilis)
OR AB = (Chickenpox) OR AB = ("West Nile Virus") OR AB = ("Yellow
Fever") )NOT (AB = (animal) OR AB = (zoonotic) OR AB = (wildlife)
OR AB =(Parasites) OR AB =(Parasitology))))
```

Query 2:
“Space&Time” NOT “Modeling”

```
((AB=("space time") OR AB=("spatial temporal") OR
AB=(spatiotemporal ))NOT (AB = (model) OR AB = (modeling)))
```

Query 3:
“Detect Clusters or Patterns” NOT
“Diagnosis of Diseases”

```
(AB=(cluster) OR AB=(pattern) OR AB=(saABcan) OR AB=("spatial
scan") OR AB=("space-time scan") OR AB=(Kulldorf) OR
AB=("permutation scan") OR AB=(clustering) OR AB=("cluster
pattern") OR AB=("clustering patterns") OR AB=(ClusterSeer) OR
AB=(Geosurveillance) OR AB=(GIS) OR AB=("Geographic
Information Systems") OR AB=("syndromic surveillance") OR
AB=(surveillance) OR AB=(Signal) OR AB=("cumulative sum") OR
AB=(monitoring) OR AB = ("disease outbreak") OR AB = ("early
warning system") OR AB = (detect) OR AB = ("kernel density") OR AB
= ("density estimation") OR AB = ("space-time pattern") OR AB =
("spatiotemporal pattern")OR (AB=(detection) NOT (AB = (diagnosis)
OR AB = (microscopy) OR AB = (cell) )))
```

Query 4:
A set of Exclusion

```
(AB = (sequencing) OR AB = (phylodynamic) OR AB = ("nucleic acid")
OR AB = (canine) OR AB = (dog) OR AB = (dogs) OR AB =
(leptospirosis) OR AB = (boar) OR AB = (boars) OR AB = (pig) OR AB
= (pigs) OR AB = (bluetongue) OR AB = ("blue tongue") OR AB =
(poultry) OR AB = (chicken) OR AB = (flock) OR AB = (bird) OR AB =
(equine) OR AB = (horse) OR AB = (donkey) OR AB = (mules)OR
AB= ("Normalized Difference Vegetation Index") OR AB= (NDVI) OR
AB= (spike) OR AB= (membrane) OR AB= (nucleocapsid) OR AB=
(proteins) OR AB= (antibody) OR AB= (antibodies) )
```

Figure 6. Search queries.

2.1.2 Screening and selection of criteria

Two individuals (myself and my committee chair) independently screened the title and the abstract of the articles that matched our inclusion/exclusion criteria, in an effort to determine whether they needed to be fully reviewed. Articles were included if spatial and spatiotemporal analytical techniques were applied for the detection of clusters of infectious diseases prevalence rates for human populations (i.e., not animals). In other words, this dissertation mainly focuses on the space-time cluster detection of incidence, not accounting for the transmission risk generated by human movement in space and time. I used the Cohen’s Kappa Statistic to evaluate the agreements between both reviewers for the screening, resulting in a *k* value of 0.77, which is considered a substantial agreement according to (Landis & Koch, 1977). For each manuscript where there was a disagreement, the two authors discussed the validity of the article; disagreement could stem from a lack of clarity and depth in the title or in the abstract, or that the abstract was misleading (for instance several papers discussed clustering of Dengue Fever, but the focus was on mosquito distribution). For each paper that had a disagreement, both authors reviewed the full text together, and arrived in a consensus. Only after both authors agreed on all manuscripts, the full text was reviewed together by both authors to confirm that those articles met the criteria described above.

2.2 General summary

The searching and screening processes are summarized in Figure 7. Using the four search criteria, a total of n = 2,887 articles published from 1974 to 2021 were identified from the PubMed, Web of Science, and Scopus databases (n = 677, n = 540, n = 1,670 respectively). After removing duplicates (n = 811) and review articles (n = 79) returned by all the databases, that set of articles was reduced to n = 1,996 papers. We further excluded n = 1,538 articles during the screening phase; some articles were related to raster studies (n = 404) or not good sources (n = 45), while others were identified as not relevant to the topic (e.g., discussing habitats of mosquitoes; n = 1,089). We consider not good sources including articles in non-English, not accessible on-line, etc. For the latter, articles were excluded when 1) there was no evidence of using space-time cluster analysis; 2) the study dealt with non-human infectious diseases; or 3) spatial regression or modeling methods were the main methods in the paper. The remaining articles (n = 458) were fully vetted for their eligibility. Of those, n = 104 were further flagged because they were not related to our search (e.g., review papers, regression-based papers, non-human diseases, etc.). Once this process was completed, a total of n = 354 articles spanning 44 years from 1977 to 2021 were included in this literature review with n = 332 articles (94%) that could be considered as application papers and n = 22 articles that were focused on methods. The list of articles can be found in the appendix.

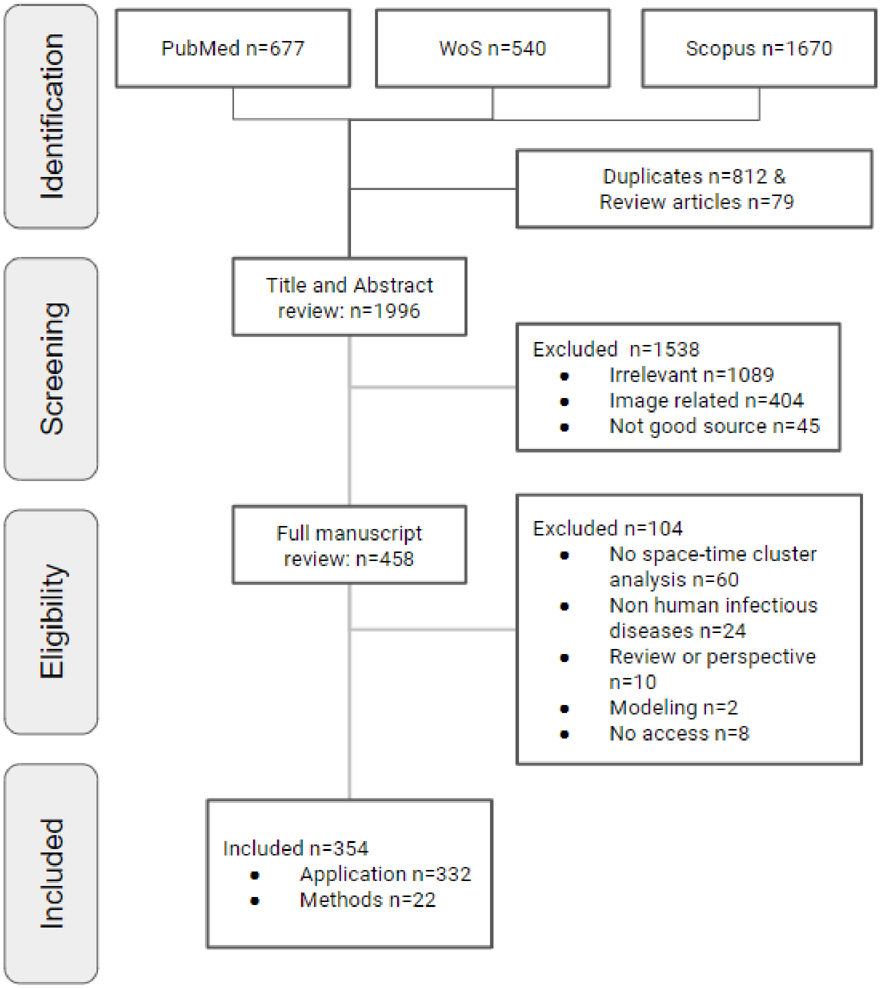


Figure 7. Searching and screening results.

2.3 Descriptive summary

This section provides a descriptive summary of our review according to disease types, study area, discipline, and data aggregation level. The number of articles published per year is summarized in Figure 8, suggesting a marked increase every year, and especially so in 2020 and

2021 many COVID-19 studies (19 out of 53 in 2020 and 40 out of 68 in 2021) contributed to this increase. I illustrate our results for papers from 1977 forward since the number of articles prior to that date were not relevant to the criteria.

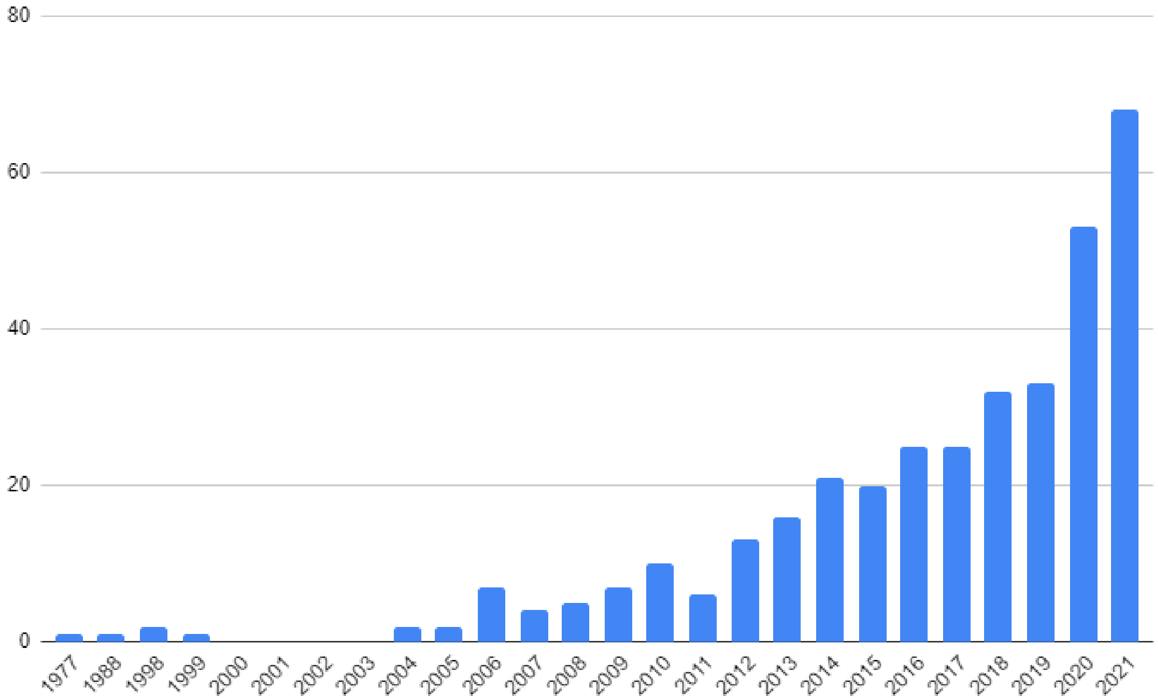


Figure 8. Published articles per year (note that from 1978 to 1988 and 1988 to 1998 there was no papers reported).

Disease types. Figure 9 summarizes the frequency of articles by disease type (Dengue Fever, COVID-19, Tuberculosis, Malaria, Sexually Transmitted Infections, and other infectious diseases), while Figure 10 reports the diseases which are most studied, year by year[4].

Overall, the largest number of articles was reported around vector-borne diseases (VBDs, n = 142). The most-reported VBD was Dengue Fever (n = 83), a disease caused by the Dengue virus, spreading from humans to humans through infected mosquitoes. Other VBDs with the same pathways, such as Chikungunya (n = 2) and Zika (n = 2), received less attention. A few

---

[4] Figures 10 and 12 were started in 2004 because of the gap in the literature, and only three articles were searched before 2000.

studies reported the space-time prevalence of multiple VBDs in the same paper, including Dengue Fever, Chikungunya, and/or Zika. Malaria, another significant vector-borne disease caused by a parasite, received less attention (n = 31) than Dengue Fever. From Figure 10, the number of studies related to Dengue Fever and malaria peaked in 2017 and 2018 but has decreased since then.

Articles related to airborne diseases (n = 139 including COVID-19, Tuberculosis, Influenza, and Respiratory Infections) formed the second most significant category. Two airborne diseases were ranked among the top five, namely COVID-19 (n = 60) and tuberculosis (n = 47). Influenza (n = 25) and Respiratory Infection (n = 7) were also ranked among the top ten of most documented infectious diseases. While the earliest articles on COVID-19 were published in 2020, the number of studies in 2021 was at least five times that for any other diseases (see Figure 10).

STIs (n = 37) include infection from human immunodeficiency virus (HIV), Syphilis, Hepatitis B, Chlamydia, and others transmitted through sexual contact. Although the number of studies was much smaller as to the other two types of infectious diseases, it shows a steady increase since 2015, according to Figure 10.

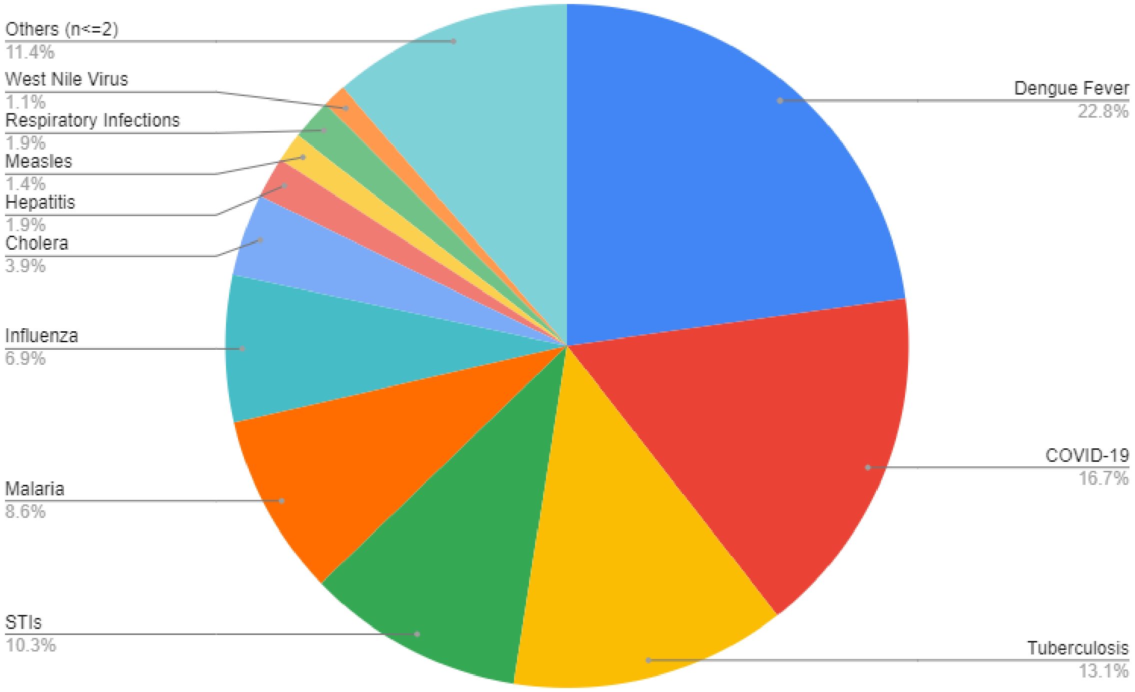


Figure 9. Number and percentage of articles related to infectious diseases. When the number of articles for a particular disease was less than three, they were categorized as "others".

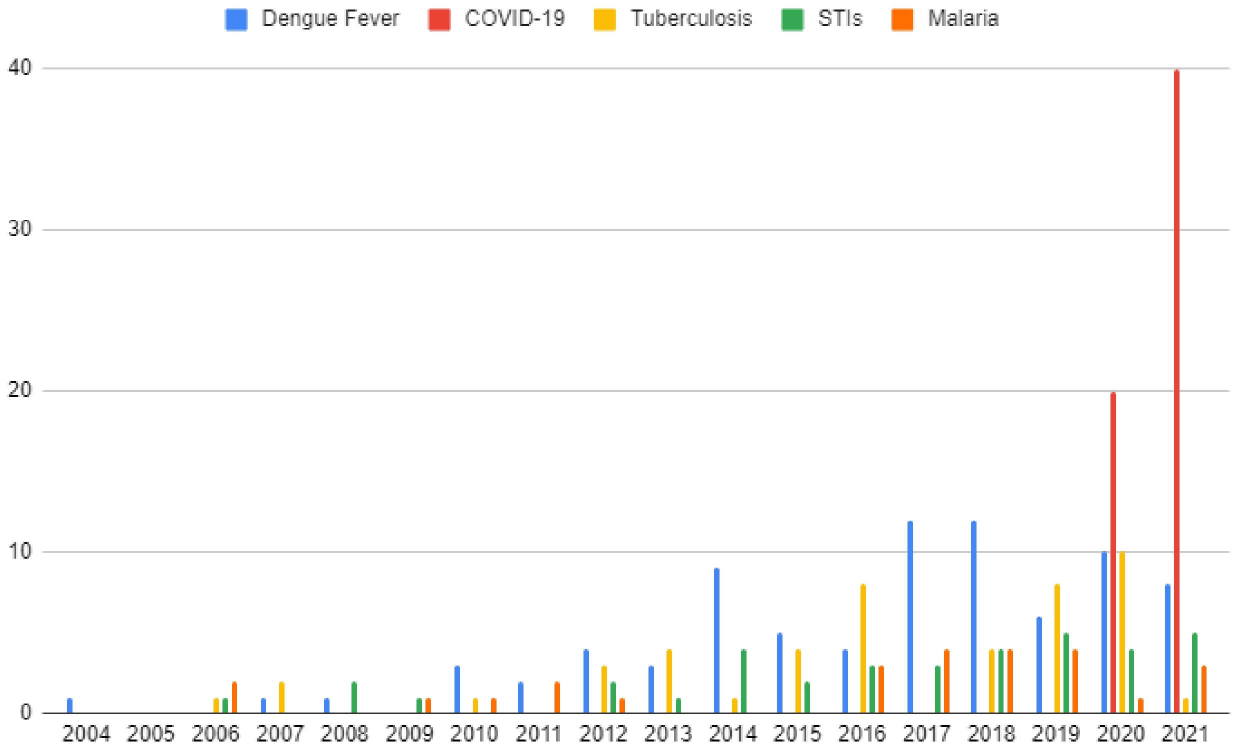


Figure 10. The five most reported diseases.

Study areas. Figure 11 maps the number of case studies based on the countries or regions where the data originated from. The top 3 countries that have experienced the most studies were China (n = 76), US (n = 35), and Brazil (n = 31). Among studies in China, more than half (n = 35) were papers on airborne transmission including COVID-19 (n = 18), followed by VBD studies (n = 17). In the US, almost two third (n = 22) of the articles focused on airborne diseases, and half of these airborne disease studies related to COVID-19 (n = 12). As to Brazil, most studies also related to airborne diseases (n = 13 including n = 8 of COVID-19) and VBD (n = 10). It is worth noting that no studies were conducted in countries or regions located in central Africa, eastern Europe, and western Asia.

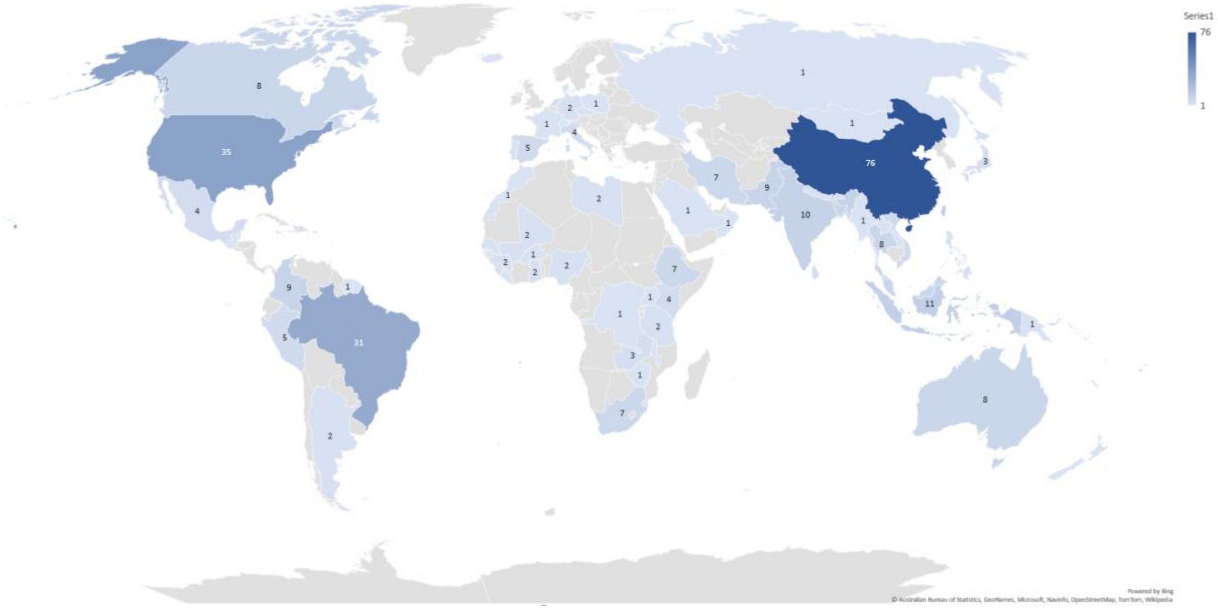


Figure 11. Variation in the number of case studies by country and region.

Discipline. We wanted to investigate the role of different disciplines for each article, because many studies are collaborative in nature. The disciplines were identified by extracting the first author’s affiliation and the type of journal. The results showed that more than 60% of articles were published by authors from epidemiology or public health, while less than 30% were from geographers or similar background. In addition, a small percentage of studies were published from other disciplines, such as bioinformatics and mathematics.

Figure 12 reflects the temporal trends of papers across the different disciplines from 2004 to 2021. The number of studies in epidemiology increased annually, while the number of studies originating from geographers experienced a sharp increase since 2020. This growth is consistent with the rise of COVID-19 studies, as 26 studies related to COVID-19 (44% of all COVID-19 studies) were published from within geography. More geographers have been involved in studies of COVID-19 and other infectious diseases, particularly in detecting spatial and temporal patterns.

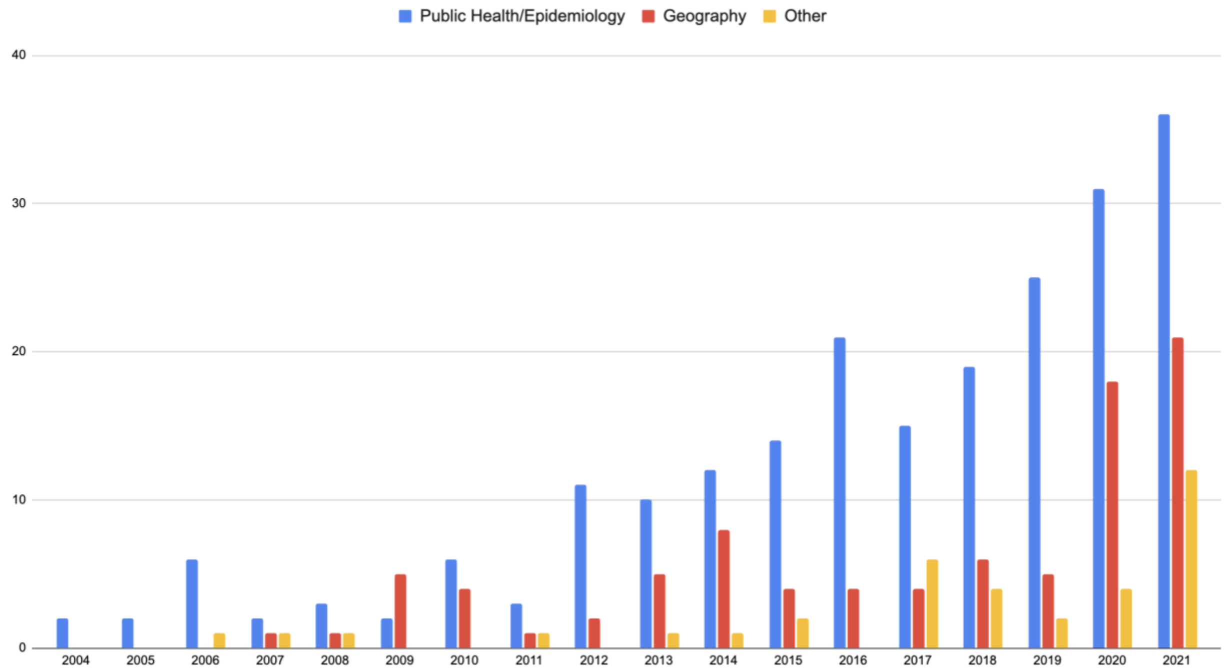


Figure 12. The number of studies by discipline and year.

Data types. About a fourth of studies used point data, while others used aggregated data. One important reason for the dominance of areal data is that such data types are more readily available, partly because of an effort to protect patient confidentiality by preventing disclosure of a patient's identity (Goovaerts, 2009). In addition, for those studies using aggregated data, the geographic scale was either at the county (27.4%) or local level (57.4%), while only 11.6% of studies were at state (10%) or country (2%) level.

2.4 Space-time cluster detection methods

This section summarizes the most common techniques of space-time cluster detection methods found in the literature (see Figure 13 for a summary of the frequencies). The most popular methodological approach in the literature was the Kulldorff's space-time scan statistics (n = 205 articles). This number is almost twice as high when compared to studies that used the (Global) Moran's *I* method (n = 79), the second most popular technique. The third and fourth most popular techniques are the Local Moran's *I* (n = 65) and the Local *Gi** (shortened for Getis-Ord, n = 44). It is worth noting that several articles compared and contrasted the results of more than one technique in the same paper. The purely spatial scan statistics was also very popular (n = 32).

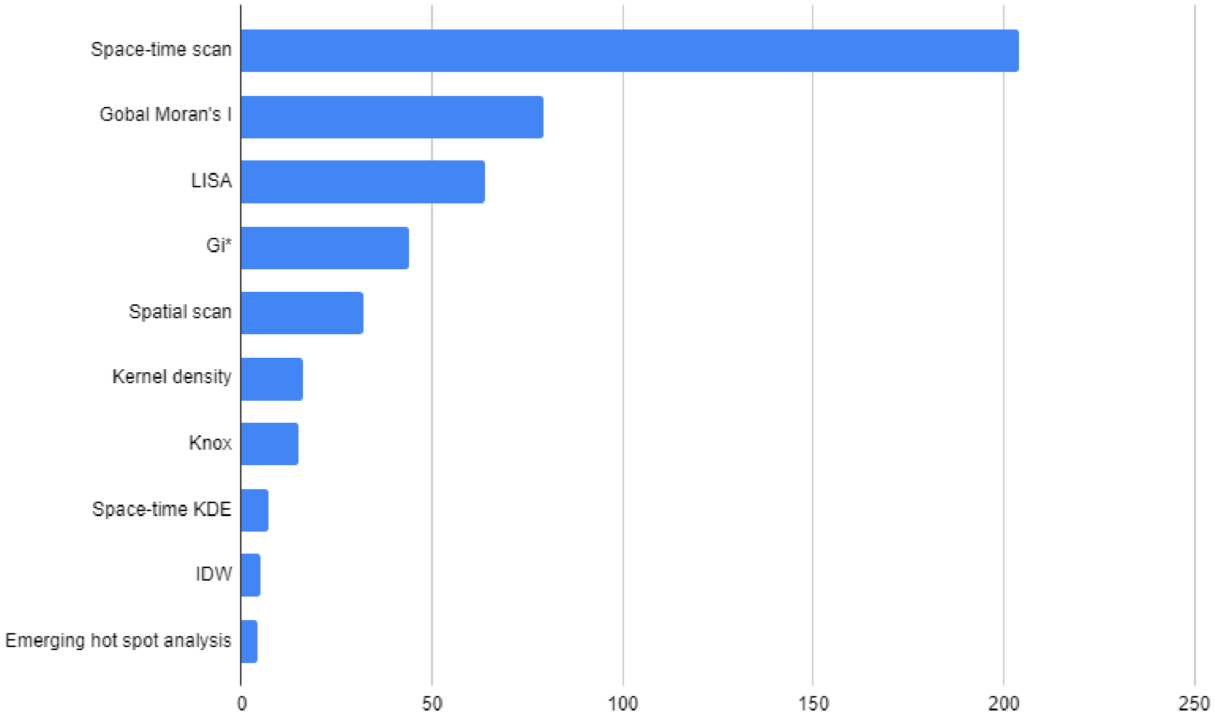


Figure 13. The number of studies by methods.

I further distinguish among two distinct approaches to perform space-time detection: either a temporal repetition of an existing spatial method or a true, space-time approach that explicitly considers space and time. Table 1 shows the number of studies using those two approaches along with the data type that was used, suggesting that more studies used true space-time methods instead of a temporal repetition of spatial methods.

Table 1 The number of studies using two types of space-time detection methods.

	Temporal repetition of spatial methods	Space-time methods	Both	Total
Point	23	48	12	83
Aggregated	87	134	55	276
Total	110	182	67	359 (n = 4 using both point and aggregated data)

2.4.1 Temporal repetition of spatial methods

A temporal repetition of a spatial method can help to identify space-time patterns of a disease over a certain time range. Generally, this approach is based on two types of methods: spatial smoothing and interpolation for point data and purely spatial cluster detection for both point data and aggregated data.

Spatial smoothing and interpolation. A total of 23 studies used interpolation and smoothing methods, including KDE, to reconstruct the spatial variation of the patterns across multiple timespans. Several studies applied interpolation methods to detect space-time clusters by generating continuous estimated surfaces from point data at multiple time intervals (See McIntosh et al., 2018; Singh & Chaturvedi, 2021). For example, de Azevedo et al. (2020) used KDE to create yearly density maps of dengue outbreaks in Brazil from 2000 to 2018 and to identify outbreaks. In an article by Pardhan-Ali et al. (2012), the authors used ordinary kriging to generate a relative risk map of notifiable gastrointestinal illness in the Northwest Territories of Canada, while relative risks were estimated using the results of the spatial scan test.

Although temporal extensions for these methods have been proposed in the literature (such as STKDE, space-time interpolation), they are computationally demanding, and despite one notable exception (Hohl et al., 2022), these approaches do not take population change into account during the period under consideration. The results of the STKDE are best visualized in a three-dimensional framework, however this is computationally demanding and can be cognitively challenging. Therefore, a simple repetition of the same method for different time ranges is generally preferred to understand the temporal variation of a disease (see de Azevedo et al., 2020; Sifuna et al., 2018).

Spatial cluster detection methods. Among that implemented a temporal repetition of spatial methods, most of them applied spatial cluster detection (e.g., LISA) instead of spatial smoothing or interpolation. Few studies applied the nearest neighbor techniques (n = 3) to detect the scale at which clusters were dominant. Some papers used *k*NN to generate spatially smoothed dengue incidence maps, but not for the cluster detection (e.g., Acharya et al., 2016).

From the literature review, n = 79 articles implemented the global Moran's *I* statistic to inform on the presence of areal clustering. For example, Lippi et al. (2020) used the Moran's *I* statistic on annual aggregated dengue cases at the level of the health districts and polyclinic administrative catchment (PAC) areas in Barbados from 2013 to 2016. Similarly, Yu et al. (2020) conducted the Moran's *I* statistic on annual aggregated rates of pulmonary tuberculosis (PTB) in counties of Chongqing, one of the biggest cities in China, from 2011 to 2018, suggesting significant clusters each year.

Several studies (n = 100) used either the LISA or Gi* algorithm to detect local clusters. In Lippi et al. (2020), the LISA algorithm identified the dynamic patterns of both high and low clusters of Dengue Fever, revealing a shift in the spatial patterns of clusters at a local level (PAC level). Yu et al. (2020) also conducted the local Gi* statistics for PTB at the county level in Chongqing, China, reporting yearly statistically significant clusters. As the Gi* statistics compares local estimates of spatial autocorrelation with global averages to detect hotspots, this method may not be suitable for small sample sizes, implying high levels of global autocorrelation (Getis & Ord, 1996).

Instead of considering spatial autocorrelation, Kulldorff's spatial scan statistic (n = 32) is applied to detect the presence of local clusters of infectious diseases (see Rocheleau et al., 2020).

Many studies used this scan statistic to estimate the relative risk (RR) or risk ratio, which can help to compare this risk of infection among areas (see Rejeki et al., 2019; Sloan et al., 2020).

2.4.2 Space-time methods

Unlike temporal repetition of spatial approaches, space-time methods include both the spatial and the temporal dimensions in statistical tests. The first literature from my search applied a trend-surface analysis by using the time as the third variable to generalize cubic surfaces to map the space-time distribution of an epidemic in a Brazilian city in 1956 (Angulo et al., 1977). Some of the most popular techniques in this category from the literature are introduced as follows.

A few studies applied the Knox test (n = 15), and in ten articles, Knox was applied to Dengue Fever (e.g., Tran et al., 2004; Vazquez-Prokopec et al., 2010; Wen et al., 2012). One possible reason is the clear space-time transmission among infected mosquitoes in dengue disease, making it easy to set the critical thresholds of space and time distance. However, because it is a global test, the Knox test is not good at visualizing cluster information on the map. Therefore, some papers have relied on additional visualization methods to display these space-time patterns. For instance, Rotela et al. (2017) used the Knox test for spatial-temporal analysis and KDE with a 300-meter radius to show the density of Dengue cases in Cordoba, a city in Argentina.

Only five studies have used the space-time *K* function, primarily due to heavy computational requirements. Hohl et al. (2016) for instance calculated spatial and temporal bandwidths using the space-time *K* function on daily Dengue Fever cases in Cali, Colombia from 2010 to 2011. They used these results (different bandwidths) as inputs to estimate the space-time kernel density, and ultimately visualized results in a three-dimensional space-time cube.

According to my review, the space-time scan statistic is the most popular technique to detect clusters for aggregated data (n = 204). In a recent study, Hohl et al. (2020) used a prospective space-time scan statistic to estimate clusters of COVID-19 in the US at the county level. One possible reason for the popularity of the scan statistic lies in its ability to incorporate multiple covariates (e.g., Whiteman et al., 2019), including population. For example, in infectious diseases, the population is a critical variable that needs to be considered in the analysis, while many other pattern detection methods cannot include this covariate directly. Also, some studies focus on interpret spatial result only, even though they used space-time methods in the studies. For example, Tadesse et al. (2013) concluded that both purely spatial and space-time scan tests detected similar and significant high-risk clusters of smear-positive TB cases in a district of Ethiopia, but no temporal information on clusters was provided in the study. In another article, even though Gurjav et al. (2015) claimed to use the retrospective space-time scan test and detect three TB clusters in Mongolia from 2006 to 2012, no temporal characteristics of the clusters were provided.

Space-time clusters can be detected or suggested by either space-time methods (e.g., the space-time scan test) or repetition of purely spatial methods (e.g., the LISA algorithm). Although both methods return geographic units considered to be in a cluster, the results are generally different. Fuentes-Vallejo (2017) implemented both the Gi* statistic and the space-time scan statistics to compare the different sensitivity of parameters to detect local clusters from these two methods. Their results show that spatial clusters (using the Gi* statistic for each year) and space-time clusters from the scan test were located in different regions, although some results were overlapping. They claimed that this different spatial and spatiotemporal clusters distribution is possibly due to different territorial dynamics.

Overall, few studies applied cluster detection approach into a web environment. Markus Reinhardt et al. (2008) launched an online GIS, EpiScanGIS, integrated SaTScan to monitor the invasive meningococcal disease in Germany, while the system was not functioning anymore. Besides EpiScanGIS, only five recent studies on COVID-19 published in 2020 and 2021 deployed their detection approach into a web-based application. Four of them used the prospective space-time scan statistics, while one used another algorithm (a modified space-time density-based spatial clustering of applications with noise). Two studies are similar in that their online systems map daily clusters as an animation based on daily COVID-19 cases (Hohl et al., 2020; Rosillo et al., 2021), and another study, instead of mapping clusters, animated daily relative risk and cluster frequency results (Lan et al., 2021). Two other studies (De Ridder et al., 2021; Güemes et al., 2021) detected space-time clusters of COVID-19 symptoms by collecting symptoms from users. But in none of these five systems were users able to conduct customized space-time analysis.

Although most studies use traditional visualization techniques (e.g., small multiples) to map space-time clusters, a few studies (n = 13) introduced novel methods to portray these clusters. The most popular method (n = 11) maps the temporal characteristics of space-time clusters using the third dimension. Among those studies, six of them used three-dimensional (3D) figures to represent cluster as cylinders with the height as the time (e.g., see Desjardins et al., 2018), and five of them rendered space-time clusters into different colors or/and transparency in the 3D volume, which have no well-defined boundary (e.g., see Kuo et al., 2018). Other than 3D methods, three studies used ring maps (Tang et al., 2019), calendar-based visualization (Wu et al., 2021), and bivariate and spike maps as different visualization of space-time clusters (Lan et al., 2021).

2.4.3 COVID-19 studies

From our systematic review, several studies in 2020 and 2021 were applied to the analysis of COVID-19 outbreaks (n = 60). The number of COVID-19 articles that used a space-time algorithm (n = 34) was nearly the same as the papers that used a repetition of a spatial clustering algorithm (n = 32); six studies use both algorithms. The number of papers was split among epidemiologists (n = 25) and geographers (n = 26), and the remaining studies (n = 9) were led by researchers from other disciplines. Most studies used aggregated data (n = 51), and 48 of those studies were conducted at the county or finer (i.e., more local) level.

2.5 Findings

This chapter presented a systematic literature review that reflected recent trends in space-time cluster detection for infectious diseases. I searched and selected 354 articles from PubMed, Web of Science, and Scopus databases. I noted important findings as followed. First, the number of articles exhibited a continuous increase from 2004 to 2021, and nearly doubled from 2019 to 2020 and 2021 due to the emergence of research related to COVID-19. Second, most of the articles were application type papers featuring spatial and spatiotemporal techniques to detect space-time clusters of infectious diseases. Third, I noted that very few studies attempted to publish their results over web-based interfaces, and the visualization results were for the most part two dimensional. Fourth, most research was focused on airborne diseases, followed by vector-borne diseases. Fifth, most studies were conducted in China, US, and Brazil. Sixth, most studies used aggregated data instead of point data. Seventh, there was also a greater number of studies using “true” space-time detection algorithms as compared to papers only using temporal repetitions of the same spatial method. Along those lines, the most popular methods were the space-time scan statistics, the global Moran’s *I*, and the LISA statistic.

2.6 Gaps in the literature

Several research gaps can be identified in the current understanding of this field. From the literature review, few online applications or platforms have implemented spatial or space-time analytical techniques to identify space-time clusters of infectious diseases. Only five online applications incorporated those techniques, and all of them are COVID-19 related with limited functionality. Some commercial GIS platforms, such as ArcGIS Online, have implemented specific spatial cluster detection techniques (e.g., the LISA algorithm) online. However, these platforms require credits to access these resources. Further, none of these applications can conduct an analysis that would account for both spatial and temporal dimensions simultaneously. In other words, users can only conduct a repetition of these purely spatial methods by performing the same analysis for different temporal intervals. Other issues that may prevent the deployment of web-based platforms lie in their scalability, computation, and high-level programming skills to develop such systems.

Another critical question requiring more attention is how to visualize space-time clusters (only 13 articles deliberately discussed this issue). Using the space-time cube as a framework, the third dimension can be used to visualize the dynamics of space-time clusters better, potentially uncovering hidden space-time patterns. New technologies for web-based data visualization, such as WebGL and D3.js, can be used to visualize space-time patterns on the internet.

As far as the literature review is concerned, no study implemented space-time analysis and visualization into one web platform. For space-time cluster detection of infectious diseases, those two components are complementary. A good analysis of space-time clusters could undermine the conclusion without proper geovisualization and vice versa. Tight-coupled systems that can handle both analysis and geovisualization are greatly needed.

I also underline several methodological concerns in applying such methods, promising to improve future studies. It is worth noting that repeating a spatial clustering algorithm across time instead of using a true space-time cluster detection technique may cause an increase of type I and type II errors (false positive and false negative, respectively), and almost half of the studies found in the literature review used a temporal repetition of a spatial clustering method alone or together with “true” space-time detection methods. In addition, the temporal range used in these approaches is usually arbitrary (e.g., week, month, trimester) based on the dataset or personal experience; analysis conducted at different scales can further exacerbate these type I and type II errors.

As mentioned earlier, space-time statistics can be applied both retrospectively and prospectively, but each approach answers different research questions. Retrospective cluster detection conducts the analysis once and identify all existed clusters during the whole study time, while a prospective method conducts the analysis on multiple time interval (e.g., daily, weekly, or yearly) to detect ‘alive clusters’ on each end date of that time interval during the study period. The retrospective method scans to detect clusters from the end of the study to the beginning, while the prospective one moves reversely through time. Mainly, retrospective methods detect patterns for a fixed dataset, while prospective methods adapt the results considering both newly available and past data. Therefore, prospective methods are more appropriate to promptly detect the dynamics of space-time clusters, especially during an outbreak that needs a rapid response. As to COVID-19, retrospective methods can determine whether it will become a seasonal recurrent disease like flu when cases are recorded for more than one year, while prospective methods can closely monitor the change in the current situation.

Some studies also overlooked the importance of interpreting the temporal information of space-time clusters. In some reviewed articles, the temporal extent of space-time clusters was not presented, although space-time scan tests were used. Without an adequate examination of the temporal characteristic of space-time clusters, their ability to offer additional insights in the temporal dimension vanishes.

Since 2011, multidisciplinary collaboration has steadily increased. In addition, the collaboration among the academy, government, and research centers represented nearly 50% of all the publications. Both results suggest that investigating the presence of infectious diseases is best tackled by a holistic team of researchers. Thus, it is also essential to incorporate multiple levels of collaboration across academics, health agencies, and other organizations. This kind of collaboration will offer theoretical evidence to support the implementation of health policies and practical experiences to guide research design and evaluation.

Several issues warrant further investigations. First, most reviewed articles did not account for the potential effects of scale. For instance, clusters identified from cases reported at the postal code level (e.g., ZIP in the US) may not be the same as if data were reported at the county level. I argue that the comparison among multiple scales for the same study region (e.g., county and census tract levels) could provide additional insights into the mechanism of the disease under study. Second, research should more explicitly discuss the temporal dimension in cluster detection, because it can reflect the cyclicity and dynamic nature of a disease. Both spatial and temporal dimensions are equally essential for the monitoring of infectious diseases. Third, more research is needed to compare the validity of clusters found from a repetitive spatial method, or from a true space-time clustering algorithm as it can affect the risk of false alarms. Fourth, other clustering techniques such as wombling (Hossain & Lawson, 2005; Hossain & Lawson, 2010; Lu & Carlin, 2005; Monir Hossain & Lawson, 2006) -which identifies various levels of cluster boundaries- are promising, but are rarely used for temporal processes, nor in infectious diseases. There is a potential to extend these approaches in time. Fifth, this literature review did not explicitly search for papers using space-time Bayesian modeling; in fact, most of the algorithms discussed in this chapter are to describe and identify space-time clusters; as such this topic falls outside of the scope of this dissertation. Sixth, I found several studies that used the space-time K function to detect space-time clustering from events, but the inhomogeneous K function (Baddeley et al., 2000) has rarely been discussed in space and time (a notable exception is (Hohl et al., 2022)); this is partly due to the difficulty to have temporally varying information on the population itself, unless the study covers a large period of time, and fine-grained population count is available. Finally, researchers should facilitate the replication of their study, either by publishing their data and developing web-based visualization solutions.

CHAPTER 3: PROBLEM STATEMENT

The literature review in Chapter 2 has revealed some important gaps in the monitoring of infectious diseases. In this dissertation, I propose to focus on four significant issues, specifically (1) the repeated use of a spatial method which do not capture the true clusters of the disease under investigation, (2) the use of rather simplistic geovisualization techniques that do not reveal the space-time patterns of the disease, (3) the lack of a tightly integrated system that can conduct both cluster detection and visualization, but also communicate with a server where infectious data information is stored and updated frequently, and finally (4) the absence of such systems on the internet. I now elaborate on these four gaps, and briefly discuss the scope and limitations of this study.

3.1 Emerging problems from the literature review

3.1.1 Reliance on non-exact space-time cluster detection algorithms

From the systematic literature review (see Chapter 2), nearly one-third of the articles have relied on a temporal repetition of purely spatial algorithms to detect space-time clusters, while the others use “exact” approaches or both. For the former, this is a critical weakness because the temporal dimension is not explicitly accounted for and can lead to false positives and possibly false negatives. The presence of false positives, which are unknown to public health experts, may lead to public health interventions where no problem exists; likewise false negatives may reduce public health efforts in areas where a problem truly exists. It is therefore critical to rely on methods that can minimize false positives and false negatives.

3.1.2 The need for robust geovisualization to reveal space-time patterns of infectious diseases

Another important but overlooked issue is how to visualize space-time patterns (e.g., clusters, relative risk) of infectious diseases. Only 13 articles deliberately discussed how to visualize space-time patterns of an outbreak. However, it is vital to capitalize on the temporal dimension because infectious diseases are not static over time. With the inherent characteristics of spatial and temporal dimensions, spatiotemporal patterns may be hidden from traditional 2D geovisualization. Although multiple static maps or basic web maps can display changes in both dimensions by displaying all maps or information simultaneously, it is impossible to present numerous maps together, such as showing all the daily or weekly maps at the county level. Space-time visualization approaches could help uncover concealed space-time patterns hidden from statistical tables and traditional maps.

3.1.3 The need for rapid surveillance at a fine temporal scale

An effective monitoring system should be able to detect and visualize clusters in both space and time, but also use the most up-to-date reported data (e.g., cases, death, hospitalizations…) that is updated on a fine temporal scale (e.g., daily). With the development of data monitoring and sharing technologies, infectious diseases data are sometimes available at the daily or weekly level and carry a massive volume of information. For instance, during the current COVID-19 pandemic, several online dashboards, or platforms, such as Johns Hopkins and New York Times, have shared and updated COVID cases and deaths worldwide, at various levels of spatial and temporal granularity. Take the dataset in the US from the Johns Hopkins dashboard as an example. This dataset has reported daily cases or deaths at the county level for most US counties since January 2019. With the advantage of daily data availability, the essential question is how to capitalize on such a rich dataset to uncover the dynamic patterns. As time is the essence in infectious disease, the monitoring system needs to regularly retrieve and analyze daily data or as soon as the data are ready. In other words, we have the opportunity to promptly discover changing patterns of the disease. By examining daily data regularly and speedily, experts have an opportunity to prepare better ahead of the next surge of infectious spreading.

3.1.4 The need for an integrated online solution

Online web-based mapping applications can facilitate the sharing of data and information among different communities, improving response time to rapidly evolving situations. From the literature review, few online applications or platforms have implemented spatial analytic techniques to identify clusters for infectious diseases. Only five online applications incorporated those techniques, and many of them consist of limited functionality. Some commercial GIS platforms, such as ArcGIS Online, have implemented some spatial cluster detection techniques (e.g., the LISA algorithm). However, they require credits to access these online resources. Furthermore, none of these applications could conduct an analysis that would account for both spatial and temporal dimensions simultaneously. In other words, if users want to conduct a space-time analysis, they need to repeat the spatial test as many as the frequency of the temporal interval. For example, for a weekly spatial analysis of COVID cases for one year, the user will need to repeat the analysis 52 times, which could be a time-consuming process. Other issues that may prevent the web-based deployment of space-time clustering tests lay in scalability[5], computation, and that the programmer who develops these applications must possess high-level programming skills.

---

[5] For instance, ArcGIS Online can conduct hotspot analysis with credits, and it will cost more credits if the number of features increases.

Furthermore, incorporating robust visualization techniques such as web-based geovisualization could provide a much-needed option for researchers who have limited knowledge in the use of geospatial technologies. Although incorporating three-dimensional visualization techniques into a web environment can be challenging, innovative techniques of web-based data visualization, such as WebGL for 3D visualization and D3.js for dynamic, interactive data visualizations make this attempt possible.

3.2 General problem statement

Those issues reveal that significant problems remain unsolved in the context of infectious diseases. My dissertation is aimed to address those issues by integrating all essential ingredients of cluster detection analysis into one tight, online system. This online health surveillance system assembles components of data retrieving, data analysis, and data and uncertainty geovisualization into a tightly coupled GIS system. This system facilitates the communication among each module with interfaces, while in a loosely coupled system, the user needs to manually switch among each module (database, analysis, and visualization). The tightly coupled GIS system is more robust than a loosely coupled system, because it integrates data, analytical tools, and visual tools together.

Taken together, my dissertation addresses these four gaps mentioned earlier, and provide the following contributions:

- Emphasis on the exact space-time cluster detection algorithm and use of the non-exact one as a complementary resource in space-time cluster detection
- Generate innovative visualization techniques for daily updated health data with spatiotemporal information

- Develop a tight, online system to facilitate experts making rapid responses during the outbreak

I evaluated the proposed system by combined a user study and an expert evaluation. The user study was conducted by collecting survey responses from graduate students, and the expert evaluation directly collected feedbacks from experts in COVID-19 data and analysis during a group interview. Collected results and feedbacks were further analyzed.

3.3 Scope and limitations of the study

Although the proposed framework could virtually be applied to any other disease, I use COVID-19 cases in the US at the county level for the 48 conterminous states to demonstrate its usefulness. The analysis module focuses on detecting both exact and inexact space-time clusters implemented with two methods (the SaTScan and LISA statistics, respectively), but other methods could also be deployed, depending on the research questions. There are other limitations to this study. First, the case dataset cannot be uploaded by users because data security is most important. Since the prototype (described in Chapter 4) is accessible to the public, it would be challenging to handle potential privacy issues, such as those mentioned in the Health Insurance Portability and Accountability Act (HIPAA) in the US (Centers for Disease Control Prevention, 2003). Second, the uploaded case information should be in the same scale as other datasets, such as population and boundaries. For example, if one user uploaded cases at the neighborhood level, then population and boundaries at the same level are also required. Third, only cases are used to demonstrate two methods for space-time cluster detection in this prototype, but deaths or hospitalizations could also be used.

CHAPTER 4: METHODOLOGY

Space-time clustering detection is essential in spatial epidemiology, especially for infectious diseases (Pfeiffer et al., 2008). As discussed earlier, the space-time scan statistics and local indicators of spatial association (LISA) are two popular methods to detect local space and space-time clusters for infectious diseases. While the LISA statistic (Anselin, 1995) is a purely spatial method that does not take the temporal information into account, many studies have used this approach, essentially repeating it over multiple time ranges to detect space-time clusters for infectious diseases (see Ghosh & Cartone, 2020; Sugumaran et al., 2009). Unlike the LISA statistic, the space-time scan statistic "scans" the data using a cylindrical window in both space and time (Kulldorff et al., 2005). Also, unlike the LISA statistics, space-time scan statistics are not restricted by administrative boundaries (Naish & Tong, 2014), because the scan statistic searches for clusters beyond the so-called 'adjacency matrix' that is central to the LISA statistic.

The space-time geovisualization of these two methods is crucial to infectious diseases such as COVID-19, but many current COVID-19 dashboards neglect this temporal component (Lan et al., 2021). Moreover, the space-time scan statistic requires an application (SaTScan) or libraries (through R, for instance) to estimate the presence and magnitude of clusters, while a commercial GIS is still needed to visualize the space-time extent of these clusters. A tight-coupling system with different modules connected into one system can address this issue. Using this approach, epidemiologists could conduct spatiotemporal analysis and uncover the underlying pattern without the need to go from one software to another (Delmelle et al., 2011). However, to the best of my knowledge, a tight coupling system for space-time clustering detection visualization is not available.

This study proposes a web-based geographic framework for the detection and visualization of space-time clusters for infectious diseases. To demonstrate the framework, I develop an automatic surveillance system that uses the prospective space-time scan statistics and the LISA algorithm at the county level in the continental US. The system retrieves daily updated COVID-19 data. I further elaborate on the objectives of this tightly-coupled system, namely: (1) to implement automatic and customized space-time clustering detection for a given geography and specific time range and (2) to generate novel 2D and 3D visual features of space-time clustering; (3) to develop a tight coupling system that incorporates daily data updated and components for the objective one and two. This system is named *US COVID-19 YuTu* and described further in the below sections.

4.1 Data

In this study, the COVID-19 Data Repository prepared by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University is used to extract daily US COVID-19 data (JHU CSSE COVID-19 Data) at both state and county levels. Halpern et al. (2021) claimed the dataset from JHU is one of the closest to the one from CDC when compared to other commonly used COVID-19 datasets.

For this study, I use reported date is from January 22, 2020 -the date CDC confirmed the first US coronavirus case in Washington state- to the latest date for which this website is updated. Table 2 is an example of daily case data retrieved on October 18, 2021. Daily COVID-19 data are extracted and updated into corresponding databases. Attributes of the COVID-19 data include federal information processing standards code (FISPS), county name, state name, date, latitude, longitude, counts of confirmed cases, and counts of deaths. In addition, the latest available population figures and boundaries are retrieved from the US Census Bureau. The population data are in the 2019 Community Survey (ACS) 5-year estimates of the resident population for both 48 states and corresponding counties, and the used boundaries data is the 2020 TIGER/Line shapefiles[6]. Rates are calculated by dividing the number of cases -averaged over seven days- in a geographic region by the population in that region.

Table 2. Example of daily case data from JHU.

UID	iso2	iso3	code3	FIPS	Admin2	State	Country	Lat	Long	1/22/2020	…	10/16/2021	10/17/2021
84001001	US	USA	840	1001	Autauga	Alabama	US	32.53953	-86.6441	0	…	9893	9901
84001003	US	USA	840	1003	Baldwin	Alabama	US	30.72775	-87.7221	0	…	37069	37087
84001005	US	USA	840	1005	Barbour	Alabama	US	31.86826	-85.3871	0	…	3554	3556
84001007	US	USA	840	1007	Bibb	Alabama	US	32.99642	-87.1251	0	…	4216	4217
84001009	US	USA	840	1009	Blount	Alabama	US	33.98211	-86.5679	0	…	10094	10102
84001011	US	USA	840	1011	Bullock	Alabama	US	32.10031	-85.7127	0	…	1517	1517
84001013	US	USA	840	1013	Butler	Alabama	US	31.753	-86.6806	0	…	3247	3248
…	…	…	…	…	…	…	…	…	…	…	…	…	…

4.2 Method

This section introduces the framework and workflow for detecting and visualizing space-time clusters of infectious diseases using COVID-19 data in the conterminous US. The framework is based on a “tight-coupling” system with customized spatial and temporal settings. It incorporates data extraction capabilities, clustering detection, and geovisualization in a web-based GIS environment, and uses a server-side (running procedure on the server) and a client-side (running procedure on the user's web browser), as illustrated in Figure 14.

---

[6] Boundaries: https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html

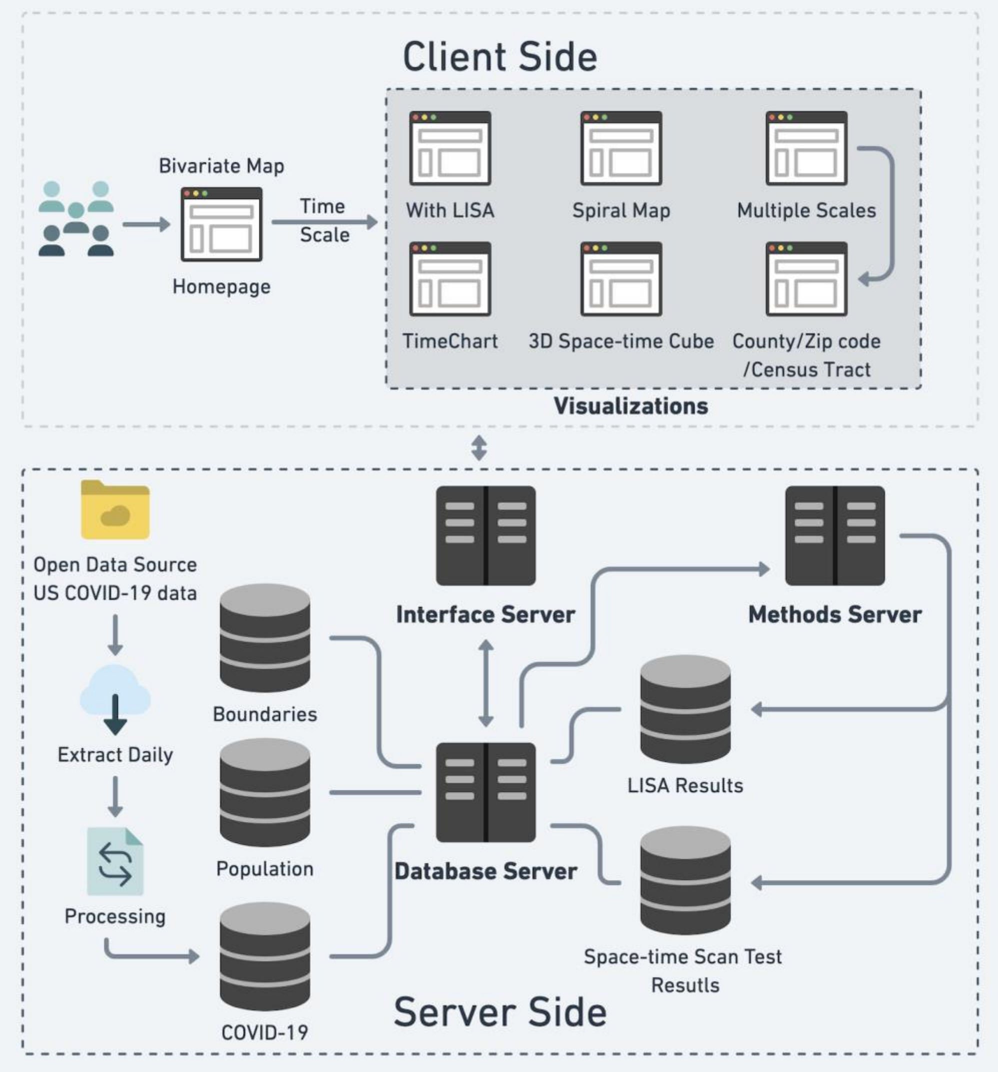


Figure 14. The framework of a tight coupling system to detect space-time clustering of COVID-19.

The **server side** is articulated around three types of servers for different purposes: a database server, an interface server, and a method server; each server in this system is represented as a docker container. The *database server* stores all the relevant input data and output results. Disease information at the county level is extracted daily from the JHU CSSE COVID-19 database and imported into our COVID-19 database using a processing script written in Python, which removes unnecessary attributes. The other two datasets are population and cartographic boundaries from the US Census Bureau. The disease detection algorithms are processed on the *methods server* and generated results. This process is repeated every night after the data is automatically retrieved and processed. The *interface server* connects the client and database sides for visualizations.

The **client side** is the graphical user interface (GUI) of the COVID prototype. It consists of basic online map functions, including zoom, pan, etc. The default homepage contains the distribution of US COVID-19 space-time clusters at the county level. The geovisualization of space-time clusters is displayed in both two and three dimensions. My framework is articulated around three modules implemented on individual servers: (a) an *analysis module* (Method Server), (b) a *visualization module* (Interface Server), and (c) a *data processing module* (Database Server).

4.2.1 Analysis module

Among various methods of disease space-time detection, the local indicators of spatial association (LISA) and space-time scan tests are two popular methods.

Local indicators of spatial association-LISA. Anselin (1995) introduced LISA as decomposition of global indicators of each individual observation, which can detect significant local clustering around an individual location and recognize the spatial nonstationary with outliers. For a region *i*, the local indicators of spatial association I_i defined as:

$$I_i = \frac{(n-1)(x_i - \underline{x})}{\sum_{j=1}^{n} (x_j - \bar{x})^2} \sum_{j=1}^{n} w_{ij}(x_j - \bar{x}) \qquad \textit{Eq. 1}$$

where x_i is the attribute of the variable of interest (here, the disease rate) in region *i*, $\bar{x}$ is the mean of x_i (*i*=1…, n), w_{ij} is the spatial weight between regions *i* and *j* (typically derived from an adjacency matrix). The results of the LISA algorithm group regions into different groups (e.g., High-High, Low-Low, High-Low, Low-High) with an associated *p*-value. When a location is categorized as High-High/Low-Low, it denotes a region exhibiting high/low rates, surrounded by other regions with similar, high values/low values. A High-Low category characterizes a county with high rates, surrounded by low rates. This could be indicative of an area experiencing a rapid increase in cases that is more than what would be expected, while surrounding regions do not experience such rapid growth. Low-High represents outlier regions of low value surrounded by high values. The LISA statistic (Anselin, 1995) is purely a cross-sectional method that does not take the temporal information into account. There are ample examples of this repetitive approach to identify clusters of infectious diseases (see Ghosh & Cartone, 2020; Sugumaran et al., 2009). However, the LISA statistic is likely to lead to the discovery of false negatives and false positives.

Space-time scan statistics. Kulldorff and Nagarwalla (1995) introduced the spatial scan statistic as a test for detecting clusters by assessing the likelihood ratios of events inside and outside of circular scanning windows, adjusted for the density of the population. The radii of the windows are varied continuously from zero to the maximum bandwidth, e.g., to a size containing a certain percentage of the population. The window with the maximum likelihood ratio is defined as a cluster, and only regions located within this window are considered to ‘belong’ to that cluster.

Kulldorff et al. (1998) further expand the spatial scan statistics to incorporate the temporal dimension by adding the circle’s height to represent the time (Figure 15). Thus, each cylinder represents the scanned geographic region (the circle projected to the area) within a temporal range (the hight of the cylinder).

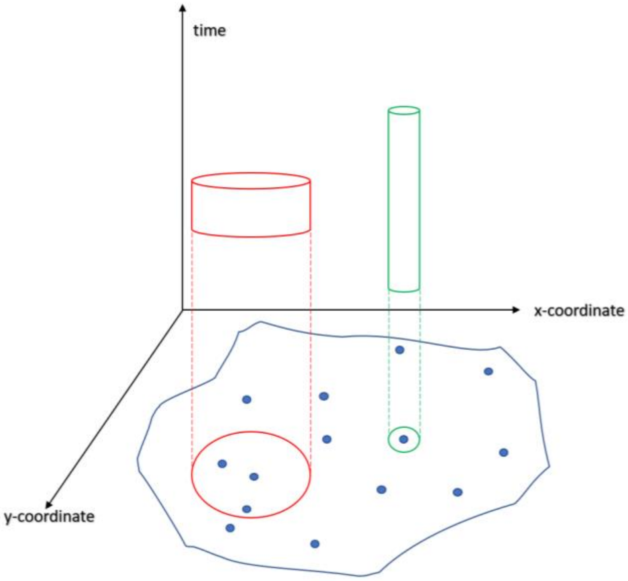


Figure 15. The illustration of space-time scan statistics.

As the statistic is designed to detect clusters, the null hypothesis H0 is that the risk of infection within a cylinder Z is similar to this risk outside the cylinder; the alternative hypothesis Ha is that the risk of infection within a cylinder Z is larger than this risk outside this cylinder. Accordingly, the expected number of cases (μ) based on the null hypothesis is

$$\mu = p \times \frac{N}{P} \qquad \textit{Eq. 2}$$

with p the population in the cylinder, N the total number of cases within the cylinder Z, and P the total population within the study area. Thus, the maximum likelihood ration to identify space-time clusters is defined as:

$$\frac{L(Z)}{L_0} = \frac{\left(\frac{n_Z}{\mu(Z)}\right)^{n_Z}\left(\frac{N - n_Z}{N - \mu(Z)}\right)^{N-n_Z}}{\left(\frac{N}{\mu(T)}\right)^{N}} \qquad \textit{Eq. 3}$$

where $L(Z)$ is the likelihood function for the cylinder Z, and L_0 is the likelihood for the null hypothesis H0, n_Z is the number of cases in the cylinder Z, $\mu(Z)$ is the number of expected cases in cylinder Z, and $\mu(T)$ is the total number of expected cases within all time periods in the study area. I reported 1) the relative risk for each location (RR of the location), defined as the estimated risk (observed/expected) within the location divided by the estimated risk outside the location, and 2) the relative risk for the cluster that the location belonged to (RR of the cluster), representing the estimated risk divided by the risk outside of the cluster. For instance, if the RR of the county is 1.4 and its RR of the cluster is 2.5, then this county is 1.4 times more likely to be exposed to COVID-19 while it also belongs to a cluster that is 2.5 times more likely than outside this cluster.

The space-time scan statistic on the other hand “scans” the data using a cylindrical window in both space and time (Kulldorff et al., 2005). Also, unlike the LISA statistics, space-time scan statistics are not restricted by administrative boundaries (Naish & Tong, 2014), because the scan statistic searches for clusters beyond the so-called ‘adjacency matrix’, which is central to the LISA statistic.

Implementation of the algorithms. The LISA and space-time scan statistics are automated and conducted every night when the JHU data are updated and retrieved. The LISA algorithm is repeatedly conducted within the system to detect geographic clusters for each day. An open-source, cross-platform python library of spatial analysis functions, including LISA, called pygeoda, is implemented and integrated for the temporal repetition of the LISA statistic. I conduct LISA on the incidence rate (7-day average cases divided by population) using a Queen contiguity matrix. For the space-time cluster detection, I run SaTScan in a batch mode approach using a discrete Poisson prospective test with a maximum spatial cluster size as 50% of the population at risk and a maximum temporal cluster size of 50 days[7]. Input files, parameter files,

---

[7] The maximum spatial cluster size (50%) is the default setting; the maximum temporal cluster size is based on our experiments with different values (1~59) for this variable using the dataset in October to December 2021 and authors’ experiences as there is no rigor rule for the parameter selection.

and batch files for the analysis in SaTScan were generated using python scripts. Both outputs from LISA and space-time scan statistics are stored in separated databases.

4.2.2 Visualization module

In the *YuTu* system, several visualizations are implemented. One of them is *the animated bivariate map* which displays two variables simultaneously (see Figure 16). This visualization displays results from the space-time scan statistics, using animated bivariate maps to visualize different cluster detection results (Lan et al., 2021). The two presented variables are the relative risk when the location with a cluster and the relative risk for this location on that day. In this interactive system, the values for each variable can be displayed by hovering over the county.

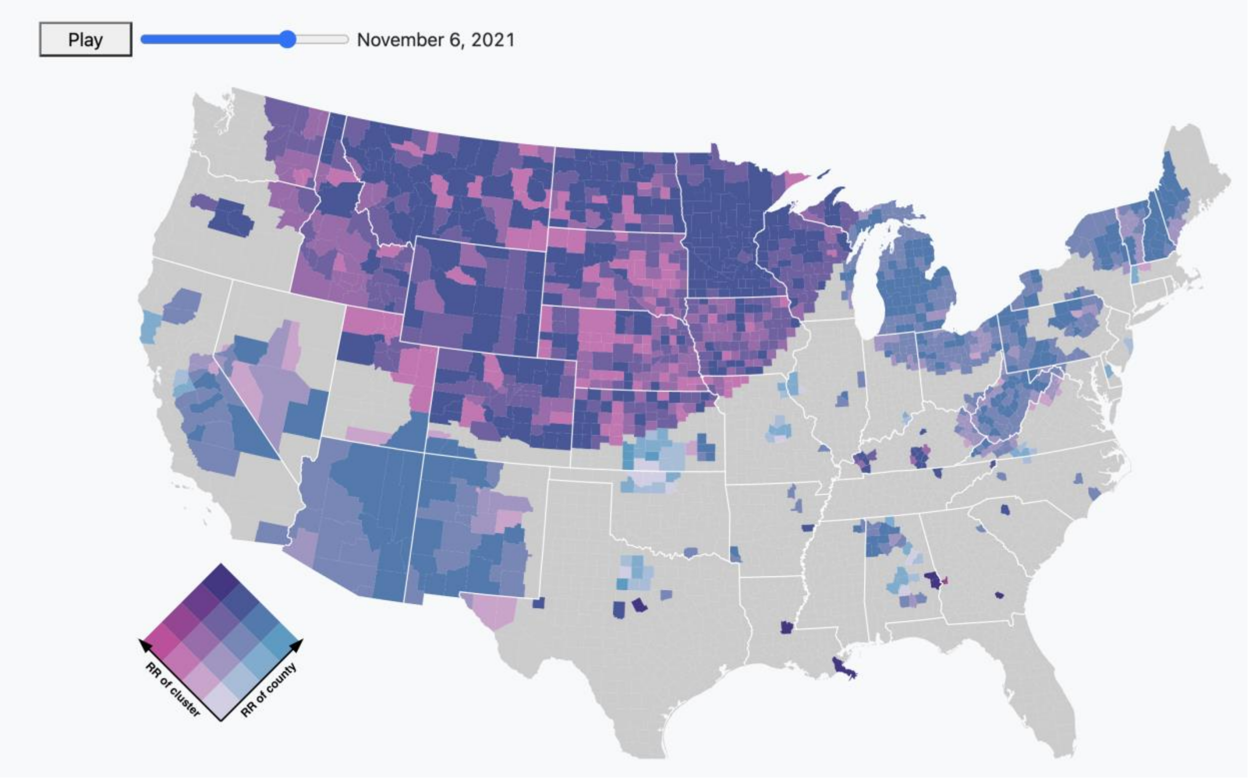


Figure 16. The animated bivariate map of space-time cluster using the space-time scan statistics.

I also complement this system with LISA results (see Figure 17). The two variables for the LISA map are the p-value and the cluster group to which a county belongs. The p-value is ranged from 0.05, 0.01, 0.001, and 0.0001, and the cluster group is ranged from Low-Low, Low-High, High-Low, to High-High. From the two maps, some areas are detected as clusters in both maps, while some regions are detected only on one map.

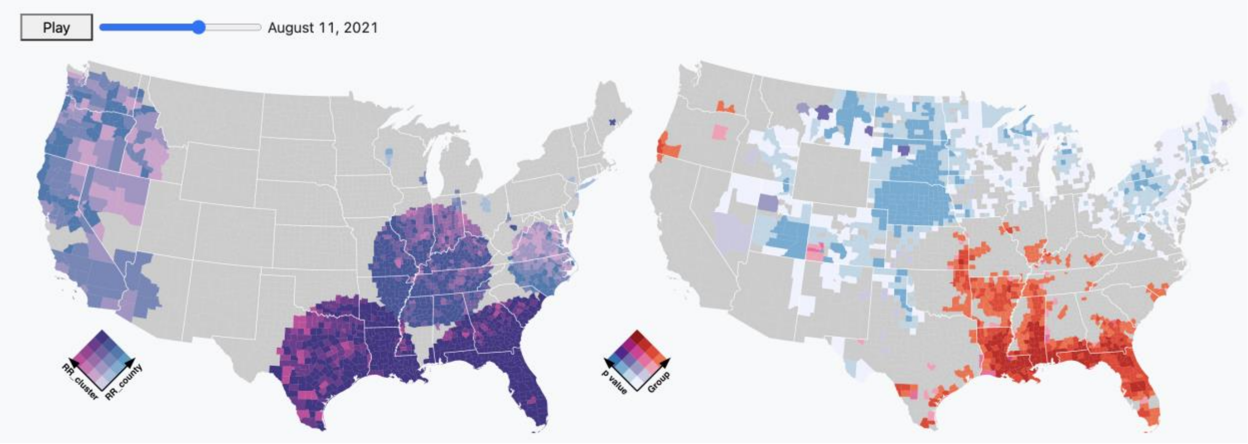


Figure 17. The animated bivariate map of space-time cluster using space-time scan statistic (left) and LISA (right) around August 11, 2022.

Although the animated bivariate map has the advantage of showing the dynamic of cluster distribution each day, it is hard to memorize the overall patterns. To complement that, other visualization solutions are incorporated to display the data in various ways, and these methods include the spiral map, the TimeChart, and the 3D space-time cube.

The spiral map (Weber et al., 2001) shows the average daily relative risk at the state level (see Figure 18). Each bar from *the spiral map* has represented this value by using both color and length. The darker and more extended the bar, the bigger its value. When one county is selected in the bivariate map, *the spiral map* is switched to the spiral of the state that this county is within. The central cartogram displays that value of each state at the latest date.

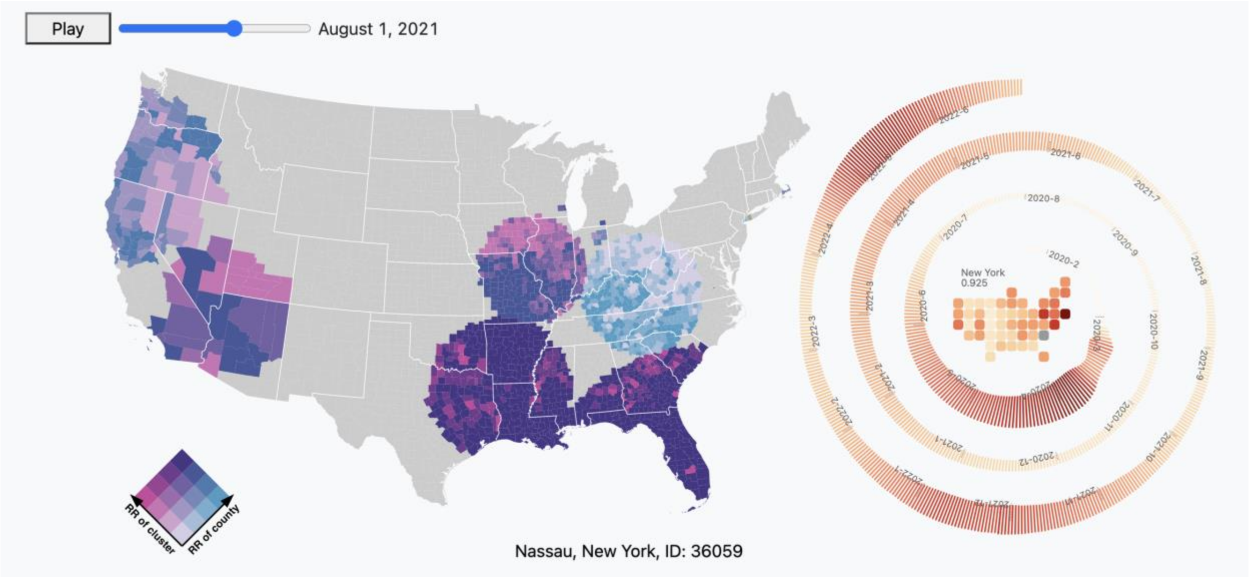


Figure 18. The animated bivariate map of space-time cluster using the space-time scan statistic (left) and a spiral map reflecting the average relative risk for each conterminous US state (right).

I use the *TimeChart* to show the results of the bivariate map in a static and linear fashion (Figure 19). When one or more counties are selected on the bivariate map, the *TimeChart* displays the chart for the selected counties. The first chart in red represents the county's relative risk (RR of the location), while the second chart in blue represents 1) the RR of the cluster that the county belonged to, and 2) the 7-day average cases for this county. In this way, animated results are linked with static and linear results to help discovering the dynamic patterns in space and time.

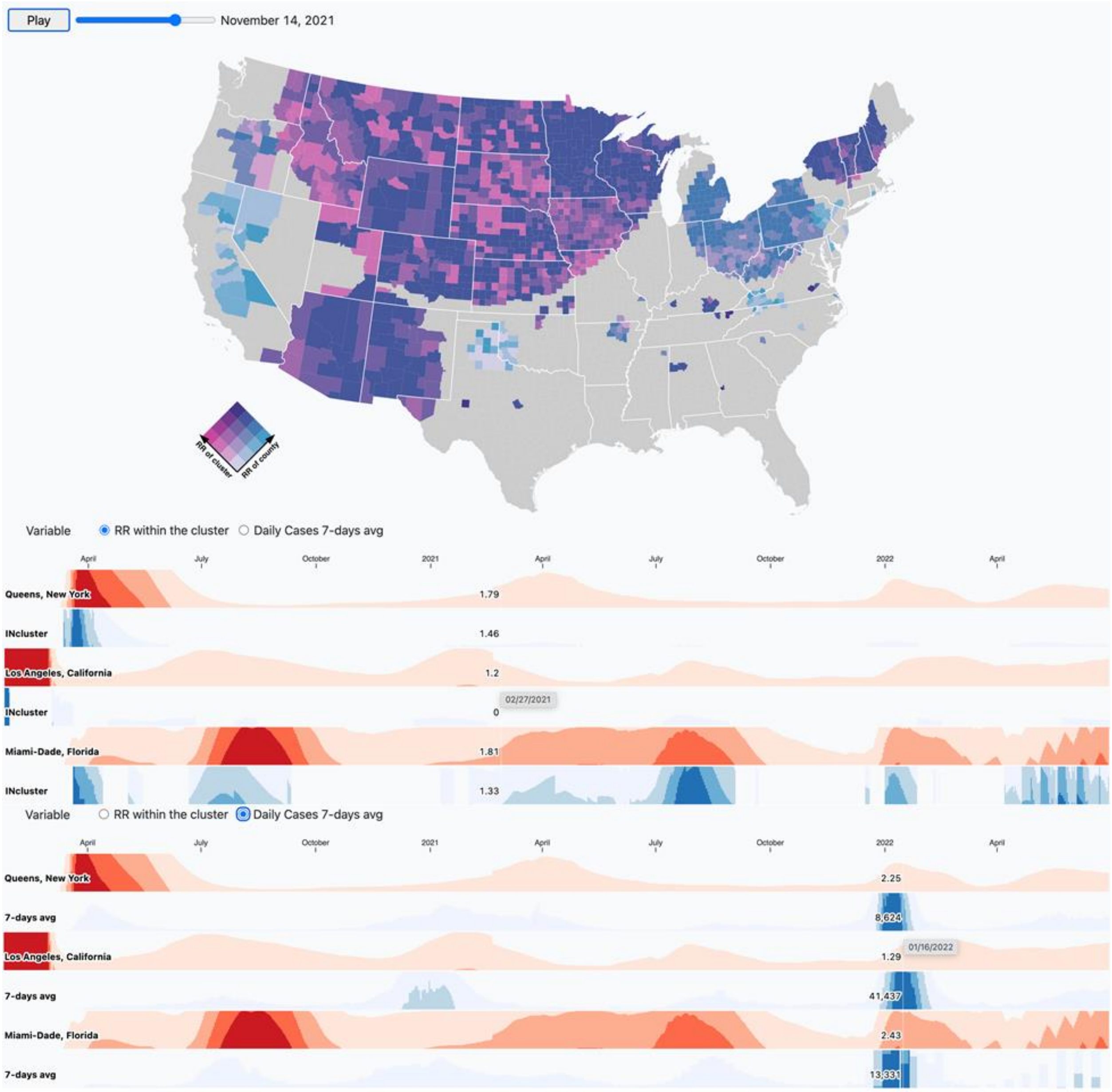


Figure 19. The animated bivariate map of space-time cluster using space-time scan statistic (top) and the *TimeChart* of different variables (bottom).

I also develop 3D web-based geovisualization, named *3D space-time cube* (Figure 20), using multiple JavaScript libraries (3D Scatter Plots Plotly and Data-Driven Documents (D3) (Bostock et al., 2011)). In this 3D plot, the x and y represent the latitude and longitude of the centroid for each county, while the z-axis represents the time. Finally, each dot is color-coded to reflect the value of its relative risk. The system also incorporates a filter that essentially masks to focus on regions is flexible given that a filter of the relative risk is offered to show more or fewer points.

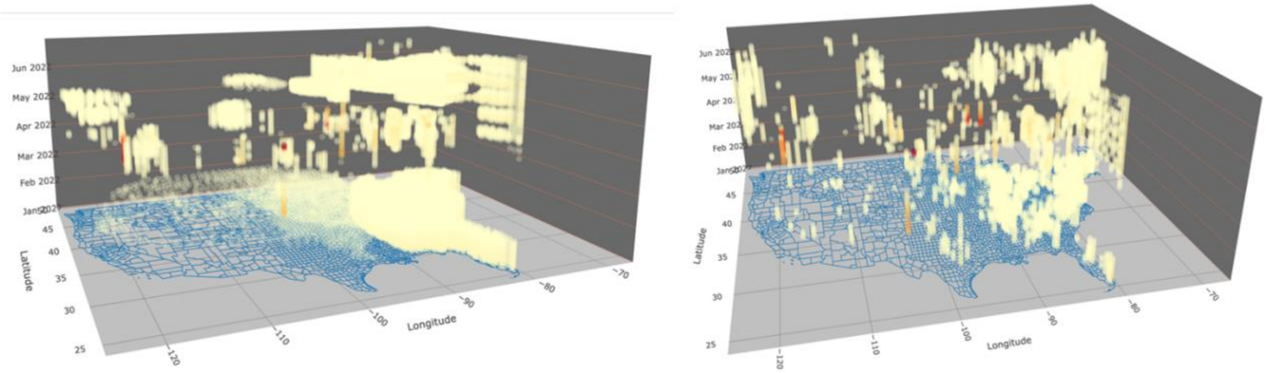


Figure 20. The *3D space-time cube* of clusters with displaying the relative risk of the cluster (left) and relative risk of the county (right).

4.2.3 Data processing module

The data processing module contains daily data retrieving and processing, data analysis, and storing, and these steps are connected to the WebGIS environment. All the data are stored in databases created and managed using PostGIS, an open-source software program that supports geographic objects.

Daily retrieved data are processed and imported into the database on the server. Python scripts are used for daily data retrieving, processing, and space-time cluster detection for all counties. Population and boundaries data are stable in years and are stored as separate databases.

4.3 Case study

I illustrate the *YuTu* system to monitor the variation of COVID-19 cases across the conterminous US. As multiple visual components display different results, I introduce serval case studies as examples to show potential ways to use this system by combining visualizations.

The animated bivariate map is intended to indicate the daily relative risks, which are the basic information for all other visuals.

4.3.1 Four waves of COVID-19 outbreaks

Four waves were identifiable from Figure 21, using the 7-day average cases in the US since the beginning of the pandemic. I selected four time intervals around the peak of each wave, that were July 21, 2020 (peak 1), January 4, 2021 (peak 2), Sep. 3, 2021 (peak 3), and January 15, 2022 (peak 4).

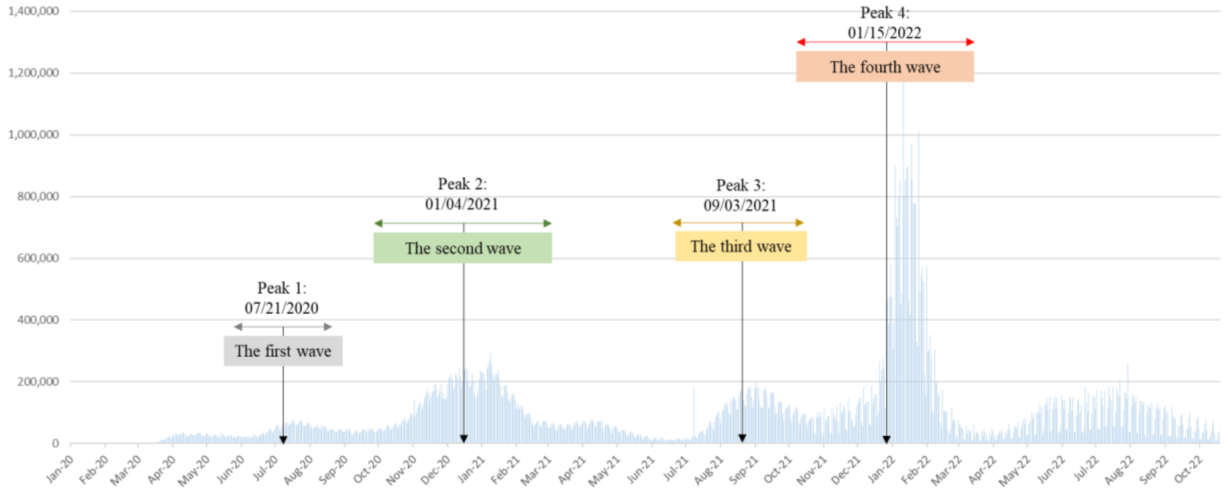
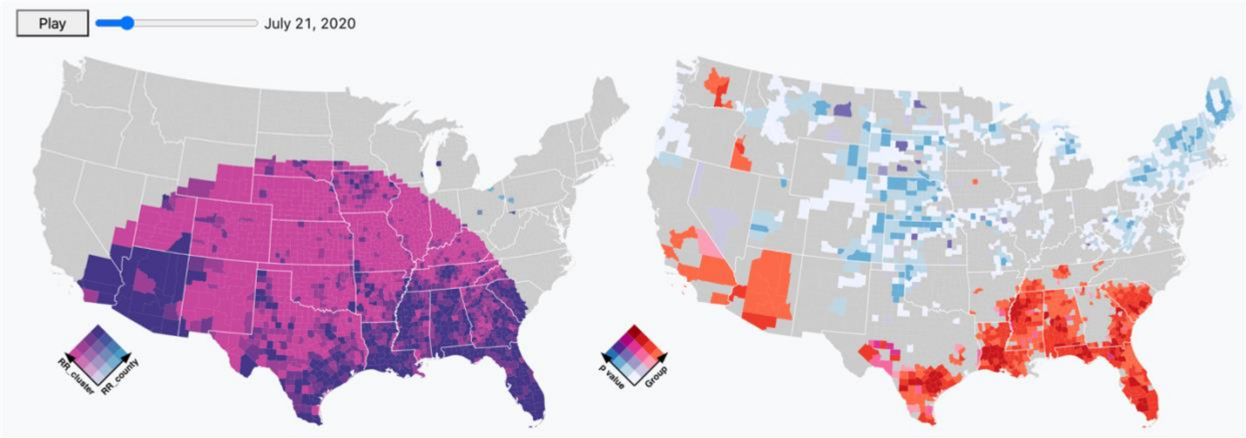


Figure 21. The four waves and their estimated peak dates using the data from WHO Coronavirus (COVID-19) Dashboard (World Health Organization, 2020).

Figure 22 shows the results of the SaTScan algorithm (a) and the LISA statistic (b) for peak 1. From the SaTScan results, one large cluster covered many counties in the south and center of the US. From the LISA results however, high-high and high-low clusters were found in the south, southeast and the southwest, while many counties in the central US belonged to groups of low-low clusters. Also, several counties in Washington state and Idaho were classified as high-high clusters by the LISA method, yet the SaTScan method did not detect these counties.

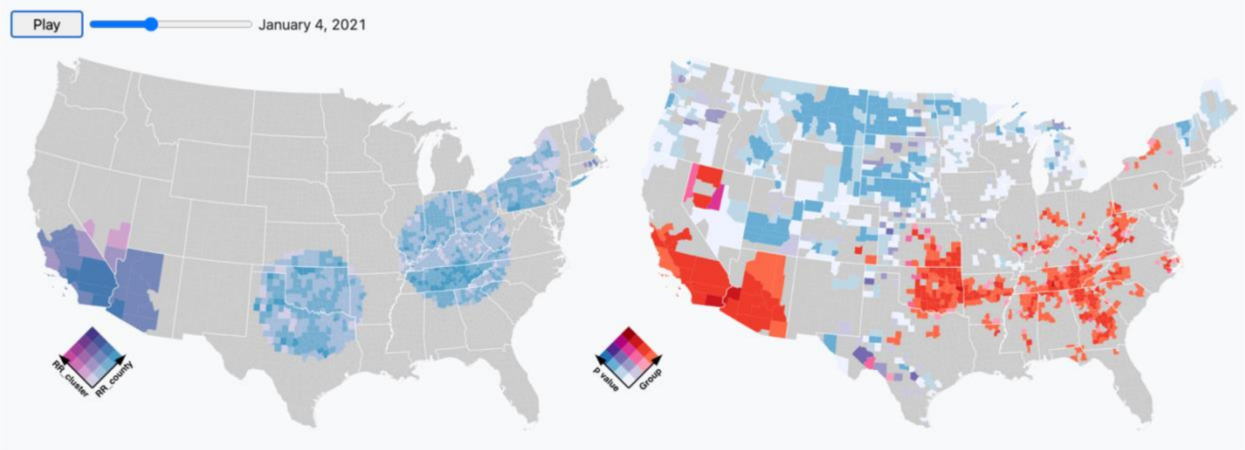


a. Using SaTScan on Jul. 21, 2020.

b. Using LISA on Jal. 21, 2020.

Figure 22. The animated bivariate maps at peak 1 using the prospective space-time scan statistics (a) and LISA (b).

Figure 23 shows the results of SaTScan (a) and LISA (b) on peak 2. However, SaTScan results showed that the RR of the clusters are not as high as during peak 1 (Figure 22). One cluster with higher relative risk was in the southwestern, including the south part of California, the west part of Arizona, and several counties in the boundary of Nevada with these two states. From the LISA results, most counties in the cluster found by SaTScan were categorized as high-high clusters with the LISA algorithm.

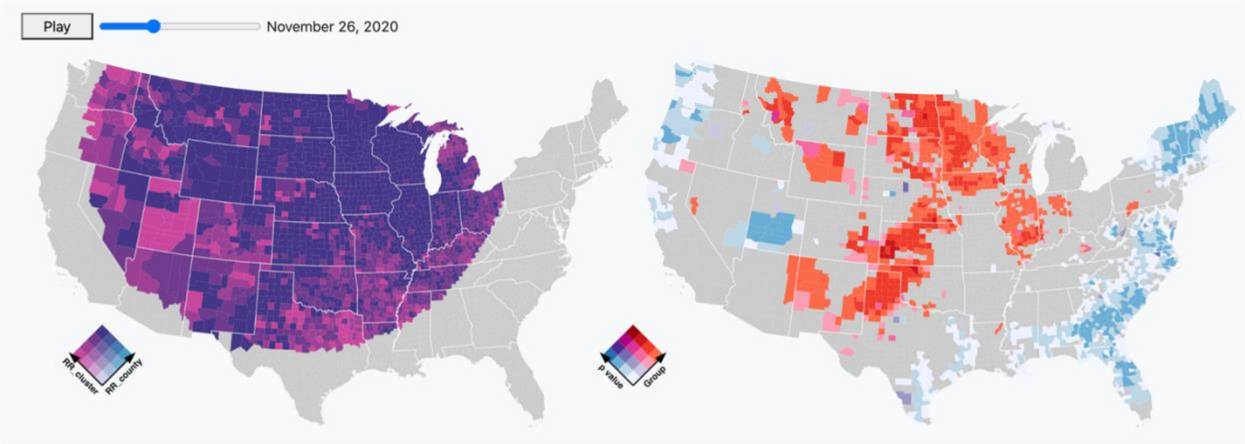


a. Using SaTScan on Jan. 4, 2021.

b. Using LISA on Jan. 4, 2021.

Figure 23. The animated bivariate maps at peak 2 using the prospective space-time scan statistics (a) and LISA (b).

Interestingly, I also found clusters with higher relative risk one month earlier than peak 2. On November 26, 2020, the SaTScan (Figure 24a) detected one very large cluster covering several counties in the North and central parts of the US. Most counties within this cluster had higher relative risks (colored in dark purple) compared with the rest of counties (colored in light purple). As to LISA results (Figure 24b), high-high and high-low were also found in the north and centre of the US.

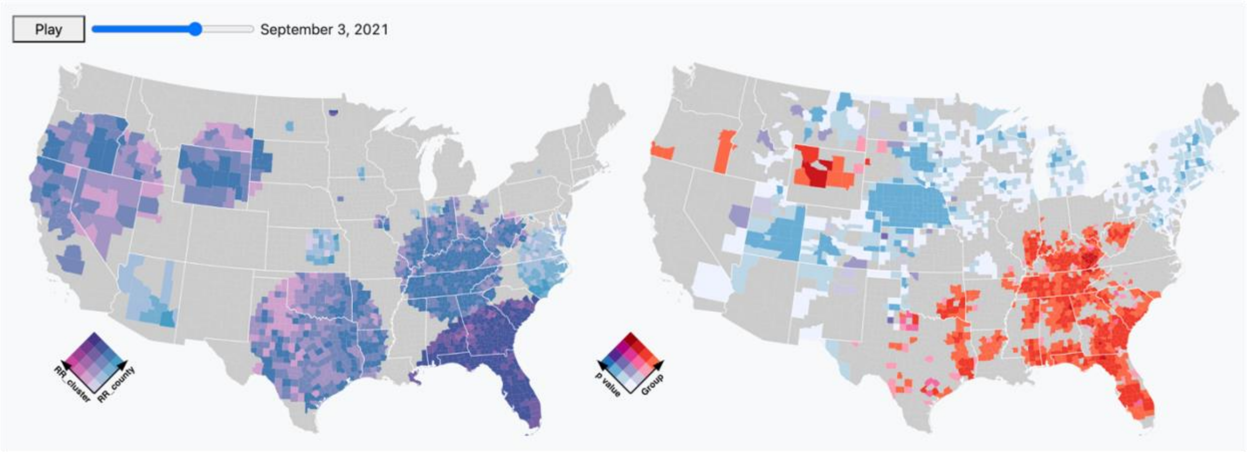


a. Using SaTScan on Nov. 26, 2020.

b. Using LISA on Nov. 26, 2020.

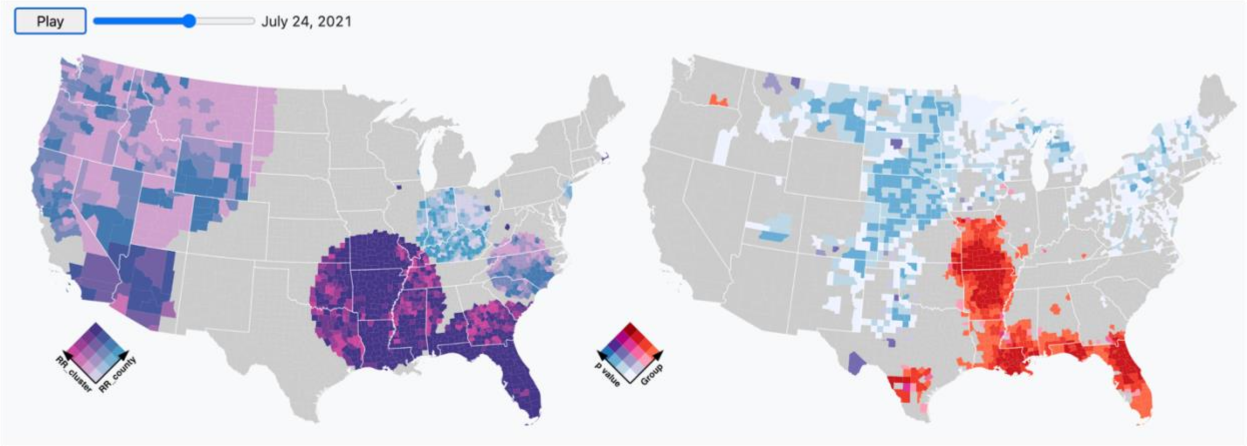
Figure 24. The animated bivariate maps on November 26, 2020, 40 days before peak 2 using the prospective space-time scan statistics (a) and LISA (b).

On peak 3 (Figure 25), multiple small clusters were shown from the SaTScan results (Figure 25a), and the cluster with the highest relative risk on that day included the whole Florida state and many counties from neighboring states. Other clusters were found in the western, central, southern sections of the US. Many counties within clusters from SaTScan were also classified as high-high clusters using LISA (Figure 25b). Similarly, clusters with higher relative risks were detected one month before peak 3 (Figure 25c&d).



a. Using SaTScan on Sep. 3, 2021.

b. Using LISA on Sep. 3, 2021.

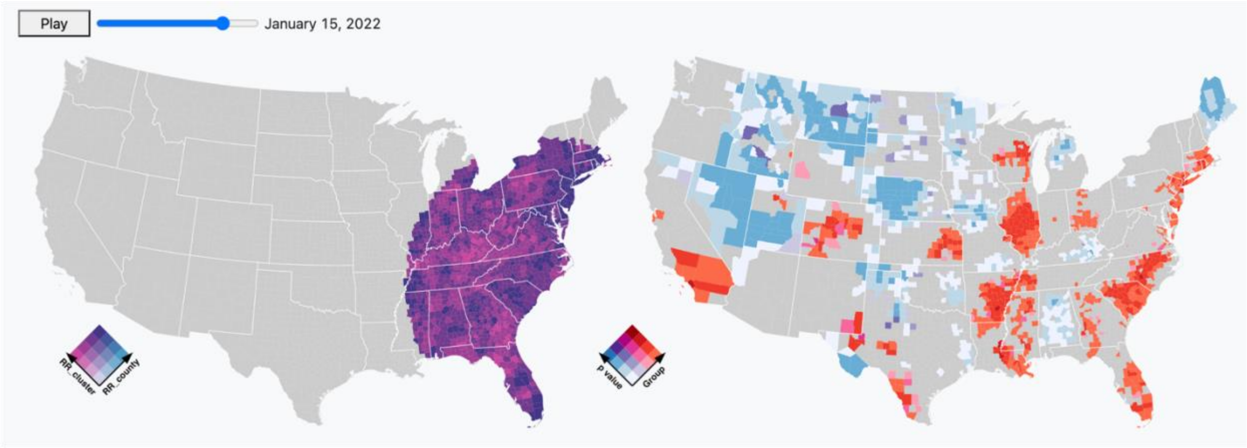


c. Using SaTScan on Jul. 24, 2021.

d. Using LISA on Jul. 24, 2021.

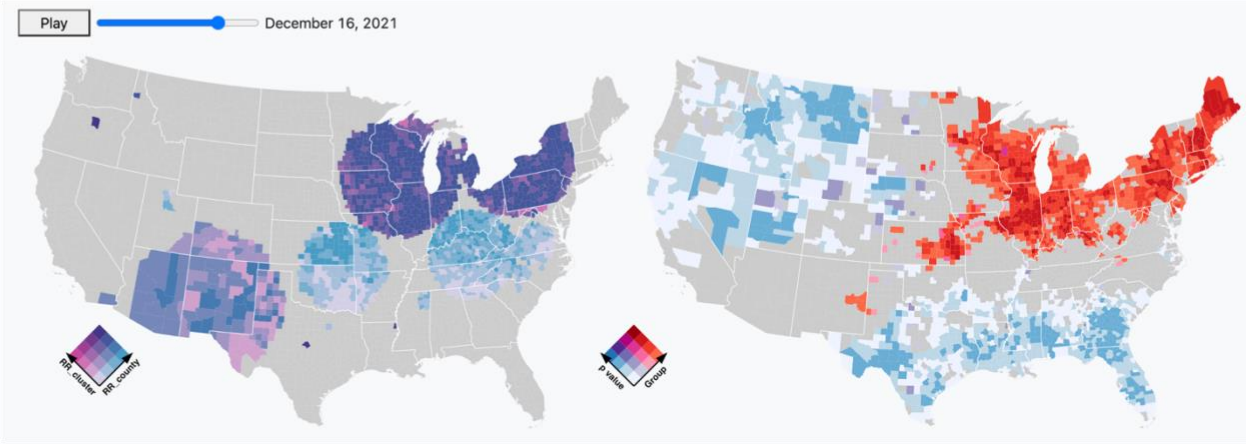
Figure 25. The animated bivariate maps at peak 3 (a & b), and the maps on July 24th, 2021, 42 days before peak 3 (c & d).

During the interval that covered peak 4, only one cluster covering many states in the east of the US was detected by the SaTScan results (Figure 26a), while high-high cluster were distributed across the US according to LISA results (Figure 26b). When looking at the results one month ago of peak 4, both SaTScan and LISA (Figure 26c&d) suggested clusters of higher relative risks or high-high values detected in the northeast of the US.



a. Using SaTScan on Jan. 15, 2022.

b. Using LISA on Jan. 15, 2022.



c. Using SaTScan on Dec. 16, 2021.

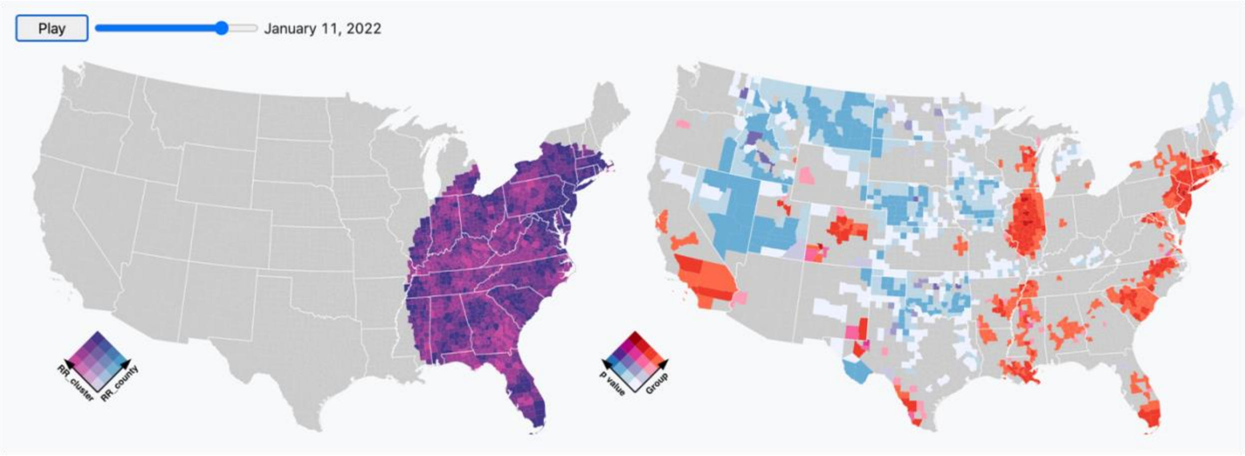
d. Using LISA on Dec. 16, 2021.

Figure 26. Animated bivariate maps at peak 4 (a and b) and the maps on December 16th, 2021, 31 days before peak 4 (c and d).

4.3.2 Comparing situations among counties

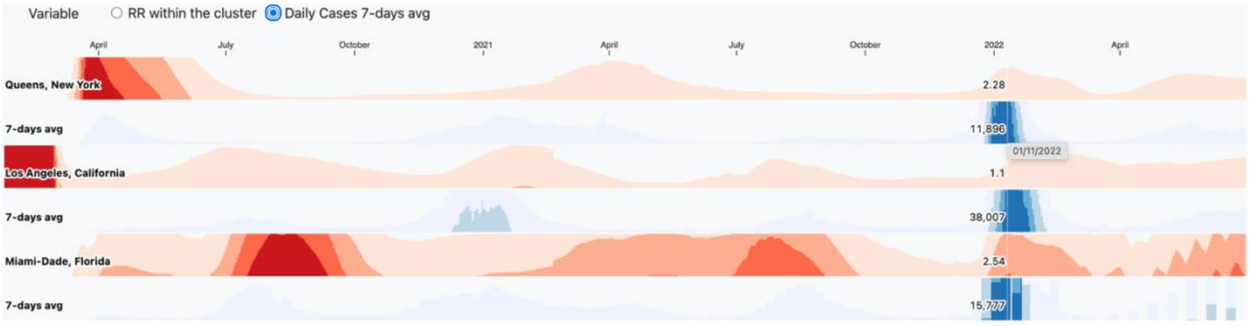
I picked three counties with relatively high population density in three different states: Los Angeles County in California, Miami-Dade County in Florida, and Queens County in New York (see Figure 27). According to the *TimeChart* functionality, all of them had a very high number of cases on January 11, 2022, around the peak day of the fourth wave. Although they had the highest cases on that day, two counties belonged to the same cluster with a relative risk of 3.7, and Los Angeles County was not within a cluster. The relative risks of those counties were

1.1 with 38,007 cases (Los Angeles County), 2.54 with 15,777 cases (Miami-Dade County), and 2.28 with 11,896 cases (Queens County).

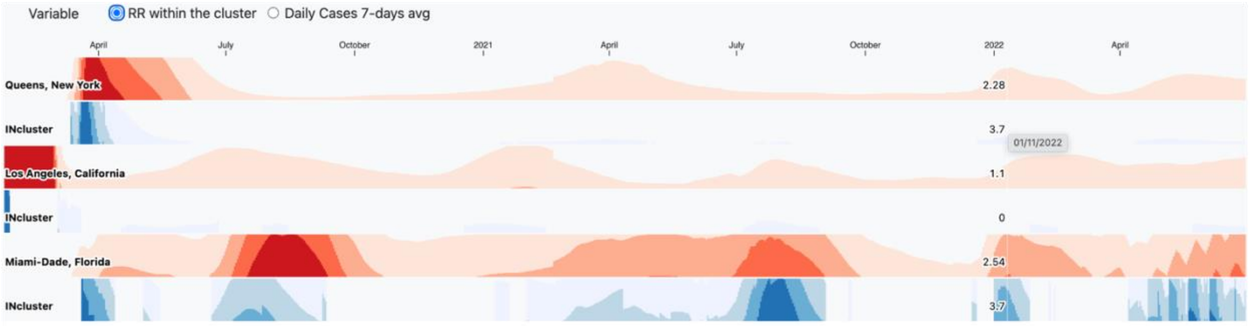


a. Using SaTScan on Jan. 11, 2022.

b. Using LISA on Jan. 11, 2022.



c. TimeCharts of selected counties showing the RR of county (red) and its 7-days avg cases (blue).



d. *TimeCharts* of selected counties showing the RR of county (red) and relative risk of clusters (blue).

Figure 27. The animated bivariate maps of selected three counties on the date that all of them reported most cases (a and b), and the TimeChart of 7-days avg cases (c), and the Timechart of relative risk of clusters (d).

4.3.3 Interpret waves using the 3D space-time cube

I also looked at the third wave from June 2020 to December 2020 using the *3D space-time cube* (Figure 28). The left side of the figure shows the extent of space-time clusters by displaying the relative risk of each cluster, and its right side shows the relative risk distribution of counites with clusters in space and time. The value was filtered to show more or fewer points. When the relative risk is equal to or larger than 2, it is clear that there is a shift from the centre and some counties in the east to the northwest of the US, and this change happened around September, which was around the peak time of the third wave. By increasing the threshold to 4 and 8, the results are clearer, and the relative risk of clusters and counties was higher before September.

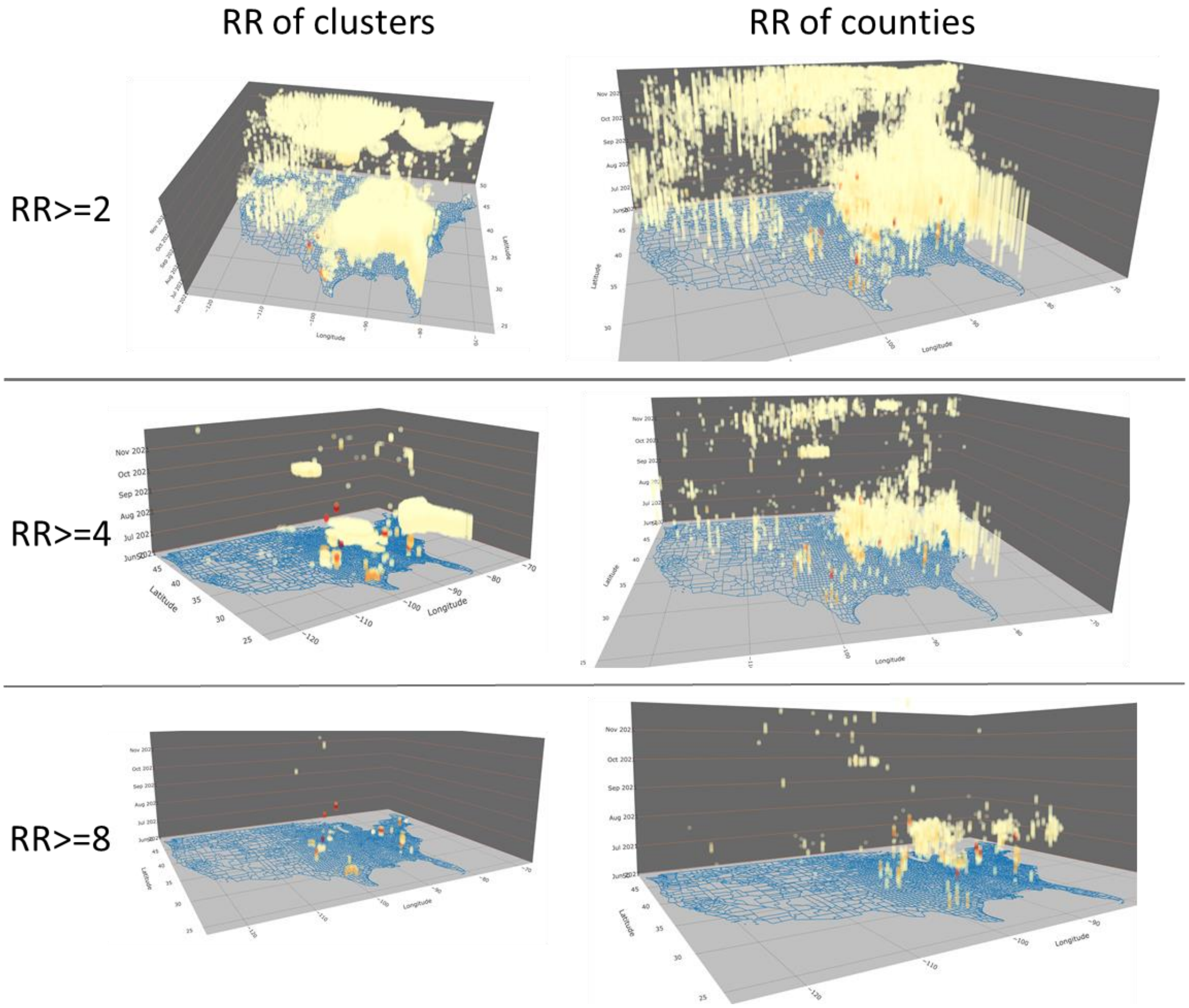


Figure 28. The 3D space-time cubes during the third wave from June 2020 to December 2020 with different threshold of relative risk.

4.3.4 Different scales using Wisconsin as an example

To explore the use of multiple scales, I also generated results of Wisconsin on three scales. According to *the spiral map*, there were two periods when the average relative risk of Wisconsin was higher (Figure 29). I selected three months from one of that two period, from December 1, 2021, to February 28, 2022, when the value peaked and then decreased.

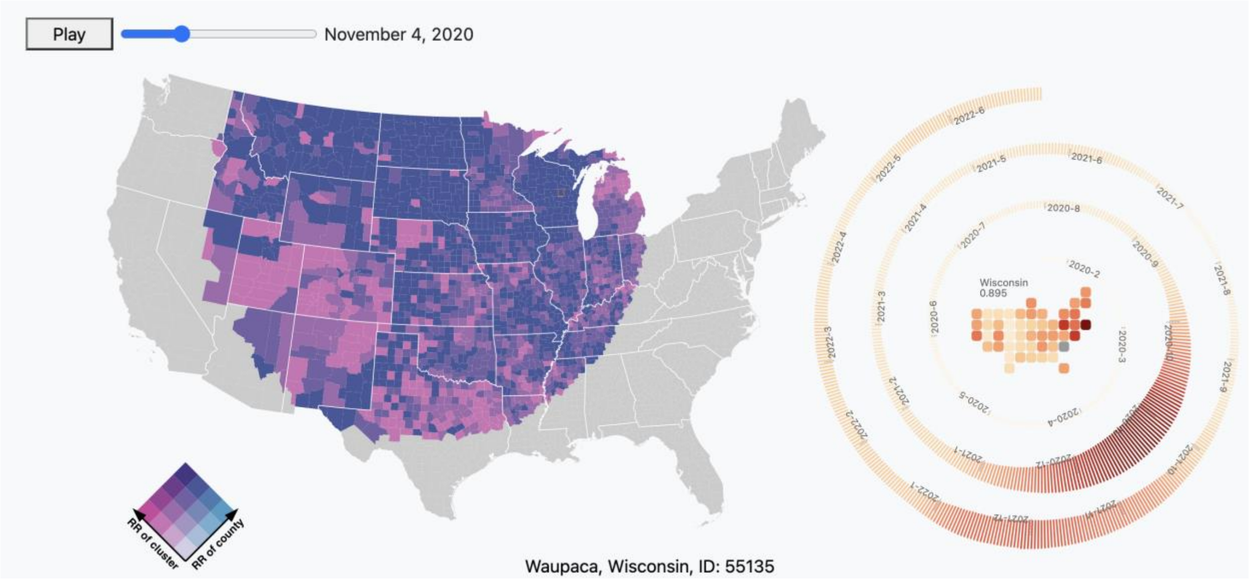
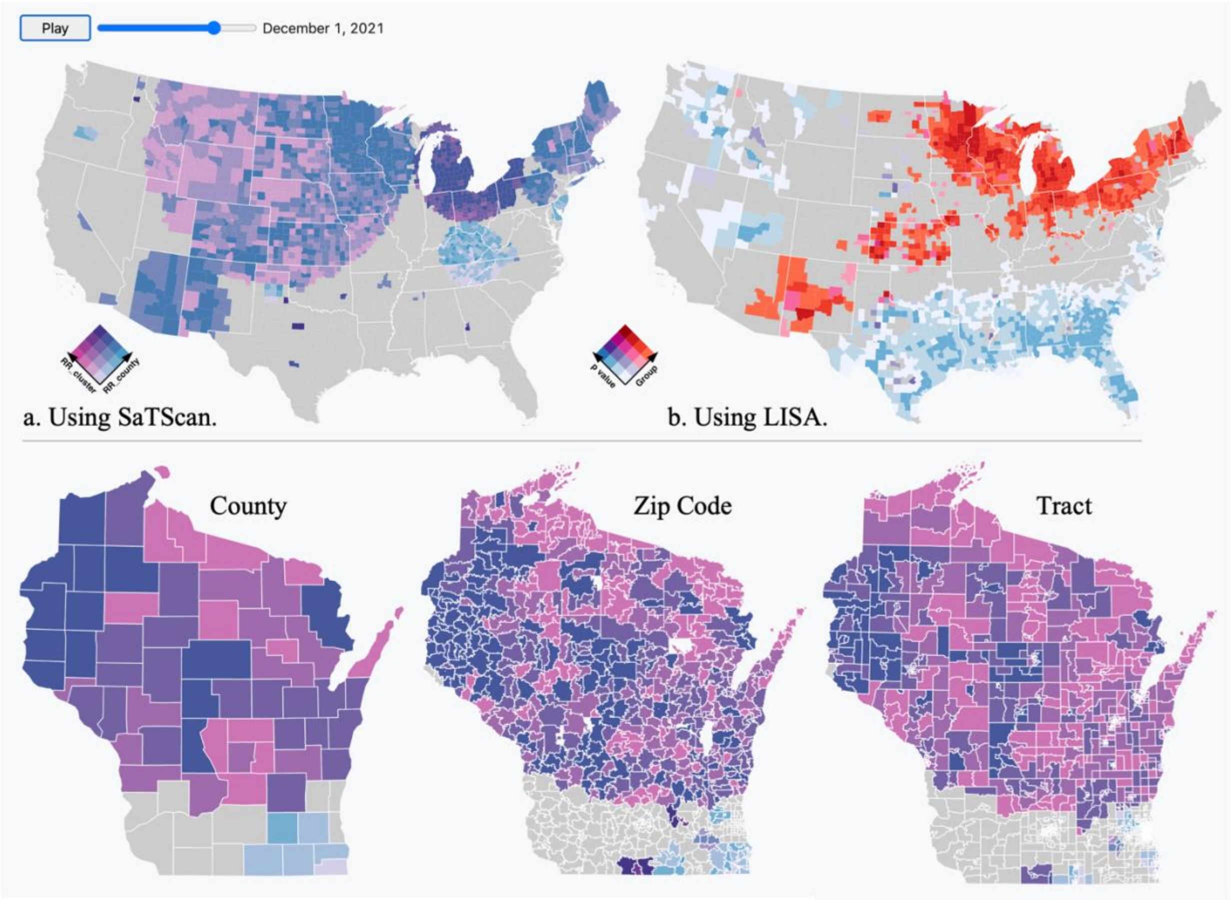


Figure 29. The animated bivariate map and the spiral map of Wisconsin.

To compare results at different scales, I included scales from the county levels with all other states, the county level with one state, the zip code level, and the census tract level (see Figure 30, Figure 31, Figure 32). From the cluster map using results of SaTScan and LISA for the entire region, most counties in Wisconsin were detected in the clusters on December 1, 2021. For SaTScan results only considered cases within the state, all three levels detected clusters in most areas except areas at the bottom, and results at three scales were similar. On January 5, 2022, most counties in Wisconsin no longer belong to a cluster from the results of SaTScan and LISA considering all states. However, clusters were detected on all three scales when only considering regions in Wisconsin, and the area that belonged to clusters was smallest at the county level and largest at the zip code level. On February 28, 2022, most counties in Wisconsin remained not within a cluster from the results of SaTScan and LISA. Clusters were only detected at the zip code level.

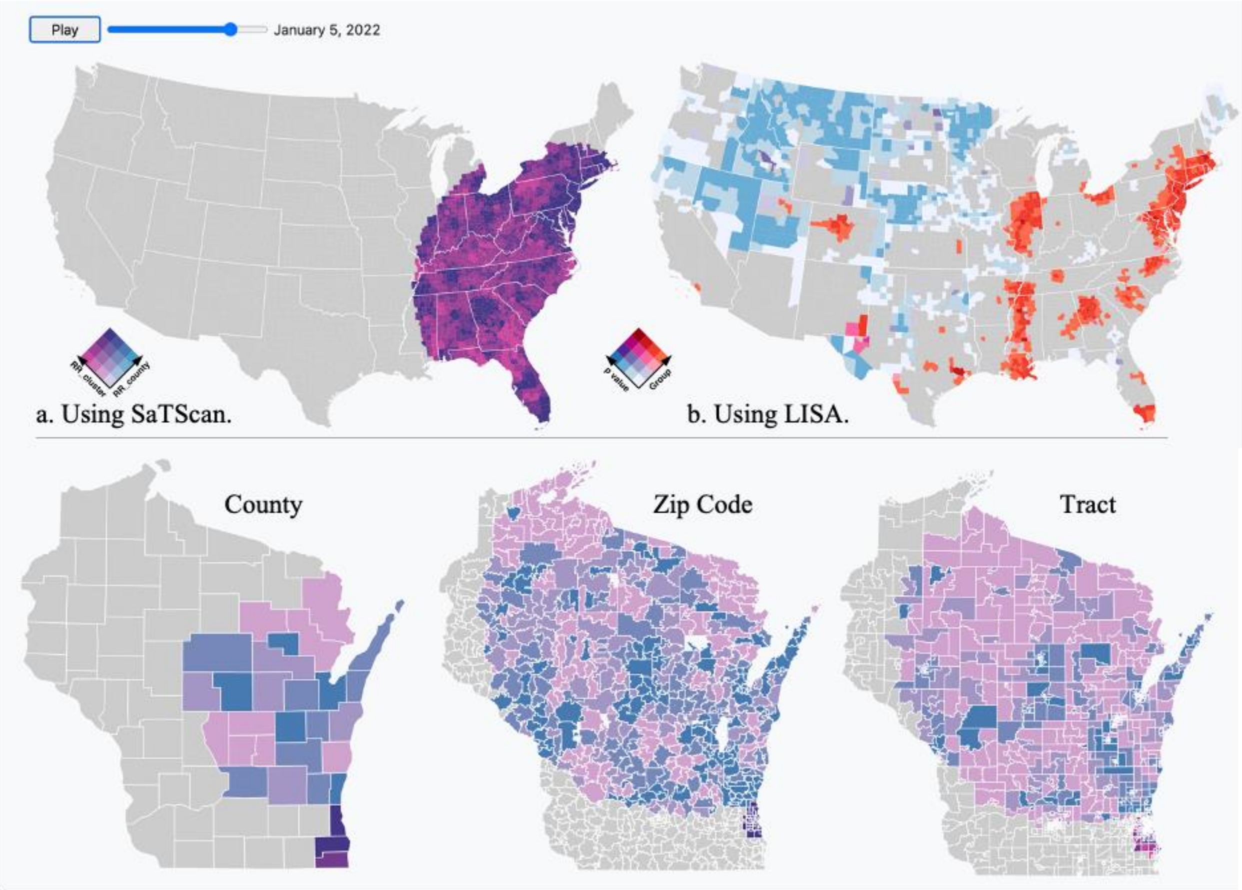


a. Using SaTScan.

b. Using LISA.

c. Using SaTScan at multiple levels.

Figure 30. The animated bivariate map of Wisconsin at multiple levels on December 1, 2021.

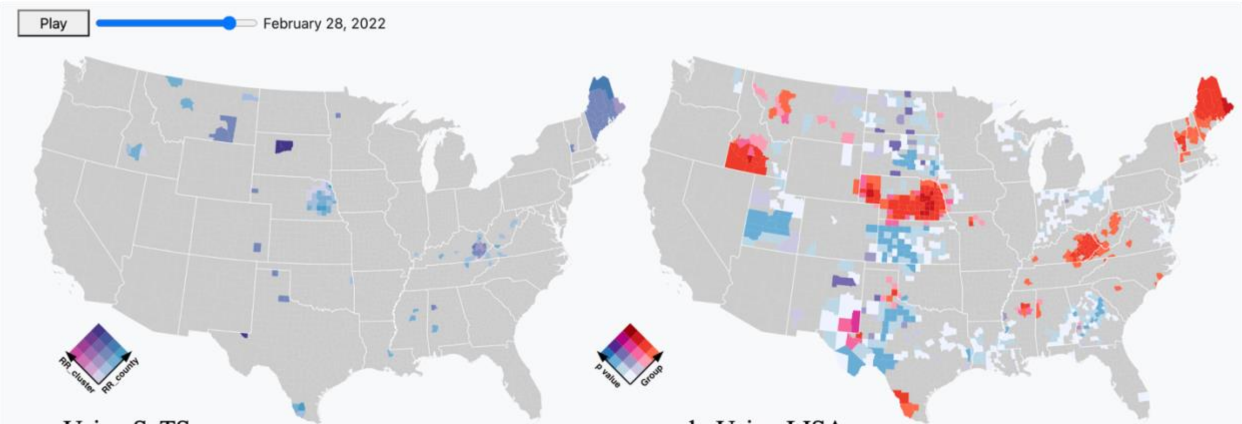


a. Using SaTScan.

b. Using LISA.

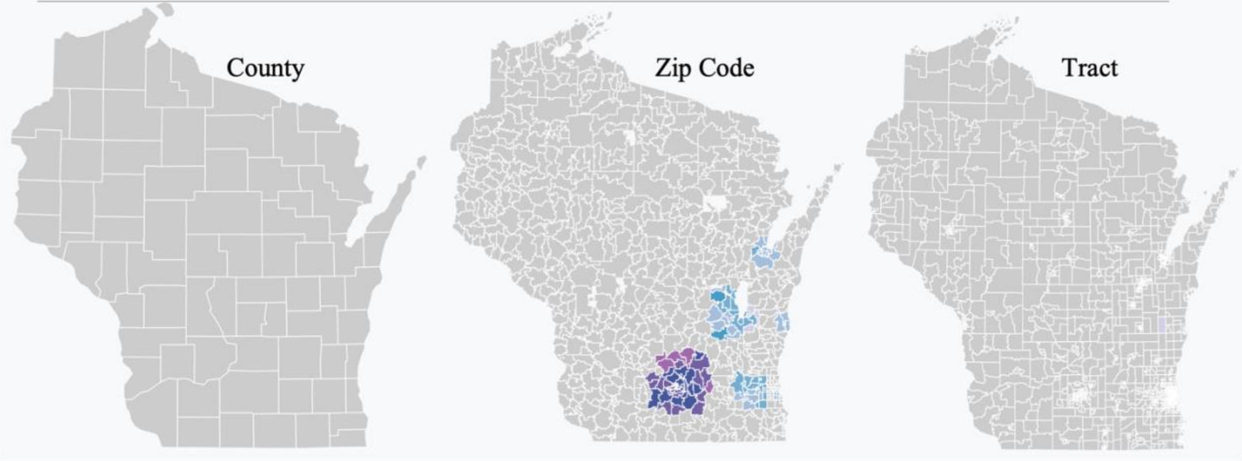
c. Using SaTScan at multiple levels.

Figure 31. The animated bivariate map of Wisconsin at multiple levels on January 5, 2022.



a. Using SaTScan.

b. Using LISA.



c. Using SaTScan at multiple levels.

Figure 32. The animated bivariate map of Wisconsin at multiple levels on February 28, 2021.

4.4 Discussion and conclusion

US COVID-19 YuTu is a health surveillance system based on space-time cluster detection analysis and visualization implemented with near real-time monitoring and novel visual features. This system emphasizes spatiotemporal analysis and representation in various ways. To illustrate the framework, the prospective space-time scan statics and the LISA algorithm are applied to detect space-time clusters daily, although other ESDA or statics methods can also be implemented. As for the visualization, *the animated bivariate map* of those two methods presents all the available space-time clusters each day, together with the *TimeChart* for linear representation and *the spiral map* for the higher level summary to complement the weakness of processing a large set of dynamic information. The *3D space-time cube* offers another relative static way to explore the same results with the ability to filter data on both the range of time and the value of the relative risk of space-time clusters. This framework can also apply multiple scales, such as from the county to the zip code level.

This system also has some limitations. First, the calibration of the parameters used in the spatial analysis requires fine tuning with epidemiologists, which may initiate some hypothesis for prediction, such as recognizing the signal before the peak of a current wave. From the introduced case study that focused on four specific waves, I found that two peaks in 2021 have higher relative risk and larger extent of clusters according to results using the prospective space-time scan statistics. Likely, this signal was detected because of the choice of time maximum range as 50 days based on the 2021 dataset. It is worth noting that the LISA result cannot detect this signal, as it only considers data for one day.

Secondly, because of the limited source of available open access datasets, this system only takes cases and deaths into account, while other COVID-19 related information, such as hospitalization rate, could also be useful to detect space-time patterns.

Thirdly, I did not consider the uncertainty in population as I use the 2019 ACS data. In the early of the system development, the 2019 ACS data was the best source in early 2021. Although the US Census Bureau released the 2020 Census later, the population after 2020 remains unknown and it is heavily impacted by the pandemic. The solution should not be simple update the population data when the latest is available. More investigations on this issue are necessary to near real-time health surveillance systems like the *US COVID-19 YuTu*.

This study presents a framework of novel surveillance systems that automatically detects space-time clusters daily, and a prototype using the framework, named *US COVID-19 YuTu*, is introduced. This COVID-19 surveillance system integrates two popular methods and a various of novel interactive visualization features in 2D and 3D within an open-source web-based GIS environment. Although the prototype focused on 48 states in the US using the space-time scan statistics and the LISA algorithm, it can apply to any other countries or regions at multiple scales using other possible algorithms. On the one hand, I hope this system can assist health policymakers in making interventions, such as slowing down the spread of COVID-19 in the US. On the other hand, I hope to inspire others to develop health surveillance systems that reflect more on the spatiotemporal analysis results instead of merely descriptive data.

CHAPTER 5: EVALUATION

When developing new visualization and analytics tools, it is important to evaluate their effectiveness (Carroll et al., 2014). The motivation for this chapter is to assess the functionality and usability of *US COVID-19 YuTu*, the proposed online surveillance. In a 2007 study, Robinson (2007) presented a design framework for exploratory geovisualization in epidemiology. The research questions in this study focused on identifying necessary features and interactions of visualizations and how epidemiologists may use the application for decision making. Likewise, my proposed online surveillance system in Chapter 4 is designed to assist epidemiologists to detect space-time patterns to assist in a response during an outbreak.

In this chapter, I summarize efforts to evaluate the proposed online surveillance system based on three specific research questions:

1. Is this a helpful system for detecting and identifying space-time patterns of infectious diseases? If so, which features of visualizations and represented information are helpful and why?
2. Does this system assist epidemiologists in making timely responses during an outbreak? If so, how could they use results or information from the system?
3. What are the strengths and weaknesses of this system compared with traditional methods epidemiologists used?

To answerer these questions, I conducted an evaluation consisting of a user study and an export evaluation. The remainder of this section presents the background information of selected qualitative methods to evaluate this system. I introduce the design and process for the evaluation in the second section and the results in the third. Lastly, I provide a summary and discussion of the evaluation.

5.1 Background

Popular methods to evaluate the effectiveness of novel visualization techniques in geospatial health are interviews, participant observations, task analysis, and a combination of these methods (Carroll et al., 2014). These qualitative methods are beneficial for collecting feedback from targeted users and connecting practice with theory. For example, Ban and Ahlqvist (2010) evaluated a geovisualization of uncertain urban ontologies, including two types of interviews, a pilot interview with students, and in-depth discussion with experts, examining different aspects of the geovisualization. They concluded that the pilot interview with closed questions was an excellent resource for collecting qualitative data and that the in-depth interview offered new information that was not indicated in the pilot interview.

Participant observation is one popular method that has been implemented in the design process of analytical and geovisualization systems. For example, Robinson et al. (2005) invited domain experts to assess an exploratory geovisualization toolkit for epidemiology, and Lloyd and Dykes (2011) collaborated with experts during the design phase of geographic visuals lasted for a the long term (3 years).

Task analysis is another widely used method for evaluating a system or application in epidemiology, and it usually combines a survey after completing an analytical task. For example, Robinson et al. (2017) evaluated a visual analytics system to detect space-time patterns using a combination of task analysis, followed by participant surveys. Their task analysis required participants to provide written responses of the patterns they identified, followed by a usability and utility survey. Similarly, Pezanowski et al. (2018) developed a user evaluation approach in which participants were required to answer questions for the evaluation after finishing a task that guided them through primary functions of an analytical application of social media. They created an online survey consisting of basic demographic questions, sample tasks to complete, and usability and utility questions. In another study, Wagner Filho et al. (2019) evaluated an immersive space-time cube geovisualization for trajectory data exploration. They introduced evaluation strategies, including measuring users' performance in relevant tasks, observing how users interact, interviewing users on their subjective experience, including both novice and domain experts.

Besides selecting qualitative methods, another vital component for conducting qualitative research is to decide on target respondents and recruitment. For instance, Anderson and Robinson (2021) recruited 320 participants using Amazon Mechanical Turk[8] to collect information on reading categorical maps, and participants were recruited without constraints of knowledge or skills. However, studies designed for professionals may require participants to demonstrate specific domain knowledge, and it may be difficult to recruit sufficient participants. For this reason, researchers generally include novice users (typically students) and domain experts (see Robinson, 2007; Robinson et al., 2017; Wagner Filho et al., 2019).

5.2 Evaluation design and process

The evaluation of *US COVID-19 YuTu* consists of a user study (task analysis integrated into a survey) and an expert evaluation (a group interview) to capture different dimensions of the same issue as the triangulation (Patton, 1999). The user study aimed at testing the functionality and general thoughts by collecting responses from graduate students using an online survey, and

---

[8] Amazon Mechanical Turk is a crowdsourcing website to remotely recruit participants to conduct tasks.

the expert evaluation further collected feedback, comments, and discussion from a group interview with experts in COVID-19 analysis. The evaluation procedure (a user study and a group interview) was approved by the University of North Carolina at Charlotte's Office of Research Protections and Integrity (Case No. IRB-22-0219).

5.2.1 User study: Survey

Survey design

The survey instrument included questions about the participant's basic background information and five groups of tasks and questions for each visualization from the prototype. The five groups are 1) *the single bivariate* map on the home page, 2) *the two bivariate* maps combining the space-time scan statistics and the LISA algorithm, 3) *the spiral map* showing results at the state level, 4) the *TimeChart* summarizing trends in a time series, and 5) *the 3D space-time cub*e representing results in a 3D dimension (see Chapter 4 for details of each visualization). To confirm that participants understood basic functionality and visualization of the evaluated system, they were required to correctly answer three questions before proceeding to the next section. Tasks and questions for each visualization can be classified into five categories: 1) the accuracy of submitted answers, 2) the confidence level of answers related to the accuracy, 3) the usefulness of this visualization, 4) the possible use of this visualization for decision making, 5) and an open-ended question to collect additional feedback.

For questions related to accuracy, participants were asked to directly ascertain values or answers to the question from the current evaluated visualization. The questions asked were similar, such as "What was the time range when more than half of the counties in California belonged to a cluster?" and "Which state had the highest average relative risk yesterday from the center of the spiral map?" Some questions in this category used a 5-point Likert scale, which asks responders to identify their level of agreement to a statement on a scale of five points: 1) strongly disagree, 2) somewhat disagree, 3) neither agree nor disagree, 4) somewhat agree, and 5) strongly agree. In this survey, all questions with the 5-point Likert scale used this scale of five points. The task asked participants to select their agreement level with a statement that described the pattern shown from the visualization. An example statement for this category is “The average relative risk in Washington state was very high at the beginning of the pandemic.”

In an effort to further interpret how confident participants were that they understood the visualization, the participants were then asked to identify their confidence level. This information is helpful to interpret how clear and easy that participants can understand the represented information from the system.

For questions that evaluated the usefulness of the visualization, participants needed to select an agreement for the following statements using the 5-point Likert scale. Questions included “I think both maps provided under the ‘@LISA’ tab allow for a comparison of clustering techniques” or “I think the ‘TimeChart’ is useful to identify the trend of different variables (e.g., relative risk).”

To evaluate and explore potential indications from the represented information of the visualization in decision making, participants were asked to imagine themselves as health policymakers and submit their agreement level on prompted statements, such as “I will rely on the information provided on both maps to make an informed decision” and “I will use the filter function from 3D space-time cube to warn the counties with a high relative risk (e.g., RR > = 4).”

At the end of each visualization evaluation, participants had the opportunity to provide overall feedback. Any suggestions and comments for the prototype, in general, were also gathered. The survey was created and distributed using Qualtrics XM, a sophisticated online survey tool. Before the survey recruitment, we also conducted a pilot test with three participants to ensure that the survey process was smooth and that no problems occurred. The participants include a post-doc in public health and two graduate students in geography.

Survey Process

For the evaluation process, each participant was required to participate in the introduction section and then complete an online survey. Twenty-four graduate students were recruited from the Department of Public Health Sciences and the Department of Geography and Earth Science at UNCC to evaluate the prototype's functionality.

As this prototype is designed to assist health experts in decision-making, the majority of participants were from the Department of Public Health Sciences (n = 18). Multiple recruitment approaches were employed to maximize the total number of participants with a background in public health. These approaches included sending recruitment information to professors in epidemiology and all graduate students in the department and hanging flyers within the department building. Each participant met with me for the introduction section. Only one participant requested to meet in person, and the others were scheduled for a Zoom meeting. The meeting usually took 15-30 minutes to introduce basic concepts and tools on the prototype and answer any questions the participants had. The survey link was sent to the participant after the meeting. We also recruited graduate students from the Department of Geography and Earth Science (n = 6).[9]

---

[9] Three participants were interested and participated in the evaluation, after I presented the prototype and introduction in front of a GIS class in the department; another three participants in geography were recruited by email, and I introduced basic concepts and tools to each of them individually.

A Qualtrics survey link was sent to the participant's email account once they obtained the basic knowledge to evaluate the prototype from an individual meeting or the presentation to the whole class. All of the surveys were taken using the participants' computers, although my personal computer was also available if needed. Only one participant reported issues accessing specific visualization, which was solved using another device. The prototype was accessible using popular browsers, such as Google Chrome, Apple's Safari, Microsoft Edge, and Firefox, without issues.

5.2.2 Expert evaluation: A group interview

Design

The expert evaluation was designed to evaluate the general prototype by experts with experience using COVID-19 data and dashboards. The group interview consisted of an introduction, group tasks, a general discussion, and a conclusion. For the introduction, all attendees, including participants and facilitators, presented relevant background information about themselves and their experiences related to COVID-19 data and dashboards. This introduction was designed to help participants learn about each other's backgrounds and encourage discussion. Three group tasks evaluated two COVID-19 dashboards or systems and my prototype by collecting feedback and making comparisons. The two COVID-19 dashboards presented were the CDC's COVID Data Tracker and the US Covid Atlas. Health officials widely use the former as it updates COVID-19 and its relevant data in accordance with the CDC. The US Covid Atlas is a near-real-time visualization tool to connect case data and community indicators from the pandemic's onset to the present. Furthermore, it incorporates maps using the LISA algorithm. Thus, the US Covid Atlas uses similar methods as my prototype. Each group task individually introduced one of the three tools (COVID Data Tracker, US Covid Atlas, US COVID-19 YuTu), and similar questions for each tool were asked during the section. The following questions were listed as examples and were adapted during the meeting based on the discussion:

1. What do you think about the effectiveness of this system for public health surveillance?
2. What do you think about the usefulness of the system's cluster detection (or could it be useful if the system does not have this function)?
3. Which information from this system would be particularly interesting to you?

The goal of the general discussion was to prompt a conversation about all three COVID-19 dashboards or tools and compare them. From the discussion, my objective was to determine which information was useful or which visualizations were useful. After the general discussion section, the group interview meeting was concluded, and participants were encouraged to provide additional comments.

Process

The group interview was held in an online Zoom format for approximately two hours, and four experts with different backgrounds in COVID-19 analysis were recruited. These four experts include one professor from the Department of Public Health Sciences at UNCC, one professor from the Department of Geography at The University of Hong Kong, and two epidemiologists from the Mecklenburg County Government in North Carolina. All of them have direct experience in COVID-19 data and analysis. Two committee members of this dissertation (Drs. Eric Delmelle and Deborah Thomas) also joined the interview with me as facilitators and notetakers. The two committee members took turns leading each section, and I was responsible for presenting the demonstration of each dashboard or tool during three group tasks. All three of us took notes during the entire meeting, which was recorded for further analysis with the participants' agreement.

5.3 Results

5.3.1 User study

Twenty-four surveys were collected and used for analysis; 75% of the participants (n = 18) were female, most of the participants (n = 21) were between the ages of 18 and 33, and 62.5% of the participants (n = 15) had a major in Public Health Science. Regarding the degree level, the number of master's students (54.17%, n = 13) was slightly higher than doctoral students. Most participants (except two) considered themselves knowledgeable in epidemiology and understood relative risk in epidemiology. Additionally, over 70% (n = 19) felt confident in interpreting clusters, while less than 60% (n = 14) felt confident in interpreting space-time clusters. Regarding interpreting results using the LISA algorithm, only approximately 45% of the participants (n = 11) considered themselves knowledgeable, although they received basic training during the introduction meeting.

In the following section, responses from the survey of the user study are demonstrated in detail.

The single bivariate map: Homepage

In the questions related to the accuracy of the single bivariate map, the results indicate that most participants understood the bivariate map by identifying the values of two variables. Seventy-five percent (n = 18) of the participants recognized the patterns from the bivariate map, and 91.67% (n = 22) identified the time range when more than half of the counties belonged to a cluster. From Figure 33, over 91% (n = 22) felt confident in their answers to the abovementioned accuracy questions. Regarding the usefulness, 83.43% of the participants (n = 20) agreed the bivariate map can help them interpret the different relative risks among counties within the same cluster (usefulness statement #1), and 91.67% of the participants (n = 22) agreed that the bivariate map is useful (usefulness statement #2). In terms of decision making, 75% of the participants (n = 18) stated that they would use the results from the bivariate map to allocate more health resources to areas with high relative risk.

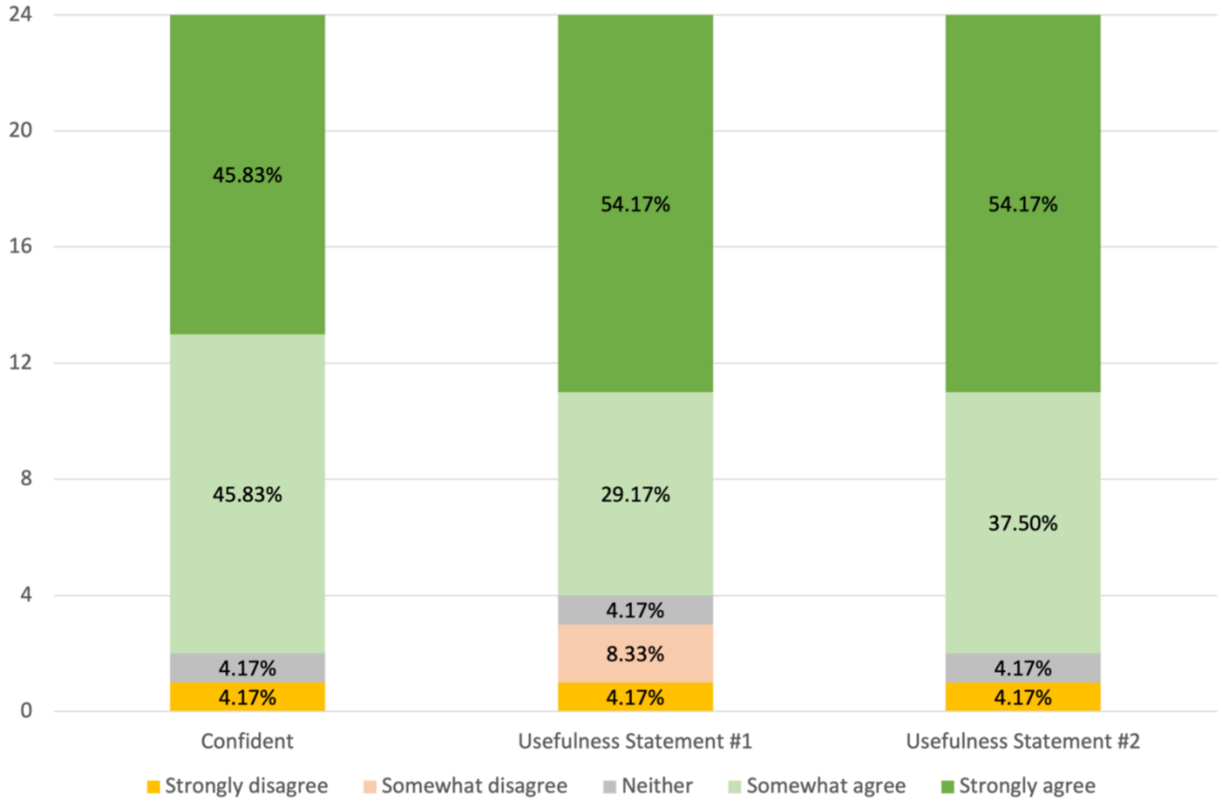


Figure 33. Agreement levels among participants for the single bivariate map.

Additional feedback and comments were also collected. Several participants mentioned that a bivariate map is a valuable tool in understanding “sensitive areas” and see their usefulness to public health professionals. An example of a response from a participant is:

> The bivariate map is useful for the comparison of within-cluster rates, which is very useful for large clusters or states with large and varying geographic areas. I see this as useful for state legislators/public health professionals, whereas other visualizations may be more helpful for between-cluster comparisons or federal public health providers.

Some participants also provided some suggestions on the design and functionality of the bivariate map. Some mentioned that the legend and its color “can be a bit confusing,” and many suggested having an additional function to be able to locate a specific date. Additionally, some of the participants think it requires more information on the website to assist the general public understanding the maps:

> Mmm, it’s okay for health professionals who have background knowledge in data science/analytics. If I was a regular person this would seem like too much or confusing – maybe add a reference key/footnotes so [people] can get a quick understanding of the maps vs. having to flip between tabs to understand the functionality of the map(s).

The combined bivariate maps: Tab “@LISA”

Regarding the accuracy of the combined bivariate maps, 91.67% of the participants (n = 22) were able to recognize the cluster indicated on the LISA map, while no clusters were detected on the left map using the SaTScan algorithm. Furthermore, participants were requested to select the agreement level for three statements that discussed the reasons for the different results from the two maps (Figure 34). Over 70% of the participants (the range is n = 18~19 for accuracy statement #1~#3) indicated that they were able to understand and interpret the differences. Regarding the usefulness, 91.67% of the participants (n = 22) agreed that these two bivariate maps combined were useful; however, 20.83% of them (n = 5) disagreed that one bivariate map is easier to interpret than two maps side-by-side (see Figure 35). Regarding the decision making question, over 70% of the participants agreed that they could use both maps for different purposes (n = 17~18 for decision statements #1 and #2) and not simply rely on the results from one single map (n = 22, decision statement #3).

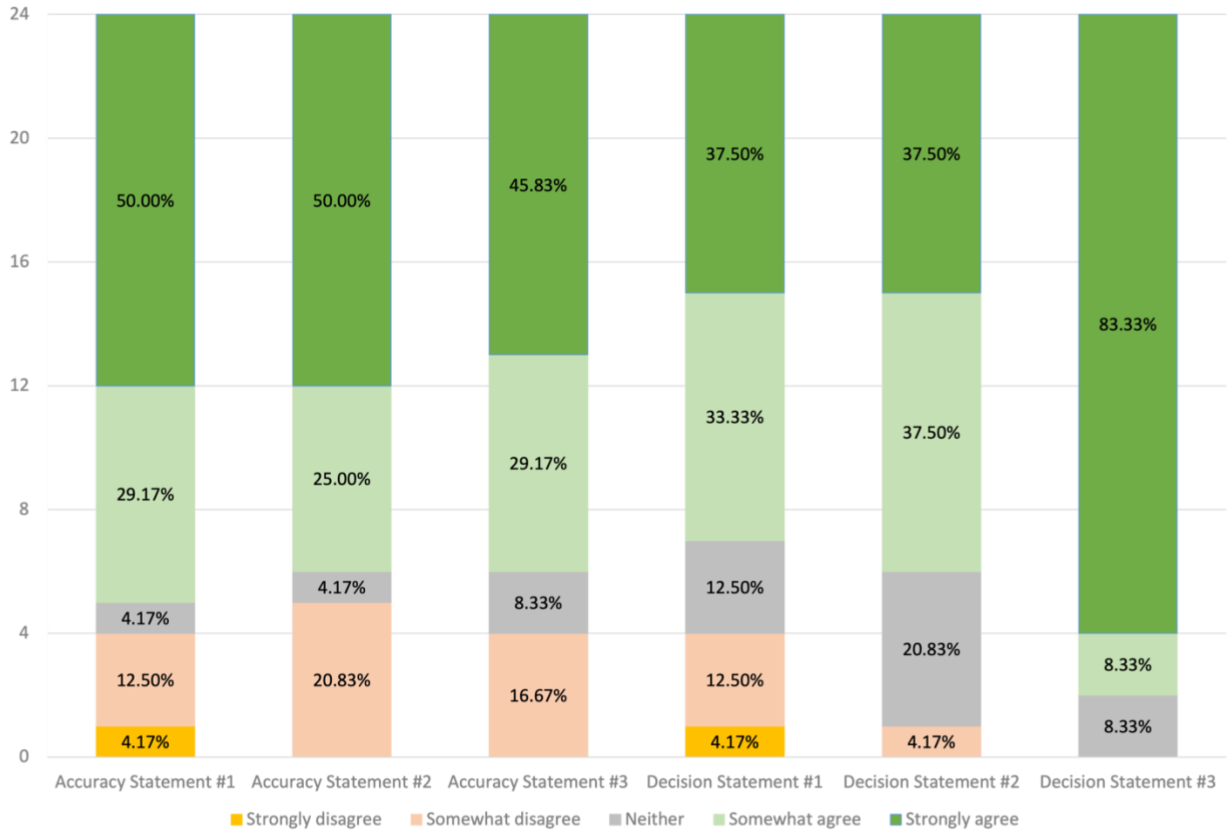


Figure 34. Agreement levels among participants for two bivariate maps, provided side-by-side.

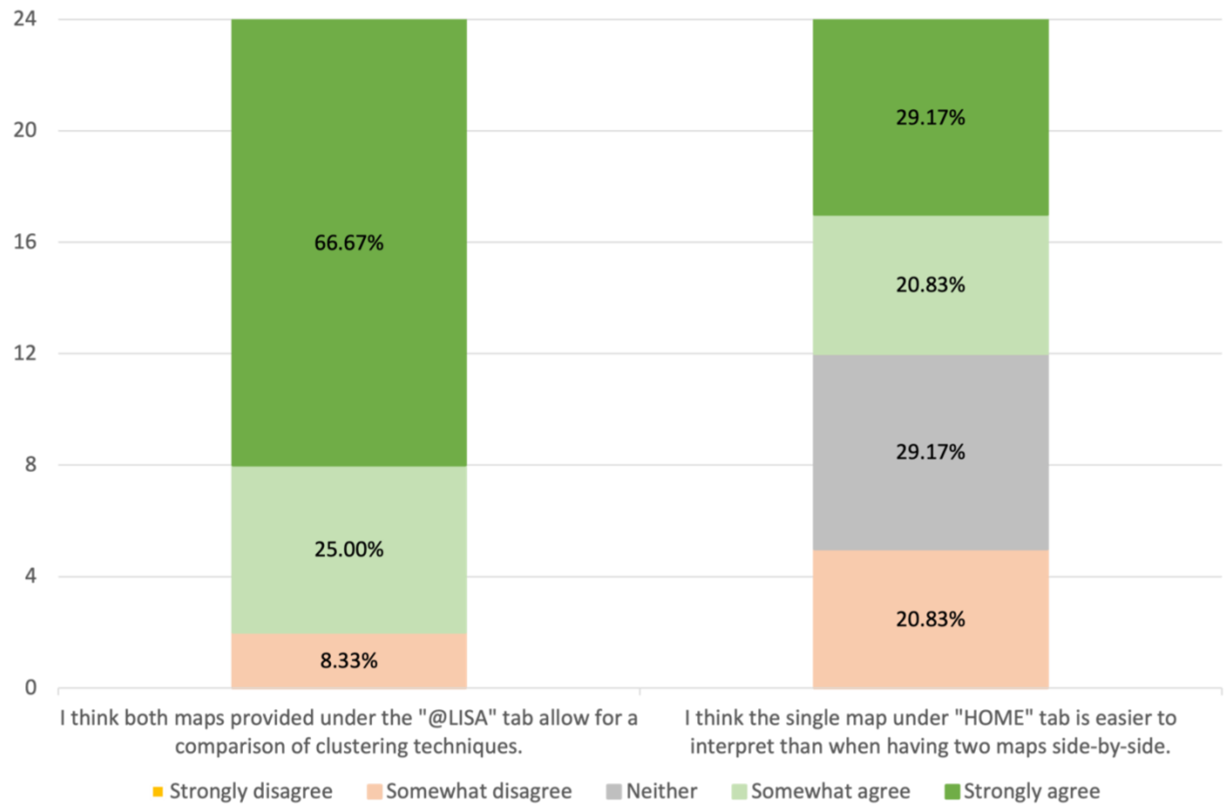


Figure 35. Usefulness of the two maps and comparison with the single bivariate map.

Although some participants thought it is “overwhelming” or “confusing” to observe two maps simultaneously, most of them believed that each map could provide helpful information from different perspectives. For example, one participant mentioned:

> If I’m a policy maker, I would probably primarily use the right map to make decisions about alerting counties and implementing policies because it shows acute emergence in the state. The left map is also useful for making longer-term decisions and seeing if a cluster may be nearing the border where we might need to prepare. It would also help to assess where we should focus resources, but the right map will tell me when those daily rates are letting up when a cluster will likely be decreasing. So, I think both are useful in policy making.

The spiral map at the state level: Tab “@State”

Regarding questions of accuracy for the spiral map, approximately 80% of the participants identified the correct value (n = 20) and date regarding the state when it had the highest average relative risk (n = 19) and the state that had the highest value using the map located in the center of the spiral (n = 19). Regarding confidence (Figure 36), although there were three missing answers, over 76% of the participants (n = 16) stated that they were confident in their responses. Concerning usefulness, over 60% of the participants agreed with the spiral map's effectiveness in connecting the bivariate map (n = 15, usefulness statement #1), exploring the latest value (n = 16, usefulness statement #2), and showing patterns at the state level (n = 17, usefulness statement #3). Regarding decision making, many of the participants (n = 17) agreed to utilize the pattern established from the visualization to decide the policy for the state (decision statement #1), while fewer (n = 14) agreed to use it to evaluate implemented policies, such as a lockdown (decision statement #2).

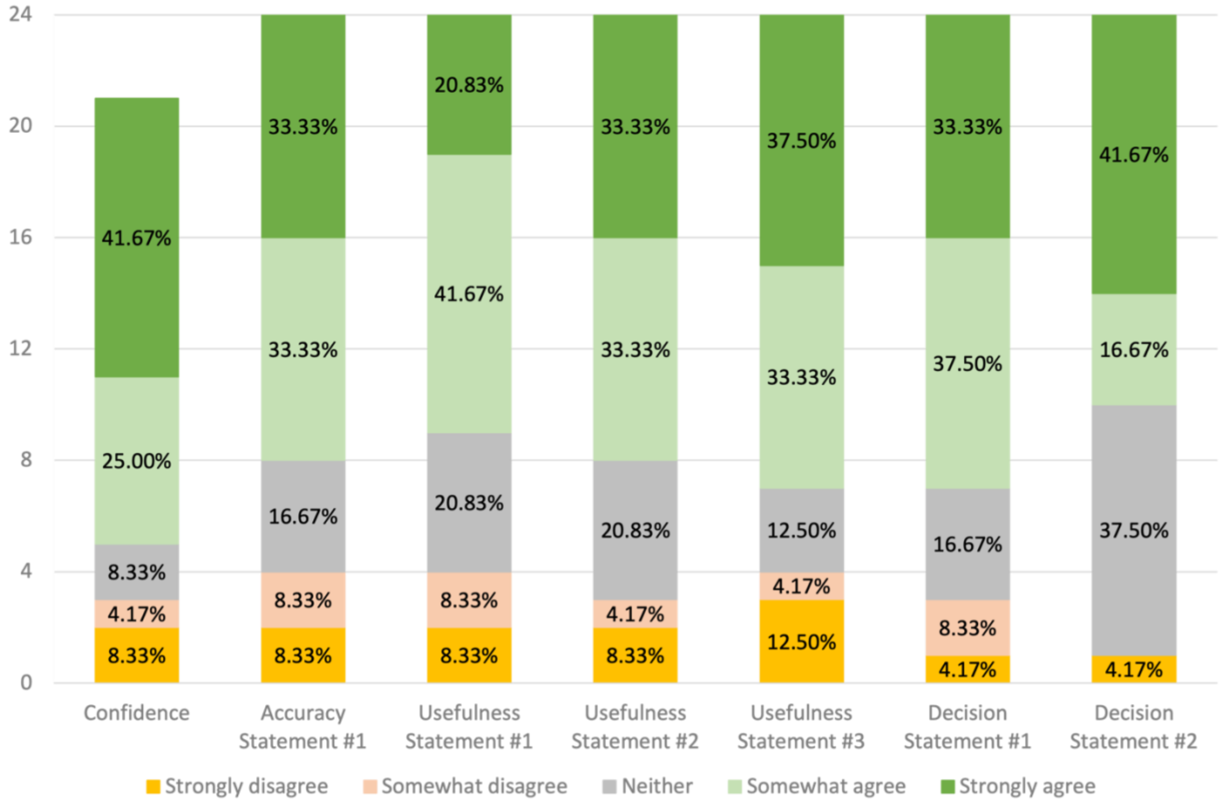


Figure 36. Agreement levels among participants for statements on the spiral map.

Some participants mentioned that the spiral map is complex or “tricky” to understand or navigate (e.g., “It is more difficult to read than the other maps”), and they suggested having some guidelines for it. However, many of them had different opinions of the visualization’s usefulness, such as “The spiral map is incredibly helpful because you can’t remember what is happening at different points in time across the bivariate map playback” and “I find the spiral map very useful for between-state comparisons and for estimating the success of COVID-19 implementations in reducing transmissions.”

Time series charts: Tab “TimeChart”

Regarding the accuracy of the *TimeChart* (Figure 37), approximately 80% of the participants (n = 19) could identify the time range with the highest value in three years for one county, while the percentage to determine the date for the highest value of another variable was lower, at approximately 45.83% (n = 11). One possible reason is that the time range was relatively short, which participants may neglect. Approximately 80% of the participants (n = 20) were able to identify which county has more often been part of a cluster when comparing two counties. Furthermore, 74% of the participants (n = 18) could identify a similar pattern between the two counties in California (accuracy statement #1), and 66.67% (n = 16) could identify the difference of relative risk between the two counties (accuracy statement #2). Concerning the confidence level, 83.33% of the participants (n = 20) felt confident in their answers. Regarding the usefulness of the visualization, 87.5% of the participants (n = 21) agreed that the TimeChart was helpful for identifying the trends of provided variables (usefulness statement #1) and comparing counties (usefulness statement #2). Concerning decision making, 87.5% of the participants (n = 21) agreed to carefully monitor the RR of the county due to the high value of this variable, while the other two values (RR of the cluster and 7-days avg. case) are decreasing according to the *TimeChart*.

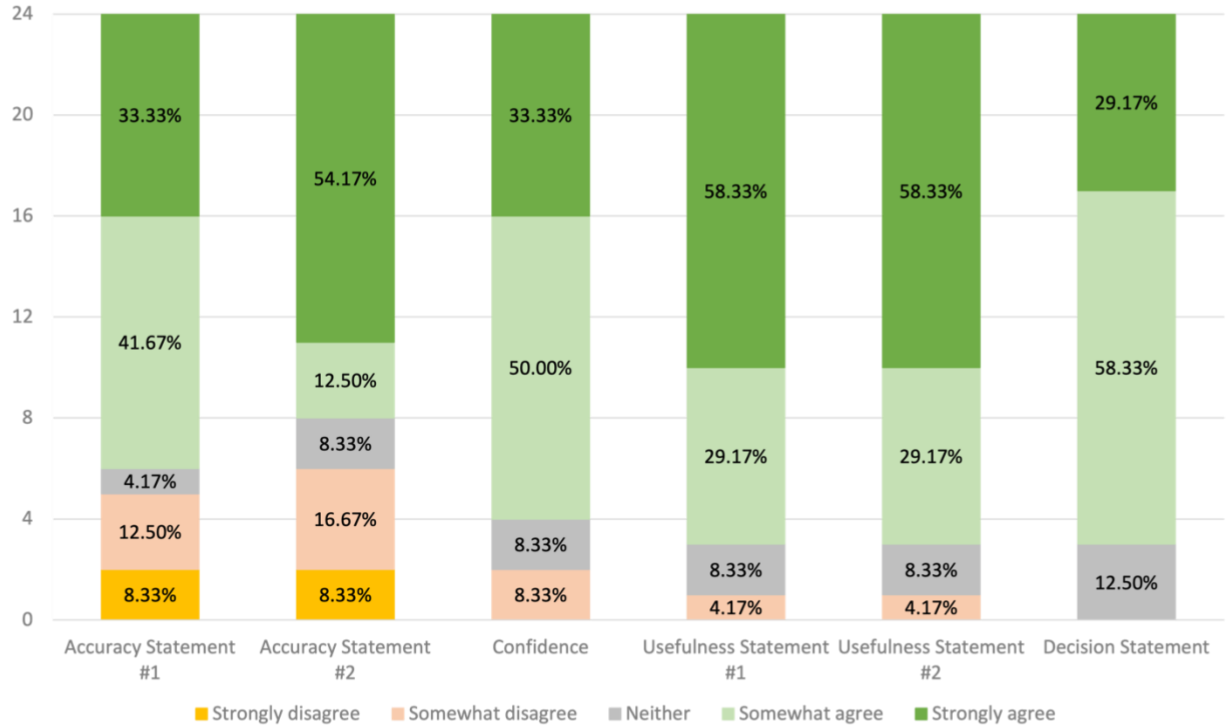


Figure 37. Agreement levels among participants for statements of the *TimeChart*.

Although some of the participants considered the *TimeChart* complicated and challenging to get the exact rate, some mentioned its usefulness in comparing multiple counties. For instance, one respondent said, “I love the ability to directly compare counties by being able to select multiple counties at a time. This function is great for public health resource allocation.”

3D space-time cube: Tab “@3D”

Regarding the accuracy of the 3D space-time cube, 87.5% of the participants (n = 21) identified the county and date with the highest relative risk among the selected space and time range. Concerning the confidence level (Figure 38), 79.17% (n = 19) felt confident in their answers. Regarding the usefulness, 75% (n = 18) agreed that the 3D space-time cube is valuable, and 50% (n = 12) agreed that it is easier to understand the dynamic of space-time clusters using the 3D visualization. However, the rate was much lower when compared with other 2D-visualization techniques. Only 25% of the participants (n = 6) agreed that the 3D visualization was more useful than the animated bivariate map, and 12.5% (n = 3) agreed that it is more useful than the *TimeChart*. For decision making, 62.5% (n = 12) indicated that they would use the filter function to identify and warn counties with a higher relative risk (decision statement #1) about the significant high risk of transmission and that they would investigate the cluster shift as shown from the 3D space-time cube (decision statement #2).

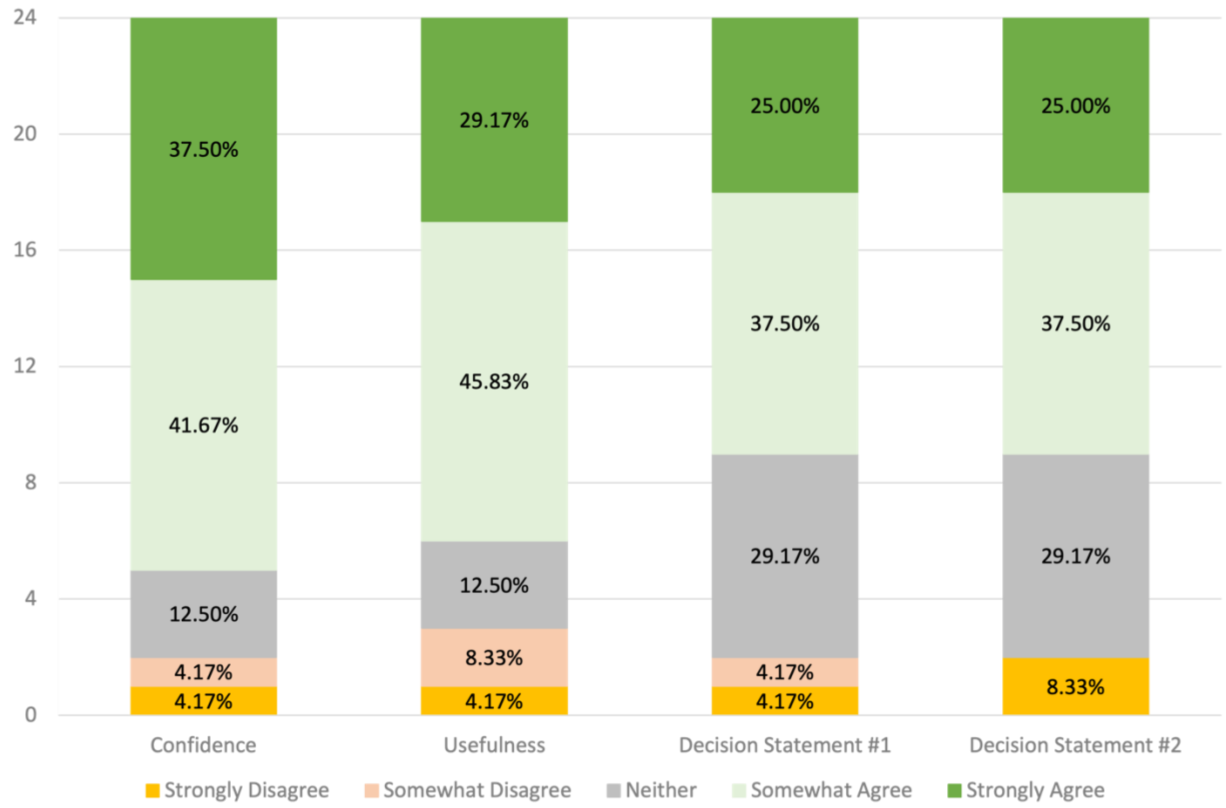


Figure 38. Agreement levels among participants for statements of the 3D space-time cube.

Although some of the participants (n = 5) thought a 3D visualization was great for representation, a higher number of the participants (n = 10) mentioned it was difficult to navigate, use, or interpret. Some of them also said that 2D visualization was more appropriate in some ways. For instance, one participant mentioned:

> I can see this type of mapping to be very useful. As it is now, it can be a little difficult to interpret compared to other forms of space-time mapping. For instance, while the orange and red dots do stand out, it is less clear than some of the 2D representations as to which county was most at risk and when.

Results from the five visualizations were summarized based on those four evaluated categories, as shown in Figure 39. When multiple questions were asked within one category, the percentages displayed in Figure 39 show the average percentages for that category. Regarding accuracy, responses with corrected answerer were in the range of 79~88% (n = 19~21), while the TimeChart was slightly lower with a rate of 71% (n = 17). For the confidence level, the lowest level was for the spiral map (67%, n = 16), and the highest was for the single bivariate map (92%, n = 22). Regarding the rate of usefulness, the one with lowest positive responses was also for the spiral map (67%, n = 16), and the highest was for the two bivariate maps (92%, n = 22). For the rate of decision making, the one with lowest positive responses was for the *3D space-time cube* (50%, n = 12), and the highest was for the *TimeChart* (92%, n = 22).

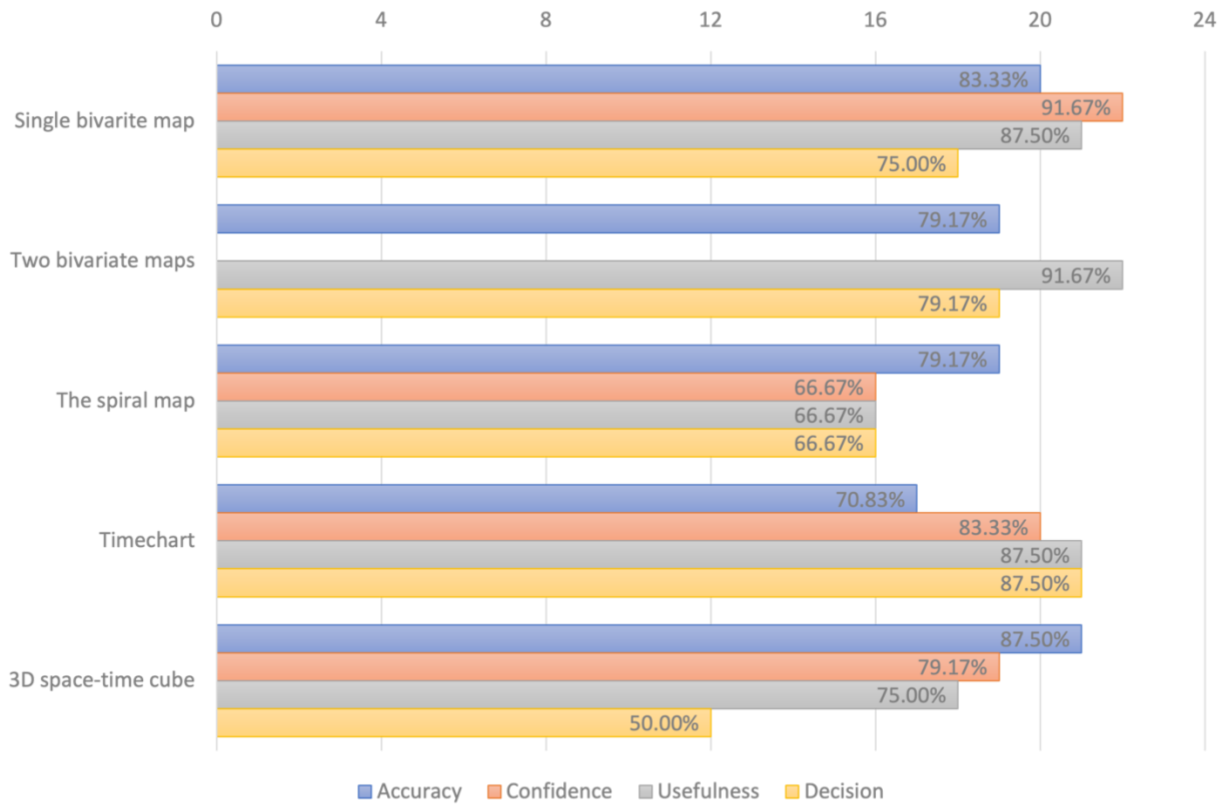


Figure 39. The summary of positive responses for all five visualizations.

The 2D visualization is also compared with *the 3D space-time cube* in Figure 40. Many participants (n = 9) disagreed that the 3D visualization is more useful than *the single bivariate* map. Similarly, a higher number of participants (n = 12) disagreed (and n = 5 strongly disagreed) that the 3D visualization is more useful than the *TimeChart*.

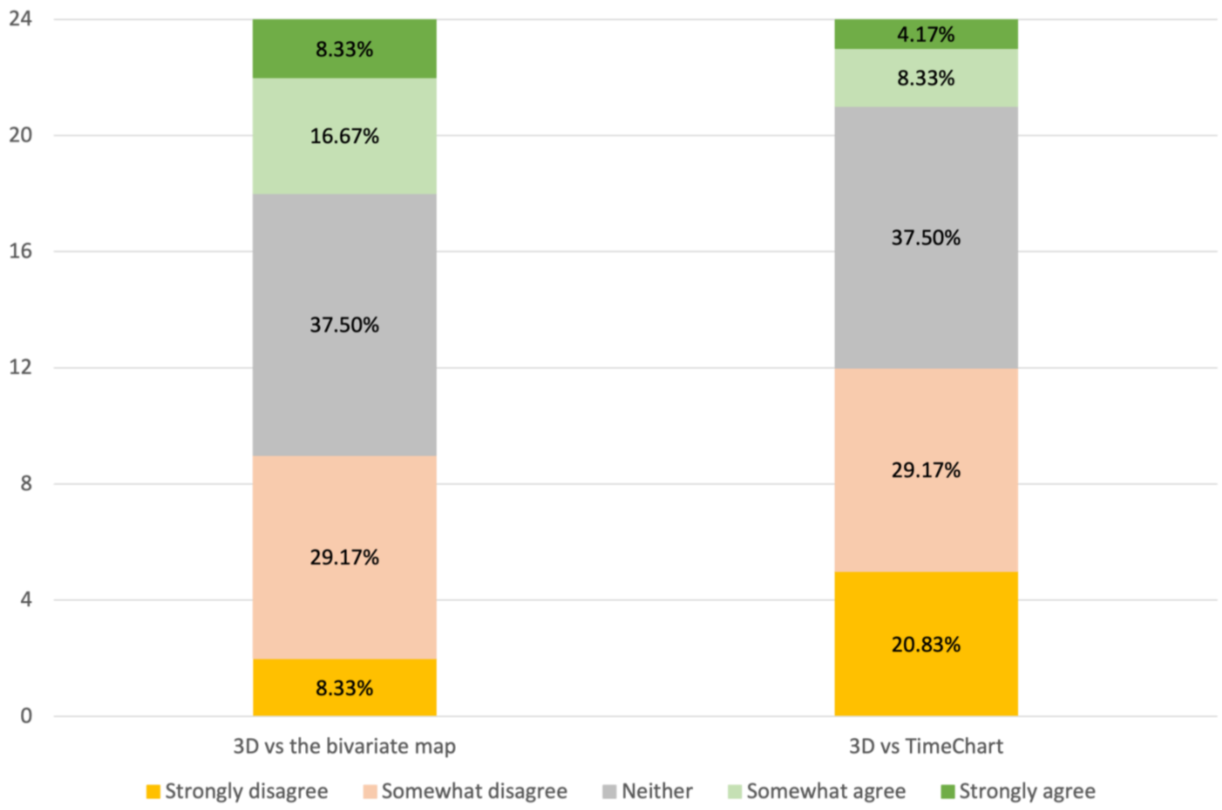


Figure 40. The results of usefulness when comparing different visualizations.

Geography student participants

The responses of geography graduate students were analyzed separately, but the response was similar to that of graduate students in public health. However, the geography graduate students' feedback that was collected through open-ended questions was more critical compared to the public health participants' feedback.

Participants were also asked to provide comments in general. Over 91% of them agreed that the introduction session and the provided information were helpful for them to interpret represented information from the visualizations and answer the questions in the survey. Some participants enjoyed the functionality in general and mentioned its usefulness. For instance, one participant said:

> The system is impressive. It needs a few enhancements to improve the user experience (UX). I suggest video tutorials, in addition to the read me button, to guide users on how to navigate the system. The time chart is the best, the dual maps are also great. The 3D is somewhat difficult to visualize but when it's few cases, it explains very well, the space-time dimension.

Another participant stated:

> I am a big fan of this dashboard overall. I think providing multiple options for representing and interpreting space-time COVID-19 data is crucial for public health practitioners and very helpful. The only map that gets a little murky is the space-time cube, but if you use the zoom function, and practice with the controls, I think it becomes [clearer].

5.3.2 Expert evaluation

During the interview, all four participants introduced their experiences with COVID-19 data and dashboards. Two health officers who worked for the COVID-19 response unit at the Mecklenburg County in North Carolina also discussed their experiences with multiple types of COVID-19 data, including cases, testing rates, and wastewater data. They mentioned that they mainly use the North Carolina COVID-19 Dashboard (e.g., to analyze neighboring counties) and that the North Carolina Department of Health and Human Services provides information about clusters and outbreaks, but they also use the CDC COVID data tracker and COVID-19 Dashboard by Johns Hopkins University. Another participant had experience with analyzing COVID-19 data in health geography and was familiar with different health data visualization techniques. The fourth participant also had experience using the CDC dashboard for research in public health.

The **first task** of the group interview essentially asked participants to discuss the CDC COVID data tracker, which is the most familiar to recruited four experts. Concerning its effectiveness, one health officer considered it "a friendly tool" that can be used to compare the rates in Mecklenburg County to surrounding counties, while the other participants mentioned its ability to discover the general pattern of COVID-19 in the US. However, participants spent more time discussing its limitations. Both health officers said that the COVID-19 data was not updated daily and discovered that it did not match their county data. The academic participant in public health stated that it is difficult to locate and download the data. As this dashboard is not capable of cluster detection, we asked the participants whether it would be helpful to implement this functionality. Both public health officers stated that it would not work or be useful, while the health geographer said it would be helpful to educate the general public with such information. Concerning the usefulness of the maps from the dashboard, one health officer claimed that the community-level map from the CDC was the most useful map as the county government did not have access to COVID-19 hospital data, including new admissions of patients and percentages of staffed inpatient beds.

The **second task** asked participants to answer questions about the US Covid Atlas. The question about its effectiveness was also discussed. Two participants liked its overlay function, which can add extra boundaries to the map to examine specific regions. Participants also mentioned that it was easy to understand the map and that the map was clear. They agreed that some content in the US Covid Atlas was suitable for the general public but that some content would have required the readers to have knowledge in geography to properly interpret the map. As this system uses the LISA algorithm to detect areas with high evaluated risks of transmission, the second question referring to the usefulness of clustering information was asked. One health officer pointed out that a cluster may be conceptualized differently based on the audience. Other participants also expressed that cluster information at the county level is not helpful for public health officers from the county because such information does not show inner county variation, which depends on the spatial resolution data. The public health researcher also pointed out the importance of such results from a temporal perspective. They said this system is good at retrospective analysis but limited to prospective studies to identify current and future outbreak clusters quickly.

The **last task** asked participants to answer questions about my prototype. First, they were asked to provide overall feedback and comments on the system; second, participants were asked to provide comments on each visualization. The participants mentioned that the prototype has many details, but they suggested defining what cluster and relative risk are using in this syterm as these terms caused confusion. One health officer pointed out the different cluster definitions in geography and epidemiology. Participants mentioned that the color scheme of bivariate maps was excellent but that there were too many colors for human eyes to distinguish. Regarding the spiral map to display values in a time-series, they liked the design and suggested adding benchmarks (e.g., the date that some health policies, such as lockdown, were implemented) to help interpretation. For another time-series visualization, the TimeChart, the participants made similar comments to the comments made about the spiral map. Concerning the 3D space-time cube, the participants indicated that it was informative but overwhelming to use in the context of public health surveillance.

In conclusion, two questions were asked: 1) Which system is the most useful? and 2) can the dashboard provide rich information to make decisions? The academic researcher in public health pointed out that each system's audience, as well as its purpose, is essential. The health geographer thought that the CDC system was not very useful and that the other two systems (the US Covid Atlas and the US COVID-19 YuTu) have nice maps and are informative. Furthermore, this expert stated that the Atlas platform was more straightforward and provided more information. Both health officers mentioned that the three systems have different purposes. They must use the CDC system because of the data instead of maps, and the Atlas system is good at combining with other information, which could be used to understand patterns behind the scenes. My prototype is suitable for in-depth analysis of specific research questions, and it could also be beneficial for providing precise definitions of terms used, such as cluster and relative risk. Generally, it was agreed that a more granular scale was needed to assist in decision making. All three systems were considered suitable for hypothesis testing but did not provide information to answer questions in decision making directly.

I further analyzed the transcriptions of the three group tasks using NVIVO, a qualitative data analysis computer software package. Each task's transcriptions were automatically generated using the software and manually checked and changed incorrect words that generated automatically to make them meaningful in the context. Then, the text of those three transcriptions was coded based on a codebook with two main types, descriptive and analytic codes. The descriptive codes included data, audience, time, and cluster; the analytical codes included pros, cons, scale, and usefulness. The coverage rate was calculated, which showed how much of the source content was coded in the transcriptions. The word clouds of the 20 most frequent words were also generated after removing irrelevant stop words (e.g., "like" and "think").

The coverage rate of coding results for the three transcriptions is shown in Table 3, and the highest coverage of each code is highlighted. The word clouds (Figure 41) were generated from the transcription of each task to further understand responses from those four experts to questions. Regarding the CDC system, the participants mostly mentioned "data" and "cluster" from the coverage table and its word cloud. This is because the CDC system is more like a "data hub" to share data, and the definition of the cluster was first discussed during this interview. Concerning the coverage of codes in the Atlas system, most discussions focused on the audience, the scale, and the scale's usefulness, which is reflected in its word cloud. The scale was discussed along with the overlay function. Regarding my prototype, significantly more discussion about time (as a temporal dimension) was identified. One reason for this focus was that the *YuTu* system focused on space-time visualization. The pros and cons of the *YuTu* system were also discussed more than the other two systems, as the purpose of this group interview was to evaluate this system.

Table 3. The coding results of transcription of three group tasks.

	Descriptive codes by coverage (%)				Analytical codes by coverage (%)			
Source	Data	Audience	Time	Cluster	Pros	Cons	Scale	Usefulness
CDC	7.77	0.82	-	21.80	0.86	11.80	0.65	4.85
Atlas	1.24	15.36	4.84	13.25	3.17	4.55	9.65	11.90
YuTu	-	4.36	26.77	3.12	5.86	20.53	-	7.91

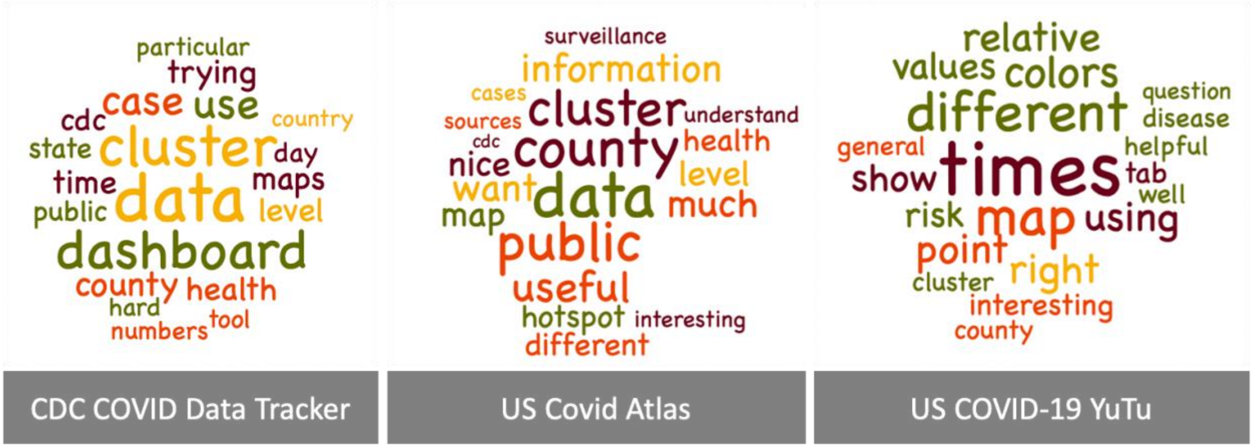


Figure 41. The word clouds for three transcriptions of group tasks.

5.4 Discussion & Conclusions

In this chapter, a user study and a group interview are introduced to evaluate my proposed system, *US COVID-19 YuTu*, to collect feedback from different perspectives. The evaluated system is implemented with several novel 2D and 3D visualizations.

According to the user study with graduate students in public health (n = 18) and geography (n = 6), the system is generally well designed as good responses were noted regarding accuracy, confidence, usefulness, and assistance with decision making. For each visualization, these four categories were evaluated. *The bivariate map* is easy to understand, while the combined maps with two different algorithms tend to be confusing. Compared with other visualizations, *the spiral map* gained fewer positive feedbacks from all three categories (confidence, usefulness, and assistance with decision making) except accuracy, which suggests a need to improve the visualization by adjusting representation and functionalities. In contrast, the *TimeChart* visualization obtained high votes in all categories except accuracy due to insufficient representation to identify patterns when the time span of a pattern (e.g., high RR) shown in the chart is relatively short. Although most participants correctly identified values and information from the *3D space-time cube*, minor participants agreed with the proposed policies derived from this visualization, suggested that it may be difficult to be utilized in decision making.

According to the group interview with experts, the missing component of the *YuTu* system is the provision of clear definitions of important concepts, such as clusters and relative risk. Furthermore, guidance on how to interpret visual representations from each visualization is also necessary. Since the recruited health officers work for the county government, their responsibility for COVID-19 monitoring mainly focuses on the local level. However, patterns indicated in the system are more beneficial to health officers who work for the state or on the national level. As a result, such patterns were not discussed during this group interview. Regarding the temporal information, experts from the meeting suggested that meaningful time points, such as the date when one COVID-19 variate dominated, should be added on the time-series visualizations. Participant responses for the 3D visualization were similar to the responses from the user study where participants felt unfamiliar and overwhelmed.

This evaluation combines two qualitative methods to assess the system. On the one hand, feedback collected from the user study with graduate students mainly assessed the functionality and usability of the proposed system by analyzing their responses. On the other hand, experts with relevant experience in such systems discussed and evaluated the prototype in depth and its potential usefulness in the real world. The user study suggests that novel visualizations are informative to represent the massive volume of space-time cluster information, which is usually limited in traditional visualizations (e.g., multiple static maps). In the group interview, the experts recognized the useful design of the system and rich details from the visualizations. However, they recommended that clear definitions of important terms and guidance for result interpretation are given, which are essential to utilize the system as a decision support tool.

The evaluation indicates that most of the participants from both groups did not prefer the 3D space-time cube due to it being overly complicated in terms of interactivity and interpretation. Although it displays the shape of space-time clusters in a natural 3D environment that cannot be achieved by the traditional 2D methods, the results suggest that further investigation of a better design of visual representation is needed.

This study also reveals the existence of disconnection between professionals and academics, such as the different understanding of the term, “cluster”. Health officers who work in the county consider a cluster to be a small group of people, such as a cluster in one high school. However, the cluster in spatial epidemiology refers to an area with an elevating risk of transmission of diseases. According to the literature search in Chapter 2, hotspots and clusters are usually interchangeable terms; however, public health professionals do not use these terms interchangeably. Thus, for systems designed for professionals, it is critical to provide a clear definition of the terms used.

Despite the strengths of this study, several limitations are also acknowledged. First, collected responses may be different even with similar questions, as it is impossible to develop questions (or statements) that is most suitable for evaluation. Although I designed most questions in the survey to be closed-ended, many questions or statement are subjective to my own perception on interpreting visual representations, especially when the answer is not number, like the 5-point Likert scale question. It is possible that participants understood the visualization and its information, but they disagreed on some provided statements because of different interpretations.

Second, although we recruited graduate students from both public health (n=18) and geography (n=6), the number of participants in each group was unequal. Because of these non-balance responses, I did not conduct a statistical test to identify whether meaningful differences existed between these two groups. This issue could be addressed by recruiting a higher number of participants from public health and geography to evaluate this system in the future.

Third, we only recruited health professionals who worked for a county instead of a state or on a national level. As a result, they are interested in data on a more granular scale, and results at the county level are not that useful for monitoring the transmission of COVID-19. The health professionals also mentioned that cluster information at the zip code level was not that useful due to the nature of human mobility (e.g., moving in and out). Thus, this limitation could be addressed by recruiting health professionals from the Department of Health and Human Services of a state or a national government in public health (e.g., CDC) to evaluate the system, as they work with data at a similar scale that the YuTu system mainly focuses on. However, it is difficult to recruit participants from the government without a connection.

Fourth, the evaluation reveals that health officials were interested in specific data for decision making, while they showed less interest in novel visualizations. The health officials mentioned that the most helpful map from the CDC dashboard is the map showing COVID-19 community level data by county, based on cases together with relevant patient data in the hospital. However, such data is not available for public access. The health officials found it hard to interpret results from novel visualizations, such as the 3D space-time cube, while it is my goal that those novel visualizations in *YuTu* system can assist in decision making. A further evaluation is necessary to be conducted to compare the usefulness of traditional visuals (e.g., basic online maps) with proposed advanced visuals.

This chapter indicates the significant role of qualitative research in developing a health surveillance system to assist decision making. The user study and the expert evaluation evaluated the *YuTu* system in terms of its usability and usefulness. Based on survey results and the group interview, information represented in the 3D space-time cube is difficult to interpret. Therefore, further research on such visualization is needed before implementing it into a surveillance system. More importantly, the usefulness of other 2D visualizations and possible required components (e.g., different data and methods) in decision making remains unclear and needs more attention in future work. Although experts did not favor one of the three evaluated systems, they expressed the importance of data and the system’s simplicity when comparing them. The profitable way to develop a health surveillance system that can assist in decision making is to collaborate with health experts from the beginning of the design phase to the end of the evaluation phase.

CHAPTER 6: GENERAL DISCUSSION AND CONCLUSIONS

In this dissertation, I proposed a web-based geographic framework to detect and visualize space-time clusters of infectious diseases. In the context of the current COVID-19 pandemic, there is a need for effective surveillance that monitors the dynamic transmission of the virus by taking advantage of data at fine spatial and temporal resolutions, as such data becomes more available than ever before due to the development of web technologies. The proposed framework is a response to this need and further explores possible visual representations of such data.

Chapter 2 identified the most popular methods used to detect space-time clusters of infectious diseases. I conducted an electronic literature search for relevant studies published before 2022 from three popular databases, including PubMed, the Web of Science, and Scopus, around this topic. The systematic review reflected recent trends in space-time cluster detection for infectious diseases. According to the results, the most popular clustering methods were the space-time scan statistics that detect regions with the maximum likelihood ratios and the LISA algorithm that identifies regions of high spatial autocorrelation. However, the implementation of such clustering methods together with representation of these analytical results into a web-based environment is missing from the literature.

Chapter 3 presented research gaps from the literature review and derived the general problem statement for this dissertation. In this chapter, I identified four critical issues that remained unclear in the literature: 1) the use of statistical methods (e.g., LISA, GI*) in clustering detection with a reflection in the temporal dimension, 2) the need for robust geovisualization to reveal space-time patterns of infectious diseases, 3) the urge for rapid surveillance on a fine temporal scale, and 4) the demand for an integrated online platform for the mentioned components. Thus, this dissertation proposed a new method to address these issues.

In Chapter 4, I introduced *YuTu*, a prototype system of the web-based analytical framework for detecting and visualizing space-time clusters using US COVID-19 data. To address the first issue identified in Chapter 3, I implemented the prospective space-time scan statistics and the LISA algorithm into the system and deliberately introduced terms of “true” (for the former method) and “not-true” (for the later method) space-time clusters to reflect the differences between these two methods. The system is articulated around five novel visualization approaches in an attempt to address the second issue, including *the animated bivariate map, the combined bivariate maps, a spiral map, the TimeChart, and the 3D space-time cube*. *The animated bivariate map* represents the results of the SaTScan algorithm and displays the values of two important variables, namely the relative risk of the cluster a county belongs to at the time of the analysis. Furthermore, this visualization is the baseline visualization linked with other visualizations to display space-time data in two factors. The second visualization is *two animated bivariate maps* side-by-side using the abovementioned two methods. In this way, clustering information that considers past days (using the SaTScan method) and information that focuses on the current day (using the LISA algorithm) are displayed together, which allow to combine this two information together to interpret the results. The third visualization linked the basic bivariate map with *a spiral map* to show each state’s daily average relative risk. Thus, the relative risk pattern of each state is shown from the start of the pandemic to the current date, and the distribution of this value at the state level is also displayed and updated every day. Users can easily understand the dynamic change of space-time clusters using the animated map, while memorizing all daily changes over the years is cognitively challenging. The fourth visualization approach combines the *TimeChart* with the animated bivariate map by displaying the pattern with all values in a continuous time-series chart. The *TimeChart* also allows users to select multiple counties to examine the change in relative risk in the same time series. Lastly, *the 3D space-time cube* displays the relative risk by applying gradient colors to the point representing each county, with the third dimension (z-axis) representing the temporal dimension. Therefore, the shape of true space-time cluster is shown, which cannot be recreated in a traditional 2D visualization.

Finally, the design of this framework addressed the third and fourth issues identified in chapter 3. Every day, the system automatically collects data for analysis and generates results for visualization by running scheduled scripts. The system is built in an open-source environment, including the database (PostGIS), the interface and its visualizations (D3), analysis procedures with multiple programming languages (e.g., Python), and other components (e.g., Docker).

The evaluation of the system was discussed in Chapter 5. To collect comprehensive feedback, I conducted two types of evaluation: a user study with graduate students in public health and geography and a group interview with experts who have experience in COVID-19 data and analysis. The user study suggested that the prototype system is generally well designed and that the various visualizations provided details and multiple ways to display and interpret space-time cluster information. Experts from the group interview pointed out the conflict regarding the definition used in the system of a cluster and demanded further information to interpret the relative risk. This evaluation result indicates the significance of communicating with audiences to agree on fundamental concepts in health surveillance systems such as *YuTu*. With the elaboration on such agreements, the target audiences (e.g., health officers) could gain the confidence to investigate the results and further utilize their interpretations in their decision making. Therefore, the user study and the expert evaluation seem important for evaluating such a system in epidemiology.

Serval issues are worth to be discussed. To begin with, it is critical to visualize space-time information in a way that is easy to understand, while the richness of such information should not be compromised. In other words, the key issue is how to represent the space-time information to be easily interpreted whilst keeping the richness of such information. The evaluation results in this dissertation show that users recognized the usefulness of space-time information from the visualizations, while some space-time representations may be overwhelming to them. The visualizations introduced in this dissertation shed light on visuals analytic system of space-time data, and creative ways to visualize high volumes of data in space and time still require investigation.

3D visualization is still unfamiliar to most people, and its implementation in health surveillance systems must be treated with caution. Both the experts and graduate students responded with concerns about transferring the identified patterns from the 3D visualization into decision making or even interpreting the patterns. However, it is still valuable to investigate such issues, as 2D can hardly display space-time data in high volumes as one single illustration simultaneously (e.g., animation can visual space-time data but not all at once). For instance, 3D visualization is capable of showing the seasonality of infectious diseases by displaying the daily incidence of infection in the US for one year at the same time, while it is challenging to represent such space-time information at once. Possible research questions could be whether users will use such a tool after training and then discern how to utilize the unique patterns in 3D. Additional immersive visualizations or real-time visualizations, including augmented reality, virtual reality, and mixed reality, could be a solution to this issue. Additionally, the evaluation of 3D representations should be included to understand the usefulness.

Another unavoidable and critical issue is the scale in spatial and temporal dimensions. The proposed framework is capable of being scaled up or down depending on the different audiences (e.g., health officials monitor COVID-19 transmission within a county or the entire US). I was attempted to include local scales (e.g., zip code) for the whole study area, while no open-source data at this level is available. With data and results at a finer scale shown in the surveillance system, it would be more beneficial for health officials to use it for decision making in the local communities (e.g., census tracts within a county). The geographic scale can also be scaled up, such as analyzing the clusters at the country level, which are mostly beneficial to health officers in the UN. Similar, the time interval can be daily, weekly, monthly, yearly, etc. It is worth investigating a combination of different time scales to uncover patterns.

When implemented with more than one method, the similarity and dissimilarity in results using different algorithms exist. In *YuTu* system, I did not discuss the accuracy of clusters from the space-time scan statistics and the LISA algorithm, as the time information is treated distinctly between these two methods. I proposed a way to use the LISA result as a complementary (spatial clustering on one day) to interpret continuous results in space and time (true space-time clusters) using the other method. However, the purely spatial scan statistic was compared to the LISA algorithm for spatial clustering by (Grubesic et al., 2014), and their conclusion is that there is no a superior method as a trade-off between log-likelihood ratio and spatial accuracy existed.

Another critical issue is the existence of edge effects when conducting a spatial analysis. Although the scan statistics scans any point within the study extent, a location on the edge is still less likely to belong to a detected cluster than a location in the interior (Gangnon, 2012). The edge effect has an influence on type I error when conducting spatial scan statistic, and Bayesian smoothing can be used to descript the spatial distribution of type I error (Guttmann et al., 2014).

Also, counties or territories from the neighbor countries have an impact on the transmission of COVID-19, while this *YuTu* system only takes accounts in the incidence in the US.

Providing an option for users to change the parameter setting accordingly could be helpful to further assist them in the rapid response of dynamic changes during an outbreak of infectious diseases. From the prototype, the maximum temporal window was selected based on one year of data, and possible signals of the peak of waves during that year were detected. Situations were changed, such as the increasing of vaccination rates and the different characteristics of variants. Thus, with the capability to change parameters of analysis that reflect situational changes, patterns that are closer to events that are happening in the real world may be detected.

Lastly, it is vital to collaborate with the target audiences (e.g., health officers work for a state government) to avoid disconnection. For instance, the target audience can offer suggestions for selecting methods and guidance to educate users to interpret results. The format of such collaboration could be different qualitative methods, such as in-depth interviews or focus groups with multiple audiences.

I suggest serval significant avenues that can be beneficial to develop similar health surveillance systems for monitoring current and future infectious diseases. First, the prevalence rates of COVID-19 used in the *YuTu* system need to be smoothed to avoid the bias in the data. For instance, a county may have a higher risk of COVID-19 infection if it has a higher percentage of senior population as they are more vulnerable to be infected. Therefore, space-time smoothing methods (e.g., Bayesian smoothing) are needed.

Second, it is important to include further analysis of uncertainty in future research. The uncertainty exists in the dataset, and it also refers to how to represent the uncertainty. In the

YuTu system, I applied a 7-day moving average to the daily COVID-19 cases, while the exact number is still hard to get as many cases are underreported by self-testing at home. Such uncertainty also exists in other used dataset, such as the population. As to representation, I did display the p-value for the LISA results as one variable from the bivariate map to reduce the uncertainty from the results, while the effectiveness of this representation remains unclear.

Third, as I discussed the importance of the collaboration with the audience, another vital issue is to decide the audience and the robustness of different scales to them. From the user study, participants were asked to evaluate the system by imaging they were health officials for the state. However, the recruited experts were from a county government, and they had limited interests on information at the county level. Thus, focus groups with health officers from other two scales (country and state) are necessary. It is also possible that the *YuTu* system could be useful to both two scales with different perceptive to interpret results.

Fourth, after the decision of the audience, it is also necessary to collocate with them on the design phase. The proposed framework and the *YuTu* system were designed and built according to gaps identified from the literature review instead of collaborating with the audience, although it is a starting point to facilitate this collaboration with proposed solutions instead of merely concepts. In the future research, based on the framework along with its visualizations, it can be applied to other infectious diseases or improve the current *YuTu* prototype. For either purpose, the collaboration with the audience (e.g., participant observation) during the (re-)design phase can gain insides of what information are useful and needed in their viewpoints.

Fifth, as to decision making, further evaluation with experts is needed to understand whether this system can lead to generating decisions in the real world. One possible way is to observe health officers if they want to use the system to make decisions in the real world.

However, it requires a long-term close collaboration with government agencies at the very beginning phase of the study.

Sixth, it is worth to investigating the way to implement changeable parameter settings from the interface. In the *YuTu* system, users can use the filter function from 3D space-time cube to focus on certain relative risk, while other parameters of two analysis methods are fixed. Although it requires advance knowledge of used algorithms, users may need to change the parameter settings to suit their needs. For instance, the temporal maximum cluster size from the space-time scan statistic is set to 50 days, while it may need to be reduced as the latest variant of the coronavirus (e.g., the Omicron variant) spreads more easily than earlier variants.

Effectively monitoring and representing the transmission of infectious diseases are critical in the health surveillance system, and my dissertation contributes to this topic. First, I conducted a systematic literature review that revealed significant gaps. Then, I attempted to address some of these gaps by proposing a framework that automatically generates clustering information in space and time and visualizes such information in serval novel ways. Furthermore, I assessed the prototype system based on the proposed framework by conducting a user study with graduate students and a group interview with experts in the domain. The findings from my dissertation not only present possible solutions to the gaps that were found but also could inspire the research community to investigate this topic further. Promising opportunities exist to improve the representation of space-time clustering information in epidemiology under the rapid evolution of web-based technologies. Lastly, I urge collaboration among and between disciplines to build health surveillance systems together in a thriving ecosystem that could evolve to suit the dynamic pattern in the transmission of infectious diseases.

REFERENCES

Aamodt, G., Samuelsen, S. O., & Skrondal, A. (2006). A simulation study of three methods for detecting disease clusters. *International Journal of Health Geographics*, *5*(1), 1-11.

Acharya, B. K., Cao, C., Lakes, T., Chen, W., & Naeem, S. (2016). Spatiotemporal analysis of dengue fever in Nepal from 2010 to 2014. *BMC Public Health*, *16*(1), 1-10.

Alimohamadi, Y., Taghdir, M., & Sepandi, M. (2020). Estimate of the Basic Reproduction Number for COVID-19: A Systematic Review and Meta-analysis. *Journal of preventive medicine and public health = Yebang Uihakhoe chi*, *53*(3), 151-157. https://doi.org/10.3961/jpmph.20.076

Anderson, C. L., & Robinson, A. C. (2021). Affective congruence in visualization design: influences on reading categorical maps. *IEEE transactions on visualization and computer graphics*.

Angulo, J. J., Haggett, P., Megale, P., & Pederneiras, C. A. (1977). VARIOLA MINOR IN BRAGANÇA PAULISTA COUNTY, 1956: A TRENDSURFACE ANALYSIS. *American Journal of Epidemiology*, *105*(3), 272-280.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, *27*(2), 93-115.

Anselin, L., Syabri, I., & Kho, Y. (2010). GeoDa: an introduction to spatial data analysis. In *Handbook of applied spatial analysis* (pp. 73-89). Springer.

Baddeley, A. J., Møller, J., & Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, *54*(3), 329-350. https://doi.org/https://doi.org/10.1111/1467-9574.00144

Ban, H., & Ahlqvist, O. (2010). User evaluation of a software interface for geovisualization and communication of uncertain urban ontologies. *Journal of Information Technology in Construction (ITcon)*, *15*(9), 122-131.

Barrows, H. H. (1923). Geography as human ecology. *Annals of the association of American Geographers*, *13*(1), 1-14.

Bivand, R. S., Pebesma, E. J., Gómez-Rubio, V., & Pebesma, E. J. (2008). *Applied spatial data analysis with R* (Vol. 747248717). Springer.

Bostock, M., Ogievetsky, V., & Heer, J. (2011). D³ data-driven documents. *IEEE transactions on visualization and computer graphics*, *17*(12), 2301-2309.

Boulos, M. K. N., & Geraghty, E. M. (2020). Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics [Editorial Material]. *International Journal of Health Geographics*, *19*(1), 12, Article 8. https://doi.org/10.1186/s12942-020-00202-8

Carroll, L., Au, A., Detwiler, L., Fu, T.-c., Painter, I., & Abernethy, N. (2014). Visualization and analytics tools for infectious disease epidemiology: a systematic review. *Journal of biomedical informatics*, *51*, 287-298. https://doi.org/https://doi.org/10.1016/j.jbi.2014.04.006

Chen, J., Roth, R. E., Naito, A. T., Lengerich, E. J., & MacEachren, A. M. (2008). Geovisual analytics to enhance spatial scan statistic interpretation: an analysis of US cervical cancer mortality. *International Journal of Health Geographics*, *7*(1), 1-18.

Cromley, E. K. (2003). GIS and disease. *Annual review of public health*, *24*(1), 7-24.

Cromley, E. K., & McLafferty, S. L. (2011). Introduction. In E. K. Cromley & S. L. McLafferty (Eds.), *GIS and public health* (pp. 2). Guilford Press.

Cuzick, J., & Edwards, R. (1990). Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *52*(1), 73-96.

de Azevedo, T. S., Lorenz, C., & Chiaravalloti-Neto, F. (2020). Spatiotemporal evolution of dengue outbreaks in Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *114*(8), 593-602.

De Ridder, D., Loizeau, A. J., Sandoval, J. L., Ehrler, F., Perrier, M., Ritch, A., Violot, G., Santolini, M., Tzovaras, B. G., Stringhini, S., Kaiser, L., Pradeau, J. F., Joost, S., & Guessous, I. (2021). Detection of Spatiotemporal Clusters of COVID-19-Associated Symptoms and Prevention Using a Participatory Surveillance App: Protocol for the @choum Study. *Jmir Research Protocols*, *10*(10), Article e30444. https://doi.org/10.2196/30444

Delamater, P. L., Street, E. J., Leslie, T. F., Yang, Y. T., & Jacobsen, K. H. (2019). Complexity of the basic reproduction number (R0). *Emerging infectious diseases*, *25*(1), 1.

Delmelle, E. (2009). Point pattern analysis. In R. Kitchin & N. Thrift (Eds.), *International encyclopedia of human geography* (Vol. 8, pp. 204-211). Elsevier.

Delmelle, E., Delmelle, E. C., Casas, I., & Barto, T. (2011). HELP: a GIS-based health exploratory analysis tool for practitioners. *Applied Spatial Analysis and Policy*, *4*(2), 113-137.

Delmelle, E., Hagenlocher, M., Kienberger, S., & Casas, I. (2016). A spatial model of socioeconomic and environmental determinants of dengue fever in Cali, Colombia [Article]. *Acta Tropica*, *164*, 169-176. https://doi.org/10.1016/j.actatropica.2016.08.028

Delmelle, E., Páez, A., & Kanaroglou, P. (2015). *Spatial analysis in health geography*. Ashgate Publishing, Ltd.

Desjardins, M. R., Whiteman, A., Casas, I., & Delmelle, E. (2018). Space-time clusters and co-occurrence of chikungunya and dengue fever in Colombia from 2015 to 2016 [Article]. *Acta Tropica*, *185*, 77-85. https://doi.org/10.1016/j.actatropica.2018.04.023

Diez Roux, A. V. (2001). Investigating neighborhood and area effects on health. *American journal of public health*, *91*(11), 1783-1789.

Diggle, P. J., Chetwynd, A. G., Häggkvist, R., & Morris, S. E. (1995). Second-order analysis of space-time clustering. *Statistical methods in medical research*, *4*(2), 124-136.

Eisen, R. J., & Eisen, L. (2014). Use of geographic information systems in infectious disease surveillance. *Concepts and Methods in Infectious Disease Surveillance, Wiley-Blackwell*, 219-229.

Farrington, C., & Beale, A. (1998). The detection of outbreaks of infectious disease. In *Geomed'97* (pp. 97-117). Springer.

Fuentes-Vallejo, M. (2017). Space and space-time distributions of dengue in a hyper-endemic urban space: the case of Girardot, Colombia. *BMC Infectious Diseases*, *17*, Article 512. https://doi.org/10.1186/s12879-017-2610-7

Gangnon, R. E. (2012). Local multiplicity adjustment for the spatial scan statistic using the Gumbel distribution. *Biometrics*, *68*(1), 174-182.

Gatrell, A. C., Bailey, T. C., Diggle, P. J., & Rowlingson, B. S. (1996). Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British geographers*, 256-274.

Getis, A., & Ord, J. K. (1996). Local spatial statistics: an overview. In P. A. Longley & M. Batty (Eds.), *Spatial analysis: modelling in a GIS environment* (pp. 261-277). John Wiley & Sons.

Ghosh, P., & Cartone, A. (2020). A Spatio-temporal analysis of COVID-19 outbreak in Italy. *Regional Science Policy & Practice*, *12*(6), 1047-1062.

Gómez-Rubio, V., Ferrándiz-Ferragud, J., & López-Quílez, A. (2005). Detecting clusters of disease with R. *Journal of geographical systems*, *7*(2), 189-206.

Goovaerts, P. (2009). Medical geography: a promising field of application for geostatistics. *Mathematical Geosciences*, *41*(3), 243-264.

Greene, S. K., Peterson, E. R., Kapell, D., Fine, A. D., & Kulldorff, M. (2016). Daily reportable disease spatiotemporal cluster detection, New York City, New York, USA, 2014–2015. *Emerging infectious diseases*, *22*(10), 1808.

Grubesic, T. H., Wei, R., & Murray, A. T. (2014). Spatial clustering overview and comparison: Accuracy, sensitivity, and computational expense. *Annals of the Association of American Geographers*, *104*(6), 1134-1156.

Güemes, A., Ray, S., Aboumerhi, K., Desjardins, M. R., Kvit, A., Corrigan, A. E., Fries, B., Shields, T., Stevens, R. D., Curriero, F. C., & Etienne-Cummings, R. (2021). A syndromic surveillance tool to detect anomalous clusters of COVID-19 symptoms in the United States. *Sci Rep*, *11*(1), 4660. https://doi.org/10.1038/s41598-021-84145-5

Gurjav, U., Burneebaatar, B., Narmandakh, E., Tumenbayar, O., Ochirbat, B., Hill-Cawthorne, G., Marais, B., & Sintchenko, V. (2015). Spatiotemporal evidence for cross-border spread of MDR-TB along the Trans-Siberian Railway line. *The International Journal of Tuberculosis and Lung Disease*, *19*(11), 1376-1382.

Guttmann, A., Li, X., Gaudart, J., Gérard, Y., Demongeot, J., Boire, J.-Y., & Ouchchane, L. (2014). Spatial heterogeneity of type I error for local cluster detection tests. *International Journal of Health Geographics*, *13*(1), 1-11.

Hernandez-Avila, J. E., Rodriguez, M. H., Santos-Luna, R., Sanchez-Castaneda, V., Roman-Perez, S., Rios-Salgado, V. H., & Salas-Sarmiento, J. A. (2013). Nation-Wide, Web-Based, Geographic Information System for the Integrated Surveillance and Control of Dengue Fever in Mexico [Article]. *Plos One*, *8*(8), 9, Article e70231. https://doi.org/10.1371/journal.pone.0070231

Hohl, A., Delmelle, E., Tang, W., & Casas, I. (2016). Accelerating the discovery of space-time patterns of infectious diseases using parallel computing. *Spatial and spatio-temporal epidemiology*, *19*, 10-20.

Hohl, A., Delmelle, E. M., Desjardins, M. R., & Lan, Y. (2020). Daily surveillance of COVID-19 using the prospective space-time scan statistic in the United States. *Spatial and spatio-temporal epidemiology*, *34*, 100354.

Hohl, A., Tang, W., Casas, I., Shi, X., & Delmelle, E. (2022). Detecting space–time patterns of disease risk under dynamic background population. *Journal of geographical systems*, *24*(3), 389-417. https://doi.org/10.1007/s10109-022-00377-7

Hohl, A., Zheng, M., Tang, W., Delmelle, E., & Casas, I. (2017). Spatiotemporal point pattern analysis using Ripley's K function. *Geospatial Data Science: Techniques and Applications. Taylor & Francis*.

Hossain, M. D. M., & Lawson, A. B. (2005). Local likelihood disease clustering: development and evaluation. *Environmental and Ecological Statistics*, *12*(3), 259-273. https://doi.org/10.1007/s10651-005-1512-9

Hossain, M. M., & Lawson, A. B. (2010). Space-time Bayesian small area disease risk models: development and evaluation with a focus on cluster detection. *Environmental and Ecological Statistics*, *17*(1), 73-95. https://doi.org/10.1007/s10651-008-0102-z

Huang, Z. J., Das, A., Qiu, Y. L., & Tatem, A. J. (2012). Web-based GIS: the vector-borne disease airline importation risk (VBD-AIR) tool [Article]. *International Journal of Health Geographics*, *11*, 14, Article 33. https://doi.org/10.1186/1476-072x-11-33

Ivanković, D., Barbazza, E., Bos, V., Fernandes, Ó. B., Gilmore, K. J., Jansen, T., Kara, P., Larrain, N., Lu, S., & Meza-Torres, B. (2021). Features constituting actionable COVID-19 dashboards: descriptive assessment and expert appraisal of 158 public web-based COVID-19 dashboards. *Journal of medical Internet research*, *23*(2), e25682.

Kirby, R. S., Delmelle, E., & Eberth, J. M. (2017). Advances in spatial epidemiology and geographic information systems. *Annals of epidemiology*, *27*(1), 1-9.

Kistemann, T., Dangendorf, F., & Schweikart, J. (2002). New perspectives on the use of Geographical Information Systems (GIS) in environmental health sciences. *International journal of hygiene and environmental health*, *205*(3), 169-181.

Knox, E. G. (1989). Detection of clusters. *Methodology of enquiries into disease clustering. London: Small Area Health Statistics Unit*, *17*, 20.

Knox, E. G., & Bartlett, M. S. (1964). The detection of space-time interactions. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *13*(1), 25-30.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics-Theory and methods*, *26*(6), 1481-1496.

Kulldorff, M. (2010). SaTScan user guide for version 9.0. In.

Kulldorff, M., Athas, W. F., Feurer, E. J., Miller, B. A., & Key, C. R. (1998). Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *American journal of public health*, *88*(9), 1377-1380. https://doi.org/10.2105/ajph.88.9.1377

Kulldorff, M., Heffernan, R., Hartman, J., Assunçao, R., & Mostashari, F. (2005). A space–time permutation scan statistic for disease outbreak detection. *Plos med*, *2*(3), e59.

Kulldorff, M., & Nagarwalla, N. (1995). Spatial disease clusters: detection and inference. *Statistics in medicine*, *14*(8), 799-810.

Kuo, F. Y., Wen, T. H., & Sabel, C. E. (2018). Characterizing Diffusion Dynamics of Disease Clustering: A Modified Space-Time DBSCAN (MST-DBSCAN) Algorithm [Article]. *Annals of the American Association of Geographers*, *108*(4), 1168-1186. https://doi.org/10.1080/24694452.2017.1407630

Kwan, M.-P. (2016). Geographies of health, disease and well-being: recent advances in theory and method.

Lan, Y., Desjardins, M. R., Hohl, A., & Delmelle, E. (2021). Geovisualization of COVID-19: State of the Art and Opportunities. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *56*(1), 2-13.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Lessler, J., Azman, A. S., McKay, H. S., & Moore, S. M. (2017). What is a hotspot anyway? *The American journal of tropical medicine and hygiene*, *96*(6), 1270.

Levine, N. (2013). Crimestat IV: A spatial statistics program for the analysis of crime incident locations, version 4.0. *Ned Levine & Associates: Houston, TX, USA*.

Lippi, C. A., Stewart-Ibarra, A. M., Romero, M., Lowe, R., Mahon, R., Van Meerbeeck, C. J., Rollock, L., Hilaire, M. G.-S., Trotman, A. R., & Holligan, D. (2020). Spatiotemporal tools for emerging and endemic disease hotspots in small areas: An analysis of dengue and chikungunya in Barbados, 2013–2016. *The American journal of tropical medicine and hygiene*, *103*(1), 149.

Lloyd, D., & Dykes, J. (2011). Human-centered approaches in geovisualization design: Investigating multiple methods through a long-term case study. *IEEE transactions on visualization and computer graphics*, *17*(12), 2498-2507.

Lu, H., & Carlin, B. P. (2005). Bayesian Areal Wombling for Geographical Boundary Analysis. *Geographical analysis*, *37*(3), 265-285. https://doi.org/https://doi.org/10.1111/j.1538-4632.2005.00624.x

Luan, H., & Law, J. (2014). Web GIS-Based Public Health Surveillance Systems: A Systematic Review [Review]. *Isprs International Journal of Geo-Information*, *3*(2), 481-506. https://doi.org/10.3390/ijgi3020481

McIntosh, A. I., Jenkins, H. E., White, L. F., Barnard, M., Thomson, D. R., Dolby, T., Simpson, J., Streicher, E. M., Kleinman, M. B., Ragan, E. J., van Helden, P. D., Murray, M. B., Warren, R. M., & Jacobson, K. R. (2018). Using routinely collected laboratory data to identify high rifampicin-resistant tuberculosis burden communities in the Western Cape Province, South Africa: A retrospective spatiotemporal analysis. *Plos med*, *15*(8), e1002638. https://doi.org/10.1371/journal.pmed.1002638

McMichael, A. J. (2004). Environmental and social influences on emerging infectious diseases: past, present and future. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1447), 1049-1058.

Monir Hossain, M., & Lawson, A. B. (2006). Cluster detection diagnostics for small area health data: with reference to evaluation of local likelihood models. *Statistics in Medicine*, *25*(5), 771-786. https://doi.org/https://doi.org/10.1002/sim.2401

Moraga, P. (2017). SpatialEpiApp: A Shiny web application for the analysis of spatial and spatio-temporal disease data. *Spatial and spatio-temporal epidemiology*, *23*, 47-57.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1/2), 17-23.

Murray, J., & Cohen, A. L. (2017). Infectious Disease Surveillance. *International Encyclopedia of Public Health*, 222-229. https://doi.org/10.1016/B978-0-12-803678-5.00517-8

Naish, S., & Tong, S. (2014). Hot spot detection and spatio-temporal dynamics of dengue in Queensland, Australia. ISPRS Technical Commission VIII Symposium,

Newsom, S. (2006). Pioneers in infection control: John Snow, Henry Whitehead, the Broad Street pump, and the beginnings of geographical epidemiology. *Journal of Hospital Infection*, *64*(3), 210-216.

Olson, K. L., Grannis, S. J., & Mandl, K. D. (2006). Privacy protection versus cluster detection in spatial epidemiology. *American journal of public health*, *96*(11), 2002-2008.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical analysis*, *27*(4), 286-306.

Padmanabhan, A., Wang, S. W., Cao, G. F., Hwang, M., Zhang, Z. H., Gao, Y. Z., Soltani, K., & Liu, Y. (2014). FluMapper: A cyberGIS application for interactive analysis of massive location-based social media [Article]. *Concurrency and Computation-Practice & Experience*, *26*(13), 2253-2265. https://doi.org/10.1002/cpe.3287

Pardhan-Ali, A., Berke, O., Wilson, J., Edge, V. L., Furgal, C., Reid-Smith, R., Santos, M., & McEwen, S. A. (2012). A spatial and temporal analysis of notifiable gastrointestinal illness in the Northwest Territories, Canada, 1991-2008. *International Journal of Health Geographics*, *11*(1), 1-10.

Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health services research*, *34*(5 Pt 2), 1189.

Pezanowski, S., MacEachren, A. M., Savelyev, A., & Robinson, A. C. (2018). SensePlace3: a geovisual framework to analyze place–time–attribute information in social media. *Cartography and Geographic Information Science*, *45*(5), 420-437.

Pfeiffer, D., Robinson, T. P., Stevenson, M., Stevens, K. B., Rogers, D. J., & Clements, A. C. (2008). *Spatial analysis in epidemiology* (Vol. 142). Oxford University Press Oxford.

Preim, B., & Lawonn, K. (2020). A survey of visual analytics for public health. Computer Graphics Forum,

Reinhardt, M., Elias, J., Albert, J., Frosch, M., Harmsen, D., & Vogel, U. (2008). EpiScanGIS: an online geographic surveillance system for meningococcal disease. *International journal of health geographics*, *7*(1), 1-7.

Reinhardt, M., Elias, J., Albert, J., Frosch, M., Harmsen, D., & Vogel, U. (2008). EpiScanGIS: an online geographic surveillance system for meningococcal disease [Article]. *International Journal of Health Geographics*, *7*, 7, Article 33. https://doi.org/10.1186/1476-072x-7-33

Rejeki, D. S. S., Fuad, A., Widartono, B. S., Murhandarwati, E. E. H., & Kusnanto, H. (2019). Spatiotemporal patterns of malaria at cross-boundaries area in Menoreh Hills, Java, Indonesia. *Malar J*, *18*(1), 80. https://doi.org/10.1186/s12936-019-2717-y

Rey, S. J., & Anselin, L. (2010). PySAL: A Python library of spatial analytical methods. In *Handbook of applied spatial analysis* (pp. 175-193). Springer.

Ripley, B. D. (1977). Modelling spatial patterns. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(2), 172-192.

Robinson, A. C. (2007). A design framework for exploratory geovisualization in epidemiology. *Information visualization*, *6*(3), 197-214.

Robinson, A. C., Chen, J., Lengerich, E. J., Meyer, H. G., & MacEachren, A. M. (2005). Combining usability techniques to design geovisualization tools for epidemiology. *Cartography and Geographic Information Science*, *32*(4), 243-255.

Robinson, A. C., Peuquet, D. J., Pezanowski, S., Hardisty, F. A., & Swedberg, B. (2017). Design and evaluation of a geovisual analytics system for uncovering patterns in spatio-temporal event data. *Cartography and Geographic Information Science*, *44*(3), 216-228.

Rocheleau, J.-P., Kotchi, S.-O., & Arsenault, J. (2020). Can local risk of West Nile virus infection be predicted from previous cases? A descriptive study in Quebec, 2011–2016. *Canadian Journal of Public Health*, *111*(2), 229-238.

Rogers, D. J., & Randolph, S. E. (2003). Studying the global distribution of infectious diseases using GIS and RS. *Nature Reviews Microbiology*, *1*(3), 231-237. https://doi.org/10.1038/nrmicro776

Rogerson, P., & Yamada, I. (2008). Introductory spatial statistics: description and interence. In *Statistical detection and surveillance of geographic clusters* (pp. 31-34). CRC Press.

Rosillo, N., Del-Aguila-Mejia, J., Rojas-Benedicto, A., Guerrero-Vadillo, M., Penuelas, M., Mazagatos, C., Segu-Tell, J., Ramis, R., & Gomez-Barroso, D. (2021). Real time surveillance of COVID-19 space and time clusters during the summer 2020 in Spain. *BMC Public Health*, *21*(1), Article 961. https://doi.org/10.1186/s12889-021-10961-z

Rotela, C., Lopez, L., Céspedes, M. F., Barbas, G., Lighezzolo, A., Porcasi, X., Lanfri, M. A., Scavuzzo, C. M., & Gorla, D. E. (2017). Analytical report of the 2016 dengue outbreak in Córdoba city, Argentina. *Geospatial health*.

Schabenberger, O., & Gotway, C. A. (2017). Introduction. In *Statistical methods for spatial data analysis* (pp. 26). CRC press.

Sifuna, P., Otieno, L., Andagalu, B., Oyieko, J., Ogutu, B., Singoei, V., Owuoth, J., Ogwang, S., Cowden, J., & Otieno, W. (2018). A Spatiotemporal Analysis of HIV-Associated Mortality in Rural Western Kenya 2011–2015. *Journal of acquired immune deficiency syndromes (1999), 78*(5), 483.

Singh, P. S., & Chaturvedi, H. K. (2021). Temporal variation and geospatial clustering of dengue in Delhi, India 2015-2018. *BMJ Open*, *11*(2), e043848. https://doi.org/10.1136/bmjopen-2020-043848

Sloan, C., Chandrasekhar, R., Mitchel, E., Ndi, D., Miller, L., Thomas, A., Bennett, N. M., Chai, S., Spencer, M., & Eckel, S. (2020). Spatial and temporal clustering of patients hospitalized with laboratory-confirmed influenza in the united states. *Epidemics*, *31*, 100387.

Smith, K. F., Goldberg, M., Rosenthal, S., Carlson, L., Chen, J., Chen, C., & Ramachandran, S. (2014). Global rise in human infectious disease outbreaks. *Journal of the Royal Society Interface*, *11*(101), 20140950.

Sonesson, C., & Bock, D. (2003). A review and discussion of prospective statistical surveillance in public health. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *166*(1), 5-21.

Sugumaran, R., Larson, S. R., & DeGroote, J. P. (2009). Spatio-temporal cluster analysis of county-based human West Nile virus incidence in the continental United States. *International Journal of Health Geographics*, *8*(1), 1-19.

Tadesse, T., Demissie, M., Berhane, Y., Kebede, Y., & Abebe, M. (2013). The clustering of smear-positive tuberculosis in Dabat, Ethiopia: a population based cross sectional study. *Plos One*, *8*(5), e65022.

Tang, J. H., Tseng, T. J., & Chan, T. C. (2019). Detecting spatio-temporal hotspots of scarlet fever in Taiwan with spatio-temporal Gi* statistic. *Plos One*, *14*(4), e0215434. https://doi.org/10.1371/journal.pone.0215434

Taylor, L. H., Latham, S. M., & Woolhouse, M. E. (2001). Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1411), 983-989.

Tran, A., Deparis, X., Dussart, P., Morvan, J., Rabarison, P., Remy, F., Polidori, L., & Gardon, J. (2004). Dengue spatial and temporal patterns, French Guiana, 2001. *Emerging infectious diseases*, *10*(4), 615.

van den Driessche, P., & Watmough, J. (2008). Further Notes on the Basic Reproduction Number. In F. Brauer, P. van den Driessche, & J. Wu (Eds.), *Mathematical Epidemiology* (pp. 159-178). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-78911-6_6

Vazquez-Prokopec, G. M., Kitron, U., Montgomery, B., Horne, P., & Ritchie, S. A. (2010). Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. *PLoS Negl Trop Dis*, *4*(12), e920.

Wagner Filho, J. A., Stuerzlinger, W., & Nedel, L. (2019). Evaluating an immersive space-time cube geovisualization for intuitive trajectory data exploration. *IEEE transactions on visualization and computer graphics*, *26*(1), 514-524.

Weber, M., Alexa, M., & Müller, W. (2001). Visualizing time-series on spirals. Infovis,

Weiss, R. A., & McMichael, A. J. (2004). Social and environmental risk factors in the emergence of infectious diseases. *Nature medicine*, *10*(12), S70-S76.

Wen, T. H., Lin, M. H., & Fang, C. T. (2012). Population Movement and Vector-Borne Disease Transmission: Differentiating Spatial-Temporal Diffusion Patterns of Commuting and Noncommuting Dengue Cases [Article]. *Annals of the Association of American Geographers*, *102*(5), 1026-1037. https://doi.org/10.1080/00045608.2012.671130

Whiteman, A., Desjardins, M. R., Eskildsen, G. A., & Loaiza, J. R. (2019). Detecting space-time clusters of dengue fever in Panama after adjusting for vector surveillance data. *PLoS neglected tropical diseases*, *13*(9), e0007266.

WHO. (2020a). *Coronavirus*. Retrieved 07/17/2020 from https://www.who.int/health-topics/coronavirus#tab=tab_1

WHO. (2020b). *Leading causes of death and disability 2000-2019: A visual summary*. Retrieved 09/24 from https://www.who.int/data/stories/leading-causes-of-death-and-disability-2000-2019-a-visual-summary

WHO. (2020c). *The top 10 causes of death*. Retrieved 09/24 from https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

WHO. (n.d.). *Infectious diseases*. Retrieved 07/01/2021 from http://www.emro.who.int/health-topics/infectious-diseases/index.html

Wichmann, O., Gascon, J., Schunk, M., Puente, S., Siikamaki, H., Gjø, I., Lopez-Velez, R., Clerinx, J., Peyerl-Hoffmann, G., & Sundøy, A. (2007). Severe dengue virus infection in travelers: risk factors and laboratory indicators. *The Journal of infectious diseases*, *195*(8), 1089-1096.

World Health Organization. (2020). *WHO Coronavirus (COVID-19) Dashboard*. Retrieved 10/22/2022 from https://covid19.who.int/

Wu, C., Zhou, M. J., Liu, P. Y., & Yang, M. J. (2021). Analyzing COVID-19 Using Multisource Data: An Integrated Approach of Visualization, Spatial Regression, and Machine Learning [Article]. *Geohealth*, *5*(8), 14, Article e2021GH000439. https://doi.org/10.1029/2021gh000439

Yu, Y., Wu, B., Wu, C., Wang, Q., Hu, D., & Chen, W. (2020). Spatial-temporal analysis of tuberculosis in Chongqing, China 2011-2018. *BMC Infectious Diseases*, *20*(1), 1-12.

Zhang, R., Li, Y., Zhang, A. L., Wang, Y., & Molina, M. J. (2020). Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proceedings of the National Academy of Sciences*, *117*(26), 14857-14863.

APPENDIX A: INCLUDE LITERRATURE OF THE SYSTEMATIC REVIEW

1. Abd Naeeim, N. S., & Abdul Rahman, N. (2021). Spatio-temporal clustering analysis of dengue disease in Peninsular Malaysia. *Journal of Public Health*, 1-11.
2. Acharya, B. K., Cao, C., Lakes, T., Chen, W., & Naeem, S. (2016). Spatiotemporal analysis of dengue fever in Nepal from 2010 to 2014. *Bmc Public Health*, *16*(1), 1-10.
3. Acharya, B. K., Cao, C., Xu, M., Chen, W., & Pandit, S. (2018). Spatiotemporal distribution and geospatial diffusion patterns of 2013 dengue outbreak in Jhapa District, Nepal. *Asia Pacific Journal of Public Health*, *30*(4), 396-405.
4. Ahmed, S., Ersbøll, A. K., Biswas, P., & Christensen, J. P. (2010). The space-time clustering of highly pathogenic avian influenza (HPAI) H5N1 outbreaks in Bangladesh. *Epidemiology & Infection*, *138*(6), 843-852.
5. Akter, R., Naish, S., Gatton, M., Bambrick, H., Hu, W., & Tong, S. (2019). Spatial and temporal analysis of dengue infections in Queensland, Australia: Recent trend and perspectives. *Plos One*, *14*(7), e0220134.
6. Al-Ahmadi, K., Alahmadi, S., & Al-Zahrani, A. (2019). Spatiotemporal clustering of Middle East respiratory syndrome coronavirus (MERS-CoV) incidence in Saudi Arabia, 2012–2019. *International Journal of Environmental Research and Public Health*, *16*(14), 2520.
7. Al-Kindi, K. M., Alkharusi, A., Alshukaili, D., Al Nasiri, N., Al-Awadhi, T., Charabi, Y., & El Kenawy, A. M. (2020). Spatiotemporal assessment of COVID-19 spread over Oman using GIS techniques. *Earth Systems and Environment*, *4*(4), 797-811.
8. Alcântara, E., Mantovani, J., Rotta, L., Park, E., Rodrigues, T., Carvalho, F. C., & Souza Filho, C. R. (2020). Investigating spatiotemporal patterns of the COVID-19 in São Paulo State, Brazil. *Geospatial Health*, *15*(2).
9. Aldstadt, J., Yoon, I. K., Tannitisupawong, D., Jarman, R. G., Thomas, S. J., Gibbons, R. V., Uppapong, A., Iamsirithaworn, S., Rothman, A. L., & Scott, T. W. (2012). Space-time analysis of hospitalised dengue patients in rural Thailand reveals important temporal intervals in the pattern of dengue virus transmission. *Tropical Medicine & International Health*, *17*(9), 1076-1085.
10. Ali, S., Corner, R. J., & Hashizume, M. (2014). Spatiotemporal analysis of dengue infection between 2005 and 2010. In *Dhaka Megacity* (pp. 367-384). Springer.
11. Allan, M., Grandesso, F., Pierre, R., Magloire, R., Coldiron, M., Martinez-Pino, I., Goffeau, T., Gitenet, R., François, G., & Olson, D. (2016). High-resolution spatial analysis of cholera patients reported in Artibonite department, Haiti in 2010–2011. *Epidemics*, *14*, 1-10.
12. Alves, A. T., & Nobre, F. F. (2014). The acquired immunodeficiency syndrome in the state of Rio de Janeiro, Brazil: a spatio-temporal analysis of cases reported in the period 2001-2010. *Geospatial Health*, *8*(2), 437-443.
13. Alves, J. D., Abade, A. S., Peres, W. P., Borges, J. E., Santos, S. M., & Scholze, A. R. (2021). Impact of COVID-19 on the indigenous population of Brazil: a geo-epidemiological study. *Epidemiology & Infection*, *149*.
14. Andrade, L. A., Gomes, D. S., Góes, M. A. d. O., Souza, M. S. F. d., Teixeira, D. C. P., Ribeiro, C. J. N., Alves, J. A. B., Araújo, K. C. G. M. d., & Santos, A. D. d. (2020). Surveillance of the first cases of COVID-19 in Sergipe using a prospective spatiotemporal analysis: the spatial dispersion and its public health implications. *Revista Da Sociedade Brasileira De Medicina Tropical*, *53*.

15. ANGULO, J. J., HAGGETT, P., MEGALE, P., & PEDERNEIRAS, C. A. (1977). VARIOLA MINOR IN BRAGANÇA PAULISTA COUNTY, 1956: A TRENDSURFACE ANALYSIS. *American Journal of Epidemiology*, *105*(3), 272-280.
16. Anis Fuad, D. (2020). Spatiotemporal proximity of rubella cases to the occurrence of congenital rubella syndrome in Yogyakarta, Indonesia. *Med J Malaysia*, *75*, 41.
17. Arauzo-Carod, J.-M., Domènech, A., & Gutiérrez, A. (2021). Do local characteristics act in a similar way for the first two waves of COVID-19? Analysis at intraurban level in Barcelona. *Journal of Public Health*, *43*(3), 455-461.
18. Areias, C., Briz, T., & Nunes, C. (2015). Pulmonary tuberculosis space-time clustering and spatial variation in temporal trends in Portugal, 2000–2010: an updated analysis. *Epidemiology & Infection*, *143*(15), 3211-3219.
19. Ashmore, P., Lindahl, J. F., Colón-González, F. J., Sinh Nam, V., Quang Tan, D., & Medley, G. F. (2020). Spatiotemporal and socioeconomic risk factors for dengue at the province level in Vietnam, 2013–2015: clustering analysis and regression model. *Tropical Medicine and Infectious Disease*, *5*(2), 81.
20. Ashok, S., Ullah, M. Z., Vadivelu, N., Islam, M. T., Nasereddin, S., & Khan, W. Z. (2021). Surveillance of COVID-19 Using Geospatial Data: An Emergency Department Perspective. *Dubai Medical Journal*, *4*(4), 330-338.
21. Atique, S., Chan, T.-C., Chen, C.-C., Hsu, C.-Y., Iqtidar, S., Louis, V. R., Shabbir, S. A., & Chuang, T.-W. (2018). Investigating spatio-temporal distribution and diffusion patterns of the dengue outbreak in Swat, Pakistan. *Journal of infection and public health*, *11*(4), 550-557.
22. Augustijn, E.-W., & Zurita-Milla, R. (2013). Self-organizing maps as an approach to exploring spatiotemporal diffusion patterns. *International Journal of Health Geographics*, *12*(1), 1-14.
23. Azage, M., Kumie, A., Worku, A., C. Bagtzoglou, A., & Anagnostou, E. (2017). Effect of climatic variability on childhood diarrhea and its high risk periods in northwestern parts of Ethiopia. *Plos One*, *12*(10), e0186933.
24. Balaji, D., & Saravanabavan, V. (2020). A geo medical analysis of dengue cases in Madurai city-Tamilnadu India. *Geojournal*, *85*(4), 979-994.
25. Balino, L. V. A., Caasi, K. S., & Addawe, R. C. (2021, 2021//). Spatio-Temporal Distribution of Dengue Infections in Baguio City, Philippines. Modelling, Simulation and Applications of Complex Systems, Singapore.
26. Banu, S., Hu, W., Guo, Y., Naish, S., & Tong, S. (2014). Dynamic spatiotemporal trends of dengue transmission in the Asia-Pacific region, 1955–2004. *Plos One*, *9*(2), e89440.
27. Bauer, D., Higgs, B. W., & Mohtashemi, M. (2007). High performance computing for disease surveillance. NSF Workshop on Intelligence and Security Informatics,
28. Bautista, C. T., Chan, A. S., Ryan, J. R., Calampa, C., Roper, M. H., Hightower, A. W., & Magill, A. J. (2006). Epidemiology and spatial analysis of malaria in the Northern Peruvian Amazon. *The American journal of tropical medicine and hygiene*, *75*(6), 1216-1222.
29. BenBella, D., & Ghosh, D. (2021). Combining geospatial analysis with hiv care continuum to identify differential hiv/aids treatment indicators in uganda. *The Professional Geographer*, *73*(2), 213-229.
30. Benincà, E., van Boven, M., Hagenaars, T., & van der Hoek, W. (2017). Space-time analysis of pneumonia hospitalisations in the Netherlands. *Plos One*, *12*(7), e0180797.

31. Bisanzio, D., Dzul-Manzanilla, F., Gomez-Dantés, H., Pavia-Ruz, N., Hladish, T. J., Lenhart, A., Palacio-Vargas, J., Gonzalez Roldan, J. F., Correa-Morales, F., & Sánchez-Tejeda, G. (2018). Spatio-temporal coherence of dengue, chikungunya and Zika outbreaks in Merida, Mexico. *Plos Neglected Tropical Diseases*, *12*(3), e0006298.
32. Blackburn, J. K., Diamond, U., Kracalik, I. T., Widmer, J., Brown, W., Morrissey, B. D., Alexander, K. A., Curtis, A. J., Ali, A., & Morris Jr, J. G. (2014). Household-level spatiotemporal patterns of incidence of cholera, Haiti, 2011. *Emerging Infectious Diseases*, *20*(9), 1516.
33. Blake, A., Keita, V. S., Sauvageot, D., Saliou, M., Njanpop, B. M., Sory, F., Sudre, B., Lamine, K., Mengel, M., & Gessner, B. D. (2018). Temporo-spatial dynamics and behavioural patterns of 2012 cholera epidemic in the African mega-city of Conakry, Guinea. *Infectious Diseases of Poverty*, *7*(1), 1-10.
34. Boitrago, G. M., Mônica, R. B., Silva, D. M., Cerroni, M. d. P., Cortez-Escalante, J. J., Almiron, M., Terabe, S. H., & Rocha, T. A. H. (2021). Restructuring of emergency services to COVID-19 in Brazil: a space-time analysis from February to August 2020. *Epidemiologia E Servicos De Saude*, *30*.
35. Bruce, A. T. I., Berra, T. Z., Dos Santos, F. L., Alves, Y. M., Souza, L. L. L., Ramos, A. C. V., Arroyo, L. H., de Almeida Crispim, J., Pinto, I. C., & Palha, P. F. (2020). Temporal trends in areas at risk for concomitant tuberculosis in a hyperendemic municipality in the Amazon region of Brazil. *Infectious Diseases of Poverty*, *9*(1), 1-14.
36. Cambou, M. C., Saad, E., McBride, K., Fuller, T., Swayze, E., & Nielsen-Saines, K. (2021). Maternal HIV and syphilis are not syndemic in Brazil: Hot spot analysis of the two epidemics. *Plos One*, *16*(8), e0255590.
37. Canal, M. R., Ferreira, E. R. d. S., Estofolete, C. F., Dias, A. M., Tukasan, C., Bertoque, A. C., Muniz, V. D., Nogueira, M. L., & Silva, N. S. d. (2018). Spatiotemporal-based clusters as a method for dengue surveillance. *Revista Panamericana de Salud Pública*, *41*, e162.
38. Cao, C., Xu, M., Chen, J., Wu, Y., & Xie, X. (2010). Space-time scan statistic based early warning of H1N1 influenza a in Shenzhen, China. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, *28*(8).
39. Carrasco-Escobar, G., Schwalb, A., Tello-Lizarraga, K., Vega-Guerovich, P., & Ugarte-Gil, C. (2020). Spatio-temporal co-occurrence of hotspots of tuberculosis, poverty and air pollution in Lima, Peru. *Infectious Diseases of Poverty*, *9*(1), 1-6.
40. Carrel, M., Emch, M., Streatfield, P. K., & Yunus, M. (2009). Spatio-temporal clustering of cholera: The impact of flood control in Matlab, Bangladesh, 1983–2003. *Health & Place*, *15*(3), 771-782.
41. Carvalho, R. M. d., & Nascimento, L. F. C. (2014). Space-time description of dengue outbreaks in Cruzeiro, Sao Paulo, in 2006 and 2011. *Revista da Associação Médica Brasileira*, *60*, 565-570.
42. Castro, M. C., Kim, S., Barberia, L., Ribeiro, A. F., Gurzenda, S., Ribeiro, K. B., Abbott, E., Blossom, J., Rache, B., & Singer, B. H. (2021). Spatiotemporal pattern of COVID-19 spread in Brazil. *Science*, *372*(6544), 821-826.
43. Chadee, D. D., & Kitron, U. (1999). Spatial and temporal patterns of imported malaria cases and local transmission in Trinidad. *Am J Trop Med Hyg*, *61*(4), 513-517. https://doi.org/10.4269/ajtmh.1999.61.513

44. Chen, J., Qiu, Y., Yang, R., Li, L., Hou, J., Lu, K., & Xu, L. (2019). The characteristics of spatial-temporal distribution and cluster of tuberculosis in Yunnan Province, China, 2005–2018. *Bmc Public Health*, *19*(1), 1-13.
45. Chen, Q., Lai, S., Yin, W., Zhou, H., Li, Y., Mu, D., Li, Z., Yu, H., & Yang, W. (2016). Epidemic characteristics, high-risk townships and space-time clusters of human brucellosis in Shanxi Province of China, 2005–2014. *Bmc Infectious Diseases*, *16*(1), 1-10.
46. Cheng, Y.-J., Norris, J., Bao, C.-J., Liang, Q., Hu, J.-L., Wu, Y., Tang, F.-Y., Liu, W.-D., Ding, K.-Q., & Zhao, Y. (2012). Geographical information systems-based spatial analysis and implications for syphilis interventions in Jiangsu province, People's Republic of China. *Geospatial Health*, *7*(1), 63-72.
47. Cherry, C. B., Griffin, M. R., Edwards, K. M., Williams, J. V., Gil, A. I., Verastegui, H., Lanata, C. F., & Grijalva, C. G. (2016). Spatial and Temporal Spread of Acute Viral Respiratory Infections in Young Children Living in High-altitude Rural Communities: A Prospective Household-based Study. *Pediatr Infect Dis J*, *35*(10), 1057-1061. https://doi.org/10.1097/inf.0000000000001234
48. Chin, W.-C.-B., Wen, T.-H., Sabel, C. E., & Wang, I. (2017). A geo-computational algorithm for exploring the structure of diffusion progression in time and space. *Scientific Reports*, *7*(1), 1-13.
49. Chiu, Y.-W., Hsu, C. E., Wang, M.-Q., & Nkhoma, E. T. (2008). Examining geographic and temporal variations of AIDS mortality: evidence of racial disparities. *Journal of the National Medical Association*, *100*(7), 788-796.
50. Choi, K.-M., Yu, H.-L., & Wilson, M. L. (2008). Spatiotemporal statistical analysis of influenza mortality risk in the State of California during the period 1997–2001. *Stochastic Environmental Research and Risk Assessment*, *22*(1), 15-25.
51. Chuang, T.-W., Ng, K.-C., Nguyen, T. L., & Chaves, L. F. (2018). Epidemiological characteristics and space-time analysis of the 2015 dengue outbreak in the metropolitan region of Tainan City, Taiwan. *International Journal of Environmental Research and Public Health*, *15*(3), 396.
52. Cos Guerra, O. d., Castillo Salcines, V., & Cantarero Prieto, D. (2021). Data mining and socio-spatial patterns of COVID-19: geo-prevention keys for tackling the pandemic.
53. Costa, M. A., & Kulldorff, M. (2014). Maximum linkage space-time permutation scan statistics for disease outbreak detection. *International Journal of Health Geographics*, *13*(1), 1-14.
54. Coura-Vital, W., Cardoso, D. T., Ker, F. T. d. O., Magalhães, F. d. C., Bezerra, J. M. T., Viegas, A. M., Morais, M. H. F., Bastos, L. S., Reis, I. A., & Carneiro, M. (2021). Spatiotemporal dynamics and risk estimates of COVID-19 epidemic in Minas Gerais State: analysis of an expanding process. *Revista Do Instituto De Medicina Tropical De Sao Paulo*, *63*.
55. Craig, M. (1988). Time-space clustering of Vibrio cholerae 01 in Matlab, Bangladesh, 1970–1982. *Social science & medicine*, *26*(1), 5-13.
56. Crighton, E. J., Elliott, S. J., Kanaroglou, P., Moineddin, R., & Upshur, R. E. (2008). Spatio-temporal analysis of pneumonia and influenza hospitalizations in Ontario, Canada. *Geospatial Health*, *2*(2), 191-202.
57. Cuadros, D. F., & Abu-Raddad, L. J. (2014). Spatial variability in HIV prevalence declines in several countries in sub-Saharan Africa. *Health & Place*, *28*, 45-49. https://doi.org/https://doi.org/10.1016/j.healthplace.2014.03.007

58. Cui, Z., Lin, D., Chongsuvivatwong, V., Zhao, J., Lin, M., Ou, J., & Zhao, J. (2019). Spatiotemporal patterns and ecological factors of tuberculosis notification: A spatial panel data analysis in Guangxi, China. *Plos One*, *14*(5), e0212051.
59. Dangisso, M. H., Datiko, D. G., & Lindtjørn, B. (2015). Spatio-temporal analysis of smear-positive tuberculosis in the Sidama Zone, southern Ethiopia. *Plos One*, *10*(6), e0126369.
60. Dass, S., Ngui, R., Gill, B. S., Chan, Y. F., Wan Sulaiman, W. Y., Lim, Y. A. L., Mudin, R. N., Chong, C. K., Sulaiman, L. H., & Sam, I.-C. (2021). Spatiotemporal spread of chikungunya virus in Sarawak, Malaysia. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *115*(8), 922-931.
61. Daw, M. A., Buktir Ali, L. A., Daw, A. M., Sifennasr, N. E., Dau, A. A., Agnan, M. M., & El-Bouzedi, A. (2018). The geographic variation and spatiotemporal distribution of hepatitis C virus infection in Libya: 2007–2016. *Bmc Infectious Diseases*, *18*(1), 1-10.
62. Daw, M. A., Daw, A. M., Sifennasr, N. E., Draha, A. M., Daw, A. A., Daw, A. A., Ahmed, M. O., Mokhtar, E. S., El-Bouzedi, A. H., & Daw, I. M. (2019). Spatiotemporal analysis and epidemiological characterization of the human immunodeficiency virus (HIV) in Libya within a twenty five year period: 1993–2017. *Aids Research and Therapy*, *16*(1), 1-9.
63. de Araújo Morais, L. R., & da Silva Gomes, G. S. (2021). Applying Spatio-temporal Scan Statistics and Spatial Autocorrelation Statistics to identify Covid-19 clusters in the world-A Vaccination Strategy? *Spatial and Spatio-Temporal Epidemiology*, *39*, 100461.
64. de Azevedo, T. S., Lorenz, C., & Chiaravalloti-Neto, F. (2020). Spatiotemporal evolution of dengue outbreaks in Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *114*(8), 593-602.
65. De Ridder, D., Loizeau, A. J., Sandoval, J. L., Ehrler, F., Perrier, M., Ritch, A., Violot, G., Santolini, M., Tzovaras, B. G., & Stringhini, S. (2021). Detection of Spatiotemporal Clusters of COVID-19–Associated Symptoms and Prevention Using a Participatory Surveillance App: Protocol for the@ choum Study. *Jmir Research Protocols*, *10*(10), e30444.
66. Delmelle, E., Casas, I., Rojas, J. H., & Varela, A. (2013). Spatio-temporal patterns of dengue fever in Cali, Colombia. *International Journal of Applied Geospatial Research (IJAGR)*, *4*(4), 58-75.
67. Delmelle, E., Dony, C., Casas, I., Jia, M., & Tang, W. (2014). Visualizing the impact of space-time uncertainties on dengue fever patterns. *International Journal of Geographical Information Science*, *28*(5), 1107-1127.
68. Desjardins, M., Whiteman, A., Casas, I., & Delmelle, E. (2018). Space-time clusters and co-occurrence of chikungunya and dengue fever in Colombia from 2015 to 2016. *Acta Tropica*, *185*, 77-85.
69. Desjardins, M. R., Hohl, A., & Delmelle, E. M. (2020). Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Applied Geography*, *118*, 102202.
70. Dhewantara, P. W., Prasetyowati, H., Ridwan, W., & Hakim, L. (2020). The application of spatiotemporal scan statistics to detect high-risk clusters for dengue fever in Jakarta, Indonesia. American Institute of Physics Conference Series,
71. Dionne, C. G., Sullivan, A. B., Norwood, T. A., MSA, Þ., Serre, M. L., & Miller, W. C. (2012). Does Core Area Theory Apply to Sexually Transmitted Diseases in Rural Environments?

72. Dismer, A. M., Lemoine, J. F., Baptiste, M. J., Mace, K. E., Impoinvil, D. E., Eng, J. V., & Chang, M. A. (2021). Detecting Malaria Hotspots in Haiti, a Low-Transmission Setting. *The American Journal of Tropical Medicine and Hygiene*, *104*(6), 2108.
73. Duczmal, L. H., Moreira, G. J., Burgarelli, D., Takahashi, R. H., Magalhães, F. C., & Bodevan, E. C. (2011). Voronoi distance based prospective space-time scans for point data sets: a dengue fever cluster analysis in a southeast Brazilian town. *International Journal of Health Geographics*, *10*(1), 1-14.
74. e Silva, M. d. A., Oliveira, C. D. L., Neto, R. G. T., & Camargos, P. A. (2016). Spatial distribution of tuberculosis from 2002 to 2012 in a midsize city in Brazil. *Bmc Public Health*, *16*(1), 1-8.
75. Edens, C., Alden, N. B., Danila, R. N., Fill, M.-M. A., Gacek, P., Muse, A., Parker, E., Poissant, T., Ryan, P. A., & Smelser, C. (2019). Multistate analysis of prospective Legionnaires’ disease cluster detection using SaTScan, 2011–2015. *Plos One*, *14*(5), e0217632.
76. Elias, J., Harmsen, D., Claus, H., Hellenbrand, W., Frosch, M., & Vogel, U. (2006). Spatiotemporal analysis of invasive meningococcal disease, Germany. *Emerging Infectious Diseases*, *12*(11), 1689.
77. Elson, R., Davies, T. M., Lake, I. R., Vivancos, R., Blomquist, P. B., Charlett, A., & Dabrera, G. (2021). The spatio-temporal distribution of COVID-19 infection in England between January and June 2020. *Epidemiology & Infection*, *149*.
78. Fang, L. Q., De Vlas, S. J., Feng, D., Liang, S., Xu, Y. F., Zhou, J. P., Richardus, J. H., & Cao, W. C. (2009). Geographical spread of SARS in mainland China. *Tropical Medicine & International Health*, *14*, 14-20.
79. Fariza, A., & Astuti, D. W. (2021). Spatial-Temporal Visualization of Dengue Haemorrhagic Fever Vulnerability in Kediri District, Indonesia, Using K-means Algorithm. 2021 International Conference on Data and Software Engineering (ICoDSE),
80. Freitas, L. P., Cruz, O. G., Lowe, R., & Sá Carvalho, M. (2019). Space-time dynamics of a triple epidemic: dengue, chikungunya and Zika clusters in the city of Rio de Janeiro. *Proceedings of the Royal Society B*, *286*(1912), 20191867.
81. Fuentes-Vallejo, M. (2017). Space and space-time distributions of dengue in a hyper-endemic urban space: the case of Girardot, Colombia. *Bmc Infectious Diseases*, *17*(1), 1-16.
82. Gaudart, J., Poudiougou, B., Dicko, A., Ranque, S., Toure, O., Sagara, I., Diallo, M., Diawara, S., Ouattara, A., & Diakite, M. (2006). Space-time clustering of childhood malaria at the household level: a dynamic cohort in a Mali village. *Bmc Public Health*, *6*(1), 1-13.
83. Gaudart, J., Rebaudet, S., Barrais, R., Boncy, J., Faucher, B., Piarroux, M., Magloire, R., Thimothe, G., & Piarroux, R. (2013). Spatio-temporal dynamics of cholera during the first year of the epidemic in Haiti. *Plos Neglected Tropical Diseases*, *7*(4), e2145.
84. Ge, E., Zhang, X., Wang, X., & Wei, X. (2016). Spatial and temporal analysis of tuberculosis in Zhejiang Province, China, 2009-2012. *Infectious Diseases of Poverty*, *5*(1), 1-10.
85. Ge, L., Zhao, Y., Zhou, K., Mu, X., Yu, H., Wang, Y., Wang, N., Fan, H., Guo, L., & Huo, X. (2016). Spatio-temporal pattern and influencing factors of hemorrhagic fever with renal syndrome (HFRS) in Hubei Province (China) between 2005 and 2014. *Plos One*, *11*(12), e0167836.
86. Ghosh, P., & Cartone, A. (2020). A Spatio-temporal analysis of COVID-19 outbreak in Italy. *Regional Science Policy & Practice*, *12*(6), 1047-1062.

87. Gómez-Barroso, D., Rodriguez-Valín, E., Ramis, R., & Cano, R. (2013). Spatio-temporal analysis of tuberculosis in Spain, 2008–2010. *The International journal of tuberculosis and lung disease*, *17*(6), 745-751.
88. Gomez-Barroso, D., Varela, C., Ramis, R., Del Barrio, J., & Simon, F. (2012). Space-time pattern of hepatitis A in Spain, 1997–2007. *Epidemiology & Infection*, *140*(3), 407-416.
89. Greene, S. K., Ionides, E. L., & Wilson, M. L. (2006). Patterns of influenza-associated mortality among US elderly by geographic region and virus subtype, 1968–1998. *American Journal of Epidemiology*, *163*(4), 316-326.
90. Greene, S. K., Kulldorff, M., Huang, J., Brand, R. J., Kleinman, K. P., Hsu, J., & Platt, R. (2011). Timely detection of localized excess influenza activity in Northern California across patient care, prescription, and laboratory data. *Statistics in Medicine*, *30*(5), 549-559.
91. Greene, S. K., Peterson, E. R., Balan, D., Jones, L., Culp, G. M., Fine, A. D., & Kulldorff, M. (2021). Detecting COVID-19 Clusters at High Spatiotemporal Resolution, New York City, New York, USA, June–July 2020. *Emerging Infectious Diseases*, *27*(5), 1500.
92. Greene, S. K., Peterson, E. R., Kapell, D., Fine, A. D., & Kulldorff, M. (2016). Daily reportable disease spatiotemporal cluster detection, New York City, New York, USA, 2014–2015. *Emerging Infectious Diseases*, *22*(10), 1808.
93. Greene, S. K., Schmidt, M. A., Stobierski, M. G., & Wilson, M. L. (2005). Spatio-temporal pattern of viral meningitis in Michigan, 1993-2001. *Journal of Geographical Systems*, *7*(1), 85-99. https://doi.org/10.1007/s10109-005-0151-x
94. Grubesic, T. H., Nelson, J. R., Wallace, D., Eason, J., Towers, S., & Walker, J. (2021). Geodemographic insights on the COVID-19 pandemic in the State of Wisconsin and the role of risky facilities. *Geojournal*, 1-23.
95. Guan, Z., Chen, C., Huang, C., Zhang, H., Zhou, Y., Zhou, Y., Wu, J., Zhou, Z., Yang, S., & Li, L. (2021). Epidemiological features and spatial–temporal distribution of visceral leishmaniasis in mainland China: a population-based surveillance study from 2004 to 2019. *Parasites & Vectors*, *14*(1), 1-11.
96. Güemes, A., Ray, S., Aboumerhi, K., Desjardins, M. R., Kvit, A., Corrigan, A. E., Fries, B., Shields, T., Stevens, R. D., & Curriero, F. C. (2021). A syndromic surveillance tool to detect anomalous clusters of COVID-19 symptoms in the United States. *Scientific Reports*, *11*(1), 1-11.
97. Gui, J., Liu, Z., Zhang, T., Hua, Q., Jiang, Z., Chen, B., Gu, H., Lv, H., & Dong, C. (2015). Epidemiological characteristics and spatial-temporal clusters of hand, foot, and mouth disease in Zhejiang Province, China, 2008-2012. *Plos One*, *10*(9), e0139109.
98. Guis, H., Clerc, S., Hoen, B., & Viel, J.-F. (2006). Clusters of autochthonous hepatitis A cases in a low endemicity area. *Epidemiology & Infection*, *134*(3), 498-505.
99. Gurjav, U., Burneebaatar, B., Narmandakh, E., Tumenbayar, O., Ochirbat, B., Hill-Cawthorne, G., Marais, B., & Sintchenko, V. (2015). Spatiotemporal evidence for cross-border spread of MDR-TB along the Trans-Siberian Railway line. *The International Journal of Tuberculosis and Lung Disease*, *19*(11), 1376-1382.
100. Gurjav, U., Jelfs, P., Hill-Cawthorne, G. A., Marais, B. J., & Sintchenko, V. (2016). Genotype heterogeneity of Mycobacterium tuberculosis within geospatial hotspots suggests foci of imported infection in Sydney, Australia. *Infection, Genetics and Evolution*, *40*, 346-351.

101. Hafeez, S., Amin, M., & Munir, B. A. (2017). Spatial mapping of temporal risk to improve prevention measures: A case study of dengue epidemic in Lahore. *Spatial and Spatio-Temporal Epidemiology*, *21*, 77-85.

102. Hass, F. S., & Jokar Arsanjani, J. (2021). The geography of the COVID-19 pandemic: A data-driven approach to exploring geographical driving forces. *International Journal of Environmental Research and Public Health*, *18*(6), 2803.

103. Hernández-Gaytán, S. I., Díaz-Vásquez, F. J., Duran-Arenas, L. G., Cervantes, M. L., & Rothenberg, S. J. (2017). 20 years spatial-temporal analysis of dengue fever and hemorrhagic fever in Mexico. *Archives of Medical Research*, *48*(7), 653-662.

104. Hervind, & Widyaningsih, Y. (2017). Dengue hemorrhagic fever and typhoid fever association based on spatial standpoint using scan statistics in DKI Jakarta. AIP Conference Proceedings.

105. Hiwat, H., Martínez-López, B., Cairo, H., Hardjopawiro, L., Boerleider, A., Duarte, E. C., & Yadon, Z. E. (2018). Malaria epidemiology in Suriname from 2000 to 2016: trends, opportunities and challenges for elimination. *Malaria Journal*, *17*(1), 1-13.

106. Hohl, A., & Chen, P. (2019). Spatiotemporal simulation: local Ripley's K function parameterizes adaptive kernel density estimation. Proceedings of the 2nd ACM SIGSPATIAL International Workshop on GeoSpatial Simulation,

107. Hohl, A., Delmelle, E., Tang, W., & Casas, I. (2016). Accelerating the discovery of space-time patterns of infectious diseases using parallel computing. *Spatial and Spatio-Temporal Epidemiology*, *19*, 10-20.

108. Hohl, A., Delmelle, E. M., Desjardins, M. R., & Lan, Y. (2020). Daily surveillance of COVID-19 using the prospective space-time scan statistic in the United States. *Spatial and Spatio-Temporal Epidemiology*, *34*, 100354.

109. Hohl, A., Delmelle, E. M., & Tang, W. (2015). Spatiotemporal domain decomposition for massive parallel computation of space-time kernel density. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*(4), 7.

110. Hohl, A., Zheng, M., Tang, W., Delmelle, E., & Casas, I. (2017). Spatiotemporal point pattern analysis using Ripley’s K function. *Geospatial Data science techniques and applications*, 155-176.

111. Horwood, P. F., Karl, S., Mueller, I., Jonduo, M. H., Pavlin, B. I., Dagina, R., Ropa, B., Bieb, S., Rosewell, A., & Umezaki, M. (2014). Spatio-temporal epidemiology of the cholera outbreak in Papua New Guinea, 2009–2011. *Bmc Infectious Diseases*, *14*(1), 1-10.

112. Hosein, S., Al-Tahir, R., & Ramlal, B. (2013). *Spatiotemporal analysis of dengue hemorrhagic fever and dengue shock syndrome incidence within Trinidad, West Indies* Proceedings of the Second ACM SIGSPATIAL International Workshop on the Use of GIS in Public Health, Orlando, Florida. https://doi.org/10.1145/2535708.2535710

113. Hu, B., Gong, J., Zhou, J., Sun, J., Yang, L., Xia, Y., & Ibrahim, A. N. (2013). Spatial-temporal characteristics of epidemic spread in-out flow—Using SARS epidemic in Beijing as a case study. *Science China Earth Sciences*, *56*(8), 1380-1397.

114. Hu, W., Clements, A., Williams, G., & Tong, S. (2011). Spatial analysis of notified dengue fever infections. *Epidemiology & Infection*, *139*(3), 391-399.

115. Huang, J., Kwan, M.-P., & Kan, Z. (2021). The superspreading places of COVID-19 and the associated built-environment and socio-demographic features: A study using a spatial network framework and individual-level activity data. *Health & Place*, *72*, 102694.

116. Huang, L., Li, X.-X., Abe, E. M., Xu, L., Ruan, Y., Cao, C.-L., & Li, S.-Z. (2017). Spatial-temporal analysis of pulmonary tuberculosis in the northeast of the Yunnan province, People's Republic of China. *Infectious Diseases of Poverty*, *6*(02), 57-67.

117. Huang, L., Michael Abe, E., Li, X.-X., Bergquist, R., Xu, L., Xue, J.-B., Ruan, Y., Cao, C.-L., & Li, S.-Z. (2018). Space-time clustering and associated risk factors of pulmonary tuberculosis in southwest China. *Infectious Diseases of Poverty*, *7*(04), 63-73.

118. Huang, Q., Hu, L., Liao, Q.-b., Xia, J., Wang, Q.-r., & Peng, H.-J. (2017). Spatiotemporal analysis of the malaria epidemic in Mainland China, 2004–2014. *The American journal of tropical medicine and hygiene*, *97*(2), 504.

119. Huang, Q., Jackson, S., Derakhshan, S., Lee, L., Pham, E., Jackson, A., & Cutter, S. L. (2021). Urban-rural differences in COVID-19 exposures and outcomes in the South: A preliminary analysis of South Carolina. *Plos One*, *16*(2), e0246548. https://doi.org/10.1371/journal.pone.0246548

120. Huang, Z. (2021). Spatiotemporal Evolution Patterns of the COVID-19 Pandemic Using Space-Time Aggregation and Spatial Statistics: A Global Perspective. *ISPRS International Journal of Geo-Information*, *10*(8), 519.

121. Hui, F.-M., Xu, B., Chen, Z.-W., Cheng, X., Liang, L., Huang, H.-B., Fang, L.-Q., Yang, H., Zhou, H.-N., & Yang, H.-L. (2009). Spatio-temporal distribution of malaria in Yunnan Province, China. *Am J Trop Med Hyg*, *81*(3), 503-509.

122. Hundessa, S. H., Williams, G., Li, S., Guo, J., Chen, L., Zhang, W., & Guo, Y. (2016). Spatial and space-time distribution of Plasmodium vivax and Plasmodium falciparum malaria in China, 2005–2014. *Malaria Journal*, *15*(1), 1-11.

123. Hussain, S., Mubeen, M., Ahmad, A., Fahad, S., Nasim, W., Hammad, H. M., Shah, G. M., Murtaza, B., Tahir, M., & Parveen, S. (2021). Using space-time scan statistic for studying the effects of COVID-19 in Punjab, Pakistan: a guideline for policy measures in regional agriculture. *Environmental Science and Pollution Research*, 1-14.

124. Jalilian, F. A., Parvin, M., Olfatifar, M., Erfani, H., & Bathaei, J. (2018). The critical role of injecting drug users on the spatial distribution of hepatitis C virus; a study in the West of Iran. *Gastroenterology and Hepatology from Bed to Bench*, *11*(Suppl 1), S129.

125. Jat, M. K., & Mala, S. (2017). *Application of GIS and Space-Time Scan Statistic for Vector Born Disease Clustering* Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance, New Delhi AA, India. https://doi.org/10.1145/3047273.3047361

126. Jaya, I., Andriyana, Y., Tantular, B., & Ruchjana, B. (2019). Spatiotemporal Dengue Disease Clustering by Means Local Spatiotemporal Moran's Index. IOP Conference Series: Materials Science and Engineering,

127. Jeefoo, P., Tripathi, N. K., & Souris, M. (2011). Spatio-temporal diffusion pattern and hotspot detection of dengue in Chachoengsao province, Thailand. *International Journal of Environmental Research and Public Health*, *8*(1), 51-74.

128. Jiang, L., Li, Z., Huang, J., Liu, B., Yang, Y., Lin, L., Wang, C., Xie, X., Peng, X., & Xu, W. (2019). Spatiotemporal clusters of HIV/AIDS infections caused by drug use and heterosexual contact in Ruili city, China 1989–2016. *Bmc Infectious Diseases*, *19*(1), 1-9.

129. Jones, S. G., & Kulldorff, M. (2012). Influence of Spatial Resolution on Space-Time Disease Cluster Detection. *Plos One*, *7*(10), e48036. https://doi.org/10.1371/journal.pone.0048036

130. Kammerer, J. S., Shang, N., Althomsons, S. P., Haddad, M. B., Grant, J., & Navin, T. R. (2013). Using statistical methods and genotyping to detect tuberculosis outbreaks. *International Journal of Health Geographics*, *12*(1), 1-8.
131. Kan, C.-C., Lee, P.-F., Wen, T.-H., Chao, D.-Y., Wu, M.-H., Lin, N. H., Huang, S. Y.-J., Shang, C.-S., Fan, I.-C., & Shu, P.-Y. (2008). Two clustering diffusion patterns identified from the 2001–2003 dengue epidemic, Kaohsiung, Taiwan. *The American journal of tropical medicine and hygiene*, *79*(3), 344-352.
132. Kan, Z., Kwan, M.-P., Wong, M. S., Huang, J., & Liu, D. (2021). Identifying the space-time patterns of COVID-19 risk and their associations with different built environment features in Hong Kong. *Science of the Total Environment*, *772*, 145379.
133. Kan, Z., Kwan, M. P., Huang, J., Wong, M. S., & Liu, D. (2021). Comparing the space-time patterns of high-risk areas in different waves of COVID-19 in Hong Kong. *Transactions in Gis*, *25*(6), 2982-3001.
134. Kang, D., Choi, H., Kim, J.-H., & Choi, J. (2020). Spatial epidemic dynamics of the COVID-19 outbreak in China. *International Journal of Infectious Diseases*, *94*, 96-102.
135. Kayembe, H. C. N., Linard, C., Bompangue, D., Muwonga, J., Moutschen, M., Situakibanza, H., & Ozer, P. (2021). The spread of cholera in western Democratic Republic of the Congo is not unidirectional from East–West: a spatiotemporal analysis, 1973–2018. *Bmc Infectious Diseases*, *21*(1), 1-11.
136. Kazazian, L., Lima Neto, A. S., Sousa, G. S., Nascimento, O. J. d., & Castro, M. C. (2020). Spatiotemporal transmission dynamics of co-circulating dengue, Zika, and chikungunya viruses in Fortaleza, Brazil: 2011–2017. *Plos Neglected Tropical Diseases*, *14*(10), e0008760.
137. Kejžar, N., & Lusa, L. (2020). Classification of weekly provincial overall age-and gender-specific mortality patterns during the COVID-19 epidemics in Italy. *Epidemiologia e prevenzione*, *44*(5-6 Suppl 2), 271-281.
138. Khademi, N., Reshadat, S., Zangeneh, A., Saeidi, S., Ghasemi, S., Rajabi-Gilan, N., & Zakiei, A. (2017). A comparative study of the spatial distribution of HIV prevalence in the metropolis of Kermanshah, Iran, in 1996− 2014 using geographical information systems. *Hiv Medicine*, *18*(3), 220-224.
139. Khalique, F., Shaheen, R., & Khan, S. A. (2020). Spatio-Temporal Investigations of Dengue Fever in Pakistan Through an HL7 Based Public Health Framework for Hotspot Analysis. *IEEE Access*, *8*, 199980-199994. https://doi.org/10.1109/ACCESS.2020.3027234
140. Kiani, B., Raouf Rahmati, A., Bergquist, R., Hashtarkhani, S., Firouraghi, N., Bagheri, N., Moghaddas, E., & Mohammadi, A. (2021). Spatio-temporal epidemiology of the tuberculosis incidence rate in Iran 2008 to 2018. *Bmc Public Health*, *21*(1), 1-20.
141. Kim, S., Kim, M., Lee, S., & Lee, Y. J. (2021). Discovering spatiotemporal patterns of COVID-19 pandemic in South Korea. *Scientific Reports*, *11*(1), 1-17.
142. Kim, Y. H., Ahn, H. J., Kim, D., Hong, S. J., Kim, T. S., & Nam, H. W. (2021). Recent Spatial and Temporal Trends of Malaria in Korea. *Korean J Parasitol*, *59*(6), 585-593. https://doi.org/10.3347/kjp.2021.59.6.585
143. Kleinman, K., Abrams, A., Katherine Yih, W., Platt, R., & Kulldorff, M. (2006). Evaluating spatial surveillance: detection of known outbreaks in real data. *Statistics in Medicine*, *25*(5), 755-769.

144. Kulldorff, M., Heffernan, R., Hartman, J., Assuncao, R., & Mostashari, F. (2005). A space-time permutation scan statistic for disease outbreak detection [Article]. *Plos Medicine*, *2*(3), 216-224, Article e59. https://doi.org/10.1371/journal.pmed.0020059

145. Kuo, F.-Y., Wen, T.-H., & Sabel, C. E. (2018). Characterizing diffusion dynamics of disease clustering: a modified space-time DBSCAN (MST-DBSCAN) algorithm. *Annals of the American Association of Geographers*, *108*(4), 1168-1186.

146. Ladoy, A., Opota, O., Carron, P.-N., Guessous, I., Vuilleumier, S., Joost, S., & Greub, G. (2021). Size and duration of COVID-19 clusters go along with a high SARS-CoV-2 viral load: A spatio-temporal investigation in Vaud state, Switzerland. *Science of the Total Environment*, *787*, 147483.

147. Lai, W. T., Chen, C. H., Hung, H., Chen, R. B., Shete, S., & Wu, C. C. (2018). Recognizing spatial and temporal clustering patterns of dengue outbreaks in Taiwan. *BMC Infect Dis*, *18*(1), 256. https://doi.org/10.1186/s12879-018-3159-9

148. Lan, Y., Desjardins, M. R., Hohl, A., & Delmelle, E. (2021). Geovisualization of COVID-19: State of the Art and Opportunities. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *56*(1), 2-13.

149. Lazarus, N. W. (2021). Multivariate analysis of the dengue virus in Sri Lanka using the ordination method. *Geojournal*, *86*(1), 281-302. https://doi.org/10.1007/s10708-019-10069-3

150. Le Viet, T., Choisy, M., Bryant, J. E., Vu Trong, D., Pham Quang, T., Horby, P., Nguyen Tran, H., Tran Thi Kieu, H., Nguyen Vu, T., & Nguyen Van, K. (2015). A dengue outbreak on a floating village at Cat Ba Island in Vietnam. *Bmc Public Health*, *15*(1), 1-8.

151. Leal, P. R., de Paula, R. J., Guimarães, S., & Kampel, M. (2021). Sociodemographic and spatiotemporal profiles of hepatitis-A in the state of Pará, Brazil, based on reported notified cases. *Geospatial Health*, *16*(2).

152. Lee, S. S., & Wong, N. S. (2011). The clustering and transmission dynamics of pandemic influenza A (H1N1) 2009 cases in Hong Kong. *Journal of Infection*, *63*(4), 274-280.

153. Leveau, C. M., Uez, O., & Vacchino, M. N. (2015). Spatiotemporal trends of cases of pandemic influenza A (H1N1) pdm09 in Argentina, 2009-2012. *Revista Do Instituto De Medicina Tropical De Sao Paulo*, *57*, 133-138.

154. Li, F., He, F., Sun, J., Zhai, Y., Jiang, J., & Lin, J. (2019). Spatial and temporal analysis of severe fever with thrombocytopenia syndrome in Zhejiang Province, China, 2011-2015. *The Journal of Infection in Developing Countries*, *13*(01), 35-43.

155. Li, J., Kolivras, K. N., Hong, Y., Duan, Y., Seukep, S. E., Prisley, S. P., Campbell, J. B., & Gaines, D. N. (2014). Spatial and temporal emergence pattern of Lyme disease in Virginia. *Am J Trop Med Hyg*, *91*(6), 1166-1172. https://doi.org/10.4269/ajtmh.13-0733

156. Li, L., Xi, Y., & Ren, F. (2016). Spatio-temporal distribution characteristics and trajectory similarity analysis of tuberculosis in Beijing, China. *International Journal of Environmental Research and Public Health*, *13*(3), 291.

157. Li, Q., Cao, W., Ren, H., Ji, Z., & Jiang, H. (2018). Spatiotemporal responses of dengue fever transmission to the road network in an urban area. *Acta Tropica*, *183*, 8-13.

158. Li, T., Cheng, Q., Li, C., Stokes, E., Collender, P., Ohringer, A., Li, X., Li, J., Zelner, J. L., & Liang, S. (2019). Evidence for heterogeneity in China's progress against pulmonary tuberculosis: uneven reductions in a major center of ongoing transmission, 2005–2017. *Bmc Infectious Diseases*, *19*(1), 1-11.

159. Li, X., Chen, D., Zhang, Y., Xue, X., Zhang, S., Chen, M., Liu, X., & Ding, G. (2021). Analysis of spatial-temporal distribution of notifiable respiratory infectious diseases in Shandong Province, China during 2005–2014. *Bmc Public Health*, *21*(1), 1-13.

160. Li, Z., Fu, J., Lin, G., & Jiang, D. (2019). Spatiotemporal variation and hotspot detection of the avian influenza A (H7N9) virus in China, 2013–2017. *International Journal of Environmental Research and Public Health*, *16*(4), 648.

161. Li, Z., Yin, W., Clements, A., Williams, G., Lai, S., Zhou, H., Zhao, D., Guo, Y., Zhang, Y., & Wang, J. (2012). Spatiotemporal analysis of indigenous and imported dengue fever cases in Guangdong province, China. *Bmc Infectious Diseases*, *12*(1), 1-9.

162. Liao, W.-B., Ju, K., Gao, Y.-M., & Pan, J. (2020). The association between internal migration and pulmonary tuberculosis in China, 2005–2015: a spatial analysis. *Infectious Diseases of Poverty*, *9*(1), 1-12.

163. Lin, C.-H., Schiøler, K. L., Jepsen, M. R., Ho, C.-K., Li, S.-H., & Konradsen, F. (2012). Dengue outbreaks in high-income area, Kaohsiung City, Taiwan, 2003–2009. *Emerging Infectious Diseases*, *18*(10), 1603.

164. Ling, C. Y., Gruebner, O., Krämer, A., & Lakes, T. (2014). Spatio-temporal patterns of dengue in Malaysia: combining address and sub-district level. *Geospatial Health*, *9*(1), 131-140.

165. Lippi, C. A., Stewart-Ibarra, A. M., Romero, M., Lowe, R., Mahon, R., Van Meerbeeck, C. J., Rollock, L., Hilaire, M. G.-S., Trotman, A. R., & Holligan, D. (2020). Spatiotemporal tools for emerging and endemic disease hotspots in small areas: An analysis of dengue and chikungunya in Barbados, 2013–2016. *The American Journal of Tropical Medicine and Hygiene*, *103*(1), 149.

166. Liu, K., Sun, J., Liu, X., Li, R., Wang, Y., Lu, L., Wu, H., Gao, Y., Xu, L., & Liu, Q. (2018). Spatiotemporal patterns and determinants of dengue at county level in China from 2005–2017. *International Journal of Infectious Diseases*, *77*, 96-104.

167. Liu, S., Qin, Y., Xie, Z., & Zhang, J. (2020). The spatio-temporal characteristics and influencing factors of covid-19 spread in Shenzhen, China—An analysis based on 417 cases. *International Journal of Environmental Research and Public Health*, *17*(20), 7450.

168. Liu, W., Wang, D., Hua, S., Xie, C., Wang, B., Qiu, W., Xu, T., Ye, Z., Yu, L., & Yang, M. (2021). Spatiotemporal analysis of COVID-19 outbreaks in Wuhan, China. *Scientific Reports*, *11*(1), 1-9.

169. Liu, W., Yang, K., Qi, X., Xu, K., Ji, H., Ai, J., Ge, A., Wu, Y., Li, Y., & Dai, Q. (2013). Spatial and temporal analysis of human infection with avian influenza A (H7N9) virus in China, 2013. *Eurosurveillance*, *18*(47), 20640.

170. Liu, Y., Li, X., Wang, W., Li, Z., Hou, M., He, Y., Wu, W., Wang, H., Liang, H., & Guo, X. (2012). Investigation of space-time clusters and geospatial hot spots for the occurrence of tuberculosis in Beijing. *The International Journal of Tuberculosis and Lung Disease*, *16*(4), 486-491.

171. Low, G. K., Papapreponis, P., Isa, R. M., Gan, S. C., Chee, H. Y., Te, K. K., & Hatta, N. M. (2018). Geographical distribution and spatio-temporal patterns of hospitalization due to dengue infection at a leading specialist hospital in Malaysia. *Geospatial Health*, *13*(1).

172. Ma, Q., Gao, J., Zhang, W., Wang, L., Li, M., Shi, J., Zhai, Y., Sun, D., Wang, L., & Chen, B. (2021). Spatio-temporal distribution characteristics of COVID-19 in China: a city-level modeling study. *Bmc Infectious Diseases*, *21*(1), 1-14.

173. Majid, N. A., & Rasdi, R. M. (2020). Dengue Hotspot Detection in Bangi, Selangor, Malaysia. *IOP Conference Series: Earth and Environmental Science*, *540*(1), 012041. https://doi.org/10.1088/1755-1315/540/1/012041
174. Majumdar, S. (2021). Spatiotemporal pattern and hotspot detection of malaria using spatial analysis and GIS in West Bengal: an approach to medical GIS. In *Healthcare Paradigms in the Internet of Things Ecosystem* (pp. 83-100). Elsevier.
175. Mala, S., & Jat, M. K. (2019). Geographic information system based spatio-temporal dengue fever cluster analysis and mapping. *The Egyptian Journal of Remote Sensing and Space Science*, *22*(3), 297-304.
176. Malvisi, L., Troisi, C. L., & Selwyn, B. J. (2018). Analysis of the spatial and temporal distribution of malaria in an area of Northern Guatemala with seasonal malaria transmission. *Parasitology research*, *117*(9), 2807-2822.
177. Manabe, T., Phan, D., Nohara, Y., Kambayashi, D., Nguyen, T. H., Van Do, T., & Kudo, K. (2021). Spatiotemporal distribution of COVID-19 during the first 7 months of the epidemic in Vietnam. *Bmc Infectious Diseases*, *21*(1), 1-8.
178. Manabe, T., Yamaoka, K., Tango, T., Binh, N. G., Co, D. X., Tuan, N. D., Izumi, S., Takasaki, J., Chau, N. Q., & Kudo, K. (2015). Chronological, geographical, and seasonal trends of human cases of avian influenza A (H5N1) in Vietnam, 2003–2014: a spatial analysis. *Bmc Infectious Diseases*, *16*(1), 1-8.
179. Mao, Q., Zeng, C., Zheng, D., & Yang, Y. (2019). Analysis on spatial-temporal distribution characteristics of smear positive pulmonary tuberculosis in China, 2004–2015. *International Journal of Infectious Diseases*, *80*, S36-S44.
180. Mao, Y., He, R., Zhu, B., Liu, J., & Zhang, N. (2020). Notifiable respiratory infectious diseases in China: a spatial–temporal epidemiology analysis. *International Journal of Environmental Research and Public Health*, *17*(7), 2301.
181. Mao, Y., Zhang, N., Zhu, B., Liu, J., & He, R. (2019). A descriptive analysis of the Spatio-temporal distribution of intestinal infectious diseases in China. *Bmc Infectious Diseases*, *19*(1), 1-14.
182. Marek, L., Tuček, P., & Pászto, V. (2015). Using geovisual analytics in Google Earth to understand disease distribution: a case study of campylobacteriosis in the Czech Republic (2008–2012). *International Journal of Health Geographics*, *14*(1), 1-13.
183. Martins-Melo, F. R., Ramos Jr, A. N., Cavalcanti, M. G., Alencar, C. H., & Heukelbach, J. (2017). Reprint of "Neurocysticercosis-related mortality in Brazil, 2000–2011: Epidemiology of a neglected neurologic cause of death". *Acta Tropica*, *165*, 170-178.
184. Masrur, A., Yu, M., Luo, W., & Dewan, A. (2020). Space-time patterns, change, and propagation of COVID-19 risk relative to the intervention scenarios in Bangladesh. *International Journal of Environmental Research and Public Health*, *17*(16), 5911.
185. McIntosh, A. I., Jenkins, H. E., White, L. F., Barnard, M., Thomson, D. R., Dolby, T., Simpson, J., Streicher, E. M., Kleinman, M. B., & Ragan, E. J. (2018). Using routinely collected laboratory data to identify high rifampicin-resistant tuberculosis burden communities in the Western Cape Province, South Africa: A retrospective spatiotemporal analysis. *Plos Medicine*, *15*(8), e1002638.
186. Mee, P., Collinson, M. A., Madhavan, S., Root, E. D., Tollman, S. M., Byass, P., & Kahn, K. (2014). Evidence for localised HIV related micro–epidemics associated with the decentralised provision of antiretroviral treatment in rural South Africa: a spatio–temporal analysis of changing mortality patterns (2007–2010). *Journal of global health*, *4*(1).

187. Meiring, S., Cohen, C., de Gouveia, L., du Plessis, M., Kularatne, R., Hoosen, A., Lekalakala, R., Lengana, S., Seetharam, S., & Naicker, P. (2019). Declining incidence of invasive meningococcal disease in South Africa: 2003–2016. *Clinical Infectious Diseases*, *69*(3), 495-504.

188. Melchior, L. A. K., & Neto, F. C. (2016). Spatial and spatio-temporal analysis of malaria in the state of Acre, western Amazon, Brazil. *Geospatial Health*, *11*(3).

189. Mercado, C. E. G., Lawpoolsri, S., Sudathip, P., Kaewkungwal, J., Khamsiriwatchara, A., Pan-Ngum, W., Yimsamran, S., Lawawirojwong, S., Ho, K., & Ekapirat, N. (2019). Spatiotemporal epidemiology, environmental correlates, and demography of malaria in Tak Province, Thailand (2012–2015). *Malaria Journal*, *18*(1), 1-15.

190. Mesquita, C. R., Enk, M. J., & Guimarães, R. J. d. P. S. (2021). Spatial analysis studies of endemic diseases for health surveillance: Application of scan statistics for surveillance of tuberculosis among residents of a metropolitan municipality aged 60 years and above. *Ciencia & Saude Coletiva*, *26*, 5149-5156.

191. Meyers, D. J., Hood, M. E., & Stopka, T. J. (2014). HIV and hepatitis C mortality in Massachusetts, 2002–2011: spatial cluster and trend analysis of HIV and HCV using multiple cause of death. *Plos One*, *9*(12), e114822.

192. Mo, C., Tan, D., Mai, T., Bei, C., Qin, J., Pang, W., & Zhang, Z. (2020). An analysis of spatiotemporal pattern for COIVD-19 in China based on space-time cube. *Journal of Medical Virology*, *92*(9), 1587-1595.

193. MohammadEbrahimi, S., Mohammadi, A., Bergquist, R., Dolatkhah, F., Olia, M., Tavakolian, A., Pishgar, E., & Kiani, B. (2021). Epidemiological characteristics and initial spatiotemporal visualisation of COVID-19 in a major city in the Middle East. *Bmc Public Health*, *21*(1), 1-18.

194. Moirano, G., Richiardi, L., Calzolari, M., Merletti, F., & Maule, M. (2020). Recent rapid changes in the spatio-temporal distribution of West Nile Neuro-invasive Disease in Italy. *Zoonoses and Public Health*, *67*(1), 54-61.

195. Morato, D. G., Barreto, F. R., Braga, J. U., Natividade, M. S., Costa, M. d. C. N., Morato, V., & Teixeira, M. d. G. L. C. (2015). The spatiotemporal trajectory of a dengue epidemic in a medium-sized city. *Memorias Do Instituto Oswaldo Cruz*, *110*, 528-533.

196. Morrison, A. C., Getis, A., Santiago, M., Rigau-Perez, J. G., & Reiter, P. (1998). Exploratory space-time analysis of reported dengue cases during an outbreak in Florida, Puerto Rico, 1991-1992. *Am J Trop Med Hyg*, *58*(3), 287-298. https://doi.org/10.4269/ajtmh.1998.58.287

197. Mulatti, P., Mazzucato, M., Montarsi, F., Ciocchetta, S., Capelli, G., Bonfanti, L., & Marangon, S. (2015). Retrospective space-time analysis methods to support west nile virus surveillance activities. *Epidemiology & Infection*, *143*(1), 202-213.

198. Muttitanon, W. (2021). Clustering Analysis Influenza Disease to Identify Spatio-Temporal Spread Pattern in Thailand. *International Journal of Geoinformatics*, *17*(5).

199. Mylona, E. K., Shehadeh, F., Kalligeros, M., Benitez, G., Chan, P. A., & Mylonakis, E. (2020). Real-time spatiotemporal analysis of microepidemics of influenza and COVID-19 based on hospital network data: colocalization of neighborhood-level hotspots. *American Journal of Public Health*, *110*(12), 1817-1824.

200. Naim, M., Sahani, M., Hod, R., Hidayatulfathi, O., Idrus, S., Norzawati, Y., Hazrin, H., Tahir, A., Wen, T., & King, C. (2014). Spatial-temporal analysis for identification of

vulnerability to dengue in Seremban district, Malaysia. *International Journal of Geoinformatics*.

201. Naish, S., Dale, P., Mackenzie, J. S., McBride, J., Mengersen, K., & Tong, S. (2014). Spatial and temporal patterns of locally-acquired dengue transmission in northern Queensland, Australia, 1993–2012. *Plos One*, *9*(4), e92524.
202. Naish, S., & Tong, S. (2014). Hot spot detection and spatio-temporal dynamics of dengue in Queensland, Australia. Proceedings of the ISPRS Technical Commission VIII Symposium [International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives, Volume XL-8],
203. Naqvi, S. A. A., Sajjad, M., Waseem, L. A., Khalid, S., Shaikh, S., & Kazmi, S. J. H. (2021). Integrating Spatial Modelling and Space-time Pattern Mining Analytics for Vector Disease-Related Health Perspectives: A Case of Dengue Fever in Pakistan. *International Journal of Environmental Research and Public Health*, *18*(22), 12018.
204. Nascimento, M. L. F. (2020). A multivariate analysis on spatiotemporal evolution of Covid-19 in Brazil. *Infectious Disease Modelling*, *5*, 670-680.
205. Nazia, N., Ali, M., Jakariya, M., Nahar, Q., Yunus, M., & Emch, M. (2018). Spatial and population drivers of persistent cholera transmission in rural Bangladesh: Implications for vaccine and intervention targeting. *Spatial and Spatio-Temporal Epidemiology*, *24*, 1-9.
206. Ngwa, M. C., Ihekweazu, C., Okwor, T., Yennan, S., Williams, N., Elimian, K., Karaye, N. Y., Bello, I. W., & Sack, D. A. (2021). The cholera risk assessment in Kano State, Nigeria: A historical review, mapping of hotspots and evaluation of contextual factors. *Plos Neglected Tropical Diseases*, *15*(1), e0009046.
207. Nimbalkar, P. M., & Tripathi, N. K. (2016). Space-time epidemiology and effect of meteorological parameters on influenza-like illness in Phitsanulok, a northern province in Thailand. *Geospatial Health*, *11*(3).
208. Noé, A., Zaman, S. I., Rahman, M., Saha, A. K., Aktaruzzaman, M., & Maude, R. J. (2018). Mapping the stability of malaria hotspots in Bangladesh from 2013 to 2016. *Malaria Journal*, *17*(1), 1-21.
209. Nsoesie, E. O., Ricketts, R. P., Brown, H. E., Fish, D., Durham, D. P., Ndeffo Mbah, M. L., Christian, T., Ahmed, S., Marcellin, C., & Shelly, E. (2015). Spatial and temporal clustering of chikungunya virus transmission in Dominica. *Plos Neglected Tropical Diseases*, *9*(8), e0003977.
210. Olfatifar, M., Karami, M., Hosseini, S. M., & Parvin, M. (2016). Clustering of pulmonary tuberculosis in Hamadan province, west of Iran: a population based cross sectional study (2005-2013). *Journal of Research in Health Sciences*, *16*(3), 166.
211. Oliveira, L. R. d., Santos, E. S. d., & Souto, F. J. D. (2020). Syphilis in pregnant women and congenital syphilis: spatial pattern and relationship with social determinants of health in Mato Grosso. *Revista Da Sociedade Brasileira De Medicina Tropical*, *53*.
212. Onozuka, D., & Hagihara, A. (2007). Geographic prediction of tuberculosis clusters in Fukuoka, Japan, using the space-time scan statistic. *Bmc Infectious Diseases*, *7*(1), 1-9.
213. Owusu-Edusei, K., & Owens, C. J. (2009). Monitoring county-level chlamydia incidence in Texas, 2004–2005: application of empirical Bayesian smoothing and Exploratory Spatial Data Analysis (ESDA) methods. *International Journal of Health Geographics*, *8*(1), 1-6.
214. Owusu, C., Desjardins, M. R., Baker, K. M., & Delmelle, E. (2019). Residential mobility impacts relative risk estimates of space-time clusters of chlamydia in Kalamazoo County, Michigan. *Geospatial Health*, *14*(2).

215. Pai, S., Polgreen, P. M., Segre, A. M., Sewell, D. K., & Pemmaraju, S. V. (2020). Spatiotemporal clustering of in-hospital Clostridioides difficile infection. *Infection Control & Hospital Epidemiology*, *41*(4), 418-424.

216. Panahi, M. H., Parsaeian, M., Mansournia, M. A., Khoshabi, M., Gouya, M. M., Hemati, P., & Fotouhi, A. (2020). A spatio-temporal analysis of influenza-like illness in Iran from 2011 to 2016. *Medical Journal of the Islamic Republic of Iran*, *34*, 65.

217. Pardhan-Ali, A., Berke, O., Wilson, J., Edge, V. L., Furgal, C., Reid-Smith, R., Santos, M., & McEwen, S. A. (2012). A spatial and temporal analysis of notifiable gastrointestinal illness in the Northwest Territories, Canada, 1991-2008. *International Journal of Health Geographics*, *11*(1), 1-10.

218. Park, S., Nam, H.-S., & Na, B.-J. (2021). Evaluating the spatial and temporal patterns of the severe fever thrombocytopenia syndrome in Republic of Korea. *Geospatial Health*, *16*(2).

219. Parra-Amaya, M. E., Puerta-Yepes, M. E., Lizarralde-Bejarano, D. P., & Arboleda-Sánchez, S. (2016). Early detection for dengue using local indicator of spatial association (LISA) analysis. *Diseases*, *4*(2), 16.

220. Pasaribu, A. P., Tsheten, T., Yamin, M., Maryani, Y., Fahmi, F., Clements, A. C., Gray, D. J., & Wangdi, K. (2021). Spatio-temporal patterns of dengue incidence in Medan city, North Sumatera, Indonesia. *Tropical Medicine and Infectious Disease*, *6*(1), 30.

221. Pearson, A. L., Kingham, S., Mitchell, P., & Apparicio, P. (2013). Exploring hotspots of pneumococcal pneumonia and potential impacts of ejecta dust exposure following the Christchurch earthquakes. *Spatial and Spatio-Temporal Epidemiology*, *7*, 1-9.

222. Pellini, A. C. G., Chiaravalloti-Neto, F., & Zanetta, D. M. T. (2020). AIDS in men in the city of São Paulo, 1980–2012: spatial and space-time analysis. *Revista De Saude Publica*, *54*, 96.

223. Petersen, J., Gibin, M., Sile, B., & Simms, I. (2016). Identifying and interpreting spatiotemporal variation in diagnoses of infectious syphilis among men, England: 2009 to 2013. *Sexually Transmitted Infections*, *92*(5), 380-386.

224. Pijnacker, R., Mughini-Gras, L., Vennema, H., Enserink, R., Van den Wijngaard, C., Kortbeek, T., & van Pelt, W. (2016). Characteristics of child daycare centres associated with clustering of major enteropathogens. *Epidemiology & Infection*, *144*(12), 2527-2539.

225. Pinchoff, J., Chipeta, J., Banda, G. C., Miti, S., Shields, T., Curriero, F., & Moss, W. J. (2015). Spatial clustering of measles cases during endemic (1998–2002) and epidemic (2010) periods in Lusaka, Zambia. *Bmc Infectious Diseases*, *15*(1), 1-8.

226. Pinchoff, J., Henostroza, G., Carter, B. S., Roberts, S. T., Hatwiinda, S., Hamainza, B., Hawela, M., & Curriero, F. C. (2015). Spatial patterns of incident malaria cases and their household contacts in a single clinic catchment area of Chongwe District, Zambia. *Malaria Journal*, *14*(1), 1-7.

227. Pishgar, E., Fanni, Z., Tavakkolinia, J., Mohammadi, A., Kiani, B., & Bergquist, R. (2020). Mortality rates due to respiratory tract diseases in Tehran, Iran during 2008–2018: a spatiotemporal, cross-sectional study. *Bmc Public Health*, *20*(1), 1-12.

228. Portella, T. P., & Kraenkel, R. A. (2021). Spatial-temporal pattern of cutaneous leishmaniasis in Brazil. *Infectious Diseases of Poverty*, *10*(03), 47-57.

229. Qi, C., Zhu, Y., Li, C., Hu, Y., Liu, L., Zhang, D., Wang, X., She, K., Jia, Y., & Liu, T. (2020). Epidemiological characteristics and spatial− temporal analysis of COVID-19 in Shandong Province, China. *Epidemiology & Infection*, *148*.

230. Qin, Q., Guo, W., Tang, W., Mahapatra, T., Wang, L., Zhang, N., Ding, Z., Cai, C., Cui, Y., & Sun, J. (2017). Spatial analysis of the human immunodeficiency virus epidemic among men who have sex with men in China, 2006–2015. *Clinical Infectious Diseases*, *64*(7), 956-963.
231. Rahman, M., Islam, A., & Islam, M. (2021). Geospatial modelling on the spread and dynamics of 154 day outbreak of the novel coronavirus (COVID-19) pandemic in Bangladesh towards vulnerability zoning and management approaches. *Modeling earth systems and environment*, *7*(3), 2059-2087.
232. Ramírez, I. J., Lee, J., & Grady, S. C. (2018). Mapping Multi-Disease Risk during El Niño: An Ecosyndemic Approach. *Int J Environ Res Public Health*, *15*(12). https://doi.org/10.3390/ijerph15122639
233. Reinhardt, M., Elias, J., Albert, J., Frosch, M., Harmsen, D., & Vogel, U. (2008). EpiScanGIS: an online geographic surveillance system for meningococcal disease. *International Journal of Health Geographics*, *7*(1), 1-7.
234. Rejeki, D., Nurhayati, N., & Aji, B. (2021). A spatiotemporal analysis of dengue hemorrhagic fever in Banyumas, Indonesia. *Int. J. Public Health Sci.(IJPHS)*, *10*, 231-240.
235. Rejeki, D. S. S., Fuad, A., Widartono, B. S., Murhandarwati, E., & Kusnanto, H. (2019). Spatiotemporal patterns of malaria at cross-boundaries area in Menoreh Hills, Java, Indonesia. *Malaria Journal*, *18*(1), 1-12.
236. Ren, H., Ning, W., Lu, L., Zhuang, D., & Liu, Q. (2015). Characterization of dengue epidemics in mainland China over the past decade. *The Journal of Infection in Developing Countries*, *9*(09), 970-976.
237. Ren, J., Ling, F., Sun, J., Gong, Z., Liu, Y., Shi, X., Zhang, R., Zhai, Y., Chen, E., & Chen, Z. (2018). Epidemiological profile of dengue in Zhejiang Province, southeast China. *Plos One*, *13*(12), e0208810.
238. Reyes-Castro, P. A., Harris, R. B., Brown, H. E., Christopherson, G. L., & Ernst, K. C. (2017). Spatio-temporal and neighborhood characteristics of two dengue outbreaks in two arid cities of Mexico. *Acta Tropica*, *167*, 174-182.
239. Riyanto, I. A., Susianti, N. A., Sholihah, R. A., Rizki, R. L. P., Cahyadi, A., Naufal, M., Ramadhan, F., Ramadan, V. K., & Risky, A. S. (2020). The spatiotemporal analysis of dengue fever in Purwosari district, Gunungkidul Regency, Indonesia. *The Indonesian Journal of Geography*, *52*(1), 80-91.
240. Rocheleau, J.-P., Kotchi, S.-O., & Arsenault, J. (2020). Can local risk of West Nile virus infection be predicted from previous cases? A descriptive study in Quebec, 2011–2016. *Canadian Journal of Public Health*, *111*(2), 229-238.
241. Romanyukha, A. A., Karkach, A. S., Borisov, S. E., Belilovsky, E. M., Sannikova, T. E., & Krivorotko, O. I. (2020). Small-scale stable clusters of elevated tuberculosis incidence in Moscow, 2000–2015: Discovery and spatiotemporal analysis. *International Journal of Infectious Diseases*, *91*, 156-161.
242. Rosillo, N., Del-Águila-Mejía, J., Rojas-Benedicto, A., Guerrero-Vadillo, M., Peñuelas, M., Mazagatos, C., Segú-Tell, J., Ramis, R., & Gómez-Barroso, D. (2021). Real time surveillance of COVID-19 space and time clusters during the summer 2020 in Spain. *Bmc Public Health*, *21*(1), 1-11.
243. Rotela, C., Lopez, L., Céspedes, M. F., Barbas, G., Lighezzolo, A., Porcasi, X., Lanfri, M. A., Scavuzzo, C. M., & Gorla, D. E. (2017). Analytical report of the 2016 dengue outbreak in Córdoba city, Argentina. *Geospatial Health*, *12*(2).

244. Rouamba, T., Nakanabo-Diallo, S., Derra, K., Rouamba, E., Kazienga, A., Inoue, Y., Ouédraogo, E. K., Waongo, M., Dieng, S., & Guindo, A. (2019). Socioeconomic and environmental factors associated with malaria hotspots in the Nanoro demographic surveillance area, Burkina Faso. *Bmc Public Health*, *19*(1), 1-14.
245. Saavedra-Nieves, P. (2021). Nonparametric estimation of highest density regions for COVID-19. *Journal of Nonparametric Statistics*, 1-20.
246. Sadeq, M., & Bourkadi, J. E. (2018). Spatiotemporal distribution and predictors of tuberculosis incidence in Morocco. *Infectious Diseases of Poverty*, *7*(1), 1-13.
247. Saita, S., Silawan, T., Parker, D. M., Sriwichai, P., Phuanukoonnon, S., Sudathip, P., Maude, R. J., White, L. J., & Pan-Ngum, W. (2019). Spatial heterogeneity and temporal trends in malaria on the Thai–Myanmar border (2012–2017): a retrospective observational study. *Tropical Medicine and Infectious Disease*, *4*(2), 62.
248. Sakai, T., Suzuki, H., Sasaki, A., Saito, R., Tanabe, W., & Taniguchi, K. (2004). Geographic and temporal trends in influenzalike illness, Japan, 1992-1999 [Article]. *Emerging Infectious Diseases*, *10*(10), 1822-1826. https://doi.org/10.3201/eid1010.040147
249. Salway, T., Gesink, D., Lukac, C., Roth, D., Ryan, V., Mak, S., Wang, S., Newhouse, E., Hayden, A., & Bharmal, A. (2019). Spatial-temporal epidemiology of the syphilis epidemic in relation to neighborhood-level structural factors in British Columbia, 2005–2016. *Sexually Transmitted Diseases*, *46*(9), 571-578.
250. Salway, T., Gesink, D., Lukac, C., Roth, D., Ryan, V., Mak, S., Wang, S., Newhouse, E., Hayden, A., Bharmal, A., Hoyano, D., Morshed, M., Grennan, T., Gilbert, M., & Wong, J. (2019). Spatial-Temporal Epidemiology of the Syphilis Epidemic in Relation to Neighborhood-Level Structural Factors in British Columbia, 2005-2016. *Sex Transm Dis*, *46*(9), 571-578. https://doi.org/10.1097/olq.0000000000001034
251. Santos, C. V. B. d., Cavalcante, J. R., Pungartnik, P. C., & Guimarães, R. M. (2021). Space-time analysis of the first year of COVID-19 pandemic in the city of Rio de Janeiro, Brazil. *Revista Brasileira de Epidemiologia*, *24*.
252. Santos, J. A., Santos, D. T., Arcencio, R. A., & Nunes, C. (2021). Space-time clustering and temporal trends of hospitalizations due to pulmonary tuberculosis: potential strategy for assessing health care policies. *European Journal of Public Health*, *31*(1), 57-62.
253. Santos, J. P. C. d., Honório, N. A., & Nobre, A. A. (2019). Definition of persistent areas with increased dengue risk by detecting clusters in populations with differing mobility and immunity in Rio de Janeiro, Brazil. *Cadernos De Saude Publica*, *35*.
254. Santos Neto, M., Yamamura, M., Garcia, M. C. d. C., Popolin, M. P., Rodrigues, L. B. B., Chiaravalloti Neto, F., Fronteira, I., & Arcêncio, R. A. (2015). Pulmonary tuberculosis in São Luis, State of Maranhão, Brazil: space and space-time risk clusters for death (2008-2012). *Revista Da Sociedade Brasileira De Medicina Tropical*, *48,* 69-76.
255. Sartorius, B., Kahn, K., Vounatsou, P., Collinson, M., & Tollman, S. (2010). Space and time clustering of mortality in rural South Africa (Agincourt HDSS), 1992–2007. *Global Health Action*, *3*(1), 5225.
256. Sasaki, S., Suzuki, H., Fujino, Y., Kimura, Y., & Cheelo, M. (2009). Impact of drainage networks on cholera outbreaks in Lusaka, Zambia. *American Journal of Public Health*, *99*(11), 1982-1987.
257. Saule, E., Panchananam, D., Hohl, A., Tang, W., & Delmelle, E. (2017). Parallel space-time kernel density estimation. 2017 46th International Conference on Parallel Processing (ICPP),

258. Schellenberg, J. A., Newell, J. N., Snow, R. W., Mung'ala, V., Marsh, K., Smith, P. G., & Hayes, R. J. (1998). An analysis of the geographical distribution of severe malaria in children in Kilifi District, Kenya [Article]. *International Journal of Epidemiology*, *27*(2), 323-329. https://doi.org/10.1093/ije/27.2.323

259. Schleihauf, E., Watkins, R. E., & Plant, A. J. (2009). Heterogeneity in the spatial distribution of bacterial sexually transmitted infections. *Sexually Transmitted Infections*, *85*(1), 45-49.

260. Sebu, A. T., Genati, K. T., Bekalo, D. B., & Deressa, T. K. (2020). Spatiotemporal dynamics of tuberculosis in East Hararge Zone, Oromiya Region, Ethiopia. *Spatial Information Research*, *28*(6), 623-633.

261. Seyoum, D., Yewhalaw, D., Duchateau, L., Brandt, P., Rosas-Aguirre, A., & Speybroeck, N. (2017). Household level spatio-temporal analysis of Plasmodium falciparum and Plasmodium vivax malaria in Ethiopia. *Parasites & Vectors*, *10*(1), 1-11.

262. Shabani, J., Lutambi, A. M., Mwakalinga, V., & Masanja, H. (2010). Clustering of under-five mortality in Rufiji Health and Demographic Surveillance System in rural Tanzania. *Global Health Action*, *3*(1), 5264.

263. Shaffer, J. G., Touré, M. B., Sogoba, N., Doumbia, S. O., Gomis, J. F., Ndiaye, M., Ndiaye, D., Diarra, A., Abubakar, I., & Ahmad, A. (2020). Clustering of asymptomatic Plasmodium falciparum infection and the effectiveness of targeted malaria control measures. *Malaria Journal*, *19*(1), 1-13.

264. Shan, X., Wang, Y., Song, R., Wei, W., Liao, H., Huang, H., Xu, C., Chen, L., & Li, S. (2020). Spatial and temporal clusters of avian influenza a (H7N9) virus in humans across five epidemics in mainland China: an epidemiological study of laboratory-confirmed cases. *Bmc Infectious Diseases*, *20*(1), 1-7.

265. Shanks, G., Waller, M., & Smallman-Raynor, M. (2013). Spatiotemporal patterns of pandemic influenza-related deaths in Allied naval forces during 1918. *Epidemiology & Infection*, *141*(10), 2205-2212.

266. Shariati, M., Mesgari, T., Kasraee, M., & Jahangiri-Rad, M. (2020). Spatiotemporal analysis and hotspots detection of COVID-19 using geographic information system (March and April, 2020). *Journal of Environmental Health Science and Engineering*, *18*(2), 1499-1507.

267. Sharma, K. D., Mahabir, R. S., Curtin, K. M., Sutherland, J. M., Agard, J. B., & Chadee, D. D. (2014). Exploratory space-time analysis of dengue incidence in Trinidad: a retrospective study using travel hubs as dispersal points, 1998–2004. *Parasites & Vectors*, *7*(1), 1-11.

268. Shi, J., Gao, X., Xue, S., Li, F., Nie, Q., Lv, Y., Wang, J., Xu, T., Du, G., & Li, G. (2021). Spatio-temporal evolution and influencing mechanism of the COVID-19 epidemic in Shandong province, China. *Scientific Reports*, *11*(1), 1-16.

269. Shiode, N., Shiode, S., Rod-Thatcher, E., Rana, S., & Vinten-Johansen, P. (2015). The mortality rates and the space-time patterns of John Snow's cholera epidemic map. *International Journal of Health Geographics*, *14*(1), 1-15.

270. Shobugawa, Y., Wiafe, S. A., Saito, R., Suzuki, T., Inaida, S., Taniguchi, K., & Suzuki, H. (2012). Novel measurement of spreading pattern of influenza epidemic by using weighted standard distance method: retrospective spatial statistical study of influenza, Japan, 1999–2009. *International Journal of Health Geographics*, *11*(1), 1-9.

271. Sifuna, P., Otieno, L., Andagalu, B., Oyieko, J., Ogutu, B., Singoei, V., Owuoth, J., Ogwang, S., Cowden, J., & Otieno, W. (2018). A spatiotemporal analysis of hiv-associated mortality in rural western kenya 2011–2015. *Journal of Acquired Immune Deficiency Syndromes (1999)*, *78*(5), 483.

272. Sifuna, P. M., Ouma, C., Atieli, H., Owuoth, J., Onyango, D., Andagalu, B., & Cowden, J. (2019). Spatial epidemiology of tuberculosis in the high-burden counties of Kisumu and Siaya, Western Kenya, 2012-2015. *Int J Tuberc Lung Dis*, *23*(3), 363-370. https://doi.org/10.5588/ijtld.18.0245

273. Silva, A. P. d. S. C., Souza, W. V. d., & Albuquerque, M. d. F. P. M. d. (2016). Two decades of tuberculosis in a city in Northeastern Brazil: advances and challenges in time and space. *Revista Da Sociedade Brasileira De Medicina Tropical*, *49*, 211-221.

274. Singh, P. S., & Chaturvedi, H. K. (2021). Temporal variation and geospatial clustering of dengue in Delhi, India 2015–2018. *Bmj Open*, *11*(2), e043848.

275. Sloan, C., Chandrasekhar, R., Mitchel, E., Ndi, D., Miller, L., Thomas, A., Bennett, N. M., Chai, S., Spencer, M., & Eckel, S. (2020). Spatial and temporal clustering of patients hospitalized with laboratory-confirmed influenza in the united states. *Epidemics*, *31*, 100387.

276. Smith, M. K., Searle, K. M., Yang, W., Rapheal, E., Wang, C., Zhao, P., Yang, L., Huang, S., & Yang, B. (2021). Spatiotemporal analysis of 11 years of Chlamydia trachomatis data from southern China. *The Lancet Regional Health-Western Pacific*, *11*, 100143.

277. Soto-Calle, V., Rosas-Aguirre, A., Llanos-Cuentas, A., Abatih, E., DeDeken, R., Rodriguez, H., Rosanas-Urgell, A., Gamboa, D., Erhart, A., & Speybroeck, N. (2017). Spatio-temporal analysis of malaria incidence in the Peruvian Amazon Region between 2002 and 2013. *Scientific Reports*, *7*(1), 1-13.

278. Stach, A. (2021). Temporal variation of spatial autocorrelation of COVID-19 cases identified in Poland during the year from the beginning of the pandemic. *Geographia Polonica*, *94*(3), 355-380.

279. Suchar, V. A., Aziz, N., Bowe, A., Burke, A., & Wiest, M. M. (2018). An exploration of the spatiotemporal and demographic patterns of Ebola Virus Disease epidemic in West Africa using open access data sources. *Applied Geography*, *90*, 272-281.

280. Sugumaran, R., Larson, S. R., & DeGroote, J. P. (2009). Spatio-temporal cluster analysis of county-based human West Nile virus incidence in the continental United States. *International Journal of Health Geographics*, *8*(1), 1-19.

281. Sun, W., Xue, L., & Xie, X. (2017). Spatial-temporal distribution of dengue and climate characteristics for two clusters in Sri Lanka from 2012 to 2016. *Scientific Reports*, *7*(1), 1-12.

282. Sun, Z., Zhang, Z., Liu, Q., Lyu, B., Fang, X., Wang, S., Xu, J., Xu, L., & Xu, B. (2020). Identifying the spatiotemporal clusters of plague occurrences in China during the Third Pandemic. *Integrative Zoology*, *15*(1), 69-78.

283. Tadesse, S., Enqueselassie, F., & Hagos, S. (2018). Spatial and space-time clustering of tuberculosis in Gurage Zone, Southern Ethiopia. *Plos One*, *13*(6), e0198353.

284. Tadesse, T., Demissie, M., Berhane, Y., Kebede, Y., & Abebe, M. (2013). The clustering of smear-positive tuberculosis in Dabat, Ethiopia: a population based cross sectional study. *Plos One*, *8*(5), e65022.

285. Tang, J.-H., Tseng, T.-J., & Chan, T.-C. (2019). Detecting spatio-temporal hotspots of scarlet fever in Taiwan with spatio-temporal Gi* statistic. *Plos One*, *14*(4), e0215434.

286. Tang, X., Geater, A., McNeil, E., Deng, Q., Dong, A., & Zhong, G. (2017). Spatial, temporal and spatio-temporal clusters of measles incidence at the county level in Guangxi,

China during 2004–2014: flexibly shaped scan statistics. *Bmc Infectious Diseases*, *17*(1), 1-10.

287. Telle, O., Vaguet, A., Yadav, N., Lefebvre, B., Daudé, E., Paul, R. E., Cebeillac, A., & Nagpal, B. (2016). The spread of dengue in an endemic urban milieu–the case of Delhi, India. *Plos One, 11*(1), e0146539.

288. Thakar, V. (2020). Unfolding events in space and time: Geospatial insights into COVID-19 diffusion in Washington State during the initial stage of the outbreak. *ISPRS International Journal of Geo-Information*, *9*(6), 382.

289. Thanh Toan, D. T., Hu, W., Quang Thai, P., Ngoc Hoat, L., Wright, P., & Martens, P. (2013). Hot spot detection and spatio-temporal dispersion of dengue fever in Hanoi, Vietnam. *Global Health Action*, *6*(1), 18632.

290. Tiwari, N., Adhikari, C., Tewari, A., & Kandpal, V. (2006). Investigation of geo-spatial hotspots for the occurrence of tuberculosis in Almora district, India, using GIS and spatial scan statistic. *International Journal of Health Geographics*, *5*(1), 1-11.

291. Tlou, B., Sartorius, B., & Tanser, F. (2017). Space-time variations in child mortality in a rural South African population with high HIV prevalence (2000–2014). *Plos One*, *12*(8), e0182478.

292. Tomita, A., Smith, C. M., Lessells, R. J., Pym, A., Grant, A. D., de Oliveira, T., & Tanser, F. (2019). Space-time clustering of recently-diagnosed tuberculosis and impact of ART scale-up: evidence from an HIV hyper-endemic rural South African population. *Scientific Reports*, *9*(1), 1-9.

293. Touray, K., Adetifa, I., Jallow, A., Rigby, J., Jeffries, D., Cheung, Y., Donkor, S., Adegbola, R., & Hill, P. (2010). Spatial analysis of tuberculosis in an urban west African setting: is there evidence of clustering? *Tropical Medicine & International Health*, *15*(6), 664-672.

294. Tran, A., Deparis, X., Dussart, P., Morvan, J., Rabarison, P., Remy, F., Polidori, L., & Gardon, J. (2004). Dengue spatial and temporal patterns, French Guiana, 2001 [Article]. *Emerging Infectious Diseases*, *10*(4), 615-621. https://doi.org/10.3201/eid1004.030186

295. Tuite, A., Guthrie, J., Alexander, D., Whelan, M., Lee, B., Lam, K., Ma, J., Fisman, D., & Jamieson, F. (2013). Epidemiological evaluation of spatiotemporal and genotypic clustering of Mycobacterium tuberculosis in Ontario, Canada. *The International journal of tuberculosis and lung disease*, *17*(10), 1322-1327.

296. Ullah, S., Daud, H., Dass, S. C., Fanaee-T, H., Kausarian, H., & Khalil, A. (2020). Space-time clustering characteristics of tuberculosis in Khyber Pakhtunkhwa Province, Pakistan, 2015–2019. *International Journal of Environmental Research and Public Health*, *17*(4), 1413.

297. Ullah, S., Daud, H., Dass, S. C., Fanaee-T, H., & Khalil, A. (2018). An Eigenspace approach for detecting multiple space-time disease clusters: Application to measles hotspots detection in Khyber-Pakhtunkhwa, Pakistan. *Plos One*, *13*(6), e0199176.

298. Ullah, S., Daud, H., Dass, S. C., Khan, H. N., & Khalil, A. (2017). Detecting space-time disease clusters with arbitrary shapes and sizes using a co-clustering approach. *Geospatial Health*, *12*(2).

299. Umer, M. F., Zofeen, S., Majeed, A., Hu, W., Qi, X., & Zhuang, G. (2018). Spatiotemporal clustering analysis of malaria infection in Pakistan. *International Journal of Environmental Research and Public Health*, *15*(6), 1202.

300. Valson, J. S., & Soman, B. (2017). Spatiotemporal clustering of dengue cases in Thiruvananthapuram district, Kerala. *Indian Journal of Public Health*, *61*(2), 74.

301. Van Aar, F., Den Daas, C., Van Der Sande, M., Soetens, L., De Vries, H., & Van Benthem, B. (2017). Outbreaks of syphilis among men who have sex with men attending STI clinics between 2007 and 2015 in the Netherlands: a space-time clustering study. *Sexually Transmitted Infections*, *93*(6), 390-395.

302. Van Den Wijngaard, C. C., Van Asten, L., Van Pelt, W., Doornbos, G., Nagelkerke, N. J., Donker, G. A., van der Hoek, W., & Koopmans, M. P. (2010). Syndromic surveillance for local outbreaks of lower-respiratory infections: would it work? *Plos One*, *5*(4), e10406.

303. Vazquez-Prokopec, G. M., Kitron, U., Montgomery, B., Horne, P., & Ritchie, S. A. (2010). Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. *Plos Neglected Tropical Diseases*, *4*(12), e920.

304. Venkat, A., Falconi, T. M. A., Cruz, M., Hartwick, M. A., Anandan, S., Kumar, N., Ward, H., Veeraraghavan, B., & Naumova, E. N. (2019). Spatiotemporal patterns of cholera hospitalization in Vellore, India. *International Journal of Environmental Research and Public Health*, *16*(21), 4257.

305. Verma, A., Schwartzman, K., Behr, M. A., Zwerling, A., Allard, R., Rochefort, C. M., & Buckeridge, D. L. (2014). Accuracy of prospective space-time surveillance in detecting tuberculosis transmission. *Spatial and Spatio-Temporal Epidemiology*, *8*, 47-54.

306. Vissoci, J. R. N., Rocha, T. A. H., Silva, N. C. d., de Sousa Queiroz, R. C., Thomaz, E. B. A. F., Amaral, P. V. M., Lein, A., Branco, M. d. R. F. C., Aquino, J., & Rodrigues, Z. M. R. (2018). Zika virus infection and microcephaly: Evidence regarding geospatial associations. *Plos Neglected Tropical Diseases*, *12*(4), e0006392.

307. Wahnich, A., Clark, S., Bloch, D., Kubinson, H., Hrusa, G., Liu, D., Rakeman, J. L., Deocharan, B., Jones, L., & Slavinski, S. (2018). Surveillance for mosquitoborne transmission of Zika virus, New York City, NY, USA, 2016. *Emerging Infectious Diseases*, *24*(5), 827.

308. Wang, P., Ren, H., Zhu, X., Fu, X., Liu, H., & Hu, T. (2021). Spatiotemporal characteristics and factor analysis of SARS-CoV-2 infections among healthcare workers in Wuhan, China. *Journal of Hospital Infection*, *110*, 172-177.

309. Wang, Q., Dong, W., Yang, K., Ren, Z., Huang, D., Zhang, P., & Wang, J. (2021). Temporal and spatial analysis of COVID-19 transmission in China and its influencing factors. *International Journal of Infectious Diseases*, *105*, 675-685.

310. Wang, T., Xue, F., Chen, Y., Ma, Y., & Liu, Y. (2012). The spatial epidemiology of tuberculosis in Linyi City, China, 2005–2010. *Bmc Public Health*, *12*(1), 1-8.

311. Wang, Y., Liu, Y., Struthers, J., & Lian, M. (2021). Spatiotemporal characteristics of the COVID-19 epidemic in the United States. *Clinical Infectious Diseases*, *72*(4), 643-651.

312. Wangdi, K., Kaewkungwal, J., Singhasivanon, P., Silawan, T., Lawpoolsri, S., & White, N. J. (2011). Spatio-temporal patterns of malaria infection in Bhutan: a country embarking on malaria elimination. *Malaria Journal*, *10*(1), 1-9.

313. Wangdi, K., Penjor, K., Lawpoolsri, S., Price, R. N., Gething, P. W., Gray, D. J., Da Silva Fonseca, E., & Clements, A. C. (2021). Space-time clustering characteristics of malaria in Bhutan at the end stages of elimination. *International Journal of Environmental Research and Public Health*, *18*(11), 5553.

314. Weiss, K., Karuchit, S., Pattanasin, S., Chitwarakorn, A., Wimonsate, W., Suksamosorn, J., Visavakum, P., Sukwicha, W., Ungsedhapand, C., & Dunne, E. F. (2020). Spatial

characteristics of men who have sex with men and transgender women attending HIV voluntary counselling and testing in Bangkok, Thailand, 2005–2015. *Geospatial Health*, *15*(1).

315. Wen, L., Li, C., Lin, M., Yuan, Z., Huo, D., Li, S., Wang, Y., Chu, C., Jia, R., & Song, H. (2011). Spatio-temporal analysis of malaria incidence at the village level in a malaria-endemic area in Hainan, China. *Malaria Journal*, *10*(1), 1-7.
316. Wen, T.-H., Lin, M.-H., & Fang, C.-T. (2012). Population movement and vector-borne disease transmission: differentiating spatial–temporal diffusion patterns of commuting and noncommuting dengue cases. *Annals of the Association of American Geographers*, *102*(5), 1026-1037.
317. Wen, T.-H., Lin, N. H., Chao, D.-Y., Hwang, K.-P., Kan, C.-C., Lin, K. C.-M., Wu, J. T.-S., Huang, S. Y.-J., Fan, I.-C., & King, C.-C. (2010). Spatial–temporal patterns of dengue in areas at risk of dengue hemorrhagic fever in Kaohsiung, Taiwan, 2002. *International Journal of Infectious Diseases*, *14*(4), e334-e343.
318. Wen, T.-h., Lin, N. H., Lin, K. C.-m., Fan, I.-c., Su, M.-d., & King, C.-c. (2007). A Spatial-Temporal Approach to Differentiate Epidemic Risk Patterns. In *GIS for Health and the Environment* (pp. 214-227). Springer.
319. Wen, T.-H., & Tsai, Y.-S. (2015). Analyzing the Patterns of Space-Time Distances for Tracking the Diffusion of an Epidemic. In M.-P. Kwan, D. Richardson, D. Wang, & C. Zhou (Eds.), *Space-Time Integration in Geography and GIScience: Research Frontiers in the US and China* (pp. 269-282). Springer Netherlands. https://doi.org/10.1007/978-94-017-9205-9_15
320. Wijayanti, S. P. M., Nurlaela, S., Octaviana, D., Putra, F. A., Nurhayati, S., & Sulistyawati, S. (2019). Dengue virus transmission during outbreak within endemic area in Indonesia: A spatial and temporal analysis. *Annals of Tropical Medicine and Public Health*, *22*, 158-164.
321. Wilson, J. G., Ballou, J., Yan, C., Fisher-Hoch, S. P., Reininger, B., Gay, J., Salinas, J., Sanchez, P., Salinas, Y., & Calvillo, F. (2010). Utilizing spatiotemporal analysis of influenza-like illness and rapid tests to focus swine-origin influenza virus intervention. *Health & Place*, *16*(6), 1230-1239.
322. Wiru, K., Oppong, F. B., Gyaase, S., Agyei, O., Abubakari, S. W., Amenga-Etego, S., Zandoh, C., & Asante, K. P. (2021). Geospatial analysis of malaria mortality in the kintampo health and demographic surveillance area of central Ghana. *Annals of Gis*, *27*(2), 139-149.
323. Wolf, A., Padayatchi, N., Naidoo, K., Master, I., Mathema, B., & O'donnell, M. R. (2020). Spatiotemporal Clustering of Multidrug-Resistant and Extensively Drug-Resistant Tuberculosis Is Associated With Human Immunodeficiency Virus Status and Drug-Susceptibility Patterns in KwaZulu-Natal, South Africa. *Clinical Infectious Diseases*, *70*(10), 2224-2227.
324. Wu, C., Zhou, M., Liu, P., & Yang, M. (2021). Analyzing COVID-19 using multisource data: An integrated approach of visualization, spatial regression, and machine learning. *Geohealth*, *5*(8), e2021GH000439.
325. Wu, H., Wu, C., Lu, Q., Ding, Z., Xue, M., & Lin, J. (2021). Spatiotemporal analysis and the characteristics of the case transmission network of 2019 novel coronavirus disease (COVID-19) in Zhejiang Province, China. *Plos One*, *16*(9), e0257587.

326. Xia, J., Huang, X., Sun, L., Zhu, H., Lin, W., Dong, X., Wu, D., Qiu, J., Zheng, L., & Cao, M. (2018). Epidemiological characteristics of malaria from control to elimination in Hubei Province, China, 2005–2016. *Malaria Journal*, *17*(1), 1-10.
327. Xiao, D., Long, Y., Wang, S., Wu, K., Xu, D., Li, H., Wang, G., & Yan, Y. (2012). Epidemic distribution and variation of Plasmodium falciparum and Plasmodium vivax malaria in Hainan, China during 1995–2008. *The American journal of tropical medicine and hygiene*, *87*(4), 646.
328. Xie, Z., Qin, Y., Li, Y., Shen, W., Zheng, Z., & Liu, S. (2020). Spatial and temporal differentiation of COVID-19 epidemic spread in mainland China and its influencing factors. *Science of the Total Environment*, *744*, 140929.
329. Xiong, Y., Wang, Y., Chen, F., & Zhu, M. (2020). Spatial statistics and influencing factors of the COVID-19 epidemic at both prefecture and county levels in Hubei Province, China. *International Journal of Environmental Research and Public Health*, *17*(11), 3903.
330. Xu, F., & Beard, K. (2021). A comparison of prospective space-time scan statistics and spatiotemporal event sequence based clustering for COVID-19 surveillance. *Plos One*, *16*(6), e0252990.
331. Xu, M., Cao, C., Zhang, X., Lin, H., Yao, Z., Zhong, S., Huang, Z., & Shea Duerler, R. (2021). Fine-scale space-time cluster detection of COVID-19 in Mainland China using retrospective analysis. *International Journal of Environmental Research and Public Health*, *18*(7), 3583.
332. Xu, X., Zhou, G., Wang, Y., Hu, Y., Ruan, Y., Fan, Q., Yang, Z., Yan, G., & Cui, L. (2016). Microgeographic heterogeneity of border malaria during elimination phase, Yunnan Province, China, 2011–2013. *Emerging Infectious Diseases*, *22*(8), 1363.
333. Yang, S., Gao, Y., Luo, W., Liu, L., Lei, Y., & Zhang, X. (2019). Spatiotemporal distribution of tuberculosis during urbanization in the new urban area of Nanchang City, China, 2010–2018. *International Journal of Environmental Research and Public Health*, *16*(22), 4395.
334. Yang, W., Deng, M., Li, C., & Huang, J. (2020). Spatio-temporal patterns of the 2019-nCoV epidemic at the county level in Hubei Province, China. *International Journal of Environmental Research and Public Health*, *17*(7), 2563.
335. Yeboah-Manu, D., Asare, P., Asante-Poku, A., Otchere, I., Osei-Wusu, S., Danso, E., Forson, A., Koram, K., & Gagneux, S. (2016). Spatio-temporal distribution of Mycobacterium tuberculosis complex strains in Ghana. *Plos One*, *11*(8), e0161892.
336. Yiannakoulias, N., & Svenson, L. (2009). Differences between notifiable and administrative health information in the spatial–temporal surveillance of enteric infections. *International Journal of Medical Informatics*, *78*(6), 417-424.
337. Ying, R., Fekadu, L., Schackman, B. R., & Verguet, S. (2020). Spatial distribution and characteristics of HIV clusters in Ethiopia. *Tropical Medicine & International Health*, *25*(3), 301-307.
338. Yu, Y., Wu, B., Wu, C., Wang, Q., Hu, D., & Chen, W. (2020). Spatial-temporal analysis of tuberculosis in Chongqing, China 2011-2018. *Bmc Infectious Diseases*, *20*(1), 1-12.
339. Yuan, F.-s., Liu, L., Liu, L.-h., Zeng, Y.-l., Zhang, L.-l., He, F., Liu, X.-j., Li, J.-m., Liu, Q., & Xu, M.-j. (2021). Epidemiological and spatiotemporal analyses of HIV/AIDS prevalence among older adults in Sichuan, China between 2008 and 2019: A population-based study. *International Journal of Infectious Diseases*, *105*, 769-775.

340. Yue, Y., Liu, X., Xu, M., Ren, D., & Liu, Q. (2019). Epidemiological dynamics of dengue fever in mainland China, 2014–2018. *International Journal of Infectious Diseases*, *86*, 82-93.

341. Zaragoza Bastida, A., Hernández Tellez, M., Bustamante Montes, L. P., Medina Torres, I., Jaramillo Paniagua, J. N., Mendoza Martínez, G. D., & Ramírez Durán, N. (2012). Spatial and temporal distribution of tuberculosis in the State of Mexico, Mexico. *The Scientific World Journal*, *2012*.

342. Zhang, J.-C., Liu, W.-D., Liang, Q., Hu, J.-L., Norris, J., Wu, Y., Bao, C.-J., Tang, F.-Y., Huang, P., & Zhao, Y. (2014). Spatial distribution and risk factors of influenza in Jiangsu province, China, based on geographical information system. *Geospatial Health*, *8*(2), 429-435.

343. Zhang, W.-Y., Wang, L.-Y., Ding, F., Hu, W.-B., Soares Magalhaes, R. J., Sun, H.-L., Liu, Y.-X., Liu, Q.-Y., Huang, L.-Y., & Clements, A. C. (2013). Scrub typhus in mainland China, 2006–2012: the need for targeted public health interventions. *Plos Neglected Tropical Diseases*, *7*(12), e2493.

344. Zhang, X., Rao, H., Wu, Y., Huang, Y., & Dai, H. (2020). Comparison of spatiotemporal characteristics of the COVID-19 and SARS outbreaks in mainland China. *Bmc Infectious Diseases*, *20*(1), 1-7.

345. Zhang, X., Tang, W., Li, Y., Mahapatra, T., Feng, Y., Li, M., Chen, F., Li, P., Xing, J., & Qian, S. (2017). The HIV/AIDS epidemic among young people in China between 2005 and 2012: results of a spatial temporal analysis. *Hiv Medicine*, *18*(3), 141-150.

346. Zhang, Y.-H., Ge, L., Liu, L., Huo, X.-X., Xiong, H.-R., Liu, Y.-Y., Liu, D.-Y., Luo, F., Li, J.-L., & Ling, J.-X. (2014). The epidemic characteristics and changing trend of hemorrhagic fever with renal syndrome in Hubei Province, China. *Plos One*, *9*(3), e92700.

347. Zhang, Y., Li, L., Dong, X., Kong, M., Gao, L., Dong, X., & Xu, W. (2014). Influenza surveillance and incidence in a rural area in China during the 2009/2010 influenza pandemic. *Plos One*, *9*(12), e115347.

348. Zhang, Y., Shen, Z., Ma, C., Jiang, C., Feng, C., Shankar, N., Yang, P., Sun, W., & Wang, Q. (2015). Cluster of human infections with avian influenza A (H7N9) cases: a temporal and spatial analysis. *International Journal of Environmental Research and Public Health*, *12*(1), 816-828.

349. Zhang, Y., Xiao, Q., Zhou, L., Ma, D., Liu, L., Lu, R., Yi, D., & Yi, D. (2015). The AIDS epidemic and economic input impact factors in Chongqing, China, from 2006 to 2012: a spatial-temporal analysis. *Bmj Open*, *5*(3), e006669.

350. Zhang, Y., Zhang, M., Kang, D., Sun, W., Yang, C., & Wei, R. (2021). Spatio-temporal analysis of bacillary dysentery in Sichuan province, China, 2011–2019. *Bmc Infectious Diseases*, *21*(1), 1-10.

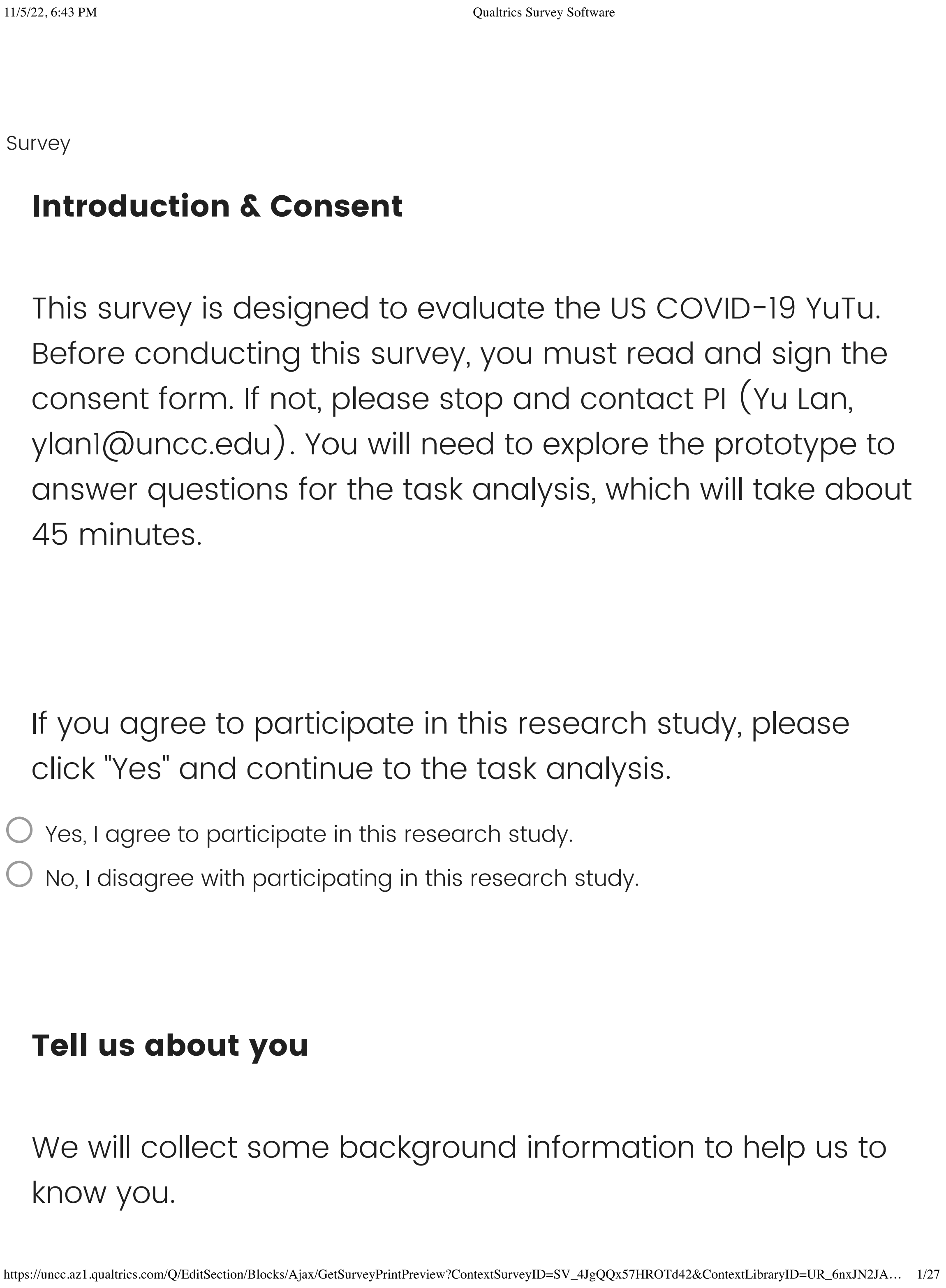
351. Zhu, B., Fu, Y., Liu, J., & Mao, Y. (2017). Notifiable sexually transmitted infections in China: epidemiologic trends and spatial changing patterns. *Sustainability*, *9*(10), 1784.

352. Zhu, B., Liu, J., Fu, Y., Zhang, B., & Mao, Y. (2018). Spatio-temporal epidemiology of viral hepatitis in China (2003–2015): implications for prevention and control policies. *International Journal of Environmental Research and Public Health*, *15*(4), 661.

353. Zhu, H., Zhao, H., Ou, R., Zeng, Q., Hu, L., Qiu, H., Sharma, M., & Ye, M. (2020). Spatiotemporal Epidemiology of Varicella in Chongqing, China, 2014–2018. *International Journal of Environmental Research and Public Health*, *17*(2), 662.

354. Zhu, Y., Xu, Q., Lin, H., Yue, D., Song, L., Wang, C., Tian, H., Wu, X., Xu, A., & Li, X. (2013). Spatiotemporal analysis of infant measles using population attributable risk in Shandong Province, 1999–2008. *Plos One*, *8*(11), e79334.

APPENDIX B: THE ONLINE SURVEY



11/5/22, 6:43 PM Qualtrics Survey Software

Survey

Introduction & Consent

This survey is designed to evaluate the US COVID-19 YuTu. Before conducting this survey, you must read and sign the consent form. If not, please stop and contact PI (Yu Lan, ylan1@uncc.edu). You will need to explore the prototype to answer questions for the task analysis, which will take about 45 minutes.

If you agree to participate in this research study, please click "Yes" and continue to the task analysis.

○ Yes, I agree to participate in this research study.
○ No, I disagree with participating in this research study.

Tell us about you

We will collect some background information to help us to know you.

https://uncc.az1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_4JgQQx57HROTd42&ContextLibraryID=UR_6nxJN2JA… 1/27

To which gender identity do you most identify?

○ Male
○ Female
○ Non-binary / third gender
○ Prefer not to say

Choose your age group.

○ 18-24
○ 25-33
○ 34-44
○ 45-54
○ 55-65
○ 66 and above

My major is

○ Health Administration
○ Public Health (Sciences)
○ Health Services Research
○ ☐ Others

I am pursuing a degree at the level of

○ Master's
○ Doctoral
○ Others

I consider myself knowledgeable in epidemiology.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

I consider myself knowledgeable in understanding relative risk in epidemiology.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

I consider myself knowledgeable in interpreting clusters in epidemiology.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

I consider myself knowledgeable in interpreting space-time clusters in epidemiology.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

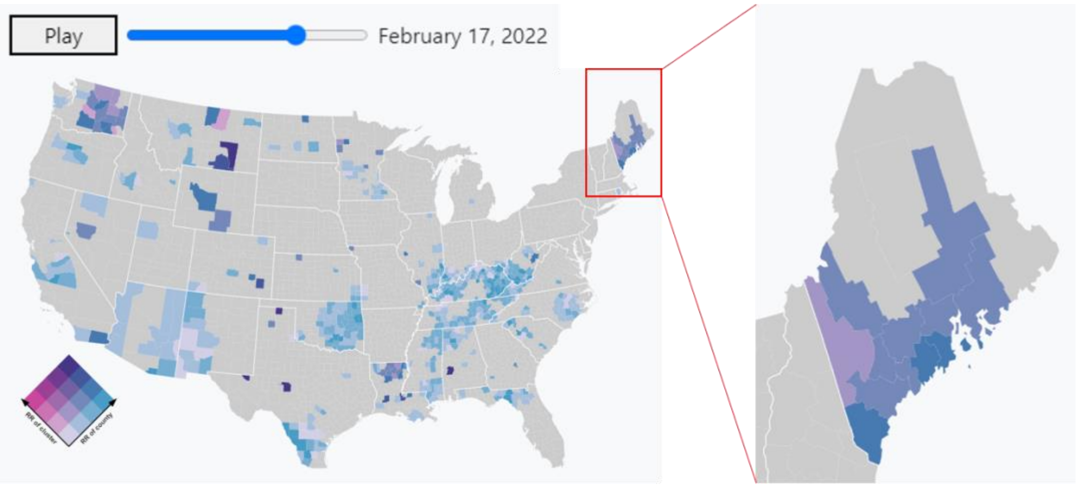
I consider myself knowledgeable in interpreting results of local indicators of spatial autocorrelation (LISA) in epidemiology.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

Warm-up

Here are a few warm-up exercises that ensure you understand the basics of the bivariate map.

Please open the US COVID YuTu for the survey. On the Home page, go to February 17, 2022 using the slide. Look at the cluster in the state of Maine, as shown in the screenshot.


Play
February 17, 2022
RR of cluster
RR of county

[illegible]

○ [illegible]

○ [illegible]

○ [illegible]

[illegible]

☐ [illegible]

☐ [illegible]

☐ [illegible]

☐ [illegible]

☐ [illegible]

☐ [illegible]

Click on the "@LISA" tab. Look at the same date and same regions from the right map (using LISA).

What were the p value and the group that the county we selected for the previous question? Check two answers, one for each column.

☐ 0.01
☐ 0.001
☐ 0.0001

☐ Low-High
☐ High-Low
☐ High-High

Bivariate Map

The task analysis will ask you to perform the required task and answer task-relevant questions. Please open the US COVID YuTu for the survey, which only includes data from 01/20/2022 - 02/28/2022 for some web pages.

On the "Home" page, interact with the cluster map using space-time scan statistics (SaTScan) by playing the animation or hovering to get details. Find the California

State (center in the screenshot). Please answer the following questions.

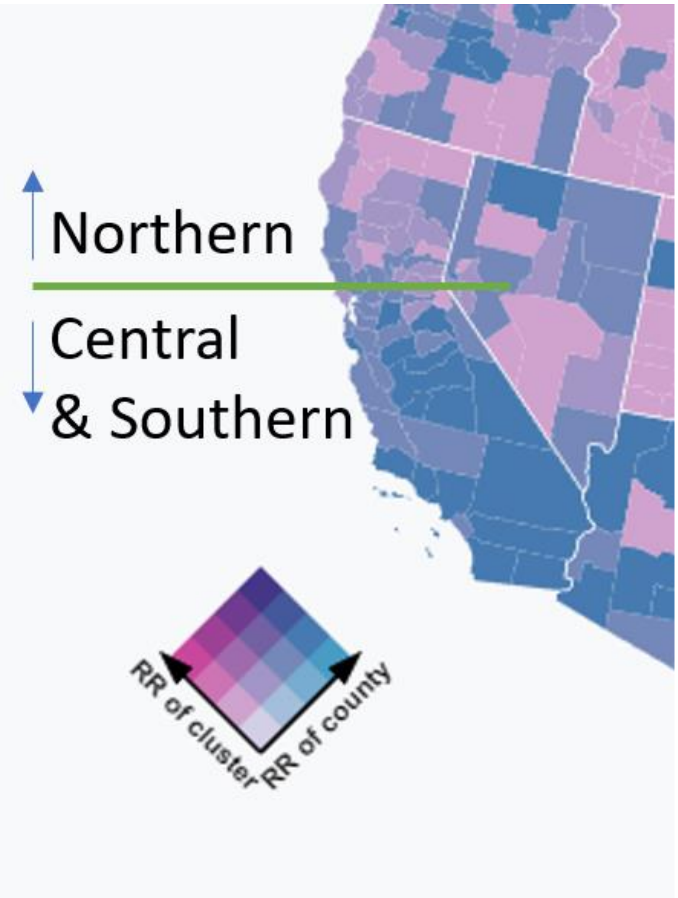
What was the time range when more than half of counties in California belongs to a cluster?

○ January 20, 2022 ~ February 14, 2022

○ January 28, 2022 ~ February 13, 2022

○ January 25, 2022 ~ February 15, 2022

On February 4th, 2022 as shown in the screenshot, what patterns do you recognize from the bivariate map?



○ Counties in Northern had a higher relative risk compared with other counties.

○ Counties in Central and Southern had a higher relative risk compared with other counties.

○ All the counties had the same relative risk (RR of county) as they were in the same cluster.

○ ☐ Others

Please select the answer that you agree most.

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I felt confident in my answers to the previous two questions.	○	○	○	○	○
The bivariate map helps me to interpret the different relative risks among counties within the same cluster.	○	○	○	○	○
I think the "Bivariate Map" with results from SaTScan is useful.	○	○	○	○	○

Imagine you are a health policy maker in California State, which of the following statements do you agree with the most?

○ I would allocate more health resources to Northern California.

○ ✋⬧ □◆●♎ ♋●□♏♐⧫♏ ❍ □□♏ ♒♏♋⧫♒□♏⬧ □◆⧫□♏⬧ ⧫□👍 ♏■⧫♐●♋■♎ 💧□◆⧫♒♏□■ 👍 ♋●⌧□■⌘♋📭

○ ✋⬧ □◆●♎ ♋●□♏♐⧫♏ ♏❑◆♋●●♒♏♋⧫♒□♏⬧ □◆⧫□♏⬧ ⧫□ ♋●●⧫♒♏ ♏□◆■⧫♓♏⬧📭

✌■◻⧫♒□⧫♒♏⬧ ♋♌□◆⧫ ⧫♒♏ ♌♓♐♐♏♏ ❍ ♋□✍

LISA Map

☝ □ ⧫□ ⧫♒♏ ⬧ ♏♌□♋♑♏ □⌧✂ ☹💧✌✂⧫♋♌📫♋■♎ ♏⌧◻●□♏ ⧫♒♏ ⬧ ♓ ♓●♋⧫◻ ♋■♎ ♎♓⬧ ♓ ♓●♋⧫◻♌♏⬧ ♏■ ⧫♒♏ ●♏⌧♋■♎ ♐♓♑♒⧫❍ ♋□⬧ 🖃♐●♏♋⬧♏ ♋■⬧ ⬧ ♏□⧫♒♏ ⌧♐●□⬧ ♓♑◆□⧫♏⬧ ⬥♓□⬧ ♌◻♏□❍ ♌♓■♓♑⧫ □ ❍ ♋□⬧ 📭

⌖ ♒♏■ ♏□❍ □♋♐♓♑⧫♒♏ ❍ ♋□ □■ ⧫♒♏ ♐♓♑♒⧫☎☹💧✌✆♋■♎ ⧫♒♏ □■♏ □■ ⧫♒♏ ●♏⌧☎♋⬧ ♏♋■✆📫 ♏ ♐♏♋⬧♏ ⧫♒♋⧫⧫♒♏◻♎□ ■□⧫♋● ♋◻ ♏□■♏♋♎♏🖃⬧ ♓♑⧫♒♏⬧ ⧫♋♏ □⌧👍 ♋●⌧□■⌘♋📭for which of these dates ♎♓♎◻□⧫■□⧫♓♏ ♏●⧫ ⬧♏□■ ⧫♒♏ ♐♓♑♒⧫❍ ♋□📫♌⧫⧫■□ ♏●⧫ ⬧♏□■ ⧫♒♏ ●♏⌧❍ ♋□✍📬

○ January 20, 2022 ~ January 27, 2022

○ February 1, 2022 ~ February 14, 2022

○ January 20, 2022 ~ February 14, 2022

What are the reasons that the results from the two maps are different?

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
The algorithm used to calculate space-time cluster as shown on the left map uses 50 days of data; as such the change in Relative Risk is not as sudden.	○	○	○	○	○
The algorithm used to calculate the spatial cluster as shown on the right map uses data for one day. As such, the change in risk can be abrupt.	○	○	○	○	○
The right map could help me to identify when a cluster is emerging, while the algorithm on the left map cannot detect it yet	○	○	○	○	○

The left map shows true space-time clusters, while the right map shows the repetitions of spatial clusters at the current date. Imagine you are a health policy maker for the state of California, what do you think about the following statements?

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I will use the right map to alarm counties located within a high-high cluster, even though the left map did not detect any clusters within the region quite yet.	○	○	○	○	○
I will use the left map to monitor the dynamic of space-time clusters, even though counties do not belong to a high-high cluster as displayed on the right map.	○	○	○	○	○
I will rely on the information provided on both maps to make an informed decision.	○	○	○	○	○

Do you agree with these statements?

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I think both maps provided under the “LISA” tab allow for a comparison of clustering techniques.	○	○	○	○	○
I think the single map under the “HOME” tab is easier to interpret than when having two maps side-by-side.	○	○	○	○	○

Any thoughts about this set of maps?

Spiral Map

Go to the webpage of "@State". The spiral map of average relative risk of states. Explore the spiral map by clicking both the bivariate map and the spiral map. Then click any county in Washington state from the left bivariate map or select the state in the center of the spiral map. Washington is located at the northwestern corner of the 48 conterminous states.

What was the highest average relative risk in the State of Washington and what date was it on? Check two answers, one for each column.

☐ the highest average relative risk was 3.7413	☐ 2020-01-24
☐ the highest average relative risk was 5.4231	☐ 2020-06-23
☐ the highest average relative risk was 5.8464	☐ 2020-03-05

Which state has the highest average relative risk yesterday from the center of the spiral map?

☐

I felt confident in my answers to the above two questions.

Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

Do you agree with these statements?

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
The average relative risk in Washington state was very high at the beginning of the pandemic.	○	○	○	○	○
The connection between the bivariate map with the spiral map is useful to explore the average relative risk at the state level.	○	○	○	○	○
The center state map from the spiral is useful for me to explore the latest average relative risk at the state level compared with other states.	○	○	○	○	○
I think the “spiral" map is useful to show the patterns at the state level.	○	○	○	○	○

Imagine you are a health policy maker for the Centers for Disease Control and Prevention (CDC) and could decide on policy decisions for the entire US, what do you think about the following statements?

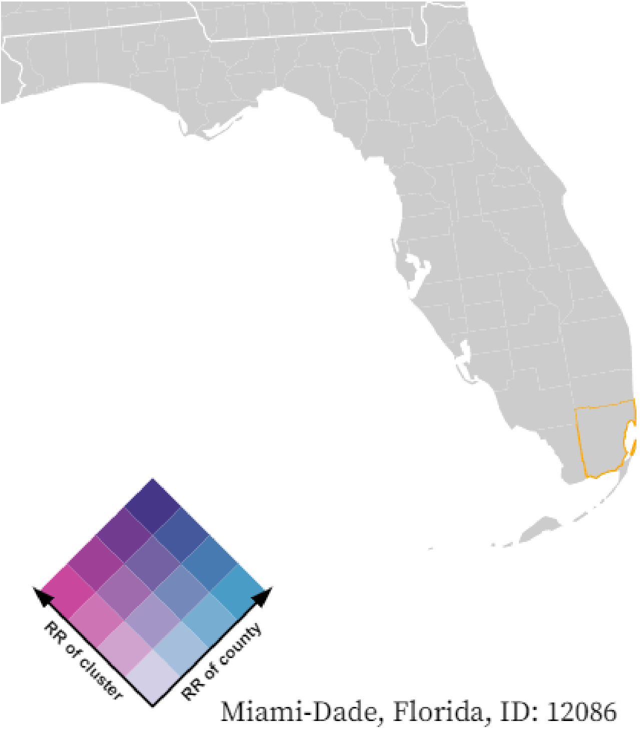
	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
In the state of Washington, I would keep the current health policy on COVID-19 as the average relative risk seems to be stable at a lower level (less than 1).	○	○	○	○	○
The spiral map allows us to compare and contrast how well some states respond to some new policies (e.g. lockdown, loosening social distancing interventions).	○	○	○	○	○

Any thoughts about the spiral map?

☐

Time Chart

Go to the tab "@TimeChart". This visualization shows the linear chart of variables during the whole study time. Click on the Miami-Dade County on the bivariate map. Miami-Dade County is located in the south of the state of Florida.



Looking at the red-colored time chart -which summarizes the relative risk in that county-, in which months did Miami-Date experienced its highest value?

○ July and August 2020
○ July and August 2021
○ July and August 2022

Make sure the "RR within the cluster" is selected (default). The blue colored time chart summarizes the relative risk of the *cluster* to which the county belongs at a particular time.

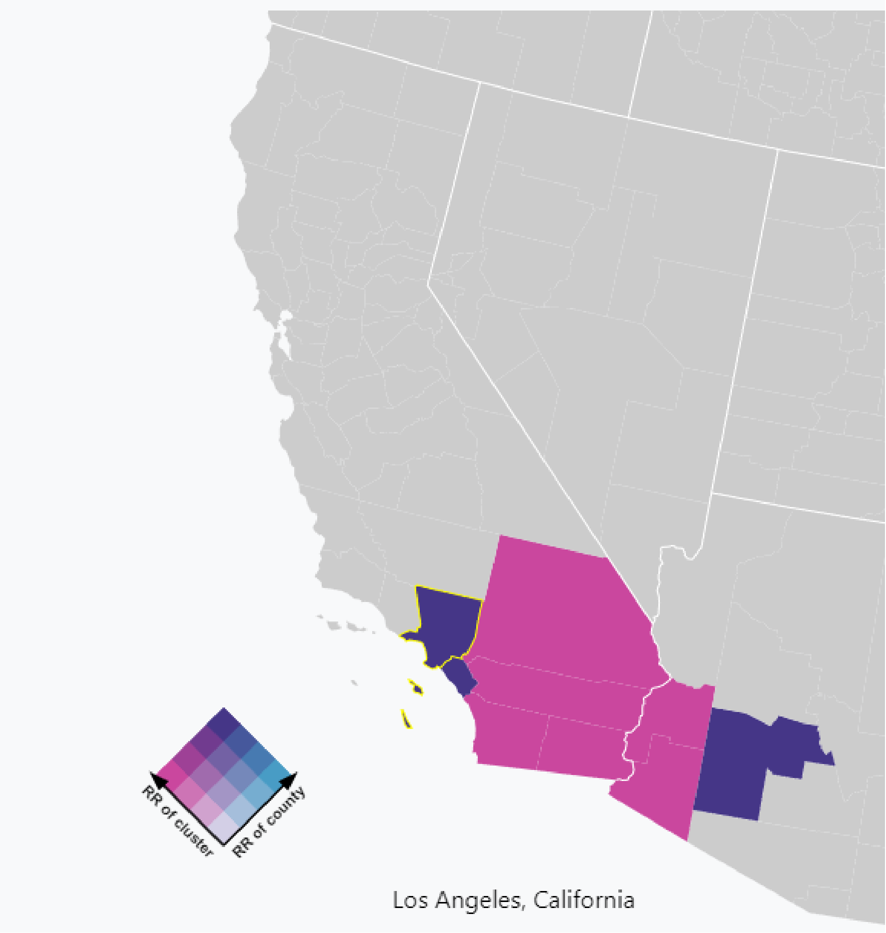
At what time was the RR of the *cluster* the highest in the blue color time chart?

○ 2022-07-01

○ 2021-08-02

○ 2020-03-21

Click on the Los Angeles County, California. It is located to the south of the State of California (highlighted in yellow in the following screenshot).



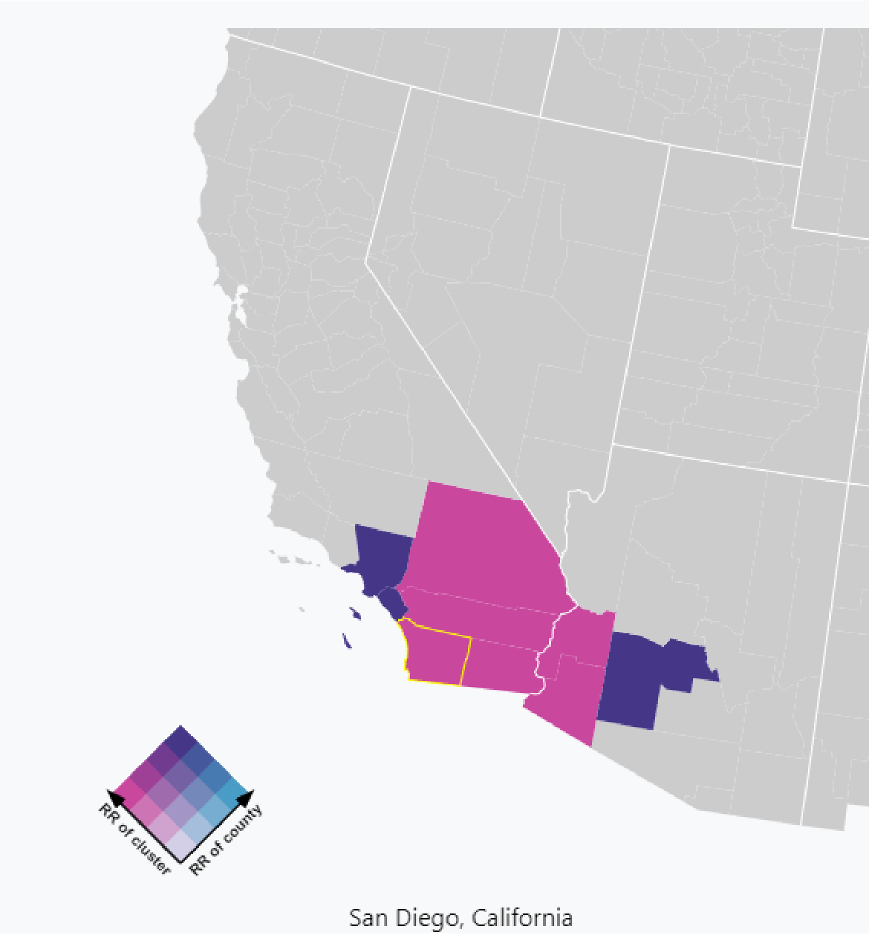
Make sure the "RR within the cluster" is selected (default). Only focus on the blue time chart.

Which of the two counties has more often been part of a cluster?

○ Los Angeles County

○ Miami-Dade County

Click on the San Diego County, California. It is also located in the southern of California (highlighted in yellow in the following screenshot).



Look at the changes by switching between "Daily Cases 7-days avg" and "RR within the cluster". From the TimeChart, what patterns do you find between those two counties in California?

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
If you toggle back and forth between the variables "Daily Cases 7-days avg" and "RR within the cluster" (blue chart), the temporal patterns are very similar overall.	○	○	○	○	○
As to RR of location (red chart), Los Angeles had the highest RR at the beginning of the pandemic, while San Diego had the highest RR around August 2022.	○	○	○	○	○

Please select the answer that best represents your feeling.

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I felt confident in my answers to the above three questions.	○	○	○	○	○
I think the "TimeChart" is useful to identify the trend of different variables (e.g. relative risk).	○	○	○	○	○
I think the "TimeChart" is useful to compare counties	○	○	○	○	○

Imagine you are a health policy maker in Los Angeles County, what do you think about the following statement?

I will carefully monitor the RR of the county (red chart) as it is still high although the RR within the cluster and daily cases -as represented by the blue charts- are decreasing.

Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

Any thoughts on the TimeChart?

3D

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
○	○	○	○	○

[illegible]

what is the value (check two answers, one for each column)?

☐ Loving County, TX	☐ rr_l = 11.23
☐ Dewey County, SD	☐ rr_l = 20.09
☐ Brown County, IL	☐ rr_l = 25.22

I felt confident in my answers to the above question.

Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

Do you agree with the following statements?

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I think the "3D space-time cube" is useful.	○	○	○	○	○
I think the "3D space-time cube" is more useful than the animated bivariate map.	○	○	○	○	○

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I think the “3D space-time cube” is more useful than the bivariate map combined with timecharts (from "@TimeChart" tab).	○	○	○	○	○

Imagine you are a health policy maker for CDC, what do you think about the following policies?

	Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I will use the filter function to 3D space-time cube to warn the counties with a higher relative risk (e.g., RR>=4).	○	○	○	○	○
I will investigate why space-time clusters were shifted from Eastern to Western.	○	○	○	○	○

Any thoughts about the 3D space-time cube?

Do you agree that the provided information (ReadMe) is helpful to interpret results and answer the questions?

Strongly Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

Please provide any suggestions or comments that you have for this system.

Powered by Qualtrics

APPENDIX C: IRB APPROVAL NOTICE

CHARLOTTE
RESEARCH AND ECONOMIC DEVELOPMENT

To:	Yu Lan Graduate School
From:	Office of Research Protections and Integrity
RE:	Notice of Exemption with Limited Review Approval
Approval Date:	25-Aug-2022
Exemption Category:	2~3
Study #:	IRB-22-0219
Study Title:	Evaluation of A Web-based Geographic Framework to Detect and Visualize Space-time Clusters of Infectious Diseases

This submission has been reviewed by the Office of Research Protections and Integrity (ORPI) and was determined to meet the Exempt category cited above under 45 CFR 46.104(d). In addition, this Exemption has received Limited Review by the IRB under 45 CFR 46.111(a)(7). This determination has no expiration or end date and is not subject to an annual continuing review. However, you are required to obtain IRB approval for all changes to any aspect of this study before they can be implemented and to comply with the Investigator Responsibilities detailed below.

Important Information:

1. Face masks are optional on UNC Charlotte's campus. This includes classrooms and other academic spaces. Researchers conducting HSR activities in other locations must continue to adhere to local and state requirements in the setting where the research is conducted.
2. Face masks are still required in healthcare settings. Researchers conducting HSR activities in these settings must continue to adhere to face coving requirements.
3. Organizations, institutions, agencies, businesses, etc. may have further site-specific requirements such as continuing to have a mask requirement, limiting access, and/or physical distancing. Researchers must adhere to all requirements mandated by the study site.

Your approved study documents are available online at Submission Page.

Investigator's Responsibilities:

1. Amendments must be submitted for review and the amendment must be approved before implementing the amendment. This includes changes to study procedures, study materials, personnel, etc.

2. Data security procedures must follow procedures as approved in the protocol and in accordance with OneIT Guidelines for Data Handling.
3. Promptly notify the IRB (uncc-irb@uncc.edu) of any adverse events or unanticipated risks to participants or others.
4. Five years (5) following this approval, ORPI will request a study status update.
5. Be aware that this study is now included in the Office of Research Protections and Integrity (ORPI) Post-Approval Monitoring program and may be selected for post-review monitoring at some point in the future.
6. Reply to ORPI post-review monitoring and administrative check-ins that will be conducted periodically to update ORPI as to the status of the study
7. Complete the Closure eform via IRBIS once the study is complete

Please be aware that additional approvals may still be required from other relevant authorities or "gatekeepers" (e.g., school principals, facility directors, custodians of records).

This study was reviewed in accordance with federal regulations governing human subjects research, including those found at 45 CFR 46 (Common Rule), where applicable.

APPENDIX D: CONCENT FORM FOR STUDENTS

UNIVERSITY OF NORTH CAROLINA
CHARLOTTE
Department of Geography and Earth Sciences
9201 University City Boulevard, Charlotte, NC 28223-0001

Consent to Participate in a Research Study

Title of the Project: Evaluation of A Web-based Geographic Framework to Detect and Visualize Space-time Clusters of Infectious Diseases

Principal Investigator: Yu Lan, Ph.D. Candidate, University of North Carolina at Charlotte and Center for Applied Geographic Information Science, USA

Co-investigators: Eric Delmelle, Ph.D., University of North Carolina at Charlotte, USA;
Deborah Thomas, Ph.D., University of North Carolina at Charlotte, USA

Study Sponsor: None

You are invited to participate in a research study. Participation in this research study is voluntary. The information provided is to help you decide whether or not to participate. If you have any questions, please ask.

Important Information You Need to Know

- The purpose of this research is to evaluate a prototype of a web-based geographic framework to detect and visualize space-time clusters of infectious diseases.

- We are asking graduate students from UNCC to participate in a survey that participants will follow the instruction to use the prototype and answer questions after exploring some functionalities of the prototype.

- You have the right to refuse participation at any time and the information collected from you will be deleted from the investigation.

- Please read this form and ask any questions you may have before you decide whether to participate in this research study.

Why are we doing this study?

Infectious diseases pose a significant threat to public health worldwide as evidenced by the recent coronavirus 2019 (COVID-19) pandemic. Despite significant human losses, the advent of web-accessed, map-based “data dashboards” that monitored disease outbreaks, proved essential in managing public health responses. In many cases, the backend of these dashboards employed basic mapping functionality, displaying counts or rates. As the pandemic advanced, the identification of elevated rates was increasingly important in the geographical allocation of public health resources. However, such maps miss the opportunity to provide accurate information to policy decision makers such as the rate of disease spread, cyclicity, direction, intensity and the risk of diffusion to new regions. Space-time geoanalytics, when coupled with rich visualizations, can address these shortcomings. Moreover, when implemented over the web, such space-time functionality can be accessed from virtually anywhere. Thus, my dissertation presents a web-based geographic framework for detecting and visualizing explicit space-time clusters of infectious diseases. As a proof of concept, I will populate the framework with COVID-19 county-level data for the 48 contiguous states in the US, and demonstrate data retrieval and storage, space-time cluster detection analysis, and 3D visualization within an open-source WebGIS environment. This tightly coupled approach coupling approach will facilitate the detection of space-time clusters of diseases in a computationally acceptable timeframe.

Why are you being asked to be in this research study?

You are being asked to be in this study because you are graduate students, who are capable to use the prototype and interpret results as it was designed to assist health policymakers during the pandemic.

What will happen if I take part in this study?

If you choose to participate you will be asked to take a survey with multiple questions. Questions include your gender, age group, major, degree level, the knowledge level in related background, questions related to prototype, and usability and user experience. There will be no time limit, and it usually will take 0.5~1 hours.

What benefits might I experience?

You will not benefit directly from being in this study. Others might benefit because it is important to understand the space-time clusters of COVID-19. The results can improve educational resources and improve health policy to help protect the community from disease.

What risks might I experience?

There are no risks if you decide to participate in this study.

How will my information be protected?

This collected response from the survey will be stored and analyzed. We are not collecting your name, address, or any information that will identify you. You will be assigned an identification number for reference for the analysis. While the study is active, the researchers listed above will only have access to the results; while the analysis results will be stored in a password-protected Google Drive that can be accessed by the primary researcher. Only the research team will have routine access to the study data. Other people with approval from the Investigator, may need to see the information we collect about you. Including people who work for UNC Charlotte and other agencies as required by law or allowed by federal regulations. Your individual privacy will be maintained in all written and published material resulting from the research.

How will my information be used after the study is over?

After this study is complete, study data may be shared with other researchers for use in other studies or as may be needed as part of publishing our results. The data we share will NOT include information that could identify you.

Will I receive an incentive for taking part in this study?

You will be paid a gift card ($25) to participate in this study, and it will be given to you after you submit all your responses.

What other choices do I have if I don't take part in this study?

You do not have to participate in this study. If you are interested in the results of this study, you can contact the principal investigator.

What are my rights if I take part in this study?

It is up to you to decide to be in this research study. Participating in this study is voluntary. Even if you decide to be part of the study now, you may change your mind and stop at any time. You do not have to answer any questions you do not want to answer.

Who can answer my questions about this study and my rights as a participant?

For questions about this research, you may contact Yu Lan – email: ylan1@uncc.edu, phone: +1(704)499-0107 and Dr. Eric Delmelle – email: Eric.Delmelle@uncc.edu.

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Office of Research Protections and Integrity at +1 (704)-687-1871 or uncc-irb@uncc.edu.

Consent to Participate

If you agree to participate in this research study, please reply to the email and paste the following sentences in your replied email.

"I have read and I understand the provided information and have had the opportunity to ask questions. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason and without cost."

Be sure that any questions have been answered clearly to you and that you have a thorough understanding of the study.

Please ask any questions regarding this study's objectives and your participation.

If you have further questions that come up later, please feel free to ask a member of the research team.

APPENDIX E: CONCENT FORM FOR EXPERTS

UNIVERSITY OF NORTH CAROLINA
CHARLOTTE
Department of Geography and Earth Sciences
9201 University City Boulevard, Charlotte, NC 28223-0001

Consent to Participate in a Research Study

Title of the Project: Evaluation of A Web-based Geographic Framework to Detect and Visualize Space-time Clusters of Infectious Diseases

Principal Investigator: Yu Lan, Ph.D. Candidate, University of North Carolina at Charlotte and Center for Applied Geographic Information Science, USA

Co-investigators: Eric Delmelle, Ph.D., University of North Carolina at Charlotte, USA;
Deborah Thomas, Ph.D., University of North Carolina at Charlotte, USA

Study Sponsor: None

You are invited to participate in a research study. Participation in this research study is voluntary. The information provided is to help you decide whether or not to participate. If you have any questions, please ask.

Important Information You Need to Know

- The purpose of this research is to evaluate a prototype of a web-based geographic framework to detect and visualize space-time clusters of infectious diseases.

- We are asking experts in public health to participate in a semi-structured group interview that participants will answer questions and have a discussion with other participants after a presentation of this system in the beginning.

- You have the right to refuse participation at any time and the information collected from you will be deleted from the investigation.

- Please read this form and ask any questions you may have before you decide whether to participate in this research study.

Why are we doing this study?

Infectious diseases pose a significant threat to public health worldwide as evidenced by the recent coronavirus 2019 (COVID-19) pandemic. Despite significant human losses, the advent of web-accessed, map-based “data dashboards” that monitored disease outbreaks, proved essential in managing public health responses. In many cases, the backend of these dashboards employed basic mapping functionality, displaying counts or rates. As the pandemic advanced, the identification of elevated rates was increasingly important in the geographical allocation of public health resources. However, such maps miss the opportunity to provide accurate information to policy decision makers such as the rate of disease spread, cyclicity, direction, intensity and the risk of diffusion to new regions. Space-time geoanalytics, when coupled with rich visualizations, can address these shortcomings. Moreover, when implemented over the web, such space-time functionality can be accessed from virtually anywhere. Thus, my dissertation presents a web-based geographic framework for detecting and visualizing explicit space-time clusters of infectious diseases. As a proof of concept, I will populate the framework with COVID-19 county-level data for the 48 contiguous states in the US, and demonstrate data retrieval and storage, space-time cluster detection analysis, and 3D visualization within an open-source WebGIS environment. This tightly coupled approach coupling approach will facilitate the detection of space-time clusters of diseases in a computationally acceptable timeframe.

Why are you being asked to be in this research study?

You are being asked to be in this study because you are experts in public health, who are capable to use interpret results as it was designed to assist health policymakers during the pandemic.

What will happen if I take part in this study?

If you choose to participate you will be asked to participate a semi-structed group interview with other recruited experts. The interview will be held as a zoom meeting, and it will be recorded and later transcribed for subsequent analysis. The interview should last around one hour.

What benefits might I experience?

You will not benefit directly from being in this study. Others might benefit because it is important to understand the space-time clusters of COVID-19. The results can improve educational resources and improve health policy to help protect the community from disease.

What risks might I experience?

There are no risks if you decide to participate in this study.

How will my information be protected?

This collected response from the survey will be stored and analyzed. We are not collecting your name, address, or any information that will identify you. You will be assigned an identification number for reference for the analysis. While the study is active, the researchers listed above will only have access to the results; while the analysis results will be stored in a password-protected Google Drive that can be accessed by the primary researcher. Only the research team will have routine access to the study data. Other people with approval from the Investigator, may need to see the information we collect about you. Including people who work for UNC Charlotte and other agencies as required by law or allowed by federal regulations. Your individual privacy will be maintained in all written and published material resulting from the research.

How will my information be used after the study is over?

After this study is complete, study data may be shared with other researchers for use in other studies or as may be needed as part of publishing our results. The data we share will NOT include information that could identify you.

Will I receive an incentive for taking part in this study?

There are no incentives and you will not be paid to participate in this study.

What other choices do I have if I don't take part in this study?

You do not have to participate in this study. If you are interested in the results of this study, you can contact the principal investigator.

What are my rights if I take part in this study?

It is up to you to decide to be in this research study. Participating in this study is voluntary. Even if you decide to be part of the study now, you may change your mind and stop at any time. You do not have to answer any questions you do not want to answer.

Who can answer my questions about this study and my rights as a participant?

For questions about this research, you may contact Yu Lan – email: ylan1@uncc.edu, phone: +1(704)499-0107 and Dr. Eric Delmelle – email: Eric.Delmelle@uncc.edu.

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Office of Research Protections and Integrity at +1 (704)-687-1871 or uncc-irb@uncc.edu.

Consent to Participate

If you agree to participate in this research study, please reply to the email and paste the following sentences in your replied email.

"I have read and I understand the provided information and have had the opportunity to ask questions. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason and without cost."

Be sure that any questions have been answered clearly to you and that you have a thorough understanding of the study.

Please ask any questions regarding this study's objectives and your participation.

If you have further questions that come up later, please feel free to ask a member of the research team.

APPENDIX F: THE CODEBOOK OF SCRIPTS FROM THE GROUP INTERVIEW

Note: There are three scripts (files) for each group task.

Codes

Name	Description	Files	References
audience	Discussion on the audience of the evaluated dashboard or system	3	15
cluster	Discussion on cluster	3	17
cons	Discussion on the disadvantages of the evaluated dashboard or system	3	25
data	Discussion on the data from the evaluated dashboard or system	2	8
maps	Discussion on the maps from the evaluated dashboard or system	2	3
pros	Discussion on the advantages of the evaluated system or dashboard	3	15
scale	Discussion on the scales from the evaluated system or dashboard	2	9
time	Discussion on the temporal information or representation from the evaluated system or dashboard	2	7
usefulness	Discussion on the usefulness of the evaluated system or dashboard	3	20