AGILE: GIScience Series, 2, 18, 2021. https://doi.org/10.5194/agile-giss-2-18-2021
Proceedings of the 24th AGILE Conference on Geographic Information Science, 2021.
Editors: Panagiotis Partsinevelos, Phaedon Kyriakidis, and Marinos Kavouras.
This contribution underwent peer review based on a full paper submission.

Public health needs GIScience (like now)

Justine I. Blanford[a] (corresponding author), and Ann M. Jolly[b]

j.i.blanford@utwente.nl, ann@contagionconsulting.com

[a]Geo-Information Science and Earth Observation (ITC) , University of Twente, Enschede, Netherlands

[b]Contagion Consulting, Ottawa, Canada

Abstract. During the last 20 years we have seen the re-emergence of diseases; emergence of new diseases in new locations and witnessed outbreaks of varying intensity and duration. Spatial epidemiology plays an important role in understanding the patterns of disease and how they change over time and across space.

The aim of this paper is to bring together a public health and geospatial data science perspective to provide a framework that will facilitate the integration of geographic information and spatial analyses at different stages of public health response so that these data and methods can be effectively used to enhance surveillance and monitoring, intervention strategies (planning and implementation of a response) and facilitate both short- and long-term forecasting.

To demonstrate elements of this framework and how it can be utilized, we selected three case studies ranging from the current the global COVID-19 Coronavirus pandemic of 2020 to more historical examples such as the John Snow Cholera outbreak of 1854 and the Ebola outbreak of 2014 in West Africa.

A variety of methods including spatial descriptive statistics, as well as methods for analysing patterns were used. The examples we provide can reveal sources of infection, connectivity between locations, delineate zones of containment and show the spread of an outbreak globally and locally across space and time.

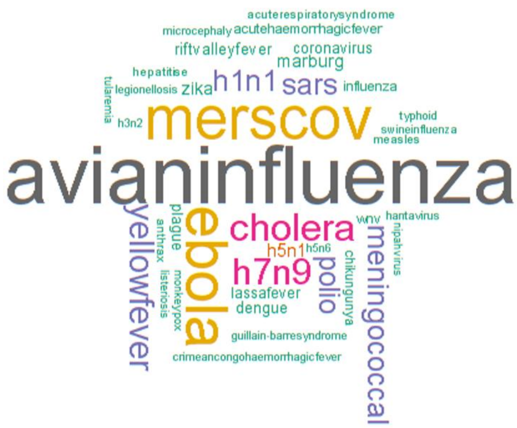
Keywords: geospatial technologies; public health; epidemiology; data science; geography; infectious disease, education

1 Introduction

During the past 20 years we have seen the emergence and re-emergence of many diseases (Figure 1), many in new locations including the current pandemic, COVID-19 (Figure 2).

A key component of staying healthy is to minimize our risk of getting sick. To do so, we want to know how to avoid getting ill by understanding where a disease is, when it is present, if there is a temporal component to its incidence, and what preventative measures we can take to stop us from becoming ill.

Figure 1: Summary of disease outbreaks over the last 20 years (2000-2020) based on the World Health Organization Disease Outbreak News (DONs) Reports (1). The word cloud shows disease based on the number of reports that contain that disease name in the summary report title provided by the DONs (N> 5 reports). Analysis was conducted in R. Diseases include those that may have occurred for a variety of reasons (8, 9) as summarized by the four categories.



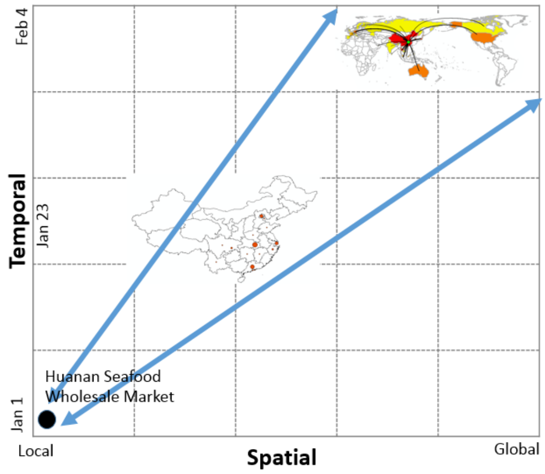
Includes:

- ***emergence of new diseases*** in new locations
- ***evolution of disease*** resulting in the emergence of new pathogens and resistance.
- ***re-emergence of eliminated diseases*** in the same or nearby locations.
- ***regular occurrence of diseases*** in the same location

Patterns influencing health and disease in the environment are complex and require an understanding of the ecology of the disease (agent, host, environment), how these interact in space and time), and how diseases may move through the landscape (mobility, connectivity, and dispersion pathways) so that we can respond (plan and implement control and prevention), and recover (seek diagnosis, prevent, and provide treatment) in a timely manner. This requires understanding the interplay between diseases, their environments, and their hosts (ecology of disease) and how these may change risk over time. We need to think simultaneously about how a disease agent and the host interact at various spatial and temporal scales in a dynamically changing environment and what the outcome of such changes may be.

In recent years, access to novel data sources has been increasing with the availability of new devices that enable data to be collected easily, alongside point-of-care diagnostics, at a precise location in time. These technologies range from mobile device add-ons (e.g. spectrometer), mobile apps, wearable technologies (e.g. GPS watches, Fitbits) and remote sensors (e.g. Wi-Fi loggers collecting a variety of environmental data; unmanned aerial vehicles (UAVs)), many of which have built-in GPS-enabled devices. In the COVID-19 era, apps specifically to help notify people of possible exposures using Bluetooth technology have been developed and are now in use (11). In addition, apps for restaurants and other social venues for patrons to register are being used to facilitate contact tracing (14). Through these data collection avenues, we are able to provide richer and more diverse sources of information about ourselves and the environments in which we live than ever before. Although we have moved into an era of digital exploration, there remain many challenges in using, analysing, integrating, and applying these data, particularly when they vary in quality and availability (both in terms quantity and at rapidity) (15, 16). Leveraging these data and technologies together with existing surveillance methods of humans and animals will be useful for improving our understanding of the mechanisms influencing health across different spatial and temporal scales; enhancing diagnostics and predictions; as well as developing preventative strategies. Furthermore, with increased mobility and the influence of external factors, such as changes in climate and globalization, we need to integrate multiple types of geographic data that capture not only the physical environment, but also human and social environments (e.g. perception, cultural, economic, political). This will facilitate a better understanding of what is happening at a local level, with regional level influences, as illustrated by the recent swift global distribution of the novel coronavirus ((SARS-CoV-2) also known as COVID-19) (17) (Figure 2).

Figure 2: Distribution of the novel coronavirus (COVID-19) 2020 locally, regionally and globally (Data source: (2); ESRI country boundaries).



Coronavirus outbreak of 2020 – In 2020, 75,765 cases of the coronavirus (COVID-2019) were confirmed globally resulting in 2,129 deaths (as of Feb 20th, 2020) (2). The source of the infection was the Huanan Seafood Wholesale Market, Wuhan, China which was shut down in early January to prevent further transmission (2). By Feb 4^{th,} the virus had been transported globally to 24 countries.

Approaches to disease mapping and spatial epidemiology range in complexity from the creation of simple maps (e.g. John Snow's Cholera map of 1854 (5)), graduated points (Figure 2) to deterministic, correlative, geostatistical and geocomputational modelling techniques as summarized in Table 1. For examples (see (18-20); and malaria maps using different methods that include: Suitability analysis (21); Bayesian geostatistical methods ((22); Geocomputational methods with host-pathogen-environment models (23)).

Table 1: Summary of how geospatial information and spatial data methods have been used in health studies (compiled from a variety of sources: (24-30), (31); (32) including COVID-19 (33)).

Type	**Purpose**
	Create / transform
Various	Create geographic data to enable for the visualization of disease risk. Various methods have been used that include conversion of data, transformation of data, geocoding, georeferencing, spatial join, aggregation of data or projection of data.
	Visualization
Carto-graphic maps	Disease maps provide a rapid visual summary of complex geographic information and may identify subtle patterns in the data that are missed in tabular presentations. • Mapping of disease incidence by points or areas (e.g. political boundaries (ward, county, district, province/state, country) to show where and when disease risk and health issues are prevalent • Presence/absence; Counts/Rates (mortality, confirmed cases). • Dot maps; graduated symbols; choropleth maps; density estimation maps

Web-based mapping	Use of the web-mapping tools and dashboards to map disease location and allow for interaction with the data and attributes • Geovisualization, interactive dashboard analytics (e.g. COVID-19 Dashboard used by World Health Organization (WHO) and Johns Hopkins University)
Explore spatially explicit relationships, and evaluate and analyse spatial relationships	
Integration of geographic information and exploration of relationships	Examine where transmissions are taking place in relation to different geographies and information (see cartographic maps, web-based mapping and spatial methods)
Correlation studies	• Examine variations in disease incidence/risk in relation to different geographies • hypothesis-generating, as the unit of observation is the geographic group rather than the individual and associations observed at the group level • useful for developing and exploring hypotheses of public health importance
Cluster Analysis	• Evaluate whether features are clustered, dispersed, or random. • Identify statistically significant hot spots, cold spots, or spatial outliers (where a disease cluster implies an excess of cases above some background rate bounded in time and space) • Useful for searching for unusual patterns • A variety of methods are available (e.g. Morans' I, LISA, Getis Ord, Ripley K, SatScan)
Connectivity	• Physical connectivity ○ Transportation networks (road, rail, flight, water) • Social Networks ○ Dedicated Social Network Analysis to understand how places are connected beyond just the physical connectivity. • Phylogeography ○ Provide information on the genetic similarity and/or evolution of organisms through space and time ○ Useful for identifying source of infection and the role of place, events and networks in the diffusion of diseases
Neighbourhood structure and composition	• the structure and composition of the landscape surrounding focal sites are important for understanding heterogeneity and the influence of variations in local biotic and abiotic features in disease prevalence, risk and diffusion due to the interaction of different populations
Health infrastructure Planning & Accessibility	• Combine location information with population information to assess availability of health facilities. • Combine distance-allocation models with location of health facilities to determine physical accessibility. • Useful for planning of health infrastructure needs (e.g. vaccination programs, availability of health care, accessibility to health care)
Spatiotemporal dynamics of disease	• retrospective analyses of spatiotemporally dynamic epidemics to understand what factors govern the spatial pattern and rate of spread of diseases. • characterize spatial variation in contemporaneous (static) ecological risk of infection and potential causes of that variation. Ecological risk can be defined as the probability of infections risk? • Evaluate dynamically changing risk and/or spatial relationships
GeoAI: Machine Learning, Deep Learning	• Useful for sifting through large quantities of data both historically and in real-time to identify patterns, assess similarity and correlation • Used for syndromic surveillance, analysis of symptoms, sentiment and perceptions. • Assessment of predicted change • Used for understanding behaviour and connectivity between places through analysis of mobility data (travel (flight, train, bus, bikeshare)), mobile phone, GPS data, payment data, social media and other volunteered information (traffic))
Modelling: Simple to advanced geocomputational methods.	
Suitability Mapping	• Determine suitability of environment for disease vectors or pests of disease. Useful when data is limited. Parameter estimations are subjective in nature. • **Multi-criteria decision analysis (MCDA) or decision science:** is used to logically evaluate and compare multiple criteria that may be conflicting. ○ Variety of methods can be used ranging from simple Boolean logic to more complex decision analysis (analytical hierarchical process (AHP), fuzzy logic, weighted overlay) • **Niche Modelling:** Variety of methods and tools are available (e.g. ecological niche models)
Spatially Explicit Models	• **Spatial Interpolation and Smoothing Methods:** Interpolation and smoothing methods applied to spatial epidemiology, are useful for improving estimation of risk across a surface by creating a continuous surface from sampled data points (filling in where data are unobserved) or to smooth across polygons (aggregate data). ○ Variety of methods ranging in complexity are available (Inverse Distance Weighted (IDW), Spline, Natural Neighbor, Trend (polynomial), Kriging (Geostatistical method) • **Mathematical Models:** Useful for determining risk and changing risks; impact of interventions on disease transmission where multiple scenarios can be studied and compared. Geocomputation allows for flexible, spatial simulation, but can be computationally intensive.
Spatial Regression	Standard statistical regression models are not appropriate for analyzing spatially dependent data. Instead, several spatially regression methods have been developed. • **Spatial autoregressive models.** Simultaneous autoregressive (SAR) models are frequentist approaches designed to address spatial autocorrelation. They incorporate spatial autocorrelation using neighborhood matrices that specify relationships between neighboring data points. • **Bayesian regression models.** Bayesian regression models provide an alternative to SAR models. can be used to estimate the effects of potential risk factors related to a disease by including fixed covariates along with the random effects. • **Geographically Weighted Regression** (GWR), models spatially varying relationships using a local linear regression model.
Decision Support Systems	

	• Can exploit multiple technologies (geographical information systems, statistical and mathematical models, decision-support modules), multiple data sources and permit widespread dissemination of epidemiological data. • Spatial simulation; geocomputation

However, with big data analytics (34, 35), GeoAI (36) and increased access to geographic data, much more can be done with existing surveillance data. For food-, water- and air-borne infections, residential addresses and zip codes of people reporting symptoms and pathogens; stratified by age and sex, can be mapped in space and time to examine incidences of infections within precise geographic areas (e.g. tuberculosis in South Africa (37, 38) and for targeted responses (e.g. vaccine deployment for cholera in KolKata (39))). Geographic cluster detection even for infections with person to person spread, such as sexually transmitted and blood-borne are meaningful, as studies have demonstrated surprisingly dense clustering of street involved people who sell sex (e.g. (40)).

To accomplish these different tasks, public health epidemiologists require sufficient training in concepts of geography and a variety of methodologies and techniques (e.g. (41-45)) including spatial analytical (28) and web-mapping methods, which are still largely absent from many educational curricula, with only brief mentions of these methods and tools (41-47). Although there has been an increase in the inclusion of data science in the health sciences (e.g. (48)), spatial analysis and the ability to examine disease incidences within geographic contexts is still largely missing (49) as highlighted in the recent article (48) on data science for public health that does not include any reference to spatial data science. This is hindering the ability to incorporate crucial, process-based understandings of health events within the context of different geographies which may influence disease outcomes (49). Geographies may include population (e.g. density, lifestyle, demographic characteristics); physical environment (e.g. land use, climate [temperature, wind, precipitation], topography, water bodies, soil type); mobility (e.g., transportation nodes, infrastructure); health facilities (e.g. location, type, availability, and accessibility) or human and social geographies such as boundaries, places of interest, social venues, cultural locations, and activity spaces. Integrating these with disease analyses will enhance public health planning and intervention (28, 49).

2 Methodology

As technologies continue to evolve and different geographic data becomes available, how can we better incorporate these into a process that can help public health practitioners evaluate disease and health risks both in the short and long term? Essentially, how do we train epidemiologists in geography and geospatial technologies and methods? To address this, we have centred our evaluation around a public health response cycle that encompasses several steps important for investigating, evaluating, and managing disease incidence and outbreaks, as described in (42, 50-52) and summarized in Table 2a-c from a number of different reviews. We further demonstrate how different spatial and mapping methods and analyses may be used by providing several case studies that range from local outbreaks to a global pandemic. These include the John Snow Cholera outbreak of 1854, the Ebola outbreak of 2014 in West Africa and the ongoing global COVID-19 pandemic that started in 2020.

2.1 Ecology of Disease - Detecting an outbreak or health event through surveillance:

The initial stage of the cycle consists of detection where ideally, an outbreak or health event is discovered through consistent monitoring, and an unexpectedly high number of people in a small geographic location (e.g. one city or hospital) are diagnosed with it. Surveillance is defined as the collection, compilation and analysis of health conditions which includes dissemination of information to those who need to know, including health care staff and policy makers (53). Mandated by law for many infectious diseases, demographic, locating, laboratory and clinical data on people who have the condition (known as cases) are collected by health care and laboratory professionals who notify local, national and international (e.g. WHO (54)) public health agencies (55). Criteria for what constitutes a case of the disease under surveillance are published by state, provincial or federal, or international authorities and usually include a positive laboratory test for the pathogen and signs and symptoms consistent with infection. As soon as the number of cases rises above the epidemic threshold, based on past mean rates and standard deviations, a potential outbreak exists, which is verified after a preliminary check for issues such as possible laboratory or data entry errors. Many surveillance systems, particularly for infectious diseases, contain minimum data to describe the affected people by person, place, and time. Age and sex of infected cases is tabulated and graphed, together with their residential addresses; dates of; onset, presentation at a clinic, specimen collected, and results reported to the public health department (e.g. DONs (1)).

2.2 Developing an understanding of the ecology of a disease.

These data, coupled with laboratory results on the pathogen identified ,are usually sufficient to form sound hypotheses as to source and exposure (50). Through the inclusion of geography, they allow for geographic visualizations and spatial analyses to be performed in GIS (Geographic Information Systems) and other such software packages. Through these methods and other case data, public health staff are able to identify clusters that highlight outliers or hotspots, examine interactions and relationships through the integration of different types of data (environment, host, pathogen) as well as compare cases with the rest of the population stratified by different attributes such as geography, time, symptoms, age, or sex.

Table 2a: Ecology of Disease: A breakdown of the different steps important for investigating, evaluating, and managing disease incidence/outbreaks and the spatial analysis methods that are useful at each stage (adapted from (42, 51))

1. Surveillance and monitoring: data collection	**Collect data from authoritative and non-authoritative sources, geocode/geo-reference cases, structure and manage data.**
2. Establish the existence of a disease/ outbreak and describe cases and how cases may be related	• **Where cases are located?** Visualise case distribution (confirmed, suspected, dead) and spatial limits of disease/outbreak (e.g. dot map; intensity maps (Kernel density Estimates (KDE)); thematic maps; Thiessen polygons) • **Are cases clustered?** Identify and confirm clustering (e.g. Kernel density estimates (KDE), Ripley K, Nearest Neighbour analysis); Moran's I, Getis-Ord G) and where significant clusters/outliers/hotspots are located (Local Indicators of Spatial Association (LISA)) • **How are cases related?** Context mapping analysis: integration of geographic data to assess where the cases are in relation to different points of interest (POIs) (e.g. topological analyses, overlays, surface analysis; descriptive statistical analysis), distance between cases and POIs, distance (e.g. buffer, cost-distance analysis), connectivity between places (e.g. network analysis)

3. Examine disease patterns and interactions develop hypotheses	• **Where are the transmission zones and pathways?**
	  ◦ Visualise distribution of cases in relation to known risk factors or potential sources (e.g. rate map (change maps (increase, decrease, unchanged)); thematic maps/choropleth maps) and symptoms or other characteristics (gender, age, socioeconomic status, profession, social behaviour etc.)
	  ◦ Identify center of outbreak (e.g. spatial mean, median center)
	  ◦ Identify and locate significant clusters (e.g. LISA; Getis Ord Gi* statistic, spatial scan statistic; hierarchical clustering; machine learning (Random Forest))
	  ◦ Identify high-risk areas (e.g. attack rates in zones at different distances from potential sources (cost distance analyses; KDE, LISA, Getis Ord Gi statistic, geostatistical analysis)
	  ◦ Use maps to assist with active case finding and locate areas of similarity or defined distances or defined accessibility pathways
	• **Why? How are cases related to transmission zones and pathways?**
	  ◦ Identify significant trends in attack rates with distance from potential sources (e.g. linear regression of log-transformed attack rates) and incorporate different factors (environmental
	  ◦ Describe progression of outbreak through directional spread using standard deviation ellipse; space-time maps; animations at different time intervals and using different visualizations); (SATScan); LISA analysis at different time intervals; rates of change and diffusion; network pathway
	  ◦ Connectivity and interactions (e.g. map phylogenetic data, social network graphs)
	  ◦ Assess context: examine hotspot and outlier areas with additional geographic data to assess where cases are in relation to different points of interest; population characteristics, network pathways, etc.
	• **Develop models to capture disease dynamics and interactions?**
	  ◦ Model concentrations of infections to understand transmission dynamics (e.g. (geo)computational and simulation modelling; compartmental models; geostatistical models; agent-based models)

2.3 Response - prevention planning and implementation of interventions to minimize risk, enable for recovery and treatment:

Once we understand the ecology of the disease, the next stage of an outbreak or health event is to develop a response that includes implementing prevention measures that range from educating the public and health officials, to infrastructure needs such as providing sanitation, developing new vaccinations or the placement of new health facilities. In the last stage of the response cycle, surveillance for all pathogens of public health importance continues after prevention measures have been taken, to ensure that no new cases arise and to detect new outbreaks (Table 2b).

Table 2b: Response: A breakdown of the different steps important for investigating, evaluating, and managing disease incidence/outbreaks and the spatial analysis methods that are useful at each stage (adapted from (42, 51))

4. Response: prevention measures	• Forecasting and prediction of outbreak: Identify geographic areas at risk of future outbreaks (e.g. risk mapping)
	• Short and long term planning and implementation:
	  • Spatial targeting of interventions (e.g. containment/isolation; barriers; vaccination campaign; health facilities and treatment centers; mobile hospitals; installation of clean (running) water or sanitation systems; placement of ultraviolet lights (e.g. protect from TB in overcrowded shelters); placement of needle exchanges clean needles, drug equipment)
	  • Policy development and implementation

2.4. Communication – informing the public

During each of these stages, communication strategies are important to ensure up-to-date information is provided (Table 2b). This can take many different forms ranging from published documents (1, 2) to interactive web maps (56, 57) that are updated in real-time (e.g. COVID-19 Dashboard provided by WHO (58); Johns Hopkins (59)) or at other time intervals (e.g. weekly (60) or adhoc (e.g. CDC Travel Recommendation Map (61)) depending on needs.

Table 2c: Communication: A breakdown of the different steps important for investigating, evaluating, and managing disease incidence/outbreaks and the spatial analysis methods that are useful at each stage

5. Communication	Use maps (static and dynamic interactive web maps) and other visualization dashboards to communicate areas of risks; provide updates of disease outbreak/event to the public; provide results to health officials/policymakers.

3 Case Studies

To demonstrate how different spatial and mapping analyses may be incorporated at each of the different steps of this framework, we provide several examples ranging from local outbreaks to a global pandemic. These include the John Snow cholera outbreak of 1854, the Ebola outbreak of 2014 in West Africa and the global COVID-19 Coronavirus pandemic of 2019-ongoing.

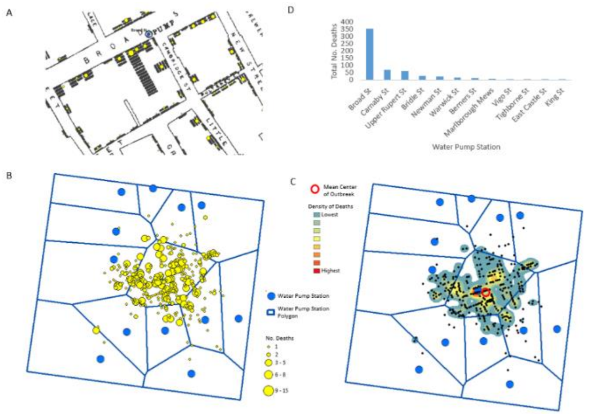
3.1 Software and Data Availability Sub-Section

All data used during each of these analyses are available in the public domain and are listed in Table 3. All analyses were completed in ArcGIS and Excel.

Table 3: Data Sources used for the case studies

Case Study Example	**Data Source**
Cholera 1854	Digitized from John Snow's map.
Ebola 2014-2016 Outbreak	(62-64)
COVID-19 in NL	Data are available from (65) at 2 week intervals
Global COVID-19 Data	WHO (2); ECDC; JHU (66); PA Health Data (67)
Country Boundary Data	ESRI country boundaries

Figure 4: Cholera Outbreak of 1854: Where were the cholera deaths located? How did these deaths relate to the environment and each other? **View the location of deaths:** Visualize the distribution of deaths (A) example of the Cholera Map of 1854 with digitized points, (B) in relation to the water pumps (B) and assess where the mean centre of the outbreak (C) and where the highest density of deaths occurred (C,D). **Summarize deaths by water pump:** Thiessen polygons were used to create boundaries for each pump, where all areas inside the boundary are closest to a single pump. This was used to find the total number of deaths closest to a particular pump (B) and summarized in (D). Kernel Density Estimates (KDE) was used to aggregate points to create a continuous surface to show where the highest number of deaths occurred and possible zone of containment. Analysis performed in ESRI ArcGIS 10.8.



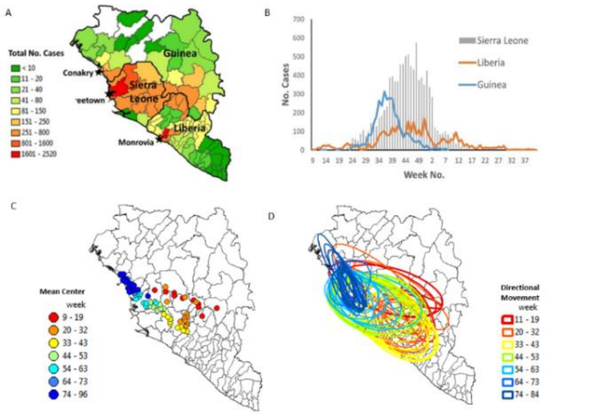
Description of outbreak: 500 deaths were detected 250 yards from the Cambridge & Broad Street intersection in 10 days.

Ecology of the Disease - Determine sources of infection:

1. **Visualize and examine outbreak cases:** Map the location of all infected cases to determine the relationships between them and the environment in which they are interacting. Examine how close the cases are to each other. Determined if cases were clustered together and identify common activity spaces and potential sources of infection.
2. **Collect more data:** Conduct in-depth interviews of ill and well people to obtain further information on all possible hypothetical exposure locations to the pathogen.
3. **Identify source of infection: single vs multiple sources of infection:** From the interviews/questionnaires and maps, identify additional potential sources of infection.
   a. **Hypothesis:** That contaminated water from Soho caused deaths from cholera (5).
   b. **Hypothesis:** That contaminated water from the Broad street pump caused cholera in Golden Square (5).

Findings: The majority of the deaths occurred in the area closest to the Broad Street pump.
Response: Request the parish officers to stop the water supply of Broad Street Pump by removing the pump handle
Continued surveillance: To ensure no new cases, and detect new ones, Snow went back 2 – 3 weeks later (5).

Figure 5: Ebola Outbreak of 2014-2016: Location of Ebola cases were obtained from the WHO. Weekly Ebola cases were reported at district levels for each of the countries, Guinea, Sierra Leone, and Liberia. (A) **Where Ebola cases were reported**: Choropleth maps were used to capture the total number of cases within the outbreak area to show areas with the highest number of cases. (B) Epi curve showing the number of cases for each of the countries. (C) **Who was affected when? Weekly distribution of cases:** The mean center for each week was mapped to show where the mean center of the outbreak was recorded over time and (D) the directional movement of the outbreak was determined using the standard deviational ellipse. Spatial analyses were performed in ESRI ArcGIS 10.8.



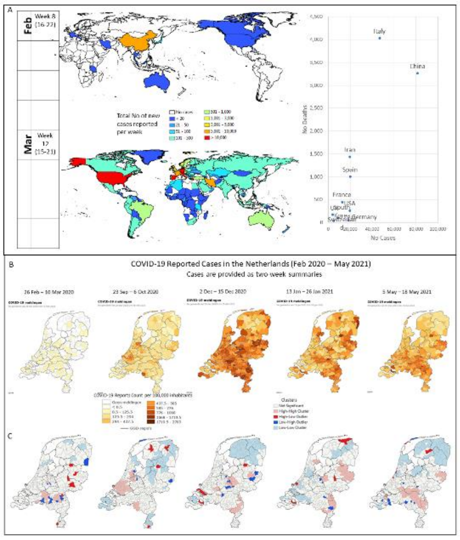
Description of outbreak: Originally an 18-month-old child playing beneath a bat infested tree in Meliandou, Guinea, a small settlement of 31 people. Several relatives, midwives and traditional healers in Meliandou developed fever, vomiting and black stools, diarrhoea, and dehydration. It was thought to be cholera, until it spread to 4 other places, and WHO was alerted on 13 March 2014. The investigation started and Ebola was identified 21 March 2014 (4).

Ecology of the Disease - Determine sources of infection: Originally found in bats, Ebola may contaminate fruit and places where children play, then transmits person to person by direct contact through broken skin; mucous membrane body fluids, contact with contaminated items, clothes, bedding, and medical equipment, infected bats, non-human primates, and sex with an infected person. Ebola is new in West Africa where populations are more urban (6).

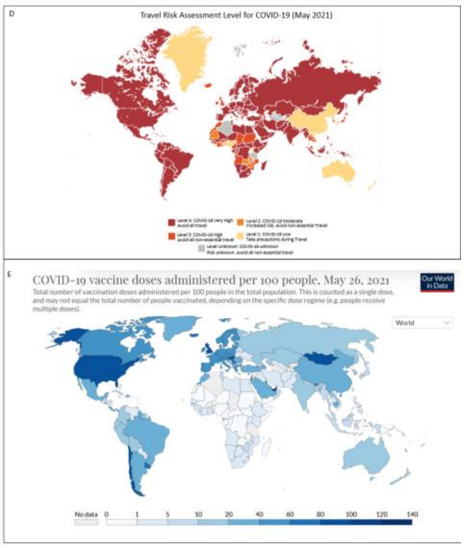
1. **Visualize and examine outbreak cases:** Map the location of Ebola cases over time to assess the spatial distribution of cases and spread of disease.
2. **Collect more data:** Collect detailed information of cases, where and when they occurred and of their contacts through contact tracing. On Jan 24, the head of the health post in Meliandou informed public health about 5 people with diarrhoea who died; the disease appeared similar to cholera, so nothing was done. Then MSF investigated again on Jan 27th, and also indicated cholera. The Guinea Ministry of health issued an alert March 13; WHO Africa investigated 14 – 25 March and found cases in three different places linked to the largest city with health care closest to Meliandou (6)
3. **Identify source of infection: single vs multiple sources of infection.** There does not seem to have been any.

Response: Investigations into contacts and cases, safe burial for those who died (mandatory cremation); quarantine those affected to a crowded slum of 75,000 people; closure of markets; restriction of movement of patients and contacts, and curfews (12)
Findings: weak health systems, undetected cases migrated to Sierra Leone and Liberia; crowding of cases (12)
Continued surveillance: To ensure no new cases.

Figure 6: Communicating risk and response: (A) Maps showing areas of risk at week 8 and week 12 during the pandemic and how the centre of risk changed from China to Europe (see graph). (B) an interactive weekly local risk of COVID-19 in the Netherlands of two-week summaries of reported cases (Source: (65)) and (C) shows the changing areas of risk using the cluster and outlier analysis (Anselin Local Moran's I) with spatial relationship defined as contiguity (edges and corners). (C) Shows the same information in B but highlights clustering (e.g. high-high: high incidence rates surrounded by high incidence rates; low-low: low incidence rates surrounded by low incidence rates; high-low and low-high: dissimilar areas or outliers where there are areas of high incidence rates surrounded by areas of low incidence rates and vice-versa). Analysis for (C) were performed in ESRI ArcGIS 10.8.



(D) global COVID-19 travel risk map (Source: (61)); (E) global vaccination updates (68) (Source: (69)).



Description of outbreak: Unusual pneumonia was detected in 27 people in Wuhan, China, most of whom were vendors at a seafood and wildlife market as of Jan 2 (3).

Ecology of the Disease - Determine sources of infection:

1. **Visualize and examine outbreak cases:** Map the location of all infected cases to determine what relationships exist with each other and the environment in which they are interacting. Examine how close the cases are to each other. Determine if more of these cases are clustered together than expected by chance, given random placement, allowing for sex and age. Identify overlapping activity spaces and common "hang out" locations. Add context by mapping where the infected are in relation to other places in the area frequented by those that are ill. Identify common features within the area of interest (e.g. food sources, markets).
2. **Collect more data:** 121 contacts being observed by physicians, Jan 3 (7). Conduct in depth interviews with those that are ill and those that are well to obtain further information on all possible hypothetical exposure locations to the pathogen. Obtain detailed data on symptoms, clinic visits and hospitalisations; places visited just before each person became ill (e.g. restaurants, parties, day trips, markets) along with interactions with animals and where these took place.
3. **Identify source of infection: single vs multiple sources of infection:** From the in-depth interviews/questionnaires and maps, identify additional potential sources where respiratory disease may have been acquired. Source identified as a coronavirus (10).
    a. **Hypothesis:** Transmission by person to person is most likely given the number of cases in Japan South Korea and the number of confirmed health care workers that are infected.
    b. **Hypothesis:** Mode of transmission is by droplet, and/or contaminated surfaces.

Response: Close the market in Wuhan. Implement social-distancing measures; temperature checks on travellers into Hong Kong (13); create technological apps to monitor the situation; develop and roll-out a vaccine to reduce infections.
Findings: Ongoing. From the time the market closed to the isolation of infectious people and the implementation of social distancing, it reportedly took 5 weeks for no new locally transmitted cases to emerge. Since then, monitoring has continued with various closures and lockdowns to manage cases locally and at a country level.
Continued surveillance: To ensure no new cases. Surveillance is ongoing as variants emerge. Surveillance is ongoing of vaccinations rollouts and coverage.

4 Discussion

By their very nature, the geospatial sciences are interdisciplinary, central to everything we do, and to everything with which we interact. Maps and geospatial technologies have been useful for showing where disease outbreaks may be taking place; identifying potential sources of infection and determining who may be affected when and where. However, the steep learning curve associated with using many GIS packages has resulted in its slow uptake in many fields (70). As we enter the digital (data) revolution and the age of web mapping (70); it will become critical to develop ways that integrate these methods and data so as to enhance communication efforts (71), sharing of sensitive data (see (72, 73)) and analytical capabilities. Examples of these include better integration of geographic analysis with other types of data such as phylogenetic data (74) (75); clustering methods (76) and forecasting in real-time (77) at all stages of public health surveillance, planning and response. This has been highlighted by the many analyses, maps and interactive dashboards that have been created during the COVID-19 pandemic (78, 79); including identifying hotspots (80), modelling risk (81) (82) and spread (83) as well as integrating environmental data to examine factors influencing COVID-19 (84) and the need for demographic characteristics (85) to better assess who may be at risk when.

As we move forward, we need to develop new methods and integrate Geography, GIScience and Spatial Data Science into the core curriculum of public health to provide a unified approach across space and time so that we can improve how we monitor and manage health and well-being and are better prepared for the next outbreak.

5 References

1. WHO. Disease outbreaks by year1996-2020 Feb 8 2020. Available from: https://www.who.int/csr/don/archive/year/en/.
2. WHO. Novel Coronavirus (2019-nCoV) situation reports Report 1 (Jan 21, 2020) -present (Feb 20, 2020)2019. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.
3. ProMED. ProMED: ProMED International Society for Infectious Diseases. 2020. [June 14 2020]. Available from: https://promedmail.org/promed-post/?id=20200102.6866757.
4. WHO. Origins of the 2014 Ebola epidemic2015 June 14 2020. Available from: https://www.who.int/csr/disease/ebola/one-year-report/virus-origin/en/.
5. Snow J. On the Mode of Communication of Cholera. 2nd ed. London, UK1855.
6. WHO. Factors that contributed to undetected spread of the Ebola virus and impeded rapid containment2015 June 14 2020. Available from: https://www.who.int/csr/disease/ebola/one-year-report/factors/en/.
7. ProMED. ProMED: ProMED International Society for Infectious Diseases. 2020. [June 14 2020]. Available from: https://promedmail.org/promed-post/?id=6869668.
8. Fauci AS, Morens DM. The perpetual challenge of infectious diseases. N Engl J Med. 2012(366):454–61.
9. Horby PW, Hoa NT, Pfeiffer DU, Wertheim HFL. Drivers of Emerging Zoonotic InfectiousDiseases. Yamada A, Kahn LH, Kaplan B, Monath TP, Woodall J, Conti L, editors. Toyko, Japan: Springer; 2014. 253 p.
10. ProMED. ProMED: ProMED International Society for Infectious Diseases. 2020. [June 14 2020]. Available from: https://promedmail.org/promed-post/?id=6876648.
11. Canada Go. App: COVID Alert today. Canada: Government of Canada; 2020.
12. WHO. One year into the Ebola epidemic: a deadly, tenacious and unforgiving virus2015 June 14 2020. Available from: https://www.who.int/csr/disease/ebola/one-year-report/introduction/en/.
13. ProMED. ProMED: ProMED International Society for Infectious Diseases. 2020. [June 14 2020]. Available from: https://promedmail.org/promed-post/?id=6874277.
14. Paglinawan D. As Canadian restaurants reopen, diners asked for personal info to help with contact tracing2020 Aug 12, 2020. Available from: https://globalnews.ca/news/7246079/coronavirus-restaurants-contact-tracing/.
15. Goodchild M. The quality of big (geo) data. Dialogues in Human Geography. 2013;3(3):280-4.
16. Goodchild M. GIS in the Era of Big Data. Cybergeo : European Journal of Geography [Internet]. 2017 May 22, 2017; (Jun 2, 2017). Available from: http://cybergeo.revues.org/27647.
17. Decaro N, Martella V, Saif LJ, Buonavoglia C. COVID-19 from veterinary medicine and one health perspectives: What animal coronaviruses have taught us. Research in Veterinary Science. 2020(131):21-3.
18. Kraemer MUG, Hay SI, Pigott DM, Smith DL, Wint GRW, Golding N. Progress and Challenges in Infectious Disease Cartography. Trends in parasitology. 2016;32(1):19-29.
19. Hay SI, Battle KE, Pigott DM, Smith DL, Moyes CL, Bhatt S, et al. Global mapping of infectious disease. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2013;368(1614):20120250.
20. Hay SI, Tatem AJ, Graham AJ, Goetz SJ, Rogers DJ. Global environmental data for mapping infectious disease distribution. Adv Parasitol. 2006;62:37-77.
21. Craig MH, Snow RW, Le Sueur D. A climate-based distribution model of malaria transmission in Africa. Parasitology Today. 1999;15(10).
22. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, et al. A World Malaria Map: Plasmodium falciparum Endemicity in 2007. PLoS medicine. 2009;6(3):e1000048.
23. Blanford JI, Blanford S, Paaijmans K, Schreiber K, Crane R, Mann M, et al. Spatial and temporal distribution of malaria parasite development rate using different measures of environmental temperature. Scientific Reports. 2013;3.
24. Elliott P, Wartenberg D. Spatial epidemiology: current approaches and future challenges. Environmental health perspectives. 2004;112(9):998-1006.
25. Ostfeld RS, Glass GE, Keesing F. Spatial epidemiology: an emerging (or re-emerging) discipline. Trends Ecol Evol. 2005;20(6):328-36.
26. Dummer TJ. Health geography: supporting public health policy and planning. CMAJ. 2008;178(9):1177-80.
27. Clements ACA, Pfeiffer DU. Emerging viral zoonoses: Frameworks for spatial and spatiotemporal risk assessment and resource planning. Vet J. 2009;182(1):21-30.
28. Auchincloss AH, Gebreab SY, Mair C, Diez Roux AV. A review of spatial methods in epidemiology, 2000-2010. Annual review of public health. 2012;33:107-22.
29. Cromley EK, McLafferty SL. GIS and Public Health. 2nd ed. New York: Guilford Press; 2012.
30. Kirby RS, Delmelle E, Eberth JM. Advances in spatial epidemiology and geographic information systems. Annals of epidemiology. 2017;27(1):1-9.
31. Rushton G. Public health, GIS, and spatial analytic tools. Annual review of public health. 2003;24:43-56.
32. Mandl KD, Overhage M, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing syndromic surveillance: A practical guide informed by the early experience. Journal of the American Medical Informatics Association. 2004;11(2):141-50.
33. Franch-Pardo I, Napoletano BM, Rosete-Verges F, Billa L. Spatial analysis and GIS in the study of COVID-19. A review. Science of the Total Environment. 2020(739):1-10.
34. Dolley S. Big Data's Role in Precision Public Health. Frontiers in Public Health. 2018;6(68).
35. Hay SI, George DB, Moyes CL, Brownstein JS. Big data opportunities for global infectious disease surveillance. PLoS medicine. 2013;10(4):e1001413.
36. Kamel Boulos MN, Peng G, VoPham T. An overview of GeoAI applications in health and healthcare. International journal of health geographics. 2019;18(1):7.
37. Bainomugisa A, Pandey S, Donnan E, Simpson G, Foster J, Lavu E, et al. Cross-border movement of highly drug-resistant Mycobacterium tuberculosis from papua new guinea to Australia through torres strait protected zone, 2010-2015. Emerging infectious diseases. 2019;25(3):406–15.

38. Peterson ML, Gandhi NR, Clennon J, Nelson KN, Morris N, Ismail N, et al. Extensively drug-resistant tuberculosis 'hotspots' and sociodemographic associations in Durban, South Africa. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2019;23(6):720-7.
39. You YA, Ali M, Kanungo S, Sah B, Manna B, Puri M, et al. Risk map of cholera infection for vaccine deployment: the eastern Kolkata case. PloS one. 2013;8(8):e71173.
40. Logan JJ, Jolly AM, Blanford JI. The Sociospatial Network: Risk and the Role of Place in the Transmission of Infectious Diseases. PLoS One. 2016;11(2):e0146915.
41. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Philadelphia: Wolters Kluwer Lippincot Williams & Wilkins 2012.
42. CDC. Principles of Epidemiology in Public Health Practice. Atlanta, USA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2012. Available from: http://www.cdc.gov/ophss/csels/dsepd/SS1978/SS1978.pdf.
43. Celentano DE, Szklo M. Gordis Epidemiology Philadephia: Elsevier; 2018.
44. Aschengrau A, Seage GR. Essentials of epidemiology in public health: Jones & Bartlett Learning; 2020.
45. Friis RH, Sellers TA. Epidemiology for public health practice. Burlington, MA: Jones & Bartlett Learning; 2014.
46. Young T, Naude C, Brodovcky T, Esterhuizen T. Building capacity in Clinical Epidemiology in Africa: experiences from Masters programmes. BMC Med Educ. 2017;17(1):46.
47. Bonita R, Beaglehole R, Kjellstrom T. Basic epidemiology. Geneva, Switzerland: World Health Organization (WHO); 2006. Available from: http://apps.who.int/iris/bitstream/10665/43541/1/9241547073_eng.pdf.
48. Goldsmith J, Sun Y, Fried LP, Wing J, Miller GW, Berhane K. The Emergence and Future of Public Health Data Science. Public Health Reviews. 2021;42:1-6.
49. Mclafferty S. Disease cluster detection methods: recent developments and public health implications. Annals of GIS. 2015;21(2):127-33.
50. Morgan O. How decision makers can use quantitative approaches to guide outbreak responses. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2019(374).
51. Smith CM, Le Comber SC, Fry H, Bull M, Leach S, Hayward AC. Spatial methods for infectious disease outbreak investigations: systematic literature review. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2015;20(39).
52. Polonsky JA, Baidjoe A, Kamvar ZN, Cori A, Durski K, Edmunds WJ, et al. Outbreak analytics: a developing data science for informing the response to emerging pathogens. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2019;374(1776):20180276.
53. Thacker SB, Berkelmann RL. Public health surveillance in the United States. Epidemiologic Reviews. 1988;10:164-90.
54. Bloom DE, Cadarette D. Infectious Disease Threats in the Twenty-First Century: Strengthening the Global Response. Frontiers in immunology. 2019;10:549.
55. Jolly AM, Logan JJ. Notifiable disease databases for client management and surveillance. In: Noughabi EA, Raahemi B, Albadvi A, Far BH, editors. Handbook of Research on Data Science for Effective Healthcare Practice and Administration. Hershey, PA, USA: IGI Global; 2017. p. 386-416.
56. WHO. Novel coronavirus (COVID-19) situation as of 20 February 2020, 16:00 (CET)2020 Feb 20 2020. Available from: http://who.maps.arcgis.com/apps/opsdashboard/index.html#/c88e37cfc43b4ed3baf977d77e4a0667.
57. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious diseases. 2020.
58. WHO. WHO Coronavirus(COVID-19) Dashboard2021 May 20 2021. Available from: https://covid19.who.int/.
59. JHU. COVID-19 Dashboard2021 May 20 2021. Available from: https://coronavirus.jhu.edu/map.html.
60. RIVM. Actuele informatie over COVID-192021 Feb 27, 2021. Available from: https://www.rivm.nl/coronavirus-covid-19/actueel.
61. CDC. COVID-19 Travel Recommendations by Destination2021 May 21 2021. Available from: https://www.cdc.gov/coronavirus/2019-ncov/travelers/map-and-travel-notices.html.
62. WHO. Ebola data and statistics: Guinea2016 May 10 2021. Available from: http://apps.who.int/gho/data/node.ebola-sitrep.ebola-country-GIN?lang=en
63. WHO. Ebola data and statistics: Liberia2016 May 10 2021. Available from: http://apps.who.int/gho/data/node.ebola-sitrep.ebola-country-LBR?lang=en.
64. WHO. Ebola data and statistics: Sierra Leone2016 May 10 2021. Available from: http://apps.who.int/gho/data/node.ebola-sitrep.ebola-country-SLE?lang=en.
65. RIVM. Weekcijfers COVID-19/Weekly COVID-19 figures 2021 May 21 2021. Available from: https://www.rivm.nl/en/coronavirus-covid-19/weekly-covid-19-figures.
66. CSSE, cartographer COVID-19 Dashboard. Maryland, USA: Center for Systems Science and Engineering (CSSE) at Johns Hopkins University; 2020.
67. PAHealth. COVID-19 Data for Pennsylvania2020. Available from: https://www.health.pa.gov/topics/disease/coronavirus/Pages/Coronavirus.aspx#map.
68. Mathieu E, Ritchie H, Ortiz-Ospina E, al. e. A global database of COVID-19 vaccinations. . Nat Hum Behav. 2021.
69. OWID_Oxford. Coronavirus (COVID-19) Vaccinations. Our World in Data [Internet]. 2021 May 26 2021. Available from: https://ourworldindata.org/covid-vaccinations.
70. Kar B, Sieber R, Haklay M, Ghose R. Public Participation GIS and Participatory GIS in the Era of GeoWeb. The Cartographic Journal. 2016;53(4):296-9.
71. George DB, Taylor W, Shaman J, Rivers C, Paul B, O'Toole T, et al. Technology to advance infectious disease forecasting for outbreak management. Nature communications. 2019;10(1):3932.
72. Gilmore RO, Adolph KE, Millman DS. Curating identifiable data for sharing: The databrary project. Scientific Data Summit (NYSDS), New York; 14-17 Aug, 2016; New York, USA: IEEE; 2016.
73. Jacquez GM, Essex A, Curtis A, Kohler B, Sherman R, Emam KE, et al. Geospatial cryptography: enabling researchers to access private, spatially referenced, human subjects data for cancer control and prevention. Journal of geographical systems. 2017;19(3):197-220.
74. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics. 2018;34(23):4121-3.
75. Lu L, Lycett SJ, Leigh Brown AJ. Determining the phylogenetic and phylogeographic origin of highly pathogenic avian influenza (H7N3) in Mexico. PloS one. 2014;9(9):e107330.
76. Shi X, Wang S. Computational and data sciences for health-GIS. Annals of GIS. 2015;21(2).
77. Lu FS, Hattab MW, Clemente CL, Biggerstaff M, Santillana M. Improved state-level influenza nowcasting in the United States leveraging Internet-based data and network approaches. Nature communications. 2019;10(1):147.
78. Kamel Boulos MN, Geraghty EM. 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. International journal of health geographics. 2020;19(1):8.
79. Franch-Pardo I, Napoletano BM, Rosete-Verges F, Billa L. Spatial analysis and GIS in the study of COVID-19. A review. Sci Total Environ. 2020;739:140033.
80. Desjardins MR, Hohl A, Delmelle EM. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters Applied Geography. 2020;118.
81. Abdulkareem SA, Augustijn EW, Filatova T, Musial K, Mustafa YT. Risk perception and behavioral change during epidemics:

Comparing models of individual and collective learning. PloS one. 2020;15(1):e0226483.
82. Gilbert M, Pullano G, Pinotti F, Valdano E, Poletto C, Boelle PY, et al. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. Lancet. 2020;395(10227):871-7.
83. Gatto M, Bertuzzo E, Maria L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. PNAS. 2020;117(19):10484–1049.
84. Briz-Redon A, Serrano-Aroca A. A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain. Science of the Total Environment. 2020.
85. Pourghasemi HR, Pouyan S, Heidari B, Farajzadeh Z, Shamsi SRF, Babaei S, et al. Spatial modelling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between 19 February to 14 June 2020). Int J Infect Dis. 2020.