Application of Data and Geospatial Analysis in Energy and Health Systems

Ali Faghani
faghani.al@gmail.com

ABSTRACT

APPLICATION OF DATA AND GEOSPATIAL ANALYSIS IN ENERGY AND HEALTH SYSTEMS

Ali Faghani, MS
Department of Mechanical Engineering
Northern Illinois University, 2021
Dr. Mahdi Vaezi, Director

Having sustainable energy and health systems are the main factors in the vision plan of every country. Both of these systems are correlated with a variety of frameworks, including social, physical, technological, political, and economic factors. Therefore, different types of analytics methods can be implemented to develop the required assessments for those who make plans since understanding the effect of such factors individually, also their interactions, and the overall effect is crucial. With this regard, applications of data analysis and geospatial techniques in both energy, and health systems have recently gained attention.

The proposed research here deployed advanced data analytics methods and Geographic Information Systems (GIS) to study energy and health systems, obtain understandings and make predictions, also analyze the corresponding data based on their spatial location and organizing multiple layers of information into visualizations. The proposed research is comprised of two sections. First, Geographic Information Systems and different data analytics methods were used to evaluate the potential of Municipal Solid Waste (MSW) as a renewable energy source in the state of Illinois. Our results demonstrated that Illinois is capable of producing 6,295,385.77 MWH annual energy using incineration technology from MSW. Also, using Anaerobic Digestion (AD)

technology in MSW management would enable the state to be capable of producing more than 1,140,493,710,450.00 Litres biogas per year.

Second, we expanded the application of data and geospatial analysis in the health system and deployed advanced data analytics methods, geographical information system (GIS), and predictive epidemiological models to analyze the anti-contagion policies implemented by the states across the country to slow the spread of COVID-19. Also, by implementing a meta-analysis in conjunction with multi-criteria decision-making methods, a Lung Cancer Risk Index (LCRI) was produced representing the probability of individuals getting lung cancer. The methods that have been developed for the extended applications of data and geospatial analysis in health can be used for various complex decision making and index generating purposes in engineering disciplinary such as additive manufacturing to evaluate the effect of process factors (e.g., injection, concentration, material characteristics, speed, temperature and so forth) individually and collectively to optimize the process and increase the performance.

NORTHERN ILLINOIS UNIVERSITY

DEKALB, ILLINOIS

AUGUST 2021

APPLICATION OF DATA AND GEOSPATIAL ANALYSIS IN ENERGY AND HEALTH SYSTEMS

BY

ALI FAGHANI

A THESIS SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

MASTER OF SCIENCE

DEPARTMENT OF MECHANICAL ENGINEERING

Thesis Director:

Mahdi Vaezi

ACKNOWLEDGEMENTS

I would first like to express sincere gratitude to my research supervisor Dr. Mahdi Vaezi for his invaluable advice, continuous support, and patience during my MS degree. His immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life.

My gratitude extends to Dr. M. Courtney Hughes of the College of Health and Human Sciences at Northern Illinois University (NIU) for her valuable guidance and knowledge; also, for providing the directions and suggestions for conducting chapter 3 of this research.

I would also like to thank Dr. Margaret Wright Geise of the Cancer Center at the University of Illinois at Chicago (UIC) for offering very helpful feedback as well as guides and suggestions on chapter 3.

I also would like to express gratitude to my thesis committee Dr. Tariq Shamim and Dr. Ting Xia for their time, support, encouragement and considerations.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

DEDICATION

I dedicate this thesis to my parents, Hossein and Azam, for their unconditional love, support, and encouragement.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Background

The aim of this thesis is to investigate the applications of data and geospatial analysis in both mechanical engineering (energy management) and health systems. In this regard, using data and geospatial analysis we evaluated the potential of the state of Illinois in achieving recovery energy in both forms of electricity and biofuels out of Municipal Solid Waste (MSW). Also, we extended the application of data and geospatial analysis in health systems and evaluated the effect of non-pharmaceutical interventions in control the spread of infectious disease (COVID-19) across the united states. Besides we combined a meta-analysis and multi-criteria decision making method to generate an index for lung cancer. Our developed methodology in chapter 3 of this thesis can be implemented for complex engineering problems (e.g., additive manufacturing, biofuel process) to investigate the effect of different parameters and generate required indices.

1.1.1 Energy System

The world is undergoing a sustainable development to address concerns such as energy security and global warming. In this regard, global endeavors have been undertaken to minimize anthropogenic greenhouse gas emissions (GHG), which are the major contributor to global climate change. Because of urbanization demand and population growth, energy consumption is predicted to rise by 28% between 2015 and 2040, especially in developing economies and fossil fuel-based power plants continue to remain the world's principal source of energy [1]. Under such circumstances, renewable energy needs to grow with a faster pace in order to prevent global warming consequences by replacing the conventional fuels and meet energy requirements.

Due to the huge demand for energy and importance of energy security, most of developed countries tend to introduce and utilize the alternative energies (e.g., renewable energy) to maximize their energy potential and reduce their vulnerability in providing the long-term energy supply. In this regard, managing the energy sources plays a significant role in energy sustainability and security. Energy management decisions are typically complex operations that depend on a variety of theoretical frameworks, including social, physical, technological, political, and economic factors. Therefore, different types of analytics methods can be implemented to develop the required assessment and accordingly make the appropriate decisions. In this regard, advanced analytical methods (e.g., negative binomial regression, Poisson regression, meta-analysis) and geographical information system (GIS) can be utilized as a powerful tool to study the reported data, obtain understandings and make predictions, analyze the data based on their spatial locations and organize multiple layers of information into visualizations.

Recently, application of geographical information system (GIS) in different aspects of energy management has attracted numerous attentions. A review on literature demonstrated that GIS is widely utilized in the exploration and development of renewable energy resources. It is also a reliable tool in locating the optimal places and determining the ideal corridors for resource transfer and distribution. Solar, wind, geothermal, hydrogen, tidal, wave, hydropower, biomass, nuclear, and fossil fuels are some of the resources that used in the energy industry in correlated with their locations.

Following a GIS based approach Siyal et al. [2] conducted a research and assessed the wind energy considering geographic and environmental restrictions in Sweden. Given the system performance, topographic restrictions, environmental, and land use constraints, they calculated the wind energy potential for the study region. Their findings demonstrated that Sweden has enough wind energy potential and land available for wind energy installations to meet the country's projected renewable energy ambitions. Application of Geographic Information Systems (GIS) in evaluating the solar potential has been investigated by several studies [3, 4]. For instance, Groppi et al. [5] studied the solar energy potential and energy consumption in urban areas. They mentioned that The evaluation's major goal was to strengthen urban buildings' independence from fossil fuels. Their findings indicated a significant association between reduction of non-renewable thermal energy demand and enhancing the buildings efficiency.

Municipal solid waste (MSW) management is becoming a dominant issue in metropolitan areas as a result of the rapid growth in population. The rapid pace of municipal solid waste generation, the complexity of created garbage, and the paucity of land are all concerns that require very immediate attention. Performing waste management analysis needs advanced analytical tools since it involves numerous data and calculations. Therefore, owing to the capability of GIS as a powerful tool, its application in waste management has attracted researcher’s attentions [6-10]; however, more investigation are required to be conducted. Due to the importance of waste management and recent demand for the application of data and geospatial analysis in that field, we evaluated the potential of the state of Illinois in producing energy from MSW which has been elaborated in chapter 2.

1.1.2 Health System

Due to the fast acceleration of technology and information technology nowadays, using new approaches to record health-related data and accordingly perform data analysis has attracted a lot of attention [11]. Thanks to advances in information technology coupled with medical fields, the bioinformatics plays a key role in enhancing the efficiency of health systems [12, 13]. In fact, health as one of the most important aspects of human life has rooted in different environmental, social, individual and biological factors and for those who make the main health decisions, the effect of these factors separately, their interaction and also their whole effect on the health status of the society is crucial [14, 15]. Depends on the availability of data, type of the data and data collection method, the appropriate for data managing and analyzing would be defined.

Selecting the appropriate data analytics method is a key to provide reliable results in public health research studies. Depending on the research topic, the type of data (e.g. continuous, ordinal, binary), and the study design, the most appropriate statistical methodology can be applied and accordingly the corresponding variables will be defined and evaluated [16]. Public health research employs a wide range of statistical techniques, owing to its strong epidemiological and biostatistical basis. The review of the implemented analytics methods on health studies indicates that descriptive statistics (e.g. means, standard deviations and percentages), contingency table analyses (e.g., chi-square tests, mcNemar’s test), Pearson’s correlation, linear and logistic regression and t-tests are frequently used [17, 18]. With the help of data and statistical analysis, the association between factors (individual, social, demographical, biological and geographical) would be evaluated. Also, for those factors that are statistically significant, the correlations would be developed which would obtain a better understanding for the health policymakers and as a result will improve the public health system.

Besides data analysis methods and techniques, geographic information system (GIS) has been highlighted as a powerful tool in evaluating the health system, health status and health decision making [19-21]. A review of literature indicates that geospatial analysis can be used as an effective means to approach a variety of policy, and planning issues in public health which is capable of describing, analyzing, modeling, and visualizing health and place issues in different areas [22-25]. There is evidence that demonstrate how GIS has influenced health policy in depth. Disease mapping and disease modeling are of the most common applications of GIS in public health that play a pivotal role in disease presentation [26, 27]. Predicting the future spread of disease, identifying the factors that may foster or inhibit the disease transmission, evaluating the implemented health policies and highlighting the high-risk areas for disease prevention are other applications of using GIS in public health [28-31]. Furthermore, risk analysis is another application of GIS that allows researchers and practitioners to easily and objectively link many various sources of environmental exposure to people's residence locations throughout time [32]. The purpose of these assessments is to identify critical needs, improve control effectiveness, and prevent outbreaks and epidemics. In many circumstances, GIS is used in conjunction with epidemiological knowledge of illness outbreaks to help avoid additional victims [33].

Although data analytics methods and geographic information system (GIS) can separately be deployed to provide the required results, combining the advanced statistical models and geographic information system (GIS) as a powerful tool is capable of integrating different types of data and analyzing them based on their spatial location which can play a pivotal role in health decision making [23, 34, 35]. This combination organizes different information layers into visualizations presenting by a broad range of maps and is a reliable tool for the prediction of disease patterns and parasite ecology associations [36, 37]. Providing suitable ground to make the appropriate decision in the health-related fields is rely on firstly the effective parameters which cause the health issues, their reliability, the implemented health policies, and also the accurate analysis of these amount of data and their corresponding locations.

In this thesis, by deploying both data analytics method and geographic information system (GIS), two different applications have been discussed. First, using advanced analytics methods (Negative Binomial Regression (NBR) and Poisson Regression (PR)) in conjunction with GIS, a comprehensive study was conducted and the effect of anti-contagion policies on control of COVID-19 across the united states has been investigated. Second, following the systematic review method (PRISMA), the major and minor modifiable risk factor of lung cancer were discerned and implementing a meta-analysis and using Analytical Hierarchy Process (AHP), a novel lung cancer index was generated.

1.2 Objectives of the Research

While the overall objective of the research was to use data and geospatial analysis in energy and health systems by providing details to corresponding policymakers, the specific objectives of the research are:

a) To highlight the potential capacity of MSW as a source of energy in the state of Illinois
b) To develop comprehensive Municipal Solid Waste (MSW) management framework/model for assessing the MSW management options for the state of Illinois.

c) To generate the corresponding maps that will help governments, county administrators, city councils, private organizations, investors, landfill owners etc. make informed decisions about diverting MSW from landfills and converting into energy.

d) To develop a predictive epidemiological model (modified SEIR) and combined it with statistical and geospatial analyses to analyze the anti-contagion policies implemented by the states across the country to slow the spread of COVID-19.

e) To generate the policy ratio index (PRI) representing the performance of implemented policy in response to covid-19 for all the 50 states of the US.

f) To assess the potential associations of policies, individually and collectively, with COVID-19 incidence and mortality, and evaluate the impact of every policy.

g) To present all the major and minor modifiable long cancer risk factors by conducting a comprehensive study.

h) To produce a Lung Cancer Risk Index (LCRI) representing the probability of developing lung cancer for individuals using advanced analysis (meta-analysis) and multi-criteria decision-making methods (e.g., AHP).

i) To present the overall effect of each of modifiable lung cancer risk factors while simultaneously considering all the risk factors.

1.3 Limitations of the Study

For the first part of the thesis (developing a Municipal Solid Waste (MSW) framework/model for then state of Illinois):

- The accuracy in presented data was limited to the most updated geographical information available for the study area.

For the second part of the thesis (evaluating the effectiveness of implemented policies on control the spread of COVID-19 in the US):

- The lack of accurate data about COVID-19 cases, especially asymptomatic cases, in the U.S. [38], posed a challenge to verifying the developed epidemiologic model's case and death values. For the third part (lung cancer modifiable risk index):

1.4 Organization of the Thesis

The thesis consists of four chapters, two of which are based on submitted papers. This thesis is a consolidation of papers, each chapter of which is intended to be read independently. As a result, some concepts and data are repeated. The current chapter provides a background on the application of data and geospatial analysis on both energy and health systems. The importance of both data analytics and GIS (individually and in conjunction with each other) in energy and health studies have been elaborated.

Chapter two is focused on developing a Municipal Solid Waste (MSW) management framework/model for the state of Illinois to enable the policy makers, governments and county administrators make informed decisions about diverting MSW from landfills and converting into energy. Chapter three is an extended application of data and geospatial analysis in health systems and investigates two different applications. First, the association between implemented anti-contagion policies and the spread of COVID-19 across the United States. and second, generating a cancer risk index by implementing multi-criteria decision-making analysis in conjunction with meta-analysis. Finally, chapter four demonstrates the conclusions and provides recommendations for future research.

CHAPTER 2: APPLICATION OF DATA AND GEOSPATIAL ANALYSIS IN MECHANICAL ENGINEERING (ENERGY MANAGEMENT)

2.1 Introduction

The world is going through an energy transition from fossil fuels to renewables in order to address energy security and global warming issues. Biomass feedstock, including agricultural and forest harvesting residues and municipal solid wastes (MSW), are among widely-available renewable resources of energy. Biomass can substitute fossil fuels and prevent burning- and landfilling-related GHG emissions, also water and land pollutions, associated with traditional handling of residues and wastes, particularly MSW, thus promoting sustainable development [36].

The volume and complexity of waste produced has increased as a result of economic expansion, urbanization, and improved living conditions in urban centers. In developing countries, urbanization is increasingly increasing and solid waste management is a major priority for all modern societies. Waste management activities are impacted by rising solid waste generation rates and disposal costs, health and environmental considerations, landfilling capacity, new laws, political situation, and public attitudes [39]. In order to limit the environmental impact of solid waste, strategic waste management, pollution control technology, and waste handling and disposal legislation have been developed with the help of implementing and deploying powerful tools such as geographical information system [40].

We performed a literature review to provide a better understanding of applications of GIS in waste management. Literature demonstrated that GIS has been effectively used for siting the recycling drop-off facilities [41]; Waste collection system [42]; Municipal solid waste collection [43]; Suitability for landfill [44]; Solid waste collection routes optimization [45]; Disposal site selection [46]; Selection of landfills [47]; Selection of dumpsites and transport routes [48] and numerous other applications. All those research studies indicate the importance of waste management and the potential of deploying GIS to evaluate the available and produced waste and provide layer of information. Thus, we started collecting the required data for the state of Illinois to evaluate the potential of waste management in the state.

In the State of Illinois, based on a study by the Illinois Recycling Association [49], 19.3 M tons of MSW (including residential; industrial, commercial, and institutional; and construction and demolition wastes) was generated in 2015 (19% more than the average waste per capita in the United States [50]), with 12.1 M tons of it (~62%) was landfilled in 35 locations. The total GHG emissions produced from the annual landfilled MSW was estimated at 2,516,928 M ton CO_2e. This is equivalent to the annual GHG emissions from approximately 461,000 passenger vehicles, or the carbon sequestered annually by 17,600 acres of forest preserved from deforestation. Moreover, multitude of other long-term problems such as leachate management and land use change have to be addressed with landfilled waste.

On the other hand, coal contributes about 40% in generating electricity in Illinois; ranking the 5^{th} state in the nation in net coal imports by weight [51]. The coal plants are responsible for more than 80% of the asthma-triggering sulfur dioxide, and produced about 60% of the carbon dioxide and smog-forming nitrogen dioxide emitted by the companies across the State of Illinois in 2018 [52]. Diverting MSW from landfills to conversion facilities to produce value-added products such as electricity and biofuels (e.g., co-combustion of coal and MSW in power plants, and producing synthetic gas via gasification or anaerobic digestion of MSW, etc.) could potentially reduce air, land and water pollutions associated with landfills as well as coal-based power plants, and ensure long term sustainable, secure and clean source of energy for the State of Illinois.

2.2 Methods

2.2.1 Description of Study Area

The State of Illinois is located in the Midwest region of the United States and is one of the states in the Great Lakes region of North America. The state's total land is 157,913.4 April 2020. The state is 95.9% land and 4.1 % water, with a population density of 232.0 individuals per square mile, ranking it as the 18th most populous state in the United States. The population of Illinois is mainly concentrated in the state's north east, particularly in the Chicago metropolitan area. Chicago is the state's largest city. The states subdivided into 102 counties as showed in Figure 1.

Figure 1: Map of the study area

2.2.2 Data Collection and Spatial Information

Detail spatial information is essential to successfully manage the municipal solid waste system. This data corresponds to the geographic region of the study territory, as well as spatial data associated to waste collecting protocols. To the best of our knowledge, there are no data on MSW to value added availability for energy production purposes in the state of Illinois. Accordingly, the first stage in this study was data collection on landfill distribution and spatial locations. Next, more information about the converting those waste to energy has been provided. Figure 2 represents both location and capacity of landfills in the state of Illinois. Also, the disposal volume distribution of Illinois has been shown in Figure 3.

2.2.3 Methodology Used

To perform this chapter, we utilized ArcGIS 10.1 software, which employs GIS technology to process geospatial data [17]. We used high-resolution data in either vector (point, line, or polygon) or raster (cells with a resolution of 30 x 30 m) format. GIS is a powerful tool for importing, managing, and analyzing geographically based data. Three steps are involved in implementing the method for this chapter. First, spatial database of study area was provided. Next, deploying the GIS techniques, reallocation of MSW in study area investigated and then, the estimation of the amount of waste to add value and siting the corresponding facilities were followed. The process developed in ArcGIS to identify collection points is iterative and has the key criteria of high waste to added potential. The flowchart of the iterative process is depicted in Figure 4.

Figure 2: Location and capacity of landfills in Illinois

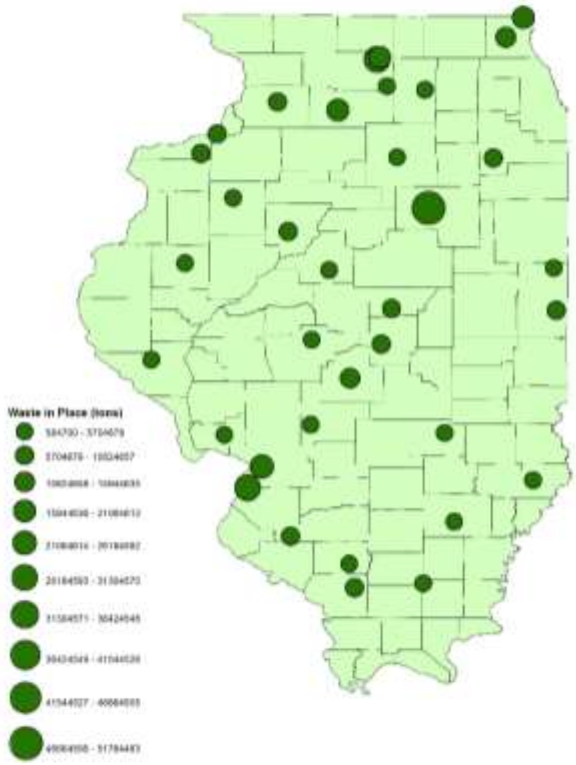


Figure 3: Disposal volume distribution in Illinois

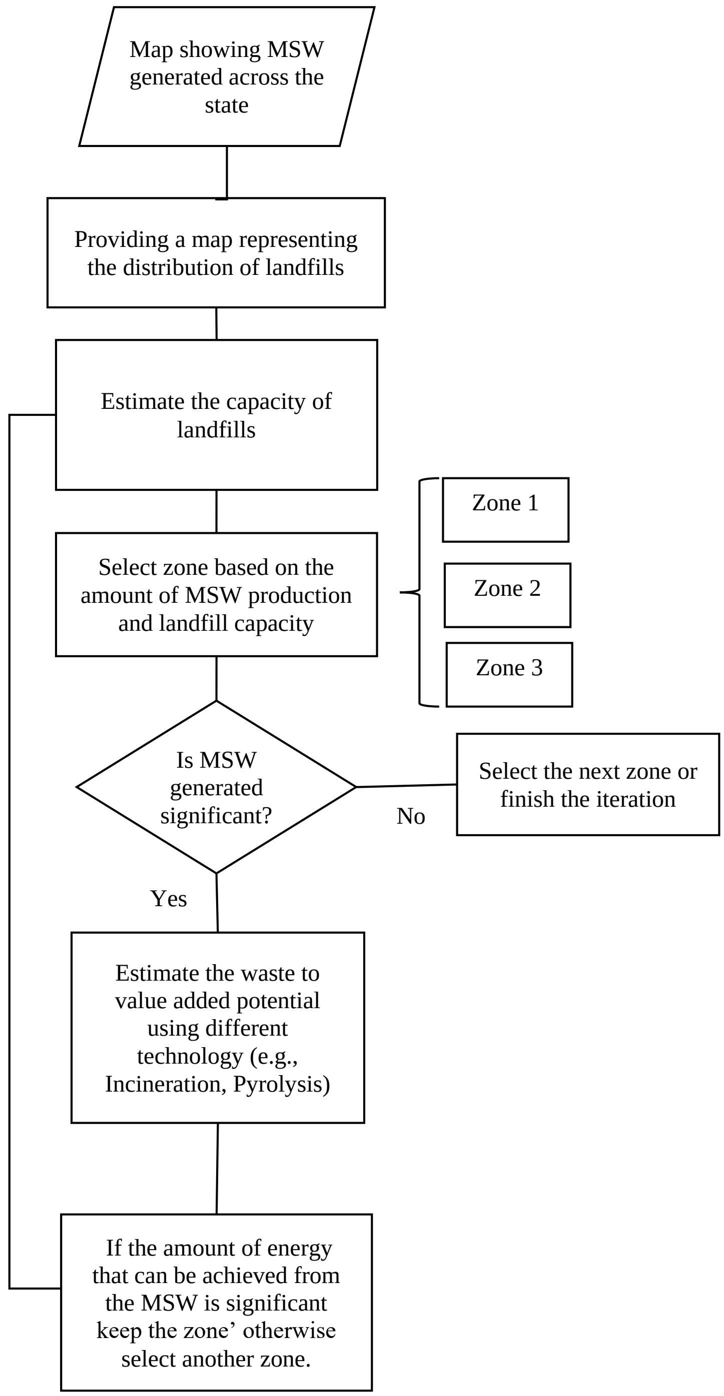


Figure 4: Iteration process for waste management in the state

Initially, the information on landfill availability at the smallest territory sub-division of the study region is combined with the MSW map, to provide the spatial location of the main sources and facilities. A 3 by 3 miles' grid was laid over the map to divide the map area into small quadrants. The area was calculated in each cell and then multiplied by the approximate waste production index for each region to provide a better picture of waste production. To perform the iterative process, the cell values in an area with a 5 miles' radius were summed up in ArcGIS, and the result was assigned to the cell in the center. The cell sum was calculated for each quadrant in the study area. Figure 5 shows a representation of the cell sum operation. The map obtained in GIS (after the cell sum operation) gives information on each cell's potential to waste to added values. From this map, a zone with high capacity was selected in every iteration.

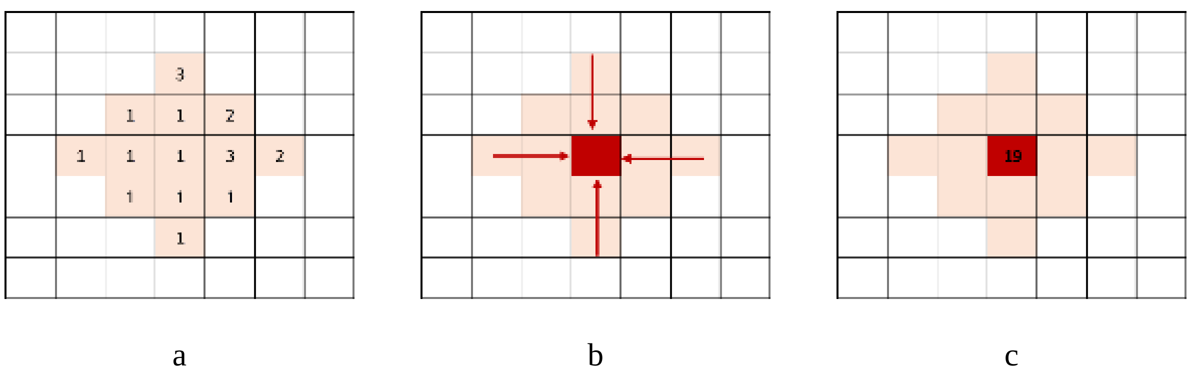


Figure 5: Cell sum method; a) assigned the border values, b) initial values in each cell, c) the middle cell value

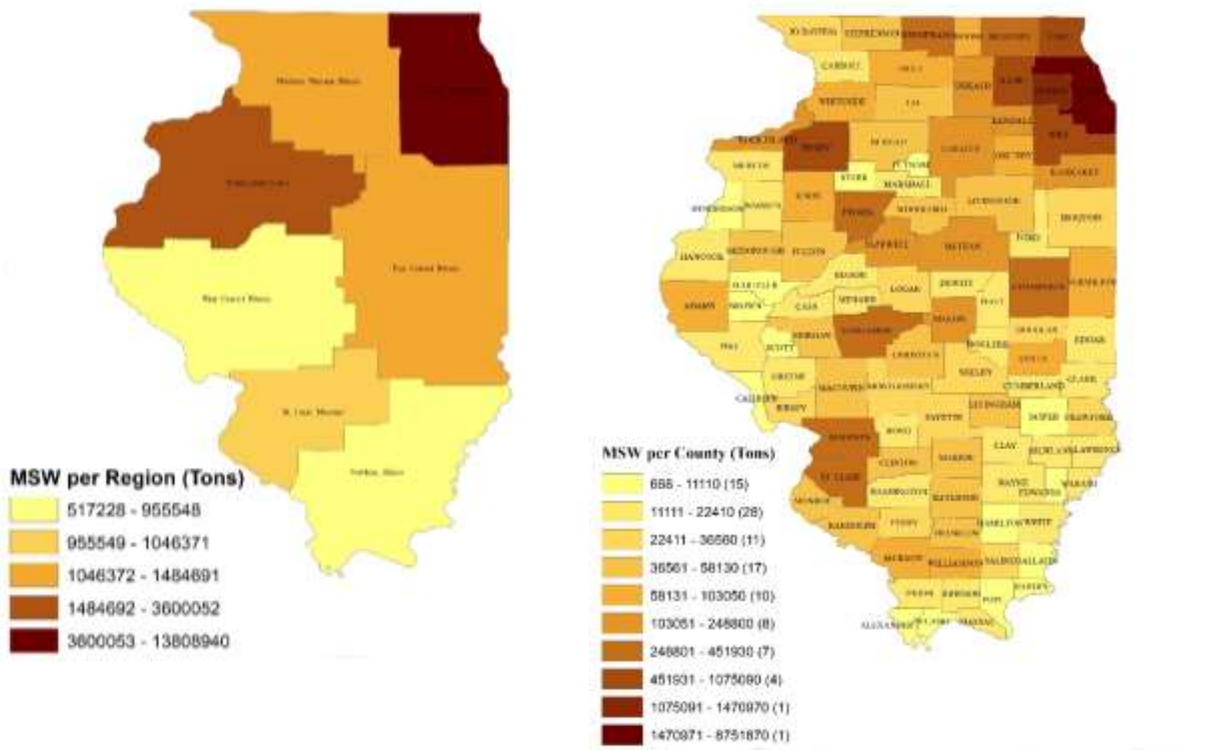
2.3 Results

MSW is made up of a variety of energy-dense items such as paper, plastics, yard waste, and wood-based products. In the United States, around 85 pounds of MSW can be burned as fuel to generate power. Our collected data represents the huge potential of the state of illinois in creating the required facilities to develop waste to added value processes. We evaluated the MSW generation in 7 different regions of the state and the total percentage of MSW generated for each of those region which has been presented in Table 1. Also, Figure 6 (a) demonstrated the portion of MSW generated in those Regions. Our results revealed that the annual amount of MSW generation for the state of Illinois is more than 17,438,741.75 tones. Figure 6 (b) showed the distribution of MSW generation across the state in a county scale.

Table 1: The total percentage of MSW generated

Region	Northwestern Illinois	Chicago Metropolitan	Peoria/Quad Cities	East Central Illinois	West Central Illinois	Metropolitan East St. Louis	Southern Illinois
MSW Generated (% of Total MSW)	5.6	71.5	5.3	6.1	3.8	5	2.7

Having an appropriate evaluation on the distribution of MWS, providing the detail about the composition of the MSW in the state is important. The composition of generated MSW has a direct impact on selecting the appropriate methods, procedure and technology in converting MSW to energy or added value products. Our results for the state of Illinois showed that the weight percentage of composition is for Construction and Demolition (C&D) and papers are more than other MSW types (26.7 and 24.8%, respectively). In this regard, the weight percentage of composition for organics, plastics, textiles, metals, glasses, inorganics, beverage containers and household hazardous waste were estimated to be 20, 10.7, 3.8, 4.1, 2.9, 5.9, 0.2 and 0.9, respectively. Figure 7 Demonstrated the variation in MSW composition for different regions of the state of Illinois. As it expected, the Chicago Metropolitan has the highest MSW generated in the state.



a b

Figure 6: MSW generated a) in 7 regions across the state of Illinois, b) per county

Due to the potential of the state of Illinois in producing energy out of MSW, we conducted a research to provide the most appropriate technology (technical and cost efficient). Waste-to-energy systems or technologies come in a variety of shapes and sizes. The mass-burn method is the most frequent in the United States, in which unprocessed MSW is burned in a big incinerator with a boiler and a generator to generate energy There are other which remove the majority of non-combustible components in order to produce refuse-derived fuel.

Our review showed that Incineration, Pyrolysis and Anaerobic digestion are of those efficiency that can be deployed in order to convert MSW to different forms of energy. Using life cycle assessment, Dong et al. [53] conducted a research and compared waste-to-energy technologies of gasification and incineration. According to their results, environmental impacts of gasification are lower than incineration technology. They mentioned that the quality of the incoming MSW, and process emission level at stack are highly correlated with the performance of incineration process.

Czajczyńska et al. [54] performed a comprehensive study and investigated the potential of pyrolysis processes in waste management. They mentioned that determining the appropriate pyrolysis method is depending on the principles of pyrolysis, most recent developments, various process conditions, and residues. The main aim of their research was to investigate the link between the pyrolysis conditions, the chemical and mineralogical composition and the advantages of pyrolysis in the waste management. Anaerobic digestion is another technology that used to achive recovery energy from MSW. This process is a net energy-producing process which produces energy in the form of biogas and compared to untreated organic waste on land, causes less environmental pollution. Anaerobic digestion attracted a lot of attentions [55-58]. Table 2 Demonstrated the above-mentioned technology, their efficiency, and both input and output.

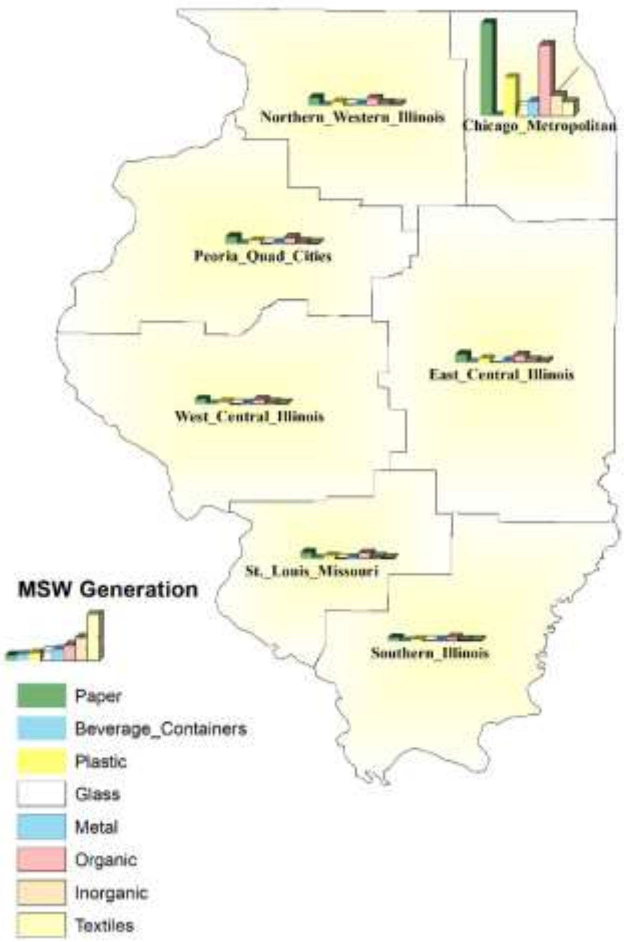


Figure 7: Composition of MSW for different regions of Illinois

Table 2: Technology from MSW to energy generation

Technology	Description	Efficiency (MWh/ton)	Input	Output
Incineration	MSW burnt in a boiler at 1000 – 2000 C	0.5	Mixed MSW	Electricity and heat
Pyrolysis	Decomposition of organic waste in absence of oxygen at 200 – 300 C	0.3	Sorted MSW	Liquid oil, Char, gas
Anaerobic digestion	Biological Process of breakdown of Organic MSW	0.15	Sorted MSW	Electricity, Heat, LNG

Due to the importance of abovementioned technology, following our methodology, we estimated the amount of energy that can be recovered by developing and building the appropriate facilities in the state of Illinois. Figure 8 represents the potential of incineration process for the state of Illinois. Our results demonstrated the total electricity output to be 6,295,385.77 MWH and the total amount of waste generated in 7 different regions that was 17,438,741.75 tonnes the incineration process in the state.

Following the same approach, results for the state of Illinois indicated the total yields of pyrolytic oil to be 190,082.29 MJ. Our analysis demonstrated the total weight of biomass was 9,504,114.25 tonnes. The potential capacity of pyrolytic in generating electricity across the state has been presented in Figure 9.

Moreover, we evaluated the potential of achieving recovery energy from MSW using Anaerobic Digestion (AD) technology. Our results revealed that the annual amount of biomass for the state of Illinois is 9,504,114.25 tonnes and using Anaerobic Digestion facilities, the total biogas yields would be a much as 1,140,493,710,450.00 Litres per year. The potential capacity to produce biogas through Anaerobic digestion technology in 7 regions across the state of Illinois has been shown in Figure 10.

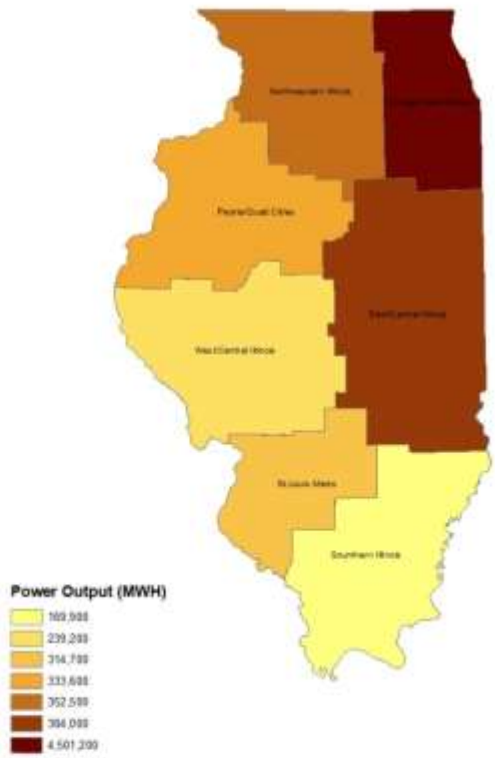


Figure 8: Potential capacity to generate electricity in 7 regions across the state of Illinois

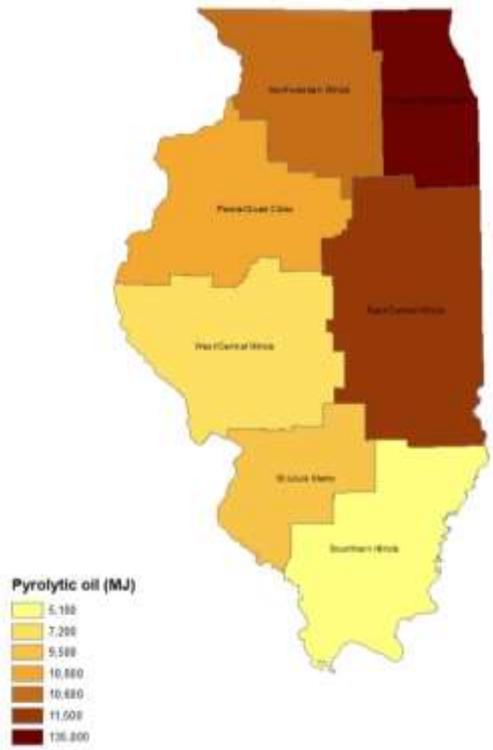


Figure 9: Potential capacity to generate electricity in 7 regions across the state of Illinois via Pyrolysis

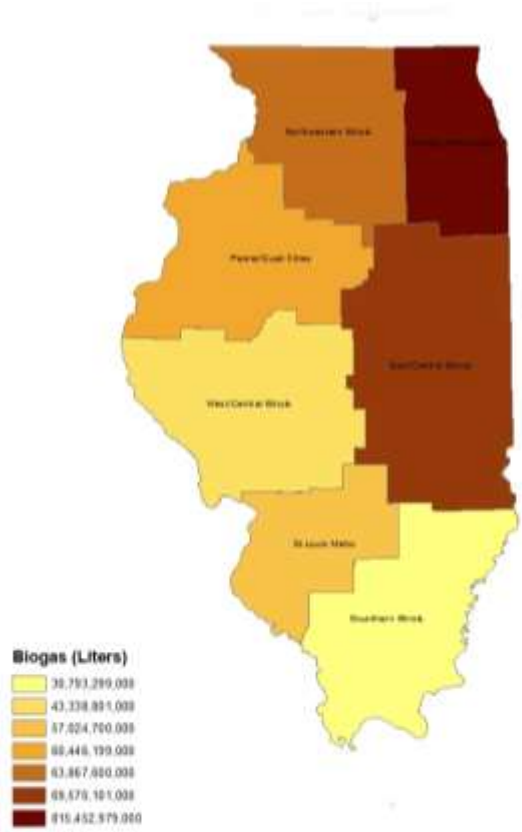


Figure 10: Potential capacity to produce biogas through Anaerobic digestion technology in 7 regions across the state of Illinois

There exists a potential to produce 1,000 Mm^3/yr of CH_4 through anaerobic digestion of organic MSW (which could be subsequently used in gas turbines to generate electricity), and 2,000 GWh electricity via waste-to-energy (WTE) plants using MSW across the state of Illinois. this preliminary study demonstrates the vast potential across the State of Illinois to produce clean energy from MSW. In addition to environmental benefits, this will promote clean energy technologies, create job opportunities, diversify energy resources and help toward a more sustainable economic development of the State of Illinois.

CHAPTER 3: EXTENDED APPLICATION OF DATA AND GEOSPATIAL ANALYSIS IN HEALTH SYSTEM

In this chapter, two different applications of data and geospatial analysis in the health system are discussed.

PART A: ASSOCIATION OF ANTI-CONTAGION POLICIES WITH THE SPREAD OF COVID-19 IN THE UNITES STATES

3.1 Introduction

In December 2019, an unknown infectious disease outbreak in China, identified as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), raised worldwide concern [59, 60]. The novel coronavirus pandemic accounts for 176,694,000 infections and has claimed the lives of 3,830,000 individuals globally as of 16 June 2021. In the United States (U.S.), there were nearly 34,380,000 confirmed cases and over 615,000 deaths of people who tested positive for SARS-CoV-2 as of 16 June 2021 [61]. The first COVID-19 case in the U.S. was reported in Washington State on 22 January 2020 [62, 63]. In the absence of reliable pharmaceutical interventions, in the first several months of the pandemic to combat the virus, governments worldwide implemented anti-contagion policies to curb transmission. States and federal governments implemented public health interventions, including social distancing, travel restrictions, and business closures, to reduce the growth rate of COVID-19 across the nation (see Table 3) [64-70]. Although previous research shows some policies' negative impact on the economy was significant [71-73], the association of all the major anti-contagion policies with COVID-19 spread has not been fully examined at a state and national level.

To study the impact of policy implementation, investigating the evolution of COVID-19 is crucial [74, 75]; thus, well-known compartmental mathematical models (e.g., SIR, SEIR) have often been applied [76-79]. Using those models, previous studies predicted the spread of COVID-19 in Europe [80], China [81], Italy [82], Germany [83], Iran [84], and various communities in South Korea, India, Australia, and the U.S. [85]. However, these research studies deployed simplified models incapable of producing reliable estimates since such models consider the transmission rate to be constant, imposing limitations on predicting other parameters (e.g., daily susceptible and infected cases) [86]. Furthermore, COVID-19 epidemiological studies mostly used reported infected cases as the response value, which comprises primarily diagnosed symptomatic infections but does not take diagnosed asymptomatic and non-diagnosed cases into account [87, 88]. In investigating the effect of policies, previous studies mainly focused on evaluating one [88-92] or a few implemented orders [93-95]. They failed to report all policies' overall associations, or the most important policies in effect, with COVID-19 spread.

To address these shortcomings for the U.S., we first modified the SEIR model by updating the model parameters on a daily basis using the daily death statistics reported for every state. Accordingly, we used our predicted number of daily infected cases (included all the infected categories) for every state produced by the modified SEIR model as the response value and the implemented anti-contagion policies as predictors in a comprehensive statistical analysis.

Table 3: Most frequently implemented anti-contagion policies across the U.S. Table is sorted based on the implementation date.

No.	Policies	Description	First time implemented in
1	Relaxed regulations to become a caregiver	Extends the licensing for doctors and assistants, removes requirements for medical students to join the workforce	29 February WA
2	Suspended provisions requiring in-person notarization of legal documents	Digital notarization is allowed for documents	1 March NY
3	Insurance coverage for all diagnosis testing and partial treatment for COVID-19	Insurance companies cover telehealth meetings and co-pay for COVID-19 testing and treatment	5 March CA-WA-NV
4	Temporarily suspend evictions	Renters and homeowners cannot lose their house or apartment due to COVID-19 complications	6 March IN
5	Restaurant dine-in restrictions	Restaurants can only do delivery and takeout	9 March RI
6	Mandatory quarantine for travelers into the state	Anyone traveling into the state must avoid contact with others for 14 days	11 March AK
7	Prohibiting visitation in hospitals and extended living facilities	Visitors aren't allowed in hospitals or nursing homes	13 March NH
8	Schools closure	K-12 schools closed for the remainder of the school year	16 March AK, AZ, DE, FL, IA, KY, LA, MD, MI, MT, NV, NH, NM, NC, ND, OR, PA, SC, UT, VA, WV, WY
9	Travel restriction/advisory to/from states	Warning against or restrictions to travel between states	17 March AK
10	Ceasing non-emergency medical and dental procedures	Any surgeries, evaluations, etc. that aren't to save lives are canceled	18 March OH
11	Non-essential business closure (stay-at-home order)	Any business that is not essential should be closed	19 March CA
12	Safer-at-home order	Usually enacted as a less extreme stay-at-home, non-essential businesses can open	19 March CA
13	Mandatory social distancing protocols for businesses	If businesses are to open, they must provide personal protection equipment for workers, adhere to social distancing policies, restrict the number of people inside the facilities, and alert the Public Health Department if workers test positive. Also, all workers who can perform their jobs from home should work remotely	24 March VA

(Continued on following page)

Table 3 continued

14	Social distancing/gatherings and meetings restrictions	Gatherings with members outside the household are prohibited	4 April AL
16	Face covering requirement	Everyone over an acceptable age is required to wear a cloth face covering in outdoor and indoor (public) areas	17 April NY

As a result, we evaluated the individual and overall effects of the most frequently implemented anti-contagion policies. For the first time in the COVID-19 literature, we present a novel index (policy ratio) highlighting the associations between policy implementation and COVID-19 spread for every state of the U.S. To perform the present research, we selected the most frequently implemented COVID-19 anti-contagion policies in the U.S. The study data includes policy activity, policy implementation duration, and the number of implemented policies in each of the 50 U.S. states for the study period, as presented in Figure 11.

3.2 Methods

3.2.1 Data Collection

We selected the most frequently implemented COVID-19 anti-contagion policies in the U.S. out of more than 50 policies issued from February 29, 2020 (see Table 3). The study data includes policy activity in each of the 50 U.S. states from March 1, through July 31, 2020 [96-98] (see Figure 1). We chose the end of July as the cut-off time because COVID-19 policy activity became relatively stable in the U.S.

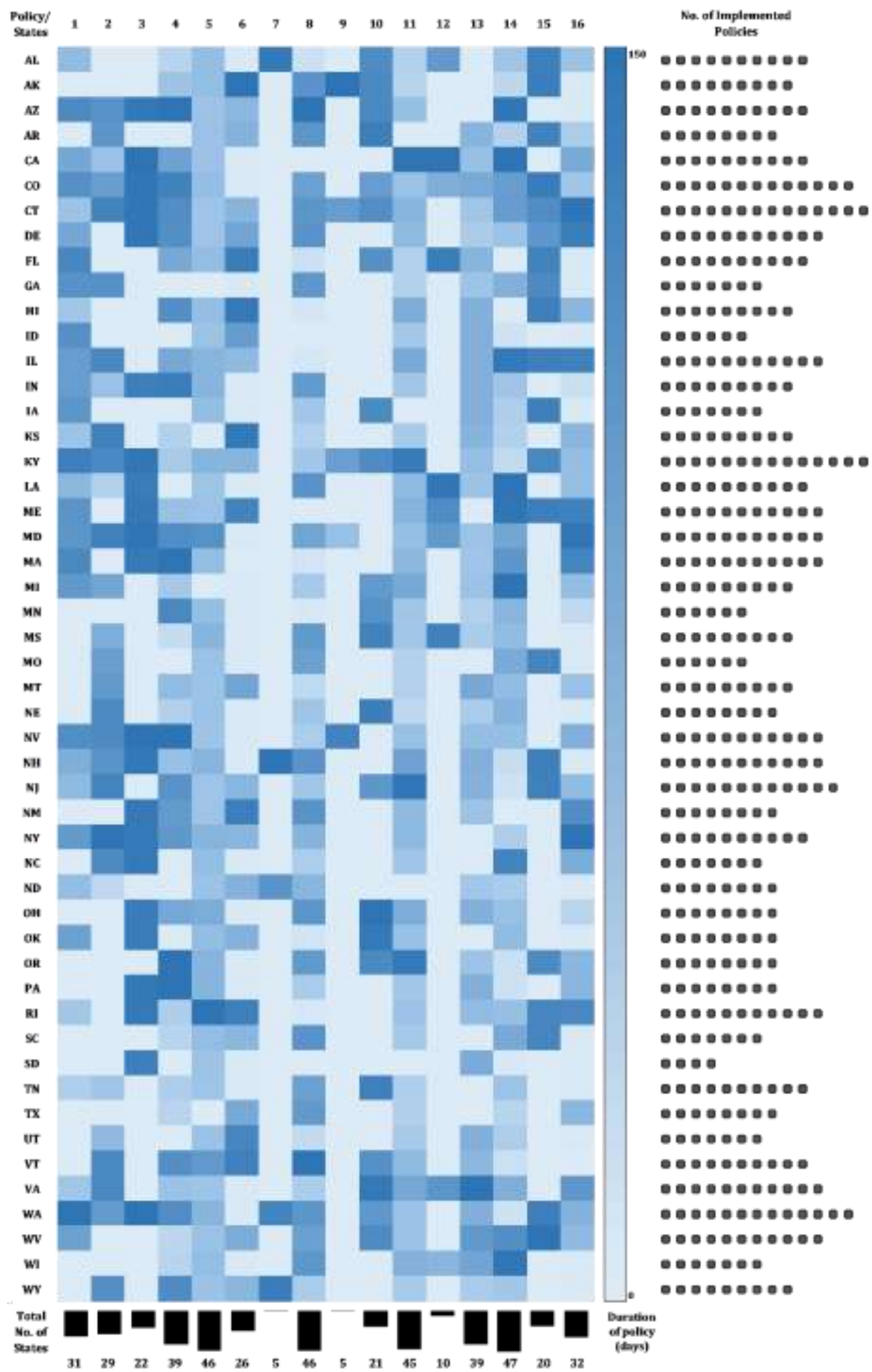


Figure 11: The duration of every policy in every state during the study period. States are ordered alphabetically based on their abbreviations. The policies listed can be found in Table 1.

The number of daily COVID-19 cases, deaths, and tests were obtained and cross-checked from multiple sources, including the Centers for Disease Control and Prevention (CDC) [99], Johns Hopkins Coronavirus Resource Center [100], The New York Times COVID data repository [101], and Worldometers data center [102] to ensure data integrity and consistency. The study used COVID-19 outcome measures (cases and deaths) starting two weeks after each policy was implemented and suspended to account for the response lag.

3.2.2 Epidemiological Analysis

Compartmental models are frequently used to model infectious diseases [103]. One such model, the Susceptible-Infectious-Recovered (SIR) model, has been used since the beginning of the COVID-19 pandemic to predict the spread [85, 104-106] and simulate the progress of COVID-19 [107, 108]. However, the SIR model fails to consider the latent phase—when the individual is infected but not yet infectious—which is an important period in the case of COVID-19 [109]. Adding a latent/exposed population can incorporate the latent phase within the SIR model. In this way, infected individuals move from susceptible to exposed to infected [110, 111]. As such, the Susceptible-Exposed-Infectious-Recovered (SEIR) model is defined by four coupled nonlinear ordinary differential equations (ODEs) as Eqs. 1-4 [112]:

$$\frac{dS(t)}{dt} = -\frac{\beta I(t)}{N} S(t) \tag{1}$$

$$\frac{dE(t)}{dt} = \frac{\beta I(t)}{N} S(t) - \sigma E(t) \tag{2}$$

$$\frac{dI(t)}{dt} = \sigma E(t) - \gamma I(t) \tag{3}$$

$$\frac{dR(t)}{dt} = \gamma I(t) \tag{4}$$

where *N, S, E, I,* and *R* represent the number of individuals in the population who are susceptible, exposed, infectious, and removed (recovered/deceased), respectively. Also, β, σ, and γ are the contact, infection, and recovery rates, respectively. In the present study, to obtain a more accurate result using the SEIR model, we recall the ODEs (Eqs. 1-4) in python™ and initialize the parameters E_{t-1}, I_{t-1}, and R_{t-1}. In this regard, for every simulation, we assumed the number of susceptible individuals equals the corresponding regional population collected from the U.S. Census Bureau [113]. Also, the number of exposed, infected, and recovered individuals on the first day of the study period was set at 1, 0, and 0, respectively.

To provide a reliable result for each state, the general SEIR model needed to be trained for every state using the corresponding reported state data. Otherwise, there would have been no differentiation among the contact rates (β). To train the model based on the available death data, we added the case fertility rate (CFR, the proportion of people who die from a specified disease among all individuals diagnosed with the disease over a certain period). Based on the CFR estimations for COVID-19 reported in the literature, we initialized the CFR to be 0.01% of the total population [114, 115]. Later, we implemented an optimizer using the least squares method to minimize the difference between the predicted and actual daily death rate for every state. In reality, due to the implementation of anti-contagion policies, the contact rate (β) and basic reproduction number (R_0, the number of secondary cases an individual would produce in a completely susceptible population [116, 117]) are not constant. Therefore, we defined β as time-dependent, representing the effective reproduction number (R_t, the number of secondary cases an individual would produce at any specific time [117, 118]). Having a time-dependent contact rate will increase the reliability of the model with the real-life situation [119]. Moreover, to estimate the values of incubation periods (i.e., the period of the days from the time the individual is exposed to the virus to the onset of symptoms) and infectious periods (i.e., the period in which an individual is infectious), we examined several studies [120-124] and considered the corresponding values to be 3 days and 10 days, respectively. Accordingly, both infection (σ) and recovery rates (γ), which are defined as the reciprocal of incubation and infectious periods, were calculated. As a result, the daily total number of susceptible, exposed, infected, and recovered cases for every state were estimated. We used the total numbers of daily infected cases for every state as response values to perform the statistical analysis.

There are four categories of infected cases, including non-diagnosed symptomatic, non-diagnosed asymptomatic, diagnosed symptomatic, and diagnosed asymptomatic [125]. According to previous studies [126-128] and the CDC [129], most people who were infected with COVID-19 were asymptomatic, and daily reported cases mainly included those symptomatic since they were the individuals more likely to get tested [130]. To perform comprehensive modeling, our modified SEIR model considers all those categories and represents the total number of daily infected cases for every state. We used this number as the response values for performing our statistical analysis. Also, to evaluate the overall effect of anti-contagion policy implementation, we used the SEIR model in two different scenarios. First, we considered every state's total population to predict the daily number of infected cases if there were no policies in effect. Next, we evaluated the daily number of infected cases with the policies in effect. Implementing the anti-contagion policies impacts both the number of daily cases and daily deaths. Since the reported number of infected cases (i.e., number of infected cases reported on data sources for each state) was not an accurate representation of the total infected cases, we used the daily reported deaths data to optimize the SEIR model. As a result, having the data of both scenarios, we calculated the policy ratio, or the average ratio of total infected cases when no policy was in effect compared to when policies were implemented, as Eq. 5:

$$Policy\ Ratio\ (PR) = \frac{\sum_{March\ 1}^{July\ 31} \frac{I_{np} - I_P}{I_{np}}}{period\ of\ time} \tag{5}$$

where I_{np} is the total number of daily infected cases when there is no policy in effect and I_P is the total number of daily infected cases when policies are implemented. The larger policy ratio represents the greater potential effect of policies in controlling COVID-19.

Figure 12 shows the flow diagram of the modified SEIR model developed in this study.

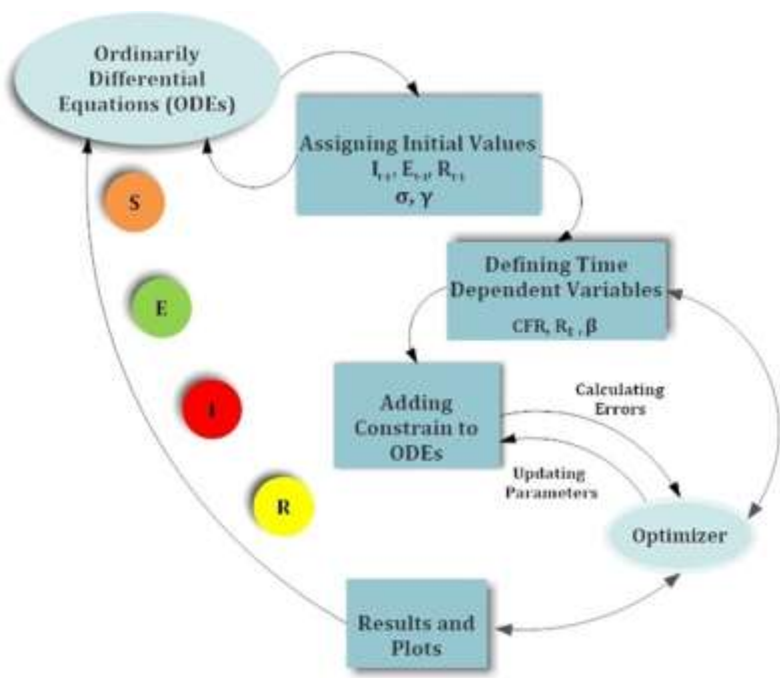


Figure 12: Flow diagram of modified SEIR model.

3.2.3 Statistical and Geospatial Analysis

To perform the statistical analysis, we defined the predictors (implemented anti-contagion policies) as categorical variables either in-effect or suspended. We used our modified SEIR model outcomes (i.e., the daily number of total infected cases) as response values to investigate every policy's potential effect in each U.S. state using statistical analysis. To find a statistically significant association between our predictors and response values, the appropriate regression modeling, either Negative Binomial Regression Modeling (NBRM) or Poisson Regression Modeling (PRM), was selected depending on data equidispersion or overdispersion [131, 132]. All the analyses were performed considering 95% confidence intervals (95% CI). We evaluated both tolerance and the variance inflation factors to diagnose the collinearity in multiple regression by observing the R^2 of regressing one predictor on all other predictors throughout the analysis. Accordingly, we removed all statistical noise (i.e., random irregularity). To investigate the effect of any probable noise and outlier, we examined the models using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). All statistical analyses were performed using SPSS (International Business Machines Corporation (IBM), Armonk, NY, USA). As a result, after removing noises and checking all the criteria (e.g., p-value (<0.05), CI (95%)), we determined which policies and to what extend decreased the number of daily infected cases for every state of the U.S. throughout the study period. Accordingly, we demonstrated the most effective anti-contagion policies in control of COVID-19 across the nation.

To find the associations between population density, area land, the effective reproduction number and policy ratio we used the geospatial analysis and superimposed different layers of data. I this regard, we collected the data and created the required layers using ArcGIS. Next, having the results of the epidemiological model, we ran the statistical analysis to predict the total number of infected cases. later, we created a layer of data using ArcGIS representing the status of the pandemic across the nation. Accordingly, having other layers of data including, population, state's area in squared miles and state land types (urban, rural) we generated the map for population density. Also, we followed the same approach and prepared the layers of data for demographic and socioeconomic factors as well.

In the next step, after completing our epidemiological and statistical analysis or both with policy and without policy implemented scenarios, we calculated the policy ratio for each state and accordingly prepared the corresponding layer of data. We also generated the map representing the variation of the effective reproduction number for all the 50 states of the U.S. After preparation of all of these separate layers of data as the map layer, we overlaid all the layers (superimposed) and evaluated whether there is any significant association between those factors. Our results showed the population density is highly correlated with the policy ratio. Also, we found a significant association between policy ratio and both the number of infected cases and deaths (see Results for more details).

3.3 Results

3.3.1 <u>The Most Effective Policies</u>

Our findings reveal that among all policies, mandatory quarantine upon entering a state (order No. 6 (see Table 3)), businesses implementing social distancing protocols (order No. 13), and mandating mask use (order No. 16) are the policies associated with reducing COVID-19 spread. There was a significant association between policy implementation and reduction in the total number of infected cases in the country for 55%, 75%, and 45% of the states that implemented mandatory quarantine upon entering a state, businesses implementing social distancing protocols, and mandating mask use, respectively. To reveal specific policy action that may help explain slowing the spread of infection, Figure 13 demonstrates the states which showed a significant decrease in the number of infected cases after implementing orders 6, 13, and 16. The findings suggested that implementing those policies is associated with an average 40% reduction in the total number of infected cases. Zeroing in on one state, for example, New York, shows that mandating mask use was associated with a 66% decrease in the total number of infected cases. Similarly, implementing the mandatory quarantine upon entering a state was associated with a 48% reduction in the total number of infected cases in New York. Table 4 presents the details of the multivariable binomial regression analyses, including the incidence rate ratio (i.e., the exponents of coefficients in the multiplicative Poisson model) and standard error for orders 6, 13, and 16.

3.3.2 Novel Policy Ratio Index

We defined and calculated a policy ratio (the average ratio of total infected cases when no policy was in effect compared to when policies were implemented - see Methods) for every state to represent the overall association of policy implementation with controlling the spread of COVID-19. Figure 14 demonstrates the value of the policy ratio calculated for each U.S. state. Alaska had the greatest impact (policy ratio: 3666), and South Dakota had the least impact (policy ratio: 17) from policy implementation, respectively.

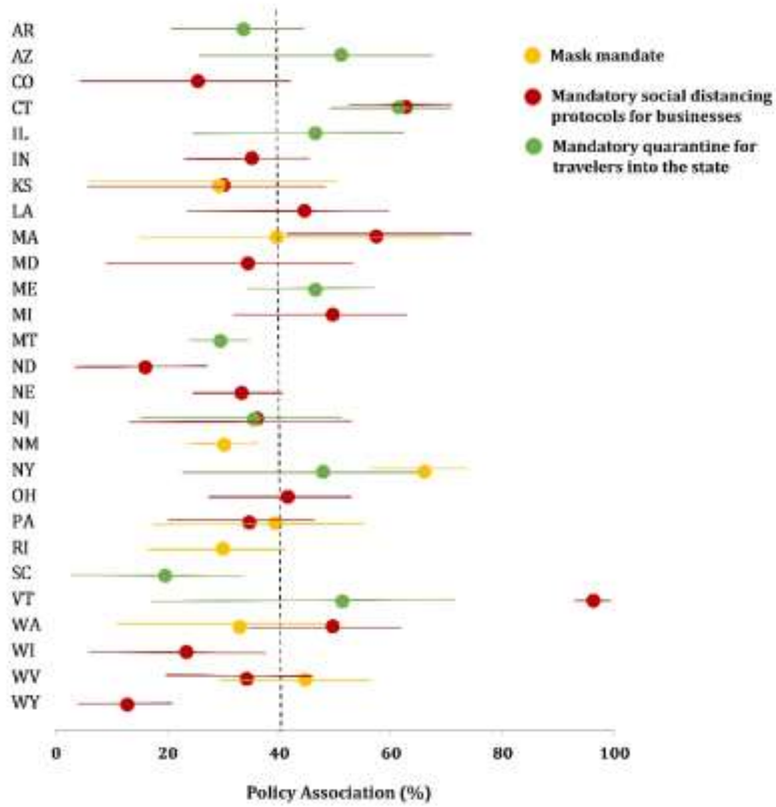


Figure 13: The association of each of three policies (mandating mask use, businesses implementing social distancing protocols, and mandatory quarantine upon entering a state) with a reduction in the number of COVID-19 infected cases. The colored lines represent the upper and lower bounds of the association. Also, the vertical dashed line demonstrates the average value of association for all states shown here.

Table 4:Multivariable Negative Binomial Regression (MNBR) analysis on COVID-19 infected cases (incidence rate ratio for the orders 6, 13, and 16; Exhibit is sorted alphabetically based on the states' abbreviations)

State	Order No.	IRR	SE	Sig	CI (95%)	
					L	H
AR	6	0.66	0.092	0.000	0.56	0.80
AZ	6	0.49	0.211	0.001	0.32	0.74
CO	13	0.74	0.128	0.021	0.58	0.96
CT	13	0.37	0.127	0.000	0.29	0.48
CT	6	0.39	0.140	0.000	0.29	0.51
IL	6	0.53	0.176	0.000	0.38	0.76
IN	13	0.65	0.087	0.000	0.55	0.77
KS	16	0.71	0.181	0.057	0.50	1.01
KS	13	0.70	0.153	0.019	0.52	0.94
LA	13	0.55	0.165	0.000	0.40	0.77
MA	16	0.60	0.176	0.004	0.43	0.85
MA	13	0.43	0.163	0.000	0.31	0.59
MD	13	0.65	0.171	0.013	0.47	0.91
MA	6	0.54	0.106	0.000	0.43	0.66
MI	6	0.50	0.155	0.000	0.37	0.68
MT	6	0.71	0.038	0.000	0.66	0.76
ND	13	0.84	0.072	0.015	0.73	0.97
NE	13	0.67	0.061	0.000	0.59	0.75
NJ	13	0.64	0.156	0.004	0.47	0.87
NJ	6	0.65	0.141	0.002	0.49	0.85
NM	16	0.70	0.045	0.000	0.64	0.76
NY	16	0.34	0.130	0.000	0.26	0.44
NY	6	0.52	0.200	0.001	0.35	0.77
OH	13	0.58	0.110	0.000	0.47	0.73
PN	16	0.61	0.158	0.002	0.45	0.83
PN	13	0.65	0.100	0.000	0.54	0.79
RI	16	0.70	0.090	0.000	0.59	0.84
SC	6	0.80	0.095	0.022	0.67	0.97
VT	13	0.04	0.179	0.000	0.03	0.05
VT	6	0.49	0.273	0.008	0.28	0.83
WA	16	0.67	0.142	0.005	0.51	0.89
WA	13	0.50	0.136	0.000	0.39	0.66
WI	13	0.77	0.104	0.010	0.62	0.94
WV	16	0.56	0.123	0.000	0.44	0.71
WV	13	0.66	0.103	0.000	0.54	0.80
WY	13	0.87	0.049	0.005	0.79	0.96

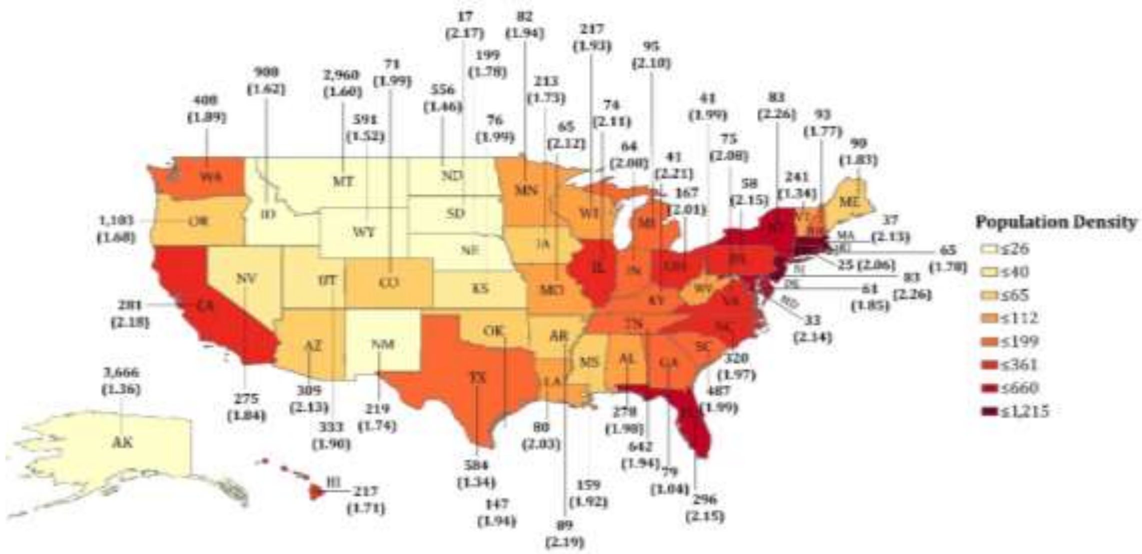


Figure 14: The distribution of policy ratio and population density over the study period (1 March to 31 July 2020. The average effective reproduction number (Rt) for every state is presented in parenthesis

Figure 15 compares the cumulative number of infected cases considering two different scenarios—one if there were no policies in effect (Figure 15a) and one with policies implemented (Figure 15b). Considering our calculated policy ratio and the cumulative number of infected cases reported by the end of July [101], we estimated the total number of infected cases when no policies were in effect. We estimate that implementing policies was associated with an average 58% reduction in the total number of infected cases (average of the column titled, "Estimated reduction in the number of infected cases due to implementing policies (%)" in Table 5). The anti-contagion policies were associated with nearly 10.8 million fewer Americans becoming infected by 31 July 2020.

According to research by Ioannidis [133], the COVID-19 infection fatality ratio (IFR; i.e., the proportion of deaths among all infected individuals) for the locations with mortality rates less than the global average (i.e., < 118 deaths/million) and high death rate (i.e., more than 500 deaths per million) is 0.20 and 0.57%, respectively. Johns Hopkins Coronavirus Resource Center [100] reported that the mortality rate of COVID-19 in the U.S. is more than 500 deaths per million. Therefore, considering the total number of infected cases when no policies were in effect (predicted by the present study) and 0.57% as the IFR, implementing the examined policies was associated with 61,560 fewer deaths nationwide as of 31 July 2020. Based on other estimations of IFR by Russel et al. [134] at 1.3% for all the ages combined and Brazeau et al. [135] at 1.15% (for high-income countries), the number of fewer deaths nationally associated with implementing the policies may have been as many as 140,000. Table 5 presents the cumulative number of reported and predicted infected cases and deaths considering two policy implementation scenarios.

Figure 15: The cumulative number of infected cases: a) no policies; b) policies in effect

Table 5: Evaluating the total number of infected cases and deaths when policies were in effect vs. no-policy scenario (1 March 2020 through 31 July 2020)

State	Number of predicted infected cases when no policy was in effect	Number of reported infected cases when policies were in effect	Estimated reduction in the number of infected cases associated with implementing policies (%)	Number of predicted deaths (using IFR = 0.57%) when no policy was in effect	Number of predicted deaths (using IFR = 1.15%) when no policy was in effect	Number of predicted deaths (using IFR = 1.3%) when no policy was in effect	Number of reported deaths
AL	331,681	87,723	74	1,891	3,814	4,312	1,580
AK	138,407	3,675	97	789	1,592	1,799	21
AZ	712,450	174,108	76	4,061	8,193	9,262	3,695
AR	80,448	42,511	47	459	925	1,046	453
CA	1,916,071	502,273	74	10,922	22,035	24,909	9,222
CO	80,431	46,948	42	458	925	1,046	1,841
CT	62,317	49,810	20	355	717	810	4,432
DE	23,769	14,788	38	135	273	309	585
FL	1,861,756	470,378	75	10,612	21,410	24,203	6,842
GA	306,548	171,342	44	1,747	3,525	3,985	3,674
HI	6,638	2,088	69	38	76	86	25
ID	210,317	20,853	90	1,199	2,419	2,734	193
IL	313,968	180,701	42	1,790	3,611	4,082	7,703
IN	111,056	67,800	39	633	1,277	1,444	2,965
IA	140,251	44,753	68	799	1,613	1,823	872
KS	49,502	28,123	43	282	569	644	358
KY	43,770	30,981	29	249	503	569	753
LA	209,009	116,394	44	1,191	2,404	2,717	3,949
ME	7,422	3,912	47	42	85	96	123

(Continued on following page)

Table 5 continued

MD	118,255	88,907	25	674	1,360	1,537	3,493
MA	160,540	117,612	27	915	1,846	2,087	8,609
MI	176,558	90,752	49	1,006	2,030	2,295	6,453
MN	99,424	54,503	45	567	1,143	1,293	1,640
MS	151,879	58,747	61	866	1,747	1,974	1,663
MO	84,372	51,045	40	481	970	1,097	1,305
MT	121,693	3,977	97	694	1,399	1,582	60
NE	78,374	26,211	67	447	901	1,019	338
NV	180,528	48,142	73	1,029	2,076	2,347	831
NH	12,728	6,583	48	73	146	165	415
NJ	239,825	183,535	23	1,367	2,758	3,118	15,819
NM	65,739	20,600	69	375	756	855	642
NY	765,994	419,723	45	4,366	8,809	9,958	32,372
NC	514,268	122,433	76	2,931	5,914	6,685	1,947
ND	42,466	6,473	85	242	488	552	107
OH	243,230	91,159	63	1,386	2,797	3,162	3,489
OK	89,995	36,456	59	513	1,035	1,170	541
OR	222,801	18,510	92	1,270	2,562	2,896	325
PA	184,196	116,787	37	1,050	2,118	2,395	7,261
RI	31,443	19,022	40	179	362	409	1,007
SC	522,328	89,016	83	2,977	6,007	6,790	1,712
SD	10,219	8,764	14	58	118	133	130
TN	765,659	103,144	87	4,364	8,805	9,954	1,047
TX	3,020,528	441,688	85	17,217	34,736	39,267	7,265
UT	174,314	40,249	77	994	2,005	2,266	308
VT	4,825	1,414	71	28	55	63	57
VA	157,735	89,888	43	899	1,814	2,051	2,174
WA	298,581	58,726	80	1,702	3,434	3,882	1,654
WV	9,342	6,642	29	53	107	121	116

(Continued on following page)

Table 5 continued

WI	181,027	57,020	69	1,032	2,082	2,353	943
WY	18,837	2,726	86	107	217	245	26
Total	15,353,514	4,539,615	70[1]	87,514	176,563	199,597	153,035

[1] Represents the overall reduction in the number of infected cases considering the total predicted number of infected cases when no policy was in effect, and the total reported number of infected cases when policies were in effect.

We evaluated the relation between our policy ratio and both normalized reported number of cases and deaths across the nation. Results demonstrated a significant negative association between policy ratio and those two criteria. This means the higher policy ratio is associated with fewer infected cases and deaths, which confirms the association of policy implementation with reducing COVID-19 spread (p-value < 0.05). Additionally, we investigated the relationship between the state policy ratio and population density. The correlation analysis exhibited a significant negative association (p-value < 0.05), demonstrating that policies have a greater association with reducing COVID-19 spread in the states with less population density. To test this correlation, using the modified SEIR model developed here, we calculated the effective reproduction number (R_t) over the study period in the states with greatest population density, including New Jersey (2.26), Rhode Island (1.78), and Massachusetts (2.13), and the least population density, including Montana (1.60), Wyoming (1.52) and Alaska (1.36). Our findings indicated that, on average, the effective reproduction number was 40% higher in states with the largest population density. This confirms an increase in the probability of becoming infected in more populated areas and, accordingly, a decrease in the potential of anti-contagion policies controlling COVID-19 in these areas. This finding is consistent with other research study by Hu et al. that showed that contact rates tend to increase with density [136]. Figure 14 presents the population density distribution, the policy ratio, and the average effective reproduction number for each state.

3.4 Discussion

We have presented COVID-19 case outcomes as they relate to the top 16 anti-contagion policies implemented in each of the U.S. states from 1 March 2020 through 31 July 2020. This study differs from other COVID-19 epidemiology and policy studies in that we modified the SEIR model and combined it with a comprehensive statistical analysis to better capture symptomatic and asymptomatic cases and examine, for the first time, the association of all the major policies implemented in the U.S. states with COVID-19 spread during the first six months of the pandemic. Additionally, we predicted the total number of COVID-19 infected cases and deaths if no policies were implemented to highlight the potential impact of anti-contagion policies in controlling COVID-19 in the U.S. Moreover, we calculated the average effective reproduction number for every U.S. state, indicating the number of secondary infections likely to occur from a single infection in every state.

Our findings show the policies of mandating mask use, businesses implementing social distancing protocols, and mandatory quarantine upon entering a state were associated with an average 40% reduction in the total number of infected cases. Additionally, policies implemented across the states may have saved nearly 10.8 million people from being infected. Considering different IFRs reported by other research studies and our estimated number of infected cases, results demonstrated that policies may have been associated with 140,000 fewer deaths nationwide. Note that previous studies based reported IFR values based on symptomatic cases only [137]. Yet, the majority of COVID-19 infected cases are asymptomatic [99]. Therefore, using the previously reported values of IFR, we expected the number of estimated deaths (Table 5) to be more than the reported number of deaths for all the U.S. states; however, it is not. The lower than expected reported number of deaths could be due to various other factors, such as population density, age, race, and ethnicity. We also calculated the average effective reproduction number for all U.S. states, and our results demonstrated a direct association between the population density and the effective reproduction number. This finding is not surprising given that when individuals are in closer proximity to one another, the droplet transmission and airborne transmission increase causing an acceleration in the spread of the virus [138, 139]. The policies of mask-wearing and businesses implementing social distancing found potentially effective in this study are consistent with other U.S. anti-contagion policy observational studies [140, 141]. However, our study provides even greater evidence for these policies because our epidemiological and statistical analysis accounted for many infected cases not included in other studies, which depended only on reported COVID-19 rates.

Several policies examined in this study did not show significant associations with reducing the number of infected cases across the U.S. One such policy that stands out is school closure which has sparked much debate throughout the COVID-19 pandemic [142]. Research based on previous influenza viruses indicated children would be major spreaders of the coronavirus [143]. However, the evidence produced since states first implemented school closure in mid-March 2020 supports our findings and indicates children are at significantly less risk for COVID-19 infection [144, 145]. However, children's role in spreading the disease remains unclear [146].

Further evidence that school closure may not be significantly linked to increasing infection rates can be seen in Europe when, throughout the fall of 2020, in contrast to the widespread canceling of in-person education in the U.S., schools remained opened using safety precautions like mask-wearing and ventilation without it significantly accelerating disease spread [147]. Experts have criticized school closures, arguing that closing schools may have caused social, economic, and health problems even more common and more severe than those due to COVID-19 [148], including increased risk of loneliness, addiction to videogames and binge-watching, alteration of circadian rhythms, direct or assisted domestic violence, and academic achievement gaps [149-151]. In learning from past pandemic research, a 2014 review by Mangtani [152] commissioned by the Department of Health in England (now known as the Department of Health and Social Care) concluded that "the benefit of school closure in reducing clinically important outcomes needs to be balanced against secondary adverse effects."

Social distancing (i.e., prohibiting gatherings with members outside the household), non-essential business closure (stay-at-home order), and safer-at-home orders were three other prominent policy areas not associated with a significant reduction in COVID-19 spread. One reason for these findings may be the difficulty with grouping the many types and levels of social distancing measures or restricting people to their homes into just one or two categories. The various aspects such as geographic level of examination, COVID-19 incidence rates at the time of implementation, duration, frequency, and intensity of such orders make evaluative studies on these policies difficult to compare to one another. For example, Thu et al. [153] reported a wide variation in the effectiveness of social distancing measures between ten highly infected countries. Additionally, the timing of social distancing, business closure, and safer-at-home orders in terms of how soon after the first reported case and how frequently (continuous versus intermittent) the policy is implemented can impact their effectiveness in reducing infection spread [154, 155].

The lack of accurate data about COVID-19 cases, especially asymptomatic cases, in the U.S. [38], posed a challenge to verifying our epidemiologic model's case and death values. While we used existing data to estimate these values, we can further analyze their precise association with policy implementation as more accurate retrospective case data becomes available over time. A limitation of this study is that the data does not account for the interactions between policies. We recommend future investigation into the interplay of policies to determine potential synergies and conflict, rollout approach, and communication strategies to emphasize prioritization. This analysis focuses mainly on COVID-19 infected cases and also considers deaths. However, we did not examine hospitalizations. Future investigations should include hospitalizations to represent severity and, together with health system capacity data, identify inadequacies in medical care to inform medical response-related policies.

The ultimate goal of this research is to help inform discussion about further policy actions state and national governments can take to curb infection spread now and during future pandemics. With 3 out of 16 (19%) of the examined policies accounting for the majority (about 75%) of positive policy-related impact on COVID-19, Pareto's principle [156], or the "80/20 rule," surfaces here. As the country continues reopening businesses and schools, decision-makers should focus on policies emphasizing mask-wearing, social distancing, and quarantining travelers as they may have the greatest chance of preventing the spread of COVID-19.

PART B: CONSTRUCTION A NOVEL INDEX OF LUNG CANCER RISK BASED ON MODIFIABLE FACTORS

3.5 Introduction

In the United States, cancer is the second leading cause of death [157], with lung cancer accounting for almost one-quarter of cancer-related deaths [158]. The American Cancer Society estimates that 235,760 new lung cancers will be diagnosed in 2021 and that this disease will claim the lives of more than 131,000 men and women in this country [159]. Numerous studies have examined the predominant risk factors for developing lung cancer [160-166]. The main risk factors are smoking [167], radon exposure [168], genetic susceptibility [169], gender [170], air pollution [171], body mass index [172], diet [173], indoor air pollution [174], passive smoking [175] and occupational exposure[176]. Of these, tobacco smoking, radon exposure, exposure, secondhand smoke, and alcohol consumption are modifiable lung cancer risk factors [177] (behaviors and exposures that can be changed [178]). Furthermore, evidence indicates the air pollution may also be considered a modifiable lung cancer risk factor [179, 180].

Our literature review demonstrated the association of each of the modifiable risk factors with developing lung cancer as follows: (1) tobacco smoking, the leading cause of lung cancer with 22-fold higher risk for current smokers [181, 182]; (2) radon exposure, the cause of 13.4% of lung cancer deaths [183]; (3) outdoor air pollution, associated with an increased risk of lung cancer in urban areas due to higher amounts of suspended particles; (4) indoor air pollution, cooking fumes is one example leading to a 5 to 20% increase in lung cancer risk [184-186]; (5) secondhand smoking, accounts for 3000 deaths per year due to lung cancer in the US [187, 188] and comprises one-third of lung cancer cases among nonsmokers [189]; (6) occupational exposure: overall contribution to lung cancer is relatively small [186] at between 9 and 15% [187]; (7) alcohol consumption, as one of the modifiable lung cancer risk factor may increase lung cancer risk after controlling for cigarette smoking [190-192].

Although many studies have investigated associations between individual risk factors and lung cancer risk or mortality [169, 193-201], less is known about how these factors interact to cause lung cancer or influence disease progression. Among the few published studies on the interaction between lung cancer risk factors, they focus on only a subset of risk factors. For example, one study showed that more than 85% of radon-induced lung cancer deaths were among smokers [202]. Ahrendt et al. [203] demonstrated a higher risk of lung cancer by examining the contribution of alcohol consumption and smoking. A study by Osann [204] found a significant interaction between smoking and family history. Another study [205] showed a significant interaction between history of other lung disease, secondhand smoke, smoking, and lung cancer.

For many study designs used in lung cancer risk factor research (e.g., case-control study, cohort study), there are limitations in providing a sample population that includes all major risk factors [206]. A significant challenge for such study designs is obtaining a sample population with a smoking history, occupational exposure and pollution, and other risk factors. To the best of our knowledge, no study has simultaneously investigated all of the modifiable risk factors of lung cancer. While several studies have demonstrated the individual and combined association between two or a few major lung cancer risk factors, little is known about the association of each risk factor while considering all the major modifiable risk factors in the causation of lung cancer.

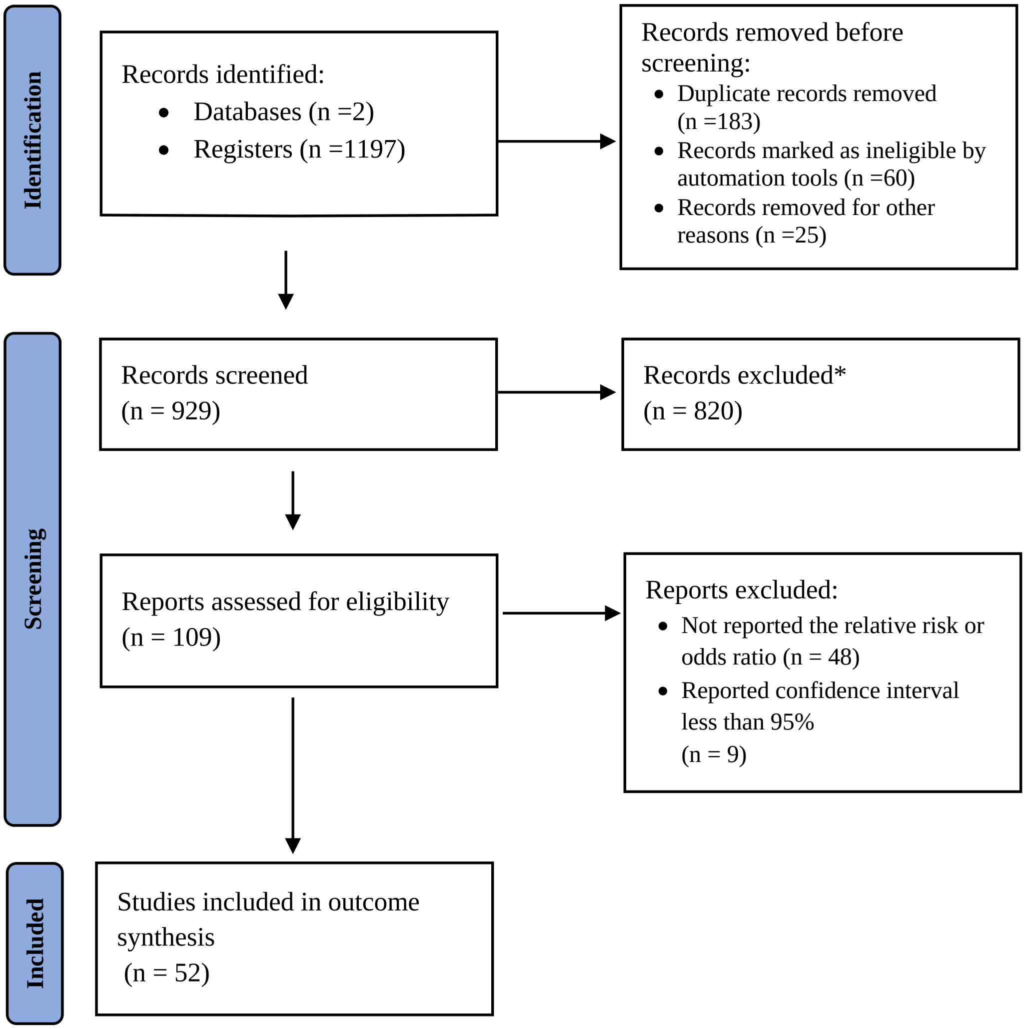
Developing a predictive model considering the major and minor modifiable risk factors of lung cancer is the scope of the present study. In this research, we performed a comprehensive review to discern the modifiable risk factors of lung cancer. We then collected the corresponding quantitative data (odds ratios (OR) and relative risks (RR)) to evaluate the overall effect size (i.e., a quantitative value measuring the strength of the relationship between two variables [207]) for each of the modified risk factors. To perform, we ran a meta-analysis on our collected data. Later, having those values, we developed a system to assign weights to risk factors and sorted them based on their degree of importance using Analytical Hierarchy Process (AHP); a the well-known multi-criteria decision making methods [208]. Consequently, we developed a lung cancer risk index based on modifiable factors. This index development methodology is intended to guide researchers in developing health indices based on multiple risk factors. Additionally, this index can inform public health officials and policymakers in decisions related to resource allocation for lung cancer prevention.

3.6 Methods

3.6.1 Search Strategy and Study Selection

Using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [209], we conducted database searches (from January 1990 up to April 2021) in PubMed (including MEDLINE) and Google Scholar for full-length published articles. We utilized the following keywords strings to capture relevant studies: “lung cancer” in conjunction with ”smoking,” “passive smoking,” “secondhand smoke,” “environmental tobacco smoking,” “radon, ” “occupational,” “air pollution,” “alcohol consumption,“ and “risk factors.” We included any article appearing in our keyword string search in initial abstract review. We reviewed titles and abstracts to ensure that the keywords were elaborated upon through the papers. The study inclusion criteria were (1) randomized controlled trial, prospective cohort study, retrospective cohort studies, case-cohort study, or nested case–control study; (2) reported the relative risk (RR) or odds ratio (OR) associated with increased risk; and (3) reported the 95% confidence intervals (CIs). We assessed the quality of the articles included in the present study using appraisal checklists provided by JBI (formerly known as "Joanna Briggs Institute"); an international organization focused on improving evidence as it relates to the feasibility, appropriateness, meaningfulness, and effectiveness of healthcare interventions [210].

The initial search yielded 1197 papers. After we removed 268 due to being duplicates or otherwise ineligible, the titles and abstracts for each of these studies were reviewed. After excluding 820 studies after the abstract reviews, at least two of the researchers reviewed the remaining studies’ full text. The authors further excluded 48 studies due to not reporting the relative risk or odds ratio and nine studies that reported CIs that were less than 95%. We included the remaining 52 articles in the analysis. The selection process suggested by PRISMA guidelines [211] is detailed in Figure 16.



* Papers were excluded based on the abstract and/or title. Those papers that did not meet the inclusion criteria (see methods) and did not present any quantitative value for risk factors.

Figure 16: Flowchart of search methodology and article selection

3.6.2 Quantitative Association of Each Risk Factor

We reviewed 52 papers that were qualified after applying the appraisal checklist (see search strategy and study selection) to understand how, quantitatively, those risk factors contribute to the occurrence of lung cancer. We presented the contribution of important risk factors to lung cancer using OR and RR are presented (Table 6).

3.6.3 Analytical Methods

3.6.3.1 Meta-analysis

To develop a more accurate estimate of the effect magnitude and establish statistical significance with studies having conflicting or disparate results, we developed a meta-analysis to obtain a pooled estimate of lung cancer risk factors from the collected studies. In this regard, the association between risk of smoking, occupational and radon exposure, secondhand smoking, indoor and outdoor air pollution, alcohol consumption, and occurrence of lung cancer was derived as a weighted average of study-specific estimates of the OR, using inverse variance weights [212]. The potential for publication bias was evaluated by funnel plots and the methods described by Egger et al. [213] and Begg et al. [214]. We analyzed the studies using a random-effects model [215] and considered heterogeneity and within-study variance. We evaluated heterogeneity using Cochrane's Q-statistic [216] and the I^2 statistic tests [217].

Table 6: Point estimates of the most important modifiable lung cancer risk factors

Risk factor	Gender	Type/Content	Measure	Point estimate	95% CI Lower limit	95% CI Upper limit	Ref
Smoking	Male		OR	7.82	4.59	13.30	[218]
Smoking	Female		OR	11.76	7.50	18.42	[218]
Smoking (United States smokers)	Both		OR	40.4	21.8	79.6	[219]
Smoking (Japanese smokers)	Both		OR	3.5	1.60	7.50	[219]
Smoking	Male		OR	9.6	5.64	16.30	[220]
Smoking	Female		OR	27.9	14.9	52.0	[220]
Smoking	Female		RR	3.4	1.75	6.61	[221]
Smoking	Male		RR	4.39	3.92	4.92	[222]
Smoking	Female		RR	2.79	2.44	3.20	[222]
Smoking	Both		OR	14.9	12.3	18.1	[223]
Smoking	Male		OR	5.0	2.0	12.7	[224]
Smoking	Both		OR	9.40	6.9	12.8	[225]
Smoking	Female		OR	13.6	12.3	15.1	[226]
Smoking	Male		OR	11.3	10.2	12.4	[226]
Smoking	Female		OR	8.94	7.54	10.6	[227]
Smoking	Female		OR	4.87	1.34	17.75	[174]
Radon exposure	Both	(200+ Bq/m^3)	OR	1.29	0.98	1.70	[228]
Radon exposure	Both	(150-199 Bq/m^3)	OR	1.19	0.86	1.66	[228]
Radon exposure	Both	(100-149 Bq/m^3)	OR	1.22	0.88	1.69	[228]
Radon exposure - North America	Both	(up to 100 Bq/ m^3)	OR	1.106	1	1.28	[229]

(Continued on following page)

Table 6 continued

Radon exposure - China	Both	(up to 100 Bq/m^3)	OR	1.139	1.01	1.37	[229]
Radon exposure	Both	(100 Bq/m^3)	OR	1.33	1.01	1.36	[230]
Radon exposure	Both		OR	1.73	1.27	2.35	[231]
Radon exposure - Germany	Both	(up to 80 Bq/m^3)	OR	1.59	1.08	2.27	[197]
Radon exposure - Germany	Both	(above 80 Bq/m^3)	OR	1.93	0.99	3.77	[197]
Radon exposure	Both	50-80 Bq/m^3	RR	1.08	0.79	1.47	[232]
Radon exposure	Both	80-140 Bq/m^3	RR	1.18	0.86	1.61	[232]
Radon exposure	Both	above 140 Bq/m^3	RR	1.44	1	2.06	[232]
Occupational exposure	Both	(welding fumes)	OR	2.50	1.0	6.5	[225]
Occupational exposure	Both	(asbestos)	OR	1.50	0.5	4.4	[225]
Occupational exposure	Both	(wood dust)	OR	1.90	1.2	3.1	[225]
Occupational exposure	Both	(diesel exhaust)	OR	3.10	2.1	4.5	[225]
Occupational exposure	Both		OR	1.60	1.4	2.1	[161]
Occupational exposures	Both		OR	2.10	1.3	3.3	[161]
Occupational exposure	Both	(crystalline silica)	OR	1.37	1.14	1.65	[233]
Occupational exposure	Both	(crystalline silica)	OR	1.41	1.22	1.62	[234]
Occupational exposure	Both	(diesel exhaust)	OR	1.43	1.23	1.67	[234]
Occupational exposure	Both	(polycyclic aromatic hydrocarbons)	OR	1.53	1.14	2.04	[234]
Occupational exposure	Both	(asbestos)	OR	1.78	0.94	3.36	[235]
Occupational exposure (Low concentrations)	Both	(asbestos)	OR	1.17	0.92	1.5	[236]
Occupational exposure (medium or high concentrations)	Both	(asbestos)	OR	2.16	1.21	3.88	[236]
Second hand smoking (highly exposed)	Both		RR	2.01	1.33	2.60	[237]

(Continued on following page)

Table 6 continued

Second hand smoking	Both		OR	1.26	1.06	1.47	[238]
Second hand smoking	Female		OR	1. 31	0.99	1.72	[239]
Second hand smoking	Both		RR	1.08	0.6	1.94	[240]
Second hand smoking	Both		RR	1.05	0.6	1.86	[241]
Second hand smoking	Female		OR	2.95	1.32	6.57	[174]
Second hand smoking	Both		OR	1.57	1.07	2.31	[242]
Second hand smoking	Both		RR	1.9	1.0	3.50	[243]
Second hand smoking	Female		RR	1.3	1.0	1.7	[244]
Second hand smoking	Female		OR	1.3	0.7	1.5	[245]
Outdoor air pollution	Both	(pesticides)	OR	5.10	3.1	8.3	[225]
Outdoor air pollution	Both		RR	1.25	1.18	1.32	[246]
Outdoor air pollution	Both	(mortality)	RR	1.23	1.16	1.30	[246]
Outdoor air pollution	Both	(diesel)	OR	3.10	2.1	4.5	[225]
Outdoor air pollution	Both		OR	1.46	0.89	2.40	[247]
Outdoor air pollution (low concentrations)	Both	(nitrogen dioxide)	OR	1.14	0.78	1.67	[247]
Outdoor air pollution (high concentrations)	Both	(nitrogen dioxide)	OR	1.30	1.02	1.66	[247]
Outdoor air pollution	Both	(PM2.5)	OR	1.29	0.95	1.76	[248]
Outdoor air pollution	Both	(nitrogen dioxide)	OR	1.34	1.07	1.69	[248]
Outdoor air pollution	Both	(nitrogen dioxide)	OR	1.3	1.02	1.66	[249]
Outdoor air pollution	Both	(PM10)	OR	1.05	0.65	1.69	[249]
Outdoor air pollution	Both	(sulfur dioxide)	OR	1.15	0.92	1.43	[249]
Outdoor air pollution	Both	(PM2.5)	RR	1.14	1.04	1.23	[250]
Outdoor air pollution	Both	(NOx)	RR	1.08	1.02	1.15	[251]
Outdoor air pollution	Both	(SOx)	RR	1.01	0.94	1.08	[251]

(Continued on following page)

Table 6 continued

Outdoor air pollution	Both	(PM10)	RR	1.66	1.21	2.27	[252]
Outdoor air pollution	Both	(NOx)	RR	1.10	0.97	1.23	[253]
Outdoor air pollution	Both	(SOx)	RR	1.01	0.98	1.03	[253]
Indoor air pollution	Both	HAP exposure	OR	1.77	1	3.14	[254]
Indoor air pollution	Both	(PAH25)	OR	2.21	1.67	2.87	[255]
Indoor air pollution	Both	(NO2)	OR	2.06	1.19	3.49	[255]
Indoor air pollution (ex-smokers)	Both		OR	4.30	2.7	6.8	[185]
Indoor air pollution	Both	(cooking only)	OR	1.15	0.97	1.37	[256]
Indoor air pollution	Both	(heating and cooking)	OR	1.17	1.01	1.37	[256]
Indoor air pollution	Female	wood or straw as cooking fuels	OR	1.77	1.08	2.91	[242]
Indoor air pollution	Both	Coal consumption (heating and cooking)	OR	1.29	1.03	1.61	[257]
Indoor air pollution	Female	cooking oil	OR	2.54	1.40	4.30	[258]
Indoor air pollution	Female	Coal consumption (heating and cooking)	OR	1.3	0.3	5.80	[245]
Alcohol consumption	Both	(>60 g/day)	OR	1.44	1.01	2.07	[190]
Alcohol consumption	Both	(>20 g/day)	OR	1.42	1.06	1.90	[196]
Alcohol consumption	Both		OR	1.60	1.00	2.04	[259]
Alcohol consumption	Both	(white wine)	OR	1.20	1.01	1.42	[260]
Alcohol consumption	Both	(alcoholism)	RR	2.40	2.29	2.51	[261]
Alcohol consumption	Both	Above 9 drinks/month	RR	1.1	0.7	1.6	[259]
Alcohol consumption	Both	Above 1 drink/day	RR	1.9	1.0	3.4	[262]
Alcohol consumption	Both	Above 0.5 drink/week	RR	1.1	0.6	2.1	[263]
Alcohol consumption	Both	Above 7 drinks/week	RR	1.2	0.8	1.7	[264]

Considering the confidence intervals (CIs), we calculated all studies' standard errors (SEs). We used the OR, logOR, and the corresponding SEs as data points for conducting the meta-analysis. For each risk factor, all the studies were plotted in order of decreasing variance of the logOR, where the horizontal lines represent the 95% CIs. All statistical manipulations were conducted using the meta-analysis package for R (metaphor Version 2, MA, USA). We used the results of our meta-analysis as an input to the following AHP analysis to increase the accuracy of outcomes since the meta-analysis could remove noise and decrease biases [265, 266].

3.6.3.2 Analytic Hierarchy Process (AHP)

AHP is known as one of the most widely used Multi-Criterion Decision Making (MCDM) methods [267]. As an analytical model, AHP has been implemented in healthcare and the number of studies applying AHP in healthcare has increased since 2005 [268]. Since AHP is capable of quantitatively prioritizing our risk factors by producing the weights for all of lung cancer major and minor modifiable risk factors, we selected it as our analytics method in the present study.

For each risk factor, we used the overall effect size calculated from the meta-analysis for smoking, occupational and radon exposure, secondhand smoking, indoor and outdoor air pollution, and alcohol consumption. To illustrate we analyzed the studies (Table 6; Fig. 16) using a random-effects model (see Meta-analysis) and imported the outcomes of that model (overall odds ratio) as our input variables in AHP for each of the risk factors. For the purpose of our analysis, since the odds ratio is a good approximation of relative risk when the outcome is rare, we consider odds ratio to represent relative risks. This approximation has been used in prior studies [269]. Having the odds ratios (calculated from meta-analysis) for all the risk factors, we normalized the ratios using Eq. 6:

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{6}$$

where $X_{normalized}$ is the normalized ratio, X is the average ratio of each risk factor, and X_{min} and X_{max} are the minimum and maximum ratio of the same risk factor, respectively. We sorted all the ratios between 0 and 1 to establish priority amongst risk factors using pairwise comparison. In this regard, having the normalized values of ratios and using the assessment matrix, we created the pair-wise comparison matrix (i.e., a matrix to perform the process of comparing risk factors in pairs to evaluate their relative importance). To define the assessment matrix in AHP for pairs with equal, weak, obvious, intense, and extreme importance, the values of 1, 3, 5, 7, and 9 were assigned, respectively [270]. Also, values of 2, 4, 6 and 8 were assigned for intermediate importance.

To illustrate, as the first step, the relative importance of smoking versus radon exposure, occupational exposure, secondhand smoking, outdoor air pollution, indoor air pollution and alcohol consumption were assigned considering the assessment matrix. This step was then repeated for all risk factors. Next, an n by n matrix was created where n represented the number of modified risk factors. Next, having one of the eigenvalues of the matrix, we defined the eigenvectors and solved the linear system considering the matrix coefficient using Eq. 7.

$$AX = \lambda X \; or \; (A - \lambda I_n)X = 0 \tag{7}$$

where A was the comparison matrix of order n and λ is one of its eigenvalues. X represented the eigenvector of A associated to λ, and $A - \lambda I_n$ represented the matrix coefficient. It should be noted that to calculate the eigen values and vectors of the matrix, MATLAB (MathWorks, Massachusetts, USA) is frequently used which is recommended by authors. Eventually, the derived eigenvector was used to specify the weights of each risk factor where the eigenvector represents the coefficient of the index (see Association of modifiable risk factors in developing lung cancer using AHP). As a result, the contribution of each risk factors to causation of lung cancer was estimated.

To check the reliability of our results we evaluated the Consistency Ratio (*CR*). In this regard, we first calculated the Consistency Index (CI_1) using Eq. 8:

$$CI_1 = (\lambda max - n)/(n - 1) \tag{8}$$

where λmax was the maximum eigenvalue and n represented the order of the matrix. Accordingly, the Consistency Ratio was calculated by dividing the Consistency Index (CI_1) by the index for the corresponding random matrix (RI) using Eq. 9:

$$CI_1 = CI_1/RI \tag{9}$$

The values for *RI* has been presented by Saaty considering the matrix size [271]. Also, Saaty [271] suggested that the *CR* needs to be less than 0.1 to produce consistent results. It should be noted that, although AHP has been used in numerous MCDM, it has the limitation that imposes a single cut point on the data. This means a continuous number cannot be assigned to the index coefficients (A_1 to A_7).

3.7 Results

3.7.1 Overall Effect Size of Lung Cancer Modifiable Risk Factors

We evaluated the association of modifiable risk factors in developing lung cancer by running a meta-analysis on the data collected (Table 6). Our results indicated the overall effect size for smoking, radon exposure, indoor air pollution, secondhand smoke, exposure to cancer-causing agents, alcohol consumption, and outdoor air pollution were 8.63, 1.24, 1.76, 1.43, 1.60, 1.45, and 1.25, respectively. Therefore, smoking and radon exposure were the most and least important modifiable risk factors that increase the probability of lung cancer.

3.7.2 Association of Modifiable Risk Factors in Developing Lung Cancer Using AHP

We evaluated the association of each risk factor in developing lung cancer by implementing AHP and using the overall OR effect sizes calculated through meta-analysis for each of the modifiable lung cancer risk factors (Table 7). The consistency ratio of the present study was 0.07, i.e., within recommended range of smaller than 0.1, which demonstrated the consistency of the analysis.

Having the final weight for all the modifiable risk factors, we produced a Lung Cancer Risk Index (LCRI) representing the probability of getting lung cancer as follows:

$$LCRI = 0.461A_1 + 0.185A_2 + 0.132A_3 + 0.078A_4 + 0.076A_5 + 0.038A_6 + 0.030A_7 \quad (9)$$

where A_1 to A_7 represent smoking, indoor air pollution, exposure to cancer-causing agents (occupational exposure), alcohol consumption, secondhand smoke, outdoor air pollution, and radon exposure, respectively. It should be noted that A_1 to A_7 take 0 or 1 only, where 0 indicated the corresponding risk factor was not in effect and 1 indicated the corresponding risk factor was in effect. For instance, in developed countries, heat source such as coal doesn't apply for cooking; therefore, in the case of utilizing the index to evaluate the risk of lung cancer for individuals living in those countries, the A_2 can be considered to be 0. The $LCRI$ could have any value between 0 to 1 in which 0 means no lung cancer risk and 1 represents the highest risk of lung cancer occurrence. The index represents the risk of lung cancer occurrence for individuals.

Table 7: Average odds ratio and final weights of modifiable risk factors of lung cancer

Risk factor	**Overall effect size**	**Final Weight (%)**
Smoking	8.63	46.1
Indoor air pollution	1.76	18.5
Occupational exposure	1.60	13.2
Alcohol consumption	1.45	7.8
Second hand smoking	1.43	7.6
Outdoor air pollution	1.25	3.8
Radon exposure	1.24	3.0

CHAPTER 4: CONCLUSION AND RECOMMENDATIONS FOR FUTURE RESEARCH

4.1 Conclusion

The aim of this thesis was to investigate the applications of data and geospatial analysis in both mechanical engineering (energy management) and health systems. In this regard, using data and geospatial analysis we evaluated the potential of the state of Illinois in achieving recovery energy in both forms of electricity and biofuels out of Municipal Solid Waste (chapter 2) (MSW). Our results demonstrated that Illinois is capable of producing 6,295,385.77 MWH annual energy using incineration technology from MSW. Also, using Anaerobic Digestion (AD) technology in MSW management enables the state to be capable of producing more than 1,140,493,710,450.00 Litres biogas per year. Moreover, our study showed that the total weight of biomass (MSW composite) is 9,504,114.25 tonnes per year and using the pyrolytic technology, more than 190,082.29 MJ can be annually produced.

We demonstrated that there exists a potential to produce 1,000 Mm^3/yr of CH_4 through anaerobic digestion of organic MSW (which could be subsequently used in gas turbines to generate electricity), and 2,000 GWh electricity via waste-to-energy (WTE) plants using MSW across the state of Illinois. This study study demonstrates the vast potential across the State of Illinois to produce clean energy from MSW. In addition to environmental benefits, this will promote clean energy technologies, create job opportunities, diversify energy resources and help toward a more sustainable economic development of the State of Illinois.

In addition, we extended the application of data and geospatial analysis in health system to analyze the impact of anti-contagion policies implemented by the states across the country to slow the spread of COVID-19. Also, by implementing a meta-analysis in conjunction with multi-criteria decision-making methods, a Lung Cancer Risk Index (LCRI) was produced representing the probability of individuals to get lung cancer. The methods that have been developed for the extended applications of data and geospatial analysis in health systems can be used for various complex decision making and index generating purposes in engineering disciplinary such as additive manufacturing to evaluate the effect of process factors (e.g., injection content, speed, temperature) individually and collectively to optimize the process and increase the performance.

4.2 Recommendations for Future Research

The ultimate goal of this research project was to develop comprehensive MSW management framework/model for assessing the MSW management options that provides the user the possibility to compare various scenarios in terms of economic feasibility as well as environmental assessments. The result of this study can help governments, county administrators, city councils, private organizations, investors, landfill owners etc. make informed decisions about diverting MSW from landfills and converting into energy. We recommend future investigation to used our procedure and results and add more layer of data and information to find the best location for creating facilities to achieve energy from MSW. Furthermore, the proposed research has broad applicability to the other states and worldwide, hence sustained research, development and outreach activities are recommended.

REFERENCES

1. Schiffer, H.-W., T. Kober, and E. Panos, *World energy council's global energy scenarios to 2060.* Zeitschrift für Energiewirtschaft, 2018. **42**(2): p. 91-102.

2. Siyal, S.H., et al., *Wind energy assessment considering geographic and environmental restrictions in Sweden: A GIS-based approach.* Energy, 2015. **83**: p. 447-461.

3. Arnette, A.N. and C.W. Zobel, *Spatial analysis of renewable energy potential in the greater southern Appalachian mountains.* Renewable Energy, 2011. **36**(11): p. 2785-2798.

4. Van Hoesen, J. and S. Letendre, *Evaluating potential renewable energy resources in Poultney, Vermont: A GIS-based approach to supporting rural community energy planning.* Renewable energy, 2010. **35**(9): p. 2114-2122.

5. Groppi, D., et al., *A GIS-based model to assess buildings energy consumption and usable solar energy potential in urban areas.* Sustainable cities and society, 2018. **40**: p. 546-558.

6. Ghose, M., A.K. Dikshit, and S. Sharma, *A GIS based transportation model for solid waste disposal–A case study on Asansol municipality.* Waste management, 2006. **26**(11): p. 1287-1293.

7. Vijay, R., et al., *GIS-based locational analysis of collection bins in municipal solid waste management systems.* Journal of Environmental Engineering and Science, 2008. **7**(1): p. 39-43.

8. Singh, A., *Remote sensing and GIS applications for municipal waste management.* Journal of environmental management, 2019. **243**: p. 22-29.

9. Sumathi, V., U. Natesan, and C. Sarkar, *GIS-based approach for optimized siting of municipal solid waste landfill.* Waste management, 2008. **28**(11): p. 2146-2160.

10. Sarptas, H., N. Alpaslan, and D. Dolgen, *GIS supported solid waste management in coastal areas.* Water Science and Technology, 2005. **51**(11): p. 213-220.

11. Berrigan, D. and N.A. Berger, *Geospatial Approaches to Energy Balance and Breast Cancer*. Vol. 15. 2019: Springer.

12. Bates, D.W., *The quality case for information technology in healthcare.* BMC medical informatics and decision making, 2002. **2**(1): p. 7.

13. Clauser, S.B., et al., *Improving modern cancer care through information technology.* American journal of preventive medicine, 2011. **40**(5): p. S198-S207.

14. Orton, L., et al., *The use of research evidence in public health decision making processes: systematic review.* PloS one, 2011. **6**(7).

15. Church, J., et al., *Citizen participation in health decision-making: past experience and future prospects.* Journal of public health policy, 2002. **23**(1): p. 12-32.

16. Kelley, T.L., *Statistical method*. 1923: Macmillan.

17. Karran, J.C., E.E. Moodie, and M.P. Wallace, *Statistical method use in public health research.* Scandinavian journal of public health, 2015. **43**(7): p. 776-782.

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

19. Wang, F., *Why public health needs GIS: A methodological overview.* Annals of GIS, 2020. **26**(1): p. 1-12.

20. Franch-Pardo, I., et al., *Spatial analysis and GIS in the study of COVID-19. A review.* Science of The Total Environment, 2020. **739**: p. 140033.

21. Khashoggi, B.F. and A. Murad, *Issues of healthcare planning and GIS: a review.* ISPRS International Journal of Geo-Information, 2020. **9**(6): p. 352.

22. Wilkinson, P., et al., *GIS in public health.* GIS and Health, 1998. **6**: p. 179-190.

23. McLafferty, S.L., *GIS and health care.* Annual review of public health, 2003. **24**(1): p. 25-42.

24. Gattrell, A. and M. Loytonen, *GIS and Health*. Vol. 6. 1998: Houghton Mifflin Harcourt.

25. Cromley, E.K. and S.L. McLafferty, *GIS and public health*. 2011: Guilford Press.

26. Wall, P.A. and O.J. Devine, *Interactive analysis of the spatial distribution of disease using a geographic information systems.* Journal of Geographical Systems, 2000. **2**(3): p. 243-256.

27. Garnelo, L., L.C. Brandão, and A. Levino, *Dimensions and potentialities of the geographic information system on indigenous health.* Revista de saude publica, 2005. **39**: p. 634-640.

28. Myers, M.F., et al., *Forecasting disease risk for increased epidemic preparedness in public health.* Advances in parasitology, 2000. **47**: p. 309-330.

29. Kalipeni, E. and L. Zulu, *Using GIS to model and forecast HIV/AIDS rates in Africa, 1986–2010.* The Professional Geographer, 2008. **60**(1): p. 33-53.

30. Dom, N.C., et al., *Integration of GIS-Based Model with Epidemiological data as a Tool for Dengue Surveillance.* EnvironmentAsia, 2017. **10**(2).

31. Aghajani, J., P. Farnia, and A.A. Velayati, *Impact of geographical information system on public health sciences.* Biomedical and Biotechnology Research Journal (BBRJ), 2017. **1**(2): p. 94.

32. Jarup, L., *The role of geographical studies in risk assessment.* Spatial epidemiology: Methods and applications, 2000: p. 415-33.

33. Cromley, E. and S. McLafferty, *GIS and public health. 2002.* New York: Guilford, 2002.

34. Driedger, S.M., et al., *Correction: Using participatory design to develop (public) health decision support systems through GIS.* International Journal of Health Geographics, 2007. **6**(1): p. 53.

35. Nykiforuk, C.I. and L.M. Flaman, *Geographic information systems (GIS) for health promotion and public health: a review.* Health promotion practice, 2011. **12**(1): p. 63-73.

36. García-Rangel, S. and N. Pettorelli, *Thinking spatially: the importance of geospatial techniques for carnivore conservation.* Ecological informatics, 2013. **14**: p. 84-89.

37. Higgs, G., *A literature review of the use of GIS-based measures of access to health care services.* Health Services and Outcomes Research Methodology, 2004. **5**(2): p. 119-139.

38. Schneider, E.C., *Failing the Test—The Tragic Data Gap Undermining the US Pandemic Response.* New England Journal of Medicine, 2020.

39. Chambal, S., M. Shoviak, and A.E. Thal, *Decision analysis methodology to evaluate integrated solid waste management alternatives.* Environmental Modeling & Assessment, 2003. **8**(1): p. 25-34.

40. Fiorucci, P., et al., *Solid waste management in urban areas: development and application of a decision support system.* Resources, conservation and recycling, 2003. **37**(4): p. 301-328.

41. Twumasi, Y.A. and E.C. Merem, *Using remote sensing and GIS in the analysis of ecosystem decline along the River Niger Basin: the case of Mali and Niger.* International Journal of Environmental Research and Public Health, 2007. **4**(2): p. 173-184.

42. Chang, N.-B., H. Lu, and Y. Wei, *GIS technology for vehicle routing and scheduling in solid waste collection systems.* Journal of environmental engineering, 1997. **123**(9): p. 901-910.

43. Le Hoang, S., *Optimizing municipal solid waste collection using chaotic particle swarm optimization in GIS based environments: a case study at Danang city, Vietnam.* Expert systems with applications, 2014. **41**(18): p. 8062-8074.

44. Javaheri, H., et al., *Site selection of municipal solid waste landfills using analytical hierarchy process method in a geographical information technology environment in Giroft.* Journal of environmental health science & engineering, 2006. **3**(3): p. 177-184.

45. Malakahmad, A., et al., *Solid waste collection routes optimization via GIS techniques in Ipoh city, Malaysia.* Procedia Engineering, 2014. **77**: p. 20-27.

46. Şener, Ş., E. Sener, and R. Karagüzel, *Solid waste disposal site selection with GIS and AHP methodology: a case study in Senirkent–Uluborlu (Isparta) Basin, Turkey.* Environmental monitoring and assessment, 2011. **173**(1): p. 533-554.

47. Dutta, D. and S. Goel, *Applications of remote sensing and GIS in solid waste management–A review*, in *Advances in solid and hazardous waste management*. 2017, Springer. p. 133-151.

48. Adeofun, C., et al., *Application of remote sensing and geographic information system for selecting dumpsites and transport routes in Abeokuta, Nigeria.* COLERM Proceedings, 2012. **1**: p. 264-278.

49. *Illinois commodity/waste generation and characterization study update*. 2015, Illinois Recycling Association IL, USA.

50. IEC. *Reduce-reuse-recycle*. 2020 [cited 2020 June 25]; Available from: https://ilenviro.org/issues/reduce-reuse-recycle/.

51. *State energy price and expenditure estimates 1970 through 2012*. 2014, U.S. E.I.A.: Washington, DC.

52. *Data indicates 4 remaining Illinois coal plants among worst polluters*. 2019, JournalStar, The Associated Press.

53. Dong, J., et al., *Comparison of waste-to-energy technologies of gasification and incineration using life cycle assessment: Case studies in Finland, France and China.* Journal of Cleaner Production, 2018. **203**: p. 287-300.

54. Czajczyńska, D., et al., *Potential of pyrolysis processes in the waste management sector.* Thermal Science and Engineering Progress, 2017. **3**: p. 171-197.

55. Baere, L.d. and B. Mattheeuws, *Anaerobic digestion of msw in EUROPE.* BioCycle, 2010. **51**(2): p. 24-26.

56. Macias-Corral, M., et al., *Anaerobic digestion of municipal solid waste and agricultural waste and the effect of co-digestion with dairy cow manure.* Bioresource technology, 2008. **99**(17): p. 8288-8293.

57. Rivard, C., et al., *Anaerobic digestion of municipal solid waste.* Applied Biochemistry and Biotechnology, 1995. **51**(1): p. 125-135.

58. Holtman, K.M., et al., *Pilot scale high solids anaerobic digestion of steam autoclaved municipal solid waste (MSW) pulp.* Renewable Energy, 2017. **113**: p. 257-265.

59. Wu, F., et al., *A new coronavirus associated with human respiratory disease in China.* Nature, 2020. **579**(7798): p. 265-269.

60. Tang, X., et al., *On the origin and continuing evolution of SARS-CoV-2.* National Science Review, 2020.

61. *Centers for Disease Control and Prevention; COVID Data Tracker*. 2021 [cited 2021 Feb. 1]; Available from: https://covid.cdc.gov/covid-data-tracker/#datatracker-home.

62. Holshue, M.L., et al., *First case of 2019 novel coronavirus in the United States.* New England Journal of Medicine, 2020.

63. Stokes, E.K., et al., *Coronavirus disease 2019 case surveillance—United States, January 22–May 30, 2020.* Morbidity and Mortality Weekly Report, 2020. **69**(24): p. 759.

64. Porcher, S., *Response2covid19, a dataset of governments' responses to COVID-19 all around the world.* Scientific Data, 2020. **7**(1): p. 1-9.

65. Hale, T., et al., *Variation in government responses to COVID-19.* Blavatnik school of government working paper, 2020. **31**: p. 2020-11.

66. COVID, T.I., et al., *Modeling COVID-19 scenarios for the United States.* Nature medicine, 2020.

67. Hsiang, S., et al., *The effect of large-scale anti-contagion policies on the COVID-19 pandemic.* Nature, 2020. **584**(7820): p. 262-267.

68. Korolev, I., *Challenges with Estimating the Effect of Anti-Contagion Policies on the COVID-19 Pandemic.* Available at SSRN 3659623, 2020.

69. Hennekens, C.H., et al., *The emerging pandemic of coronavirus and the urgent need for public health leadership.* The American journal of medicine, 2020.

70. Fauci, A.S., H.C. Lane, and R.R. Redfield, *Covid-19—navigating the uncharted*. 2020, Mass Medical Soc.

71. Wang, S. and D. Ramkrishna, *A model to rate strategies for managing disease due to COVID-19 infection*. Scientific reports, 2020. **10**(1): p. 1-10.

72. Baker, S.R., et al., *Covid-induced economic uncertainty*. 2020, National Bureau of Economic Research.

73. Demirguc-Kunt, A., M. Lokshin, and I. Torre, *The sooner, the better: The early economic impact of non-pharmaceutical interventions during the COVID-19 pandemic*. World Bank Policy Research Working Paper, 2020(9257).

74. Askitas, N., K. Tatsiramos, and B. Verheyden, *Estimating worldwide effects of non-pharmaceutical interventions on COVID-19 incidence and population mobility patterns using a multiple-event study*. Scientific Reports, 2021. **11**(1): p. 1-13.

75. Aravindakshan, A., et al., *Preparing for a future COVID-19 wave: insights and limitations from a data-driven evaluation of non-pharmaceutical interventions in Germany*. Scientific reports, 2020. **10**(1): p. 1-14.

76. Liao, Z., et al., *TW-SIR: time-window based SIR for COVID-19 forecasts*. Scientific reports, 2020. **10**(1): p. 1-15.

77. Bizzarri, M., et al., *New statistical RI index allow to better track the dynamics of COVID-19 outbreak in Italy*. Scientific Reports, 2020. **10**(1): p. 1-13.

78. Rădulescu, A., C. Williams, and K. Cavanagh, *Management strategies in a SEIR-type model of COVID 19 community spread*. Scientific reports, 2020. **10**(1): p. 1-16.

79. Liu, M., R. Thomadsen, and S. Yao, *Forecasting the Spread of COVID-19 under Different Reopening Strategies*. medRxiv, 2020.

80. Bhanot, G. and C. DeLisi, *Predictions for Europe for the Covid-19 pandemic from a SIR model*. medRxiv, 2020.

81. Peng, L., et al., *Epidemic analysis of COVID-19 in China by dynamical modeling*. arXiv preprint arXiv:2002.06563, 2020.

82. Mangoni, L. and M. Pistilli, *Epidemic analysis of Covid-19 in Italy by dynamical modelling*. Available at SSRN 3567770, 2020.

83. Barbarossa, M.V., et al., *Modeling the spread of COVID-19 in Germany: Early assessment and possible scenarios*. Plos one, 2020. **15**(9): p. e0238559.

84. Zahiri, A., S. RafieeNasab, and E. Roohi, *Prediction of peak and termination of novel coronavirus covid-19 epidemic in Iran*. MedRxiv, 2020.

85. Cooper, I., A. Mondal, and C.G. Antonopoulos, *A SIR model assumption for the spread of COVID-19 in different communities.* Chaos, Solitons & Fractals, 2020. **139**: p. 110057.

86. Wang, J., *Mathematical models for COVID-19: Applications, limitations, and potentials.* Journal of public health and emergency, 2020. **4**.

87. Chaudhry, R., et al., *A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes.* EClinicalMedicine, 2020. **25**: p. 100464.

88. Dandekar, R. and G. Barbastathis, *Quantifying the effect of quarantine control in Covid-19 infectious spread using machine learning.* medRxiv, 2020.

89. Karaivanov, A., et al., *Face masks, public policies and slowing the spread of COVID-19: evidence from Canada*. 2020, National Bureau of Economic Research.

90. VoPham, T., et al., *Effect of social distancing on COVID-19 incidence and mortality in the US.* MedRxiv, 2020.

91. Chinazzi, M., et al., *The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak.* Science, 2020. **368**(6489): p. 395-400.

92. Fowler, J.H., et al., *The effect of stay-athome orders on COVID-19 cases and fatalities in the United States.* medRxiv, 2020.

93. Giordano, G., et al., *Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy.* Nature medicine, 2020. **26**(6): p. 855-860.

94. Zhang, Y., et al., *The impact of social distancing and epicenter lockdown on the COVID-19 epidemic in mainland China: A data-driven SEIQR model study.* MedRxiv, 2020.

95. Cacciapaglia, G. and F. Sannino, *Interplay of social distancing and border restrictions for pandemics (COVID-19) via the epidemic Renormalisation Group framework.* arXiv preprint arXiv:2005.04956, 2020.

96. *STATESIDE: Introducing State Snapshot: A COVID-19 Report*. 2020 [cited 2020 Jul. 31]; Available from: https://www.stateside.com/blog/introducing-state-snapshot-covid-19-report.

97. *COVID-19 Policy Dashboard*. 2020 [cited 2020 Jul. 31]; Available from: https://www.multistate.us/research/covid/public/.

98. *Kaiser Family Foundation: State COVID-19 Data and Policy Actions*. 2020 [cited 2020 Jul. 31]; Available from: https://www.kff.org/coronavirus-covid-19/issue-brief/state-covid-19-data-and-policy-actions/.

99. *Centers for Disease Control and Prevention*. 2020 [cited 2020 Jul. 31]; Available from: https://www.cdc.gov/.

100. *Johns Hopkins Coronavirus Resource Center*. 2020 [cited 2020 Jul. 31]; Available from: https://coronavirus.jhu.edu/.

101. *The New York Times Coronavirus (Covid-19) Data in the United States*. 2020 [cited 2020 Jul. 31]; Available from: https://github.com/nytimes/covid-19-data/.

102. *Worldometer - real time world statistics*. 2020 [cited 2020 Jul. 31]; Available from: https://www.worldometers.info/coronavirus/.

103. Tolles, J. and T. Luong, *Modeling Epidemics With Compartmental Models.* Jama, 2020.

104. Anand, N., et al., *Predicting the Spread of COVID-19 Using $$ SIR $$ SIR Model Augmented to Incorporate Quarantine and Testing.* Transactions of the Indian National Academy of Engineering, 2020. **5**(2): p. 141-148.

105. Chen, Y.-C., et al., *A Time-dependent SIR model for COVID-19 with undetectable infected persons.* IEEE Transactions on Network Science and Engineering, 2020.

106. Jia, W., et al., *Extended SIR prediction of the epidemics trend of COVID-19 in Italy and compared with Hunan, China.* 2020.

107. Das, S., *Prediction of covid-19 disease progression in india: Under the effect of national lockdown.* arXiv preprint arXiv:2004.03147, 2020.

108. Nguemdjo, U., et al., *Simulating the progression of the COVID-19 disease in Cameroon using SIR models.* PloS one, 2020. **15**(8): p. e0237832.

109. Kaddar, A., A. Abta, and H.T. Alaoui, *A comparison of delayed SIR and SEIR epidemic models.* Nonlinear Analysis: Modelling and Control, 2011. **16**(2): p. 181-190.

110. Biswas, M.H.A., L.T. Paiva, and M. De Pinho, *A SEIR model for control of infectious diseases with constraints.* Mathematical Biosciences & Engineering, 2014. **11**(4): p. 761.

111. Sun, C. and Y.-H. Hsieh, *Global analysis of an SEIR model with varying population size and vaccination.* Applied Mathematical Modelling, 2010. **34**(10): p. 2685-2697.

112. Li, M.Y. and J.S. Muldowney, *Global stability for the SEIR model in epidemiology.* Mathematical biosciences, 1995. **125**(2): p. 155-164.

113. *The U.S. Census Bureau* 2020 [cited 2020 Jul. 31]; Available from: https://www.census.gov/data.html.

114. Carcione, J.M., et al., *A simulation of a COVID-19 epidemic based on a deterministic SEIR model.* arXiv preprint arXiv:2004.03575, 2020.

115. Wilson, N., et al., *Case-fatality risk estimates for COVID-19 calculated by using a lag time for fatality.* Emerging infectious diseases, 2020. **26**(6): p. 1339.

116. Dietz, K., *The estimation of the basic reproduction number for infectious diseases.* Statistical methods in medical research, 1993. **2**(1): p. 23-41.

117. Daley, D.J. and J. Gani, *Epidemic modelling: an introduction*. Vol. 15. 2001: Cambridge University Press.

118. Nishiura, H. and G. Chowell, *The effective reproduction number as a prelude to statistical estimation of time-dependent epidemic trends*, in *Mathematical and statistical estimation approaches in epidemiology*. 2009, Springer. p. 103-121.

119. Grimm, V., et al., *Estimating the time-dependent contact rate of SIR and SEIR models in mathematical epidemiology using physics-informed neural networks.* 2020.

120. Lauer, S.A., et al., *The Incubation Period of Coronavirus Disease.* 2020.

121. Lauer, S.A., et al., *The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application.* Annals of internal medicine, 2020. **172**(9): p. 577-582.

122. Lei, S., et al., *Clinical characteristics and outcomes of patients undergoing surgeries during the incubation period of COVID-19 infection.* EClinicalMedicine, 2020: p. 100331.

123. McAloon, C., et al., *Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research.* BMJ open, 2020. **10**(8): p. e039652.

124. Organization, W.H., *Transmission of SARS-CoV-2: implications for infection prevention precautions: scientific brief, 09 July 2020*. 2020, World Health Organization.

125. Giordano, G., et al., *Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy.* Nature Medicine, 2020: p. 1-6.

126. Kim, G.-u., et al., *Clinical characteristics of asymptomatic and symptomatic patients with mild COVID-19.* Clinical Microbiology and Infection, 2020.

127. Wu, Z. and J.M. McGoogan, *Asymptomatic and pre-symptomatic COVID-19 in China*. 2020, BioMed Central.

128. Yang, R., X. Gui, and Y. Xiong, *Comparison of clinical characteristics of patients with asymptomatic vs symptomatic coronavirus disease 2019 in Wuhan, China.* JAMA Network Open, 2020. **3**(5): p. e2010182-e2010182.

129. *Centers for Disease Control and Prevention: COVID-19 Science Update released*. 2020 [cited 2020 Nov. 17]; Available from: https://www.cdc.gov/library/covid19/111720_covidupdate.html/.

130. Buitrago-Garcia, D., et al., *Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis.* PLoS medicine, 2020. **17**(9): p. e1003346.

131. Lawless, J.F., *Negative binomial and mixed Poisson regression.* The Canadian Journal of Statistics/La Revue Canadienne de Statistique, 1987: p. 209-225.

132. Consul, P. and F. Famoye, *Generalized Poisson regression model.* Communications in Statistics-Theory and Methods, 1992. **21**(1): p. 89-109.

133. Ioannidis, J., *The infection fatality rate of COVID-19 inferred from seroprevalence data.* medRxiv, 2020.

134. Russell, T.W., et al., *Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020.* Eurosurveillance, 2020. **25**(12): p. 2000256.

135. Brazeau, N., et al., *Report 34: COVID-19 infection fatality ratio: estimates from seroprevalence.* 2020.

136. Hu, H., K. Nigmatulina, and P. Eckhoff, *The scaling of contact rates with population density for the infectious disease models.* Mathematical biosciences, 2013. **244**(2): p. 125-134.

137. Basu, A., *Estimating The Infection Fatality Rate Among Symptomatic COVID-19 Cases In The United States: Study estimates the COVID-19 infection fatality rate at the US county level.* Health Affairs, 2020: p. 10.1377/hlthaff. 2020.00455.

138. Bahl, P., et al., *Airborne or droplet precautions for health workers treating COVID-19?* The Journal of infectious diseases, 2020.

139. Van Doremalen, N., et al., *Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1.* New England Journal of Medicine, 2020. **382**(16): p. 1564-1567.

140. Lyu, W. and G.L. Wehby, *Community Use Of Face Masks And COVID-19: Evidence From A Natural Experiment Of State Mandates In The US: Study examines impact on COVID-*

19 growth rates associated with state government mandates requiring face mask use in public. Health affairs, 2020. **39**(8): p. 1419-1425.

141. Rubin, D., et al., *Association of social distancing, population density, and temperature with the instantaneous reproduction number of SARS-CoV-2 in counties across the United States.* JAMA network open, 2020. **3**(7): p. e2016099-e2016099.

142. *Forbes; The Frustrating Coronavirus School Opening Debate*. 2020 [cited 2020 Dec. 7]; Available from: https://www.forbes.com/sites/mikemcshane/2020/12/07/the-frustrating-coronavirus-school-opening-debate/?sh=5b9ecfe42123.

143. Esposito, S., N. Cotugno, and N. Principi, *Comprehensive and safe school strategy during COVID-19 pandemic.* Italian journal of pediatrics, 2021. **47**(1): p. 1-4.

144. Viner, R.M., et al., *Susceptibility to SARS-CoV-2 infection among children and adolescents compared with adults: a systematic review and meta-analysis.* JAMA pediatrics, 2020. **175**: p. 143-156.

145. Gudbjartsson, D.F., et al., *Spread of SARS-CoV-2 in the Icelandic population.* New England Journal of Medicine, 2020. **382**(24): p. 2302-2315.

146. Di Domenico, L., et al., *Expected impact of school closure and telework to mitigate COVID-19 epidemic in France.* EPIcx Lab, 2020. **15**.

147. *The Washington Post: Europe's schools still open, still relatively safe, through covid-19 second wave*. 2020 [cited 2020 Dec. 1]; Available from: https://www.washingtonpost.com/world/europe/europe-schools-covid-open/2020/12/01/4480a5c8-2e61-11eb-9dd6-2d0179981719_story.html/.

148. Esposito, S. and N. Principi, *School closure during the coronavirus disease 2019 (COVID-19) pandemic: an effective intervention at the global level?* JAMA pediatrics, 2020. **174**(10): p. 921-922.

149. Poletti, M. and A. Raballo, *COVID-19 and effects of school closure for children and their families: a deafening silence.* JAMA Pediatr. Forthcoming, 2020. **175**: p. 210-212.

150. Van Lancker, W. and Z. Parolin, *COVID-19, school closures, and child poverty: a social crisis in the making.* The Lancet Public Health, 2020. **5**(5): p. e243-e244.

151. Armitage, R. and L.B. Nellums, *Considering inequalities in the school closure response to COVID-19.* The Lancet Global Health, 2020. **8**(5): p. e644.

152. Mangtani, P., *Impact of school closures on an influenza pandemic: scientific evidence base review*. 2014.

153. Thu, T.P.B., P.N.H. Ngoc, and N.M. Hai, *Effect of the social distancing measures on the spread of COVID-19 in 10 highly infected countries.* Science of The Total Environment, 2020. **742**: p. 140430.

154. Medline, A., et al., *Evaluating the impact of stay-at-home orders on the time to reach the peak burden of Covid-19 cases and deaths: does timing matter?* BMC public health, 2020. **20**(1): p. 1-7.

155. Kissler, S.M., et al., *Social distancing strategies for curbing the COVID-19 epidemic.* medRxiv, 2020.

156. Dunford, R., Q. Su, and E. Tamang, *The pareto principle.* 2014.

157. *Centers for Disease Control and Prevention: Leading Causes of Death*. Available from: https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm.

158. *Cancer Facts & Figures 2020*. Available from: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2020/cancer-facts-and-figures-2020.pdf.

159. *American Cancer Society*. 2021 [cited 2021 April, 20]; Available from: https://www.cancer.org/.

160. Malhotra, J., et al., *Risk factors for lung cancer worldwide.* European Respiratory Journal, 2016. **48**(3): p. 889-902.

161. Brenner, D.R., et al., *Lung cancer risk in never-smokers: a population-based case-control study of epidemiologic risk factors.* BMC cancer, 2010. **10**(1): p. 1-9.

162. Ridge, C.A., A.M. McErlean, and M.S. Ginsberg. *Epidemiology of lung cancer*. in *Seminars in interventional radiology*. 2013. Thieme Medical Publishers.

163. Couraud, S., et al., *Lung cancer in never smokers–a review.* European journal of cancer, 2012. **48**(9): p. 1299-1311.

164. Weiderpass, E., *Lifestyle and cancer risk.* Journal of preventive medicine and public health, 2010. **43**(6): p. 459-471.

165. Dresler, C., *The changing epidemic of lung cancer and occupational and environmental risk factors.* Thoracic surgery clinics, 2013. **23**(2): p. 113-122.

166. Samet, J.M., et al., *Lung cancer in never smokers: clinical epidemiology and environmental risk factors.* Clinical Cancer Research, 2009. **15**(18): p. 5626-5645.

167. Wynder, E.L. and D. Hoffmann, *Smoking and lung cancer: scientific challenges and opportunities.* Cancer Research, 1994. **54**(20): p. 5284-5294.

168. Gawełek, E., B. Drozdzowska, and A. Fuchs, *Radon as a risk factor of lung cancer.* Przeglad epidemiologiczny, 2017. **71**(1): p. 90-98.

169. Kanwal, M., X.J. Ding, and Y. Cao, *Familial risk for lung cancer.* Oncology letters, 2017. **13**(2): p. 535-542.

170. Gasperino, J., *Gender is a risk factor for lung cancer.* Medical hypotheses, 2011. **76**(3): p. 328-331.

171. VENA, J.E., *Air pollution as a risk factor in lung cancer.* American Journal of Epidemiology, 1982. **116**(1): p. 42-56.

172. Kabat, G.C., A.B. Miller, and T.E. Rohan, *Body mass index and lung cancer risk in women.* Epidemiology, 2007: p. 607-612.

173. Key, T.J., et al., *The effect of diet on risk of cancer.* The Lancet, 2002. **360**(9336): p. 861-868.

174. Behera, D. and T. Balamugesh, *Indoor air pollution as a risk factor for lung cancer in women.* JAPI, 2005. **53**: p. 190-192.

175. Taylor, R., F. Najafi, and A. Dobson, *Meta-analysis of studies of passive smoking and lung cancer: effects of study type and continent.* International journal of epidemiology, 2007. **36**(5): p. 1048-1059.

176. Gustavsson, P., et al., *Occupational exposure and lung cancer risk: a population-based case-referent study in Sweden.* American journal of epidemiology, 2000. **152**(1): p. 32-40.

177. Świątkowska, B., *Modifiable risk factors for the prevention of lung cancer.* Reports of Practical Oncology & Radiotherapy, 2007. **12**(2): p. 119-124.

178. *Cancer Care Ontario*. 2021 [cited 2021 May 5]; Available from: https://www.cancercareontario.ca/en/data-research/view-data/modifiable-risk-factors#:~:text=Modifiable%20risk%20factors%20are%20behaviours,e.g.%2C%20diabetes%2C%20cardiovascular%20disease)%3A.

179. Fajersztajn, L., et al., *Air pollution: a potentially modifiable risk factor for lung cancer.* Nature Reviews Cancer, 2013. **13**(9): p. 674-678.

180. Wu, S., et al., *Evaluating intrinsic and non-intrinsic cancer risk factors.* Nature communications, 2018. **9**(1): p. 1-12.

181. Koop, C.E., *Reducing the Health Consequences of Smoking 25 Years of Progress: Report of the Surgeon General*. 1995: DIANE Publishing.

182. Shopland, D.R., *Tobacco use and its contribution to early cancer mortality with a special emphasis on cigarette smoking.* Environmental Health Perspectives, 1995. **103**(suppl 8): p. 131-142.

183. Clement, C.H., et al., *Lung cancer risk from radon and progeny and statement on radon.* Annals of the ICRP, 2010. **40**(1): p. 1-64.

184. Metayer, C., et al., *Cooking oil fumes and risk of lung cancer in women in rural Gansu, China.* Lung Cancer, 2002. **35**(2): p. 111-117.

185. Seow, A., et al., *Fumes from meat cooking and lung cancer risk in Chinese women.* Cancer Epidemiology and Prevention Biomarkers, 2000. **9**(11): p. 1215-1221.

186. Gao, Y.-t., *Risk factors for lung cancer among nonsmokers with emphasis on lifestyle factors.* Lung cancer, 1996. **14**: p. S39-S45.

187. Alberg, A.J. and J.M. Samet, *Epidemiology of lung cancer.* Chest, 2003. **123**(1): p. 21S-49S.

188. Sun, S., J.H. Schiller, and A.F. Gazdar, *Lung cancer in never smokers—a different disease.* Nature Reviews Cancer, 2007. **7**(10): p. 778-790.

189. Cersosimo, R.J., *Lung cancer: a review.* American journal of health-system pharmacy, 2002. **59**(7): p. 611-642.

190. Bagnardi, V., et al., *Alcohol consumption and lung cancer risk in the Environment and Genetics in Lung Cancer Etiology (EAGLE) study.* American journal of epidemiology, 2010. **171**(1): p. 36-44.

191. Wakai, K., et al., *Alcohol drinking and lung cancer risk: an evaluation based on a systematic review of epidemiologic evidence among the Japanese population.* Japanese journal of clinical oncology, 2007. **37**(3): p. 168-174.

192. Bandera, E.V., J.L. Freudenheim, and J.E. Vena, *Alcohol consumption and lung cancer: a review of the epidemiologic evidence.* Cancer Epidemiology and Prevention Biomarkers, 2001. **10**(8): p. 813-821.

193. Turner, M.C., et al., *Radon and lung cancer in the American Cancer Society cohort.* Cancer Epidemiology and Prevention Biomarkers, 2011. **20**(3): p. 438-448.

194. Yun, Y.H., et al., *Relative and absolute risks of cigarette smoking on major histologic types of lung cancer in Korean men.* Cancer Epidemiology and Prevention Biomarkers, 2005. **14**(9): p. 2125-2130.

195. Nikić, D. and A.M. Stanković, *Air pollution as a risk factor for lung cancer.* Archive of Oncology, 2005. **13**(2): p. 79-82.

196. Brenner, D.R., et al., *Alcohol consumption and lung cancer risk: A pooled analysis from the International Lung Cancer Consortium and the SYNERGY study.* Cancer epidemiology, 2019. **58**: p. 25-32.

197. Kreienbrock, L., et al., *Case-control study on lung cancer and residential radon in western Germany.* American Journal of Epidemiology, 2001. **153**(1): p. 42-52.

198. Pesch, B., et al., *Cigarette smoking and lung cancer—relative risk estimates for the major histological types from a pooled analysis of case–control studies.* International journal of cancer, 2012. **131**(5): p. 1210-1219.

199. Lee, T. and F. Gany, *Cooking oil fumes and lung cancer: a review of the literature in the context of the US population.* Journal of immigrant and minority Health, 2013. **15**(3): p. 646-652.

200. Stockwell, H.G., et al., *Enviromental tobacco smoke and lung cancer risk in nonsmoking women.* JNCI: Journal of the National Cancer Institute, 1992. **84**(18): p. 1417-1422.

201. Miller, A.B., et al., *Fruits and vegetables and lung cancer: Findings from the European Prospective Investigation into Cancer and Nutrition.* International journal of cancer, 2004. **108**(2): p. 269-276.

202. Lantz, P.M., D. Mendez, and M.A. Philbert, *Radon, smoking, and lung cancer: the need to refocus radon control policy.* American journal of public health, 2013. **103**(3): p. 443-447.

203. Ahrendt, S.A., et al., *Alcohol consumption and cigarette smoking increase the frequency of p53 mutations in non-small cell lung cancer.* Cancer Research, 2000. **60**(12): p. 3155-3159.

204. Osann, K.E., *Lung cancer in women: the importance of smoking, family history of cancer, and medical history of respiratory disease.* Cancer research, 1991. **51**(18): p. 4893-4897.

205. Wang, S.-y., et al., *A comparative study of the risk factors for lung cancer in Guangdong, China.* Lung Cancer, 1996. **14**: p. S99-S105.

206. Spitz, M.R., et al., *A risk model for prediction of lung cancer.* Journal of the National Cancer Institute, 2007. **99**(9): p. 715-726.

207. Kelley, K. and K.J. Preacher, *On effect size.* Psychological methods, 2012. **17**(2): p. 137.

208. Saaty, T.L., *Decision-making with the AHP: Why is the principal eigenvector necessary.* European journal of operational research, 2003. **145**(1): p. 85-91.

209. Liberati, A., et al., *The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration.* Journal of clinical epidemiology, 2009. **62**(10): p. e1-e34.

210. *JBI's Critical Appraisal Tools*. 2021 [cited 2021 1 March]; Available from: https://jbi.global/critical-appraisal-tools.

211. Page, M.J., et al., *PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews.* bmj, 2021. **372**.

212. Kleinbaum, D., L. Kupper, and H. Morgenstern, *Principles and quantitative methods. Epidemiologic research.* New York: Van Nostrand Reinhold Company, 1982. **427**: p. 450.

213. Egger, M., et al., *Bias in meta-analysis detected by a simple, graphical test.* Bmj, 1997. **315**(7109): p. 629-634.

214. Begg, C.B. and M. Mazumdar, *Operating characteristics of a rank correlation test for publication bias.* Biometrics, 1994: p. 1088-1101.

215. DerSimonian, R. and N. Laird, *Meta-analysis in clinical trials.* Controlled clinical trials, 1986. **7**(3): p. 177-188.

216. Cochran, W.G., *The combination of estimates from different experiments.* Biometrics, 1954. **10**(1): p. 101-129.

217. Higgins, J.P., et al., *Measuring inconsistency in meta-analyses.* Bmj, 2003. **327**(7414): p. 557-560.

218. Remen, T., et al., *Risk of lung cancer in relation to various metrics of smoking history: a case-control study in Montreal.* BMC cancer, 2018. **18**(1): p. 1-12.

219. Stellman, S.D., et al., *Smoking and lung cancer risk in American and Japanese men: an international case-control study.* Cancer Epidemiology and Prevention Biomarkers, 2001. **10**(11): p. 1193-1199.

220. Risch, H.A., et al., *Are female smokers at higher risk for lung cancer than male smokers? A case-control analysis by histologic type.* American journal of epidemiology, 1993. **138**(5): p. 281-293.

221. Katsouyanni, K., et al., *A case-control study of air pollution and tobacco smoking in lung cancer among women in Athens.* Preventive Medicine, 1991. **20**(2): p. 271-278.

222. Wakai, K., et al., *Tobacco smoking and lung cancer risk: an evaluation based on a systematic review of epidemiological evidence among the Japanese population.* Japanese journal of clinical oncology, 2006. **36**(5): p. 309-324.

223. Boffetta, P., et al., *Cigar and pipe smoking and lung cancer risk: a multicenter study from Europe.* Journal of the National Cancer Institute, 1999. **91**(8): p. 697-701.

224. Armadans-Gil, L., et al., *Cigarette smoking and male lung cancer risk with special regard to type of tobacco.* International journal of epidemiology, 1999. **28**(4): p. 614-619.

225. Luqman, M., et al., *Risk factors for lung cancer in the Pakistani population.* Asian Pacific Journal of Cancer Prevention, 2014. **15**(7): p. 3035-3039.

226. Brownson, R.C., J.C. Chang, and J.R. Davis, *Gender and histologic type variations in smoking-related risk of lung cancer.* Epidemiology, 1992: p. 61-64.

227. Agudo, A., et al., *Lung cancer and cigarette smoking in women: A multicenter case-control study in Europe.* International journal of cancer, 2000. **88**(5): p. 820-827.

228. Krewski, D., et al., *A combined analysis of North American case-control studies of residential radon and lung cancer.* Journal of Toxicology and Environmental Health, Part A, 2006. **69**(7-8): p. 533-597.

229. H. Lubin, J., *Studies of radon and lung cancer in North America and China.* Radiation protection dosimetry, 2003. **104**(4): p. 315-319.

230. Lubin, J.H., et al., *Risk of lung cancer and residential radon in China: pooled results of two studies.* International Journal of Cancer, 2004. **109**(1): p. 132-137.

231. Lorenzo-González, M., et al., *Lung cancer and residential radon in never-smokers: A pooling study in the Northwest of Spain.* Environmental research, 2019. **172**: p. 713-718.

232. Lagarde, F., et al., *Residential radon and lung cancer among never-smokers in Sweden.* Epidemiology, 2001: p. 396-404.

233. Cassidy, A., et al., *Occupational exposure to crystalline silica and risk of lung cancer: a multicenter case-control study in Europe.* Epidemiology, 2007: p. 36-43.

234. Brüske-Hohlfeld, I., et al., *Occupational lung cancer risk for men in Germany: results from a pooled case-control study.* American journal of epidemiology, 2000. **151**(4): p. 384-395.

235. Pintos, J., et al., *Occupational exposure to asbestos and man-made vitreous fibers, and risk of lung cancer: evidence from two case-control studies in Montreal, Canada.* Journal of occupational and environmental medicine, 2008. **50**(11): p. 1273-1281.

236. Villeneuve, P.J., et al., *Occupational exposure to asbestos and lung cancer in men: evidence from a population-based case-control study in eight Canadian provinces.* BMC cancer, 2012. **12**(1): p. 1-10.

237. Stayner, L., et al., *Lung cancer risk and workplace exposure to environmental tobacco smoke.* American Journal of Public Health, 2007. **97**(3): p. 545-551.

238. Hackshaw, A.K., M.R. Law, and N.J. Wald, *The accumulated evidence on lung cancer and environmental tobacco smoke.* Bmj, 1997. **315**(7114): p. 980-988.

239. Seki, T., et al., *Cigarette smoking and lung cancer risk according to histologic type in Japanese men and women.* Cancer science, 2013. **104**(11): p. 1515-1522.

240. Kabat, G.C., S.D. Stellman, and E.L. Wynder, *Relation between exposure to environmental tobacco smoke and lung cancer in lifetime nonsmokers.* American journal of epidemiology, 1995. **142**(2): p. 141-148.

241. Nyberg, F., et al., *Environmental tobacco smoke and lung cancer in nonsmokers: does time since exposure play a role?* Epidemiology, 1998: p. 301-308.

242. Sobue, T., *Association of indoor air pollution and lifestyle with lung cancer in Osaka, Japan.* International Journal of Epidemiology, 1990. **19**(Supplement_1): p. S62-S66.

243. Jee, S.H., H. Ohrr, and I.S. Kim, *Effects of husbands' smoking on the incidence of lung cancer in Korean women.* International Journal of Epidemiology, 1999. **28**(5): p. 824-828.

244. Brownson, R.C., et al., *Passive smoking and lung cancer in nonsmoking women.* American Journal of Public Health, 1992. **82**(11): p. 1525-1530.

245. Ko, Y.-C., et al., *Risk factors for primary lung cancer among non-smoking women in Taiwan.* International journal of epidemiology, 1997. **26**(1): p. 24-31.

246. Ribeiro, A.G., et al., *Incidence and mortality risk for respiratory tract cancer in the city of São Paulo, Brazil: Bayesian analysis of the association with traffic density.* Cancer epidemiology, 2018. **56**: p. 53-59.

247. Vineis, P., et al., *Air pollution and risk of lung cancer in a prospective study in Europe.* International journal of cancer, 2006. **119**(1): p. 169-174.

248. Hystad, P., et al., *Long-term residential exposure to air pollution and lung cancer risk.* Epidemiology, 2013: p. 762-772.

249. Vineis, P., et al., *Lung cancers attributable to environmental tobacco smoke and air pollution in non-smokers in different European countries: a prospective study.* Environmental Health, 2007. **6**(1): p. 1-7.

250. Pope Iii, C.A., et al., *Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution.* Jama, 2002. **287**(9): p. 1132-1141.

251. Nafstad, P., et al., *Lung cancer and air pollution: a 27 year follow up of 16 209 Norwegian men.* Thorax, 2003. **58**(12): p. 1071-1076.

252. Abbey, D.E., et al., *Long-term inhalable particles and other air pollutants related to mortality in nonsmokers.* American journal of respiratory and critical care medicine, 1999. **159**(2): p. 373-382.

253. Nyberg, F., et al., *Urban air pollution and lung cancer in Stockholm.* Epidemiology, 2000: p. 487-495.

254. Raspanti, G.A., et al., *Household air pollution and lung cancer risk among never-smokers in Nepal.* Environmental research, 2016. **147**: p. 141-145.

255. Vermeulen, R., et al., *Constituents of household air pollution and risk of lung cancer among never-smoking women in Xuanwei and Fuyuan, China.* Environmental health perspectives, 2019. **127**(9): p. 097001.

256. Bruce, N., et al., *Does household use of biomass fuel cause lung cancer? A systematic review and evaluation of the evidence for the GBD 2010 study.* Thorax, 2015. **70**(5): p. 433-441.

257. Kleinerman, R., et al., *Lung cancer and indoor air pollution in rural China.* Annals of epidemiology, 2000. **10**(7): p. 469.

258. Ko, Y.-C., et al., *Chinese food cooking and lung cancer in women nonsmokers.* American journal of epidemiology, 2000. **151**(2): p. 140-147.

259. Bandera, E.V., et al., *Alcohol consumption and lung cancer in white males.* Cancer Causes & Control, 1992. **3**(4): p. 361-369.

260. Ruano-Ravina, A., A. Figueiras, and J. Barros-Dios, *Type of wine and risk of lung cancer: a case-control study in Spain.* Thorax, 2004. **59**(11): p. 981-985.

261. Boffetta, P., et al., *Risk of cancers of the lung, head and neck in patients hospitalized for alcoholism in Sweden.* British journal of cancer, 2001. **85**(5): p. 678-682.

262. Carpenter, C.L., H. Morgenstern, and S.J. London, *Alcoholic beverage consumption and lung cancer risk among residents of Los Angeles County.* The Journal of nutrition, 1998. **128**(4): p. 694-700.

263. Hu, J., et al., *Risk factors for lung cancer among Canadian women who have never smoked.* Cancer detection and prevention, 2002. **26**(2): p. 129-138.

264. Benedetti, A., M.-E. Parent, and J. Siemiatycki, *Consumption of alcoholic beverages and risk of lung cancer: results from two case–control studies in Montreal, Canada.* Cancer Causes & Control, 2006. **17**(4): p. 469-480.

265. Borenstein, M., et al., *Introduction to meta-analysis*. 2021: John Wiley & Sons.

266. Hunter, J.E. and F.L. Schmidt, *Methods of meta-analysis: Correcting error and bias in research findings*. 2004: Sage.

267. Mu, E. and M. Pereyra-Rojas, *Understanding the analytic hierarchy process*, in *Practical decision making*. 2017, Springer. p. 7-22.

268. Schmidt, K., et al., *Applying the Analytic Hierarchy Process in healthcare research: A systematic literature review and evaluation of reporting.* BMC medical informatics and decision making, 2015. **15**(1): p. 1-27.

269. Matakidou, A., T. Eisen, and R. Houlston, *Systematic review of the relationship between family history and lung cancer risk.* British journal of cancer, 2005. **93**(7): p. 825-833.

270. Saaty, T.L., *Decision making with the analytic hierarchy process.* International journal of services sciences, 2008. **1**(1): p. 83-98.

271. Saaty, T.L., *Group decision making and the AHP*, in *The analytic hierarchy process*. 1989, Springer. p. 59-67.