Research Design and Methodology III

Derrick Snow

Omega Graduate School

February 17, 2025

Professor

Dr. Taladay

Assignment

**Assignment #2 – Developmental Readings**

1. Create Developmental Readings from seminal sources and scholarly peer-reviewed

journal articles. Review instructions for Assignment #3, the course essential elements, and course readings to identify selections of books and journals to create entries.

1. Refer to the "Student Guide to Developmental Readings" in the General Helps folder for updated information on sample comments, the grading rubric, and key definitions related to developmental readings.

**Source One:** Embury, J., Tsou, M.-H., Nara, A., & Oren, E. (2022). A Spatio-Demographic Perspective on the Role of Social Determinants of Health and Chronic Disease in Determining a Population’s Vulnerability to COVID-19. *Preventing Chronic Disease*, *19*, 210414. <https://doi.org/10.5888/pcd19.210414>

**Comment 1:**

**Quote/Paraphrase:** “We obtained data sets from the San Diego County Open Data Portal (11), aggregated to SRAs, containing 2017 rates for hospitalization, emergency department discharge, and death per 100,000 residents for coronary heart disease (CHD), diabetes, hypertensive diseases (hypertension, hypertensive heart disease, hypertensive chronic kidney disease, and hypertensive encephalopathy), mental illness, and pulmonary disease.” (Embury, 2022, p. 2)

**Essential Element:** Sampling Techniques

**Additive/Variant Analysis:** The use of rates in measuring illness and disease in healthcare research is additive to my understanding.

**Contextualization:** This article is a good example of the process of collecting archival/secondary data. In a lot of social research data is collected by the researcher, surveys, observations-participation, and so forth collecting primary data. For an individual researcher, they could collect a minimum of data given their resource constraint. However, given this era of technology and society, there is so much data being collected virtually. This provides such a great data to conduct social research, one just needs to understand where the data is coming in and where to go to obtain the data. This data that is collected through technology is very rich for researchers, because it includes a great number of participants and variables on those participants and the data has a great deal of reliability and validity given they are factual accounts of reality, such as a person’s diagnosis.

**Comment 2:**

**Quote/Paraphrase:** “Socioeconomic variables were chosen for further analysis if the Pearson correlation P values were less than or equal to 0.05 for all stages, with 2 exceptions for variables with P values equal to 0.07 during 1 or 2 of the stages. The Pearson correlation coefficient is commonly used in medical research to test the strength of linear relationships between 2 variables.” (Embury, 2022, p. 3)

**Essential Element:** Data Analysis

**Additive/Variant Analysis:** The use of correlational studies is additive to my understanding of healthcare research.

**Contextualization:** In this research, the researchers used statistical analysis to evaluate variables to continue the research. It is a good example of how the research process really works, it is a continuous process of using statistics to understand the data and evaluate the data to make further decisions on the testable research data, such as what to do with outliers, and so forth. This article informed us that the “Pearson correlation coefficient is commonly used in medical research”, where even though there are many types of medical research, such as clinical trials, correlation research is very important to find variables that are associated with medical phenomena such as illness. This is important because medical phenomena is not in isolation, but have influences, such as social and economic, and environmental factors such as poverty, smoking and body mass index. It is really challenging to find direct causation, where there are probably many factors that may cause illness, diseases, genes, and social, economic, and environmental factors. Nevertheless, if we know what is correlated with given illnesses, we could implement changes in those other factors to then, in turn, reduce the prevalence of a given illness.

**Source Two:** Gobaud, A. N., Mehranbod, C. A., Dong, B., Dodington, J., & Morrison, C. N. (2022). Absolute versus relative socioeconomic disadvantage and homicide: A spatial ecological case–control study of US zip codes. *Injury Epidemiology*, *9*(1), 7. <https://doi.org/10.1186/s40621-022-00371-z>

**Comment 3:**

**Quote/Paraphrase:** “The unit of analysis for this ecological case–control

study were 2017 ZIP codes. Though an imperfect proxy for neighborhoods, we chose ZIP codes in 2017 as the unit of analysis because of data availability. ZIP codes eligible for inclusion were in the 34 US states and four counties in California that participated in the National Violent Death Reporting System (NVDRS) during that year (n = 23,949). The Centers for Disease Control and

Prevention (CDC) created the NVDRS in 2002 to collect data on all types of violent deaths—including homicides—in all settings for all age groups (State Profileset al. 2019).(Gobaud, et al., 2022, p.3)

**Essential Element:** Sampling Techniques

**Additive/Variant Analysis:** The use of zip codes, to attach attribute data to a spatial location in order to conduct geohealth research is additive to my understanding.

**Contextualization:** With geohealth research one is interested in understanding health and illness from a geographical perspective. There is always a geographic aspect of health and illness because people are located geographically. With our modern computing tools such as GIS, we can not only map health and illness, but we can also conduct spatial statistical analysis. With the use of Electronic Health Records (EHR), there is a wealth of medical information that could be analyzed, in addition there is demographic information and geographic information associated with patients, therefore we can link health data with geographic data such as zip code. In addition to patients addresses, address location of the health clinic or hospital. Thus, given this connection with health data and geographic data, researchers can conduct insightful geohealth research, which provides an additional context to understanding health and illness of people.

**Source Three:** Habibi, R., Alesheikh, A. A., & Bayat, S. (2022). An event-based model and a map visualization approach for spatiotemporal association relations discovery of diseases diffusion. *Sustainable Cities and Society*, *87*, 104187. <https://doi.org/10.1016/j.scs.2022.104187>

**Comment 4:**

**Quote/Paraphrase:** “In this study, the data on rate of COVID-19 cases per 100,000 people COVID-19 were obtained from the NYC Health Department repository on GitHub (NYC-Health, 2022). This dataset was stratified by week and modified ZIP Code Tabulation Area (ZCTA) for 75 weeks from August 2, 2020, to January 8, 2022.” (Habibir et al, 2022, p. 87)

**Essential Element:** Data Analysis

**Additive/Variant Analysis:** The concept and the utilization of Zip Code Tabulation Area (ZCTA) for enabling geo shape files of small geographic areas is additive to my understanding of spatial data.

**Contextualization:** In healthcare research, rates are often used to yield standardized data given a population base. For example, death rates are often reported as age-adjusted death rates, to account for the age at death. This provides standardized values where researchers could utilize to understand health and illness phenomena in context of phenomenon. With the study of examining the relationship between land use and illness, counts of illness need to be converted into rates to put illness in context of the patient population. The method for standardizing illness data I will use is to use the patient population of the health centers as the denominator and the count of the number of patients within the illness/disease categories as the numerator. This will be additional work, but it is all part of data preparation that is done in archival research. This will give the rate of illness in a given health center area. Then the primary service area of the health center given by the main zip codes of patients of the health center will provide the geolocation data to place the rates of illness on the map. The overlay of these service areas on top of the land use spatial data will provide the context of the research to identify which illnesses are associated with different types of land use.

**Source Four:** Zeng, S., Pelzer, K. M., Gibbons, R. D., Peek, M. E., & Parker, W. F. (2022). Association of Zip Code Vaccination Rate With COVID-19 Mortality in Chicago, Illinois. *JAMA Network Open*, *5*(5), e2214753. <https://doi.org/10.1001/jamanetworkopen.2022.14753>

**Comment 5:**

**Quote/Paraphrase:** “We performed 3 robustness checks for the Poisson regression. First, we changed the exposure to the 6-week lagged percentage of residents fully vaccinated. Next, we used vaccination group as a categorical treatment variable to assess for nonlinear effects of vaccination rate on mortality. Lastly, we used mortality data from the Cook County Medical Examiner Case Archive,19 which provided residential addresses that we geocoded to census tracts. Because vaccination data are collected by the CDPH and only available at the zip code level, we constructed a data set that contained weekly deaths and vaccination coverage within the geographic overlap of a census tract and zip code with population estimated from the Crosswalk Files from the US Department of Housing and Urban Development.20 We performed all analyses with R statistical software version 4.0.5 (R Project for Statistical Computing).” (Zengt et al., 2022, p. 3)

**Essential Element:** Data Analysis

**Additive/Variant Analysis:** The use of various regression modeling, including logistic regression used in geohealth research is additive to my understanding of statistical procedures used in healthcare research.

**Contextualization:** The research of finding correlations between illness and land use, is a bit different from other geohealth correlation studies, so it has been a challenge to find the best statistical method to conduct the research from reviewing literature. Land use is a spatial variable and categorical, with different land use associated with different land use codes (i.e. 11, 12, 21) and these codes are associated with different land use types and definitions. For example,

11 Open Water- areas of open water, generally with less than 25% cover of vegetation or soil.

12 Perennial Ice/Snow- areas characterized by a perennial cover of ice and/or snow, generally greater than 25% of total cover.

21 Developed, Open Space- areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20% of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.

22 Developed, Low Intensity- areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20% to 49% percent of total cover. These areas most commonly include single-family housing units.

Land use is not really ordinal data but could possibly be treated as such for statistical analysis purposes. But I think I will treat it as nominal and utilize a form of logistic regression with the illness data as interval-ratio data as being the rate of a given illness or perhaps conducting Poisson regression, as they did in the above article.

**Source Five:** Benavidez, G. A., Zahnd, W. E., Hung, P., & Eberth, J. M. (2024). Chronic Disease Prevalence in the US: Sociodemographic and Geographic Variations by Zip Code Tabulation Area. *Preventing Chronic Disease*, *21*, 230267. <https://doi.org/10.5888/pcd21.230267>

**Comment 6:**

**Quote/Paraphrase:** “Data on ZCTA-level chronic disease prevalence estimates came from the 2020 Centers for Disease Control and Prevention’s (CDC’s) PLACES data set. The PLACES project is designed to generate health data for small geographic units across the country (ie, county, census tract, and ZCTA) to enable informed decisions when planning public health interventions.” (Benavidez et al., 2024, p.2)

**Essential Element:** Literature Review Outline

**Additive/Variant Analysis:** The concept of trying to use the lowest level of geographical data in order to capture the most detail type of information is additive to my knowledge of geohealth research.

**Contextualization:**  In this research, they used data derived health data from health surveillance systems for smaller geographic areas, such as ZTCAs from the CDC. This would also be a good data set to use to evaluate the correlation of land use and illnesses. The CDC PLACES data set has fewer illness groupings than I would like, and not exact data that comes out of EHR systems. However, the CDC PLACES data is publicly available and has less data prep required such would be the case with EHR data. In my test run of a sample set of data, I will do a comparison with PLACES data to the EHR data among the comparable diseases. This article included a good statement about utilizing small-area health data to “enable informed decisions when planning public health interventions”, I too hold this view, it is important to get to smaller levels of analysis to be able to see and understand the health differences of people within the smaller geographic area, so changes could be made to help communities become more healthier.

**Comment 7:**

**Quote/Paraphrase:** “Sociodemographic and economic characteristics have been well studied and have shown consistently to be associated with various chronic diseases (25,26). Although our study corroborates these previous studies, we provide greater insight into the characteristics and the socioeconomic position of communities with a high prevalence of multiple chronic diseases.” Benavidez et al., 2024, p. 5)

**Essential Element:** Literature Review Outline

**Additive/Variant Analysis:** The use of social, demographic, economic and environmental to enhance the understanding of health and illness in research is additive to my understanding.

**Contextualization:** This quote from the research discusses the sociodemographic and socioeconomic factors of health. These are known as social drivers of health, which is a major component of health. Utilizing land use as a spatial feature in my intended research, I will be able to get to these sociodemographic and socioeconomic factors of health, because land use of a given area informs us what the social and economic situation of that given land use. Such as agriculture areas, we know that some people will be agricultural workers and others may work in support of the agriculture area, like truck drivers, agri-business, and so forth. Census data on income and poverty data could also be included in the research to further examine the sociodemographic and socioeconomic factors of health in these given land use areas. I am seeking to do a nationwide study, but I believe I will need to do a stratified analysis by region because though the land use may be the same in different regions, the regional difference may be enough to suggest that the research should examine nationally but by regions of the U.S.

**Source Six:** Lee, E. K., Donley, G., Ciesielski, T. H., Freedman, D. A., & Cole, M. B. (2023). Spatial availability of federally qualified health centers and disparities in health services utilization in medically underserved areas. *Social Science & Medicine*, *328*, 116009. <https://doi.org/10.1016/j.socscimed.2023.116009>

**Comment 8:**

**Quote/Paraphrase:** “For linking zip code-level outcome data and FQHC availability, information on FQHCs’ patient volume and prevalence estimates for

routine care visits in the past twelve months were available at the zip code-level for year 2019 and were linked with the exposure variable available at the zip-code level. We further examined exposed areas based on FQHC availability distributions: Low-availability or 1 FQHC site per zip code, medium-availability or 2–4 FQHC sites per zip code, and high availability or ≥5 FQHC sites per zip code.” (Lee et al, 2023, p. 3)

**Essential Element:** Data Analysis

**Additive/Variant Analysis:** This quote on the use of Federally Qualified Health Centers (FQHCs) data for healthcare research is additive to my understanding.

**Contextualization:** This quote presents what the researchers are doing to link attribute health data to the spatial, with a spatial connector such as zip codes. The Federally Qualified Health Centers (FQHC) are community health centers that serve many underserved patient populations, such as low-income, no insurance, or underinsured, and receive federal grants to serve these populations. Given that FQHCs are community health centers, they also serve any member of the community with their healthcare needs, even if they are high-income and have health insurance. The federal agency Health Resources and Services Administration (HRSA) oversees the nation’s health center network, which serves 1 in 11 Americans; this is a great number of Americans healthcare data available through EHR data that is mandatory and reported annually to HRSA. This is a rich data set to use for several reasons, including that it would get at people who may work close to the land use of that given community and the working class, who may be impacted by the environmental, social, and economic realities of a given land use. Therefore, I believe that utilizing HRSA EHR data is the best data set to use in identifying correlations of illness with land use.

**Source Seven:** Delmelle, E. M., Desjardins, M. R., Jung, P., Owusu, C., Lan, Y., Hohl, A., & Dony, C. (2022). Uncertainty in geospatial health: Challenges and opportunities ahead. *Annals of Epidemiology*, *65*, 15–30. <https://doi.org/10.1016/j.annepidem.2021.10.002>

**Comment 9:**

**Quote/Paraphrase:** “The other disadvantage of online geocoders, perhaps more important but less discussed, is confidentiality. In public health and criminology, confidentiality and privacy are significant considerations. The risk of confidentiality breach increases when the dataset containing information on individuals, such as addresses, is passed on to online service providers [54] . As a result, protected health information (PHI) can be leaked, violating the Health Insurance Portability and Accountability Act in the US (HIPAA –US Depart- ment of Health and Human Services)”. (Delmelle et al, 2022, p. 17).

**Essential Element:** Ethical Research Criteria

**Additive/Variant Analysis:** The importance of patient privacy in healthcare research is additive to my understanding of handling healthcare data in an ethical manner

**Contextualization:** In working with healthcare data, there are regulations on how to handle the patient healthcare information and the protection of patient privacy and identifiable information known as PHI (protected health information) and is governed by the Health Insurance Portability and Accountability Act (HIPAA). Healthcare researchers need to understand these regulations and requirements in using healthcare data in their studies. One of the great features of GIS is the ability to geocode data, which means placing a spatial point feature on a map at an exact street address location. With healthcare data, this practice is not common for patient-level data because of data privacy issues. One could geocode health center locations and other important factors such as hospitals and other care locations. Therefore, in my intended research, the lowest geographic level for healthcare data is at the zip code level, and even that is not commonly accessible through online data but may need to be requested by the researcher from the custodian agency. This is an important factor in my research where land use is categorized in low-level geography, even smaller than at the zip code level. However, in my research, I will need to associate a single land use category to a service area composed of a few zip code areas. It may be possible to create a combination of land use to a service area, such as 50% Cultivated Crops and 50% Developed, Open Space.

**Source Eight:** Swales, L., Gooden, A., & Thaldar, D. (2024). The anatomy of a data transfer agreement for health research. *Frontiers in Pharmacology*, *15*, 1332700. <https://doi.org/10.3389/fphar.2024.1332700>

**Comment 10:**

**Quote/Paraphrase:** “It is a written agreement that facilitates the lawful transfer of data between parties. Typically, an agreement of this type will seek to comply with applicable legislation. Additionally, a DTA will regulate other important legal issues such as ownership of data, intellectual property, the terms of the agreement and how it will terminate, liability, dispute resolution, and whether any consideration is payable.” (Swales, Gooden, & Thaldar, 2024, p. 1)

**Essential Element:** Ethical Research Criteria

**Additive/Variant Analysis:** The utilization of Data Use Agreements and Data Transfer Agreements between researchers and organizations is additive to my knowledge of the process of obtaining archival data.

**Contextualization:** In this article, the author addresses the concepts of researchers' responsibility to handle, use, and protect data they may obtain through agencies and organizations. Known as Data use Agreements or Data Transfer Agreements. These agreements state the intended and restricted use of the data that is being obtained. This is an important duty researchers have to use the data in an ethical and regulatory manner. Many online datasets have acknowledged agreements for the researchers to review and agree to before they can access the data. If requesting data from an agency or organization that is not publicly accessible, the researchers must complete data use agreements as part of requesting the use of the data. Data use agreements may be mostly standard, but depending upon the data may have specific requirements. I will submit a request to HRSA to obtain illness data at the health center level and will need to complete a data use agreement. Then I use the health center illness data and the available health center zip code data to associate illness rates to the zip code group level.

**Source Nine:** National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research.

**Comment 11:**

**Quote/Paraphrase:** “Three principles, or general prescriptive judgments, that are relevant to research involving human subjects are identified in this statement. Other principles may also be relevant. These three are comprehensive, however, and are stated at a level of generalization that should assist scientists, subjects, reviewers and interested citizens to understand the ethical issues inherent in research involving human subjects. These principles cannot always be applied so as to resolve beyond dispute particular ethical problems. The objective is to provide an analytical framework that will guide the resolution of ethical problems arising from research involving human subjects.” (The Belmont Report, 1979, p.1)

**Essential Element:** Ethical Research Criteria

**Additive/Variant Analysis:** The ethical concerns and treatment of human research subjects in research are additive to my understanding of ethical research conduct.

**Contextualization:** The Belmont Report by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, are principles set forth to guide researchers in the ethical conduct of dealing with human subjects in research. Participants have rights and ethical status that human treatment is required for all researchers to follow. In the twentieth century, there were many scientific advances. However, there were some unethical practices that were used, such as in Nazi Germany, where German scientists used humans as scientific subjects with total disregard for human values. However, this was not the only place unethical scientific research was being performed, such as in the United States medical research. The Belmont Report provides three ethical principles for the researcher to follow; respect for persons, beneficence, and justice. These are then used in the application of informed consent, assessment of risks and benefits, and selection of subjects.

**Source Nine:** Walshe, C., & Brearley, S. (Eds.). (2020). *Handbook of Theory and Methods in Applied Health Research: Questions, Methods and Choices*. Edward Elgar Publishing.

**Comment 12:**

**Quote/Paraphrase:** “There is, therefore a need to have some understanding of different research designs beforehand. Critical appraisal tools help you to formulate your views and opinions on a piece of research you read about in a structured, more objective manner. The tools can also enable tables to be constructed; assisting with the structure and organization of how the assessments are written up.” (Walshe & Brearley, 2020, p.11).

**Essential Element:** Literature Review Outline

**Additive/Variant Analysis:** The concept of using tools in your literature review process is additive to my understanding of the research process.

**Contextualization:** In this book they take the reader through the healthcare research process, including the literature review, which is a vital step in the research process. In this section, the authors inform us that there are literature review tools that can be used to help us through the process and organize our literature review, such tools as Zotero and many other online tools. However, you can also use internal tools to keep records of things from the articles, and an organized fashion for quick retrieval, such as spreadsheet based tools. No matter the tool used, the researcher should feel comfortable and easily function within the tool to help organize, summarize, and utilize the products of the literature review.

**Works Cited**

Benavidez, G. A., Zahnd, W. E., Hung, P., & Eberth, J. M. (2024). Chronic Disease Prevalence in the US: Sociodemographic and Geographic Variations by Zip Code Tabulation Area. *Preventing Chronic Disease*, *21*, 230267. <https://doi.org/10.5888/pcd21.230267>

Delmelle, E. M., Desjardins, M. R., Jung, P., Owusu, C., Lan, Y., Hohl, A., & Dony, C. (2022). Uncertainty in geospatial health: Challenges and opportunities ahead. *Annals of Epidemiology*, *65*, 15–30. <https://doi.org/10.1016/j.annepidem.2021.10.002>

Embury, J., Tsou, M.-H., Nara, A., & Oren, E. (2022). A Spatio-Demographic Perspective on the Role of Social Determinants of Health and Chronic Disease in Determining a Population’s Vulnerability to COVID-19. *Preventing Chronic Disease*, *19*, 210414. <https://doi.org/10.5888/pcd19.210414>

Gobaud, A. N., Mehranbod, C. A., Dong, B., Dodington, J., & Morrison, C. N. (2022). Absolute versus relative socioeconomic disadvantage and homicide: A spatial ecological case–control study of US zip codes. *Injury Epidemiology*, *9*(1), 7. <https://doi.org/10.1186/s40621-022-00371-z>

Habibi, R., Alesheikh, A. A., & Bayat, S. (2022). An event-based model and a map visualization approach for spatiotemporal association relations discovery of diseases diffusion. *Sustainable Cities and Society*, *87*, 104187. <https://doi.org/10.1016/j.scs.2022.104187>

Lee, E. K., Donley, G., Ciesielski, T. H., Freedman, D. A., & Cole, M. B. (2023). Spatial availability of federally qualified health centers and disparities in health services utilization in medically underserved areas. *Social Science & Medicine*, *328*, 116009. <https://doi.org/10.1016/j.socscimed.2023.116009>

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research.

Swales, L., Gooden, A., & Thaldar, D. (2024). The anatomy of a data transfer agreement for health research. *Frontiers in Pharmacology*, *15*, 1332700. <https://doi.org/10.3389/fphar.2024.1332700>

Walshe, C., & Brearley, S. (Eds.). (2020). *Handbook of Theory and Methods in Applied Health Research: Questions, Methods and Choices*. Edward Elgar Publishing.

Zeng, S., Pelzer, K. M., Gibbons, R. D., Peek, M. E., & Parker, W. F. (2022). Association of Zip Code Vaccination Rate With COVID-19 Mortality in Chicago, Illinois. *JAMA Network Open*, *5*(5), e2214753. <https://doi.org/10.1001/jamanetworkopen.2022.14753>