ERIC Notebook

Second Edition

Ecologic Studies

Second Edition Authors:

Lorraine K. Alexander, DrPH

Brettania Lopes, MPH

Kristen Ricchetti-Masterson, MSPH

Karin B. Yeatts, PhD, MS

Ecologic studies are studies in which the unit of observation is a group, not separate individuals, for one or more study variables. For example, exposure and risk factors are known only at the group level, such as the average air pollution concentration in different cities. The occurrence of the health outcome may also be only known at the group level, such as overall mortality rates from chronic lung disease in the same cities with measured levels of air pollution. Ecologic studies may be used to generate hypotheses of an association between exposure and a health outcome, but these studies cannot confirm causation. This is because, for example, we do not know whether those individuals who died in a particular city under observation had a higher exposure than individuals who remained alive.

Ecologic fallacy

When dealing with group level information, it is important to be aware of what is called the ecologic fallacy. This fallacy results from concluding that because an association exists between exposure and a health outcome at the group level, it therefore exists at the individual level. The cause of this fallacy is that we do not know the link between exposure and the health outcome among individuals within each group. For example, we don't know the number of diseased persons who were exposed or not exposed in the high exposure group or in the low exposure group.

Description of the ecologic fallacy

In ecologic studies, only information on aggregate measures, such as the average exposure in City A and the death rate in City A can be known. At the individual level, however, we can, for example, determine the proportion of people who died within each of the categories of exposure (low or high).

Suppose air pollution is higher in Baltimore than in Tampa, but mortality from lung disease is lower in Baltimore than in Tampa. It would be fallacious to conclude that air pollution protects against lung disease deaths. It is possible that persons dying of lung disease in Tampa may have moved from cities with high air pollution or that another risk factor for lung disease – such as smoking – is more prevalent in Tampa than Baltimore. We do not know the cumulative exposures of cases and non-cases in either city. The heterogeneity of lifetime air pollution exposure among individuals in each city makes the average exposure unrepresentative of the distribution of exposure among individuals in the population.

Using the previous example, in-migrants to Tampa had previously experienced higher levels of exposure to air pollution, thus causing the aggregate level of lung disease deaths to appear higher in Tampa. The aggregate level of air pollution in Tampa does not allow us to see that there are the varying levels of exposure between lifelong residents and in-migrants, this shows how the ecologic fallacy can occur.

Examples of questions investigated by ecologic studies include:

- Is the ranking of cities by air pollution levels associated with the ranking of cities by mortality from cardiovascular disease, adjusting for differences in average age, percent of the population below poverty level, and occupational structure?
- Have new car safety measures such as passenger air bags made a difference in motor vehicle fatality rates in areas with different laws over the same period of time?
- Has introduction of newer cars with safety features such as multiple airbags made a difference in motor vehicle fatality rates in areas with a different distribution of newer versus older cars, over the same period of time?
- Are daily variations in mortality in Boston related to daily variations in particle air pollution, adjusting for season of year and temperature?
- What are the long-term time trends (1950-2012) for mortality from the major cancers in the U.S., Canada, and Mexico?

Advantages of ecologic studies

Aggregate data on exposure and health outcomes are often publicly available in state and national databases, such as the US census or from the Center for Disease Control and Prevention. Agencies of the state and federal government collect considerable data reported at the aggregate level on the economy, the environment, and the health and wellbeing of the population. Data are regularly obtained on air quality, water quality, weather conditions, the size of the population, the status of the economy, and the health of the population through surveys such as the National Health Interview Survey, National Health and Nutrition Examination, and Behavioral Risk Factor Survey. Data on the vital status of the population are obtained via birth and death registries and cancer and birth defects registries. These publicly available records provide databases for linking health outcomes with characteristics of the population, the environment, and the economy at the aggregate level.

Aggregate level data can conveniently be obtained by researchers at a low cost and can be useful for evaluating the impact of community-level interventions. Examples of interventions that may be evaluated through ecologic study designs include fluoridation of water, seat belt laws, and mass media health campaigns. Aggregate level information can be compared before and after the intervention to determine the effects of an intervention at the community level.

In addition, minimal within-community differences between exposures may exist; however exposures may differ substantially between communities, cities, states, and countries. Examples of small within-community exposure differences but large between-community differences include:

- Quality of drinking water
- Concentration of certain air pollutants such as ozone and fine particles
- Average fat content of diet (larger differences between countries than between individuals within the same city)
- Cumulative exposure to sunlight (larger differences by latitude [north-south] of residence than among individuals at the same latitude)

Ecologic studies are also useful for studying the effect of short-term variations in exposure within the same community, such as the effect of temperature on mortality.

Types of ecologic study designs

There are three main types of ecologic study designs: cross-sectional ecologic studies, time-trend ecologic studies, and solely descriptive ecologic studies.

Cross-sectional ecologic studies compare aggregate exposures and outcomes over the same time period. An example of this study design is an investigation comparing bladder cancer mortality rates in cities with surface drinking water sources that contain chlorine by-products compared to rates in cities with ground drinking water sources that contain little or no chlorine by-products.

Time-trend ecologic studies compare variations in aggregate exposures and outcomes over time within the same community. A study investigating whether hospital admissions for cardiac disease in Los Angeles increase on days when carbon monoxide levels are higher would be an example of this type of study.

Solely descriptive ecologic studies investigate disease or risk factor differences between communities at the same time, or within the same community over time. This type of study design would be used to investigate the following questions: What are the differences in lung cancer mortality among cities in North Carolina? What is the secular trend of lung cancer mortality between 1960 and 2010 for the entire state of North Carolina?

Ecologic Study Designs at a Glance

Study type	Design	Time frame
Cross-sectional	Across communities	Same time period
Time-trend	Within the same community	Over time
Descriptive	Across communities or Within the same community	At a point in time or Over time

Limitations of ecologic studies

Ecologic studies are subject to numerous biases and limitations. Most notably, these study designs are subject to the ecologic fallacy, which occurs by inferring that associations at the aggregate level are true at the individual level. Ecologic studies are also more often subject to confounding bias than are individual risk studies. *Confounding* is a mixing of the effects of other risk factors with the exposure of interest. Confounding bias may occur in an ecologic study if the confounding factor is correlated with the background rate of disease (the disease rate among unexposed persons in each study community). *Cross-level bias* occurs when the confounding factor is associated with the background rate of disease differentially across groups. The ecologic fallacy may occur as a result of cross-level bias.

Example

Suppose the association between average fat consumption and breast cancer risk is examined across communities in the US. Certain communities may have a larger percentage of women with a genetic predisposition to breast cancer than other communities. Suppose these same communities containing large percentages of women with a genetic predisposition to breast cancer are also communities with a high per capita dietary fat consumption. The results of the study will show a strong correlation between average dietary fat consumption and breast cancer mortality. The association will be inflated due to the confounding factor, genetic predisposition to breast cancer. This bias occurred because the background rate of breast cancer incidence

Time-trend ecologic studies are further limited in that an investigator cannot be confident that exposure preceded the outcome. Migration into and out of communities can also bias the interpretation of ecologic results.

Practice Questions

Answers are at the end of this notebook

1) Researchers study the community of one town in North Carolina over a 10 year period. They conduct an ecologic study and collect data on the rate of automobile accidents each year among teenagers and the percentage of teens in the town who report using their phones for texting each year. Based on their data, the researchers conclude that teenagers who report texting on their phones are more likely to be in a car accident. Which of the following are true about the researchers' conclusion? Choose all that apply.

a) The researchers' conclusion is valid

b) The researchers have incorrectly used group-level data to draw conclusions about individual teenagers

c) The researchers do not know if the teens involved in car accidents are the same teens who have reported texting on their phones, therefore, their conclusion is not valid

d) The researchers' conclusion does not account for what other factors, if any, are related to texting and related to being in a car accident

2) Reported cases of the flu are higher in city A than city B. Vaccination rates for the flu are lower in city A than city B. Which of the following are reasons why would it be incorrect to simply assume that higher vaccination in city B is what is causing city B to have fewer reported cases of the flu? Choose all that apply.

a) City A and city B may have different strains of the flu

b) City A and city B may have different proportions of people in their populations who are especially vulnerable to the flu (e.g. the elderly, children and pregnant women)

c) City A and city B may have differences in health care accessibility, leading to differences in testing for the flu and diagnosis of the flu

d) City A and city B may have different climates, leading to differences in how/where people come into contact with each other. This may affect flu transmission rates

References

Dr. Carl M. Shy, Epidemiology 160/600 Introduction to Epidemiology for Public Health course lectures, 1994-2001, The University of North Carolina at Chapel Hill, Department of Epidemiology

Rothman KJ, Greenland S. Modern Epidemiology. Second Edition. Philadelphia: Lippincott Williams and Wilkins, 1998.

The University of North Carolina at Chapel Hill, Department of Epidemiology Courses: Epidemiology 710, Fundamentals of Epidemiology course lectures, 2009-2013, and Epidemiology 718, Epidemiologic Analysis of Binary Data course lectures, 2009-2013.

Acknowledgement

The authors of the Second Edition of the ERIC Notebook would like to acknowledge the authors of the ERIC Notebook, First Edition: Michel Ibrahim, MD, PhD, Lorraine Alexander, DrPH, Carl Shy, MD, DrPH, and Sherry Farr, GRA, Department of Epidemiology at the University of North Carolina at Chapel Hill. The First Edition of the ERIC Notebook was produced by the Educational Arm of the Epidemiologic Research and Information Center at Durham, NC. The funding for the ERIC Notebook First Edition was provided by the Department of Veterans Affairs (DVA), Veterans Health Administration (VHA), Cooperative Studies Program (CSP) to promote the strategic growth of the epidemiologic capacity of the DVA.

Answers to Practice Questions

1. Answer choices b, c and d are correct. In the example, the researchers have used overall data on car accidents for a town and overall data on the prevalence of teens who reported texting to draw a conclusion about individuals. The researchers actually do not have any data on whether the teens involved in car accidents were the same teens who reported texting. In addition, there may be other factors involved that explain the study results such as another confounding variable that is related to being in a car accident and related to reporting texting.

2. All answer choices are correct. City B may have a different strain of the flu circulating, which could affect transmission rates, morbidity/mortality and whether or not an infected person seeks medical care. There may be more people in city A who are especially vulnerable to the flu and these vulnerable groups may experience greater illness, leading to more visits to a health provider and more diagnoses. If health care accessibility varies between the 2 cities and/or knowledge of testing and diagnosis for the flu differs between the 2 cities, this may affect whether infected persons seek health care and may affect which tests they receive. Such variability may lead to the appearance of greater or fewer cases. Finally, the climate may influence social mixing patterns leading to greater virus transmission in the city where people spend more time together indoors.