Observational Studies *vs.* Designed Experiments

MATH 130, *Elements of Statistics I*

J Robert Buchanan

Department of Mathematics

Fall 2023

Objectives

Once a research question has been formulated researchers must collect data. Data may be collected through observation or experimentation.

During this lesson we will learn to:

- distinguish between an observational study and a designed experiment, and
- explain the various types of observational studies.

Types of Studies

Definition

An **observational study** measures the characteristics of a population by studying individuals in a sample, but does not attempt to manipulate or influence the variable(s) of interest.

Definition

A **designed experiment** applies a treatment, an **explanatory variable** to individuals (referred to as **experimental subjects** or **units**) and attempts to isolate the effects of treatment on a **response variable**.

Example: Observational Study

Researchers Joachim Schüz and associates wanted “to investigate cancer risk among Danish cellular phone users who were followed for up to 21 years.” To do so, they kept track of 420,095 people whose first cellular telephone subscription was between 1982 and 1995. In 2002, they recorded the number of people out of the 420,095 people who had a brain tumor and compared the rate of brain tumors in this group to the rate of brain tumors in the general population.

They found no significant difference in the rate of brain tumors between the two groups. The researchers concluded “cellular telephone was not associated with increased risk for brain tumors.” (Source: Joachim Schüz *et al*. “Cellular Telephone Use and Cancer Risk: Update of a Nationwide Danish Cohort,” *Journal of the National Cancer Institute* 98(23): 1707-1713, 2006)

Example: Designed Experiment

Researchers Joseph L. Roti and associates examined “whether chronic exposure to radio frequency (RF) radiation at two common cell phone signals - 835.62 megahertz, a frequency used by analogue cell phones, and 847.74 megahertz, a frequency used by digital cell phones - caused brain tumors in rats. The rats in group 1 were exposed to the analogue cell phone frequency; the rats in group 2 were exposed to the digital frequency; the rats in group 3 served as controls and received no radiation. The exposure was done for 4 hours a day, 5 days a week for 2 years. The rats in all three groups were treated the same, except for the RF exposure.

After 505 days of exposure, the researchers reported the following after analyzing the data. “We found no statistically significant increases in any tumor type, including brain, liver, lung or kidney, compared to the control group.” (Source: M. La Regina, E. Moros, W. Pickard, W. Straube, J. L. Roti Roti. “The Effect of Chronic Exposure to 835.62 MHz FMCW or 847.7 MHz CDMA on the incidence of Spontaneous Tumors in Rats.” Bioelectromagnetic Society Conference, June 25, 2002.)

Response and Explanatory Variables

In both studies, the goal of the research was to determine if radio frequencies from cell phones increase the risk of contracting brain tumors.

Whether or not brain cancer was contracted is the **response variable**. The level of cell phone usage is the **explanatory variable**.

In research, we wish to determine how varying the amount of an explanatory variable affects the value of a response variable.

Causation *vs.* Association

Remark: observational studies are useful for determining whether there is a relation between two variables, but a designed experiment is required to isolate the cause of the relation.

Causation *vs.* Association

Remark: observational studies are useful for determining whether there is a relation between two variables, but a designed experiment is required to isolate the cause of the relation.

Definition

Confounding occurs when the effects of two or more explanatory variables cannot be separated. Therefore, any relation that may exist between an explanatory variable and the response variable may be due to some other variable or variables not accounted for in the study.

Definition

A **lurking variable** is an explanatory variable that was not considered in a study, but that affects the value of the response variable in the study. In addition, lurking variables are typically related to any explanatory variables considered in the study.

Example

Studies have observed that women taking hormone replacement therapy (HRT) seem to have a lower risk of heart disease. The researchers observed women who decided for themselves whether or not to take HRT.

Perhaps women who choose to take HRT are healthier to begin with (lurking variable) and thus are already at a lower risk of heart disease.

Example

Studies have observed that women taking hormone replacement therapy (HRT) seem to have a lower risk of heart disease. The researchers observed women who decided for themselves whether or not to take HRT.

Perhaps women who choose to take HRT are healthier to begin with (lurking variable) and thus are already at a lower risk of heart disease.

Observational studies do not allow a researcher to claim causation, only association.

Sources of Data

- Census: a list of all individuals in a population along with certain characteristics of each individual.

Sources of Data

- Census: a list of all individuals in a population along with certain characteristics of each individual.
- Existing sources.

Sources of Data

- Census: a list of all individuals in a population along with certain characteristics of each individual.
- Existing sources.
- Survey sampling.

Sources of Data

- Census: a list of all individuals in a population along with certain characteristics of each individual.
- Existing sources.
- Survey sampling.
- Designed experiments.

Types of Observational Studies

Cross-sectional: collect information about individuals at a specific point in time or over a very short interval of time.

Case-control: **retrospective** studies that require individuals to look back in time or researchers to look at existing records.

Cohort: record observations of a group of individuals over a (long) period of time. Since the data are collected over time, these are **prospective** studies.

Example

Scientists were interested in determining if abdominal obesity is related to coronary artery calcification (CAC). The scientists studied 2,951 participants in the Coronary Artery Risk Development in Young Adults Study to investigate a possible link. Waist and hip girths were measured in 1985-1986, 1995-1996, and in 2000-2001. CAC measurements were taken in 2001-2002.

Example

Scientists were interested in determining if abdominal obesity is related to coronary artery calcification (CAC). The scientists studied 2,951 participants in the Coronary Artery Risk Development in Young Adults Study to investigate a possible link. Waist and hip girths were measured in 1985-1986, 1995-1996, and in 2000-2001. CAC measurements were taken in 2001-2002.

- What type of observational study was this?
- What is the response variable?
- What is the explanatory variable?

Example

Scientists were interested in determining if abdominal obesity is related to coronary artery calcification (CAC). The scientists studied 2,951 participants in the Coronary Artery Risk Development in Young Adults Study to investigate a possible link. Waist and hip girths were measured in 1985-1986, 1995-1996, and in 2000-2001. CAC measurements were taken in 2001-2002.

- What type of observational study was this? (cohort)
- What is the response variable?
- What is the explanatory variable?

Example

Scientists were interested in determining if abdominal obesity is related to coronary artery calcification (CAC). The scientists studied 2,951 participants in the Coronary Artery Risk Development in Young Adults Study to investigate a possible link. Waist and hip girths were measured in 1985-1986, 1995-1996, and in 2000-2001. CAC measurements were taken in 2001-2002.

- What type of observational study was this? (cohort)
- What is the response variable? (CAC)
- What is the explanatory variable?

Example

Scientists were interested in determining if abdominal obesity is related to coronary artery calcification (CAC). The scientists studied 2,951 participants in the Coronary Artery Risk Development in Young Adults Study to investigate a possible link. Waist and hip girths were measured in 1985-1986, 1995-1996, and in 2000-2001. CAC measurements were taken in 2001-2002.

- What type of observational study was this? (cohort)
- What is the response variable? (CAC)
- What is the explanatory variable? (waist and hip girth)

Observational Study or Designed Experiment?

Researchers wanted to assess the long-term psychological effects on children evacuated during World War II. They obtained a sample of 169 former evacuees and a control group of 43 people who were children during the war but were not evacuated. The subjects' mental states were evaluated using questionnaires. It was determined that the psychological well being of the individuals was adversely affected by evacuation. (Source: Foster D, Davies S, and Steele H (2003) The evacuation of British children during World War II: a preliminary investigation into the long-term psychological effects. Aging & Mental Health (7)5.)

Observational Study or Designed Experiment?

Researchers wanted to assess the long-term psychological effects on children evacuated during World War II. They obtained a sample of 169 former evacuees and a control group of 43 people who were children during the war but were not evacuated. The subjects' mental states were evaluated using questionnaires. It was determined that the psychological well being of the individuals was adversely affected by evacuation. (Source: Foster D, Davies S, and Steele H (2003) The evacuation of British children during World War II: a preliminary investigation into the long-term psychological effects. Aging & Mental Health (7)5.)

(Observational study, case-control)

Observational Study or Designed Experiment?

Xylitol has proven effective in preventing dental caries (cavities) when included in food or gum. A total of 75 Peruvian children were given milk with and without xylitol and were asked to evaluate the taste of each. Overall, the children preferred the milk flavored with xylitol. (Source: Castillo JL, *et al*. (2005) Children's acceptance of milk with xylitol or sorbitol for dental caries prevention. BMC Oral Health (5)6.)

Observational Study or Designed Experiment?

Xylitol has proven effective in preventing dental caries (cavities) when included in food or gum. A total of 75 Peruvian children were given milk with and without xylitol and were asked to evaluate the taste of each. Overall, the children preferred the milk flavored with xylitol. (Source: Castillo JL, *et al*. (2005) Children's acceptance of milk with xylitol or sorbitol for dental caries prevention. BMC Oral Health (5)6.)

(Designed experiment)

Observational Study or Designed Experiment?

A total of 974 homeless women in the Los Angeles area were surveyed to determine their level of satisfaction with the healthcare provided by shelter clinics versus the healthcare provided by government clinics. The women reported greater quality satisfaction with the shelter and outreach clinics compared to the government clinics. (Source: Swanson KA, Andersen R, Gelberg L (2003) Patient satisfaction for homeless women. Journal of Women's Health (12)7.)

Observational Study or Designed Experiment?

A total of 974 homeless women in the Los Angeles area were surveyed to determine their level of satisfaction with the healthcare provided by shelter clinics versus the healthcare provided by government clinics. The women reported greater quality satisfaction with the shelter and outreach clinics compared to the government clinics. (Source: Swanson KA, Andersen R, Gelberg L (2003) Patient satisfaction for homeless women. Journal of Women's Health (12)7.)

(Observational study, cross-sectional)

Observational Study or Designed Experiment?

The Cancer Prevention Study II (CPS-II) is funded and conducted by the American Cancer Society. Its goal is to examine the relationship among environmental and lifestyle factors on cancer cases by tracking approximately 1.2 million men and women. Study participants completed an initial study questionnaire in 1982 providing information on a range of lifestyle factors such as diet, alcohol and tobacco use, occupation, medical history, and family cancer history. These data have been examined extensively in relation to cancer mortality. Vital status of study participants is updated biennially. Cause of death has been documented for over 98% of all deaths that have occurred. Mortality follow-up of the CPS-II participants is complete through 2002 and is expected to continue for many years. (Source: American Cancer Society)

Observational Study or Designed Experiment?

The Cancer Prevention Study II (CPS-II) is funded and conducted by the American Cancer Society. Its goal is to examine the relationship among environmental and lifestyle factors on cancer cases by tracking approximately 1.2 million men and women. Study participants completed an initial study questionnaire in 1982 providing information on a range of lifestyle factors such as diet, alcohol and tobacco use, occupation, medical history, and family cancer history. These data have been examined extensively in relation to cancer mortality. Vital status of study participants is updated biennially. Cause of death has been documented for over 98% of all deaths that have occurred. Mortality follow-up of the CPS-II participants is complete through 2002 and is expected to continue for many years. (Source: American Cancer Society)

(Observational study, cohort)