It's just what the doctor ordered: observational studies.

Steve Simon, P.Mean Consulting
- Facebook: www.facebook.com/pmean
- LinkedIn: www.linkedin.com/in/pmean
- Main website: www.pmean.com
- Newsletter: www.pmean.com/news
- Twitter: twitter.com/profmean

You can get a copy of this handout at:
- www.pmean.com/webinars/20120519/observational.pdf

2. Abstract

- Abstract: An observational study is a study where the researchers do not directly intervene, but instead let the patients and/or their doctors choose the treatment. Observational studies also arise when a group is intact at the start of the study. There are four types of observational studies: cohort studies, case-control studies, cross-sectional studies, and historical control studies. While observational studies are generally considered to be less authoritative than randomized studies, with careful selection of the control subjects, observational studies can still provide persuasive results.

3. Objectives

In this class you will learn how to:
- list the four common types of observational studies,
- distinguish between cohort and case-control studies, and
- explain the limitations of historical control studies.

4. Sources

Part of the material for this webinar comes from:
- Simon SD. Statistical Evidence in Medical Trials, What Do the Data Really Tell Us? 2006. Oxford University Press: Oxford, England.
- Stats #32b: Statistical Evidence: Apples or Oranges? Randomized studies.
  - http://www.childrens-mercy.org/stats/training/hand32b.asp

5. Pop quiz #1

Which of the following is NOT an observational design?
1. Case-control study
2. Cohort study
3. Cross-sectional study
4. Historical control trial
5. Randomized control trial
6. Don't know/not sure

6. Pop quiz #2

Which type of study is best for evaluating rare diseases:
1. Case-control study
2. Cohort study
3. Cross-sectional study
4. Historical control trial
5. Randomized control trial
6. Don't know/not sure

7. Pop quiz #3

The historical control design is considered a weak form of evidence except when:

1. the disease being studied is rare
2. the exposure is too risky to allow random assignment
3. the mortality/morbidity rate is close to 100%
4. there is strong evidence of covariate imbalance
5. those who don't understand history are doomed to repeat it.
6. don't know/not sure

8. Observational studies

- There are many situations where randomization is not ethical, practical, or possible. This includes setting with:
  - a dangerous exposure,
  - limited financial resources,
  - strong patients/physicians preferences
  - groups that already exist

9. Observational studies

Observational studies are those studies where the researcher can't/won't assign patients to treatment/control groups. There are four major flavors for observational studies:

1. cohort studies,
2. case control studies,
3. cross-sectional studies, and
4. historical controls studies.

10. Cohort studies

In a cohort study, a group of patients has a certain exposure or condition. They are compared to a group of patients without that exposure or condition. Does the exposed cohort differ from the unexposed cohort on an outcome of interest?

11. Cohort studies

Example: In a study of suicide among Swedish men in the Swedish military service conscription register (Gunnell 2005), 987,308 men registered between 1968 and 1994 were divided into nine groups on the basis of four intelligence tests. These men were also linked to a Swedish cause of death register which identified a total of 2,811 suicides among these men. For each of the four intelligence tests, men scoring lower tended to have a higher rate of suicide.

12. Cohort studies

Example: In a study of psychotic symptoms in young people, a sample of young adults aged 14–24 years were divided into a group of 320 with admitted use of cannabis and a group of 2,117 did not admit to cannabis use. Both groups were followed four years later for psychotic symptoms.

13. Cohort studies

Cohort studies are intuitively appealing and selection of a control group is usually not too difficult. You have to be wary of covariate imbalance, but do not worry about every possible covariate imbalance. You should look for large imbalances, especially for covariates which are closely related to the outcome variable.

14. Cohort study

When you are studying a very rare outcome, the sample size may have to be extremely large. As a rough rule of thumb, you need to observe 25–50 outcomes in each group in order to have a reasonable level of precision. So when a condition occurs only once in every thousand patients, a cohort study would require tens of thousands of patients.

15. Cohort study

You want to avoid 'leaky groups' in a cohort design. If the exposure group includes some unexposed patients and the control group includes some exposed patients, then any effect you are trying to detect will be diluted.

Examples:

- Equating caffeine consumption with coffee drinking.
- Measuring dietary consumption of individuals through family shopping data.

16. Case-control study

A case-control study selects patients on the basis of an outcome, such as development of breast cancer, and are compared to a group of patients without that outcome.

17. Case-control study

Example: In a study of asthma deaths (Anderson 2005), researchers selected 532 patients who died between 1994 and 1998 with asthma mentioned in part I of the death certificate. For each asthma death, a similar asthma admission (without death) was identified at the same hospital, with a similar admission date and a similar age..

18. Case-control study

Example: In a study of vascular dementia (Chan Carusone 2004), researchers selected 28 patients with vascular dementia who were enrolled in the Geriatric Clinic at Henderson Hospital in Hamilton, Ontario, between July 1999 and October 2001. They also selected controls from a list of all caregivers at that clinic, regardless of the diagnosis of their spouse or family member, as long as the caregiver did not have any signs of dementia or stroke. Caregivers were matched by age (within 5 years) and sex. The researchers tested both cases and controls for Chalamydia.

19. Case-control study

A case-control study is very efficient in studying rare diseases. With this design, you round up all of the limited number of cases of the disease and then find a comparable control group. By contrast, a cohort design has to round up far more exposures to ensure that a handful of them will develop the rare disease.

20. Case-control study

The case-control study is always retrospective because the outcome in a case-control study has already occurred. Retrospective studies usually have more problems with data quality because our memory is not always perfect. What is worse is that sometimes the ability to remember is sharply influenced by the outcome being studied.

21. Case-control study

In a case-control study, it is often very hard to find a good control group. You want to find controls that are identical to the cases in all aspects except for the outcome itself. What does it mean to be exactly like a lung cancer patient, except for the lung cancer?

22. Case-control study

Finally, the case-control design just does not sit well with your intuition. You are trying to find factors that cause an outcome, so you are sampling from the causes while a cohort design samples from the effects. Don't let this bother you too much, though. The mathematics that justify the case-control design were developed half a century ago (Cornfield 1951).

23. Case-control design

The careful use of the case-control design has helped answer important clinical questions which could not have been answered by other research designs. Case-control designs, for example, established the use of aspirin as a cause of Reye's syndrome (Monto 1999). It is hard to imagine how a randomized trial for Reye's syndrome could have been done.

24. Cross-sectional design

In contrast to the cohort and the case-control design, the cross-sectional study select on the basis of neither exposure nor outcome. With the cross-sectional design, you select a single group of patients and simultaneously assess both their exposure variables and their outcome variables. Typically, there are multiple exposures and multiple outcomes in a cross-sectional study.

25. Cross-sectional study

Example: In a study of intimate partner violence (Malcoe 2004), 312 Native American women attending a tribally operated clinic filled out a survey form. The survey included a modified Conflict Tactics Scale to assess whether the women experienced verbal or psychological aggression, or physical or sexual assault. The survey also asked about educational attainment, employment status, receipt of food stamps, and other questions to help determine their socioeconomic status. Since both the outcome (intimate partner violence) and the exposure (socioeconomic status) were determined at the same time, this represents a cross-sectional survey.

26. Cross-sectional study

Example: In a study of respiratory problems (Salo 2004), 5,051 seventh grade students in Wuhan, China, completed a self-administered questionnaire.These students were classified according to six respiratory outcomes (wheezing with colds, wheezing without colds, bringing up phlegm with colds, bringing up phlegm without colds, coughing with colds, coughing without colds) and two exposure variables (coal burning for cooking and cleaning, and smoking in the home). Students were not randomly assigned to an exposure; so this is an observational study. Both the outcome variables and the exposure variables were assessed at a single point in time, so this represents a cross-sectional study.

27. Cross-sectional study

Since there is no separation in time between assessment of exposure and assessment of outcome, you often cannot determine which came first. This loss of temporality makes it difficult to infer a cause-and-effect.

28. Cross-sectional study

A hypothetical example of patient height (Mann 2003), describes how a cross-sectional study might notice a negative association between height and age. Could this be because people shrink as they age, or perhaps successive generations of people are taller because of the improvements in nutrition, or perhaps taller people just die earlier? With a cross-sectional study, you cannot easily disentangle these alternate explanations.

29. Cross-sectional study

Cross-sectional studies are fast as you do not have to wait around to see what happens to the patients. These studies also allow you to easily explore relationships between multiple exposure variables and/or multiple outcome variables. But unlike the cohort design, which is useful for rare exposures, or the case-control design, which is useful for rare outcomes, the cross-sectional study is only effective if both the exposure and the outcome are relatively common events.

30. Historical controls study

In a historical controls study, researchers will assign all of the research subjects to the new therapy. The outcomes of these subjects are compared to historical records representing the standard therapy.

31. Historical controls study

Example: In a study of the rapid parathyroid hormone test (Johnson 2001), 49 patients undergoing parathyroidectomy received the rapid test. These patients were compared to 55 patients undergoing the same procedure before the rapid test was available. This is an observational study because the calendar, not the researchers, determined which test was applied. This particular observational study is a historical controls design because the control group represents patients tested before the availability of the rapid test.

32. Historical controls study

The very nature of a historical controls study guarantees that there will be a major covariate imbalance between the two groups. Thus, you have to consider any factors that have changed over time that might be related to the outcome. To what extent might these factors affect the outcome differentially?

33. Historical controls study

For the most part, historical controls are considered one of the weakest forms of evidence. The one exception is when a disease has close to 100% mortality. In that situation, there is no need for a concurrent control group, since any therapy that is remotely effective can readily be detected. Even in this situation, you want to be sure there is a biological basis for the treatment and that the disease group is homogeneous.

34. Practice exercises

- For each of the following abstracts, categorize the research studies as one of the following:
  - case-control study
  - cohort study
  - cross-sectional study
  - historical control study

35. Conclusion

Observational studies are used when randomization is not possible, practical, or ethical. Cohort designs select patients on the basis of their exposure. Case-control designs select patients on the basis of their outcome. Selecting appropriate controls in a case-control design is difficult, but this design is efficient when studying a rare disease.

36. Conclusion

Cross-sectional studies select a single group of patients and classify them by multiple exposures and multiple outcomes. Because there is not always an obvious time order in the data collection, it is easy in a cross-sectional study to confuse causes and effects. Historical control studies provide an intervention to all new patients and compare them to previous medical records. Historical control studies always have a serious covariate imbalance, but are still useful when studying a condition that has close to 100% morbidity/mortality.

37. Repeat of pop quiz #1

Which of the following is NOT an observational design?

1. Case-control study
2. Cohort study
3. Cross-sectional study
4. Historical control trial
5. Randomized control trial
6. Don't know/not sure

38. Repeat of pop quiz #2

Which type of study is best for evaluating rare diseases:

1. Case-control study
2. Cohort study
3. Cross-sectional study
4. Historical control trial
5. Randomized control trial
6. Don't know/not sure

39. Repeat of pop quiz #3

The historical control design is considered a weak form of evidence except when:

1. the disease being studied is rare
2. the exposure is too risky to allow random assignment
3. the mortality/morbidity rate is close to 100%
4. there is strong evidence of covariate imbalance
5. those who don't understand history are doomed to repeat it.
6. don't know/not sure

40. Conclusion

Where do you go from here?

1. Don't pretend that you are a professional statistician, no matter how well I taught this course.
2. But, you should be a much better consumer of Statistics.
3. You are in a better position to raise questions that your customers need to ask when they read a paper.

Sign up for The Monthly Mean, www.pmean.com/news