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

Steve Simon
P.Mean Consulting
www.pmean.com

2. Why do I offer this webinar for free?

I offer free statistics webinars partly for fun and partly to build up goodwill for my consulting business,

- www.pmean.com/consult.html.

Also see my Facebook and LinkedIn pages

- www.facebook.com/pmean
- www.linkedin.com/in/pmean

I provide a free newsletter about Statistics, The Monthly Mean.

- www.pmean.com/news
- www.facebook.com/group.php?gid=302778306676

3. 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.

4. 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.

5. 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

6. 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

7. 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

8. 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

9. 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

10. 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.

11. 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?

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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..

19. 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.

20. 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.

21. 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.

22. 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?

23. 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).

24. 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.

25. 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.

26. 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.

27. 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.

28. 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.

29. 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.

30. 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.

31. 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.

32. 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.

33. 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?

34. 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.

35. 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

40. 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.

41. 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.

42. 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

43. 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

44. 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

1. Body fatness during childhood and adolescence and incidence of breast cancer in premenopausal women: a prospective cohort study. Heather J Baer, Graham A Colditz, Bernard Rosner, Karin B Michels, Janet W Rich-Edwards, David J Hunter and Walter C Willett. Breast Cancer Research 2005, 7:R314-R325 doi:10.1186/bcr998.

Introduction Body mass index (BMI) during adulthood is inversely related to the incidence of premenopausal breast cancer, but the role of body fatness earlier in life is less clear. We examined prospectively the relation between body fatness during childhood and adolescence and the incidence of breast cancer in premenopausal women. **Methods** Participants were 109,267 premenopausal women in the Nurses' Health Study II who recalled their body fatness at ages 5, 10 and 20 years using a validated 9-level figure drawing. Over 12 years of follow up, 1318 incident cases of breast cancer were identified. Cox proportional hazards regression was used to compute relative risks (RRs) and 95% confidence intervals (CIs) for body fatness at each age and for average childhood (ages 5–10 years) and adolescent (ages 10–20 years) fatness. **Results** Body fatness at each age was inversely associated with premenopausal breast cancer incidence; the multivariate RRs were 0.48 (95% CI 0.35–0.55) and 0.57 (95% CI 0.39–0.83) for the most overweight compared with the most lean in childhood and adolescence, respectively (P for trend < 0.0001). The association for childhood body fatness was only slightly attenuated after adjustment for later BMI, with a multivariate RR of 0.52 (95% CI 0.38–0.71) for the most overweight compared with the most lean (P for trend = 0.001). Adjustment for menstrual cycle characteristics had little impact on the association. **Conclusion** Greater body fatness during childhood and adolescence is associated with reduced incidence of premenopausal breast cancer, independent of adult BMI and menstrual cycle characteristics. http://breast-cancer-research.com/content/7/3/R314

2. **Impact of a nurses' protocol-directed weaning procedure on outcomes in patients undergoing mechanical ventilation for longer than 48 hours: a prospective cohort study with a matched historical control group**. Jean-Marie Tonnelier, Gwenaël Prat, Grégoire Le Gal, Christophe Gut-Gobert, Anne Renault, Jean-Michel Boles and Erwan L'Her. Critical Care 2005, 9:R83-R89 doi:10.1186/cc3030. **Introduction** The aim of the study was to determine whether the use of a nurses' protocol-directed weaning procedure, based on the French intensive care society (SRLF) consensus recommendations, was associated with reductions in the duration of mechanical ventilation and intensive care unit (ICU) length of stay in patients requiring more than 48 hours of mechanical ventilation. **Methods** This prospective study was conducted in a university hospital ICU from January 2002 through to February 2003. A total of 104 patients who had been ventilated for more than 48 hours and were weaned from mechanical ventilation using a nurses' protocol-directed procedure (cases) were compared with a 1:1 matched historical control group who underwent conventional physician-directed weaning (between 1999 and 2001). Duration of ventilation and length of ICU stay, rate of unsuccessful extubation and rate of ventilator-associated pneumonia were compared between cases and controls. **Results** The duration of mechanical ventilation (16.6 ± 13 days versus 22.5 ± 21 days; P = 0.02) and ICU length of stay (21.6 ± 14.3 days versus 27.6 ± 21.7 days; P = 0.02) were lower among patients who underwent the nurses' protocol-directed weaning than among control individuals. Ventilator-associated pneumonia, ventilator discontinuation failure rates and ICU mortality were similar between the two groups. **Discussion** Application of the nurses' protocol-directed weaning procedure described here is safe and promotes significant outcome benefits in patients who require more than 48 hours of mechanical ventilation. http://ccforum.com/content/9/2/R83

3. **Extravascular lung water in patients with severe sepsis: a prospective cohort study**. Greg S Martin, Stephanie Eaton, Meredith Mealer and Marc Moss. Critical Care 2005, 9:R74-R82 doi:10.1186/cc3025. **Introduction** Few investigations have prospectively examined extravascular lung water (EVLW) in patients with severe sepsis. We sought to determine whether EVLW may contribute to lung injury in these patients by quantifying the relationship of EVLW to parameters of lung injury, to determine the effects of chronic alcohol abuse on EVLW, and to determine whether EVLW may be a useful tool in the diagnosis of acute respiratory distress syndrome (ARDS). **Methods** The present prospective cohort study was conducted in consecutive patients with severe sepsis from a medical intensive care unit in an urban university teaching hospital. In each patient, transpulmonary thermodilution was used to measure cardiovascular hemodynamics and EVLW for 7 days via an arterial catheter placed within 72 hours of meeting criteria for severe sepsis. **Results** A total of 29 patients were studied. Twenty-five of the 29 patients (86%) were mechanically ventilated, 15 of the 29 patients (52%) developed ARDS, and overall 28-day mortality was 41%. Eight out of 14 patients (57%) with non-ARDS severe sepsis had high EVLW with significantly greater hypoxemia than did those patient with low EVLW (mean arterial oxygen tension/fractional inspired oxygen ratio 230.7 ± 36.1 mmHg versus 341.2 ± 92.8 mmHg; $P < 0.001$). Four out of 15 patients with severe sepsis with ARDS maintained a low EVLW and had better 28-day survival than did ARDS patients with high EVLW (100% versus 36%; $P = 0.03$). ARDS patients with a history of chronic alcohol abuse had greater EVLW than did nonalcoholic patients (19.9 ml/kg versus 8.7 ml/kg; $P < 0.0001$). The arterial oxygen tension/fractional inspired oxygen ratio, lung injury score, and chest radiograph scores correlated with EVLW ($r2 = 0.27$, $r2 = 0.18$, and $r2 = 0.28$, respectively; all $P < 0.0001$). **Conclusions** More than half of the patients with severe sepsis but without ARDS had increased EVLW, possibly representing subclinical lung injury. Chronic alcohol abuse was associated with increased EVLW, whereas lower EVLW was associated with survival. EVLW correlated moderately with the severity of lung injury but did not account for all respiratory derangements. EVLW may improve both risk stratification and management of patients with severe sepsis. http://ccforum.com/content/9/2/R74

4. **Breast implants following mastectomy in women with early-stage breast cancer: prevalence and impact on survival**. Gem M Le, Cynthia D O'Malley, Sally L Glaser, Charles F Lynch, Janet L Stanford, Theresa HM Keegan and Dee W West. Breast Cancer Res 2005, 7:R184-R193 doi:10.1186/bcr974. **Background** Few studies have examined the effect of breast implants after mastectomy on long-term survival in breast cancer patients, despite growing public health concern over potential long-term adverse health effects. **Methods** We analyzed data from the Surveillance, Epidemiology and End Results Breast Implant Surveillance Study conducted in San Francisco–Oakland, in Seattle–Puget Sound, and in Iowa. This population-based, retrospective cohort included women younger than 65 years when diagnosed with early or unstaged first primary breast cancer between 1983 and 1989, treated with mastectomy. The women were followed for a median of 12.4 years (n = 4968). Breast implant usage was validated by medical record review. Cox proportional hazards models were used to estimate hazard rate ratios for survival time until death due to breast cancer or other causes for women with and without breast implants, adjusted for relevant patient and tumor characteristics. **Results** Twenty percent of cases received postmastectomy breast implants, with silicone gel-filled implants comprising the most common type. Patients with implants were younger and more likely to have in situ disease than patients not receiving implants. Risks of breast cancer mortality (hazard ratio, 0.54; 95% confidence interval, 0.43–0.67) and nonbreast cancer mortality (hazard ratio, 0.59; 95% confidence interval, 0.41–0.85) were lower in patients with implants than in those patients without implants, following adjustment for age and year of diagnosis, race/ethnicity, stage, tumor grade, histology, and radiation therapy. Implant type did not appear to influence long-term survival. **Conclusions** In a large, population-representative sample, breast implants following mastectomy do not appear to confer any survival disadvantage following early-stage breast cancer in women younger than 65 years old. http://breast-cancer-research.com/content/7/2/R184