Putting your life in the hands of a coin: Randomized trials.

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/20121003

2. Abstract

- Abstract: In research studies that compare a treatment group and a control group, you need to assess whether the comparison is a fair comparison—an apples to apples comparison. Randomization is a simple method that insures that patients assigned to the treatment group are comparable to patients assigned to the control group. There are, however, practical and ethical constraints that can sometimes prevent the use of randomization.

3. Objectives

In this class you will learn how to:
- describe how covariate imbalance can threaten the validity of a research study,
- explain how randomization prevents covariate imbalance, and
- understand the practical and ethical limitations to randomized 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.
- Simon SD. Is the randomized clinical trial the gold standard of research?. J Androl. 2001 Nov-Dec;22(6):938-43.
- Stats #32a: Statistical Evidence: Apples or Oranges? Randomized studies.
  - http://www.childrens-mercy.org/stats/training/hand32a.asp

5. Pop quiz #1

When the demographic profile of the patients in your treatment group differ sharply from the profile of patients in your control group, you have:
1. covariate imbalance,
2. observational data,
3. response bias,
4. spectrum bias,
5. stratified data,
6. don't know/not sure

6. Pop quiz #2

Randomization is not practical:
1. when doctors believe that the new treatment is superior to the current standard
2. when patients have a strong preference for a particular treatment
3. when the experiment requires deliberate exposure of patients to something that is known to be harmful
4. randomization is impractical for all of the above situations
5. randomization can be applied easily in all of the above situations
6. don't know/not sure

7. Pop quiz #3

The following approaches are credible alternatives to randomization:

1. alternating between treatment and control
2. assigning all new patients to the treatment group and choosing controls from a medical database
3. assigning treatment group on the basis of the last digit of your birthday
4. letting the doctor choose whether a patient gets into the treatment group or the control group
5. none of these approaches is as effective as randomization
6. don't know/not sure

8. Covariate imbalance

Almost all research involves comparison. Do women who take Tamoxifen have a lower rate of breast cancer recurrence than women who take a placebo? Do left-handed people die at an earlier age than right-handed people? Are men with severe vertex balding more likely to develop heart disease than men with no balding?

9. Covariate imbalance

When you make a comparison between a treatment group and a control group, you want a fair comparison. You want the control group to be identical to the treatment group in all respects, except for the treatment in question. You want an apples-to-apples comparison.

10. Covariate imbalance

Sometimes, however, you get an unfair comparison, an apples-to-oranges comparison. The control group differs on some important characteristics that might influence the outcome measure. This is known as covariate imbalance. Covariate imbalance is not an insurmountable problem, but it does make a study less authoritative.

11. Covariate imbalance

Women who take oral contraceptives appear to have a higher risk of cervical cancer. But covariate imbalance might be producing an artificial rise in cancer rates for this group. Women who take oral contraceptives behave, as a group, differently than other women.

12. Covariate imbalance

For example, women who take oral contraceptives have a larger number of pap smears. This is probably because these women visit their doctors more regularly in order to get their prescriptions refilled and therefore have more opportunities to be offered a pap smear. This difference could lead to an increase in the number of detected cancer cases. Perhaps the other women have just as much cancer, but it is more likely to remain undetected.

13. Covariate imbalance

- There are many other variables that influence the development of cervical cancer: age of first intercourse, number of sexual partners, use of condoms, and smoking habits. If women who take oral contraceptives differ in any of these lifestyle factors, then that might also produce a difference in cervical cancer rates.
- The possibility that oral contraceptives causes an increase in the risk of cervical cancer is quite complex; a good summary of all the issues involved is available at:
  - www.jhuccp.org/pr/a9/a9chap5.shtml.

14. Randomization

One way to avoid most of the problems with imbalanced covariates is to use randomization. Randomization is the assignment of treatment groups through the use of a random device, like the flip of a coin or the roll of a die, or numbers randomly generated by a computer. Randomization is not always possible, practical, or ethical. But when you can use randomization, it greatly adds to the credibility of the research study.

15. Randomization

In a randomized study, the researchers have a high degree of control over the patients. They decide who gets what. This is a hallmark of a randomized design and it only can occur when the patients and/or their doctors have no say in the assignment. This is an incredible gift that patients in a research study offer you. They sacrifice their ability to choose between two therapies and instead let that choice be decided by the flip of a coin.

16. Randomization

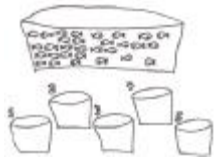
Randomization helps ensure that both measurable and immeasurable factors are balanced out across both the standard and the new therapy, assuring a fair comparison. Used correctly, it also guarantees that no conscious or subconscious efforts were used to allocate subjects in a biased way.

17. Randomization

Randomization relies on the law of large numbers. With small sample sizes, covariate imbalance may still sneak in. A study examining the probability of covariate imbalance (Hsu 1989) showed that total sample sizes less than 10 could have a 50% chance or higher of having a categorical covariate with levels twice as large in one group than the other. This study also showed that total sample sizes of 40 or greater would have very little chance of such a serious imbalance.

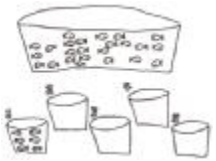
18. A fishy story about randomization

Study of fish exposed to various levels of a chemical



19. A fishy story about randomization

The researchers caught the first 20 fish and put them in the first tank,...



20. A fishy story about randomization

then put the next 20 fish in a second tank,...

21. A fishy story about randomization

the next 20 in the third tank,...

22. A fishy story about randomization

the next 20 in the fourth tank,...

23. A fishy story about randomization

and the last twenty in the fifth tank.

24. A fishy story about randomization

When they were done, they found that mortality was related not to the chemical concentration but to the order in which the tanks were filled.

25. Concealed allocation

Another important aspect of randomization is concealed allocation, which is withholding the randomization list from those involved with recruiting subjects. This concealment occurs until after subjects agree to participate and the recruiter determines that the patient is eligible for the study. Only then is a sealed envelope opened that reveals the treatment status. Concealed allocation can also be done through a special phone number that the doctor calls to discover the treatment status.

26. Concealed allocation

If the randomization list is not concealed, doctors have the ability to consciously or unconsciously influence the composition of the groups. They can do this by applying exclusion criteria differentially or by delaying entry of a certain healthier (or unhealthier) subject so he/she gets into the 'desirable' group. Unblinded allocation schemes tend, on average to overstate the effectiveness of the new therapy by 30–40% (Schulz 1996).

27. Ethical and practical constraints on randomization

There are many situations where randomization is not practical or possible. Sometimes patients have a strong preference for one particular treatment and would not consider the possibility of being randomized into a different treatment. Surgery is one area with strong patient preferences especially for newer approaches like laparoscopic surgery (Lefering 2003).

28. Ethical and practical constraints on randomization

Randomization is also problematic for interventions that are already known to be effective. While further research would help better define these advantages, you cannot ask half of your patients to sacrifice the benefits of the new intervention.

29. Ethical and practical constraints on randomization

Randomization also does not work when you are studying noxious agents, like second-hand cigarette smoke or noisy workplaces. It would be unethical to deliberately expose people to any of these agents, so we have use non-randomized studies of people who are unfortunate enough to be trapped in settings with noxious agents.

30. Ethical and practical constraints on randomization

Sometimes researchers just do not want to go to the effort of randomizing. If you assign the treatment or therapy, rather than letting the patients and their doctors choose, you have to expend a lot of energy. Is it worth the effort? It is usually faster and cheaper to use existing nonrandomized databases. You get a lot larger sample size for your money. Depending on the situation, that might be enough to counterbalance the advantages of randomization.

31. Variations on randomization

There are three variations to randomization where the researchers have control over treatment assignment, but they use something other than a table of random numbers for the assignment. The first approach, minimization, is a credible and reasonable choice, but the other two approaches, alternating assignment and haphazard assignment, do not have much to recommend them.

32. Variations on randomization

An alternative, when the researchers have sufficient control, is to allocate the assignments so that at each step, the covariate imbalance is minimized.

So if the treatment group has a slight surplus of older patients and the next patient to join the study is also older than average, then that patient would be assigned to the control group so as to reduce the age discrepancy.

33. Variations on randomization

Another approach used in place of randomization is to alternate the assignment, so that every even patient is in the treatment group and every odd patient is in the control group. Alternating assignment was popular in trials before World War II; it was felt that researchers would not understand and not tolerate randomization (Yoshioka 1998).

34. Variations on randomization

Alternating assignment seems on the surface to be a good approach, but it can sometimes lead to trouble. This is especially true when one patient has a direct or indirect influence on the next patient. You may have seen this level of influence if you grow vegetables in a garden. If you have a row of cabbages, for example, you will often see a pattern of big cabbage, little cabbage, big cabbage, little cabbage, etc.

35. Variations on randomization

What happens, if the cabbages are planted a bit too closely, is that one of the cabbages will grow just a bit faster at first. It will extend into the neighboring cabbage's territory, stealing some of the nutrients and water, and thus growing even faster at the expense of the neighbor. If you assigned a fertilizer to every other cabbage, you would probably see an artificial difference because of the alternating pattern in growth within a row.

36. Variations on randomization

Haphazard assignment uses some arbitrary value like a birthdate or social security number to assign patients to groups. Often it is the evenness/oddness of the arbitrary number that determines the treatment assignment. For example, patients born on even-numbered dates would be assigned to the treatment group and those born on odd-numbered dates would be assigned to the control group. An arbitrary or haphazard number is never going to be as good as a purely random number. The haphazard assignment will always cast a shadow of doubt over the research study.

37. Practice exercises

For each of the following abstracts, randomization was NOT used. Explain why it would be impractical or unethical to conduct a randomized experiment in each of these settings.

38. Conclusion

Randomization is the use of a random device to assign patients to a treatment group or control group. When the sample size is sufficiently large, randomization prevents covariate imbalance in your experiment. Randomization is not practical if patients have a strong preference for a particular treatment and is unethical if it forces some patients to endure a harmful exposure.

39. Repeat of pop quiz #1

When the demographic profile of the patients in your treatment group differ sharply from the profile of patients in your control group, you have:

1. covariate imbalance,
2. observational data,
3. response bias,
4. spectrum bias,
5. stratified data,
6. don't know/not sure

40. Repeat of pop quiz #2

Randomization is not practical:

1. when doctors believe that the new treatment is superior to the current standard
2. when patients have a strong preference for a particular treatment
3. when the experiment requires deliberate exposure of patients to something that is known to be harmful
4. randomization is impractical for all of the above situations
5. randomization can be applied easily in all of the above situations
6. don't know/not sure

41. Repeat of pop quiz #3

The following approaches are credible alternatives to randomization:

1. alternating between treatment and control
2. assigning all new patients to the treatment group and choosing controls from a medical database
3. assigning treatment group on the basis of the last digit of your birthday
4. letting the doctor choose whether a patient gets into the treatment group or the control group
5. none of these approaches is as effective as randomization
6. don't know/not sure

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

- 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.

43. 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.

44. 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

45. Pop quiz #4

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

46. Pop quiz #5

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

47. Pop quiz #6

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

48. 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

49. 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.

50. 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?

51. 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.

52. 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.

53. 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.

54. 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.

55. 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.

56. 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.

57. 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..

58. 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.

59. 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.

60. 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.

61. 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?

62. 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).

63. 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.

64. 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.

65. 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.

66. 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.

67. 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.

68. 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.

69. 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.

70. 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.

71. 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.

72. 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?

73. 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.

74. 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

75. 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.

76. 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.

77. Repeat of pop quiz #4

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

78. Repeat of pop quiz #5

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

79. Repeat of pop quiz #6

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

80. P-values and confidence intervals

- In this hour, you will learn how to:
  - distinguish between statistical significance and clinical significance;
  - define and interpret p-values; and
  - describe the advantages of confidence intervals.

81. Outline

1. Pop quiz
2. Definitions
3. What is a p-value?
4. Practice exercises
5. What is a confidence interval?
6. More practice exercises
7. Repeat of pop quiz

Please feel free to ask questions at anytime

82. Pop quiz #7

A research paper computes a p-value of 0.45. How would you interpret this p-value?

1. Strong evidence for the null hypothesis
2. Strong evidence for the alternative hypothesis
3. Little or no evidence for the null hypothesis
4. Little or no evidence for the alternative hypothesis
5. None of these answers are correct.
6. I do not know the answer.

83. Pop quiz #8

A research paper computes a confidence interval for a relative risk of 0.82 to 3.94. What does this confidence interval tell you.

1. The result is statistically significant and clinically important.
2. The result is not statistically significant, but is clinically important.
3. The result is statistically significant, but not clinically important.
4. The result is not statistically significant, and not clinically important.
5. The result is ambiguous.
6. I do not know the answer.

84. Definitions

- Population

85. Definitions

- Population
- Sample

86. Definitions

- Population
- Sample
- Type I error

87. Definitions

- Population
- Sample
- Type I error
- Type II error

88. What is a p-value?

- A p-value is a **measure of how much evidence we have against the null hypothesis**.
- The smaller the p-value, the more evidence we have against H0.

89. What is a p-value?

- Suppose that a drug company alleges that **only 50% of all patients who take a certain drug will have an adverse event of some kind**. You believe that the adverse event rate is much higher. **In a sample of 12 patients, all twelve have an adverse event**.
- P-value = 0.000244.

90. What is a p-value?

A **small p-value** means **lots of evidence against the null hypothesis**.

A **large p-value** means **little or no evidence against the null hypothesis**.

A p-value is NOT the probability that the null hypothesis is true.

91. What is a p-value?

A large p-value should not automatically be construed as evidence in support of the null hypothesis.

Instead of just the p-value, look for

1. a **power calculation**; and/or
2. a **confidence interval**.

Also be cautious about a small p-value.

92. Practice exercise: interpret the p-values shown in the abstract.

Use the PICO format.

- P = Patient population
- I = Intervention
- C = Control/Comparison group
- O = Outcome

93. What is a confidence interval?

- A confidence interval is a range of values that tries to quantify **uncertainty associated with the sampling process**.
- Consider it as a **range of plausible values**.

94. What is a confidence interval?

- Wide interval = poor precision
- Narrow interval = good precision
- It is unethical to conduct research if you know that your confidence interval will be so wide as to be uninformative.

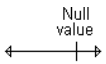
95. What is a confidence interval?

- Consider a recent study of homoeopathic treatment of pain and swelling after oral surgery.
  - P= patients undergoing oral surgery
  - I=homeopathic treatment
  - C=patients taking placebo.
  - O=swelling after 3 days.
- Homoeopathy led to 1 mm less swelling on average.
- The 95% confidence interval, however, ranged from -5.5 to 7.5 mm.

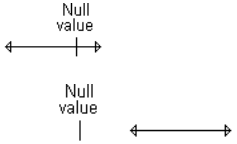
96. What is a confidence interval?

- Look for two things:
  1. Does the interval contain a value that implies no change or no effect?

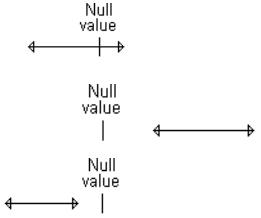
97. No change,…



98. No change, a positive change,…



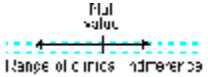
99. No change, a positive change, and a negative change



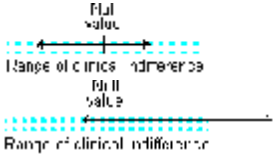
100. What is a confidence interval?

- Look for two things:
  1. Does the interval contain a value that implies no change or no effect?
  2. Does the interval lie entirely inside the range of clinical indifference?

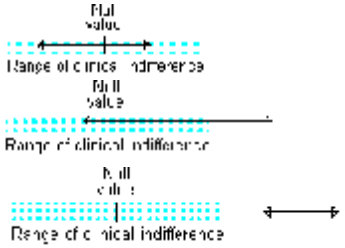
101. A true null finding,…



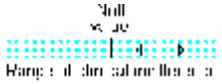
102. A true null and an ambiguous finding, …



103. A true null, an ambiguous, and a true positive finding.



104. Statistical significance without practical significance



105. Practice exercise: interpret the confidence interval(s) shown in the abstract.

106. Repeat of pop quiz #7

A research paper computes a p-value of 0.45. How would you interpret this p-value?

1. Strong evidence for the null hypothesis
2. Strong evidence for the alternative hypothesis
3. Little or no evidence for the null hypothesis
4. Little or no evidence for the alternative hypothesis
5. More than one answer above is correct.
6. I do not know the answer.

107. Repeat of pop quiz #8

A research paper computes a confidence interval for a relative risk of 0.82 to 3.94. What does this confidence interval tell you.

1. The result is statistically significant and clinically important.
2. The result is not statistically significant, but is clinically important.
3. The result is statistically significant, but not clinically important.
4. The result is not statistically significant, and not clinically important.
5. The result is ambiguous.
6. I do not know the answer.

108. How bad is it, really? Measures of risk.

- The odds ratio and the relative risk are both measures of risk used for binary outcomes, but sometimes they can differ markedly from one another. The relative risk offers a more natural interpretation, but certain research designs preclude its computation.

109. Objectives

In this class you will learn how to:

- compute an odds ratio and a relative risk from a two by two table;
- list the types of research designs where the relative risk should not be computed, and

110. Sources

Part of the material for this webinar comes from:

- Simon SD. Understanding the odds ratio and the relative risk. J Androl. 2001 Jul-Aug;22(4):533-6.
- Stats: Odds ratio versus relative risk (January 9, 2001).
  - http://www.childrens-mercy.org/stats/journal/oddsratio.asp

111. Very bad joke

A doctor is advising her patient about the risks of an upcoming surgery. She warned that the probability that the patient would die during surgery was 60%. Then she looked up an said, no wait, the risk is twice as big in your demographic group. The chances that you will die during surgery is actually 120%. The patient seemed a bit confused. I know what a 100% risk of mortality would be—I'm a goner. But what would a 120% risk of mortality be? The doctor replied, that is a fate worse than death.

112. Pop quiz #9

A relative risk should not be computed for the following design because the prevalence of the disease is artificially constrained.

1. Case-control design
2. Cohort design
3. Cross-sectional design
4. Historical control design
5. Don't know/Not sure

113. Pop quiz #10

The odds ratio and the relative risk are close to one another when

1. The prevalence of the disease is low
2. The prevalence of the disease is high
3. The sample size is small
4. The sample size is large
5. Don't know/Not sure

114. What are odds ?

If you head south from Kansas City on Highway 71, you will encounter a town called "Peculiar". This town is very proud of its name and has a sign which says "Welcome to Peculiar, where the odds are with you."

Mathematicians and gamblers use odds frequently but the concept may be alien to most of the rest of the public. Odds is the ratio of successes to failures.

115. What are odds ?

"If there is a 50-50 chance that something will go wrong, then nine times out of ten it will." (Paul Harvey).

In this silly example a 50-50 chance means one success for every failure or 1 to 1 odds. This is sometimes called even odds.

Nine times out of ten means one success for every nine failures or one to nine odds.

116. What are odds ?

To be perfectly accurate, you should specify whether you are talking about the odds of success or the odds of failure, but in most setting, it should be obvious from the context.

If your odds of winning the lottery are a million to one, that means either that:

- One million people win for every person that loses, or
- One person winds for every million that lose.

117. What are odds ?

If you know the probability of a success, you can calculate the odds using the formula

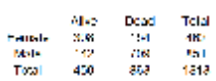
- Odds = prob / (1- prob).
- For example, a probability of 0.25 corresponds to an odds of 0.25 / (1-0.25) = 0.25 / 0.75 = 1 / 3. This means that for every single success, there are three failures.

If you know the odds, then you can calculate the probability of success using the formula

- Prob = Odds / (1 + Odds).
- For example, if the odds are 3 to 1, then prob = 3 / (1 + 3) = 3 / 4.

118. Odds ratio/relative risk

Consider the following data on survival of passengers on the Titanic. Clearly, a male passenger on the Titanic was more likely to die than a female passenger. But how much more likely? You can compute the odds ratio or the relative risk to answer this question.

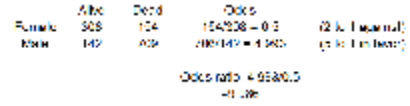


119. Odds ratio/relative risk

The odds ratio compares the relative odds of death in each group.

- For females, 2 to 1 odds against dying
- For males, almost 5 to 1 in favor of death

The odds ratio is approximately 10.

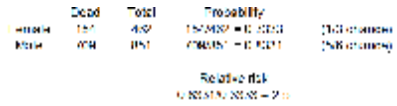


120. Odds ratio/relative risk

The relative risk (sometimes called the risk ratio) compares the probability of death in each group rather than the odds.

- The females probability of death is 1/3 (2/6).
- The male probability of death is 5/6.

The relative risk of death is 2.5



121. Odds ratio/relative risk

There is quite a difference. Both measurements show that men were more likely to die. But the odds ratio implies that men are much worse off than the relative risk. Which number is a fairer comparison?

122. Odds ratio/relative risk

There are three issues here:

1. The relative risk measures events in a way that is interpretable and consistent with the way people really think.
2. The relative risk, though, cannot always be computed in a research design.
3. Also, the relative risk can sometimes lead to ambiguous and confusing situations.

123. Repeat of Pop quiz #9

A relative risk should not be computed for the following design because the prevalence of the disease is artificially constrained.

1. Case-control design
2. Cohort design
3. Cross-sectional design
4. Historical control design
5. Don't know/Not sure

124. Repeat of Pop quiz #10

The odds ratio and the relative risk are close to one another when

1. The prevalence of the disease is low
2. The prevalence of the disease is high
3. The sample size is small
4. The sample size is large
5. Don't know/Not sure

125. Putting it all together: Meta-analyses and systematic overviews

This class helps you assess the quality of a systematic overview or meta-analysis. In this class you will learn how to: recognize sources of heterogeneity in meta-analysis; identify and avoid problems with publication bias; and explain the ethical concerns with failure to publish and with duplicate publication.

This material is derived mainly from Chapter 5 of **Statistical Evidence in Medical Trials**.

126. Outline

1. Pop quiz
2. Introduction and motivating example
3. Were apples combined with oranges?
4. Were some apples left on the tree?
5. Repeat of pop quiz

Note: there are also issues involving study quality (were all of the apples rotten?) and practical significance (did the pile of apples amount to more than just a hill of beans?) but we will not have time to discuss those issues today.

127. Pop quiz #11

A funnel plot is useful for assessing

1. heterogeneity
2. publication bias
3. study quality
4. not sure/don't know

128. Pop quiz #12

Cochran's Q and I^2 are measures of

1. heterogeneity
2. publication bias
3. study quality
4. not sure/don't know

129. Putting it all together: Meta-analyses and systematic overviews

- When there are multiple research studies evaluating a new intervention, you need to find a way to assess the cumulative evidence of these studies. You can do this informally, but medical researchers now use a formal process, known as meta-analysis. Meta-analysis, involves the quantitative pooling of data from two or more studies.

130. Introduction

- More recently, another term, systematic overview, has come into favor. A systematic overview involves the careful review and identification of all research studies associated with a topic, but it may or may not end up pooling the results of these studies. So meta-analysis represents a subset of all the systematic overviews.

131. Motivating example

- In 1992, the British Medical Journal published a controversial meta-analysis. This study (Carlsen 1992) reviewed 61 papers published from 1938 and 1991 and showed that there was a significant decrease in sperm count and in seminal volume over this period of time. For example, a linear regression model on the pooled data provided an estimated average count of 113 million per ml in 1940 and 66 million per ml in 1990.

132. Motivating example

- Several researchers (Olsen 1995; Fisch 1996) noted heterogeneity in this meta-analysis, a mixing of apples and oranges. Studies before 1970 were dominated by studies in the United States and particularly studies in New York. Studies after 1970 included many other locations including third world countries. Thus the early studies were US apples. The later studies were international oranges. There was also substantial variation in collection methods, especially in the extent to which the subjects adhered to a minimum abstinence period.

133. Motivating example

- The original meta-analysis and the criticisms of it highlight both the greatest weakness and the greatest strength of meta-analysis. Meta-analysis is the quantitative pooling of data from studies with sometimes small and sometimes large disparities. Think of it as a multicenter trial where each center gets to use its own protocol and where some of the centers are left out.

134. Motivating example

- On the other hand, a meta-analysis lays all the cards on the table. Sitting out in the open are all the methods for selecting studies, abstracting information, and combining the findings. Meta-analysis allows objective criticism of these overt methods and even allows replication of the research.

135. Motivating example

- Contrast this to an invited editorial or commentary that provides a subjective summary of a research area. Even when the subjective summary is done well, you cannot effectively replicate the findings. Since a subjective review is a black box, the only way, it seems, to repudiate a subjective summary is to attack the messenger.

136. Were apples combined with oranges?

- Meta-analyses should not have too broad an inclusion criteria. Including too broad a range of studies can lead to problems with heterogeneity (mixing apples and oranges).

137. First example of heterogeneity

- In a meta-analysis looking at antiretroviral combination therapy (Jordan 2002), both short-term and long-term outcomes were examined. A plot of duration of trial versus the log odds ratio showed that shorter duration trials of zidovudine had substantial evidence of effect (odds ratios much smaller than 1) but that the largest duration studies had little or no evidence of effect (odds ratios very close to 1).

138. Second example of heterogeneity

- Example: In a meta-analysis looking at dust mite control measures to help asthmatic patients (Gotzsche 1998), the studies exhibited heterogeneity across several factors.

139. Second example of heterogeneity

- Type of intervention:
  - six examined chemical interventions,
  - thirteen examined physical interventions,
  - four examined a combination approach.
- Research design:
  - nine of these trials were crossovers,
  - fourteen had a parallel control group.
- Blinding
  - seven studies had no blinding,
  - three studies had partial blinding,
  - thirteen studies used a double blind.

140. Second example of heterogeneity

- Age of patients
  - nine studies the average age of the patients was only 9 or 10 years,
  - nine other studies had an average age of 30 or more,
  - five studies had a greater mix of ages.
- Duration
  - eleven studies lasted eight weeks or less,
  - five studies lasted a full year,
  - seven studies had an intermediate duration

141. Possible sources of heterogeneity

- This list is adapted from Horwitz 1987
  - Inclusion/exclusion criteria
  - Geographical limitations
  - Independent versus matched controls
  - Dose/timing of drug administration
  - Length of follow-up
  - Drop-out rates
  - Allowable physician discretion
  - Outcome measure

142. Measuring heterogeneity

- Cochran's Q: A value close to the number of studies is good, but a value much larger is bad.
- I^2: ranges between 0% and 100%, larger values indicating greater heterogeneity.
- Many researchers recommend a qualitiative assessment of heterogeneity.

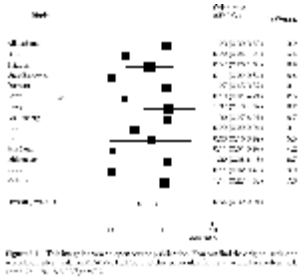
143. Forest plot

- The forest plot provides a graphical summary of the studies. This plot can be used to evaluate heterogeneity.
  - Location of square represents the point estimate,
  - Size of square represents weight associated with that estimate, and
  - Lines drawn to upper and lower confidence limits.

144. Forest plot

- Look for marked departures from a normal random scatter:
  - Most studies cluster together, but one or two outlying studies (but okay if outlying studies have small sample sizes).
  - Bimodal patterns (e.g., half the studies show a strong effect, half show little or no effect).

145. Forest plot example



146. Handling heterogeneity

- There are several common approaches for coping with heterogeneity
  - Strict inclusion/exclusion criteria
  - Sensitivity/subgroup analysis
  - Meta-regression
  - "Just say no"

147. Example of strict inclusion/exclusion criteria

- A meta-analysis of topical NSAIDs for musculoskelatal pain (Mason 2004) identified 60 target papers, but for 12 of the papers, there was no data that could be extracted for a meta-analysis. An additional 23 studies were removed based on the following exclusion criteria:
  - no studies for mouth or eye diseases;
  - no studies where fewer than 10 patients were randomized to the treatment;
  - no studies where treatment occurred less frequently than daily;
  - no observational studies; and
  - no unblinded studies.

148. First example of strict sensitivity/subgroup analysis

- In a study of extra corporeal shock wave therapy for plantar heel pain (Thomson 2005), six studies met the researchers inclusion criteria, but one study did not report a standard deviation for the outcome measure. The authors were forced to estimate what the standard deviation should be for this study. As a quality check, they also ran a meta-analysis without this study and found that a modest effect in favor of the therapy was no longer statistically significant.

149. Second example of strict sensitivity/subgroup analysis

- In a study of topical NSAIDs for osteoarthritis and tendinitis (Mason 2004), researchers identified 25 trials relating to efficacy or harm, including 14 placebo-controlled trials. These studies varied substantially in
  - quality scores,
  - number of patients studied,
  - type of outcome measure (physician determined versus self report) and
  - condition being treated (osteoarthritis versus other musculoskeletal conditions).
- But when the results were tabulated separately for low and high quality scores, small and large studies, etc., there were no statistically significant differences.

150. Meta-regression

- You can use meta-regression to try to adjust for heterogeneity in a metaanalysis. In meta-regression, each study becomes a data point, and various study characteristics, such as the severity of illness at baseline, the dose of the medication being given, etc. become independent variables. This is an approach that would work very similarly to the adjustment for covariates in a regression model. The result, meta-regression, is an area of active research and looks to be a promising way to handle heterogeneity in a more rigorous fashion.

151. Example of meta-regression

- In a study of diagnostic tests for endometrial hyperplasia (Clark 2004), researchers identified 27 studies using miniature endometrial biopsy devices or ultrasonography. In some of the studies, verification of the diagnosis was delayed by more than 24 hours. Although the ability to discriminate between diseased and healthy patients was present in most studies, the discriminatory power, as measured by the diagnostic odds ratio was four times weaker among studies with delayed verification than studies with no delay.

152. "Just say no"

- If the degree of heterogeneity is too extreme, you should just say no and refuse to run a meta-analysis. You can still discuss the studies in a qualitative fashion, but do not try to compute an overall estimate of effect because that estimate would be meaningless.

153. Example of "Just say no"

- In a systematic review of beta-2 agonists for treating chronic obstructive pulmonary disease (Husereau 2004), researchers identified 12 studies. But the authors could not pool the results because they
  - "found that even commonly measured outcomes, such as FEV1, could not be combined by meta-analysis because of differences in how they were reported. For example, in the six trials comparing salmeterol with placebo, FEV1 was reported as a mean change in percent predicted, a mean change overall, a mean difference between trial arms, no difference (without data), baseline and overall FEV1 (after 24 hrs without medication) and as an 0 to 12 hour area-under-the-curve (FEV1-AUC) function. We were not successful in obtaining more data from study authors. We also had concerns about the meta-analysis of data from trials of parallel and crossover design and differences in spirometry protocols including allowable medications. Therefore, we decided on a best evidence synthesis approach instead."

154. Were some apples left on the tree?

- Publication bias: the tendency on the parts of investigators, reviewers, and editors to submit or accept manuscripts for publication based on the direction or strength of the study findings. There is solid empirical evidence (e.g., Dickersin 1990) that negative studies are less likely to be published.

155. Ethical concerns with failure to publish

- Researchers who fail to publish their research, however, are behaving unethically (Chalmers 1990). These research studies almost always use human volunteers. These volunteers might be participating because they need the money or perhaps they are curious about the scientific process. But many of them volunteer because they want to help others who have the same disease or condition. These volunteers submit themselves willingly to some level of inconvenience, and possibly additional pain and risk. If you ask these volunteers to make this sacrifice, but you do not publish, you have abused their good will.

156. Should unpublished studies be included?

- The inclusion of unpublished studies, however, is controversial. At least one reference (Cook 1993), has argued that unpublished studies have failed to meet a basic quality screen, the peer review process. Including studies that have not been peer reviewed will lower the overall quality of the meta-analysis. This opinion, however, is in the minority, and most experts in meta-analysis suggest that you include unpublished studies if you can find them. Failure to include unpublished studies can lead to serious bias.

157. Duplicate publication

- Duplicate publication is the flip side of the same coin. The data from some studies may appear twice (or even three times) in the peer-reviewed literature, without appropriate attribution. If you double count these studies accidentally, you will produce a biased result because duplicate publications are more likely to be positive.

158. Ethical concerns with duplicate publication

- Duplicate publication raises serious ethical issues:
  - Violation of copyright
  - Padding of resumes
  - Abuse of volunteer services of referees/editors
  - Taking page space away from other deserving publications.
- There are reasonable justifications for duplicate publication, such as translating a publication into English to insure a wider dissemination of the research findings. These exceptions, however, would always have an obvious citation of the original source.

159. Example of duplicate publication

- In 84 studies of the effect of ondansetron on postoperative emesis, 14 (17%) were second or even third time publications of the same data-set (Tramer 1997). The duplicate studies had much larger effects and adding the duplicates to the originals produced an overestimation of treatment efficacy of 23%. Tracking down the duplicate publications was quite difficult. More than 90% of the duplicate publications did not crossreference the other studies. Four pairs of identical trials were published by completely different authors without any common authorship.

160. Don't rely exclusively on Medline

- While a Medline search is a very effective way to identify published research, it should not be the only source of publications for a meta-analysis. There are many important journals which are not included in Medline. It is hard to get an accurate count of how many journals do NOT appear in Medline, but the numbers appear to be substantial. You might suspect that journals indexed by Medline are more prestigious and more likely to publish positive findings than other journals, but I am unaware of any data to substantiate this. Still, a search that included only Medline articles would be considered grossly inadequate in most situations.

161. Don't rely English-language only publications

- Some meta-analyses restrict their attention to English language publications only. While this may seem like a convenience, in some situations, researchers might tend to publish in an English language journal for those trials which are positive, and publish in a (presumably less prestigious) native language journal for those trials which are negative (Gregoire 1995). Restrictions to English language only publications is especially troublesome for complementary and alternative medicine, since so much of this research appears in non-English language journals.

162. Using a funnel plot to detect publication bias

- The most common approach to evaluate publication bias is to use a funnel plot. The funnel plot displays
  - the results of the individual studies (e.g. the log odds ratio) on the horizontal axis,
  - the size of the study (or sometimes the standard error of the study) on the vertical axis.
- Often a reference line is drawn at the value that represents no effect.

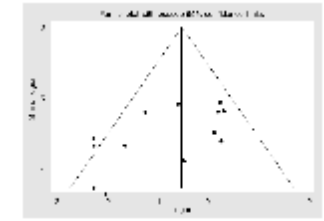
163. Using a funnel plot to detect publication bias

- The rationale behind this plot
  - big studies get published no matter what the result
  - smaller studies are subject to publication bias
- If there is no publication bias, then the funnel plot should show symmetry for both small sample sizes and large sample sizes, though you should expect to see less variation as the sample size increases. This leads to a funnel shape.

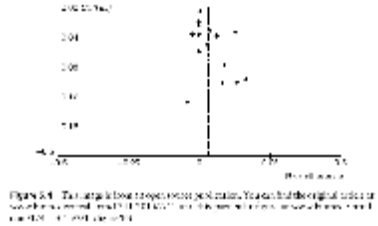
164. Example of a funnel plot

- The rationale behind this plot
  - big studies get published no matter what
  - smaller studies are subject to publication bias
- If there is no publication bias, then the funnel plot should show symmetry for both small sample sizes and large sample sizes, though you should expect to see less variation as the sample size increases. This leads to a funnel shape.
- Although funnel plots are commonly used, there is some suggestion that they are not effective.

165. Funnel plot example showing symmetry



166. Funnel plot example showing possible publication bias



167. How to avoid or minimize problems with publication bias

1. Use several bibliographic databases, not just Medline.
2. Search through registries of clinical trials.
3. Hand search through specialized journals
4. Examine bibliographies of articles found on first pass through.
5. Examine “gray literature” (presentations, dissertations, etc.)
6. Send out letter to prominent leaders in the area asking for help.

168. Repeat of pop quiz #11

A funnel plot is useful for assessing

1. heterogeneity
2. publication bias
3. study quality
4. not sure/don’t know

169. Repeat of pop quiz #12

Cochran's Q and I^2 are measures of

1. heterogeneity
2. publication bias
3. study quality
4. not sure/don't know

170. 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

PLoS One. 2011 Mar 29;6(3):e18227.

Psoriasis and hypertension severity: results from a case-control study.

Armstrong AW, Lin SW, Chambers CJ, Sockolov ME, Chin DL.

Department of Dermatology, University of California Davis, Sacramento, California, United States of America. aprilarmstrong@post.harvard.edu

Abstract

BACKGROUND: Epidemiologic studies have provided new insights into the association between psoriasis and cardiovascular diseases. Previous population studies have examined hypertension frequency in psoriasis patients. However, the relationship between severity of hypertension and psoriasis has not been characterized.

OBJECTIVE: We sought to investigate whether patients with psoriasis have more difficult-to-manage hypertension compared to non-psoriatic hypertensive patients.

APPROACH: We performed a case-control study using the University of California Davis electronic medical records. The cases were defined as patients diagnosed with both psoriasis and hypertension, and controls were defined as patients with hypertension and without psoriasis. In this identified population, 835 cases were matched on age, sex, and body mass index (BMI) to 2418 control patients.

KEY RESULTS: Treatment with multiple anti-hypertensives was significantly associated with the presence of psoriasis using univariate ($p < 0.0001$) and multivariable analysis, after adjusting for diabetes, hyperlipidemia, and race ($p < 0.0001$). Compared to hypertensive patients without psoriasis, psoriasis patients with hypertension were 5 times more likely to be on a monotherapy antihypertensive regimen (95% CI 3.607.05), 9.5 times more likely to be on dual antihypertensive therapy (95% CI 6.68-13.65), 16.5 times more likely to be on triple antihypertensive regimen (95% CI 11.01-24.84), and 19.9 times more likely to be on quadruple therapy or centrally-acting agent (95% CI 10.58-37.33) in multivariable analysis after adjusting for traditional cardiac risk factors.

CONCLUSIONS: Psoriasis patients appear to have more difficult-to-control hypertension compared to non-psoriatic, hypertensive patients.

PMID: 21479272 [PubMed - indexed for MEDLINE] PMCID: PMC3066207 Free PMC Article

Ann Indian Acad Neurol. 2012 Jan;15(1):19-22.

Low cholesterol as a risk factor for primary intracerebral hemorrhage: A case-control study.

Valappil AV, Chaudhary NV, Praveenkumar R, Gopalakrishnan B, Girija AS.

Department of Neurology, Malabar Institute of Medical Sciences, Calicut, Kerala, India.

Abstract

INTRODUCTION: An inverse association between serum cholesterol and the risk of hemorrhagic stroke has been noted in epidemiological studies. We performed a case-control study to assess the relationship between primary intracerebral hemorrhage (ICH) and low serum cholesterol.

MATERIALS AND METHODS: Prospectively recruited fully evaluated patients with ICH were compared with a control group based in a primary care practice, i.e. age- and sex-matched individuals attending the routine preventive health check-up. Low cholesterol was defined by the sex-specific lowest quintile of the population.

RESULTS: The proportion of ICH patients with low cholesterol was significantly higher than the controls (68% vs. 43%). Mean total cholesterol was also signficantly low in ICH patients compared with controls (177 mg/dL vs. 200 mg/dl; P-value = 0.0006). Low-density lipoprotein cholesterol (LDL-c) and triglycerides were also significantly low in ICH patients compared with controls. Mean LDL-C in the ICH patient group was 114 mg/dL, whereas it was 128.5 mg/dL in the control group (P-value = 0.016). There was no significant difference in the high-density lipoprotein (HDL) levels in both groups. In a subgroup analysis, both men and women in the ICH group had a significantly low mean cholesterol compared with the control group. Although lower mean cholesterol was seen in both young and older individuals in the ICH group than in controls, the difference was significant only in the older group (age >45 years). In multivariate analysis, presence of low cholesterol remained a significant predictor of hemorrhage. The odds ratio of low cholesterol in the hemorrhage cases was 2.75 (95% CI = 1.44-5.49) unadjusted and 2.15 (1.13-4.70) adjusted for age and hypertension.

CONCLUSIONS: This study confirms an increased risk of primary ICH associated with low cholesterol both in men and women, especially in older individuals.

PMID: 22412267 [PubMed - in process] PMCID: PMC3299064

J Foot Ankle Res. 2012 Apr 16;5(1):9. [Epub ahead of print]

Comparison of shoe-length fit between people with and without diabetic peripheral neuropathy: a case-control study.

McInnes AD, Hashmi F, Farndon LJ, Church A, Haley M, Sanger DM, Vernon W.

Abstract

ABSTRACT:

BACKGROUND: Amongst the many identified mechanisms leading to diabetic foot ulceration, ill-fitting footwear is one. There is anecdotal evidence that people with diabetic peripheral neuropathy wear shoes that are too small in order to increase the sensation of fit. The aim of this study was to determine whether people with diabetic sensory neuropathy wear appropriate length footwear.

METHODS: A case-control design was used to compare internal shoe length and foot length differences between a group of people with diabetes and peripheral sensory neuropathy and a group of people without diabetes and no peripheral sensory neuropathy. Shoe and foot length measurements were taken using a calibrated Internal Shoe Size Gauge(R) and a Brannock Device(R), respectively.

RESULTS: Data was collected from 85 participants with diabetes and 118 participants without diabetes. The mean difference between shoe and foot length was not significantly different between the two groups. However, a significant number of participants within both groups had a shoe to foot length difference that lay outside a previously suggested 10 to 15 mm range. From the diabetic and non-diabetic groups 82% (70/85) and 66% (78/118), respectively had a foot to shoe length difference outside this same range.

CONCLUSIONS: This study shows that although there is no significant difference in shoe-length fit between participants with and without neuropathy, a significant proportion of these populations wear shoes that are either too long or too short for their foot length according to the 10 to 15 mm value used for comparison. The study has highlighted the need for standardised approaches when considering the allowance required between foot and internal shoe length and for the measurement and comparison of foot and shoe dimensions.

PMID: 22507446 [PubMed - as supplied by publisher] Free full text

LinkOut - more resources

PLoS Med. 2011 Dec;8(12):e1001141. Epub 2011 Dec 6.

Rotating night shift work and risk of type 2 diabetes: two prospective cohort studies in women.

Pan A, Schernhammer ES, Sun Q, Hu FB.

Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA.

Abstract

BACKGROUND: Rotating night shift work disrupts circadian rhythms and has been associated with obesity, metabolic syndrome, and glucose dysregulation. However, its association with type 2 diabetes remains unclear. Therefore, we aimed to evaluate this association in two cohorts of US women.

METHODS AND FINDINGS: We followed 69,269 women aged 42-67 in Nurses' Health Study I (NHS I, 1988-2008), and 107,915 women aged 25-42 in NHS II (1989-2007) without diabetes, cardiovascular disease, and cancer at baseline. Participants were asked how long they had worked rotating night shifts (defined as at least three nights/month in addition to days and evenings in that month) at baseline. This information was updated every 2-4 years in NHS II. Self-reported type 2 diabetes was confirmed by a validated supplementary questionnaire. We documented 6,165 (NHS I) and 3,961 (NHS II) incident type 2 diabetes cases during the 18-20 years of follow-up. In the Cox proportional models adjusted for diabetes risk factors, duration of shift work was monotonically associated with an increased risk of type 2 diabetes in both cohorts. Compared with women who reported no shift work, the pooled hazard ratios (95% confidence intervals) for participants with 1-2, 3-9, 10-19, and ≥20 years of shift work were 1.05 (1.00-1.11), 1.20 (1.14-1.26), 1.40 (1.30-1.51), and 1.58 (1.43-1.74, p-value for trend <0.001), respectively. Further adjustment for updated body mass index attenuated the association, and the pooled hazard ratios were 1.03 (0.98-1.08), 1.06 (1.01-1.11), 1.10 (1.02-1.18), and 1.24 (1.13-1.37, p-value for trend <0.001).

CONCLUSIONS: Our results suggest that an extended period of rotating night shift work is associated with a modestly increased risk of type 2 diabetes in women, which appears to be partly mediated through body weight. Proper screening and intervention strategies in rotating night shift workers are needed for prevention of diabetes.

Comment in

PLoS Med. 2011 Dec;8(12):e1001138.
PLoS Med. 2011 Dec;8(12):e1001152.

PMID: 22162955 [PubMed - indexed for MEDLINE] PMCID: PMC3232220 **Free PMC Article**

Diabetes Care. 2008 Mar;31(3):470-5. Epub 2007 Dec 10.

Breast-feeding and type 2 diabetes in the youth of three ethnic groups: the SEARCh for diabetes in youth case-control study.

Mayer-Davis EJ, Dabelea D, Lamichhane AP, D'Agostino RB Jr, Liese AD, Thomas J, McKeown RE, Hamman RF.

Center for Research in Nutrition and Health Disparities and Epidemiology and Biostatistics, University of South Carolina, Columbia, South Carolina 29208, USA. ejmayer@gwm.sc.edu

Abstract

OBJECTIVE: To evaluate the hypothesis that breast-feeding is associated with reduced type 2 diabetes among African-American, Hispanic, and non-Hispanic white youth, mediated in part by current weight status.

RESEARCH DESIGN AND METHODS: The SEARCH Case-Control Study, an ancillary study to SEARCH for Diabetes in Youth, was conducted in two of six SEARCH clinical sites. Eighty youth with type 2 diabetes aged 10-21 years were included. Nondiabetic control participants were recruited from primary care provider offices (n = 167). Breast-feeding information was recalled by biological mothers.

RESULTS: Prevalence (%) of breast-feeding (any duration) was lower among youth with type 2 diabetes than among control subjects (19.5 vs. 27.1 for African Americans, 50.0 vs. 83.8 for Hispanics, and 39.1 vs. 77.6 for non-Hispanic whites). The overall crude odds ratio for the association of breast-feeding (ever versus never) and type 2 diabetes was 0.26 (95% CI 0.15-0.46). Results were similar by race/ethnic group (P value for interaction = 0.17). The odds ratio for the association after adjusting for 12 potential confounders was 0.43 (0.19-0.99). When current BMI z-score was added to the model, the odds ratio was attenuated (0.82 [0.30-2.30]), suggesting possible mediation through current childhood weight status. Analyses that incorporated duration of breast-feeding, adjusted for potential confounders, provided evidence for dose response (test for trend, P value <0.0001), even after inclusion of BMI z-score.

CONCLUSIONS: Breast-feeding appears to be protective against development of type 2 diabetes in youth, mediated in part by current weight status in childhood.

PMID: 18071004 [PubMed - indexed for MEDLINE]

J Health Care Poor Underserved. 2010 Feb;21(1 Suppl):71-94.

Psychosocial determinants of mammography follow-up after receipt of abnormal mammography results in medically underserved women.

Fair AM, Wujcik D, Lin JM, Zheng W, Egan KM, Grau AM, Champion VL, Wallston KA.

Department of Surgery, Meharry Medical College, Nashville, TN 37208, USA. afair@mmc.edu

Abstract

This article targets the relationship between psychosocial determinants and abnormal screening mammography follow-up in a medically underserved population. Health belief scales were modified to refer to diagnostic follow-up versus annual screening. A retrospective cohort study design was used. Statistical analyses were performed examining relationships among sociodemographic factors, psychosocial determinants, and abnormal mammography follow-up. Women with lower mean internal health locus of control scores (3.14) were two times more likely than women with higher mean internal health locus of control scores (3.98) to have inadequate follow-up (OR=2.53, 95% CI=1.12-5.36). Women with less than a high school education had lower cancer fatalism scores than women who had completed high school (47.5 vs. 55.2, p-value=.02) and lower mean external health locus of control scores (3.0 vs. 5.3) (p-value<.01). These constructs have implications for understanding mammography follow-up among minority and medically underserved women. Further comprehensive study of these concepts is warranted.

PMID: 20173286 [PubMed - indexed for MEDLINE] PMCID: PMC2927817 Free PMC Article

Ophthalmology. 2010 Feb;117(2):199-206.e1. Epub 2009 Dec 23.

Central and peripheral visual impairment and the risk of falls and falls with injury.

Patino CM, McKean-Cowdin R, Azen SP, Allison JC, Choudhury F, Varma R; Los Angeles Latino Eye Study Group.

Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California 90033, USA.

Abstract

OBJECTIVE: To evaluate whether central (CVI) and peripheral visual impairment (PVI) are independent risk factors for falls and falls with injury 4 years later.

DESIGN: Population-based, prospective cohort study.

PARTICIPANTS: A population-based sample of 3203 adult Latinos.

METHODS: Baseline presenting binocular central distance acuity was measured and impairment was classified as mild (20/40-20/63) or moderate/severe (<or=20/80). Peripheral visual impairment was classified as mild (-6 dB < mean deviation < -2 dB in worse eye), moderate/severe (mean deviation <or=-6 dB in worse eye).

MAIN OUTCOME MEASURES: Falls and falls with injury in the past 12 months were assessed by self-report at the 4-year follow-up visit.

RESULTS: Out of 3203 individuals, 19% reported falls and 10% falls with injury 4 years after the baseline examination; participants with falls were more likely to be >or=60 years of age, be female, report lower income, have >2 comorbidities, report alcohol use, report wearing bifocal glasses, and report obesity. Among those who reported falls, 7% had CVI (visual acuity >20/40) compared with 4% who did not report falls; and 49% had PVI (mean deviation < -2 dB) compared with 39% of those who did not report falls (both P<0.0001). After adjusting for confounders, moderate to severe CVI and PVI were associated with increased risk for falls (odds ratio [OR], 2.36; 95% confidence interval [CI], 1.02-5.45; P(trend) = 0.04; and OR, 1.42; 95% CI, 1.06-1.91l P(trend) = 0.01, respectively) and with falls with injury (OR, 2.76; 95% CI, 1.10-7.02; P(value) = 0.03; and OR, 1.40; 95% CI, 0.94-2.05 P(trend) = 0.04, respectively).

CONCLUSIONS: Both CVI and PVI were independently associated with increased risk for falls and falls with injury 4 years after the initial examination in a dose-response manner. Although vision-related interventions for preventing falls have mainly focused on correcting CVI, this study suggests that targeting both central and peripheral components may be necessary to effectively reduce rates of falls and falls with injury related to vision loss.

PMID: 20031225 [PubMed - indexed for MEDLINE] PMCID: PMC2819614 Free PMC Article

J Cardiovasc Nurs. 2009 Jan-Feb;24(1):21-8; quiz 29-30.

Predictors of implantable cardioverter defibrillator shocks during the first year.

Dougherty CM, Hunziker J.

Biobehavioral Nursing and Health Systems, University of Washington School of Nursing, Seattle, Washington 98195, USA. cindyd@u.washington.edu

Abstract

The purpose of this study was to predict implantable cardioverter defibrillator (ICD) shocks using demographic and clinical characteristics in the first year after implantation for secondary prevention of cardiac arrest. A prospective design was used to follow 168 first-time ICD recipients over 12 months. Demographic and clinical data were obtained from medical records at the time of ICD insertion. Implantable cardioverter defibrillator shock data were obtained from ICD interrogation reports at hospital discharge, 3, 6, and 12 months. Logistic regression was used to predict ever receiving an ICD shock using background characteristics. Patients received an ICD for secondary prevention of sudden cardiac arrest, they were 64.1 years old, 89% were white, 77% were male, with a mean (SD) ejection fraction of 33.7% (14.1%). The cumulative percentage of ever receiving an ICD shock was 33.3% over 1 year. Three variables predicted shocks in the first year: history of chronic obstructive pulmonary disease (COPD) (odds ratio [OR], 4.42; 95% confidence interval [CI], 1.2-16.4; P = .03), history of congestive heart failure (OR, 3.55; 95% CI, 1.4-9.3; P = .01), and documented ventricular tachycardia (VT) at the time of ICD implant (OR, 10.05; 95% CI, 1.8-55.4; P = .01). High levels of anxiety approached significance (OR = 2.82; P = .09). The presence of COPD, congestive heart failure, or VT at ICD implant was a significant predictor of receiving an ICD shock in the first year after ICD implantation. Because ICD shocks are distressing, painful, and associated with greater mortality, healthcare providers should focus attention on prevention of shocks by controlling VT, careful management of HF symptoms, reduction of the use of short acting beta agonist medications in COPD, and perhaps recognizing and treating high levels of anxiety.

PMID: 19114797 [PubMed - indexed for MEDLINE] PMCID: PMC2673010 Free PMC Article

Crit Care. 2011;15(3):R123. Epub 2011 May 11.

Treatment effects of recombinant human soluble thrombomodulin in patients with severe sepsis: a historical control study.

Yamakawa K, Fujimi S, Mohri T, Matsuda H, Nakamori Y, Hirose T, Tasaki O, Ogura H, Kuwagata Y, Hamasaki T, Shimazu T.

Department of Emergency and Critical Care, Osaka General Medical Center, 3-1-56 Bandai-Higashi, Sumiyoshi-ku, Osaka, Japan. k.yamakawa0911@gmail.com

Abstract

INTRODUCTION: Cross-talk between the coagulation system and inflammatory reactions during sepsis causes organ damage followed by multiple organ dysfunction syndrome or even death. Therefore, anticoagulant therapies have been expected to be beneficial in the treatment of severe sepsis. Recombinant human soluble thrombomodulin (rhTM) binds to thrombin to inactivate coagulation, and the thrombin-rhTM complex activates protein C to produce activated protein C. The purpose of this study was to examine the efficacy of rhTM for treating patients with sepsis-induced disseminated intravascular coagulation (DIC).

METHODS: This study comprised 65 patients with sepsis-induced DIC who required ventilatory management. All patients fulfilled the criteria of severe sepsis and the International Society on Thrombosis and Haemostasis criteria for overt DIC. The initial 45 patients were treated without rhTM (control group), and the following 20 consecutive patients were treated with rhTM (0.06 mg/kg/day) for six days (rhTM group). The primary outcome measure was 28-day mortality. Stepwise multivariate Cox regression analysis was used to assess which independent variables were associated with mortality. Comparisons of Sequential Organ Failure Assessment (SOFA) score on sequential days between the two groups were analyzed by repeated measures analysis of variance.

RESULTS: Cox regression analysis showed 28-day mortality to be significantly lower in the rhTM group than in the control group (adjusted hazard ratio, 0.303; 95% confidence interval, 0.106 to 0.871; $P = 0.027$). SOFA score in the rhTM group decreased significantly in comparison with that in the control group ($P = 0.028$). In the post hoc test, SOFA score decreased rapidly in the rhTM group compared with that in the control group on day 1 ($P < 0.05$).

CONCLUSIONS: We found that rhTM administration may improve organ dysfunction in patients with sepsis-induced DIC. Further clinical investigations are necessary to evaluate the effect of rhTM on the pathophysiology of sepsis-induced DIC.

PMID: 21569368 [PubMed - indexed for MEDLINE] PMCID: PMC3218981 Free PMC Article

Chest. 1994 Jan;105(1):76-82.

Early sepsis treatment with immunoglobulins after cardiac surgery in score-identified high-risk patients.

Pilz G, Kreuzer E, Kääb S, Appel R, Werdan K.

Department of Medicine I, Grosshadern University Hospital, University of Munich, Germany.

Erratum in
Chest 1994 Jun;105(6):1924.

Abstract

In patients at risk for sepsis after cardiac surgery, the efficacy of intravenous immunoglobulin (Ig) treatment was compared with a historical control population, equivalent in patient characteristics and disease severity. Using APACHE II scores, especially in the high-risk group (IgG), we could discriminate between low-risk patients (score < 19; mortality 1 percent) and the small groups at risk (score 19 to 23) and high risk (score > or = 24) with a significantly higher mortality (14 percent and 76 percent, respectively) [corrected]. Subsequently, among 1,341 consecutive patients we prospectively identified and treated (IgG n = 41 IgGMA: n = 25) these at-risk groups. In contrast to controls (risk: n = 21; high-risk; n = 21), we found a marked fall in APACHE II scores, especially in the high-risk group (IgG, n = 26: p < 0.05; IgGMA, n = 13: p = 0.08) [corrected]. In this group, Ig therapy produced higher (p < 0.05) response rates (score decrease 7 within 4 days: IgG: 54 percent, IgGMA: 62 percent; controls: 19 percent) and reduced mortality (IgG: 46 percent, IgGMA: 46 percent; controls: 76 percent), statistically significant (p < 0.05) for Ig treatment overall. Thus, early Ig treatment improves disease severity and may improve prognosis in prospectively score-identified high-risk postcardiac surgical patients.

PMID: 8275790 [PubMed - indexed for MEDLINE] Free full text

Pediatrics. 2011 May;127(5):e1112-9. Epub 2011 Apr 18.

Computerized physician order entry with decision support decreases blood transfusions in children.

Adams ES, Longhurst CA, Pageler N, Widen E, Franzon D, Cornfield DN.

Center for Excellence in Pulmonary Biology, Stanford University School of Medicine, 770 Welch Rd, Suite 350, Stanford, CA 94304, USA. eloa@stanford.edu

Abstract

OBJECTIVE: Timely provision of evidence-based recommendations through computerized physician order entry with clinical decision support may improve use of red blood cell transfusions (RBCTs).

METHODS: We performed a cohort study with historical controls including inpatients admitted between February 1, 2008, and January 31, 2010. A clinical decision-support alert for RBCTs was constructed by using current evidence. RBCT orders resulted in assessment of the patient's medical record with prescriber notification if parameters were not within recommended ranges. Primary end points included the average pretransfusion hemoglobin level and the rate of RBCTs per patient-day.

RESULTS: In total, 3293 control discharges and 3492 study discharges were evaluated. The mean (SD) control pretransfusion hemoglobin level in the PICU was 9.83 (2.63) g/dL (95% confidence interval [CI]: 9.65-10.01) compared with the study value of 8.75 (2.05) g/dL (95% CI: 8.59-8.90) (P < .0001). The wards' control value was 7.56 (0.93) g/dL (95% CI: 7.47-7.65), the study value was 7.14 (1.01) g/dL (95% CI: 6.99-7.28) (P < .0001). The control PICU rate of RBCTs per patient-day was 0.20 (0.11) (95% CI: 0.13-0.27), the study rate was 0.14 (0.04) (95% CI: 0.11-0.17) (P = .12). The PICU's control rate was 0.033 (0.01) (95% CI: 0.02-0.04), and the study rate was 0.017 (0.007) (95% CI: 0.01-0.02) (P < .0001). There was no difference in mortality rates across all cohorts.

CONCLUSIONS: Implementation of clinical decision-support alerts was associated with a decrease in RBCTs, which suggests improved adoption of evidence-based recommendations. This strategy might be widely applied to promote timely adoption of scientific evidence.

PMID: 21502229 [PubMed - indexed for MEDLINE] Free full text