How do I test my hypothesis?

Introduction to Outbreak Investigation
Module 7 – Analytical Epidemiology

Jason Lombard, Julie Lenoch and Kathe Bjork

Outline

- Objectives for this module
- Review the steps in an outbreak (epidemiological) investigation
- Analytical epidemiology
  - Type of studies
- Hierarchy of evidence
- Data sources
- Quiz

Analytical epidemiology

Learning objectives for this module:

1. Understand the role of analytical epidemiology in an outbreak investigation
2. Become familiar with the types of analytical studies and which type to use depending on the situation.
3. Understand the benefits and limitations of analytical epidemiology.
4. Review sources of data.

9 steps of outbreak investigation

- Confirm the existence of an outbreak
- Confirm the diagnosis
- Develop a case definition
- Establish disease monitoring and surveillance
- Perform descriptive epidemiology
- Generate hypotheses
- ✓ **Perform analytical epidemiology**
- Implement control and prevention measures
- Communicate findings

Analytical epidemiology

An analytical study is typically done after a descriptive study is completed

- Remember descriptive studies: ecologic, case report, case series and surveys
- Hypothesis generation/ development
- Analytical epidemiology will evaluate/ test the hypothesis

Epidemiology

Describes:

- Who?
  what population?
- What?
- Where?
- When?
- Why?
- How?

Descriptive epidemiology

Describes:

- Who?
  what population?
- What?
- Where?
- When?

Analytic epidemiology

Cause =
BVD vaccination

Effect=
Bovine neonatal pancytopenia

Describes:

- Who?
  what population?
- What?
- Where?
- When?
- Why?
- How?

Analytical epidemiology benefits

- Useful to *identify the cause* of the disease but not always conclusive
  - Provides evidence

- Establishes estimates of *risk* (odds ratio or relative risk) of a particular exposure leading to the disease of interest

Analytical epidemiology limitations

- More consuming than descriptive epidemiology – resource intensive (time, personnel, etc.)
- Requires hypothesis, case definition and potential exposure(s)
- Studies also require unaffected animals which make studies more costly

Purpose of analytical studies

A study considers two types of variables: exposure and outcome

- Exposure refers to factors that might influence one's risk of disease
- Outcome refers to case definitions

Purpose of analytical studies

- To statistically show a relationship between exposure and disease outcome

Types of epidemiology studies

- Descriptive epi: studies the occurrence and distribution of disease

- **Analytical epi:**
  - **Observational: More in depth studies to validate (or reject) a hypothesis about the occurrence of disease**

  - Experimental Epi: Deliberate control of an exposure to evaluation the response and monitor it's effect

Observational studies

- Investigators try not to influence the natural course of events, but make observations with attention paid to exposure and outcome
- These type of studies “allow nature to take its course” and the results are obtained
- Because researchers are not controlling all the risk factors or exposures, confounding may be an issue

Analytical studies

Prospective vs. Retrospective

- Prospective: Only the exposure has occurred at the start of the study
- Retrospective: Both the exposure and outcome have occurred when study begins, these require prerecorded data or secondary sources of data

Observational studies

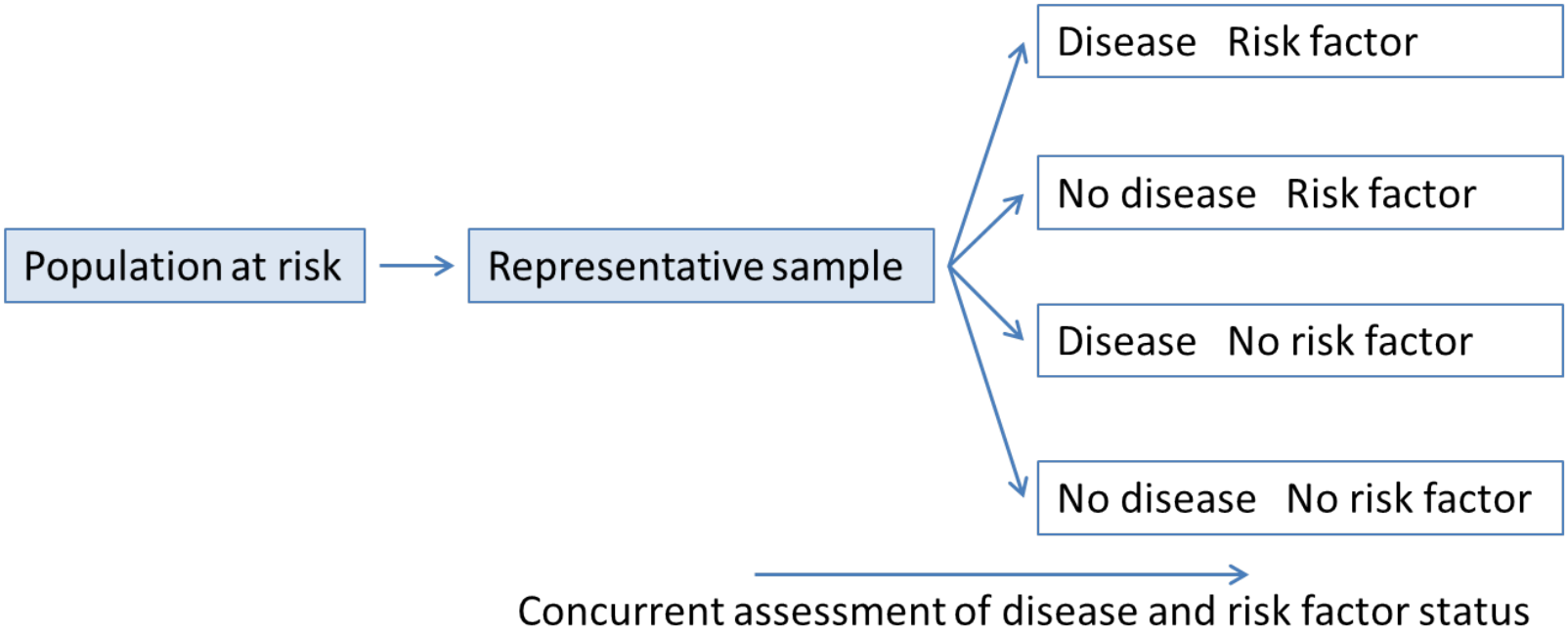
Three main types:

1. **Cross-sectional study**
2. Cohort study
3. Case-control study

Cross-sectional studies

- Primarily used to determine prevalence
- Both exposure and disease status are determined at the same time in the sample population
- Temporality is not established based on a cross-sectional study

Cross-sectional Study Design



Study population=Survivors

Prevalence of fecal shedding of *Salmonella* spp in dairy herds

Carla L. Huston, DVM, PhD; Thomas E. Wittum, PhD; Brenda C. Love, DVM, PhD, DACVM; James E. Keen, DVM, PhD

Objective—To estimate prevalence of *Salmonella* spp in Ohio dairy farms and to identify potential risk factors for fecal shedding of salmonellae.

Design—Cross-sectional study.

Sample Population—105 Ohio dairy farms.

Procedure—Individual fecal samples from all mature cows in study herds were tested for *Salmonella* spp by use of standard bacteriologic culture procedures. Herds were identified as infected if at least 1 cow was shedding *Salmonella* spp. Information regarding herd characteristics, management practices, and health history were collected. Potential risk factors for herd-level *Salmonella* infection were identified.

Results—In 31% of the study herds (95% confidence interval, 22 to 40%), at least 1 cow was shedding *Salmonella* spp. Six percent of 7,776 fecal samples contained *Salmonella* organisms; prevalence within infected herds ranged from < 1 to 97%. Herd size, use of free stalls for lactating and nonlactating cows, and use of straw bedding in nonlactating cows were significantly associated with fecal shedding of *Salmonella* spp, as determined by use of univariate analysis. By use of multivariate analysis, large herds were more likely to be infected than smaller herds; however, no other factors were associated with *Salmonella* infection after adjustment for herd size.

Conclusions and Clinical Relevance—Subclinical shedding of *Salmonella* spp is common in Ohio dairy herds, although we could not identify specific interventions that may influence the prevalence of *Salmonella* spp on dairy farms. It appears that large herd size and intensive management may provide an environment conducive to *Salmonella* shedding and chronic dairy herd infection. (*J Am Vet Med Assoc* 2002;220:645–649)

culling rates, reduced feed efficiency, decreased weight gain, and decreased milk production.

Cattle infected with *Salmonella* spp may shed large numbers of the organism in their feces either continuously or intermittently, with subclinical shedding of the organism being more common than clinical disease.[5,6] Stress, as is often encountered during calving, lactation, and intensive rearing, may induce fecal shedding and propagate the infection within a herd.[7,8] In addition, the organism survives for prolonged periods in animals and the environment, which can lead to chronic herd infection.[9,10]

Dairy cattle infected with *Salmonella* spp can pose a substantial risk to public health when they are culled and enter the food supply as beef products. The 1996 USDA survey of dairy cattle, National Animal Health Monitoring System (NAHMS) Dairy '96, reported that 14.9% of cull dairy cows were shedding *Salmonella* spp, and 66.7% of dairy cull markets had at least 1 animal shedding *Salmonella* spp. In addition, several documented cases of human illness have been traced back to dairy farms.[11,12] Thus, the control of *Salmonella* infections on dairy farms may have important public health implications.

Little is known about the actual prevalence of *Salmonella* infection in dairy herds or about herd characteristics or management practices that may protect against or predispose to infection. Identification of preventive practices and risk factors for dairy operations may lead to herd intervention strategies that protect human and animal populations. Thus, the objectives of this study were to estimate the prevalence of *Salmonella* infection among dairy herds in Ohio and identify potential herd-level risk factors for fecal shedding of this organism.

Materials and Methods

Study population—Whole-herd fecal samples from dairy

Cross-sectional study

- The familiar 2 X 2 table

	Diseased	Non Diseased	Total
Exposed	a	b	Total Exposed
Non Exposed	c	d	Total Non Exposed
Total	Total Diseased	Total Non Diseased	Total of All in study group

Cross-sectional study- results

- Results from a cross-sectional study allow calculation of **prevalence and odds ratios**

- **Prevalence in population**
  - the number of cases at a single point in time in the population of interest.

$$Prevalence = \frac{a + c}{a + b + c + d}$$

Cross-sectional study- Results

- Results from a cross-sectional study allow calculation of **Prevalence and Odds ratios**

- **Odds ratio**
  - the ratio of the odds of an event occurring in an exposed population to the odds of it occurring in the unexposed population

$$Odds\ ratio\ (OR) = \frac{a\ x\ d}{b\ x\ c}$$

Cross-sectional studies

- ***For cross-sectional studies***: OR is a ratio of the odds of disease in exposed animals to the odds of disease in unexposed.

- Interpretation of OR
  - OR > 1: Increased odds of disease among those with exposure
  - OR < 1: Decreased odds, or protective effects from disease, among those exposed
  - OR = 1: No association between disease and exposure

Cross-sectional study advantages

- Easy to implement
  - Study population is chosen
  - Exposure and disease status are obtained at one point in time
  - Outcome measured is prevalence of exposures and disease
- Short time frame
- Lower cost
- Can assess multiple exposures and outcomes

Cross-sectional study disadvantages

- Population sampling and response rate
  - Selection bias
  - External validity
- Reverse-causation
  - Unable to determine if the exposure occurred before the outcome
- Survival bias
  - May impact response rate
- Not good for rare diseases
  - Random selection

Observational studies

Three main types:

1. Cross-sectional study
2. **Cohort study**
3. Case-control study

Cohort studies

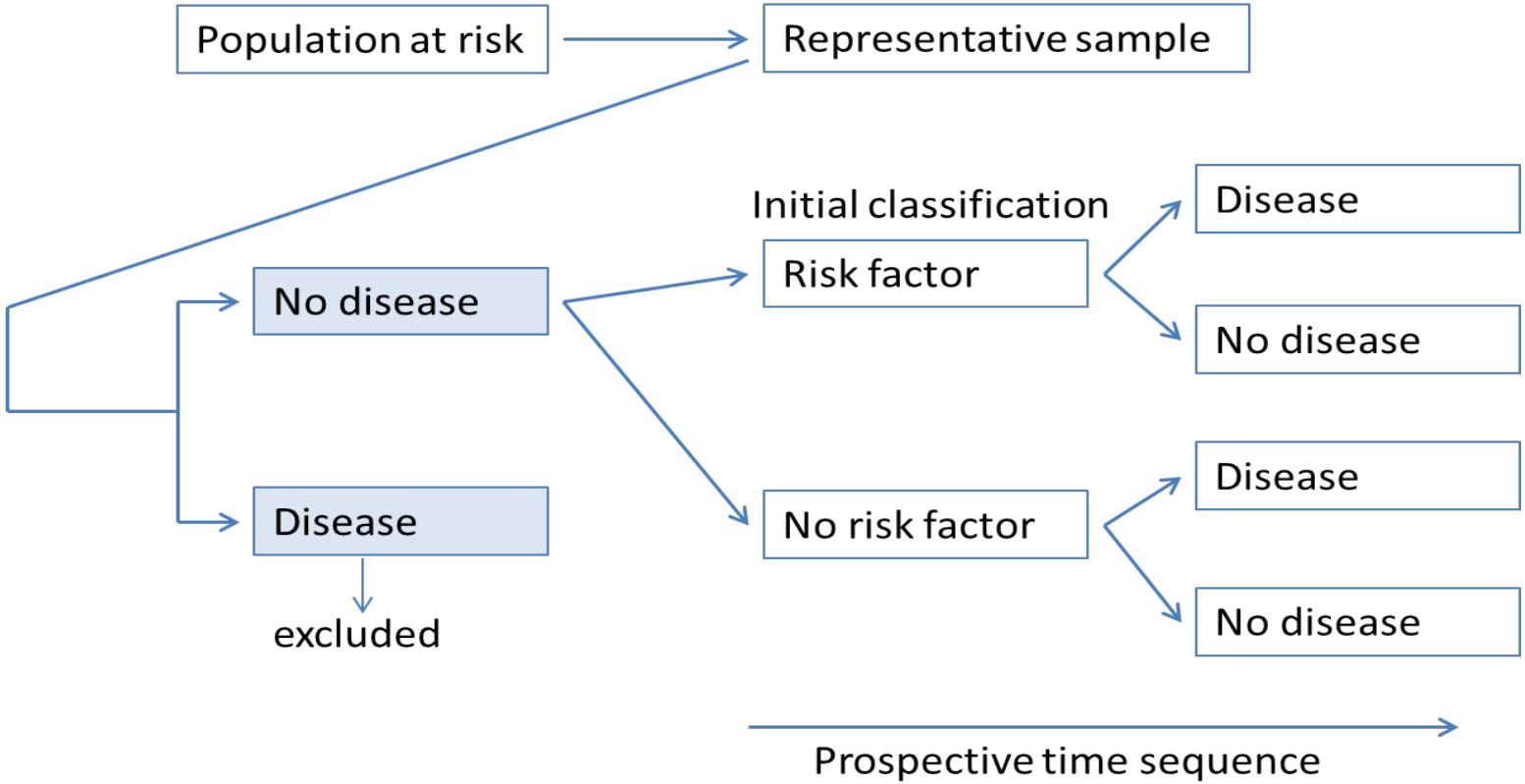
- Study population is grouped by *exposure status* in a non-diseased population
- Diseased animals are excluded from enrollment
- Groups are studied to determine if individuals develop the outcome of interest (i.e., disease)

Cohort studies

- *Prospective* cohort study: Exposure is known, participants are followed over a future time period to determine if they develop the disease

- *Retrospective* cohort study: Cohort is compiled from past history/records of exposure and followed forward.

Cohort Study Design



Study population=Disease-free subjects

A cohort study to examine maternally-associated risk factors for bovine spongiform encephalopathy

J. W. Wilesmith, G. A. H. Wells, J. B. M. Ryan, D. Gavier-Widen, M. M. Simmons

Veterinary Record (1997) **141**, 239-243

This long-term cohort study, initiated in July 1989, was designed to examine maternally-associated risk factors for bovine spongiform encephalopathy (BSE), forming part of the epidemiological research programme to assess the risks of non-feedborne transmission of BSE. In this study, the incidence of BSE in offspring of cows which developed clinical signs of BSE is compared with that in offspring, born in the same calving season and herd, of cows which had reached at least six years of age and had not developed BSE. All offspring were allowed to live to seven years of age. The results indicate a statistically significant risk difference between the two cohorts of 9·7 per cent and a relative risk of 3·2 for offspring of cows which developed clinical BSE. However, there is some evidence that this enhanced risk for offspring of BSE cases declined the later the offspring was born, but was increased the later the offspring was born in relation to the stage of the incubation period of the dam. The results presented cannot distinguish between a genetic component and true maternal transmission or a combination of both risks, but they do not indicate either that the BSE epidemic will be unduly prolonged or that the future incidence of BSE in Great Britain will increase significantly.

Although a necessarily long-term cohort study had been initiated in July 1989 (Wilesmith 1996a), the opportunity to conduct a further within-herd matched case-control study of cases born after this ban was taken in 1994 to examine the potential roles of maternal and/or horizontal transmission of BSE. The results of this study indicated that neither of these means of transmission could explain the majority of the cases in animals born after the feed ban (Hoinville and others 1995). The offspring of animals that were subsequently affected with BSE were not found significantly more often among the cases. Additional studies indicated that observance of the feed ban had been imperfect resulting in a continued, but much reduced, risk from the feedborne source, notably as a result of cross-contamination within feedmills (Wilesmith 1996b, c). This paper describes the initial results of the cohort study to examine the possibility of maternally-associated risk factors in the occurrence of BSE.

Materials and methods

Study animals

The main BSE epidemiological database (Wilesmith and others 1992b) was used to identify offspring of confirmed cases of BSE

Cohort study

- The familiar 2 X 2 table

	Diseased	Non Diseased	Total
Exposed	a	b	Total Exposed
Non Exposed	c	d	Total Non Exposed
Total	Total Diseased	Total Non Diseased	Total of All in study group

Cohort studies–results

- *Incidence* is the number of new cases per population in a given time period. It is calculated as the number of new cases of disease divided by the total number of animals in the study during the study period.

$$Incidence = \frac{a+c}{a+b+c+d} \text{ per study period}$$

Incidence is a rate (time is important) as opposed to prevalence which is a proportion

Cohort studies–results

- *Relative risk* is the probability of developing the disease if exposed compared to the probability if not exposed.

$$Relative\ risk\ (RR) = \frac{a\ /\ (a+b)}{c\ /\ (c+d)}$$

- If RR > 1, exposed have more probability of developing disease than the unexposed
- If RR < 1, exposed have less probability of developing disease than the unexposed
- If RR = 1, exposed and unexposed have the same risk of developing disease

Cohort studies–advantages

- Allows calculation of disease incidence
- Follows logical clinical course of disease
  - "If exposure occurs, will disease follow?"
- May establish a risk associated with the exposure of interest
- Results are high on hierarchy of evidence
  - High "proof" of causal association

Cohort studies–disadvantages

- More resource intensive than other studies
- Studies may take a long time for disease to occur
- Inappropriate for rare diseases
- Prospective cohort generally not feasible during an outbreak; retrospective may be feasible

Observational studies

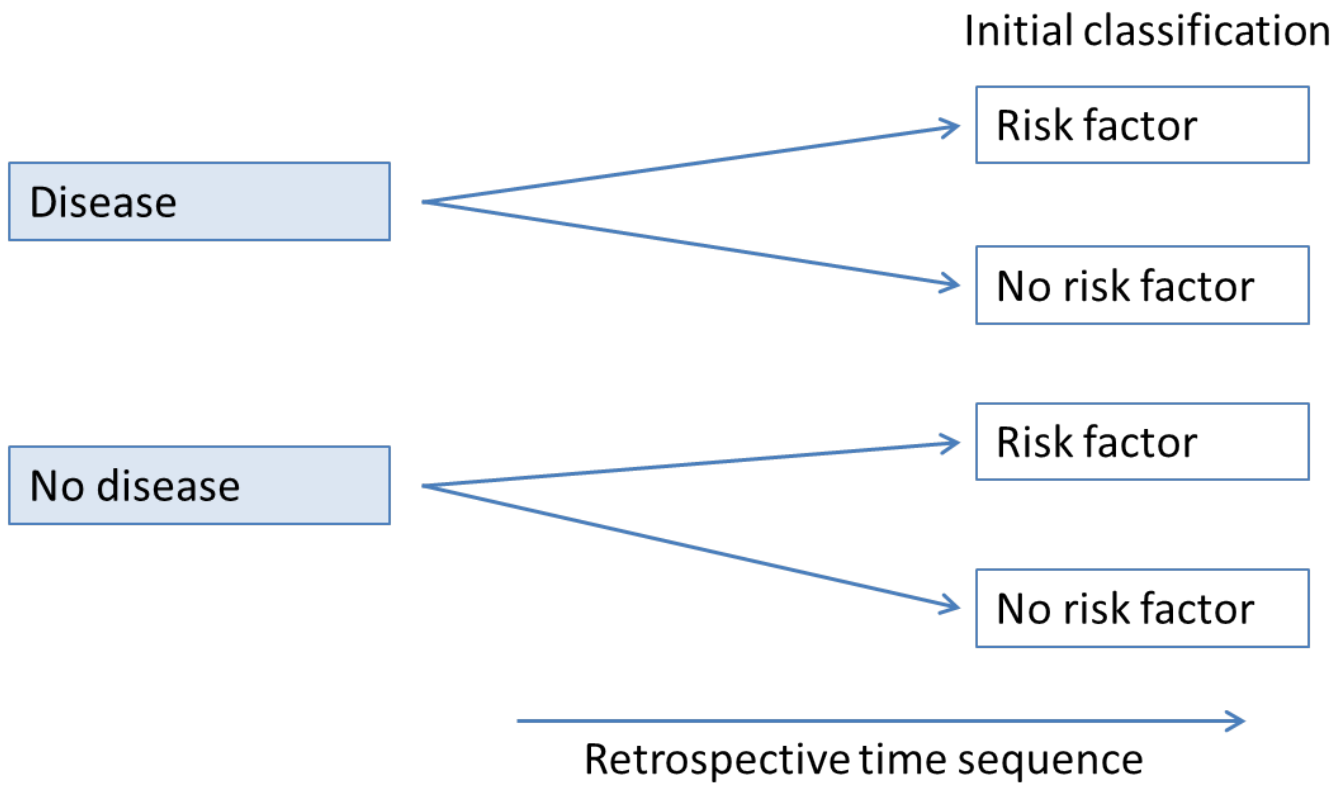
Three main types:

1. Cross-sectional study
2. Cohort study
3. **Case-control study**

Case-control study

- The disease outcome defines the study group
- Cases (based on case definition) have the disease
- Controls (based on case definition) do not have the disease
- The 2 groups are compared based on previous exposures

Case-control Study Design



Study population=Cases and controls

J Vet Intern Med 2013;27:339–346

Case–Control Study of a Multistate Equine Herpesvirus Myeloencephalopathy Outbreak

J.L. Traub-Dargatz, A.M. Pelzel-McCluskey, L.H. Creekmore, S. Geiser-Novotny, T.R. Kasari, A.M. Wiedenheft, E.J. Bush, and K.E. Bjork

Background: A large multistate outbreak of equine herpesvirus myeloencephalopathy (EHM) occurred in May 2011 among horses that participated in a competitive event.

Objective: To identify EHM risk factors among horses with a common exposure venue.

Animals: A total of 123 horses: 19 horses with EHM, 14 equine herpesvirus-1 cases with no reported neurologic signs, and 90 control horses.

Methods: EHM case survey data were compared with data from EHV-1 cases with no neurologic signs and healthy controls using univariable and multivariable methods.

Results: Significant factors associated with higher risk for EHM compared with EHV-1 cases with no neurologic signs were (1) greater number of biosecurity risks at the event, (2) female sex, (3) increasing number of classes competed in at the event, and (4) an interaction between sex and number of classes competed in. In the EHM versus controls comparison, in addition to sex and biosecurity risks, factors associated with higher EHM risk included EHV-1 vaccination in the 5 weeks before the event and increasing number of events attended in April 2011; zinc dietary supplementation was associated with decreased risk. An interaction between sex and the number of events attended in April 2011 also was significant.

Conclusions and Clinical Importance: Findings from this study suggest that dietary zinc supplementation may be associated with decreased risk of EHM. Several factors were associated with increased risk of EHM. Additional investigations of factors associated with risk of EHM are warranted to evaluate the importance of these factors in this complex disease of horses.

Key words: EHM risk; Neurologic disease.

Case-control study

- Will lead you to the familiar 2 X 2 table

	Diseased	Non Diseased	Total
Exposed	a	b	Total Exposed
Non Exposed	c	d	Total Non Exposed
Total	Total Diseased	Total Non Diseased	Total of All in study group

Case-control studies

- Disease prevalence cannot be calculated because the population at risk is not known

- OR is the measure used with case-control studies

$$Odds\ ratio\ (OR) = \frac{a\ x\ d}{b\ x\ c}$$

Case-control studies

- ***For case-control studies***: OR is a ratio of the odds of exposure in cases to the odds of exposure in controls.

- Interpretation of OR
  - OR > 1: Increased odds of exposure among those with outcome
  - OR < 1: Decreased odds, or protective effects, among those with outcome
  - OR = 1: No association between exposure and outcome

Case-control study advantages

- Shorter time frame (disease has already occurred)
- Fewer resources than a cohort study
- Smaller sample sizes may be okay
- Good for rare disease

Case-control study disadvantages

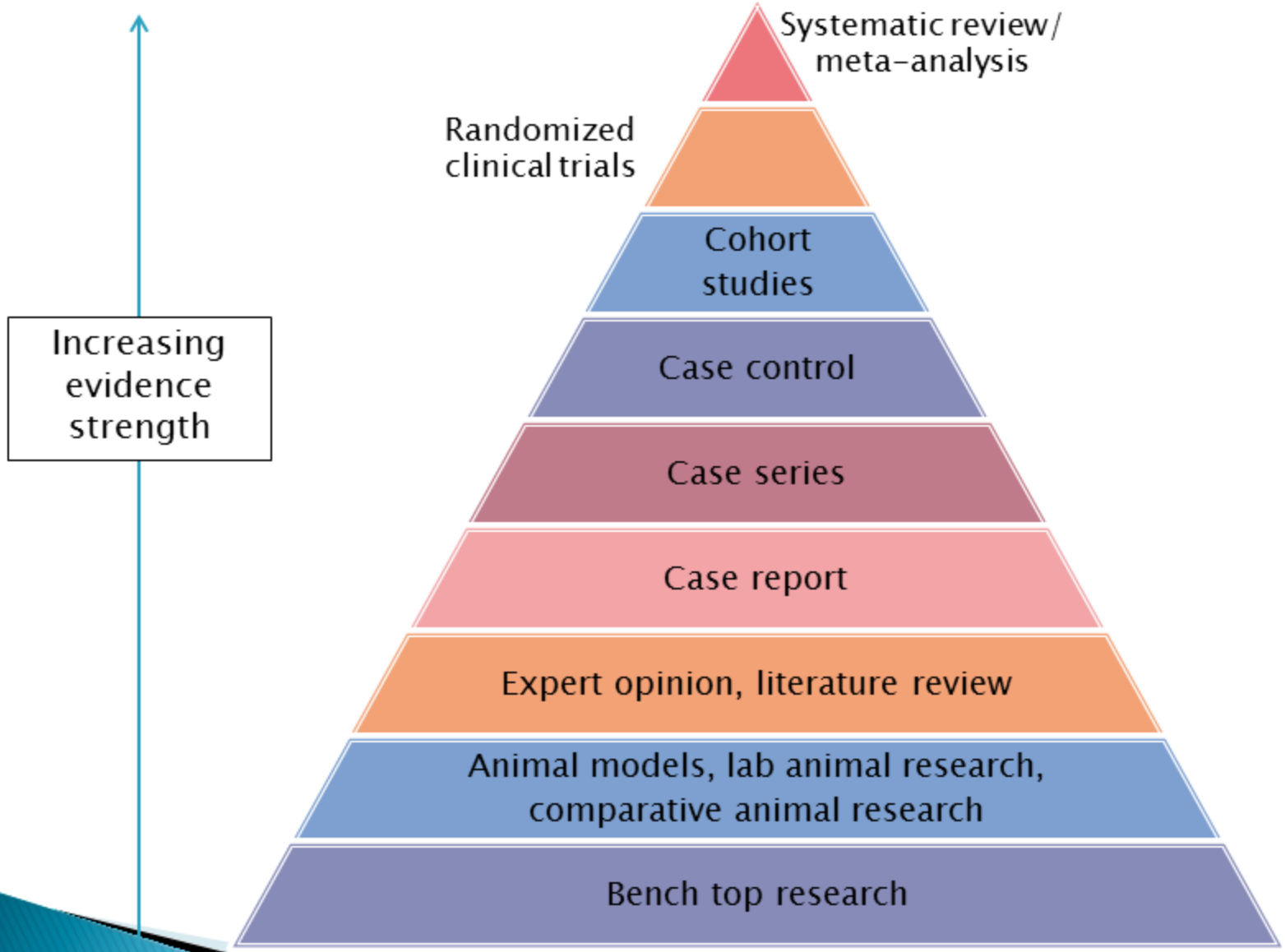
- Difficulty finding appropriate controls
- Possible survivor bias
  - Those who die quickly can't be counted as cases
- Retrospective time sequence
- Difficult for rare exposures
- Generally restricted to one outcome

Analytic- observational studies

Summary

1. Cross-sectional study
2. Cohort study
3. Case-control study

Hierarchy of Evidence



So, you want to conduct a study?

- What type of study?
  - Cross-sectional
  - Case-control
  - Cohort

- Determined by your objectives and resources

Cross-sectional study

- Sampling is done without regard to disease status
- Relatively quick and easy to perform
- Estimate prevalence
- Associations between diseased and non-diseased populations
- No causation determination – temporal issue

Case-control study

- Sampling is done based on disease status
- Relatively quick and easy to perform
- Can’t determine prevalence
- Best design for rare diseases
- Selection of control population is critical

Case-control study (con't)

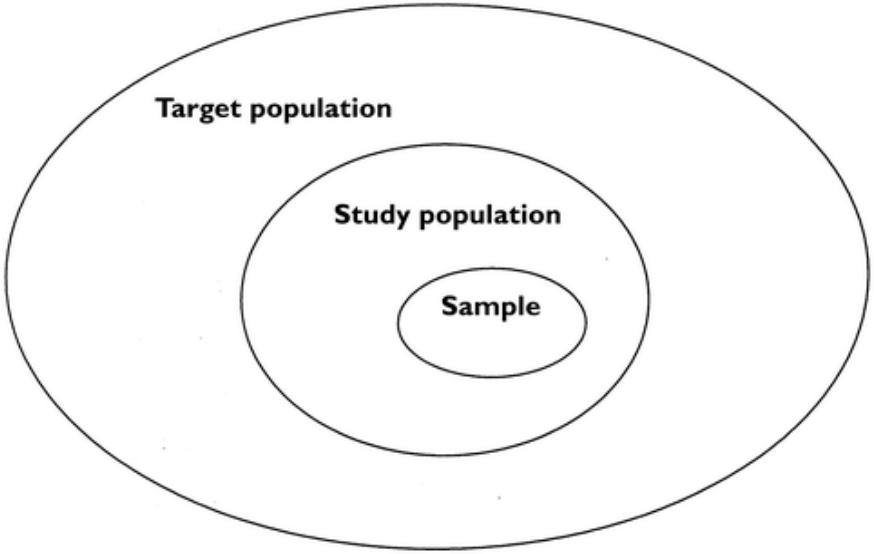
- Associations between diseased and non-diseased populations
- Can't determine causation alone but can provide evidence
- Potential biases
  - Selection
  - Recall

Cohort study

- Grouping is assigned based on exposure
- Most resource intensive design
- Can estimate incidence
- Evidence of risk factors and subsequent causation
- Can’t determine prevalence of exposure

What is your population of interest?

- Target population
  - Inference
- Study population
  - Sample drawn
- Sample
  - Herd, county, state, region, national?



Sample size

- Determined prior to the study
- Power – ability to detect a difference
  - Hypothesis: Group A > Group B
- Seek the expertise of a statistician to determine your sample size or use a sample size calculator

Are the data already available?

Examples

- National Agricultural Statistics Service (NASS)
  - Demographic, production, price data, etc.
- National Animal Health Monitoring System (NAHMS)
  - Production practices, disease prevalence, etc.
- Veterinary Services data
- Food Safety Inspection Service
  - Slaughter plants
  - Condemnation data

Are the data already available?

- Be creative!
  - Certificates of Veterinary Inspection (CVI)
  - Dairy Herd Information Association
  - Diagnostic labs

Conducting a new survey

- Can be conducted via:
  - Face-to-face interviews
  - Telephone
  - Internet
  - Mail
- Methodology can influence response rates
- May be resource intensive

Conducting a new survey

- NAHMS can assist in providing/writing questions
- Avoid open-ended/fill-in-the-blank questions
- Office of Management and Budget approval needed if more than 9 interviews conducted

9 steps of outbreak investigation

- Confirm the existence of an outbreak
- Confirm the diagnosis
- Develop a case definition
- Establish disease monitoring and surveillance
- Perform descriptive epidemiology
- Generate hypotheses
- ✓ Perform analytical epidemiology
- Implement control and prevention measures
- Communicate findings

Analytical epidemiology – references

- Dohoo I, Martin W, and Stryhn H (2003) Veterinary Epidemiologic Research, VER Inc.
- Gordis, L. (2000). Epidemiology: 2nd Edition. W.B. Saunders Company: Philadelphia, PA.
- Huston CL, Wittum TE, Love BC, Keen JE. 2002. Prevalence of fecal shedding of *Salmonella spp* in dairy herds. J Am Vet Med Assoc. 220:645-9.
- Mann CJ. 2003. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. Emerg Med J. 20:54-60. Review..
- North Carolina Center for Public Health Preparedness, Public Health Information Network Series 1. Retrieved from http://www.vdh.virginia.gov/EPR/Training.asp
- Rothman KJ, Greenland S & Lash TL (2008). Modern Epidemiology. 3rd Edition. Lippincott Williams & Wilkins: Philadelphia, PA.
- Szklo M & Nieto FJ (2000). Epidemiology, Beyond the Basics. Aspen Publishers: Gaithersburg, MD.
- Traub-Dargatz JL, Pelzel-McCluskey AM, Creekmore LH, Geiser-Novotny S, Kasari TR, Wiedenheft AM, Bush EJ, Bjork KE. 2013. Case-control study of a multistate equine herpesvirus myeloencephalopathy outbreak. J Vet Intern Med. 27:339-46.
- Wilesmith JW, Wells GA, Ryan JB, Gavier-Widen D, Simmons MM. 1997. A cohort study to examine maternally-associated risk factors for bovine spongiform encephalopathy. Vet Rec. 141:239-43.