# Epidemiologic Studies

When and Why?

# DESCRIPTIVE EPIDEMIOLOGY

- Definition:
  - Any general observational study of the occurrence of disease or other health-related event in a population (animal or human)
- Characteristics can be described under the general headings: animal, place, time

# DESCRIPTIVE EPIDEMIOLOGY

- Types of descriptive studies
  1. Surveys
  2. Case reports or case studies
  3. Monitoring or surveillance programs

# ANALYTIC EPIDEMIOLOGY

- Definition:
  - Design (and analysis) of studies that will allow valid *comparisons* to be made between different groups.
- Purpose:
  - The comparisons are made to elucidate associations between risk factors (determinants) and diseases (or other events).
  - Ultimate goal is to define causal relationships that can be used to design treatment, control and prevention strategies.

# STUDY DESIGNS

- Three most commonly used designs
  - Cross-sectional
  - Case-control
  - Cohort

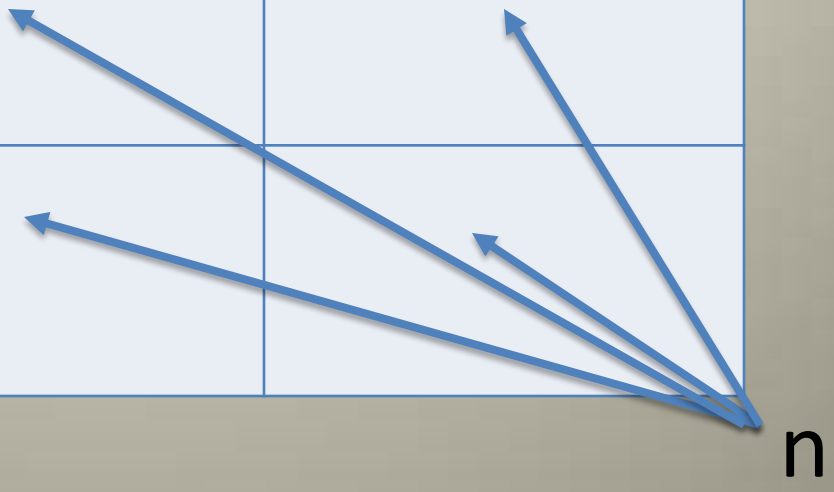
- Sometimes additional terms *retrospective* or *prospective* are used to describe how the data were/will be collected in time

# Cross-sectional study (survey)

- Select sample of flocks without knowledge of disease or risk factor history
- Determine NCD  status (positive or negative) based on testing a sample of birds from the flock
- Determine risk factor history by owner interview – usually just at the time the flock is visited

# Cross-sectional study

| Risk factor | NCD positive | NCD negative |
|---|---|---|
| Positive (eg. Exotic birds) (exposed) | | |
| Negative (eg. Exotic birds) (not exposed) | | |



# Example: Cross-sectional study

| Risk factor | NCD positive | NCD negative |
|---|---|---|
| Exotic bird introductions | 10 | 35 |
| No exotic bird introduction (not exposed) | 5 | 50 |

100

# Case-control study

- Select flocks with confirmed NCD outbreaks
  - using a standard case definition as *cases*
- Select flocks without confirmed NCD as *controls*
- Interview owners to obtain histories of potential exposure to risk factors
  - e.g. live bird introductions, lack of vaccination, introduction of exotic birds, etc.
- Numbers of cases and controls usually equal but sometimes 2x or 3x more controls

# Case-control study

| Risk factor | NCD positive (Cases) | NCD negative (Controls) |
|---|---|---|
| Positive (eg. exotic birds) (exposed) | | |
| Negative (eg. no exotic birds) (not exposed) | | |
| | $n_1$ | $n_2$ |

# Example: Case-control study

| Risk factor | NCD positive (Cases) | NCD negative (Controls) |
|---|---|---|
| Exotic bird introductions | 32 | 10 |
| No exotic bird introductions | 18 | 40 |
| | 50 | 50 |

# Cohort study

- Select flocks *with* a history of the risk factor of interest
  - e.g. live bird introductions, Introduction of exotic birds
- Select flocks *without* a history of the risk factor of interest
- After a suitable time period, determine NCD status (positive or negative) based on testing a sample of birds from the flock

# Cohort study

| Risk factor | NCD positive | NCD negative |
|---|---|---|
| Positive (exposed) | | |
| Negative (not exposed) | | |

# Example: Cohort study

| Risk factor | NCD positive | NCD negative | |
|---|---|---|---|
| Exotic bird introductions | 40 | 10 | 50 |
| No exotic bird introduction (not exposed) | 15 | 35 | 50 |

# MEASURES OF ASSOCIATION

- Relative risk and attributable risk
  - Applicable only to cohort studies where incidence is measured
- Odds ratio
  - Applicable to cross-sectional, case-control and cohort studies

# "ASSOCIATION"

- Means a relationship, a linkage in occurrence, or a dependency.

- Examples of associations

# "ASSOCIATION"

How can an association be evaluated?

- Observation (the eyeball test)

or

- Statistics (incorporates uncertainty associated with sampling and biologic variation)

# "CORRELATION"

- Measure of linear relationship between 2 ordinal or continuous variables.
- All significant correlations are associations, but you should not think of the reverse as being true because not all associations should be measured as correlations.

- Examples of correlations

# CAUSAL ASSOCIATION

- Factor X "causes" Y if a change imposed directly on results in a change in Y.
  - Difficult to prove, hence we make judgments based on 4 pieces of information:

1. Chance that the observed association occurred just because of random variation (the P value);
2. Possibility of intrinsic non-causal relationship ("night-and-day go together");
3. Chance that there was bias [systematic--not random--error] in the study design; and
4. "Surgeon General's Criteria".

# 1. Random variation

Examples

a) Biologic variation within/among animals
b) Imprecision in measuring devices

Improving precision (reliability)

i) Better measuring devices, standardize measurement methods
ii) Sampling subset of population only
iii) Repeating measurements - use mean of 2 or more measurements on a single animal/sample
iv) Increasing sample sizes for estimation of mean response

# Evaluation of random variation

a) Hypothesis testing - relate observed diff. between groups to the predicted or expected variation.

Null hypothesis (Ho) vs alternate hypothesis (Ha)

**"P-value"** is the probability (P for "probability") that there could be "this much or more observed relative difference" if the Ho were true. The smaller the P value, the less likely it is that observed relative difference is <u>just</u> due to chance.

# Decisions – correct or incorrect

b) Errors in making decisions

--complete the table

- possibilities are
  - power,
  - type II error,
  - type I error,
  - confidence

| | | Truth | |
|---|---|---|---|
| | | Different | Not different |
| Decision | Reject Ho (difference) | | |
| | Accept Ho (no. difference) | | |

# Decisions – correct or incorrect

b) Errors in making decisions
   - power,
   - type II error,
   - type I error,
   - confidence

| | | Truth: Different | Truth: Not different |
|---|---|---|---|
| Decision | Reject Ho (difference) | Power (80%) | Type I error (5%) |
| | Accept Ho (no. difference) | Type II error (20%) | Confidence (95%) |

# 2. Intrinsic non-causal relationships

Certain things just go together!

<u>Examples</u>

a) Black and white cows as beigng dairy cattle
b) Sheep and wool production

# 3. Bias

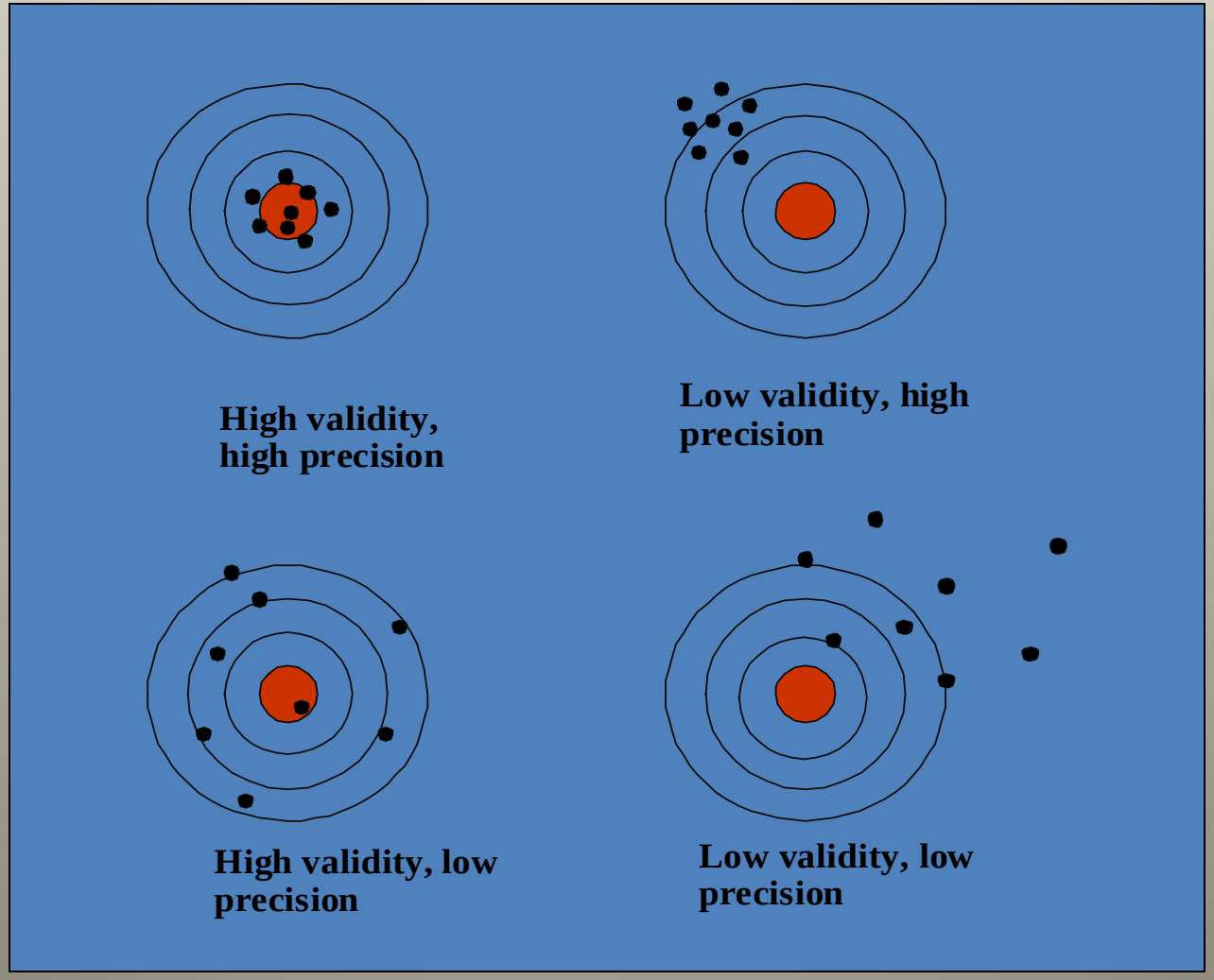
- Bias:
- Systematic error in the data
  - (measurement, collection, analysis)
  - NOT random variation or imprecision.

Bias is caused by messed-up study designs.

**"Validity"** means lack of bias.

- A valid study is one that obtained the correct risk estimates

# 3. Validity vs precision



# 3. Bias

The only thing worse than a small amount of bad (biased) data is a large amount of bad data -- why is this true?

Impact of biases

1. Make factor seem important when it is not
2. Make factor seem unimportant when it really is
3. Under or overestimate the "true risk"

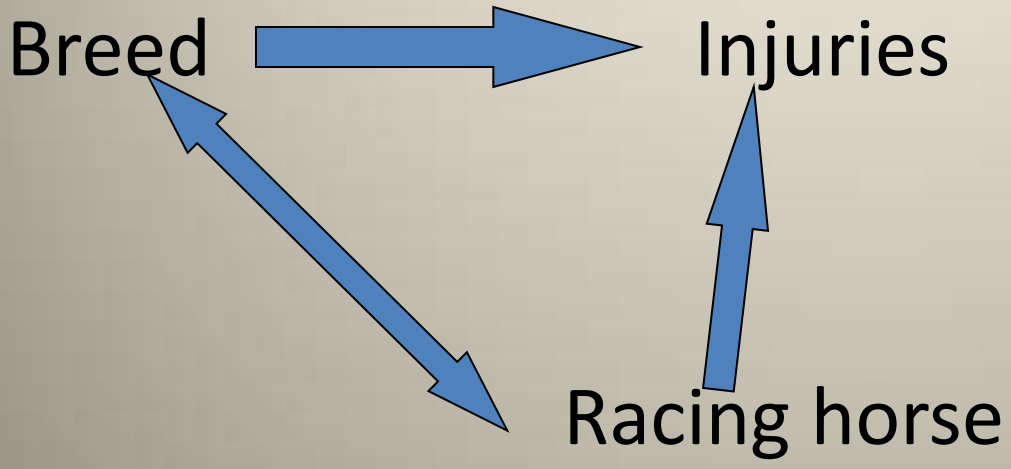
# Bias categories

a) Selection bias - bias in sample selection or allocation

b) Information bias - bias in data gathering or measurement.

c) Confounding - bias due to failure to account for a 3rd unknown variable in the design or analysis

# Confounding



# Confounding

- Variable (e.g. breed/purpose) must have a different distribution in the populations being compared
- Must be a risk factor for disease (ie. must affect the rate being calculated)
- Factor must not be on causal pathway

# Some strategies to reduce bias

1. Selection bias
   - Random selection or allocation
   - Formal clearly defined criteria

2. Information bias
   - Blinding of measurer
   - Improved measuring device
   - Standardized protocol
   - Objective vs subjective criteria
   - Use of prospective vs retrospective studies

# Some strategies to reduce bias

3. Confounding
    - Randomization (random allocation)
    - Restriction to certain groups
    - Matching
    - Statistical adjustment
        - Stratified analysis
        - Modeling

# 4. The Surgeon-General's Criteria

- *Koch's postulates* - developed for highly virulent infectious agents.
- Did not consider the influence of environmental and management factors nor were they applicable to non-infectious disease.
- The Surgeon-General put together newer criteria on which to base decisions about disease causation.

# 4. The Surgeon-General's Criteria

These criteria are:

a) Time sequence
b) Strength of association
c) Dose-response relationship
d) Consistency of findings
e) Biologic plausibility
f) Specificity
g) Analogy

# a) Time sequence

Time sequence - A cause must always occur before its effect.

Choice of study design influences the ability to do this.

List some reasons why it might be difficult to establish the temporal sequence in some studies:

# b) Strength of Association

- Strength of association usually is measured by a statistic such as a correlation coefficient, a relative risk, odds ratio, or an attributable risk.

- Strength is **NOT** measured by the size of the P value (as long as the statistic is significant).

# b) Strength of Association

**RR: The relative risk (also called "risk ratio", "relative incidence rate ratio", etc.)**

Ratio of the incidence (IR) in the exposed group to the incidence in the unexposed group.

# RELATIVE RISK

- Tells you how many more times likely disease was in the risk-factor positive (exposed) group compared with the risk-factor negative (non-exposed) group
- RR ranges from 0 to ∞; equal risk = 1
- If you get a RR between 0 and 1, the RR is telling you that exposure was protective

# Relative Risk Calculation

| | Disease Status | | | |
|---|---|---|---|---|
| | | + | - | Total |
| Risk factor | + | a | b | a+b |
| | - | c | d | c+d |
| | | | | |

IR exp = a/(a+b);  IR non-exp = c/(c+d)

RR = a/(a+b) ÷ c/(c+d)

# Example

| | NCD status | | | |
|---|---|---|---|---|
| | | + | - | Total |
| Poor biosecurity | + | 12 | 61 | 73 |
| | - | 1 | 73 | 74 |
| | | | | |

IR exp = 12/73 = 0.164

IR non-exp = 1/74 = 0.0135

RR = 12.1

# c) Dose-response relationship

- A demonstrable dose-response relationship (linear or curvilinear) between the risk factor is--like strength of association--strong evidence for causation if present, but only indeterminate evidence if absent.
- For the factor, confinement status of the herd, it might be difficult to do this --- may be possible to have an intermediate category of partial confinement between total confinement and no confinement.

# c) Dose-response relationship

| Example | Confinement | PRV incidence% |
|---|---|---|
| | Total | 16.4 |
| | Partial | 3.4 |
| | Nil | 1.4 |

List some circumstances where it might be difficult (or impossible!) to show a dose-response relationship between a factor and a disease

# d) Consistency of association upon replication

- Have several studies found a relationship between confinement and PRV risk?
  - this could of course be the first study of this relationship.

# e) Biologic plausibility

- Is there an underlying mechanism that makes biologic sense?

# f) Specificity of the association

- Specificity refers to the extent of "1- to -1" correspondence between the cause and the effect. Perfect specificity would imply that the risk factor has no effect other than the one being studied.
- Also, perfect specificity would imply that the risk factor was both a **NECESSARY CAUSE** (the disease can't happen without the risk factor) and a **SUFFICIENT CAUSE** (the disease always occurs if the risk factor is present.

# g) Analogy

- It's easier to believe in the causal nature of an association if the situation is analogous to another one which we already know to be causal.

# OBSERVATIONAL STUDY DESIGNS

- A chief advantage of observational studies is that they are directed towards the species of concern in its natural environment.
- This greatly reduces the problems associated with extrapolation of results from a particular study to the target population.
- It also allows the investigator to test a much broader range of hypotheses than would be possible under controlled experimental conditions.

# OBSERVATIONAL STUDY DESIGNS

- A major disadvantage of observational studies is the possibility of biases especially confounding which may distort the true relationships between variables.
- Missing data are also a common problem in prospective studies because animals are sold, die, or lose their identification.