

[Welcome](#) [Module Index](#) [Glossary](#) [Epiville Chamber of Commerce](#) [About this site](#) [Requirements](#)

Glossary

[A](#) [B](#) [C](#) [D](#) [E](#) [G](#) [I](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [V](#)

A

Adjusted Effect Estimates

Calculation of the effect estimate (in this case odds ratio) necessarily implies pooling the data together and producing an average estimate of risk. However, there might be situations when crude estimates are misleading. In particular, if there is a strong confounding in the data. There are several ways to deal with confounding in the data. At the study analysis stage, we can stratify the data into several groups based on the level of the confounding variable, calculate estimates of effect in each strata and then combine them by using a Mantel-Haenszel procedure. This statistic combines information across partial tables and enables you to calculate one common OR, as opposed to many for each strata.

Administrative Work (Study-Specific)

1. Obtain permission from Glop Industries regarding use of their internal data.
2. Get approval from your Institutional Review Board (IRB) - this will ensure that the study adheres to the ethical principles of conducting public health research and that the rights of the study participants are protected.
3. Prepare a budget and get funding.
4. Develop an operations manual to be used by all study personnel which will describe the standardized procedures for collecting and managing data.
5. Design a consent form for study participants that clearly indicating the study's goals, risks, benefits, expenses that participants might incur if they decide to participate in the study, how participants might benefit from their participation, and how you plan to make use of the data once they are collected.
6. Design a structured questionnaire that will be used to collect data from the study participants.
7. Hire and train interviewers.
8. Design a data management plan (e.g., how and where paper will be stored, when and how they will be entered into the computer database, how the data cleaning will be performed, and how often checks of the data will be performed to spot possible problems in the study).
9. Design a data analysis plan and propose how you will publicize the findings.

Association

Statistical dependence between two or more variables (Aschengrau & Seage, p. 184)

Attack Rates

Attack rate is really the proportion of exposed people that become ill. There are two types of attack rate: primary attack rate and secondary attack rate. An attack rate is also an incidence rate (we discussed this in Study 1 of the SARS exercise). An attack rate is used when the occurrence of disease among a population at risk increases dramatically over a short period of time.

B

Blinding (referred to as "masking" in Aschengrau & Seage)

The process of preventing researchers and subjects participating in an experimental study from gaining knowledge of their treatment status.

- In a "single-blind" experimental study, the study subjects are unaware of their treatment assignment.

- In a "double-blind" experimental study, both the study subjects and the investigators monitoring the effects of the treatment are unaware of the treatment assignment
- In a "triple-blind" experimental study, the study subjects, the investigators administering the treatment, and the investigators monitoring the effects of the treatment are unaware of the treatment assignment.

Source: Aschengrau & Seage, p.184; Rothman and Greenland, p.69

C

Case-Control Study

The observational epidemiologic study of persons with the disease (or other outcome variable) of interest and a suitable control (comparison, reference) group of persons without the disease (Last, p.22).

Case Definition

The case definition is the list of specific criteria used to decide whether or not a person has the disease of concern.

Data collection should not start until the case definition has been established because the case definition determines the data needed in order to classify persons as affected or unaffected.

The case definition is based on:

1. Clinical criteria (signs and symptoms found upon physical examination of cases)
2. Distribution of cases with person, place, and time (PPT)

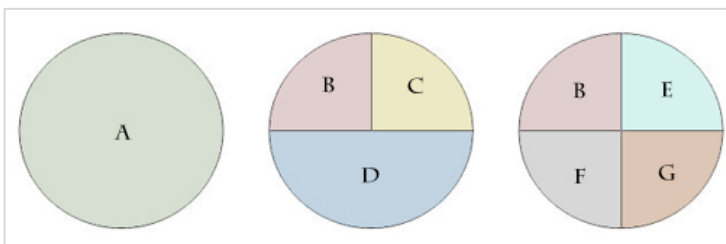
Case-fatality Ratio (Giesecke, pp.11-12)

The term is used almost exclusively in infectious disease epidemiology, while people outside our field tend to use the term 'lethality' (or inversely the 'rate of survival', which has a more positive ring to it). They both mean the same thing, namely the proportion of people who will die of a certain disease out of those who contract it.

Causal Heuristic

Sufficient-Component Cause Model: Causal Heuristic

A *heuristic* is a learning tool used to simplify concepts. In epidemiology, the sufficient-component cause model described by Ken Rothman is an example of a heuristic which shows the multicausal nature of disease. Under this model, a disease can be caused by any completed "pie," which is itself comprised of component causes of the disease under investigation. Causes which are present in every "pie" are called "necessary" causes, causes which are the only component of a "pie" are called "sufficient" causes, and causes which are neither necessary nor sufficient are called "component" causes.



In the above diagram, cause A is a sufficient cause, while the remaining causes are neither necessary nor sufficient component causes.

Causation

Causal agent is an event, condition, or characteristic that preceded the disease event and without which the disease event either would not have occurred at all or would not have occurred until some later time (Rothman & Greenland). While many theories of causation have been proposed, the most popular one postulates that 'a cause is something that makes a change' (Susser). Thus, causation is a theory describing the relationship between the cause and the outcome.

Central Limit Theorem

Given a population of any non-normal functional form with a mean μ and finite variance σ^2 , the sampling distribution of \bar{x} , computed from samples of size n from this population, will have a mean μ and a variance σ^2/n and will be approximately normally distributed when the sample size is large. (Daniel, p.130)

Cohort Study

The analytic method of epidemiologic study in which subsets of a defined population can be identified as exposed to a factor(s) hypothesized to influence the probability of occurrence of a given disease or other outcome (Last, p.33)

- **Ambidirectional Cohort:** A cohort study that shares components of both the prospective and retrospective designs (Aschengrau & Seage, pp. 206-208)
- **Prospective Cohort (a.k.a., Concurrent Cohort Study):** Participants are grouped on the basis of past or current exposure and are followed into the future in order to observe the outcome(s) of interest. (Aschengrau & Seage, pp. 206-208)
- **Retrospective Cohort:** Both the exposures and outcomes have already occurred by the time the study begins (Aschengrau & Seage, pp. 206-208)

Concordant Pairs

Case-control pairs in which both the case and control were exposed ('a' cell in a 2x2) or in which both were unexposed ('d' cell in a 2x2)

Confidence Interval

The computed interval with a given probability, e.g., 95%, that the true value of a variable is contained within that interval (Last, p. 37). Thus, if we were to repeat our study 100 times, 95 times the estimated measure of effect will lie within this interval.

Confounding

Distortion of the estimated effect of an exposure on an outcome, caused by the presence of an extraneous factor associated both with the exposure and outcome but not an intermediate step in the causal pathway between exposure and outcome (Last, p. 37)

Coronavirus

Coronaviruses were first isolated from chickens in 1937. After the discovery of Rhinoviruses in the 1950's, ~50% of colds still could not be ascribed to known agents. In 1965, Tyrrell and Bynoe used cultures of human ciliated embryonal trachea to propagate the first human coronavirus (HCoV) in vitro. There are now approximately 15 species in this family, which infect not only man but cattle, pigs, rodents, cats, dogs and birds (some are serious veterinary pathogens, especially chickens).

Correlation

The degree to which variables change together. A correlation coefficient indicates the degree to which two variables have a linear relationship (Last, p. 41)

Counterfactual

In this context, it is the condition that is counter to the factual condition, the "counterfactual." Exposed people have been exposed. That is the fact. The experience of exposed people were they not exposed is the counterfactual. Although we can't observe the counterfactual scenario, what we always want to know is the difference between the fact and the counterfactual. That is the difference in the disease experience of the exposed given that they are exposed (the fact) and the disease experience of the exposed had they not been exposed (the counterfactual).

(Aschengrau & Seage, pp. 212,288)

Cumulative Incidence

Incidence is defined as the number of individuals who fall ill with a certain disease during a defined time period, divided by the total population. In most instances, incidence is calculated from clinical cases, but by following people with serological test it is possible to detect the subclinical cases, and thus to obtain an incidence figure for the true number of infections. (Giesecke, p. 8-9)

D

Descriptive Analysis

Descriptive analysis is concerned with describing the general characteristics of the distribution of a disease. Descriptive studies often provide the first important clues about possible determinants of a disease and are primarily useful for the formulation of hypotheses that can be tested subsequently.

Detectable pre-clinical phase

The stage in a disease from the pathological onset to the first appearance of signs and symptoms during which it can be identified by a screening test. Thus the detectable pre-clinical phase is a function of both the latent period of a disease and technical capabilities of a screening test. (Aschengrau & Seage, p. 416)

Discordant Pairs

Case-control pairs in which the case was exposed and the control was not ('b' cell in a 2x2), or pairs in which the control was exposed and the case was not ('d' cell in a 2x2).

DNA adducts

The initial step in chemical carcinogenesis is characterized by attachment of the chemical to DNA to produce DNA adducts. The chemicals can alter the structure of the DNA and in turn, the biological processing of the DNA by cellular proteins governing replication, transcription and repair. If not repaired or repaired incorrectly, these modifications may eventually lead to mutations and ultimately cancer, especially if the adduct is located in an oncogene or tumor suppressor gene. Thus, different DNA adducts can affect cellular responses to DNA damage.

E

Ecological Fallacy

The bias that may occur because an association observed between variables on an aggregate level does not necessarily represent the association that exists at an individual level.

Ecological Study

A study in which the units of analysis are populations or groups of people, rather than individuals.

Eligibility Criteria

An explicit statement of the conditions under which persons are admitted to an epidemiologic study (Last, p. 58).

Elimination

Reduction of the incidence of infection (disease) caused by a specific agent below detectable levels in a defined geographic area.

Endemic

When an infection disease lingers at around the same incidence for a long time, it is considered endemic. (Giesecke, p. 19)

Epidemic

Analysis of disease incidence by person, place, and time is used to determine if an epidemic is occurring. An epidemic is defined as " the occurrence in a community or region of cases of an illness, specific health-related behavior, or other health related events clearly in excess of normal expectancy. The community or region and time period in which the cases occur are specified precisely. (Giesecke, p. 19)

Equipoise

A guiding principle in human experimental research which states that, in order for it to be ethical to assign individuals to an exposure (or the lack of an exposure, in the case of the placebo arm of a trial), there must be genuine confidence that a treatment may be worthwhile to justify administering it to some individuals, balanced by genuine reservations about the treatment to justify withholding the treatment from others. (Aschengrau & Seage, p. 155)

Eradication

Termination of all transmission of infection by extermination of the infectious agent through surveillance and containment.

Exchangeability

Exchangeability occurs when the unexposed group is a good proxy (i.e., approximation) for the disease experience of the exposed group had they not been exposed. Of course, we can't know what the disease frequency in the exposed group would have been if they had not been exposed (this is the unobservable "counterfactual"), so instead we choose an unexposed group as a substitute for it. In a case-control study, we select subjects on disease status, not exposure status, so while we conceptually want the unexposed group to represent the disease experience of the exposed group (had they not been exposed), we have to think a little differently. Specifically, the control group should represent the exposure distribution in the underlying source population from which the cases arose. If you over-sample one of the following 2 control groups (E+/D-, E-/D-) in the underlying population (i.e., sample *dependent* on exposure), then you will end up with a control group that does not represent the prevalence of the exposure in that underlying population, and therefore your effect estimate (OR) will be biased. This is why control group selection is a very tricky and important part of the case-control studies.

Experimental Study Design

Experimental studies are characterized by the investigator assigning an exposure of interest to individuals or populations for the purposes of comparing the effect of the exposure on an outcome of interest (Aschengrau & Seage, p. 142). The active manipulation of the exposure by the investigator is the hallmark that distinguishes experimental designs from observational ones (Aschengrau & Seage, p. 169). The most common type of experimental design used in epidemiology is the Randomized Trial.

External Validity (generalizability)

The ability to generalize the results from a given study to populations beyond the study subjects. Evaluation of external validity requires review of the study methods, the makeup of the study population, and subject-matter knowledge such as the biological basis of the association (Aschengrau & Seage, p. 262).

G

Gold standard

(Jargon) A method, procedure, or measurement that is widely accepted as being the best available. Often used to compare with new methods.

Group-Level Variables

Variables defined at a population or aggregate level (e.g., poverty index).

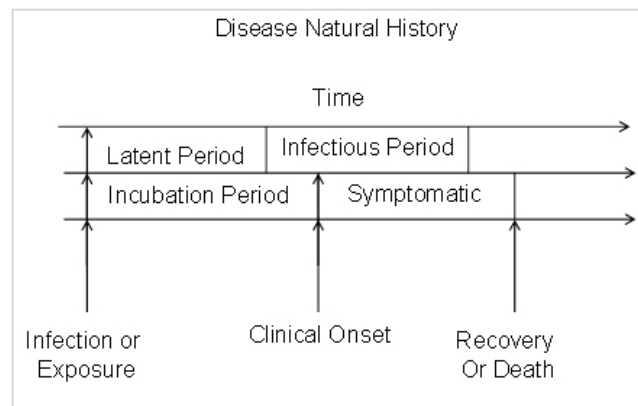
I

Incidence proportion

The cumulative proportion of a population that becomes newly diseased over a specified period of time (Aschengrau & Seage, p. 42). In epidemiological studies, incidence proportion is used synonymously with risk, attack rate and cumulative incidence.

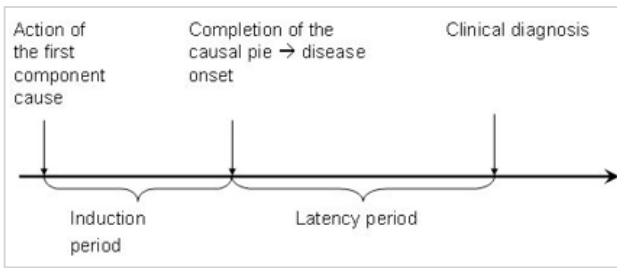
Incubation period

Different diseases have different incubation periods. There is no such thing as a precise incubation period but rather a range of incubation periods that is characteristic of a particular disease.



Induction time

There are two ways of looking at induction time. Since the exposure may need to accumulate to a certain threshold, other factors have to be present before the disease can occur. Accumulation takes a period of time termed 'the induction period.' Alternatively, the exposure may be the first event in a series of causal events that must occur for the disease to develop. For example, Susser Syndrome may be the result of a) a susceptible individual's accumulated exposure to a chemical, leading to b) a genetic damage which leads to c) a decrease of a certain neurotransmitter. This process may take months or years. Individual susceptibility varies based on specific biological/ physiological factors. (Giesecke, p.176-177)



Informed Consent

The process of gaining the agreement of eligible individuals to participate in a study. During this process, investigators describe the nature and objectives of the study, the tasks required for participants, and the benefits and risks of participating. After this information is disseminated, investigators ask for the consent of the potential participants (Aschengrau & Seage, p. 143)

Intent-to-Treat Analysis

A method of analysis in a randomized trial that compares the outcome of interest between study groups based on the treatment to which they were randomized, regardless of whether the individuals actually took their assigned treatment or not (Aschengrau & Seage, p. 145).

Internal Validity

Refers to the ability of a study to make the inference that an observed association between exposure and outcome is a plausibly causal relationship (Shadish, Cook, and Campbell, p. 37). Common threats to the internal validity of epidemiological study designs are bias, confounding, and random error (Aschengrau & Seage, p. 262).

Isolation

An emergency measure used in outbreaks of a highly lethal and contagious disease in which all individuals known to be infected with the disease are moved into hospital environments designed to treat the illness while minimizing the chance of spreading the infectious agent.

K

Kappa statistic

A measure of the degree of nonrandom agreement between observers or measurements of the same categorical variable. (Last) Complete agreement corresponds to $K = 1$, and lack of agreement corresponds to $K = 0$.

L

Latent period

Interval between disease onset and clinical diagnosis (Aschengrau & Seage, p. 222).

Lognormal Distribution

The lognormal distribution is frequently used in Biostatistics and Epidemiology. A random variable x is said to have the lognormal distribution, with parameters μ and σ , if $\ln(x)$ has the normal distribution with mean μ and standard deviation σ . Log of time distribution of

cases (x-axis) turns the epidemic curve into normal distribution with many valuable statistical properties.

Loss to Follow-Up

Study subject(s) who cannot or do not complete participation in a study for various reasons (Last, p. 108).

M

Matching

In this study, it is necessary to perform a statistical analysis appropriate for a matched case-control study since the authors matched on age (in decades), sex, hospital, and hospital- room status (private, semiprivate, or ward). Since cases and controls are matched on these factors, you can no longer elucidate the effects of these variables. Cases can be matched individually or frequency matched on particular variable. In the first scenario, one or more controls are selected to match a particular case on a set of variables, while in the second the controls are selected in such a way that their distribution on a set of variables resembles that of the cases.

Matched Case-Control Study

The observational epidemiologic study of persons with the outcome of interest and a suitable control (comparison, reference) group of persons without the disease that are chosen based on specific factors of interest (i.e., potential confounding variables by which the investigators match controls to cases [e.g., gender, age, race/ethnicity, etc.]) (Last, p. 22)

Modes of Transmission

Several different classifications exist for the routes of transmission of the different infections. These have been generated mostly for the purpose of grouping similar disease together in handbooks on preventive measures, and none of them is entirely satisfactory. Common classifications include person-to-person spread, air-borne, water-borne, food-borne, and vector-borne infections. (Giesecke, pp. 16-17)

Multivariate Analysis

The process of creating mathematical models to assess the association(s) between exposures, outcomes, and confounders (See Aschengrau & Seage, pp. 299-300). The mathematical model employed depends on the data in your study and include: linear regression, logistic regression, Cox proportional hazard, and Poisson models. Typically, investigators evaluate potential confounders first by stratification. If, upon stratification, the crude estimate is changed by 10% or more, the variable is retained and controlled for using multivariate techniques.

N

Negative predictive value (predictive value negative)

The proportion of individuals without preclinical disease who test negative. $\text{Number of individuals who test negative and do not have preclinical disease} / \text{number of individuals who test negative}$ (Aschengrau & Seage, p. 423)

Neoplastic

A disease or lesion characterized by abnormal new growth of tissue, often used as a synonym for 'malignant tumor.'

Non-neoplastic

A non-cancerous, non-malignant, or benign disease or lesion.

O

Observation Bias

A flaw in measuring exposure or outcome data that result in different quality (accuracy) of information between comparison groups. (Aschengrau & Seage, p. 270) Different types of observation bias include:

1. **Recall bias:** a result of differential level of accuracy in the information provided by compared groups. (Aschengrau & Seage, p. 271)
2. **Interviewer bias:** systematic differences in soliciting, recording, or interpreting information that occurs in studies using in-person or telephone interviews.
3. **Losses to follow-up:** When a subject whose outcome is unknown can no longer be followed in the course of a study, that individual is considered lost to follow-up.
4. **Differential and non-differential misclassification**

Observational Study Design

Observational studies refer to the broad class of epidemiological study designs characterized by the fact that exposure is not assigned by the investigator. Rather, the investigator passively observes as nature takes its course (Aschengrau & Seage, pp. 140-141).

Two subsets of observational study are descriptive and analytical studies

- Descriptive observational studies are not designed to test a causal hypothesis, but collect information on the distribution of disease patterns in terms of the characteristics of person, place, and time (Aschengrau & Seage, p. 99).
- Analytical observational studies assemble a population for study for the purposes of testing a causal hypothesis concerning the relationship between exposure and disease. These include the cohort and case-control studies.

Outbreak

An epidemic limited to a localized increase in the incidence of a disease, e.g., in a village, town, or closed institution.

Outbreak Management

The process of anticipating, preventing, preparing for, detecting, responding and controlling outbreaks in order to minimize their health and economic impact.

Outcome

All the possible results that may stem from exposure to a causal factor, or from preventive or therapeutic interventions (Last, p. 129).

P

Pandemic

A worldwide epidemic involving millions of people.

Person-Time

A measurement combining persons and time (days, months, years etc.) as the denominator in incidence and mortality rates when, for varying periods, individual subjects are at risk for developing disease or dying (Last, p. 134).

Placebo Control

An inactive treatment administered to the comparison arm of a placebo-controlled randomized trial which is designed to match as closely as possible the experience of the comparison group with that of the active treatment group. Placebos aid in the masking of subjects and investigators by attempting to prevent them from knowing their true treatment status (Aschengrau & Seage, p. 184).

Population at Risk

All the inhabitants of a given area that have the potential to develop the outcome of interest.

Positive predictive value (predictive value positive)

The proportion of individuals with a positive test who have preclinical disease. $\text{Number of individuals who test positive and have preclinical disease} / \text{number of individuals who test positive}$ (Aschengrau & Seage, p. 423).

PPT (Person, Place, and Time)

PPT data is crucial to determine whether someone is, or is not, a suspected case. Key questions in determining PPT are:

1. **Person:** Was the person connected to other suspected cases of the disease?
2. **Place:** Was the person connected by place to where the other cases came from?
3. **Time:** Were the symptoms temporally related to other known cases of the disease?

Prevalence

Prevalence is the total number of people who have that disease at a specific time, divided by the total population. (Giesecke, p. 9).

Primary prevention

The maintenance of health through individual or community efforts so that the disease process never starts (Aschengrau & Seage, p. 413).

Proprioception

Proprioception refers to the unconscious perception of movement and spatial orientation. Individuals without it lack the ability to direct the body's various parts to move and must resort to using visual clues. Impaired proprioception is routinely tested by police officers via a "field sobriety test," which asks individuals suspected of being under the influence of alcohol or other drugs to perform a set of tasks (such as touching their nose with their finger) with their eyes closed. In more severe cases, individuals lacking proprioception are unable to stand or walk or move their limbs without conscious effort.

Q**Quarantine**

An emergency measure used in outbreaks of highly lethal and contagious diseases in which all individuals suspected of having contact with infected individuals are kept apart from the general population and carefully monitored for signs of the disease.

R

Randomization

The process of allocating individuals to groups (i.e., exposed and unexposed) by chance (Last, p. 150). If the sample size is large enough, randomization helps balance the distribution of known and unknown confounders between study groups (See Aschengrau & Seage, p. 22).

Randomized Trial

An experimental study design in which exposure is randomly assigned, and in which the frequency of the outcome of interest is compared between one or more groups receiving an experimental treatment and a group receiving a placebo or the current standard of care.

Rate Ratio

- A rate is an expression of the frequency with which an event occurs in a defined population in a specified period of time.
- The numerator of a rate is the number of events (e.g., disease occurrences), while the denominator is usually person-time.
- Using the 2x2 table (Please see Aschengrau & Seage, Chapter 3):

	Disease +	Disease -	Person-Time (PT)
Exposure +	a	--	PT exposed
Exposure -	c	--	PT unexposed
Total	a+c	--	Total PT

Rate of Disease in the Exposed: $a/\text{Person-Time exposed}$

Rate of Disease in the Unexposed: $c/\text{Person-Time unexposed}$

- The rate ratio is the ratio of two rates, usually exposed/unexposed:
Rate Ratio: $[a/\text{Person-time exposed}]/[c/\text{Person-time unexposed}]$
- Interpretation: During time t , the rate of developing disease was (n times higher/ n times less likely) in exposed persons compared with non-exposed persons in this population per person-time of observation.

Reliability

The degree of stability exhibited when a measurement is repeated under identical conditions. Reliability refers to the degree to which the results obtained by a measurement, procedure can be replicated.

Reproductive Rate

(also basic reproduction number (R_0))

The average number of new infections that one infectious case generates during his/her infectious lifetime in a community of susceptible individuals.

Restriction

Eligibility criteria for study participation is limited to specified categories of a confounder (e.g., between the ages of 25 and 35, women only).

Specified categories of a confounder (e.g., between the ages of 25 and 35, women only).

Risk Difference

Using the 2×2 table (Please see Aschengrau & Seage, Chapter 3):

	Disease +	Disease -	Total
Exposure +	a	b	a+b
Exposure -	c	d	c+d
Total	a+c	b+d	a+b+c+d

Probability of Disease in the Exposed: $a/a+b$

Probability of Disease in the Unexposed: $c/c+d$

The risk difference is the absolute difference between two risks, usually exposed - unexposed:

$$RD = [a/a+b] - [c/c+d]$$

The risk difference measures clinical and public health importance of the causal relationship.

Risk Ratio (Relative Risk)

a. Risk is the average probability that an event (disease) will occur given particular exposure (Last, p. 159).

b. Using the 2×2 table (Please see Aschengrau & Seage, Chapter 3):

	Disease +	Disease -	Total
Exposure +	a	b	a+b
Exposure -	c	d	c+d
Total	a+c	b+d	a+b+c+d

Probability of Disease in the Exposed: $a/a+b$

Probability of Disease in the Unexposed: $c/c+d$

c. The risk ratio is the ratio of two risks, usually exposed/unexposed: $RR = [a/a+b]/[c/c+d]$

d. Interpretation: During time t, the risk of developing disease was (n times higher/n times less likely) in the exposed persons compared with non-exposed persons in this population.

S

Secondary Attack Rate

Secondary attack rate refers to the spread of disease in a family, household, dwelling unit, dormitory, or similar circumscribed group. The spread of infection from an index case (the initial case, i.e. the case that introduced the organism into the population) to the attending medical staff is called secondary attack rate. It is a good measure of person-to-person spread of disease after the disease has been introduced into a population.

Screening

The presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures which

can be applied rapidly. Screening tests sort out apparently-well persons who probably have a disease in a pre-symptomatic stage from those who probably do not.

Secondary prevention

The reduction in the expression and severity of clinical disease among asymptomatic individuals (Aschengrau & Seage, p. 414).

Selection Bias

An error due to selection of cases and controls based on differing criteria that are related to exposure status, or selection (or follow-up) of exposed and unexposed individuals in a way that is related to the development of the outcome. (Aschengrau & Seage, p. 263)

Different types of selection bias include:

1. **Control Selection bias:** A result of selecting controls from a different source population than the cases. (Aschengrau & Seage, pp. 264-266)
2. **Self-selection bias:** A type of bias which can result from differential rates of participation between cases and eligible controls. (Aschengrau & Seage, p. 266)
3. **Differential surveillance, diagnosis, or referral bias:** can result from a tendency to hospitalize patients differentially based on their exposure status. Ex. Oral contraceptives and thromboembolism. (Aschengrau & Seage, p. 267)

Sensitivity

The probability that a test correctly classifies as positive individuals who have preclinical disease. $\text{Number of individuals with preclinical disease who test positive} / \text{number of individuals with preclinical disease}$ (Aschengrau & Seage, p. 421).

Source Population

The underlying cohort, or study base, representing the group of subjects that gives rise to cases. Controls should be selected to represent the proportions of exposed and non-exposed persons in the source population to help lessen the introduction of bias.

Specificity

The probability that a test correctly classifies individuals without preclinical disease as negative. $\text{Number of individuals without preclinical disease who test negative} / \text{number of individuals without preclinical disease}$ (Aschengrau & Seage, p. 421).

Standardized Incidence Ratio (SIR)

The ratio of the number of incident cases of a specified condition in the study population to the incident number that would be expected if the study population had the same incidence rate as the standard or population for which the incidence rate is known (Last, p. 172).

Stratified Analysis

Stratification is used both to evaluate and control for confounding and requires separating your sample into subgroups, or strata, according to the confounder of interest (e.g., by age, gender, race/ethnicity, etc.). Because each stratum is homogeneous with regard to the confounder of interest, one can then evaluate the association between exposure and disease within each stratum (e.g., the odds ratio for women only and the odds ratio for men only). Once you have conducted stratified analyses you have the option to combine your data to derive a summary, or pooled estimate. One of the most common techniques for pooling data is the Mantel-Haenszel procedure. This statistic combines information across partial tables and enables you to calculate one common odds ratio, as opposed to many for each strata (See Aschengrau & Seage, pp. 297-300).

Spot Map

Map showing the geographic location of people with a specific attribute, e.g., cases of a disease or elderly persons living alone. The making of a spot map is common procedure in the investigation of a localized outbreak of disease

Surveillance

The continuous collection and analysis of data, with or without subsequent action. Several types of surveillance systems used include; centralized disease registries, microbial laboratory monitoring systems, hospital discharge notes, etc. The main task of a surveillance system is to allow for the detection of unexpected changes in disease incidence. (Giesecke, pp. 148-159)

T

Tertiary prevention

The slowing or blocking of the progression of a disease among individuals for whom a clinical diagnosis has been made (Aschengrau & Seage, p. 415).

Tumor registry

Tumor registries are established with the sole purpose of gathering and disseminating current epidemiologic data on all primary tumors, usually malignant, but sometimes benign and malignant, for the purposes of accurately describing the incidence and survival patterns, evaluating diagnosis and treatment, facilitating etiologic studies, establishing awareness of the disease, and ultimately, for the prevention of all tumors. In the U.S. there is no one unified tumor registry as is the case in some European countries. However, the Surveillance, Epidemiology, and End Results (SEER) Program of the [National Cancer Institute](#) is an authoritative source of information on cancer incidence and survival in the United States. The SEER Program currently collects and publishes cancer incidence and survival data from 14 population-based cancer registries and three supplemental registries covering approximately 26 percent of the US population. Information on more than 3 million in situ and invasive cancer cases is included in the SEER database, and approximately 170,000 new cases are added each year within the SEER coverage areas. The [SEER Registries](#) routinely collect data on patient demographics, primary tumor site, morphology, stage at diagnosis, first course of treatment, and follow-up for vital status. The SEER Program is the only comprehensive source of population-based information in the United States that includes stage of cancer at the time of diagnosis and survival rates within each stage. The mortality data reported by SEER are provided by the [National Center for Health Statistics](#).

Types of epidemic

- **Point source** - An epidemic in which all cases are infected at the same time, usually from a single source or exposure. (Giesecke, pp. 135-136)
- **Continuous source** - An epidemic in which the causal agent (e.g. polluted drinking water, spoiled food) is infecting people who come into contact with it, over an extended period of time. (Giesecke, pp. 135-136)
- **Person-to-Person** (a.k.a. Propagated) - An epidemic in which the causal agent is transmitted from person to person, allowing the epidemic to propagate and spread as a function of R_0 . (Giesecke, pp. 135-136)

V

Validity

Validity, measurement. An expression of the degree to which a measurement measures what it purports to measure. Validity, study. The degree to which the inference drawn from a study is warranted when account is taken of the study methods, the representativeness of the study sample, and the nature of the population from which it is drawn.

References

- Last, JM (Ed). A Dictionary of Epidemiology, Fourth Edition. Oxford University Press, 2001.
- Giesecke, J. Modern Infectious Disease Epidemiology, Second Edition. Oxford University Press, 2002.
- Aschengrau, A, Seage, GR. Essentials of Epidemiology in Public Health, Second Edition. Jones and Bartlett Publishers, 2008.
- Rothman, K., Greenland, S. Modern Epidemiology, Second Edition. Lippincott, Williams, & Wilkins, 1998.

