The epidemiologic approach: Steps to public health action

Center for Infectious Disease Preparedness
UC Berkeley School of Public Health
URL: http://www.idready.org

Updated June 2006

Created using freely available, open source software:
http://www.openoffice.org

What is epidemiology?

Epidemiology is the study of the *distribution* and *determinants* of health-related states or events in specified populations, and the application of this study to the control of health problems.

The epidemiologic approach: Steps to public health action

SURVEILLANCE
- Detect outbreaks & threats
- Find cases for intervention
- Monitor trends
- Direct interventions
- Evaluate interventions
- Generate hypotheses

DESCRIPTIVE
- What (case definition)
- Who (person)
- Where (place)
- When (time)
- How many (measures)

ANALYTIC
- Why (Causes)
- How (Causes)

MEASURES
- Count
- Time
- Rate
- Risk/Odds
- Prevalence

STUDY DESIGN
- Design
- Implementation
- Analysis
- Interpretation
- Reporting

THREATS TO VALIDITY
- Chance
- Bias
- Confounding

INFERENCES
- Epidemiologic
- Causal

ACTION
- Clinical
- Behavioral
- Community
- Environmental

Notes:

What's is public health surveillance?

"Public health surveillance is the ongoing, systematic collection, analysis, interpretation, and dissemination of data about a health-related event for use in public health action to reduce morbidity and mortality and to improve health."

4 PMID: 15129191

Goals of surveillance system that guide public health actions

- Detect outbreaks
- Detect public health threats
- Detect infectious cases (case finding)
- Monitor trends in a target population
- Monitor exposed individuals for symptoms
- Monitor treated individuals for complications
- Direct public health interventions
- Evaluate public health interventions
- Generate hypotheses for further evaluation

Types of disease surveillance

- *Passive* surveillance
  - *Passive* surveillance
    - Title 17 reportable diseases in California
  - *Enhanced passive* surveillance
  - *Stimulated passive* surveillance
    - Anthrax surveillance in New Jersey, Delaware, Pennsylvania, 2001 (Emerg Infect Dis. 2002;8:1073)
- *Active* surveillance
  - California Emerging Infections Program (CEIP)

Notes:

Sources of surveillance data

- Mortality data
  - Death registry, medical examiner
- Morbidity data
  - Legally reportable diseases, including cancer
- Birth registry
- Hospital discharge diagnoses (utilization data)
- Special surveys (NHANES, NHIS, CHIS)

Sources of surveillance data (continued)

- Outbreak reporting
- "Sentinel" systems
  - Influenza-like illness reporting/testing (Kaiser)
  - CDC Acute Hepatitis Sentinel County Study
  - West Nile Virus Surveillance Program
    - Encephalitis case surveillance
    - Larval and adult mosquito testing
    - Sentinel chicken testing
    - Dead bird surveillance

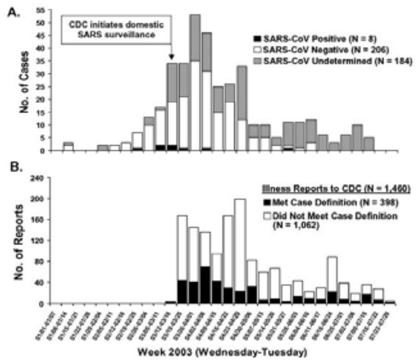
The case definition

- Inclusion criteria
  - Clinical criteria (symptoms, signs)
  - Epidemiologic criteria (person, place, and time)
  - Laboratory criteria
- Case classification
  - Suspect
  - Probable
  - Confirmed
- Exclusion criteria (for suspect and probable)
- Consider operating characteristics (sensitivity, specificity, predictive value)

Notes:

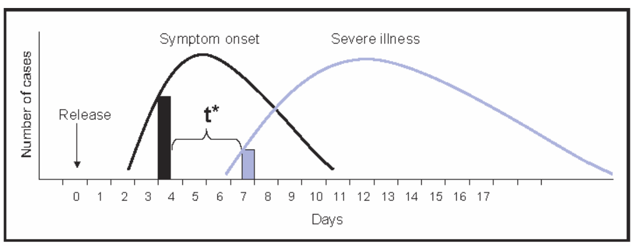
U.S. severe acute respiratory syndrome (SARS) cases reported to CDC, January–July 2003.

Displayed in Figure B is the number of unexplained respiratory illness reports received by CDC by week of illness report (N = 1,460). However, only 398 met the U.S. SARS case definition and is displayed by week of illness onset in Figure A.



Emerg Infect Dis. 2004 Feb;10(2):185

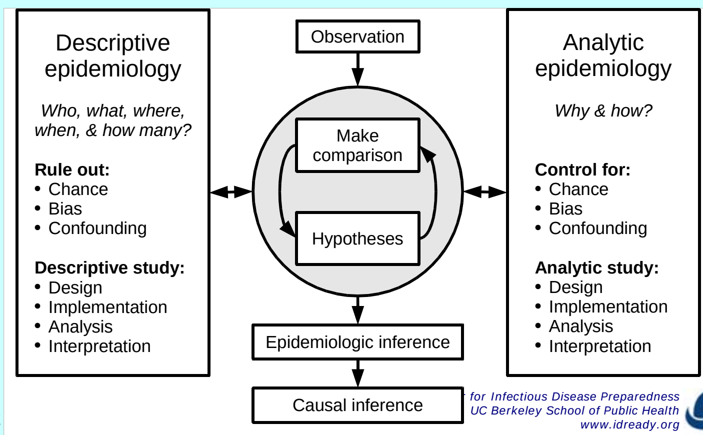
Syndromic surveillance — rationale for early detection



* t = time between detection by syndromic (prediagnostic) surveillance and detection by traditional (diagnosis-based) surveillance.

Centers for Disease Control and Prevention. Syndromic Surveillance: Reports from a National Conference, 2003. MMWR 2004:53(Suppl).

Inferences in epidemiology



Notes:

Descriptive epidemiology

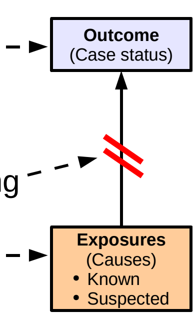
- Study of the *distribution* of health-related states or events
  - For a specific outcome, make comparisons and note differences across one or more dimensions (e.g. time series)
  - Seek known explanations to account for observed differences (rule out chance, bias, confounding as explanations)
  - Draw conclusions from descriptive study (epidemiologic inference #1)

Analytic epidemiology

- Study of the *determinants* of health-related states or events
  - Generate hypotheses from descriptive studies
  - Design and conduct studies to test hypotheses (control for chance, bias, confounding)
  - Draw conclusions from analytic study (epidemiologic inference #2)

Descriptive epidemiology in outbreak investigation

- Describe cases
  - What (case definition)
    - How many? (measures of occurrence)
    - When? (time)
    - Where? (place)
    - Who? (person)
- Rule out chance, bias, confounding
- Generate new hypotheses, if necessary



Notes:

Describing disease occurrence

- Measures of occurrence
  - Count, time, rate, risk, prevalence
- Describing disease occurrence by
  - Person (Age, ethnicity, sex/gender, risk factors)
  - Place (Country, state, county, city, census tract)
  - Time (Calendar time, seasonality)

Epidemiologic measures: Overview

- Types of measures
- How we combine numbers
  - Measures of occurrence
  - Measures of association
  - Measures of attribution (not today)

Types of measures

- Quantitative
  - Continuous numbers
  - Discrete numbers
    - Integers (..., -4, -3, -2, -1, 0, 1, 2, 3, 4, ...)
    - Counting (natural) numbers (0, 1, 2, 3, 4, ...)
- Qualitative
  - Nominal categorical (non-ordered)
    - Outcomes are usually binary (event vs. no event)
  - Ordinal categorical (ordered)

Notes:

Epidemiologic Measures: How we combine numbers

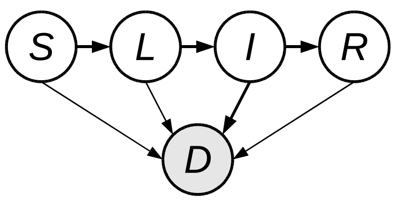
Rate	$r = \frac{\Delta x}{\Delta t}$	Change in *x* per change in *t*
Proportion	$p = \frac{a}{a+b}$	Numerator is part of the denominator
	$P = \left(\frac{a}{a+b}\right) \times 100$	Percent
Ratio	$R = \frac{x}{y}$	*x* and *y* are different and are compared

Center for Infectious Disease Preparedness
UC Berkeley School of Public Health
www.idready.org

19

Epidemiologic Measures: Measures of occurrence

- Count
- Time
- Rate
- Risk or Odds
- Prevalence



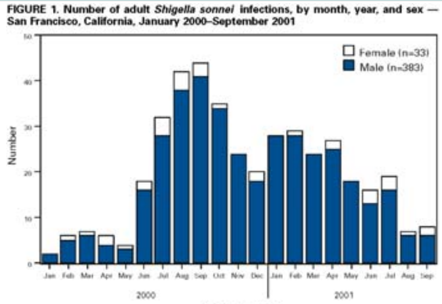
S = Susceptible
L = Latent infection
I = Infective
R = Recovered
D = Dead

Center for Infectious Disease Preparedness
UC Berkeley School of Public Health
www.idready.org

20

Community outbreak of shigellosis, San Francisco, 2000

FIGURE 1. Number of adult *Shigella sonnei* infections, by month, year, and sex — San Francisco, California, January 2000–September 2001



During June–December 2000, 230 cases of culture-confirmed* *S. sonnei* infection were reported to the San Francisco Department of Public Health; an average of 21 cases (range: 13–29 cases) occurred during the same period from 1996 to 1999. [MMWR 2004;50(42):922]

*Residents of San Francisco County aged >=15 years.

Center for Infectious Disease Preparedness
UC Berkeley School of Public Health
www.idready.org

21

Notes:

Epidemiologic Measures: Measures of occurrence - Rate

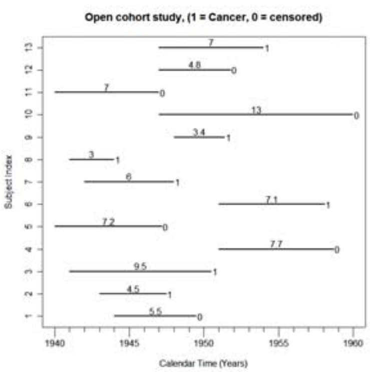
$$rate = \frac{Number\ of\ events}{Person\text{-}time\ at\ risk}$$

$$Period\ rate_{1989-1991} = \frac{D_{1989} + D_{1990} + D_{1991}}{K_{1989} + K_{1990} + K_{1991}}$$

$D_x = Events\ in\ year\ x$

$K_x = Midyear\ population\ estimate\ in\ year\ x$

Measures of occurrence – Cancer rates in an open cohort study, Person-time data



$$rate = \frac{\#cases}{\sum person\text{-}time_i}$$

$$= \frac{7\ cases}{85.7\ person\text{-}years}$$

$$= 0.08168028$$

$$= 8.2\ cases\ per\ 100\ py$$

Measures of occurrence - Period rates: Female breast cancer deaths, San Francisco

Annual period rates

CATEGORIES	YEAR 1989	1990	1991
Breast cancer deaths	125	130	131
Female population	361,975	361,401	366,613
Rate per 100,000 per year	34.5	36.0	35.7

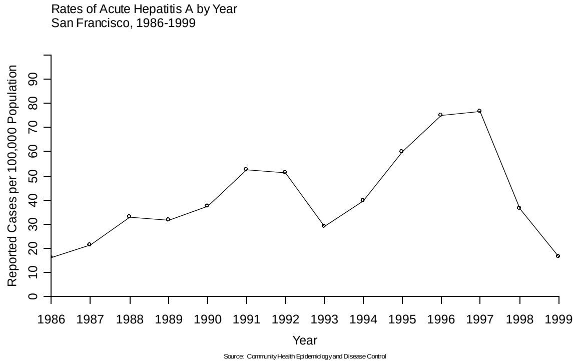
Period rates

$$r_{1989-1991} = \frac{125 + 130 + 131}{361{,}975 + 361{,}401 + 366{,}613}$$

$$= 35.4\ per\ 100{,}00\ per\ year$$

Notes:

Rates of acute hepatitis A, San Francisco, 1986-1999



Center for Infectious Disease Preparedness
UC Berkeley School of Public Health
www.idready.org

Epidemiologic Measures: Measures of occurrence - Risk

- Risk is a probability
- Probability models used to calculate risk
  - Binomial model (our focus)
  - Hazard-based models (not today)
    - Constant hazard (constant rate) model
    - Non-constant hazard model
      - Exponential formula method (fixed time intervals)
      - Kaplan-Meier method (time-to-event data)

Center for Infectious Disease Preparedness
UC Berkeley School of Public Health
www.idready.org

Measures of occurrence – Risk estimation with binomial data

$$R(0,t) = \frac{\#events\ in\ (0,t)}{Population\ at\ risk\ at\ time\ 0} = \frac{18}{111} = 0.1622$$

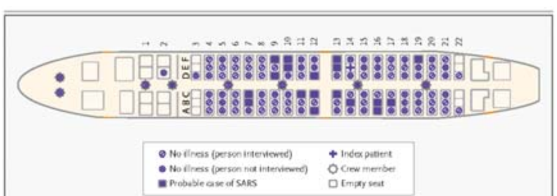


Figure 2. Schematic Diagram of the Boeing 737-300 Aircraft on Flight 2 from Hong Kong to Beijing.
Two flight attendants and two Chinese officials also reportedly had illness that met the WHO criteria for a probable case of SARS. The flight attendants are shown here as members of the crew. The seat locations of the two Chinese officials are unknown, and they are not included in the diagram.

NEJM 2003, 349;25:2416

Center for Infectious Disease Preparedness
UC Berkeley School of Public Health
www.idready.org

Notes:

Epidemiologic Measures: Measures of occurrence - Odds

Odds is a transformed risk (R) estimate

$$Odds(0,t) = \frac{R(0,t)}{1 - R(0,t)}$$

Odds from binomial data

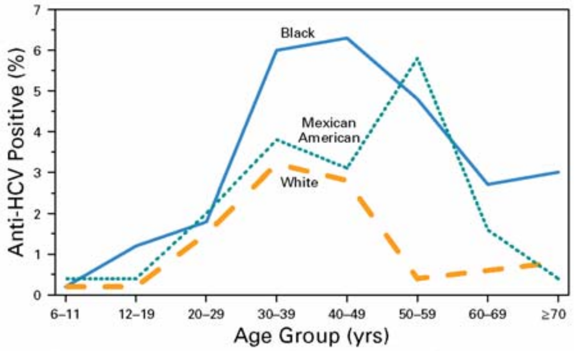
$$Odds(0,t) = \frac{\frac{a}{(a+b)}}{\frac{b}{(a+b)}} = \frac{a}{b}$$

Measures of occurrence - Prevalence

Point prevalence

$$P = \frac{\text{Number of existing cases}}{\text{Number in total population}}, \quad \text{at a point in time}$$

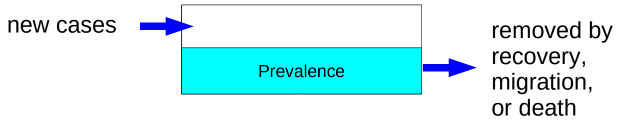
Prevalence of hepatitis C virus (HCV) infection by age and race/ethnicity — United States, 1988–1994



Source: Third National Health and Nutrition Examination Survey, CDC.

Notes:

Measures of occurrence: Prevalence, rate, and duration



Under steady state,

$$\text{Prevalence odds} = \text{Rate} \times \text{duration}$$

$$\frac{P}{1-P} = r\,d$$

$$P \approx r\,d$$

Measures of association

Measures of occurrence	Measures of association
Rate	Rate ratio
Risk	Risk ratio
Odds	Odds ratio

Outbreak of nodding off at CIDP lectures

The Centers for Disease Control (CDC) is responding to an outbreak of nodding off at training conferences provided by a west coast CDC Center for Public Health Preparedness. The UC Berkeley Center for Public Health Preparedness trainees are nodding off during lectures at an alarming rate. Local public health authorities have requested assistance from the CDC to determine the cause of this problem.

Notes:

Outbreak of nodding off at CIDP lectures

Based on initial interviews, the CDC field investigators suspect that the lectures given by Dr. Tomás Aragón, Director of the Center, is putting his students to sleep. At the next 3-day conference, investigators set up a surveillance system to measure students nodding off and who was lecturing when the nodding off occurred.

Outbreak of nodding off at CIDP lectures

Investigators set up hidden cameras in the lecture hall with unobstructed views of all the students and the lecturers. A case of nodding off was classified as probable if the student appeared to be sleeping with their eyes closed for 2 or more minutes; a case was classified as confirmed if the student snored audibly for 15 or more seconds. All student attendance times, and who was lecturing and for how long was measured.

Outbreak of nodding off at CIDP lectures

Investigators hypothesize that students exposed to Dr. Aragón were at higher risk for nodding off.

The 3-day conference had 100 attendees. There were 6 hours of lecture per day. Dr. Aragón lectured 6 of 18 hours. Risk of nodding off was estimated as the proportion that nodded off at each lecture. Rates were also calculated.

Notes:

Outbreak of nodding off at CIDP lectures: Risk ratio & Odds ratio

	Exposure Tomás	Exposure Other
Individuals nodding off	19	12
Cases of any nodding off	21	13
Number started conference	100	100

$R_1 = \frac{19}{100} = 0.19$ $\quad R_0 = \frac{12}{100} = 0.12$

Why is this analytic approach not recommended?

$RR = \frac{R_1}{R_0} = 1.58$ $\quad OR = \frac{R_1/(1-R_1)}{R_0/(1-R_0)} = 1.72$

$p\ value = 0.24$

Outbreak of nodding off at CIDP lectures: Risk ratio & Odds ratio

	Exposure Tomás	Exposure Other
Cases of nodding off	21	13
No. started lecture	559	1151

$R_1 = \frac{21}{559} = 0.0376$ $\quad R_0 = \frac{13}{1151} = 0.0113$

$RR = \frac{R_1}{R_0} = 3.33$ $\quad OR = \frac{R_1/(1-R_1)}{R_0/(1-R_0)} = 3.42$

$p\ value = 0.0006$

Outbreak of nodding off at CIDP lectures: Rate ratio

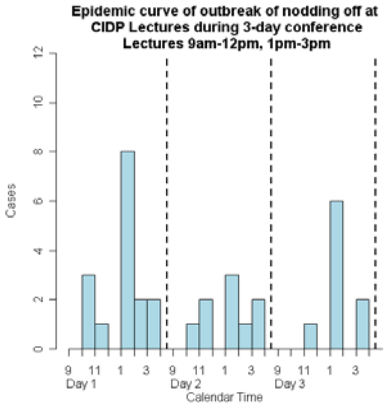
	Exposure Tomás	Exposure Other
Cases of nodding off	21	13
Person-hours at lecture	522.7	1047.3

$r_1 = \frac{25}{699.4}$ $\quad r_1 = \frac{8}{876.4}$ $\quad rr = \frac{r_1}{r_0} = 3.24$

$p\ value = 0.0008$

Notes:

Epidemic curve of nodding off



For 3 days, Dr. Aragón lectured the hour before *and* the hour after lunch.

What else could explain the nodding off?

What is this called?

Epidemiologic jargon!!!

- Ambiguous terminology
  - Attack rate
  - Case fatality rate, Case fatality ratio
  - Survival rate
  - Prevalence rate
- The many lives of "incidence"
  - Incidence
  - Incidence rate
  - Incidence density
  - Cumulative incidence

Epidemiologic study designs

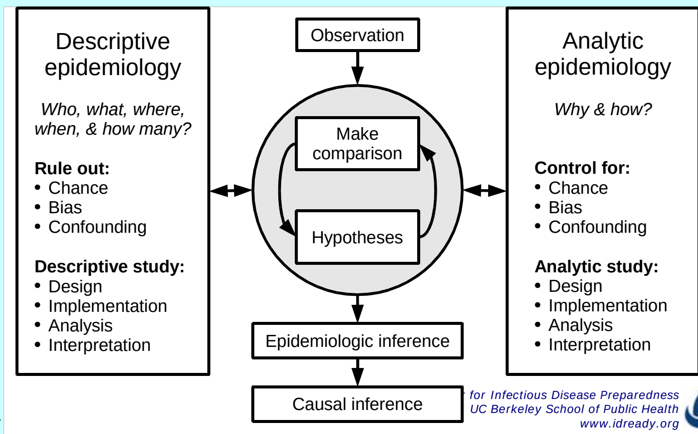
- Study designs
  - Experimental (Randomized control trial)
  - Observational (Cohort, Case-control)
- Steps
  - Design (study protocol)
  - Implementation (operations manual)
  - Analysis
  - Interpretation

Notes:

Threats to validity

- <u>Chance</u>: Random error
  - Confidence interval (precision)
  - P value (Pr{$X \geq x$}, under null hypothesis)
- <u>Bias</u>: Systematic error
  - Selection bias
  - Measurement bias (information bias)
- <u>Confounding</u>

Inferences in epidemiology



Public health action

- Clinical
- Behavioral
- Community
- Environmental

Adapted from Haddix AC, et al. Prevention Effectiveness: A Guide to Decision Analysis and Economic Evaluation. Oxford University Press 2003, 2nd Edition

Notes: