Epidemiology and Evidence-Based Medicine

Charles DiMaggio, PhD, MPH, PA-C
Columbia University
Departments of Anesthesiology and Epidemiology

Course Objectives

- Define epidemiology and its applications
- Identify the most commonly used study designs
- Understand the importance of measurement
- Apply coursework in evaluating the medical literature
- Define evidence-based medicine
- Create a searchable clinical questions
- Search medical databases to answer clinical queries.

Course Overview

- 1. Hx / ID / Biostatistics
- 2. Study Designs: RCTs, Cohorts
- 3. Study Designs: Ecologic, X-Sectional
- 4. Case-Control, Bias, Confounding
- 5. EBM: Intro and Concepts
- 6. EBM: Sources and Searches
- 7. Dx: Sens / Spec, PPV, NPV, LR
- 8. Tx: NNT
- 9. Oral Presentations

What is Epidemiology?

- What epidemiologists do
- Epidemiologists count
  - Define a population
  - Count cases of diseases in the population and compute rates
  - Compare those rates to another population
  - Make inferences regarding patterns and suggest interventions

Case 1

- August 2nd, 1976
- Robert Craven, CDC viral diseases branch, recieves report of 2 cases of severe respiratory illness (1 fatality) from Philadelphia ICN
- By August 3rd, 71 more cases, 18 deaths
- Legionella pneumophelia

Case 2

- October 30th, 1989
- New Mexico MD informs state DOH of 3 cases of severe myalgia and marked eosonophelia ➔ EMS
- Intense investigation reveals common vehicle: L-tryptophan supplements
- Remove from shelves

- Epidemiologists are like clinicians for a community
  - Gather information, make informed diagoses, suggest and implement interventions

What is Epidemiology?

- The study of the distribution and determinants of disease in human populations.
- Study
  - Methods are intended to be scientific (basic science of public health)
  - “Epidemiology is reasoned argument.”
- Distribution
  - Descriptive Epi – person, place, time
  - Look for patterns among different groups
- Determinants
  - Epi Triad – Agent, Host, Environment
  - Causality – Criteria (AB Hill), Induction vs. Deduction (Popper)
- Populations
  - Probability (chance) and Statistics
  - Study Designs

Study

- Basic Science of Public Health
- Quantitative, based on principles of statistics and research methodologies
- Methods are intended to be scientific (Reasoned Argument)

Distribution

- Frequencies and patterns of health events in groups
- Descriptive epidemiology – person, place and time
- E.g. age, sex, pre-existing conditions (COPD, DM, smoking, SES) how dx made (lab, culture, clinically) location (restaurant, gathering)

Determinants

- Search for Causes and Risk Factors
- Analytic Epidemiology
- Causality, inductive vs. deductive reasoning, Popperian refutation
- What is a cause and how do we know it?
- What is a disease?
  - ID’s, chronic diseases, injuries, disasters

Populations

- Distinguishing characteristic of epidemiology
- Need for specialized study designs
- Medicine is a social science, and politics nothing but medicine on a grand scale. (Virchow, 1848)

A Brief 3000-Year Timeline of Epi

- David and King Nebuchadnezzar
- James Lind and Scurvy
  - HMS Salisbury, 1740-1744
- John Snow and Cholera
  - 19th Century London (~1856)
  - the importance of rates and comparisons
    - impressions are not good enough
    - not a matter of "luck"
    - begins with observation (like any good detective work)

Snow on Cholera

WATER COMPANY	NUMBER OF HOUSES	DEATHS FROM CHOLERA	DEATHS PER 10,000 HOUSES
Southwark and Vauxhall	40,046	1263	315
Lambeth	26,107	98	37
Rest of London	256,423	1422	59

Lessons from Snow

- Need numbers; impressions not good enough
- Can intervene well before the actual causative variable is fully characterized
  - Smoking, HIV
- Diligence and perseverance

Austin Bradford Hill

- “Nature makes the experiments, and we watch and understand them if we can”
- “The highest returns can be reaped by imagination in combination with a logical and critical mind, a spice of ingenuity coupled with an eye for the simple and humdrum, and a width of vision in pursuit of facts that is allied with an attention to detail that is almost nauseating”

Evidence-Based Medicine

- What’s old is new again.
- Classic epidmiologic principles applied to clinical care (Clinical Epidemiology)

Concepts in Infectious Disease Epidemiology

- Kuru
  - Neurodegenerative Disease (Scrapie, BSE)
  - Fore Tribe Papua New Guinea
    - Women and children
    - Temporally related to deaths
    - Cultural practices?
  - Prions?

Commonly Used Terms

- Epidemic
- Outbreak
- Cluster
- Endemic
- Pandemic

Epidemic vs. Outbreak

- Epidemic - the occurrence of more cases of disease than would normally be expected in a specific place or group of people over a given period of time
- Outbreak – basically the same thing, but may have less serious connotations in public’s mind

Cluster

- Group of cases in a particular place and time; may or may not be more than expected
- Aim of investigation is to determine if there is an increases rate
- Sometimes used incorrectly in place of epidemic or outbreak

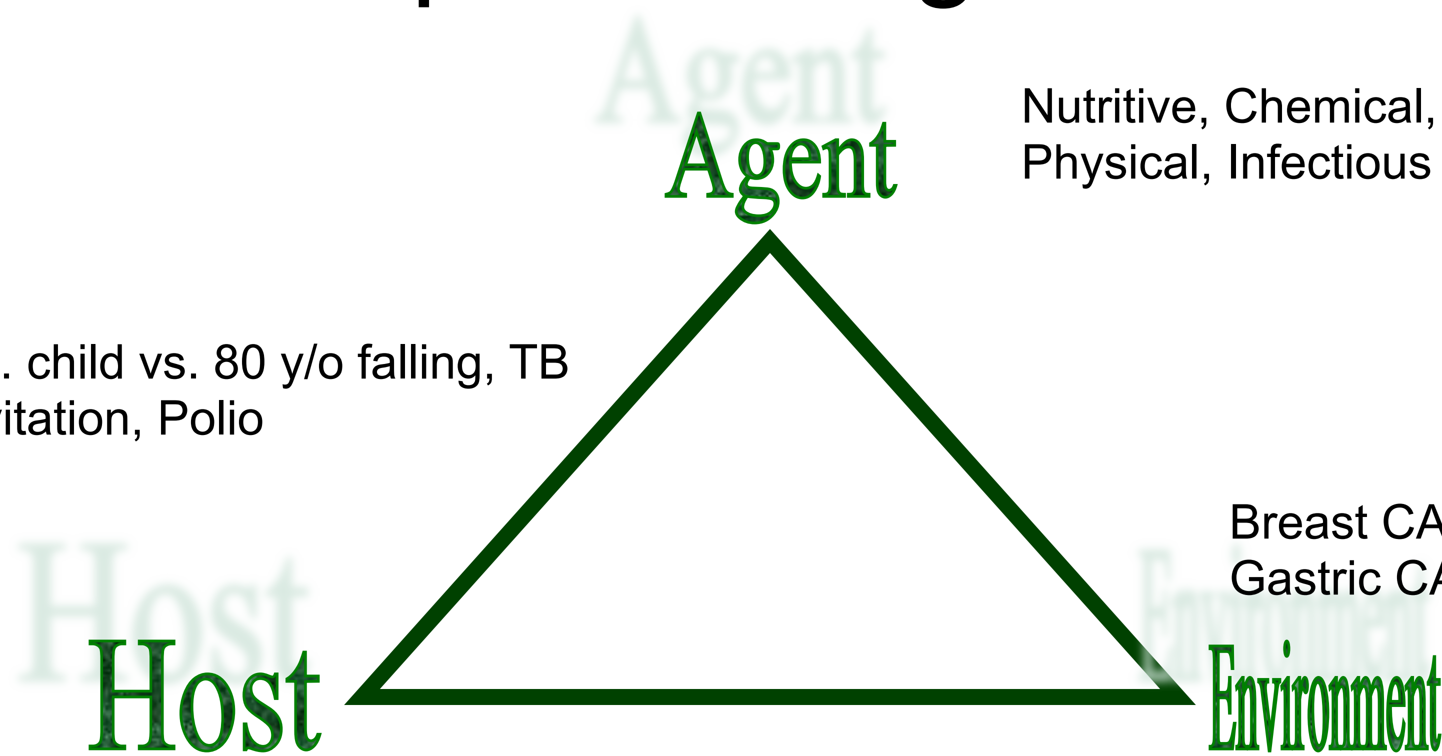
- Endemic – high background rate of disease
- Pandemic – widespread, often global disease

The Epidemiologic Triad

Nutritive, Chemical,
Physical, Infectious

(e.g. child vs. 80 y/o falling, TB
Cavitation, Polio

Breast CA on LI?
Gastric CA Asia



Inborn, Acquired,
Behavioral

Much Overlap e.g. Kuru

Descriptive Epidemiology gives clues to interventions.

Bubonic Plague (14th Century)

- Mortality in Florence up to 70% (Host)
  - Why not the other 30%?
  - Why humans not dogs?
- Mostly in the Summer (Agent)
  - Why not in winter?
- Mostly in the cities (Environment)
  - Why not in countryside?

Agent

- Entity necessary to cause disease in susceptible person
  - Infectious
    - viruses (HIV), bacteria (TB), protozoa (malaria), rickettssia (Rocky Mountain Spotted Fever)
  - Chemical - Tylenol, CO, heroin
  - Physical – cars, ionizing radiation
  - Nutrition
    - ↑ cholesterol ➔ CAD, ↓ protein ➔ kwashiorkor

Terms associated with infectious diseases

- Infectivity - the capacity to cause infection in a susceptible host.
- Pathogenicity - the capacity to cause disease in a host
- Virulence - the severity of disease that the agent causes in the host.

Host

- Individual susceptibility is key to disease process
  - E.g TB cavitation due to immune process; consequences of falls vary by age
- Factors may be
  - Inborn – thallasemia and Mediterranean descent; gender and MI
  - Aquired – immunlogical experience and age
  - Behavioral – cigarettes, exercise
- Host status
  - susceptible, immune or infected

Environment

- Often among the most challenging clues
  - can be misleading
- Wide variety of potential factors
  - physical, climatologic, biologic, social, and economic
  - E.g MVC – speed, weather, road conditions, local law enforcement, community views of drinking
  - Stomach CA in Japan due to pickled and smoked foods?

Breast Cancer on Long Island

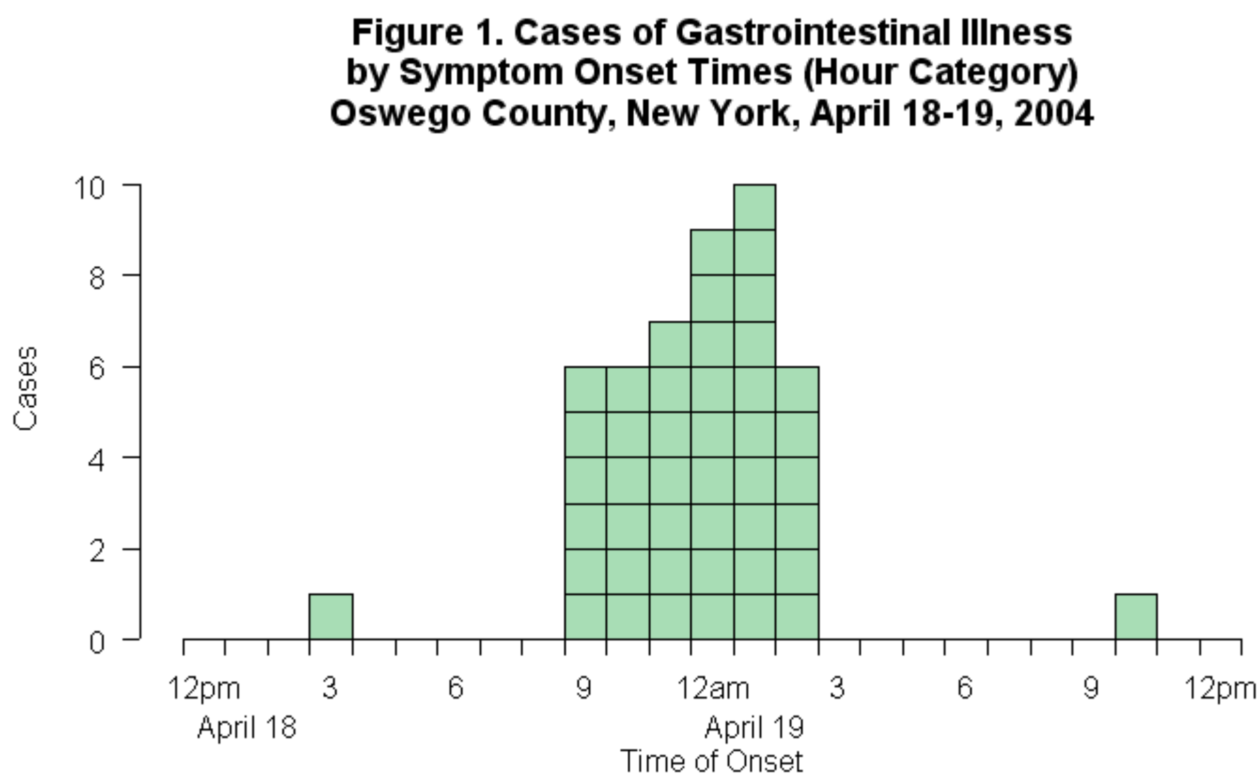
- "I have just come from the breast cancer capital of the world, and that is Long Island." 1991 News Conference, prior to $30 Million 1993 Congressional authorization to study breast cancer on LI

- Statistical anomoly? Pesticides in water? Fatty foods?
- 1.1% Percentage above the national average of breast cancer rates in Nassau County, LI. (115.6/100,000 vs. 114.3 cases / 100,000, 1994-1998)

Mode of transmission

- Direct
  - contact with soil, plants, other people
- Indirect
  - Airborne – agent carried from source to host on air particle
  - Vector borne – tranmitted by live vehicle e.g. tse tse fly (African Trypanosomiasis) reduvid bug (American Trypanosomiasis) anopheles mosquito (malaria) ticks (Lyme disease)
  - Vehicle borne – inanimate objects, e.g bedding, surgical instruments

Time → Epidemic Curve



Selected GI Differential Dx's

T	Predominant Sx	Organism or toxin
	Upper gastrointestinal tract symptoms (nausea, vomiting) predominate	
2-4 h	N/V, retching, diarrhea, abdominal pain, prostration.	*Staphylococcus aureus* and its enterotoxins
	Lower GI tract symptoms (abdominal cramps, diarrhea) predominatet	
18-36 h	Abdominal cramps, diarrhea, vomiting, fever, chills, malaise, nausea, headache, possible. Sometimes bloody or mucoid diarrhea.	*Salmonella* species (including *S. arizonae*), *Shigella*, enteropathogenic *Escherichia coli*, other *Enterobacteriacae*,
	Neurological symptoms (visual disturbances, vertigo, tingling, paralysis)	
12-36 h	Vertigo, double or blurred vision, loss of reflex to light, difficulty in swallowing. speaking, and breathing	*Clostridium botulinum* and its neurotoxins
	Allergic symptoms (facial flushing, itching) occur	
1 h	Numbness around mouth, tingling sensation, flushing, dizziness, headache, nausea.	Monosodium glutamate
	Generalized infection symptoms (fever, chills, malaise, prostration, aches, swollen lymph nodes) occur	
9 day	Gastroenteritis, fever, edema about eyes.	*Trichinella spiralis*

January 1992

Measuring Disease (Frequency)

"…when you can measure what you are speaking about, and express it in numbers, you know something about it…"

Lord Kelvin


BRITISH SLAUGHTER IN INDIA

Why Numbers Count

NYT, 14 August, 2005

Police Deaths In Car Crashes Are Increasing

WASHINGTON, Aug. 13 (AP) — Officer David Scott was pursuing a robbery suspect when he swerved to avoid a car. His police cruiser crossed into oncoming traffic and was hit by a pickup truck.

Officer Scott and Officer Yamil Baez-Santiago, his rookie partner in the Police Department in Clarksville, Tenn., were killed.

The incident is part of a worrisome trend, as more police officers are being killed in traffic. In 1999 and 2003, car crashes topped guns as the No. 1 killer of on-duty officers.

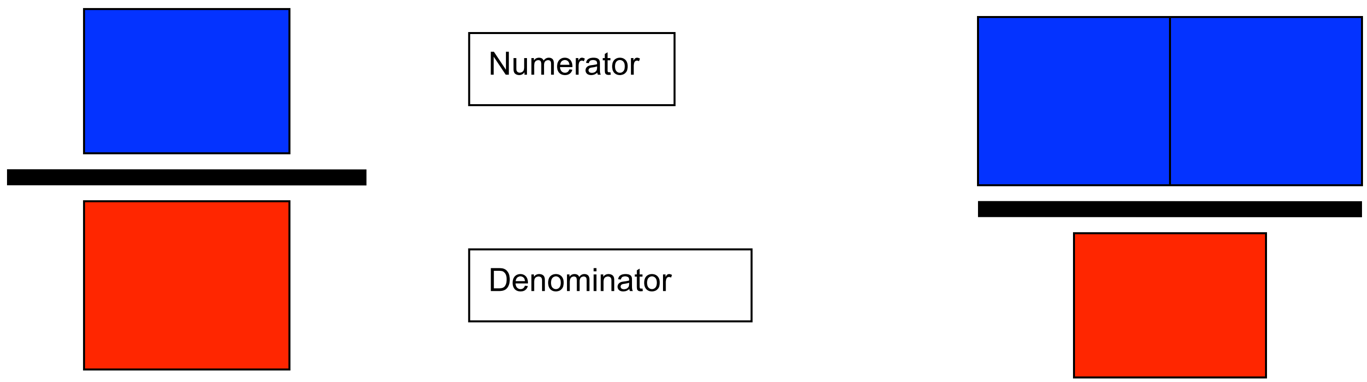
The National Law Enforcement Officers Memorial Fund said that in the decade ended last year, 477 officers died in auto accidents, up from 369 in the previous decade and 342 in the decade before that.

Some of the deaths occur because there are more officers on the roads — 52 of every 100 county and city officers in 2003, compared with 49 in 1997, according to the Bureau of Justice Statistics. There also are more deaths from high-speed chases, the Transportation Department reports.

A Memorial Fund spokesman, Bruce Mendelsohn, points to the increasing number of civilian drivers using cellphones and other devices that can distract them.

Ratio

numerator and denominator are separate

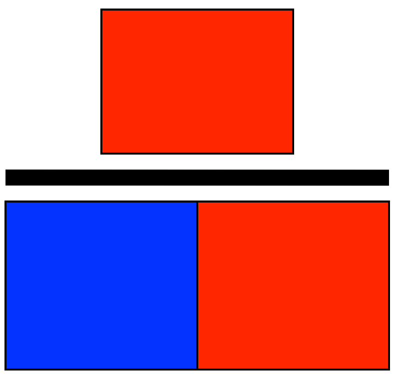


1:1 **2:1**

Odds: # chances for vs. # chances against

Proportion

numerator is included in the denominator



½ = 0.50 = 50%

probability or risk: fraction (0→1) of # chances for over total # chances

Odds ➔ Probability

Odds = Probability / 1 – Probability

Probability = Odds / 1+ Odds

e.g. 4 marbles: 1 blue, 3 red

Probability = ¼ = 0.25

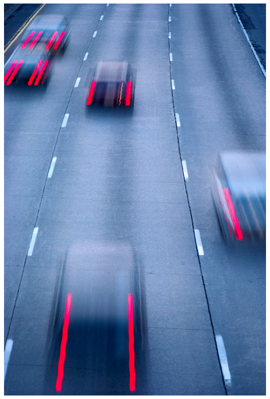
Odds = 1:3 = 0.333

Probability = Odds / 1+ Odds = 0.333 / 1.333 = 0.25

Odds = Probability / 1 – Probability = 0.25 / 0.75 = 0.33

Rate

- a measure of change per unit of another quantity (time)
- In epidemiology, often measure people (ill or dead) per year
  - N.B. many numbers called rates are actually proportions e.g. infant mortality



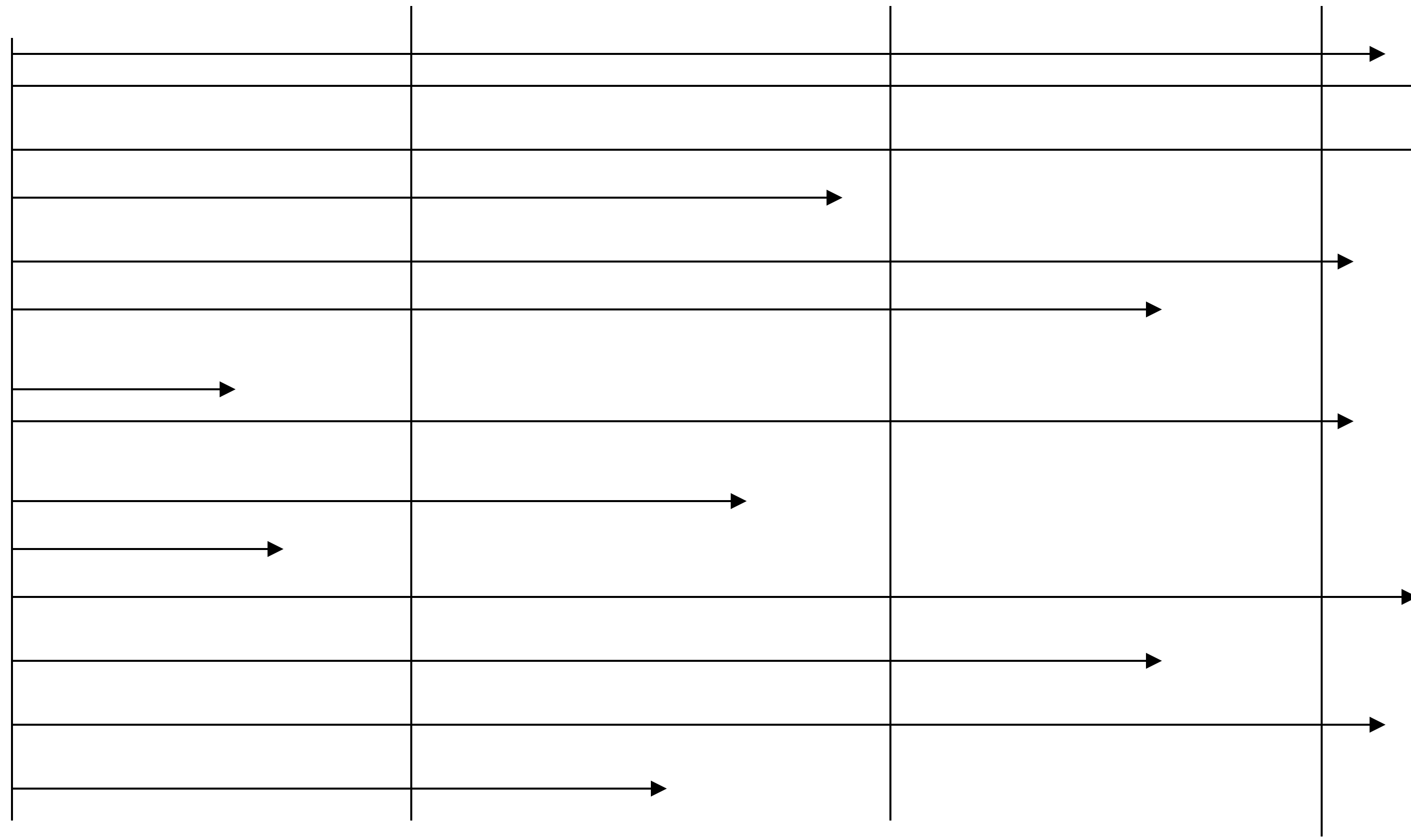
Miles (first unit) per hour (second unit)

Relative Rates

- Compare one rate to another
- Often compare the experience of one group of people with a particular exposure with that of a group that lacks that exposure
- E.g. # Lung CA deaths per year in smokers vs. non-smokers

How do epidemiologists measure time?

- Person Years



14 Persons

$7 + (\frac{1}{2})\,7 =$ 10.5 Person-Years

Difference between Rates and Risks

- 100 persons alive at beginning of year
- 40 die over the course of the year
- Risk (proportion) = 40/100 = 0.40
- How many person years?
  - 60 + (½) 40 = 80 person years
- Rate = 40/80 = .50

Pay attention to units of measurement

Incidence and Prevalence

- 1. Prevalence - measure all current cases of disease
- 2. Incidence -measures the rapidity with which a disease occurs or the frequency of addition of ***new*** cases

Incidence – Prevalence Bias

	TIME ON AFDC		
	1-2 YRS	3-7 YRS	>7 YRS
PERCENT WHO HAVE EVER RECEIVED AFDC	30%	40%	30%
PERCENT RECEIVING AFDC AT PARTICULAR TIME	7%	28%	65%

Incidence Rate and Cumulative Incidence

- Incidence Rate - “measure of the instantaneous force of disease”

- Cumulative Incidence – proportion who become ill (or die) during a specified time interval; it is a measure of average risk, dimensionless, from 0 to 1 (like a true probability)

Measures of Effect:

- ***Absolute*** = differences
  - I(E) – I(e)
- ***Relative*** = ratios
  - I(E) – I(e) / I(e) = I(E) / I(e) – 1
  - I(E) / I(e)
- Attributable Proportion – proportion of the diseased for whom exposure is a component cause
  - I(E) – I(e) / I(E)
  - = RR – 1 / RR

Absolute vs. Relative Effects

Rates per 100,000

	SMOKERS	NON SMOKERS	RELATIVE RISK	RATE DIFFERENCE
Lung Cancer	48.33	4.49	10.8	43.84
Coronary Artery Disease	294.67	169.54	1.7	125.13

Measure of Effect Driven by Aim of Research (Etiology or Public Health)

Standardized Rates: A way to compare two populations

Country A: (population = 7,496,000; deaths = 73,555)

- **Crude death rate** = 7,496,000 / 73,555 x 1000 = **9.8 per 1000**

Country B: (population = 1,075,000; deaths = 7,871)

- **Crude death rate** = 1,075,000 / 7,871 x 1000 = **7.3 per 1000**

Which is the healthier country?

Country A=Sweden Country B=Panama

Leveling the Field

- The standardized rate is a weighted average of the category specific rates
- Weights taken from “standard” population
  - 4 steps:
  - 1) calculate age-specific rates for each population
  - 2) multiply those rates by age distribution standard population
  - 3) add the total deaths
  - 4) divide by population of the standard

The Health of Nations: Sweden vs Panama

	YEARLY MORTALITY PER 1000	
AGE	SWEDEN	PANAMA
0-29	1.1	5.3
30-59	3.6	5.2
>60	45.7	41.6

Choose standard age distribution

AGE	WEIGHT
0-29	3, 145, 000
30-59	3, 057, 000
>60	1, 294,000

Multiply age-specific mortality rates* by the standard

AGE	SWEDEN	PANAMA
0-29	1.1 x 3145000 = 3459.5	5.3 x 3145000 = 16668.5
30-59	3.6 x 3057000 = 11005.2	5.2 x 3057000 = 15896.4
>60	45.7 x 1294000 = 59135.8	41.6 x 1294000 = 53838.4

*Rates per 1000 e.g. 1.1 = 0.0011

Compare Rates

- Add deaths divide by total (standard) population
- **Sweden = 73599.5 / 7, 496, 000 x 1000 = 9.8 per 1000**
- **Panama = 86403.3 / 7, 496, 000 x 1000 = 11.5 per 1000**

A word about standards

- Choice may be arbitrary or hypothetical
  - e.g. Sweden (note crude=standard)
  - US Standard
  - Caution comparing standardized rates (same standard?)
- You 'standardize' by the variable you use to categorize
  - E.g. 'age standardized' rates

Direct vs. Indirect Standardization

- Direct:
  - Rates Weighted by Standard Population
- Indirect
  - Populations Weighted by Standard Rates
  - Results in ‘Expected’ Rate of Occurrence
  - Standardized Mortality Ratio (SMR) = Observed Rate / Expected Rate X 100
  - Greater than 100 = More than Expected
  - Caution comparing SMRs that use different standards

Indirect Standardization

Age Group	Person-Years	US Cancer Mortality/ 100,000	Expected Cancer Deaths
15-24	1250	9.9	0.1
25-34	3423	17.7	0.6
35-44	3275	44.5	1.5
45-54	2028	150.8	3.1
55-64	1144	409.4	4.7

The Importance of Significance

- Much of biostatistics is concerned with how likely or unlikely it is that our results are due to chance alone
- Much of statistical chance is based on sample size.

P Values

- What is the probability that our results are due to chance?
- Presented as Decimal. The smaller the p value, the less likely the results were due to chance.
- E.g. $p=0.1$ means that given our null hypothesis (H_0) we can expect our result to occur 1 time out of 10 (10%) by chance alone
- $P=0.05$ the (arbitrary) cut off for statistical significance
- Again, the smaller the p value, the greater the significance

Confidence Intervals

- Gives same information as P values i.e. whether the results are likely due to chance
- Gives the information using a relevant metric
- Tells us something about how *important* the results are
- Form: Point Estimate ± Critical Value x Standard Error e.g. z (95% CI x, y)
- Caution: CI includes zero (absolute effects) or one (relative effects)

Significant. But important?

- Treatment prolongs life. ($p<0.0001$)
- Average increased life expectancy 3 days (95% CI 2.5, 4.5)
  - Cost?
  - Side Effects
  - Compared to What? (no treatment, alternative treatments)
- Average increased life expectancy 3 days (95% CI -2.5, 6.5)

More fun with Confidence Intervals

- RR = 2.5 (95% CI 1.5, 3.5)
  - Our result was 2.5, but it could have, with 95% statistical certainty (confidence) been anywhere between 1.5 and 3.5
- RR = 2.5 (95% CI 0.5, 4.5)
  - One is the loneliest number.

Chronic Disease Epidemiology: Eras and Paradigms

- Sanitary Statistics and Miasma
- Germ Theory and Infectious Disease
- Transition to Chronic Disease Epidemiology
- Study Designs that Arose from Chronic Disease Epidemiology

Sanitary Statistics and Miasma (18^{th}-19^{th} Centuries)

- Arose from social concern
  - Problems (and solutions) societal and environmental
- Foul emanations from ground or water
  - Theory incorrect but approach worked
- William Farr – statistics
  - "Death rate is a fact…"
- Edwin Chadwick poverty ➔ disease

Germ Theory and Infectious Disease (19th – 20th Centuries)

- Lab based, narrow – 'silver bullet'
- Downfall of population/environmental epi
  - 1546 – Hieronymous Fracastori
  - 1849 - John Snow
  - 1865 – Louis Pasteur
  - 1882 – Robert Koch

Henle-Koch Postulates

- Epitome of infectious disease era
  - 1) occur in every case
  - 2) occur in no other disease
  - 3) induce the disease when introduced into a “virgin” organism
- **Viruses?**
- **Chronic Diseases?**

Chronic Disease Epidemiology

- WWII watershed
- “epidemic transition”
- “diseases of civilization”
- “Black Box Epidemiology
- “web of causation”
- Richard Doll and Sir Austin Bradford Hill smoking and lung cancer

Study Designs that Arose from Chronic Disease Epidemiology

Complex
Valid

- 1. Experiments and Quasi-Experiments
- 2. Cohort Studies
- 3. Case-Control Studies
- 4. Cross Sectional Field Study
- Milestones:
  - Doll and Hill – Lung CA
  - 1951 – Cornfield – OR can approximate the RR
  - 1959 – Mantel-Haenszel –multivariate to C-C
  - 1973 – first text for 2x2 table by Joe Fleiss

The Four-Fold Table

	Disease	No Disease
Exposed	A (Exposed and Diseased)	B (Exposed, not Diseased)
Not Exposed	C (Not Exposed, Diseased)	D (Not Exposed, No Disease)