Epidemiology 101

May 13, 2026

Kate Smith, MD, MPH

Executive Director

ksmith@delamed.org

What is public health?

- The science and art of **preventing disease, prolonging life, and promoting health** through organized efforts and informed choices of society, organizations (public and private), communities, and individuals

- Deals with **preventive** (rather than curative) aspects of health
- Deals with **population-level** (rather than individual level) health issues

Health Care Provider

- **Head injury –** rule out concussion
- **Knee injury –** rule out fracture, clean wound and bandage
- **Check for ticks (?)**

Public Health Professional

- Is the road safe for bikers?
- Are there helmet laws there?
- Does she/her family have access to a PCP?
- Tall grass may have ticks – is Lyme's disease a problem there?

The Fence or the Ambulance

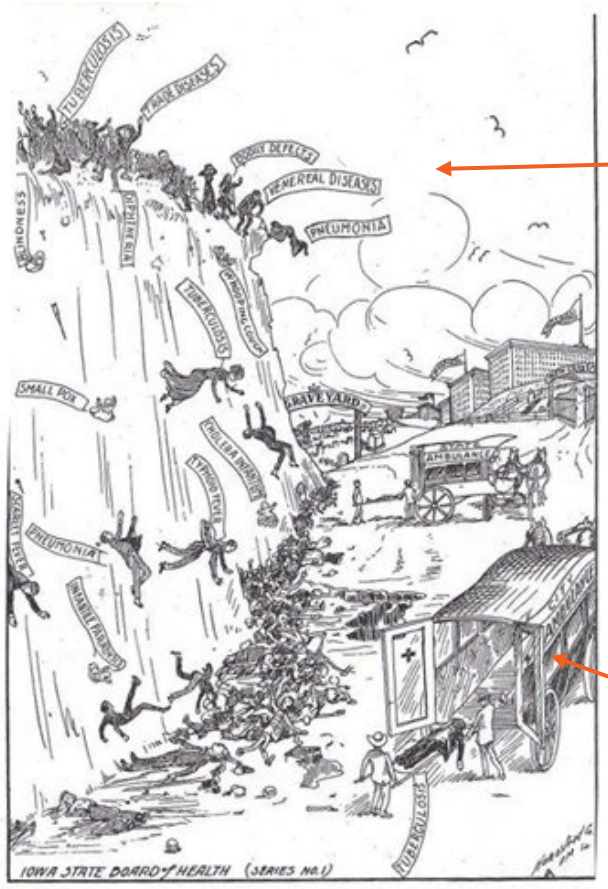


FIG. 2.—Illustrating the dangerous cliff without a fence.

Upstream

Downstream

Joseph Malines

Twas a dangerous cliff, as they freely confessed,
though to walk near its crest was so pleasant;
but over its terrible edge there had slipped
a duke and full many a peasant.
So the people said something would have to be done,
but their projects did not at all tally;
some said, 'Put a fence 'round the edge of the cliff,'
some, 'An ambulance down in the valley.'

But the cry for the ambulance carried the day,
for it spread through the neighboring city;
a fence may be useful or not, it is true,
but each heart became full of pity
for those who slipped over the dangerous cliff;
And the dwellers in highway and alley
gave pounds and gave pence, not to put up a fence,
but an ambulance down in the valley.

What is Epidemiology?

The study of the distribution and determinants of health-related states or events in specified populations

- The fundamental science of public health

Epidemiology:
The study of the distribution and determinants of health-related states or events in specified populations

Health:

"A state of complete physical, mental, and social well-being, not merely the absence of disease or infirmity"

World Health Organization

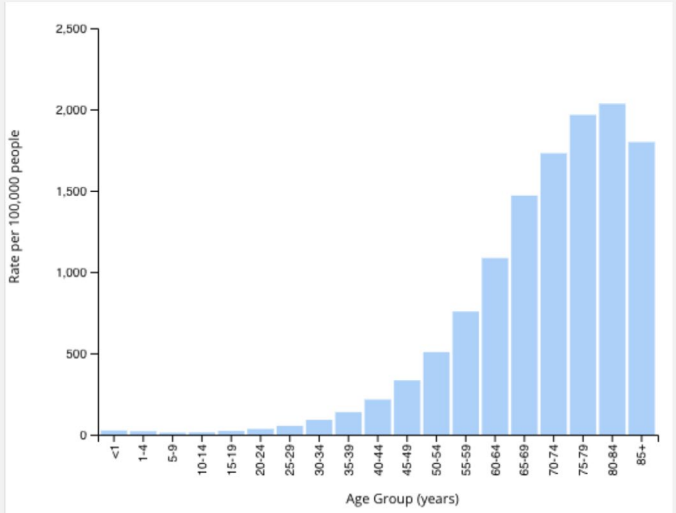
Epidemiology:
The study of the distribution and determinants of health-related states or events in specified populations

Distribution:

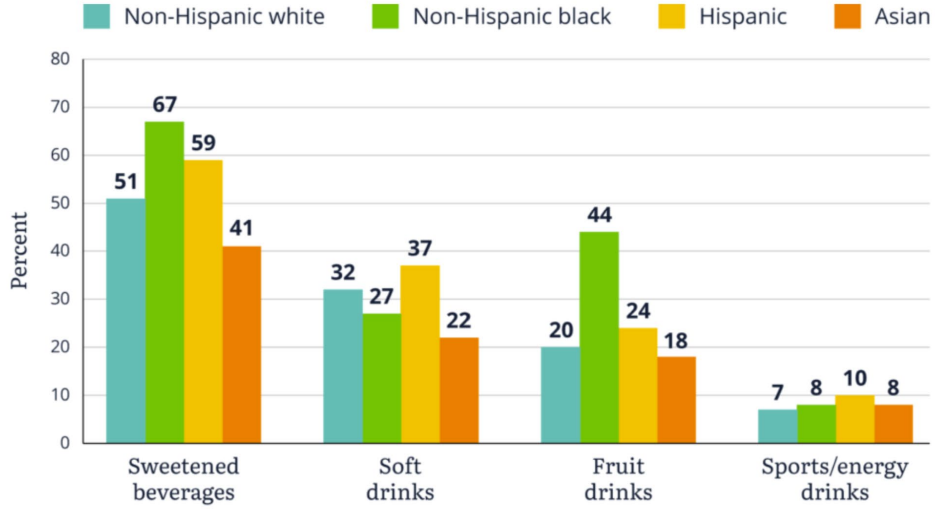
- A description involving some/all of the following characteristics:
  - Who (person) is getting the disease?
  - Where (place) is the disease occurring?
  - When (time) is the disease occurring?

People

Rate of New Cancers By Age Group (years), United States, 2020



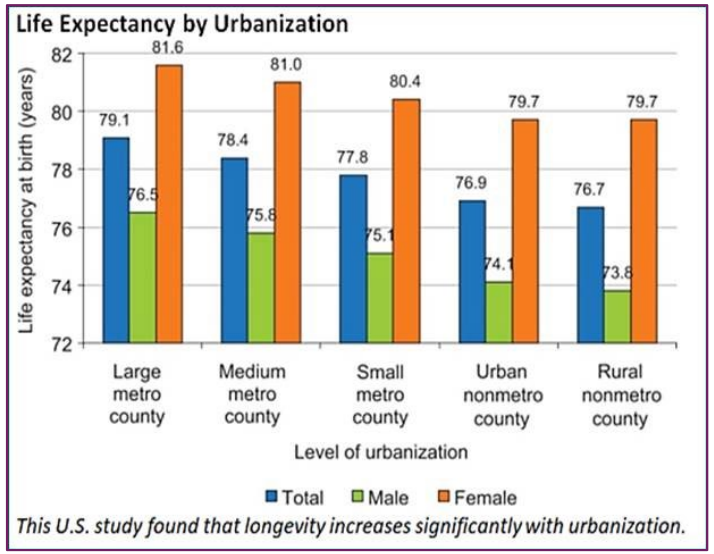
What types of drinks do children consume on any given day (by race and ethnicity)?

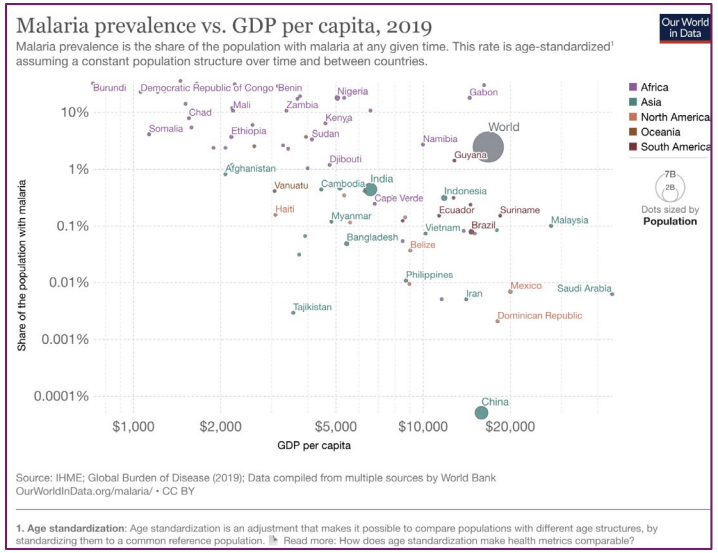


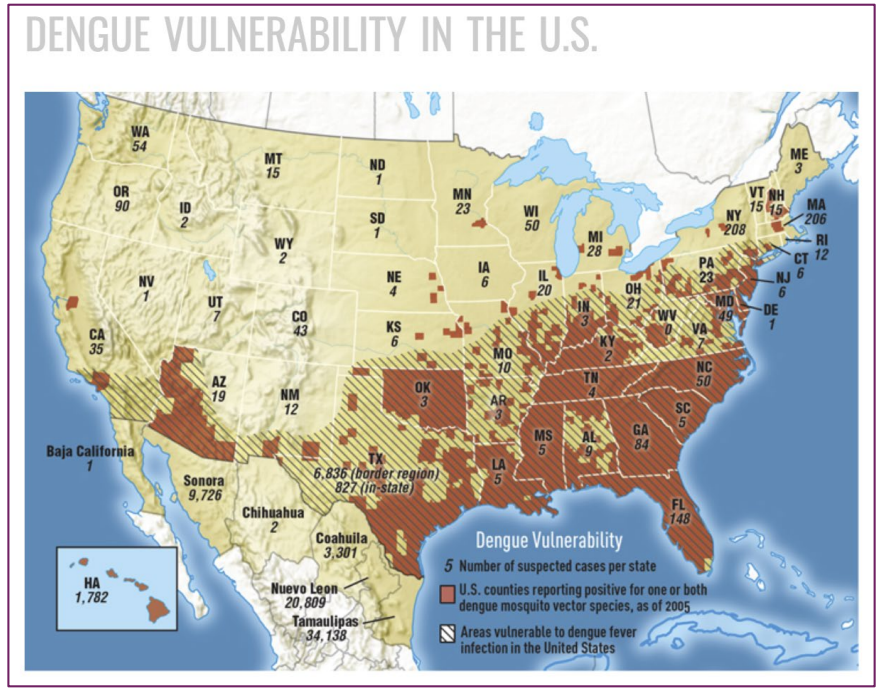
Estimated New Cases

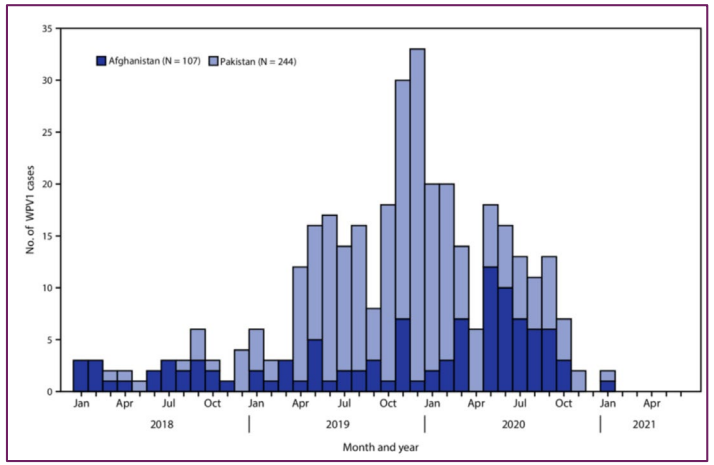
Males			Females		
Prostate	174,650	20%	Breast	268,600	30%
Lung & bronchus	116,440	13%	Lung & bronchus	111,710	13%
Colon & rectum	78,500	9%	Colon & rectum	67,100	8%
Urinary bladder	61,700	7%	Uterine corpus	61,880	7%
Melanoma of the skin	57,220	7%	Melanoma of the skin	39,260	4%
Kidney & renal pelvis	44,120	5%	Thyroid	37,810	4%
Non-Hodgkin lymphoma	41,090	5%	Non-Hodgkin lymphoma	33,110	4%
Oral cavity & pharynx	38,140	4%	Kidney & renal pelvis	29,700	3%
Leukemia	35,920	4%	Pancreas	26,830	3%
Pancreas	29,940	3%	Leukemia	25,860	3%
All Sites	**870,970**	**100%**	**All Sites**	**891,480**	**100%**

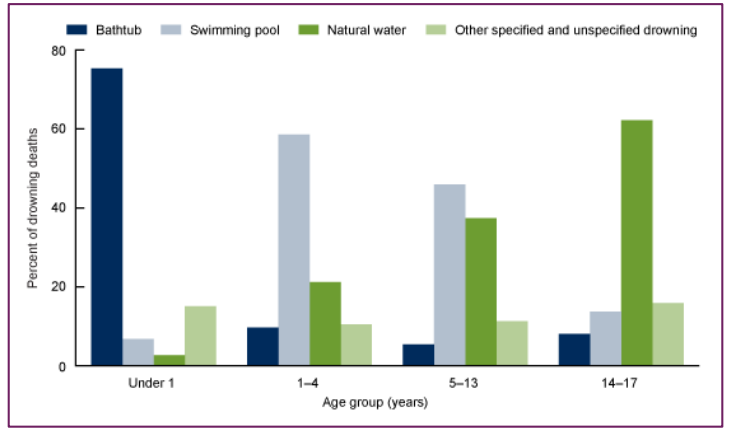
Places


Life Expectancy by Urbanization
Life expectancy at birth (years)
82
80
78
76
74
72
81.6
79.1
76.5
81.0
78.4
75.8
80.4
77.8
75.1
79.7
76.9
74.1
79.7
76.7
73.8
Large metro county
Medium metro county
Small metro county
Urban nonmetro county
Rural nonmetro county
Level of urbanization
Total
Male
Female
This U.S. study found that longevity increases significantly with urbanization.


Malaria prevalence vs. GDP per capita, 2019
Malaria prevalence is the share of the population with malaria at any given time. This rate is age-standardized[1] assuming a constant population structure over time and between countries.
Our World in Data
Share of the population with malaria
10%
1%
0.1%
0.01%
0.001%
0.0001%
$1,000
$2,000
$5,000
$10,000
$20,000
GDP per capita
Africa
Asia
North America
Oceania
South America
Dots sized by Population
Source: IHME; Global Burden of Disease (2019); Data compiled from multiple sources by World Bank
OurWorldInData.org/malaria/ • CC BY
1. Age standardization: Age standardization is an adjustment that makes it possible to compare populations with different age structures, by standardizing them to a common reference population. Read more: How does age standardization make health metrics comparable?


DENGUE VULNERABILITY IN THE U.S.
Dengue Vulnerability
5 Number of suspected cases per state
U.S. counties reporting positive for one or both dengue mosquito vector species, as of 2005
Areas vulnerable to dengue fever infection in the United States
Baja California 1
Sonora 9,726
Chihuahua 2
Coahuila 3,301
Nuevo Leon 20,809
Tamaulipas 34,138
TX 6,836 (border region) 827 (in-state)
HA 1,782

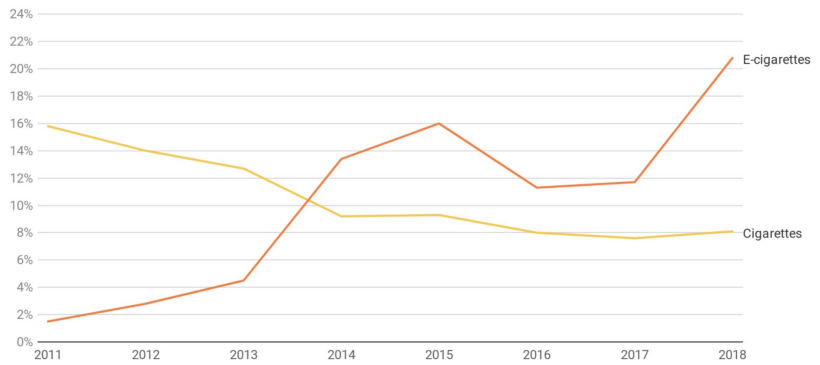

Afghanistan (N = 107)
Pakistan (N = 244)
No. of WPV1 cases
35
30
25
20
15
10
5
0
Jan
Apr
Jul
Oct
Jan
Apr
Jul
Oct
Jan
Apr
Jul
Oct
Jan
Apr
2018
2019
2020
2021
Month and year


Bathtub
Swimming pool
Natural water
Other specified and unspecified drowning
Percent of drowning deaths
80
60
40
20
0
Under 1
1–4
5–13
14–17
Age group (years)

Time

High School Students Smoking Less, Vaping More

The National Youth Tobacco Survey, published by the Centers for Disease Control and Prevention, quantified the surging popularity of e-cigarettes among high school students.



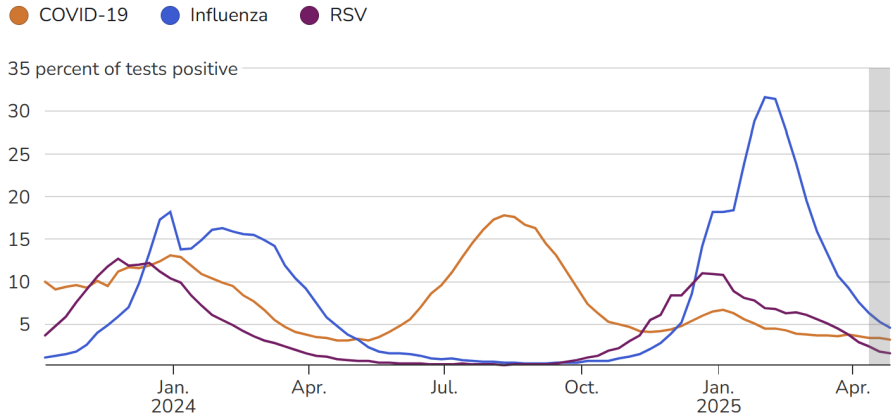
These numbers represent the percentage of high school students who reported having used cigarettes or e-cigarettes within the previous 30 days.

Credit: Harriet Blair Rowan/California Healthline

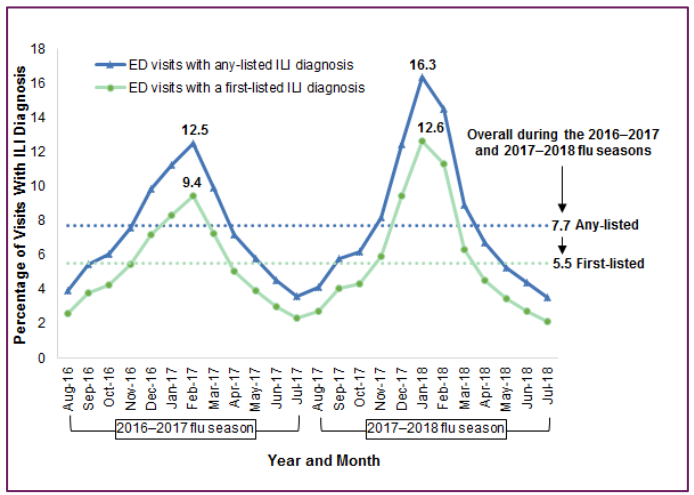
Source: Centers for Disease Control and Prevention • Created with Datawrapper

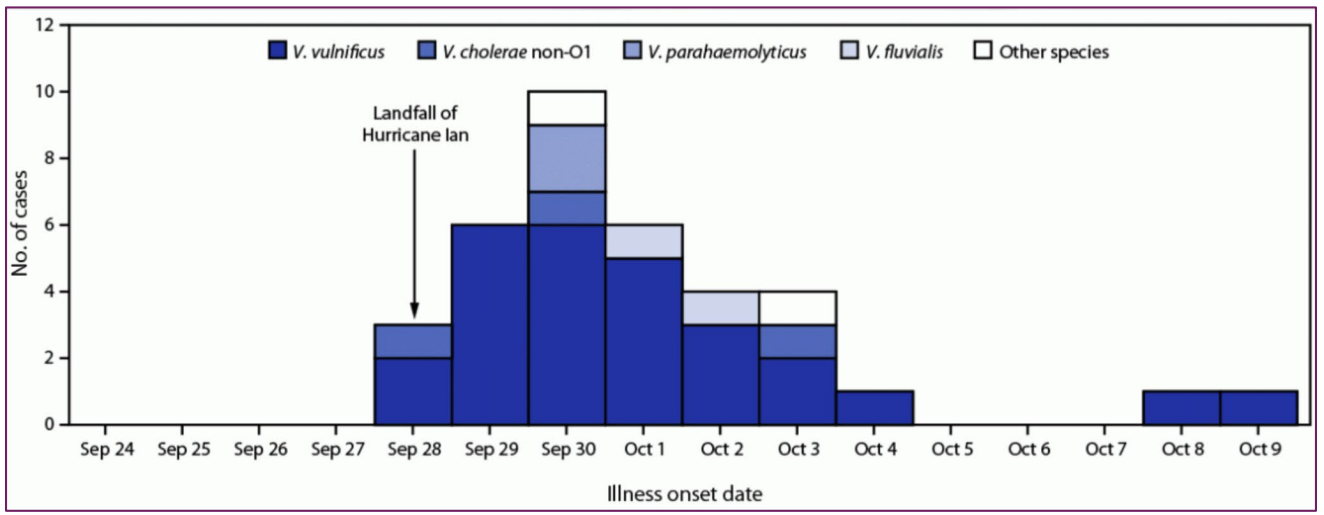
Percent of Tests Positive for Respiratory Viruses

Weekly percent of tests positive for the viruses that cause COVID-19, influenza, and RSV at the national level. Preliminary data are shaded in gray. Refer to data notes for more details.



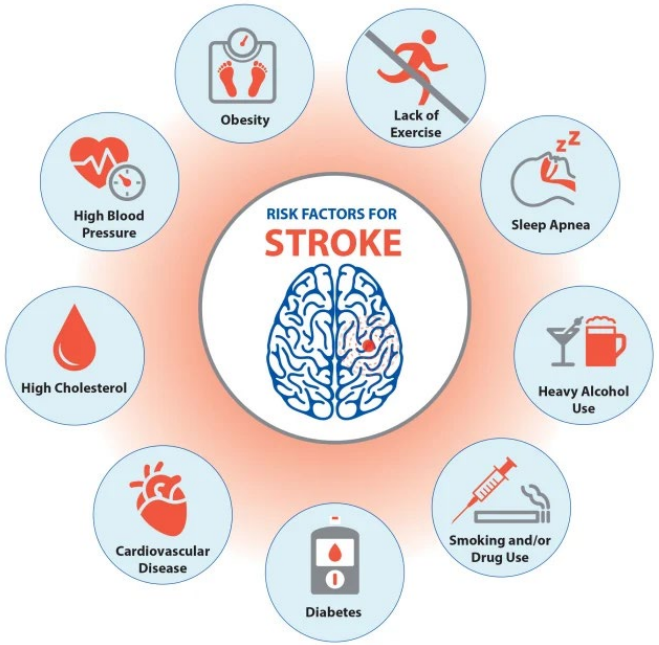
Data last updated on May 1, 2025 and presented through April 26, 2025. View this dataset on data.cdc.gov.





Epidemiology:
The study of the distribution and determinants of health-related states or events in specified populations

Determinants:



- Factors that “cause” health-related states OR prevent their occurrence

Social Determinants (Drivers) of Health

- Cancer screenings
- Vaccinations
- Family planning
- Treatments for SUD, STIs
- Prescriptions
- Access to care

- Neighborhood safety
- Walkability
- Air & water quality
- Green areas
- Transportation

- How many HS graduates?
- How many reading at grade level?
- Do early childhood education programs exist?
- Are students with disabilities in regular education programs?

- Mental health
- Adverse childhood events
- Literacy & health literacy
- Recidivism rate
- Foster care
- Food insecurity

- How many employed/in school?
- Work related injuries
- People living in poverty
- Disabilities

Conditions in which people are born, grow, work, live, play, pray, and age.

Epidemiology:
The study of the distribution and determinants of health-related states or events in specified populations

Population:

- A group of individuals having a quality or characteristic in common

Examples:

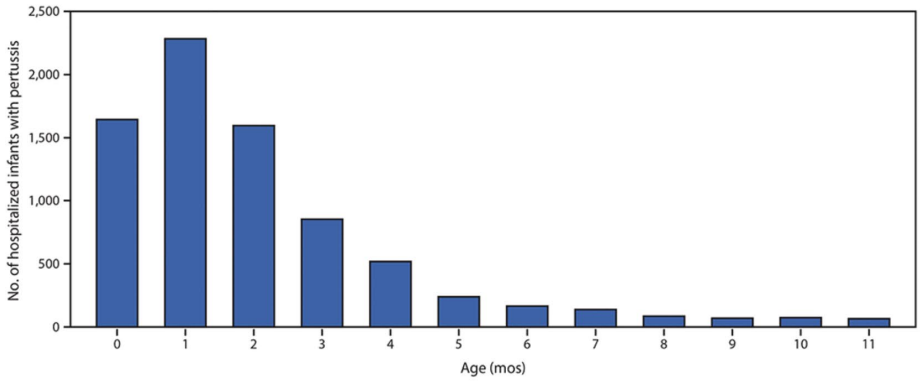
- Residents of the United States
- African-American mothers in Baltimore
- Superintendents of public school districts
- Women with stage 3 breast cancer
- Mini Medical School Students

Applications of Epidemiology

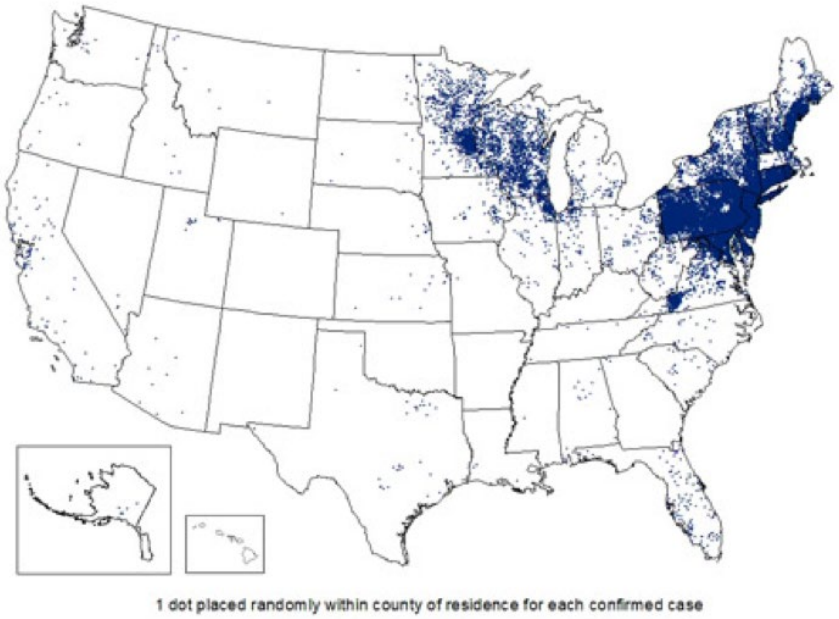
1. Determine the **extent** of disease in a community
2. Identify the **cause** of a disease and **factors** that increase or decrease a person's risk for getting that disease
3. **Evaluate** interventions designed to affect health-related states or events.

Extent of Disease

FIGURE 2. Number of infants with pertussis who were hospitalized, by age in months (N = 7,731) — National Notifiable Diseases Surveillance System, United States, 2010–2017



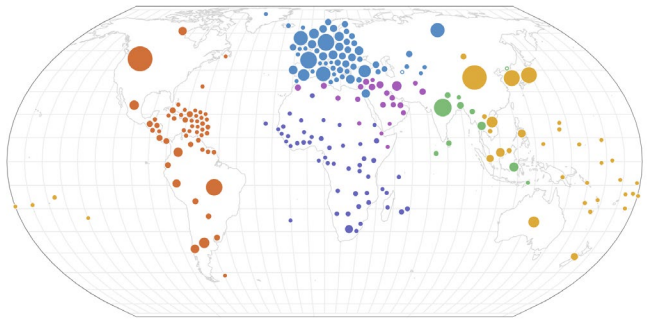
Reported Cases of Lyme Disease, USA 2017



1 dot placed randomly within county of residence for each confirmed case

Number of COVID-19 cases reported to WHO (cumulative total)

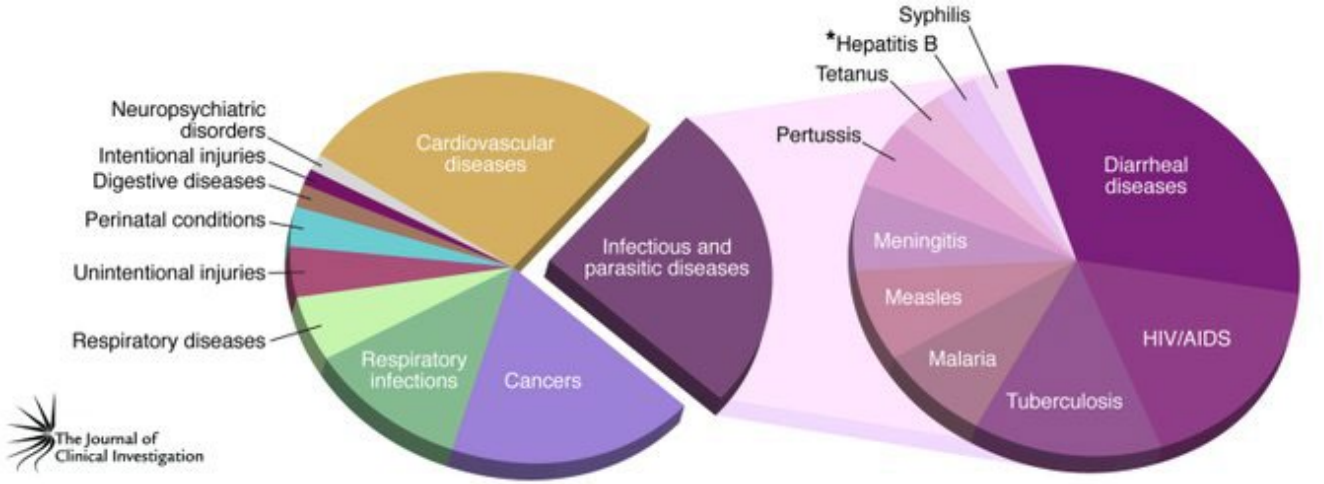
World

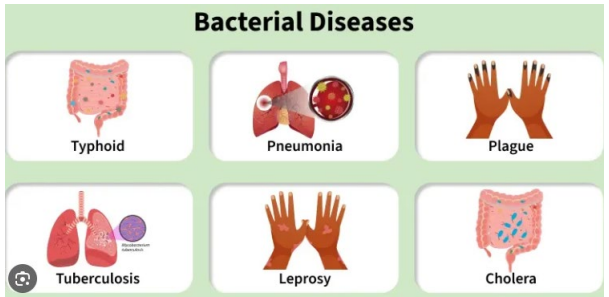


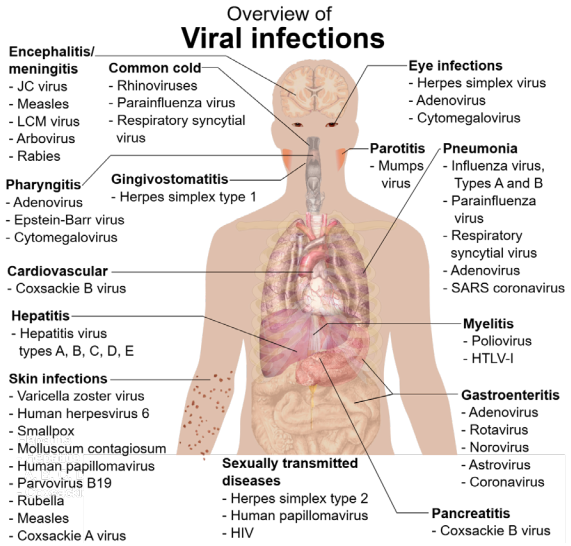
WHO Regions

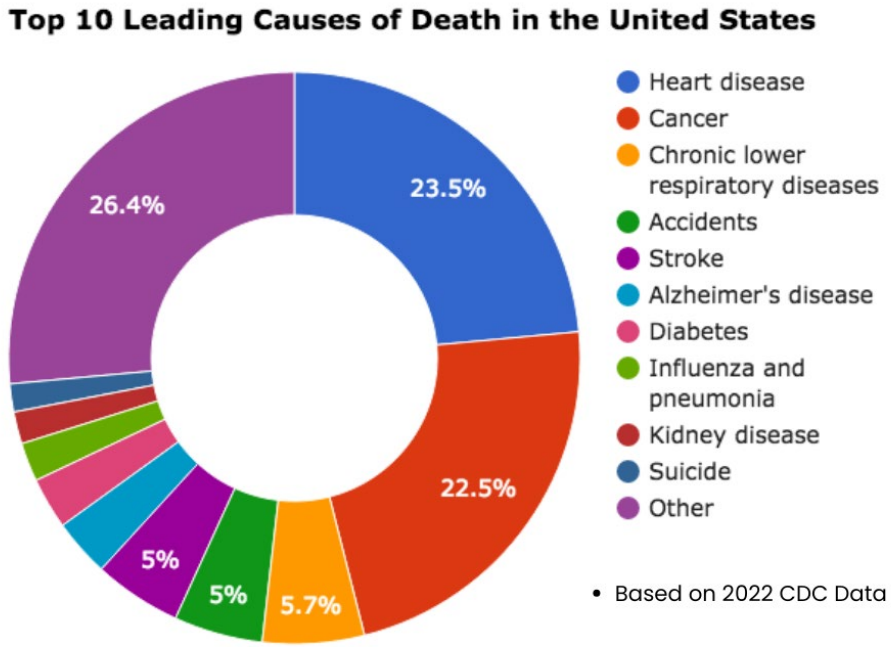
Africa Americas Eastern Mediterranean Europe South-East Asia Western Pacific

Cause of Disease





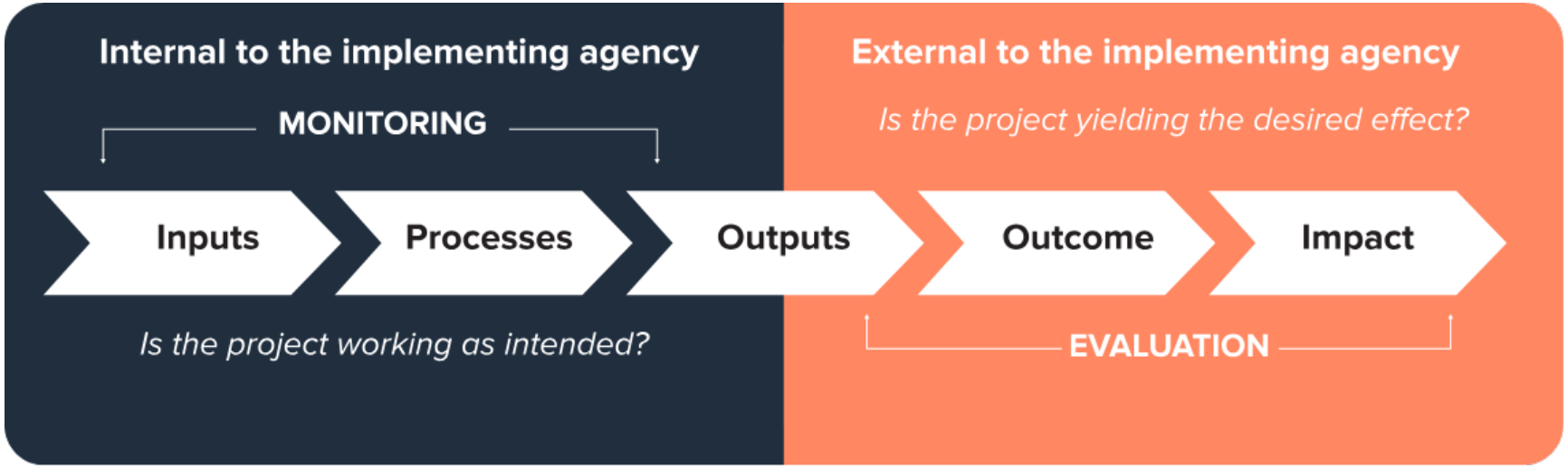




- Based on 2022 CDC Data

Evaluate Interventions

- Provides the basis for public policy, regulatory decisions

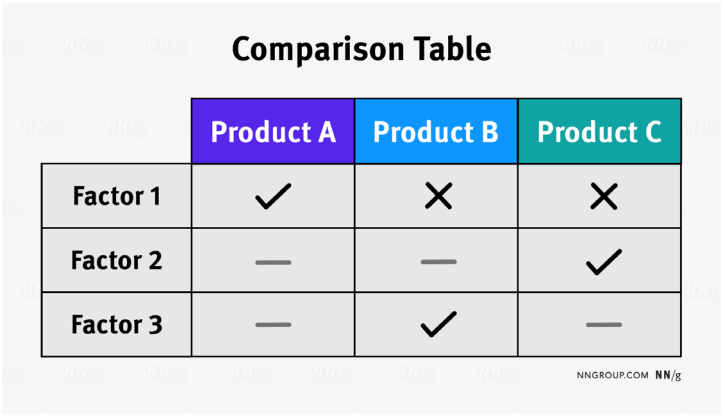


Describing Disease

Descriptions Should Include

Counts

Timing



Comparisons

Prevalence

The total number of cases in a population,

- at a specified time (point prevalence) or
- over a specified period of time (period prevalence)

Measures burden of disease

$$= \frac{\text{All the people who are sick**}}{\text{the whole population}}$$

***or healthy, or admitted to the hospital, or are happy … whatever we're studying*

Example

500 men with lung cancer were asked about their smoking habits. 461 men smoked at least one pack of cigarettes per day.

What is the prevalence of smoking in this group of men?

$$= \frac{\text{461 men who smoked}}{\text{500 men total in the study}}$$

= 0.922 or 92%

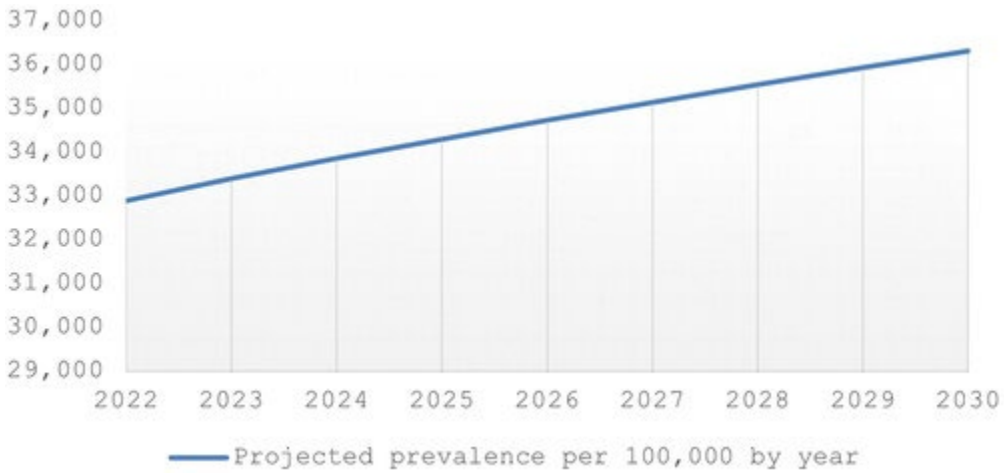
Numbers Less than 1

- Decimals can be confusing
- In public health, we want our numbers to be used and understood
- Prevalence is a proportion, and can be represented by:
  - A decimal (0.922)
  - Percent (92.2%)
  - OR multiplied to express the number of cases in a given population size

    922 smokers per 1,000 people

The Problem with Prevalence

- If we monitor prevalence over time, it is hard to measure the success of public health efforts
- Prevalence = total cases
  - the number will likely always increase (although it may potentially slow down)
- There is something more useful than prevalence…



Amyotrophic Lateral Sclerosis (ALS)
Estimated Prevalence, 2022-2030

Incidence

The number of new cases that occur during a specified time period, in a population at risk for becoming a case

Measures risk – something with a high incidence is a problem right now

$$= \frac{\text{\# new cases}}{\text{population at risk}}$$

Example

- What is the weekly incidence of flu among the UD football team for each of the first two weeks of September?

  100 people on the team
  - Wk 1: 7 players diagnosed
  - Wk 2: 31 players diagnosed

Week 1

Prevalence = 7 cases / 100 people

P = 0.07 (or 7% or 70 per 1,000)

Incidence = 7 cases / 100 people

I = 0.07

Week 2

Prevalence = (7 + 31) cases / 100

P = 0.38

Incidence = 31 cases / (100-7) people

I = 31 / 93 = 0.33 (or 33%)

Example

The town of Smithville (population 6,222) is concerned about high blood pressure (HBP). In January of 2018, they conducted a survey asking all residents, “Do you have HBP?” They found that 352 residents had HBP.

What is the prevalence of HBP in Smithville in January 2018?

= (352 cases) / (6,222 people)

= 0.06 = 6% = 60 per 1,000 people

Example

In June 2018, Smithville Public Health Department conducted a public health campaign to test blood pressure in the community. 125 residents were **newly diagnosed** with HBP.

What is the incidence of HBP in Smithville in June 2018?

= 125 new cases / (6,222 people – 352 previously diagnosed)

= 125 / 5,870

= 0.02 or 2% or 20 cases per 1,000 people

What is the prevalence of HBP in Smithville in June 2018?

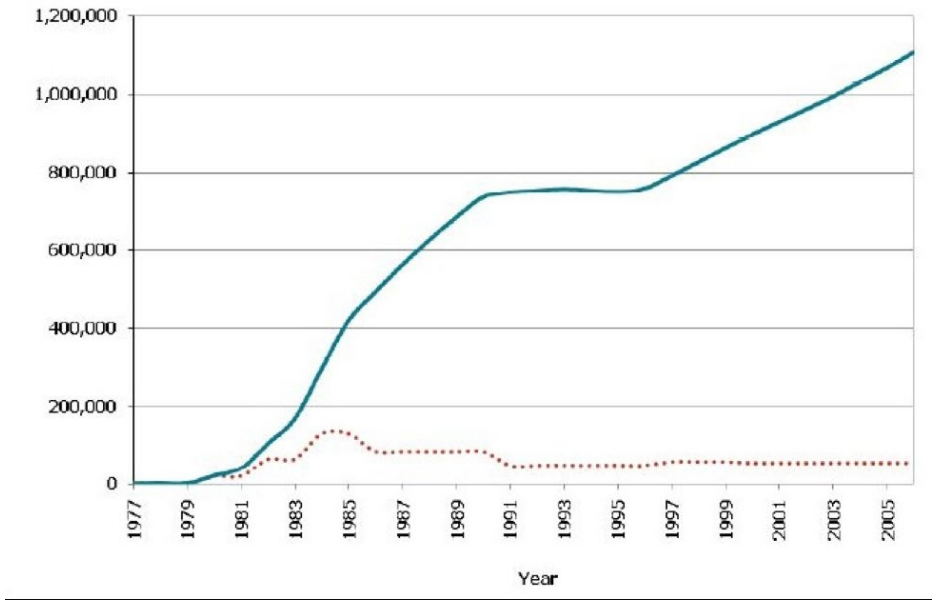
= (352 previous cases + 125 new cases) / 6,222 people

= 477 / 6,222

= 0.076 or 7.6% or 76 cases per 1,000 people

How are prevalence and incidence linked?

- Prevalence = Incidence x Duration



HIV Cases in the US
Blue line = prevalence, Red line = incidence

Prevalence = Incidence x Duration

	Prevalence	Incidence	Duration
A new treatment helps people with the flu recover faster	↓	Same	↓
A new strain of the flu kills people quickly	↓	Same	↓
A preventive hand washing program is introduced	↓	↓	Same
A new viral strain is more infectious than previous strains and spreads faster	↑	↑	Same
A new strain of the flu lasts for 2 weeks instead of 1	↑	Same	↑

Prevalence vs. Incidence

Prevalence

- Total cases
- Plan routine public health efforts
- Identify locations to focus on
- Track changes over time

Incidence

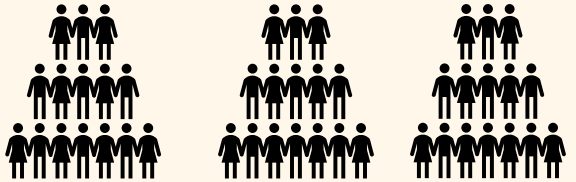
- New cases
- Plan urgent public health efforts
- Learn about potential causes
- Identify the end of an outbreak

Descriptive Epidemiology

Hypothesis Generating (What is happening?)

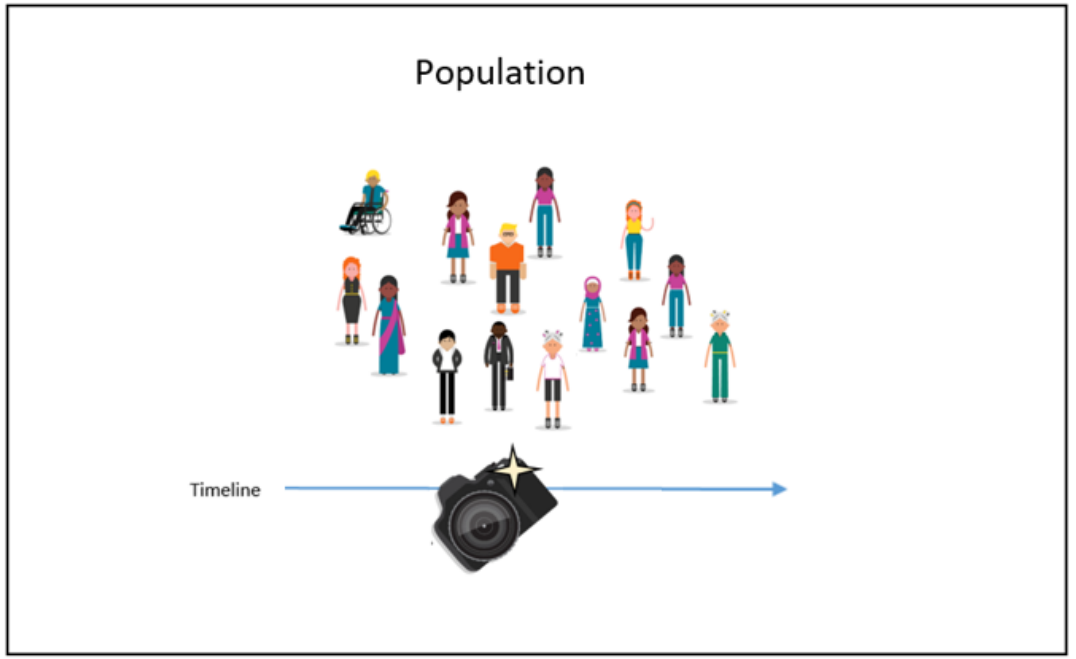
Descriptive Study Designs

Hypothesis Generating

Case Report	Case Study	Cross Sectional
		
“That was really weird! Has that ever happened to anyone else?”	“I’ve seen it a few times! Here’s what it looked like. I wonder how many other people have it?”	“Hmm, let’s design a study to see what’s going on.”
One person’s disease experience	Study of a small group of people’s disease experience	Collecting a “snapshot” from a large group of people at one time

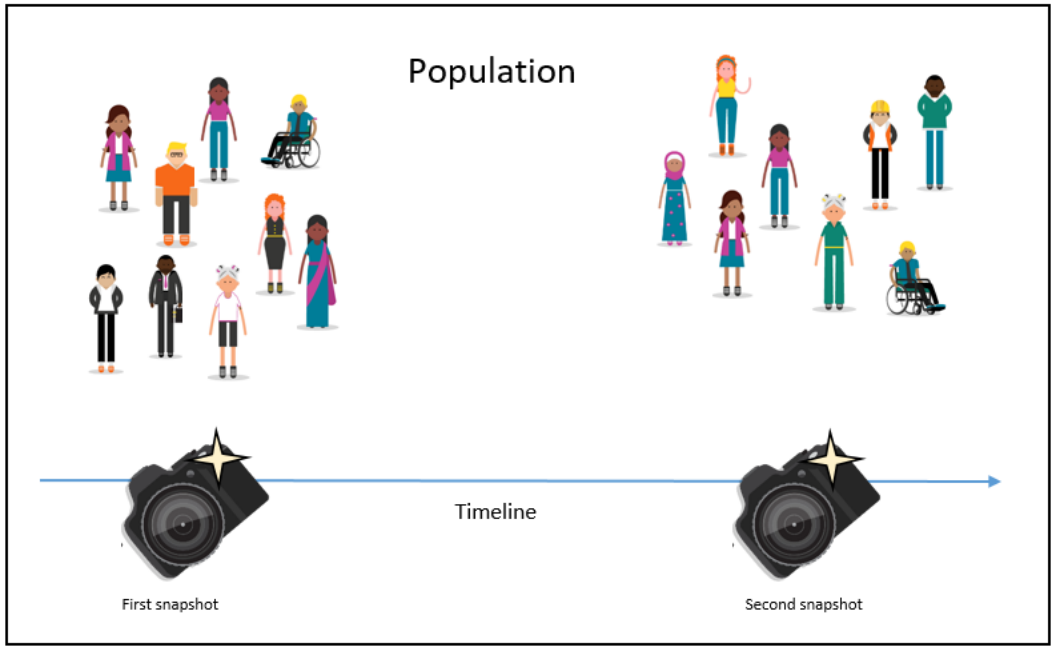
Cross Sectional Study

- Collects a lot of information from a large group of people at a single point in time
- Theory
  - Collects prevalence data
  - Looks at everything concurrently
  - Estimates magnitude and/or distribution of a health issue
  - Observe patterns



Surveillance Studies

- Cross Sectional done multiple times at different points in time with different people
- Theory
  - Collects prevalence data
  - Looks at everything concurrently
  - Estimates magnitude and/or distribution of a health issue
- Observe patterns
- Different group of people each time



Descriptive Studies are Best For

- Answering questions about **incidence/prevalence**

  How common is [disease / risk factor]?

- Establishing what is "**normal**" for a specific demographic at a specific time

  What risk factors are associated with [outcome]?

- Justifying further research on the topic

Analytic Epidemiology

Hypothesis Testing

Hypothesis Testing

- Is [exposure] associated with [outcome]?
- Positively associated: ⬇⬇ ⬆⬆
- Negatively associated: ⬇⬆ ⬆⬇

- Is eating raw eggs associated with muscle development?

Exposures vs. Outcomes

Exposures

(Things we can Assess)

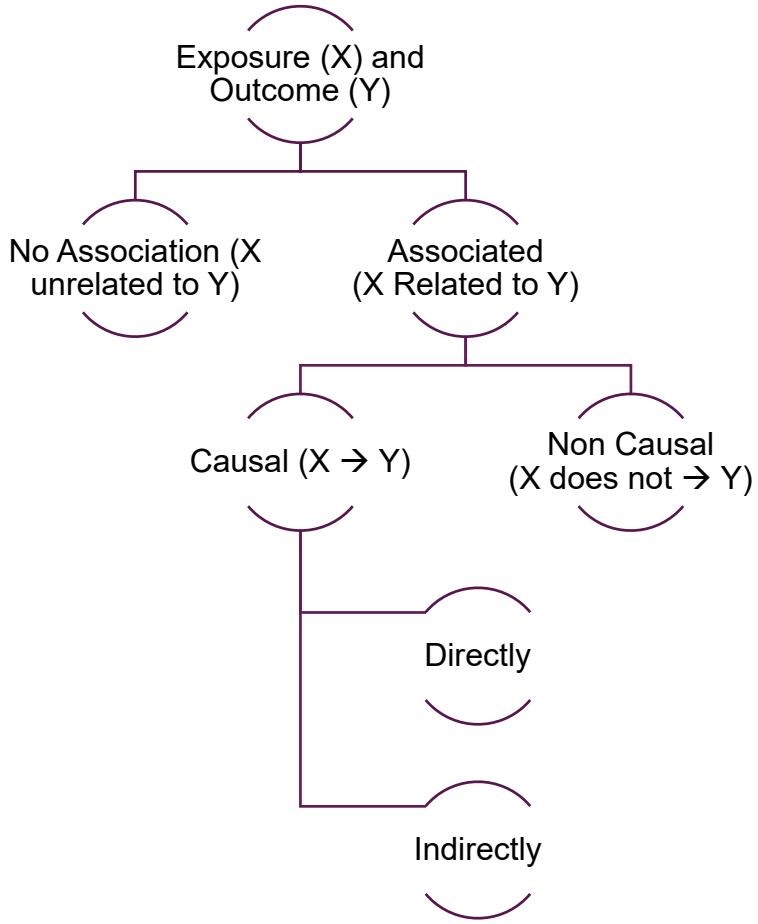
- Occupational groups
- Groups undergoing medical treatment
- Groups with unusual diet or lifestyle factors
- Professional groups
- Students/Alumni
- Geographically defined areas

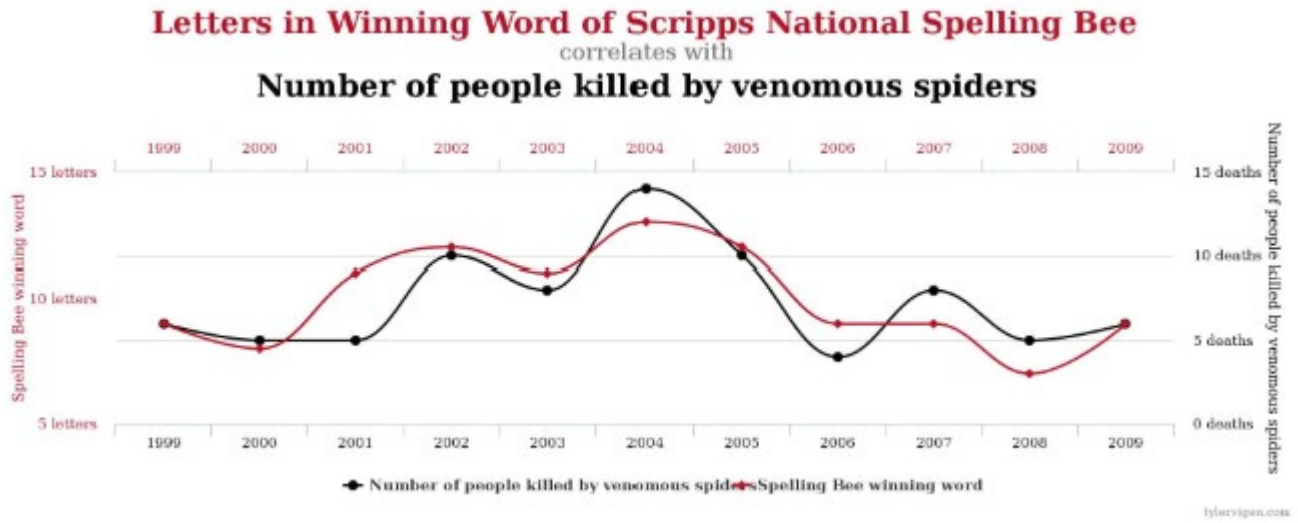
Outcomes

(e.g., diagnoses, death, hospital stays)

- Death certificates
- Billing records
- Questionnaires
- Medical Exams

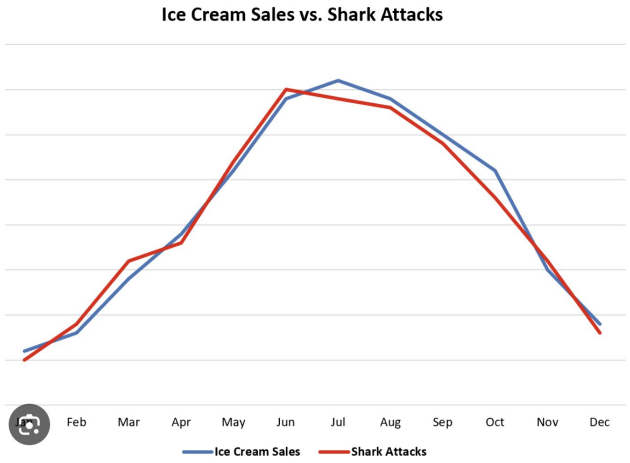
Relationships Between Exposure (X) and Outcome (Y)

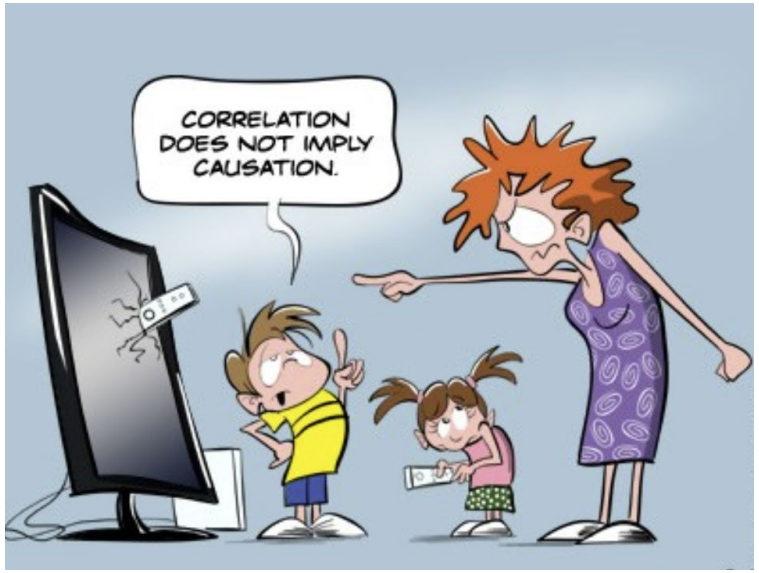


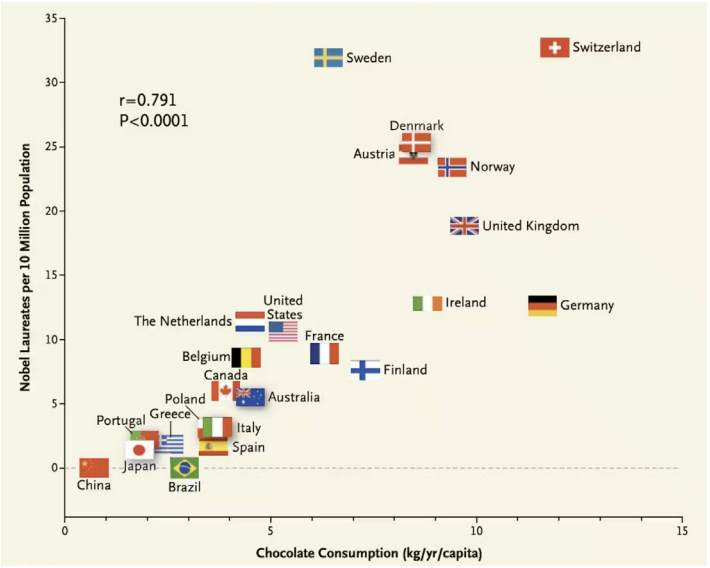


Correlation does NOT imply causation!

Correlation does NOT imply Causation

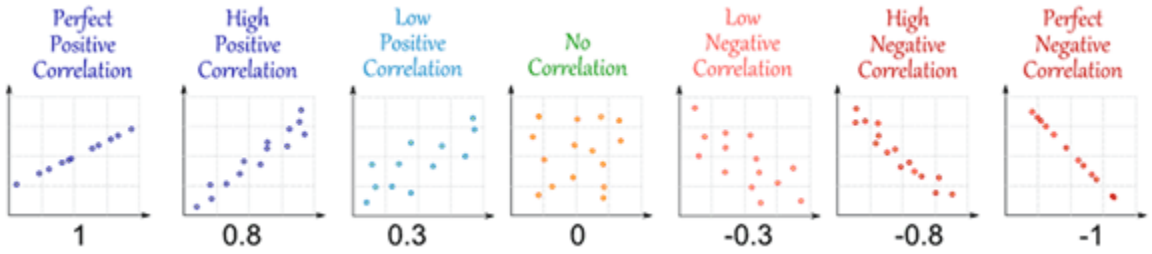






Hill's Causality Criteria (4 of 8)

- **Strength of Association**
  (if A → B, when A ↑ B will ↑)



- **Consistency**
  (if A → B, we will see it across multiple studies)

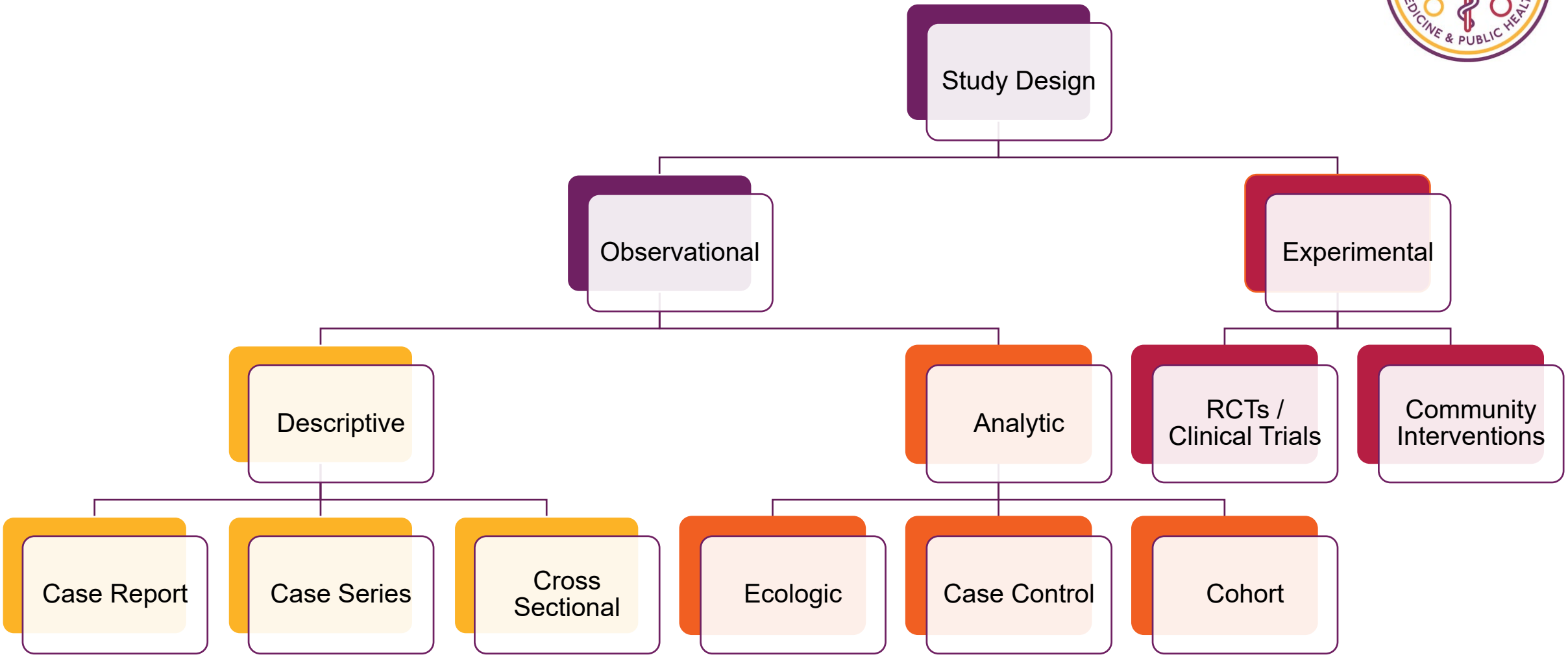
- **Temporality**
  (if A → B, A will always be present before B)

- **Biologic Gradient**
  (if a little of something is bad, a lot of something will be very bad; aka: dose-response relationship)

Table 1.

Studies that fail to support an association between measles-mumps-rubella vaccine and autism.

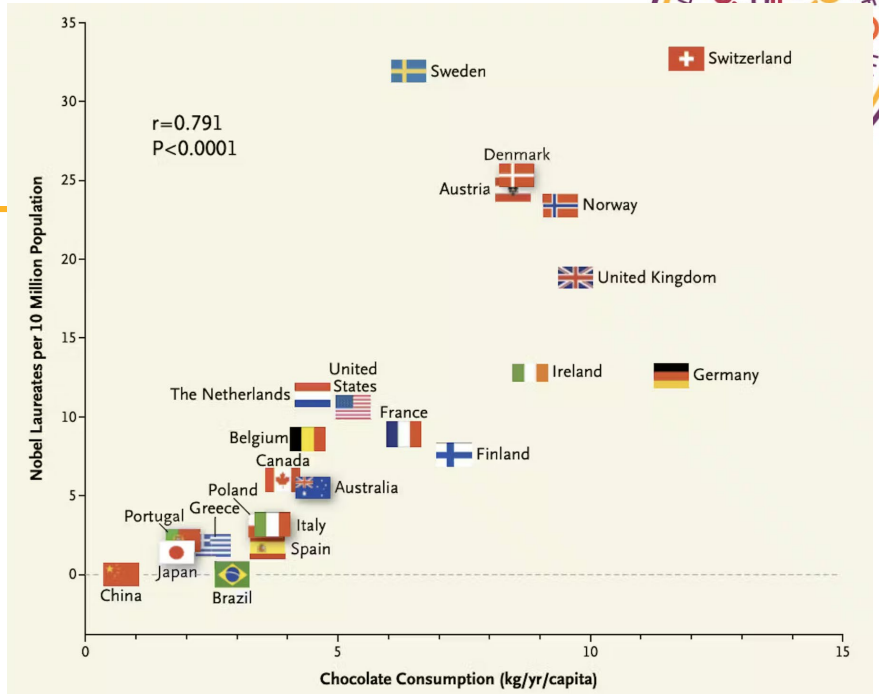
Source	Study design	Study location
Taylor et al., 1999 [5]	Ecological	United Kingdom
Farrington et al., 2001 [6]	Ecological	United Kingdom
Kaye et al., 2001 [7]	Ecological	United Kingdom
Dales et al., 2001 [8]	Ecological	United States
Fombonne et al., 2006 [9]	Ecological	Canada
Fombonne and Chakrabarti, 2001 [10]	Ecological	United Kingdom
Taylor et al., 2002 [11]	Ecological	United Kingdom
DeWilde et al., 2001 [12]	Case-control	United Kingdom
Makela et al., 2002 [13]	Retrospective cohort	Finland
Madsen et al., 2002 [14]	Retrospective cohort	Denmark
DeStefano et al., 2004 [15]	Case-control	United States
Peltola et al., 1998 [16]	Prospective cohort	Finland
Patja et al., 2000 [17]	Prospective cohort	Finland


Study Design
Observational
Experimental
Descriptive
Analytic
RCTs / Clinical Trials
Community Interventions
Case Report
Case Series
Cross Sectional
Ecologic
Case Control
Cohort

Analytic & Experimental Studies

Ecologic Studies

- Looks at groups (not individuals)
- Explores associations at group level
  - Is cancer mortality associated with income?
  - Often used when individual data are unavailable
  - Associations, NOT causes
- Strengths
  - Quick, data from multiple sources



› Health Aff (Millwood). 2020 Nov;39(11):1984-1992. doi: 10.1377/hlthaff.2020.01040. Epub 2020 Aug 27.

Community-Level Factors Associated With Racial And Ethnic Disparities In COVID-19 Rates In Massachusetts

Jose F Figueroa [1], Rishi K Wadhera [2], Dennis Lee [3], Robert W Yeh [4], Benjamin D Sommers [5]

Example: Ecologic

- Variables
  - Aggregate (summaries)
  - Environmental (mean temp)
  - Global (population density)

Health Aff (Millwood). 2020 Nov;39(11):1984-1992. doi: 10.1377/hlthaff.2020.01040. Epub 2020 Aug 27.

Community-Level Factors Associated With Racial And Ethnic Disparities In COVID-19 Rates In Massachusetts

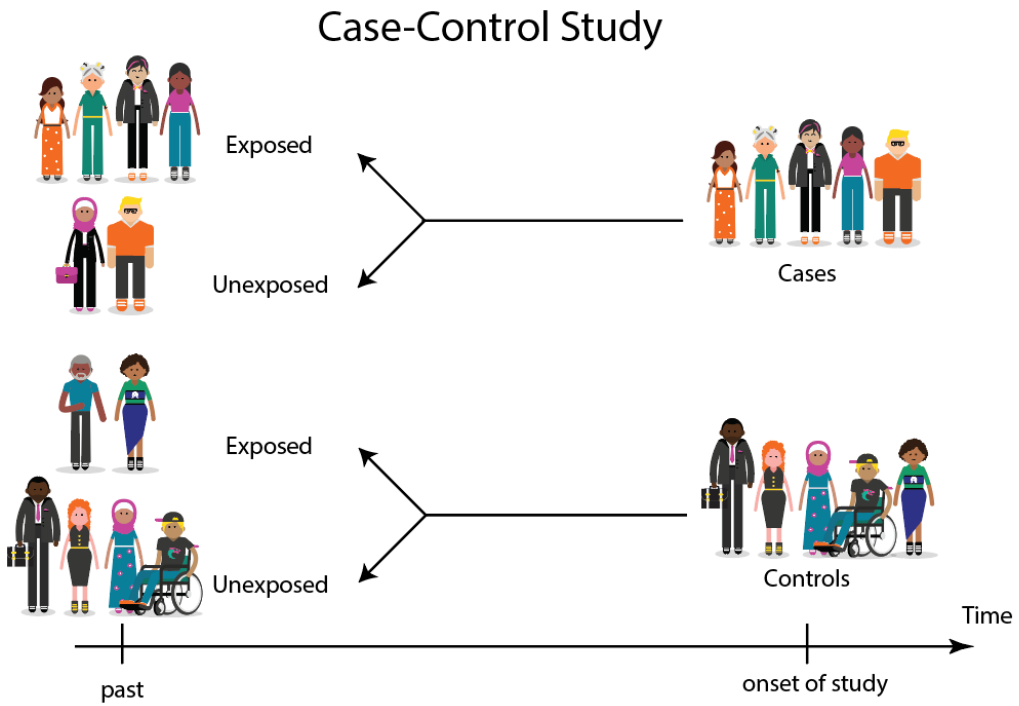
Jose F Figueroa [1], Rishi K Wadhera [2], Dennis Lee [3], Robert W Yeh [4], Benjamin D Sommers [5]

- Weaknesses
  - We cannot draw conclusions about individuals from group data ("Ecologic Fallacy")
  - Imprecise
  - Often relies on others' data

"A 10-percentage point increase in the Black non-Latino population was associated with an increase of 312.3 COVID-19 cases per 100,000 population."

…does NOT mean that a Black, non-Latino person is more likely to get COVID.

Case Control Study



- Subjects are defined based on presence or absence or an outcome of interest ('disease')
- Study is retrospective (looks backward)
- Theory
  - Find a group of sick people
  - Find a similar group of not sick people
  - Compare previous exposures to see if we can find a pattern that might explain why some are sick and some aren't
- Strengths
  - Good for studying rare diseases / outcomes
  - Multiple exposures can be studied
  - Inexpensive, fast
  - Calculate Odds Ratio – how strongly the exposure is associated with the outcome (probability an event will occur)
- Weaknesses
  - Cannot measure incidence (temporality)
  - Selecting appropriate control groups is hard
  - Self-reported exposure data

Example: Case Control

What is the association between poor diet and suffering from a torn ACL?

Outcome

Exposure

Cases: people with a torn ACL

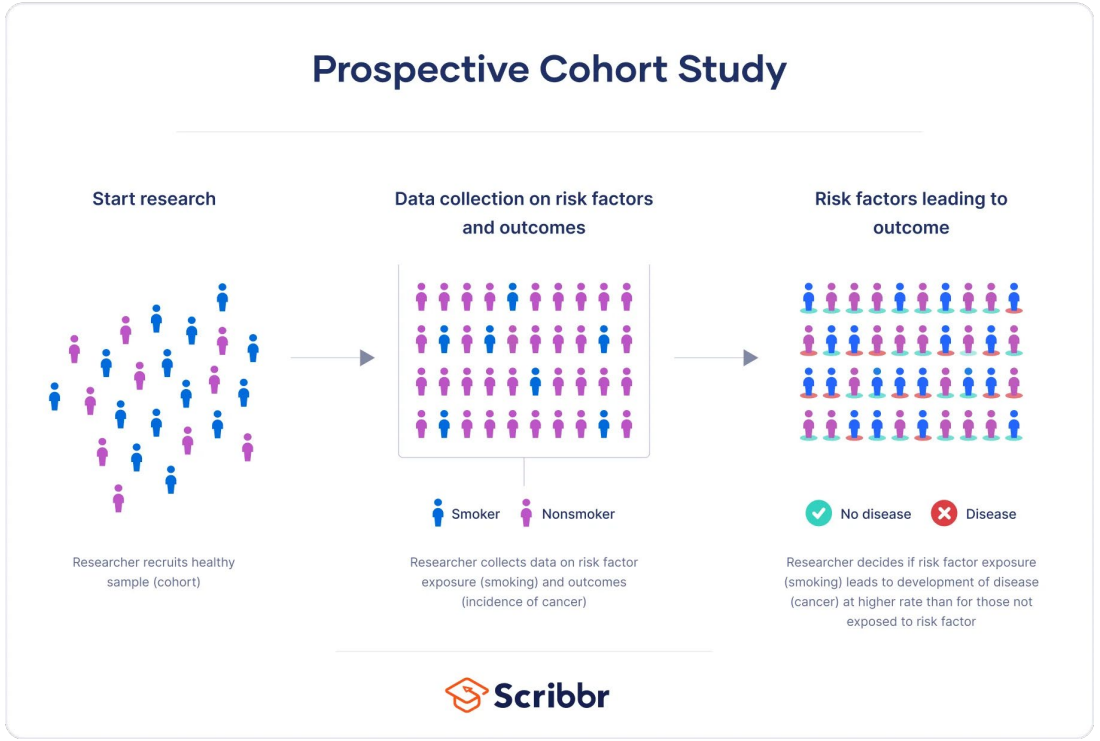
Controls: people without a torn ACL

Note: the only difference between the cases & controls should be the thing you're testing for

- Cases and controls could both be recruited from the same orthopedic surgery practice;
- Cases with an ACL tear, controls with a shoulder issue

Cohort Study

- Subjects are defined by presence or absence of an exposure
- Study is prospective (looks forward)
- Theory
  - Find a group of exposed people
  - Find a similar group of not exposed people
  - Follow them for a period of time to see who develops the outcome of interest
- Strengths
  - Good for rare exposures, evaluate multiple outcomes
  - Temporal relationship (causality)
  - Can calculate incidence, relative risk (of getting disease/outcome)
- Weakness
  - Expensive, time consuming, loss to follow up
  - Inefficient for rare diseases/outcomes



Example: Cohort

Researchers recruited 938 individuals to be in a study looking at the relationship between smoking and lung cancer. The baseline questionnaire asked subjects about their smoking status. Researchers then followed the individuals for 10 years to see who develops lung cancer.

Outcome of interest

At baseline, 438 subjects reported smoking, while 500 reported not smoking.

Exposure

No exposure

We can determine incidence (and causality)!

Framingham Heart Study

Objective:
To study the impact of several factors on incidence of cardiovascular disease

Exposures:
Blood Pressure
Smoking
Body Weight
Diabetes
Exercise
Etc.

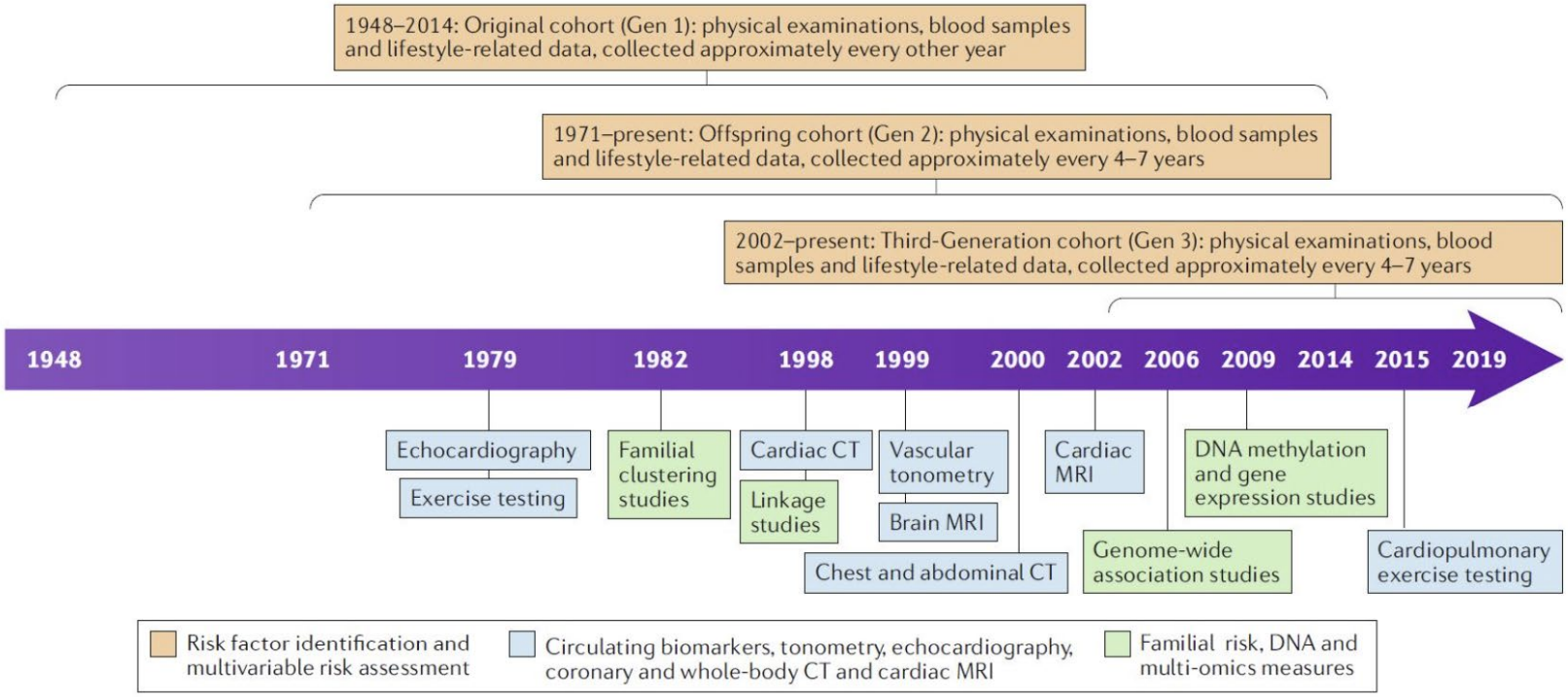


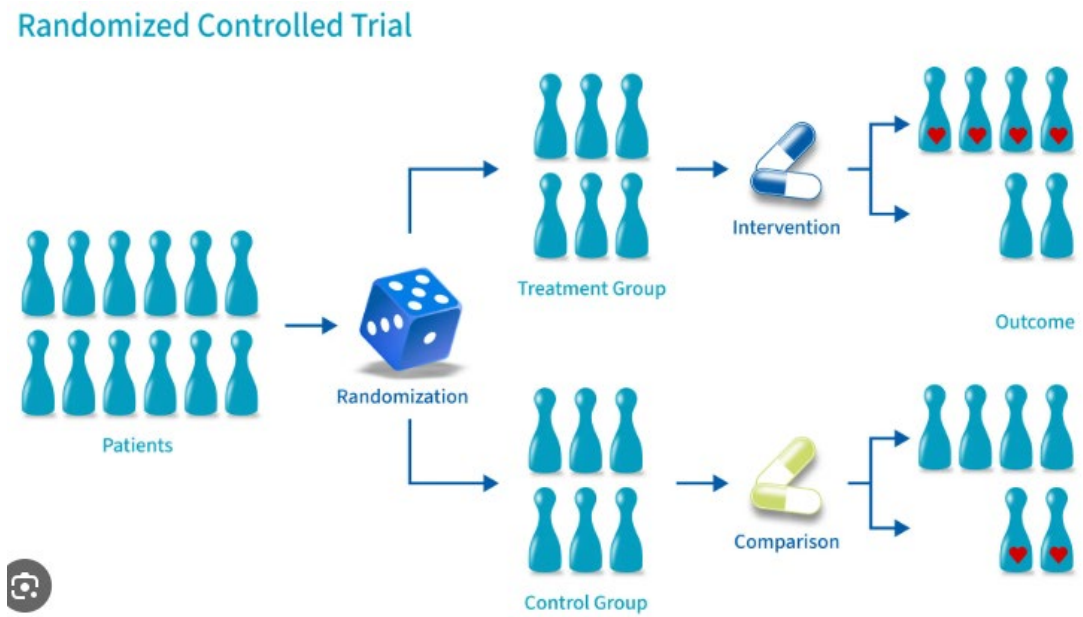
Fig. 1 | **Timeline of the Framingham Heart Study.** The timeline shows the temporal enrolment of the three generations of participants (Gen 1, Gen 2 and Gen 3) and when various measures and heritability studies were introduced to the Framingham Heart Study.

c/o Nature Reviews Cardiology

Outcomes

1948	Framingham Heart Study begins
1956	Findings on progression rheumatic heart disease reported
1957	High blood pressure and high cholesterol levels shown to increase likelihood of heart disease
1959	Some heart attacks discovered to be “silent” (causing no pain)
1960	Cigarette smoking found to increase the risk of heart disease
1961	The term “risk factor” introduced
1961	Cholesterol level, blood pressure, and electrocardiogram abnormalities found to increase the risk of heart disease
1967	Physical activity found to reduce the risk of heart disease and obesity to increase the risk of heart disease
1970	High blood pressure found to increase the risk of stroke
1970	Atrial fibrillation increases stroke risk 5-fold
1976	Menopause found to increase the risk of heart disease
1978	Psychosocial factors found to affect heart disease
1988	High levels of HDL cholesterol found to reduce risk of death
1994	Enlarged left ventricle (one of two lower chambers of the heart) shown to increase the risk of stroke
1996	Progression from hypertension to heart failure described
1998	Framingham Heart Study researchers identify that atrial fibrillation is associated with an increased risk of all-cause mortality
1998	Development of simple coronary disease prediction algorithm involving risk factor categories to allow physicians to predict multivariate coronary heart disease risk in patients without overt CHD
1999	Lifetime risk at age 40 years of developing coronary heart disease is one in two for men and one in three for women
2001	High-normal blood pressure is associated with an increased risk of cardiovascular disease, emphasizing the need to determine whether lowering high-normal blood pressure can reduce the risk of cardiovascular disease
2002	Lifetime risk of developing high blood pressure in middle-aged adults is 9 in 10
2002	Obesity is a risk factor for heart failure
2004	Serum aldosterone levels predict future risk of hypertension in non-hypertensive individuals
2005	Lifetime risk of becoming overweight exceeds 70 percent, that for obesity approximates 1 in 2
2006	The National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health announces a new genome-wide association study at the Framingham Heart Study in collaboration with Boston University School of Medicine to be known as the SHARe project (SNP Health Association Resource)

Randomized Controlled Trial



- Theory
  - Get a group of healthy people together that are as alike as possible
  - Randomized them into groups and assign exposure status randomly
    - Test / Treatment Group – gets the intervention
    - Control / Placebo Group – gets standard of care or placebo
    - Don't tell them which group they're in
  - Follow in time to see what their outcomes are
  - Strengths
    - Certainty exposure → outcome (causality)
    - All characteristics are randomly assigned
    - Blinding
  - Weaknesses
    - Expensive, time consuming
    - Can be unethical to withhold treatments
    - Can't randomly assign all exposures (smoking, pollution, etc.)

Example: RCTs

- Exposures
  - New drugs
  - New treatments
  - New technology
  - New methods of prevention
  - New screening programs

2. Materials and Methods

Patients at the lymphoedema treatment clinic at Ste. Croix Hospital in Leogane, Haiti, were eligible to participate if they lived within a 10 km radius of the hospital, were competent in lymphoedema self-care, and provided informed consent.

Research staff interviewed 200 patients with lymphoedema of the leg using a standardized questionnaire and randomly assigned them to receive a free supply of antibacterial soap, containing 1.2% triclocarban, or plain soap (Supplementary File 1). The unmarked bars of soap were packaged in individual boxes labeled only with the patient study identification number. Patients were requested to wash their legs daily for the next 12 months using only this soap and were provided antifungal and antiseptic creams for skin lesions as appropriate.

At the time of enrollment, during the spring of 2001, patients estimated the number of ADLA episodes during the previous 12 months. Lymphoedema stage was clinically assessed.[2] Clinic-based research staff made monthly visits to the patients' home or workplace to assess ADLA frequency during the previous month, assess the presence of skin lesions, monitor proficiency in washing, collect information on compliance with lymphoedema self-care and resupply soap and other needed supplies (e.g., antifungal creams).

Blinding or Masking

- Placebo Effect
  - People can have a response (positive and negative) to things they think have an effect
  - Even with a placebo, people can report improvement, side effects, etc.
  - The brain is weird
- Blinding
  - Single blind – participants are unaware of what group they're in
  - Double blind – neither participants nor investigators are aware of who has been assigned to which group

Analytic Study Break Down

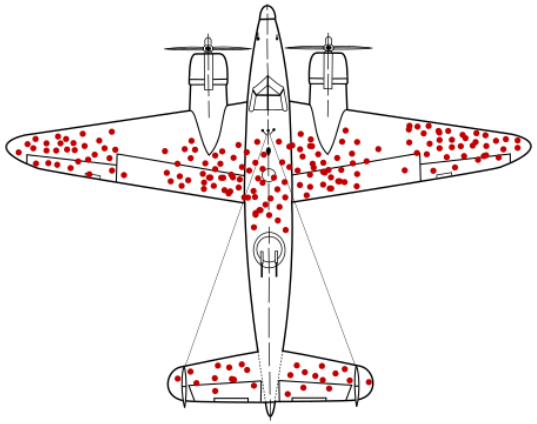
RCTs are called the “gold standard” of study design, but not everything can or should be studied using an RCT.

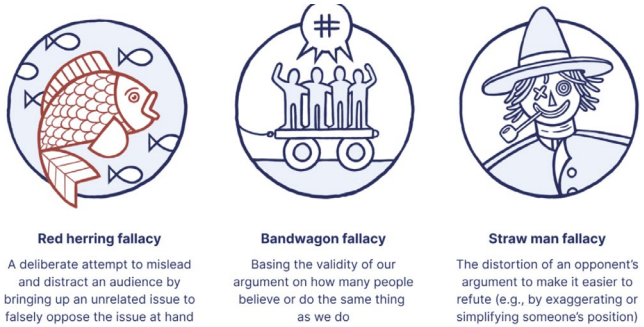
Attribute	Case Control	Cohort	RCT
Good for?	Rare outcomes	Rare exposures	Fair comparisons
Starts with?	Outcome assessment	Exposure assessment	Randomization
Looks	Backward	Forward	Forward
Calculate incidence?	No	Yes	Yes
Shows causality?	No	Can	Can

But wait, there's more!

- Epidemiology also looks at
  - Efficacy of screening tests
    - If you get a positive test result, do you actually have cancer / are pregnant, etc.?
    - False positives and false negatives
  - Confounders
    - Is the exposure just guilty by association and something else is actually causing the outcome?
  - Types of Bias
    - Availability – using information that's most easily/immediately accessible
    - Confirmation – data supports preconceived ideas
    - Conformity – change ideas to fit with the group
    - Sampling – smaller populations that don't represent the whole
    - **Survivorship – only evaluating surviving data, and not the data as a whole**
    - Recall – outcomes may color subjects' recollections of prior events
  - Logical Fallacies
    - Cherry picking data
    - False balance
    - Red herring
    - Straw man
    - Circular reasoning

	Test says you don't have it	Test says you do have it
You really don't have it	TRUE NEGATIVE	FALSE POSITIVE
You really do have it	FALSE NEGATIVE	TRUE POSITIVE





Red herring fallacy

A deliberate attempt to mislead and distract an audience by bringing up an unrelated issue to falsely oppose the issue at hand

Bandwagon fallacy

Basing the validity of our argument on how many people believe or do the same thing as we do

Straw man fallacy

The distortion of an opponent's argument to make it easier to refute (e.g., by exaggerating or simplifying someone's position)

Questions?