Measuring disease occurrence & association

Slides developed by Madhukar Pai, MD, PhD

Email: madhukar.pai@mcgill.ca

McGill

Lecture adapted and presented by Ngozi Erondu MPH, PHD

Email: Ngozierondu@gmail.com

CHATHAM HOUSE

Objectives

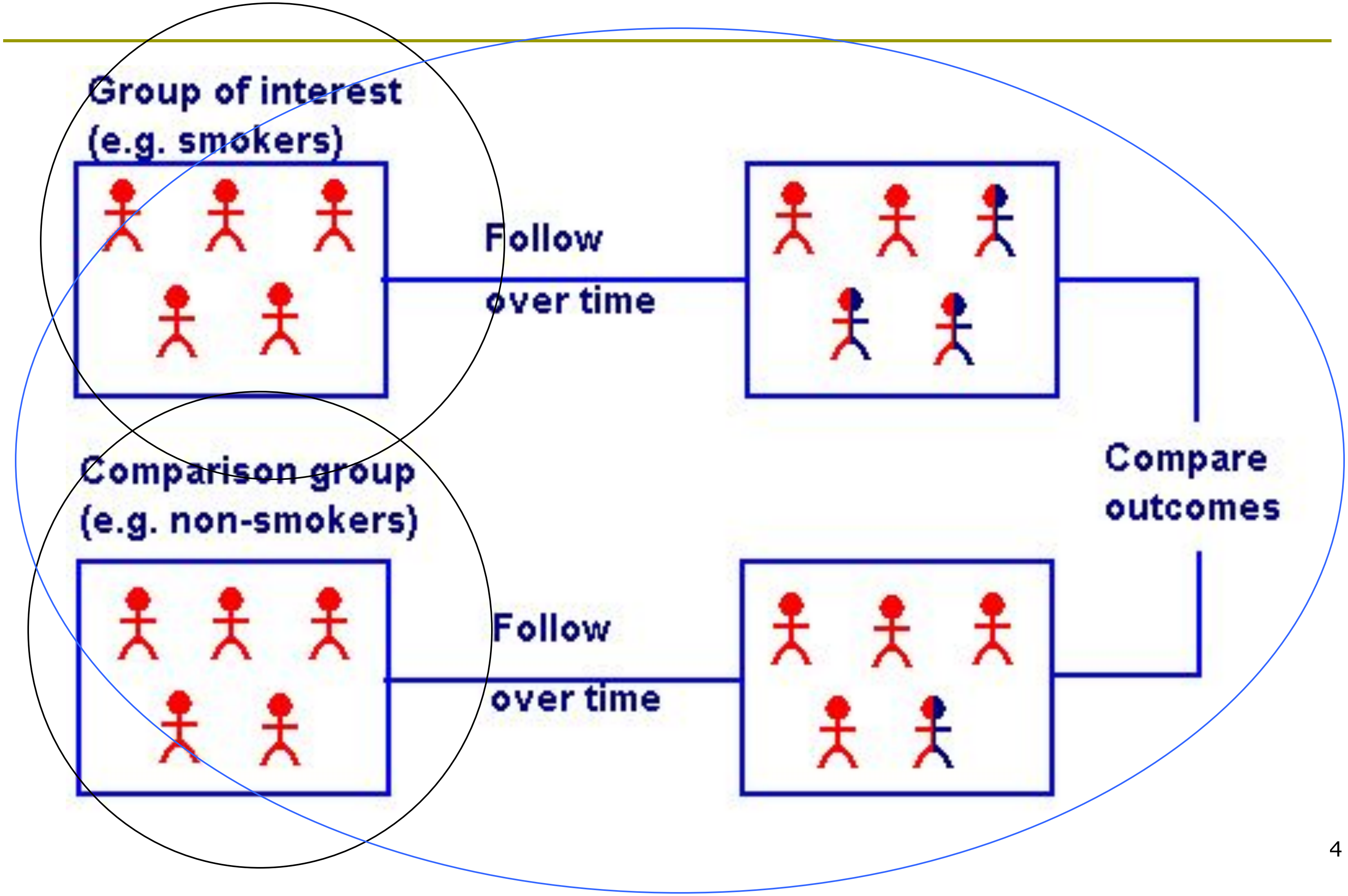
- Define commonly used terms and concepts in measuring and estimating disease occurrence and association
- Simplify interpretation of these measurements
- Provide examples of correct reporting of these measurements and epidemiological information

The concept of 'Cohort'

- Derived from Latin word 'cohorts' meaning enclosure, company, or crowd

- An epidemiological cohort is a group of people in a defined population that with something in common, such as
  - Geography (E.g. country, city)
  - Exposure (E.g. behavior such as smoking)
  - Outcome (E.g. disease such as lung cancer)
  - Occupation (e.g. Health care workers)

Cohort


Group of interest
(e.g. smokers)
Follow
over time
Comparison group
(e.g. non-smokers)
Follow
over time
Compare
outcomes

Morbidity and mortality

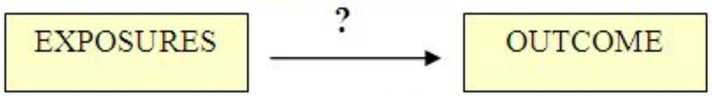
- **Morbidity** - any departure, subjective or objective, from a state of physiological or psychological well-being. It encompasses disease, injury, and disability.

- **Mortality** - is related to the number of deaths caused by the health event under investigation. It can
  be communicated as a rate or as an absolute number. A mortality rate is a measure of the frequency of occurrence of death in a defined population during a specified interval.

Both can be represented or estimated using different measures

Source: Centers for Disease Control and Prevention (CDC), Principles of Epidemiology and public health, 3rd Edition

AN OVERVIEW OF MEASUREMENTS IN EPIDEMIOLOGY [VER 3, 2007]

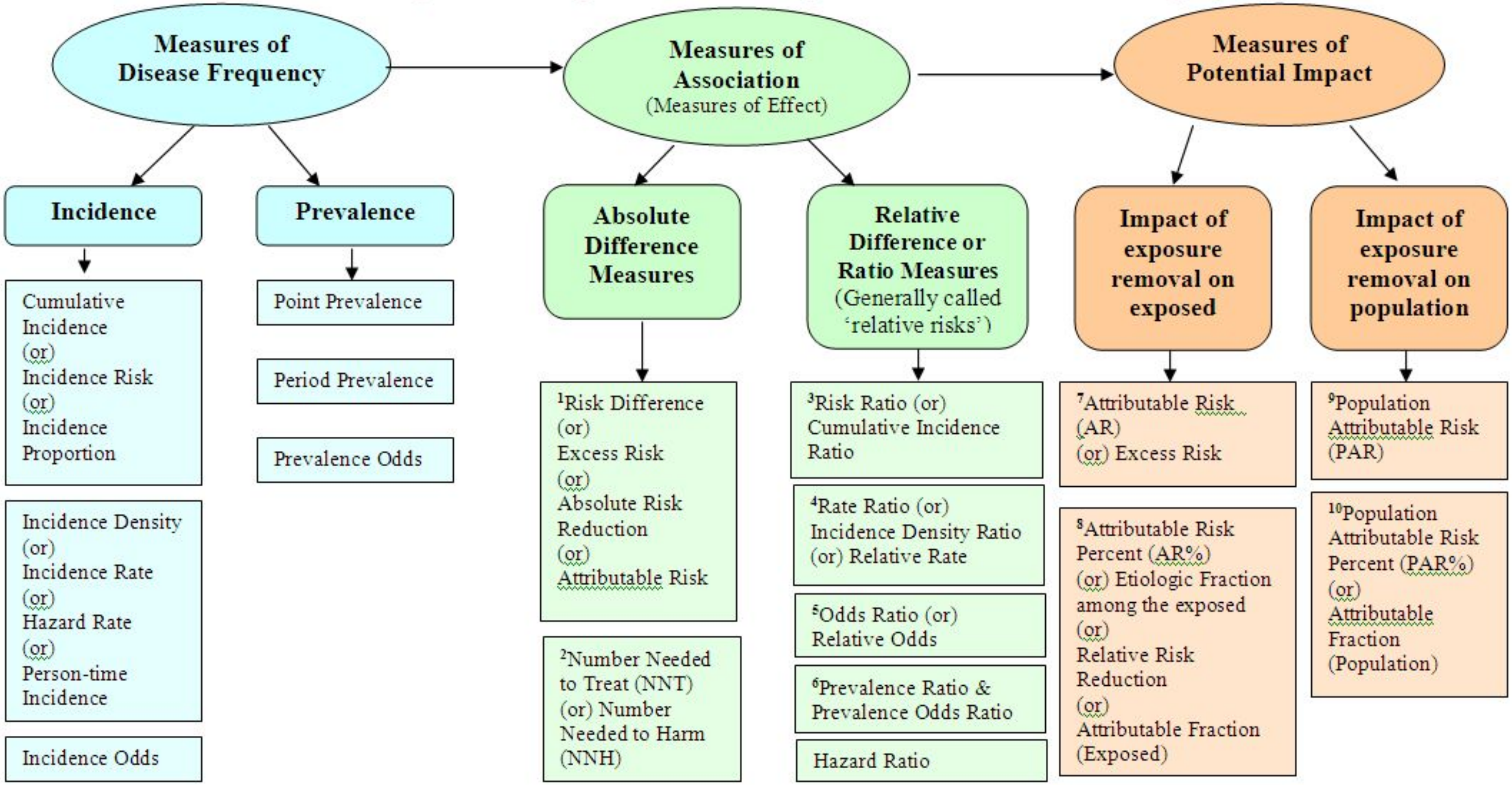


EXPOSURES

?

OUTCOME

Epidemiology is about identifying associations between exposures and outcomes. To identify any association, exposures and outcomes must first be measured in a quantitative manner. Then rates of occurrence of events are computed. These measures are called "*measures of disease frequency.*" Once measured, the association between exposures and outcomes are then evaluated by calculating "*measures of association or effect.*" Finally, the impact of removal of an exposure on the outcome is evaluated by computing "*measures of potential impact.*" In general, measures of disease frequency are needed to generate measures of association, and both these are needed to get measures of impact. There is some overlap between these measures, and terminology is poorly standardized.



Measures of Disease Frequency

Incidence

Cumulative Incidence (or) Incidence Risk (or) Incidence Proportion

Incidence Density (or) Incidence Rate (or) Hazard Rate (or) Person-time Incidence

Incidence Odds

Prevalence

Point Prevalence

Period Prevalence

Prevalence Odds

Measures of Association (Measures of Effect)

Absolute Difference Measures

[1]Risk Difference (or) Excess Risk (or) Absolute Risk Reduction (or) Attributable Risk

[2]Number Needed to Treat (NNT) (or) Number Needed to Harm (NNH)

Relative Difference or Ratio Measures (Generally called 'relative risks')

[3]Risk Ratio (or) Cumulative Incidence Ratio

[4]Rate Ratio (or) Incidence Density Ratio (or) Relative Rate

[5]Odds Ratio (or) Relative Odds

[6]Prevalence Ratio & Prevalence Odds Ratio

Hazard Ratio

Measures of Potential Impact

Impact of exposure removal on exposed

[7]Attributable Risk (AR) (or) Excess Risk

[8]Attributable Risk Percent (AR%) (or) Etiologic Fraction among the exposed (or) Relative Risk Reduction (or) Attributable Fraction (Exposed)

Impact of exposure removal on population

[9]Population Attributable Risk (PAR)

[10]Population Attributable Risk Percent (PAR%) (or) Attributable Fraction (Population)

Rates, Ratios, Proportions

- Three general classes of mathematical parameters.

- Often used to relate the number of cases of a disease [numerator] or health outcome to the size of the source population [denominator] in which they occurred.

- Numerator (“case”) has to be defined
- Denominator (“population size”) has to be defined
  - Epidemiologists have been referred to as “people in search of the denominator”!

Ratio

- Obtained by dividing one quantity by another. These quantities may be related or may be totally independent.

- Usually expressed as: $\frac{x}{y} \times 10^{n}$

Example: Number of stillbirths per thousand live births.

$$\frac{\#\,\text{stillbirths}}{\#\,\text{live births}} \times 1000$$

- **“Ratio” is a general term that includes Rates and Proportions.**

- Dictionary: “The value obtained by dividing one quantity by another.” [Porta 2008]

Kleinbaum et al. ActivEpi
www.activepi.com

Proportion

- A ratio in which the numerator (x) is included in the denominator (y)
- Expressed as: $\frac{x}{y}\times 10^n$ where, 10^n is often 100.

Example: The number of fetal deaths out of the total number of births.

$$\frac{\#\text{ of fetal deaths}}{\text{live births} + \text{fetal deaths}}\times 100$$

- **Answer often read as a percent.**
- Dictionary: “A type of ratio in which the numerator in included in the denominator.” [Porta 2008]

Kleinbaum et al. ActivEpi

Risk

Probability that an individual with certain characteristics such as:
Age, Race, Sex

will experience a health status change over a specified follow-up period (i.e. risk period)

Dictionary: “Probability that an event will occur within a stated period of time.” [Porta 2008]

$0 \leq \text{RISK} \leq 1$

$0\% \leq \text{percentage} \leq 100\%$

Assumes:
Does not have disease at start of follow-up.
Does not die from other cause during follow-up (no competing risks).

Risk is often used for prediction at the individual level

Rate

- A measure of how quickly something of interest happens (time is automatically captured)
- Expressed as: $\frac{x}{y} \times 10^{n}$

Example: The number of new cases of Parkinson's disease which develops per 1,000 person-years of follow-up.

$$\frac{\text{\# of new cases of Parkinson's disease}}{\text{Total time disease - free subjects observed}} \times 1000$$

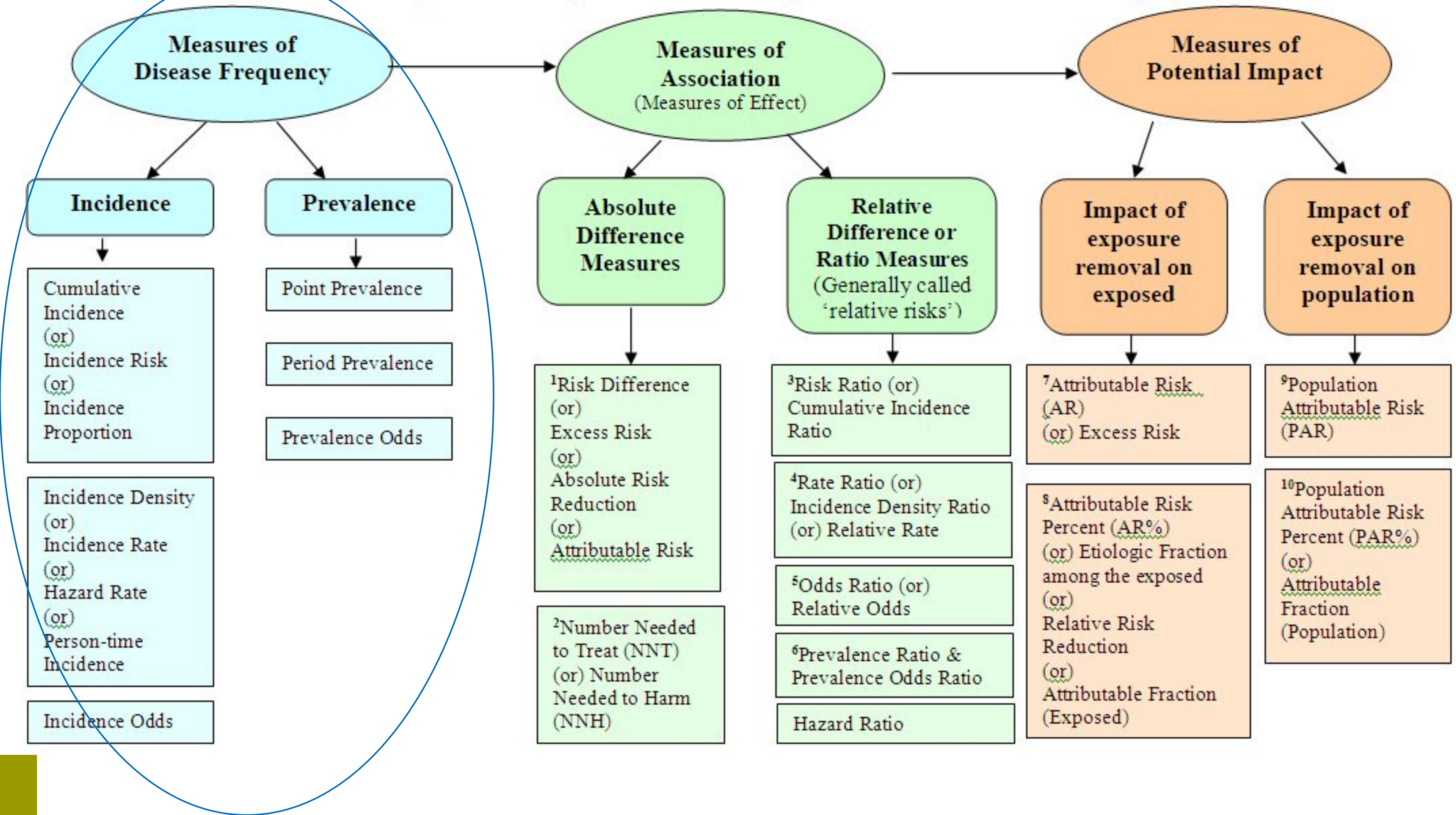
- Time is already in the denominator
- Place and population must be specified for each type of rate.
- In a rate, numerator is not a subset of the denominator
- **Rate is not a proportion (and cannot be a %)**

Kleinbaum et al. ActivEpi

AN OVERVIEW OF MEASUREMENTS IN EPIDEMIOLOGY [VER 3, 2007]

EXPOSURES → ? → OUTCOME

Epidemiology is about identifying associations between exposures and outcomes. To identify any association, exposures and outcomes must first be measured in a quantitative manner. Then rates of occurrence of events are computed. These measures are called *"measures of disease frequency."* Once measured, the association between exposures and outcomes are then evaluated by calculating *"measures of association or effect."* Finally, the impact of removal of an exposure on the outcome is evaluated by computing *"measures of potential impact."* In general, measures of disease frequency are needed to generate measures of association, and both these are needed to get measures of impact. There is some overlap between these measures, and terminology is poorly standardized.



Measures of Disease Frequency

Incidence

Cumulative Incidence (or) Incidence Risk (or) Incidence Proportion

Incidence Density (or) Incidence Rate (or) Hazard Rate (or) Person-time Incidence

Incidence Odds

Prevalence

Point Prevalence

Period Prevalence

Prevalence Odds

Measures of Association (Measures of Effect)

Absolute Difference Measures

[1]Risk Difference (or) Excess Risk (or) Absolute Risk Reduction (or) Attributable Risk

[2]Number Needed to Treat (NNT) (or) Number Needed to Harm (NNH)

Relative Difference or Ratio Measures (Generally called 'relative risks')

[3]Risk Ratio (or) Cumulative Incidence Ratio

[4]Rate Ratio (or) Incidence Density Ratio (or) Relative Rate

[5]Odds Ratio (or) Relative Odds

[6]Prevalence Ratio & Prevalence Odds Ratio

Hazard Ratio

Measures of Potential Impact

Impact of exposure removal on exposed

[7]Attributable Risk (AR) (or) Excess Risk

[8]Attributable Risk Percent (AR%) (or) Etiologic Fraction among the exposed (or) Relative Risk Reduction (or) Attributable Fraction (Exposed)

Impact of exposure removal on population

[9]Population Attributable Risk (PAR)

[10]Population Attributable Risk Percent (PAR%) (or) Attributable Fraction (Population)

Measures of Disease Frequency

- **Incidence** (I): Measures new cases of a disease that develop over a period of time.
  - Very helpful for etiological/causal inference
  - Difficult to estimate
  - Implies follow-up over time (i.e. cohort design)

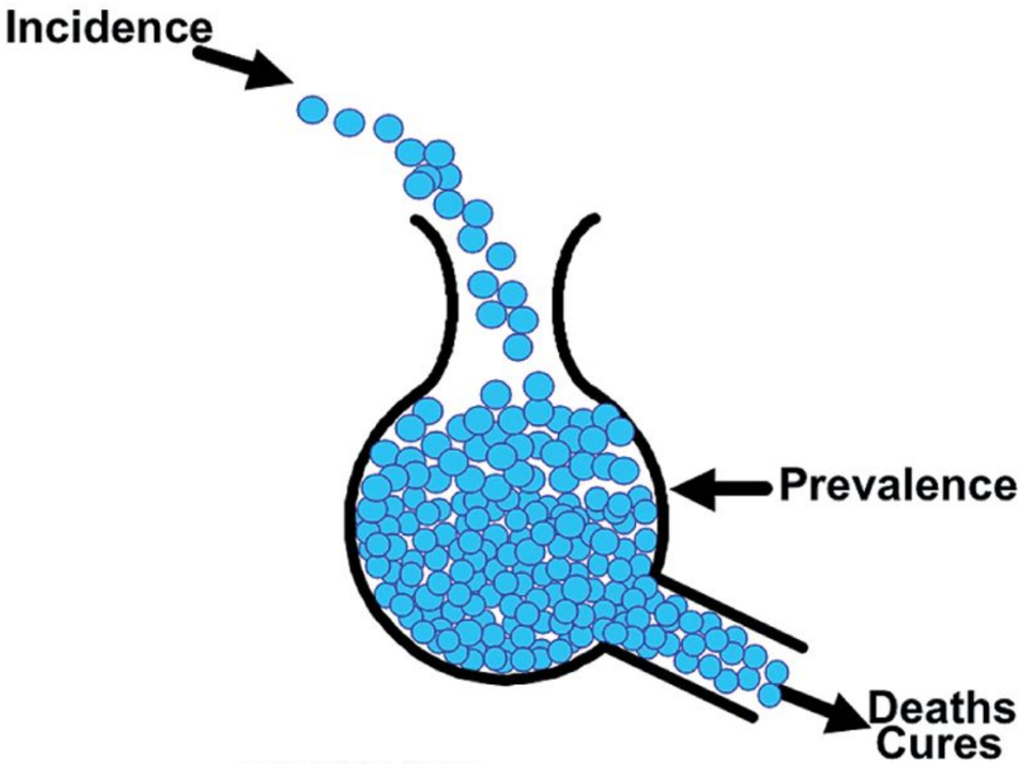
Think “new”

- **Prevalence** (P): Measures existing cases of a disease at a particular point in time or over a period of time.
  - Very helpful for quantifying disease burden (e.g. public health)
  - Relatively easy to estimate
  - Implies a cross-sectional design

Think “current state”

Kleinbaum et al. ActivEpi

Prevalence vs. Incidence



Gordis: Epidemiology, 4th Edition.
Copyright © 2008 by Saunders, an imprint of Elsevier, Inc. All rights reserved

Relationship between incidence and prevalence: IV.

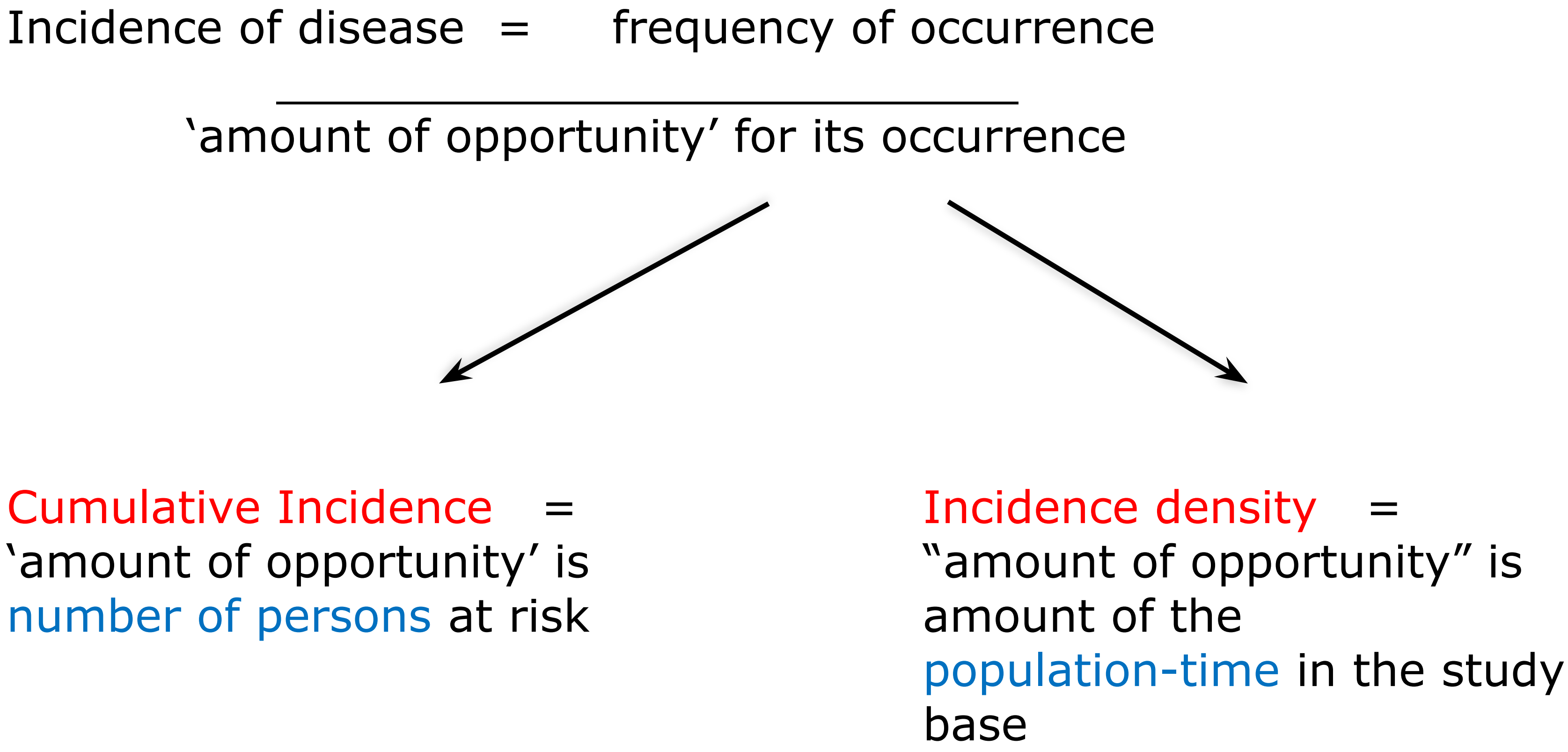
Prevalence can be viewed as describing a pool of disease in a population.

Incidence describes the input flow of new cases into the pool.

Deaths and cures reflects the output flow from the pool.

Prevalence = Incidence Rate X Average Duration

Incidence measures (big picture)



Incidence of disease = frequency of occurrence / ‘amount of opportunity’ for its occurrence

Cumulative Incidence = ‘amount of opportunity’ is number of persons at risk

Incidence density = “amount of opportunity” is amount of the population-time in the study base

Adapted from OS Miettinen, Epidemiological research: terms and concepts, Springer, 2011

Cumulative Incidence

$$CI = \frac{I}{N}$$

I = # of new cases during follow-up
N = # of disease-free subjects at start of follow-up (they should be ‘at risk’)

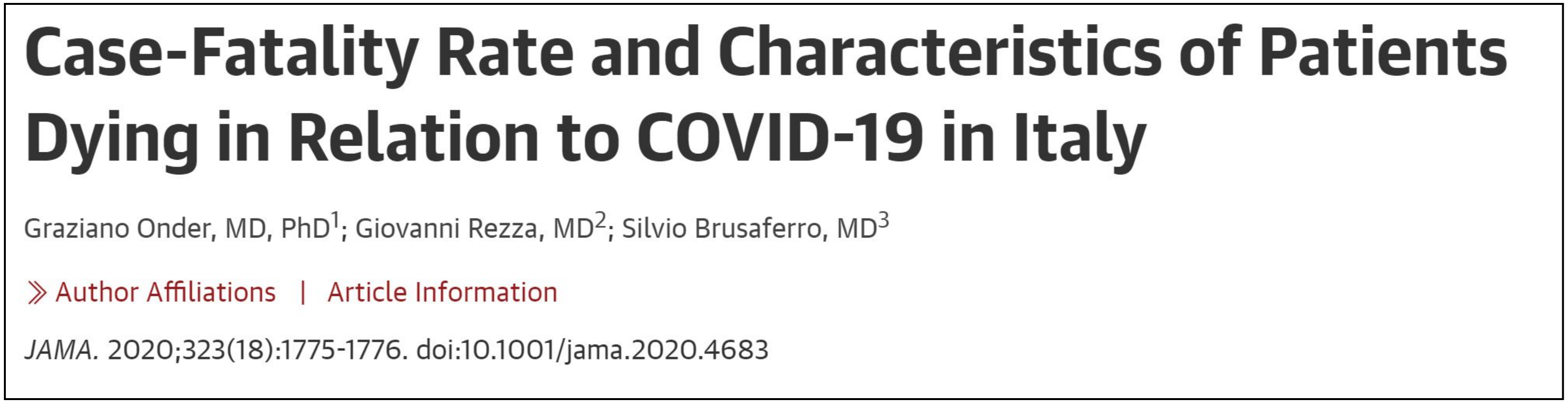
Measures the frequency of addition of new cases of disease and is always calculated for a given period of time (e.g. **annual** incidence)

- Must always state the time period (e.g. attack “rate” calculated for an outbreak)

- Most common way to estimate risk

- Not great if population changes a lot (e.g. attrition, competing risk)

Example



Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy

Graziano Onder, MD, PhD[1]; Giovanni Rezza, MD[2]; Silvio Brusaferro, MD[3]

» Author Affiliations | Article Information

JAMA. 2020;323(18):1775-1776. doi:10.1001/jama.2020.4683

- The fatality rate was defined as number of deaths in persons who tested positive for SARS-CoV-2 divided by number of SARS-CoV-2 cases.
- 1625 deaths
- 22,512 persons with confirmed COVID-19 in Italy
- CFR = 1625/22512 = 7.2%
- 95% confidence interval: 6.9% to 7.6%

Incidence density (incidence rate)

$$IR = \frac{I}{PT}$$

I = # of new cases during follow-up
PT = total time that disease–free individuals in the cohort are observed over the study period (total person-time experience of the cohort).

Describes how rapidly health events are occurring in a population of interest

Dictionary: “The average person-time incidence rate” [Porta, 2008]

Measures the rapidity with which new cases are occurring in a population

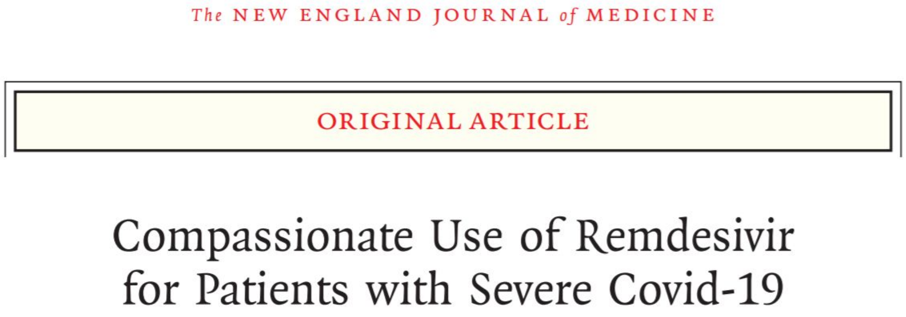
Most sophisticated form of measuring incidence [most difficult as well]

- Accounts for losses, competing risks, dynamic turn-over, differential follow-up time, changes in exposures over time

- *hazard function (in survival analysis) is the event rate at time *t* conditional on survival until time *t* [hazard rate is something like an instantaneous rate]

Kleinbaum et al. ActivEpi

Example



The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Compassionate Use of Remdesivir for Patients with Severe Covid-19

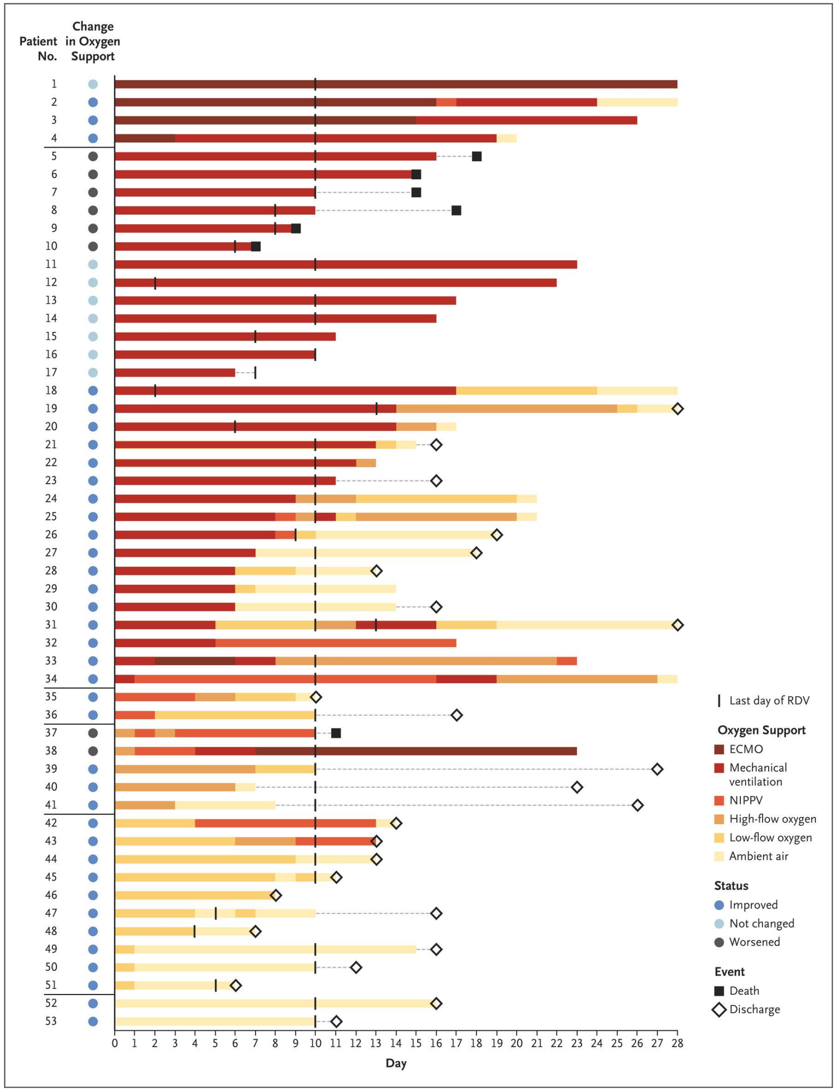
53 patients got the drug; follow-up was to continue through at least 28 days after the beginning of Rx with remdesivir or until discharge or death.

7 patients died

Cumulative incidence: 7/53 (13%)

Incidence density: 7/1120 person-days

= 0.63 deaths per 100 person-days
= 6.3 deaths per 1000 person-days



Summary: Risk vs Rate

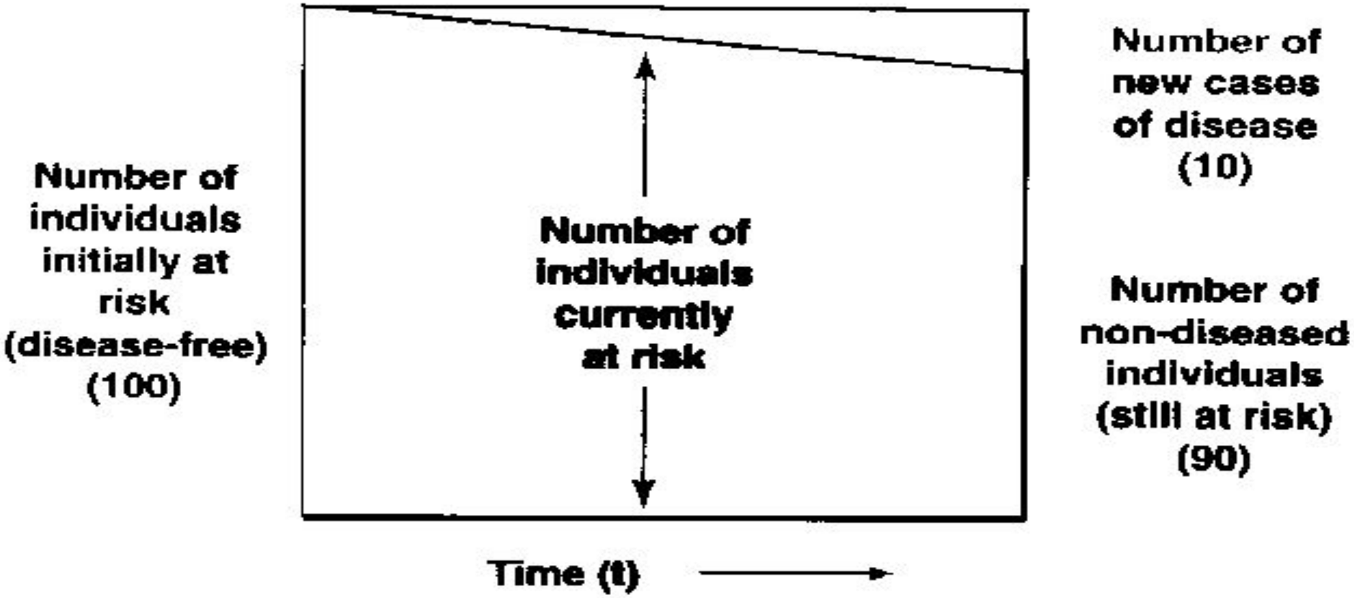
RISK

- E.g. Cumulative incidence
- Proportion (always between 0 and 1)
- Probability that an individual will develop a disease during a specific period
- Use for individual prognosis
- More assumptions
- Cannot handle variable follow-up times, attrition, competing risks
- Easy to compute in a fixed cohort with few losses; but gets difficult with open populations with longer follow up and losses

RATE

- E.g. Incidence density
- Non-negative and no upper bound
- Describes how rapidly new events occur in a specific population
- Use for etiological comparisons
- Fewer assumptions
- Can handle variable follow-up times, attrition, competing risks
- Can be computed even with open populations with losses and longer follow up

Risk vs Odds


Number of individuals initially at risk (disease-free) (100)
Number of individuals currently at risk
Number of new cases of disease (10)
Number of non-diseased individuals (still at risk) (90)
Time (t)

Thus, it is possible to calculate the risk and the odds of developing the disease during the study period as:

Risk = 10/100 = 0.10 = 10%

Odds of disease = 10/90 = 0.11 = 11%

Dictionary: “Odds is the ratio of the probability of occurrence of an event to that of non-occurrence.” [Porta, 2008]

Source: Silva 1999

Prevalence

- Measures existing cases of a health condition
  - Inherently biased towards inclusion of “survivors”

- Primary outcome of a cross-sectional study (e.g. sample surveys)

- Two types of Prevalence
  - Point prevalence
  - Period prevalence

Point Prevalence

$$P = \frac{C}{N}$$

C = # of observed cases at time t

N = Population size at time t

Measures the frequency of disease at a given point in time

Dictionary: “A measure of disease occurrence: the total number of individuals who have an attribute or disease at a particular time (or period) divided by the population at risk of having the disease at that time or midway through the period. It is a proportion, not a rate.” [Porta 2008]

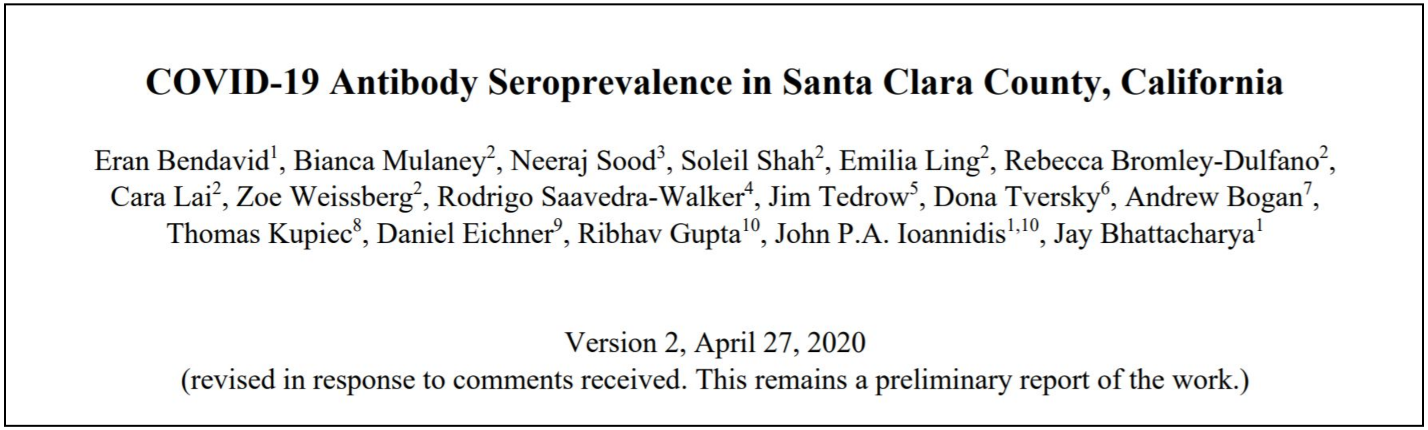
Period Prevalence

$$PP = \frac{C + I}{N}$$

- C = the # of prevalent cases at the beginning of the time period.
- I = the # of incident cases that develop during the period.
- N = size of the population for this same time period.

Example: one year prevalence: proportion of individuals with the disease at any time during a calendar year. It includes cases arising before and during the year. Denominator is total population during the time period.

Point Prevalence Example



COVID-19 Antibody Seroprevalence in Santa Clara County, California

Eran Bendavid[1], Bianca Mulaney[2], Neeraj Sood[3], Soleil Shah[2], Emilia Ling[2], Rebecca Bromley-Dulfano[2], Cara Lai[2], Zoe Weissberg[2], Rodrigo Saavedra-Walker[4], Jim Tedrow[5], Dona Tversky[6], Andrew Bogan[7], Thomas Kupiec[8], Daniel Eichner[9], Ribhav Gupta[10], John P.A. Ioannidis[1,10], Jay Bhattacharya[1]

Version 2, April 27, 2020
(revised in response to comments received. This remains a preliminary report of the work.)

- April 3rd and 4th, 2020, researchers did serologic testing for SARS-CoV-2 antibodies in 3,330 adults and children in Santa Clara County
- Total number of positive cases by antibodies = 50
- Crude point prevalence = 50/3330 = 1.5% (95 CI 1.1-2.0%)

Prevalence

Useful for:

- Assessing the health status of a population.
- Planning health services.
- Often the only measure possible with chronic diseases where incident cases cannot be easily detected (e.g. prevalence of hypertension)

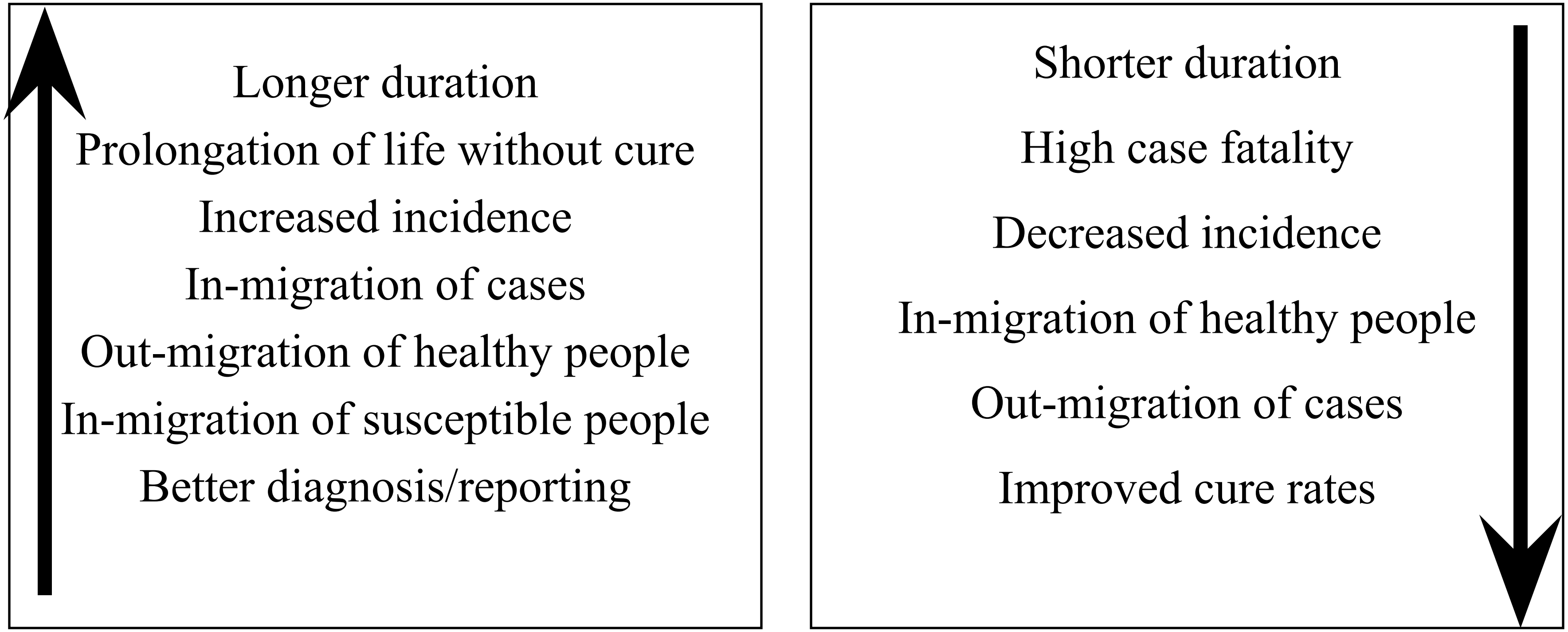
Not very useful for:

- Identifying risk factors (etiology): confusion between risk factors for survival vs. risk factors for developing disease
- Makes no sense for conditions that are acute and short duration (e.g. diarrhea)

Kleinbaum et al. ActivEpi

What factors can affect prevalence?

Prevalence



Longer duration

Prolongation of life without cure

Increased incidence

In-migration of cases

Out-migration of healthy people

In-migration of susceptible people

Better diagnosis/reporting

Shorter duration

High case fatality

Decreased incidence

In-migration of healthy people

Out-migration of cases

Improved cure rates

Source: Beaglehole, 1993

Be critical when reviewing results

- Crude rates
- Confounding factors
- Confidence intervals

Crude vs. adjusted rates

- Crude rates are useful, but not always comparable across populations
- Example: crude death rate in Sweden is higher than in Panama
  Why?
- Confounding by age
- Age standardization is nothing but adjustment for confounding by age

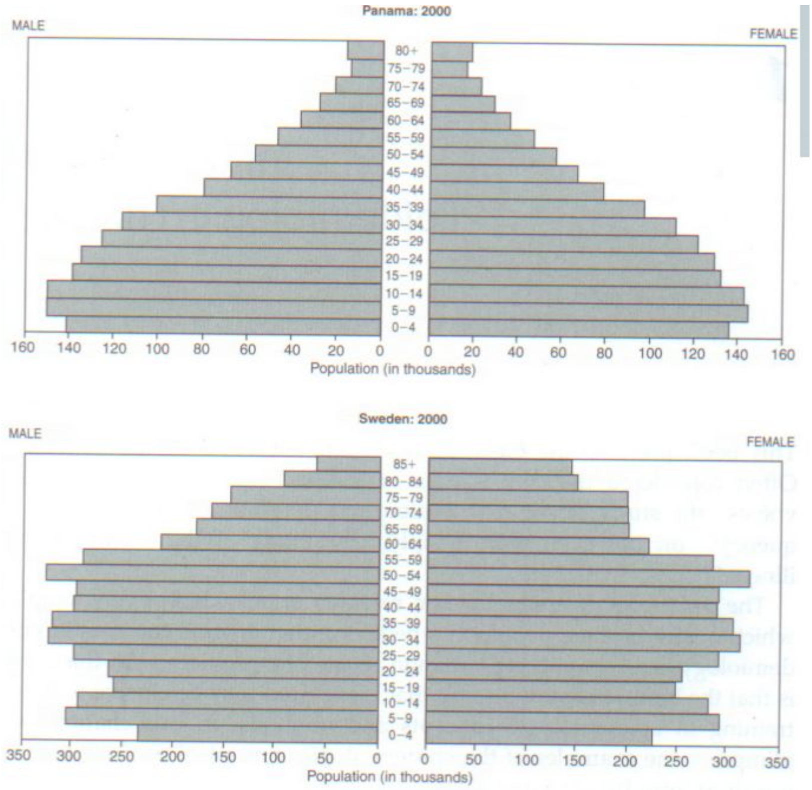


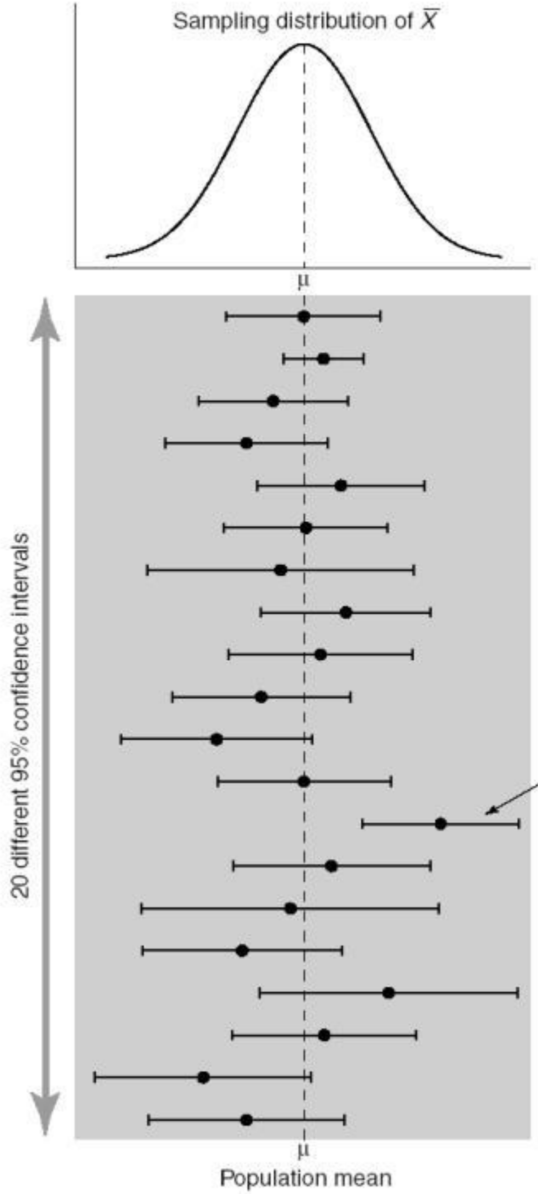
Figure 1–1. Age distribution of the populations of Panama and Sweden (population pyramids). Source: U.S. Census Bureau, International Data Base.

Rothman KJ, 2002

Let’s imagine the true population seroprevalence of Covid-19 in Santa Clara county is 5% (population mean)

Let’s say 100 samples were taken in that county and 100 estimates and confidence intervals were constructed

95% of the intervals will capture the true population prevalence of 5%



Look at this interval.
It "missed" the
population parameter!

FIGURE 16.2 A sampling distribution of the mean (based on all possible samples of size 100) and an illustration of the 95 percent confidence intervals for twenty possible samples. The width of the intervals will be slightly different because they are estimated from different random samples. In the long run, 95 percent of confidence intervals will capture the population mean.

http://www.southalabama.edu/coe/bset/johnson/dr_johnson/index.htm

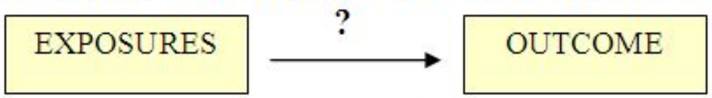
What are 95% confidence intervals?

- The interval computed from the sample data which, were the study repeated multiple times, would contain the true effect 95% of the time

- Incorrect Interpretation: "There is a 95% probability that the true effect is located within the confidence interval."
  - This is wrong because the true effect (i.e. the population parameter) is a constant, not a random variable. Its either in the confidence interval or it's not. There is no probability involved (in other words, truth does not vary, only the confidence interval varies around the truth).

Useful reading: Primer on 95% CI by American College of Physicians

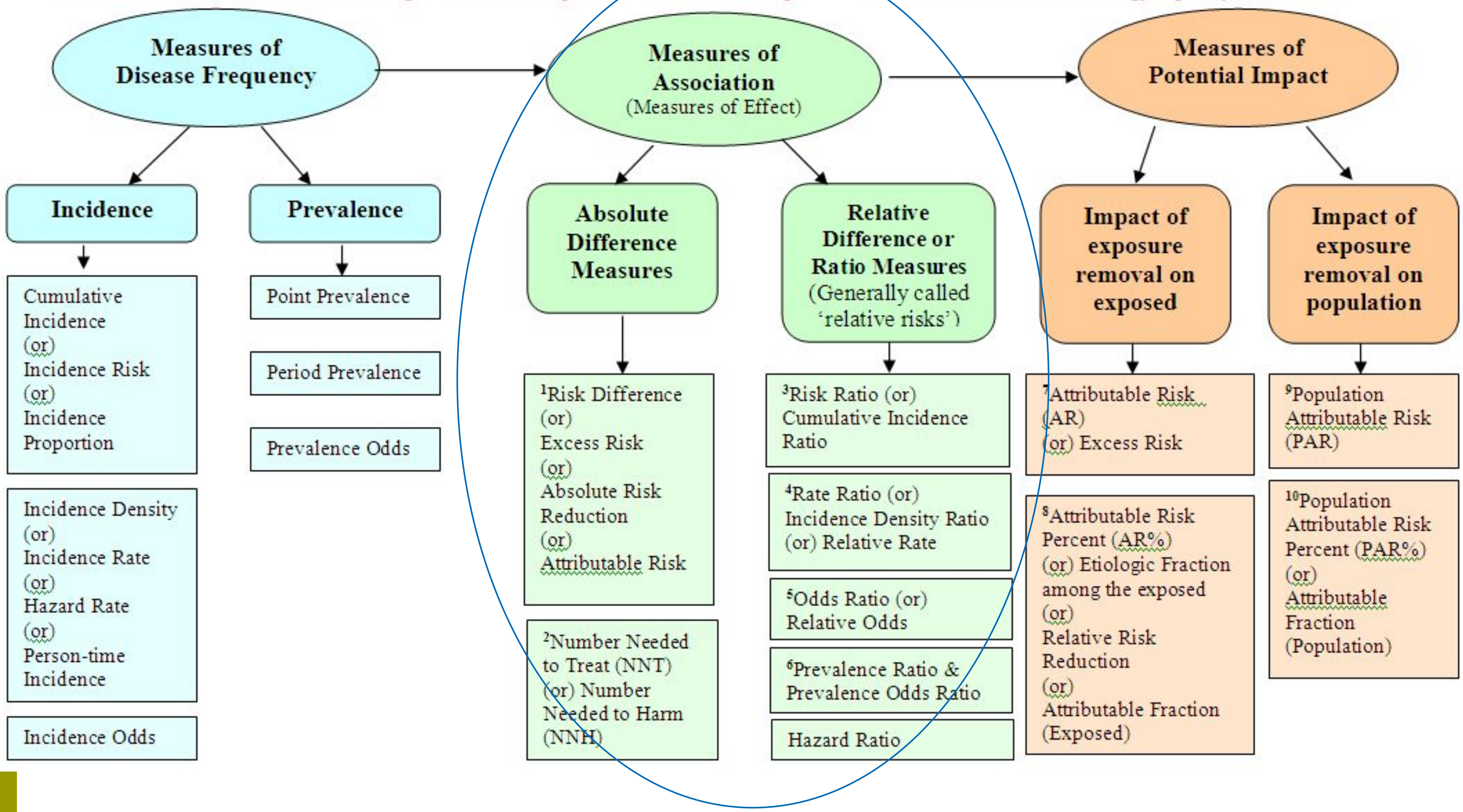
AN OVERVIEW OF MEASUREMENTS IN EPIDEMIOLOGY [VER 3, 2007]


EXPOSURES

?

OUTCOME

Epidemiology is about identifying associations between exposures and outcomes. To identify any association, exposures and outcomes must first be measured in a quantitative manner. Then rates of occurrence of events are computed. These measures are called "*measures of disease frequency*." Once measured, the association between exposures and outcomes are then evaluated by calculating "*measures of association or effect*." Finally, the impact of removal of an exposure on the outcome is evaluated by computing "*measures of potential impact*." In general, measures of disease frequency are needed to generate measures of association, and both these are needed to get measures of impact. There is some overlap between these measures, and terminology is poorly standardized.


Measures of Disease Frequency

Incidence

Cumulative Incidence (or) Incidence Risk (or) Incidence Proportion

Incidence Density (or) Incidence Rate (or) Hazard Rate (or) Person-time Incidence

Incidence Odds

Prevalence

Point Prevalence

Period Prevalence

Prevalence Odds

Measures of Association (Measures of Effect)

Absolute Difference Measures

[1]Risk Difference (or) Excess Risk (or) Absolute Risk Reduction (or) Attributable Risk

[2]Number Needed to Treat (NNT) (or) Number Needed to Harm (NNH)

Relative Difference or Ratio Measures (Generally called 'relative risks')

[3]Risk Ratio (or) Cumulative Incidence Ratio

[4]Rate Ratio (or) Incidence Density Ratio (or) Relative Rate

[5]Odds Ratio (or) Relative Odds

[6]Prevalence Ratio & Prevalence Odds Ratio

Hazard Ratio

Measures of Potential Impact

Impact of exposure removal on exposed

[7]Attributable Risk (AR) (or) Excess Risk

[8]Attributable Risk Percent (AR%) (or) Etiologic Fraction among the exposed (or) Relative Risk Reduction (or) Attributable Fraction (Exposed)

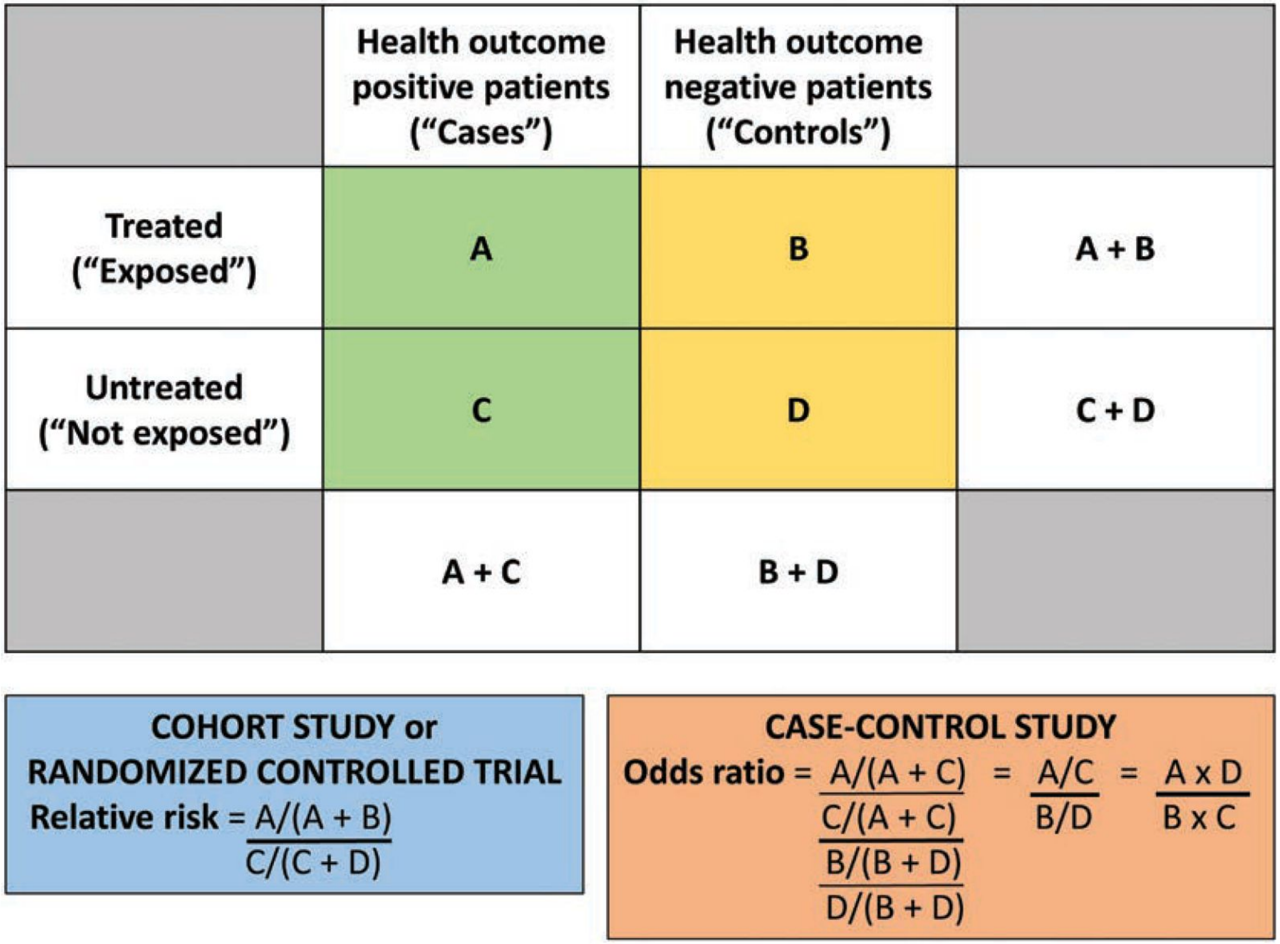
Impact of exposure removal on population

[9]Population Attributable Risk (PAR)

[10]Population Attributable Risk Percent (PAR%) (or) Attributable Fraction (Population)

MEASURES OF EFFECT

The famous epi 2 x 2 table!



	Health outcome positive patients ("Cases")	Health outcome negative patients ("Controls")	
Treated ("Exposed")	A	B	A + B
Untreated ("Not exposed")	C	D	C + D
	A + C	B + D	

COHORT STUDY or RANDOMIZED CONTROLLED TRIAL

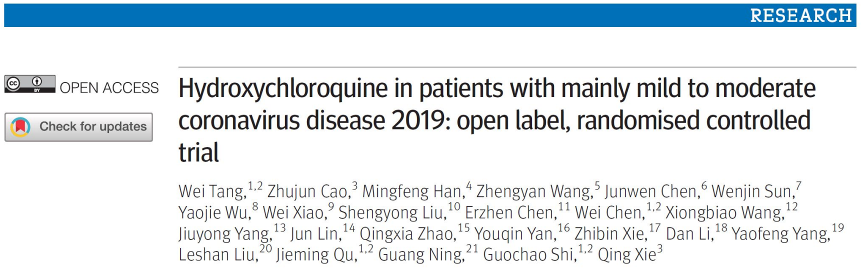
Relative risk = $\frac{A/(A + B)}{C/(C + D)}$

CASE-CONTROL STUDY

Odds ratio = $\frac{\frac{A/(A + C)}{C/(A + C)}}{\frac{B/(B + D)}{D/(B + D)}} = \frac{A/C}{B/D} = \frac{A \times D}{B \times C}$

Vetter TR, 2017

Example: Measures of effect in RCTs



RESEARCH

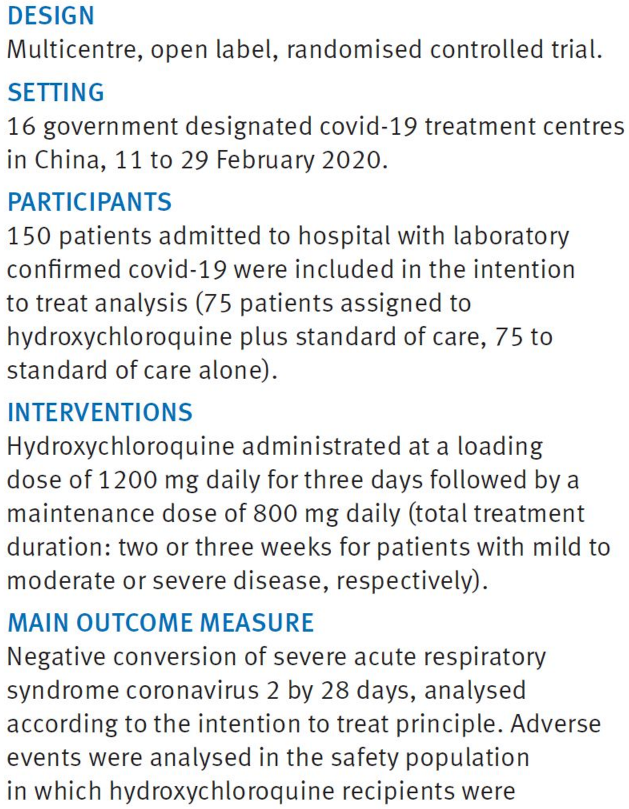
OPEN ACCESS

Check for updates

Hydroxychloroquine in patients with mainly mild to moderate coronavirus disease 2019: open label, randomised controlled trial

Wei Tang,[1,2] Zhujun Cao,[3] Mingfeng Han,[4] Zhengyan Wang,[5] Junwen Chen,[6] Wenjin Sun,[7] Yaojie Wu,[8] Wei Xiao,[9] Shengyong Liu,[10] Erzhen Chen,[11] Wei Chen,[1,2] Xiongbiao Wang,[12] Jiuyong Yang,[13] Jun Lin,[14] Qingxia Zhao,[15] Youqin Yan,[16] Zhibin Xie,[17] Dan Li,[18] Yaofeng Yang,[19] Leshan Liu,[20] Jieming Qu,[1,2] Guang Ning,[21] Guochao Shi,[1,2] Qing Xie[3]

75 patients in HCQ arm
75 patients in standard of care arm



DESIGN

Multicentre, open label, randomised controlled trial.

SETTING

16 government designated covid-19 treatment centres in China, 11 to 29 February 2020.

PARTICIPANTS

150 patients admitted to hospital with laboratory confirmed covid-19 were included in the intention to treat analysis (75 patients assigned to hydroxychloroquine plus standard of care, 75 to standard of care alone).

INTERVENTIONS

Hydroxychloroquine administrated at a loading dose of 1200 mg daily for three days followed by a maintenance dose of 800 mg daily (total treatment duration: two or three weeks for patients with mild to moderate or severe disease, respectively).

MAIN OUTCOME MEASURE

Negative conversion of severe acute respiratory syndrome coronavirus 2 by 28 days, analysed according to the intention to treat principle. Adverse events were analysed in the safety population in which hydroxychloroquine recipients were

Measures of effect

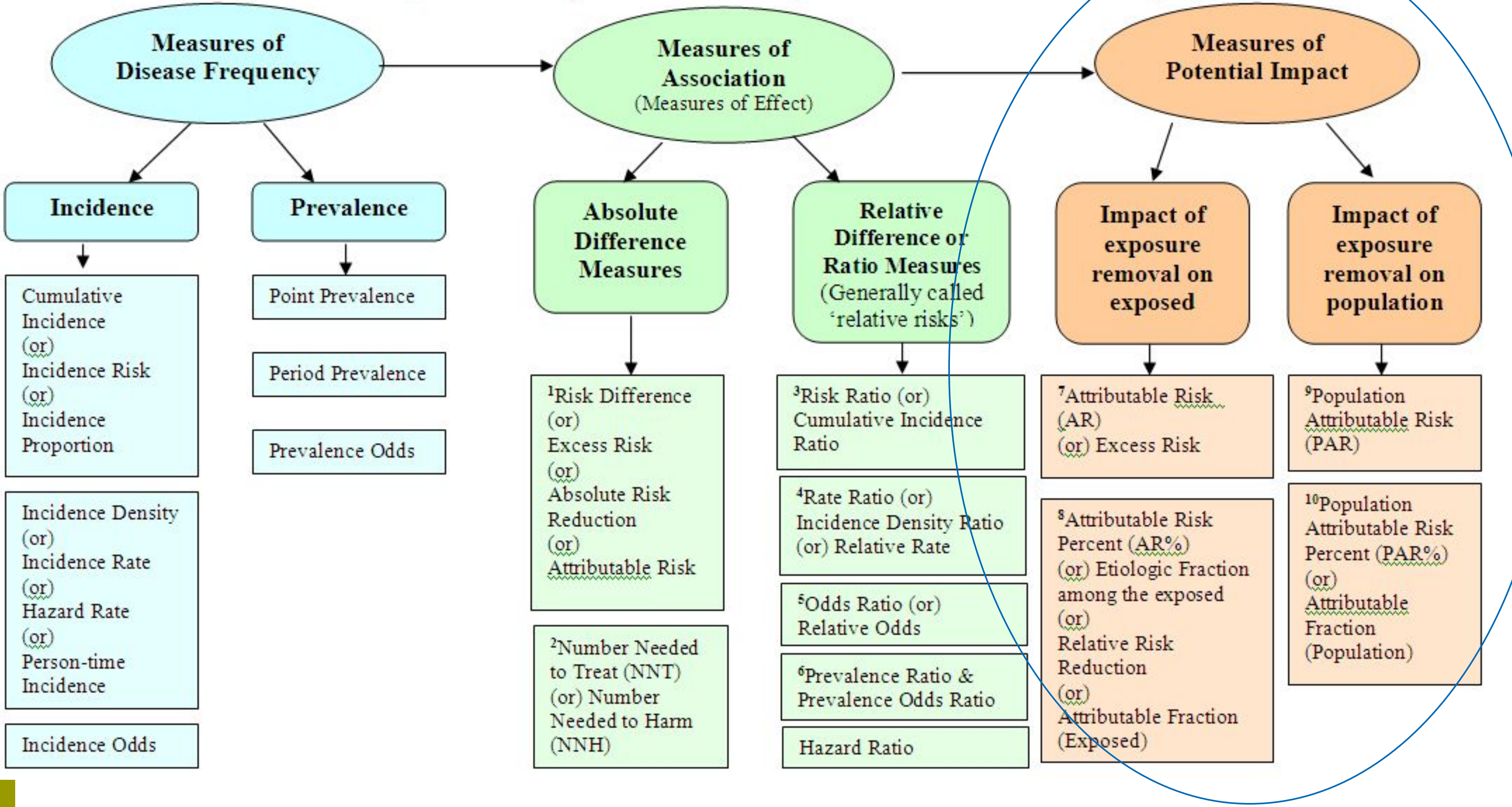
	Covid test becomes negative	Covid test does not become neg	Row total (Margins)
HCQ + standard of care	53	22	75
Standard of care	56	19	75
Column total (Margins)	109	41	150

Cumulative incidence in HCQ group = 70.6%
Cumulative incidence in SOC group = 74.6%
Risk Ratio = 0.94 (95% CI 0.78, 1.15)
Risk difference = -4%
Odds ratio (OR) = 0.81

AN OVERVIEW OF MEASUREMENTS IN EPIDEMIOLOGY [VER 3, 2007]

$$\begin{tikzcd}\boxed{\text{EXPOSURES}} \arrow[r, "?"] & \boxed{\text{OUTCOME}}\end{tikzcd}$$

Epidemiology is about identifying associations between exposures and outcomes. To identify any association, exposures and outcomes must first be measured in a quantitative manner. Then rates of occurrence of events are computed. These measures are called *"measures of disease frequency."* Once measured, the association between exposures and outcomes are then evaluated by calculating *"measures of association or effect."* Finally, the impact of removal of an exposure on the outcome is evaluated by computing *"measures of potential impact."* In general, measures of disease frequency are needed to generate measures of association, and both these are needed to get measures of impact. There is some overlap between these measures, and terminology is poorly standardized.



Measures of Disease Frequency

Incidence

Cumulative Incidence (or) Incidence Risk (or) Incidence Proportion

Incidence Density (or) Incidence Rate (or) Hazard Rate (or) Person-time Incidence

Incidence Odds

Prevalence

Point Prevalence

Period Prevalence

Prevalence Odds

Measures of Association (Measures of Effect)

Absolute Difference Measures

[1]Risk Difference (or) Excess Risk (or) Absolute Risk Reduction (or) Attributable Risk

[2]Number Needed to Treat (NNT) (or) Number Needed to Harm (NNH)

Relative Difference or Ratio Measures (Generally called 'relative risks')

[3]Risk Ratio (or) Cumulative Incidence Ratio

[4]Rate Ratio (or) Incidence Density Ratio (or) Relative Rate

[5]Odds Ratio (or) Relative Odds

[6]Prevalence Ratio & Prevalence Odds Ratio

Hazard Ratio

Measures of Potential Impact

Impact of exposure removal on exposed

[7]Attributable Risk (AR) (or) Excess Risk

[8]Attributable Risk Percent (AR%) (or) Etiologic Fraction among the exposed (or) Relative Risk Reduction (or) Attributable Fraction (Exposed)

Impact of exposure removal on population

[9]Population Attributable Risk (PAR)

[10]Population Attributable Risk Percent (PAR%) (or) Attributable Fraction (Population)

MEASURES OF POTENTIAL IMPACT

Measures of potential impact

- Impact of removing exposure in:
  - Exposed people (e.g. only smokers) = attributable risk (also called risk reduction)
  - All people (entire population – made up of both exposed and unexposed people) = population attributable risk

After accounting for background risk, how much excess risk can removal of exposure bring?

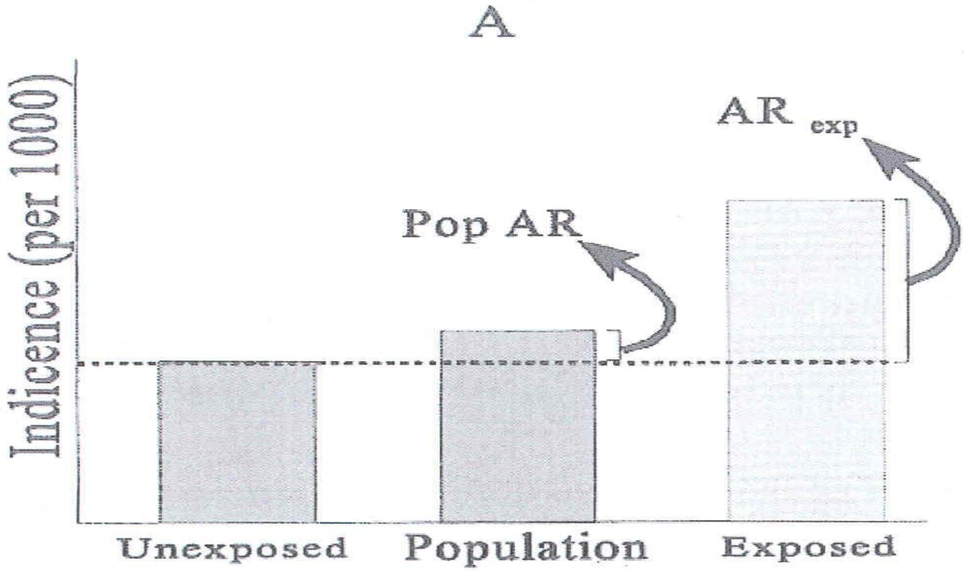
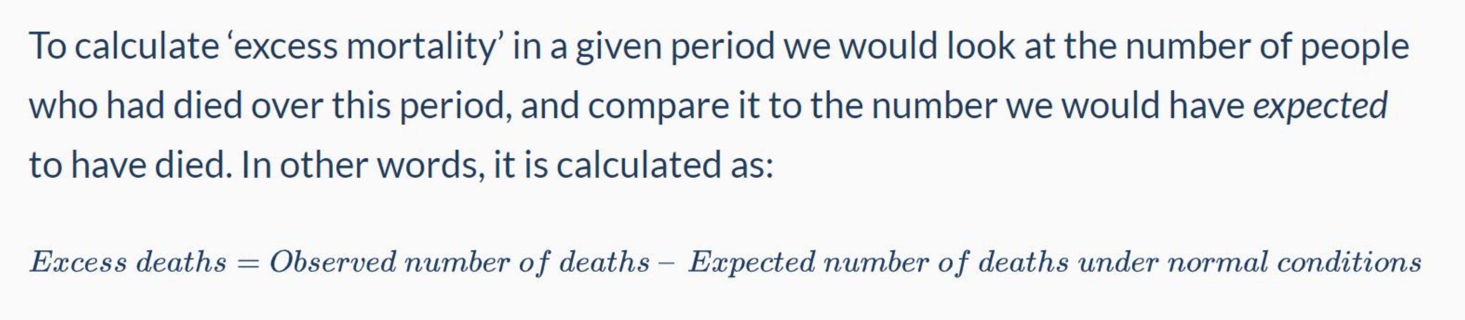


Figure 3-2: Szklo and Nieto, Epidemiology Beyond the Basics, 2000

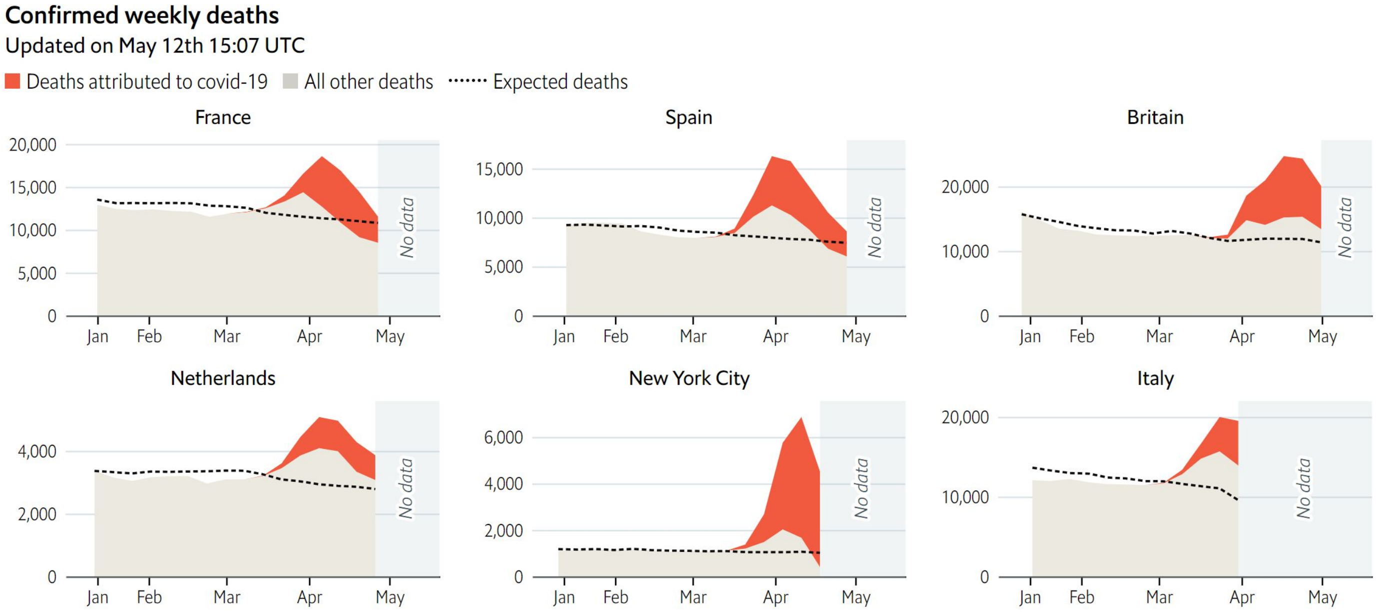
Lots of data show ~80% of lung cancer deaths are attributable to smoking (PAR)

Excess mortality due to Covid-19



To calculate 'excess mortality' in a given period we would look at the number of people who had died over this period, and compare it to the number we would have *expected* to have died. In other words, it is calculated as:

$$Excess\ deaths = Observed\ number\ of\ deaths - Expected\ number\ of\ deaths\ under\ normal\ conditions$$



Confirmed weekly deaths

Updated on May 12th 15:07 UTC

France

Spain

Britain

Netherlands

New York City

Italy

https://www.economist.com/graphic-detail/2020/04/16/tracking-covid-19-excess-deaths-across-countries