UNIVERSITY OF KWAZULU-NATAL
INYUVESI YAKWAZULU-NATALI

Introduction to Epidemiology

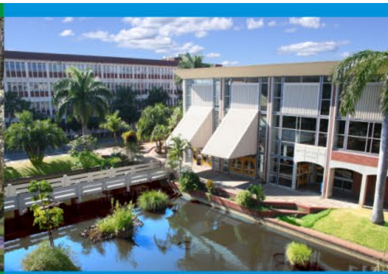
Dr S. Singh
Department of Public Health Medicine
2024


EDGEWOOD CAMPUS


HOWARD COLLEGE CAMPUS


NELSON R MANDELA SCHOOL OF MEDICINE


PIETERMARITZBURG CAMPUS


WESTVILLE CAMPUS

UKZN INSPIRING GREATNESS

OBJECTIVES OF THE LECTURE

- To understand the definition and objectives of Epidemiology
- To learn about the basic concepts in Epidemiology
- Key concepts of Disease Occurrence
  - Counts
  - Prevalence
  - Incidence
  - Mortality rate
  - Morbidity measures
- Causality

Epidemiology *vs* biostatistics *vs* data science

■ Epidemiologists receive substantial training in the science of study design, measurement, and the art of causal inference *including statistical methodology*	■ Biostatisticians are well versed in the theory and application of statistical methodological techniques *including study design and causal inference*	■ Data scientists receive equivalently rigorous training in computational and visualization approaches for high-dimensional data *including statistical methodology*

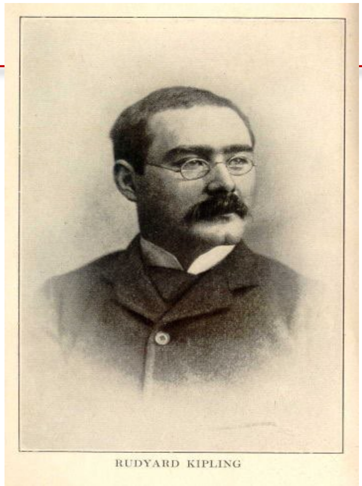
Collaboration and cross-training provides opportunity to share and learn:

- Constructs
- Frameworks
- Theories
- Methods

→ fresh and innovative perspectives for tackling challenging problems in health and health care

Goldstein et al. On the Convergence of Epidemiology, Biostatistics, and Data Science. Harv Data Sci Rev. 2020 PMCID: PMC8734556

EPIDEMIOLOGY



"I keep six honest men,

(they taught me all I know),

Their names are **What** and **Why** and

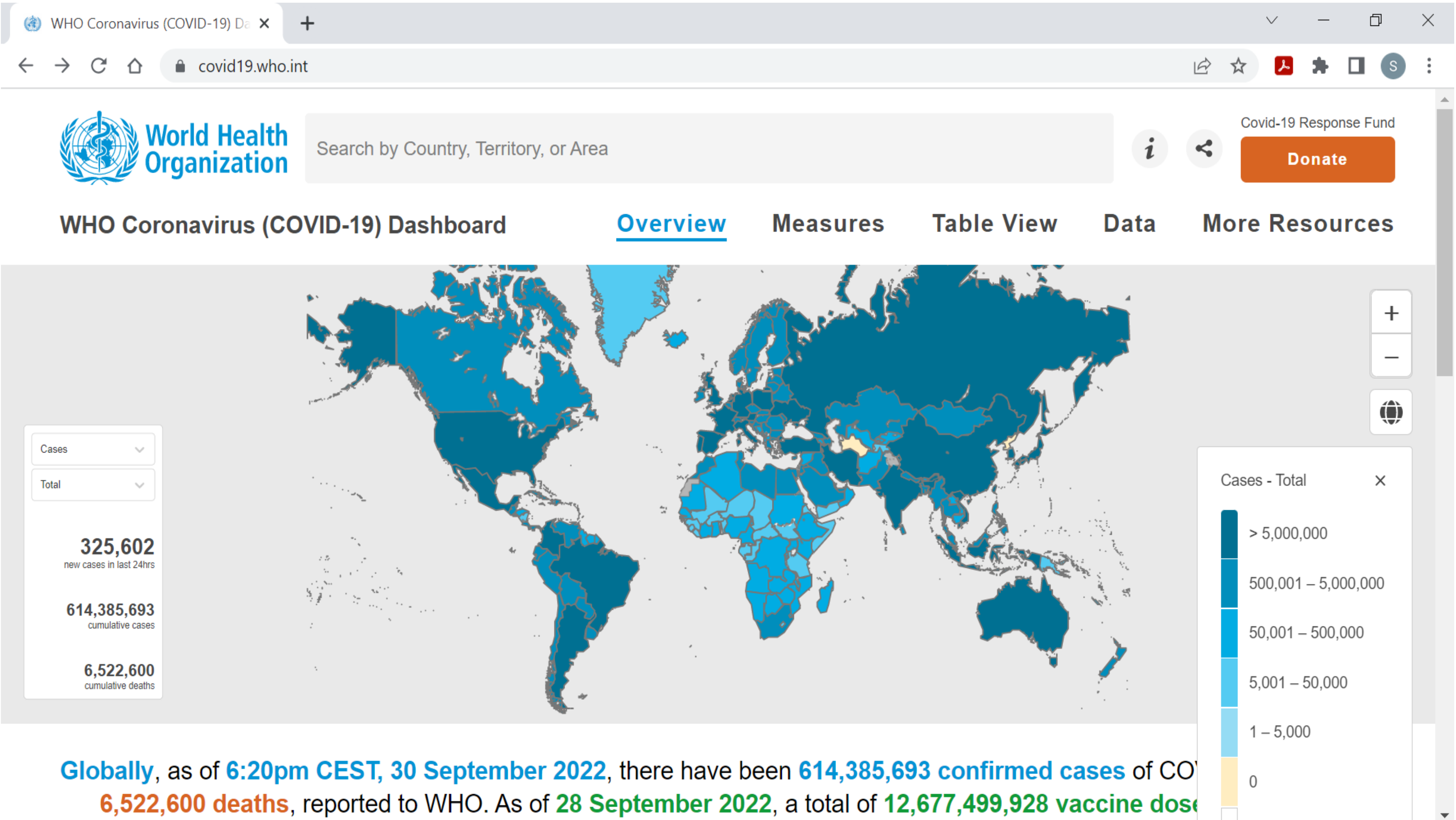
When, and **How** and **Where** and **Who**"

Rudyard Kipling (from: *Just So Stories* 1902)

What is Epidemiology

The study of the **distribution** and **determinants** of health-related state or events in **specified populations** and the **application of study** to the **control** of health problems (*Last, 1995*)

DISTRIBUTION



https://covid19.who.int/

DETERMINANT OF HEALTH ?

- A factor which affects human health
  - Biological
  - Behavioural
  - Social
  - Structural

DETERMINANTS OF HEALTH ????

factors that influence the occurrence of disease and other health-related events

Defining a population of interest

- A population is a collection of individuals, at moments in time, defined by at least one organizing characteristic
- The definition of a population has implications for **analysis**, **interpretation**, and **generalizability** of results from epidemiologic studies

Two axes of population definitions

1. Populations are defined by eligibility criteria
2. Populations are defined by whether individuals move in and out of eligibility; i.e., populations can be dynamic or stationary

.

What could be eligibility criteria?

1. Geographic area and time period of interest
2. Characteristics of persons, events, or exposures for which health-related factors are of interest
3. Factors that promote successful study completion

POPULATION AT RISK

HEALTH CARE WORKERS

TB

- Loss of weight
- Night sweats
- Loss of appetite
- Cough

COVID

- Loss of taste
- Fever
- Cough
- Body pain

OBJECTIVES / USES OF EPIDEMIOLOGY

1. Causation
2. Burden of disease (description of health)
3. Natural history of disease
4. Evaluation of preventative / therapeutic measures
5. Policy and interventions

Main roles of epidemiological research

- 1. Descriptive role (surveillance and observation)– a measure of the changing burden of disease within and between populations– explain local disease patterns – measure of outcome occurrence
- 2. Explanatory and analytic role (hypothesis testing and experiments)– Explains the cause(s) of disease states (even when the biology is not fully understood)- “aetiological epidemiology”– Assesses the effectiveness of interventions and modes of healthcare delivery– measure of exposure effect
- FUNDAMENTAL TO PUBLIC HEALTH

Why is it important to measure health and disease?

- Understanding and assessing the health status of a population and its changes with time
- Need to identify which interventions would have the greatest effect,
- To identify emerging trends
- Anticipate future needs,
- Assist in determining priorities for expenditures,
- To provide information for education to the public,
- Help in setting health research agendas.

- The practice of epidemiology relies on a systematic approach.
- The epidemiologist:
  - Counts cases or health events, and describes them in terms of time, place, and person;
  - Divides the number of cases by an appropriate denominator to calculate rates; and
  - Compares these rates over time or for different groups of people.
- Before counting cases, however, the epidemiologist must decide what a case is.
- This is done by developing a case definition

Defining Health and Disease

- Definitions of health states = “disease present” or “disease absent”
- requires a definition of “normality” and “abnormality.”
- In clinical practice the definition of "a case" generally assumes that, for any disease, people are divided into two discrete classes - the affected and the non-affected

Defining Health and Disease

- *Statistical* - "Normal" may be defined as being within two standard deviations of the age specific mean, as in conventional laboratory practice.
- *Clinical* - Clinical importance may be defined by the level of a variable above which symptoms and complications become more frequent.

Defining Health and Disease

- *Prognostic*- Some clinical findings such as high systolic blood pressure or poor glucose tolerance may be symptomless and yet carry an adverse prognosis.
- "Prognosticate abnormal" is then definable by this level.
- *Operational*- A person may be symptom free yet benefit from treatment or alternatively may have an increased risk that cannot be remedied.

Case Definition/ Diagnostic criteria

- A case definition is a set of standard criteria for classifying whether a person has a particular disease, syndrome, or other health condition.
- The definition of a ‘case’ is critical in planning an epidemiologic investigation.
- Case definitions can include a degree of certainty (e.g. probable or confirmed, etc.) or specify the method to be used in assessing whether or not criteria are met.

Case Definition/ Diagnostic criteria

- Standard case definitions
  - ensure that every case is equivalent, regardless of when or where it occurred, or who identified it.
  - allow for comparison of the number of a disease that occurred in one time or place with the number that occurred at another time or another place.
  - when a difference is observed it is likely to be real rather than the result of variation in how cases are classified.

Case Definition/ Diagnostic criteria

- Measuring disease frequency in populations requires *prior stipulation of which clinical, laboratory, epidemiologic or quantitative criteria indicate the presence of the disease.*
- Problem: may change quite rapidly as knowledge increases or diagnostic techniques improve; they also often change according to the context in which they are being used.

Case Definition

- A case definition consists of clinical criteria and, sometimes, limitations on time, place, and person.
- The clinical criteria usually include confirmatory laboratory tests, if available, or combinations of symptoms (subjective complaints), signs (objective physical findings), and other findings.

Case Definitions

- **Clinically compatible case**:
  - a clinical syndrome generally compatible with the disease,
  - A general clinical impression that this is a case of disease.
- **Confirmed case**: The case meets established criteria.
- **Laboratory-confirmed case**: a case that is confirmed by one or more of the laboratory methods listed in the case definition under laboratory criteria for diagnosis.

Case definition

- **Epidemiologically linked case**: a case in which
  - a) the patient has had contact with one or more persons who either have/had the disease or have been exposed to a point source of infection (i.e., a single source of infection, such as an event leading to a food borne-disease outbreak, to which all confirmed case-patients were exposed)
  - b) transmission of the agent by the usual modes of transmission is plausible.

Case Definition

- **Possible/Suspected**
- A case with the clinical criteria as described in the case definition without epidemiological or laboratory evidence of the disease in question.
- The definition of a possible case has high sensitivity and low specificity.
- It allows for detection of most cases but some false positives cases will be included into this category.

Case Definition

- **Probable**
  - A probable case is usually a case with clinical criteria and an epidemiological link as described in the case definition.
  - Laboratory tests for probable cases are specified only for some diseases.
- **Confirmed**
  - A confirmed case should be laboratory confirmed and may fulfil the clinical criteria or not as described in the case definition.
  - The definition of a confirmed case is highly specific and less sensitive; therefore most of the collected cases will be true cases

Rule of Thumb

- To investigate a highly infectious, transmissible or serious and deadly disease, casting a broad net will capture all suspected and probable cases.
- On the other hand, if declaring an individual to be a 'case' is likely to result in imposing severe restrictions, such as closing schools or removing a product from the market, the case definition should be more stringent.

Measles Case Definition

- **Suspected:**
  - any febrile illness that is accompanied by rash and that does not meet the criteria for probable or confirmed measles or any other illness.

Measles case definition

- **Probable**: In the absence of a more likely diagnosis, an illness characterized by
  - Generalized rash lasting ≥3 days; and
  - Temperature ≥101°F or 38.3°C; and
  - Cough, coryza, or conjunctivitis; and
- No epidemiologic linkage to a confirmed case of measles; and
- Non contributory or no serologic or virologic testing.

Measles case definition

- **Confirmed**: Laboratory confirmation by any of the following
  - Positive serologic test for measles immunoglobulin M antibody;
  - Significant rise in measles antibody level by any standard serologic assay;
  - Isolation of measles virus from a clinical specimen; or
  - Detection of measles-virus specific nucleic acid by polymerase chain reaction
  - Note: A laboratory-confirmed case does not have to have generalized rash lasting ≥3 days; temperature ≥101°F or 38.3°C; cough, coryza, or conjunctivitis.
  - OR
- An illness characterized by
  - Generalized rash lasting ≥3 days; and
  - Temperature ≥101°F or 38.3°C; and
  - Cough, coryza, or conjunctivitis; and
  - Epidemiologic linkage to a confirmed case of measles.

Measuring health and disease

- The relative importance (burden) of different diseases in a population depends on:
  - their frequency (incidence or prevalence),
  - severity (the mortality and extent of serious morbidity),
  - consequences (health, social, economic),
  - and the type of people affected (gender, age).

Individual level measure

- **Count**:
  - measure of disease frequency
  - number of cases or other health outcomes being studied.

EPIDEMIOLOGY

Scientific basis of public health

↓

Measure of mortality and morbidity

Risk (Cumulative incidence) / Prevalence / Incidence

Rates | **Ratio** | **Proportion**

Ratio,

- Ratio is one number divided by another number (numerator doesn't have to be included in denominator—and vice versa)
- **Example: males/females**

Proportion

- Proportion is a ratio in which the numerator is included in the denominator
  - It has no dimension because the unit of the numerator cancels out the unit of the denominator

Example: males/total births

Rates

- Rate is a ratio
  - The numerator is the number of events
    - The unit is event (i.e., number of cases)
  - The denominator is the sum of follow-up time contributed by the people at risk of the event
    - The unit is time or, more accurately, person-time to account for duration of time of follow-up
  - the number at risk of the event during that time period
  - A rate may or may not be a proportion
- **Example: (deaths in 1999/population in 1999) x 1,000**

- Rates tell us how fast the disease is occurring in a population.
- Proportions tell us what fraction of the population is affected.

RISK VERSUS 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

– up 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

Incidence and prevalence

- Two measures overcome many of the limitations of a simple count of cases - incidence and prevalence
- **Prevalence** tells us about the proportion of cases among the total population at any given time
- **Incidence** tells us the probability of a new onset of disease among those at risk for developing the illness

Measures of disease occurrence/frequency

- There are three commonly used measures of disease occurrence:
  - cumulative incidence,
  - Incidence density,
  - and prevalence

Importance of measuring disease frequency

- Prior to any clinical trial:
  How frequently does the disease occur?
- Provide big picture information about a disease, framing public health questions and guiding resource allocation.
- Describe the *absolute* risk of a disease.
- Can be categorized, or *stratified* to gain insight into the pathogenesis (mechanism) of disease

Prevalence

- **Prevalence** is the proportion of population with the disease at a specific point or over a specified time
- Prevalence (%) =

$$\frac{\text{number of people with disease}}{\text{number of people in the population}} \times 100\%$$

Prevalence

- Prevalence is a proportion
- It measures the extent (amount) of the event (disease) in the population in a specified time
- The numerator includes both new and existing cases of disease
- Time is a descriptor of the measure but is not a part of the denominator (does not use person-time)
- Sometimes, the term **prevalence rate** is used even though it is a proportion Example: prevalence rate of HIV/AIDS in Botswana in 2003 = 37.3%

Prevalence

- Prevalence doesn’t really tell us when the disease developed or how long someone’s had it.
- The numerator is a mix of people with different durations of disease.
- Can’t measure risk

Prevalence

- Prevalence is really a good measure of the burden of disease in a population if it is relatively stable, and it's a chronic condition.
- Prevalence is a good measure of disease for us when we are talking about allocating health resources.
- If we know that 20% of our community has obesity or is obese, then we can allocate health resources accordingly

Point and Period Prevalence

- Prevalence is not a rate
- *Point prevalence* measures the frequency of all current events (old and new) at a given instant in time
- *Period prevalence* measures the frequency of all current events (old and new) for a prescribed period of time

Point and Period Prevalence

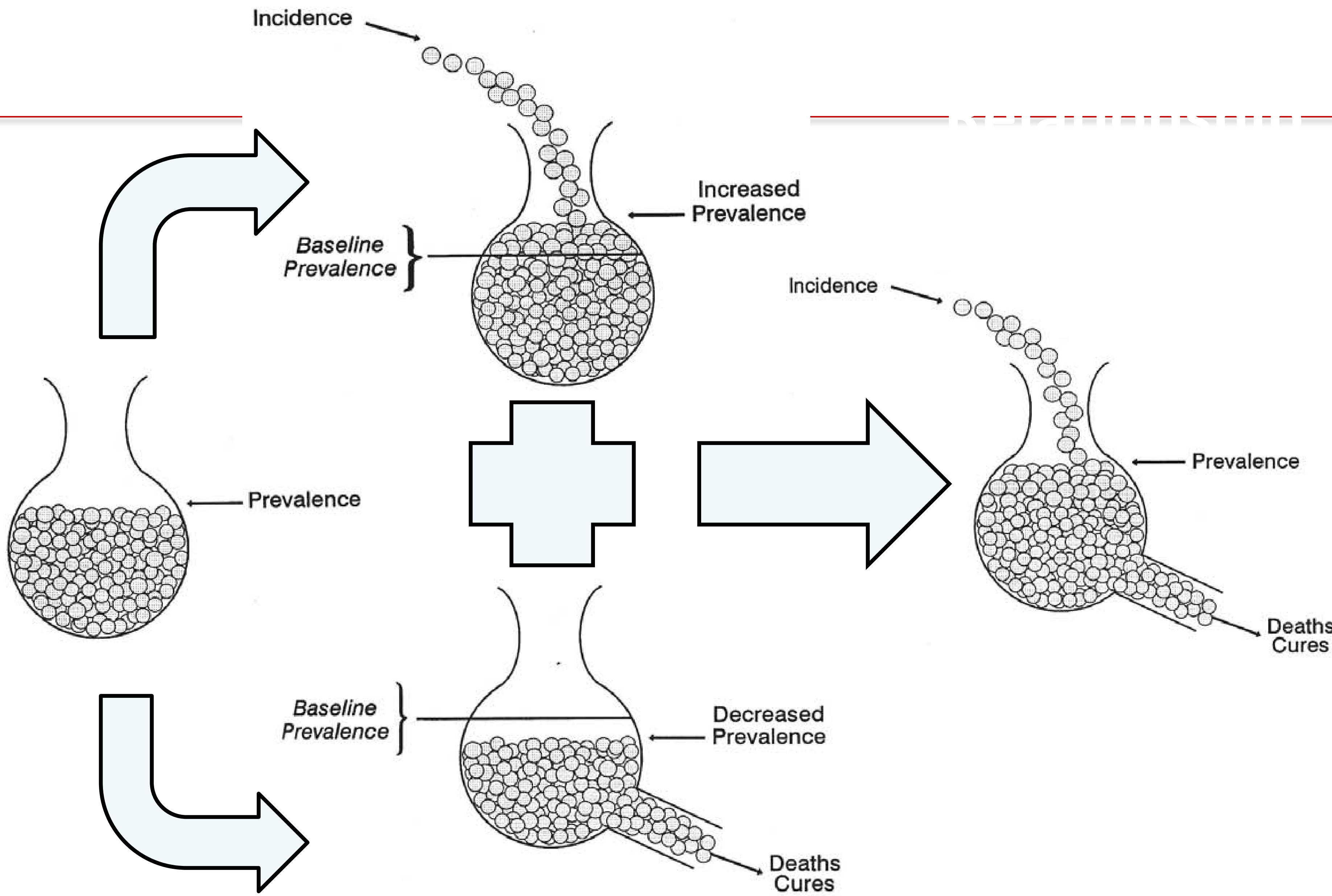
- Examples of point and period prevalence and cumulative incidence in interview studies of asthma
  - "Do you currently have asthma?" Point prevalence
  - "Have you had asthma during the last *n* years?" Period prevalence
  - "Have you ever had asthma?" Cumulative or life-time incidence

Incidence & Prevalence are related

Prevalence = Incidence X Duration

– Higher incidence results in higher prevalence

– Longer duration results in higher prevalence


Incidence
Increased Prevalence
Baseline Prevalence
Prevalence
Incidence
Prevalence
Deaths
Cures
Baseline Prevalence
Decreased Prevalence
Deaths
Cures

Factors affected prevalence rate

Prolongation of life of patients without cure

-Increase in new cases (increase in incidence)

-In-migration of cases

-Out-migration of susceptible people

-Improved diagnosis facilities (better reporting)onger duration of the disease

-Prolongation of life of patients without cure

-Increase in new cases (increase in incidence)

-In-migration of cases

-Out-migration of susceptible people

-Improved diagnosis facilities (better reporting)

High case-fatality rate from disease

-Decrease in new cases (decrease in incidence)

-In-migration of healthy people

-Out-migration of cases

-Improved cure rate of cases

Incidence

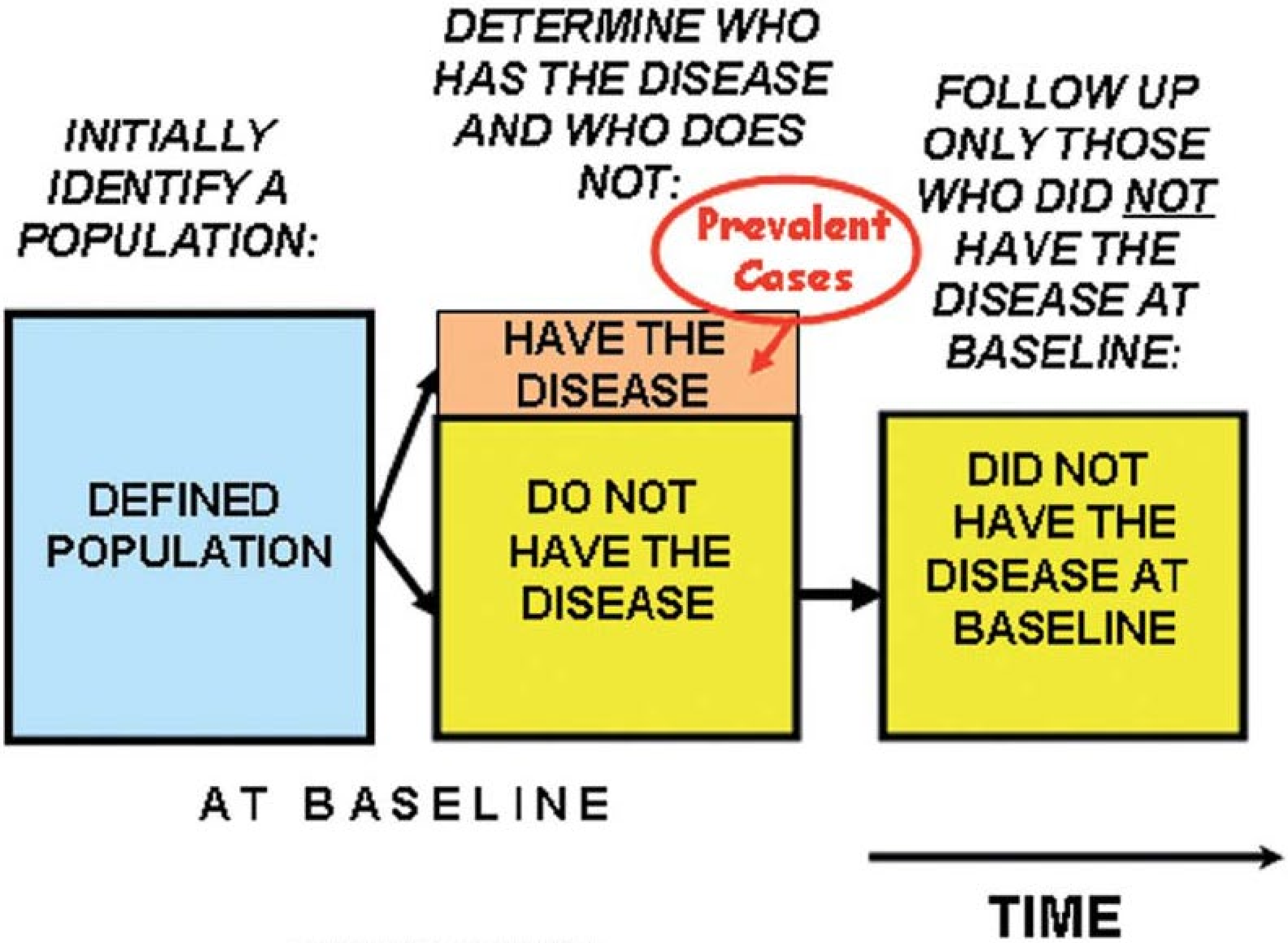
- Quantifies development of disease
- is the number of new cases of a disease occurring in an at-risk population during a defined time interval
  - Example: 80% of all incident TB cases were found in 22 countries
- provide a measure of risk= probability that an individual will develop a disease over a specified time.

Use of incidence rates

- To monitor changes over time in disease.
- Make comparisons between groups
  - age-specific rates
  - gender or sex-specific rates.
- Incidence can be used for goal setting.

Incidence

- Cumulative Incidence Rate / Incidence Proportion/Attack Rate
- **Incidence density rate** (also known as *incidence rate; person-time rate, force of morbidity/mortality, hazard rate, disease intensity*)



INITIALLY IDENTIFY A POPULATION:
DETERMINE WHO HAS THE DISEASE AND WHO DOES NOT:
FOLLOW UP ONLY THOSE WHO DID NOT HAVE THE DISEASE AT BASELINE:
DEFINED POPULATION
HAVE THE DISEASE
Prevalent Cases
DO NOT HAVE THE DISEASE
DID NOT HAVE THE DISEASE AT BASELINE
AT BASELINE
TIME

Incidence Proportion

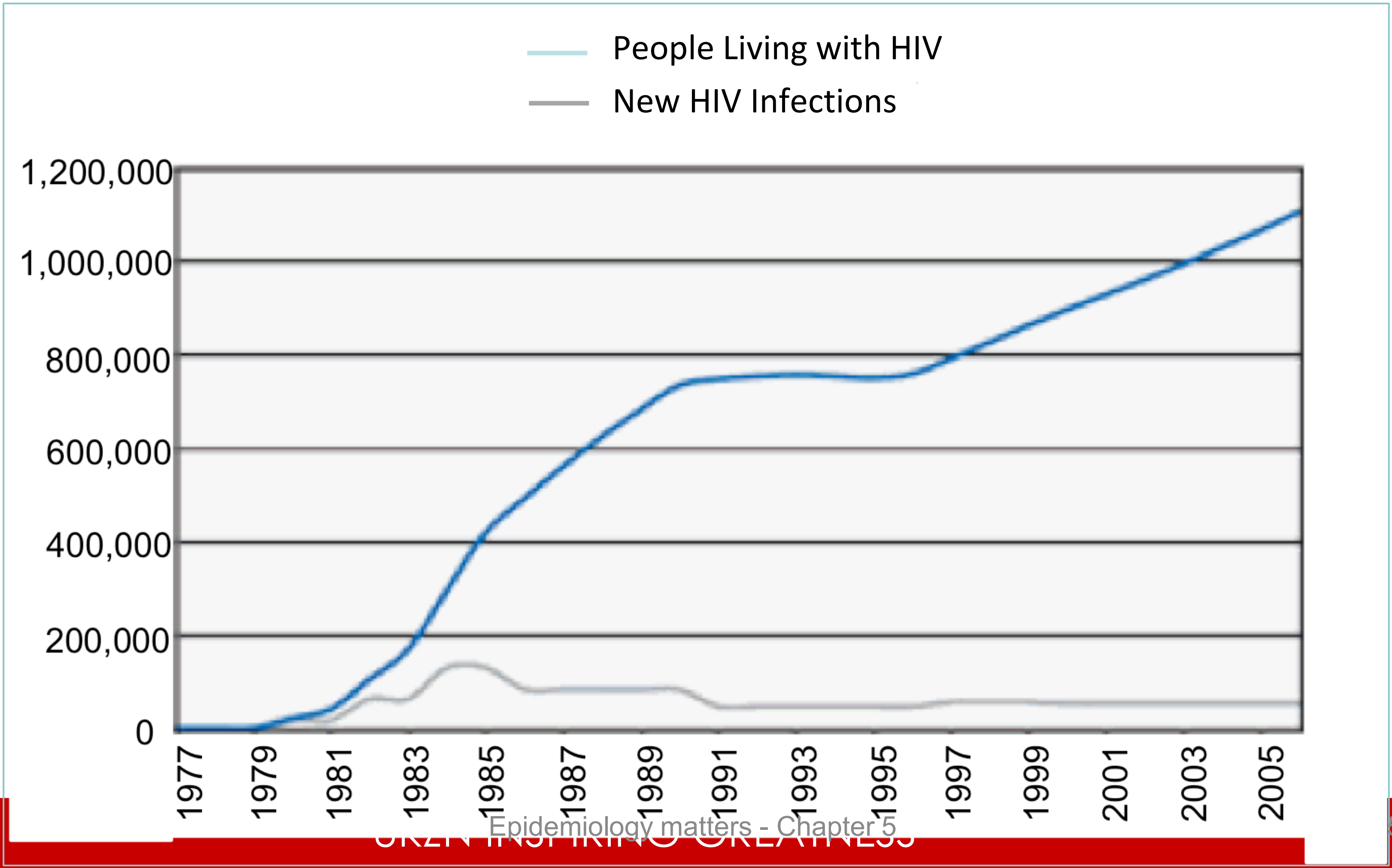
$$\text{Incidence per 1,000} = \frac{\text{Number of NEW cases of a disease occurring in the population during a specified period of time}}{\text{Number of persons at risk of developing the disease during that period of time}} \times 1{,}000$$

- a definition of the onset of the event,
- a defined population,
- and a particular period of time.

The critical point is *new cases* of disease—the disease must develop in a person who did not have the disease previously.

Examples of the relation between incidence and prevalence

Impact of a new treatment that prolongs life with the disease but does not cure it



Attack Rate

- Same as incidence proportion
  - It implies rate but is not actually a rate
- Similar to an incidence rate,
- used often in outbreak settings,
- nature of the disease is such that the population is being observed for a very short period of time.
- when you know there has been a specific exposure.
- Attack rates can be used to compare the risk of disease in groups with different exposures.

Cumulative Incidence/Risk

- Assumes a fixed or closed cohort (no exits allowed)
- For brief specified periods of time, e.g. an outbreak, commonly called an Attack "Rate"
- In reality, attrition is a huge problem (losses to follow-up, deaths, competing risks)
- Formula does not reflect continually changing population size for dynamic cohorts (open populations).
- Does not allow subjects to be followed for different time periods.
- In real life, one has to deal with losses, competing risks, attrition, dynamic cohorts, and differential follow-up time!!
- So, rate becomes more relevant

Incidence Density/Rate

$$\text{Incidence} \quad \frac{\text{No. of new cases}}{\text{Persons-time}} \Rightarrow \text{Person-year}$$

($\sum$person-time at risk)

incidence density

force of morbidity

force of mortality in reference to deaths

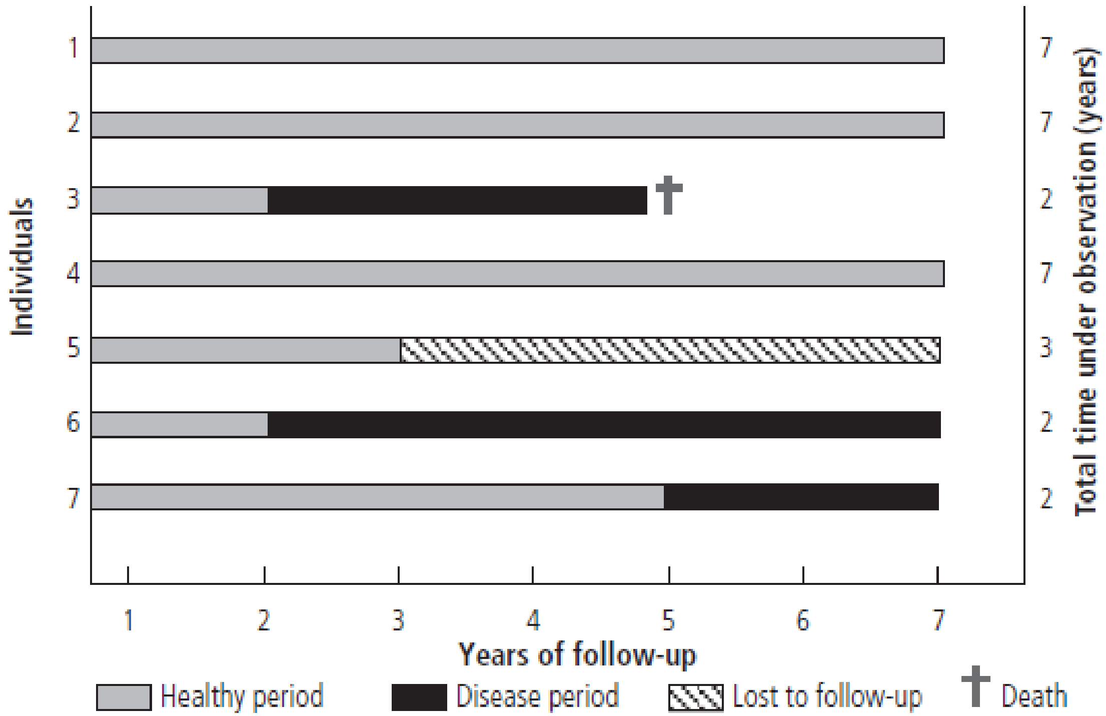
hazard rate

disease intensity

Incidence Density Rate

- the number of persons who newly experience the outcome during a specified period of time divided by the sum of the time that each member of the population is at-risk
- This is also presented with upper and lower bounds.
- The numerator is the same as the numerator of incidence proportion
- The denominator accumulates time at risk of the event
- It is not just the number of people at risk

Calculating Incidence Density



Calculating Incidence

- **Cumulative Incidence**: is the number of new events in the population at risk (3) divided by the number of people in the same population free of the disease at the beginning of the period (7), i.e. 43 cases per 100 persons

Incidence density

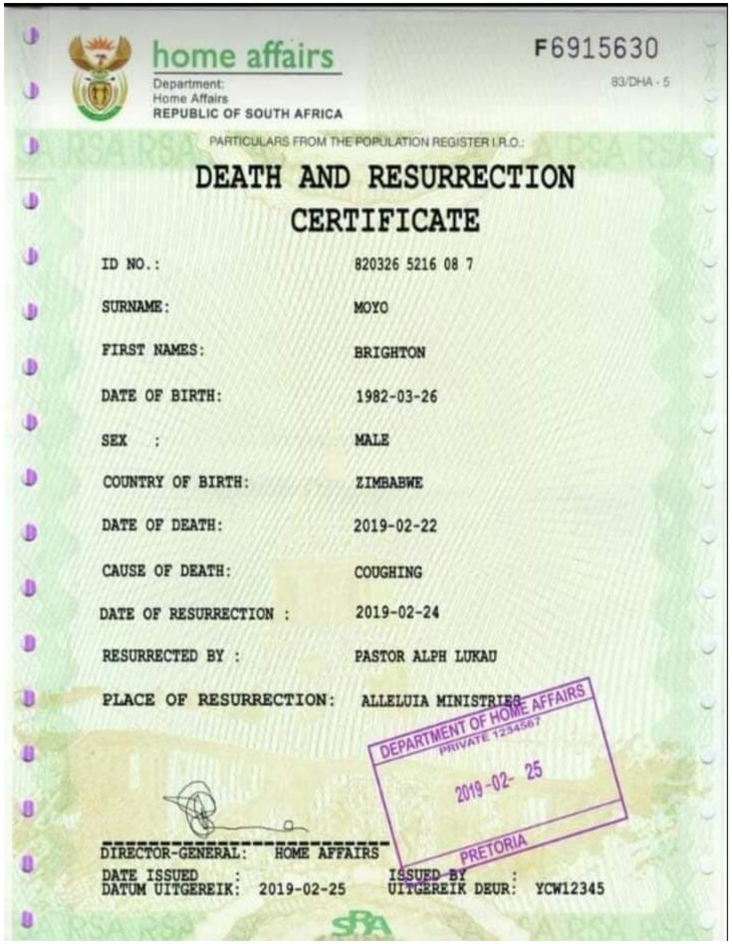
- the incidence of the disease during the seven-year period is the number of
- new events (3) divided by the sum of the lengths of time at risk of getting the
- disease for the population (33 person-years), i.e. 9.1 cases per 100 person years;

	Incidence	Prevalence
Numerator	Number of **new** cases of disease during a specified period of time	Number of **existing** cases of disease at a given point of time
Denominator	Population at risk	Population at risk
Focus	Whether the event is a new case Time of onset of the disease	Presence or absence of a disease Time period is arbitrary; rather a "snapshot" in time
Uses	Expresses the risk of becoming ill The main measure of acute diseases or conditions, but also used for chronic diseases More useful for studies of causation	Estimates the probability of the population being ill at the period of time being studied. Useful in the study of the burden of chronic diseases and implication for health services

Note: If incident cases are not resolved, but continue over time, then they become existing (prevalent) cases. In this sense, prevalence = incidence x duration.

Mortality

- Death is the foundation of all vital statistics in most countries, laws require registration of deaths
- Death certificate completed
- Registered at Home Affairs
- Analysed by STATSSA



Rates

3 main types

- Crude rate
- Specific rate
- Standardized (adjusted) rate

Mortality Rate

- the number of all deaths in one year divided by the number of people in that population at midyear.
- For vital statistics purpose, the midpoint (midyear) population is used with the assumption that:
- The denominator thus becomes person-year, and mortality rate can be considered as a “rate” and not a “proportion”
  - Even though the number of persons at midyear is used in the calculation

Mortality Rate

$$\text{Annual mortality rate from all causes (per 1,000 population)} = \frac{\text{20 deaths from all causes in one year}}{\text{12,000 persons in the population at midyear}} \times 1{,}000$$

$$= 1.7 \text{ per } 1{,}000$$

Formula for mortality rate

$$\text{Annual mortality rate from all causes (per 1,000 population)} = \frac{\text{Total number of deaths from all causes in one year}}{\text{Number of persons in the population at midyear}} \times 1{,}000$$

Does NOT account for differences of age, sex, etc. in any aspect of death

Info needed:

- total deaths
- total population
- a given period of time

Mortality Rate

$$\text{Annual mortality rate from all causes (per 1,000 population)} = \frac{\text{20 deaths from all causes in one year}}{\text{12,000 persons in the population at midyear}} \times 1{,}000$$

$$= 1.7 \text{ per } 1{,}000$$

Age Specific mortality rate

$$\text{Annual mortality rate from all causes for children under age 10 (per 1,000 population)} = \frac{\text{Total number of deaths from all causes in one year in children under age 10}}{\text{Number of children in the population under age 10 at midyear}} \times 1{,}000$$

Cause Specific mortality rate

$$\text{Annual mortality rate from lung cancer (per 1,000 population)} = \frac{\text{Total number of deaths from lung cancer in one year}}{\text{Number of persons in the population at midyear}} \times 1{,}000$$

Age and Cause Specific mortality rate

$$\text{Annual mortality rate from leukemia for children under age 10 (per 1,000 population)} = \frac{\text{Total number of deaths from leukemia in one year in children under age 10}}{\text{Number of children in the population under age 10 at midyear}} \times 1{,}000$$

Case Fatality Rate

$$\textbf{Case fatality rate (\%)} = \frac{\text{Number of individuals dying during a specified period of time after disease onset or diagnosis}}{\text{Number of individuals with the specified disease}} \times 100$$

Measure	Numerator	Denominator	10^n
Crude death rate	Total number of deaths during a given time interval	Mid-interval population	1,000 or 100,000
Cause-specific death rate	Number of deaths assigned to a specific cause during a given time interval	Mid-interval population	100,000
Neonatal mortality rate	Number of deaths among children < 28 days of age during a given time interval	Number of live births during the same time interval	1,000
Infant mortality rate	Number of deaths among children < 1 year of age during a given time interval	Number of live births during the same time interval	1,000
Maternal mortality rate	Number of deaths assigned to pregnancy-related causes during a given time interval	Number of live births during the same time interval	100,000

Morbidity and Disability

- The problem with mortality-based indicators is that they “note the dead and ignore the living” (Kaplan, 1990).
- Measurements of morbidity are problematic
- there is no clearly defined endpoint, such as death
- several components of morbidity and disability need to be assessed: duration, severity, and consequences

Measures of population health

- years of potential life lost (YLL) based on the years of life lost through premature death (before an arbitrarily determined age);
- healthy life expectancy (HALE);
- disability-free life expectancy (DFLE);
- quality-adjusted life years (QALYs);
- disability-adjusted life years (DALYs).

Healthy Life year

- The *healthy life year* (HeaLY) is a composite measure that combines the amount of healthy life lost due to morbidity with that lost due to death

Years of potential life lost (YPLL)

- **Years of potential life lost** measures the impact of mortality on society
- It is calculated by summing the years that individuals would have lived had they experienced normal life expectancy and had not died from the particular disease
- Often, age 65 (or 75) is used in the calculation
- For example, a person who died at age 30 from heart disease will contribute 65–30=35 YPLL
- YPLL is weighted more by premature deaths, while **crude mortality** is weighted by the larger number of deaths in older people

$$\text{YPLL rate per100,000} = \frac{\text{sum}\,(65 - \text{age at death})}{\text{number of people 65 and younger}} \times 100{,}000$$

Disability adjusted life years

- DALYs for a disease or health condition are calculated as the sum of the Years of Life Lost (YLL) due to premature mortality in the population and the Years Lost due to Disability (YLD) for incident cases of the health condition:
- years of lost life (YLL) – calculated from the number of deaths at each age
- multiplied by the expected remaining years of life according to a global standard life expectancy
- years lost to disability (YLD), where the number of incident cases due to injury and illness is multiplied by the average duration of the disease and a weighting factor reflecting the severity of the disease on a scale from 0 (perfect health) to 1 (dead).

CHALLENGES

- Defining the “at risk” population
- Defining a case
- Recording movement within a population under study
- Specifying time periods

Causal Effects

- Key premise: Diseases have causes
- Some causes can be partly or fully eliminated→ prevent some cases of disease
- Problem: Most epidemiological studies are observational and measure associations

Association ≠ Causation

Epidemiological Triad

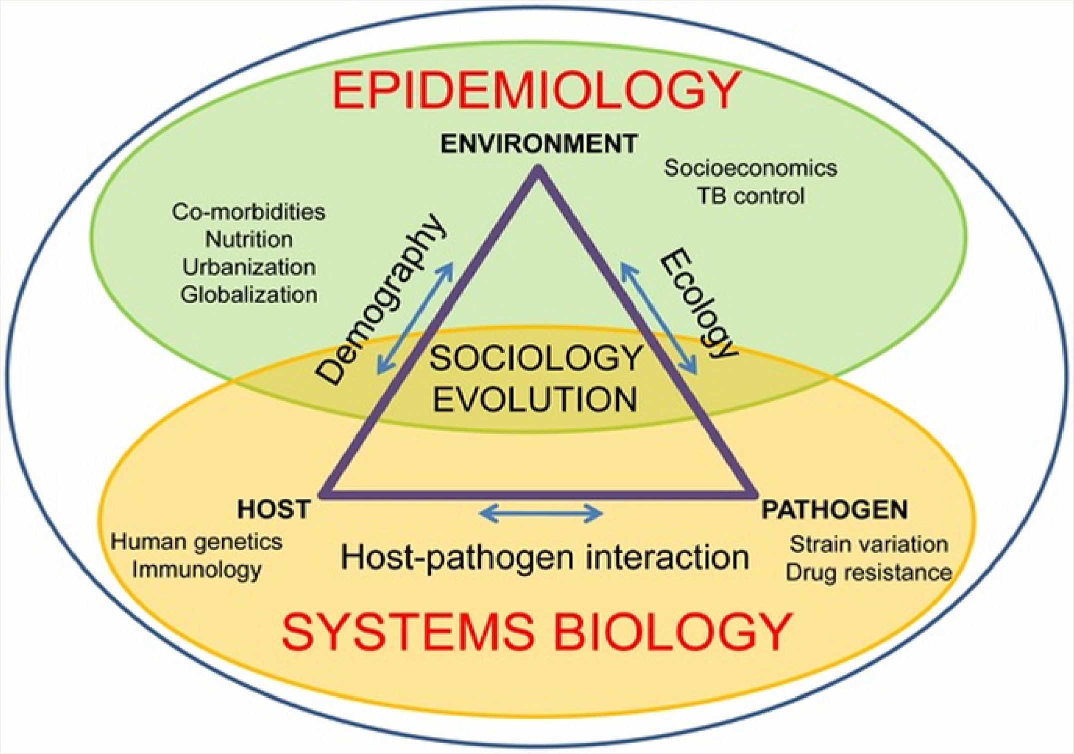
- The epidemiologic triangle or triad is the traditional model of infectious disease causation.
- It has three components: an external agent, a susceptible host, and an environment that brings the host and agent together.
- In this model, the environment influences the agent, the host, and the route of transmission of the agent from a source to the host

Epidemiological triad

"Systems epidemiology approach"

- **Pathogen (exposure)**
- Host
- Environment

- Epidemiology addresses the burden of the disease and the social, economic, and ecological causes of its frequency and distribution
- Systems biology integrates approaches that address the host, the pathogen, and interactions between the two



Comas I, Gagneux S (2009) The Past and Future of Tuberculosis Research. PLOS Pathogens 5(10): e1000600. https://doi.org/10.1371/journal.ppat.1000600

Agent factors

- **Agent** originally referred to an infectious microorganism—virus, bacterium, parasite, or other microbe.
- must be present for disease to occur.
- the concept of agent has been broadened to include chemical and physical causes of disease.
- These include chemical contaminants, such as the l-tryptophan contaminant responsible for eosinophilia myalgia syndrome, and physical forces, such as repetitive mechanical forces associated with carpal tunnel syndrome.

Host factors

- Host factors are intrinsic factors that influence an individual's exposure, susceptibility, or response to a causative agent.
- Age, race, sex, socioeconomic status, and behaviors (smoking, drug abuse, lifestyle, sexual practices and contraception, eating habits)
- Age, genetic composition, nutritional and immunologic status, anatomic structure, presence of disease or medications, and psychological makeup are some of the host factors which affect a person's susceptibility and response to an agent.

Environmental factors

- Environmental factors are extrinsic factors which affect the agent and the opportunity for exposure.
- Generally, environmental factors include physical factors such as geology, climate, and physical surroundings (e.g., a nursing home, hospital); biologic factors such as insects that transmit the agent; and socioeconomic factors such as crowding, sanitation, and the availability of health services

Sufficient and Necessary cause

- Two more specific notions of cause:
- – The cause is sufficient to bring about the effect
- – The cause is necessary to bring about the effect
- Neither works perfectly, but each is suggestive of important features of causation and help us understand how to test causal claims

Necessary cause

- If disease does not develop without the factor being present, then we term the causative factor "**necessary**
- Such a view implies: (a) that the cause is defined univocally and is easily identifiable; (b) that the disease can be also defined univocally, i.e., it is not a complex and variable constellation of symptoms.

EXAMPLE OF NECESSARY CAUSE

If outcomes are defining in terms of causes, the cause is necessary by definition. For example, the tubercle bacillus is necessary for tuberculosis by the definition of tuberculosis. Etiologic (as contrasted to manifestational) classification of diseases often produce necessary causes. Hepatitis B once looked to be a necessary cause of hepatocellular carcinoma. But now we see that Hepatitis C may produce it too.

NECESSARY CAUSE

(e.g. the tubercle bacillus and tuberculosis)

	HAS DISEASE	FREE OF DISEASE
HAS EXPOSURE	YES	YES
DOES NOT HAVE EXPOSURE	NO	YES

Necessary cause

- smallpox is a clear-cut disease entity, easy to define and diagnose; it is due to a single necessary virus (no smallpox develops in the absence of the specific virus);
- and clear proof of the causal link has come from the disappearance of smallpox after large scale vaccination.

Necessary cause

- infectious agents such as viruses, bacteria or parasites;
- environmental agents such as sun-rays or allergens (e.g. pollen, dust-mites);
- industrial agents such as chemicals (e.g. nicotine) or radiation (e.g. mobile phones);
- genetic factors such as chromosomal abnormalities;
- physical factors such as violence or car accidents;
- psychological factors such as stress or abuse.

The Sufficient-Component Cause Model

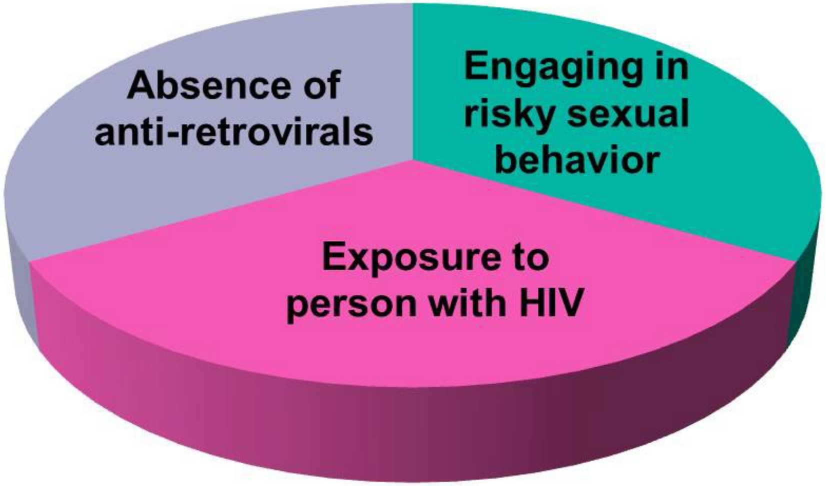
- If disease does not develop without the factor being present, then we term the causative factor "**necessary**."
- If the disease always results from the factor, then we term the causative factor "**sufficient**

Sufficient cause

- A "sufficient cause," which means a complete causal mechanism, can be defined as a set of minimal conditions and events that inevitably produce disease; "minimal" implies that all of the conditions or events are necessary to that occurrence

Sufficient cause

- a complete causal mechanism" that "inevitably produces disease."
- Consequently, a "sufficient cause" is not a single factor, but a minimum set of factors and circumstances that, if present in a given individual, will produce the disease.



Component cause

- Component causes: together they constitute a sufficient cause for the outcome in question. In CDs, this may include the biological agent as well as environmental conditions (e.g. TB, measles, ARF/RHD).
- In NCDs, this may include a whole range of genetic, environmental as well as personal / psychosocial / behavioral characteristics (e.g. diabetes, cancers, IHD)

THANK YOU

Acknowledgments of Slides:

- Prof S. Naidoo
- Prof O. Mahomed

REFERENCES

1. Principles of Epidemiology in Public Health Practice – Centers for Disease Control
2. EPI Matters - Katherine M. Keyes and Sandro Galea