

Introduction to Epidemiology

Midterm Study Notes // Lectures 1 to 4, and 6

LECTURE ONE

EPIDEMIOLOGY

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

- Draws on methods from other scientific fields – biostatistics, sociology, informatics, economics, and behavioural sciences
- Directs practical and appropriate public health action

Areas of Focus

- **Distribution**
 - o Frequency and pattern of health events in a population
- **Determinants**
 - o Any factor, weather event, characteristic, or other definable entity, that brings about a change in a health condition or other defined characteristic
- **Specified Populations**
 - o Clinician = patient; Epidemiologist = community
- **Health-related states of events**
 - o Communicable disease, chronic disease, injuries, birth defects, occupational health, environmental health, behaviours, genetic markers

Investigate: environmental exposures, infectious diseases, injuries, non-infectious disease, natural disasters, terrorism, behaviours and policies

DEFINITIONS

Endemic	the constant presence of a disease or infectious agent within a given geographic area or population group
Epidemic	The occurrence of more case of disease than expected in a given area or among a specific group of people over a particular period of time
Outbreak	Synonymous with epidemic; sometimes the preferred word, as it may escape sensationalism associated with the word epidemic
Pandemic	An epidemic occurring over a very wide area (several countries or continents) and usually affecting a large proportion of the population
Epizootic	The occurrence of more cases of disease than expected in a given area or among a specific animal population over a particular period of time

ORIGINS OF EPIDEMIOLOGY

Hippocrates (460 – 377 BC)

- Explained disease occurrence from a rational rather than supernatural viewpoint
- “on airs, waters, and places” – suggested environmental factors (water quality/air/food) were implicated in causation of diseases
- **Real contribution to epidemiology**
 - o The need for observation and empirical evidence
 - o Observation that the location, time of year, and the characteristics of the patient had an important impact on the risk of illness

John Graunt Bills of Mortality (1620 – 1674)

- First to use quantitative methods to describe population vital stats by organizing mortality data into a mortality table
- Recorded descriptive characteristics of birth and death data
- Known as the ‘Columbus of Statistics’

Bernardino Ramazzini (1633 – 1714)

- Founder of field of occupational medicine
- Created elaborate descriptions of diseases among many types of workers
- Pioneer in the field of ergonomics

Smallpox (Variola)

- Viral disease responsible for epidemics since ancient times
- **Variolation**
 - o In china, a person was given a mild smallpox which protected them from all types
 - o Dangerous and caused deaths, but also became popular among Arab physicians

Doctor Edward Jenner infected a farmer’s boy with the cowpox and later exposed him to smallpox; the boy remained well, and “vaccination” became widespread and replaced variolation

Doctor John Snow (1813 – 1858)

The Father of Epidemiology

- **1854** – severe cholera outbreak in Soho district of London, England near Broad street
- Violent and sudden – 12.8% death rate in some subdivisions
- **Cholera**
 - o Now known to be an a bacterial **enteric** disease
 - o Characterized by profuse rice-water stool, nausea and vomiting

Robert Koch (1843 – 1910)

German physician who verified that a human disease was caused by a specific living organism

- Isolated the bacteria that causes anthrax (*Bacillus anthracis*), cholera, and tuberculosis
- Koch’s postulates
 - o The disease must be observed in every case of the disease
 - o It must be isolated and grown in pure culture
 - o The pure culture must, when inoculated into a susceptible animal, reproduce the disease
 - o The organism must be observed in, and recovered from, the experimental animal

THE CONTAGION THEORY

Also known as the germ theory

- In 1870s, **Koch** and **Pasteur** both put to rest the popular concept of disease due to “spontaneous generation”, and instead due to microbes
- Marks the beginning of modern microbiology

IMPORTANT CONTRIBUTORS

Hippocrates

- 400BC brought diseases from “supernatural” to natural/environmental; place and time and people as important

Graunt

- 1630; serious collection of mortality data

Ramazzini

- 1690; observations began ergonomics and occupational hygiene

Jenner

- 1800; recognized vaccination better than variolation and ‘formalized’ it

Snow

- 1854; used Epi observations to state water cause of cholera

Semmelweis

- 1850; contaminated hands as cause of childbed fever and hand-washing as essential preventative measure

Pasteur

- 1870; demonstrated microbial cause of many diseases and how they can be prevented

Koch

- 1870; link between specific diseases and specific organisms

CONTEMPORARY ERA – 1940 TO PRESENT

Framingham Study (1948)

- Named for Framingham, Massachusetts
- Prospective cohort study of more than 6,000 people to determine causes of cardiovascular disease

Smoking and health, Doll and Hill (1962)

- Showed smoking as cause of lung cancer
- Prospective study of physicians

Smallpox eradication (1977)

Current Uses of Epidemiology

- Community health assessments
- Making individual decisions
- Completing the clinical picture
- Searching for causes/risk factors

Core Functions in Epidemiology

- Public health surveillance
 - o *The systematic collection, analysis, interpretation, and dissemination of health data on an ongoing basis, to gain knowledge of the pattern of disease occurrence and potential in a community, in order to control and prevent disease in the community*
- Field investigations
 - o *The application or practice of epidemiology to control and prevent health problems*
- Research & evaluation
 - o Investigate risk factors, look for associations
 - o Evaluate the effectiveness of public health programs and activities
- Policy development
 - o Regularly provide input, testimony, and recommendations regarding disease control strategies, reportable disease regulations, and health-care policy

LECTURE TWO

MEASURES

Of Disease Occurrence

- Quantify exposures and outcomes, calculate rates

Of Association

- Evaluate association between exposure and outcome

Of Potential Impact

- Determine what would happen to outcome if the exposure were removed

COUNTING CASES

Cases of death, illness, disease, health-related events are counted to determine the extent of a health problem

- A **count** or **frequency** is a useful number but it cannot be tracked over time because the population which generated it may have increased or shrunk over time
- It cannot be compared with numbers from other populations because their populations may be much bigger or smaller
 - o We need a **COMPARABLE** number
- Cases can be incident or prevalent

Incident (New Cases)

- changes in the status of an individual – from well to ill, uninfected to infected, alive to dead
- determined by following individuals over time and counting those who change their status
- must be counted over a specified period of time
- needed to study the etiology of the disease and to monitor preventive efforts

Prevalent (Existing Cases)

- represent the existing status of an individual – well, ill, uninfected, infected, alive, dead
- determined by taking measurement on people at one point in time
- key in health planning

RISK & INCIDENCE

Risk

The probability that an event will occur

- estimated by observing events among a population during a specified time period
- always a value between 0 and 1
 - o usually expressed as a %
- assumes that the individual
 - o doesn't have the disease at start of follow-up period
 - o doesn't die from other causes during follow-up

Incidence

*Number of **new** cases of a disease among:*

- total number of disease-free individuals at the start of follow-up period
- total time that disease-free individuals are observed during follow-up period

$$\frac{\text{Number of NEW cases of diseases during a period (I)}}{\text{population at the beginning of the period (N)}} = \text{Cumulative incidence or attack rate}$$

Attack Rate

A term used in an outbreak situation

Rates in an Outbreak Scenario	Numerator	Denominator
Overall attack rate	Total # of new cases	Total population
Food-specific attack rate	# of persons who ate a specific food and became ill	Total # or persons who ate the food
Secondary attack rate	# of cases among contacts of primary cases	Total # of contacts

Some may remain disease free for longer than others for various reasons:

- develop outcome of interest, moved away/lost to follow-up, death from unrelated causes, refusals etc.
 - o = **CENSORED**
 - o the denominator changes as persons enter/leave the study (**open/dynamic population**)
 - o a person-time data is used to account for this
 - **Person-time** presents the sum of each individual's time at-risk (a person is at risk as long as they remain disease-free)
 - Provides an estimate of the actual time at-risk of all persons in the study
 - Number of NEW cases of disease (I)
total person-time of observation (P/T) ← incidence density or person time incidence

- Practice question on slide 27 and 28

EQUATIONS COVERED ABOVE:

Cumulative incidence: (I/N)

- I = number of new cases during a period
- N = population at beginning of period

Incidence density: (I/PT)

- I = number of new cases of disease
- PT = total person-time of observation

PREVALENCE

Prevalence = Incidence x Duration – $P = I \times D$

Rates

Number of **existing** cases of a disease in a given population

- At a specific point in time (point prevalence)
 - o # of cases in defined population at one point in time
persons in defined population at same point in time
- Over specified time period (period prevalence)
 - o # cases in defined population during specified period
of persons in defined population during specified period
- Prevalence is a proportion, not a rate; has no units

Use of Prevalence

Useful for:

- Assess health status of a population, planning health services, chronic disease

Note useful for:

- Identifying risk factors, conditions that are acute and/or of short duration

Factors affecting prevalence

This part is mainly common sense: prevalence will increase with health people migrating out and with increased incidence of a disease, whereas the opposite will be true to decrease the prevalence

- Example question on slide 47 of week 2 ppt

Measure	Numerator	Denominator
<i>Incidence proportion</i>	Number of new cases of disease during specified time interval	Population at start of time interval
<i>Secondary Attack Rate</i>	Number of new cases of disease among contacts	Total number of contacts
<i>Incidence rate (or person-time rate)</i>	Number of new cases of disease during specified time interval	Summed person-years of observation or average population during time interval
<i>Point prevalence</i>	Number of current cases (new and pre-existing) at a specified point in time	Population at the same specified point in time
<i>Period prevalence</i>	Number of current cases over a specified period of time	Average or mid-interval population

LECTURE THREE

MORE IMPORTANT RATES

Mortality Rates: we are measuring these as incidence rates (the incidence of death)

Crude (Annual) Mortality Rate

- $\frac{\text{Total \# deaths (all causes) in a year}}{\text{\# persons in population midyear}}$
 - o **EXAMPLE:**
 $\frac{1,892,897 \text{ deaths in USA in 1998}}{212,946,226 \text{ pop. 1 Jul 1998}}$ ← July is roughly the middle of the year hence it's usage

Age/Sex Specific Mortality Rates

- $\frac{\text{Deaths among (x) in (y)}}{\text{midyear pop. of (x) in (y)}} \times 1000$
 - o **(x)** would be the specific demographic (sex, age, etc.)
 - o **(y)** would be the year
 - **EXAMPLE:**
 - $\frac{\text{Deaths among males in 2002}}{\text{midyear population of males in 2002}} \times 1000$
 - this is a sex-specific mortality rate

Cause Specific Mortality Rate

- $\frac{\text{Total no. deaths (assigned to specific cause) in year}}{\text{no. persons in populations midyear}} \times 100,000$

Case-Fatality Rate

- $\frac{\text{No. deaths in specified period}}{\text{no. cases with the disease}} \times 100$

Proportionate Mortality

- $\frac{\text{No. deaths from (cause) in 2005}}{\text{total deaths in 2005}} \times 100$

Mortality Rate = Attack Rate x Case-Fatality Rate

- Attack rate = incidence rate (often used for small groups, acute disease and outbreaks, but essentially same)

***** Week 4 has a clarification PowerPoint on some of the equations given *****

LECTURE FOUR

*Counts and rates must be organized by **person, place, and time**; it allows us to identify and show epidemiologically relevant patterns*

ORIENTING THE DATA – PERSON

Characterize by Attributes

- Need to collect the following information for each case:
 - o Identifying
 - Name, date of birth (DOB), contact information, must be kept confidential
 - o Demographic
 - Age, gender, occupation/place of occupation, geographic location – describes 'at risk' pop.
 - o Clinical
 - Date of symptom onset & date of report, symptoms, severity of illness, laboratory findings
 - *****MAJOR SYMPTOMS: diarrhea, vomiting, headache, fever**
 - o Risk factors

- Usually restricted to general and well-established risk factors early in investigation
 - Food history, recent travel, contact with animals, etc.
 - Contacts
 - Who are the contacts? Are they ill?
 - Cluster distributions can suggest common exposures & transmissions
- Typically collected using generic or disease specific case report form

When orienting the data by person, calculate the frequency distributions of demographic clinical and risk factor characteristics of cases, as well as make comparisons to historical or baseline data

- Must calculate rates in order to make valid comparisons between populations

REVIEW OF KEY CONCEPTS

Counts

- Number of cases
- Numerator data alone is not very informative
 - Also need:
 - Who/what is in the denominator?
 - In what time period did the cases occur
- Need to calculate proportions, ratios, and rates

Proportion

- Division of two numbers
- Numerator is always included in the denominator
- Quantities have to be of the same nature
- Always ranges between 0 and 1
 - Usually expressed as a %
 - **EXAMPLE:** proportion of green M&Ms in a big of M&Ms

Ratio

- Division of two numbers
- Numerator NOT included in denominator
- Quantities can be of different natures

Rate

- Division of two numbers
- Time is included in denominator
 - Measure the frequency with which an event occurs in a defined population
- Always non-negative but no upper limit
 - May be expressed as a power of 10

Important Measures

- **Prevalence Rate:** number of existing cases in a population at a given point in time or over a specified period of time – associated with burden
- **Incidence Rate:** number of new cases in a population over a specified period of time – associated with risk

Age

- Represents
 - The condition of the host and its susceptibility to disease
 - Differing intensities of exposure to causative agents
 - Passage of time

ORIENTING THE DATA – PLACE

- What kind of place?
- Need to distinguish between place of onset, exposure, case identification
- Map the cases by location (spot map vs area map)
- Data can be expressed as counts or rates
 - Rates allow comparisons between populations but may need to be age standardized
- Illustrates geographical extend of outbreak and may provide clues to the source or mode of transmission

ORIENTING THE DATA – TIME

- Construct an epidemic curve ('epi' curve)
 - o Illustrates distribution of cases over time
- Epi curves can provide information on outbreaks
 - o Pattern of spread, magnitude, time trend, outliers, type of transmission, time of exposure
 - o Plot the number of cases on the y-axis
 - o The date of symptom onset on the x-axis
 - Use time intervals that are 1/4 the average incubation period (if incubation known)

There's more on drawing epi curves, but it's probably not needed for a multiple choice exam

- o if you know the disease and have a **point source outbreak (PSO)**, you can use the epi curve to identify the most likely time period of exposure that led to the outbreak
 - **PSO** – persons are exposed to the same source over a brief time, such as through a single meal or at an event
 - **EXAMPLE:** Aisha has Tuberculosis and sneezes all over the big bowl of popcorn everyone is sharing – the point source is the bowl of popcorn, and since no one saw her sneeze and continued to eat from the bowl, they are the persons exposed to the same source
 - **Thanks Aisha**

PATTERNS OF SPREAD

Common source outbreak

- a single source of pathogen results in exposure:
 - o at one point in time (**Point Source Outbreak**)
 - sharp upslope (# of affected increased immediately/rapidly)
 - gradual downslope (cases will occur within one incubation period but those affected will heal more gradually in comparison to how they were infected)
 - o at several points in time (**Intermittent source**)
 - multiple peaks
 - o over a continuous period of time (**Continuous Source**)
 - gradually rises and sometimes plateaus as opposed to peaking
 - lasts more than one incubation period

Propagated outbreak

- spread from person-to-person via an intermediate host
 - o **EXAMPLE:** one person has the flu and decides to drink from a water fountain by putting their entire damn mouth on the spout. After this, the water coming out from it will be perpetually contaminated with the harmful bacteria which will in turn infect other users of the fountain. In this case, the fountain spout is the intermediate host and transfers it person-to-person

Mixed Epidemics

Outbreaks with several modes of transmission

- Common source of initial point source, outbreak followed by secondary transmission from person-to-person spread
- Infected individuals serve as sources of the agent to infect others

Calculating Exposure Period

- Assume point source outbreak and known agent
- If incubation period (IP) is known, epi-curve suggests a time period of exposure
- Cases will be distributed across range of the incubation period
- Earlier cases = shorter IP
- Later cases = longer IP

TRENDS

Secular

The rate of disease over many years

- Can suggest key events, improvements in control or other factors
- Typically a line graph

Cyclical

Increases and decreases in the frequency of a disease over a period of several years or within a year

- Description by season, month, day, or time of day, can be important for some diseases
- Chicken pox is seasonal (March/May), Influenza peaks in February
 - o Seasonal patterns often seen in foodborne illness outbreaks
 - o Weather or event-related
 - Increase of colds during colder months

*****Go over the *Epi Curve and Central Tendency* PowerPoint in week 4 for Mean/Median/Mode*****

LECTURE FIVE (WEEK SIX)

DISEASE TRANSMISSION

Infectious vs non-infectious

- Similar use of surveillance
- Similar study methods
- Similar methods of collecting data and analysis
- Similar methods of conducting investigations

CASES

Primary case: *patient zero* – the initial patient in the population of an epidemiological investigation

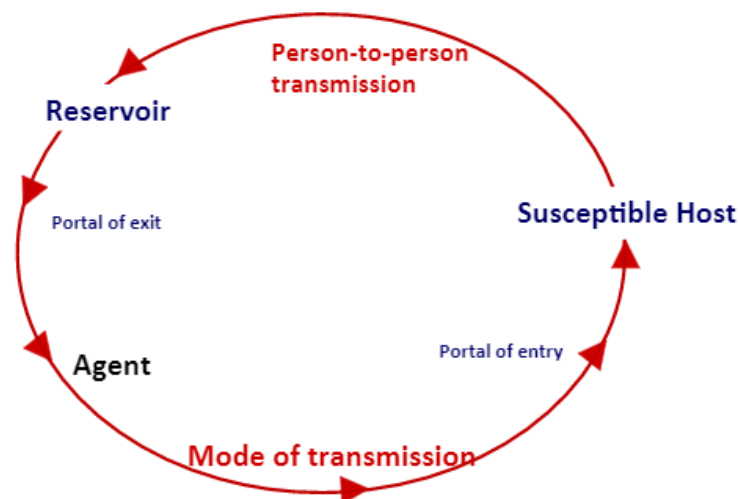
- A more general definition: the first case of a condition or syndrome

Secondary case: occurrence of a disease in a subject due to close contacts of a primary-case patient 24 hours after onset of illness in the primary case

- **EXAMPLE:** Sana develops Ebola because she's gross. As the first instance of Ebola, she is patient zero; she then proceeds to cough on Danny a few hours later – Danny is the secondary case.

Index case: the first identified case in a group of related cases of a particular communicable or heritable disease

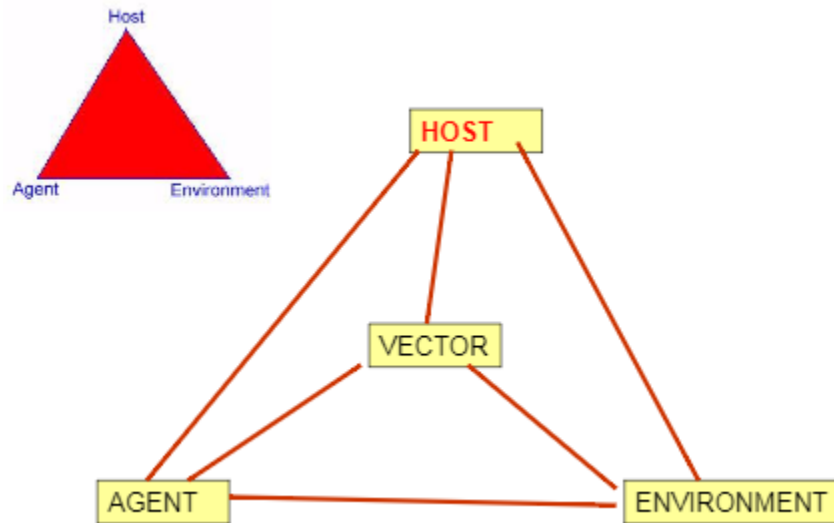
CHAIN OF TRANSMISSION



DISEASE OCCURRENCE IN POPULATIONS

- **Sporadic:** occasional cases occurring at irregular intervals
- **Endemic:** continuous occurrence at an expected frequency over a certain period of time and in a certain geographical location
- **Epidemic or outbreak:** occurrence in a community or region of cases of an illness with a frequency clearly in excess of normal expectancy
- **Pandemic:** epidemic involves several countries or continents affecting a large population

EPIDEMIOLOGICAL "TRIAD"



*remove one of these things from the triad and the transmission is usually broken

MODES OF TRANSMISSION

Direct Transmission

- Direct and essentially immediate transfer of infectious agents to a receptive portal of entry
- Contact

Indirect Transmission

- Vehicle-borne: contaminated inanimate materials or objects
 - o Think of the water fountain example
- Vector-borne: mechanical and biological
- Airborne: dissemination of microbial aerosols to a portal of entry, usually the respiratory tract