Epidemiology

Health Science 2003
Ali Walker

Epidemiology = study of the distribution & determinants of health or events in human populations and the application of this study to control health problems

Endemic = infectious & chronic disease @ places
Epidemic = infectious & chronic disease @ times
Pandemic = worldwide epidemic

(Porta, 2008, as cited in Webb and Bain, 2011)

Assumptions of Epidemiology

diseases occur randomly within populations but follow predictable patterns (CDC, 2012)

Truth is diseases have causes & patterns but protective determinants of disease are *not known*

4 key Epidemiologic aspects

Health [morbidity - illness] [mortality] cancer / heart disease / stroke

chronic disease infectious disease injury/disability

Population not individuals

demographic transition = high birth & deaths >>> lower rates

epidemiological transition = chronic diseases causes of morbidity & mortality

Distribution = person/ place /time

Determinants = factors that influence health state/events of individuals & communities.

Biological - *bacteria, viruses, fungi*

Enviro - *air, water, chemical, structural*

Individual - *lifestyle/choices/behavior*

Social - *poverty, discrimination, income inequality*

DISEASE = any health phenomenon of interest

Social Determinants of Health

Individual Level income | education | housing | food security | discrimination | trauma

Social level country's per capita | unemployment rate | neighborhood quality | income inequality

Disease measurements

Symptoms - subjective indicators reported by person

signs - disease indicators apparent to a doctor

tests - tool used to determine diagnosis

Clinical Endpoints of Disease: death disability recovery (remission)

Counts = # of cases of morbidity or mortality

Ratios | Proportions | Rates

Ratios compare 2 values which *may* be related A / B 754 men / 771 women = 97.8%

Proportions compare parts to a whole A / A+B 7543 men, 189 are 50+ yrs, 189 / 7543= 2.4%

Rates type of ratio but measures time

(2009): 238,418 people died, population at 2009 midpoint was 33,580,000

Crude Mortality Rate = 238418 / 33 580 000 *1000 = 7.1 deaths per 1000 people in 2009

Measuring the burden of disease

Prevalence what proportion of population has/ had the disease (new+existing cases)

Prevalence = # of persons with disease / total # population

Point Prevalence - cases @ point in time Dec.15, 2005. 507 people had the flu, population=32,359,000

Period Prevalence - cases over period of time Nov.1-Dec31, 951 people had the flu, N=32,359,000

Lifetime Prevalence - cases over lifespan phone survey of 1877 people, 1766 had the flu in their lifetime

Factors that influence Prevalence:

increase: longer duration of disease/ improved diagnostics/ new cases (incidence)

decrease: high fatality rate/ cure / less new cases/ shorter duration of disease

Incidence # of NEW cases of disease during time period

measures how quickly people are getting disease/ risk of developing it

Cumulative Incidence = # NEW cases over time period / population @ risk, same time

19,300 NEW diabetics, 206,000 diabetes cases @ end of 2009, n = 3,632,000

1. 206,000 (d) - 19300 (new) = 186,000 (diabetics)
2. 3,632,00 (n) - 186,700 (diabetics) = 3,445,300 (@ risk)
3. 19,300 (new) / 3,445,300 (@ risk) * 10,000 = 5.6 cases per 10,000 people

Incidence Rate = # NEW cases / # people * time [person-years @ risk]

in a study: participant #1 followed for 4 yrs, #2 = 6yrs, #3 = 14yrs .... and #1, #2 & #7 = diagnosis

... Total person-years = 57 pyrs and Incidence rate = 3 (new) / 57 pyrs *1000 == 53 cases per 1000 pyrs

midpoint (mean) population

start of year population + end of year population * 1 / 2

Lethbridge population start of 2013 = 89,074 @ end of 2013: 90,417

Surveillance ! needed for descriptive epidemiology (morbidity/ mortality), health data, risk factors, identify populations @risk

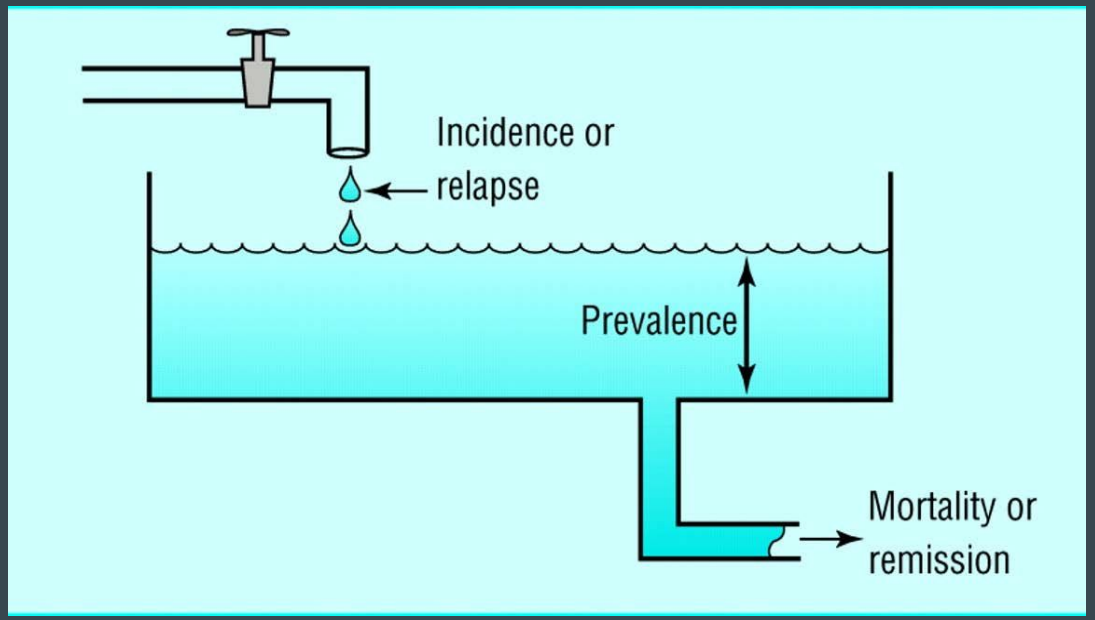
Passive surveillance = mandated reporting, health care system

Active surveillance = field trips! interview physicians, patients, medical records

Types: surveys, data collection= mortality, morbidity

Incidence & Prevalence

Prevalence = Incidence Rate * duration of disease



MORTALITY

MORTALITY !

(Ctrl + a) mortality = death from all causes

(Ctrl + a) = # deaths in 1 yr / # ppl @risk

Cause - specific mortality = death by disease type

deaths in 1 yr by disease type / # ppl @ risk

Case FATALITY ! = # deaths by case in time period
diagnosed cases same time *100

Crude Mortality Rates = # of death but not accurate for comparing populations

* age is most important mortality predictor

time MUST be used in rate

Case FATALITY vs. Cause - specific mortality

In a population of 100,000, say 5 people get Rabies.

In 1 year 4 people have died from rabies

Cause - specific mortality = 4 deaths / 100,000 = .004%

Case FATALITY ! = 4 deaths / 5 cases = 80% !

so, don't worry *unless* you get rabies

Mortality

Proportionate Mortality =

deaths by specific disease (time)
total deaths (same time)

the major causes of death in a population

Survival rate

cases alive (time period) post diagnosis
cases diagnosed (start of time) *100

Maternal Mortality Ratio

indicates the health of a population: health care access, economics, inequalities

deaths by childbirth in 1 yr / # of livebirths in 1 yr *100,000 births

Infant Mortality Rate (IMR)

death of infants <1 yr / # of livebirths in same year *1000 births

Now to the burden of disease …

the loss of health in society resulting from disease/injury

life expectancy = # of years a person is expected to live

Potential Years of Life Lost (PYLL) = death before age 75, [age @ death - 75]

Disability adjusted life yrs = PYLL + Yrs Lost from disability

Health adjusted life expectancy = # of yrs @ full health

60 yrs @ full health = 50 yrs in full health + 20 yrs @ 50%

Epidemiology study designs

Hierarchy of study designs

1. Systematic Reviews
2. Experiments: Community & Clinical Trial
3. [observation - Analytic] Cohort & Case-Control
4. [observation - Descriptive] Cross-section & Ecological populations
5. [observation - Descriptive] Case report/series

Descriptive = Who| What| Where| When *what is*
Analytical = How + Why, *ID's cause*

Cross-sectional study individuals exposures + outcome, prevalence

Population >> sample >> exposures | outcome

Cons: selection bias, **no temporal sequence, no rare outcomes**

Pros: cheap, simple

Case-Control = Odds Ratio

determines how disease & non-disease groups differ based on past exposures

Population >> sample >> disease cases || disease controls << >> PAST

selecting: Cases based on registries & health records

Cons: selection/ recall/ interviewer **bias**, **no temporal seq**

Controls *sample *must* be from same population

Exposures for only 1 outcome, many factors

Pros: quick, many exposures + rare outcomes,

Experimental studies analytical, evidence for causation

sample population >> exposed | not exposed >> expos.-outcome | non-exposed outcome

compare therapy vs. no therapy| therapy vs. placebo | therapy A vs. B

Preventative trials / intervention trials (high risk groups)/ therapy trials (disease)

Controls may be historical/ non-random / ***random*** is best no bias & confounders

Blinding = random selection

Single-blind study = participants don't know what group they're in

Double-blind study = participants & experimenters don't know group assignments

triple-blind study = participants,experimenters & data scientists don't know groups

Phases of trials:

Phase 0 = animals or lab trials
Phase 1 = drug safety
Phase 2 = efficacy in large groups
Phase 3 = efficacy + safety, placebo
Phase 4 = long term effects, any adverse effects

Crossover - Random Clinical Trials

participants get intervention AND the placebo @ random, patients are self(controls)
“washout period” is needed between interventions

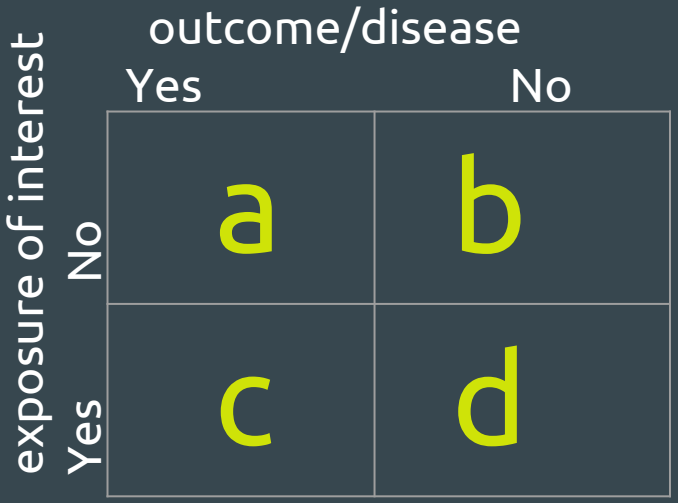
Ethics in experimental research :
1. informed consent 2. no holding effective treatment 3. must protect the participants

@random Pros : causal + high confidence level, less bias
@ random Cons : costly, lengthy, hard to find samples, ethics

Community Trials Pros: real estimates of exposure, random, policy review
Cons: hard to blind study, outcomes from other factors, ecological fallacy

Odds, ratios & risks

what are the odds? & odds ratio



odds ratio = $\frac{a*d}{b*c}$

Association between exposure & disease in

case control study = Odds Ratio

odds = individuals with attribute ÷ individuals w/o attribute
odds ratio = exposure and disease association

Relative Risk = Cohort study (a/a+b) ÷ (c/c+d)

odds ratio interpretation

NULL (no effect)

← 0.**3** -- *0.4* -- *0.5* -- *0.6* -- 0.7 -- 0.8 -- 0.9 --[1]-- 1.1 -- 1.2 -- 1.3 -- 1.4 -- 1.**5** →

------ mod-------

1- *0.6* = 40% ***LESS LIKELY*** to get disease

1.4 = 40% **more** likely to get disease

odds ratio <1		odds ratio >1
0.3	= **strong** association =	5.0+
0.4 - 0.6	= *moderate* association =	2.0 - 4.0
0.7+	= weak association =	< 2.0

so if odds ratio = 6.3, means 6.3 times ***more likely*** to get outcome

if odds ratio = 0.43, 1 - 0.43 == 0.57, means *57%* ***less likely*** for outcome

Attributable Risk

the incident difference between the exposed and unexposed

AR = incident rate(exposed) - incident rate (*un*exposed)

AR = cumulative incidence (exposed) - cumulative incidence (*un*exposed)

melanoma in indoor tanners = 49.6 cases per 100,000 non-tanners = 17.7 per 100,000
AR = 49.6 - 17.7 == 31.0 difference between exposed & unexposed
… means tanners **have extra 31 cases** of melanoma

injury/death in Italy (5 yrs Cumulative Incidence) = 0.8 per 100,000
injury/death in UK (5 yrs Cumulative Incidence) = 0.3 per 100,000
AR = .8 - .3 == .5 per 100,000

Attributable Fraction (AF)

is the % difference of disease in the exposed group which would've NOT occurred

AF = Attributable Risk (AR) / incidence (exposed) *100

Example:
Group 1 AR = 48 per 100,000 cases of disease and incidence of exposed = 58 per 100,000
AF = 0.00048 / 0.00058 = 82% means 82% of disease is found in group 1

AR = 31.0 per 100,000 and disease in tanners = 49.6 per 100,000
AF = 0.00031 / 0.000496 = 62.5% means 62.5% of disease is found in tanners group

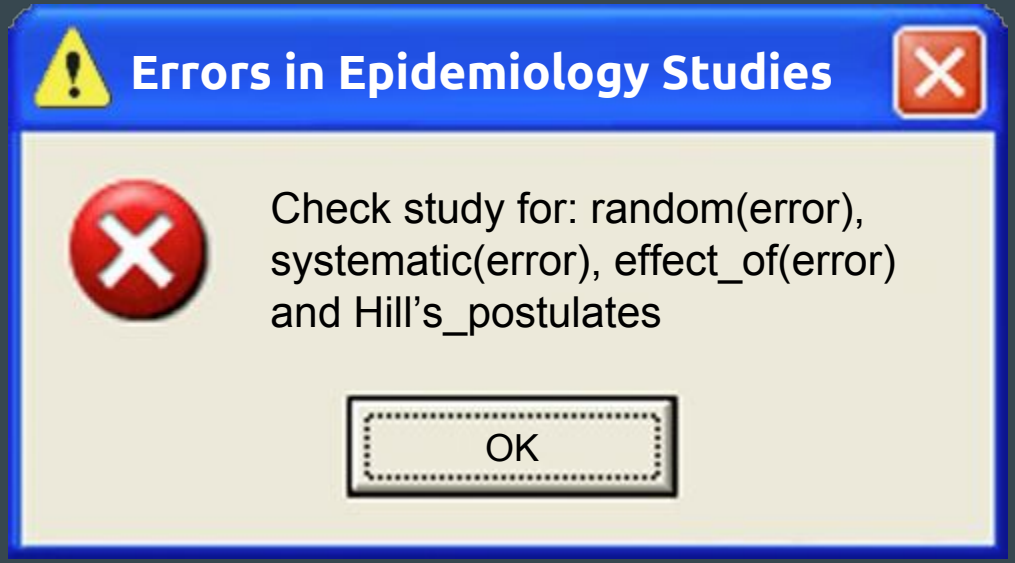
Population Attributable Risk (PAR)

compares amount of disease in whole population with amount of disease in unexposed

PAR = incidence in population − incidence of *un*exposed

Population Attributable Fraction = PAR / incidence population

Causation & Errors



Epidemiology process of Errors

Association?
    Bias ?
        Chance ?
            Confounding ?
                Causation ?
                    return (Action)

Internal validity = degree to which a study is *Free* from error

External validity = extent a study can be applied to a **broader** population

Sources of Error: Bias & Chance

Precision = *ideal, low random error & systematic error

Accurate & not precise = some systematic error, random error

Precise & not accurate = some random error, systematic error

Inaccurate & not precise = random & systematic error

P + A
A -P
P -A
-A -P

systematic error = selection bias + measurement bias

Random Error chance

value of sample diverges due to chance alone from population value

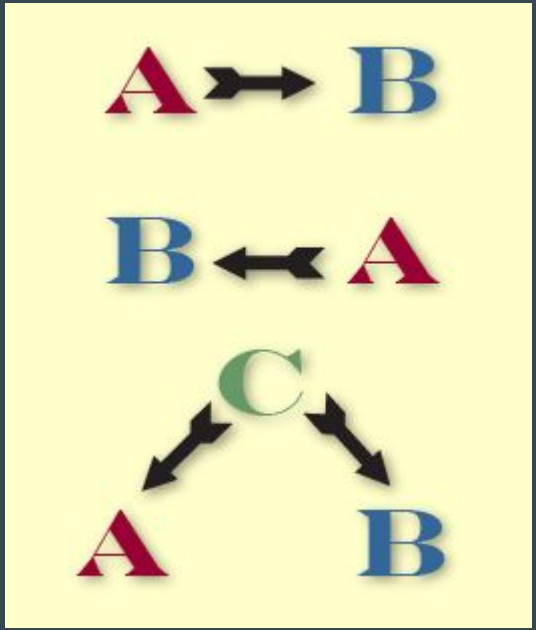
Confidence Interval helps determine if association is statistically significant

Systematic Error (bias)

Selection bias = selecting participants differ from selected & targeted

Measurement bias = process of data collecting
(recall bias, reporting bias, miscalculation bias)

Causation !



Cause an event/ condition/ characteristic or combo of these that play a role in producing an occurrence of a disease

Component cause = factor that contributes towards disease causation but NOT sufficient to cause disease on its own

Sufficient cause = a factor or combo that will produce disease

Necessary cause = any agent that is *required* for development of a disease

Causal relationships:

1. necessary & sufficient
2. necessary but not sufficient
3. sufficient but not necessary
4. not sufficient nor necessary

Hill's Postulates for Causation **STDREB**

1. Strength of association (cause & effect: risk ratio or RR)
2. Temporal sequence = exposure *before* disease (prevalence or incidence?)
3. Dose response relationship = risk increases with exposure level
4. Repetition = *association between exposure & outcome (same exposure & outcome)*
5. Experimental evidence = random + controlled trials, * experimental study
6. Biological plausibility = evidence for bio association

Random Clinical Trials = **strong** causation "proof"
Cohort studies = moderate "proof"
Case control = moderate "proof"
Cross-sectional & Ecological studies = weak "proof"

Prevention

Prevention

Primary = prevent disease exposure from occurring, reduce the incidence rates
immunizations, safe drinking water Upstream = *primordial prevention*

Barriers of implementation = financial/ cultural/ social/ ethical

Secondary = screening, detection, reduce morbidity, mortality that already exists
screen for cholesterol, hearing loss

Tertiary = limit disease progression, rehab to enhance quality of life

Use Population Attributable Fraction to target high yield intervention

Strategies for prevention

High risk approach for vulnerable individuals: ID people, control level of exposure to cause or protect against exposure outcome.

[drug users → needle exchange → vaccine for Hep B]

Best for clinical practice, community level problem is ID minority & control isolation

Pro: benefits those likely to have future health problems

Cons: no reduction of Incidence, little benefit to population

Strategies for prevention

Mass strategy (Geoffrey Rose) - disease & exposures reflect society behavior, aim to reduce health risks of population

Immunization, fluoridation of water, seat belts

Attributable burden = amount of disease above exposure reference level

Attributable fraction = proportion of disease attributed to risk factor

Pro high-risk: best for individuals, cost effective Cons high risk: expensive, temp.

Pro pop.: radical, social risk reduction Cons pop.: no benefit at individual

Screening

Screening will not diagnose disease but reduce morbidity & mortality + disability.

Protect population from exposure - *immigrants screened for HIV, Hep B.*

Screen people free from disease, selected by health service

Disease process: Critical Point = irreversible Bio onset → CP1 → early diagnosis → CP2 (screen) →clinical diagnosis → CP3 → outcome

Epidemiologists evaluate screening tests & programs: *Clinical trials, ecological studies, Case-control studies*

Screening the disease

Disease should be severe & common public health problem

Screening best if disease has long duration between 1st signs & symptoms

Screening dichotomous variable = infectious disease w/ known agent

Screening continuous variable = weight, alcoholism, blood pressure

Screening the test

The test that enables detection of disease *before* time of diagnosis

Must be accurate, sensitive - ID's all people with disease, specific - ID only those with that disease, all w/o should get test negative & safe

Social considerations = health problems should be important to population, cost benefit ratio, public acceptance
Scientific evidence = prevalence is high in population, early detection improves outcomes
Ethical considerations = is it equitable, etc
Small risk vs. large benefit from diagnosis + social/cultural acceptance + simple & cheap

validity(accuracy) = ability to distinguish who has disease vs. "gold standard"

reliability(precision) = repeat same results

Impossible for screen to be unreliable but accurate (not precise & valid)

Screening approaches

Target screening = aimed at groups w/ high risk

Mass screening = population based

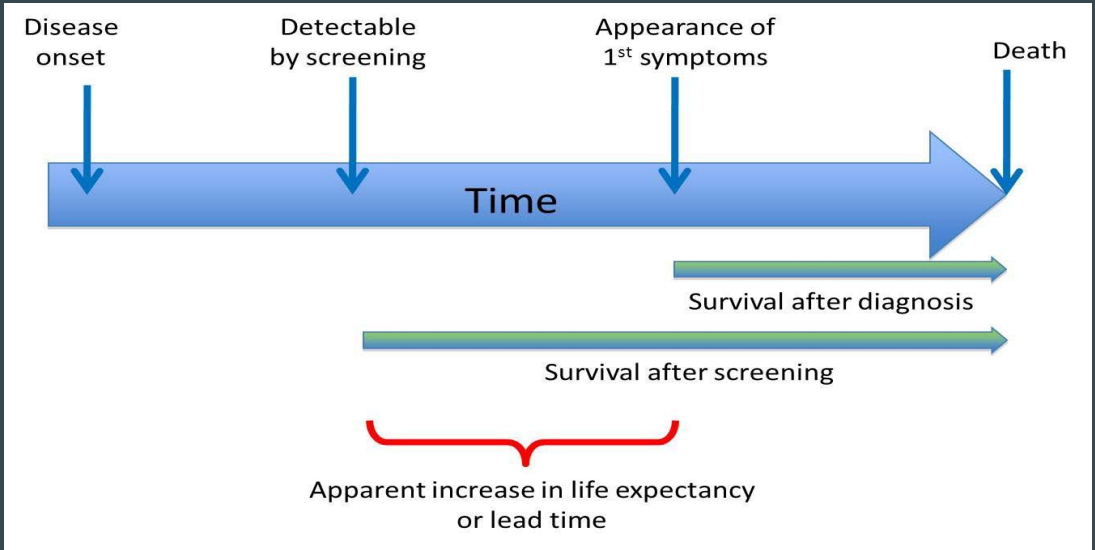
Case finding = early detection at doctor's office

3 BIAS in screening: Volunteer bias | lead-time bias | length time bias

Volunteer bias *Wealthy* | *healthy* | *wise* | *family history* go to screenings

Lead-time bias delay of disease onset and symptoms then diagnosis, *illusion of better outcome (survival rate) because of earlier detection when there could be no benefit at all from screenings compared to non-screened group*

Survival proportion = # cases alive / # cases diagnosed *100 (5yr period) ** ignores new diagnosed

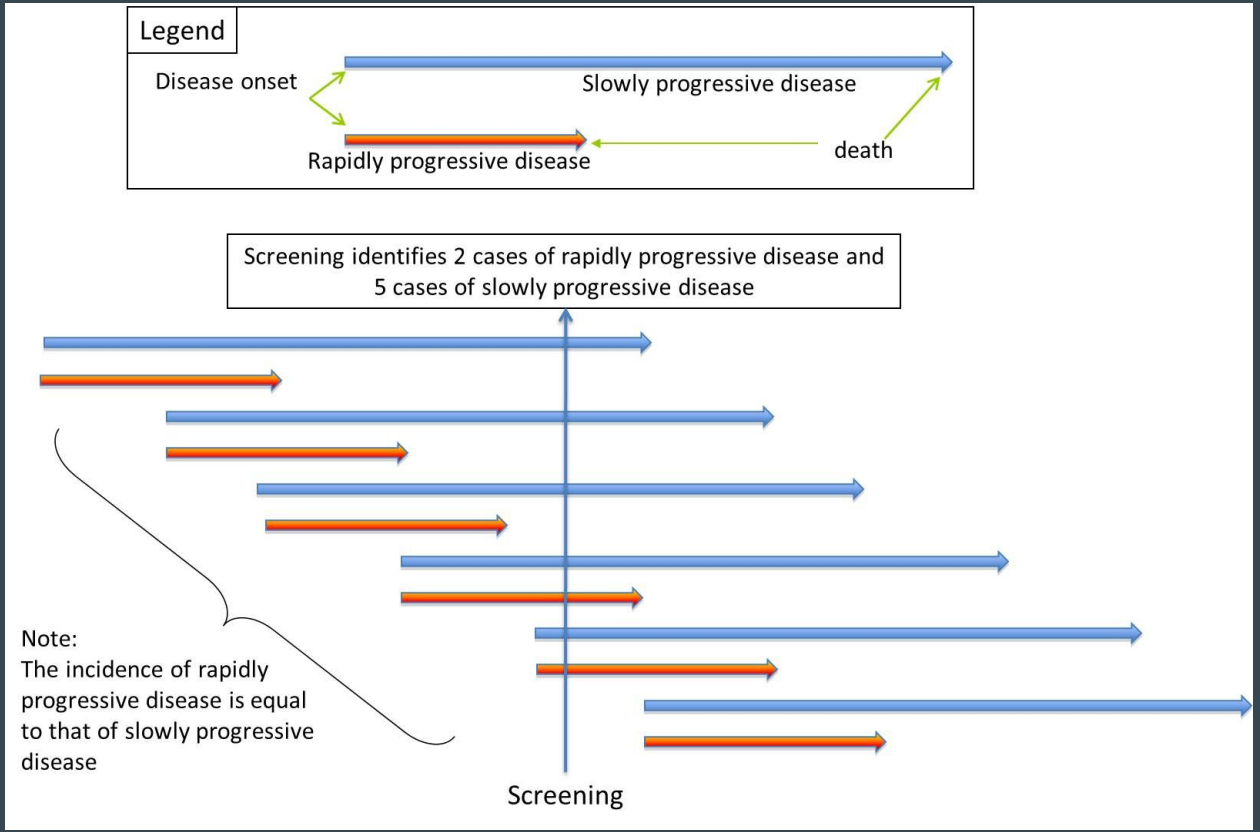


Length time bias

Disease not equal in everybody

Screen for slow disease = pos. outcomes & detection

`Impression:` screening looks effective than reality due to detection



Test quality: sensitive & specific

Test results	Disease status: Positive	Disease status: Negative
Positive	True Pos (a)	False Pos (b)
Negative	False Neg (c)	True Neg (d)

Sensitivity = 152/160 = 95%
correctly ID for disease

Specificity = 714/840 = 85%
CORRECTLY ID for True Negative
Screening incorrectly ID 15% of healthy people as having disease

Sensitivity = how well a test classifies people with disease as sick. % of those test positive

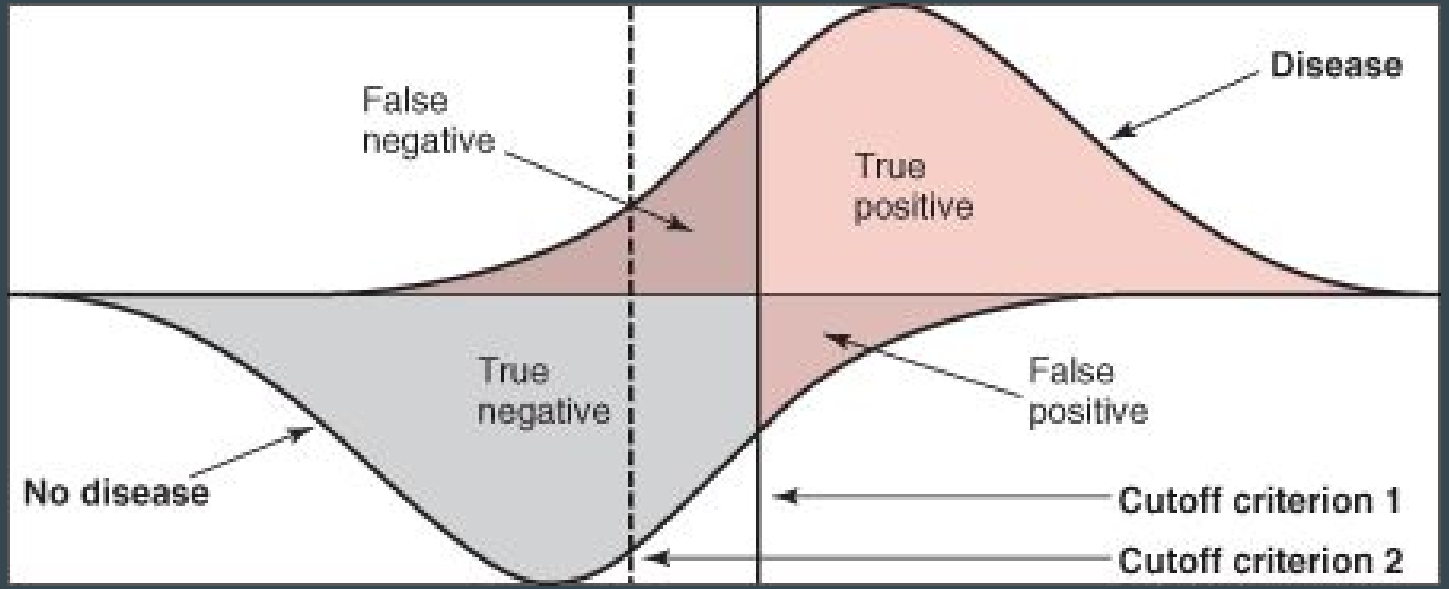
Sensitivity % = True Pos / all disease
= a / a+c *100

Specificity = how well a test classifies people without disease as healthy, those who test negative

Specificity % = True Neg / all -disease
= d / b+d *100

Cut-points on sensitivity & specificity

Screening for continuous variables, cut points determine when a test result is considered positive and negative, which affects sensitivity & specificity



Pos or Neg Predictive Values

positive or negative predictive values (PPV or NPV) really measure how well the test works in a given population, given **Prevalence rates** in population

PPV % = True Pos / all pos *100 = a / a+b *100. (% of people who really have disease)

NPV % = True Neg / all neg *100 = d / c+d *100 (% of people tested: don't have disease)

Measures the accuracy of test (sensitivity & specificity) and disease Prevalence

Predictive Values - example

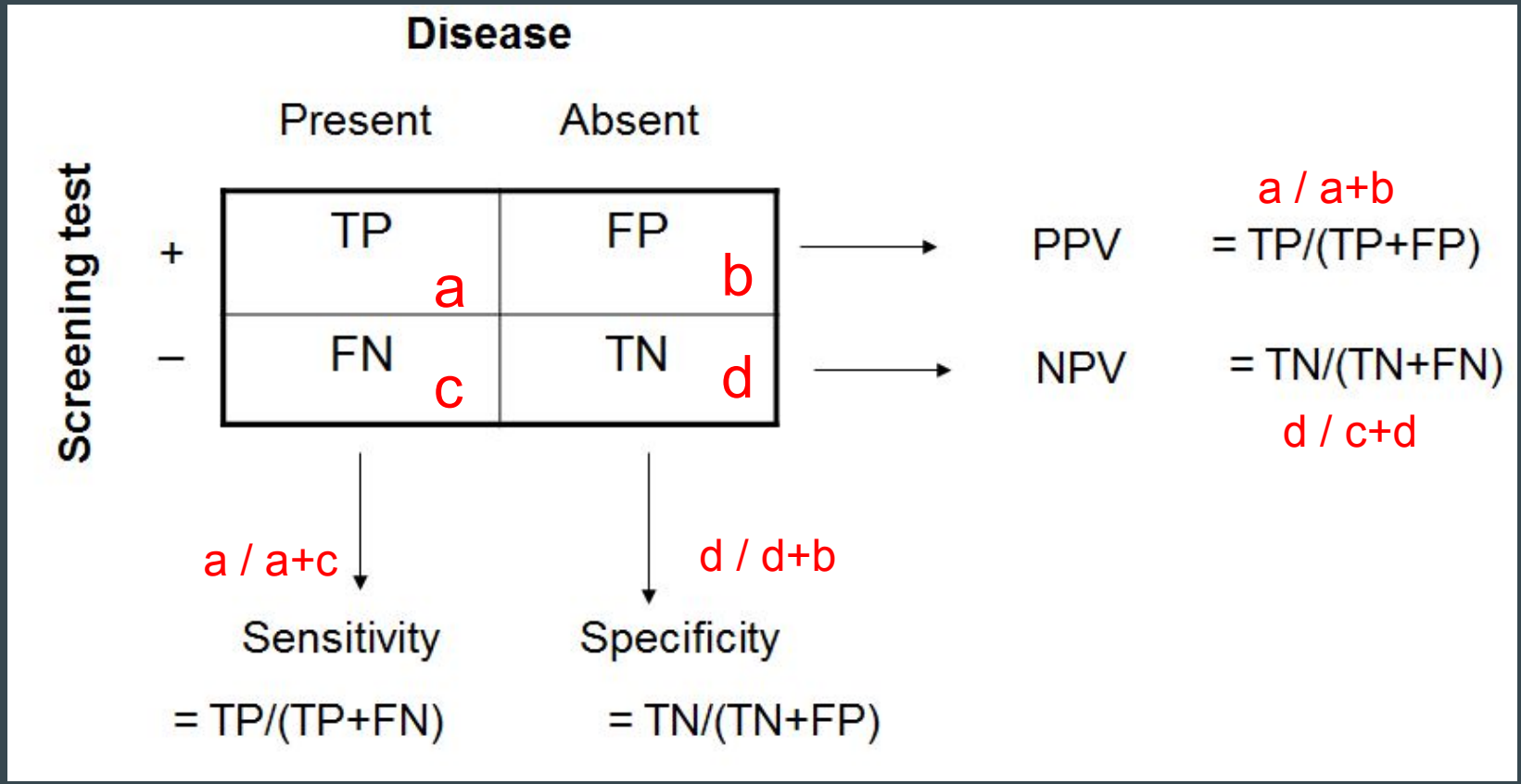
IV drug user	HIV status: Pos	Neg	Total
Pos	5473	22	5495
Neg	27	4478	4505
	5500	4500	10,000

PPV = 5473 / 5495 =99.6%
NPV = 4478 / 4505 =99.4%

New blood donors	HIV status: Pos	Neg	Total
Pos	4	50	54
Neg	0	9946	9946
	4	9996	10,000

PPV = 4 / 54 =7.4% (93% false positives)
NPV = 9946 / 9946 = 100%

Blood donors: 4 people of 10,000 really have HIV, 50 would falsely test positive


Disease
Present
Absent
Screening test
+
–
TP
a
FP
b
FN
c
TN
d
PPV
= TP/(TP+FP)
a / a+b
NPV
= TN/(TN+FN)
d / c+d
a / a+c
Sensitivity
= TP/(TP+FN)
d / d+b
Specificity
= TN/(TN+FP)

Infectious diseases

Terms

EPIDEMIC [outbreak] = excess of disease than expected

ENDEMIC = constant disease in geo area or population group

- `Holoendemic` = infects everyone, mostly kids
- `Hyperendemic` = high incidence & prevalence in all ages

PANDEMIC = disease gone international

CLUSTER = grouping of uncommon disease in a space/time [enviro driven]

Infectious Agents Bacteria viruses fungi protozoa (parasites)

Characteristics: IPVA

Infectivity = ability of agent to invade & multiply in the host
Pathogenicity = power of illness in the host
Virulence = ability to produce serious illness or death
Antigenicity = ability to use antibody production in host

Portals of entry: I.I.A.Vt.S.B

Inhalation / ingestion / absorption / vertical transmission / sexual contact/ blood

Infestation = external | Infection = internal

Reservoirs: Human (STD) | Animal (sheep-anthrax) | Enviro (soil-fungus)

Attack rate proportion # exposed & sick / # people exposed *100

Case fatality # deaths from cases (time) / # cases same time *100

Incubation period = $t_{initial}$ infection ⟶ onset of disease

Latent infection = agent in host & being transmitted with(out) signs

Clinical disease = signs & symptoms

Subclinical disease = no signs & symptoms

SEVERITY OF DISEASE DEPENDS ON **HOST** preventing agent entrance & immune response

Vaccination for disease control

Host immunity = disease makes antibodies to destroy agent

Active immunity = disease trigger immune response or via vaccines -- for life

Passive immunity = acquire antibodies via vertical transmission or antiglobulins -- short lived protection

HERD immunity = protection for all by immunity of many

Enviro determinants = ability to survive & thrive

Physical enviro = climate change, creation of breeding sites

Social enviro = burial practices, overcrowding, travel

Political enviro = war, globalization, economics

Infectious disease Transmission - contact (DID)

Direct = person to person

Indirect = agent on fomite (object) & survives long time

Droplet = cough/ sneeze from infected person to host

Non-contact transmission (AVV)

Airborne = aerosols via respiratory tract, living long time
Vehicle = food, water, soil
Vector = transmission agent from living organism

Infectious disease

Control = reduce epi rates (incidence, prevalence, morbidity, mortality)

Elimination = incidence is zero in geo area

Eradication = reduce world incidence to zero

Extinction = agent no longer exists naturally

Point-source epidemics = sudden exposure to common agent, cases defined by symptoms

Reproduction number R_0

Average # of secondary cases of disease from primary source in population

R_0 = 1 means each cases replaces itself, disease is endemic

R_0 > 1 means each primary case produces more than second cases, epidemic

R_0 < 1 means no replacement of disease, `elimination of disease`

Quarantine = restricts movement of people exposed

Isolation = separate infected from uninfected

Conditions for epidemics

Re-emerging infectious agent

Environments that enable transmission

Increase in susceptible hosts

Failure to control

6 epidemic management steps:

1. Confirm epidemic exists
2. Case definitions & case counts
3. Extent of epidemic
4. Define population @risk
5. Hypothesis & test source
6. Plan long term prevention & control

ENVIRONMENTAL EPIDEMIOLOGY & GIS

Health & Environment

Global burden of disease (death & disability) attributable to environment = 25% and 19% of cancers from environmental exposures

Environment = *physical, chemical, biological, social, political and built environments*

Air pollution and mortality rates

Host characteristics : CSI-EP

- Children - unable to process & remove chemicals
- Seniors - “body burden” from lifetime of exposures & weak immune system
- Individuals - weak immune system from disease or physical alterations
- Exposure status - location, population & level of exposure
- Pregnant women - chemicals cross placental barrier

Exposure agents

Environment Agent EXposures = hazard waste, air pollution, water quality & radiation

Chemical agents - pesticide, asbestos, cleaners, paints

Metallic compounds - natural & byproduct of metal refinement

Ionizing radiation - natural & synthetic, nuclear exposure & health outcomes, radon

Electric & magnetic (non-ionizing radiation) - microwave, clocks, cellphones, electric line

Allergens & molds - in air & environment

Sick building syndrome - caused by poor air ventilation, low humidity, poor lighting

Toxicology concepts

Dose response curves - association between amount of exposure in host and health effects

Threshold - lowest dose at which a response may occur, subthreshold = no effect

Latency - time period between initial exposure and measurable response

Synergism - combined effect of several exposures, (asbestos + smoking = lung cancer)

Canadian environment

Strengths - forest resources, threatened species, electric power, water quality

Weaknesses - marine environment, water consumption & waste, greenhouse gas, nitrous oxide = pollution

Geographic Information Systems

WHERE people live/ agents of disease are/ delivery of health service is needed

Reasons for Geo variation of disease: local difference in reporting & detection, social demographics & exposures, local environmental factors

GIS = collection of computer software, data, analysis, spatial/temporal reference points

Applications: disease surveillance & mapping, risk analysis of exposures, health access planning, community health profiling

Social epidemiology

Social epidemiology study of social distribution & determinants of health

Health outcomes determinants

Genetics (Biology)

Behaviour

Environment

Social determinants of health

Individual level = income, housing, employment, education, childhood, food security

Group level = GDP, economics of country, community, income inequality

Pathways SDOH influence health

Material deprivation = lacking resources or conditions to participate in society

Low socioecon. status >> few choice/resources >> low health

Biologic (psychosocial) = uncontrolled stress/discrimination & unhealthy coping skills

High stress >> psych impact >> bio reaction >> bad coping skills = ill health

Life course :

Latency model = exposures during critical/sensitive times can cause biological changes that impact disease outcomes (child development & nutrition)

Cumulative & chain of risk model = exposures across entire life & accumulate

Epigenetics

stress

Sympathetic nervous system & Hypothalamic Pituitary Adrenal Axis (HPAA)

Stress >> brain >>

pituitary >> adrenal cortex >> glucocorticoids (cortisol) {15 seconds}

sympathetic nervous system >> adrenal **medulla** >> nor/ epinephrine (adrenaline)

Chronic stress = imbalance of sympathetic nervous system, impairs cognition, blood sugar, bone density, blood pressure, low immune system