Introduction to clinical epidemiology and research methodology

Chaiyawat Suppasilp, M.D.

Department of Clinical Epidemiology and Biostatistics

Faculty of Medicine Ramathibodi Hospital Mahidol University

August 2021

Content

- Introduction to Clinical Epidemiology

- Steps of Doing Research

- Measurements in Epidemiology

- Study design

- Evidence-based Medicine

Introduction to Clinical Epidemiology

- Definitions

- The purpose of Clinical Epidemiology

- Basic principles

Definition

Clinical Epidemiology = Clinical medicine + Epidemiology

Epidemiology

Epidemiology is the study (scientific, systematic, and data-driven) of
the <u>distribution</u> (frequency, pattern) and
<u>determinants</u> (causes, risk factors)
of <u>health-related</u> states and events (not just diseases)
in <u>specified populations</u> (neighborhood, school, city, state, country,
global).
It is also the application of this study to the <u>control</u> of health problems

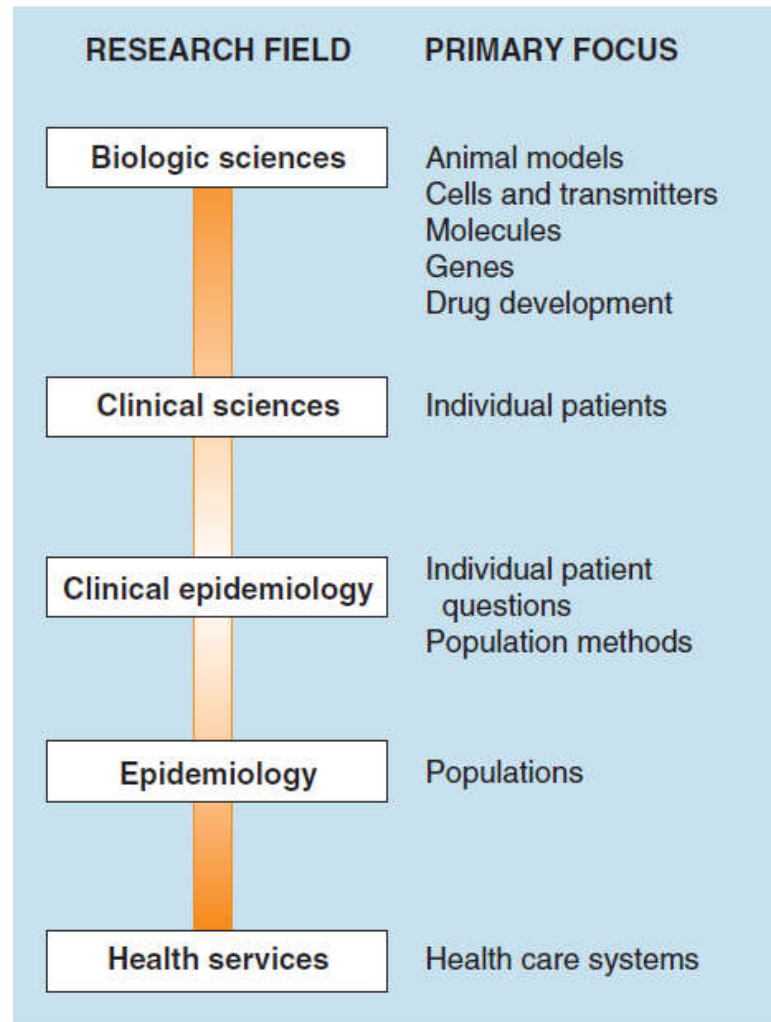(Last JM, editor. Dictionary of epidemiology. 4th ed. New York: Oxford University Press; 2001. p. 61.)

Definition

Clinical Epidemiology

Clinical epidemiology is the science of making predictions about <u>individual patients</u> by counting clinical events (the 5 Ds) in groups of similar patients and using strong <u>scientific methods</u> to ensure that the predictions are accurate.

Clinical Epidemiology is the <u>application of epidemiological methods</u> to the care of individual patients (to the practice of clinical medicine).

The health sciences and their complementary relationships.

RESEARCH FIELD	PRIMARY FOCUS
Biologic sciences	Animal models Cells and transmitters Molecules Genes Drug development
Clinical sciences	Individual patients
Clinical epidemiology	Individual patient questions Population methods
Epidemiology	Populations
Health services	Health care systems

Clinical epidemiology 5ed, Robert H Fletcher
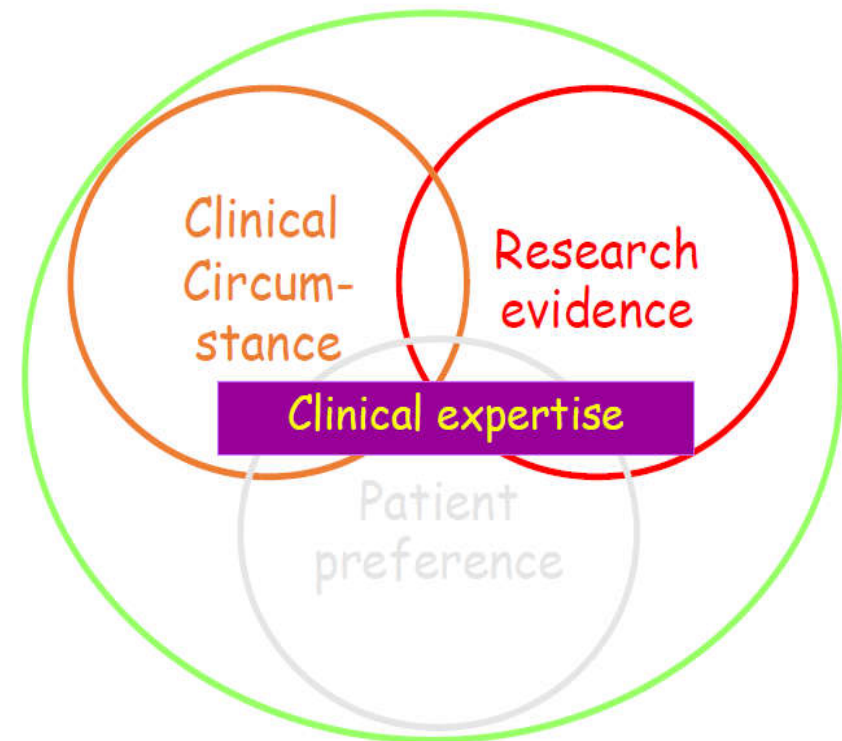
5

Definition

Evidence-based Medicine

"Expertise in integrating

    1. Best research evidence

    2. Clinical Circumstance

    3. Patient values

in clinical decisions"

(Haynes, Devereaux, & Guyatt, 2002)

Clinical Circum-stance

Research evidence

Clinical expertise

Patient preference

The purpose of Clinical Epidemiology

To foster methods of clinical observation and interpretation that lead to valid conclusions and better patient care.

Basic principles:

- Observations should address questions facing patients and clinicians

- Results should include patient-centered health outcomes (the 5 Ds)

- Etc.

Basic principles: Outcome of Disease

Outcomes of Disease (the 5 Ds)	
Death	A bad outcome if untimely
Disease	A set of symptoms, physical signs and laboratory abnormalities (the patient's experience of disease)
Discomfort	Symptoms such as pain, nausea, dyspnea, itching, and tinnitus
Disability	Impaired ability to go about usual activities at home, work, or recreation
Dissatisfaction	Emotional reaction to disease
Destitution	The financial cost of illness (for individual patients or society)

Clinical epidemiology 5ed, Robert H Fletcher

Basic principles: Variables

Variables

Attributes of patients and clinical event that vary and can be measured

Types of variables

- Independent variable: Predictor, purported cause

- Dependent variable: Outcome, possible effect

- Extraneous variable: Covariates, may affect the relationship between independent and dependent variable

Basic principles: Population & Sample
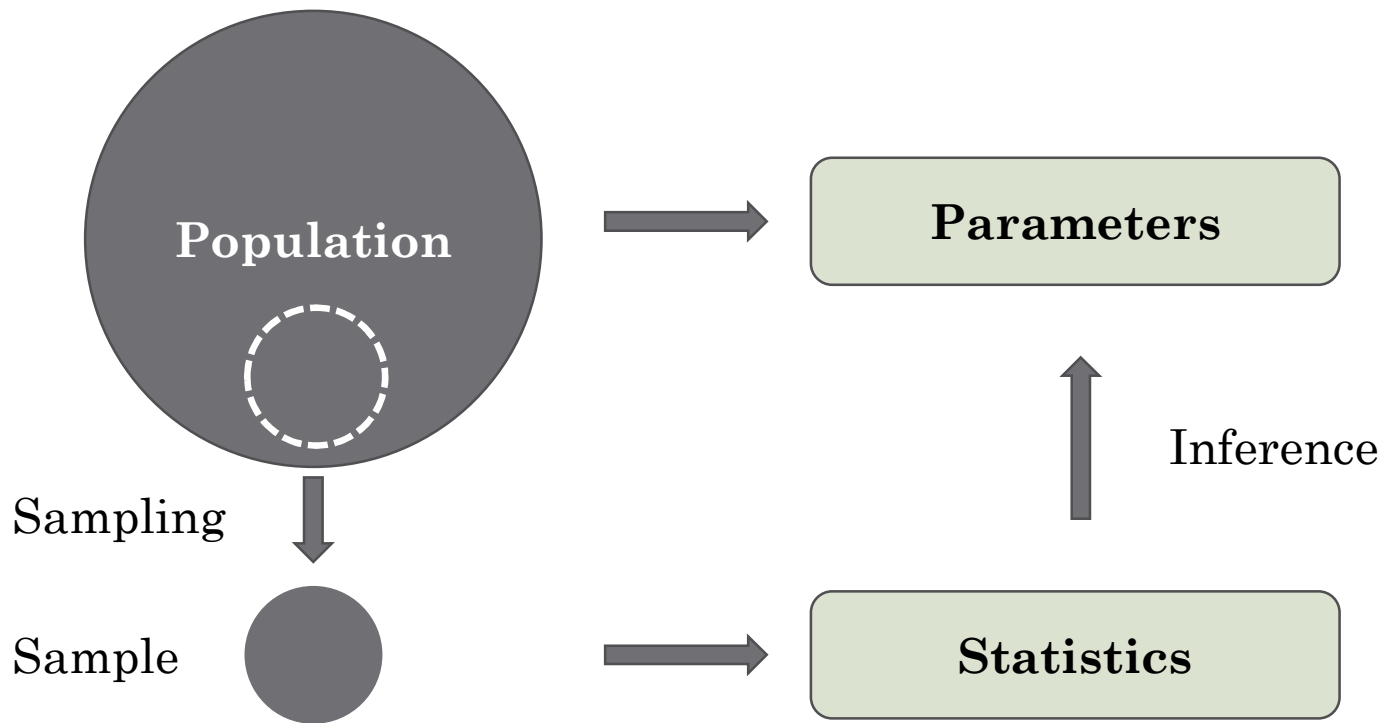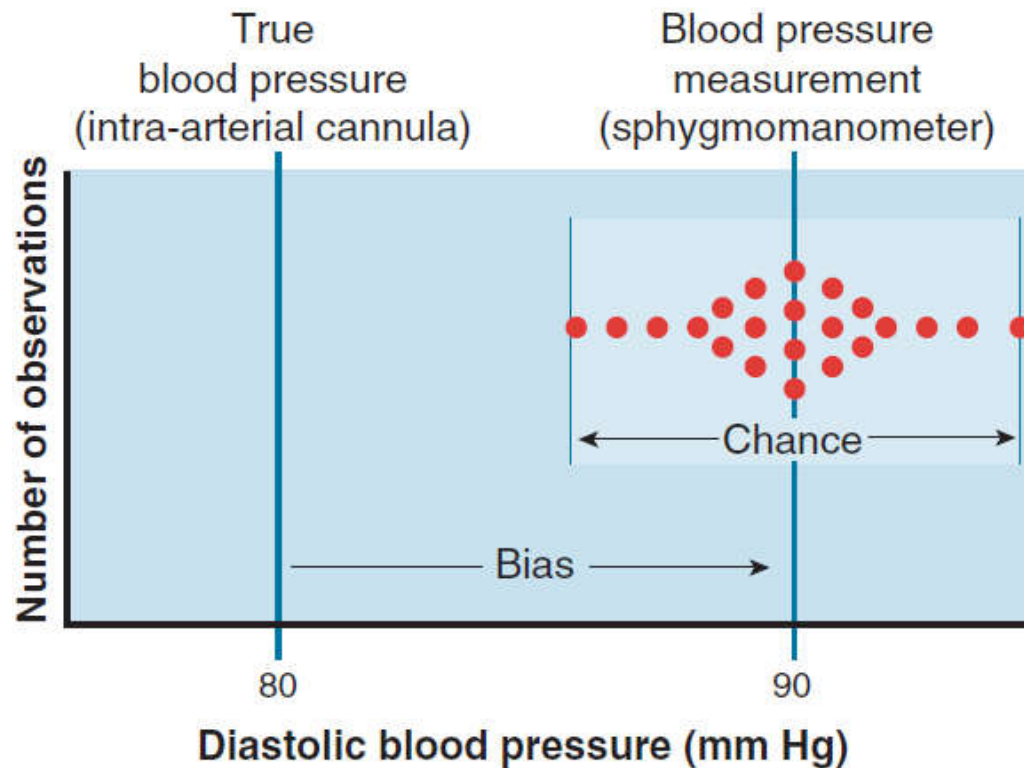
Basic principles: Error

The difference between the retained value and the true value.

Types of error

- Random error (by Chance)

- Systematic error (Bias)

> A <u>Process at any stages</u> of inference
> (Design, Conduct, Analysis, Interpretation, and Conclusion)
> tending to produce results that <u>depart systematically from true values</u>.

Basic principles: Error

True blood pressure (intra-arterial cannula)

Blood pressure measurement (sphygmomanometer)

Number of observations

Chance

Bias

80

90

Diastolic blood pressure (mm Hg)

Bias (Systematic error)

| Bias in Clinical Observation	
Selection bias	Occurs when comparisons are made between groups of patients that differ in determinates of outcome other than the one under study.
Measurement bias	Occurs when the methods of measurement are dissimilar among groups of patients
Confounding	Occurs when two factors are associated (travel together) and the effect of one is confused with or distorted by the effect of the other

Selection Bias

A nonrandom <u>selection of study participants</u> leads to erroneous conclusions or method or conduct to <u>absence of comparability</u> between groups being studied.

Common selection biases

- Berkson Bias (Hospital case differ than community population)

- Unmasking bias/Ascertainment bias/Surveillance bias

- Healthy worker effect (EGAT Good vs Poor)

- Volunteer Bias (Healthy or diseases sample e.g. MRI brain)

- Non-Response Bias (e.g. Questionnaire sexual issue, confidential issue)

Example 1: Berkson bias

It can arise when the sample is taken not from the general population, but from a subpopulation.
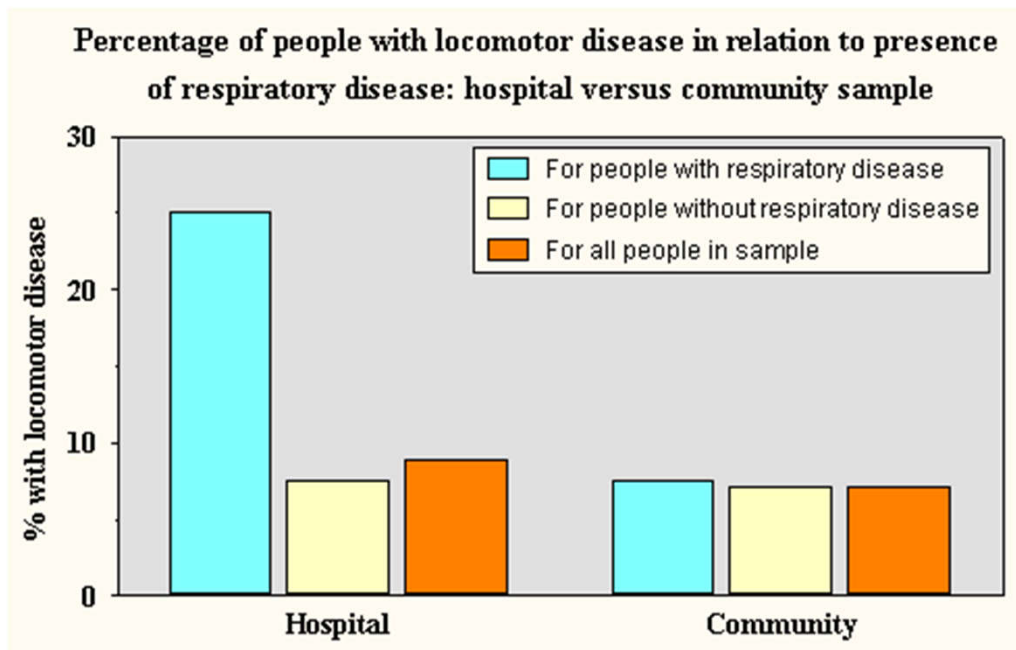
In hospital: OR = 4.06

In community: OR = 1.06

Sackett (1979)

Measurement Bias

- Occurs when the method of measurement leads to systematically incorrect results
- Incorrect determination of exposure or outcome, or both
- Gathering information in different way

Common measurement biases

- Recall bias

- Observer bias

- Attention bias (Hawthorn effect)

- Insensitive measurement bias

- Lead time bias

- Response bias

Example 2: Lead-time bias

Lead time bias refers to the phenomenon where early diagnosis of a disease falsely makes it look like people are surviving longer. This occurs most frequently in the context of screening.
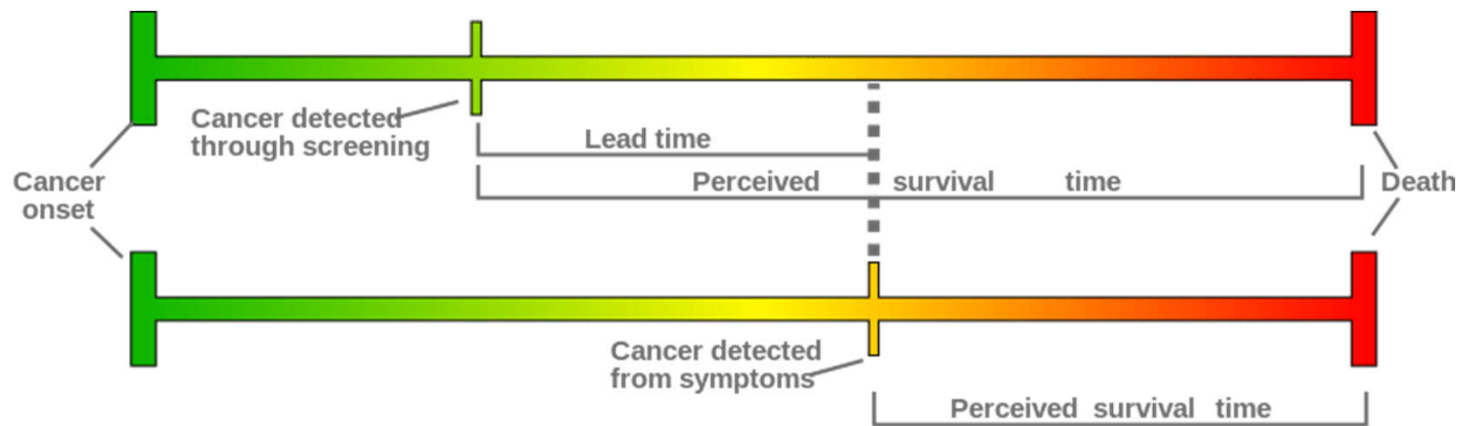
Confounding (Bias)

A confounding variable is <u>associated with the exposure</u> and it <u>affects outcome</u>, but it is **not** in an intermediate link <u>in the chain of causation</u> between exposure and outcome

- **A priority criteria of confounder**

Confounding Factor
Smoking

Study Factor
Hot Tea drinking

Outcome
Gastric cancer
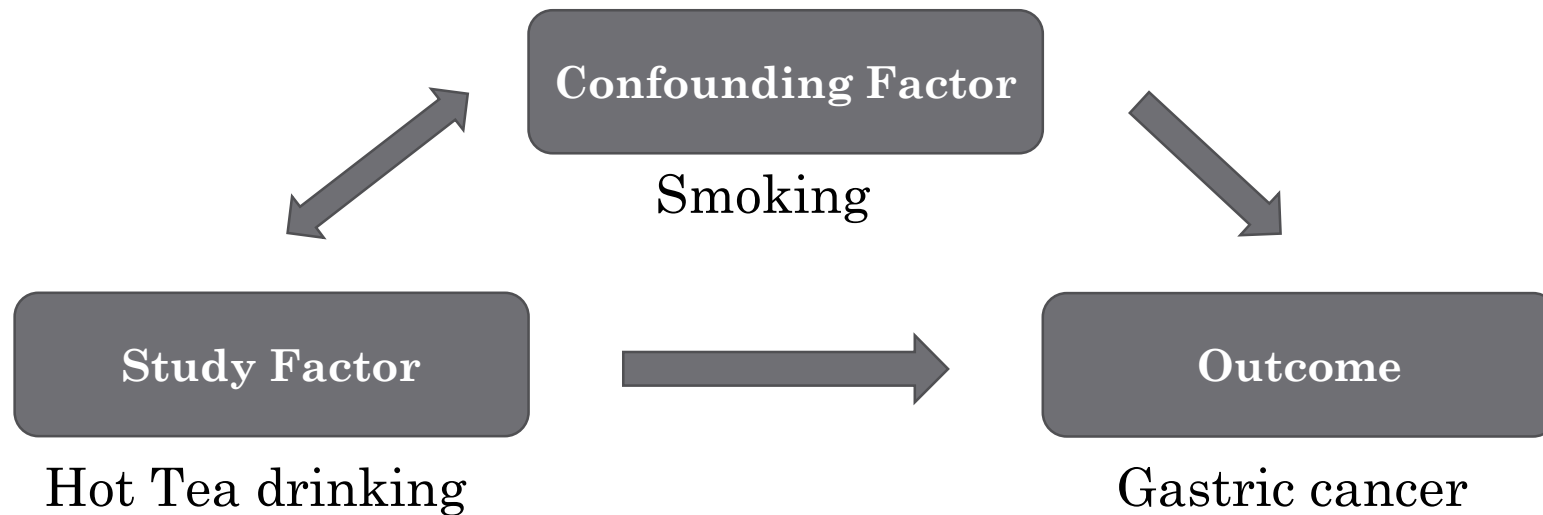
Example 3: Confounding

Cohort study of Hot tea drinking and Gastric cancer

	Tea +	Tea -	
CA	27	14	41
No CA	48	67	115
Total	75	81	156

Relative Risk = 2.1

Non-Smoker

	Tea +	Tea -	
CA Stomach	1	2	3
No CA	24	48	72
Total	25	50	75

Relative Risk = 1 **No association**

Smoker

	Tea+	Tea -	
CA Stomach	26	12	38
No CA	24	19	43
Total	50	31	81

Relative Risk = 1.23 **Confounder**

Confounding (Bias)

Control and minimized of confounder

Design Phase

- Exclusion (Restriction)
- Matching
- Randomization and blind assessed

Analysis Phase

- Adjustment (standardized)
- Stratification
- Multivariate Analysis
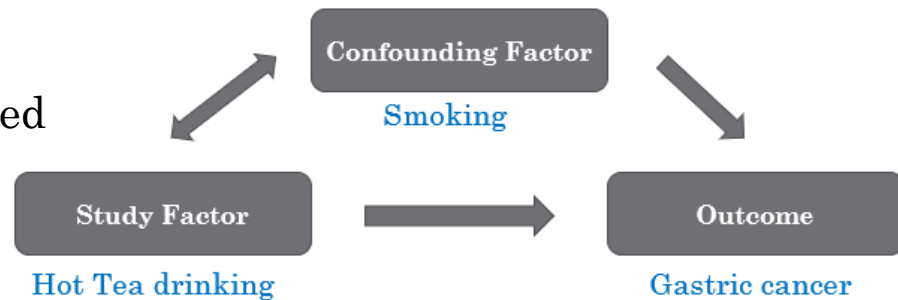
Random error (Chance)

A portion of variation in a measurement that has no apparent connection to any other measurement or variable, generally regarded as due to chance.

- Type I Error (Alpha): Rejecting the null hypothesis when it is true (False Positive)

- Type II Error (Beta): Accepting the null hypothesis when it is false (False Negative)

H_0: $u_{SBP(A)} = u_{SBP(B)}$
H_1: $u_{SBP(A)} \neq u_{SBP(B)}$

Statistical Testing	In Population Different exist (+)	No Different exist (-)
Different (+) Reject Null Hypothesis	Power 1-Beta	False Positive (alpha) Type I Error
No-Different (-) Accept Null Hypothesis	False Negative(Beta) Type II Error	True Negative 1-alpha

Random error (Chance)

Sources of error

- Small sample size

- High variation in Samples/Subjects

- Measurement-influenced errors
  - One-time measurement (or too many measurement)
  - Non-standardized measurement
  - Unreliable measurement (no calibration)

Random error (Chance)

How to assess random error

P-value

A numeric representative of the degree to which random variation alone could account for the difference observed between groups or data being compared e.g. $P < 0.05$, $P < 0.01$

Confidence interval

Provide a plausible range within which the true association lies and provide all the information in P-value and more.

Random error (Chance)

Strategies to reduce random error

- Appropriated sample size (Not largest sample size)

- Precise measurement endpoint

- Standardizing aspect of the protocol which impact on patient to patient variations

- Collecting data on key prognostic factors

- Choosing a homogenous group of patient

- Choosing the most appropriated design

Basic principles: Internal & External validity

- **Internal validity** is the degree to which the results of a study are <u>correct for the sample</u> of patients being studied.

- **External validity** (generalizability) is the degree to which the results of an observation hold <u>true in other settings.</u>

All patients with the condition of interest

INTERNAL VALIDITY

SAMPLE

Sampling

SAMPLE

Selection bias

Measurement and confounding bias

Chance

CONCLUSION

?

??

EXTERNAL VALIDITY (generalizability)

Clinical epidemiology 5ed, Robert H Fletcher

25
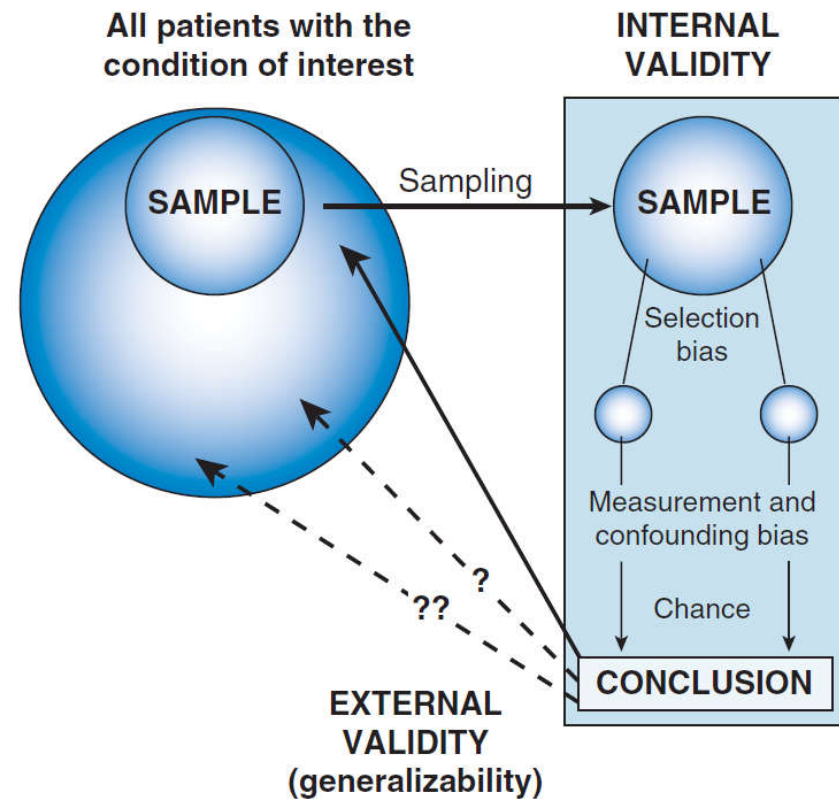
Steps of Doing Research

Steps of Doing Research

1. Research questions

2. Reviews and literature searching

3. Create study design:
   - Methodology, Sample size, Measurement etc.
   - Protocol wringing, Ethic submission

4. Perform Data correction
   - Select database program: Epidata, Excel, SQL
   - Design database and variables

5. Data management (entry, validation)

6. Statistical analysis

7. Results/ Conclusion (+/- publication)

Research question

- Good research question:
  - Relevant
  - Interesting
  - Focused

- Component of Research question
  - **P**opulation (Patients, Problems)
  - **I**ntervention (exposure)
  - **C**ontrol (comparison)
  - **O**utcome

Clinical scenario 1

- A 67-year-old woman was found to have HT 12 years ago.

- Her blood pressure is uncontrolled in last year of follow-up even 4 anti-hypertensive agents were prescribed.

- Her serum creatinine is 1.6 mg/dl (eGFR 33 cc/min/1.73m2).

- You wonder this patient may have renal artery stenosis and ischemic nephropathy.

- You ask your advisor and he suggested to perform MRA.

- But you think about CTA.

Research question 1

What is the accuracy of MRA compared with CTA in diagnosed of renal artery stenosis among uncontrolled hypertension population?

PICO

Patient

Intervention or Exposure

Comparison

Outcome

Measurements in Epidemiology

Measurement in Epidemiology

- Measures of disease frequency

- Measures of association

- Measures of potential impact

Measurement in Epidemiology

34
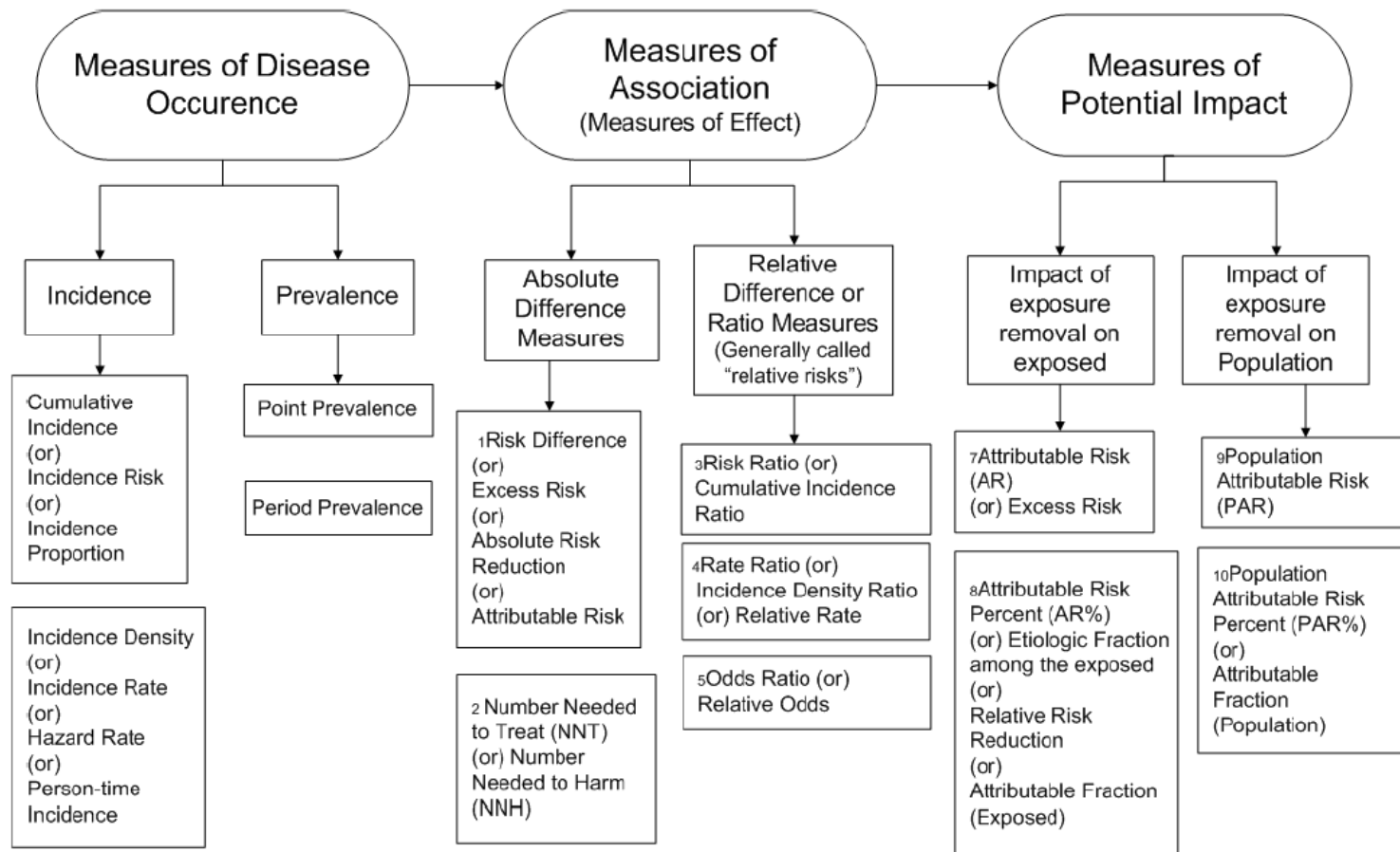
Measures of disease frequency

Measures of disease frequency in mathematical quantity

- Count

- Fraction
  - Rate
  - Ratio
  - Proportion (percentage)

Measures of disease frequency in epidemiology

- Prevalence

- Incidence

Count

- Simplest & most basic measure –absolute number of persons who have disease or characteristic of interest.

- Useful for health planners & administrators: for allocation of resources (e.g. quantity of ORS needed by diarrheal cases)

Limitations

- It depends on the size of the population at risk of the disease in an area. (e.g. the bigger group, the higher is the expected number of cases)

- The duration of observation also affects the frequency of cases. (e.g. the longer the observation period, the more cases can occur.)

Incidence and Prevalence

Incidence(I): Measures **new cases** of a disease that develop over a period of time.

- Cumulative incidence = incidence proportion (incidence)
- Incidence rate = incidence density rate (hazard rate)

Prevalence(P): Measures **existing cases** of a disease at a particular point in time or over a period of time.

- Point prevalence
- Period prevalence

Incidence proportion

$$\frac{\text{Number of new cases of disease or injury during specified period}}{\text{Size of population at start of period}}$$

In 2001, among 5,572 women aged 20-39 years who were sex workers, there were 45 HIV positive cases during 2002-2005

SO, The cumulative incidence of HIV positive cases during these 4 years

=

Incidence rate

Hypothetical cohort of 12 initially disease-free subjects followed over a 5-year period from 1990 to 1995.

x: disease   O: death

$$\hat{IR} = \frac{I}{PT} = \frac{5}{25\ PY} = 0.20$$

39
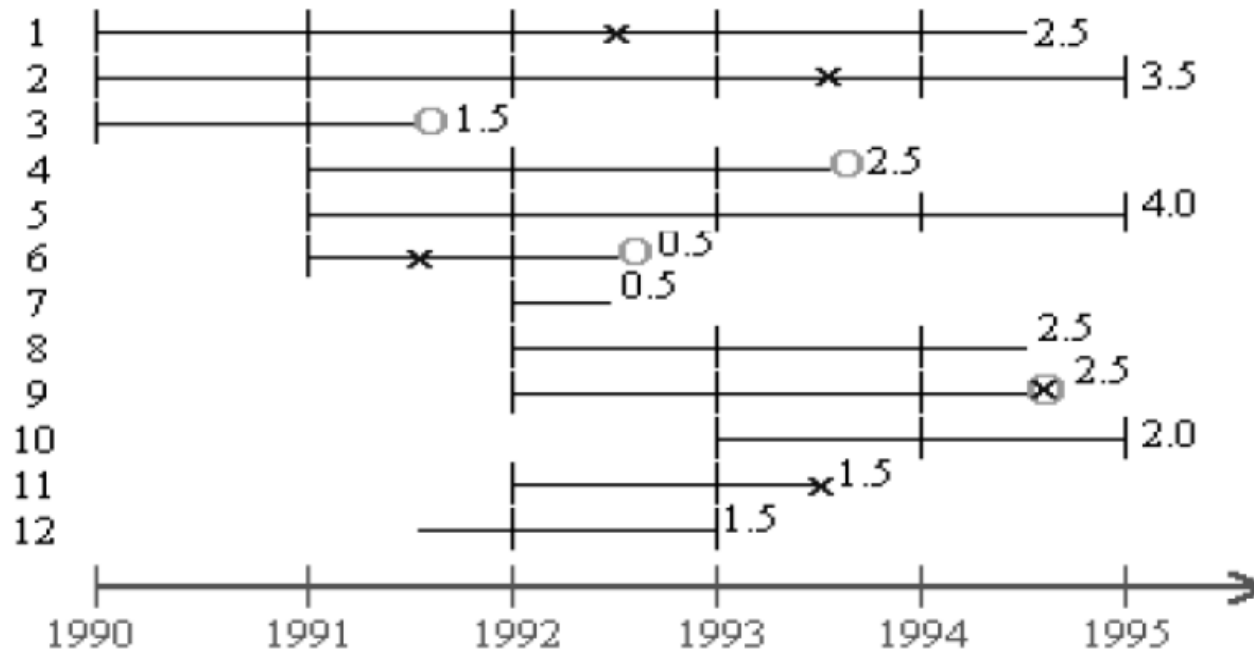
Prevalence

$$\frac{\text{All new and pre-existing cases during a given time period}}{\text{Population during the same time period}}$$

In a survey of 1,150 women who gave birth in Ramathibodi hospital in 2000, a total of 468 reported taking a multivitamin at least 4 times a week during the month before becoming pregnant.

The prevalence of frequent multivitamin use in this group

=

Prevalence and Incidence

- Incidence measure appearance of the disease
  Prevalence measure the existing of the disease

- Incidence means NEW
  Prevalence means ALL

- Prevalence = Incidence X duration
  (the relationship is most apparent in a stable, chronic disease)

- Prevalence is based on both incidence and duration of illness. High prevalence of a disease within a population might reflect high incidence or prolonged survival without cure or both. Conversely, low prevalence might indicate low incidence, a rapidly fatal process, or rapid recovery.

41
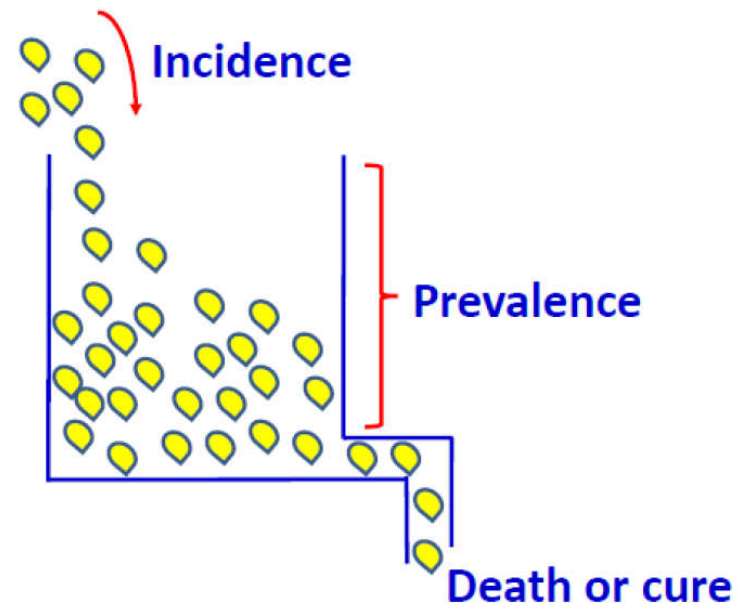
Example:

- From total of 1000 cases, there are 20 and 10 new cancer cases detected in the first and second year, respectively.

- Find the cumulative incidence of the first years

- Find the cumulative incidence of the second years

- Find the cumulative incidence of these two years

Measures of association

- Absolute Difference
  - Risk Difference or
    Excess Risk or
    Absolute Risk Reduction or
    Attributable Risk
  - Number needed to treat (NNT)
  - Number needed to harm (NNH)

- Relative Difference (Ratio measurement or Relative risk)
  - Risk Ratio (Relative Risk)
  - Odds Ratio (Relative Odds)

43

Risk vs Odds
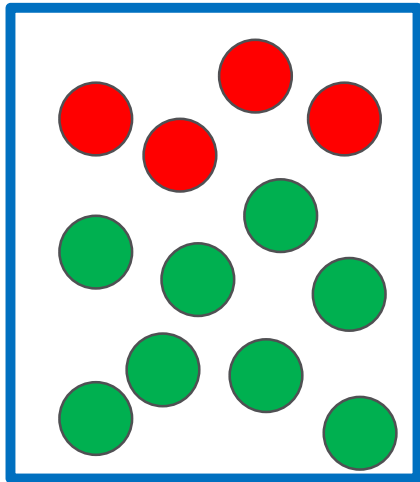
$$Odds = \frac{P(event)}{1 - P(event)}$$

$$Risk = \frac{number\ of\ cases\ of\ disease}{number\ of\ people\ at\ risk}$$

$$Risk(Red) =$$

$$Risk(Green) =$$

$$Odds = \frac{Chance\ that\ event\ occurs}{Chance\ that\ event\ dose\ not\ occur}$$

$$Odds(Red) =$$

$$Odds(Green) =$$

Risk ratio (RR) vs Odds Ratio (OR)

$$Risk\ Ratio = \frac{Risk\ of\ disease\ in\ Exposed, Treatment\ Group}{Risk\ of\ disease\ in\ Non-exposed, Control\ Group}$$

$$Odds\ Ratio = \frac{Odds\ of\ disease\ in\ Exposed, Treatment\ Group}{Odds\ of\ disease\ in\ Non-exposed, Control\ Group}$$

Example

	Disease	No Disease	Total
Smoking	100	1,900	2,000
Non-smoking	80	7,920	8,000
Total	180	9,820	10,000

- Risk of having the disease in Exposed Group =

- Risk of having the disease in Non-exposed Group =

- Risk Ratio (Relative Risk) =

- Meaning:
  = The risk of developing the disease in smoking group is 5 times of non-smoking group
  = The risk of developing the disease in smoking group is 4 times higher than non-smoking group

Example

	Disease	No Disease	Total
Smoking	100	1,900	2,000
Non-smoking	80	7,920	8,000
Total	180	9,820	10,000

- Odds of having the disease in Exposed Group =

- Odds of having the disease in Non-exposed Group =

- Odds Ratio (Relative Odds) =

- Meaning:
  = The odds of developing the disease in smoking group is 5.2 times of non-smoking group
  = The odds of developing the disease in smoking group is 4.2 times higher than non-smoking group

OR vs RR

- If the interested event is rare (or very rare), OR and RR are similar.
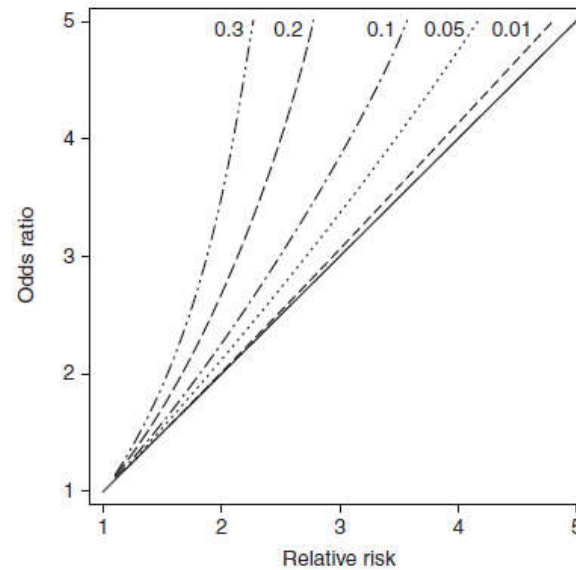
Figure 3.1. The odds ratio against the relative risk. The fanned lines are each drawn at a specific value (as so labelled) of the risk of disease in subjects who are unexposed to the risk factor. The diagonal solid line is the line of equality between the odds ratio and the relative risk.

Why we use Odds??

- Multiple logistic regression, a frequently used multivariate technique, calculates adjusted ORs and not RRs.

- The odds ratio would be substantially larger than the relative risk.

- Meta-analysis: each study has their own prevalence, using Odds is better.

- In retrospective (case-control) studies, where the total number of exposed people is not available, RR cannot be calculated, and OR is used as a measure of the strength of association between exposure and outcome.

- By contrast, in prospective studies (cohort studies), where the number at risk (number exposed) is available, either RR or OR can be calculated.

Measure of potential impact

- Reflects the burden that an exposure contribute to the frequency of disease in the population

- Impact of exposure removal

- Two concepts
  - Attributable risk among exposed
  - Population attributable risk

Attributable Risk (AR)

- Quantifies disease burden in exposed group attributable to exposure.

- Provides answer to
  - What is the risk which can be attributed to the exposure?
  - What is the excess risk due to the exposure?

- Calculated as risk difference (RD)

Incidence

$I_{exposed} - I_{unexposed}$
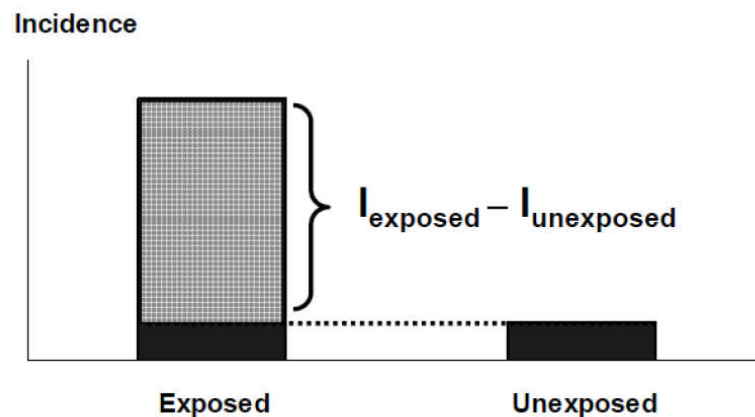
Exposed   Unexposed

I = Incidence

Population Attributable Risk (PAR)

- Excess risk of disease in total population attributable to exposure.

- Reduction in risk which would be achieved if population entirely unexposed.

- Helps determining which exposures relevant to public health in community.

- PAR =AR*Prevalence of exposure in the population

Risk

$$I_{population} - I_{unexposed}$$

Population   Unexposed
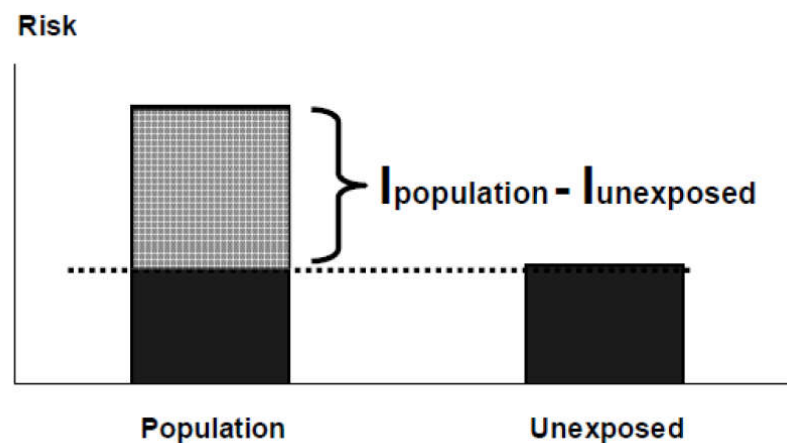
Attributable Risk fraction

- Attributable risk in the exposed group

$$AR\% = \frac{I_{exposed} - I_{unexposed}}{I_{exposed}}$$

- Attributable risk in the total population

$$PAR\% = \frac{I_{population} - I_{unexposed}}{I_{population}}$$

Example

- Consider a cohort study of risk of ischemic stroke, taken in 1 year, with 500 subjects with atrial fibrillation (AF) controlled against 500 subjects without AF.

- Given the proportion of AF in general population is 30%.

- The results are summarized as follow:

	Ischemic stroke: present	Ischemic stroke: absent	Total
AF	2	498	500
No AF	1	499	500
Total	3	997	1000

Study design

Types

By study design

- Experimental, Observational study
- Descriptive, Analytic study

By category of clinical question

- Risk
- Prognosis
- Therapy
- Diagnosis

Classification of study design

Observational study

• Descriptive or case series

• Cross-sectional study

• Case control studies(retrospective)

• Cohort studies(prospective)

• Historical cohort studies (retrospective)

Experimental study

• Controlled trials

• Studies with no controls

Systematic Reviews/Meta analysis

Hierarchy of evidences

Systematic Reviews

Randomized Controlled Trial (RCT)

Cohort Study

Case-Control study

Cross-Sectional study

Case Report
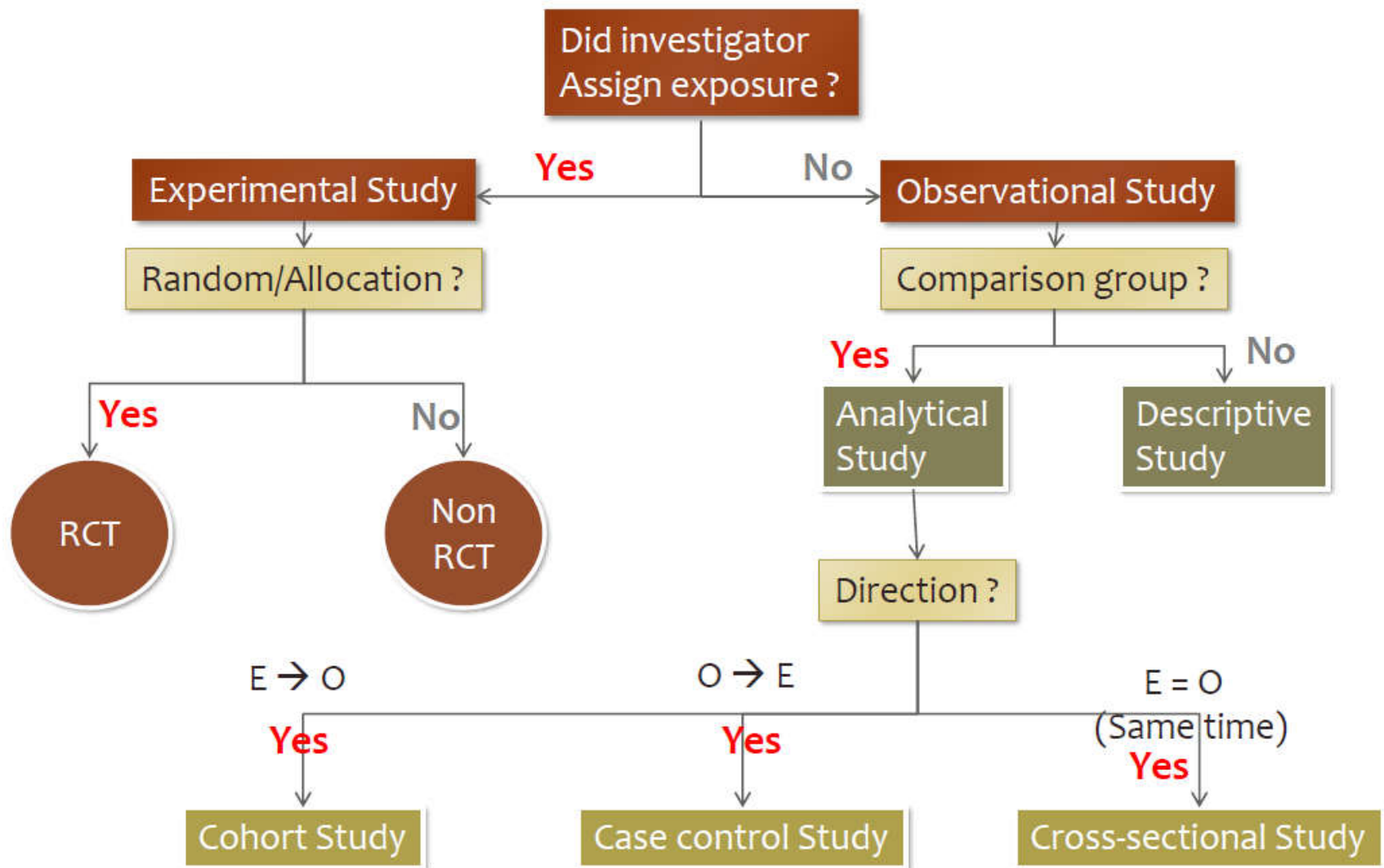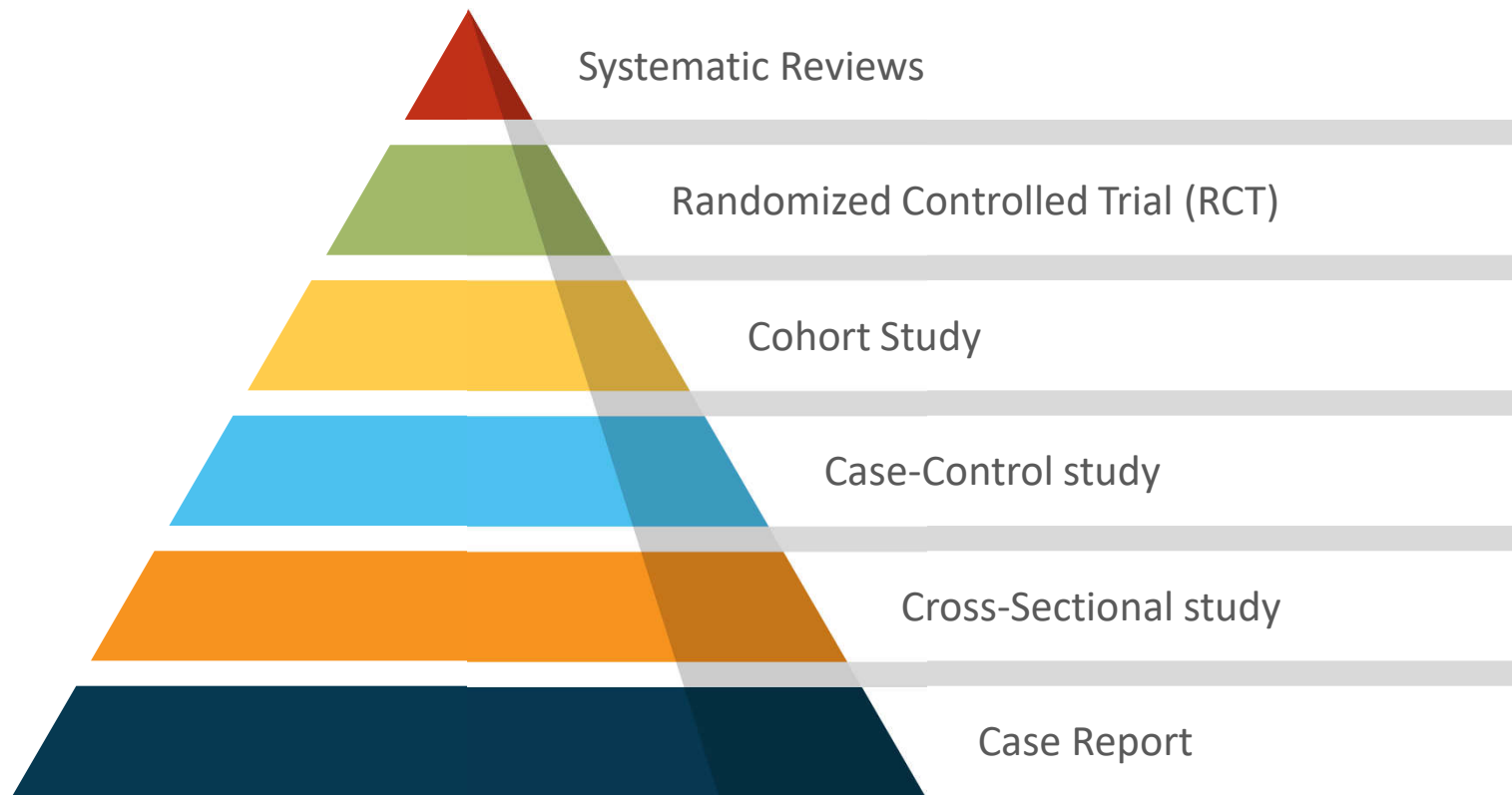
Cross-Sectional study

- Study about the characteristics of a population at one point in time
  (like a photo "snapshot")
  (also for a short period of time)

- No intervention

- Usually a prevalence study (Not incidence study)

- Hardly to control bias

- Can establish **association <u>but NOT "causation"</u>**

Present
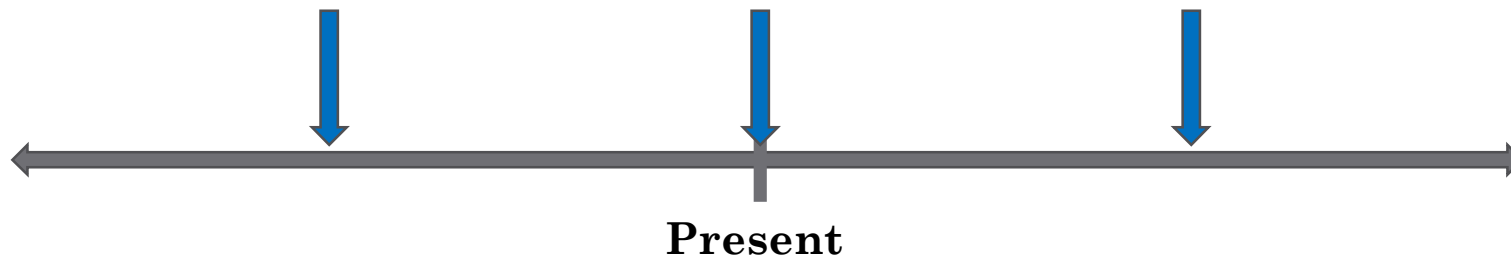
Cross-Sectional study

Advantages	Disadvantages
Simple (no follow-up)	Can establish **association but NOT "causation"**
Inexpensive	Hardly to control bias: • Confounding • Recall bias • Incidence-prevalence bias
No drop out	
No intervention/exposure	

Case-Control study

- Study Subjects Participants <u>selected based on outcome status</u>
  - Case-subjects have outcome of interest
  - Control-subjects do not have outcome of interest
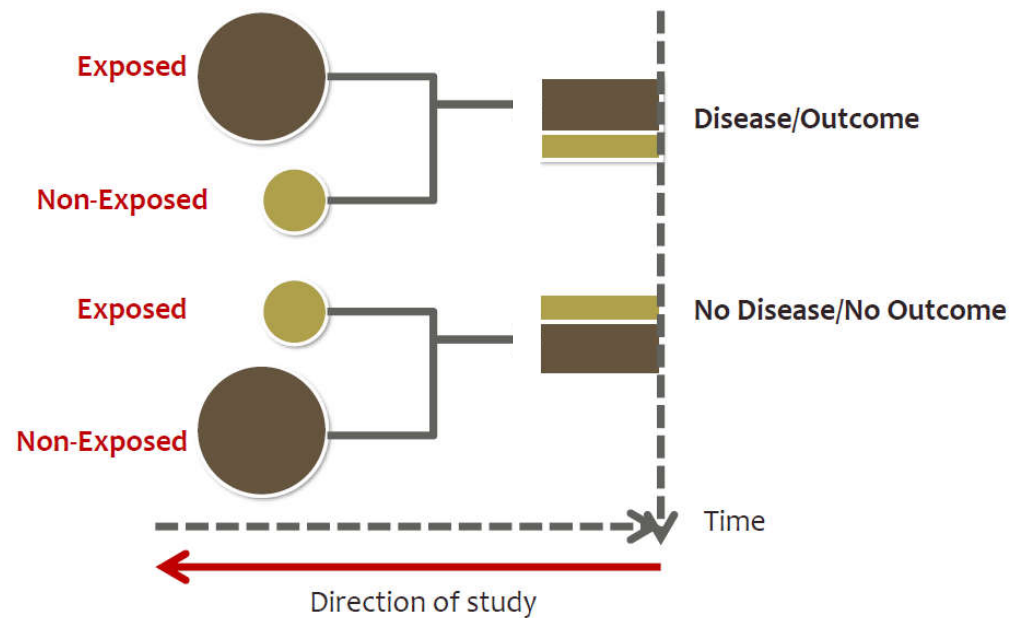
Case-Control study

- Study Subjects Participants selected based on outcome status
  - Case-subjects have outcome of interest
  - Control-subjects do not have outcome of interest

- Usually Retrospective study

- Used for studying rare diseases

- For multiple exposures that may be related to a single outcome

- Use Odds ratio (OR) Not Relative risk (RR)
  (in rare disease, it is ok to use RR, will be discuss later)

Case-Control study

Advantages	Disadvantages
Quickly and inexpensive	Recall Bias
Feasible for rare disorder or long follow-up	More effect of confounder
May required fewer subjects	Difficult to find control group
	OR is not easy to understand by physicians.

Cohort Study

- Participants <u>classified according to exposure</u> status and <u>followed-up</u> over time to ascertain outcome.
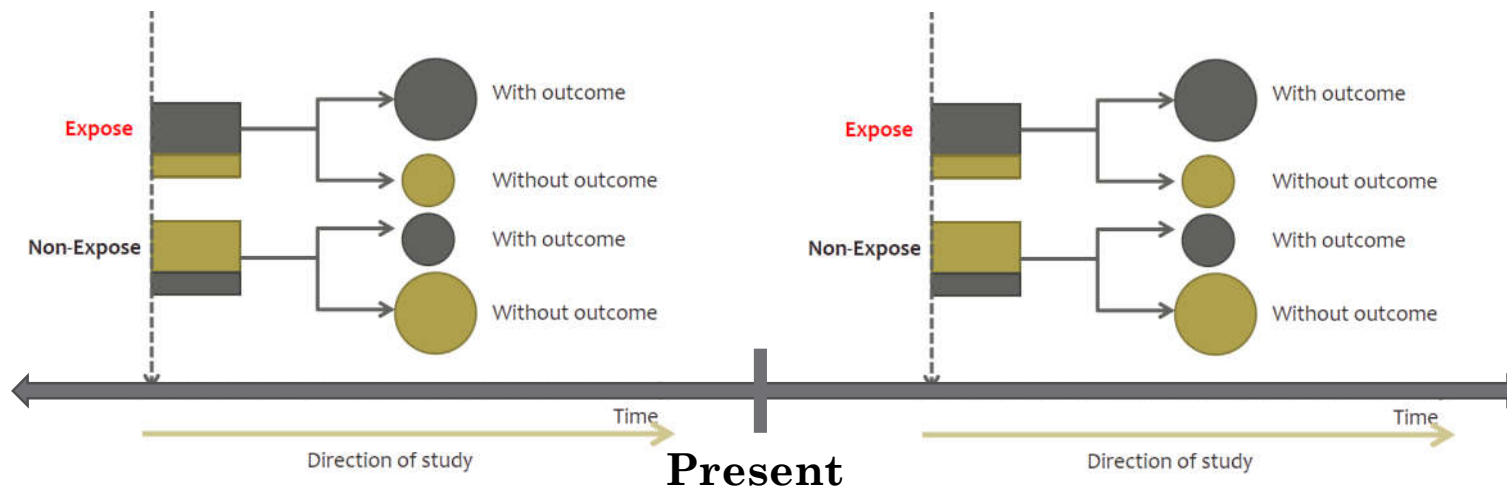
- Ensures temporality (exposure occurs before observed outcome)
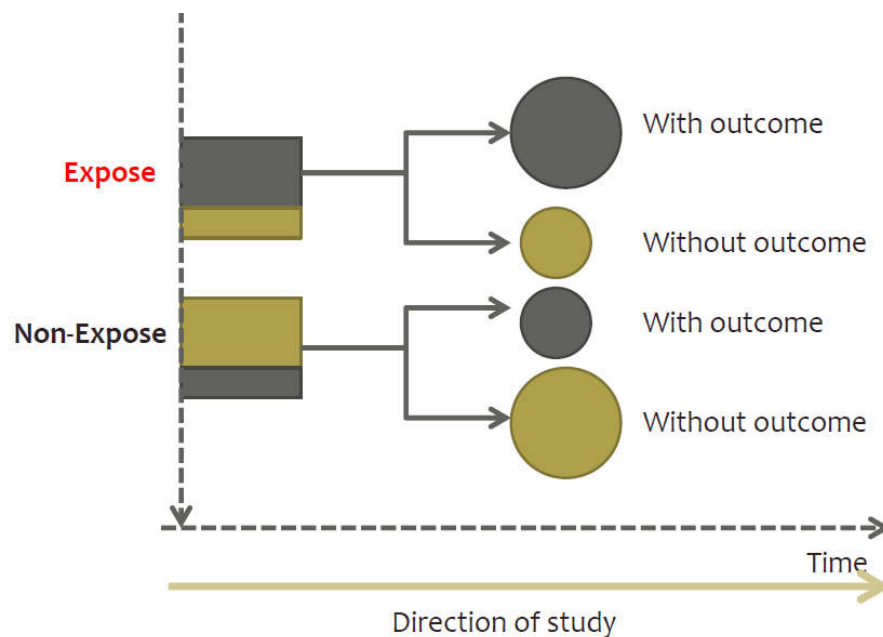
Cohort Study

Types of cohort study

Historical cohort studies (retrospective) **Cohort studies(prospective)**

Cohort Study

Advantages	Disadvantages
Can be matched	Relatively expensive
Can establish temporal association	Hard to blind
Can be standardized in eligible criteria & outcome assessment	Long follow-up period for rare disorder
	Difficult to find controls and confounders
	Follow-up need manpower

Randomized Controlled Trial (RCT)

- Give intervention/ exposure/ treatment to participants

Randomized Control Trial
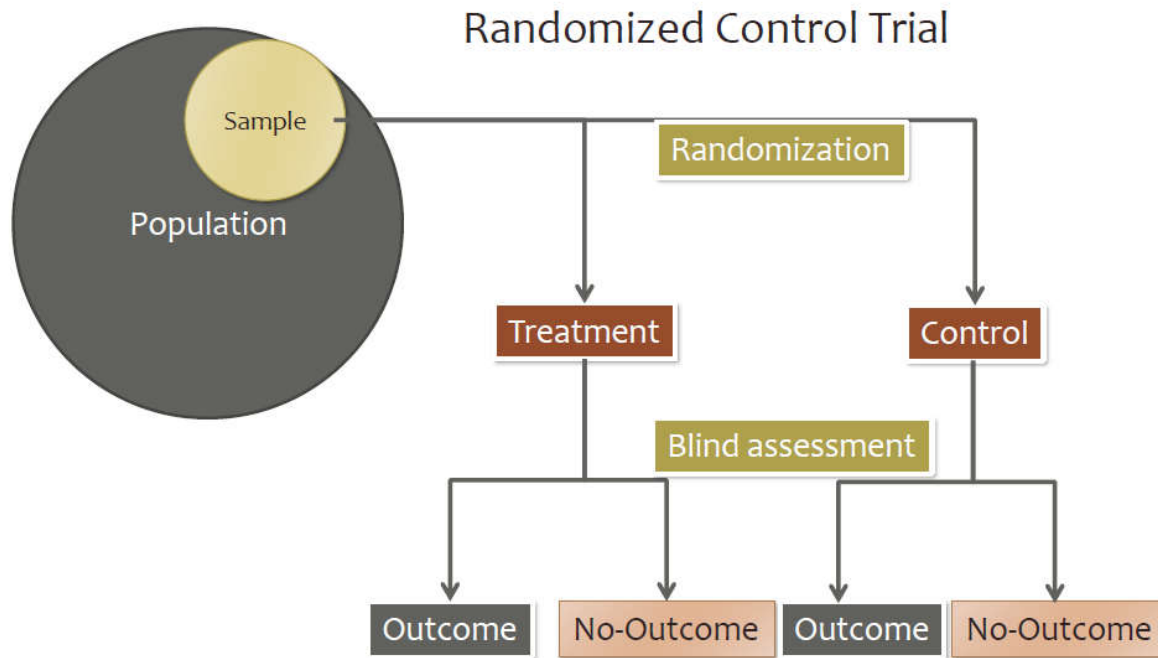
Randomized Controlled Trial (RCT)

Advantages	Disadvantages
Confounding and variables can be balance by randomization	High cost in term of time and money
Blinding of subjects, medical staff and investigators are achievable (Selection bias is minimized, but volunteer bias is <u>not</u> minimized)	Dropout or loss to follow up are common event
Considered to be the best sort of experimental study	Need time to final results

Why do not always perform RCT?

Hierarchy of evidences

- Systematic Reviews
- Randomized Controlled Trial (RCT)
- Cohort Study
- Case-Control study
- Cross-Sectional study
- Case Report
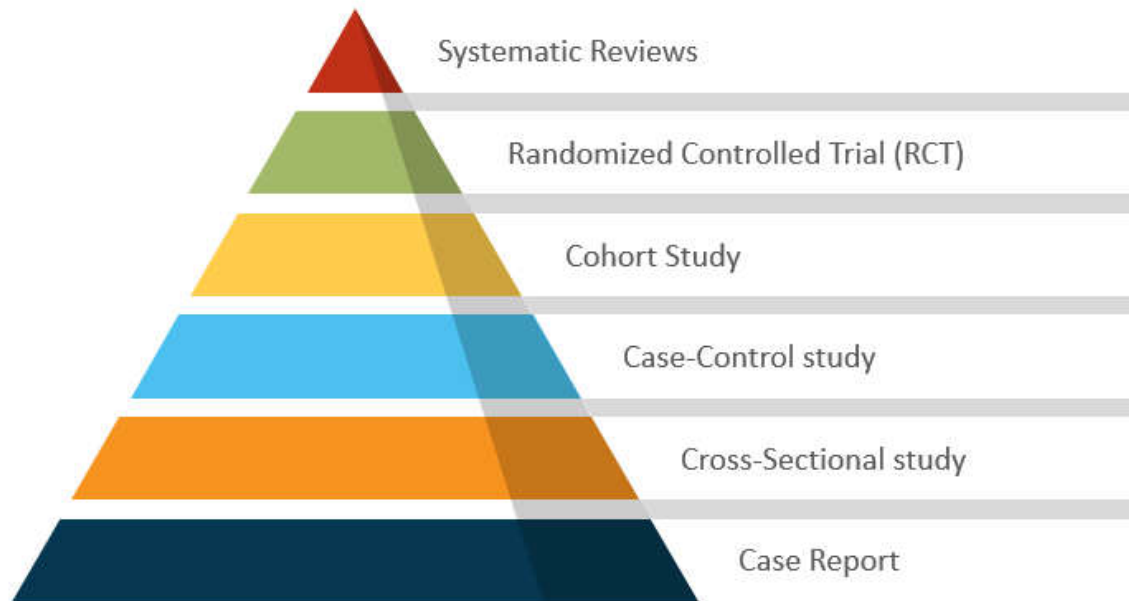
Example 01

- A study has been conducted to assess an effect of alcohol consumption on the risk of liver cancer using data registry of health of workers of one company between 2000-2010.

- Alcohol consumption was collected at baseline and follow up until the occurrences of liver cancer documented in medical records.

- What is an appropriate study design?
  - A cross-sectional study
  - A prospective cohort study
 - A retrospective cohort study
  - A case-control study
  - A clinical trial

Example 02

- 120 women with endometrial cancer and another 480 women with no apparent disease were contacted and asked whether they had ever used estrogen. Each woman with cancer was matched by age, race, weight, and parity to a woman without disease.

- What is an appropriate study design?
  - A cross-sectional study
  - A prospective cohort study
  - A retrospective cohort study
  - A case-control study
  - A clinical trial

OR vs RR

- In retrospective (case-control) studies, where the total number of exposed people is not available, RR cannot be calculated, and OR is used as a measure of the strength of association between exposure and outcome.

- By contrast, in prospective studies (cohort studies), where the number at risk (number exposed) is available, either RR or OR can be calculated.

case-control

cohort studies

73
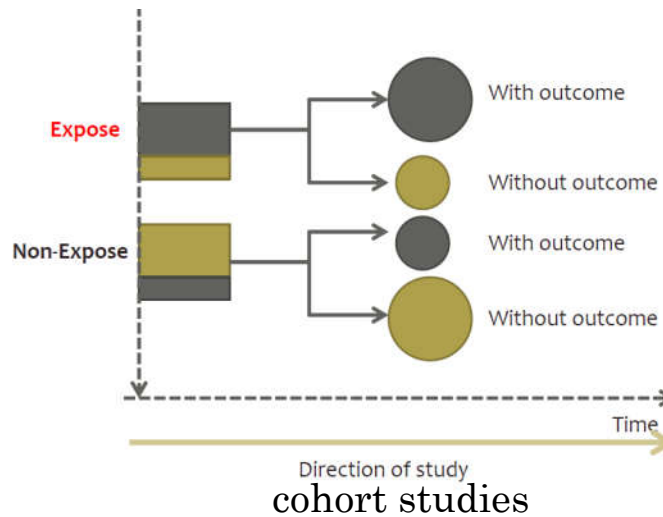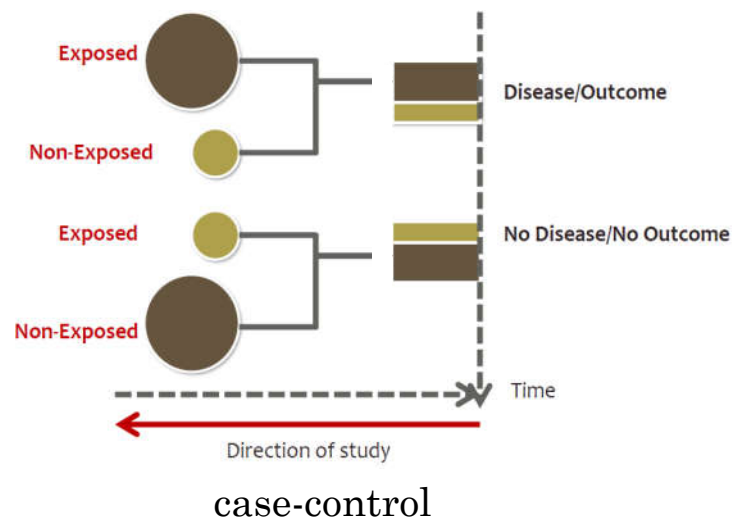
Clinical question and Study design

- Diagnosis Demonstrate that new diagnosis test is valid/reliable, preferred "cross sectional study"

- Causation or Risk Determine that agent is related to development of illness, preferred "Cohort or case-control study"

- Therapy Testing the efficacy of intervention preferred "RCT"

- Prognosis determine what happen to someone with some stage of disease, preferred "Prospective Cohort study"
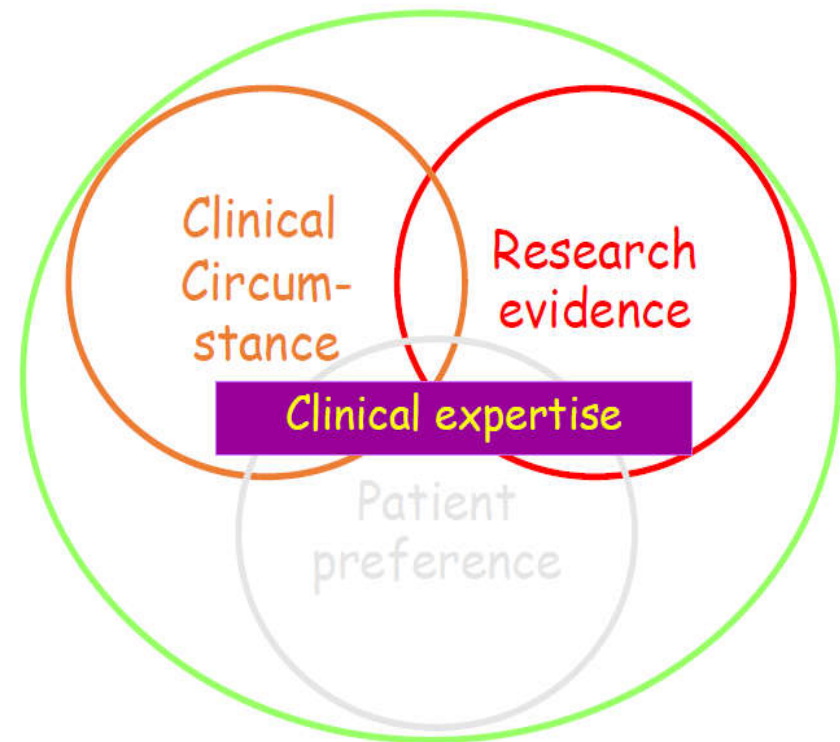
Evidence-based Medicine

Evidence-based Medicine

"Expertise in integrating

      1. Best research evidence

      2. Clinical Circumstance

      3. Patient values

in clinical decisions"

(Haynes, Devereaux, & Guyatt, 2002)
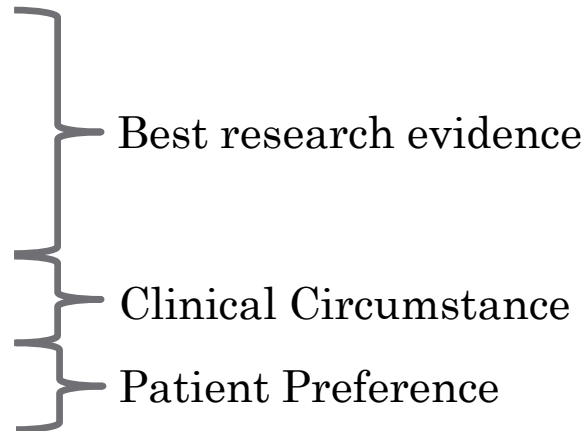
How do we actually practice EBM?

5 A's of EBM

Step 1: Ask answerable question

Step 2: Find Articles

Step 3: Critical Appraisal the evidence

Step 4: Apply

Step 5: Assess patient preference

Best research evidence

Clinical Circumstance

Patient Preference

Critical appraisal

Board issues need to be considered.

- Are the results of the study valid?

- What are the results?

- How can you apply the results to patient care?

Critical appraisal: Tools

CEBM

(The Centre for Evidence-Based Medicine develops, promotes and disseminates better evidence for healthcare.)

mnemonics

Question: PICO

Method: RAMMbo

- Recruitment
- Allocation
- Maintenance
- Measurements

Author:			Ref:	
Description			**Numbers**	

		Description	CER (%)	IER (%)
Patients				
Intervention				
Comparator				
Outcomes 1				
Outcomes 2				

Randomized		
Ascertainment		
Measures		

RDifference	CER – EER	
RRR	RD/CER	
NNT	1/RD	

Clinical Bottom-line:

Further Actions:

Appraisal checklist -RAMMbo

Was the Study valid?

Recruitment: Who did the subjects represent?

Allocation: Was the assignment to treatments randomized?
Were the groups similar at the trial's start?

Maintenance: Were the groups treated equally?
Were outcomes ascertained & analyzed for most patients?

Measurements: Were patients and clinicians "blinded" to treatment? OR
Were measurements objective & standardized?

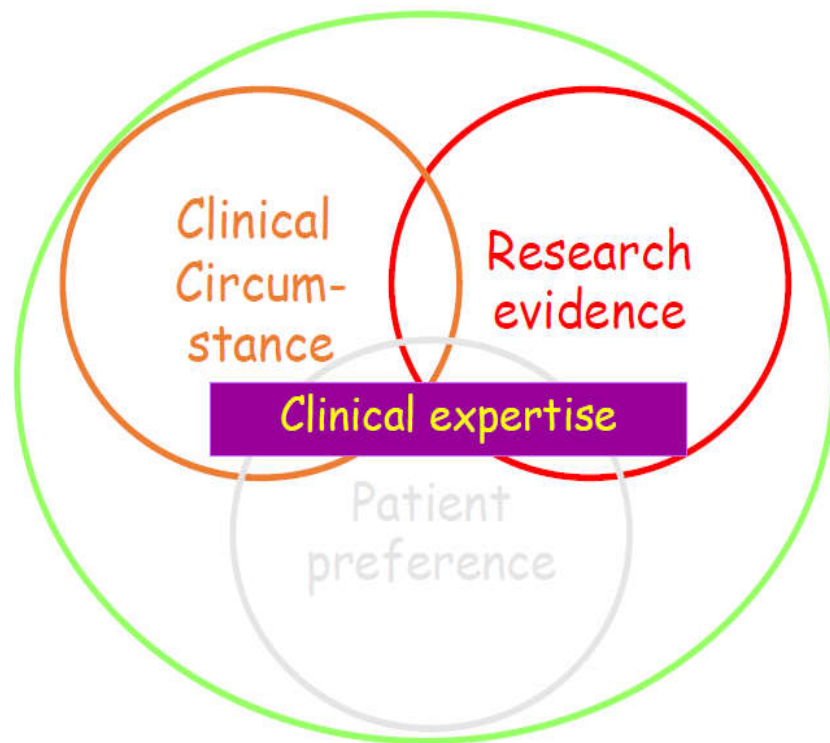Study statistics (p-values & confidence intervals)

Critical appraisal: Tools

CASP
Critical Appraisal
Skills Programme

www.casp-uk.net

info@casp-uk.net

Summertown Pavilion, Middle
Way Oxford OX2 7LG

How to use this appraisal tool: Three broad issues need to be considered when appraising a trial:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

https://casp-uk.net/casp-tools-checklists/

EBM is lifelong learning process

Clinical
Circum-
stance

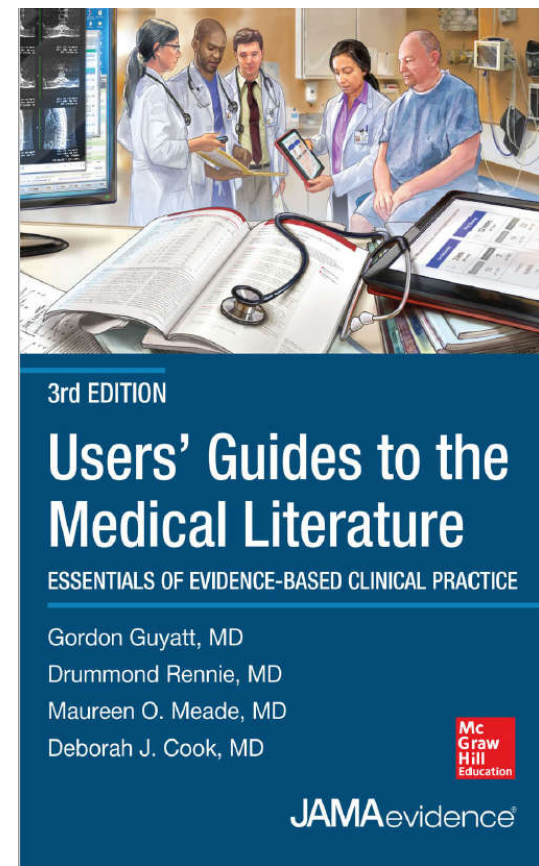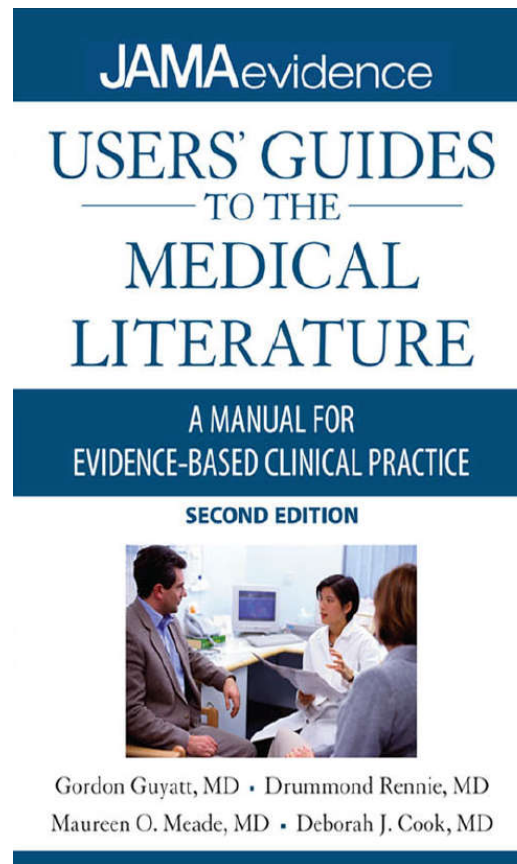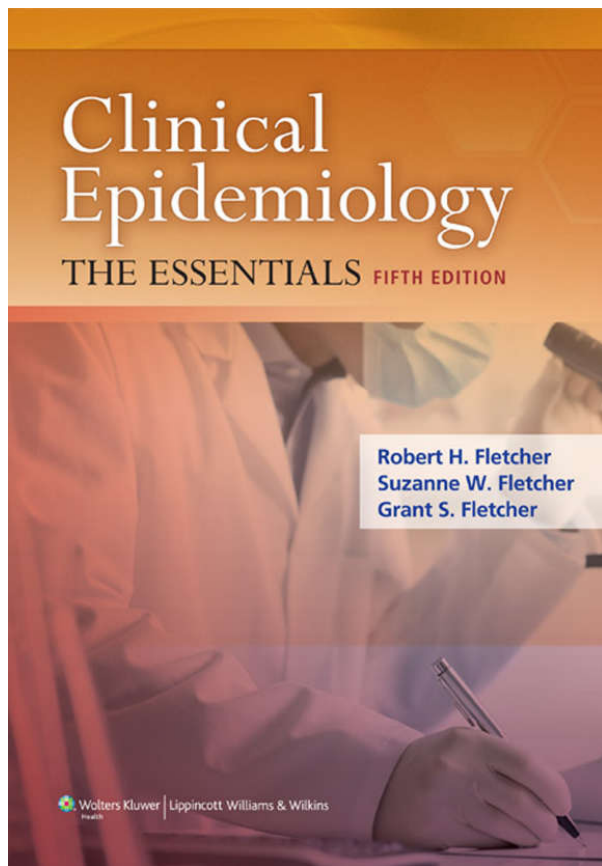Research
evidence

Clinical expertise

Patient
preference

Dr. Gordon Guyatt

Acknowledgement

- Prof.Ammarin Thakkinstian, Ph.D.

- Assoc.Prof. Atiporn Ingsathit, M.D., Ph.D.

- Asst.Prof. Pawin Numthavaj, M.D., Ph.D.

- Aj. Chusak Okascharoen, MD, PhD

- Aj. Kunlawat Thadanipon, M.D., Ph.D.

References

Thank you

Chaiyawat Suppasilp