UNIVERSITY OF KWAZULU-NATAL ™

INYUVESI YAKWAZULU-NATALI

Intro. to Study Designs & Observational Descriptive study design

Themba Ginindza, PhD, MSc, MPH

Professor/Epidemiologist/MLCCP Lead PI/CCPAC PI
Director: Cancer & Infectious Diseases Epidemiology Research Unit (CIDERU)
Department of Public Health Medicine, University of KwaZulu-Natal
Tel: +27 31 260 4214 | Mob:+27 719 1111 79
Email: Ginindza@ukzn.ac.za

Entire class fails

[EDITOR'S NOTE: I found this in my news feed on one of the social media sites and thought how practical a lesson this would be. Even I would have learned from it as dense as I am].

TEACHER FAILS ENTIRE CLASS: An economics professor at a local college made a statement that she had never failed a single student before, but had recently failed an entire class. That class had insisted that socialism worked and that no one would be poor and no one would be rich, a great equalizer.

The professor then said, "OK, we will have an experiment in this class on this plan". All grades will be averaged and everyone will receive the same grade so no one will fail and no one will receive an A.... (substituting grades for dollars - something closer to home and more readily understood by all).

After the first test, the grades were averaged and everyone got a B. The students who studied hard were upset and the students who studied little were happy. As the second test rolled around, the students who studied little had studied even less and the ones who studied hard decided they wanted a free ride too so they studied little.

The second test average was a D! No one was happy. When the 3rd test rolled around, the average was an F.

As the tests proceeded, the scores never increased as bickering, blame and name-calling all resulted in hard feelings and no one would study for the benefit of anyone else.

To their great surprise, ALL FAILED and the professor told them that socialism would also ultimately fail because when the reward is great, the effort to succeed is great, but when government takes all the reward away, no one will try or want to succeed.

These are possibly the 5 best sentences you'll ever read and all applicable to this experiment:

1. You cannot legislate the poor into prosperity by legislating the wealthy out of prosperity.

2. What one person receives without working for, another person must work for without receiving.

3. The government cannot give to anybody anything that the government does not first take from somebody else.

4. You cannot multiply wealth by dividing it!

5. When half of the people get the idea that they do not have to work because the other half is going to take care of them, and when the other half gets the idea that it does no good to work because somebody else is going to get what they work for, that is the beginning of the end of any nation.

***

Outline

- Introduction
- Elements of study designs
- Selection of research design
- Validity
- Study designs
- Types of Epidemiological study designs
  - Analysis
  - Strength & Weakness

**"It is much more important to know what sort of a patient has a disease,
than what sort of disease a patient has"**

(William Osler)

Introduction

- Research approach & Research/study design: terms frequently used interchangeably **yet different**
  - ***Research/study design***: is a broader plan to conduct a study.
  - ***Research approach*** is an important element of the research design, which governs it.

The Approach

- It involves the description of the plan to investigate the **Research Problem** under study in a structured **(quantitative)**, unstructured **(qualitative)** or a combination of the 2 approaches **(quantitative-qualitative integrated approach)**
- Approach helps to decide about the presence or absence as well as manipulation & control over variables
- The approach of research study depends on several factors but primarily on the nature of **Research Problem** under study
- At this stage of the research study, **conceptual Framework** may or may not be incorporated

Study design

- The **FRAMEWORK/GUIDE** used for planning, implementation & analysis of a study.
- A **SYSTEMATIC PLAN** of what is be done, how it will be done, & how analyses will be done.
- Is the **MASTER PLAN** specifying the methods and procedures for collecting and analysing data
- Defined as a **BLUE PRINT** to conduct a research study, which involves the descriptive of research approach, study setting, sampling size determination, sampling technique, tools and method of data collection and analysis to answer specific research questions or for Hypothesis testing.
- A **PLAN** of how, when & where data are obtained and analysed.

Elements of Study design

- **Population, sample size & Sampling technique**:
  - Study design also provide the investigator/researcher with directions about population, sample size & sampling technique, to be used for the research study.
  - **Example:** in a Quantitative study design, an investigator gets a directive that the population will be pregnant women attending ANC (group) and the study will include a **sample size** of **X** selected through **systematic** sampling technique.

Elements of Study design

- **Time, Place & source Data collection**:
  - **Time**: specifying days, months, years of study
  - **Place**: location (Study settings): Study Area vs Settings
  - The **sources of the requisite data** are the other important constituents essential to ensure effective planning to conduct a research study

Elements of Study design

- **Tools and methods of data collection**:
  - It involves the description of different tools and methods of data collection

  Example:
  - Questionnaires
  - Interviews
  - Direct observation
  - Any other method that suit the particular approach of research as well as nature of the phenomenon under study.

Elements of Study design

Data Analysis Methods:

- **Data-driven** and relies on a **systematic** and **unbiased** approach to **collect**, **analyze**, and **interpret** data
- Research design **MUST** also include the description of the methods of data analysis – either quantitative or qualitative data analysis techniques.
- That helps the researcher/investigator to collect relevant data, which later can be analysed as per research design plan.
- Without a formal plan of data analysis an investigator may collect irrelevant data, which can later become difficult to analyse ***(GIGO).***

Data Analysis (Guided by Statistical Analysis Plan)

A. Univariable

- Descriptive
- Describes the study population (generalizability); compare intended vs actual population
- Precedes bivariable and multivariable data analyses

B. Bivariable

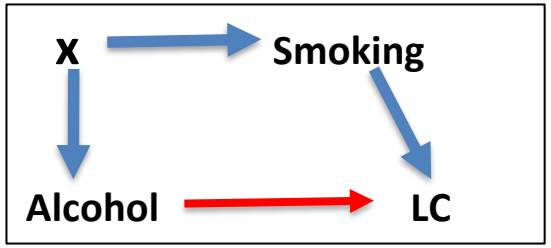
- Simple, unadjusted associations between two variables
- 2x2, 2xX, Chi-Square, T-tests
- Check test assumptions; Expected and Low Values
- Avoid bivariable regression models until the end
- Multiple testing (a-priori analyses otherwise report as exploratory)

	LC+	LC-	
Alc+	60	10	70
Alc-	9	50	59
	69	60	129

Data Analysis (Guided by Statistical Analysis Plan)

C. Multivariable

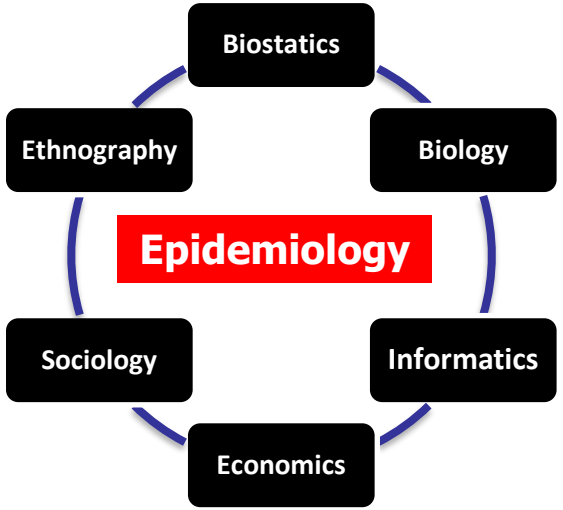
- $Y_{ii} = \alpha_i + E_1\beta_1 + E_2\beta_2 + E_3\beta_3 + E_4\beta_4 + E_5\beta_5$ (Risk Factor)
- $Y_{ii} = \alpha_i + E_1\beta_1 + C_1\beta_2 + C_2\beta_3 + C_3\beta_4 + E_1{*}C_2\beta_5$ (Hypothesis Driven)
- Exposure, Confounders, Mediators, Effect Measure Modifiers
- Should know what the results will be even before running models
- Model assumptions – Confounder Criteria – Confounding - Modification
- Study type & measures of association
- Sensitivity analysis
- Model fitness



Data Analysis Methods

- Draws on methods from **various scientific fields** such as biostatistics, informatics, biology, economics, social and behavioral sciences.

- Considers the role of **genetics/biology, individual behavior, community, socio-economic, cultural (Anthropology), environmental and historical factors** on health.



DECISION: SELECTION OF RESEARCH DESIGN

Epidemiology In action

- **Def. 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.

- Study designs are plans and procedures for research that extend the decisions from broad assumptions to detailed methods of data collections and analysis.

- Investigators must select the most appropriate design in order to achieve the aims and objectives of the study.

FIRST STEPS

- The selection of study design largely depends on the:
  - Nature of the research problem
  - What is the purpose of your investigation?
    - Health challenges you observe; review of the literature
  - Availability of Resources (Time, Cost, Materials, expertise of the investigator (**4Ms**))
  - Subjects/participants accessibility
  - Research ethics
- Do you have a validated instrument to collect the data?
  - Have the validated instrument been piloted and training provided?

Factors affecting selection of study design

- **Nature of Research Problem**
  - Is the most important factor, which helps the investigator to decide about the selection of study design.
  - Based on the nature of the research problem, investigator deicide whether it should be investigated through **experimental or non-experimental**.

- **Purpose of the study**
  - Study may be conducted for the purpose of prediction, descriptive, exploration, or correlation of the research variables.
  - Therefore, the purpose helps the investigator to choose the suitable study design.

Count……

- **Investigator's expertise**
  - Selection of study design is largely influenced by the investigator's ***Knowledge & experience***, because they avoid using those designs wherein they lack confidence, relevant expertise
- **Investigator's interest and motivation**
  - ***Interest*** and ***motivation*** levels help investigator.
  - Therefore, the purpose helps the investigator to choose the suitable study design(s).
  - Most motivated investigators ***always analyse most aspects*** of study design before selected one or combination.
  - **NB:** ***Casual*** investigators may choose study design (s) that may lead to ***Failure***

Count……

- **Research Ethics and Principles**
  - The incorporation and application of ethical considerations and principles in study design are vital.
  - This includes **moral obligations** such as;
    - Respect & Rights for study subjects/participants.
    - Informed consent.
    - Potential risks or discomfort on participants: Biological risks, Psychological risks, Social Risks, Legal risks, Financial risks, Other risks **(BREC-Form).**
  - Selection is significantly influenced by the **ethics of the study design**.
    - e.g. An investigator may want to conduct a study through a certain experimental approach, but problems of ethical approval may stop the investigator to carry on with the study and investigator may opt or settle for another available possible deign.

Count……

- **Study population/participants**
  - Sample size (number) and availability of the participants does influence the selection of the research design. If only few subjects are involved, an in-depth qualitative investigator may opt for Qualitative design.

- **Resources : (4Ms)**
  - **NO** investigator can conduct research without resources
  - Money, equiments, facilities & support from colleagues
  - However, some studies may require a lot of resources as compared to others. Therefore, the selection of a study design may be affected by the availability of resources.

Count……

- **Time**
  - Time is also a major deciding factors for the selection of study design.
    - **Ex.,** an Investigator needs more time to conduct longitudinal/cohort studies, while Cross-sectional studies can be conducted in shorter time.

- **Users of the findings**
  - Study design also use various methods of data collection and data analysis.
  - Therefore, while choosing a study design, investigator **Must** ensure that study design is as appropriate for the users of the study findings as possible, so that that maximum advantage of the results can be obtained.

Count……

- **Possible Control on Extraneous variables**
  - An efficient design can maximize results, decrease errors, and control pre-existing or impaired conditions that may affect study outcome.
  - The maximize efforts of the investigator should maximize control.
  - Therefore, possible control over the extraneous variables may affect the selection of a study design.

    Ex., An investigator wants to conduct a study through TRUE- Experimental design but because of inability to control selected extraneous variables, other similar design has to be opted for, such as quasi-experimental or pre-experimental design

Validity of study design

- Two important CRITERIA for evaluating the credibility & dependability of the research results:
  - **Internal Validity**
  - **External Validity**

Internal Validity

- Validates whether the **independent** variables actually made a difference
- Internal validity refers to the extent to which it is possible to make an ***INFERENCE*** that the independent variable is ***TRULY*** influencing the **dependent** variable (Campbell et al, 1963)
- In the internal validity, the **independent** variable is responsible for variation in **dependent** variable.
- Internal Validity demands a tighter control over study to maximize the effectiveness of the results

Count……

- Internal validity is helpful in making the inference that the **independent** variable influences the **dependent** variable
- ***SIX*** major extraneous variables have been identifies which can jeopardize the internal validity (knows as ***Threats to internal validity***):
  - History
  - Maturation of subjects (age)
  - Testing
  - Instrumentation changes
  - Mortality
  - Selection Bias

Mortality

- The threats of History occurs when some event beside the experimental treatment occurs during the course of the study, and this events even influences dependent variables.
  - E.g., you are conducting a health education programme on the importance of Cervical cancer Screening (CCS), while recently a Public figure is diagnosed to be suffering from cervical cancer.
  - It catches media attention: Medical experts are interviewed, & importance CCS is supported.
  - All major TV channels and newspapers starts reporting on the importance of CCS
  - Results: CCS activity has improved, you as an Investigator may be able to conclude if the change in behaviour is due to your intervention (Health education programme) or it is due to the diagnosed of affliction of the public figure and subsequent media coverage.

Maturation of subjects

- When a research (e.g. experimental research) is carried on for along period of time over a group of subjects, they may be changes in the subjects in different ways, e.g. in children there is increase in weight, height, etc.

  - **Example.**; An investigator is interested in assessing the effect of particular nutritional protocol on the weight and height of the malnourished children

  - If this experiment is conducted for a very long period, it is difficult to make out whether the effect on weight and height is due to maturation or nutritional protocol

Testing

- It refers to the effect of taking a pre-test of subjects' performance post-test.
- The effect of taking a pre-test may sensitize an individual and improve the score of the post-test.
- Individuals generally score higher when they take test a 2nd time regardless of the treatment

Instrumentation change

- Instrumentation is a threat that related to Measurement.
- This bias reflects changes in measuring instruments or methods of measurements between two points of data collection.
- Instruments like weighing scale, tape measure, thermometer, sphygmomanometer, etc. Should be checked for their accuracy at regular intervals, & same instruments should be used throughout the study to minimize the instrument-related error of the internal validity.

Mortality and Mobility

- Mortality can lead to loss or dropout study participants during the study period.
- If the participants who remain in the study or join later are not similar to those who dropped out, the results could be affected.
- **Ex.**, In Cohort Study: some participants who participated in the 1st follow-up might not be available in the next follow-up.

Selection Bias

- If the participants are not selected randomly for participation in groups, then there is a possibility that the groups which will be compared may not be equivalent (not representative of the population intended to be analysed).

External Validity

- It refers to the extent to which the results can be generalized to the larger population **(GENERALIZABILITY OF RESULTS)**.
- External validity researches **under what conditions** & in **which type of participants** the same results can be expected to be **replicated**, or whether the same intervention can be **applicable** in another setting & with different participants.
- It explores the generalization beyond specific experiment, to check if the results come out to be the same with other settings or with other participants/population.

Factors affecting External Validity

- **Hawthorne effect**
  - Subject may behave in a particular manner because they are aware that they are being observed.
  - Participants have the knowledge that they are involved in a study, thus affecting the result.
- **Experimental effect**
  - It's a threat to study results when investigator's characteristics, mannerisms, behaviour may influence subject behaviour.
    - **E.g**.: investigator characteristics: facial expressions, clothes, age, gender, body built.
  - Thus, the way researcher dresses up or his or her gender can influence the way in which respondents answer research questions

Factors affecting External Validity

- **Reactive effect of pre-test**
  - It occurs when participants have been sensitized to the treatment because of taking pre-test.
  - Participant might not respond to the treatment in the manner they finally do if they had not received the pre-test.
    - E.g.:, An Investigator wants to conduct a study to assess the effect of a health education programme on the awareness of HIV/AIDS among youth.
    - Instead, he/she conduct pre-test to collect baseline data before the intervention.
    - This pre-test may sensitize the participants to learn more about HIV/AIDS irrespective of health education is provided or not to theme

Novelty Effect

- When treatment is new, investigators & participants might behave in different ways.
- They may be enthusiastic about new methods of new methods of doing things.
- Once treatment is more familiar & as the novelty wears off, results might be different.

Factors affecting External Validity

- **People**
  - E.g., people with specific race, such as whites have high prevalence of coronary artery disease compare to indigenous Africans.
  - Therefore, a generalization made for whites will not be applicable to indigenous Africans. hence, this is a threat to external validity (limit the generalizability of the findings)

Place

- E.g. people living in high altitudes have high haemoglobin (Hb) levels because of the higher altitudes the requirement of oxygen is more, due to which there is more production of red blood cells (RBCs). Therefore, you can't compare people living in higher altitudes to those in lower altitudes (thus affect generalizability).

- **Time**
  - If study was carried out of a community in 1995 & then again in 2010, the results of these 2-studies would be different. Therefore, older results cannot be generalized over periods of time as societies & circumstances constantly change.

STUDY DESIGNS

- Generally research designs are classified into TWO Broad Categories
  - Qualitative design
  - **Quantitative design** ***** **Epidemiological Study Designs**

Quantitative/Qualitative

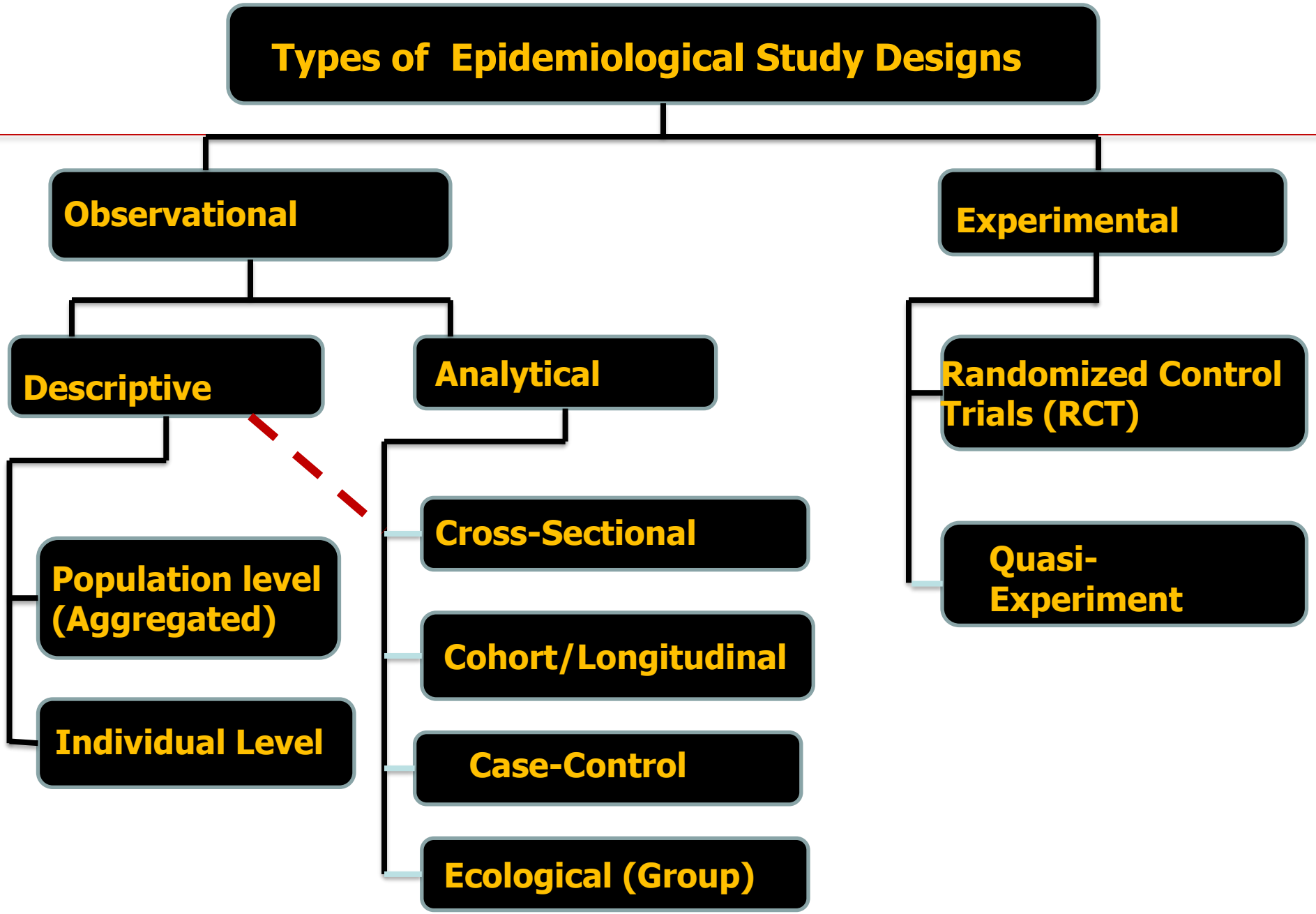
- **Quantitative research:** involves measurement of outcomes using numerical data under standardized conditions
  - May be used along the continuum of research.
- **Qualitative research:** is concerned with narrative information under less structured conditions that often takes the research context into account
  - Descriptive and exploratory research.
  - Purposes: describing conditions, exploring associations, formulating theory, generating hypotheses.

Assumptions in Scientific Research

- Nature is orderly and regular.
- To some extent, events are consistent and predictable.
- Events or conditions have one or more causes that can be discovered
  - This enables establishing cause and effect relationships

Properties of scientific method

- **Systematic**
  - Use of orderly procedures to ensure reliability
  - Logical sequence is used from problem identification, through data collection, analysis, & interpretation
- **Empirical**
  - Documentation of objective data through direct observation (or other systematic methods)
  - Findings are grounded in the objective observation of phenomena rather than the personal bias or subjective belief of the researcher
- **Control**
  - In order to understand how one phenomenon relates to another, factors are controlled that are not directly related to the variables in question
  - Investigators have confidence in their research outcomes to the extent that they control extraneous influences.


Types of Epidemiological Study Designs
Observational
Experimental
Descriptive
Analytical
Randomized Control Trials (RCT)
Cross-Sectional
Population level (Aggregated)
Quasi-Experiment
Cohort/Longitudinal
Individual Level
Case-Control
Ecological (Group)

Observational studies

Descriptive

- Suggest hypotheses
- Resource allocation
- Natural history

Analytic

- Test hypotheses
- Assess association/causation

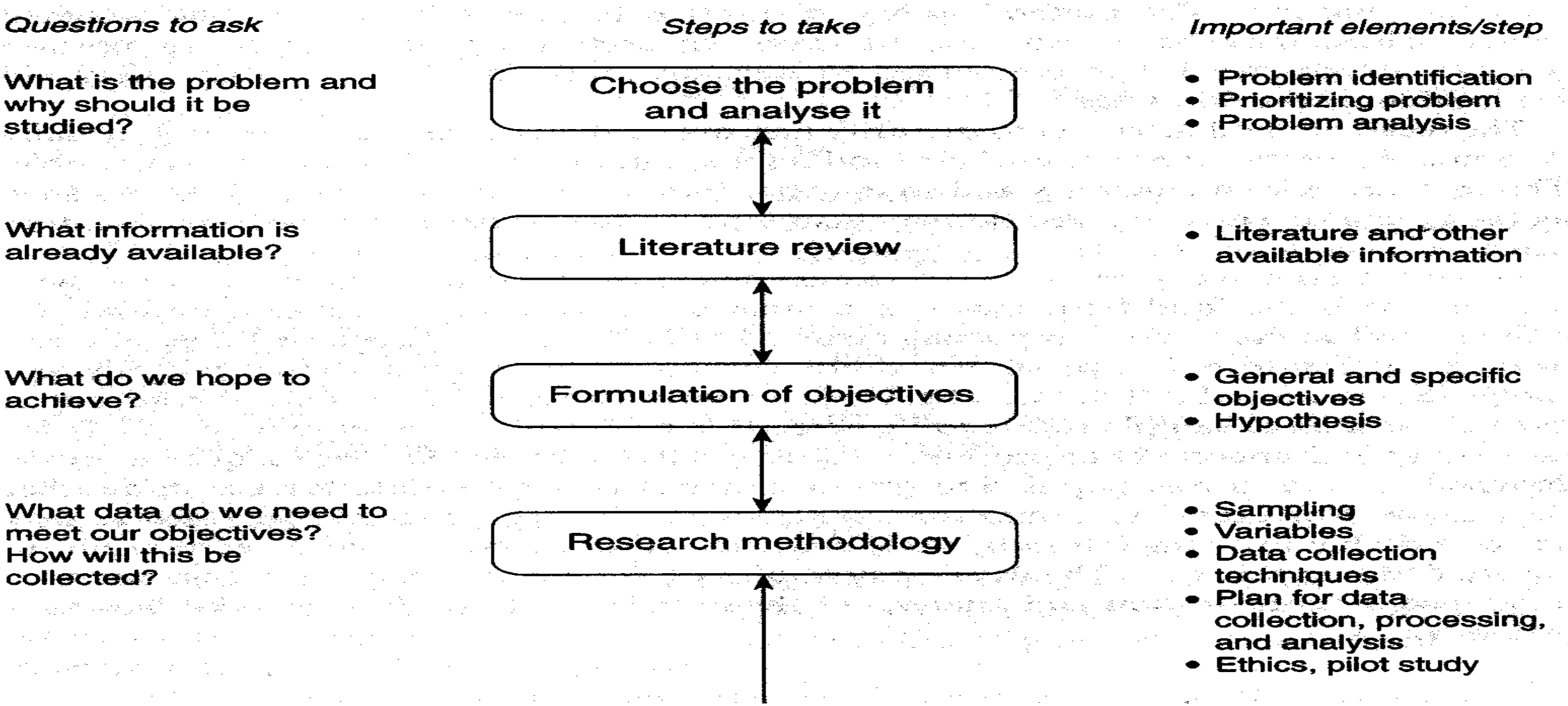
Descriptive study/research

- **Descriptive**: investigator attempts to describe a group of individuals on a set of variables or characteristics.
  - Enables classification and understanding.
  - **Methods**: surveys, case study, qualitative, developmental (natural history of something, patterns of growth and change), normative, evaluation.

Exploratory study/Research

- Investigator examines a problem of interest and explores its dimensions, including how it relates to other factors.
  - Proven relationships between the problem and other factors can lead to predictive models.
  - Correlational studies, cohort and case control, secondary analysis, historical research.

Steps in cross-sectional studies



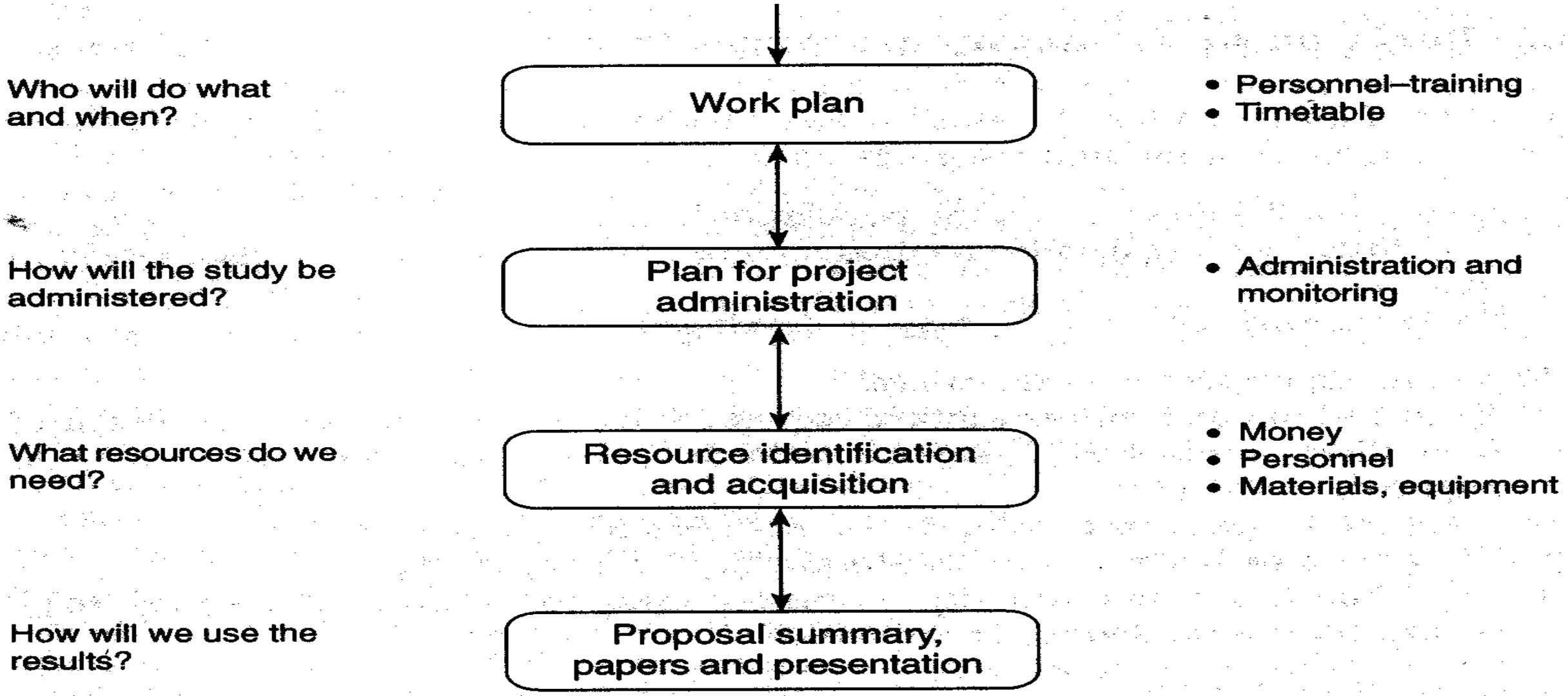


Fig. 13.5 Steps in the design of a cross-sectional study. (Modified from Varkevisser *et al.*[23])

Experimental study/Research

- Provides a basis for comparing 2 or more conditions.
- Controls or accounts for the effects of extraneous factors, providing the highest degree of confidence in the validity of outcomes.
- Enables the researcher to draw meaningful conclusions about observed differences.
- **Types:** Randomized controlled trials, single subject designs, sequential clinical trials, evaluation research, quasi-experimental research, meta-analysis.

Design options in epidemiologic research

Type of study	Alternate name	Unit of study
Observational studies		
▪ ***Descriptive studies***		
▪ ***Analytical studies***		
Ecological	Correlational	Populations
Cross-sectional	Prevalence	Individuals
Case-Control	Case-Reference	Individuals
Cohort	**Follow-up/ Longitudinal**	**Individuals**
Experimental/ intervention Studies		
Randomized Controlled Studies	Clinical Trial	Patients
Field Trial		Healthy person
Community Trial	Community intervention studies	Communities

Qualitative Study

- Seeks to describe how individuals perceive their own experiences within a social context
- Emphasizes in-depth, nuanced understanding of human experience and interactions
- Methods include in-depth interviews, direct observations, examining documents, focus groups
- Data are often participants' own words and narrative summaries of observed behavior

Qualitative Study

Strengths

- ❖ Data based on the participants' own categories of meaning
- ❖ Useful for studying a limited number of cases in depth or describing complex phenomena
- ❖ Provides understanding and description of people's personal experiences of phenomena
- ❖ Can describe in rich detail phenomena as they are embedded in local contexts
- ❖ The researcher can study dynamic processes (i.e., document sequential patterns/change)

Weaknesses

- ❖ Knowledge produced might not generalize to other people or other settings
- ❖ It is difficult to make quantitative predictions
- ❖ It might have lower credibility with some administrators and commissioners of programs
- ❖ Takes more time to collect and analyze the data when compared to quantitative research
- ❖ The results are more easily influenced by the researcher's personal biases and idiosyncrasies

FINAL WORDS

- World is not linear
- We seek to publish positive findings but negative findings are also important
- Publications: mindful of stigmatizing language ("HIV infected/positive")
- Peer-reviewed publications almost never reach the community we seek to inform
- **Epidemiology** is not simple but it is crucial to public health
- You may not have the expertise but partners next to you do, but you need to ask (Partnership/Collaboration is key)

References

- Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to Epidemiology: Open University Press;2005
- Kelsey JL, Whitemore AS, Evans AS, Thompson WD. Methods in Observational Epidemiology; 2nd Edition:Oxford University Press; 1996
- Campbell & Stanley (1963, 1966) introduced the terms ***Internal validity*** and ***external validity*** while the later Cook & Campbell (1979) edition added *statistical conclusion validity* and *construct validity* (Rosenthal & Rosnow, 2008). The first version of this work also introduced the term *quasi-experiment*.

UNIVERSITY OF KWAZULU-NATAL ™

INYUVESI YAKWAZULU-NATALI

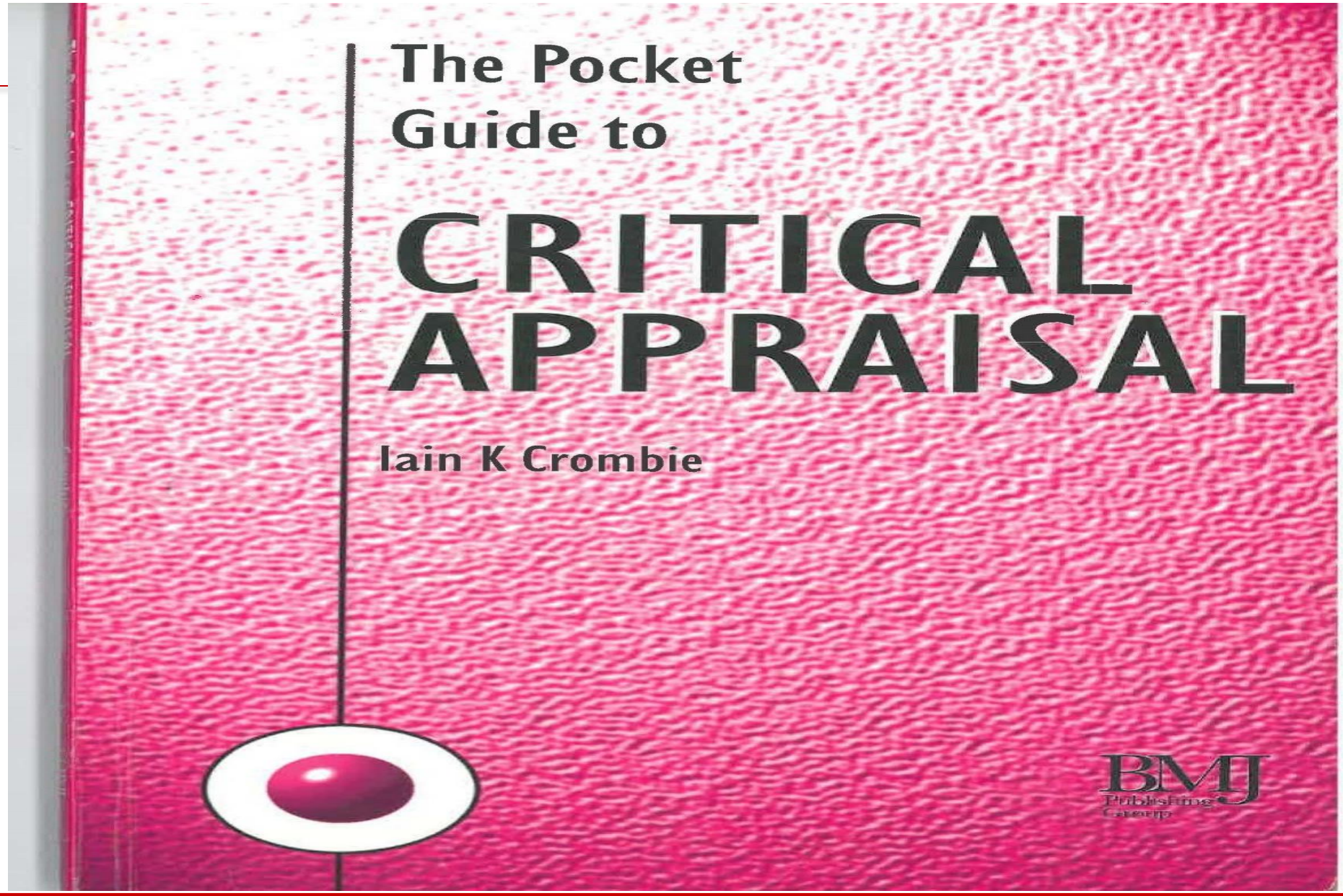
Causal Inference in Epidemiology

Themba G. Ginindza, PhD, MPH, MSc-Epi

Professor/Epidemiologist/MLCCP Lead PI/CCPAC PI
Director: Cancer & Infectious Diseases Epidemiology Research Unit (CIDERU)
Discipline of Public Health Medicine, University of KwaZulu-Natal
Tel: +27 31 260 4214 | Mob:+27 719 1111 79
Email: Ginindza@ukzn.ac.za

Whenever You Feel Like Giving Up ... Ask yourself if you are giving up on or transforming your vision?

Whenever You Feel Like Giving Up….
ASK Yourself, Why Did You Start?


The Pocket
Guide to
CRITICAL
APPRAISAL
Iain K Crombie
BMJ
Publishing
Group

Outline

- Background
  - Scientific philosophy
  - Historical development of Disease Causation
- Definitions
- Levels/types of causality
- Bradford-Hills criteria

Background

Significance of Epidemiology:

- To practice medicine and public health effectively; Evidence/Knowledge on 3 fundamentals type of professional knowing "gnosis" is needed

Dia-gnosis	Etio-gnosis	Pro-gnosis	For individual (clinical medicine)
Dia-gnosis	Etio-gnosis	Pro-gnosis	For Community (Public & Community Health)

- **Of the 3 types of knowing ("gnosis") Etio-gnosis (Causality) is the central concern of Epidemiology**
  - **The principal aim of Epidemiology is to Assess the Cause of Disease**

Significance of Epidemiology

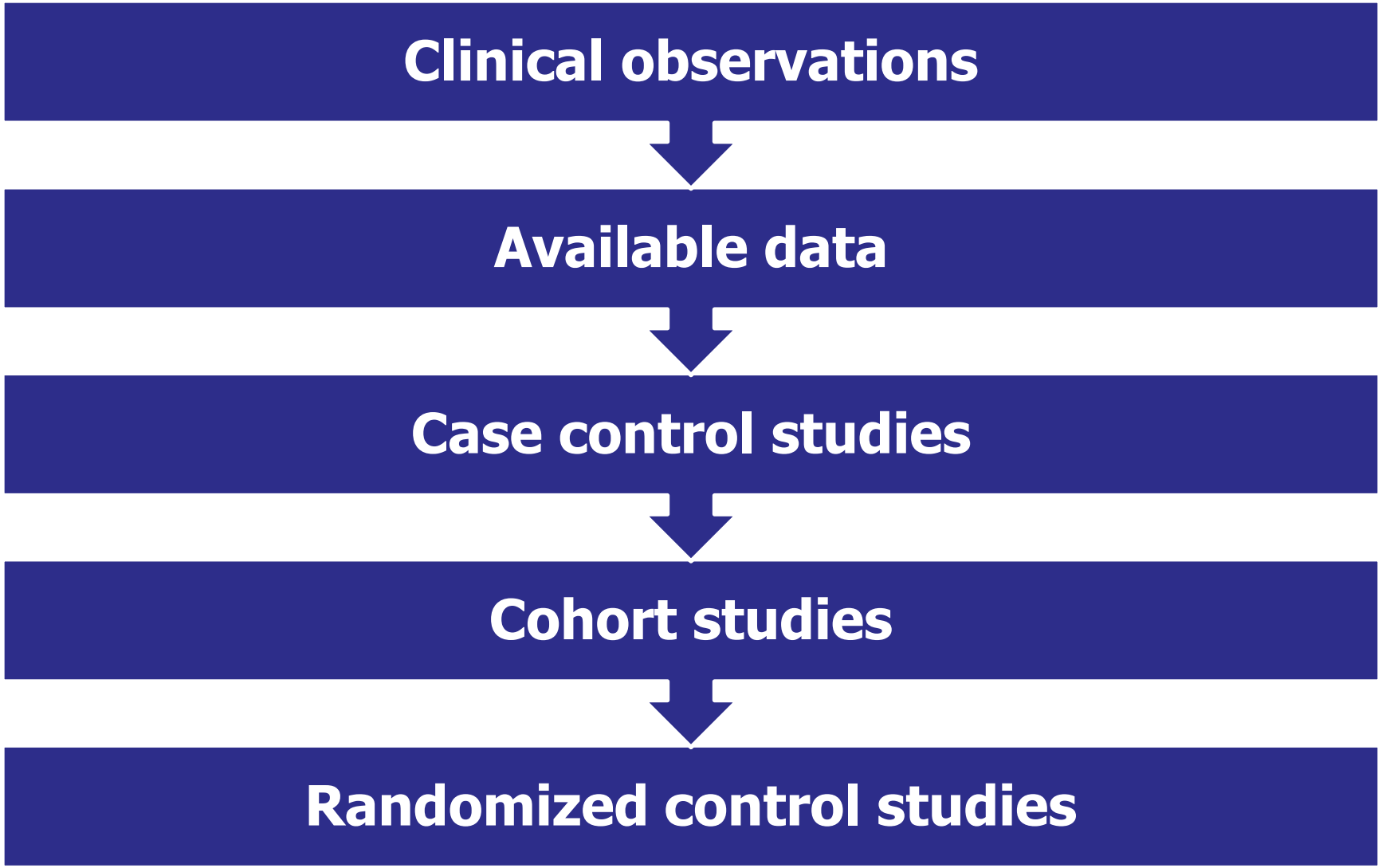
- Epidemiology aims at establishing variational causal claims
  - To study the variability of disease due to variation of exposure.
  - Variation guides causal reasoning.
- Epidemiologist search for causes, want to make asymmetrical statements that have direction.
  - What causes and prevent disease?
  - They seek to establish that an Independent variable ***X*** causes changes in the Dependent variable ***Y*** and not the reverse.
- The central problem of cohort studies is to cope with the change that occurs with the passage of ***Time***, Yet the study of causal involves the detection of change in a ***Dependent*** variable by change in an ***Independent*** variable.

Significance of Epidemiology

Establish variation using:

- Descriptive epidemiology (studies)
  - Study of distribution of diseases
- Analytical epidemiology (studies)
  - Study of determinants of diseases
    - Association of exposure to disease
- Social Epidemiology
  - Study of relations between social factors and disease in populations
  - Ethinicity, gender, socio-economic status

APPROACHES FOR STUDYING DISEASE ETIOLOGY



Methods: Observational studies consider

1. **Exposure**: Does/Does not cause disease
2. "The risk of disease is ***X*** times greater among exposed individuals than un-exposed individuals

NB

To establish **(1)** we need to Establish **(2)** First

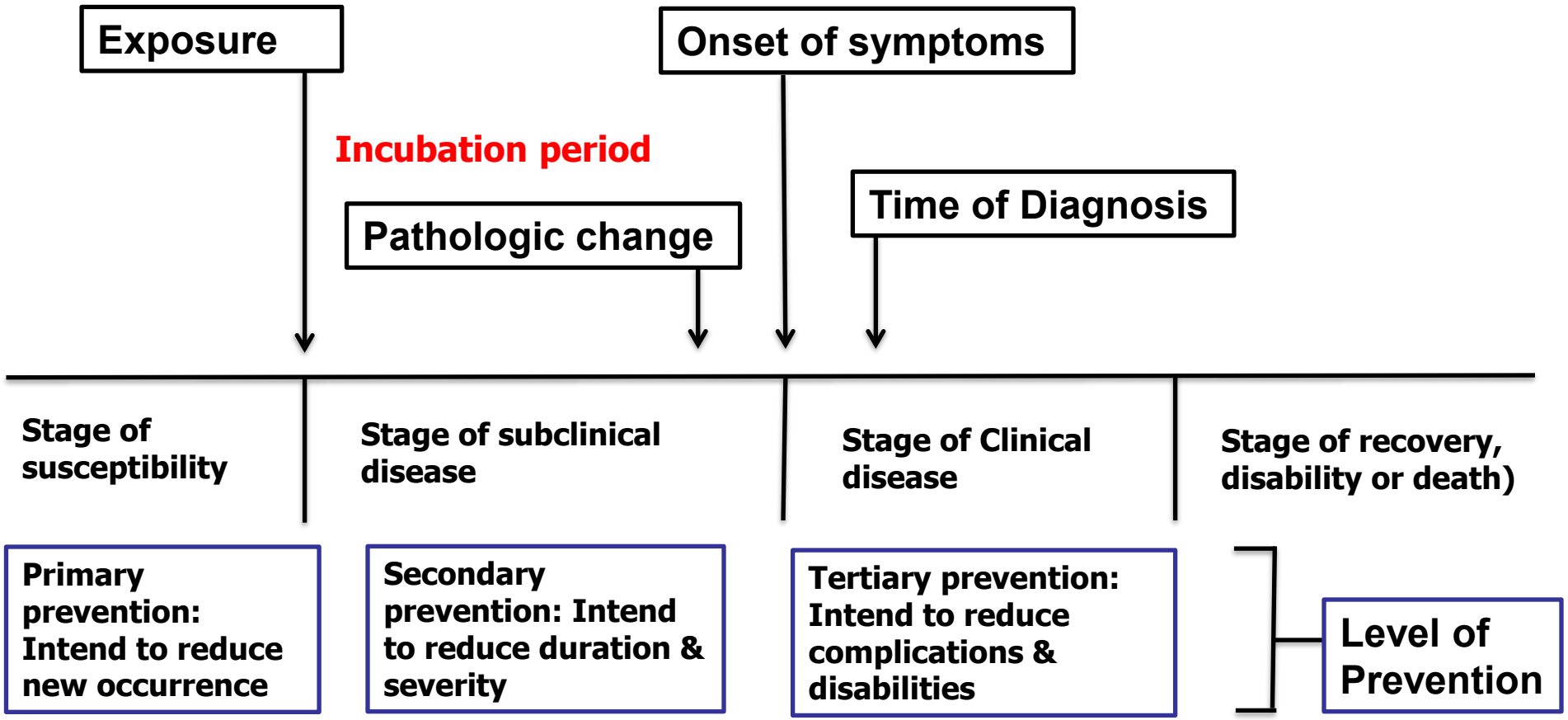
- ❖ Causal relations are established through comparative statements

Epidemiological Studies (Re-Cap)

- ***Cohort:*** Compare exposed with un-exposed individuals
- ***Case-Control:*** Compare diseased with non-diseased individuals
- ***Cross-sectional:*** compare various individuals characteristics at a specific point of time
  - Measure exposure and outcome at the same time

Natural History of Disease

- Refers to the progression of a disease in an individual over times
- Includes all disease-related phenomena from before initiation of diseases (stage of susceptibility) until resolution of the disease (stage of recovery, disability or death

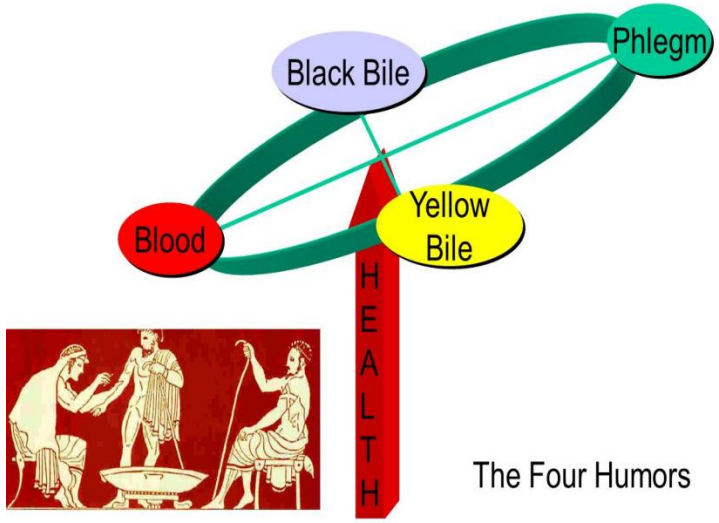


Historical development of diseases

4th Century BC: Hippocrates: Four Vital Humors

- Historically, there have been many efforts to account for the occurrence of disease outcome. **Religion** often attributed disease outbreaks or other misfortunes to divine retribution – punishment for mankind's sin
- **Hippocrates** promoted the concept that disease is a result of an imbalance among four vital "humors" with us:
  - Yellow Bile
  - Black Bile
  - Phlegm
  - Blood

  Caused by changes in season, air, wind, water and stars



The Four Humors

Miasmas (1800's)

- Disease rose from foul clouds low on earth's surface

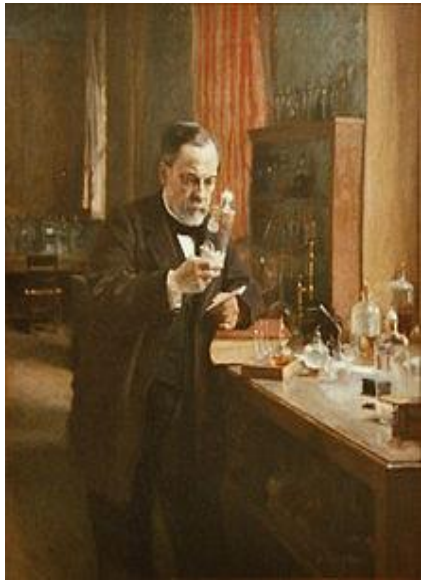
- Miasmas were toxic vapors or gases that emerged from dirty pools or swamps or filth, and it was believed that if one inhale the vapors, disease would result

Germ Theory (1878)

- Louis Pasteur and others introduced the **Germ Theory** in **1878**
- In **1890** Robert Koch proposed **Specific Criteria** that should be met before concluding that a disease was caused by a particular bacterium

Koch's Postulates criteria:

- The bacteria must be **present** in every case of the disease.
- The bacteria must be **isolated** from the host with the disease and grown in pure culture.
- The specific disease must be **reproduced** when a pure culture of the bacteria is **inoculated** into a healthy susceptible host.
- The bacteria must be **recoverable** from the experimentally infected host



Limitations of Koch Postulate

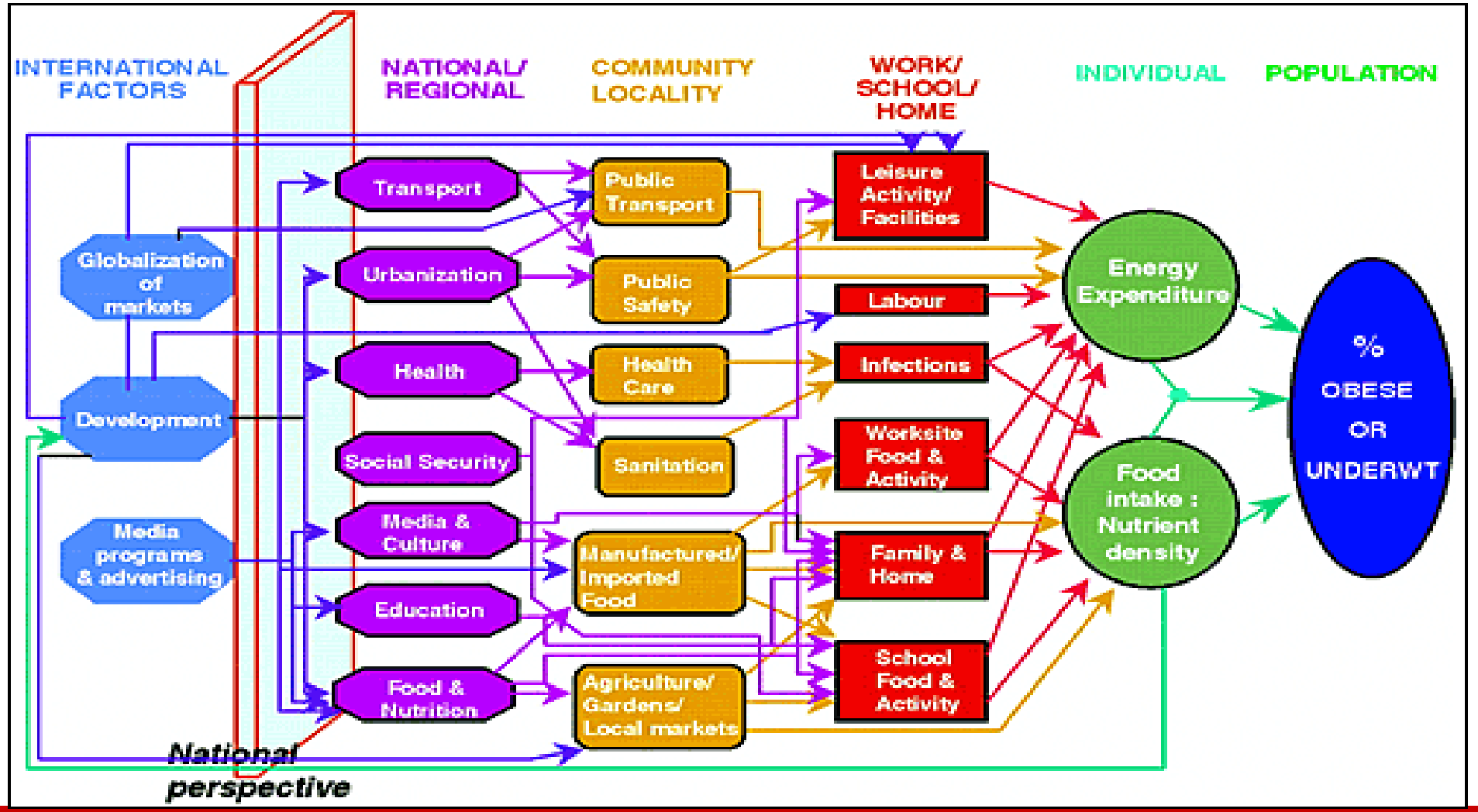
- Non communicable disease
- One to one relation are **rare biology**
- Disease production may require **co-factors**
- Always it is not possible to isolate organism from diseased person
- Viruses **cannot** be cultured like bacteria because they need living cells in which to grow.
- Always infection **does not produce disease**
- Pathogenic microbes **may be present** without clinical disease (sub-clinical infections, carrier states)

Webs of causation

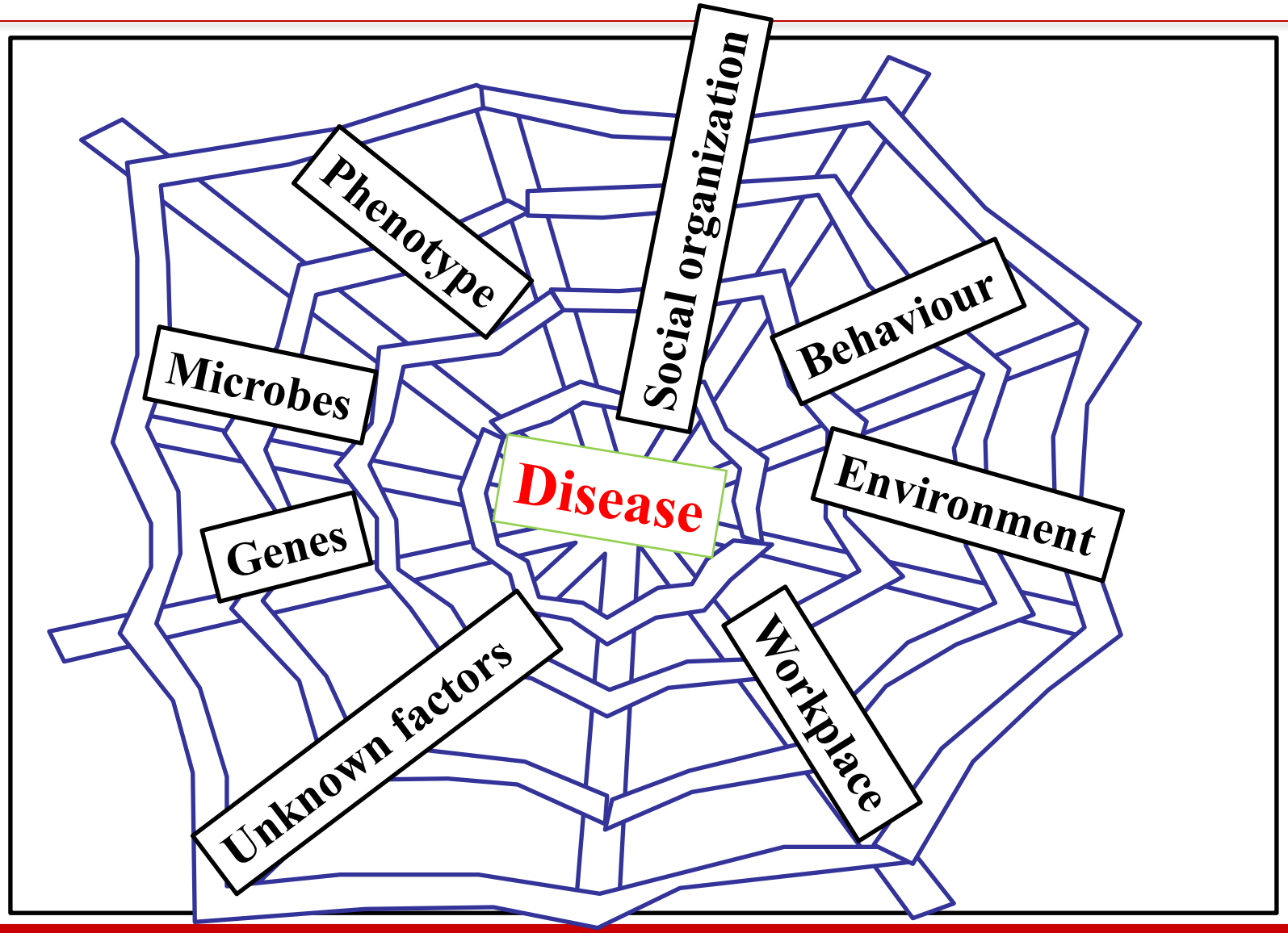
- The germ theory didn't provide insights regarding the causes of chronic diseases, and over time it became increasingly apparent that for most diseases there were many contributory factors
- Researchers began thinking about complex "***Webs***" of causation
- Useful for understanding the causes on non-infectious diseases
- Complex of web **interconnected** host and environmental, social and psychological factors
  - Illustrates the interconnectedness of possible causes
- Causes of disease are **interacting**
- Multiple causes and pathways of disease
- **Chain of causation**

Web causation

A web of causation for obesity in the context of socio-ecologic perspective



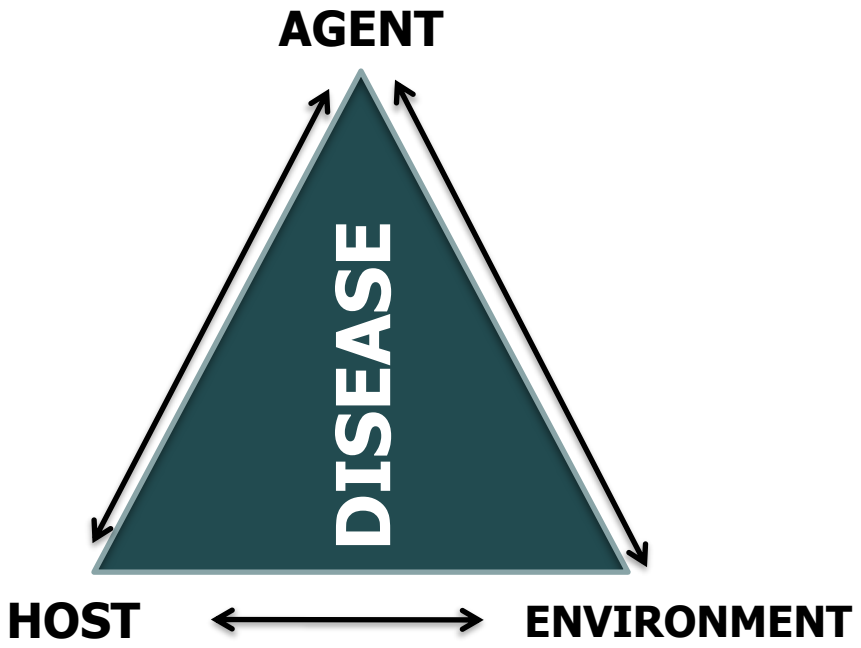
Web of Causation



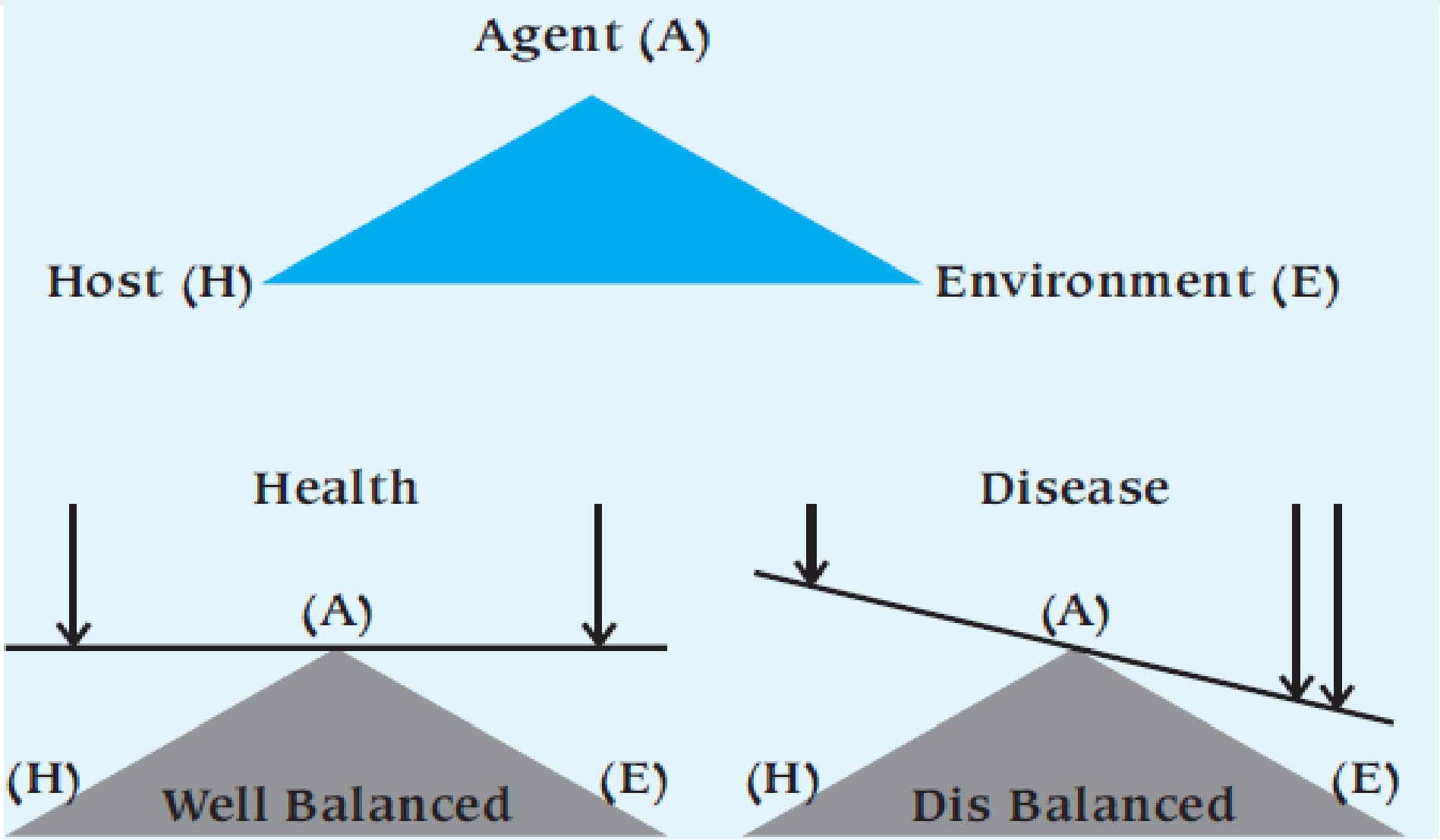
Epidemiologic Triad of disease (1920-1950s)

Disease is the result of forces within a dynamic system consisting of:

- **Agent of infection**
- **Host**
- **Environment**



Epidemiological Triad



Classic Epidemiologic Theory

- **Agents**
  - Living organisms
  - Exogenous chemicals
  - Genetic traits
  - Psychological factors and stress
  - Nutritive elements
  - Endogenous chemicals
  - Physical forces

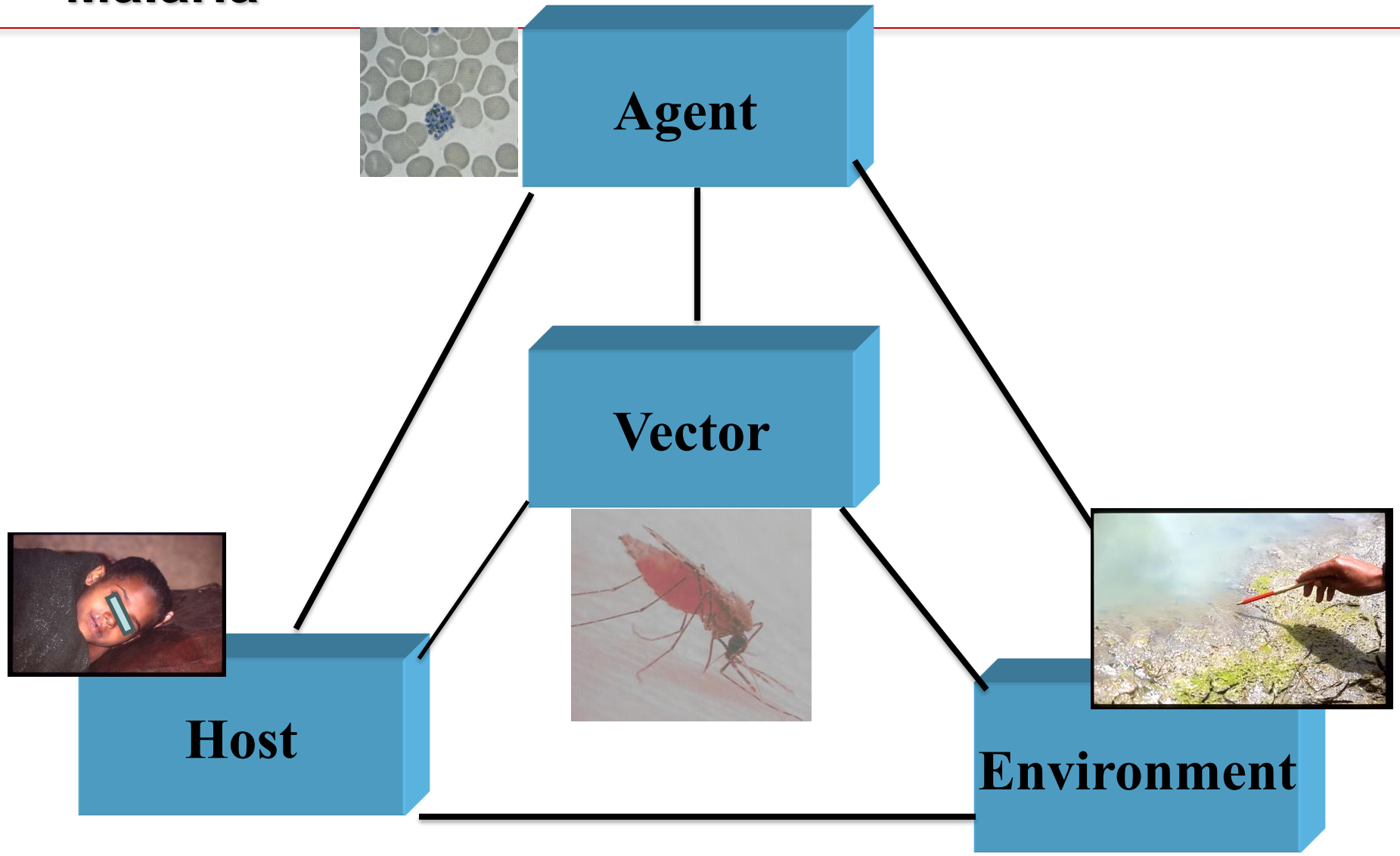
- Agents have characteristics such as **infectivity**, **pathogenicity** and **virulence** (ability to cause serious disease)
  - They may be transmitted to host via vectors

Classic Epidemiologic Theory

- **Host factor:**
  - Immunity and immunologic response
  - Host behaviour

- **Environmental factors**
  - Physical environment (heat, cold, moisture)
  - Biologic Environment (flora, fauna)
  - Social environment (economic, political, culture)

Malaria



Era of Epidemiology

Era	Medical Paradigm	Preventative Approach
Sanitary	Miasma: Environmental pollutants	Drainage, sewage, sanitation
Infectious disease epidemiology	Germ theory: single agent for a single diseases	Vaccines, Isolations, quarantines & Antibiotics
Chronic disease epidemiology	Black box: exposure related to outcome without necessity for intervening factors or pathogenesis Multi-causal	Lifestyle & Environmental modification
Eco-Epidemiology	Societal perspective Globalisation Post genome	Inter govt agencies, fiscal and environmental policies Gene therapy

Era of epidemiology: The Sanitary era

- Cause of disease at a societal level
- Mechanism – Miasma
  - Setting up a system that routinely recorded the cause of death **(Farr)**
- Foul emanations from soil, air and water
- **Solutions:** Drainage, clean water supplies, sanitation

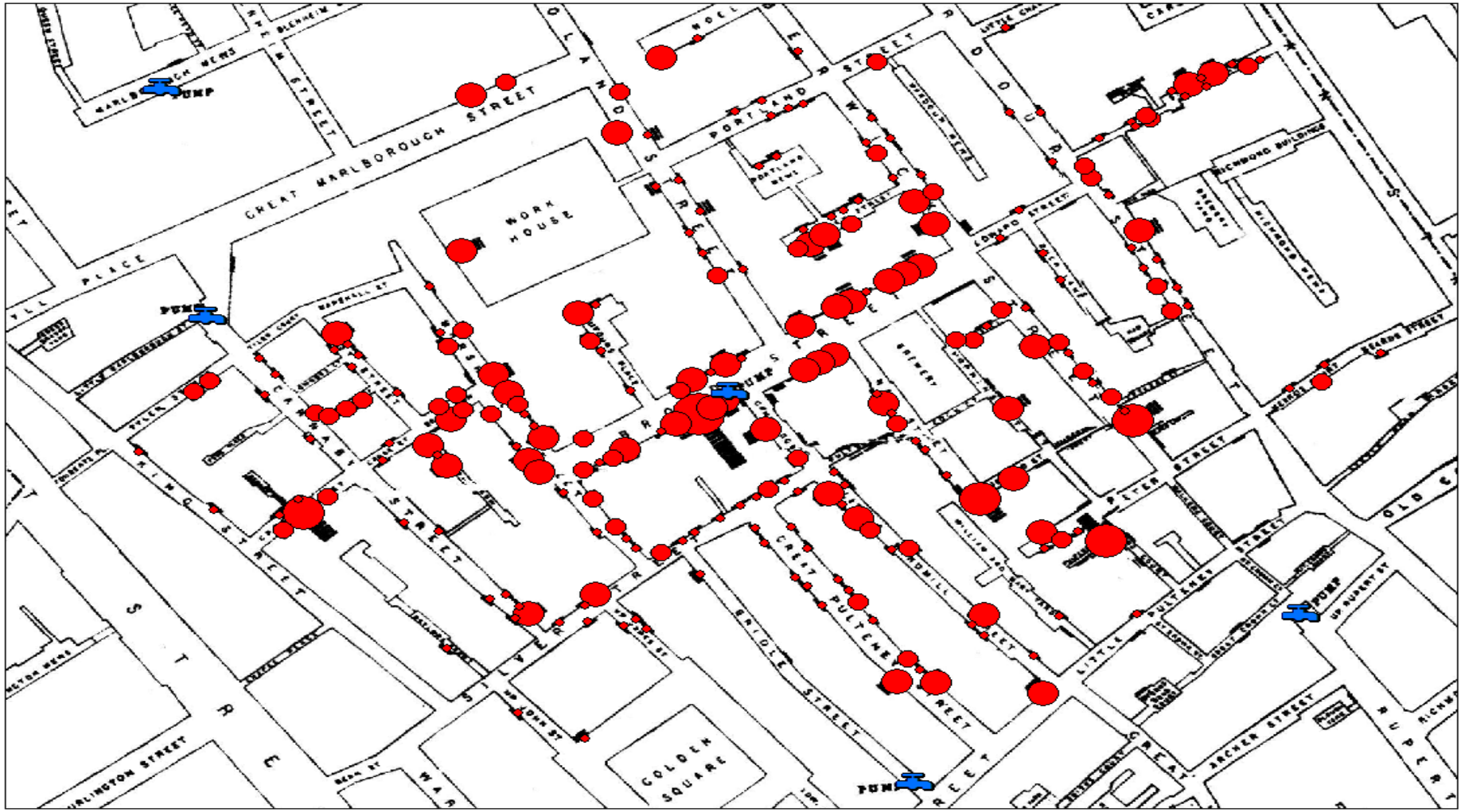
Nightingale

Edwin Chadwick

William Farr

John Snow

John Snow Discovery : Cholera outbreak

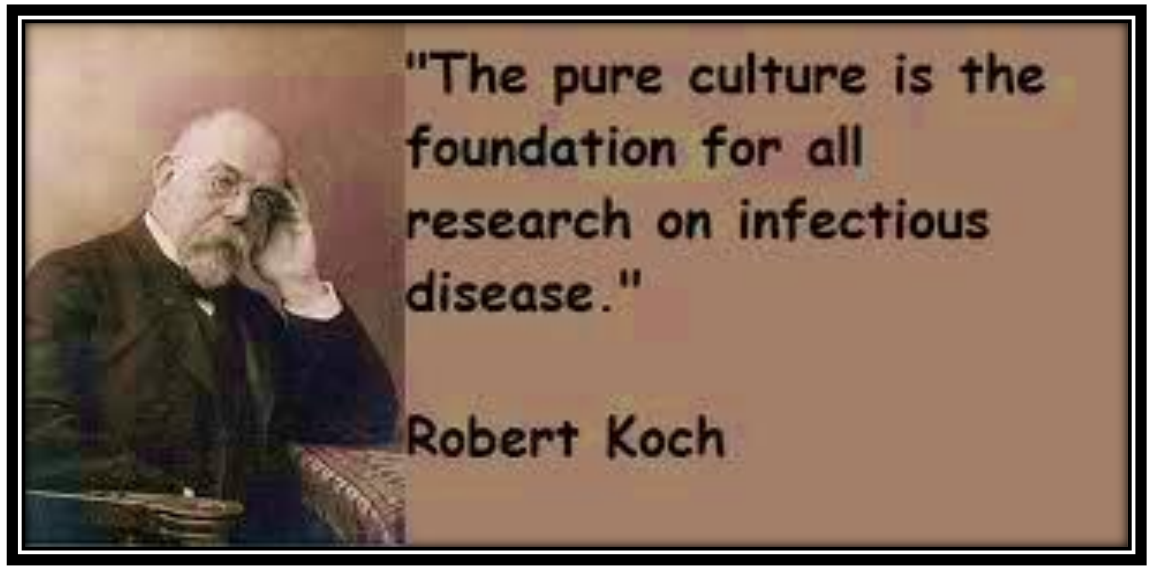


The Broad street cholera outbreak was a severe outbreak that occurred near Broad street (now renamed Broadwick street) in Soho district of London, England, 1854

Broadwick street showing the John Snow memorial and public house

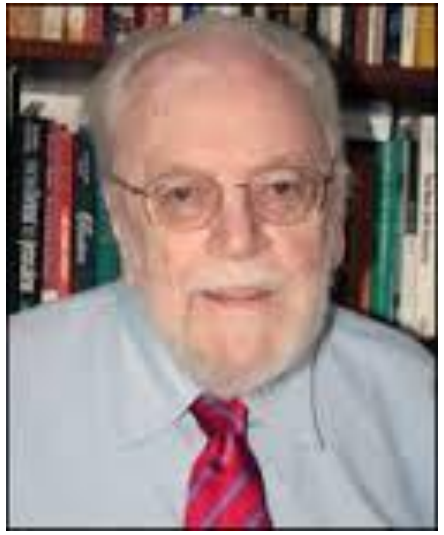
Infectious Disease Era (paradigm: Germ Theory)

- Discovery of causal agent of Antrax, Tuberculosis & Cholera (by Robert Koch)
  - Vibro cholera (**1854**)
  - Bacillus Antrax (**1877**)
  - Mycobacterium Tuberculosis(**1882**)



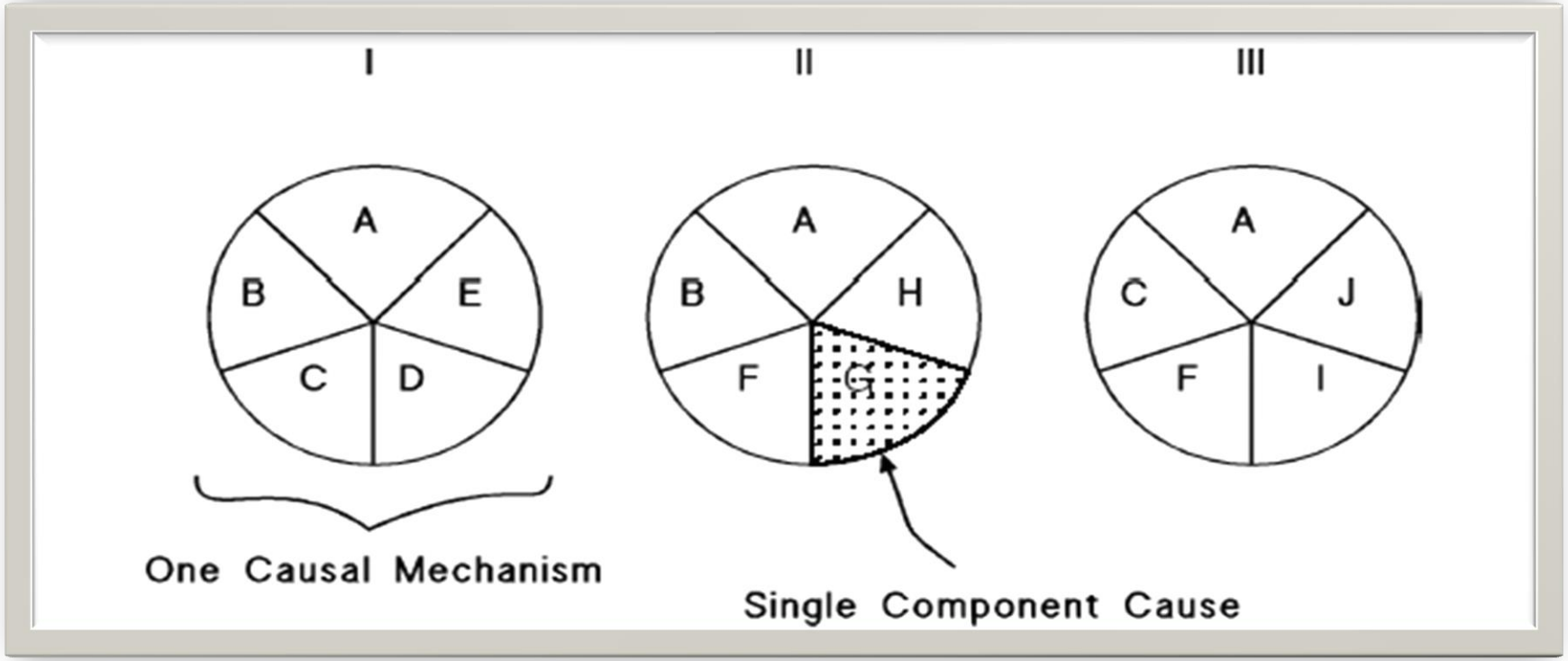
Risk factors Era (Paradigm: Black Box)

- Web causation (Macmahon, 1960)
  - All factors are at the same level
  - Diseases can be preventable by cutting a few strands of the web
  - Does not elucidate societal forces or air related to health



Brian Macmahon

Causal Pie



- **Component causes** (each slice of pie)
- **Necessary causes** (A: in very pie, disease cannot occur without it)
- **Sufficient causes** (a complete pie so disease will occur
- **Prevention**: Remove at least one slice of the pie

Implications of the causal pie

- The "**Strength"** of a component cause depends on the **prevalence in the population** of the other component causes in the causal mechanism
  - A factor will appear to have a strong effect if the other component causes are common in the population
  - A factor will appear to have a weak effect if the other component causes are rare in the population

- **No component** cause/acts alone to produce disease
- Every case of a disease is the result of **multiple component** causes acting jointly in a causal mechanism ("***Pie***")

- The sum of the fractions of a disease attributable to each of its causes ***does not equal to 100%*** but instead has no up limit

Cause

Why worry about Causes

- So that we can **intervene**
- So that we can **reduce or prevent** disease occurrence

Epidemiology Goal

- To learn causes of diseases and factors that could prevent or delay disease development

Cause & Effect

Definitions

Cause:

- An event, condition or characteristic that preceded the disease event and without which the disease event would not have occurred at all or would not have occurred until some later time ***(Rothman & Greenland,1998)***

Definitions (cont.)

- **Cause:** is an event, condition, characteristic (or a combination)which plays an important role/ regular/ predicable change in occurrence of the outcome (e.g. smoking and lung cancer)
  - Cause may be "***genetic***" and or "***environmental***" (e.g. many NCDs including: diabetes, cancers etc)

- **Causal inference:** A process of determining causal and preventive factors

causality

- **Emerson's view on causality**

"Shallow men believe in luck.
Strong men believe in cause and effect"

Ralph W.Emerson's 1803-1882

- Causality is in our psychological habit of witnessing effects that regularly followed causes in times and space ***(David Hume, 1772)***

Causal thinking in health science

A determinant (of health) can be an factor, whether an event, characteristic, or other definable entity so long as it brings about change for better or for worse in a health condition.

Mervyn susser,1973

A factor is a causal of a certain disease when alterations in the frequency or intensity of this factor, without associated alterations in any other factor, are followed by changes in the frequency of occurrence of the disease, after the passage of a certain time period.

Pagona Lagiou,2005

Causal Relationships

- A relationship is considered causal whenever evidence indicates that the factors form part of the complex circumstances which increase the probability of occurrence of disease and that a reduction of one or more of these factors decreases the frequency of disease

- Causal pathway may be **Direct** or **Indirect**

Causal relationship

- **Direct** causation, **X** causes **Y** without intermediate effects

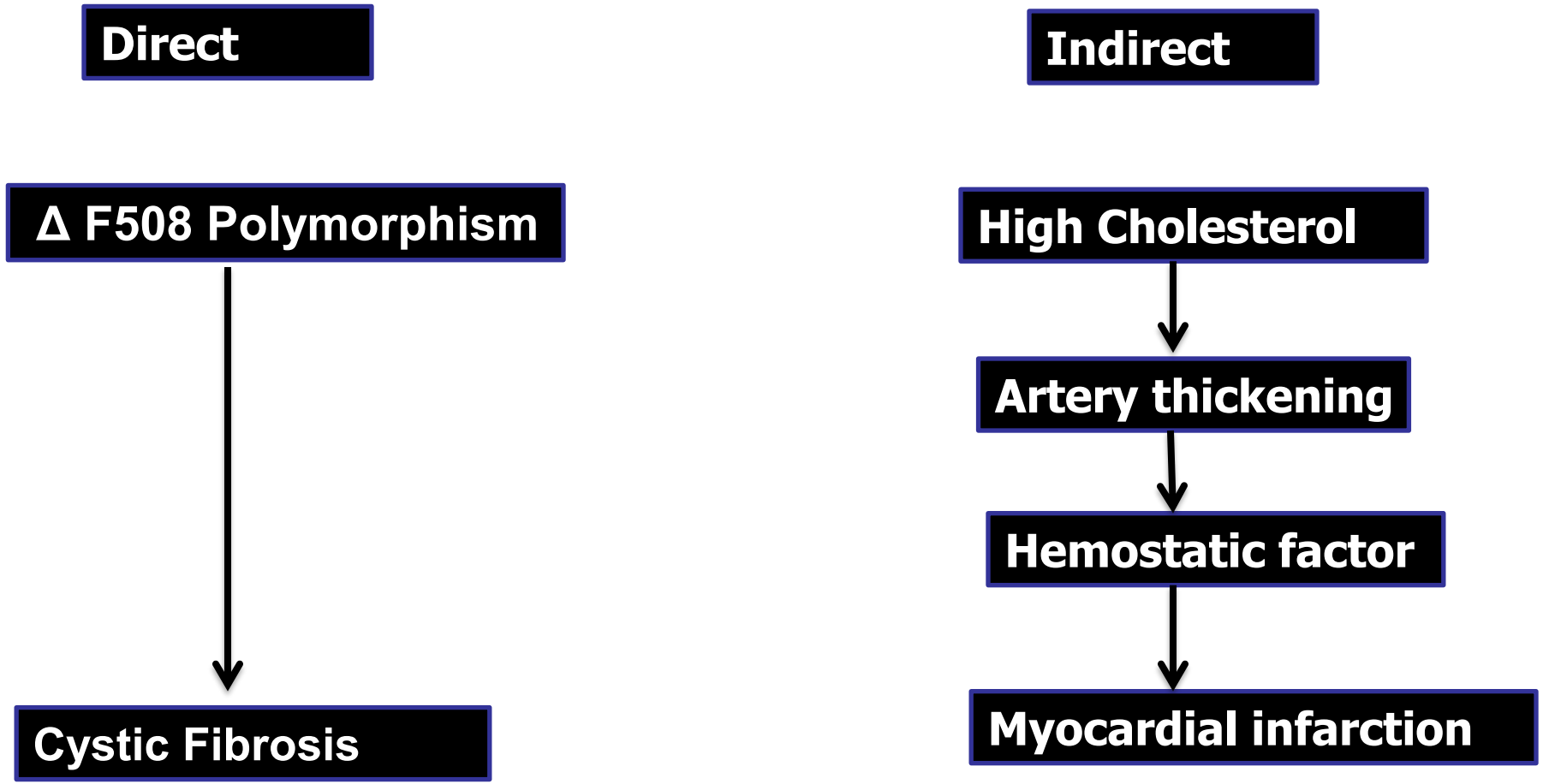
Exposure to the risk factor causes disease directly

- **Indirect** causation, **X** causes **Y**, but with intermediate effects (**I** causes **X**, which in turn causes **Y**

Exposure influences intermediate factor (intervening variable) which in turn cause disease

- In human biology, intermediate steps are virtually always present in any causal process

Direct vs Indirect Causation

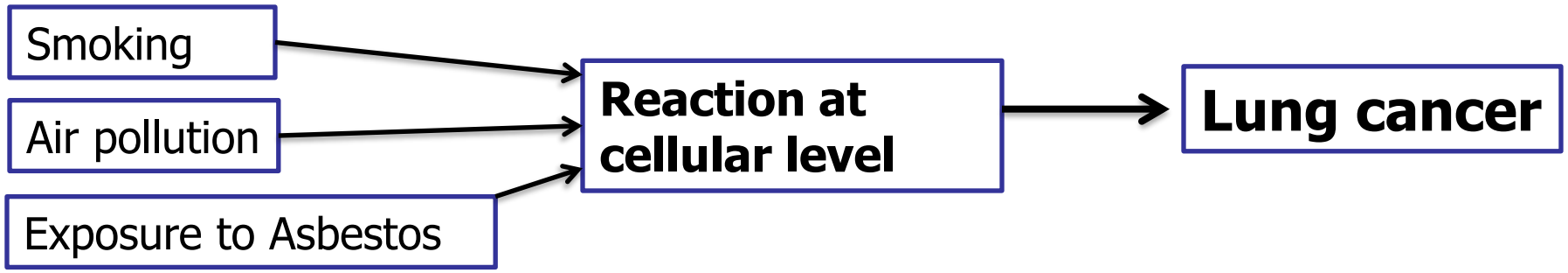


Direct causation

- One-to-one causal association

  X causes **Y** (e.g. Rubeola virus → Measles)

- Multifactorial causation
  - Sufficient and necessary cause
  - Web causation
  - Multiple factor leads to the diseases
  - Common in non-communicable diseases



Multifactorial causation

Exposure to Asbestos	History of tobacco use	Lung cancer death rate per 100 000
No	No	11
Yes	No	58
No	Yes	123
Yes	Yes	602

Age-standardized lung cancer deaths rates (per 100 000 population) in relation to tobacco use and occupational exposure to asbestos dust

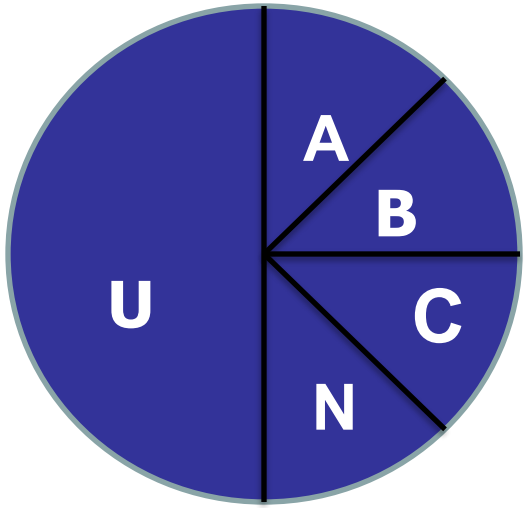
Types of causal relationship

- **Sufficient and but not Necessary:** The factor alone can cause disease, but so can other factors in its absence
  - Benzene or radiation can cause leukemia without the presence of the other factor
- **Neither sufficient nor necessary (supporting component cause:** the factor cannot cause disease on its own, no is it the only factor can cause that disease
  - This is the probable model for chronic disease relationships
  - **Necessary causes + component causes = Sufficient cause**

Types of causal relationship

- **Necessary and Sufficient:** without the factor disease never develops
  - With the factor, disease always develops (this situation rarely occurs)
- **Necessary but not sufficient:** the factor in and of itself is not enough to cause disease
  - Multiple factors are required, usually in a specific temporal sequence (such as carcinogenesis)

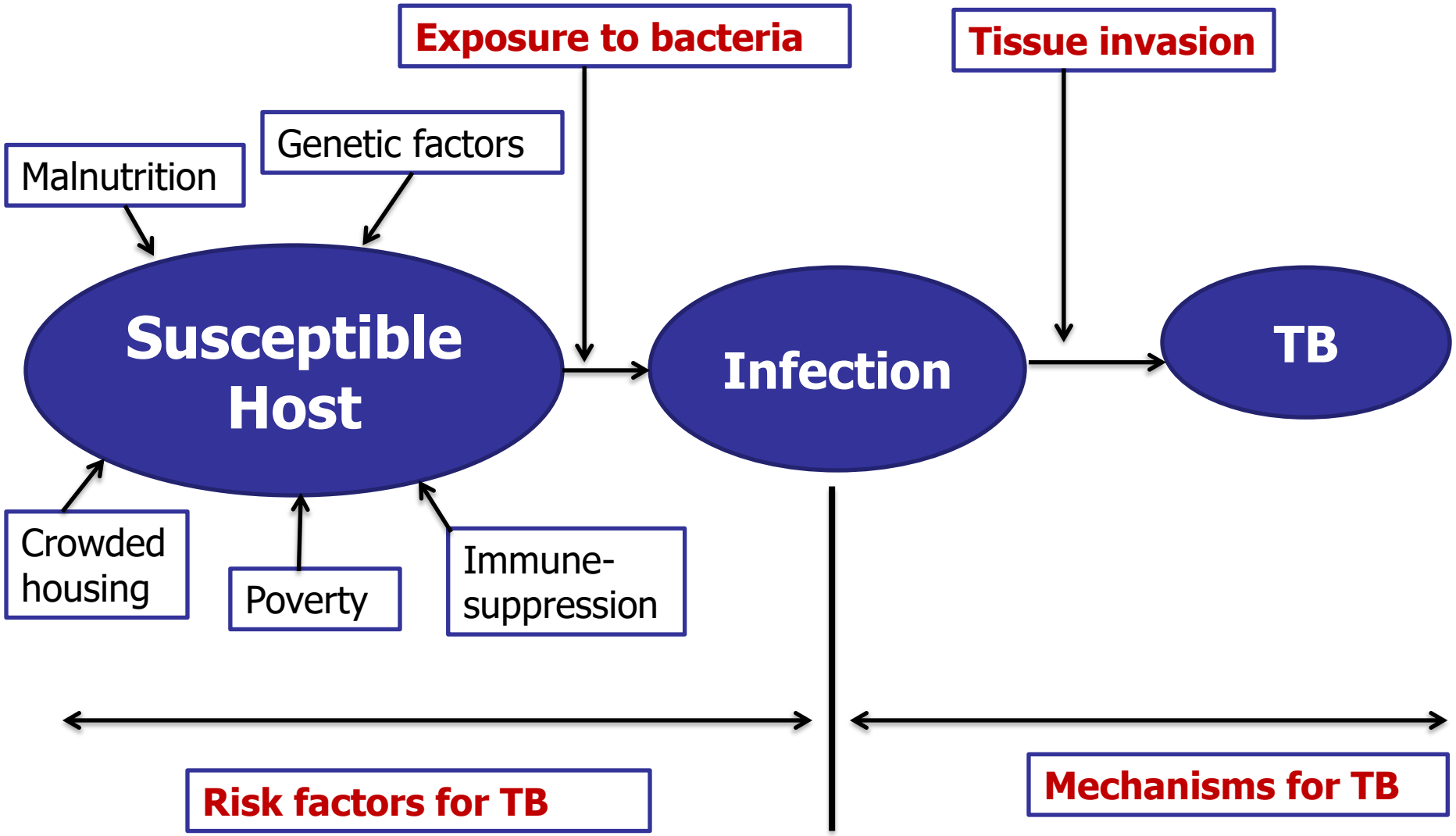
Sufficient & Necessary cause



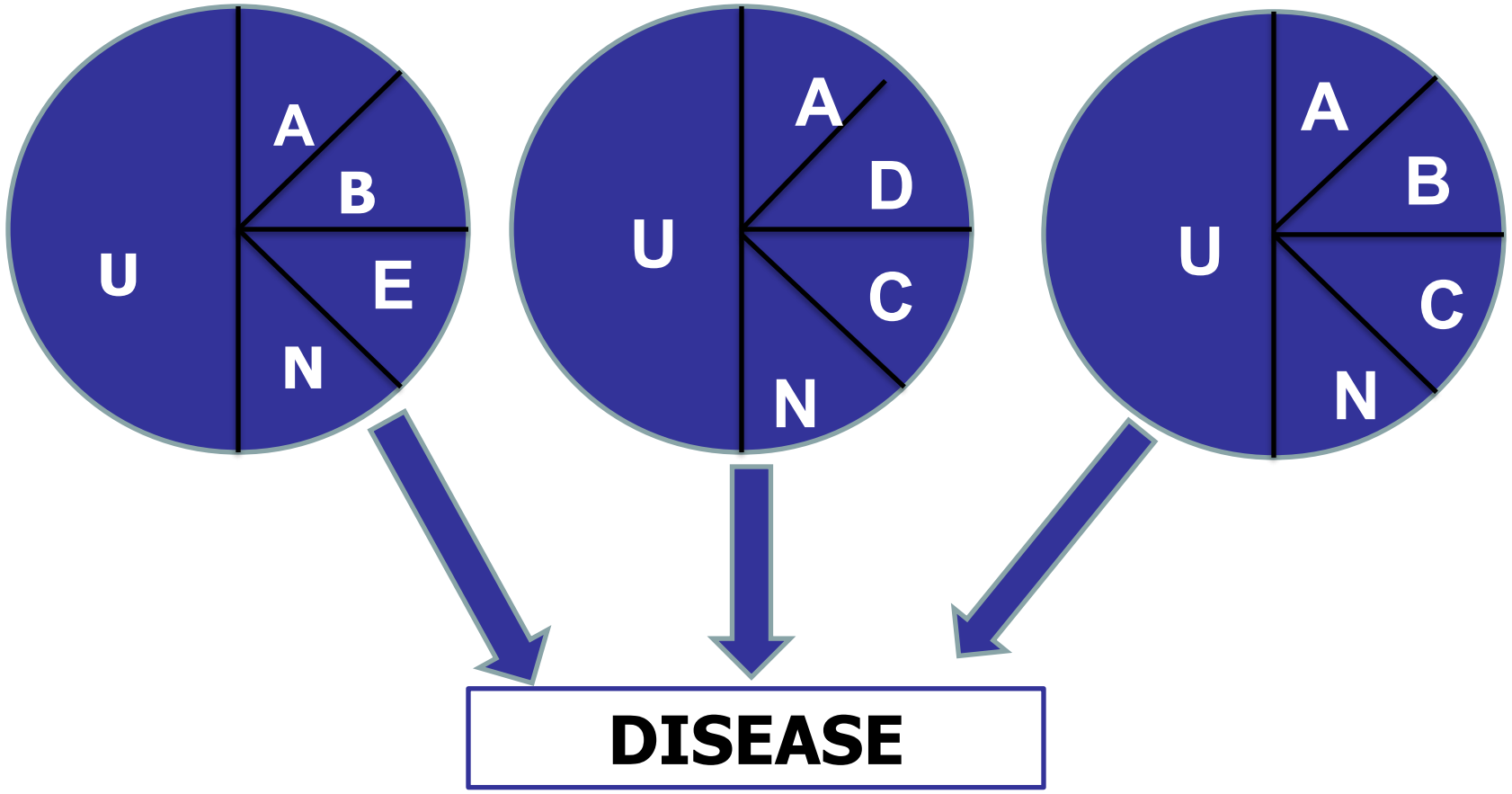
A,B,C,N – known components (causes)
U – unknown component (cause)
N - necessary cause

known component + unknown component cause + necessary cause = sufficient cause

Example: causes of TB



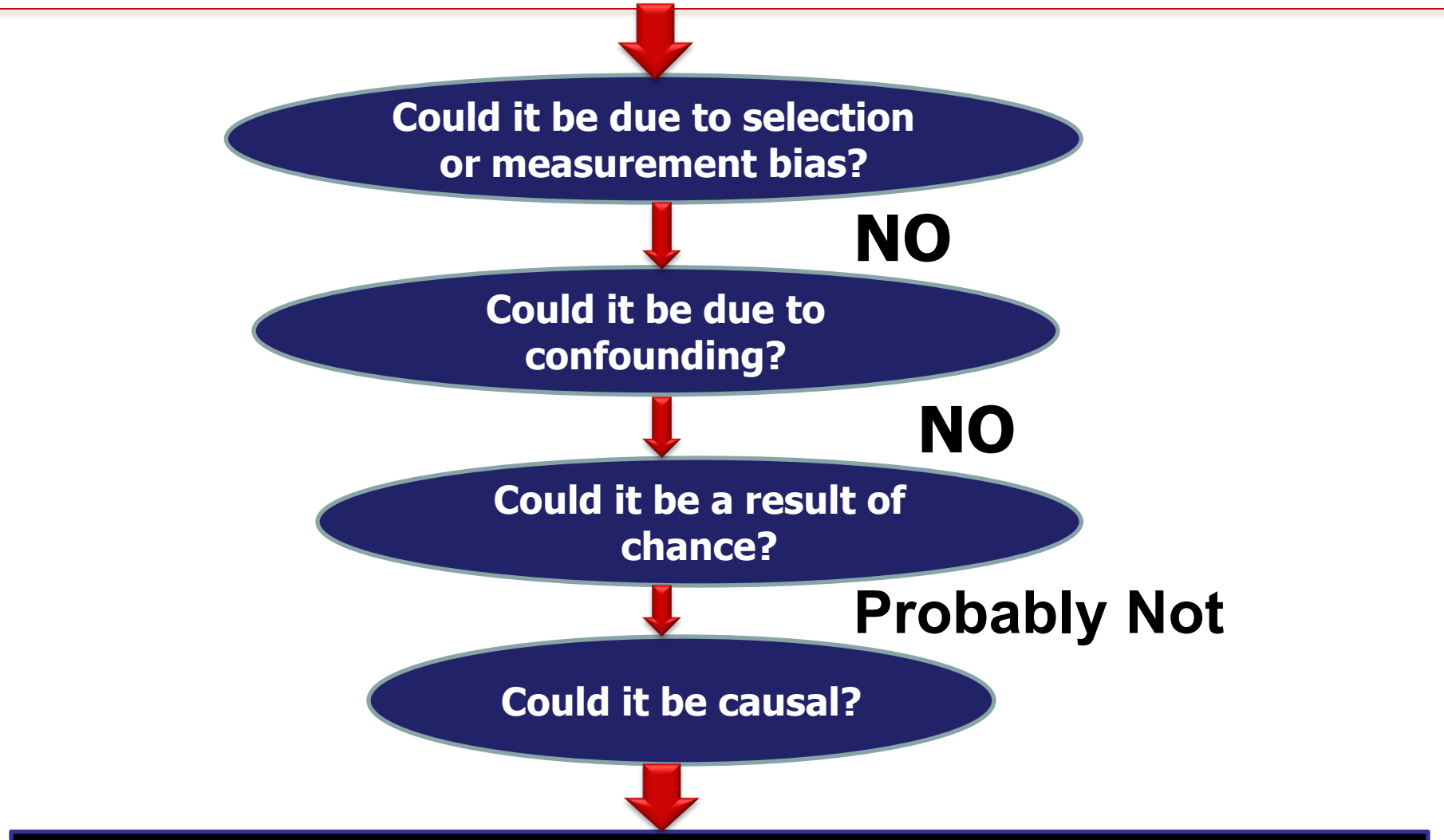
There may be number of sufficient causes for single disease in various combination of component causes, necessary causes



Factors in Causation

- All may be necessary but **rarely sufficient** to cause a particular disease or state
  - **Predisposing:** age, gender, or previous illness may create a state of susceptibility to a disease agent
  - **Enabling:** low income, poor nutrition bad housing or inadequate medical care may favour the development of disease
    - Conversely, circumstances that assist in recovery or in health maintenance may be enabling
  - **Precipitating:** exposure to a disease or noxious agent
  - **Reinforcing:** repeated exposure or undue work or stress may aggravate an established disease or state

How to establish the cause of a disease?
OBSERVED ASSOCIATION?



Criteria for Causal Association

Surgeon General's Report (1964)

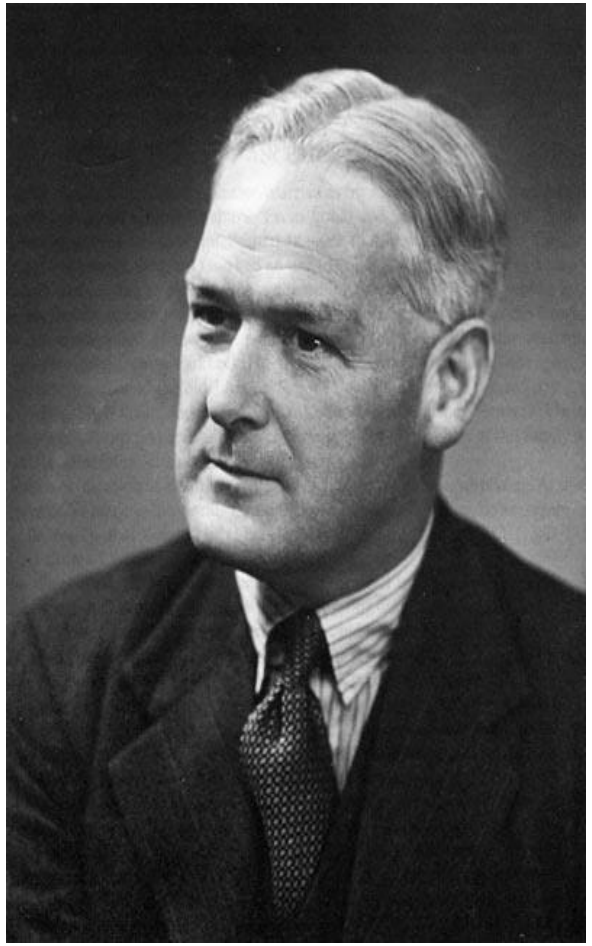
1. Consistency
2. Strength
   — Dose-response
3. Specificity
4. Temporality
5. Coherence

Hill's Criteria (1965)

1. Strength
2. Consistency
3. Specificity
4. Temporality
5. Biological gradient
6. Plausibility
7. Coherence
8. Experiment
9. Analogy

Bradford-Hill's Criteria(1897-1991)

The first complete statement of the epidemiologic criteria of a causality is attributed to **Austin Hill (1897-1991)**



1) Consistency (on replication)
2) Strength (of association)
3) Specificity
4) Dose response relationship
5) Temporal relationship (directionality)
6) Biological plausibility (evidence)
7) Coherence
8) Experiment
9) Analogy

Applying guidelines (Hills criteria/Guidelines for causation) and making judgment regarding causation

Consistency	Have similar result been shown in other studies?
Plausibility	Is the association consistent with other knowledge? (mechanism of action; evidence from experimental animals
Temporal relationship	Does the cause precede the effect? (essential)
Strength	What is the strength of the association between the cause and effect? (Relative Risk)
Dose response relationship (Biological gradient)	Is increased exposure to the possible cause associated with increased effect?
Reversibility	Does the removal of possible cause lead to reduction of disease risk?
Study design	Is the evidence based on a strong study design?
Judging the evidence	How many lines of evidence lead to the conclusion

1. Consistency

- **Is the same association found in many studies?**
- **Different persons** in **Different places**, in **Different circumstances** & **Times** by **Different method** (by various types of studies) is established the **Same result** by **several studies**

Example:

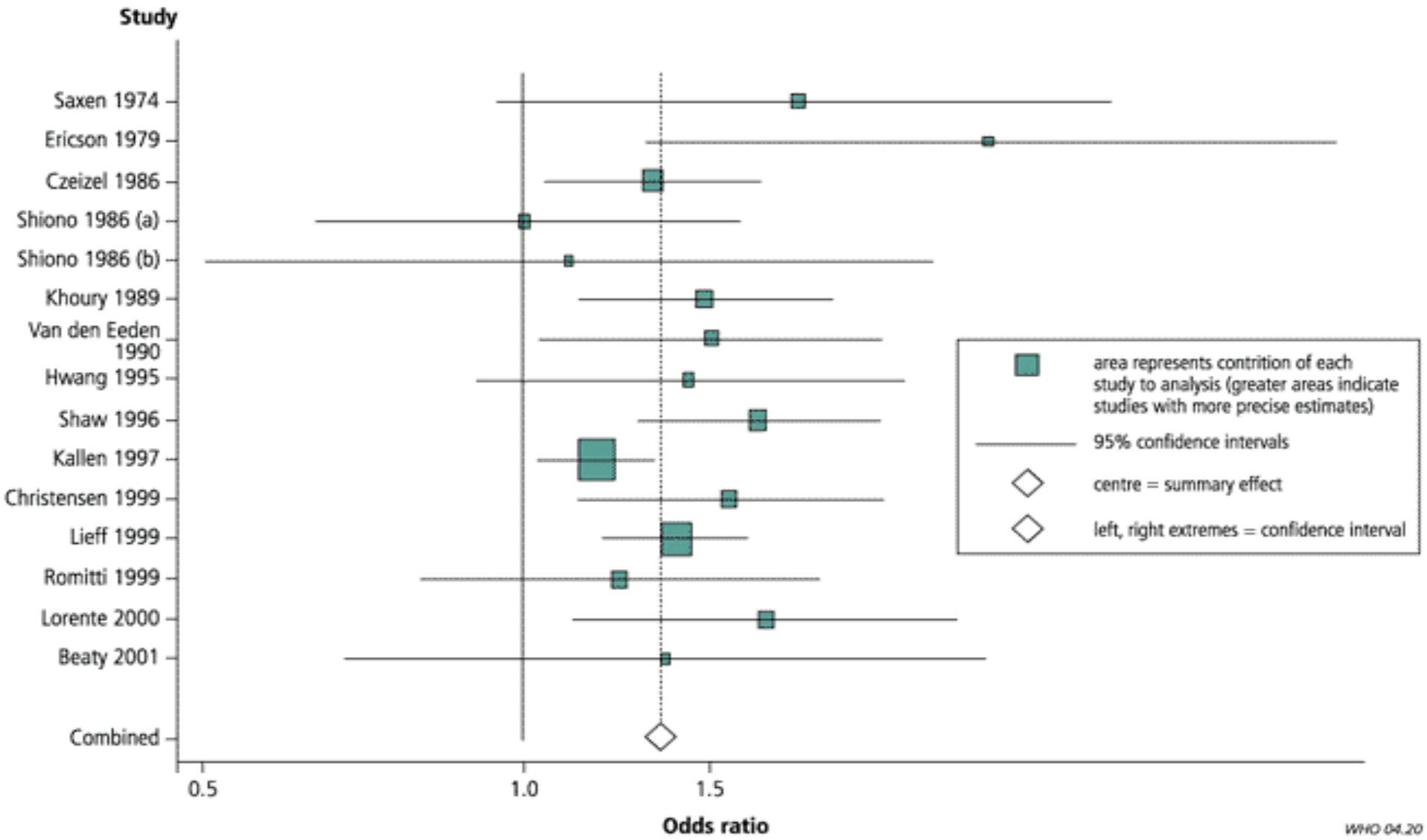
- **Cigarette smoking and lung cancer**
  - **Over 50 respective studies and at least 9 prospective studies have shown the association**

1. Consistency

- **Meta-analysis** is an good method for testing consistency
  - It summarizes odds ratios from various, excludes bias

- Consistency could either mean:
  - **Exact replication** (as lab sciences, impossible in epidemiological studies)
  - **Replication** under **similar** circumstances (possible)

Fig. 1. **Results of fixed effects meta-analysis of the relative risk of cleft lip, with or without cleft palate, in the offspring of mothers who smoked during pregnancy compared with the offspring of mothers who did not smoke** (see ref. *11, 14, 20–31*)



Little J., 2004. Tobacco smoking and oral cleft: meta-analysis

2. Strength of Association

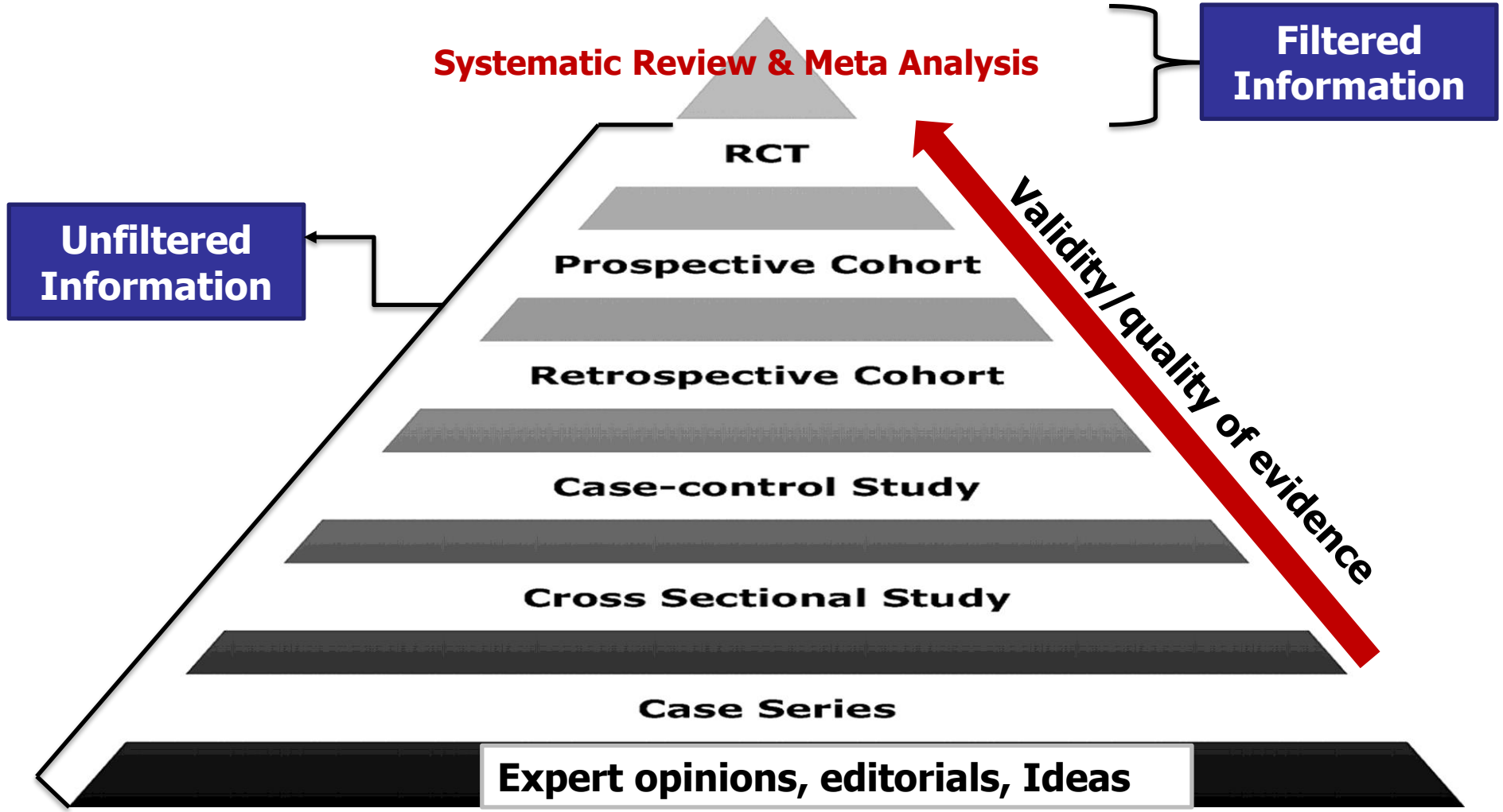
Is the Association Strong? Measured by

- Effect measure (Rate ratio, Risk ratio or Odds Ratio)
  - Away from unity (the higher, the stronger the association
  - The stronger the association the more likely it is to be truly causal
    - Less likely to be due to confounding or bias (known or unknown confounders)
  - Weak association may be causal
    - Measurement error dilutes association
- **P-value** (95% Confidence level): less than 0.05 (the smaller, the stronger the association

Strength of Association

Risk Ratio	Interpretation
< 1	Protective
0.9 – 1.1	No Association
1.2 – 1.6	Weak Causal Association
1.7 – 2.5	Moderate Causal Association
> 2.6	Strong Causal Association

Figure. The pyramid of evidence.



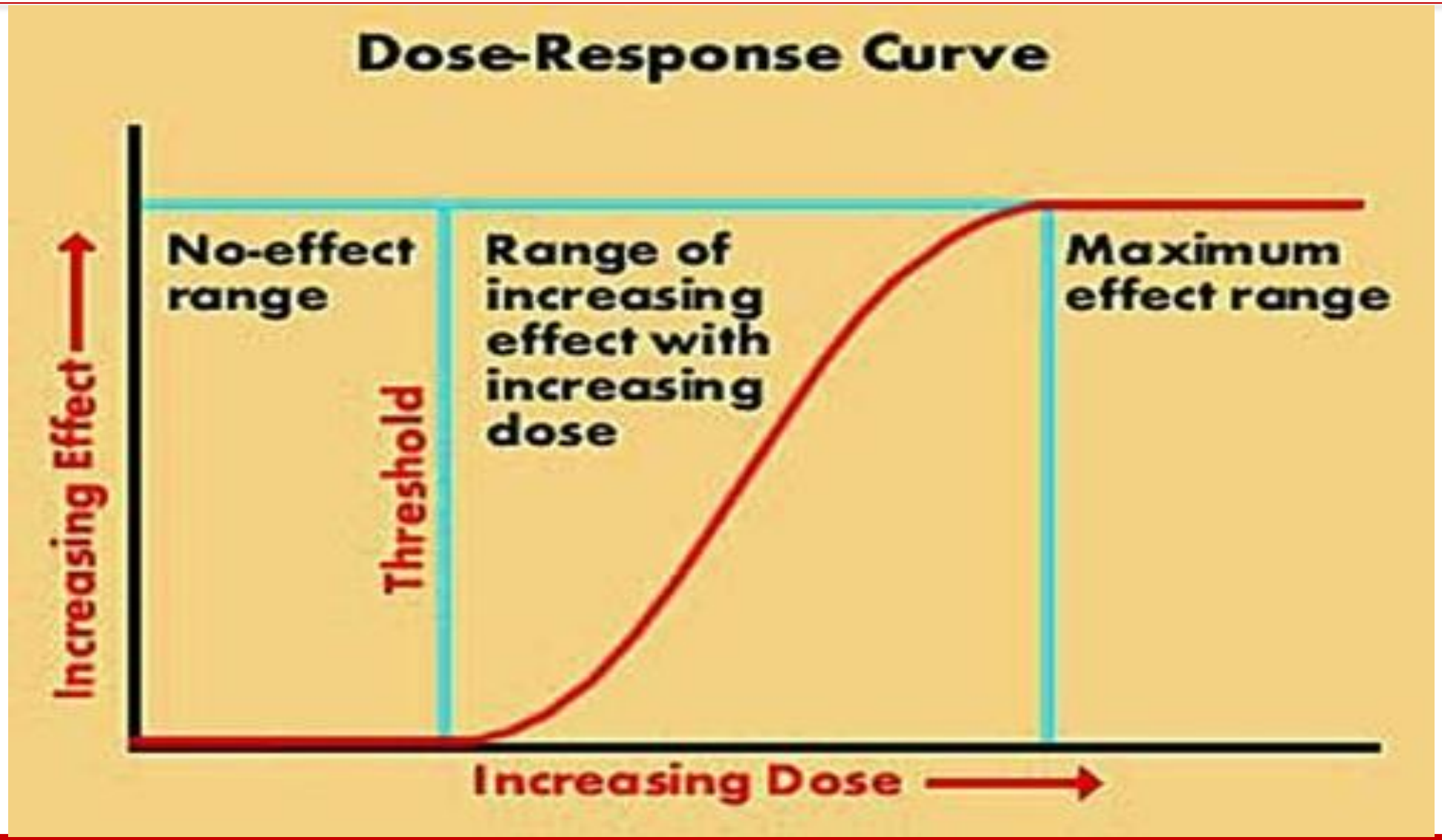
Michael Ho et al. Circulation. 2008;118:1675-1684 & M Hassan Murad et al. Evid Based Med doi:10.1136/ebmed-2016-110401

3. Specificity

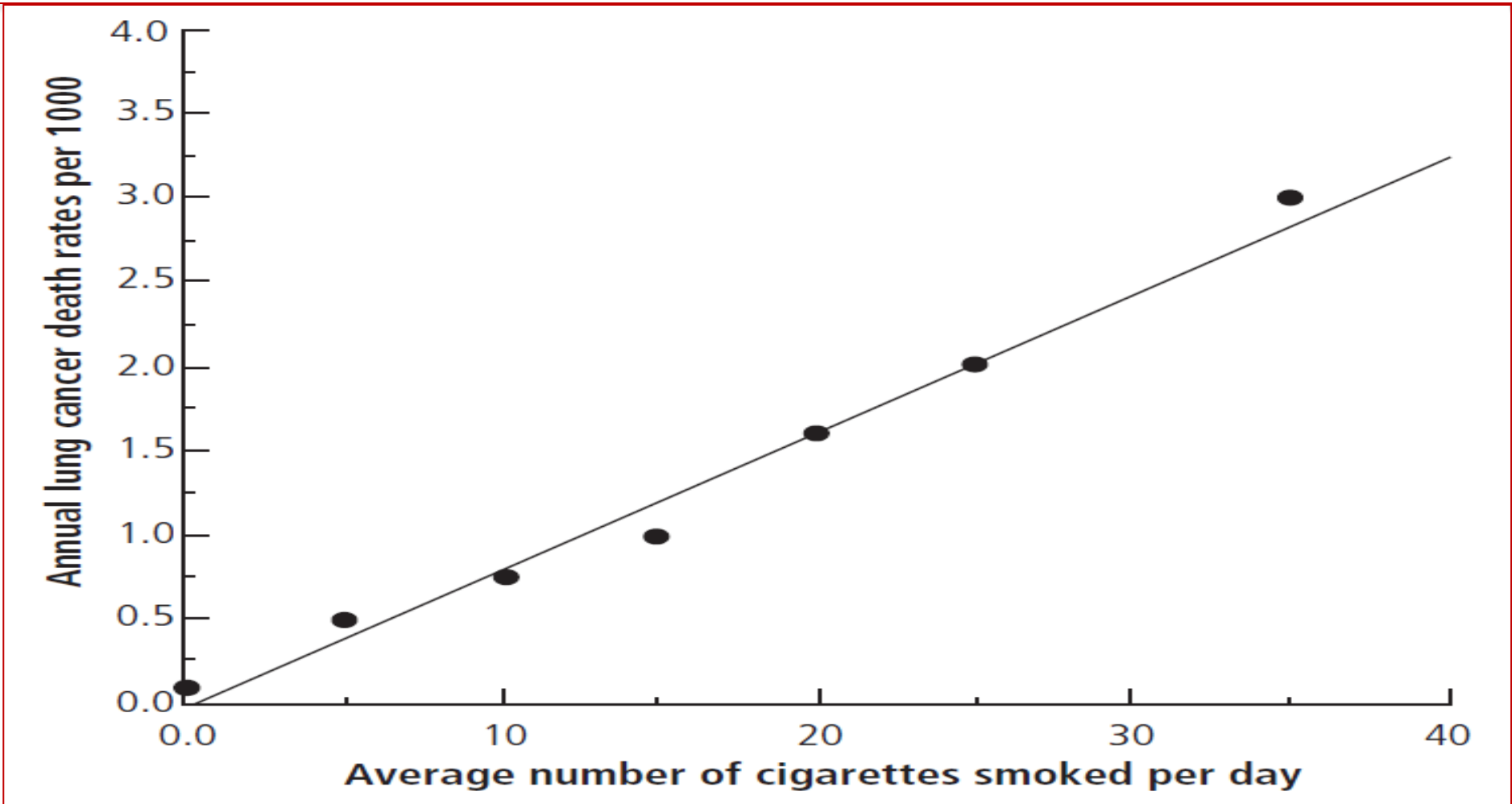
- Occurs when a single factor is associated with a single outcome
  - One to one association

- **Weakest** of all guidelines: Increasingly irrelevant to current models of disease causation
  - Single factor may have many outcomes
    - E.g. smoking is a causal of **Ca** in: lung, oral cavity, larynx, kidney, Oesophagus

4. Dose-Response Relationship (The Biological Gradient)



4. Dose-Response Relationship (The Biological Gradient)



Death rates from lung cancer (per 1000) by number of cigarettes smoked, British male doctors, 1951 –1961

4. Dose-Response Relationship (The Biological Gradient)

- As level of exposure increases, the risk of disease increases
- Observed an increase in the magnitude of risk of outcome with the magnitude of exposure
- Unlikely to be explained by bias or undetected confounding
  - But liver cirrhosis shows strong dose relationship with smoking probably confounded by alcohol
- Lack of biological gradient does not rule out causality
  - Threshold effect
  - **J** or **U** shaped relationship e.g. alcohol

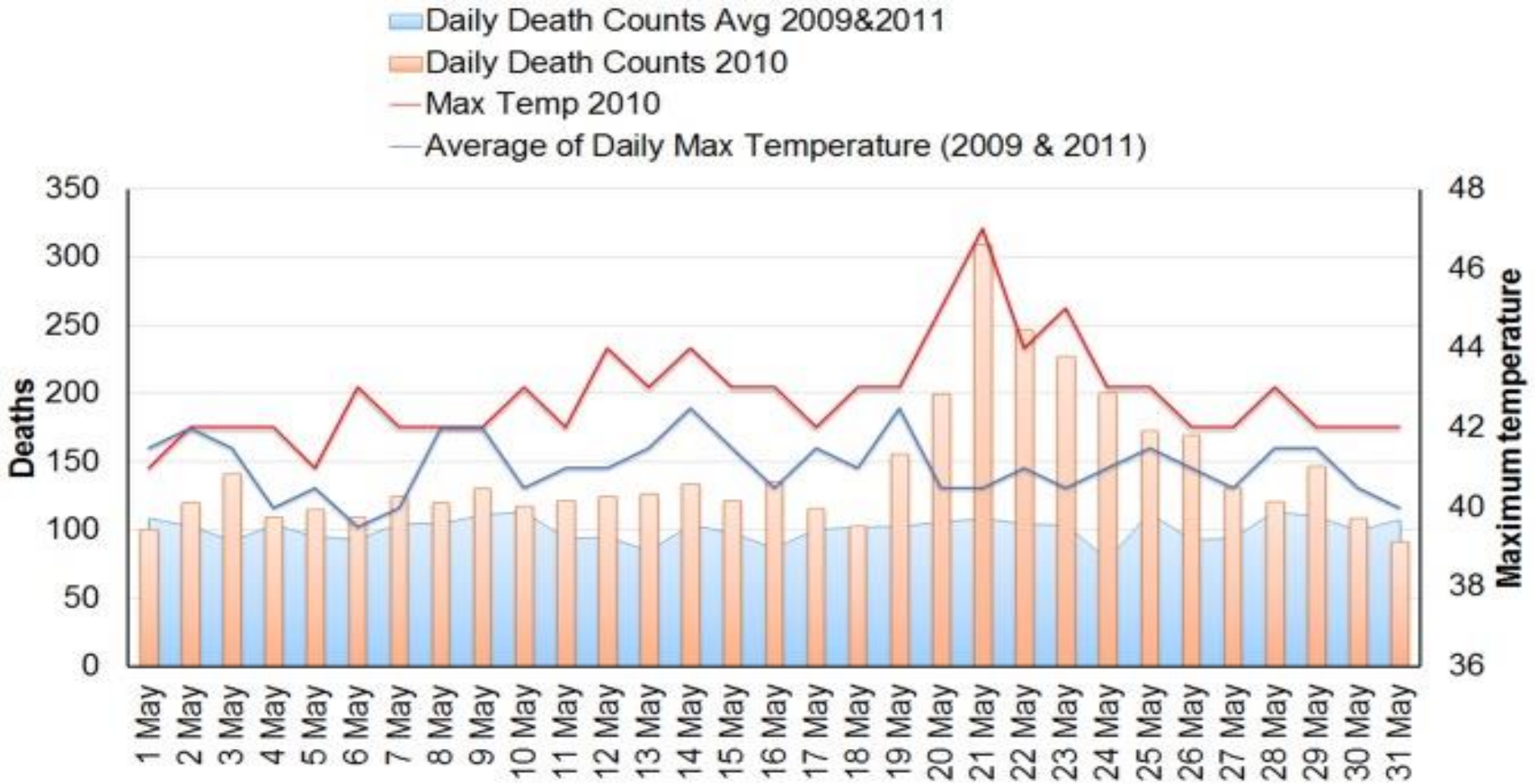
5. Temporal relationship (Directionality)

- The relationship with **Time** (***Time Order***) and **Direction**
  - Drinking contaminated water → Occurrence of diarrhoea

- However many chronic cases, because of insidious onset and ignorance of precise induction period, it become hard to establish a temporal sequence as which comes first – the suspected agent or the disease

- Cause must precede the effect(essential)

5. Temporal relationship (Relationship with Time)



Azhar G.S,. et al, (2014) Heat-Related Mortality in India: Excess All-Cause Mortality Associated with the 2010 Ahmedabad Heat Wave, PLOS one Vol.9, issue 3

6. Biological plausibility

- Consistency with biological knowledge of the day
  - Smoking causing lung cancer
  - Smoking causes skin cancer?

- Lack of plausibility may simply reflect lack of scientific knowledge

7. Coherence

- Cause and effect relationship does not conflict with known natural history and biology of disease
- **Theoretical:** Compatible with pre-existing theory
- **Factual:** compatible with pre-existing knowledge
- **Biological:** compatible with current biological knowledge from other species or other level of organisation
- **Statistical:** compatible with a reasonable statistical model (e.g. dose-response)
- **Correct temporal relationship is essential**; then greatest weight may be given to plausibility, consistency and the dose-response relationship. The likelihood of a causal association is heightened when many different types of evidence leading to the same conclusion

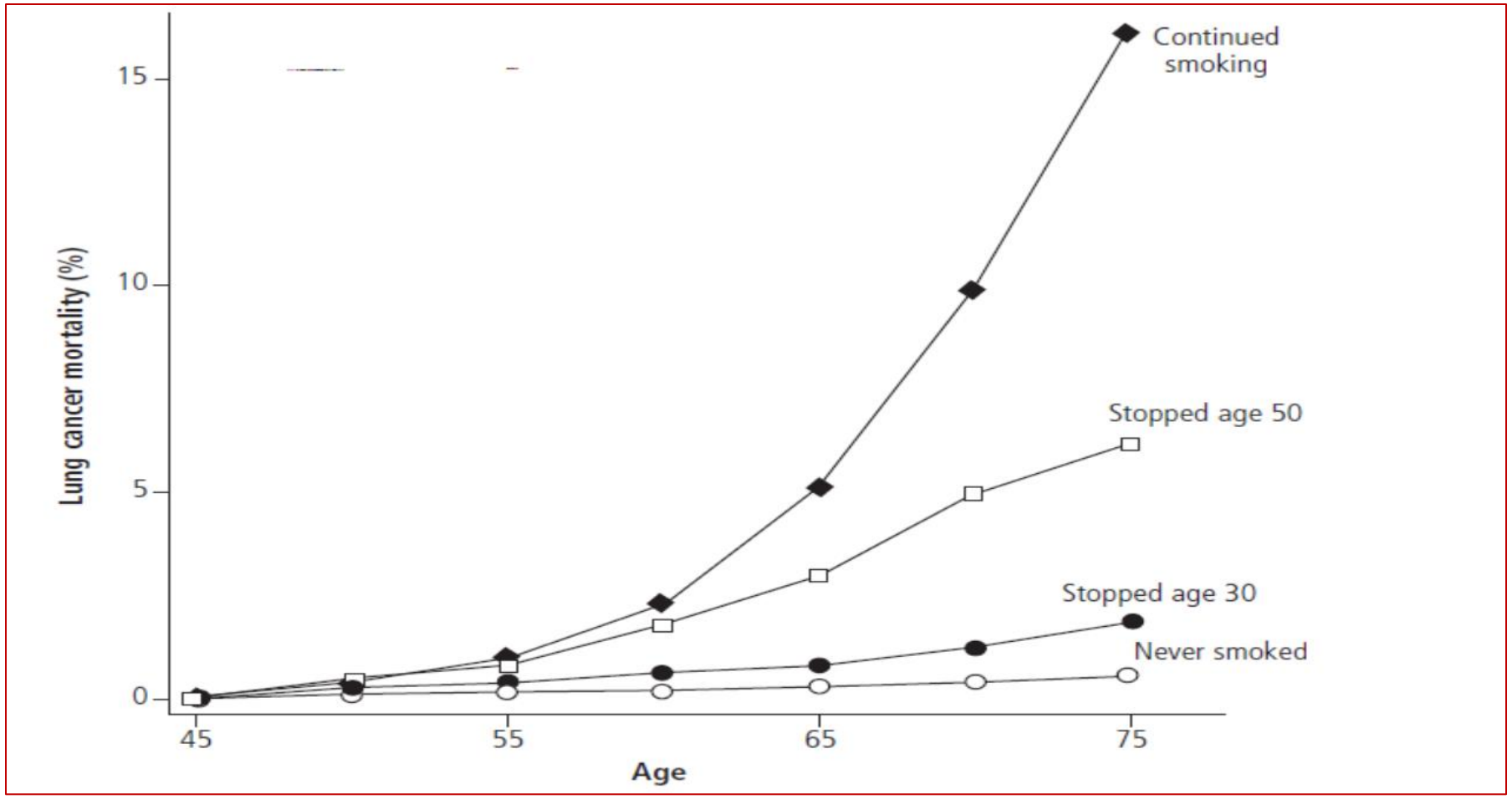
8. Experiment

- **RCT**, the "**Gold Standard**" for determining causal association

- **Reversibility:** Removal of exposure leads to a reduction in the risk of the outcome

8. Experiment: Removal (Reversibility)

Stopping works: cumulative risk of lung cancer mortality



9. Analogy

- Judging by Analogy
- Similarities between observed association and other associations

DERIVING CAUSAL INFERENCES: EXAMPLE

- Until the 1980s, the major causes of peptic ulcer disease were considered to be **stress and lifestyle factors**, including smoking

- In 1984, **Australian physicians** Drs. Barry J. Marshall and J. Robin Warren reported that they had observed small curved bacteria colonizing the lower part of the stomach in patients with gastritis and peptic ulcers

- After several attempts, Marshall succeeded in cultivating a hitherto unknown bacterial species (later named *Helicobacter pylori*) from several of these biopsies

He drank a Petri dish containing cultured *H. pylori*, expecting to develop, perhaps years later, an ulcer. He was surprised when, only three days later, he developed vague nausea and halitosis

- Together they found that the organism was present in almost all patients with gastric inflammation or peptic ulcer

- Many of these patients had biopsies performed which showed evidence of inflammation present in the gastric mucosa close to where the bacteria were seen

- Based on these results, they proposed that *Helicobacter pylori* is involved in the etiology of these diseases. It was subsequently shown that the ulcer was often not cured until *Helicobacter pylori* had been eliminated.

Assessment of the Evidence Suggesting *Helicobacter pylori* as a Causative Agent of Duodenal Ulcers

1. Temporal relationship

- *Helicobacter pylori* is clearly linked to chronic gastritis. About 11% of chronic gastritis patients will go on to have duodenal ulcers over a 10-year period.

2. Strength of the relationship

- *Helicobacter pylori* is found in at least 90% of patients with duodenal ulcer.

  In at least one population reported to lack duodenal ulcers, a northern Australian aboriginal tribe that is isolated from other people, it has never been found.

3. Dose-response relationship

- **Density** of *Helicobacter pylori* per square millimeter of gastric mucosa is higher in patients with duodenal ulcer than in patients without duodenal ulcer

4. Replication of the findings

- Many of the observations regarding *Helicobacter pylori* have been replicated repeatedly

5. Consideration of alternate explanations

- Data suggest that smoking can increase the risk of duodenal ulcer in *Helicobacter pylori*–infected patients but is not a risk factor in patients in whom *Helicobacter pylori* has been eradicated

6. Biologic plausibility

- Although originally it was difficult to envision a bacterium that infects the stomach antrum causing ulcers in the duodenum, it is now recognized that *Helicobacter pylori* has binding sites on antral cells and can follow these cells into the duodenum.
- *Helicobacter pylori* also induces mediators of inflammation.
- *Helicobacter pylori*–infected mucosa is weakened and is susceptible to the damaging effects of acid.

7. Cessation of exposure

- Eradication of *Helicobacter pylori* heals duodenal ulcers at the same rate as histamine receptor antagonists.
- Long-term ulcer recurrence rates were zero after *Helicobacter pylori* was eradicated using triple-antimicrobial therapy, compared with a 60% to 80% relapse rate often found in patients with duodenal ulcers treated with histamine receptor antagonists.

8. Specificity of the association

- Prevalence of *Helicobacter pylori* in patients with duodenal ulcers in 90% to 100%. However, it is found in some patients with gastric ulcer and even in asymptomatic individuals.

9. Analogy (Consistency with other knowledge)

- Prevalence of *Helicobacter pylori* infection is the same in men as in women. The incidence of duodenal ulcer, which in earlier years was believed to be higher in men than in women, has been equal in recent years.
- The prevalence of ulcer disease is believed to have peaked in the latter part of the 19th century, and the prevalence of *Helicobacter pylori* may have been much higher at that time because of poor living conditions.

Causal Criteria of Bradford-Hills

Criteria	Problems with the criteria
1. Strength	Strength depends on the prevalence of other causes and, thus, is not a biologic characteristic; could be confounded.
2. Consistency	Exceptions are understood best with hindsight.
3. Specificity	A cause may have many effects.
4. Temporality	It may be difficult to establish the temporal sequence between cause and effect.
5. Dose response relationship (Biological gradient)	Could be confounded; threshold phenomena would not show as progressive relation.
6. Plausibility	Too objective.
7. Coherence	How does it differ from consistency or plausibility?
8. Experimental evidence	Not always available.
9. Analogy	Analogies abound.

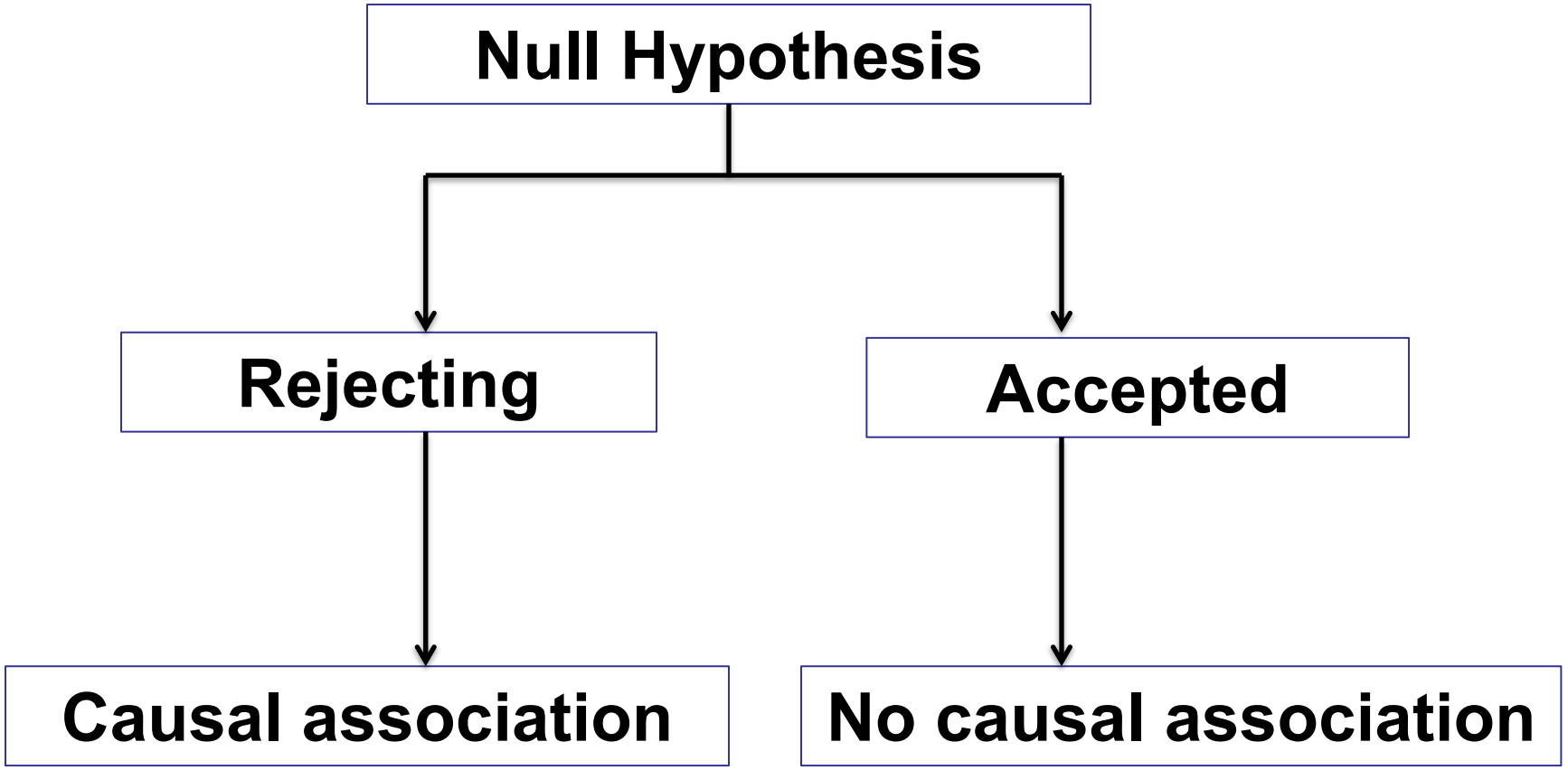
Comparing Rules of Evidence

Criminal Law	Causation
Criminal present at scene of crime	Agent present in the disease
Premeditation	Causal events precede onset of disease
Accessories involved in the crime	Co-factors and/or multiple causality involved
Severity of crime to state of victim	Susceptibility and host response determine severity
Motivation –there must be gain to the criminal	The role of the agent in the disease must make biologic and common sense
No other suspect could have committed the crime	No other agent could have cause disease under the circumstances given
Proof of guilty must be established beyond a reasonable doubt	Proof of causation must be established beyond reasonable doubt or role of chance

Analytical Methods

- **Measures of association / strength of association**
- **Testing hypothesis of association**
- **Controlling confounders**

Testing hypothesis of Association



Controlling confounders

- At **Time designing** of epidemiological study or while carrying study
  - Randomisation
  - Restriction
  - Matching

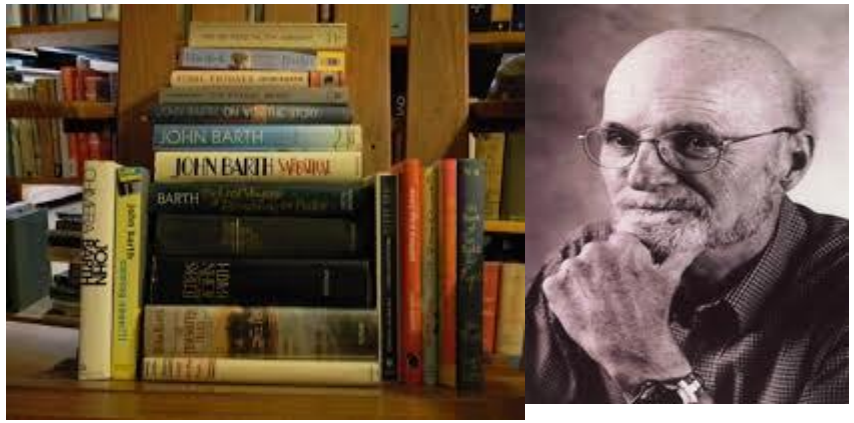
- **At Analysis Stage**
  - Stratification
  - Adjustment
  - Statistical modelling

Measure of association / strength of association

- Ratio measures
  - Relative risk
  - Odds ratio

- Difference measures
  - Attributable risk
  - Population Attributable risk

“The world is Richer with Associations than Meanings, and it’s the part of wisdom to differentiate the two”



John Barth (novelist

References

1. Hill AB. The environment and disease: Association or Causation? Proc R. Soc.Med.1965;58:295-300
2. Susser MW. What is a cause and how do we know one? A grammar for pragmatic Epidemiology. American Journal 1991;133:635-648
3. Gordis L. Epidemiology (2009) (4th Ed) chapter 14: From Association to causation: Deriving Inferences from Epidemiologic studies
4. Beaglehole R, Bonita R., (2006) Basic Epidemiology. Delhi: AITBS publisher & Distributor
5. Last J., A dictionary ofEpidemiology, 4th Ed.
6. Galea S, et al.,Causal thinking & Complex sytem approach in epidemiology. International Journal of Epidemiology, 2010 Feb; 39 (1):97-106
7. Rothman KJ, Greenlan S,et al., Modern Epidemiology. 3rd Ed. New Delhi