SYNOPSIS IN GENERAL EPIDEMIOLOGY

First Edition, 2024

Authors

Saad Hamdallah Kharabsheh
MD, MSc Epi., Board C.M, FRIPHH

Mohannad Abed Al Fattah Al Nsour
MD, MSC, FETP, PhD, FFPH

SYNOPSIS IN GENERAL EPIDEMIOLOGY

First Edition, 2024

رقم الإيداع لدى دائرة المكتبة الوطنية (2024/7/3875)

Primary indexing data for the book

Book Title	Synposis in General Epidemiology
Author	Kharabsheh, Sa'ad H.S
Author (others)	AlNsour, Mohannad Abed Alfattah Ahmad
Publication data	Amman: Mohannad Abed Alfattah Ahmad AlNsour, 2024
Physical description	219 Pages
Classification number	614.4
Descriptors	/Epidemics/Infectious Diseases/Health Protection/
Edition data	First Edition

يتحمل المؤلف كامل المسؤولية القانونية عن محتوى مصنفه ولا يعبّر هاذا المصنف عن رأي دائرة المكتبة الوطنية أو أي جهة حكومية أخرى.

Biography of the Authors

First Author: Dr. Saad Hamdallah Kharabsheh

Qualifications:

1. Fellow of the Royal Institute of Public Health and Hygiene (FRIPHH), UK, 2000
2. Jordanian Board in Community Medicine/Epidemiology, 1995
3. Master's degree in Epidemiology, Ain-Shams University, Egypt, 1987
4. MBChB, Baghdad University, Iraq, 1975

Work experience

- Minister of health, Jordan, 2006–2007
- Under-secretary of Health, Ministry of Health, Jordan, 2003–2006
- General Director of Primary Health Care, Ministry of Health, Jordan, 1999–2003
- Director of Disease Prevention & Control, Ministry of Health, Jordan, 1992–1999
- Practicing physician in the field of primary healthcare and public health, 1976–1991

Research experience

Conducted and participated in many published and unpublished research activities

International experience

- Worked several times as a short-term consultant or temporary advisor with the World Health Organization (WHO), UNICEF, International Organization of Migration, and Non-governmental organizations in different fields of health.
- Awarded the Dr. Shousha Foundation Prize for 2006 of WHO for his work and achievement in promotion of health programs in Jordan and countries of the Eastern Mediterranean Region.

Second Author: Mohannad Abed Al Fattah Al Nsour

Qualifications:

1. Fellowship through Distinction, Faculty of Public Health (FPH), United Kingdom, 2017 to present
2. PhD in Epidemiology, International Agency for Research on Cancer (IARC) and Glasgow University, Scotland, UK, 2009–2014
3. Jordanian Board, Community Medicine/Epidemiology, Jordan Medical Council, Jordan, 2005
4. MSc. in Epidemiology (Scholarship Student), American University of Beirut, Lebanon, 2003–2005
5. FETP, Field Epidemiology Training Program Graduate, Jordan, 2001–2003
6. Diploma in Community Medicine and Epidemiology, Jordan, 2000–2001
7. Bachelor of Medicine, Bachelor of Surgery (MB, BSc.), Pirogov Medical Institute, Odesa, Ukraine, 1990–1997

Work experience:

- Executive Director – Global Health Development (GHD),|Eastern Mediterranean Public Health Network (EMPHNET), Jordan, 2009–present
- Consultant with CDC for the FETP in Eastern Mediterranean Region, Jordan, 2008–2009
- Director, Field Epidemiology Training Program (FETP), Ministry of Health, Jordan, 2006–2009
- Head of the Surveillance Department, Balqa Health Directorate, Jordan, 2005–2006
- Resident Jordan Field Epidemiology Training Program, CDC affiliated, Jordan, 2001–2003
- Head of Surveillance Department, Ministry of Health, Balqa Health Directorate, Jordan, 1999–2001

International experience:

Dr. Al Nsour serves on several regional and global initiatives, association and networks as he is the vice-chair of the Steering Committee for the Global Outbreak Alert and Response Network (GOARN), and a member in the Steering Committee of Roadmap on Public Health and Emergency Workforce (WHO), FETP Enterprise's Strategic Leadership Group (SLG), IHR Review Committee of IHR amendments, Middle East and North Africa Health Policy Forum (MENA HPF), and many others. He is also the founder of the International Academy of Public Health (IAPH).others. He is also the founder of the International Academy of Public Health (IAPH).

Preface to the First Edition

Although the medical and public health literature is rich in the publications of many textbooks of epidemiology, public health, and community medicine, this book's authors felt, based on their years of practical experience as field epidemiologists in the public health field, that it was time to prepare an easily understandable book in the language of lectures which focuses on the principles and outlines of general epidemiology. It should be understandable to under- and post-graduate students studying in various public health fields. It should also suit all types of health workers and health practitioners of different specialties and help them to become oriented and familiar with the general principles of disease occurrence, methods of epidemiological studies, public health surveillance, and epidemic investigation. This little book will act as a handbook for many practitioners and especially those who are undertaking training in the field of epidemiology and public health.

The COVID-19 pandemic taught us that every health practitioner, particularly those working in the public health sector, should have a background and knowledge of epidemiological measures and investigations. The book consists of ten chapters covering most general epidemiology areas as outlined in the table of contents. Examples and exercises were included in each chapter to simplify the text and make it easily understandable.

At the end, the book also contains a list of self-assessment questions with their answers to help the reader better understand the different subjects and to give students an insight into their preparation to enter epidemiology examinations. This text has been primarily written with the needs of medical students in mind, however, students of other health professions such as nursing, dentistry, pharmacy, and veterinary medicine should find it suitable for their needs as well.

Upon completion of this book, the reader should be able to calculate and interpret basic epidemiologic measures, recognize the strengths and limitations of various study designs, understand the concepts of variability and bias, and critique published epidemiologic studies. In addition, they should have acquired basic skills in epidemiologic surveillance and epidemic investigation.

The guiding principle in this book's development is the presentation of epidemiology in a manner that is both understandable and interesting to the reader. The scope of topics is limited to core principles and concepts, thus reducing the text's overall length. In contrast to other introductory books, special and detailed emphasis was devoted to enabling various public health workers, including general practitioners, to acquire the essential skills to be efficient members of the epidemiologic investigation team and apply these skills in conducting public health surveillance and epidemic investigation. Figures are used extensively to promote comprehension and retention of the material, and only essential formulas with illustrative calculations are included in the text. Many references in the form of textbooks or relevant websites were utilized to enrich the preparation of this work as well.

Finally, we hope that those who read this book will find it a useful introduction to the subject and that it stimulates their desire to learn more about it.

Acknowledgement

First of all, we sincerely thank God who gave us the power and patience over a period of two years' time to accomplish the hard work needed for the completion of this book.

The authors gratefully acknowledge and appreciate the contributions of many colleagues (whose names are listed in the following list of reviewers) who spent time reviewing the different chapters of this book for their valuable comments and advice, they were very generous with their time and talents in revising the manuscript. Their comments, suggestions and views helped a lot in the enrichment of the scientific content of the text and improvement of the language of this manuscript. They have often suggested specific comments that have helped clarify many of the concepts discussed.
We highly appreciate the efforts of technical and administrative staff of EMPHNET; namely; Farah Khalifeh, Rawan Al Araj, Mohammad Al Zraiqi, Lara Kufoof and Isra Noufal, and Aliakbar Mohsin who helped a lot in editing and designing the text of this book.

List of Reviewers

1. Abdulwahed Al Serouri Yemen Filed Epidemiology Program, Ministry of Public Health and Population, Sana'a, Yemen
2. Anwar Batieha, Prof. of Epidemiology and Public Health, Faculty of Medicine, Jordan University of Science and Technology
3. Biagio M. Pedalino, MD, Dr. PH, EPIET (EU-FETP), Global Senior Consultant in Public Health & Field Epidemiology
4. Dr. Carl Reddy, Training Programs in Epidemiology and Public Health Interventions Network (TEPHINET), a program of The Task Force for Global Health, Atlanta, USA
5. Elfatih Mohamed Malik (NBBS, MD, FPH-UK), Associate Professor, Fellow at Faculty of Medicine University of Khartoum
6. Faris Lami, MBChB, PhD, FFPH, International Academy of Public Health
7. Professor Dr. Mahmudur Rahman, MBBS (Bangladesh), MPHM (Thailand), PhD (Cambridge, UK), Country Director at Eastern Mediterranean Public Health Network, Bangladesh office
8. Hanan F. Abdul Rahim, M.Sc., PhD, CPH, Associate Professor (Epidemiology), Department of Public Health, College of Health Sciences, QU Health, Qatar University
9. Maha El Rabbat, Prof. of Public Health, Faculty of Medicine, Cairo University
10. Mohamed Chahed, Professor of Epidemiology and Public Health, Faculty of Medicine of Tunis, El Manar University, Tunisia
11. Dr. Mohammed Youbi, MD-MPH, Director of Epidemiology and Disease Control, Ministry of Health and Social Protection, Morocco
12. Dr. Rana Jawad Asghar, Chief Executive Officer of GHSI. He is among the WHO IHR 2005 Experts in Epidemiology and Surveillance and US CDC EIS alumni and adjunct Prof Epi at University of Nebraska Medical Center
13. Scott J.N. McNabb, Ph.D., MS, Research Professor, Emory University, Rollins School of Public Health, Managing Partner, Public Health Practice, LLC.
14. Yousef Khader, CEAE Director, Center of Excellence for Applied Epidemiology at Eastern Mediterranean Public Health Network (EMPHNET)

Dedication

This book is dedicated to our parents who worked hardly to form our future and to the distinguished persons who taught us the science of epidemiology.

Contents

Abbreviations

NGOs: Non-governmental organizations
CDR: Crude death rate
SBR: Stillbirth rate
CMF: Comparative mortality figure
SMR: Standardized mortality ratio
IMR: Infant mortality rate
CBR: Crude birth rate
TFR: Total fertility rate
GRR: Gross reproduction rate
NRR: Net reproduction rate
GFR: General fertility rate
GMFR: General marital fertility rate
ASFR: Age-specific fertility rate
ASMFR: Age-specific marital fertility rate
TMFR: Total marital fertility rate
GRR: Gross reproductive rate
SAABR: Sex age adjusted birth rate
UN: United Nations
CWR: Child-women ratio
CVD: Cardiovascular disease
NCDs: Non-communicable diseases
OCs: Oral contraceptives
PCP: Pneumocystis carinii pneumonia
AIDS: Acquired immune deficiency syndrome
USA: United States of America
CHD: Coronary heart disease
UK: United Kingdom
OR: Odds ratio
RR: Relative risk
CRTs: Cluster randomized trials
CI: Confidence interval
FGDs: Focus group discussions
TB: Tuberculosis
DOTS: Directly observed therapy for the treatment of tuberculosis
PKU: Phenyl ketone urea
HIV: Human immune deficiency virus
PPV: Positive predictive value
NPV: Negative predictive value
HIT: Herd immunity threshold
HIL: Herd immunity level
MMR: Measles, mumps, and rubella
R naught or R0: Basic reproduction number
R: Effective reproductive number
SARS: Severe acute respiratory syndrome
PHEIC: Public health emergency of international concern
WHO: World Health Organization
FAO: Food and Agricultural Organization
WOAH: World Organization for Ani

Chapter 1

Introduction to Epidemiology

Diseases (or other health events) do not occur at random. They occur in certain people for particular reasons, and they do not occur in other people. Diseases (or other health events) have causal and preventative factors that can be identified.

The word "epidemiology" comes from the Greek language, in which epi means "upon," demos denotes "the population" and logy means "the study of". Therefore, classical epidemiology is population-oriented and explores the community origins of health problems.

Epidemiology was originally defined as the scientific study of epidemics, later, the definitions broadened, reflecting the growing concerns of epidemiologists with all types of diseases and the factors that can influence human health. The following is a recent definition agreed upon by an international panel: Epidemiology is "the study of the distribution and determinants of health-related states or events in specified human populations and its application to the control of health problems" [1].

Epidemiology is the basic science of public health. It approaches problems systematically and quantitatively, uses rigorous techniques and methods to produce valid (accurate) and reliable (precise) findings. In epidemiology, we compare the observed number of cases of a disease during a certain period with the expected number based on historical patterns.

The Five Ws of Journalism for epidemiology include: What, Who, Where, When, and Why. Because epidemiology is a quantitative science, the What includes not only what is the problem, i.e., what is the disease, but also what is the size of that problem, i.e., how much disease or how many cases. We must answer all five of the Ws to provide a complete description of a health problem in the community.

What and How Many = Diagnosis and quantity. Who = Persons affected by the disease. Where = Place of occurrence of the disease. When = Distribution of disease by time. Why/How = Cause, risk factors, and modes of transmission. What, How Many, Who, Where, and When provide answers relating to the distribution of the disease or the health-related event. This is descriptive epidemiology. Why and How provide answers relating to the determinants of the disease or the health-related event. This is analytical epidemiology.

In epidemiology, the key to identifying causes and risk factors is to use a comparison group. For example, Why did 30 of 100 people in one building become ill? What made these 30 individuals sick? An initial step in answering this question is comparing the 30 sick individuals to the 70 healthy individuals. Possible reasons for why the 30 became sick may be poor air ventilation – they all work in the same part of the building. Water may be another reason – they all drink from the same water fountain. Food may be a reason – they all ate the same thing in the cafeteria. By interviewing everyone and comparing the sick versus the non-sick people, we can get a very good clue about the

cause of the disease. If both cases and non-cases drank water but only cases ate fish, then fish is the leading suspect. If both cases and non-cases ate fish but only cases drank the water, then water becomes the leading suspect. You need a comparison group to help you sort out which exposures might have been associated with illness.

Historically, the original focus of epidemiology was on infectious diseases and specifically on epidemics of infectious diseases. The current domain has become much broader to cover acute and chronic diseases (communicable and non-communicable diseases, injuries, disabilities, behaviors, social, genetics, and any other health related conditions).

The unit of observation in epidemiology is groups rather than individuals. That is, it focuses on populations rather than individuals. An epidemiologist's "patient" is the population or the community [2, 3].

Clinician versus Epidemiologist for a Patient with Diarrhea

Patient with Diarrhea	**Clinician**	**Epidemiologist**
Main Goal	Diagnosis and treatment of the case	Determine the cause and prevent other cases.
Questions Asked	What is wrong with the patientpatient? What treatment is appropriate?	What are the exposures and sources? Who else was exposed? Is there a potential for spread?
Action	Treatment	Control and prevention measures

Although ology means "the study of," the practice of epidemiology, particularly field epidemiology, is not just "the study of." Like medicine, epidemiology is both a science and a practice. Epidemiologists use knowledge of the distribution and determinants of health-related states and events in populations to take action (to control and ideally to prevent, health problems) [4].

Let's get back to the definition of epidemiology, which is the study of the distribution and determinants of diseases, health related states, or events in specified human populations and its application to the control of health problems.

Distribution deals with time, place, person = Descriptive epidemiology
Determinants deal with cause, risk factors, or etiology = Analytical epidemiology
Population ► Public health
Application ► Information for action

Uses of Epidemiology:

Epidemiology and the information generated by epidemiologic methods have been used in many ways. The common uses are to [5]:

Determine the magnitude and trends

- Identify the etiology or cause of disease
- Determine the mode of transmission
- Identify risk factors and susceptibility

- Determine the role of the environment
- Evaluate the impact of the control measures

Epidemiologic Functions:

The following are the main tasks of epidemiology in public health practice [6]:

- Public health surveillance
- Outbreak investigations
- Analytic studies
- Data analysis
- Evaluation of disease control programs
- Communication
- Management and teamwork
- Policy development

Field Epidemiology:

Field epidemiology is the use of epidemiology as a tool to design, evaluate, or improve interventions to protect the health of a population or it is the practice of epidemiology in the field — in the community — commonly in a public health service [1].

- Field epidemiology is usually the domain of epidemiologists working for ministries of health (or other governmental agencies) or allied organizations (e.g., NGOs).
- The focus is on issues of public health importance and/or public health concern. That is, the issue requires an urgent response or is a priority.
- The goal is to provide information for decision-making, i.e., field epidemiology is not an academic exercise.
- Tasks include monitoring the health of the population, using surveillance or other tools to identify health problems or potential problems, conducting field assessments and investigations, and recommending or implementing interventions to address those problems.
- In general, field epidemiology requires field work. For field investigations, data collection occurs in the field, even if data cleaning and final analysis occurs in an office.
- Sometimes but not always, a prompt answer is needed.
- Field epidemiology is often a team activity, e.g., multiple epidemiologists, or collaboration with laboratorians, sanitarians, etc.
- Field epidemiology activities often have numerous stakeholders (multiple government agencies at different levels of government and/or with different responsibilities (e.g., human vs. agricultural, animal, business, or tourism interests), affected populations, health workers, etc.) with whom epidemiologists must engage, particularly to implement interventions that are both effective and acceptable to the diverse groups with different interests.

Historic Evolution of Epidemiology

During the mid-1800s, John Snow was an anesthesiologist who is now regarded as the “father of field epidemiology.”

Twenty years before the development of the microscope, Snow conducted several investigations of cholera epidemics in London both to discover the cause of disease and to prevent its recurrence. These investigations illustrate the classic sequence from descriptive epidemiology to hypothesis generation to hypothesis testing (analytic epidemiology) to application and action.

Snow conducted one of his now famous studies in 1854 when a cholera epidemic erupted in the Golden Square of London [7]. He began his investigation by determining where persons with cholera in this area had lived and worked and he marked each residence on a map of the area. Because Snow believed that water was a source of infection for cholera, he marked the location of water pumps on his map. He looked for a relationship between the distribution of households with cases of cholera and the location of pumps. Clearly Pump A, on Broad Street, was the pump most central to the distribution of cases. Snow gathered information on where persons with cholera had obtained their water. Consumption of water from the Broad Street pump was the one common factor among cholera patients. After Snow presented his findings to municipal officials, the pump handle was removed, and the outbreak ended.

Deaths from Cholera per 10,000 Houses, by Source of Water Supply, 9 July - 26 August 1854

	Number of Houses	Number of Deaths from Cholera	Deaths/10000 Houses
Southwark& Vauxhall	40,046	1,263	315
Lambeth	26,107	98	37
Rest of London	256,423	1,422	59

Source: Snow J. Snow on cholera. London: Humphrey Milford: Oxford University Press; 1936.

The cholera death rate was more than five times higher in districts served only by the Southwark and Vauxhall Company (intake downstream from London) than in those served only by the Lambeth Company (intake upstream from London). Interestingly, the mortality rate in districts supplied by both companies fell between the rates for districts served exclusively by either company. Additional studies in the districts served by both companies confirmed a much higher cholera mortality rate in households served by Southwark and Vauxhall than in households served by the Lambeth Company. Thus, with no knowledge of the existence of microorganisms, Snow demonstrated through epidemiologic studies that water could serve as a vehicle for transmitting cholera and that epidemiologic information could be used to direct prompt and appropriate public health action.

In the mid- and late-1800s, epidemiological methods began to be applied in the investigation of disease occurrence. At that time, most investigators focused on acute infectious diseases. In the 1930s and 1940s, epidemiologists extended their methods to noninfectious diseases.

Epidemiology has been applied to the entire range of health-related outcomes, behaviors, and even knowledge and attitudes. The studies by Doll and Hill linking lung cancer to smoking [8], and the study of cardiovascular disease among residents of Framingham, Massachusetts [9] are two examples of how pioneering researchers have applied epidemiologic methods to chronic disease since World War II.

Measures of Disease Occurrence and Frequency

1. Counts
2. Ratio
3. Proportion
4. Rates

5. Prevalence
6. Incidence/risk
7. Other measures

RATIO: A fraction in which the numerator is not part of the denominator. e.g., fetal death ratio: fetal deaths/live births. Here, fetal deaths are not included among live births, by definition.

PROPORTION: A fraction in which the numerator is part of the denominator. e.g., fetal death rate: fetal deaths/all births. Here, all births include both live births and fetal deaths.

- Synonyms for proportions are a risk and, (if expressed per 100) a percentage.
- Most fractions in epidemiology are proportions.

RATE: Ideally, a proportion in which change over time is considered, but in practice, often used interchangeably with proportion, without reference to time, (as mentioned for fetal death rate).

Frequently Used Measures of Morbidity

Prevalence rate is of two types:

1. Point prevalence rate: the number of current cases (new and preexisting) of a disease of interest at a given point in time in a specified population at the same specified point in time.
2. Period prevalence rate: the number of current cases (new and preexisting) of a disease over a specified period of time in a specified mid-interval population or the average population. For example: Annual prevalence rate.

When the type of prevalence rate is not specified it is usually point prevalence.

The incidence rate is divided into two types:

1. Cumulative incidence rate
2. Incidence density

Cumulative incidence rate: Number of new cases of a disease occurring over a specified period of time in a population at risk at the beginning of the interval.

Example of cumulative incidence rate: If we count all new cases of influenza occurring in a university undergraduate from September 1, 1997–August 31, 1998, and we take as the denominator all undergraduates enrolled in September 1, 1997, we would be describing the cumulative incidence rate of influenza.

Attack rate is another concept of cumulative incidence usually used in outbreaks.

Secondary attack rate: Number of new cases of a disease among contacts divided by the number of contacts.

Incidence density: Number of new cases of disease occurring over a specified period of time in a population at risk throughout the interval. The numerator does not differ between the two types of incidences. However, the denominator can differ in incidence density from cumulative incidence because in the previous example it takes account of.

- Students who left university during the year

- Students who died
- Students who had influenza once and will not have it again during the same season
- Students who entered the university later in the year

Incidence density requires us to add up the period of time each individual was present in the population and was at risk of becoming a new case of disease. Incidence density characteristically used as the denominator person-years at risk. (Time period can be years, months, days, or even hours, depending on the disease process being studied.)

Uses of incidence and prevalence:

1. Incidence is generally used for acutely acquired diseases, while prevalence is used for more permanent states, conditions or attributes of ill-health.
2. Incidence is more important when thinking of the etiology of the disorder, while prevalence is important when thinking of societal burden of the disorder including the costs and resources consumed as a result of the disorder.
3. Incidence always requires a duration, while prevalence may or may not.
4. In incidence, the unit of analysis is the event, while it is a person in prevalence. Thus, incidence may exceed 100% (e.g., annual incidence of colds) unless a convention is adopted to count only the first episodes of an illness that can occur more than once.
5. Prevalence can never exceed 100%.
6. Incidence generally requires an initial disease-free interval before counting starts, because incidence is measured only in those at-risk of the disease.

Relationship between incidence and prevalence: In a steady state (i.e., if incidence is not changing and the population is stable), the prevalence rate is approximately = incidence rate times the duration of disease ($P = I \times D$).

Example

Tuberculosis in New York City: Tuberculosis is a reportable condition and all diagnosed cases must be reported to the Department of Health.

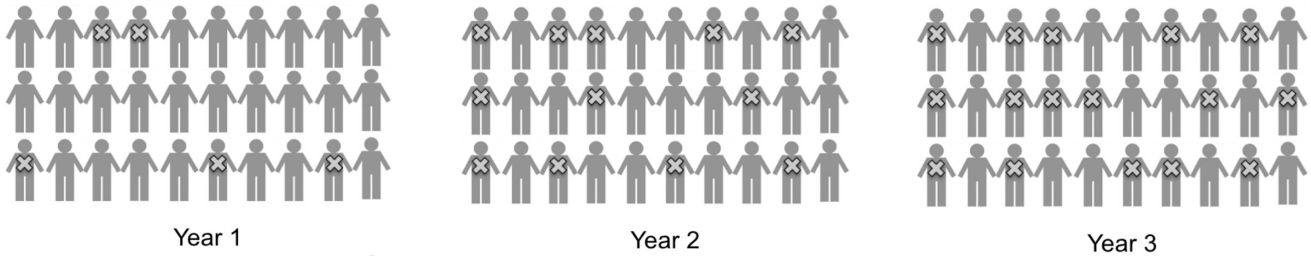
In 2011, there were 689 new cases of tuberculosis in New York City [10]. This count provides an absolute number of the burden of disease. However, counts have limited utility for two reasons:

1. The burden of disease in the population is very different if the population size is different (e.g. 100,000 versus 1,000,000).
2. 2. Some people are not at risk for developing a new onset of tuberculosis in 2011 (due to pre-existing infection), thus we need to know not only the size of the total population, but the size of the total population at risk.

Incidence and prevalence are two measures that overcome many of the limitations of a simple count of cases. The prevalence tells us about the proportion of cases among the total population at any given time and the incidence tells us the probability of a new onset of disease among those at risk for developing the illness.

Exercise

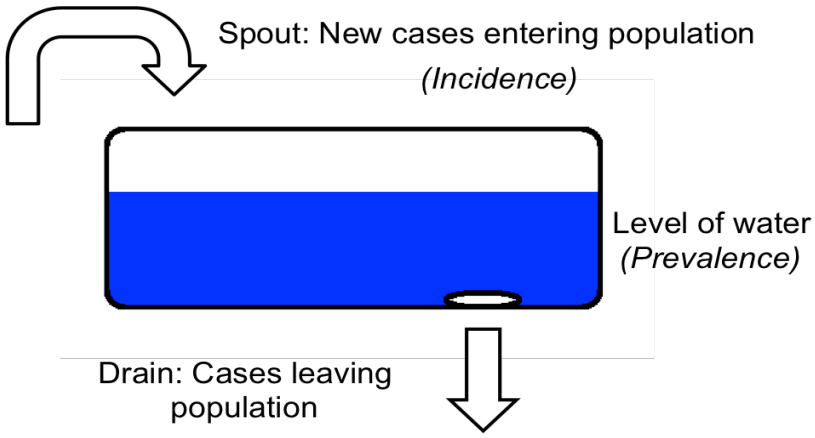
Disease occurrence in a sample of the population below over time: In Year 1, 5 individuals developed



the outcome. In Year 2, an additional 7 people developed the outcome. Year 3, an additional 4 people developed the outcome.

What is the prevalence of the disease in Year 2?
What is the numerator?
5 cases in Year 1 + 7 cases in Year 2 = 12
What is the denominator?
Total sample size = 30
Prevalence = 12/30 = 0.4
The prevalence of the disease in Year 2 is 40%.
What is the prevalence of the disease in Year 3?
What is the numerator?
5 cases in Year 1 + 7 cases in Year 2 + 4 cases in Year 3 = 16
What is the denominator?
Total sample size = 30
Prevalence = 16/30 = 0.533
The prevalence of the disease in Year 3 is 53.3%.
So, for prevalence, we need a numerator (number of existing cases), a denominator (total sample size), and a time period of interest. The time period should be specified as much as possible.
Incidence is perhaps the most widely used tool in epidemiology, it goes by many names. The most common alternative name is “risk,” and less commonly, “incidence proportion.”
Numerator = number of new cases
Denominator = population at risk of becoming new cases over a specific time period
What is the incidence of disease in Year 2?
What is the numerator?
7 new cases in Year 2
What is the denominator?
25 people at risk (5 people already developed the disease in Year 1 and are thus not at risk.)
Incidence = 7/25 = 0.28.
The incidence (risk) of disease in Year 2 is 28%.
What is the incidence of disease in Years 2 and 3?
What is the numerator?
7 new cases in Year 2 + 4 new cases in Year 3 = 11
What is the denominator?
25 people at risk (5 people already developed the disease in Year 1 and are thus not at risk.)

Incidence = 11/25 = 0.44
The incidence (risk) of disease in Years 2 and 3 is 44%.

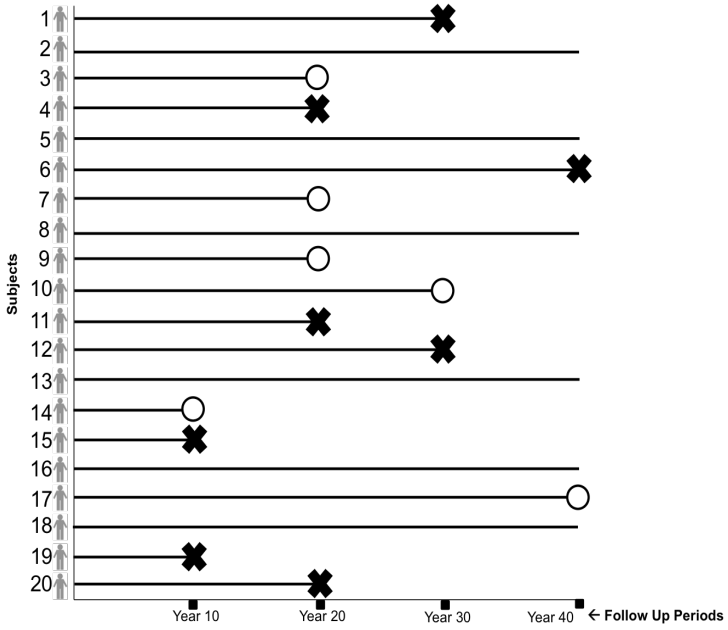


Understanding Incidence and Prevalence – The Bathtub Example

So, for incidence, we need a numerator (number of new cases), a denominator (total sample size at risk), and a period of interest. The time period should again be specified as much as possible.
Examples of the relation between incidence and prevalence:

- High incidence, steady prevalence: Highly contagious infectious disease with very short duration or a high case-fatality
- Low incidence, high prevalence: Diseases with long duration such as arthritis, diabetes, Crohn's disease, and other chronic illnesses

Prevalence is affected by incidence and duration. If a disease has a short duration, prevalence = incidence, assumes that incidence is constant over time. If a disease has a long duration, in general, prevalence is bigger than incidence.
Incidence rates are commonly used in prospective studies in which some people are lost over time. To estimate the rate over the study's time frame, we need to know how much total time each person



Understanding Person Years: Person-time and Disease Status Among 20 Subjects Followed for Forty Years

contributed to the study follow-up before they either developed the outcome or dropped out, we term the total time that each person contributed as person-time.
Person 2 stayed in the study for all 40 years and did not develop the outcome.
Person 10 dropped out of the study in Year 30.

Subject	Years Contributed	Developed Disease? (0 no/1 yes)
1	25	1
2	40	0
3	10	0
4	15	1
5	40	0
6	35	1
7	10	0
8	40	0
9	10	0
10	20	0
11	15	1
12	25	1
13	40	0
14	0	0
15	5	1
16	40	0
17	30	0
18	40	0
19	5	1
20	15	1

Person 19 developed the outcome at Year 10.
Person-time and disease status among 20 subjects followed for forty years.

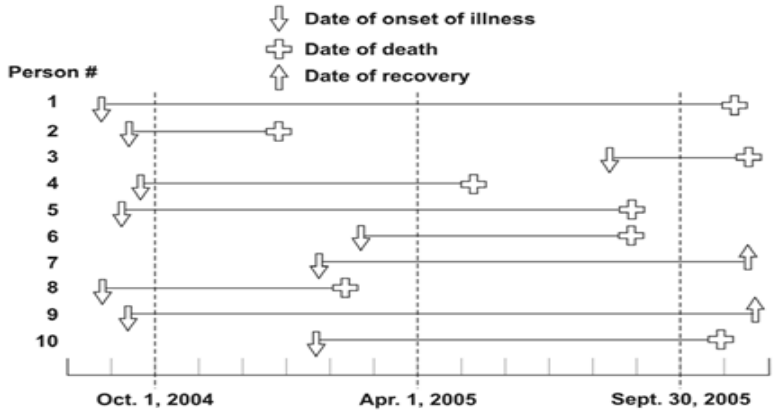
Calculating the Incidence Rate (Incidence Density)

The numerator is the number of cases= 8, the denominator is the total person-time= 440
In our example: 8/440 = 0.18, or a rate of 18 cases per 1,000 person-years of observation.

The incidence rate can be interpreted as the number of expected cases in every set of 1,000 person years. That is, if we were to observe 1,000 people for 1 year, we would expect 18 cases. If we were to observe 500 people for 2 years, we would still expect 18 cases.

The assumption underlying this is that the incidence rate is constant over time, so for every year in which 1,000 person years are observed an additional 18 cases will be expected. Given this assumption, the incidence rate tells us the average number of cases per a specified set of person time.

Exercise on Incidence and Prevalence



Cases of Illness from October 1, 2004–September 30, 2005

Calculate the incidence rate from October 1, 2004, to September 30, 2005, using the midpoint population (population alive on April 1, 2005) as the denominator. Express the rate per 100 population.
Incidence rate numerator = number of new cases between October 1 and September 30= 4 (the other 6 all had onsets before October 1, and are not included)
Incidence rate denominator = April 1 population = 18 (persons 2 and 8 died before April 1). Incidence rate = (4 / 18) × 100 = 22 new cases per 100 population.

Calculate the point prevalence on April 1, 2005. Point prevalence is the number of persons ill on the date divided by the population on that date. On April 1, seven persons (persons 1, 4, 5, 7, 9, and 10) were ill.

Point prevalence = (7 / 18) × 100 = 38.89%.

Calculate the period prevalence from October 1, 2004, to September 30, 2005. The numerator of period prevalence includes anyone who was ill at any time during the period. In Figure 3.1, the first 10 persons were all ill at some time during the period.

Period prevalence = (10 / 20) ×100 = 50.0%.

Measures of Mortality

Mortality analysis begins with good quality data on deaths and population. This data is conventionally obtained from vital registration systems and population censuses respectively. The crude death rate and the specific death rates (age, sex, and cause) are simple measures of mortality. The other measures are based on the life tables.

1. **Crude death rate (CDR)**
   The crude death rate is calculated by dividing the number of registered deaths for all causes in a year by the mid-year population for the same year. The rate is expressed as per 1,000 population. The crude death rate = Total number of deaths during a certain year/Total mid-year population multiplied by 1000.

   This rate has a simple interpretation, for it gives the number of deaths that occur on average per 1,000 people in the community. Further, it is relatively easy to compute, requiring only the total population size and the total number of deaths. It is a true probability rate. It represents an estimate of the chance of dying for a person belonging to the given population. However, it has some serious drawbacks. In using the crude death rate, we ignore the fact that the chance of dying is not the same for the young and the old or for males and females, and the fact that it may also vary with respect to race, occupation, or locality of dwelling.

2. **Specific death rates**
   - Cause-specific death rate = Total number of deaths due to a particular cause during a certain year/Total mid-year population multiplied by 1000.
   - Cause and sex-specific death rate = Total number of deaths due to a particular cause among males (or females) during a certain year/Total mid-year population of males (or females) multiplied by 1000.

- Age specific death rate = Total number of deaths for a certain age group during a certain year/ Total mid-year population of that age group multiplied by 1000.
- The age-cause-specific death rates = Total number of deaths due to a particular cause among a certain age group during a certain year/Total mid-year population of the same age group multiplied by 1000.
- Infant mortality rate = Number of deaths among children under 1 year of age during a certain year/number of live births for that year multiplied by 1000.
- Child mortality rate = Total number of deaths of children aged 1 to less than 5 years in a given year and geographical region/population of the same age in that year and geographical regions multiplied by 1000.
- Under 5-year mortality = Number of deaths of children aged less than 5 years in a given year and region/Total population aged less than 5 years in the same given year and region multiplied by 1000.
- Neonatal mortality rate = Deaths of infants aged less than one month (or less than 4 weeks) in a given year and region/total live birth in the same given year and region multiplied by 1000.
- Post neonatal mortality rate = Number of deaths of newborns between 4 weeks and less than one year of age in a given year and region/ total live birth in the same given year and region multiplied by 1000.
- Fetal death: It is known as the death prior to the complete expulsion or extraction from its mother of a product of conception at any point of time of pregnancy.
- The stillbirth rate (SBR) is defined as the number of babies born with no signs of life at 28 weeks or more of gestation, per 1,000 total births.

3. **Proportionate mortality ratio**
   Deaths due to a specific cause/total deaths multiplied by 100. The % sign comes from the fact that this figure is usually expressed as a percentage.
4. **Maternal mortality ratio**
   A maternal death is defined by the World Health Organization as follows: "The death of a woman while pregnant or during delivery or within 42 days of termination of pregnancy, irrespective of the duration and the site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes." It can be calculated by dividing the number of maternal deaths during a certain year in a certain geographic area by total number of live births in the same year and geographic area multiplied by 100,000.
5. **Case fatality rate**
   Also called case fatality risk or case fatality ratio, it is the proportion of people who die from a specified disease among all individuals diagnosed with the disease over a certain period of time. Case fatality rate is typically used as a measure of disease severity and is often used for prognosis (predicting disease course or outcome). Comparatively high rates are indicative of relatively poor outcomes. It can also be used to evaluate the effect of new treatments, with measures decreasing as treatments improve. Case fatality rates are not constant; they can vary among populations and over time, depending on the interplay between the causative agent of disease, the host, and the environment as well as the available treatments and quality of patient care.

Standardization of Rates

Although age-specific death rates provide the most appropriate basis for comparing the mortality experience of different populations, it is useful to have a single overall measure of mortality which,

unlike the crude death rate, adjusts for the effect of difference in age distribution of the populations to be compared [11].

Direct standardization: In this method, the distributions of the compositional variables (age, sex, marital status etc.) of the populations that are being compared are made identical and the standardized rates are calculated such that the difference between them is only due to the variation in the age-specific rates of their population.

When we want to compare the mortality experience between two regions in a country or between two countries by using a single overall measure of mortality (summary measure), we choose a standard population with a known age distribution, then we apply the age-specific mortality rates of the two regions to be compared to the corresponding age groups of the standard population, yielding the number of deaths which would occur in that standard population if it was subject to the mortality rates prevailing in each region. So, the effect of difference in age distribution of the two regions has been eliminated by the use of standard population.

Data needed: If we want to standardize for age as an example, we need the age distribution of the standard population, and the age-specific death rates in the populations to be compared.

We multiply the age-specific mortality rates of the other population under study to the number of persons in each age group of the standard population. This way, we will get the expected deaths for each age group of each population.

Hypothetical example

Commenting on crude death rates, Population B seems to have a higher crude death rate than Population A. The crude death rate is not a valid measure for comparing the risk of death between countries or between regions within the same country when their age composition is different.

As an example, let us say we chose a reference (standard) population whose age distribution structure is known, and we have two populations, A and B which have known age-specific death rates and we need to come up with a valid summary measure of mortality to compare the risk of death between these two populations by adjusting for the difference in the composition of the populations to be compared. The calculation will therefore be as follows in the table below:

Age Groups of Standard Population	**Number in Each Age Group of Standard Population**	**Population A's Age-Specific Death Rate/1000**	**Population A's Expected Deaths**	**Population B's Age-Specific Death Rate/1000**	**Population B's Expected Deaths**
0-24	11000	1.94	21.34	2.31	25.41
25-49	17000	5.45	92.65	7.14	121.38
50-74	20000	41.11	822.2	36.36	727.2
75 and above	3000	83.33	249.99	95.0	285.0
Total	51000		1186.18		1158.99

Step 1: We multiply Population A's age-specific death rates by the corresponding number of populations in each group of the standard population (multiply third column of the table by the second column) and we come up with expected deaths in the fourth column.

Step 2: We multiply Population B's age-specific death rates by the corresponding number of populations in each group of the standard population (multiply fifth column of the table by the second column) and we come up with expected deaths in the sixth column.

Step 3: We divide the total expected deaths for Population A in column 4 over the total standard population in column 2: 1186.18/51000 multiplied by 1000 = 23.3/1000 population, is the standardized (adjusted) death rate of Population A.

The standardized (adjusted) death rate of Population B is calculated by dividing the expected deaths in column 6 over the total standard population in column 2 =1158.99/51000 multiplied by 1000 = 22.7/1000 population.

Although the crude death rate of Population B was higher than the crude death rate of Population A, after the adjustment, the adjusted rate for Population B became smaller than that of Population A. The reason behind this was the difference in population age composition between Populations A and B.

An alternative mortality measure is the comparative mortality figure (CMF). It is calculated as follows:

- Suppose D is the number of observed (actual) deaths in the standard population.
- D1 is the expected deaths of Population A, and D2 is the expected deaths of Population B.
- CMF of Population A = D1/D multiplied by 100
- CMF of Population B = D2/D multiplied by 100

If CMF for Population A is less than 100, it indicates that mortality conditions are better in Population A compared with the standard population, while a CMF greater than 100 indicates that they are worse.

Indirect method of standardization

The indirect method of standardization is commonly used when age-specific rates of a standard population are applied to the corresponding age groups of the populations of interest to yield the number of deaths expected if each population had experienced the mortality conditions of the standard population.

- If D1e = expected deaths in Population A, and D2e = expected deaths in Population B. And if D1o = observed (actual) deaths in Population A, and D2o = observed (actual) deaths in Population B, then:
- The standardized mortality ratio (SMR) for Population A = Observed deaths in A/Expected deaths in A multiplied by 100. Or D1o/D1e*100.
- The same calculation goes for the SMR of Population B.

Example

If we want to compare the mortality experience between two populations, A & B by using the indirect method, we may choose a third population as a standard population with known age-specific death rates, or if one of the populations A & B has a known age-specific death rates, we can use that population as a standard to the second population. The table below explains this, where Population A was used as a standard population for Population B. The calculation will be as follows:
Step 1: We multiply the second column by third column = the observed deaths for Population A (in column 4). This is equivalent to the expected deaths of the same population.

Step 2: We multiply the fifth column by second column = the expected deaths for Population B (in column 6).

Step 3: Observed deaths for Population B is known to be = 15300, while expected deaths were calculated to be = 9540.

Step 4: Standard mortality ratio (SMR) for Population B = Observed deaths/Expected deaths = 15300/9540 multiplied by 100 = 160.4. The SMR for Population A (standard population) = 147000/147000 multiplied by 100 = 100. This means that the number of observed deaths in Population B is 60% higher than the number we would expect if Country B had the same mortality experience as Country A. So, the risk of mortality is higher in Population B than Population A.

Age Group	**Country A's (Standard Population) Age Specific Death Rate**	**Population of Country A**	**Country A's Age Specific Death Rate * Population of Country A**	**Population of Country B**	**Country A's Age Specific Death Rate * Population of Country B**
0-29	0.0012	6000,000	0.0012*6000000 = 7200	1,500,000	0.0012*1,500,000 = 1800
30-59	0.0036	5,500,000	0.0036*5500000 = 19800	550,000	0.0036*550,000 = 1980
60 and Above	0.048	2,500,000	0.048*2500000 = 120000	120,000	0.048*120,000 = 5760
Total Observed Deaths			147000		15300
Total Expected Deaths			147000		9540
Standardized Mortality Ratio (SMR) = Observed/ Expected*100			147000/147000*100 = 100		15300/9540*100 = 160

Life Expectancy

Definition: "The average number of years an individual is expected to live." This term is generally used to refer to the expectation of life at birth (the average number of years of life that a newborn infant is expected to live). However, expectation of life can be calculated for any age. Expectations at age 20, for example, indicate the average number of remaining years of life for those who have attained the age of 20 years.

Life expectancy is a very important summary measure for comparing death rates within and among countries and over time. One practical application of this measure is its use for life insurance purposes.

Life expectancy is calculated from tables known as demographic life tables. These tables are constructed from current age-specific death rates as if these rates remain unchanged throughout the lifetime of the cohort. That is, life expectancy for infants born in 1990 is calculated from 1990 age-specific death rates even though the 1990 birth cohort will, as it ages, be subjected to the age-specific death rates prevailing in 2000, 2010, and so on. This limits the accuracy with which demographic life tables can predict life expectancy over the lifetime of the cohort.

Construction of life table

Suppose a cohort of 100,000 male births in a particular year. A number of the cohort will die during the first year of life (infant mortality). The number who die could be estimated by means of the current infant mortality rate (IMR). The average IMR for that year was 20.78 per 1000, so that out of 100,000 male births, the expected number of deaths can be estimated as 2078. Thus, of the original cohort, 97922 may be expected to survive to the age of 1 year. The figures are shown in the columns headed **Ix** and **dx** in the table below.

Now, how many of the cohort will survive to the age of 2 years? This can be estimated by using the actual (1970–1972) specific mortality rates for the country for the age group 1 and under 2 years of age. It is estimated that 136 of the cohort will die between the ages of 1 and 2, leaving 97786 to survive until the age of 2. Another 87 will die between the ages of 2 and 3 leaving 97699 to survive until the age of 3, and 59 will die between the age of 3 and 4 leaving 97640 to survive until the age of 4.

For each age group, the cohort is subjected to a specific mortality rate for that age group. Eventually, the cohort will die off. The number who die at each age is determined by the specific mortality rates, and these are usually based on the average mortality rates for the most recent period for which accurate statistics are available.

The interpretation of the figures in column **Lx:**

Suppose the whole cohort had survived to age 1 if the IMR is zero (which is unrealistic). In this case, the total number of years lived by the cohort between birth athe nd the age of 1year would be 100000 (each member would have lived for one year). However, in the real situation since IMR will never reach zero, only 97922 of the cohort lived for one year and 2078 lived for only a part of the first year. It is estimated based on a precise method of calculation (not explained here). Those who die after the age of one year live 6 months on average for each member. This does not apply to those who die during the first year since most of them usually die during the first month of life.

Accordingly, those who died during the first year (2078) account only for 276 years of life (each member lives on average 13.3% of a year), when we add this to the figure 97922, it will give 98198 years lived by the cohort at age one year as shown in column **Lx**.

97922+276(13.3% of 2078) = 98198 = Total number of years lived by the cohort between the ages of 0-1

97786+(136/2) = 97854 = Total number of years lived by the cohort between the ages of 1 and 2.
97699+(87/2) = 97743 = Total number of years lived by the cohort between the ages of 2 and 3.
97640+(59/2) = 97670 = Total number of years lived by the cohort between the ages of 3 and 4.

Column **Tx**:
The figure 6876850 represents the life span of all the members of the cohort at all ages (the sum of all the figures in column **Lx** when the complete life table for all ages is constructed). Remember that the data in the table represents only 5 age groups.

Now, to fill in the **Tx** column, we follow the below calculation till we reach the end of the table for all ages (remember again this table is just a sample table for 5 age groups only used for the sake of this exercise). We subtract the figures in column Lx from the figures in column **Tx** as follows:

6876850 minus 6876850 = 6778652
6778652 minus 97854 = 6680798
6680798 minus 97743 = 6583055
6583055 minus 97670 = 6485385

The figures in column **Ex0** represent the life expectancy rate for each age. These figures came from dividing figures in column **Tx** over figures in column **Ix**:
6876850 ÷100000 = 68.76 = Life expectancy at birth.
6778652÷97922 = 69.22 = Life expectancy at the age of 1 year.
6680798÷97786 = 68.32 = Life expectancy at the age of 2 years.
And so on for the rest of the table.

Note: The reader may be surprised to note that the mean expectation of life at the age of 1 year exceeds the mean at the age of birth. This is due to the effect of infant mortality on the life expectancy of the cohort at birth. However, the expectation of life from age 1 onwards declines as expected.

A Sample of the Irish Life Table for Males During 1970–1972

Age X	Survivor Cohort at Age x (Ix)	Deaths at Age x (dx)	Number of Years Lived by Cohort Between Successive Ages (Lx)	Tx	Ex0
0	100000	2078	6876850	6876850	
1	97922	136	97854	6778652	
2	97786	87	97743	6680798	
3	97699	59	97670	6583055	
4	97640	57	97611	6485386	

Life tables can be constructed for different populations and comparisons made on the basis of life expectancy. Separate tables can be constructed for males and females, for different areas of the country (urban and rural), and different occupations.

The most obvious limitation of life tables is the use of prevailing age-specific mortality rates to calculate the expected mortality experience of the cohort.

Fertility Measures

Types of fertility measures

There are two broad types of fertility measures and the analysis of fertility is basically carried out in two ways; one is from a period perspective and the other from a cohort perspective[12]. The events that occur in a given period (calendar years) are studied in relation to the durations of exposure of the population during that period. In the cohort, the events and duration of exposure are studied for well-defined cohorts as they move over time. The term "cohort" indicates a group of people who have had a similar experience at the same time. Two types of cohorts are generally used in demography – birth cohorts and marriage cohorts.

- **Period measures:** They are related to a period and based on data on the number of births in that period. These include the Crude Birth Rate (CBR), General Fertility Rate (GFR), and Child Women Ratio (CWR).
- **Cohort measures:** In any sample fertility survey, a question is usually asked about the number of children ever born (CEB) to women up to a time in the reproductive age groups. Using this approach, fertility is estimated indirectly on the basis of age and sex distribution of the population. These include Total Fertility Rate (TFR), Gross Reproduction Rate (GRR), and Net Reproduction Rate (NRR).

The second categorization of measures of fertility is:

- Direct measures of fertility
- Indirect measures of fertility

Direct measures of fertility: In these methods, data on live births are directly used. Some direct measures of fertility are described below.

1. Crude Birth Rate (CBR): This is defined as the ratio of the total number of live births in a year in a specified area divided by the total mid-year population of the same specific area in the same year multiplied by a constant K. CBR = B/P*1000 Where B = the total number of live births in a year, P = the total population in the middle of the year and K is constant, usually 1000.
   Advantages and disadvantages of CBR: It is an important measure of fertility; it directly links fertility to the growth rate of population. Computation of CBR is easy and quick and requires minimal data. CBR also indicates the level of fertility in a population. A major weakness of CBR is that it is not very sensitive to small fertility changes; in fact, it tends to minimize them. CBR is affected by many factors: age, sex, and marital status. It is also influenced by the age structure of the population, and by the level of fertility and age pattern of fertility.
2. General Fertility Rate (GFR): The relative frequency of childbirth varies significantly with the age of parents. The age at which maximum fertility occurs may be different for males and females. Furthermore, fertility is highest among couples who have established some type of cohabitation (legal marriage or common law marriage) than among persons not in such a union (single). Conversely, specific fertility rates are given separately for female parents and male parents.

Usually, children are born to women between the ages of 15 and 45 years, which is known as the reproductive age group. The fertility rate for this group, called the "General Fertility Rate" (GFR), is calculated as the ratio of total number of yearly births to the total number of females (mid-year population) of childbearing ages (15–44 or 15–49 years).

GFR = Number of live births during a year/Mid-year female population aged 15-49 **X** 1000.

The purpose of having a GFR is to restrict the denominator to potential mothers.

3. General Marital Fertility Rate (GMFR): Besides age, marital status is an important factor in fertility. In almost all societies in the world, birth is allowed only in a marital bond. Therefore, it may be more appropriate to consider only currently married women, and not all women, in the reproductive ages. It is calculated from the following expression:
   GMFR = Live births in a year/married women aged 15-49 X 1000
   Although it is a refinement over CBR, GFR also suffers from certain limitations. The measure considers the entire female population in the reproductive ages as a homogeneous group, whereas the fecundity of women is not uniform during the period. Thus, GMFR must also be considered as a crude rate.
4. Age-Specific Fertility Rate: The Age-Specific Fertility Rate (ASFR) addresses the limitations of GMFR. ASFR is calculated in the following manner:
   ASFR = Live births to women in a certain age group during a year/mid-year female population of the same age group X 1000.
   The reproductive age interval 15–49 can be either divided into single or five years or wider intervals and rates can be made specific to each age group. Because of the wide variations in fertility by age, age specific fertility rates have been found to be very useful. Generally, five-year age groups of women are used for calculating the ASFR resulting in seven numbers, one for each age group 15–19, 20–24, ----, 45–49. The general pattern of the ASFR is, the rate increases to a maximum between ages 20–29 and then decreases slowly to reach zero by age 50.
5. Age-Specific Marital Fertility Rate (ASMFR): One must note that the measure ASFR can be used with reference to only currently married women in an age group. Thus, it becomes necessary to introduce an Age-Specific Marital Fertility Rate (ASMFR), which can be expressed as:
   ASMFR = Live births to married women in a certain age group during a year/mid-year married female population of the same age group X 1000.
   Since there is a possibility of greater incidence of unmarried women in the early age groups, and divorced, separated, and widowed women in the older age of the reproductive age span, ASMFR provides a more realistic picture of fertility levels in a population. It is also possible to compute the total marital fertility rate (TMFR), which is equivalent to the TFR for a married woman.
   TMFR = Sum of the Age-Specific Marital Fertility Rates for all age groups (15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49) x 5/1000 , where the number 5 represents the width of class interval.
6. Total Fertility Rate: Usually, the reproductive age span is divided into age groups in five-year intervals. Thus, there would be six or seven groups, depending on the upper limit of the reproductive age span. The use of age-specific fertility rates in comparison between two or more populations is a cumbersome exercise. Thus, we use the Total Fertility Rate (TFR), which is a summary measure of ASFR, to facilitate comparison. TFR is calculated as we calculated the TMFR above by multiplying the sum of ASFR by the width of the age group, and then dividing the product by 1,000. The following is the formula:
   TFR = Sum of the age-specific fertility rates for all age groups (15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49) x 5/1000, where number 5 represents the width of class interval. Thus, TFR refers to the total number of children a woman will produce during her childbearing age span if she is subjected to a fertility schedule as prescribed by the age-specific fertility rates.

7. Gross Reproductive Rate (GRR): Total fertility includes all births, both males and females. The GRR shows how many girls babies (potential future mothers) would be born to 1000 women passing through their childbearing years. It represents the average number of daughters who would replace their mothers, assuming that the age and sex-specific fertility rate for the current period were to continue indefinitely. GRR indicates the number of daughters that every woman is likely to bear during her entire childbearing age span, if she is subjected to a fertility schedule as prescribed by given sex and age specific fertility rates. Also considered a replacement index, this measure is generally used while comparing current fertility in different populations. Calculation of GRR requires data on the number of live births by sex along with distribution of women in different age groups in the childbearing age span. If the data is available, GRR can also be worked out by simply multiplying the TFR by the femininity ratio (the ratio between the number of female babies born and the total live births in a population). For instance, if 105 male babies are born for every 100 female babies in a certain country, the femininity ratio is 0.4878 (i.e., 100/205).

   GRR = TFR X Femininity Ratio

   As with TFR, GRR also assumes that women in the reproductive age group will survive till the end of their child-bearing period.

   GRR thus indicates the number of daughters a woman is expected to produce if there is no attrition in the cohort due to mortality. This is, however, not a realistic assumption.
8. Net Reproduction Rate (NRR): The Net Reproduction Rate (NRR), a refinement over GRR, with a component of mortality built into it, allows for decrease due to deaths among mothers. Thus, NRR is the number of daughters ever born to a woman, if she gives birth according to the given schedule of age-specific fertility rates, and experiences given age-specific mortality rates up to the end of her reproductive span. NRR measures the extent to which a woman will replace herself with female babies under predetermined schedules of fertility and mortality.
9. Sex Age Adjusted Birth Rate (SAABR): Another measure that reduces the effects of age structure to a minimum and hence, facilitating comparison of the fertility levels of two or more populations is Sex Age Adjusted Birth Rate (SAABR). The United Nations defines it as "the number of births per 1,000 of a weighted sums of the number of women in various five-year age groups from 15 to 44."

   The UN recommends a standard set of weights (1, 7, 7, 6, 4, and 1) corresponding to the six five-year age groups in the reproductive age span from 15 to 44 years. These weights are roughly proportional to the typical relative fertility rates of various age groups. These weights were derived based on a study of 52 nations having varying levels of fertility.

   SAABR is calculated from the following formula:

   SAABR = B / [(1xW1) + (7xW2) + (7xW3) + (6xW4) + (4xW5) + (1xW6)].

   Where B is the number of live births in a calendar year.

   W1, W2… W6 are the numbers of women in the six five-year age groups in the reproductive age span. (1, 7, 7, 6, 4, and 1) are the UN's recommended standard set of weights corresponding to the six five-year age groups.

Indirect measures

In addition to the direct measures discussed above, there are some indirect measures of fertility, which are useful particularly when data on live births are not readily available or are not reliable. These measures arrive at estimates of fertility indirectly using data on age-sex structure, marital

status, and cross-classified by age and sex. Child Women Ratio and Female Mean Age at Marriage are the most commonly used indirect measures.

1. Child–Women Ratio (CWR): This is a ratio between women and children in a population. It is expressed in terms of the number of children below five years of age per thousand females of the reproductive age group (15–49 years). The formula is as follows:
   CWR = (P0–4 / W 15–44 or 49) K. Where, P0–4 is the number of children in the age groups 0–4 years (under 5 years), W15–44 or 49 is the number of women of the childbearing age 15–44 or 15–49. K is usually taken as 100.

   As P0–4 is the survivors of the children born over the preceding five years, and not the total births, CWR is affected by infant and child mortality. Hence, it is not a very accurate measure of fertility. Nevertheless, it may be used as a relative measure to study the fertility performance of different sections of the same population.

2. Mean Age at Marriage: Age at marriage is said to have significant bearing on the fertility performance of women in a population. If the age at marriage is low, women start bearing children at an early age. However, when the age at marriage is raised, the reproductive span is reduced and the overall fertility level is low. Mean age at marriage, therefore, is taken as a proximate indicator of fertility levels.

Chapter 2

Epidemiologic Transition

Concept and Evolution

Epidemiologic transition is the process by which the pattern of mortality and disease in a population is transformed from one of high mortality among infants and children and episodic famine and epidemics affecting all age groups to one of degenerative and human-made diseases (such as those attributed to smoking) affecting principally the elderly.

It is generally believed that epidemiologic transitions prior to the 20th century (i.e., those that took place in today's industrialized countries) were closely associated with rising standards of living, nutrition, and sanitation. In contrast, those occurring in developing countries beginning in the 20th century have been more or less independent of such internal socioeconomic development and more closely tied to organized healthcare and disease control programs developed and financed internationally. There is no doubt that 20th- and 21st-century declines in mortality in developing countries have been far more rapid than those that occurred in the 19th century in industrialized countries.

Epidemiological transition is a theory which "describes changing population patterns in terms of fertility, life expectancy, mortality, and leading causes of death" [13, 14]. For example, a phase of development marked by a sudden increase in population growth rates brought by improved food security and innovations in public health and medicine can be followed by a re-leveling of population growth due to subsequent declines in fertility rates. Such a transition can account for the replacement of infectious diseases by chronic diseases over time due to increased life span as a result of improved healthcare and disease prevention. This theory was originally suggested by Abdel Rahim Omran in 1971. [14]

Phases of the Epidemiological Transition of Mortality

According to Omran, the epidemiological transition of mortality is divided into three phases.

1. The Age of Pestilence and Famine: Mortality is high and fluctuating, precluding sustained population growth, with low and variable life expectancy vacillating between 20 and 40 years. It is characterized by an increase in infectious diseases, malnutrition and famine, and was common during the Neolithic age. Before the first transition, our hominid ancestors were hunter-gatherers and foragers, a lifestyle partly enabled by a small and dispersed population. However, unreliable and seasonal food sources put communities at risk for periods of malnutrition.
2. The Age of Receding Pandemics: Mortality progressively declines, with the rate of decline accelerating as epidemic peaks decrease in frequency. Average life expectancy increases steadily from about 30 to 50 years. Population growth is sustained and begins to become exponential.
3. Age of Degenerative and Man-Made Diseases: Mortality continues to decline and eventually approaches stability at a relatively low level. Mortality is increasingly related to degenerative diseases, cardiovascular disease (CVD), cancer, violence, accidents, and substance abuse, some of these due primarily to human behavior patterns. The average life expectancy at birth rises gradually until it exceeds 50 years. It is during this stage that fertility becomes the crucial factor in population growth.

In 1998 Barrett et al. proposed the following additional two phases:

- The Age of Declining CVD Mortality, Aging and Emerging Diseases: Technological advances in medicine stabilize mortality and the birth rate levels off. Emerging diseases become increasingly lethal due to antibiotic resistance, new pathogens like Ebola or Zika, and mutations that allow old pathogens to overcome human immunity.
- The Age of Aspired Quality of Life with Persistent Inequalities: The birth rate declines as lifespan is extended, leading to an age-balanced population. Socioeconomic, ethnic, and gender inequalities continue to manifest differences in mortality and fertility.

During these two additional phases, cardiovascular diseases diminish as a cause of mortality due to changes in culture, lifestyle and diet, and diseases associated with aging increase in prevalence. In the final phase, disease is largely controlled by those with access to education and healthcare, but inequalities persist.

The epidemiological transition occurs when a country undergoes the process of transitioning from developing nation to developed nation status. The development of modern healthcare and medicine, such as antibiotics, drastically reduces infant mortality rates and extends average life expectancy which, coupled with subsequent declines in fertility rates, reflects a transition to chronic and degenerative diseases as more important causes of death.

The theory of epidemiological transition uses patterns of health and disease as well as their forms of demographic, economical, and sociological determinants and outcomes.
Omran's first phase occurs when the human population sustains cyclic, low-growth, and mostly linear, up-and-down patterns associated with war, famine, epidemic outbreaks, as well as small golden ages and localized periods of prosperity. In early pre-agricultural history, infant mortality rates were high and average life expectancy low.

The second phase involves improved nutrition as a result of stable food production along with advances in medicine and the development of healthcare systems. Mortality in Western Europe and North America was halved during the 19th century due to closed sewage systems and clean water provided by public utilities, with a particular benefit for children of both sexes and to females in the adolescent and reproductive age periods, probably because the susceptibility of these groups to infectious and deficiency diseases is relatively high. An overall reduction in malnutrition enabled populations to better resist infectious disease. Treatment breakthroughs of importance included the initiation of vaccination during the early nineteenth century, and the discovery of penicillin in the mid-20th century, which led respectively to a widespread and dramatic decline in death rates from previously serious diseases such as smallpox and sepsis. Population growth rates surged in the 1950s, 1960s and 1970s to 1.8% per year and higher, with the world gaining 2 billion people between 1950 and the 1980s. A decline in mortality without a corresponding decline in fertility leads to a population pyramid assuming the shape of a bullet or a barrel, as young and middle-age groups comprise equivalent percentages of the population.

The third phase occurs when human birth rates drastically decline from highly positive replacement rates to stable replacement numbers. Several European nations' replacement rates have even become negative. This transition generally represents the net effect of individual choices on family size and

the ability to implement those choices. Omran provides three possible factors that tend to encourage reduced fertility rates.

1. Bio-physiologic factors, associated with reduced infant mortality and the expectation of longer life.
2. Socioeconomic factors, associated with childhood survival and the economic challenges of large family size.
3. Psychological or emotional factors, where society as a whole changes its rationale and opinion on family size and parental energies are redirected to qualitative aspects of child-raising.

Impact on Fertility

Improvements in women and childhood survival that occur with the shift in health and disease patterns discussed above have distinct and seemingly contradictory effects on fertility. While better health and greater longevity enjoyed by females of reproductive age tend to enhance fertility, the reduced risks to infants and young children that occurs in the later stages of the transition tends to have the opposite effect: prolonged breastfeeding associated with reduced mortality among infants and toddlers, together with parental recognition of improved childhood survival, tend to lengthen birth intervals and depress overall reproductive rates.

Economic Impact

The transition may also be associated with demographic movements to urban areas, and a shift from agriculture and labor-based production output to technological and service-sector-based economies. This shift in demographic and disease profiles is currently underway in most developing nations, however, every country is unique in its transition speed that is based on numerous geographical and sociopolitical factors. Whether the transition is due to socioeconomic improvements (as in developed countries) or by modern public health programs (as has been the case in many developing countries), the lowering of mortality and of infectious disease tends to increase economic productivity through better functioning of adult members of the labor force and through an increase in the proportion of children who survive and mature into productive members of society.

Basic Models of Epidemiological Transition

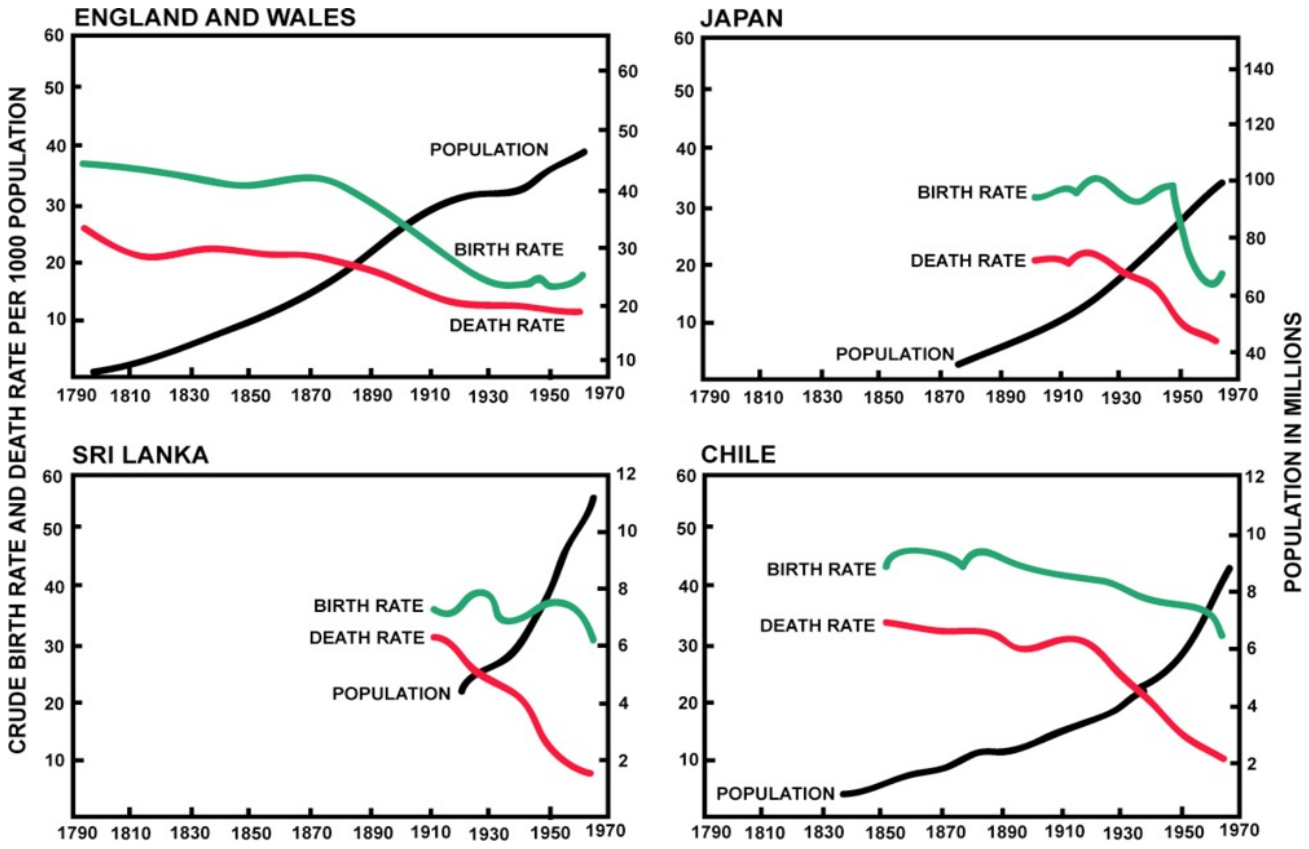
1. Classical/Western model: Countries in Western Europe typically experienced a transition that began in the late eighteenth century and lasted over 150 years to the post-World War II era. The lengthy transition allowed fertility to decline at virtually the same rate that mortality declined.
   The classical model describes the gradual, progressive transition from high mortality (above 30 per 1,000 population) and high fertility (above 40 per 1,000) to low mortality (less than 10 per 1,000) and low fertility (less than 20 per 1,000) that accompanied the process of modernization in most western European countries.
2. Accelerated model: Japan experienced a rapid transition as a result of a few decades of intensive war-driven industrialization followed by postwar occupation. The accelerated transition follows a pattern similar to the Classical/Western Model except that it occurs within a much shorter time span. China might be considered another example of this model. A major distinction of the accelerated model is that the period taken for mortality to reach the 10 per 1,000 level was much shorter than that for the classical model.
   Most of the countries fitting this model had begun a slow process of modernization prior to the drop in mortality in the twentieth century, which was determined by sanitary and medical

advances as well as by general social improvements. In these countries, national and individual aspirations favored a controlled rate of population increase and provided the intense motivation needed to lower fertility in a relatively short period of time. Abortion, especially in Japan, has played a major role in the rapid fertility transition depicted by this model.

3. Delayed model: The delayed model describes the relatively recent and yet-to-be completed transition of most developing countries. Although slow, unsteady decline in mortality began in some of these countries shortly after the turn of the century, rapid and truly substantial declines in mortality have been registered only since World War II. Both national and international programs of "population control" designed to hasten fertility decline artificially are prominent features of this model for countries where death control has far outstripped birth control.

   Despite unmistakable gains in the survival of women and children, infant and childhood mortality remains excessively high in most of these countries and in some, females of reproductive age continue to have higher mortality risks than males in the same age group. Although most countries in Latin America, Africa and Asia fit this model, important differences between these areas suggest the utility of developing sub models, particularly with regard to the varying responses of fertility and socioeconomic conditions to national development programs

The following diagram represents the three models. England and Wales represent the classical model, Japan represents the accelerated model, Sri Lanka and Chili represent the delayed model.



Birth Rates, Death Rates, and Population Size over the Last Two Centuries in Four Different Areas
Illustrating the Demographic Changes that Prompted the Development of the Epidemiological Transition Model. Modified By Omran.

Determinants of Disease

1. Eco biological: Changing patterns of immunity, vectors (such as the black rat partially responsible for spreading bubonic plague in Europe), and the movement of pathogenic organisms. These alter the frequency of epidemic infectious diseases as well as chronic infections and other illnesses that affect fertility and infant mortality.
2. Socioeconomic: Political and cultural determinants, including standards of living, health habits, hygiene and nutrition. Hygiene and nutrition are included here, rather than under medical determinants, because their improvement in western countries was largely a byproduct of social change rather than a result of medical design.
3. Medical/Public health: specific preventive and curative measures used to combat disease, including improved public sanitation, immunization and the development of decisive therapies. Medical and public health factors came into play late in the western transition but have an influence early on certain accelerated and delayed transitions.

Current Evidence

The majority of the literature on the epidemiological transition that has been published confirms the context-specific nature of the epidemiological transition – while there is an overall all-cause mortality decline, the nature of cause-specific mortality declines differs across different contexts. Increasing obesity rates in high-income countries are further confirming the epidemiological transition theory as the epidemic leads to an increase in NCDs. In low- and middle-income countries, there are signs of a protracted transition with the double burden of communicable and non-communicable diseases [15, 16].

A recent review of cause-specific mortality rates from 12 low- and middle-income countries in Asia and sub-Saharan Africa by Santosa and Byass (2016)[13]shows broadly that low- and middle-income countries are rapidly transitioning to lower total mortality and lower infectious disease mortality. A more macro-level analysis from the Global Burden of Disease data conducted by Murray and others (2015) found that while there is a global trend towards decreasing mortality and increasing NCDs prevalence, this global trend is being driven by country-specific effects as opposed to a broader transition; further, there are varying patterns within and between countries, which makes it difficult to have a single unified theory of epidemiological transition.

Stages of Epidemiologic Transition

1. Stage I: Characterized by the prevalence of infectious and parasitic diseases, accidents and animal attacks, and natural checks on population.
2. Stage II: Characterized by receding pandemics and improvements in sanitation, nutrition, and medical care, which lower the crude death rate.
3. Stage III: Characterized by prevalence of degenerative and man-made diseases, heart diseases, cancer, diabetes, obesity, and others.
4. Stage IV: Characterized by delayed degenerative diseases and extended life expectancy due to medical advances.
5. Stage V: Characterized by potential resurgence of infectious diseases due to globalization.

Stages of Health Transition as Adapted from Vorster et al. Nutrition Rev 1999

1. Early stages: Characterized by high fertility, high mortality, high prevalence of infectious diseases

and high prevalence of undernutrition.

2. Middle stages: Characterized by reduced mortality, changing age structure, receding infection, poor environmental conditions, and receding famines.
3. (Early) late stages: Characterized by reduced fertility, aging, chronic lifestyle-related diseases (non-communicable diseases) and diet chronic diseases predominate.
4. (Late) late stages: Characterized by reduced fertility, very old age, shift to other chronic diseases such as mental illnesses and osteoarticular diseases, healthy diet and lifestyle predominate.

Chapter 3

Epidemiological Studies

Researchers can select from various types of epidemiological studies. However, to choose the right design, you first need to know well each type and its advantages and disadvantages, which you consider carefully so that you can make a sensible decision.

Classification

Epidemiological studies are classified as

- Descriptive studies
- Analytical studies

Other classification [17]:

- Observational or non-intervention studies. These include descriptive studies, cohort studies, and case control studies. The investigator doesn't interfere with the exposure of the study population to the risk factor, and just observes what is happening. Another name for these studies is non-interventional studies.
- Non-observational or intervention studies. These include experimental studies and clinical trials. The investigator interferes with the exposure, and assigns one group of the study population to the exposure and doesn't expose the other group. These studies are also called interventional studies.

Descriptive studies

Descriptive studies include activities related to characterizing the patterns of disease occurrence in terms of person, place, and time, leading to generation of hypothesis. It studies the distribution of disease or the health event in the population. In descriptive epidemiology we

1. Observe
2. Count cases (events)
3. Describe a health-related event in terms of time, place, and person
4. Calculate rates
5. Compare rates
6. Develop hypotheses

Descriptive epidemiology is characterized by being inexpensive, and time saving. It describes disease patterns and it formulates research questions and hypotheses. It is unable to test hypotheses.

It is concerned with studying the distribution of the disease or the health condition in the community in terms of:

- How common is the disease?
- Who gets the disease?
- Where does the disease occur?
- When does the disease occur?

How common or unusual is disease occurrence:
Endemic: The ongoing, usual level of or constant presence of a disease within a given population or geographic area.

Epidemic: The occurrence of disease at a higher level than normally expected in a population.
Pandemic: An epidemic that is widespread across a country, continent, or possibly worldwide.
Descriptive epidemiology – Person

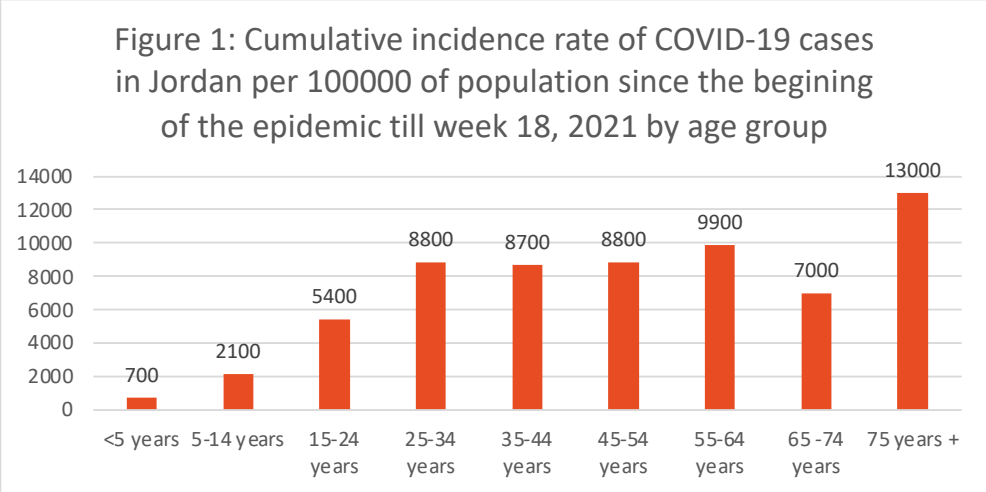
- Disease does not occur at random.
- Not all persons within a population are equally likely to develop a particular disease.

Variation of occurrence in relation to personal characteristics may reflect differences in the level of:

1. Exposure to causal factors
2. Susceptibility to causal factors
3. The need for some level of both susceptibility and exposure

The personal characteristics that are commonly examined with respect to disease occurrence are age, sex, race, marital status, education, income, occupation, and others.

Example of the Distribution of Disease by Age as One of Personal Characteristics



Source: Epidemiological bulletin on COVID-19, 2021, Ministry of Health, Jordan.

Descriptive epidemiology – Place: Where are the highest disease rates? Where are the lowest disease rates? Does the disease rate vary by country, region, etc.?

Example of the distribution of disease by place:

Cumulative Incidence of COVID-19 Cases in Jordan per 100 000 Population by Governorates from the Beginning of Epidemic Until Week 18, 2021

Governorates	Cumulative Incidence/100 000
Amman	7600.0
Irbid	5599.0
Zarqa	4640.0
Balqa	8301.0
Mafraq	4480.0

cont'd

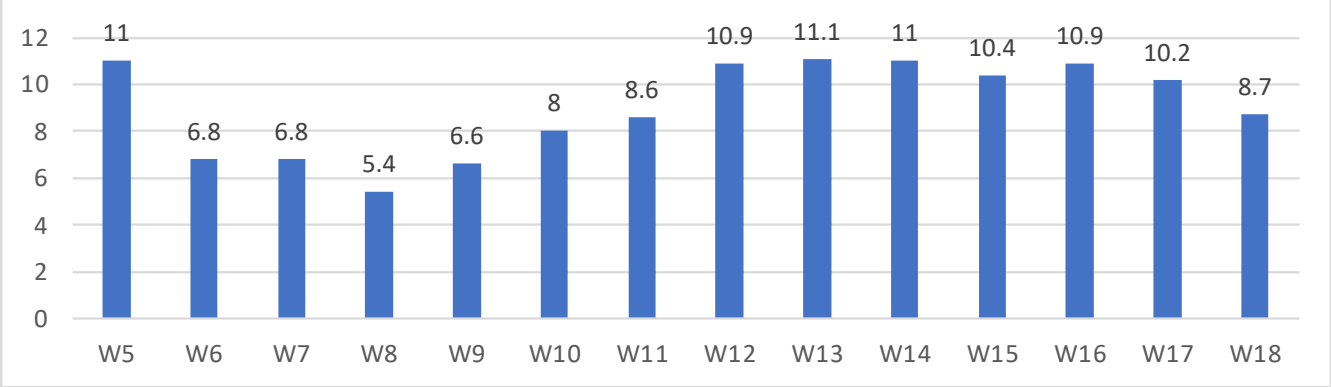
Karak	**7333.0**
Jarash	**8286.0**
Aqaba	**9953.0**
Ajloun	**9823.0**
Madaba	**8424.0**
Tafileh	**11346.0**
Maan	**5464.0**

Source: Epidemiological bulletin on COVID-19, 2021, Ministry of Health, Jordan.

Descriptive epidemiology – Time: When is the disease common? When is the disease rare? Is the frequency of the disease in the present different from its frequency in the past?

Example of the distribution of disease by time:

Case Fatality Ratios of Hospitalized Patients During the Weeks of the Second Wave of COVID-19 Epidemic in Jordan Until Week 18, 2021



Source: Epidemiological bulletin on COVID-19, 2021, Ministry of Health, Jordan.

Types of diseases trends as related to time:

1. Secular trend (long term): Occurrence of disease over a long period of time, decades, and years.
2. Seasonal trend (medium term): Variation in occurrence by season of the year, like diarrheal diseases which increase during summer.
3. Cyclic trend (short term): Like what happens during epidemics of diseases.

Uses of descriptive studies

- Display patterns of occurrence
- Focus on person, place and time
- Used for program planning
- Generate hypotheses

Types of descriptive studies

- Case reports and case series
- Descriptive incidence studies
- Cross-sectional studies (descriptive prevalence studies)
- Ecologic (correlational) studies

Case reports and case series
- They describe a profile of a case or series of cases.
- They may generate new hypotheses.
- They form an interface between clinical medicine and epidemiology.
- They only provide numerator data.
- They don't provide measures of disease occurrence.

A case report is considered to be the most basic type of descriptive study of individuals, consisting of a careful, detailed report by one or more clinicians of the profile of a single patient.

Case report example: In 1961, a case report was published of a 40-year-old pre-menopausal woman who developed a pulmonary embolism (PE) 5 weeks after beginning to use an oral contraceptive (OC) preparation to treat endometriosis [18]. PEs are not common in this age group. Could OC use be the cause? However, OC use is not uncommon in this age range. Are women who develop PEs more likely to use OCs than are women who do not use OCs?
Case series: A description of the characteristics of a number of patients with a given disease (series of case reports).

Example of case series: 1980–1981: In a 6-month period, five young, previously healthy homosexual men were diagnosed at three hospitals as having Pneumocystis carinii pneumonia (PCP). This clustering of cases was striking in that PCP was seen almost exclusively in the elderly. This unusual circumstance suggested that these individuals were suffering from an unknown underlying condition. (AIDS).

Descriptive incidence studies
They study the patterns of occurrence of incident cases (often from surveillance data) in a defined population (denominators from census) during a specified period of time and study the distribution of cases by factors of interest.
Ecologic (correlational) studies

Characteristics:
- Exposure and disease at aggregate (e.g., country) level
- Data from groups, not individuals
- Unit of observation is a population not individuals
- These studies provide a crude way of exploring associations between factors and disease
- They are considered to be hypothesis generating rather than hypothesis testing
- The group rather than the individual is the unit of comparison
- Limitation: No individual link between exposure and disease and the aggregate association may not lead to individual association
- Main advantages: It is quick to conduct, inexpensive, and it uses available data

Example 1 of ecologic study: Comparison of the trend of saccharin usage to the trend in bladder cancer rates in the United States.

Example 2 of ecologic study: During the period 1950–1969, the national cancer institute in the USA

published maps showing mortality rates for cancer by county. They found a clustering of cancer lung in the Northeast and Southeast and on the Gulf coast. An ecological study correlating these county rates with industry concentration data revealed that lung cancer mortality was elevated in counties with paper, chemical petroleum, and transportation industries. This study hypothesized that lung cancer in certain coastal areas was associated with the ship-building industry. There was a need to test this hypothesis to prove the association. Blot et al., 1978 conducted a case control study and confirmed the association between shipbuilding and lung cancer, possibly as a result of asbestos exposure.

Although correlation studies might be an inexpensive means of generating hypotheses, one should be cautious in drawing conclusions regarding individual risk based on group risk. This is because data on individual behaviors that may influence risk has not been collected. This may cause the ecological fallacy (a type of bias) [19].

Characteristics of cross-sectional studies

- It is a snapshot of a well-defined population. (The study takes place at a single point in time or a period of time.)
- Unlike longitudinal studies, which look at a group of people over an extended period, cross-sectional studies are used to describe what is happening at the present moment.
- They measure exposures and outcomes at the same time.
- This type of research is frequently used to determine the prevailing characteristics in a population at a certain point in time. For example, a cross-sectional study might be used to determine if exposure to specific risk factors might correlate with particular outcomes. Researchers might collect cross-sectional data on past smoking habits and current diagnoses of lung cancer, for example. While this type of study cannot demonstrate cause and effect, it can provide a quick look at correlations that may exist at a particular point.

For example, researchers may find that people who reported engaging in certain health behaviors were also more likely to be diagnosed with specific ailments. While cross-sectional studies cannot prove for certain that these behaviors caused the condition, such studies can point to a relationship worth investigating further.

Advantages of cross-section studies

- Capture all existing diseases (serosurveys capture asymptomatic cases). It allows researchers to look at numerous characteristics at once (age, income, gender, etc.).
- Quick and inexpensive, these studies can generate a lot of useful data.

Disadvantages of cross-section studies

- Uncertain temporal relationship: Cannot guarantee that exposure precedes the disease. They can't differentiate Cause and Effect; they don't establish the temporal sequence of events necessary for drawing causal inferences.
- Other variables can affect the relationship between the inferred cause and outcomes, and this type of research doesn't allow for conclusions about causation.
- It is affected by the survivor effect; that means, death cases and cured cases in the study population cannot be measured.

Cross-sectional survey example

In examinations carried out by the National Health Survey in the US, the prevalence of coronary heart disease (CHD) and level of serum cholesterol were determined at the same visit. The fact that those with CHD had a higher mean cholesterol level than those without CHD does not necessarily lead to the conclusion that elevated serum cholesterol increases the risk of CHD. This may well be so. But it is only by demonstrating an increase in CHD in people with previously elevated cholesterol that a causal inference about the relationship may be drawn.

Analytical studies

An analytical study attempts to establish causes or risk factors for certain problems. This is done by comparing two or more groups, some of which have or develop the problem and some of which have not.

The goal is to determine the relationship between exposure and disease with validity and precision. They assess determinants of disease, focus on risk factors, causes and analyze distribution of exposures and disease. They are used for testing hypotheses and looking for and quantifying associations. The basic design in analytic epidemiology is to examine if exposures are correlated with disease (i.e., are exposure and disease linked?). The hallmark feature that distinguishes an analytic study from a descriptive study is the comparison group in analytic studies.

To link exposure and disease, we have to answer the following questions: What is the exposure? Who are the exposed? What are the potential health effects? What approach does it take to study the relationship between exposure and effect? To examine the link of exposure to disease, there needs to be standardized evaluation of exposure, as well as disease.

The basic analytical studies consist of

1. Cross sectional studies (prevalence studies)
2. Cohort studies (also called prospective or follow up studies, they can sometimes be retrospective)
3. Case control studies (retrospective studies)
4. Experimental studies, including clinical trials

Analytical cross-sectional studies

Although cross-sectional studies are classified mostly as a type of descriptive study, sometimes they are considered as a type of analytic studies and used to test epidemiologic hypothesis when current values of the exposure variables are unalterable over time, like factors present at birth (eye color, blood group etc.) Cross-sectional studies widely used to compare different behavioral factors but with the limitation of lacking temporal relationship or causation

Measuring Association in Analytical Cross-sectional Surveys

	Ill	**Not ill**	**Total**
Exposed	a	B	a +b
Unexposed	c	D	c +d
Total	a+c	b +d	a+b+c+d

Prevalence exposed = a / (a+b)
Prevalence unexposed = c/ (c+d)
Prevalence ratio (PR) = Prevalence exposed / Prevalence not exposed
Excess prevalence (EP) =Prevalence exposed - Prevalence not exposed
PR and EP are cross-sectional analogs of relative risk and excess risk in cohort studies.

Cohort studies

Cohort studies are a type of longitudinal study. They are an approach that follows research participants over a period of time (often many years). Specifically, cohort studies recruit and follow participants who share a common characteristic, such as a particular occupation or demographic similarity. Cohort studies are a type of epidemiological research used to investigate the causes of disease and to establish links between risk factors and health outcomes. The word "cohort" means "a group of people". These types of studies look at groups of people. During the follow-up period, some of the cohort will be exposed to a specific risk factor or characteristic; by measuring outcomes over a period of time. It is then possible to explore the impact of this variable (e.g., identifying the link between smoking and lung cancer in the British Doctors Study)[8].

Cohort studies are, therefore, of particular value in epidemiology, helping to build an understanding of what factors increase or decrease the likelihood of developing disease. Cohort studies are the cornerstone of epidemiological research, providing an understanding of risk factors for disease based on findings in thousands of participants over many years.

Cohort design is a type of non-experimental or observational study design. In a cohort study, the participants do not have the outcome of interest to begin with. They are selected based on the individual's exposure status. They are then followed over time to evaluate for the occurrence of the outcome of interest.

The characteristic feature of a cohort study is that the investigator identifies subjects at a point in time when they do not have the outcome of interest and compares the incidence of the outcome of interest among groups of exposed and unexposed (or less exposed) subjects.
A well-designed cohort study can provide powerful results. In a cohort study, an outcome or disease-free study population is first identified by the exposure or event of interest and followed in time until the disease or outcome of interest occurs.

In a cohort study, two or more groups are formed based on exposure. The outcome is then determined and compared between groups. As a result, a cohort study can measure several outcomes in the same study. Cohorts may be fixed (every individual in a cohort starts at the same time and is followed up for a similar period of time) or dynamic (individuals recruited to or leave the cohort at different times). Individuals within cohorts are followed up over time, usually to determine the incidence of the condition under study. There are two types of cohort studies – prospective and retrospective (or historical). The two groups of cohorts (exposed and un-exposed) are followed prospectively over time to track the development of new disease.

The distinguishing feature of a prospective cohort study is that at the time that the investigators begin enrolling subjects and collecting baseline exposure information, none of the subjects had developed any of the outcomes of interest.

Cohort studies provide the best information about the causation of disease, because you follow persons from exposure to the occurrence of the disease. An added advantage is that you can examine a range of outcomes/diseases caused by one exposure (e.g., heart disease, lung disease, renal disease caused by smoking).

Cohort studies are an effective and robust method of establishing cause and effect. As they are usually large in size, researchers are able to draw confident conclusions regarding the link between risk factors and disease. In many cases, because participants are often free of disease at the commencement of the study, cohort studies are particularly useful at identifying the timelines over which certain behaviors can contribute to disease. However, the nature of cohort studies can cause challenges. Collecting prospective data on thousands of participants over many years (and sometimes decades) is complex, time-consuming, and expensive.

Participants may drop out, increasing the risk of bias; equally, it is possible that the behavior of participants may alter because they are aware that they are part of a study cohort. The analysis of data from these large-scale studies is also complex, with large numbers of confounding variables making it difficult to link cause and effect.

In 1951, Richard Doll and Austin Bradford-Hill [8] commenced a ground-breaking research project by writing to all registered doctors in the UK to ask about their smoking habits. The British Doctors Study recruited and followed up over 40 000 participants, monitoring mortality rates and causes of death over the subsequent years and decades. Even by the time of the first set of preliminary results in 1954, there was evidence to link smoking with lung cancer and increased mortality. Over the following decades, the study provided further definitive evidence of the health risks from smoking and was extended to explore other causes of death (e.g., heart disease). The Doctors' Health Survey is one of the largest, most ambitious and best-known cohort studies and demonstrates the value of this approach in supporting our understanding of disease risk.

The British Doctors Study is just one of many large-scale, long-term cohort studies carried out to enhance understanding of the causes of disease and to help to develop evidence-based guidelines for healthier living. For example, the Framingham Heart Study—which commenced in 1948 and is now following up a third generation that includes grandchildren of the original cohort of participants from a Massachusetts town—has provided extensive data on the risk factors for cardiovascular disease and underpinned international guidelines on prevention.

One way to make a cohort study less time-consuming is to carry it out retrospectively. This is a more pragmatic approach, as it can be completed more quickly using historical data. However, this retrospective approach increases the risk of bias in the sampling of the cohort, with greater likelihood of missing data.

Retrospective cohort studies are often used as an intermediate step between a weaker preliminary study and a prospective cohort study, as the results gleaned from a retrospective cohort study strengthen assumptions behind a future prospective cohort study.

A retrospective cohort study would be a good fit for your research if

1. A prospective cohort study is not yet feasible for the variables you are investigating.
2. You need to quickly examine the effect of an exposure, outbreak, or treatment on an outcome.
3. You are seeking to investigate an early-stage or potential association between your variables of interest.

Advantages of cohort studies

1. Clarity of temporal sequence: Cohort studies more clearly indicate the temporal sequence between exposure and outcome, because in a cohort study, subjects are known to be disease-free at the beginning of the observation period when their exposure status is established.
2. Allow calculation of incidence: Cohort studies allow you to calculate the incidence of disease, so you can calculate:
   - Absolute risk (incidence)
   - Relative risk (risk ratio or rate ratio)
   - Risk difference or attributable risk
   - Attributable proportion (attributable risk %)
3. Facilitate study of rare exposures: They are particularly useful for evaluating the effects of rare or unusual exposures, because the investigators can make it a point to identify an adequate number of subjects who have an unusual exposure, like, exposure to toxic chemicals, adverse effects of drugs (e.g., thalidomide) or treatments (e.g., radiation treatments for ankylosing spondylitis), unusual occupational exposures (e.g., asbestos, or solvents in tire manufacturing).
4. Allow examination of multiple effects of a single exposure.
5. Avoid selection bias at enrollment: Cohort studies, especially prospective cohort studies, reduce the possibility that the results will be biased by selecting subjects for the comparison group who may be more or less likely to have the outcome of interest, because in a cohort study the outcome is not known at baseline when exposure status is established. Nevertheless, selection bias can occur in retrospective cohort studies (since the outcomes have already occurred at the time of selection), and it can occur in prospective cohort studies as a result of differential loss to follow up.

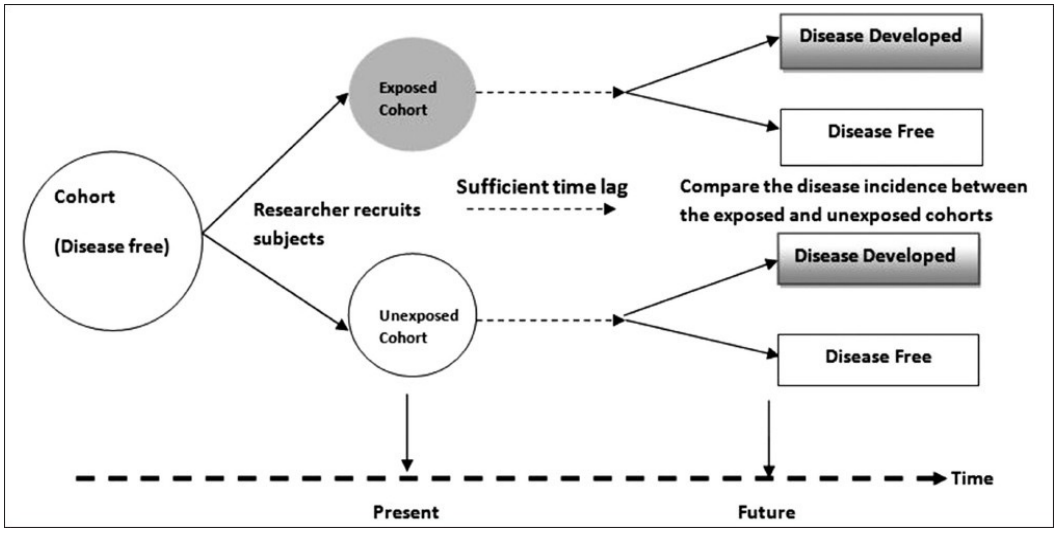
Disadvantages of prospective cohort studies

1. You may have to follow large numbers of subjects for a long time.
2. They can be very expensive and time consuming.
3. They are not good for rare diseases.
4. They are not good for diseases with a long latency.
5. Differential loss to follow up can introduce bias.
6. Many subjects are needed for rare diseases.
7. Follow-up: logistics and losses.
8. Exposure can change over time.
9. Retrospective: requires suitable records.
10. Changes in practice, usage, and exposure may make findings irrelevant.

Disadvantages of retrospective (historical) cohort studies

1. As with prospective cohort studies, they are not good for very rare diseases.
2. If one uses records that were not designed for the study, the available data may be of poor quality.

3. There is frequently an absence of data on potential confounding factors if the data was recorded in the past.
4. It may be difficult to identify an appropriate exposed cohort and an appropriate comparison group.
5. Differential losses to follow up can also bias retrospective cohort studies.



Synonyms for cohort studies

- Follow up studies
- Prospective studies
- Incidence studies
- Longitudinal studies
- Forward-looking studies
- Concurrent studies

In prospective cohort studies, when the study starts, the relevant events (exposure) may or may not have occurred, but the outcomes have certainly not yet occurred.

Exposure……. Study starts ……Time…… Disease occurrence

Study starts…… Exposure……. Time…… Disease occurrence

In retrospective cohort studies, the relevant events (both the exposures and outcomes of interest) have already occurred when the study is initiated.

Exposure……Time……Disease occurrence……Time……. Study starts

In cohort studies, the investigator:

- Does not determine the exposure.
- Enrolls subjects on the basis of exposure status (or enrolls all members of a group, then classifies by exposure). There are two common options for enrollment, option 1 – enroll an entire population. Identify which members of the population have the exposure and which do not. Option 2 – enroll subjects who are known to have been exposed. Then enroll a comparable group who do not have the exposure.
- Follows subjects over time and records occurrence of health event (outcome of interest).
- Compares rates of disease occurrence among exposed and unexposed groups of persons.
- Calculates risk ratios or rate ratios (relative risk).

- Calculates tests of significance or confidence intervals.

The most difficult aspect of conducting cohort studies is documenting disease occurrence among exposed and unexposed subjects, as study subjects can be lost to follow up, and therefore their disease status cannot be determined.
Measures of disease occurrence in cohort studies:

1. Cumulative incidence (attack rate, risk): Number of new cases at end of follow-up divided by number of disease-free persons at start of follow-up.
2. Person-time rate (incidence density): Number of new cases at the end of follow-up divided by person-time at risk (e.g., person-years of disease-free follow-up).

Cohort 2 by 2 Table

Exposure status	Disease	No disease	Total
Exposed	a	B	a+b
Unexposed	c	D	c+d

Risk (incidence) among exposed = a / a+b
Risk (incidence) among unexposed = c / c+d
Relative risk (risk ratio) = Risk Exposed / Risk Unexposed = (a / a+b) / (c / c+d)
Relative risk is the measure of association in cohort studies.

Famous Outbreak of Gastroenteritis Following a Church Picnic in Oswego County, New York

Eating Vanilla Ice Cream	Diseased	Not Diseased	Total
Yes (exposed)	43	11	54
No (not exposed)	3	18	21
Total	46	29	75

Source: Centers for Disease Control: An outbreak of gastrointestinal illness following a church supper, Atlanta, July 1976.

Incidence among exposed = 43/54 = 79.6%
Incidence among not-exposed = 3/21= 14.3%

Relative risk (risk ratio = 79.6/14.3 = 5.6, this means that those who ate vanilla ice cream were 5.6 times more likely to have the disease than those who did not eat the substance.
Attributable risk (risk excess) = Incidence (Risk) among exposed - Incidence (Risk) among not-exposed = 79.6- 14.3 = 65.3.
The attributable risk, or AR, amongst the exposed group tell us how much of the disease in that group is "added" because of the exposure.

From the AR, we know what risk is attributed to that exposure, or how "dangerous" that exposure is, and what is the excess risk due to that exposure, or how much more dangerous life has become because of that exposure.

Case Control Study (Retrospective Study)
A case control study is one that compares two groups of people – those with the disease or condition

under study (cases) and a very similar group of people who do not have the disease or condition (controls). Researchers study the medical and lifestyle histories of the people in each group to learn what factors may be associated with the disease or condition. For example, one group may have been exposed to a particular substance that the other was not. It is also called retrospective study. Case-control studies start with the identification of a group of cases (individuals with a particular health outcome) in a given population and a group of controls (individuals without the health outcome) to be included in the study, and we look back in their past to ask about specific exposures.

In a case-control study, the prevalence of exposure to a potential risk factor(s) is compared between cases and controls. [20]. If the prevalence of exposure is more common among cases than controls, it may be a risk factor for the outcome under investigation. A major characteristic of case-control studies is that data on potential risk factors are collected retrospectively and as a result may give rise to bias. This is a particular problem associated with case-control studies and therefore needs to be carefully considered during the design and conduct of the study.

Issues in the design of case-control studies

1. As with all epidemiological investigations, the beginning of a case-control study should begin with the formulation of a clearly defined hypothesis.
2. It is essential that the case definition is clearly defined at the outset of the investigation to ensure that all cases included in the study are based on the same diagnostic criteria.
3. The source of cases needs to be clearly defined. Case-control studies may use incident or prevalent cases.
4. The use of incident (newly diagnosed) cases is considered as preferential, as the recall of past exposure(s) may be more accurate among newly diagnosed cases. In addition, the temporal sequence of exposure and disease is easier to assess among incident cases.
5. The use of prevalent cases (individuals who have had the outcome under investigation for some time) may give rise to recall bias as prevalent cases may be less likely to accurately report past exposures(s). As a result, the interpretation of results based on prevalent cases may prove more problematic.
6. Cases may be recruited from a number of sources; for example, they may be recruited from a hospital, clinic, or may be population based. Population-based case control studies are generally more expensive and difficult to conduct.

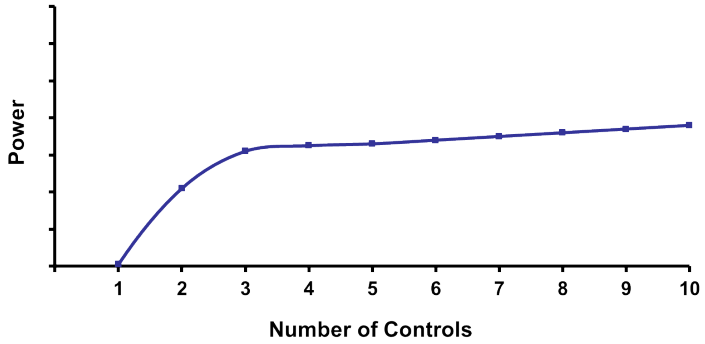
Selection of controls

A particular problem inherent in case-control studies is the selection of a comparable control group. Controls are used to estimate the prevalence of exposure in the population, which gave rise to cases. Therefore, the ideal control group would comprise a random sample from the general population that gave rise to the cases. However, this is not always possible in practice. The goal is to select individuals in whom the distribution of exposure status would be the same as that of the cases in the absence of an exposure disease association. That is, if there is no true association between exposure and disease, the cases and controls should have the same distribution of exposure. The source of controls is dependent on the source of cases. To minimize bias, controls should be selected to be a representative sample of the population which produced the cases. For example, if cases are selected from a defined population such as a general practitioner register, then controls should comprise a sample from the same register.

In case-control studies where cases are hospital-based, it is common to recruit controls from the hospital population. However, the choice of controls from a hospital setting should not include individuals with an outcome related to the exposure(s) being studied. For example, in a case-control study of the association between smoking and lung cancer, the inclusion of controls being treated for a condition related to smoking (e.g., chronic bronchitis) may result in an underestimate of the strength of the association between exposure (smoking) and outcome.

Recruiting more than one control per case may improve the statistical power of the study, though including more than 4 controls per case is generally considered to be no more efficient.

The graph below illustrates the issue of how many controls you should enroll per case. Statistical power on the Y-axis addresses the issue. If you have plenty of cases, say, 50 or more, then you should have plenty of power with 1 control per case, e.g., 50 cases and 50 controls. Note that from this curve, you get a big jump in power when you go from 1 control per case to 2 controls per case, and you get a smaller jump going from 2 to 3 controls per case. But you get much less improvement after 3 or 4 controls per case.



Measuring exposure status

Exposure status is measured to assess the presence or level of exposure for each individual for the period of time prior to the onset of the disease or condition under investigation, when the exposure would have acted as a causal factor. Note that in case-control studies the measurement of exposure is established after the development of disease and as a result is prone to both recall and observer bias. Various methods can be used to ascertain exposure status. These include:

- Standardized questionnaires
- Biological samples
- Interviews with the subject
- Interviews with spouse or other family members
- Medical records
- Employment records
- Pharmacy records
- Others

The procedures used for the collection of exposure data should be the same for cases and controls.

Common sources of bias in case-control studies

Due to the retrospective nature of case-control studies, they are particularly susceptible to the effects of bias, which may be introduced as a result of a poor study design or during the collection of

exposure and outcome data. Because the disease and exposure have already occurred at the outset of a case control study, there may be differential reporting of exposure information between cases and controls based on their disease status. For example, cases and controls may recall past exposure differently (recall bias). Similarly, the recording of exposure information may vary depending on the investigator's knowledge of an individual's disease status (interviewer/observer bias). Therefore, the design and conduct of the study must be carefully considered, as there are limited options for the control of bias during the analysis.

Selection bias is a particular problem inherent in case-control studies, where it gives rise to non-comparability between cases and controls. Selection bias in case control studies may occur when cases (or controls) are included in (or excluded from) a study because of some characteristics they exhibit which is related to exposure to the risk factor under evaluation. The aim of a case-control study is to select study controls who are representative of the population which produced the cases. Controls are used to provide an estimate of the exposure rate in the population. Therefore, selection bias may occur when those individuals selected as controls are unrepresentative of the population that produced the cases.

The potential for selection bias in case control studies is a particular problem when cases and controls are recruited exclusively from hospitals or clinics. Hospital patients tend to have different characteristics than the population, for example, they may have higher levels of alcohol consumption or cigarette smoking. If these characteristics are related to the exposures under investigation, then estimates of the exposure among controls may be different from those in the reference population, which may result in a biased estimate of the association between exposure and disease.

Berkesonian bias is a bias introduced in hospital-based case-control studies, due to varying rates of hospital admissions. As the potential for selection bias is likely to be less of a problem in population-based case-control studies, neighborhood controls may be a preferable choice when using cases from a hospital or clinic setting. Alternatively, the potential for selection bias may be minimized by selecting controls from more than one source, such as by using both hospital and neighborhood controls. Selection bias may also be introduced in case-control studies when exposed cases are more likely to be selected than unexposed cases.

Analysis of case-control studies

The odds ratio (OR) is used in case-control studies to estimate the strength of the association between exposure and outcome. Note that it is not possible to estimate the incidence of disease from a case control study.

The results of a case-control study can be presented in a 2x2 table as follows:

	Cases	**Controls**	**Total**
Exposed	a	b	a+b
Not Exposed	c	d	c+d
Total	a+c	b+d	a+b+c+d

The odds ratio (OR) is a measure of the odds of disease in the exposed compared to the odds of disease in the unexposed (controls), It is an approximation of the relative risk used in cohort studies.

The odds ratio is calculated as follows:
OR = ad/bc.
Example:

Calculation of the OR from a Hypothetical Case-control Study of Smoking and Cancer of the Pancreas Among 100 cases and 400 Controls

	Cases	Controls	Total
Smokers	60 (a)	100 (b)	160
Non-smokers	40(c)	300 (d)	340
Total	100	400	500

OR = (60 times 300) / 100 times 40 = 4.5
The OR calculated from these hypothetical data estimates that smokers are 4.5 times more likely to develop cancer of the pancreas than non-smokers.

Strengths and weaknesses of case-control studies
Strengths:
- Cost effective relative to other analytical studies such as cohort studies
- Case-control studies are retrospective, and cases are identified at the beginning of the study; therefore, there is no long follow up period (as compared to cohort studies).
- Efficient for the study of diseases with long latency periods
- Efficient for the study of rare diseases
- Good for examining multiple exposures

Weaknesses:
- Particularly prone to bias; especially selection, recall, and observer bias
- Case-control studies are limited to examining one outcome
- Unable to estimate incidence rates of disease
- Poor choice for the study of rare exposures
- The temporal sequence between exposure and disease may be difficult to determine.

Experimental (Non-observational or Interventional) Studies
Experimental study: A study in which the investigator intentionally alters one or more factors and controls the other study conditions in order to analyze the effects of doing so and in which conditions are under the direct control of the investigator. In contrast to observational studies, the investigator assigns exposure to the study subjects. Although experiments provide the strongest evidence for testing any hypothesis, they are rarely possible in the human population.

Characteristics of experimental studies [21]:
- Tests hypotheses about cause-and-effect relationships. It is the best research methodology to establish cause and effect relationships among variables.
- Two groups are compared; the experimental group receives treatment and the control group does not receive treatment or receive a placebo.
- Subjects are randomly assigned to treatment and control groups. Randomization is a process of assigning individuals to groups randomly and forms groups that are equivalent and differ only by chance. It takes place before the experiment begins. Measurements are collected at the same time for both groups.

- Experimental studies have at least one intervention and one comparison group.
- Each group is followed prospectively until there is a well-defined endpoint/outcome.
- Trials can be conducted on individuals or communities.
- Allocation of participants to the study group and control group should be through random allocation.
- Masking or blinding is necessary in these trials to avoid bias.

Types of blinding

1. Single blinded studies: Only the subject is blind regarding the group to which the subject is assigned.
2. Double blinded studies: Both the investigator and the subject are blind regarding the group to which the subject is assigned.
3. Triple blinded studies: The investigator, the subject and the person responsible for data analysis are all blind regarding the group to which the subject is assigned. Blinding is very important when the outcome is subjectively determined. If the outcome is death or stroke, blinding becomes less essential.

If groups are similar at baseline, differences in outcome can reasonably be attributed to the action of the intervention (assuming the study is well done, has high validity and high credibility in establishing causality).

Types of interventional (experimental) studies

1. True experimental study
2. Natural experiments
3. Quasi-experimental study
4. Clinical trials

Steps of true experimental studies

1. Develop research questions.
2. Write protocol.
3. Get ethical approval.
4. Enroll study sample.
5. Assign participants to exposure and control groups.
6. Monitor participants in each group for study outcome (first occurrence of disease, improvement, side effects, etc.).
7. Analyze the data.

The experimental population in which the trial is conducted must be stable/available over time to obtain complete and accurate follow-up for the duration of the trial, and the trial should yield valid results generalizable to the reference population.

After invitation of the study population to be enrolled in the trial, they should be provided with information on the purpose of the trial, study procedures, possible risks and benefits, possibility of allocation to a group receiving no treatment or usual care and conducting screening for their eligibility to enter the study according to predetermined criteria, such as absence of previous history of study end points, definite need for study treatment, or contraindication.

Eligibility criteria, which are the requirements that determine whether an individual can be included in a study, must be considered. They often include age, gender, medical conditions, previous treatment history, and other characteristics unique to the protocol. Inclusion criteria (factors that must be met for an individual to be included in a study) and exclusion criteria (factors that prevent an individual from being included in a study) should be applied carefully.
Eligibility criteria help define the study population, ensure safe and ethical research, and promote scientific validity.

The trial usually consists of two groups – the intervention group which receives the therapy or the preventive means or other intervention (as in operational research) and the comparison group which receives nothing; placebo or the usual care. The two groups are called the arms of the study.

Random allocation of participants to study arms
Randomization (random allocation) means allocation of each participant to the study group and comparison group randomly (by chance). So, each participant has the same chance to be in the intervention or in the comparison group. The purposes of random allocation are:
1. To achieve baseline comparability between the two groups (two arms of the study) of both measured and unmeasured characteristics, so difference in outcome can be attributed to difference in the intervention.
2. To remove investigator bias in assigning patients to groups.
3. To increase validity of statistical tests.

Experimental designs have numerous advantages compared to other epidemiological methods. Randomization, when used, tends to balance confounding variables across the various study groups, especially variables that might be associated with changes in the disease state or the outcome of the intervention under study. Detailed information and data are collected at the beginning of an experimental study to develop a baseline; this same type of information is also collected at specified follow-up periods throughout the study. The investigators have control over variables such as the dose or degree of intervention. The blinding process reduces distortion in assessment. And, of great value, and not possible with other methods, is the testing of hypotheses. Most important, this design is the only real test of cause–effect relations.

Advantages of experimental studies
1. They provide researchers with a high level of control.
2. Experimental research provides conclusions that are specific, and it is accepted by the medical and scientific community as the "gold standard," i.e., has great credibility.
3. Randomization minimizes confounding.
4. Experimental research allows cause and effect to be determined perfectly.
5. Possibility of selection bias is minimal since the study is under the control of the investigator, and by ensuring only one factor differs among study arms.
6. Attractive statistically since many statistical methods assume random assignments.

Disadvantages of experimental studies
1. May be complex and expensive, especially for low incidence outcomes
2. Are exposed to certain ethical problems
3. May lack generalizability/representativeness

4. Resistance to randomization by clinicians and patients
5. Administrative complexity
6. Statistical complexity for complex designs

Natural experiments

Naturally occurring circumstances in which subsets of the population have different levels of exposure to a supposed causal factor, in a situation resembling an actual experiment where human subjects would be randomly allocated to groups. Natural experiments are those events where an observer or researcher does not have control. Most public health interventions, such as the implementation of tobacco control policies, can be considered natural experiments. Natural experiments lack a random assignment which may result in multiple threats to causal inference.

Natural experiments are often used to study situations in which controlled experimentation is not possible, such as when an exposure of interest cannot be practically or ethically assigned to research subjects. Situations that may create appropriate circumstances for a natural experiment include policy changes, weather events, and natural disasters. Natural experiments are most commonly used in the fields of epidemiology, political science, psychology, and social science. Because natural experiments do not randomize participants into exposure groups, the assumptions and analytical techniques customarily applied to experimental designs are not valid for them. Rather, natural experiments are quasi experiments and must be thought about and analyzed as such. For this reason, natural experiments will never unequivocally determine causation in a given situation. Nevertheless, they are a useful method for researchers, and if used with care they can provide additional data that may help with a research question and that may not be obtainable in any other way. The major limitation in inferring causation from natural experiments is the presence of unmeasured confounding. Examples of natural experiments: Snow on cholera in the middle of 19th century, and atomic bombing in Japan in 1945.

Quasi-experimental study

Quasi-experimental designs are generally used to establish the causality (effect of independent variable on dependent variable) in situations where researchers cannot randomly assign the subjects to groups or for various reasons no control group is available for an experimental study. They involve the manipulation of independent variables to observe the effect on the dependent variable. However, they lack at least one of the two characteristics of the true experimental design – randomization or control group.

Types of quasi-experimental designs

- Non-randomized control group design: In this design, experimental and control groups are selected without randomization.The dependent variables are observed in experimental and control groups before the intervention. Later the experimental group receives treatment and after that, observation of dependent variable is carried out for both experimental and control groups to assess the effect of treatment on the experimental group.

  Example: This method was used to study the effect of integrated care on quality of work of caregivers in nursing homes. The study's purpose was to examine the implementation of integrated care in nursing homes and its effect on the quality of caregivers' work. The data was collected by a questionnaire. The results showed that the intervention was successful on the somatic wards of

the nursing home. The caregivers in these wards were more successful in creating a home-like environment for their residents

- Time-series design: This design is useful when the experimenter wants to measure the effects of a treatment over a long period. The experimenter would continue to administer the treatment and measure the effects several times during the course of the experiment. Generally it is a single-subject research, in which the researcher carries out an experiment on an individual or a small number of individuals, by alternating between administering and then withdrawing the treatment to determine the effectiveness of the intervention.

  Example 1 in time-series design: A researcher might assess pain levels of a group of patients with lower back pain. After 3 weeks of pain assessment, subjects are taught special exercises to reduce that pain. Over the next 3 weeks, pain levels would again be measured.

  Example 2 in time-series design: Measuring a child's school performance on a weekly basis, and then introducing a new teaching technique, then again measuring on a weekly basis.

Advantages of quasi-experimental design

- It is more frequently used because it is more practical and feasible to conduct.
- It is more suitable for real-world natural settings than true experimental research design.
- It allows researchers to evaluate the impact of independent variables under naturally occurring conditions.
- It may be able to establish a causal relationship. Some of the hypotheses are practically answered through this design only.

Disadvantages of quasi-experimental design

- There is no control over extraneous variables (confounders) influencing the dependent variable.
- The absence of a control group or lack of control over the research setting makes the results of this design less reliable and weak for the establishment of causal relationship between independent and dependent variables.

Human subjects protection (ethical considerations)

Ethical problems are major issues to be considered in experimental trials. The research question/s should be appropriate for the study and the approval of the institutional review board should be obtained. All study subjects must be properly informed about the various study procedures and an informed consent must be signed by all subjects before embarking on the study.

Clinical trials

A clinical trial is a prospectively planned experiment for the purpose of evaluating potentially beneficial therapies or treatment. Clinical trials are part of clinical research and at the heart of all medical advances. They look at new ways to prevent, detect, or treat diseases [22]. Treatments might be new drugs, new drug combinations, new surgical procedures or devices, or new ways to use existing treatments.

Types of clinical trials

- Therapeutic (clinical) trials: Enroll patients with an existing disease or disability to determine the ability of an agent or procedure to reduce symptoms, prevent recurrence, or decrease the risk of death from the disease.
- Preventive (prophylactic) trials: Enroll individuals without the study disease to determine the ability of agents or procedures to reduce the risk of developing disease among disease-

free individuals, like trials conducted on vaccines. They can be conducted on individuals or communities.

- Community trials: These are when the intervention is applied to communities rather than individuals, like conducting health education campaigns, or fluoridation of water supply in certain communities. The communities should be randomized. Since there is no selection of individual subjects for the study, there are savings in the costs of individual screening and enrollment. Baseline and follow-up community surveys are essential to measure the effect of the intervention. Surveillance system data that is already in place can be used for this purpose.

Clinical trial phases

Phase one

Before phase one, your product should be tested in a lab to determine whether or not there is enough evidence to say that it might provide a benefit to the patients who are going to use it.

You won't know if your product is safe or effective for sure at this point, but you should have enough data collected to suggest that it's worth looking into further. During phase one, a small group of willing participants who would be ideal patients for your product will use it and report their results back to you. These participants will be monitored closely for any signs of adverse effects or symptoms you may need to report.

Phase two

Assuming the results from phase one found that your product was effective and safe for use, you'll be ready to test it on a higher number of participants in phase two.

If your product is a medication, you will have established the appropriate dose amount in phase one. Now that your product has been tested on enough people to be reasonably sure that it's worth distributing on a larger scale, you'll move on to phase three.

Phase three

In phase three, you will test your product in more depth on much larger groups of people. You'll also monitor the side effects and the efficacy of the drug for every patient.

Placebos may be used in your trial if the product you're testing is a medication. Placebos help act as a control so you can ensure that your patient's perception and hopes don't skew the results you're getting.

Phase four

At this stage of the clinical trials, your product has received approval from the concerned agency and your marketing campaigns are underway. Studies performed in phase 4 focus on discovering the long-term effects of your product and gathering more data on how it interacts with other drugs if your product is a medication.

Designs of clinical trials

1. Parallel design
2. Crossover design
3. Cluster design
4. Factorial design

Parallel design

Parallel arm design is the most commonly used study design. In this design, subjects are randomized

to one or more study arms and each study arm will be allocated a different intervention. After randomization, each participant will stay in their assigned treatment arm for the duration of the study. A parallel design, also called a parallel group study, compares two or more treatments. Participants are randomly assigned to either group, treatments are administered, and then the results are compared. A key element of this design is randomization, which places participants randomly into a group.
It is a type of clinical study where two groups of treatments, A and B, are given so that one group receives only A while another group receives only B. Other names for this type of study include "between patient" and "non-crossover". In a parallel study, the two treatment groups can either consist of two completely separate treatments (i.e., different drugs), or simply different doses of a common drug. One major aspect of a parallel study is randomization. This ensures that the results are accurate and have a lower risk of being biased. A parallel study would be more appropriate if any concerns about carryover effects were present.

This type of study might also be more beneficial if the disease or disorder being studied has a likely chance of progression during the time in which the study takes place.

Subjects are randomized to groups, groups followed in parallel to determine effect in each (most common design). Patients are randomized to treatment group and control group and remain on the treatment or placebo throughout the duration of the trial.

Schematic presentation of parallel design
Intervention group arm----------Treatment A---------Follow up------Measuring outcome
Control group arm----------Treatment B or placebo ----------Follow up------- Measuring outcome

Crossover designs
In this design, subjects are exposed to more than one treatment, where subjects are randomly assigned to different orders of treatment. This design is more efficient in establishing the highest possible similarity among subjects exposed to different conditions, where groups compared obviously have equal distribution of characteristics. Though crossover design is considered an extremely powerful research design, sometimes it is not effective because when subjects are exposed to two different treatment/conditions, their responses to the second treatment/condition may be influenced by their experience in the first treatment/ condition.

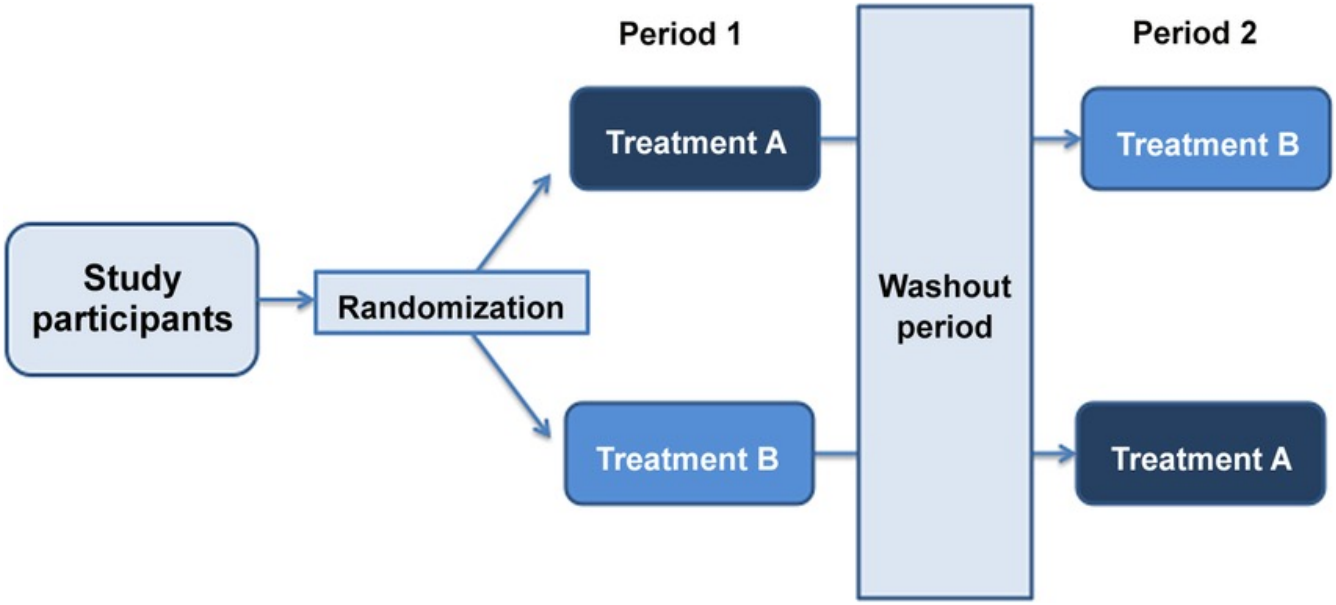
Crossover trials are designed so that each recruited subject receives both active and controlled treatments in either order for a specified duration, with a washout period between treatments when no treatment is administered. In such trials, patients act as their own control. The fundamental assumption of crossover design is that patients usually have a chronically stable condition that will not vary between when they are taking the first and second treatments.

The reason to consider a crossover design when planning a clinical trial is that it could yield a more efficient comparison of treatments than a parallel design, i.e., fewer patients might be required in the crossover design in order to attain the same level of statistical power or precision as a parallel design.

Each patient serves as his/her own matched control. Every patient receives both treatment A and B. A comparison is made of the subject's response on A vs. B.

In medical clinical trials the disease should be chronic and stable, and the treatments should not result in total cures but only alleviate the disease condition. If treatment A cures the patient during the first period, then treatment B will not have the opportunity to demonstrate its effectiveness when the patient crosses over to treatment B in the second period. Therefore, this type of design works only for chronic conditions where there is no cure such as asthma and the treatments attempt to improve quality of life. An adequate washout period is essential between periods of a crossover study.

Example: When we compare the effectiveness of a chlorhexidine mouth care protocol on subjects of group one and a saline mouth care protocol on subjects of group two. Later, the treatment is swapped, where group one receives saline and group two receives chlorhexidine. In such studies, subjects serve as their own control.



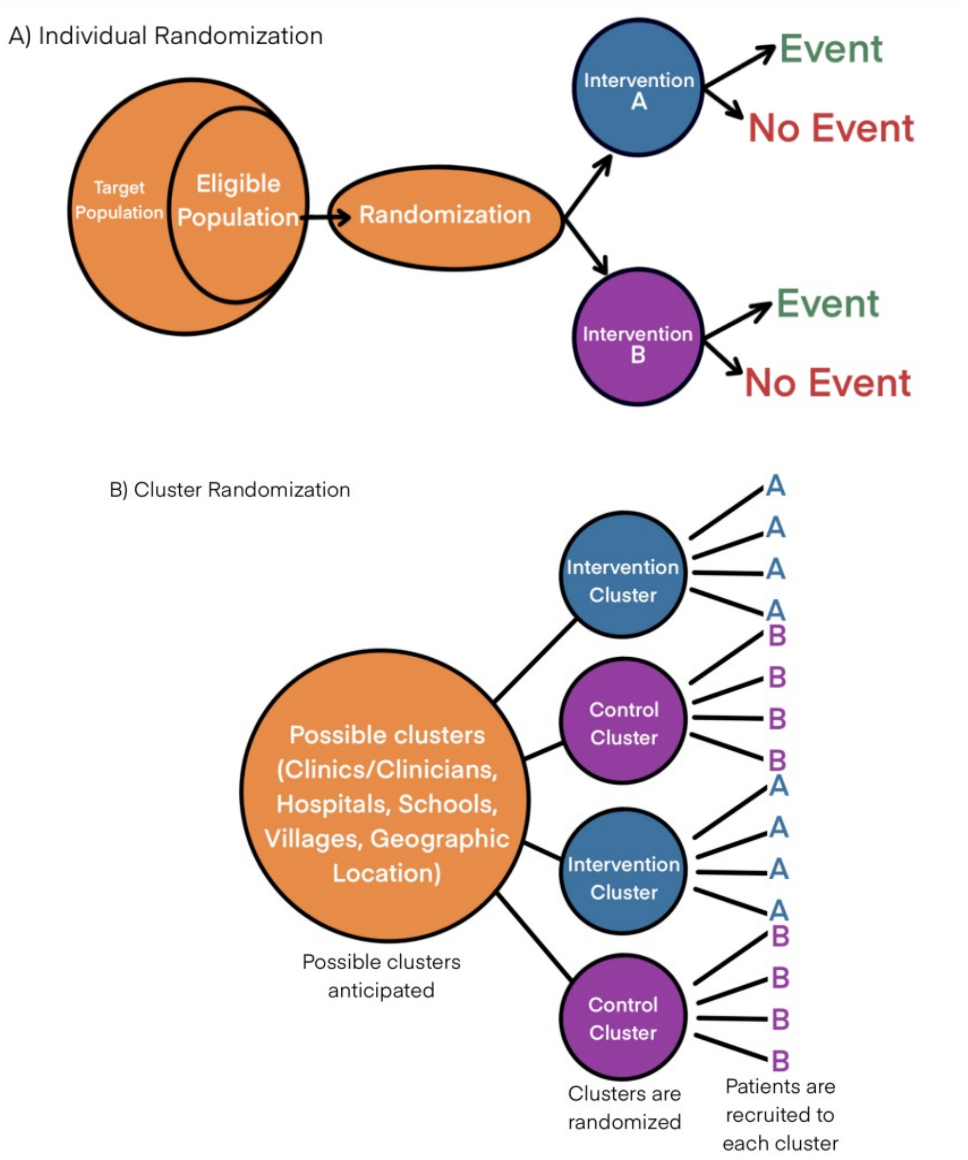
Schematic Presentation of Cross-over Design

Cluster design

Cluster randomized trials (CRTs) involve randomization of groups (clusters) of individuals to control or intervention conditions. CRT design is commonly used to evaluate non-drug interventions, such as policy and service delivery interventions. CRTs differ from individually randomized ones in that the unit of randomization is something other than the individual.

CRT is a randomized controlled trial in which pre-existing groups, called clusters of individuals, are randomly allocated to treatment arms. CRTs can be used when individual randomization to treatment arms is not possible, or the intervention is naturally applied to a whole cluster. A cluster randomized design is associated with a loss in statistical power and additional complexity in design, conduct, and analysis.

Diagram Showing Individual Randomization Versus Cluster Randomization



Example

A controlled randomized trial of typhoid vaccination was conducted in eastern Kolkata, India [23]. In this study, 60 000 residents were randomly assigned by neighborhood clusters (N = 80) to receive either typhoid (N = 40) or hepatitis (N = 40) vaccination. Of note, owing to logistics, eligibility, or consenting reasons, only about 61% of the residents were vaccinated, and the other 39% were not. Culture confirmation of subsequent enteric fever episodes was obtained, along with oral informed consent, for the entire population over the subsequent 2 years of follow-up.

The overall benefit among all residents of the typhoid vaccine clusters was 57% (95% confidence interval [CI], 37%–71%) protection from bacteremia typhoid fever. Of interest, there was evidence of benefit among residents of the typhoid vaccine–assigned neighborhoods whether typhoid vaccinated (61% [95% CI, 53%–91%]) or not (44% [2%–69%]) versus those among the hepatitis vaccine–assigned neighborhoods.

Factorial design

Factorial design is the evaluation of more than one treatment for safety and/or efficacy compared to a control. In this design, researchers manipulate two or more independent variables simultaneously

to observe their effects on the dependent variables. This design is useful when there are more than two independent variables, called factors, to be tested. This design also facilitates the testing of several hypotheses at a single time.

Typical design incorporates 2 by 2 or 2 by 3 factorial, but it can be in any combination.

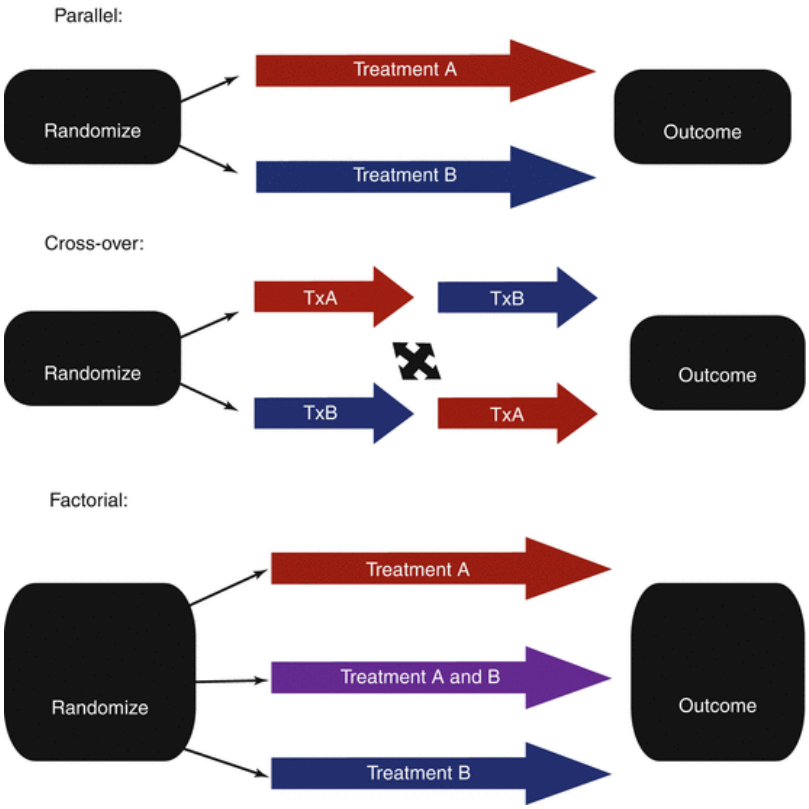


Diagram Showing Factorial Design in Comparison with Other Clinical Trial Designs

In a factorial design, two drugs or interventions can be simultaneously evaluated. With two drugs, four combinations of treatments and placebo are possible. Patients are randomly assigned to each group. For example, one group might receive drug A and drug B. Another group would receive drug A and a placebo. Another would receive drug B and a placebo, and another would receive two placebos. The factorial design can be very efficient, as data is gathered on two drugs at the same time. A drawback of the factorial design is the concern over potential drug interactions; however, this design also allows for the determination of synergies between two treatments.

In the simplest psychology experiments, researchers look at how one independent variable affects one dependent variable. But what happens if researchers want to look at the effects of multiple independent variables? This type of study that involves the manipulation of two or more variables is known as a factorial design.

For example, imagine that a researcher wants to do an experiment looking at whether sleep deprivation has a negative impact on reaction times during a driving test. If he were to only perform

the experiment using these variables – the sleep deprivation being the independent variable and the performance on the driving test being the dependent variable – it would be an example of a simple experiment. However, let's imagine that he is also interested in learning if sleep deprivation impacts the driving abilities of men and women differently. He has just added a second independent variable of interest (sex of the driver) into his study, which now makes it a factorial design.

One of the big advantages of factorial designs is that they allow researchers to look for interactions between independent variables. An interaction is a result in which the effects of one experimental manipulation depends upon the experimental manipulation of another independent variable. For example, if a researcher wants to test the effects of a memory-enhancing drug. Participants are given one of three different drug doses, and then asked to either complete a simple or complex memory task. The researchers note that the effects of the memory drug are more pronounced with simple memory tasks, but not as apparent when it comes to complex tasks. In this 3×2 factorial design, there is an interaction effect between the drug dosage and the complexity of the memory task.

A university wants to assess the starting salaries of their graduates. The study looks at graduates working in four different employment areas: accounting, management, finance, and marketing. In addition to looking at the employment sector, the researchers also look at gender. In this example, the employment sector and gender of the graduates are the independent variables, and the starting salaries are the dependent variables. This would be considered a 4×2 factorial design.

Researchers want to determine how the amount of sleep a person gets the night before an exam impacts performance on a math test the next day. But the experimenters also know that many people like to have a cup of coffee (or two) in the morning to help them get going. So, the researchers decide to look at how the amount of sleep and the amount of caffeine influence test performance. The researchers then decide to look at three levels of sleep (4 hours, 6 hours, and 8 hours) and only two levels of caffeine consumption (2 cups versus no coffee). In this case, the study is a 3×2 factorial design. Factorial experiments have rarely been used in the development or evaluation of clinical interventions. However, factorial designs offer advantages over randomized controlled trial designs, the latter being much more frequently used in such research. Factorial designs are highly efficient (permitting evaluation of multiple intervention components with good statistical power) and present the opportunity to detect interactions amongst intervention components.

Qualitative Epidemiologic Studies

Qualitative studies have considerable possibilities within the domain of healthcare research. A wide variety of phenomena that cannot be explained using the quantitative approach can be explored using a qualitative method. The major types of qualitative research designs are narrative research, phenomenological research, grounded theory research, ethnographic research, historical research, and case study research [24].

The greatest strength of the qualitative research approach lies in the richness and depth of the healthcare exploration and description it makes. In health research, these methods are considered as the most humanistic and person-centered way of discovering and uncovering thoughts and actions of human beings.

While quantitative research method uses data, which are measures of values and counts and are often described using statistical methods which in turn aids the researcher to draw inferences, qualitative research incorporates the recording, interpreting, and analyzing of non-numeric data with an attempt to uncover the deeper meanings of human experiences and behaviors.

Mixed methods research, the third methodological approach, involves the collection and analysis of both qualitative and quantitative information with an objective to solve different but related questions, or at times the same questions [25, 26].

In healthcare, qualitative research is widely used to understand patterns of health behaviors, describe lived experiences, develop behavioral theories, explore healthcare needs, and design interventions. Because of its ample applications in healthcare, there has been a tremendous increase in the number of health research studies undertaken using qualitative methodology.

According to Munhall, "Qualitative research involves broadly stated questions about human experiences and realities, studied through sustained contact with the individual in their natural environments and producing rich, descriptive data that will help us to understand those individual's experiences."[27]

The qualitative method of inquiry examines the "how" and "why" of decision making, rather than the "when," "what," and "where." Unlike quantitative methods, the objective of qualitative inquiry is to explore, narrate, and explain the phenomena and make sense of the complex reality.

The ultimate strength of the qualitative research approach lies in the richness of the data and the descriptions and depth of exploration it makes. Hence, qualitative methods are considered as the most humanistic and person-centered way of discovering and uncovering thoughts and actions of human beings.

Differences between quantitative and qualitative studies

The quantitative and qualitative forms of studies vary based on their underlying objectives. They are in no way opposed to each other; instead, these two methods are like two sides of a coin. The critical differences between them are summarized in the following table [25]:

Areas	Quantitative Study	Qualitative Study
Nature of reality	Assumes there is a single reality	Assumes existence of dynamic and multiple realities
Goal	Test and confirm hypotheses	Explore and understand phenomena
Data collection methods	Highly structured methods like questionnaires, inventories, and scales	Semi-structured like in-depth interviews, observations, and focus group discussions
Design	Predetermined and rigid	Flexible and emergent
Reasoning	Deductive process to test the hypothesis	Primarily inductive to develop a theory or hypothesis
Focus	Concerned with the outcomes and prediction of the causal relationships	Concerned primarily with processes, rather than outcomes or products

cont'd

Sampling	Rely largely on random sampling methods	Based on the purposive sampling methods
Sample size determination	Involves a-priori sample size calculation	Collect data until data saturation is achieved
Sample size		
	Relatively large	Small sample size but studied in-depth
Data analysis	Variable-based and use statistical or mathematical methods	Case-based and use non statistical descriptive or interpretive methods

Qualitative questions are exploratory and open-ended. A well-formulated study question forms the basis for developing a protocol and guides the selection of design and data collection methods.

Qualitative research questions generally involve two parts – a central question and related sub questions. The central question is directed towards the primary phenomenon under study, whereas the sub questions explore the subareas of focus. It is advised not to have more than five to seven sub questions.

A commonly used framework for designing a qualitative research question is the "PCO framework," wherein P stands for the population under study, C stands for the context of exploration, and O stands for the outcome/s of interest [28].
Example: In the question, "What are the experiences of mothers on parenting children with thalassemia?" the population is "mothers of children with thalassemia," the context is "parenting children with thalassemia," and the outcome of interest is "experiences."

Review of literature

In quantitative research, the researchers do an extensive review of scientific literature prior to the commencement of the study. However, in qualitative research, only a minimal literature search is conducted at the beginning of the study. This is to ensure that the researcher is not influenced by the existing understanding of the phenomenon under the study. The minimal literature review will help the researchers to avoid the conceptual pollution of the phenomenon being studied. Nonetheless, an extensive review of the literature is conducted after data collection and analysis.

Types of qualitative study designs

- Narrative study
- Phenomenological study
- Grounded theory study
- Ethnographic study
- Historical study
- Case study

Narrative study

Narrative study focuses on exploring the life of an individual and is ideally suited to tell the stories of individual experiences. The purpose of narrative research is to utilize story telling as a method in communicating an individual's experience to a larger audience.

Narrative study encompasses the study of individual experiences and learning the significance of those experiences. The data collection procedures include mainly interviews, field notes, letters, photographs, diaries, and documents collected from one or more individuals.

Example: Karlsson et al. undertook a narrative inquiry to "explore how people with Alzheimer's disease present their life story." Data were collected from nine participants. They were asked to describe their life experiences from childhood to adulthood, then to their current life and their views about their future life [29].

Phenomenological study

This methodology has its origin from philosophy, psychology, and education. The unit of analysis of phenomenology is the individuals who have had similar experiences of the phenomenon. Interviews with individuals are mainly considered for data collection, although documents and observations are also useful.

The phenomenological approach is further divided into descriptive and interpretive phenomenology. Descriptive phenomenology focuses on the understanding of the essence of experiences and is best suited in situations that need to describe the lived phenomenon.
Interpretive phenomenology moves beyond the description to uncover the meanings that are not explicitly evident. The researcher tries to interpret the phenomenon, based on their judgment rather than just describing it.

Example: A phenomenological study conducted by Cornelio et al. aimed at describing the lived experiences of mothers in parenting children with leukemia. Data from ten mothers were collected using in-depth semi-structured interviews and were analyzed using Husserl's method of phenomenology. Themes such as "pivotal moment in life," "the experience of being with a seriously ill child," "having to keep distance with the relatives," "overcoming the financial and social commitments," "responding to challenges," "experience of faith as being key to survival," "health concerns of the present and future," and "optimism" were derived. The researchers reported that the essence of the study was "chronic illness such as leukemia in children results in a negative impact on the child and on the mother."[30]

Grounded theory study

Grounded theory has its base in sociology. The primary purpose of grounded theory is to discover or generate theory in the context of the social process being studied. The major difference between grounded theory and other approaches lies in its emphasis on theory generation and development.
Data collection in grounded theory research involves recording interviews from many individuals until data saturation.

Example: Williams et al. conducted grounded theory research to explore the nature of the relationship between the sense of self and eating disorders. Data were collected from 11 women with a lifetime history of anorexia nervosa and were analyzed using the grounded theory methodology. Analysis led to the development of a theoretical framework on the nature of the relationship between the self and anorexia nervosa. [31]

Ethnographic study

Ethnography has its base in anthropology, where the anthropologists used it for understanding the culture-specific knowledge and behaviors.

In health sciences research, ethnography focuses on narrating and interpreting the health behaviors of a culture-sharing group. A "culture-sharing group" in an ethnography represents any "group of people who share common meanings, customs or experiences."

In health research, it could be a group of physicians working in rural care, a group of medical students, or it could be a group of patients who receive home-based rehabilitation.

To understand cultural patterns, researchers primarily observe individuals or group of individuals for a prolonged period. Ethnographers collect data using a variety of methods such as observation, interviews, audio-video records, and document reviews.

A written report includes a detailed description of the culture-sharing group with emic and etic perspectives. When the researcher reports the views of the participants it is called emic perspectives and when the researcher reports his or her views about the culture, it is termed etic.

Example: The aim of the ethnographic study by LeBaron et al. was to explore the barriers to opioid availability and cancer pain management in India. The researchers collected data from fifty-nine participants using in-depth semi-structured interviews, participant observations, and document reviews. The researchers identified significant barriers by open coding and thematic analysis of the formal interview. [32]

Historical study

Historical study is the "systematic collection, critical evaluation, and interpretation of historical evidence." The purpose of historical research is to gain insights from the past and involves interpreting past events in the light of the present.

The data for historical research are usually collected from primary and secondary sources. The primary source mainly includes diaries, firsthand information, and writings. The secondary sources are textbooks, newspapers, second or third-hand accounts of historical events, and medical/legal documents. The written report describes "what happened." "how it happened," "why it happened," and its significance and implications to current clinical practice.

Example: Lubold (2019) analyzed the breastfeeding trends in three countries (Sweden, Ireland, and the United States) using a historical qualitative method. Through analysis of historical data, the researcher found that strong family policies, adherence to international recommendations, and adoption of baby-friendly hospital initiatives could greatly enhance the breastfeeding rates. [33]

Case study

Case study research focuses on the description and in-depth analysis of the case(s) or issues illustrated by the case(s). Case studies are best suited for the understanding of case(s), thus reducing the unit of analysis into studying an event, a program, an activity, or an illness.

Observations, one-on-one interviews, artifacts, and documents are used to collect the data, and the

analysis is done through the description of the case. A written case study report includes a detailed description of one or more cases.

Example: Perceptions of post-stroke sexuality in women of childbearing age were explored using a qualitative case study approach by Beal and Millenbrunch. A semi-structured interview was conducted with a 36-year-old mother of two children with a history of acute ischemic stroke. The data were analyzed using an inductive approach. The authors concluded that "stroke during childbearing years may affect a woman's perception of herself as a sexual being and her ability to carry out gender roles." [34]

Sampling in qualitative study

Qualitative researchers widely use non-probability sampling techniques. The four widely used sampling techniques are convenience sampling, purposive sampling, snowball sampling, and intensity sampling. The selection of a sampling technique depends on the nature and needs of the study.

In convenience sampling, researchers collect data from subjects who are selected based on accessibility, geographical proximity, ease, speed, and or low cost.

Purposive sampling is a widely used sampling technique. It involves identifying a population based on already established sampling criteria and then selecting subjects who fulfill that criteria to increase credibility.

The snowball sampling method is also known as chain referral sampling or network sampling. The sampling starts by having a few initial participants, and the researcher relies on these early participants to identify additional study participants. It is best adopted when the researcher wishes to study a stigmatized group, or in cases where finding participants is likely to be difficult by ordinary means.
Intensity sampling involves identifying information-rich cases that manifest a phenomenon of interest. It requires prior information and considerable judgment about the phenomenon of interest. The researcher should do some preliminary investigations to determine the nature of the variation. Intensity sampling will be done once the researcher identifies the variation across the cases (extreme, average, and intense) and picks the intense cases from them.

Sample size determination

A-priori sample size calculation is not undertaken in the case of qualitative research. Researchers collect the data from as many participants as possible until they reach the point of data saturation. Data saturation or the point of redundancy is the stage where the researcher no longer sees or hears any new information.

Data saturation gives the idea that the researcher has captured all possible information about the phenomenon of interest. Since no further information is being uncovered as redundancy is achieved, at this point the data collection can be stopped.

Data collection in qualitative study

The various strategies used for data collection in qualitative research include in-depth interviews (individual or group), focus group discussions (FGDs), participant observation, narrative life history, document analysis, audio materials, videos or video footage, text analysis, and simple observation. Among these, the three popular methods are the FGDs, one to one in-depth interviews, and participant observation.

FGDs are useful in eliciting data from a group of individuals. They are normally built around a specific topic and are considered the best approach to gather data on an entire range of responses to a topic. Group size in an FGD ranges from 6 to 12.

Depending upon the nature of the participants, FGDs can be homogeneous or heterogeneous.

One-on-one in-depth interviews are best suited to obtaining individuals' life histories, lived experiences, perceptions, and views, particularly while exporting topics of sensitive nature. In-depth interviews can be structured, unstructured, or semi-structured. However, semi-structured interviews are widely used in qualitative research. Participant observations are suitable for gathering data regarding naturally occurring behaviors.

Operational Studies

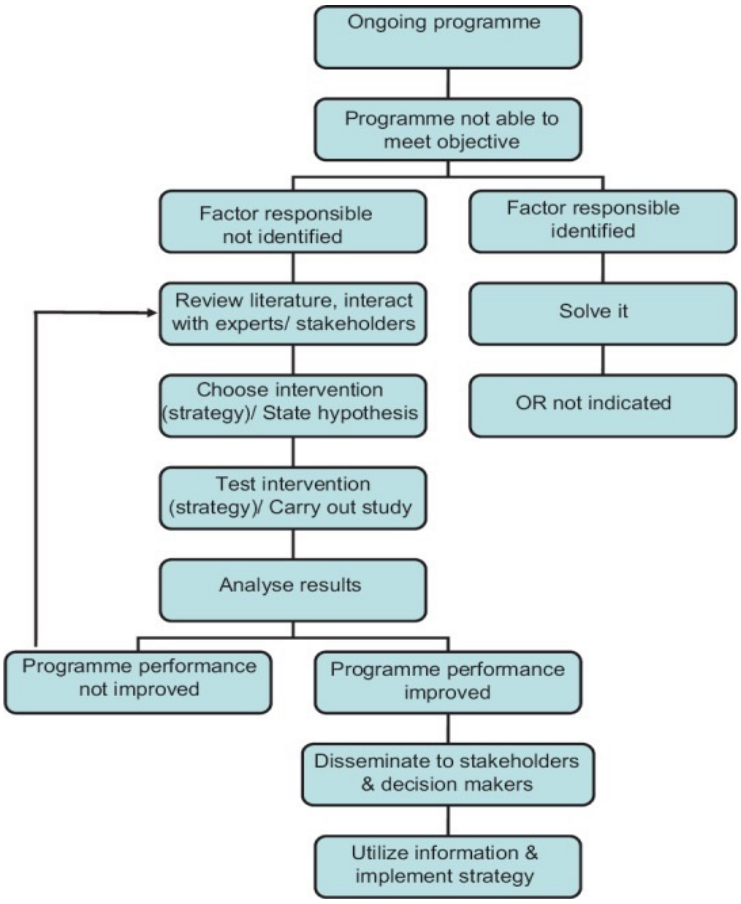
The dictionary of epidemiology defines operation research as "a systematic study of the working of a system with the aim of improvement." [1]

The International Union against TB and Lung Disease and many of its research partners define operational research as follows: "research into strategies, interventions, tools or knowledge that can enhance the quality, coverage, effectiveness or performance of the health system or program in which the research is being conducted." [35]

Characteristics of operational studies [36]

- Focuses on a specific problem in an ongoing program
- Involves research into the problem using principles of epidemiology
- Tests more than one possible solution and provides a rational basis, in the absence of complete information, for the best alternative to improve program efficiency
- Requires close interaction between program managers and researchers
- Succeeds only if the research is conducted in the existing environment and study results are implemented in true letter and spirit.

The process of operational research involves the steps in the following diagram [36]:



Operational study is different from clinical or epidemiological study in that it examines a system (in this case the healthcare system) rather than focusing on an individual or a group of individuals (as in clinical or epidemiological study where patients are examined). In addition, operational study has at its core the goal of the improvement of a system (the healthcare system). To do this, it is necessary to identify challenges in the system and evaluate or recommend solutions.

There are three basic steps that guide an operational study:

1. Spell out well-defined goals and objectives of the health program or system in question.
2. Identify, prioritize, and articulate constraints and obstacles that prevent these objectives from being achieved.
3. Develop research questions that address the constraints.

How do operational studies differ from using routine data for quality improvement?

Those working in or responsible for health services can use routine data to drive quality improvement through data analysis, identification of gaps, development of quality improvement initiatives, and monitoring whether or not these have resulted in improvements to the service. However, it is impossible to differentiate between improvement due to the intervention (the quality improvement initiative) and other factors (management interest in the problem, improved monitoring of the problem, or other changes that occur with time, for example). Although the operational study also starts with identifying problems or challenges in the health system, what differentiates operation

study from the use of routine data to drive quality improvement is that it is hypothesis driven. The hypothesis is evaluated using rigorous scientific methods that allow for analytical comparisons, so that inferences can be made about the target population and used to inform policy and practice.

The true value of operational study to health programs is the improvement of health via the impact of research results on programmatic and policy decisions and on practice. The importance of operational research is its ability to address and solve local problems in delivering quality health services.
A necessary starting point is to identify the obstacles to providing high quality services, analyze why these obstacles occur and to adopt policies and practices to overcome them.

While the concept of operational study as an essential tool for health programs is widely accepted, challenges to successful implementation of comprehensive operation study activities at country level are numerous. The following are examples of these challenges:

- Many countries still operate in the absence of a detailed, systematic research plan, with clear linkages to program priorities, thus limiting the impact of research efforts.
- Implementing research studies in the absence of a carefully conducted situation analysis prevents many countries from achieving their desired goals.
- The appropriate external sources of support – financial, technical, and research mentoring – must be in place at all stages of planning and implementation of the study. These resources must allow local partners (rather than those providing funds or external experts) to set priorities. Such resources are insufficient or absent at some or all stages of operation study implementation in many countries.
- Training in operational study methodology is required for both service providers and academics.

Proposal overview of operational study [37]

Outline and key questions

Operational studies develop in response to a problem that stakeholders wish to address. It sets out to answer the following questions:

- What is the problem and why is it important to address this problem?
- What is already known about this problem?
- What does this study aim to achieve?
- How will this be achieved – what data is required and how will it be collected and analyzed?
- How will the work be undertaken and what resources will be required?

- **The proposal outline consists of the following:**

1. Title page which contains
    - Proposal title
    - Investigator names
    - Affiliated institutions
    - Contact details
    - Total budget requested
2. Summary
3. An introduction which consists of
    - Context
    - Problem statement

- Problem analysis
- Justification

4. Defining the research in terms of
   - Research question
   - Hypothesis
   - Aims and objectives
5. Study methods involve
   - Study setting
   - Study design
   - Target and study population
   - Sampling, sample size, and power
   - Variables, definitions, and data sources
   - Data collection
   - Data management
   - Data analysis plan
   - Quality assurance
6. Ethical considerations, if any
7. Application of research findings
   - Strengths and limitations
   - Dissemination and stakeholder engagement
   - Implications for policy and practice
8. Project management that includes
   - Roles and responsibilities
   - Project timelines
   - Budget
   - Regulatory aspects
9. References
10. Appendices which include
    - Researcher's curriculum vitae
    - Data collection tools e.g., case report form
    - Data dictionary

Examples of operation study in healthcare settings [36]
Example 1: It was demonstrated in India, through an operational study, that the successful implementation of DOTS strategy throughout the country had led to a reduction in the prevalence of TB, reduction in fatalities due to TB and the release of hospital beds occupied by TB patients; and thereby a potential gain to the Indian economy.

Example 2: OR has been successfully used in hospital settings too. In Latin America, unsafe abortions used to be one of the most common causes of high maternal mortality. Billings and Bensons reviewed ten completed OR projects conducted in public sector hospitals in seven Latin American countries. Their findings indicated that sharp curettages replaced by manual vacuum aspirations for conducting abortion reduced the requirement of resources for post-abortion care, reduced costs and length of hospital stays, and reduced maternal mortality.

Chapter 4

Errors in Epidemiological Studies

Before deciding on an exposure-outcome relationship, we need to exclude the possibility that a finding can be attributed to one or more of the following:

1. Random error (chance)
2. Bias (systematic error)
3. Confounder

Random Error (Chance)

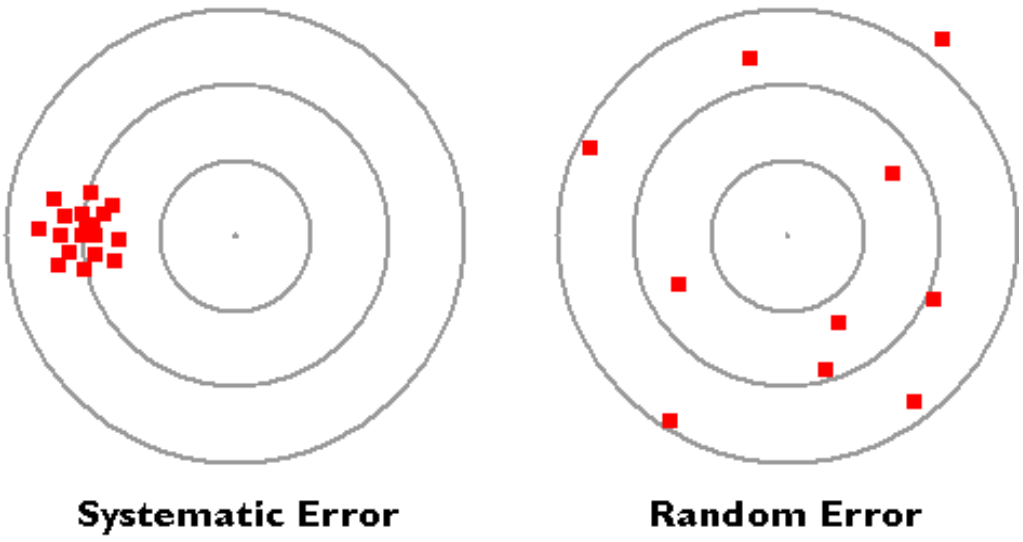
Random error is the chance difference between a particular variable's observed and true values. It usually results from the experimenter's inability to take the same measurement in exactly the same way to get exactly the same number. Random error is a statistical error that is wholly due to chance. These errors are unpredictable and can't be replicated by repeating the experiment again. Random error (also called unsystematic error, system noise or random variation) has no pattern. One minute your readings might be too small; the next they might be too large. You can't predict random error and these errors are usually unavoidable.

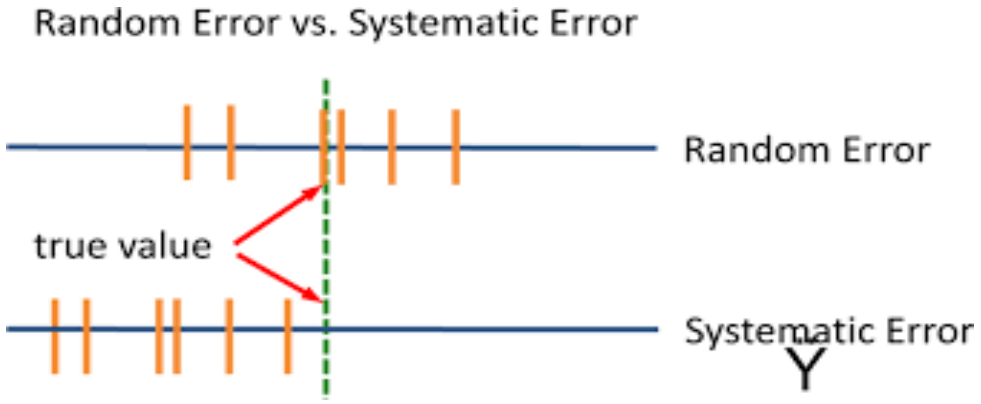
If you take multiple measurements, the values cluster around the true value. Thus, random error primarily affects precision. The main reasons for random errors are limitations of instruments, environmental factors, and slight variations in procedure.

Examples of random errors:

- When weighing yourself on a scale, you position yourself slightly differently each time
- Measuring your height is affected by minor posture changes

The following two diagrams show random error vs. systematic error:





Random error can be reduced by using an average measurement from a set of measurements or by increasing the sample size. These errors can be minimized but not eliminated. An error is considered random if the value of what is being measured sometimes goes up or sometimes goes down. A very simple example is our blood pressure. Even if someone is healthy, it is normal that their blood pressure does not remain exactly the same every time it is measured.

Chance (random error) means random variation. In statistics, random variation usually assumes that you are taking a sample from a larger population. The characteristics of the people within your sample may vary from those of all the people in the population simply by chance.

For example, take a population of 30 people (18 females and 12 males), or a ratio of 3:2. If you took a sample size of 3 from this population, you would never get a ratio of 3:2, because there aren't enough people. But it is possible to get some samples with 2 women and 1 man. It is also possible that some samples would be all men, or 2 men and 1 woman, or all women. This is what we call random variation or chance. The true ratio is 3:2, but you can get a sample with all men, which does not reflect the total population. It's a random error. The larger the sample size, the smaller the role of random error.

Statistical tests can identify whether the observed exposure-outcome relationship can possibly be due to the role of chance (random error) when the P-value exceeds 0.05 and we will state here that we cannot exclude chance as a possible explanation of the observed relationship. While if the P-value is <0.05, then we can state that chance (random error) is an unlikely explanation of the observed finding and we label that relationship (association or difference) as statistically significant.

Systematic Error (Bias)

This refers to any systematic process at any stage (of design, conduct, analysis, inference, or publication) that tends to produce results or conclusions that differ systematically from the truth. In other words, it is a systematic error that results in an incorrect estimate, e.g., of the effect of exposure on outcome, i.e., an incorrect relative risk (RR) or odds ratio (OR). Systematic errors are reproducible inaccuracies that are consistently in the same direction. Systematic errors primarily influence a measurement's accuracy or the validity of the measurement. Generally, bias occurs through two main sources – the method of selection of study subjects for the study and how information is obtained, reported, or interpreted.

Examples of systematic errors:

- Forgetting to zero a balance produces mass measurements that are always "off" by the same amount. An error caused by not setting an instrument to zero prior to its use is called an offset error.

- Measuring length with a metal ruler will give a different result at a cold temperature than at a hot temperature, due to thermal expansion of the material.
- An improperly calibrated thermometer may give accurate readings within a certain temperature range but become inaccurate at higher or lower temperatures.
- Measured distance is different using a new cloth measuring tape versus an older stretched one. Proportional errors of this type are called scale factor errors.

While random errors can be minimized by increasing sample size and averaging data, systematic errors can be reduced or avoided only when you identify the cause and correct it. Generally, bias cannot be easily rectified once it is introduced into the data, making the prevention of the bias through prediction of the type of bias and the possibility of having it the best approach to minimize its occurrence. Predicting the occurrence of a particular type of bias depends on the type of study, how the subjects were recruited, and how the information was obtained from the study participants. Suppose you read the results of a study, and the investigator reported that the risk ratio was 2.1. In other words, persons with some exposures were 2.1 times more likely to develop illness than persons without the exposure. This finding could be due to chance or due to bias (different types of bias) or they may represent a true association.

Types of systematic error (bias)

1. Selection bias
2. Information bias
3. Confounding

Selection bias is defined as an error due to a systemic difference in the enrollment of participants in a study that results in an incorrect estimate of effect. In other words, selection bias is a problem with who gets into (or how they get into) your study.

Selection bias in cohort studies

How are participants enrolled in a cohort study?

Usually, an entire group is identified (e.g., all who attended the wedding, or all who work at Factory X), then a questionnaire is distributed to find out which study subjects were exposed, and which were not exposed.

Less commonly, a group with known exposure is identified (e.g., workers at a factory exposed to Chemical Y on the production line) and a comparable group without that exposure is enrolled (e.g., clerical workers, loading workers, etc. at the same factory).

Potential bias occurs if some people are misclassified, i.e., loading workers who recently transferred to loading after working inside the factory (and being exposed to Chemical Y) for 15 years. They were exposed but are classified as unexposed because they are part of the loading comparison group. The comparison group will have an artificially increased risk, resulting in a risk ratio biased downward toward the null hypothesis.

Similarly, if all the long-term exposed workers decline to participate, and the workers with briefer exposures have not had enough time to get sick, the rate will be artificially low in the exposed group, again biasing the risk ratio toward the null hypothesis.

Potential selection bias occurs due to misclassification of exposure, or an inappropriate comparison group or differences in participation rates.

Selection bias in case-control studies

Potential bias occurs in case control studies

- If a poor case definition is used, i.e., some cases are not really cases but have a similar but unrelated disease
- If asymptomatic cases are included in the control group
- If the control group does not represent the population from which the cases came
- If an inappropriate comparison group was chosen
- If there are differences in participation rates. If exposed cases are more (or less) likely to participate.

There are several forms of selection bias. Selection bias can occur during the process of selecting and recruiting subjects into the study, or as a result of differences in participation between different groups.

Types of selection bias

- Sampling bias
  - A bias in which a sample is collected in such a way that some members of the intended population are less likely to be included than others.
- Ascertainment bias
  - Surveillance
  - Referral and admission
  - Diagnostic
- Participation bias
  - Self-selection (volunteerism)
  - Non-response and refusal
  - Healthy worker effect, survival
  - Loss to follow-up or drop-out bias

Example of Selection Bias

Alcohol Use	Cases of Cirrhosis	Trauma Cases as Control	Total
Heavy	40	30	70
Light/None	10	20	30
Total	50	50	100

How representative are hospitalized trauma patients of the population from which the cases came? What is the likely effect on the odds ratio?

In a hospital-based case-control study of cirrhosis patients to examine the association with alcohol use, investigators used trauma patients as the control group. The odds ratio for this two-by-two table = 2.7

Regarding alcohol use, how representative of the general population (i.e., catchment area) are hospitalized trauma patients? In other words, hospitalized trauma patients do not provide a reasonable

estimate of alcohol use for the general population. Alcohol use is higher in trauma patients than in the general population. The likely effect on the odds ratio is biased toward the null.

When trauma patients are used as the control group, the odds ratio was 2.7, but when a control group that was more representative of the general population (non-trauma control group) was used, the odds ratio got much higher. It became 16 as in the table below.

Alcohol Use	Cases of Cirrhosis	Controls Non-trauma
Heavy	40	10
Light/none	10	40

Diagnostic bias
This type of bias can happen when knowledge of an exposure is used as a diagnostic criterion. Example: If a pathologist knows a patient is an alcoholic, a borderline liver specimen is more likely to be labeled as alcoholic cirrhosis.

Another way of defining diagnostic bias is when exposed and unexposed patients have an unequal measurement of health outcome statuses. Diagnostic bias is usually categorized as a type of selection bias, although some authors classify it as a sub-type of information bias.

Another example of diagnostic bias is if a group of workers in the industry finds out that one of the chemicals they have been exposed to is a carcinogen, then these workers might present to a medical facility sooner, or be more likely to attend the screening than a non-exposed population.

Admission bias
This happens when exposed cases to a certain risk factor have a different likelihood of admission than non-exposed cases such as when the researcher in the case-control study is interested in studying lung cancer cases with a history of asbestos exposure. These cases are not representative of all lung cancer cases and this results in an overestimation of the rate of exposure among cases to asbestos and, as a result, an overestimation of the odds ratio.

Referral bias
This is a type of selection bias. People who are referred to studies are frequently different from those who are not, meaning that the results of a trial may not generalize well to the general population.
Neyman bias
This refers to the exclusion of individuals with severe or mild disease from the study resulting in a systematic error in the estimated association or effect of an exposure on an outcome.

Survival bias
This happens when only survivors of a highly lethal disease like Ebola haemorrhagic fever enter the study and the death cases are not considered. This will result in decreasing the rate of the disease among exposed group and underestimation of the risk ratio erroneously.

Non-response bias
This is a systematic error due to the differences in response rates of participants in a study and

happens when participation in the study is related to exposure status.
Characteristics:

- It happens when individuals chosen for the sample are unwilling or unable to participate in the survey and when respondents differ meaningfully from non-respondents.
- Non-response is often a problem with mail surveys, where the response rate can be very low.
- It can happen in all types of studies.
- In a case-control study, it is sometimes difficult to identify controls. Some don't respond either because they refuse, because they cannot be contacted, or because their exposure cannot be documented.
- The assumption that controls not included in the study (non-respondents) have the same history of exposure as controls who respond. If this is not true and non-respondents exhibit exposures or outcomes which differ from those of respondents, the exposure among controls may be either overestimated or underestimated, leading to a lower or higher odds ratio.
- In the Nurses' Health Study, out of 172,000 nurses requested to participate in the study; only 122,000 were accepted (71%). In the Framingham Study, out of 6507 residents, only 4469 (69%) were accepted to participate.
- Non-response can affect the generalizability of the study but not the validity unless the non-response is related to the exposure under the study.

Loss to follow-up selection bias

Retention of subjects may be differentially related to exposure and outcome, and this has a similar effect that can bias the results, causing either an overestimate or an underestimate of an association. In a hypothetical cohort study, investigators compared the incidence of thromboembolism (TE) in 10,000 women on oral contraceptives (OC) and 10,000 women not taking OC. TE occurred in 20 subjects taking OC and 10 subjects not taking OC, so the true risk ratio was (20/10,000)/(10/10,000) = 2.

Suppose there were substantial losses to follow-up in both groups, and a greater tendency to lose subjects taking oral contraceptives who developed thromboembolism. In other words, there was a differential loss to follow up with the loss of 12 diseased subjects in the group taking oral contraceptives, but the loss of only two subjects with thromboembolism in the unexposed group. So, in this scenario both exposure groups lost about 40% of their subjects during the follow-up period, but there was a greater loss of diseased subjects in the exposed group than in the unexposed group, and it was this differential loss to follow- up that biased the results. This biased data would give a risk ratio of 1.

If the loss to follow-up ratio is around 30-40%, the study's validity will be affected. This is a problem if the loss to follow-up differs between the group who developed the disease and the other group. To estimate the true relationship, we can calculate the relation using the most extreme situations and have the interval, i.e., assuming that all the lost to-follow-up participants had developed the outcome and then assuming that all the lost to-follow-up did not develop the outcome and identify the range of the risk ratio.

<u>Preventing loss to follow-up</u>

The only way to prevent bias from loss to follow-up is to maintain high follow-up rates (>80%). This can be achieved by:

- Enrolling motivated subjects, as people at higher risk of having the outcome but do not have it yet.
- Using subjects who are easy to track, like people of well-defined residence, employees of a certain factory, alumni members, members of certain societies and syndicates.
- Making questionnaires as easy to complete as possible.
- Maintaining participants' interest and making them feel that the study is important.
- Providing incentives.

Healthy worker effect bias

This is really a special type of selection bias that occurs in cohort studies of occupational exposures when the general population is used as a comparison group. The general population consists of both healthy and unhealthy people. Those who are not healthy are less likely to be employed, while the employed workforce tends to have fewer sick people. Moreover, people with severe illnesses would most likely be excluded from employment, but not from the general population. As a result, comparisons of mortality rates between an employed group and the general population will be biased. The "healthy worker effect" influences the interpretation of findings from occupational health research because it may diminish the validity of the data. It is a form of bias lowering the mortality rate in employed persons compared to the general population.

Many occupations, such as fire-fighters, police, and the military, have to undergo strenuous physical and endurance examinations to assess their physical health before they are hired to work in such professions. Thus, comparisons of mortality rates between an employed/occupational group and the general population will be biased, as not all in the general population will be at "risk" of being employed.

Often, occupational epidemiologists compare the experience of a known group of workers to the vital statistics data from the entire population.

Example: In Table1, suppose that 50 deaths occurred among 500 workers followed over 10 years, contributing to a total of 5,000 person-years of observation, compared with 6500 deaths occurred among 50000 of the general population observed for 10 years, contributing to a total of 500000 person-years of observation, the mortality rate for exposed workers appears to be lower than that of the general population (10/1000 vis. 13/1000), so the exposure appears protective.

Table 1

	Exposed Workers	**General Population**
Deaths	50	6500
Person-years Observation	5000	500000
Mortality	10/1000	13/1000

When we split workers and non-workers in the general population in Table 2 below, it appears that workers in the general population have the same mortality rate as that for exposed workers, indicating that exposure is neither harmful nor protective. The non-workers, at least some of whom are not working because they have disabilities or illness or other conditions that prevent them from working and increase their risk of death, do indeed have a much higher mortality rate. These people contributed

to the overall population mortality rate in Table1, making the exposed workers look healthier.

Table 2

		General Population	
	Exposed Workers	**Workers**	**Not Workers**
Deaths	**50**	**4500**	**2000**
Person-time (Years)	**5000**	**450000**	**50000**
Mortality (Cases Per Year)	**10/1,000**	**10/1,000**	**40/1,000**

HWE is a phenomenon initially observed in studies of occupational diseases: Workers usually exhibit lower overall death rates than the general population because the severely ill and chronically disabled are ordinarily excluded from employment. Most studies indicate that HWE will reduce the association between exposure and outcome by an average of 20-30%.

Berkson's bias

This is a form of selection bias that causes hospital cases and controls in a case-control study to be systematically different from one another because the combination of exposure to risk and disease occurrence increases the likelihood of being admitted to the hospital. This produces a systematically higher exposure rate among hospital patients distorting the odds ratio. Berkson's bias is a type of selection bias described by the American statistician Joseph Berkson. [38]

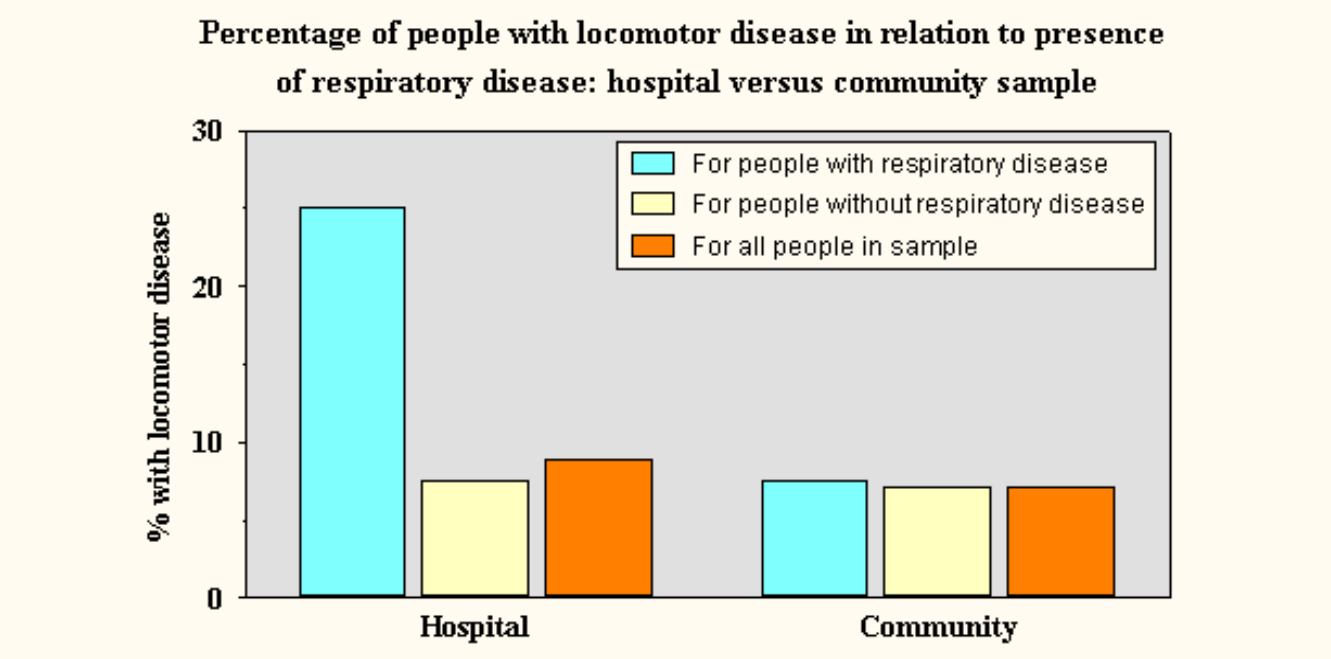
Berkson's bias can arise when the sample is taken not from the general population, but from a subpopulation. It was first recognized in case-control studies when both cases and controls are sampled from a hospital rather than from the community.

When we take the sample, we have to assume that the chance of admission to hospital for the disease is not affected by the presence or absence of the risk factor for that disease. This may not be the case, especially if the risk factor is another disease. This is because people are more likely to be hospitalized if they have two diseases, rather than only one.

The best-known example of this is given by Sackett (1979). [39] He took a random sample of 2784 people from the community and determined the presence or absence of respiratory disease and locomotor disease. He then looked at the same thing for those people within the sample who had been hospitalized in the previous six months. The results are shown in the chart below.

If we only looked at the hospital sample, we would conclude that people with the respiratory disease are much more likely to suffer from locomotor disease. In other words, there is an association between the two complaints. Moreover, any analysis of risk factors will (wrongly) suggest that the risk factors for locomotor disease are also risk factors for respiratory disease.

If, however, we look at the full community sample, we would conclude that having respiratory disease has no effect on whether or not one is likely to suffer from locomotor disease. The latter is



of course the correct conclusion. The incorrect conclusion from the hospital sample arises because people who have both diseases are more likely to be hospitalized than people who only have one.

Ways to minimize selection bias

- Select an adequate and representative sample.
- Choose groups comparably.
  In cohort studies:
  - Exposed and unexposed should have the same opportunity to become cases.
  - Have equal follow-up of exposed and not exposed.
  - Take measures to minimize losses (drop-out, attrition) from the sample.

  In case-control studies:
  - Controls should represent the case source population.
  - Choose cases and controls without regard to exposure status.
  - Use a clean case definition.
  - Use a clean control definition.

Information bias

Information bias is an error due to systematic differences in the collection of exposure or disease data that results in an incorrect estimate of effect. In other words, information bias represents a problem with the information you get from or about the people who are already in your study.

Information bias in cohort studies:

- Potential information bias occurs if some people are lost to follow-up (particularly if different for exposed vs. unexposed people).
- Drop-out, or refuse to participate after some time.
- Potential information bias occurs if the diagnoses are not made or are more likely to be made among exposed persons than among non-exposed persons.
- Potential information bias occurs if there is a poor case definition, so some people with the disease are classified as non-cases, and/or some people without the disease of interest (usually have a similar but unrelated illness) are included.
- Bias due to misclassification, i.e., loading worker who recently transferred to loading after

working inside the factory (and being exposed to Chemical Y) for 15 years. He is exposed but is classified as unexposed because he is part of the loading comparison group. The comparison group will have an artificially increased risk, resulting in RR biased downward toward the null value. Similarly, if all the long-term exposed workers decline to participate, and the workers with briefer exposures have not had enough time to get sick, the rate will be artificially low in exposed group, again biasing the RR toward the null value.

Information bias in case-control study

Potential information bias occurs if:

- Poor documentation of exposures
- Differential rigor of interviewers asking about exposures among cases and controls
- Poor recall of exposure data by cases or controls results in recall bias
- Poor definition of exposure variable/s

Types of information bias

- Questionnaire faults: When the questionnaires are developed in a way that makes them unable to collect the intended data from the respondents.
- Observer error or interviewer bias: When the interviewer asks questions differently based on knowledge of outcome status. Example: When the interviewer asks cases more aggressively about smoking than controls in a case-control study of coronary heart disease, for example. This may result in an over estimation of the odds ratio.
- Recall bias: It happens when case-patients remember exposures differently (usually better) than controls. Example: The mother of a child with a birth defect is more likely to remember even minor exposures in the past than the mother of a normal child. This leads to an over estimation of the odds ratio.
- Hawthorne effect bias: This bias occurs when people under the study behave differently because they know they are being watched. So, they try to please the observer or the interviewer. In other words, the Hawthorne effect refers to a type of reactivity in which individuals modify an aspect of their behavior in response to their awareness of being observed.

Ways to prevent information bias

- Use a standard, pretested questionnaire.
- Train interviewers well.
- Be diligent in tracking down study participants.
- Use memory recall aids.
- Verify exposure using other data sources.

What about misclassification?

Misclassification is an erroneous classification of an individual, value, or attribute into a category other than that to which it should be assigned.

- Common misclassifications in field epidemiology
  - Asymptomatic case enrolled as control, or not recognized as case (in cohort study)
  - Unrecognized exposure classified as not exposed
- Less common misclassifications in field epidemiology
  - Non-case misclassified as a case

- "Not exposed" misclassified as "exposed"

So, depending on the type of study and whether the exposure or outcome is misclassified, misclassification can be either selection bias or information bias.
In the cohort study, misclassification of exposure (how they get in the study) = selection bias.
In the cohort study, misclassification of outcome (information you get once they are in the study) = information bias.
In the case-control study, the misclassification of exposure (information you get once they are in the study) = information bias.
In the case-control study, misclassification of outcome (how they get in the study) = selection bias.

Confounding

The systematic distortion in the measure of association between exposure and health outcomes is caused by mixing the effects of exposure of primary interest with extraneous risk factors. Or it is the distortion of the estimated effect of an exposure on an outcome caused by the presence of a factor associated with both exposure and outcome. Apart from exposure and outcome, the confounder is a third factor that is associated with exposure, and independent of this association, it is a risk factor for the outcome. In other words, a confounding variable (confounder) is a factor other than the one

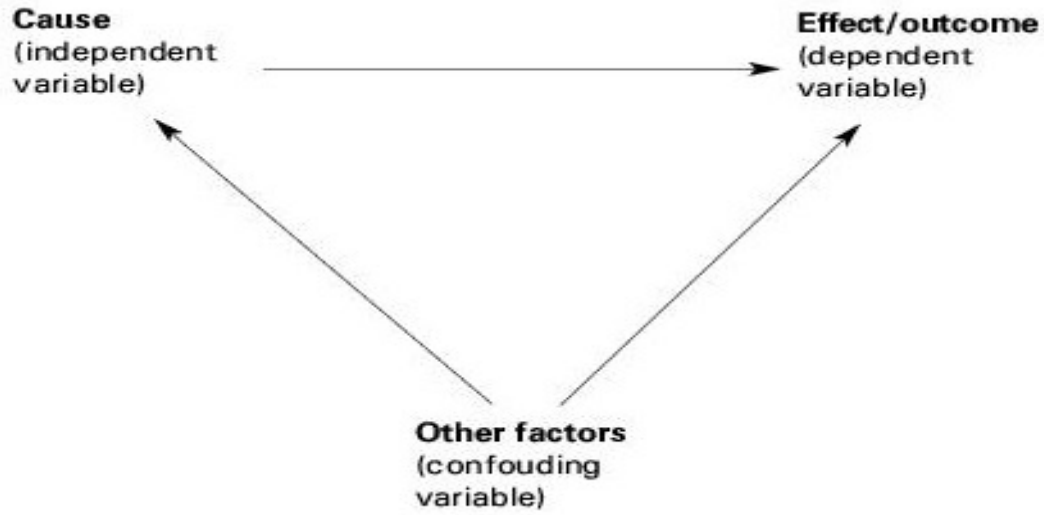


Diagram of Confounding

being studied that is associated both with the disease (dependent variable) and with the factor being studied (independent variable). A confounding variable may distort or mask the effects of another variable on the disease in question. [40].

The French translation of "confounding" is "confusion." Confounders can lead to:

- Over estimation of the true association
- Under estimation of the true association
- Can even change the direction of the observed effect

Positive confounding is when the observed association is biased away from the null hypothesis. In other words, it overestimates the effect.

Negative confounding is when the observed association is biased toward the null hypothesis. In other words, it underestimates the effect.

- Confounders are not causal of the outcome but correlates to another causal factor.
- Confounders must predict the disease independent of its association with the exposure.
- Confounders cannot be an intermediate link in the causal chain between exposure and outcome.

Example of an apparent association between owning a radio and getting malaria [41]:

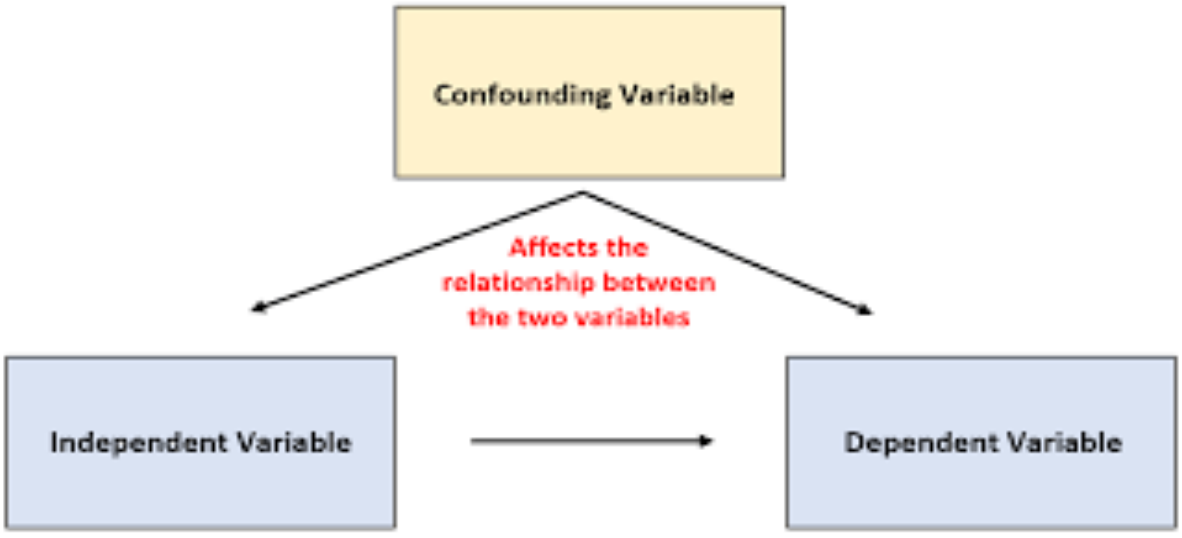
	Malaria	**No Malaria**	**Total**	**Incidence**
Owning a Radio	80	440	520	15.4%
No Radio	220	860	1080	20.4%
Total	300	1300	1600	

Risk Ratio = 0.76, 95% CI = 0.6- 0.9; $p < 0.02$

In this study, investigators found that people who owned radios were at lower risk of getting malaria than those who did not – 15% versus 20%. The risk ratio was 0.76, the confidence interval was quite narrow, and the p-value was < 0.02.

Can the reduced incidence of malaria be attributed to the radio itself? If not, what is a possible alternative explanation?

In this study, there was another hidden factor that reduced the risk of malaria among families having radios that is the use of mosquito bed nets, since families that own radios are usually wealthier and can afford screens on windows or mosquito bed nets. We saw an apparent protective effect of radios on the occurrence of malaria. This is the crude association from the 2-by-2 table. But we speculated that perhaps the real protective effect against malaria is mosquito bed nets and that perhaps people who own radios are the same people who can afford bed nets. In other words, we know that bed nets protect against malaria and that bed nets may be associated with radios. So, in this study the bed net was the confounding factor that made the apparent association between radio and malaria. The confounding factor is both related to the outcome (using mosquito bed nets protects from malaria even in families who have no radios), and owning bed nets is related to the presence of radios, since people who own radios can afford to own bed nets.

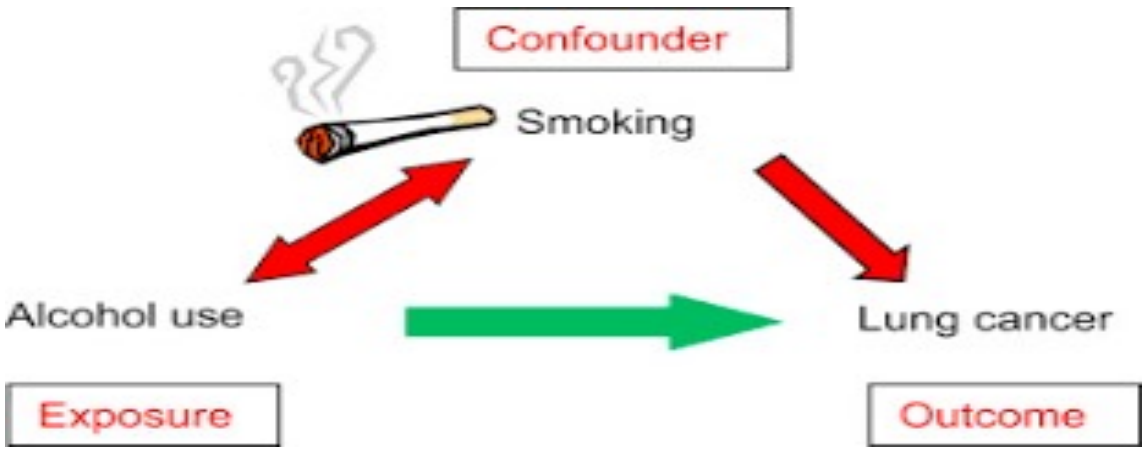


To be a confounder, a variable must be associated with the outcome, independent of the exposure (that is, even in the unexposed group) and associated with the exposure.

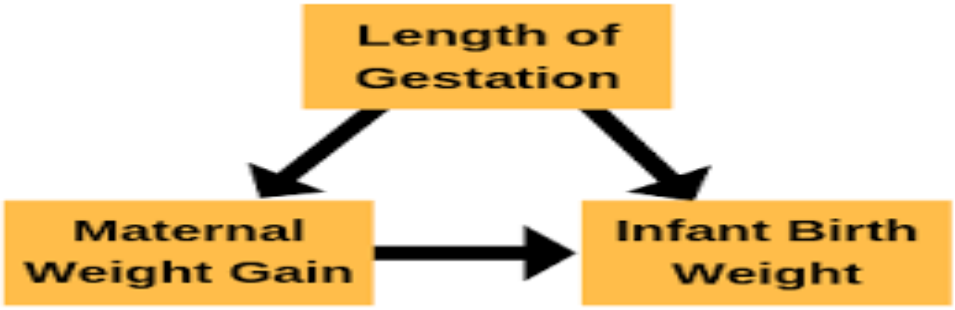
Other examples of confounders

1. Alcohol use and lung cancer

   The association between alcohol use and lung cancer is distorted because of the presence of a third variable (confounder), which is smoking. Smoking is related to the outcome, as smoking is a known risk factor for lung cancer, and it is also related to alcohol use, since alcohol drinkers are usually smokers. The apparent association between alcohol use and lung cancer becomes unrealistic when the effect of smoking is removed.



2. Maternal weight gain and infant birth weight, where the gestational length is the confounder

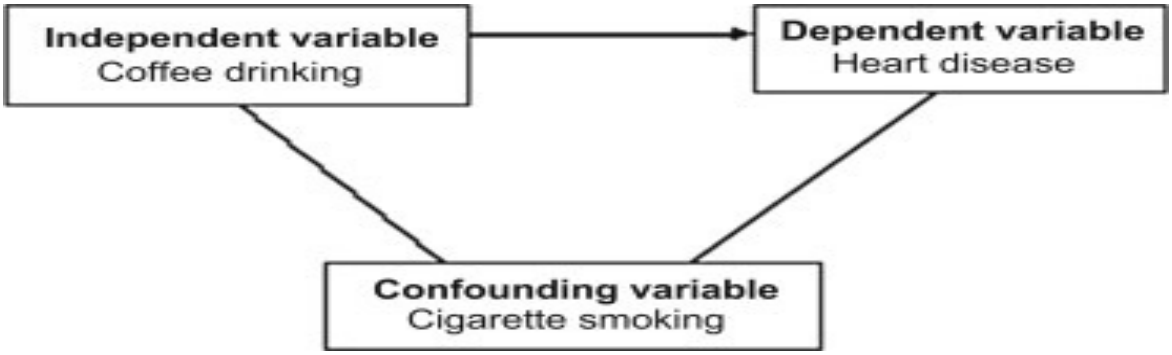


3. Coffee drinkers and heart disease, where cigarette smoking is the confounder

   The hypothesis that coffee drinkers have more heart disease than non-coffee drinkers may be influenced by another factor. Coffee drinkers may smoke more cigarettes than non-coffee drinkers, so smoking is a confounding variable in the study of the association between coffee drinking and heart disease. The increase in heart disease may be due to smoking and not coffee. More recent studies have shown coffee drinking to have substantial benefits in heart health and in the prevention of dementia.

The selection of a particular confounder will depend on:

- Knowledge of the disease



- Previous evaluation of the variable
- Investigator's judgment

Statistical significance is not reliable in deciding on confounders.
In the evaluation of the role of a confounder, it is necessary to decide:

- Its presence or absence
- The magnitude of its effect
- Its direction

Control of confounders

Several methods are available to control confounding, either through study design or during the analysis of results. There are several methods used to control the confounding in epidemiological studies. [42]

At the design stage, confounding can be controlled by:

- Randomization
- Restriction
- Matching

At the analysis stage, confounding can be controlled by:

- Stratification
- Statistical modeling

Randomization in experimental studies is the ideal method for ensuring that potential confounding variables are equally distributed among the groups being compared. The sample sizes have to be sufficiently large to avoid random maldistribution of such variables. Randomization avoids the association between potentially confounding variables and the exposure that is being considered.

Restriction is another way to control confounding at the design stage, where we limit the study to people who have particular characteristics. For example, in a study on the effects of coffee on coronary heart disease, participation in the study could be restricted to nonsmokers, thus removing any potential effect of confounding by cigarette smoking.

Matching is used to control confounding by selecting study participants so as to ensure that potential confounding variables are evenly distributed in the two groups being compared. For example, in a case-control study of exercise and coronary heart disease, each patient with heart disease can be matched with a control of the same age group and sex to ensure that confounding by age and sex does not occur. Matching has been used extensively in case-control studies but it can lead to problems in the selection of controls if the matching criteria are too strict or too numerous; this is

called overmatching. Matching can be expensive and time-consuming, but is particularly useful if the danger exists of there being no overlap between cases and controls, such as in a situation where the cases are likely to be older than the controls.

Stratification and statistical modelling

In large studies, it is usually preferable to control for confounding in the analytical phase rather than in the design phase [43]. Confounding can then be controlled by stratification, which involves the measurement of the strength of associations in well-defined and homogeneous categories (strata) of the confounding variable. If age is a confounder, the association may be measured in, say, 10-year age groups; if sex or ethnicity is a confounder, the association is measured separately in men and women or in different ethnic groups. Methods are available for summarizing the overall association by producing a weighted average of the estimates calculated in each separate stratum. Although stratification is conceptually simple and relatively easy to carry out, it is often limited by the size of the study and it cannot help to control many factors simultaneously, as is often necessary. In this situation, multivariate statistical modelling is required to estimate the strength of the associations while controlling several confounding variables simultaneously.

Effect modification

The effect of exposure on the outcome differs depending on the level of another variable called the effect modifier. Effect modification occurs when the magnitude or direction of association varies according to the presence of a third factor.

When the measure of effect or measure of association changes over the value of some other factor we say that effect is modified or in other words we have effect modification. You may also see the term interaction used to describe effect modification.

Examples of effect modification:

- Drug treatment's effect on a particular health outcome may be stronger in males than females, in which case, gender is acting as an effect modifier.
- Driving and alcohol consumption as risk factors for injury, and alcohol is likely to increase the impact of driving on the risk of injury. Alcohol is said to modify the association between driving and injury, alcohol is an effect modifier.
- Tetracycline and tooth mottling in children below 8 years. Age is the effect modifier.
- Breast feeding and diarrhea. Lack of breastfeeding is a huge problem in infants younger than 1 month, less so for infants 1 month or older. So, age is an effect modifier.
- An association that is stronger in older people than in younger people; age is an effect modifier.
- Another definition of effect modification is when the incidence rate of disease in the presence of two (or more) risk factors differs from the incidence rate expected to result from their individual effects. For example, the effects are synergistic.

In one study, investigators studied the impact of smoking on deaths from lung cancer among individuals with and without asbestos exposure. We see that the risk of death from lung cancer is increased with cigarette smoking among those with no asbestos exposure but is even more dramatically increased with cigarette smoking among those with asbestos exposure [44].

Deaths from Lung Cancer per 100, 000 among Individuals With and Without Exposure to Cigarette Smoking and Asbestos

Cigarette Smoking	Asbestos Exposure	
	No	Yes
No	11.3	58.4
Yes	122.6	601.6

Effect modification should not be confused with confounding. Effect modification doesn't distort the data like confounding does.

Steps followed during data analysis for effect modifiers:

- Conduct crude analysis (simple 2-by-2 table).
- Stratify data by the third variable.
- Calculate a measure of association for each stratum.
- Determine whether the association is consistent across strata.
- If the association is not consistent across strata, then effect modification is present, so STOP!

Example of stratification of data:

Breastfeeding and Diarrhea Among Infants Less Than One Month of Age

	Cases (diarrhea)	**Controls (No diarrhea)**
No breastfeeding	110	133
Breastfeeding	43	136
Odds ratio= 2.6, p<0.0001		

When we did the stratification by the age of infant, the odds ratio decreased markedly from 32.4 for infants below one month to 2.6 for infants one month or more. So, the age of the infant was an effect modifier.

Investigator Error

Investigator error refers to data mishandling, misanalysis, misinterpretation, or other errors made by the investigator, either knowingly or unknowingly, that results in a mistaken estimate of an exposure's effect on the risk of disease. In other words, the investigator can make a mistake at any step and get spurious results.

Finally, bias is not limited to comparative studies. It can happen in surveillance, like differential participation in reporting of notifiable diseases to the Ministry of Health by private clinics compared to public clinics or reporting of only severe and fatal diseases. This also happens in surveys (e.g., immunization coverage), whether we use convenient samples instead of probability samples, or when there are high refusal rates among the study population in the survey.

Assessing Bias in a Study

When listening to a presentation or reading an article in which data is presented to support a conclusion, one must always consider alternative explanations that may threaten the validity of the conclusions. [45], specifically, one needs to consider whether random error, bias, or confounding could have undermined the conclusions to a significant extent. Virtually all studies have potential

flaws, but carefully done studies are designed and conducted in a way that minimizes these problems so that they don't have any important effect on the conclusions. In view of this, it is always important to ask oneself:

- Given the conditions of the study, could bias have occurred?
- Is it likely that bias was present?
- If there was bias, would it bias the results toward the null hypothesis or away from the null hypothesis?
- Is the magnitude of distortion likely to be small or large?

Chapter 5

Association and Causality

Association is an identifiable relationship (statistical relationship) between two variables. A measure of association quantifies the relationship between exposure and disease among the two groups. Examples of measures of association include risk ratio (relative risk), rate ratio, and odds ratio. Statistical methods cannot establish proof of a causal relationship [46].
Two variables may be associated without a causal relationship, while causation means that there is a true causal relationship between a risk factor and disease. A cause leads to disease (the exposure produces the effect). An association alone does not make the relationship between risk factor and disease causal. The stronger the association, or magnitude of the risk, between a risk factor and outcome, the more likely the relationship is thought to be causal.

What Is the Difference Between Association and Correlation?
Association refers to the general relationship between two random variables while the correlation refers to a more or less a linear relationship between the random variables [47].
If there is a difference in disease occurrence between groups, the first question to be asked is how big is this difference and whether this difference is statistically significant. If this difference is significant, then a statistical association is said to exist between the factor and the disease. This statistical association could be:

1. Artefactual (spurious)
2. Indirect (non-causal)
3. Causal

Artefactual (spurious) association
A spurious association (or spuriousness) refers to a connection between two variables that appears to be causal but is not. With spurious association, any observed dependencies between variables are merely due to chance or are both related to some unseen confounder. One should attempt to confirm this association by replication. If such an association doesn't hold up in this replication, then it may be considered spurious. Thus, whenever a statistically significant association is found, it must be examined carefully to be sure that it is not attributable to some artifact or bias. Confirming a causal relationship requires a study that controls for all possible variables.

Indirect (non-causal) association
This is an association between a factor and a disease due to the presence of another factor i.e., common factor (confounding variable). It refers to a relationship between two variables that is not directly influenced by each other. In other words, there may be a correlation or connection between the variables, but one variable does not cause or directly impact changes in the other.
It occurs when a factor and disease are associated only because both are related to some common underlying condition.

Alteration in the indirectly associated factor will not produce alteration in the frequency of the disease unless the change affects the common underlying condition as well.

The difference between direct and indirect association is that causal effects go directly from one variable to another in the direct association (where one variable causes a change in the other and there are no intervening variables), while indirect association occurs when the relationship between two variables is mediated by one or more variables.

An example of indirect association is altitude and cholera.

Previously, there was a belief that cholera occurs more in areas of low altitudes than in high altitude areas because of fetid air in the low altitudes. Later, it was proved that this association was an indirect association because of the presence of a common underlying condition which is the water impurity in the low altitude areas. Fetid air at low altitudes is the indirect factor, while water impurity in low altitude areas is the risk factor (common underlying condition) for cholera and not the fetid air. So, low altitude is indirectly associated with cholera through the common underlying condition, which is water pollution.

Causal association

Association between two variables where a change in one makes a change in the other one happens. While all causal relationships are associational, not all associational relationships are causal, that is, correlation does not equal causation [48]. A causal relation between two events exists if the occurrence of the first causes the other. The first event is called the cause and the second event is called the effect. The correlation between the two variables does not imply causation.

Bradford Hill criteria

These criteria are a group of principles that can be useful in establishing epidemiologic evidence of a causal relationship between a presumed cause and an observed effect and have been widely used in public health research. They were established in 1965 by the English epidemiologist Sir Austin Bradford Hill [49].

4. **Strength of the association (effect size):** A weak association does not mean that there is no causal effect, though the larger the association, the more likely that it is causal. Strength of the association is measured by the relative risk (risk ratio) or the odds ratio. The larger this ratio, the greater the likelihood of this association to be causal.
5. **Consistency (reproducibility):** Consistent findings observed by different persons in different places with different samples strengthens the likelihood of an effect. An association discovered in one study persists in other studies conducted by using other methods and on different populations. Associations are observed repeatedly in different populations, different places and times, in different types of studies, conducted by different investigators.
6. **Specificity**: Causation is likely if there is a very specific population at a specific site and disease with no other likely explanation. The more specific an association between a factor and an effect is, the bigger the probability of a causal relationship. Specificity of the association becomes ideal when one manifestation follows from only one cause. E.g., Angiosarcoma of the liver and exposure to vinyl chloride, adenocarcinoma of the vagina in female offspring resulting from diethyl stilbestrol ingestion by mothers during pregnancy. Although specificity is strong evidence for causality yet lack of it is of less significance.

7. **Temporality**: The effect has to occur after the cause (and if there is an expected delay between the cause and expected effect, then the effect must occur after that delay). In other words, exposure to the factor should precede the occurrence of disease and allows for the necessary period of induction or latency.
8. **Biological gradient (dose-response relationship):** With increasing levels of exposure to the factor, there is a corresponding rise in the disease. Greater exposure should generally lead to greater incidence of the effect. However, in some cases, the mere presence of the factor can trigger the effect. In other cases, an inverse proportion is observed: greater exposure leads to lower incidence.
9. **Coherence with existing information (biological plausibility):** Coherence between epidemiological and laboratory findings increases the likelihood of an effect. However, Hill noted, "... the lack of such [laboratory] evidence cannot nullify the epidemiological effect on associations." A plausible mechanism between cause and effect is helpful (but Hill noted that knowledge of the mechanism is limited by current knowledge). Plausible association means that the association should be in line with substantive knowledge.
   Example:
   - Plausible: Higher incidence of disease in individuals who are more sexually active. The disease could be a sexually transmitted disease.
   - Not plausible: Polio vaccine reduces fertility, conflicts with current knowledge of polio vaccine production, biological mechanisms for fertility and contraception.
10. **Experiment**: "Occasionally it is possible to appeal to experimental evidence," since it gives the strongest evidence.
11. **Analogy**: The use of analogies or similarities between the observed association and any other associations.

Reverse Causality

Reverse causality (also called reverse causation), refers to a direction of cause-and-effect contrary to a common presumption.

Example 1: When lifelong smokers are told they have lung cancer or emphysema, many may then quit smoking. This change of behavior after the disease develops can make it seem as if ex-smokers are more likely to die of emphysema or lung cancer than current smokers.

Example 2: Reverse causation can occur when people change their diet or other lifestyle habits after developing a disease like a heart attack. It seems as if they are more likely to die of a heart attack than those who did not change their diet or other lifestyle habit.

Chapter 6

Design of Epidemiological Research

This design consists of nine steps. The first five steps are related to the development of the research proposal, while the other four are related to the execution and reporting of the research [50, 51].
A Research Proposal Consists of the Following Steps:

1. Research identification
2. Selection and formulation of a research problem (Step II)
3. Appraisal of existing information (literature review) and reconsideration of the problem and resources (Step III)
4. Formulation and statement of research hypotheses (Step IV)
5. Research plans for testing hypotheses (Step V)

Step I: Research identification

1. Topic
2. Principal investigator (name, degree, position, address)
3. Institutional affiliations
4. Funds required for the first year of research
5. Total funding required for the research
6. Research period

These should be on the cover page of the research proposal.

Step II: Selection and formulation of research problem

The second step in the development of a research proposal is to state the research problem in precise, clear terms. It is an integral part of selecting a research topic, and the essential basis for the construction of a research proposal (research objectives and hypotheses, methodology, work plan, and budget, etc.

The investigator should describe the problem systematically, reflect on its importance and its priority in the country and in the local area, and point out why the proposed research on the problem should be undertaken. This facilitates a peer review of the research proposal by funding agencies. How should the statement of the problem be written for a research proposal? The writing should be precise and concise but should include essential points. The information about the problem should be summarized so that the reader is not "drowned" in detail.
This step requires the investigator to:

- Choose an appropriate research topic.
- Define the nature, extent and significance of the problem.
- Frame specific research questions and the possible value of seeking answers to these questions.
- State research objectives – immediate and ultimate.
- Provide a workable definition of key terms.

Research objectives are the goal to be achieved by a research project. Differentiation between "general" and "specific" objectives may eliminate unnecessary confusion. The general objective of research is what is to be accomplished by the research project and why. Example: to determine whether or not a new vaccine should be incorporated into public health programs.

The specific objectives are, in detail, the specific aims of the research project, often breaking down what is to be accomplished into smaller logical components. In other words, specific objectives relate to the specific research questions the investigator wants to answer through a proposed study.

Example: in evaluating a new vaccine, to determine the degree of protection that is attributable to the vaccine in a study population by comparing the vaccinated and unvaccinated groups.

Step III: Appraisal of existing information (literature review) and reconsideration of the problem and resources

The third step is for the investigator to familiarize themselves with the existing knowledge about the research problem and to discover whether or not others have investigated the same or similar problems before. The investigator will be able to compare their results with those of others, to learn from their experience, their approaches and any new technique they have developed. This step is accomplished by a thorough and critical review of literature and personal communications with experts. The investigator should then decide whether or not they need statistical, methodological and technical consultations. It is at this early stage in research design, not after collection of data, that such statistical or methodological collaboration should be sought and obtained. The investigator should also reconsider their research interests and the resources available.

Summary of Step III:

- Literature review and search for other sources on the subject of the research
- Classification of existing pertinent information on the subject
- Critical appraisal of existing information
- Consideration of statistical, methodological, and technical collaboration
- Reconsideration of research needs and interests
- Consideration of available resources such as funds and personnel
- Selection of specific research questions to be answered

The source of information may include the following:

- Card catalogues of books in libraries
- Indices, such as the Index Medicus and the International Nursing Index, which identify journal articles by subject, author, and title
- Computer-based literature searches such as MEDLINE, MEDLARS and CATLINES
- Bibliographies, such as those found at the end of books, articles, and theses, or prepared as separate documents
- Statistics collected at national, provincial and/or departmental levels
- Responses to enquiries about ongoing research

Step IV: Formulation and statement of research hypotheses

The value of scientific work depends heavily on the originality and logic with which hypotheses are formulated. A hypothesis is a shrewd supposition or inference adopted to explain observations and/or to guide further investigation. It may be simply an educated answer to a specific question. The hypotheses may be derived from the body of knowledge on the subject, from the experience of the investigator or of others, or from previous research endeavors. It must be emphasized that hypotheses are not meant to be haphazard guesses but should reflect the investigator's depth of knowledge, imagination, and experience.

Step V: Research plans for testing hypotheses

The plans of testing hypotheses that have been formulated require selection of research strategy and setting, definition of the unit of observation, the methods and particulars of observation, and an outline of the sampling procedures and the choice of controls. Plan for analysis of data to be collected should also be made. Administrative details are also to be specified.

- Selection of research strategy
- Selection of research setting
- Sampling
- Definition of the unit of observation
- Controls and case allocation
- Plan for testing of equality between sample and control groups
- Plan for minimizing non-sampling errors (such as observer error, use of volunteers, paid participants); and errors of coverage, recording, or data processing
- Development of study instruments

Selection of research strategy

This is the core of research design and is probably the single most important decision the investigator has to make. The choice of strategy whether descriptive, analytical, experimental, operational, qualitative research, or a combination of these, depends on several considerations.

The specific types of studies are as follows:

1. Descriptive strategies (observational hypothesis generation rather than testing) include:
    - Descriptive cross-sectional study or population survey, e.g., malaria survey, opinion survey, knowledge-attitude-practice (KAP) survey
    - Epidemiological description of disease occurrence by person, place, and time
    - Studies of changing patterns of health and disease over time and space: the epidemiological transition
    - Community diagnosis of a health problem or assessment of needs
    - Studies of existing data: case-series, disease registries, surveillance reports
    - Studies of the natural history of disease
2. Observational analytical strategies (hypothesis testing):
    - Prospective study (cohort study)
    - Historical (or reconstructed or retrospective) cohort study, when adequate historical data or

records are available
- Retrospective study (case-control study)
- Analytical cross-sectional study
- Follow-up study (longitudinal study; repeated cross-sectional study)

3. Experimental strategies:
   - Animal studies
   - Therapeutic clinical trials
   - Prophylactic clinical trials
   - Field trials
   - Quasi-experimental studies (intervention studies, health systems research).
4. Operational strategies (observation, time-motion study).

Selection of research setting

The research setting includes all the pertinent facets of the study such as the population to be studied, the place and time of the study, the type of observation, and the collaborating institutions, if any. The investigator is entitled to choose a convenient setting, but they should be careful not to sacrifice appropriateness of methodology to convenience.

The research setting includes:
- Selection of the study population, place, and time
- Definition of the unit of observations, variables, and data to be collected
- Choice of methods and particulars of observation
- Consideration of ethical problems

Sampling

A sample is a part of the whole population (universe or reference population). Sampling is the process or technique of selecting a sample of appropriate and manageable size for study. In epidemiological investigations and in health work in general, it is almost always possible to deal with a sample drawn from reference population or universe. This universe may be a population of people, a population of cases of certain disease, the clients of a family planning clinic or recipients of certain treatment. The universe may not be people at all, as a universe of birth or death certificates or a universe of medical records, or the universe may consist of health centers, village units or hospital units.

Selection of sampling procedure:
- Probability sampling methods: Random, systematic, and stratified sampling; use of panels, area sampling, cluster sampling, others
- Non-probability sampling: Quota, and convenience sampling, others

Sample size: The sample should be of sufficient size to be dependable and to allow for statistically significant testing to be applied. Statistical methods are used in estimating the sample size.

Plan to ensure representativeness and reliability of the sample and to minimize sampling errors (such as non-response and selection).

Control and case allocation through:

- Matching
- Random allocation
- Alternation
- Population as control
- Before and after control
- Analysis of subgroups

Plans for testing of equality between sample and control groups

Many gross errors have been made in attempting to equate groups and make generalizations based on comparisons between groups that, in reality, are very different. Therefore, plans must be made for testing equality between experimental (or sample) and control groups.

Plans for minimizing non-sampling errors

Such as observer error, use of volunteers, paid participants, errors of coverage, recording, or data processing.

Development of study instruments

1. Questionnaires and interviews:
    - Preparation, precoding, and pretesting of questionnaires
    - Preparation of instruction manuals
    - Translation of questionnaires in cross-cultural studies
    - Developing editing and coding plans
    - Plan for interviews
    - Training of interviewers
2. Other methods of observation:
    - Medical examination
    - Laboratory tests
    - Screening procedures
3. Design of recording forms (which may be the same as analysis forms)
4. Plans for reliability and validity checks

Plans for data analysis

The investigator is advised to incorporate the plan of data analysis in their proposal to increase the confidence of reviewers in the competence of the research team in handling the data once they are collected. This plan includes:

- Design of analysis forms
- Selection of data handling techniques
- Choice of statistical methods to be used for each hypothesis
- Design of dummy tables and graphs (skeleton of anticipated tables and graphs)

Plans for data collection include:

- Organization of data study and collection

- Plan for personnel training if applicable
- Plan for pilot, pretesting, and feasibility studies

Budgeting includes:

- Personnel (salaries, fringe, benefits)
- Consultant fees
- Equipment and supplies
- Travel – domestic and international
- Cost of data analysis
- Miscellaneous expenditure
- Overheads

Timetable for data collection and analysis

Other administrative details for the proposed research

- Description of facilities available to the investigator (computers, office space, etc.)
- Other sources of research support
- Biographical sketches of principal investigator and co-investigators with emphasis on previous experience in fields related to the research
- Permission to investigate human subjects, if applicable

Step VI: Data collection

After a research proposal has been accepted, the investigator should proceed to collect the data following the details of the proposal. Special emphasis should be given to the following items:

- Personnel recruitment and training
- Pilot and feasibility studies and pretesting of instruments
- Conduct the study according to the selected design
- Record keeping and editing
- Problems related to data collection like sampling and non-sampling errors

Step VII: Processing, classification, and analysis of data

Includes:

- Data processing according to the plan
- Tabulation and graphing
- Analysis according to the plan
- Preparation of progress reports as required

Step VIII: Interpretation and conclusions

Careful, unbiased and critical interpretation of data requires great skill and experience on the part of the investigator. Such skill can be developed by appraising well designed epidemiological studies, by learning from experience of others and by participating in scientific conferences. This step should include:

- Accurate and unbiased evaluation of results
- Determination of causation versus association

- Determination of population to which results may be referred
- Drawing conclusions and explaining practical applications of findings and outlining future research needs

Step IX: Reporting

The general outline of the report should contain the following:

1. Title
2. Author(s)
3. Introduction which consists of background, objectives, and hypotheses
4. Material and methods
5. Results
6. Discussion
7. Summary
8. References

Chapter 7

Screening

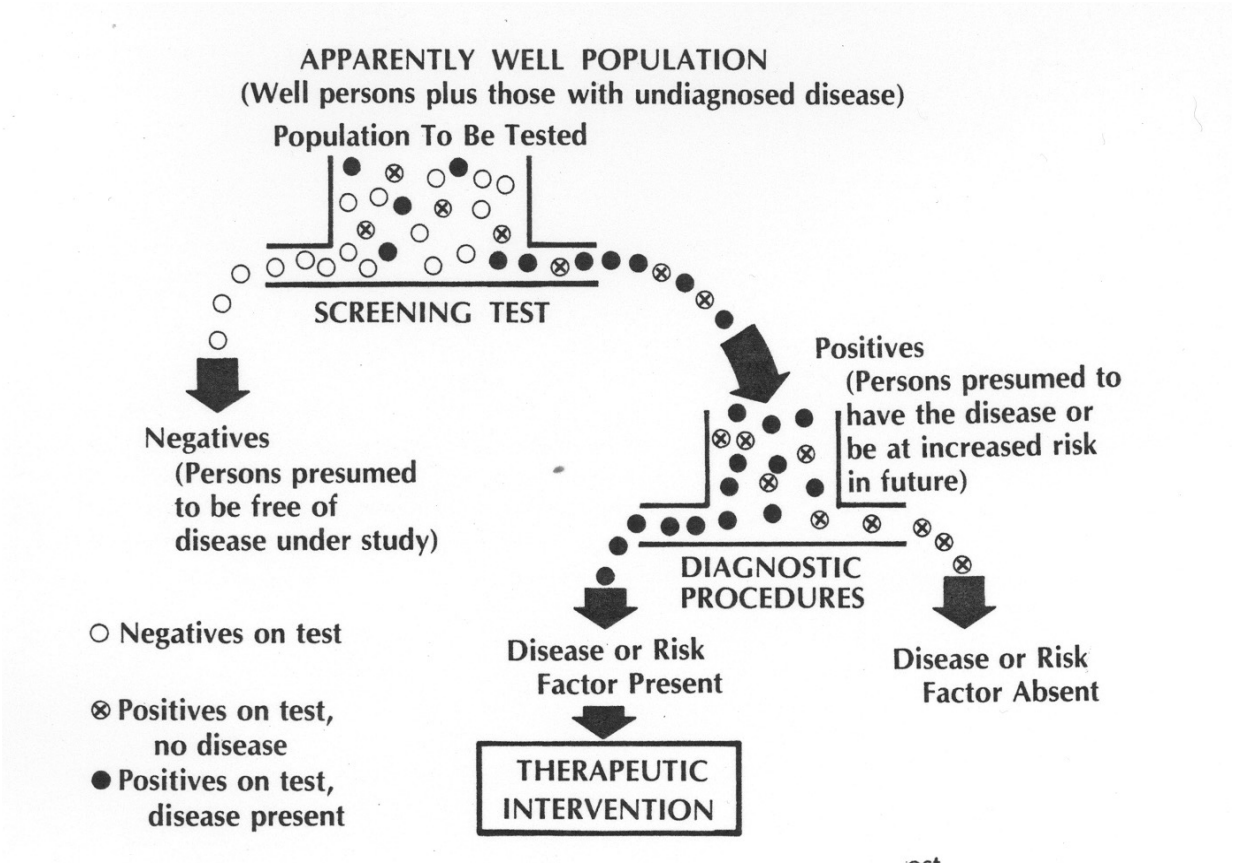
Definition

Screening is the presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures which can be applied rapidly to sort out apparently well persons who probably have a disease from those who probably don't.

Screening test is not intended to be diagnostic [52]. Persons with positive or suspicious results must be referred to their physicians for diagnosis and necessary treatment.

Screening involves testing of apparently healthy populations to identify previously undiagnosed diseases or people at high risk of developing a disease. It aims to detect early disease before it becomes symptomatic. Screening is an important aspect of prevention, but not all diseases are suitable for screening [53].

The goal of screening is to identify disease or susceptibility when it is more easily and more successfully treated, i.e., to reduce morbidity and mortality. The objective is to classify individuals as likely or unlikely to have the disease or susceptibility, those classified as likely are referred for diagnostic testing.



A Flow Diagram Clarifies the Process of Mass Screening Test [54]

Screening versus Diagnosis

Screening = testing among asymptomatic persons (although can be targeted at persons thought to be at increased risk, i.e., with family history of colon or breast cancer).

Diagnosis = testing among individuals suspected to have the disease to confirm the presence or absence of the disease.

Examples of screening tests:

- Questions through questionnaires
- Clinical examinations – e.g., screening kids for congenital hip dislocation
- Laboratory tests – e.g., fasting blood sugar
- Genetic tests – e.g., newborn screening for phenylketonuria
- Radiographic exam – e.g., mammogram

What conditions should we screen for?

Many groups have advocated for screening for one disease or another. What criteria should we use to make the decision to screen for a particular disease?

Should we routinely screen for:

- Congenital conditions, e.g., PKU?
- Various cancers, e.g., breast, colon cancer, lung cancer, ovarian, testicular cancer?
- Cardiovascular disease and stroke risk?
- HIV/AIDS?
- Scoliosis?
- Others?

Prerequisites for a Successful Screening Program

The requirements for screening have been summarized by Wilson and Jungner (1968)[55] as follows:

1. The condition is an important public health problem.
2. The natural history of the condition is understood.
3. The condition should have a recognizable latent or early asymptomatic stage.
4. There should be an accepted treatment for the condition.
5. Applying early treatment of the disease should influence the course and prognosis of the disease.
6. There should be a suitable screening test or examination with the following characteristics:
    - Relatively sensitive and specific
    - Detects disease at a latent or early symptomatic stage
    - Simple and inexpensive
    - Safe
    - Acceptable to populations and providers
7. Available facilities for diagnosis and treatment.
8. Screening costs are economically balanced with possible expenditures on medical care.
9. There should be an agreed-on policy concerning whom to treat as patients.
10. There should be an ongoing case-finding, not a one-time shot.

Intrinsic Properties for a Screening Test

- Reliability (reproducibility) of the test = extent to which repeated measurements get similar results

- Validity (accuracy) = Ability of test to detect who has the disease and who has not, usually judged against another test of greater-known accuracy ("gold standard"). Validity is measured by sensitivity, specificity, and predictive value of positive or negative test.

In an ideal test, the test values of healthy and ill subjects do not overlap as in the diagram below.

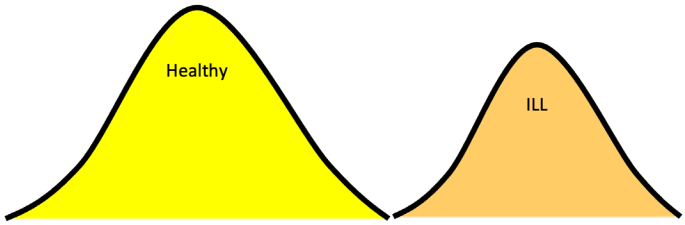


Diagram Showing No Overlap in the Test Values

However, in the real world, test values of healthy and ill subjects do overlap and lead to a problem with validity as in the diagram below.

We need a threshold (cut-off) value to distinguish "normal" from "abnormal" test values.

When the result of a screening test is compared with the result of the confirmatory (gold standard) test, the following four findings will result as shown in the following diagram:

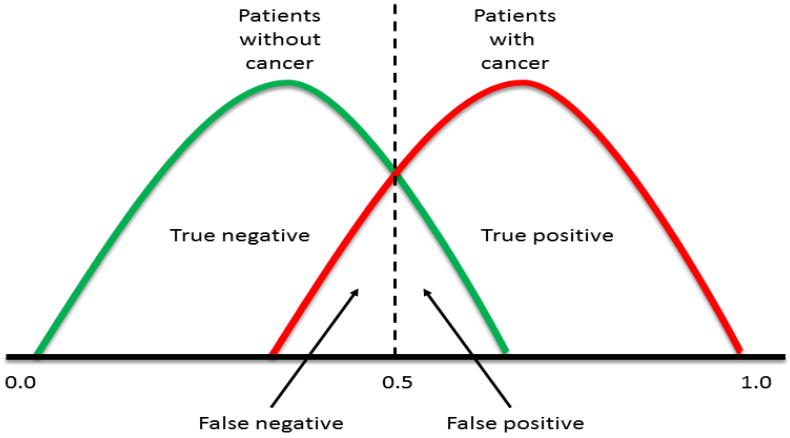


Diagram Showing Test Values Overlap

- True positive: Ill with a positive screening test
- True negative: Healthy with a negative screening test
- False positive: Healthy with positive screening test
- False negative: Ill with negative screening test

Validity

Validity refers to the accuracy of a measure or a test that means whether the results really do represent what they are supposed to measure. It has two components; sensitivity and specificity.

Sensitivity: Probability of positive test among those who have the disease or the ability of the test to correctly identify those who have the disease.

Specificity: Probability of a negative test among those without disease or ability of the test to

correctly identify those who do not have the disease.

In the table below:
Sensitivity = true positives/total with disease = a/(a+c)
Specificity = True negatives/total without disease = d/(b+d)

A Two by Two Table that Shows the Result of a Screening Test as Compared with the Result of the Confirmatory (Gold Standard) Test

	Diseased	**Not Diseased**	**Total**
Positive Screening Test	True positive (a)	False positive (b)	a+b
Negative Screening Test	False negative (c)	True negative (d)	c+d
Total	a+c	b+d	a+b+c+d

Example 1: Sensitivity and specificity
Suppose you have the figures in the table below which represent the results of comparing results of a screening test with the results of a confirmatory test (gold standard test). How can you calculate sensitivity and specificity of the screening test?

	Diseased	**Not Diseased**	**Total**
Positive Screening Test	9500 (true positives)	4500 (false positives)	14000
Negative Screening Test	500 (false negatives)	85500 (true negatives)	86000
Total	10000	90000	100000

- Sensitivity = True positives/All diseased = 9500/10000 = 95%
- Specificity = True negatives/All not diseased = 85500/90000 = 95%
- The false positive rate is calculated as FP/FP+TN, where FP is the number of false positives and TN is the number of true negatives = 4500/90000 = 5%. It's the probability that a positive result will be given when the true value is negative.
- The false negative rate is calculated as FN/FN+TP, where FN is the number of false negatives and TP is the number of true positives = 500/10000 = 5%. It's the probability that a negative result will be given when the true value is positive.

Example 2: Suppose you have been given the following data about a hypothetical screening program in a population:
Number of population screened is 50000
Sensitivity of the screening test used is known to be 90%
Specificity of the screening test used is known to be 80%
Prevalence of disease screened in the population is 5%

How can you fill in the following blank table depending on the data given above?

	Diseased	Not Diseased	Total
Positive Screening Test			
Negative Screening Test			
Total			

The answer to this exercise will be as follows:
The total population screened is 50000 which include the diseased and not diseased. Number of diseased persons = Prevalence * population = 0.05*50000 = 2500
Number of not diseased persons = Total population – Diseased persons = 50000-2500 = 47500
The number of true positives is calculated by using the sensitivity formula.
Sensitivity = true positives/all diseased
90/100= true positives/2500; true positives = 90% of 2500 = 2250
False negatives = all diseased-2250 = 2500-2250 = 250
True negatives = specificity*all not diseased = 80% of 47500 = 38000
False positives = all not diseased-true negatives = 47500-38000 = 9500

	Diseased	Not Diseased	Total
Positive Screening Test	2250	9500	11750
Negative Screening Test	250	38000	38250
Total	2500	47500	50000

Predictive value of a positive test (PVP): The likelihood among those with a positive test of having the condition.
PVP = True positives/All positives= a/a+b
Predictive value of a negative test (PVN): The likelihood of not having the condition among those with a negative test.
PVN = True negatives/All negatives = d/c+d
When we refer to the table before PVP = 2250/11750 = 19.1%
PVN = 38000/38250 = 99.3%
Effect of prevalence on PVP and PVN:
Prevalence impacts the positive predictive value (PPV) and negative predictive value (NPV) of tests. As the prevalence increases, the PPV also increases but the NPV decreases. Similarly, as the prevalence decreases, the PPV decreases while the NPV increases [56].

How much does prevalence, sensitivity, and specificity influence PVP?

Prevalence	Sensitivity	Specificity	Positive Predictive Value
1%	95%	95%	16.1%
1%	99%	95%	16.7%
1%	95%	99%	49.0%

cont'd

10%	95%	95%	67.9%
10%	99%	95%	68.8%
10%	95%	99%	91.4%

Sensitivity has little effect on predictive value positive, while prevalence and specificity has a strong effect on predictive value positive.

Can repeat testing of the same positive samples improve PVP? Let us consider the following example:

First test

Prevalence of the disease = 1%, sensitivity = 95%, specificity = 95%
The PVP= 16.1%

	Diseased	**Not Diseased**	**Total**
Screening Test Positive	950	4950	5900
Screening Test Negative	50	94050	94100
Total	1000	99000	100000

Second test

When we repeat the test on those who were positive in the first test, the following table will identify the improvement that happens on the PVP. Where the PVP increases to 78.5%.

	Diseased	**Not Diseased**	**Total**
Screening Test Positive	903	247	1150
Screening Test Negative	47	4703	4750
Total	950	4950	5900

Consequences of screening:

- True-positive results in
  - early diagnosis
  - early (possibly less radical) treatment
  - reduced morbidity, mortality, disability
  - reduced cost
- True-negative results in
  - reassurance

Yield of screening program

Yield is defined as the amount of previously unrecognized disease that is diagnosed and brought to treatment as a result of screening.

Factors that affect yield:

1. Sensitivity of the screening test: If the test has low sensitivity, it identifies a fraction of diseased individuals, and then the yield is poor, regardless of other factors.

2. Prevalence of disease: If prevalence of the unrecognized disease is high, the yield will be high.
3. Persons selected for the screening: Screening should be directed to persons with high risk. Example: Screening for diabetes is usually directed to persons over the age of 40 years, the obese, or those with a family history of diabetes.

Problems of false results in screening:

- False-positive results in
  - Unnecessary follow-up tests (inconvenience, morbidity, expense)
  - Labeling and anxiety
  - Over-treatment of questionable abnormalities
  - Fear of future tests
- False-negative results in
  - Delayed diagnosis, more advanced disease, premature death/disability
  - Disregard of early signs/symptoms
  - False reassurance
  - Exposure of others to infection

Bias in Screening

Length bias

Screening selectively identifies those with a long preclinical and clinical phase (i.e., those who would have a better prognosis regardless of the screening program) because more aggressive diseases are asymptomatic for a shorter period, screening is more likely to detect slower progressing diseases, such as slow-growing tumors, which have a better prognosis, including longer survival. Screening may thus falsely appear to improve survival.

Lead time bias

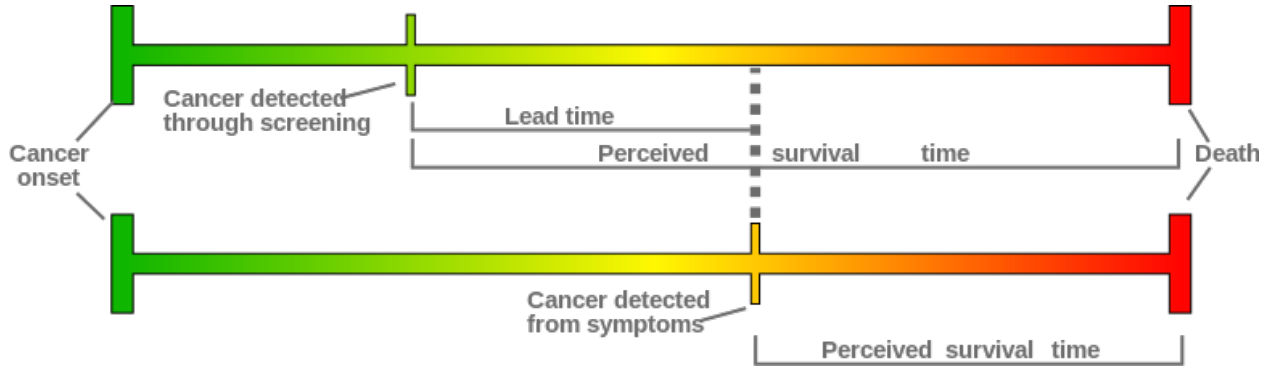
The apparently better survival that is observed for those screened is not because these patients are living longer, but instead because diagnosis is being made at an earlier point in the natural history of the disease.

Lead time: The time between early diagnosis with screening and the time in which diagnosis would have been made without screening.

Early diagnosis through screening may not necessarily prolong someone's life.
No additional life span has been gained and the patient may even be subject to added anxiety as the patient must live for longer with knowledge of the disease.

For example, the genetic disorder Huntington's disease is diagnosed when symptoms appear at around 50, and the person dies at around 65. The typical patient therefore lives about 15 years after diagnosis. A genetic test at birth makes it possible to diagnose this disorder earlier. If this newborn baby dies at around 65, the person will have "survived" 65 years after diagnosis, without having lived any longer than those diagnosed without DNA detection.

Lead time bias occurs when we conclude that persons discovered by screening survive longer than those discovered by clinical presentation.



Common Screening Tests

Fasting blood glucose for diabetes

- Blood pressure for hypertension
- PSA test for prostate cancer
- PAP smear for cervical cancer
- Mammography for breast cancer
- Fecal occult blood for colon cancer

Types of Screening

1. **Mass screening**

   The application of screening test to large, unselected population. Everyone in the group is screened regardless of the probability of having the disease or condition.

   Examples:

   - Visual defects in school children
   - Mammography in women aged 40 years or more
   - Newborn screening program

2. **High risk/selective/targeted screening**

   The screening of selected high-risk groups in the population.

   Examples:

   - Screening fetus for Down's syndrome in a mother who already has a baby with Down's syndrome
   - Screening for familial cancers, and DM
   - Screening for cancer of cervix in low SES women
   - Screening for HIV in risk groups

3. **Multipurpose (multiphasic) screening**

   The screening of a population by more than one test is done simultaneously to detect more than one disease, for example: screening of pregnant women for VDRL, HIV, HBV by serological tests.

4. **Opportunistic/case finding screening**

   There is no accurate or precise diagnostic test for the disease and where the frequency of its occurrence in the population is small. The main objective is to detect disease and bring patients to treatment. Example: Rheumatic heart disease in children.

Chapter 8

General Concepts of Disease Occurrence

Models for Disease Causation

Disease and other health events do not occur randomly in a population but are more likely to occur in some members of the population than others because of risk factors that may not be distributed randomly in the population. In epidemiology, there are several models of disease causation that help understand disease process. The most widely applied models are:

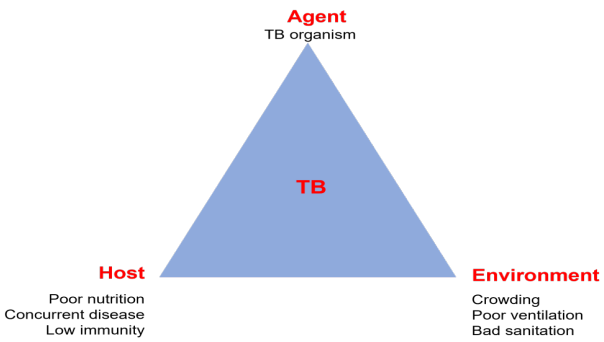
1. The epidemiological triad
2. Rothman's Causal Pies model
3. The "BEINGS" model
4. The Web of Causation model
5. The Wheel Theory model

The epidemiological triad

Among the simplest of these is the epidemiologic triad or triangle, the traditional model for infectious disease. The triad consists of an agent, a susceptible host, and an environment that brings the host and agent together. In this model, disease results from the interaction between the agent and the susceptible host in an environment that supports transmission of the agent from a source to that host.

Agent, host, and environmental factors interrelate in a variety of complex ways to produce disease.

Different diseases require different balances and interactions of these three components. Development of appropriate, practical, and effective public health measures to control or prevent disease usually requires assessment of all three components and their interactions.



Example: Epidemiological Triad of Tuberculosis

Agent: Originally referred to as an infectious microorganism or pathogen: a virus, bacterium, parasite, or other microbe. Generally, the agent must be present for disease to occur; however, the presence of that agent alone is not always sufficient to cause disease. A variety of factors influence whether exposure to an organism will result in disease, including the organism's pathogenicity (ability to cause disease) and dose.

Over time, the concept of "agent" has been broadened to include chemical and physical causes of disease or injury. These include chemical contaminants (e.g., lead) as well as physical forces (e.g., heat).

While the epidemiologic triad serves as a useful model for many diseases, it has proven inadequate for cardiovascular disease, cancer, and other diseases that appear to have multiple contributing causes without a single necessary one.

Host: Refers to the human who can get the disease. A variety of factors intrinsic to the host, sometimes called risk factors, can influence an individual's exposure, susceptibility, or response to a causative agent. Opportunities for exposure are often influenced by behaviors such as sexual practices, hygiene, and other personal choices as well as by age and sex. Susceptibility and response to an agent are influenced by factors such as genetic composition, nutritional and immunologic status, anatomic structure, presence of disease or medications, and psychological makeup.

Environment: Refers to extrinsic factors that affect the agent and the opportunity for exposure. Environmental factors include physical factors such as geology and climate, biological factors such as insects that transmit the agent, and socioeconomic factors such as crowding, sanitation, politics, health system, and the availability of health services.

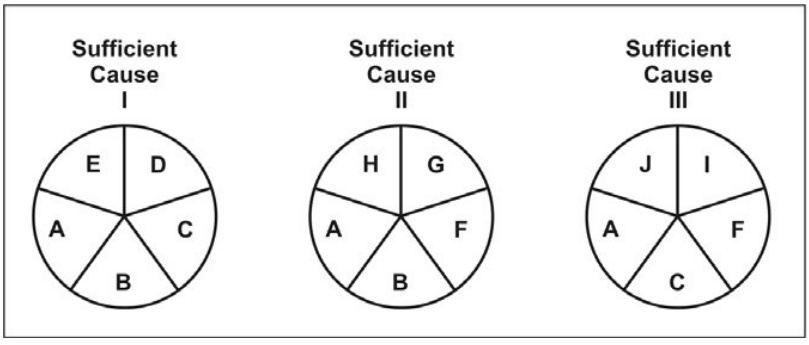
Rothman's Causal Pies model [41]

Because the agent-host-environment model did not work well for many non-infectious diseases, several other models that attempt to account for the multifactorial nature of causation have been proposed. One such model was proposed by Rothman in 1976 and has come to be known as the Causal Pies. This model is illustrated in Figure 1.17. An individual factor that contributes to cause disease is shown as a piece of a pie. After all the pieces of a pie fall into place, the pie is complete, and disease occurs.

The individual factor is called "component cause". The complete pie, which might be considered a causal pathway, is called a "sufficient cause". A disease may have more than one sufficient cause, with each sufficient cause being composed of several component causes that may or may not overlap. A component that appears in every pie or pathway is called a necessary cause, because without it, disease does not occur. Note in the figure below, the component cause A is a necessary cause because it appears in every pie.

The component causes may include intrinsic host factors as well as the agent and the environmental factors of the agent-host environment triad. A single component cause is rarely a sufficient cause by itself. For example, even exposure to a highly infectious agent such as the measles virus does not invariably result in measles disease. Host susceptibility and other host factors may also play a role.
At the other extreme, an agent that is usually harmless in healthy persons may cause devastating disease under different conditions.

Rothman's Causal Pies



Source: Rothman KJ. Causes. Am J Epidemiol 1976; 104:587–592.

For example, Pneumocystis carinii is an organism that harmlessly colonizes the respiratory tract of some healthy persons but can cause potentially lethal pneumonia in persons whose immune systems have been weakened by the human immunodeficiency virus (HIV).

The presence of Pneumocystis carinii organisms is therefore a necessary but not sufficient cause of pneumocystis pneumonia. It would be represented by component cause A in the above figure. As the model indicates, a particular disease may result from a variety of different sufficient causes or pathways.

For example, lung cancer may result from a sufficient cause that includes smoking as a component cause. Smoking is not a sufficient cause by itself, however, because not all smokers develop lung cancer. Smoking is also not a necessary cause, because a small fraction of lung cancer victims has never smoked.

Suppose Component Cause B is smoking, and Component Cause C is asbestos.

Sufficient Cause I include both smoking (B) and asbestos (C). Sufficient Cause II includes smoking without asbestos, and Sufficient Cause III includes asbestos without smoking. But because lung cancer can develop in persons who have never been exposed to either smoking or asbestos, a proper model for lung cancer would have to show at least one more Sufficient Cause Pie that does not include either component B or component C.

Public health action does not depend on the identification of every component cause. Disease prevention can be accomplished by blocking any single component of a sufficient cause, at least through that pathway. For example, elimination of smoking (component B) would prevent lung cancer from sufficient causes I and II, although some lung cancer would still occur through sufficient cause III.

The "BEINGS" model of disease causation

This model consists of a complex interplay of the following nine different factors:

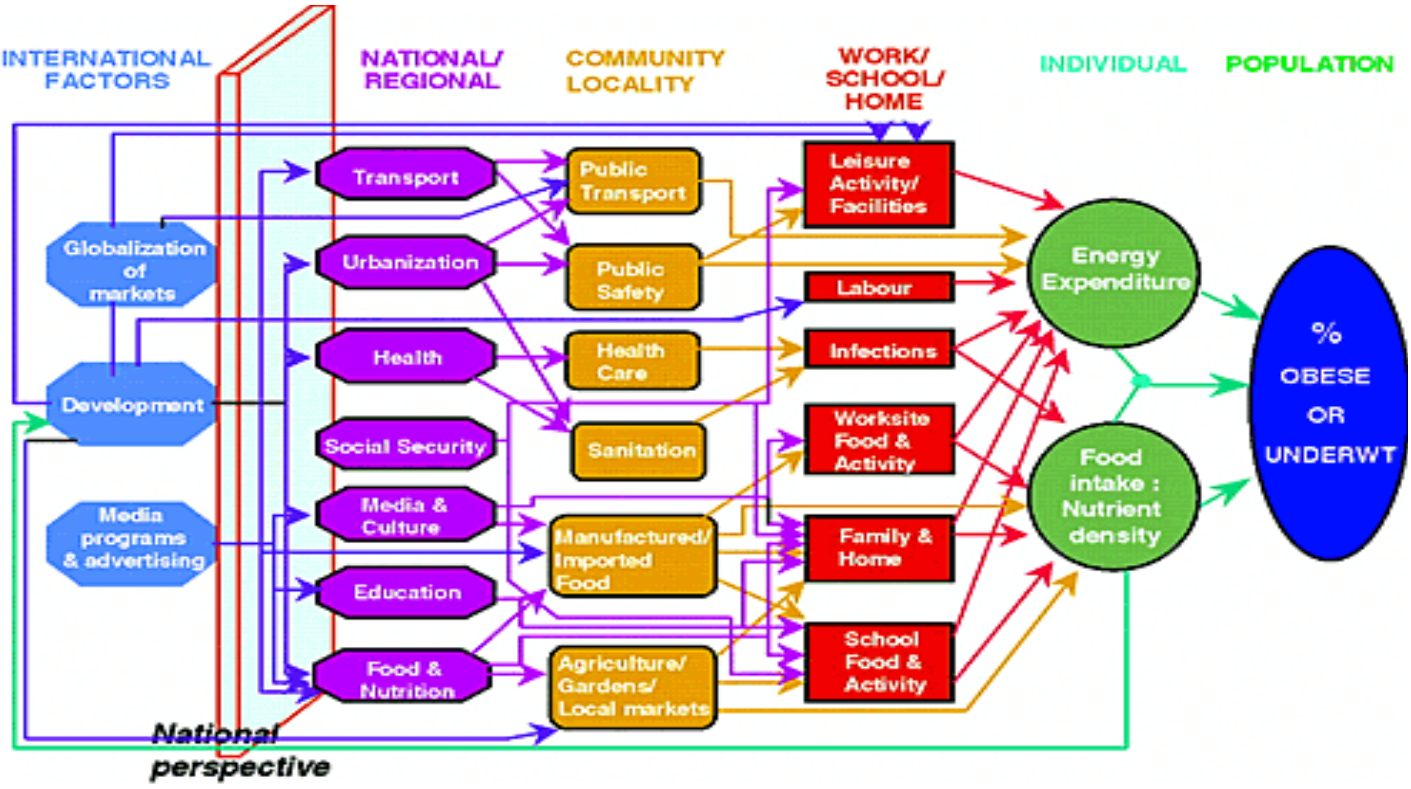
- **Biological factors in a human being**
- **Behavioral factors concerned with individual lifestyles, e.g., physical or sedentary lifestyle**
- **Environmental factors such as physical, chemical, and biological aspects of the environment**
- **Immunological factors**
- **Nutritional factors**
- **Genetic factors**
- **Social factors**
- **Spiritual factors**
- **Services factors, related to the various aspects of healthcare services**

The Web of Causation model

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. The image below summarizes a web of causation for obesity in the context of a socio-ecologic perspective. Note that some factors

are more "proximate" or immediate, such as decreased energy expenditure and increased food intake, while other factors or perhaps root causes are more "distal," such as globalization of markets, development, and advertising.

Web of Causation Diagram for Obesity



This model is ideally suited to the study of chronic disease, where the agent is often not known, and disease is the outcome of an interaction of multiple factors. The model considers all predisposing factors of any type and their complex interrelationship with one another in the causation of disease.

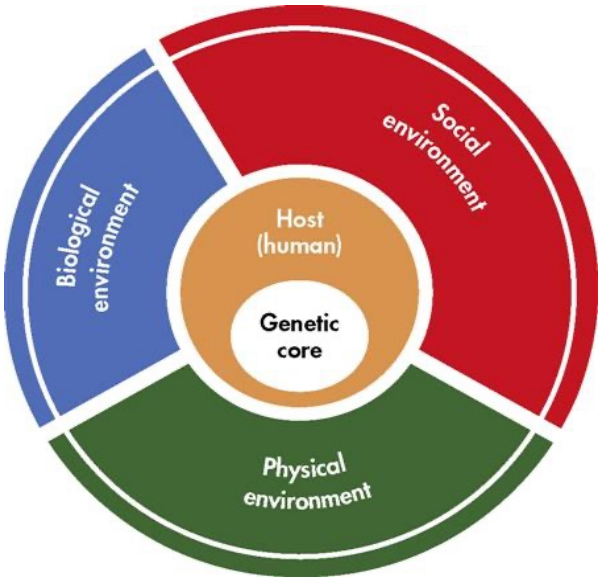
The various factors in this web are like an interacting web of a spider. Each factor has its own relative importance in causing the final departure from the state of health, as well as interacts with others, modifying one another's effect.

The Wheel Theory model

As medical knowledge advanced, an additional aspect of interest that came into play is the comparative role of genetic and environmental (i.e., extrinsic factors outside the host) factors in causation of disease.

The "triad" as well as the "web" theory do not adequately cover up this differential. To explain such relative contribution of genetic and environmental factors, the "wheel" theory has been postulated.

Wheel Model of Man-Environment Interaction



Source: Mausner and Kramer, 1985.

It emphasizes the interplay of physical, biological, and social environments in disease causation. It also brings genetics into the mix. This model de-emphasizes the agent as the sole cause of disease.

The wheel of causation model has a characteristic core which is the host, the other aspects that surround the core are biological, social, and physical determinants of disease. This model emphasizes the unity of genes and host within an interactive environmental envelope.

Natural History of Disease

The natural history of disease refers to the progress of a disease process in an individual over time, in the absence of intervention. The process begins with exposure to or accumulation of factors capable of causing disease without medical intervention, and the process ends with recovery, or disability, or death.

Understanding the progress of disease process and its pathogenetic chain of events is important for the application of preventive measures. So, if the patient develops a certain disease, we will be able to know the symptoms and signs and its duration.

The natural history of disease is best established by cohort studies. As these studies are costly, understanding of the natural history of disease is largely based on other epidemiological studies, such as cross-sectional and retrospective studies, undertaken in different population settings.

What the physician sees in the hospital is just an "episode" in the natural history of disease. The epidemiologist, by studying the natural history of disease in the community setting is in a unique position to fill the gaps in knowledge about the natural history of disease

The natural history consists of two phases

1. Pre-pathogenesis phase
2. Pathogenesis phase

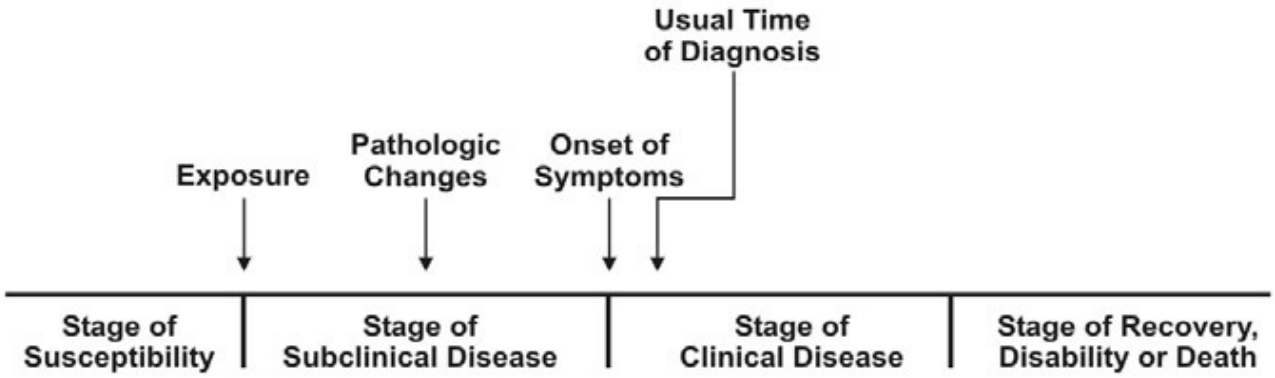
Pre-pathogenesis phase: This refers to the period prior to the onset of disease in humans. The disease agent has not yet entered the human host, but the factors which favor its interaction with the human host are already existing in the environment. This situation is frequently referred to as "man exposed to the risk of disease."

Pathogenesis phase: This phase begins with the entry of the disease "agent" in the susceptible human host. After the entry, agent multiplies and induces tissue and physiological changes, the disease progresses through the period of incubation and later through the period of early and late pathogenesis. The outcome of the disease may be recovery, disability, or death.

In chronic diseases, the early pathogenesis phase is less dramatic and is also called a pre-symptomatic phase. During the pre-symptomatic stage, there is no clinical disease. The pathological changes are essentially below the level of the "clinical horizon."

The clinical stage begins when recognizable signs or symptoms appear. By the time symptoms and signs appear, the disease phase is already well advanced into the late pathogenesis phase.

Natural History of Disease Timeline



Source: Centers for Disease Control and Prevention. Principles of Epidemiology, 2nd ed. Atlanta: U.S. Department of Health and Human Services, 1992.

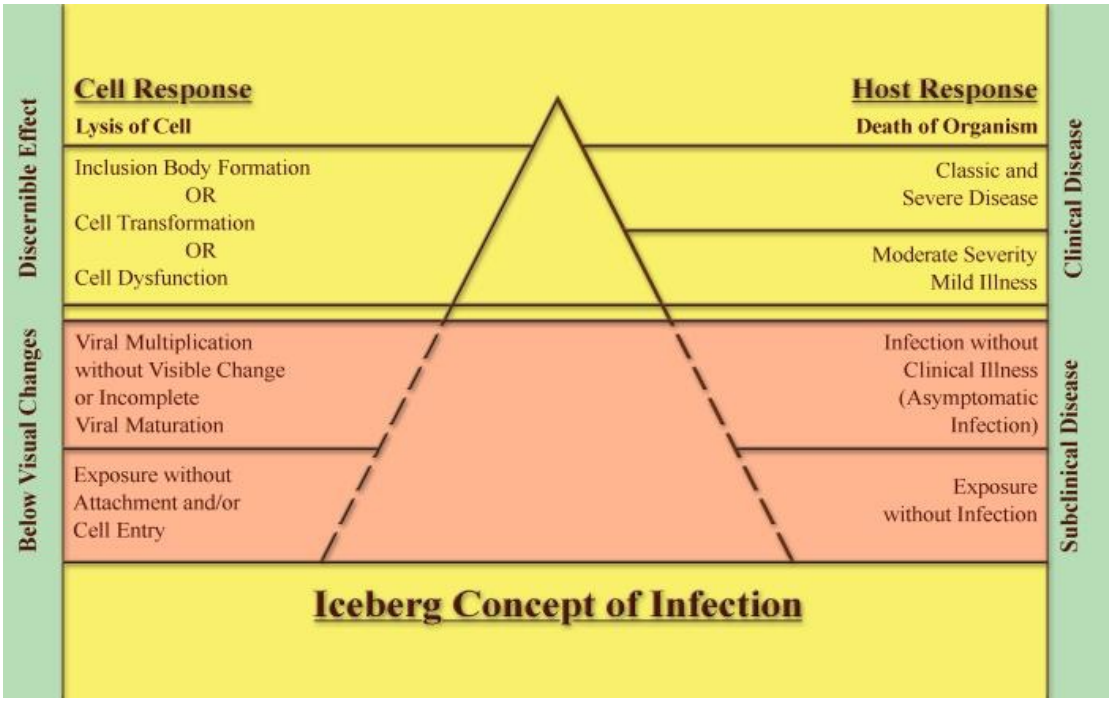
Spectrum of Disease and the Iceberg Phenomenon

The disease spectrum is a graphic representation of variations in the manifestations of disease. At one end of the disease spectrum there are sub-clinical infections which are not ordinarily identified, and at the other end there are fatal illnesses. In the middle of the spectrum lie illnesses ranging in severity from mild to severe. These different manifestations are the result of individuals' different states of immunity and receptivity.

The disease spectrum presents challenges to clinicians and to public health workers. For instance, cases of illness diagnosed by clinicians in the community often represent only the "tip of the iceberg." Many additional cases may be too early to diagnose or may remain asymptomatic or may be misdiagnosed etc. For public health workers and epidemiologists, the additional challenge is that

persons with undiagnosed infections may also be able to transmit the disease to others.

The below diagram illustrates the classic iceberg concept of infection, dividing the seen, or discernible effects of a disease process, from the effects that are unseen. The disease process begins at a point of inoculation, reaching the prodromal, or earliest onset of symptoms, and eventually becoming a full blown, observable clinical disease, where it reaches the tip of the iceberg. Those patients that are visibly infected with an organism represent just the tip of the iceberg of patients that are colonized or infected. Just because a patient is not showing signs of infection does not mean that he does not carry organisms that could be transferred to another patient if proper hygiene and other infection control precautions are not taken.



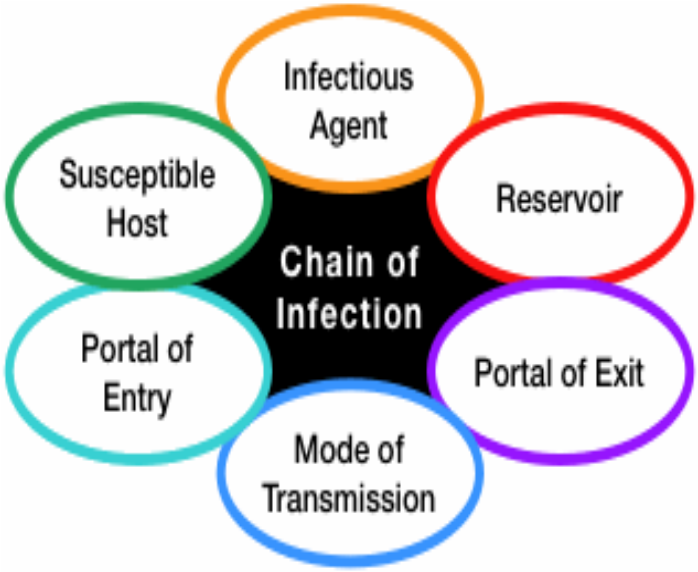
Source: CDC/ Dr. Francis T. Forrester, 1977.

Chain of Infection

The traditional epidemiologic triad model holds that infectious diseases result from the interaction of agent, host, and environment. More specifically, transmission occurs when the agent leaves its reservoir or host through a portal of exit, is conveyed by some mode of transmission, and enters through an appropriate portal of entry to infect a susceptible host. This sequence is called the chain of infection. It is usually presented by what we call the life cycle of the disease.

The spread of infection can be described as a chain with six links:

1. Infectious agent (pathogen)
2. Reservoir (the normal location or habitat of the pathogen)
3. Portal of exit from the reservoir
4. Mode of transmission
5. Portal of entry into a host
6. Susceptible host



The Six Links of the Chain of Infection

Infectious agents (pathogens)

This includes not only bacteria but also viruses, fungi, and parasites. The virulence of these pathogens depends on their number, potency, ability to enter and survive in the body, and the host's susceptibility. For example, the smallpox virus is particularly virulent, infecting almost all people exposed. In contrast, the tuberculosis bacillus infects only a small number of people, usually people with weakened immune function, or those who are undernourished and living in crowded conditions.

Viruses are intracellular parasites; that is, they can only reproduce inside a living cell. Some viruses, such as HIV, hepatitis B and C, have the ability to enter and survive in the body for years before disease symptoms occur. Other viruses, such as influenza and COVID-19, quickly announce their presence through characteristic symptoms.

Reservoir

A reservoir is any person, animal, arthropod, plant, soil, or substance (or combination of these) in which an infectious agent normally lives and multiplies. The infectious agent depends on the reservoir for survival, where it can reproduce itself in such a manner that it can be transmitted to a susceptible host. Animate reservoirs include people, insects, birds, and other animals. Inanimate reservoirs include soil, water, food, feces, intravenous fluid, and equipment.

Portal of exit

The portal of exit is the path by which a pathogen leaves the reservoir. For a human reservoir, the portal of exit can include blood, respiratory secretions, and anything exiting from the gastrointestinal or urinary tracts.

Portal of entry

Infectious agents get into the body through various portals of entry, including mucous membranes, non-intact skin, and the respiratory, gastrointestinal, and genitourinary tracts. Pathogens often enter the host's body through the same route they exited the reservoir, e.g., airborne pathogens from one person's sneeze can enter through the nose of another person.

Once a pathogen has exited the reservoir, it needs a mode of transmission to transfer itself into a host. This can be accomplished by entering the host through a receptive portal of entry. Transmission can be by direct contact, indirect contact, or through the air.

Transmission of respiratory infections such as COVID-19 is primarily via virus-laden fluid particles (i.e., droplets and aerosols) that are formed in the respiratory tract of an infected person and expelled from the mouth and nose while breathing, talking, singing, coughing, and sneezing. The competing effects of inertia, gravity, and evaporation determine the fate of these droplets. Large droplets settle faster than they evaporate and contaminate surrounding surfaces. Smaller droplets evaporate faster than they settle, forming droplet nuclei that can stay airborne for hours (becoming aerosolized) and may be transported over long distances.

Human-to-human transmission of COVID-19 occurs primarily via three routes: (1) large particles that are expelled with sufficient momentum to directly impact the recipients' mouth, nose, or conjunctiva; (2) physical contact with droplets deposited on a surface and subsequent transfer to the recipient's respiratory mucosa; and (3) inhalation of aerosolized droplet nuclei delivered by ambient air currents. The first two routes associated with large droplets are referred to as the "droplet" and "contact" routes of transmission, whereas the third is referred to as "airborne" transmission.

Airborne (aerosol) transmission

Aerosols are small particles (≤5 μm) that can rapidly evaporate in the air, leaving behind droplet nuclei that are small enough and light enough to remain suspended in the air for hours. Airborne transmission can occur when the residue of evaporated droplets from an infected person remains in the air long enough to be transmitted to the respiratory tract of a susceptible host.

There is increasing evidence that the COVID-19 coronavirus can move from person-to-person through the air, particularly in poorly ventilated, enclosed spaces. This means an infectious agent may remain infectious when suspended in the air over long distances and time.

Airborne transmission of SARS-CoV-2 is known to occur during aerosol-generating medical procedures. The scientific community has been actively discussing and evaluating whether SARS-CoV-2 may also spread through aerosols in the absence of aerosol-generating procedures, particularly in indoor settings with poor ventilation.

Comparing airborne (aerosol) transmission to droplet transmission is an important issue because, if COVID-19 is easily transmitted via airborne particles, then distancing, facemasks, and shields may not be enough to protect someone from exposure to the virus.

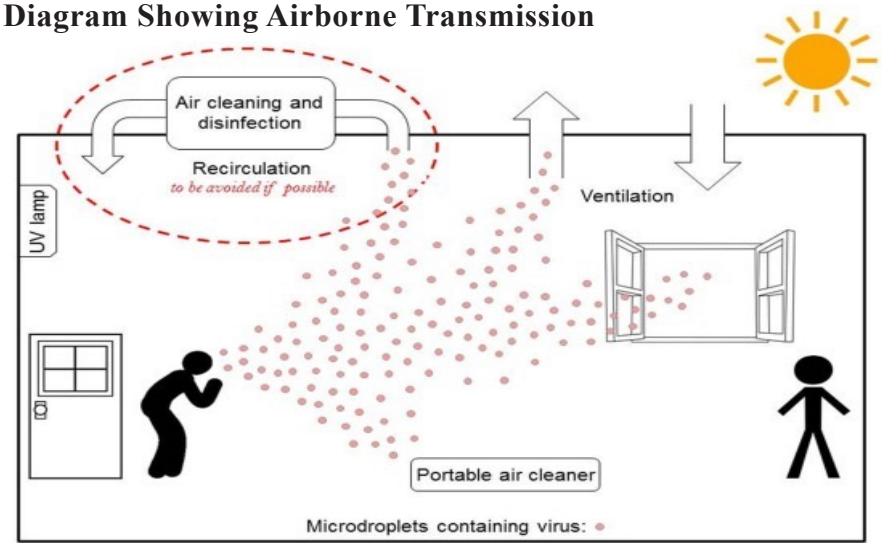
Investigators have demonstrated that speaking and coughing produce a mixture of both droplets and aerosols in a range of sizes, that these secretions can travel together for up to 27 feet, that it is feasible for SARS-CoV-2 to remain suspended in the air and viable for hours, that SARS-CoV-2 RNA can be recovered from air samples in hospitals, and that poor ventilation prolongs the amount of time that aerosols remain airborne.

During the initial isolation of thirteen individuals from the Diamond Princess cruise ship who had COVID-19, at the University of Nebraska Medical Center, researchers collected air and surface samples to examine viral shedding from isolated individuals. They detected viral contamination among all samples, supporting the use of airborne isolation precautions when caring for COVID-19 patients.

The presence of contamination on personal items was expected, particularly those items that are routinely handled by individuals in isolation, such as cell phones and remote controls, as well as medical equipment that is in near-constant contact with the patient. The observation of viral replication in cell culture for some of the samples confirms the potentially infectious nature of the recovered virus.

The transport of droplet nuclei over larger distances is primarily driven by ambient air flow, and indoor environments such as homes, offices, malls, aircraft, and public transport vehicles pose a particular challenge for disease transmission. The importance of ventilation in controlling airborne transmission of infections is well known. Indoor spaces can have extremely complex flows, due to ventilation systems and other factors that influence them.

Diagram Showing Airborne Transmission



Source: Environmental International Volume 142. CC BY-NC-ND 4.0.

Indirect contact

Indirect contact includes both vehicle-borne and vector borne contact. A vehicle is an inanimate go-between; an intermediary between the portal of exit from the reservoir and the portal of entry to the host. Inanimate objects such as cooking or eating utensils, handkerchiefs and tissues, soiled laundry, doorknobs and handles, and surgical instruments and dressings are common vehicles that can transmit infection. Blood, serum, plasma, water, food, and milk also serve as vehicles. For example, food can be contaminated by E.coli if food handlers do not practice appropriate handwashing techniques after using the bathroom. If the food is eaten by a susceptible host, such as a young child or a person with HIV/AIDS, the resulting infection can be life-threatening.

Vector-borne is transmission by an animate intermediary, an animal, insect, or parasite that transports the pathogen from reservoir to host. Transmission takes place when the vector injects salivary fluid by biting the host, or deposits feces or eggs in a break in the skin. Mosquitoes are vectors for malaria

and West Nile virus. Rodents can be vectors for Hantavirus.

Susceptible host

The final link in the chain of infection is the susceptible host who is someone at risk of infection. Infection does not occur automatically when the pathogen enters the body of a person whose immune system is functioning normally. When a virulent pathogen enters an immune-compromised person, however, infection generally follows. Whether exposure to a pathogen results in infection depends on several factors related to the person exposed (the host), the pathogen (the agent), and the environment. Host factors that influence the outcome of an exposure include the presence or absence of natural barriers, and the functional state of the immune system.

Prevention of Diseases

Prevention is the process of intercepting or opposing the "cause" or the "cycle" of a disease and thereby the disease process.

Why is it important?

1. Individual benefit: increases the survival rates and the person's productivity
2. Economic benefit: preventing the disease is less costly than treating the disease itself and possibly its complications

Successful prevention depends on:

1. Knowledge of causation
2. Dynamics of transmission
3. Identification of risk factors (smoking, hypertension, physical activity) and risk groups (family history of colon cancer; you'll do a check-up for the family to reduce the incidence of the disease)
4. Availability of prophylactic or early detection and treatment measures
5. Organization to apply these measures
6. Continuous evaluation

Levels of prevention

- Primordial prevention
- Primary prevention
- Secondary prevention
- Tertiary prevention

Primordial prevention: This is the prevention of the emergence or development of risk factors in population groups in which they have not yet appeared. For example, many adult health problems (e.g., obesity and hypertension) have their early origin in childhood, so efforts are directed towards encouraging children to adopt healthy lifestyles (e.g., physical exercise, healthy dietary habits etc.) so the prevalence of HTN and obesity will reduce when they get older. The main intervention in primordial prevention is through individual and mass education.

Primary prevention: It can be defined as "action taken prior to the onset of disease, which removes the possibility that a disease will ever occur." They are at high risk but they don't have the disease yet, so we interfere with this stage to prevent the disease from happening. It signifies intervention in

the pre-pathogenesis phase of a disease. Example: vaccination against certain diseases.

The concept of primary prevention is now being applied to the prevention of chronic diseases such as coronary heart disease, hypertension and cancer based on elimination or modification of "risk-factors" of disease.

Two types of strategies:

- Population (mass) strategy: Directed at whole population irrespective of the individual risk levels. Directed towards socio-economic, behavioral, and lifestyle changes.
- High risk strategy: Includes identification of "High risk groups" in the population and brings preventive care to these risk groups. e.g., People with a family history of hypertension, allergic disease, diabetes.

Secondary prevention: Defined as an action which stops the progress of a disease at its initial stage and prevents complications. It is applied in the early pathogenesis stage of disease. It reduces the prevalence of the disease by shortening its duration. It may also protect others in the community from acquiring the infection and thus provide, at once, secondary prevention for the infected individuals and primary prevention for their potential contacts. The specific interventions used are early diagnosis and treatment.

Early detection of health impairment is defined as the detection of disturbances of homoeostatic and compensatory mechanisms while biochemical, morphological and functional changes are still reversible. e.g., screening for disease for breast cancer (using mammography) and cervical cancer (using Pap smear). Medical examinations of school children, of industrial workers and various disease screening programs.

Tertiary prevention: These include all measures undertaken when the disease has become clinically manifest or advanced, with a view to:

1. Prevent or delay death, i.e., chemotherapy treatment for cancer patients
2. Reduce or limit impairments and disabilities
3. Minimize suffering
4. Promote the subject's adjustment to incurable conditions

Herd immunity

The term "herd immunity" was first used in 1894 by American veterinary scientist and then Chief of the Bureau of Animal Industry of the US Department of Agriculture Daniel Elmer Salmon to describe the healthy vitality and resistance to disease of well-fed herds of hogs. In 1916, veterinary scientists inside the same Bureau of Animal Industry used the term to refer to immunity arising following recovery in cattle infected with brucellosis.

By 1923 the term was being used by British bacteriologists to describe experimental epidemics with mice, experiments undertaken as part of efforts to model human epidemic disease. By the end of the 1920s the concept was used extensively – particularly among British scientists – to describe the buildup of immunity in populations to diseases such as diphtheria, scarlet fever, and influenza.

Herd immunity was recognized as a naturally occurring phenomenon in the 1930s when A. W.

Hedrick published research on the epidemiology of measles in Baltimore, and took notice that after many children had become immune to measles, the number of new infections temporarily decreased, including among susceptible children. Despite this knowledge, efforts to control and eliminate measles were unsuccessful until mass vaccination using the measles vaccine began in the 1960s.

Herd immunity has been defined by Fox 1970[57], as "the resistance of a group to invasion and spread of an infectious agent, based on the immunity of a high proportion of individual members of the group."

In the 1970s, the theorem used to calculate a disease's herd immunity threshold was developed. During the smallpox eradication campaign in the 1960s and 1970s, the practice of ring vaccination, to which herd immunity is integral, began as a way to immunize every person in a "ring" around an infected individual to prevent outbreaks from spreading.

Herd immunity (also called herd effect, community immunity, population immunity, or mass immunity) is a form of indirect protection from infectious disease that can occur with some diseases when a sufficient percentage of a population has become immune to an infection, whether through vaccination or previous natural infections or both, thus reducing the likelihood of infection for individuals who lack immunity[58].

Immune individuals are unlikely to contribute to disease transmission, thereby disrupting chains of infection, which stops or slows the spread of disease. The greater the proportion of immune individuals in a community, the smaller the probability that non-immune individuals will come into contact with an infectious individual. Some individuals cannot become immune because of medical conditions, such as an immunodeficiency or immunosuppression, and for this group herd immunity is a crucial method of protection.

Once the herd immunity threshold has been reached, the disease gradually disappears from a population. This elimination, if achieved worldwide, may result in a permanent reduction in the number of infections to zero, called eradication. Herd immunity created via vaccination contributed to the eventual eradication of smallpox in 1977 and has contributed to the reduction of other diseases. Herd immunity threshold (HIT): A critical proportion of the population who becomes immune to a certain disease, at which the disease may no longer persist in the population, or cease to be endemic. It is also called herd immunity level (HIL).

The exact herd immunity threshold (HIT) varies depending on the basic reproduction number of the disease. An example of a disease with a high threshold is measles, with a HIT exceeding 95%. The theoretical basis for herd immunity generally assumes that:

- Vaccines induce solid immunity.
- Populations mix at random.
- The pathogen does not evolve to escape the immune response.
- There is no non-human vector for the disease.

Herd immunity (or community immunity) occurs when a high percentage of the community is immune to a disease (through vaccination and/or prior illness), making the spread of this disease from person

to person unlikely. Even unvaccinated individuals (such as newborns and the immunocompromised) are offered some protection because the disease has little opportunity to spread within the community. As a result, the whole community becomes protected — not just those who are immune.

Herd immunity depends on the disease's contagiousness. Diseases that spread easily, such as measles, require a higher number of immune individuals in a community to reach herd immunity (high threshold).

If enough people are vaccinated against dangerous diseases, those who are susceptible and cannot be vaccinated are protected because the germ will not be able to "find" those susceptible individuals. There are several reasons why some individuals in the community are unprotected:
Some protection from vaccines "wanes" or "fades" after a period.

- Some individuals do not receive the complete vaccine schedule that they should receive to be fully vaccinated. For example, you need two measles, mumps, and rubella (MMR) injections to be adequately protected.
- Some individuals may only receive one dose and mistakenly believe they are protected.
- Some individuals may object because of religious beliefs, and others are fearful of potential side effects or are skeptical about the benefits of vaccines.

One of the drawbacks of herd immunity is that people who have the same beliefs about vaccinations frequently live in the same neighborhood, go to the same school, or attend the same religious services, so there could be potentially large groups of unvaccinated people close together, the so called "hard to reach populations or groups" Once the percentage of vaccinated individuals in a population drops below the herd immunity threshold, an exposure to a contagious disease could spread very quickly throughout the community.

Overshoot

The cumulative proportion of individuals who get infected during a disease outbreak can exceed HIT. This is because the HIT does not represent the point at which the disease stops spreading, but rather the point at which each infected person infects fewer than one additional person on average. Once the HIT is reached, the number of additional infections does not immediately drop to zero. The excess of the cumulative proportion of infected individuals over the theoretical HIT is known as the overshoot.

Basic Reproduction Number (R naught or R0)

The basic reproduction number, R0, of an infectious disease is the average number of secondary cases generated by a single primary case in a fully susceptible population [59]. Statistical estimation of R0 has been performed for various infectious diseases aiming towards understanding the dynamics of transmission and evolution, and designing effective public health intervention strategies [60, 61].

The basic reproduction number is one of the most fundamental and often-used metrics to study the way of communicable disease transmission. The symbol R represents the actual transmission rate of a disease and stands for reproduction. Naught, or zero, stands for the zeroth generation (patient zero). It refers to the first documented patient infected by a disease in an epidemic.

R0 is an indicator of the contagiousness or transmissibility of infectious and parasitic agents and

represents the number of new infections estimated to stem from a single case in a population that has never seen the disease before. If R0 is 2, then one person is expected to infect, on average, two new people. R0 is one of the key values that can predict whether an infectious disease will spread into a population or die out. It is used to assess the severity of the outbreak, as well as the strength of the medical and/or behavioral interventions necessary for control.

R0 values indicate if a disease will spread or decline within a community and how far and how rapidly transmission will occur. It can also inform public health policy decisions used to mitigate the spread. The higher the R0, the more likely the disease will become an epidemic.

There are three different possibilities that can be conveyed by R0:
1. If R0 is less than 1, the disease will not spread and will eventually die out.
2. If R0 is 1, the disease will remain stable in the community but will not cause an epidemic.
3. If R0 is greater than 1, the disease will spread and may cause an epidemic.

How R0 is calculated

R0 is determined using complex mathematical equations that look at data from the disease's characteristics and transmissibility, human behavior, how often sick and susceptible people are expected to come into contact with one another, and where the affected community is located. Scientists may also add educated guesses.

One of the ways epidemiologists calculate R0 is by using contact tracing data obtained at the onset of the epidemic. Once an individual is diagnosed, that person's contacts are traced and tested. R0 is then computed by averaging the number of secondary cases caused by diagnosed individuals. However, counting the number of cases of infection during an epidemic can be extremely difficult, even when public health officials use active surveillance and contact tracing to attempt to locate all infected persons.

Although measuring the true R0 value is possible during an outbreak of a newly emerging disease, there are rarely sufficient data collection systems in place to capture the early stages of an outbreak when R0 might be measured most accurately. As a result, R0 is nearly always estimated retrospectively from sero-epidemiologic data (which looks for the presence of antibodies in the blood) or by using theoretical mathematical models. The estimated values of R0 generated by mathematical models are dependent on numerous decisions made by the modeler.

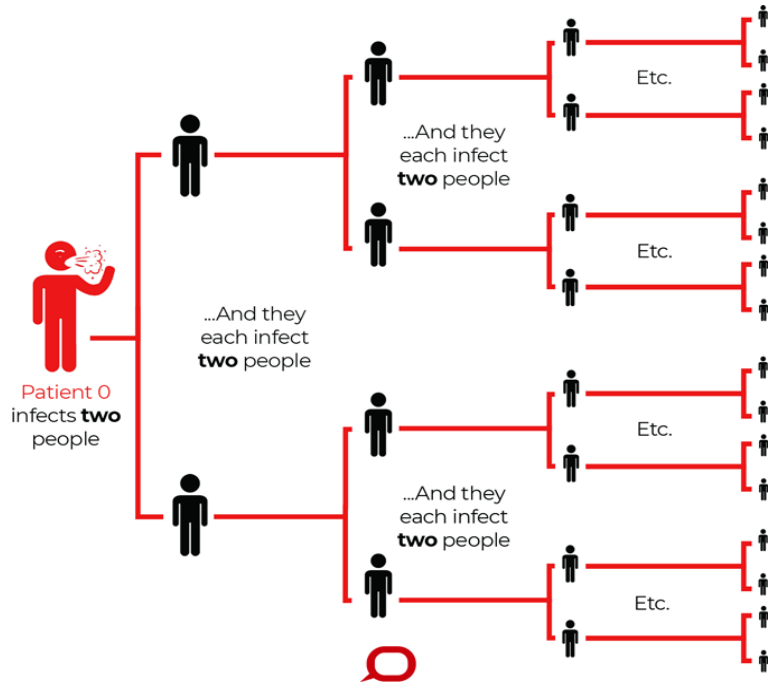
R0 is used to measure the transmission potential of a disease. The average number of secondary infections produced by a typical case of an infection is in a population where everyone is susceptible. For example, if the R0 for measles in a population is 15, then we would expect each new case of measles to produce 15 new secondary cases (assuming everyone around the case was susceptible). R0 excludes new cases produced by secondary cases.

The basic reproductive number is affected by several factors:
- The rate of contacts in the host population
- The probability of infection being transmitted during contact
- The duration of infectiousness

In general, for an epidemic to occur in a susceptible population R0 must be >1, so the number of cases increases.

Measures used successfully in previous epidemics, which have been shown to reduce the R0 of a disease, are:

- Screening
- Social distancing
- Tracking and tracing of exposed people and their contacts
- Handwashing
- Masking
- Quarantining
- Providing healthcare workers with proper protective equipment
- Vaccination



How a Virus with a Reproduction Number (R0) of 2 Spreads

Effective Reproductive Number (R)

Only in the case of a new pathogen (i.e., COVID-19, or previous recent pandemics), is a population rarely totally susceptible to an infection in the real world. Some contacts will be immune, for example due to prior infection which has conferred life-long immunity, or as a result of previous immunization. Therefore, not all contacts will become infected and the average number of secondary cases per infectious case will be lower than the basic reproduction number. This is measured by the effective reproductive rate (R). The effective reproductive number (R) is the average number of secondary cases per infectious case in a population made up of both susceptible and non-susceptible hosts. If R>1, the number of cases will increase, such as at the start of an epidemic. Where R=1, the disease is endemic, and where R<1 there will be a decline in the number of cases.

The effective reproduction number can be estimated by the product of the basic reproductive number and the fraction of the host population that is susceptible (x). So, R = R0 multiplied by x

Example: If R0 for influenza is 12 in a population where half of the population is immune, the

effective reproductive number for influenza is 12 x 0.5 = 6. Under these circumstances, a single case of influenza would produce an average of 6 new secondary cases. To successfully eliminate a disease from a population, R needs to be less than 1.

Herd immunity threshold (HIT) is calculated using the following formula:
HIT= (R0 -1) / R0 multiplied by 100. Or HIT = 1- (1/ R0) multiplied by 100.

Values of Basic Reproduction Number (R0) and Herd Immunity Thresholds (HITs) of Well-known Infectious Diseases Prior to Intervention

Disease	**R0**	**HIT**
Measles	12-18	92-94%
Chicken pox	10-12	90-92%
Mumps	10-12	90-92%
Rubella	6-7	83-86%
Polio	5-7	80-86%
Pertussis	5.5	82%
Smallpox	3.5-6	71-83%
COVID-19 (Alpha variant)	4-5	75-80%
COVID-19 (Delta variant)	5-8	80-88%
SARS	2-4	50-75%
Diphtheria	1.7-4.3	41-77%
Influenza (seasonal strains)	1.2-1.4	17-29%

Source: ***https://en.wikipedia.org/w/index.php?title=Basic_reproduction_number&action=edit&section=16***

Chapter 9

Public Health Surveillance

The word "surveillance" was originally derived from the French sur (over) and veiller (to watch). The term "surveillance" was used initially in public health to describe the close monitoring of persons who, because of exposure, were at risk for developing highly contagious and virulent infectious diseases that had been controlled or eradicated in a geographic area or among a certain population (e.g., cholera, plague, and yellow fever). These persons were monitored so that, if they exhibited evidence of disease, they could be quarantined to prevent spreading the disease to others.

Alexander Langmuir in 1963 applied surveillance for a disease to mean "the continued watchfulness over the distribution and trends of incidence [of a disease] through the systematic collection, consolidation, and evaluation of morbidity and mortality reports and other relevant data." He illustrated this application with four communicable diseases: malaria, poliomyelitis, influenza, and hepatitis [62].

The World Health Assembly in 1968 defined surveillance as "the systematic collection and use of epidemiologic information for the planning, implementation, and assessment of disease control" [63].

In the 1980s and 1990s, Thacker and others [64, 65, 66] expanded the term to include not just disease, but any outcome, hazard, or exposure. The term "surveillance" is often applied to almost any effort to monitor or determine health status, disease, or risk factors within a population.
The Centers for Disease Control and Prevention (CDC) defined epidemiologic surveillance as the "ongoing systematic collection, analysis, and interpretation of health data essential to the planning, implementation, and evaluation of public health practice closely integrated with the timely dissemination of these data to those who need to know" [67].

The Scope of Surveillance

The scope of surveillance is broad, from early warning systems for rapid response in the case of communicable diseases, to planned response in the case of chronic diseases which generally have a longer lag time between exposure and disease. Most countries have regulations for mandatory reporting of a list of diseases. These lists of notifiable diseases often include vaccine-preventable diseases such as polio, measles, tetanus and diphtheria as well as other communicable diseases such as tuberculosis, hepatitis, meningitis and leprosy etc. [68].

The following elements should be included in most definitions of surveillance:

- Systematic – organized, not haphazard
- Ongoing – not just a one-time survey
- Collection
- Analysis
- Interpretation
- Dissemination

- Of health data
- Link to public health practice (with the expectation that public health officials will look at it and take action).

Let's situate public health surveillance within the overall public health approach.

What are we trying to accomplish in public health?

We are trying to improve the health of a population. To do that, we need to know the following elements of the public health approach:

1. What's the problem?
2. What's the cause of the problem?
3. How does intervention work to address that problem?
4. How did you implement the intervention?

Public Health Surveillance has several purposes, but they all involve monitoring the health of the population. Specifically, surveillance is used to

- Assess the status of the public's health – how many people are getting sick, and from what?
- Trigger public health action – if the number of people getting sick is high, we'd better do something about it
- Define public health priorities – but we can't do everything immediately, so we can use data to set priorities for action, funding, programs, etc.
- Evaluate programs – finally, are those actions and programs having an impact, e.g., are disease rates declining? Surveillance data can tell us.

Uses of Public Health Surveillance

Surveillance is an essential feature of epidemiologic practice and may be used to [70, 71]:

1. Recognize isolated or clustered cases.
2. Describe the burden of or potential for diseases.
3. Assess the public health impact of events and assess trends.
4. Measure the causal factors of disease.
5. Monitor effectiveness and evaluate the impact of prevention and control measures, intervention strategies and health policy changes.
6. Plan and provide care.
7. In addition to estimating the magnitude of an epidemic and monitoring its trends, data can also be used to:
    - Strengthen commitment
    - Mobilize communities
    - Advocate for sufficient resources

A key principle of surveillance is to include only conditions for which surveillance can effectively lead to prevention. Another important principle is that surveillance systems should reflect the overall disease burden in the community.

Criteria for Selecting Diseases for Surveillance Include the Following [48, 58]:

- Incidence and prevalence

- Indices of severity (case-fatality ratio)
- Mortality rate and premature mortality
- Index of lost productivity (bed-disability days)
- Medical costs
- Preventability
- Epidemic potential
- Information gaps on new diseases

Sources of Data

Sources of data may be general or disease-specific, and include the following:

- Notifiable diseases or syndromes
- Mortality and morbidity reports
- Hospital records
- Laboratory diagnoses
- Outbreak reports
- Vaccine utilization
- Sickness absence records
- Blood banks
- Sentinel surveillance
- Environmental, vector, and animal surveillance
- Surveys
- Vital records
- Census data
- Registries
- Other data sources

Surveillance can collect data on any variable of the causal chain of disease, behavioral risk factors, preventive actions, cases, and program or treatment costs. The scope of a surveillance system is constrained by human and financial resources.

Surveillance relies mostly upon a routine system of reporting suspected cases within the health system, followed by validation and confirmation. Active and appropriate responses ranging from local containment measures to investigation and containment by a highly specialized team,

Prerequisites for Case Reporting

- Patient (or proxy, e.g., mother) must recognize that he/she is ill
- Patient must seek care
- Healthcare provider or lab must make an appropriate diagnosis
- Healthcare providers, labs, or institutions must report the case to the concerned health authority

Notifiable Disease System

Based on laws and regulations, the following steps should be made clear:

- Who must report, what, how, when etc.
- Usually from primary care providers, clinics, hospitals (e.g., infection control nurses) and/or laboratories to local health agencies

- Local health agencies are usually responsible for case investigations and are taking immediate action as necessary
- The local health agency forwards the report up, e.g., to the state or province, then to the national level

Regulations specify not only who must report (usually, physicians, laboratories, hospitals, clinics, sometimes school nurses and others) and the list of diseases to report, but also how to report, how promptly, etc. Some places are now accepting electronic reports via e-mail or website. Unfortunately, not every case of a reportable disease is reported.

After many years of requiring reporting of the 4 then 3 diseases that require quarantine, the WHO revised its requirements in 2005. The WHO now wants reporting of all cases of

1. Smallpox
2. Poliomyelitis (wild type)
3. Human influenza is caused by a new subtype
4. Severe acute respiratory syndrome (SARS)
5. Unexpected or "impactful" cases of cholera
6. Pneumonic plague
7. Yellow fever
8. Viral hemorrhagic fever
9. West Nile fever
10. Diseases of regional concern, e.g., dengue, Rift Valley fever, meningococcal disease
11. Any event that is a potential Public Health Emergency of International Concern (PHEIC)

Types of Surveillance

- Passive vs. active surveillance
- Case-based (individual) vs. aggregate
- Disease-specific vs. syndromic
- Population-based (comprehensive) vs. sentinel.
- Sero-surveillance
- Zoonotic surveillance

Passive surveillance versus active surveillance

Passive surveillance means the physician takes the initiative to send in the report; the health department sits back in its easy chair and waits for reports to come in. This is the most common type. It is initiated by the health provider and is usually adequate for monitoring trends over time, place and person.

For diseases of special interest, or during an outbreak or special event, the health department may solicit reports from healthcare providers by, say, calling hospitals, clinics, and physician practices once a week to ask whether they have seen any cases of disease X. This is called active surveillance.

The table below compares the number of reports received by active versus passive surveillance. Active surveillance collected about 3 times as many reports as passive surveillance for all diseases in that particular community.

Number of Reports and Ratio of Active to Passive Surveillance for 5 Selected Diseases

Disease	Active	Passive	Ratio
Hepatitis	78	27	2.9
Measles	11	8	1.4
Rubella	7	3	2.3
Salmonellosis	44	9	4.9
Total	140	47	3.0

Case-based (individual) versus aggregated surveillance

Case-based surveillance

Case-based surveillance refers to surveillance systems that collect information about each case at the individual level. This type of surveillance system has a case investigation form where information can be gathered from the patient or their family members, their medical records, and their laboratory records. At a minimum, more detailed information on person (who is infected), place (where they live, where they might have been infected), and time (when they became ill) is collected. A line list from this investigation form is created and reported up to their normal reporting channels.

In some scenarios, a case-based surveillance system might transition to aggregate as the number of cases becomes large as it overwhelms the system, like what happened during the 2009 H1N1 outbreak. In contrast, an aggregate surveillance system might become case-based temporarily in an outbreak to understand more of the epidemiology of the disease. Certain diseases, such as polio and measles, are recommended to be case based.

Example

Initially, when the United Nations (UN) development goals were established in 1990, measles was endemic in many countries, and mortality reduction was the primary goal. Given this, aggregate data were the most feasible approach and were conducted in most countries. By 2016, all six WHO regions had measles elimination goals. As measles has moved away from control toward elimination, case-based surveillance instead of aggregate surveillance is needed to ensure every case is reported and investigated.

When disease is endemic, case-based surveillance would quickly be overwhelmed given the time and resources, but as countries have fewer and fewer cases, it is relatively easier to conduct an investigation on every single case. One key advantage of case-based surveillance is that it allows one to analyze which age cohorts are being infected and their individual vaccination status to help target vaccination efforts and close existing vaccination gaps.

Aggregate surveillance

The surveillance of a disease or health event by collecting summary data on groups of cases. Aggregate surveillance data can exist in a variety of forms, but the main feature is that it lacks detailed information on specific cases. Aggregate data typically includes the number of cases (for example, the number of suspects and confirmed neonatal tetanus cases, or by age group) for a specific region and time period. This information can monitor the number of cases but lacks the individual-level data required for specific analyses.

Disease-specific versus syndromic surveillance

Surveillance networks identify and enroll cases that meet a specific case definition. Case definitions have three essential components: person, place, and time. Case definitions vary in sensitivity and specificity. Sensitive case definitions are more inclusive and are less likely to miss cases, but will include patients that do not have the disease. Specific case definitions have stricter criteria and exclude more patients that do not have the disease but can also miss patients with milder or atypical disease presentations. Both sensitive and specific case definitions can be used in infectious disease surveillance depending on the goals of surveillance. For example, sensitive case definitions may be preferred if it is important not to miss cases. In general, case definitions should be as sensitive and specific as possible. However, since a highly sensitive and specific case definition is not always possible, it is important that the case definition is at least applied systematically and consistently over the surveillance period.

Syndromic surveillance involves monitoring cases that meet a clinical case definition for the disease under surveillance, typically without laboratory confirmation [72].

This allows for rapid identification of a cluster of cases that might warrant further investigation. An example of syndromic surveillance includes acute fever/rash surveillance in many countries, which is used to monitor measles and rubella. The fever and rash could be due to a multitude of causes, and if there is an increase in the number of fever/rash cases reported, this could indicate an outbreak.

Other examples of diseases that are involved in syndromic surveillance include diarrhea and sexually transmitted diseases.

As mentioned earlier, syndromic surveillance focuses on syndromes rather than diagnosed illnesses. The advantage of this method is that reporting does not need to wait for laboratory confirmation, which can take 2 or more days. As a result, syndromic surveillance has become the preferred method for surveillance of bioterrorism. Particularly in the bioterrorism context, the term also refers to surveillance of over-the-counter medications (could a sharp rise in sales of anti-diarrheal medication in a community indicate an outbreak of gastroenteritis?), school and work absenteeism.

As field investigations are ongoing, laboratory testing can be performed on some or all of the cases identified by syndromic surveillance to determine the etiology. In the acute fever/rash surveillance system, laboratory specimens might be collected to undergo testing for measles and rubella.

Syndromic surveillance case definitions can be used in emergency or outbreak situations as an alert system to identify suspect cases that meet a broad case definition to then be further investigated.

In contrast, disease-specific or laboratory-based surveillance is based on confirmed cases in a laboratory where the etiologic agent can be identified through a variety of laboratory tests. For example, influenza surveillance networks use laboratory-confirmed influenza to determine the circulating strains to provide information for vaccine composition. The critical objective of laboratory-based surveillance is to monitor for emerging drug resistance in pathogens or shifts in serotype distribution.

Sentinel versus population-based surveillance

A sentinel surveillance site is a single or small number of health facilities that are responsible for collecting data on cases enrolled with the case definition under surveillance, including global networks surveying for diarrhea or pneumonia. Most sentinel sites do not have a predefined catchment population (or denominator to calculate incidence), and therefore data at these sites are simply numbers of cases (numerators). Sentinel site surveillance provides useful epidemiological information on proportions caused by different pathogens, age distribution, and risk factors and could also be used for monitoring trends of hospitalized cases within a health facility if health-care patterns and population have been stable. Furthermore, these data may be used in case–control studies to assess effectiveness of a vaccine or other preventive measures. Surveillance focused on one or a small number of surveillance sites often allows for gathering more data of higher quality [73].

In contrast, with population-based surveillance, every appropriate health facility reports on predefined diseases with the goal of identifying all cases in a specific geographic area. Population-based surveillance can either represent the whole country (national) or a defined subnational population area. Since the population is defined, these surveillance sites can produce rates of disease (for example, incidence and mortality rates), which allows for comparison of rates of disease between other population-based surveillance sites. Population-based surveillance is more costly than sentinel site surveillance but produces more generalizable data on incidence of disease. Population-based surveillance is surveillance that pertains to a general population (i.e., the entire population) defined by geographical boundaries. As a result, population-based surveillance tends to be:

- Representative of the population in the geographic area
- Based on the existing public health structure
- Increase potential for detection of rare diseases

Sero-surveillance

Sero-surveillance involves the use of blood specimens to determine the burden of disease or immunity gaps in a population. Sero-surveillance is frequently done as a periodic survey for multiple diseases of interest simultaneously.

Sero-surveillance cannot provide information in a timely manner; thus, an outbreak might have occurred that is discovered by sero-surveillance, but it might be potentially too late for an intervention to decrease disease transmission. Sero-surveillance is sometimes the only type of surveillance conducted for an infectious disease.

For example, hepatitis B is frequently asymptomatic in children, making evaluating the impact of vaccination efforts extremely challenging. The standard has become to perform sero-surveillance among cohorts of vaccinated children to identify the burden of the disease and determine the impact of vaccination efforts.

In some countries, national health surveys are conducted periodically and include a serologic component, allowing one to monitor trends in diseases and immunity over time.

Zoonotic surveillance

Zoonotic diseases cause disease in humans and can be challenging to control since both animals and humans can be hosts, including brucellosis, West Nile Virus, avian influenza, Ebola (and other hemorrhagic fevers), Lyme disease, SARS, and rabies. Historically, zoonotic and human disease surveillance existed separately, but there is a push to harmonize these systems to improve surveillance for diseases affecting both populations. Illness in one species might be a harbinger of illness in humans, and an integrated comprehensive surveillance system can help identify potential disease transmission that might be ongoing. For example, surveillance of Borrelia burgdorferi, the causative agent of Lyme disease in the tick population, can help public health authorities determine proper interventions to decrease the transmission from ticks to humans. One Health emphasizes the link of human health to the surrounding environment and animals. One of the mission statements of One Health is to improve the lives of all species by harmonizing both animal and human disease surveillance and control efforts. International organizations participating in One Health include WHO, the UN Food and Agricultural Organization (FAO), and the World Organization for Animal Health (WOAH).

The Process of Public Health Surveillance Includes the Following:

- Data collection
- Data tabulation/analysis
- Data interpretation
- Data dissemination
- Link to action

After morbidity, mortality, and other relevant data about a health problem have been gathered and compiled, the data should be analyzed by time, place, and person. Different types of data are used for surveillance, and different types of analyses might be needed for each. For example, data on individual cases of disease are analyzed differently than data aggregated from multiple records. The display of frequencies (counts) or rates of the health problem in simple tables and graphs, is the most common method of analyzing data for surveillance. For analysis of the majority of surveillance data, descriptive methods are usually appropriate.

Rates are useful and frequently preferred for comparing occurrence of disease for different geographic areas or periods because they take into account the size of the population from which the cases arose. One critical step before calculating a rate is constructing a denominator from appropriate population data.

For countywide rates, general population data is used. This data is usually available from the national census. In other calculations, the population at risk can dictate an alternative denominator. For example, an infant mortality rate uses the number of live-born infants; rates of surgical wound infections in a hospital requires the number of such procedures performed. In addition to calculating frequencies and rates, more sophisticated methods (e.g., space-time cluster analysis, time series analysis, or computer mapping) can be applied.

To determine whether the incidence or prevalence of a health problem has increased, data must be compared either over time or across areas. The selection of data for comparison depends on the

health problem under surveillance and what is known about its typical temporal and geographic patterns of occurrence.

For example, data for diseases that indicate a seasonal pattern (e.g., influenza and mosquito-borne diseases) are usually compared with data for the corresponding season from past years. Data for diseases without a seasonal pattern are commonly compared with data for previous weeks, months, or years, depending on the nature of the disease. Surveillance of chronic diseases typically requires data covering multiple years. Data for acute infectious diseases might only require data covering weeks or months. Data from one geographic area are sometimes compared with data from another area. For example, data from a governorate, province or state might be compared with data from adjacent governorates, provinces, states or with data from the country.

Analyzing by time

Basic analysis of surveillance data by time is usually conducted to characterize trends and detect changes in disease incidence. For notifiable diseases, the first analysis is usually a comparison of the number of case reports received for the current week with the number received in the preceding weeks. This data can be organized into a table, a graph, or both. An abrupt increase or a gradual buildup in the number of cases can be detected by looking at the table or graph.

For example, health officials reviewing the data for area F in the table below will have noticed that the number of cases of hepatitis A reported during week 4 exceeded the numbers in the previous weeks. This method works well when new cases are reported promptly.

Reported Cases of Hepatitis A, by Area and Week of Report, 1991 in Country X

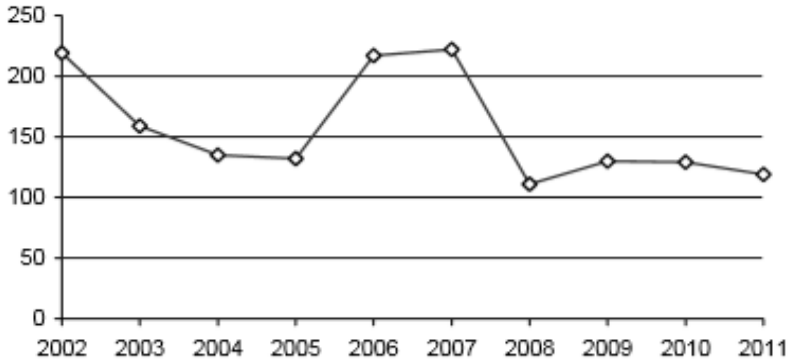
Area	Week1	Week2	Week3	Week4	Week5	Week6	Week7	Week8	Week9
A	-	-	-	1	-	-	1	-	-
B	-	-	-	-	-	-	-	-	-
C	-	-	-	2	1	-	2	-	3
D	-	-	1	3	1	1	-	-	1
E			1					2	
F	-	3	8	14	13	11	6	-	-
G	-	-	-	-	-	-	-	-	-
H	2	-	3	-	-	6	4	9	-
I	-	-	-	-	-	-	-	-	-
J	-	-	-	-	-	-	-	-	-
K	-	-	3	2	3	-	5	-	4
L	-	-	-	-	-	-	-	1	-

Another common analysis is a comparison of the number of cases during the current period to the number reported during the same period for the last 2–10 years. For example, health officials will have noted that the 11 cases reported for area F during weeks 1–3 during 1991 exceeded the numbers reported during the same 3-week period of the previous 3 years.

Analysis of long-term time trends, also known as secular trends, usually involves graphing occurrence of disease by year.

Example of secular trend:

Human Brucellosis in Jordan During the Years 2002-2011



Source: Ministry of health yearly statistical book, Jordan, 2011

Analysis by place

The usual method used is through analysis by the place of reporting, although the place where the exposure occurred is preferred. The analysis of cases by place is usually displayed in a table or a map.

The usual method of analysis is by the place of reporting, although the analysis by the place where the exposure occurred is preferred. Health departments usually analyze surveillance data by districts or by governorates. Rates are often calculated by adjusting for differences in the size of the population of different districts, governorates or other geographic areas.

Example 1 of tabulation analysis by place:

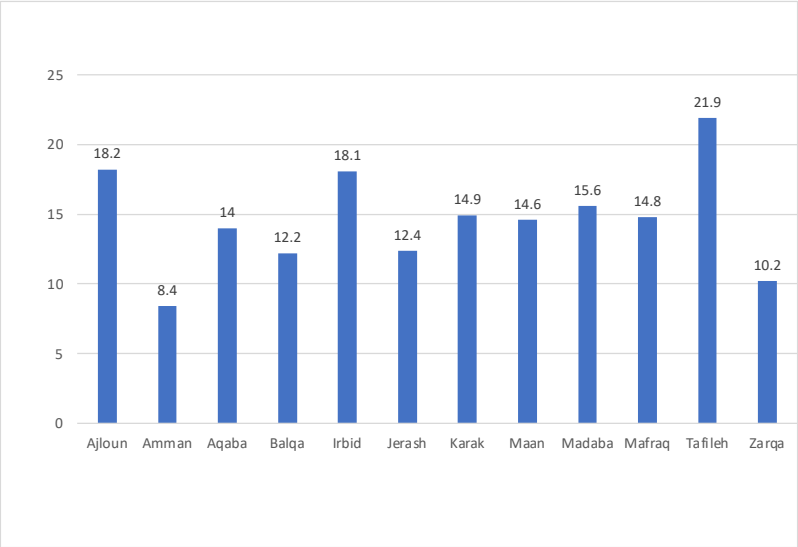
Cumulative Incidence Rates of COVID-19 per 100000 of Population in Jordan from the Beginning of the Epidemic Until the Week Number18, 2021 by Governorates

Governorate	Cumulative Incidence Rate per 100000 Population
Amman	7600.0
Irbid	5599.0
Zarqa	4640.0
Balqa	8301.0
Mafraq	4480.0
Karak	7333.0
Jarash	8286.0
Aqaba	9953.0
Ajloun	9823.0
Madaba	8424.0
Tafileh	11346.0
Maan	5464.0

Source: Ministry of health epidemiologic bulletins on COVID-19 epidemic, Jordan, 2021.

Example 2 of graphic analysis by place :

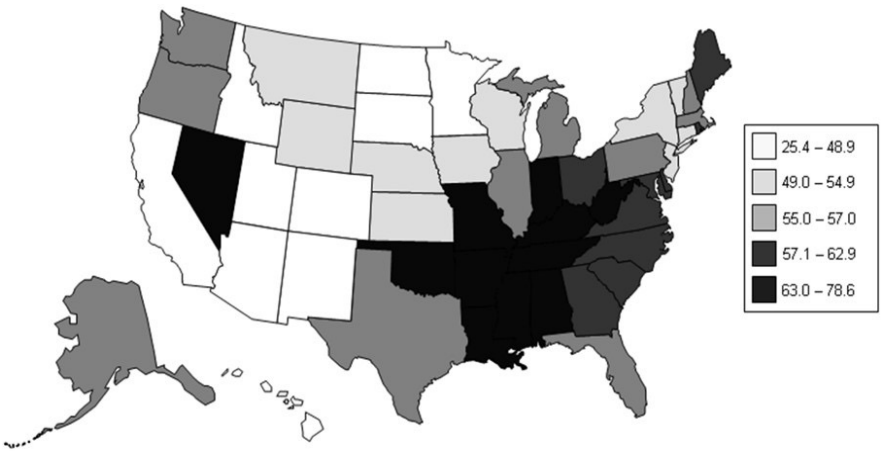
Cumulative PCR positivity rates for the weeks 5-18 of the second wave of the epidemic, 2021, Jordan by governorates



Source: Ministry of health epidemiologic bulletins on COVID-19 epidemic, Jordan, 2021.

Example 3 of mapping analysis by place:

Age-Adjusted Lung and Bronchus Cancer Mortality Rates (per 100,000 Population) by State — United States, 1998–2002



Source: National Cancer Institute [Internet] Bethesda: NCI [cited 2006 Mar 22] Surveillance Epidemiology and End Results (SEER). Available from: http://seer.cancer.gov/faststats/.

Analysis by person

Data analysis by personal characteristics include distribution of cases by age, sex, occupation, education, ethnicity, race, social class, risk factors, vaccination status etc. The most commonly collected and analyzed personal characteristics are age and sex.

Meaningful age categories for analysis depend on the disease of interest. Categories should be mutually exclusive and all-inclusive. Mutually exclusive means the end of one category cannot

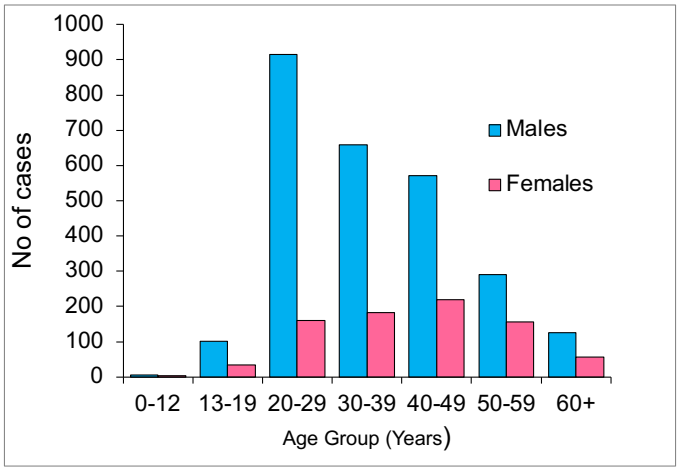
overlap with the beginning of the next category (e.g., 1–4 years and 5–9 years rather than 1–5 and 5–9).

All-inclusive means that the categories should include all possibilities, including the extremes of age (e.g., <1 year and ≥84 years) and unknowns. The standard age categories for childhood illnesses are usually <1 year and ages 1–4, 5–9, 10–14, 15–19, and ≥20 years.

The characteristic age distribution of a disease should be used in deciding the age categories, multiple narrow categories for the peak ages, and broader categories for the remainder. If the age distribution changes over time or differs geographically, the categories can be modified to accommodate those differences. To use data in the calculation of rates, the age categories must be consistent with the age categories available for the population at risk. For example, census data are usually published as <5 years, 5–9, 10–14, and so on in 5-year age groups. These denominators could not be used if the surveillance data had been categorized in different 5-year age groups (e.g., 1–5 years, 6–10, 11–15, and so forth).

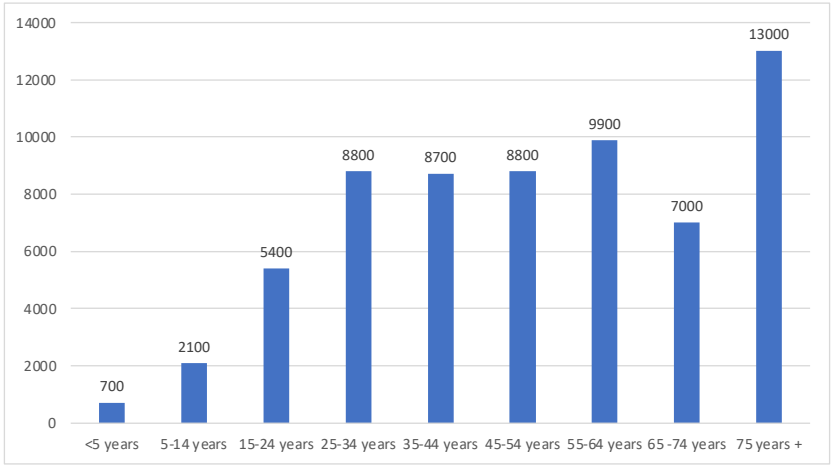
Example 1 of data analysis using personal characteristics (age and sex):

Reported Cases of Disease X by Age and Sex, Country X, Year Y (Hypothetical Data)



Example 2 of graphic analysis of data by age groups:

Cumulative Incidence Rate of COVID-19 per 100000 of Population in Jordan, Since the Beginning of the Epidemic till Week 18, 2021 by Age Groups



Source: Ministry of Health epidemiologic bulletins on COVID-19 epidemic, Jordan, 2021.

A surveillance system usually requires that each country should have a list of notifiable diseases to be reported on a regular basis to their concerned health authorities. The number of diseases in this list varies between countries depending on the types of diseases and their importance for the relevant country.

According to the WHO's International Health Regulations 2005, the following diseases or conditions should be notified by each country or territory to the WHO:

A. All cases of
   - Smallpox
   - Poliomyelitis (wild type)
   - Human influenza caused by a new subtype
   - Severe acute respiratory syndrome (SARS)

B. Unexpected or "impactful" cases of
- Cholera
- Pneumonic plague
- Yellow fever
- Viral hemorrhagic fever
- West Nile fever
- Diseases of regional concern, e.g., dengue, Rift Valley fever, meningococcal disease
- Any event that is a potential Public Health Emergency of International Concern (PHEIC)

C. After the detection of an event by the national surveillance system, the following criteria should be considered:
- Is the public health impact of the event serious?
- Is the event unusual or unexpected?
- Is there a significant risk of international spread?
- Is there a significant risk of international travel or trade restrictions?

If the event meets two or more of the above criteria, the event should be reported to the WHO.
As mentioned earlier, surveillance relies on two main methods of data collection; passive and active. Passive data collection is initiated by the healthcare provider, it relies on others to report disease, and usually is considered to be adequate for monitoring trends over time, place and person.

Active data collection requires assertive action, where the health-agency solicits information, and this is usually reserved for diseases of special interest.
How does a case get reported?
- Patient (or proxy, e.g., mother) must recognize that he/she is ill
- Patient must seek care
- Healthcare provider or laboratory must make an appropriate diagnosis
- Healthcare provider, laboratory, or institution must report the case to the concerned health authority

Reporting sources:
✓ Healthcare providers

- ✓ Laboratories
- ✓ Pharmacists
- ✓ Sick persons, family, neighbors
- ✓ Media reports
- ✓ Others

Information that should appear on the case report form includes:
- ✓ Identifying information
- ✓ Demographic information
- ✓ Clinical information
- ✓ Exposure/risk factor information
- ✓ Reporter information
- ✓ Contacts/others potentially exposed

Reporting sources:
- ✓ Healthcare providers
- ✓ Laboratories
- ✓ Pharmacists
- ✓ Sick persons, family, neighbors
- ✓ Media reports
- ✓ Others

Information that should appear on the case report form includes:
- ✓ Identifying information
- ✓ Demographic information
- ✓ Clinical information
- ✓ Exposure/risk factor information
- ✓ Reporter information
- ✓ Contacts/others potentially exposed

Organizing and displaying surveillance data

The organization and presentation of data is essential for several different reasons:

- Most data sets have too many records to summarize simply by looking at the line listing or the individual case report forms.
- Organizing and summarizing data condenses a larger amount of information into a smaller, more comprehensible amount of data.
- Organizing and summarizing data helps the investigator become familiar with the data, and helps identify problems with the data such as the number of records with missing values, illegal values (if 1=Male and 2=Female, what does 3 mean?), and outliers (is weight = 440 pounds [200 kg] real or data entry error?)
- Summarizing data helps the investigator identify patterns (mostly children or mostly adults?), trends (increasing or decreasing over time or seasonal?), relationships (were the people who ate pastries more likely to get sick than people who did not eat pastries?), and exceptions or outliers.
- Organizing and presenting data is an extremely useful way to communicate and share information with others.

Basic methods for organizing and displaying data

Data can be organized through creation of tables, graphs, and maps.

Many different types of tables can be used to organize and present data. Some of the most popular include:

- One variable table
- Two or three variable table
- For publication purposes, several simple tables, each of different variables, can be combined and presented in a single "composite" table

Example of a one-variable table, with percent and cumulative percent column:

Number of Brucellosis Cases by Age Groups, Country X, Year Y (Hypothetical Data)

Age Group	No. Cases	Percent	Cum. Pct.
15-24	214	19.4%	19.4%
25-34	396	35.9%	55.3%
35-44	282	25.5%	80.8%
45-54	168	15.2%	96.0%
55-64	44	4.0%	100.0%
Total	1,104	100.0%	

The cumulative percent or cumulative frequency, here, the percent of a given row is added together with the percentage of all previous rows. So, the 55.3% cumulative percent for the 25-34 category is the sum of 19.4% plus 35.9%. One can then say that a little more than half (55.3%) of brucellosis cases were less than 35 years of age. Thus, the median age (at 50%) is in the 25–34-year age category; probably close to the upper end of that range.

Example of two-variable table:

Number of Brucellosis Cases by Age Group and Test Results, Country X, Year Y (Hypothetical Data)

Age Group	Brucellosis +ve	Brucellosis -ve	Unknown	Total
15–24	214 (4.2%)	4,921	756	5,891
25–34	396 (9.6%)	3,749	570	4,715
35–44	282 (9.6%)	2,664	365	3,311
45–54	168 (7.6%)	2,057	237	2,462
55–64	44 (3.1%)	1,358	159	1,561
Total	1,104 (8.0%)	14,749	2,087	17,940

The two-variable table shows counts according to two variables simultaneously, with one variable along the rows and the other variable along the columns. It is also called a cross-tab or contingency table. The two variable table shows counts according to two variables simultaneously, with one variable along the rows and the other variable along the columns. When both variables of a two-variable table have only two categories each, then it is called a two-by-two table.

Example of two-by-two table:

Eating beef at a party	Ill	Not ill	Total
Ate beef			
Did not eat beef			
Total			

This table has two variables: Eating beef and disease status. Each of them has two categories.

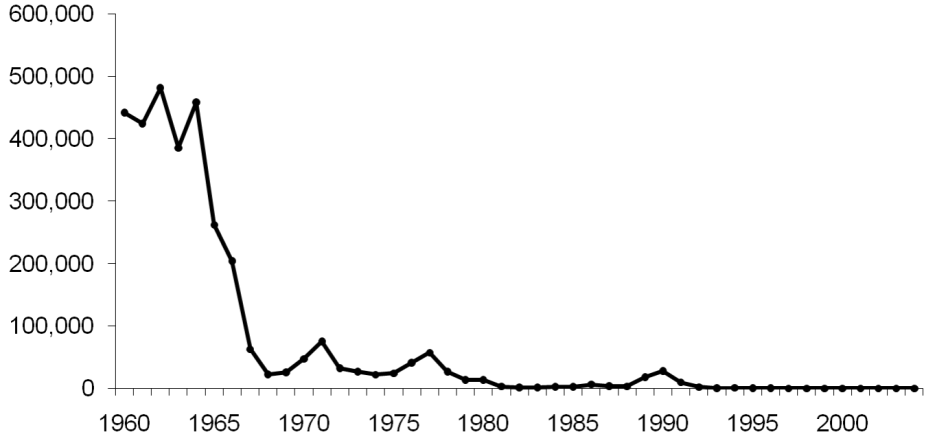
Graphs

An arithmetic-scale line graph is probably the most common type of graph used, particularly with surveillance data.

The features of an arithmetic-scale line graph are:

- It is a rectangular graph (Remember, X-axis is the horizontal axis along the bottom, Y-axis is the vertical axis along the side).
- X-axis has equal intervals, so a certain distance (such as 1 cm) represents the same number of years (e.g., 5 years), anywhere along the line.
- X-axis usually portrays time.
- Y-axis also has equal intervals.
- Y-axis portrays number, rate, or proportion, for example, number of cases of disease, or incidence rate per 100,000 population, or the percentage of the population with a particular characteristic (e.g., the percentage of the population who smoke) over time.
- Arithmetic-scale line graphs are used to portray data collected over time, i.e., to portray the time trend or pattern.
- Multiple diseases or other characteristics can be displayed on the same graph, so patterns can be compared.

Arithmetic-scale Line Graph



Now, let's focus on creating an arithmetic-scale line graph:

1. Draw x- and y-axes. Most visually appealing X: Y ratio (and best for PowerPoint, computer screens, and projection screens) is 5:3
2. X-axis: Match x-axis scale to intervals used during data collection, e.g., for what range of years do you have data?
3. Y-axis: Always start y-axis with zero.
4. Determine the range of values in the y-axis by identifying the largest value.
5. Select an interval size for y-axis that will provide enough intervals to illustrate the data in adequate detail.
6. Plot the data.
7. Create a title that describes the data, the location, and the time period.
8. Add notations, footnotes, and indicate the source of data.

Histogram

Characteristics of histogram:

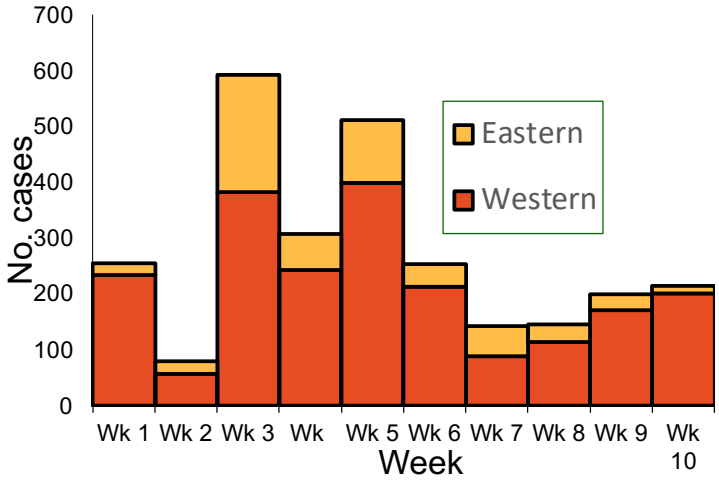
- Frequency distribution of quantitative data
- X-axis continuous, usually time (onset or diagnosis date)
- Y-axis represents frequency (number of cases)
- No spaces between adjacent columns, i.e., adjacent columns touch
- Easiest to interpret with equal class (x) intervals
- Column height is proportional to the number of observations at that interval
- "Epidemic curve" in outbreak investigations

How to make a histogram:

1. For continuous numeric data, assign equal width and non-overlapping categories.
2. Count the number of times each category appears.
3. Assign one bar for each category.
4. Make the bar height equal to the frequency for each category.
5. Include axis labels with units and a descriptive title.

Example of a simple histogram:

Diarrhea Cases by Week in Eastern and Western Governorates, Country X, Year Y (Hypothetical Data)



Example

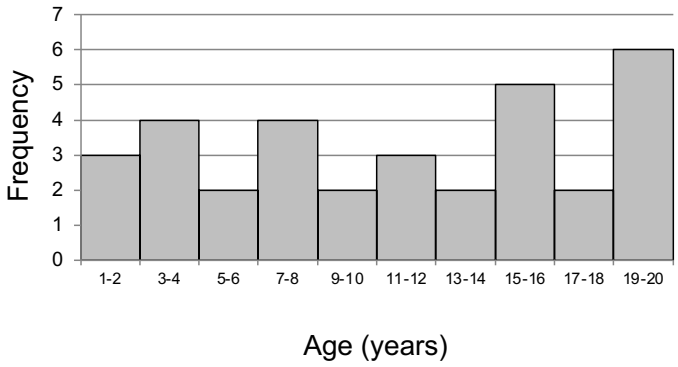
Study in Children Aged 1-20 Years Old

Ages: 1, 1, 2, 3, 3, 3, 3, 6, 6, 7, 7, 7, 7, 9, 10, 12, 12, 12, 13, 13, 15, 15, 15, 15, 16, 18, 18, 19, 19, 20, 20, 20, 20.

Possible categories:

- 1 category for each year (= 20 categories)
- 1 category for every 2 years (= 10 categories)
- 1 category for every 5 years (= 4 categories)

Example: Histogram with 2-year categories

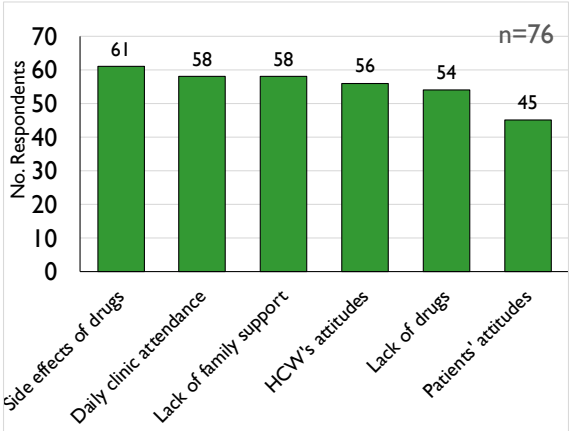


Bar charts

- Can be vertical or horizontal
- Use for variables with discrete, non-linear categories (qualitative), such as districts
- Bars have the same width
- Bars have space ("gaps") between them, since categories are not continuous
- Types – simple, grouped, stacked

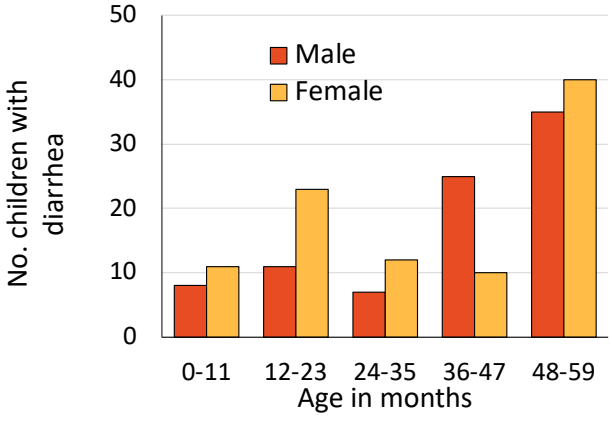
Example of simple bar chart:

Healthcare Workers Perceived Causes of Patients' Non-Adherence to TB Treatment, in a Certain Large Health Facility, 2011



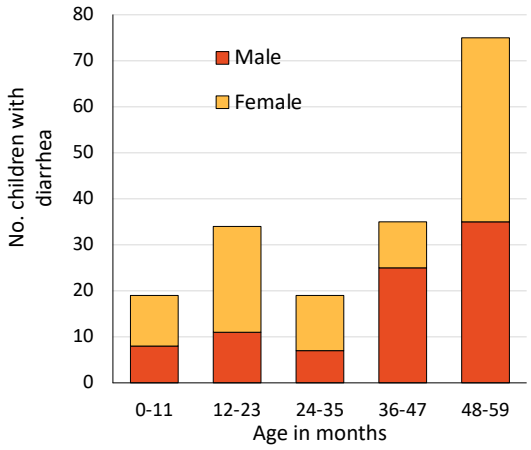
Example of a grouped bar chart:

Age and Gender Distribution of Children < 5 Years with Diarrhea, Country X, Year Y (Hypothetical Data)



Example of a stacked bar chart:

Age and Gender Distribution of Children < 5 Years with Diarrhea, Country X, Year Y ((Hypothetical Data)



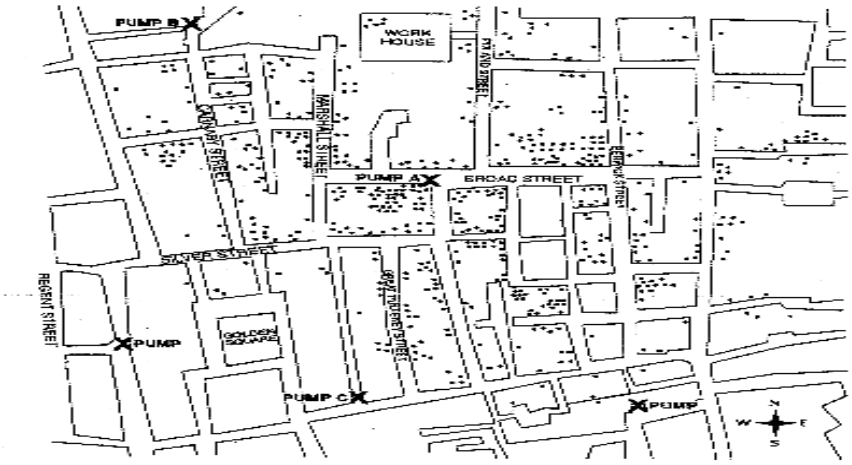
Maps

- Describe the geographic distribution of disease, services, etc.
  - Types
  - Spot or dot maps
  - Area maps
  - Others
- Symbols represent events, disease
- The size of the circle can be proportionate to the disease burden

Maps in epidemiology are visual representations of data distribution by geographical unit, providing an overview of geographical patterns. They are mainly used for explanatory purposes. Disease maps can be used to survey high risk areas, to help policy makers or to decide on resource allocation in specific areas.

Example of spot or dot map:

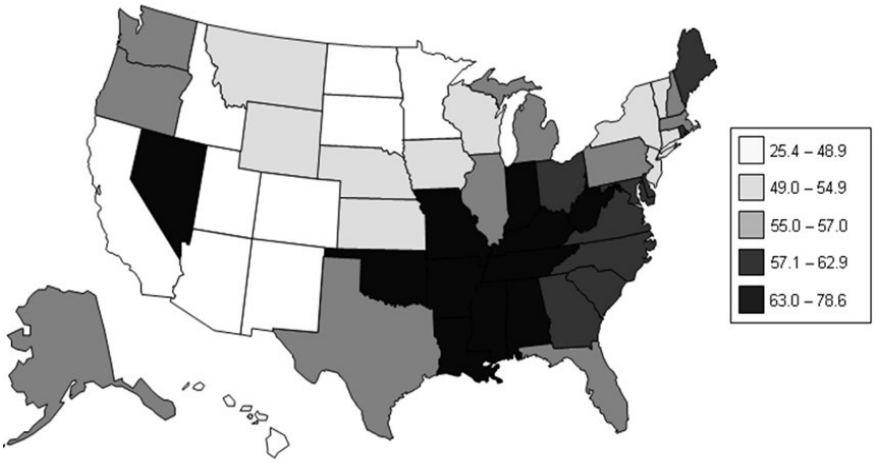
Spot Map of Deaths from Cholera in Golden Square Area, London, 1854 (Snow on Cholera). Snow Used this Spot Map to Distribute Households with Cases of Cholera in London.



Source: Snow J. Snow on cholera. London: Humphrey Milford: Oxford University Press; 1936.

Example of area map (shaded map):

Age-Adjusted Lung and Bronchus Cancer Mortality Rates (per 100,000 Population) by State — United States, 1998–2002



Source: National Cancer Institute [Internet] Bethesda: NCI [cited 2006 Mar 22] Surveillance Epidemiology and End Results (SEER). Available from: http://seer.cancer.gov/faststats/.

The advent of geographic information systems (GIS) allows more robust analysis of data by place and has moved spot and shaded maps to much more sophisticated applications. Using GIS is particularly effective when different types of information about a place are combined to identify or clarify geographic relationships.

Getting from data to reporting

To do that you need to:

- Know what data you have
- Decide which data to include in the report
- Plan on how to summarize and analyze data (analysis plan)
- Summarize each important variable by using measures of central location and frequency
- Display summary data effectively through using tables, graphs, and maps
- Package your findings into a good report

Goals of the analysis:

What are the questions you want to answer?

For example,

- How many cases of each notifiable disease?
- What is the time trend?
- How are cases distributed by place?
- What are their age and sex characteristics?
- How many reporting sources reported on time?

The analysis plan, also called the roadmap for analysis, should consider the following points:

- How to get from the data to the final report?
- The analysis plan describes:

    - What data are you using?
- Source(s) of data and variables
    - How will you look at and analyze the data?
- Summary measures, tables, figures
    - How will you compile the analyzed data into a report?
- The analysis plan is based on what question(s) you need to answer and what information you want to communicate.
- Consider the kind of data you have.
- Create table shells that are ready for analysis, except for the data.

In summary

- Data can be organized through the creation of tables, graphs, charts, and maps.
- The purpose of creating these visual displays of data is to verify and analyze the data, explore patterns and trends, and to communicate information to others.
- Always start with tables so you can see the data.
- Use appropriate titles and labels for the tables and the graphs.
- Tables can illustrate the number of people with particular characteristics and can provide valuable information about relationships between 2 variables.
- Line graphs are useful for showing patterns or trends over some variable, usually time.
- Histograms are most commonly used in epidemiology for epidemic curves (cases by time).
- Bar charts provide a visual display of data from a one-variable, but grouped bar charts can show 2 variables.
- Maps are useful for showing the geographic distribution of health events or conditions.

Interpretation of surveillance data

Systematic process for interpreting summarized data:

1. Explain epidemiologic and statistical measures in plain language.
2. Compare the observed data to the expected data.
3. Consider the quality of the data.
4. Consider possible explanations for an apparent increase in cases.
5. Make inferences about disease occurrence from summary data.

When we see an increase in the occurrence of certain disease in an area, we have to go through the following possibilities before jumping to the conclusion that there is a real increase in the incidence of the disease.

- Change in reporting procedures/change in surveillance system
- Change in case definition
- Improvements in diagnostic procedures
- Increased awareness
- Increased access to healthcare
- New physician, or clinic – may see more referred cases, may make diagnosis more often, or report more consistently
- Laboratory or diagnostic error
- Change in denominator
- A real increase in incidence

Example of interpretation:

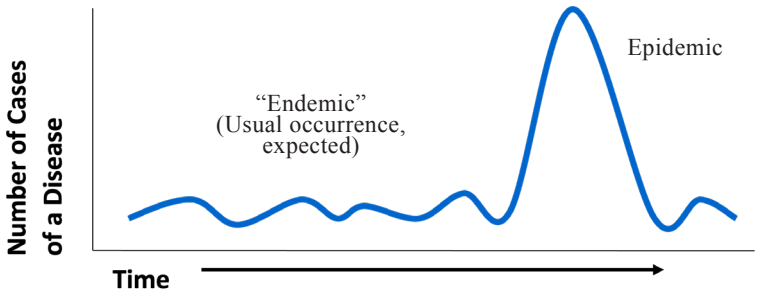
Cases of Measles by District, July, Country X, Year Y (Hypothetical Data)

District	Cases	Population	Rate per 1,000 Population
A	10	800	12.5
B	18	8,200	2.2
C	33	5,500	6.0
D	57	8,245	6.9
E	23	3,000	7.7

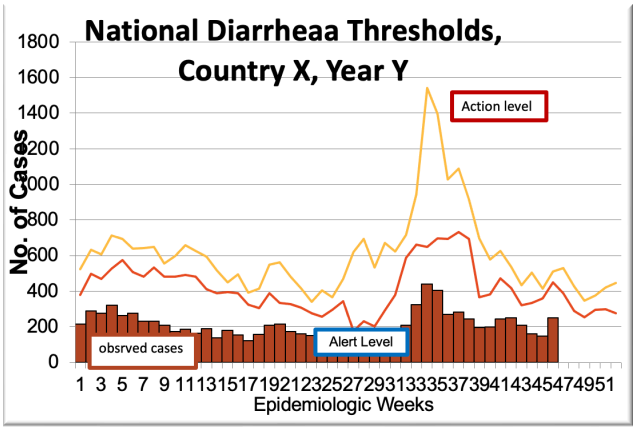
How would you explain the data on measles in the above table?
This table can be interpreted as follows: the highest incidence rate was in District A. (12.5 per 10000) and the lowest was in District B.

When you look to the curve below, which shows the occurrence of a disease over time in a certain geographical area, you can interpret that there was a usual and expected occurrence of the disease during the first part of the curve which represent endemicity of that disease in that area, and at a later time the occurrence started to increase sharply which indicates the occurrence of an outbreak or epidemic.

Endemic: The constant presence of an agent or health condition within a geographic area or population in a given period of time.
Epidemic: Occurrence of more cases of disease than expected in a given area or among a specific group of people over a particular period of time.

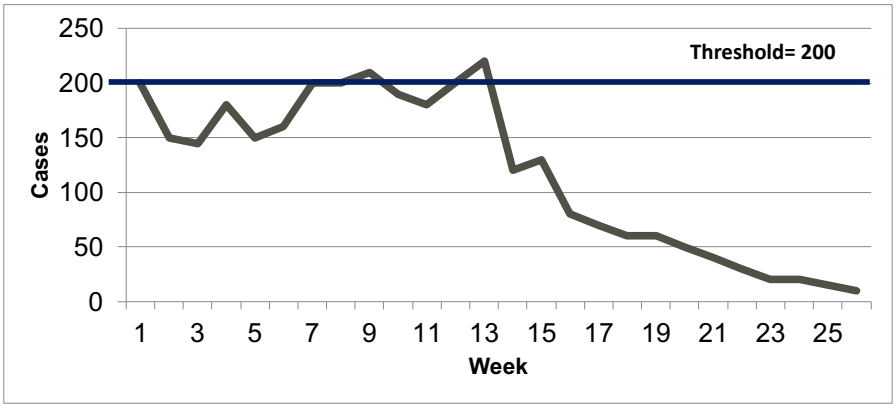


The graph below shows the importance of the threshold curve in surveillance. Threshold for a disease is a marker that indicates when action might be needed, it enables the surveillance officer to be alert at the beginning of any rise in the number of cases of the disease more than expected.

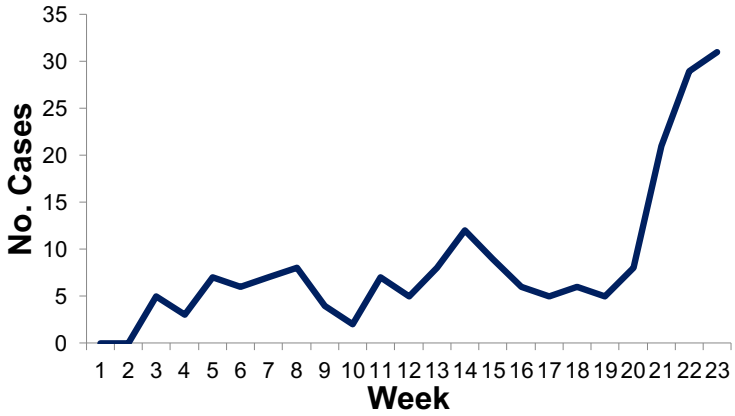


In the below graph, has influenza-like illness exceeded the threshold?
The answer is yes, the occurrence exceeded the threshold during weeks 10, 13 and 14.

Number of Cases of Influenza-like Illness Seen in Clinics, by Week, Country X, Year Y

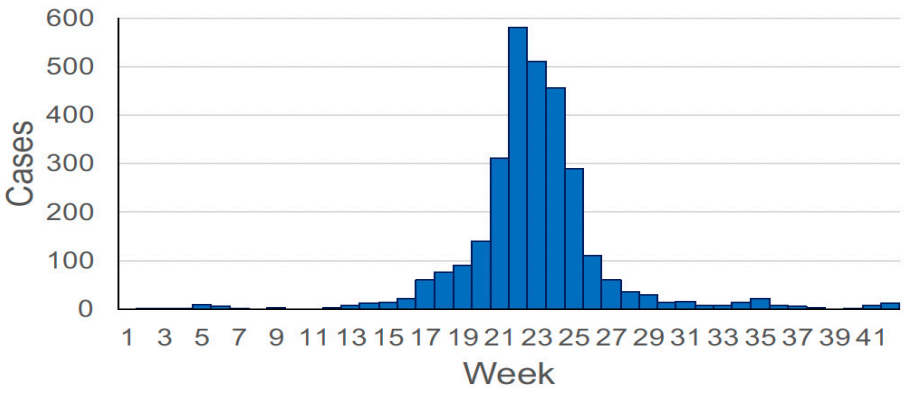


In the below line graph, what factors can explain an observed increase in the number of cases?
The answer: The increase in number of cases could be related to a real increase (start of an outbreak, or seasonality of the disease) or apparent increase (like improvement of reporting or change in reporting procedures). So, one should investigate the actual reasons.



How do you make inferences about disease occurrence from the epidemic curve below?
Answer: Depending on the epidemic curve below, most probably it is a point source epidemic.

Confirmed Cases of Hemorrhagic Fever, by Week, Country X, Year Y



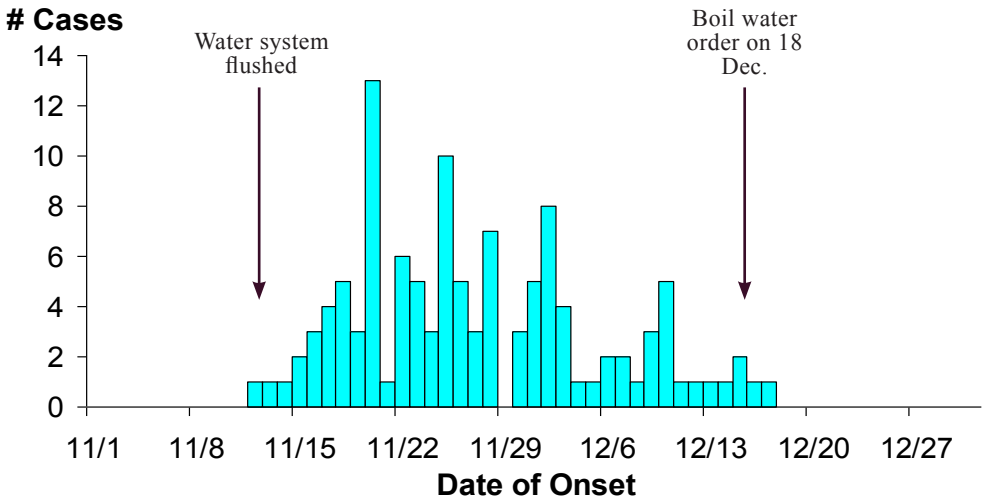
In the epidemic curve below, a boiling water order was issued by the health department on December 18. In order to know if this intervention has been effective or not, the investigator needs to continue surveillance for the occurrence of cases. He did that and found cases continued to occur despite the boil water order as in the second curve below.

The health department investigated further and found that people were forgetting to boil their water, or they didn't believe that the water needed to be boiled, or that they didn't realize that ice should be made with boiled water. As a result, public health officials concluded that the boil water order was not enough to stop the outbreak.

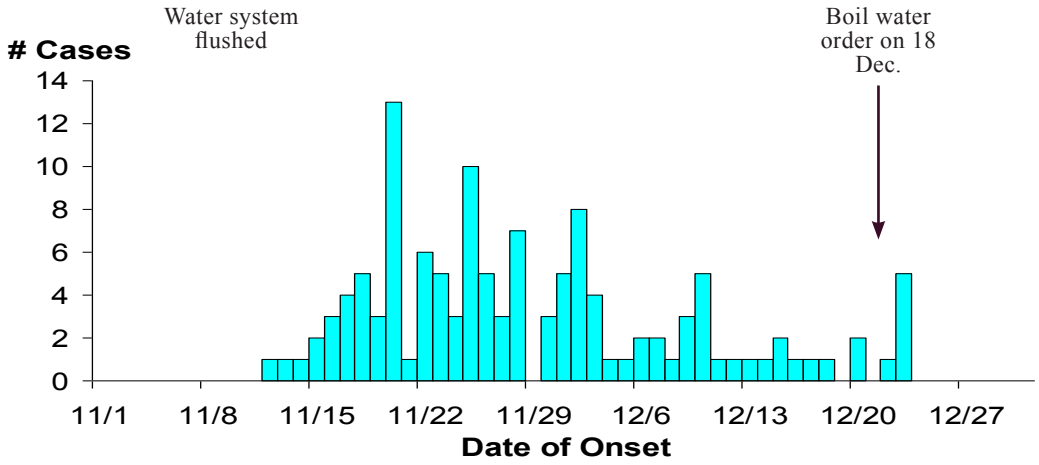
Therefore, they started chlorinating the water supply around December 23. After that, cases declined; ultimately the outbreak was interrupted. So these are examples of the link to public health action in public health surveillance.

Epidemic curve of diarrheal illness in a certain country

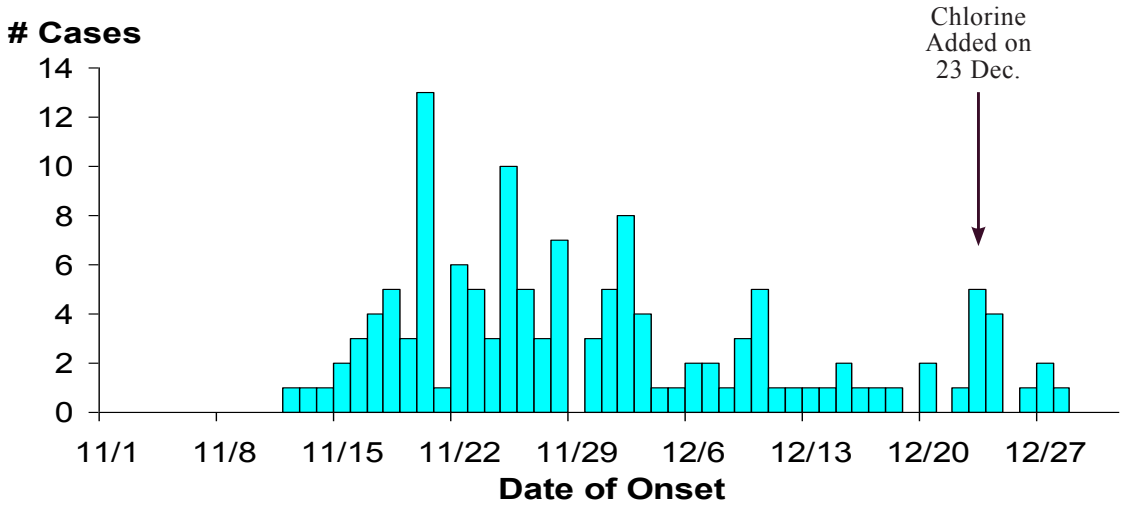
Epidemic Curve of Diarrheal Illness in a Certain Country



Epidemic Curve of Diarrheal Illness in a Certain Country after Boiling Water Order



Curve of Diarrheal Illness in a Certain Country after Chlorination of Water



In conclusion:

- Data must be interpreted to be useful.
- First, explain epidemiologic and statistical results in plain language.
- For surveillance and outbreak data, compare what you observe with what you expect.
- Consider the quality of the data.
- If observed disease occurrence is greater than expected, consider all explanations, including both true increases and artificiality.
- Use summary data to identify patterns.

Dissemination of surveillance data

For appropriate public health action to be taken, surveillance data should be disseminated to the following parties:

- Those responsible for taking action need to know and need to know in a timely fashion.
- Public health officials in neighboring areas might want to know, so they can be vigilant for cases that might occur across geographic lines.
- Other government authorities (mayor, governor, etc.) usually want to know, particularly if the situation is serious.
- Those who sent in the case reports appreciate being "in the loop", because then they know that their efforts to report are not a waste of time and that the reports they sent in are not just being filed in some dusty file room.
- Clinicians like to know what diseases are prevalent at any given time, since it may help them with a differential diagnosis.
- The public is sometimes interested.

Summary reports can be disseminated in a number of ways, depending on the audience and how timely the information needs to be. For example, your boss or the governor may want to be briefed in person, particularly about a serious case or epidemic. Other methods include newsletters, reports posted to a website, press releases, and articles published in journals.

Monitoring and evaluation of public health surveillance

The evaluation of surveillance systems should promote the best use of public health resources by ensuring that only important problems are under surveillance and that surveillance systems operate efficiently. Insofar as possible, the evaluation of surveillance systems should include recommendations for improving quality and efficiency, e.g., eliminating unnecessary duplication. Most importantly, an evaluation should assess whether a system is serving a useful public health function and is meeting the system's objectives [74].

Surveillance for a disease or other health-related problem should be evaluated periodically to ensure that it is serving a useful public health function and is meeting its objectives. Such an evaluation: (1) identifies elements of surveillance that should be enhanced to improve its attributes, (2) assesses how surveillance findings affect control efforts, and (3) improves the quality of data and interpretations provided by surveillance. Although the aspects of surveillance that are emphasized in an evaluation can differ, depending on the purpose and objectives of surveillance, the evaluation's overall scope and approach should be similar for any health-related problem. The evaluation usually begins by identifying and interviewing key stakeholders and by collecting background documents, forms, and reports. The evaluation should address the purpose of surveillance, objectives, and mechanics of conducting surveillance; the resources needed to conduct surveillance; the usefulness of surveillance; and the presence or absence of the characteristics or qualities of optimal surveillance. The outcome of the evaluation should provide recommendations for improvement.

The International Health Regulations emphasize the commitment of Member States to the goal of global health security. This will require all Member States to maintain a functional and effective surveillance and response system that is able to detect, investigate and respond to public health emergencies of national and international concern [75, 76].

Monitoring = routine and continuous tracking of planned surveillance activities.
Evaluation = periodic (e.g., annual) assessment of whether surveillance and response objectives have been achieved.

Indicators and targets:
Indicator = Statement to measure the achievement of an activity objective.
Example: Is reporting done on time?
Target = Desired level of achievement
Example: 80% of monthly reports have been sent on time to the national level.
Monitoring and evaluation should be used to develop strategies for improvement.
What indicators could monitor performance?

Health Facility Reporting Practices	District Reporting Practices	Disease/PH Action	Laboratory Practices
Proportion of health facilities submitting reports on time to district	Proportion of districts submitting reports on time to next higher level	Proportion of outbreaks notified to the next higher level within 2 days	Proportion of districts that report laboratory data
Proportion of health facilities submitting complete reports to district	Proportion of districts using line graphs or histograms	Proportion of outbreak reports that analyzed case-based data	Proportion of district laboratories visited

cont'd

Proportion of health facilities submitting zero reports for epidemic-prone diseases	Number of epidemics missed by district (detected by national level)	Proportion of confirmed outbreaks with recommended PH response	Proportion of provinces reporting analyzed laboratory data to national level
		Case-fatality rate for each epidemic-prone disease reported	

Monitoring and Quality Control:

- Good monitoring helps health staff to perform their best.
- Monitoring is a vital component of any surveillance program.
- Monitor all surveillance activities using standard performance indicators.
- District surveillance office monitor indicators of reporting on a regular basis.
- Results comes from the monitoring will help to take action to improve surveillance and response by correction of the deviation or mistakes.
- Monitoring is essential to maintain quality.
- Established indicators are available for monitoring the performance of reporting sites.

Main indicators or characteristics used for evaluation of a surveillance system

1. **Simplicity**
   The simplicity of a public health surveillance system refers to both its structure and ease of operation. Surveillance systems should be as simple as possible while still meeting their objectives.
2. **Flexibility**
   A flexible public health surveillance system can adapt to changing information needs or operating conditions with little additional time, personnel, or allocated funds. Flexible systems can accommodate, for example, new health-related events, changes in case definitions or technology, and variations in funding or reporting sources.
3. **Acceptability**
   Acceptability reflects the willingness of persons and organizations to participate in the surveillance system.
4. **Sensitivity**
   The sensitivity of a surveillance system can be considered on two levels. First, at the level of case reporting, sensitivity refers to the proportion of cases of a disease (or other health-related event) detected by the surveillance system.
   Second, sensitivity can refer to the ability to detect outbreaks, including the ability to monitor changes in the number of cases over time.
5. **Predictive Value Positive**
   Predictive value positive (PVP) is the proportion of reported cases that actually have the health-related event under surveillance.
6. **Representativeness**
   A public health surveillance system that accurately describes the occurrence of a health-related event over time and its distribution in the population by place and person.

7. **Timeliness**
   Timeliness reflects the speed between steps in a public health surveillance system.
   The amount of time between the onset of a health-related event and the reporting of that event to the public health agency responsible for instituting control and preventive measures.
8. **Stability**
   Stability refers to the reliability (i.e., the ability to collect, manage, and provide data properly without failure) and availability (the ability to be operational when it is needed) of the public health surveillance system.
9. **Data Quality**
   Data quality reflects the completeness and validity of the data recorded in the public health surveillance system. Example: Examining the percentage of "unknown" or "blank" responses to items on surveillance forms is a straightforward and easy measure of data quality. Data of high quality will have low percentages of such responses.
10. **Usefulness**
    Usefulness refers to whether surveillance contributes to prevention and control of a health-related problem. Note that usefulness can include improved understanding of the public health implications of the health problem. Usefulness is typically assessed by determining whether surveillance meets its objectives. For example, if the primary objective of surveillance is to identify individual cases of disease to facilitate timely and effective control measures, does surveillance permit timely and accurate identification, diagnosis, treatment, or other handling of contacts when appropriate?

Chapter 10

Investigation and Control of Epidemic

An epidemic/outbreak is defined as the occurrence of more cases of a disease than expected for a particular group of people in a particular place and time. Some people use the terms "outbreak" and "epidemic" interchangeably. Others use the term "epidemic" for a more widespread event and "outbreak" for a localized event.

When describing an epidemic, the time period, geographical region and particulars of the population in which the cases occur must be specified. The number of cases needed to define an epidemic varies according to the agent, the size, type and susceptibility of population exposed, and the time and place of occurrence.

The identification of an epidemic also depends on the usual frequency of the disease in the area among the specified population during the same season of the year. A very small number of cases of a disease not previously recognized in an area, but associated in time and place, may be sufficient to constitute an epidemic. For example, the first report on the syndrome that became known as AIDS concerned only four cases of Pneumocystis carinii pneumonia in young homosexual men previously this disease had been seen only in patients with compromised immune systems [77].

The dynamic of an epidemic is determined by the characteristics of its agent and its pattern of transmission, and by the susceptibility of its human hosts.

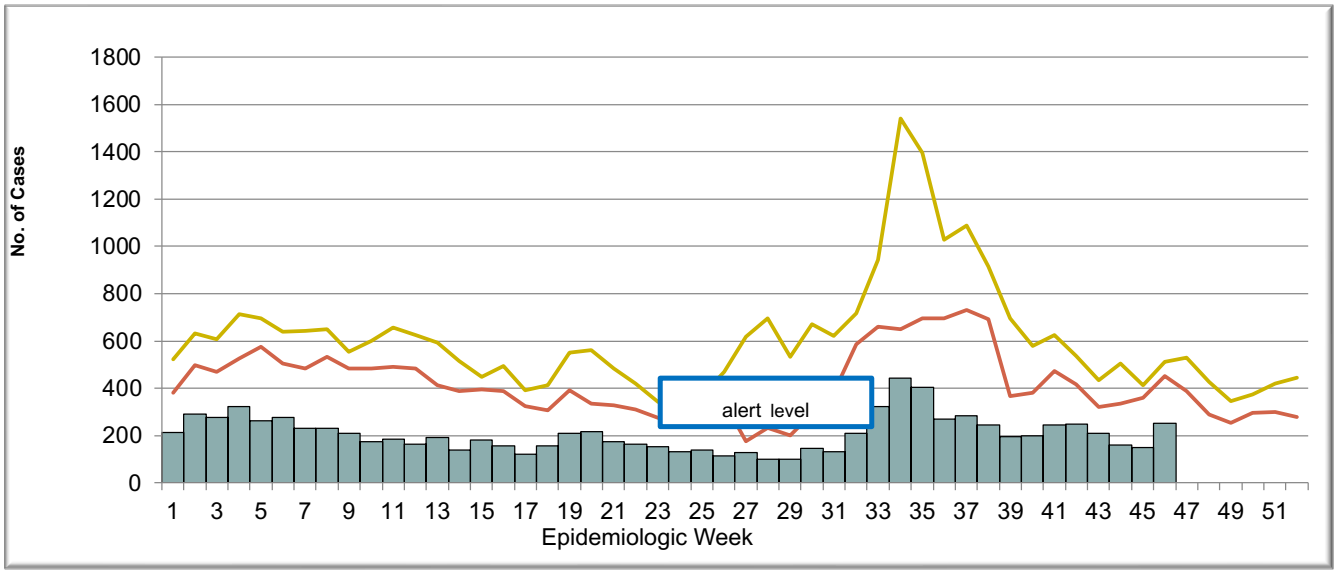
The investigation of an outbreak/epidemic consists of a set of procedures used to identify the cause responsible for the disease, the people affected, the circumstances and mode of spread of the disease, and other relevant factors involved in propagating the epidemic, and to take effective actions to contain and prevent the spread of the disease. The purpose of investigating a communicable disease epidemic is to identify its cause and the best means to control it. This requires detailed and systematic epidemiological work, in the following sequential or simultaneous steps:

- Undertaking preliminary investigation
- Identifying and notifying cases
- Collecting and analyzing data
- Managing and controlling
- Disseminating findings and follow-up

The Threshold in Epidemics

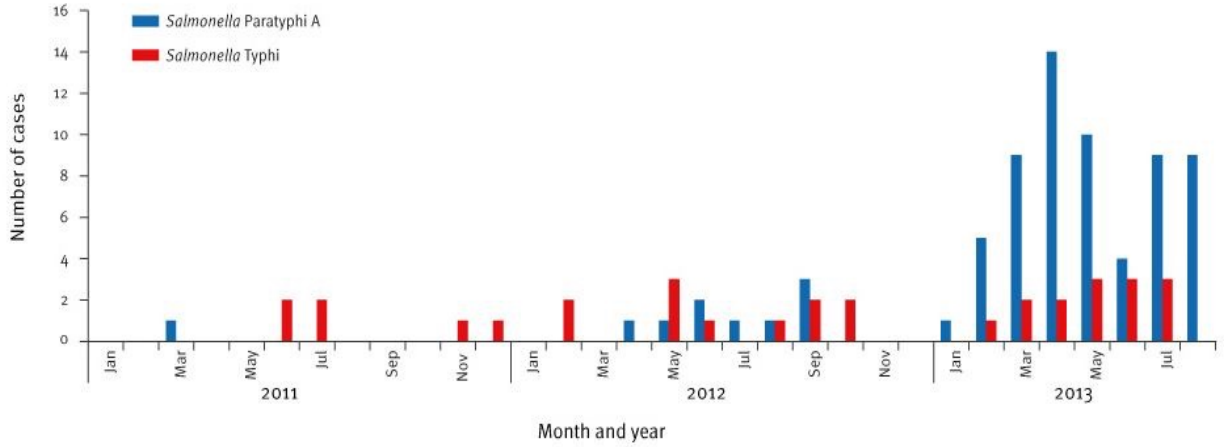
An alert threshold (or epidemic threshold) indicates the level of incidence above which a disease requires an urgent response. Each disease has a specific threshold that depends on its infectiousness, other determinants of transmission, and the degree to which it is locally endemic. Alert threshold means a level beyond which there is a risk to human health from brief exposure, and at which immediate corresponding steps shall be taken for informing and warning the population of affected areas.

The chart below gives an example about the threshold or alert level for diarrhea in a certain country, where the middle line represents the threshold level, the upper line represents the action or intervention to be taken when the observed cases exceed the threshold line. So, when the occurrence of diarrhea starts to go above the threshold level (red line in the middle) alert should be increased and action might be needed to investigate and to intervene. In this chart, the observed occurrence of diarrhea is still below the alert level (below the threshold level).



Sometimes, a threshold is simply a constant, such as 1 case of cholera or, in the below graph, 5 cases of salmonella. By looking at the graph, the cases of salmonella paratyphi A increase above the threshold during March-May and July-August. While cases of salmonella typhi remained below the threshold.

Salmonella typhi and salmonella paratyphi A occurrence in Country X, years 2011-2013
How Potential Outbreaks Are Identified



- Regular reviews of surveillance data are one way.
- Clinicians may notice and report a single unusual case or an increase in patients with a specific disease. For example, an outbreak of meningitis was identified after physicians reported the disease in several pediatric patients.

- Health clinics also take calls and reports from patients or other citizens who are concerned about an illness. Personnel can investigate these reports, particularly if they get multiple reports from different sources. An example of this type of report might be an outbreak associated with a school.
- The media may also pick up on an outbreak or disease problem as a news item that health officials were otherwise not aware of.

When should you investigate?
This depends on:
- Number of cases exceeded the threshold
- The severity of the illness
- Potential of the disease to spread
- Availability of prevention and control measures
- Availability of resources
- Sometimes you investigate for political considerations or for public relations

Why conduct an investigation?
- The most pressing reason for starting an outbreak investigation is to prevent and control the disease. Sometimes we know the disease and how it is spread (e.g., measles), and we can take immediate action (vaccination).
- Other times we need to conduct an investigation to identify the risk factors or sources, and then develop and implement measures to control these risk factors. For example, one epidemiologic investigation of hepatitis A showed that the consumption of strawberries was associated with the illness. Without knowing how the strawberries became contaminated, immediate measures could still be taken to prevent further disease by recalling the strawberries and making the public aware of the need to avoid eating strawberries.
- An investigation is also done to characterize the problem, particularly for a new disease – what is the clinical spectrum of illness? Which people are most at risk? How does the disease spread?
- Outbreak investigations may provide new research insights into a disease. As an example, the investigation of Zika virus infection in the United States in 2016 was focusing on answering questions such as, are we sure that Zika virus infection causes microcephaly? Does it cause other health problems? When during pregnancy is an infection most likely to cause microcephaly? How long does the virus persist in the body? Can it be transmitted sexually from man to woman? For how long?
- Sometimes, a field investigation may be conducted because of political pressure to investigate a problem, or to gather evidence for legal proceedings. For example, during a 2008-2009 outbreak of salmonella typhimurium in the United States, authorities learned that executives at a peanut plant may have knowingly shipped contaminated peanut butter products. Findings from the epidemiologic investigation were used to change food safety regulations and to prosecute peanut executives.
- Outbreak investigations provide opportunities for training of health department staff in methods of public health investigation and emergency response. While the costs do not justify conducting an outbreak investigation solely for the purposes of training, it is good practice to include individuals in the outbreak team who can learn from the experience.
- The balance between control measures and further investigation depends on how much is known

about the cause, the source, and the mode of transmission of the agent [78]. The table below shows how to prioritize investigation versus control measures according to the following points:

- If the source and causative agent are unknown, the investigation is of highest priority– we need to know how people are getting sick.
- When we know the source, but the agent is still unknown and both investigation and control measures are a high priority.
- If the source is unknown, but we know the causative agent, investigation is a higher priority so we can discover the source/mode of transmission.
- Once transmission is known and we know the causative agent, the highest priority is control.
- The major rule is to implement control measures as soon as possible.

Causative Agent	Source/Mode of Transmission: *Known*	Source/Mode of Transmission: *Unknown*
Known	Investigation + Control +++	Investigation +++ Control +
Unknown	Investigation +++ Control +++	Investigation +++ Control +

+++ Higher Priority
+ Lower Priority

Source: Goodman RA, Buehler JW, Koplan JP. The epidemiologic field investigation: science and judgment in public health practice. Am J Epidemiol 1990; 132:9–16.

In the table above, let's consider 4 different situations. Which box and what should you prioritize?

1. Cholera cases among persons using a well previously contaminated by an adjacent latrine — Known agent, known source (upper left), so priority is control.
2. An unknown disease with unknown source — Unknown agent, unknown source (lower right, so priority is investigation.
3. Gastroenteritis outbreak (agent not known) is associated with food served at a restaurant — Unknown agent, known source (lower left), so control measures can be implemented, but you also want to investigate to identify the cause.
4. Anthrax cases without a known source— Known agent, but unknown source (upper right), need to investigate to identify source or mode of transmission, then implement appropriate control measures.

Exceptions to the rule: If the source is suspected and still a threat to public health, take immediate control measures.

Objectives of a field investigation:

- Identify the agent, source, and/or mode of transmission.
- Characterize the extent of the outbreak, e.g., those who have been affected, who are at risk.
- Identify exposures or risk factors that increase the risk of the disease.
- Develop and implement control and prevention measures.

Once we have identified an outbreak and it meets the criteria for investigation, we have to start an investigation following the phases and steps of an outbreak investigation.

Phases of an Outbreak Investigation

An outbreak investigation has 3 general phases – the descriptive or preliminary phase, the explanatory phase, and the response phase. Each of these phases has its own steps.

The descriptive or preliminary phase steps [79]:

1. Prepare for fieldwork.
2. Confirm the existence of an outbreak.
3. Verify the diagnosis.
4. Construct a case definition.
5. Find cases systematically and record information.
6. Perform descriptive epidemiology.

During this phase, the health department staff should quickly become "experts" in the disease, learning as much as possible about the disease, its symptoms, and its possible causes and routes of transmission.

The first three steps are preparing for field work, confirming the existence of an outbreak, and verifying what it is an outbreak of. Depending on the circumstances and whether you are based in a district office or in the national office, the order can vary. They can be done simultaneously or in any order.

Once you know you have a real outbreak, the next step is to develop a case definition that will be used for additional case finding. The case definition provides the criteria for considering an individual a case.

Step 5 is to identify as many cases as possible and to compile the information into a single database. Then, once the data is compiled, you can summarize the characteristics of the cases using descriptive epidemiology.

Remember, even though this is the descriptive phase, at any time, if an obvious source of contamination is identified, immediate control measures should be taken to prevent further spread of disease. For example, if you don't have laboratory confirmation of a disease, but a water supply looks to be the likely source of contamination, you can immediately recommend using a different source, boiling water, or drinking bottled water.

The explanatory phase consists of the following 4 steps of the investigation:

7. Develop hypotheses.
8. Evaluate or test hypotheses epidemiologically.
9. Reconcile epidemiology with laboratory and environmental findings.
10. Conduct additional studies as necessary.

The explanatory phase is called explanatory because in this phase, we're trying to explain or determine scientifically what caused the outbreak. In this phase, we use the information we collected in the descriptive phase to develop hypotheses about the cause of the outbreak.

After developing hypotheses, we need to evaluate those hypotheses. To evaluate hypotheses we

sometimes simply assess the information we already have, or we can conduct some sort of analytic study.

Until now, we have been focusing on the epidemiologic approach. But often, an investigation includes laboratory and environmental studies as well. Do they all point to the same conclusions? Sometimes, additional studies are necessary.

The response phase consists of 3 steps:
11. Implement and evaluate prevention and control measures.
12. Initiate or maintain surveillance.
13. Communicate findings.

Now that we have evaluated our hypotheses and reached conclusions about the cause and source of the outbreak, we can implement control measures. But we have already discussed that we should implement control as early as possible, particularly if the source and/or mode of transmission were known.

Once we have implemented control measures, we want to ensure that those control measures work. The best way to do so is to continue surveillance. If the condition is one that is not part of routine surveillance systems, then we need to establish a new surveillance system to determine whether our control measures are working. Finally, the results of the investigation must be communicated.

Components of an outbreak field investigation
Many outbreak investigations include 3 components – an epidemiologic component, a laboratory component, and an environmental assessment. We will focus on the epidemiologic component, but remember that the other activities may be taking place at the same time, and should be done in a coordinated way.

Step I: Preparation for fieldwork
Some people have described epidemiology as a "team sport." Rarely does a single epidemiologist conduct an investigation by him/herself, without any collaboration. Epidemiology is just one of three components of the investigation. So one of the first tasks is to assemble the team that will work on the investigation.

Tasks to prepare for fieldwork:
- Form a team.
- Learn about the disease.
- Make necessary administrative, personnel, and logistical arrangements.
- Coordinate with partner agencies and local contacts.

The team should consist of a mix of individuals providing expertise on the disease as well as other areas important to an investigation. Potential team members include:
- The team leader who should have experience in outbreak investigation and public health epidemiology. Depending on the size and organization of the health department, this could be the local health director, public health nurse, epidemiologist, or environmental health specialist.

The team leader will then outline the plan for investigation, and assign roles and responsibilities to the team members.

- The epidemiologist, who is often the team leader, should have expertise in various aspects of outbreak investigations, from choosing the study design and questionnaire development to creating a database and conducting data analysis.
- A lab technician or microbiologist, usually at a state or regional public health laboratory, is important in verifying the diagnosis and sub-typing pathogens to help refine the case definition.
- Environmental health specialists (EHS) or sanitarians are important in preventing foodborne outbreaks from occurring through routine inspections of food preparation facilities, health education, and training of foodhandlers. Once an outbreak occurs, the EHS are able to identify food safety issues that may have contributed to the outbreak, such as time and temperature violations, and can assist in proper collection of food and environmental samples. An environmental health specialist may also provide guidance on food safety regulations and engineering during the outbreak. They may also be involved in water quality control and in regulating, preventing, and testing contamination of food, water, or other parts of the environment.
- Clinicians may be necessary to administer vaccines or prophylactic therapies, or assist in the collection of clinical specimens from case-patients. Clinicians may come from the health department or the local medical community.
- Zoonotic outbreaks may involve veterinarians and other scientists, such as entomologists, to provide expertise in animal reservoirs or vectors.
- Interviewers will be used to collect data, either in person or by telephone. Interviewers can come from the ranks of health department personnel, including clerical support staff. During large outbreaks, state or federal personnel or students in medicine or public health may be recruited to conduct interviews.
- Regulators from concerned institutions (as a person representing the governor, police, etc.) may be included on the outbreak team to help facilitate identification of the source of contaminated food items or water and develop prevention strategies through enforcement of safety regulations.

Before going out into the field, the team members need to learn about the disease, the outbreak, they need to learn about the area and its people and culture.They may need to contact others who have experience conducting similar investigations, perhaps they can share their questionnaires and lessons learned.

Numerous administrative and personnel arrangements must be made like transport, supplies (laboratory supplies, and clinical supplies), food and the team might need to bring personal protective equipment. Finally, particularly if the team is coming from outside the area, they need to know who to meet locally. Local public health workers or clinicians may have been involved in the initial suspicion that there was an outbreak, may have reported the suspected outbreak, may be leading a local investigation, may be contacted regarding potential cases, or may be asked by an outbreak team leader to assist in an investigation. They also know the conditions on the ground, such as who the community leaders are and if there is community resistance.

Step II: Confirm the existence of an outbreak

The next step is the need to ensure whether the increase in reported cases truly represents an outbreak. Remember the definition of outbreak -- the occurrence of more cases of a disease than expected for

a particular group of people in a particular place and time.

How do we establish that an outbreak is occurring?

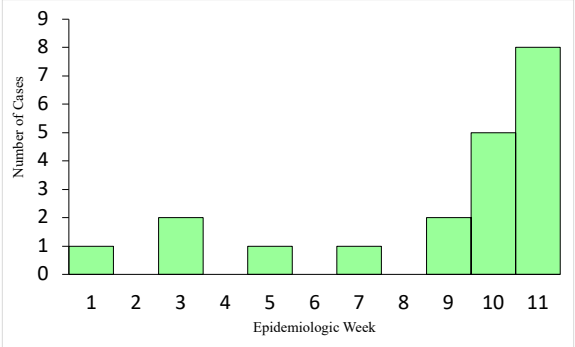
- The initial reports or complaints must be considered.
- Surveillance data or reports from clinicians or laboratory technicians should be reviewed.
- If the initial complaint was from a patient, concerned community member, or the media, the data would need to be obtained and reviewed.
- Then, confirmation that the cases have the same disease should be made. Sometimes signs or symptoms may overlap, but patients do not have the same disease.
- After ensuring that the cases are of the same disease, investigators must confirm that the number of cases exceeds what would be expected for the particular population over the specific time period.
- If there does appear to be an increase in cases, keep in mind that there could be another reason for the apparent increase. There could be a change in case definition, reporting procedures, or diagnostic tools. There could also be an increase in reporting because of media attention to an issue or because a new clinician is more likely to report a disease than his or her predecessor. Lastly, there could be a misdiagnosis or laboratory error causing the apparent increase.

The increase in disease occurrence could result from the following:

- True outbreak/epidemic
- Seasonal pattern
- Sudden increase in the size of population
- Change in reporting procedures or surveillance system
- Change in case definition
- Increase or improvement in laboratory testing/diagnostic procedure
- Increased awareness of the public and/or providers about the disease
- Increased access of the population to healthcare
- New healthcare provider, reporter
- Laboratory or diagnostic error
- Other reasons

In summary, remember, not all increases in cases represent outbreaks, but you cannot assume that it is not an outbreak unless you go through the above mentioned possibilities.

Example: In the chart below we see an increase in cases of dysentry during weeks 10 and 11. Is it an outbreak? It looks suspicious, but not certain. It would be nice to have historic data to see whether the increase in cases is true or simply the usual seasonal increase.



Number of Reported Cases of Dysentery by Epidemiologic Week, City X, Year Y

Step III: Verify the diagnosis

Once we have confirmed a true increase in cases, and an outbreak has been established, we need to confirm what disease we are dealing with and to verify which agents could be causing the outbreak.

If the laboratory has already confirmed the diagnosis, we do not need to test or have a positive laboratory result from every patient, but it is nice to have laboratory confirmation from a few.

If we do not have laboratory confirmation, we need to do some evaluation clues. Is the clinical presentation characteristic of a particular disease? What are the predominant signs and symptoms? Can they help us differentiate the diseases that are sometimes confused with one another? Even if we do not have a definitive laboratory diagnosis, do the clinical laboratory results (hematocrit, white blood cell count, urinalysis, liver function tests, etc.) help? Sometimes the course of the illness, including duration of symptoms, varies by agent and can help with the diagnosis.

For a definitive diagnosis, you will want laboratory confirmation. To find the agent, testing of clinical specimens such as stool, blood, or vomitus is invaluable. Pathogen identification will help identify the potential incubation period, and the incubation period will pinpoint at what time the exposure took place.

However, because lab results can take time, we do not need to wait for laboratory diagnosis to proceed; sometimes the investigation must move forward before a definitive diagnosis is reached. Once an agent is identified, the laboratory may be able to conduct further testing, to "fingerprint" the agent and verify that all case-patients are outbreak-related.

On the other hand, if the laboratory results are negative, do you cancel your investigation? There are several possibilities to explain why the results may be negative. Take into consideration possibilities such as that the illness could be due to an agent that we didn't test for, or that the specimen had been mishandled. Laboratory confirmation is an important clue but it's not the only clue.

Step IV: Construct a case definition

Once you have established that there is truly a notable increase in reports of the same diagnosis, then you form a case definition. The case definition of the outbreak is a bit different from the case definition used for surveillance.

Outbreak case definitions are commonly constructed using the following elements:

- Clinical criteria are based on clinical symptoms and signs that are characteristic of the disease of interest.
- Laboratory tests, both microbiological and clinical.
- Epidemiological criteria (especially for outbreaks) that specify the time period, place, person.

When developing a case definition, you must be as precise as possible in the language that is used, because other people will use the case definition, and you want everyone to make the same decision in the same way about whether a patient meets the case definition or not. The criteria should be: Objective, simple, accurate, practical and measurable where possible.

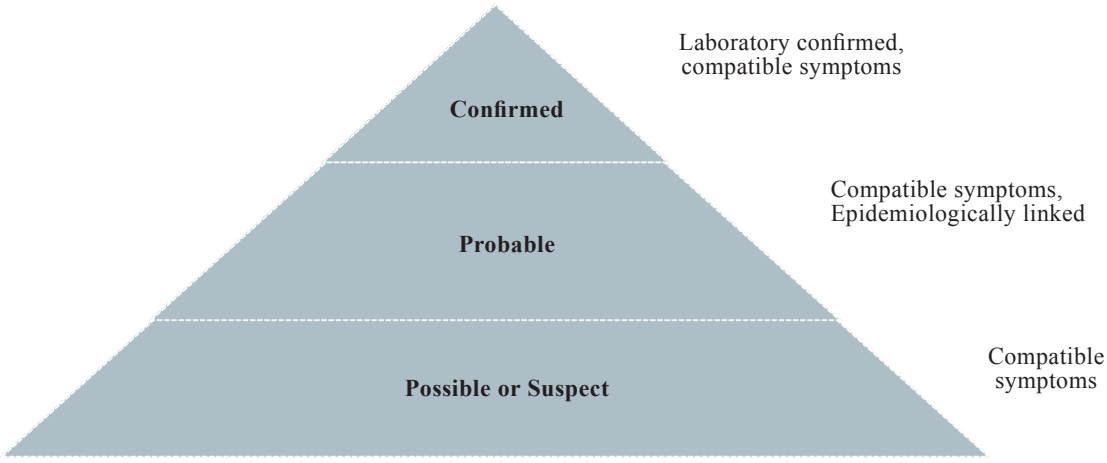
Finally, in most situations, do not include the suspected exposure in your case definition, if that

exposure is what you want to evaluate objectively. For example, if our hypothesis was that the mosquitoes causing disease were breeding in the floodwaters of a river, we would not initially want to limit the case definition to people who lived near the river, as we would be excluding the possibility that other exposures were involved before we had enough information.

Sometimes it's not clear whether a person should be a case or not. Because of this, just as in surveillance cases are classified by level of certainty, or by how much information is available. Individuals can change classification as more information becomes available.

- The confirmed case: It is the strongest level of certainty is that this is someone who has had a positive laboratory test for the disease. Laboratory confirmation is not always possible, and often isn't available early on in an investigation.
- The probable case: This means that the clinical symptoms and epidemiology are compatible with the case definition.
- The suspect case. This is a patient with compatible clinical symptoms, but may not be all of the symptoms, and a likely epidemiologic link has not yet been confirmed.

Using these levels of certainty allows investigators to include potential case-patients in their investigation even if needed laboratory confirmation is not yet available or not possible.
In a real outbreak situation, there will also be people that you cannot immediately place into one of these levels. These people are sometimes referred to as "persons under suspicion" or "persons under investigation." You need to get more information about these people to determine whether or not they are a case.



Case classification levels

Example: Outbreak case definition for cholera:
- Suspected case: A resident of area A with at least 1 episode of severe diarrhea during the period 1 Jan – 30 Apr, 2011.
- Confirmed case: Suspect case with rectal swab positive for Vibrio Cholera O1

What makes this different from a surveillance case definition?
The greatest difference is that there is a specific location and a specific time frame.

Step V: Find additional cases systematically and record the information

Once you have determined the criteria for being a case in this particular outbreak, you can begin identifying additional cases (ensuring that they meet the initial case definition) and collecting information from or about them. Often, just a few cases are reported to the health office. How would you find additional cases? Depending on the disease and the setting, you might want to:

- Contact health facilities – have other facilities seen similar cases?
- Contact laboratories – have specimens been sent for testing because of similar concerns?
- Contact community health workers – for some diseases that may not require hospitalization, community health workers may know of cases.
- Contact other districts – are you sure that the outbreak is restricted to one area? One of our objectives is to characterize the geographic extent of the outbreak, so it is worth contacting other districts.
- Talk to patients – sometimes, patients know or have heard of others with the same disease, or may have been infected by coming into contact with someone who escaped surveillance.
- Media – the media can be friend or foe by increasing awareness among the public of a health problem. More cases may seek care, but more non-cases (the worried) may seek care and overwhelm services.

For each case or suspected case, you want to collect and compile data in a single database. This database can be a piece of paper or logbook, or a computerized database using software such as Excel or Epi Info. This table shows part of a line list used by investigators of an outbreak of gastroenteritis. Not that the line list should be updated as new information, particularly laboratory results, become available.

Sample of a Line List

		Signs/Symptoms			Laboratory	Demographics	
Case #	**Date of Symptoms Onset**	**Diarrhea**	**Vomiting**	**Fever >37°C**	**Stool culture Result**	**Age**	**Gender**
1	22/10/14	Y	Y	Not done	Positive	19	M
2	25/10/14	N	Y	N	Negative	17	M
3	22/10/14	N	Y	N	Positive	23	F
4	27/10/14	Y	?	?	Pending	18	?
5	23/10/14	N	Y	N	Positive	21	M
6	21/10/14	Y	Y	Y	Not done	18	F

Step VI: Perform descriptive epidemiology

Now that you have identified cases and have basic data about them or from them, you can start describing those data in terms of what, who, when, and where.

Descriptive epidemiology encompasses the following 4 W's:

- "What" means the diagnosis disease, injury, or health condition, as a quantitative science, we also include how many cases of that disease.
- "When" means time.
- "Where" means place.
- "Who" means person.

Time: We would like to know when people became infected, but more practically, we usually only know when people became ill. Usually displayed graphically with epidemic curve ("epi curve" for short) – histogram (not bar chart) with time on X-axis, number of cases on Y-axis

Place: Similarly, we would like to know where people were exposed, but more commonly, all we know is where the case-patients live or work. Display with a map (spot map or shaded map), if possible. If not, use a table.

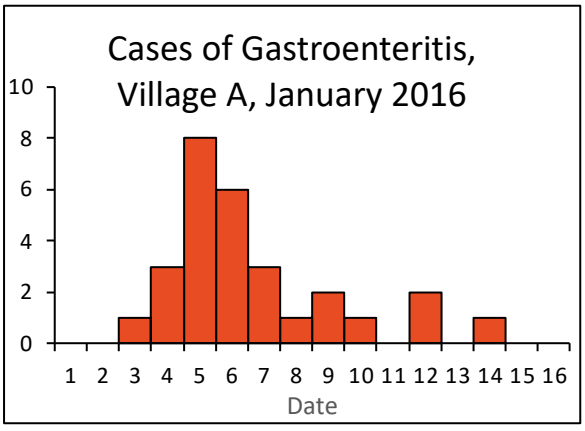
Person: Who are they? All ages, both sexes, or restricted to just some groups? Calculate rates using appropriate population denominators, if possible. Sometimes we have information on what cases may have in common, but sometimes not known until further investigation is conducted.

The traditional way to summarize an outbreak by time is to draw an epidemic curve. An epidemic curve is drawn as a histogram. The difference between a bar chart and a histogram is that bar charts have spaces between columns, but in a histogram, adjacent columns touch. The X-axis of the histogram represents the time, specifically the time or date of onset of illness. The Y-axis represents the number of cases that occurred during each time interval.

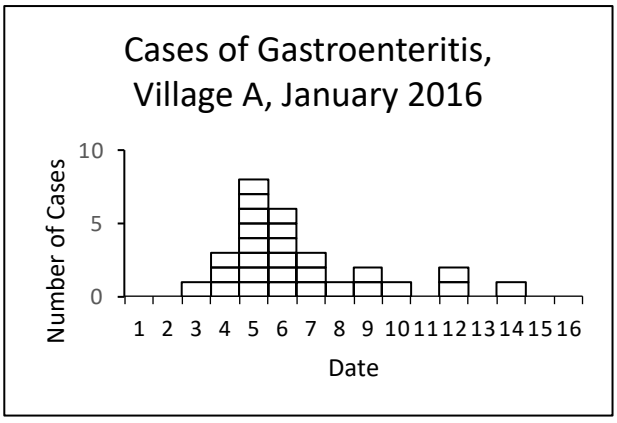
Time can be designated by weeks, or by days, or even by hour if the disease has a short incubation period. For example, Bacillus cereus, a foodborne toxin, has an incubation period of 1 to 6 hours.

Curves can be done by hand, using graph paper, or on a computer. But remember, epidemic curves are a type of histogram, so there should not be any space between the x-axis categories. Some people use columns, some like to show stacks of boxes. Either is acceptable.

Columns histogram



Stacks of boxes histogram



Example on drawing an epidemic curve:
To draw an epidemic curve, begin by looking at the dates of the onset. The X-axis should begin a few days before the first outbreak case, and should end a day or two after the last outbreak case.

The first step is to summarize the data by counting how many cases occurred on each day (see the tables below). No cases occurred during the period 1-8 October, one case occurred on 9 October, no cases on 10 October, one case on 11October and so on. The first case of this outbreak occurred on 9th of October and the last case occurred on the 25th of October.

What range would we suggest for the X-axis? In other words, when would we start the X-axis? We start in early October, before cases begin, continuing to the last date for which data is available. Accordingly, the X-axis includes dates from 1 October to 26 October.

What range do we suggest for the Y-axis? The maximum number of cases in one day was 13 on the 15th of October. The Y-axis should range from 0 to 14 or 15.

Dates of Onset of 56 Persons Having Disease X, District Y, October 2015

9 Oct	14 Oct	15 Oct	16 Oct	17 Oct	19 Oct
11 Oct	14 Oct	15 Oct	16 Oct	17 Oct	20 Oct
13 Oct	14 Oct	15 Oct	16 Oct	17 Oct	20 Oct
13 Oct	14 Oct	15 Oct	16 Oct	17 Oct	22 Oct
13 Oct	14 Oct	15 Oct	16 Oct	17 Oct	23 Oct
14 Oct	15 Oct	15 Oct	16 Oct	17 Oct	25 Oct
14 Oct	15 Oct	15 Oct	16 Oct	18 Oct	
14 Oct	15 Oct	15 Oct	16 Oct	18 Oct	
14 Oct	15 Oct	16 Oct	16 Oct	18 Oct	
14 Oct	15 Oct	16 Oct	17 Oct	19 Oct	

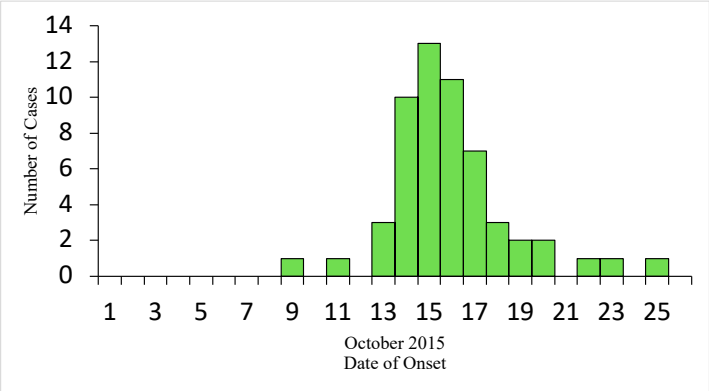
Date of onset	Number of cases
1-8 Oct.	0
9 Oct.	1
10 Oct.	0
11 Oct.	1
12 Oct.	0
13 Oct.	3
14 Oct.	10
15 Oct.	13
16 Oct.	11
17 Oct.	7
18 Oct.	3
19 Oct.	2
20 Oct.	2
21 Oct.	0
22 Oct.	1
23 Oct.	1
24 Oct.	0
25 Oct.	1
Total	56

Now we can add the data to the X-axis and Y-axis. The date of onset to the X-axis and the number of disease cases to Y-axis. Also we have to add axis labels and title of the histogram. Now the epidemic curve is completed. This is what your epidemic curves should look like – histogram, X-axis with pre-epidemic period till the end of the epidemic, Y-axis from 0 to slightly above largest value, labels for both axes, and a title with the disease, place, and time as in the example shown below.

Note: The epidemic curve could be in the form of a histogram or line graph.

Example of an epidemic curve:

Number of Cases of Disease X by Date of Onset, District Y, October 2015

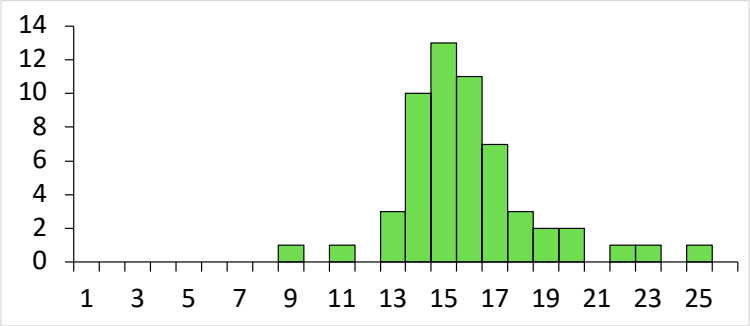


The value of an epidemic curve

- Shows the magnitude of the outbreak
- Shows the time course of the outbreak
- Can help determine the incubation period or exposure period
- Can show the pattern of spread
- Highlights outliers

We can see the magnitude of the outbreak, i.e., how many people were affected. Some epidemic curves show only a few cases, while others show hundreds or even thousands.

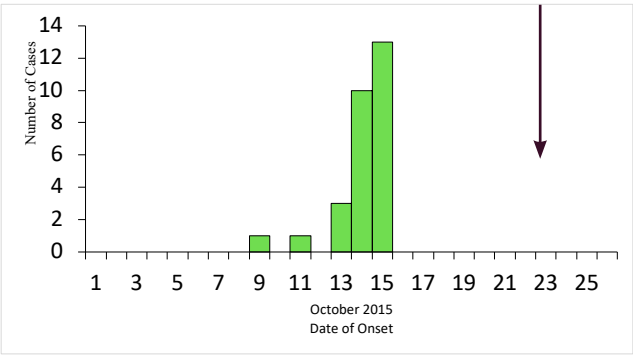
Magnitude of the outbreak



October 2015

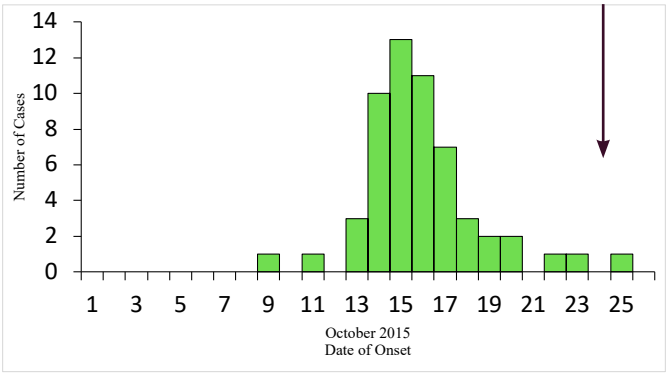
Another use is figuring out where we are in the outbreak. Look at this epidemic curve. If today was the 16th of October, would we expect more cases, or would the epidemic be mostly over?

Time course of the outbreak



Since the number of cases has climbed each day in the previous days, we would expect more cases.

Now, if today were the 26th of October, would we still be expecting many more cases?



Since the epidemic peaked 11 days ago, it appears as if it is over or almost over. Perhaps a few more cases will still occur, e.g., secondary cases or long incubation period cases, but we would not expect many.

Epidemic curves and manner of spread

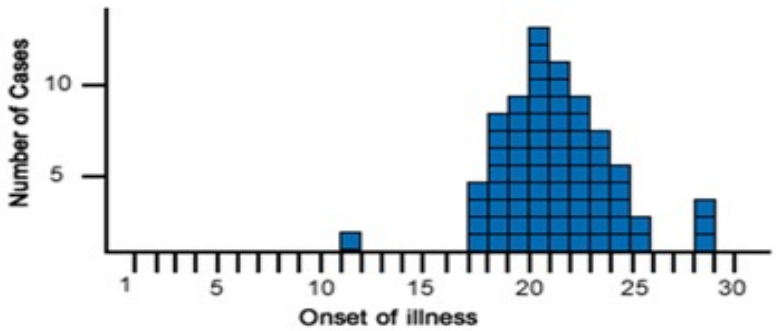
There are several outbreak patterns that can be useful in identifying the transmission method or source and predicting the future rate of infection:

1. Common source – All victims acquire the infection from the same source (e.g. contaminated water supply). There are two types:
   - Point common source – Common source outbreak where the exposure occurs in less than one incubation period.
   - Continuous common source – Common source outbreak where the exposure occurs over multiple incubation periods.
2. Propagated – Transmission occurs from person to person.
3. Intermittent– An outbreak caused by an intermittent source has intermittent cases.

Sometimes the shape of the epidemic curve can suggest the type of epidemic spread.

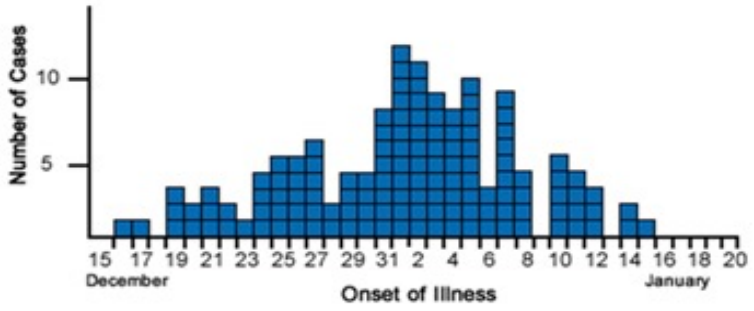
A point source outbreak resulting from exposure at a single point in time usually has a single peak, sometimes with a steeper upslope and a more gradual downslope. A point source is typical of foodborne outbreaks in which exposure occurs at a single meal, such as in a wedding. At point source outbreak, persons are exposed over a brief time to the same source, such as a single meal or event. The number of cases rises rapidly to a peak and falls gradually. The majority of cases occur within one incubation period of the disease.

Point source (single exposure) epidemic



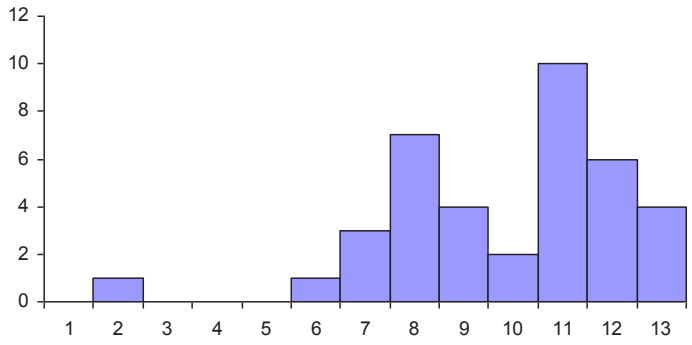
An outbreak caused by a continuous common source rises but stays high for the duration that the source is contaminated. This often occurs with cholera or other diseases whose source is a contaminated water supply. In a continuous common source outbreak, persons are exposed to the same source but exposure is prolonged over a period of days, weeks, or longer. The epi curve rises gradually and might plateau.

Continuous common source epidemic



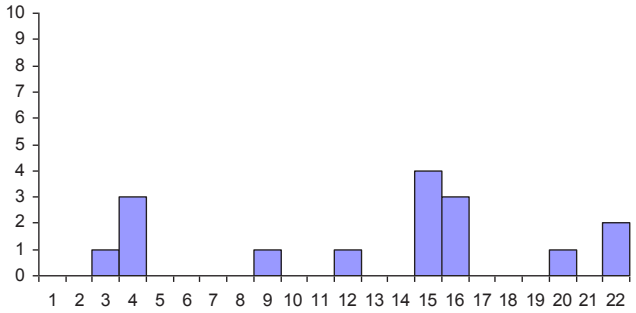
An outbreak caused by an intermittent source has intermittent cases.
An intermittent common-source outbreak means patients are exposed to the source of the disease at irregular intervals.

Intermittent source epidemic



A propagated outbreak such as measles usually has successive waves of cases.This is the typical pattern of diseases spread from person-to-person.

Propagated spread epidemic



Note that these are classic epidemic curves, and real-world epidemic curves may not be as classic as these illustrations.

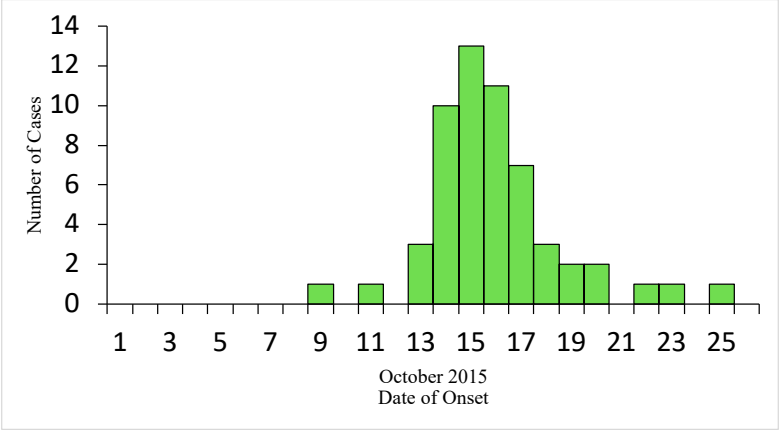
Estimation of the likely time of exposure:
Another use of an epidemic curve, particularly for a point source epidemic, is to identify the likely time of exposure. Consider a disease that has a known incubation period that ranges from 2 to 10 days, with a median or average of 4 days. The first case would have the shortest possible incubation period of 2 days. Firstly, we can subtract 2 days from the earliest case to estimate when exposure may have occurred.

In the epidemic curve below, if we assume that the cases on 9 October and 11 October are not part of the outbreak (maybe background cases) and that the first true case of the outbreak were the cases with onset on 13 October, then exposure would have occurred on or around 11 October.

If you disagree with the assumption and think that the first outbreak-associated case occurred on 9 October, then exposure would have occurred on or around 7 October.

Secondly, we subtract the median or average incubation period from the median case or peak of the epidemic. So subtract 4 days from 15 October (peak). Again, the exposure would have occurred on or around 11 October. Accordingly, we should look for a common exposure on or around 11 October, or perhaps some time between 7 October and 11 October.

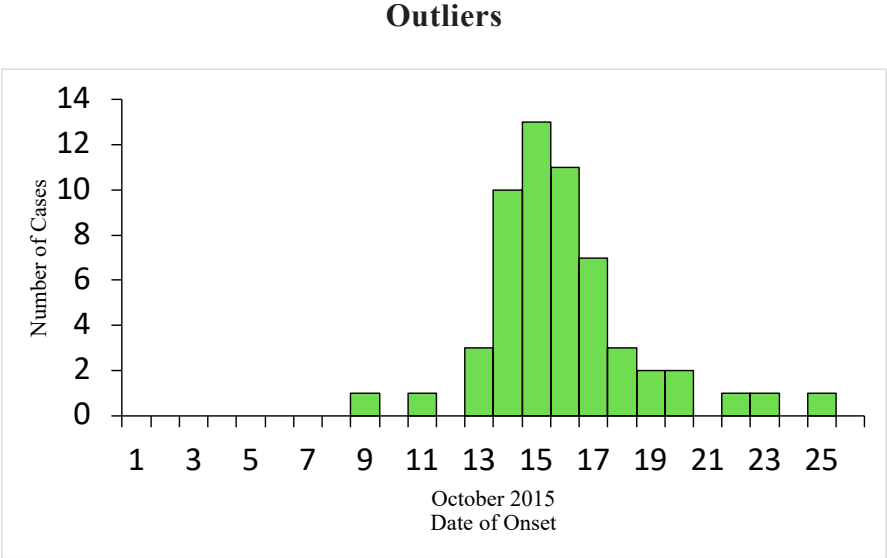
Identifying likely exposure period (point source outbreak)



The outlier cases can provide important clues to the exposure:

The outliers in the curve below are the cases occurred on 9 and 11 October before the start of the epidemic and cases occurred late on the 22, 23, and 25 October after the end of the epidemic. If cases on 9 and 11 October were part of the outbreak, how were they exposed? Perhaps they were the foodhandlers who catered the meal on October 11 and were the source of the outbreak, or perhaps they ate the contaminated food or ingredients a few days earlier than it was served to everyone else.

Now what about the late cases (late outliers)? We have to ask if someone brought home leftover food, so they were exposed later than everyone else, or perhaps they are just secondary cases or unrelated cases.

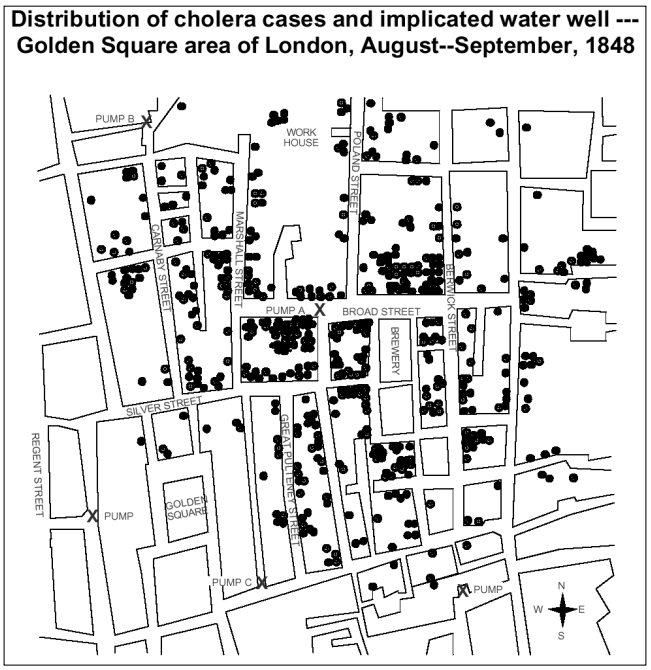


Description of the data by place

A description of the place affected by the outbreak is important in understanding the scope.We may need to describe a building, like a hospital, or the place could be a neighborhood. Often epidemiologic data is presented by the city, governorate or country.

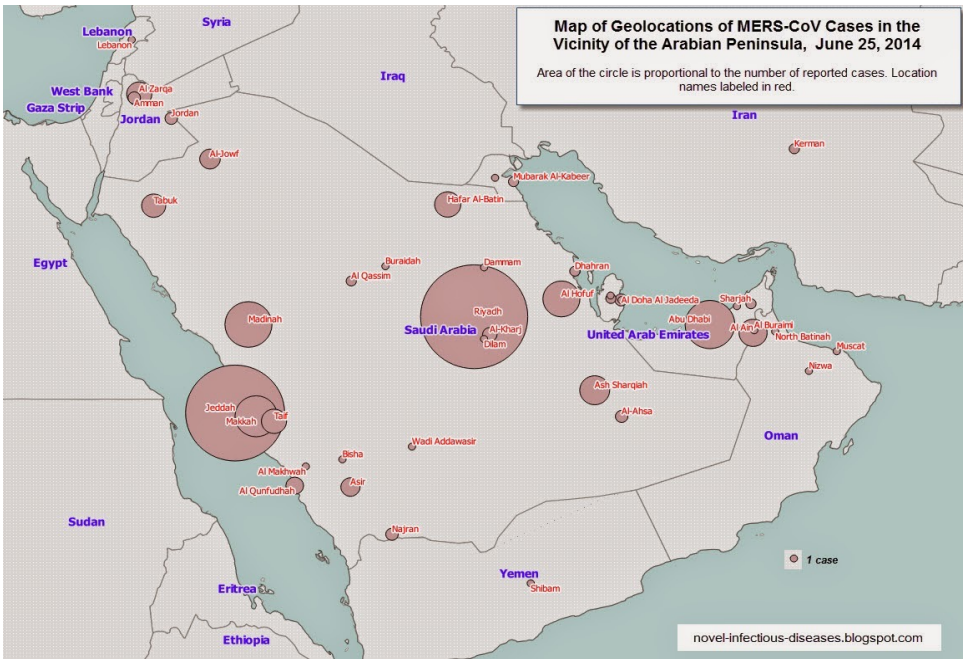
While place information can be included in text or a line listing, a map can help an investigator to better visualize the outbreak "place." Mapping the outbreak allows the investigator to assess the geographic extent of the situation and may also reveal patterns, such as clusters of cases, that may provide information about the cause or source of the outbreak.

There are two general types of maps commonly used to describe the disease. A spot map indicates a case by specific characteristic, usually where a case lives or works. An area map shows the number of cases by geographic area.



Example 1 of spot map: Developed by John Snow during his investigation on cholera epidemic in London in the middle of 19th century.

The area of the circle is proportional to the number of reported cases, affected areas labelled in red.



Example 2 of spot map: Confirmed cases of MERS-CoV in KSA, June-July 2014

This map shows visually where in the Arabian Peninsula the risk of infection is highest. This can help focus resources on prevention, surveillance, and treatment in these areas. It may also be used to help identify risk factors for the disease.

Description of the outbreak by personal characterstics

Examples of these characterstics:

- Age
- Sex
- Occupation
- Education
- Social class
- Income
- Marital status
- Race and ethnicity
- Underlying medical conditions
- Many others

The selection of variables to describe the outbreak depends on the circumstances and setting of the investigation. In most of the time, age and sex are very important variables to be considered.

Example:Distribution of brucellosis in Country X, year Y by age and sex

Age in Years	Males	Females	Total
15-24	10	22	32
25-34	45	50	95
35-44	47	25	72
45-54	17	13	30
55 and more	10	3	13
Total	129	113	242

The interpretation of the data in the above table: There are slightly more males than females with brucellosis and the peak age group for men is 35-44, while that for females is 25-34. Is there anything else we need to know? It would be better to know denominators in order to calculate rates.

The following table provides the denominators. How would we use them? We divide the case data by the denominators to calculate rates, in this situation the rate is prevalence. We had more males than female cases, but we have more males than females in the denominator.

Number of Investigated People (Denominator) for Brucellosis in Country X, Year Y by Age and Sex

Age in Years	Males	Females	Total
15-24	105	108	213
25-34	184	127	311
35-44	102	69	171
45-54	43	40	83
55 and more	44	15	59
Total	478	359	837

So, when we calculate the prevalence rates by using the data about cases and the number of people investigated in the previous two tables, the following table shows these rates by age and sex. Now we see that the prevalence of brucellosis is slightly higher in females than males. On the other hand, the peak age groups remain the same 35-44 years for males, and 25-34 years for females.

Prevalence of Brucellosis in Country X, Year Y by Age and Sex

Age in Years	Males	Females	Total
15-24	9.5%	20.4%	15%
25-34	24.5%	39.4%	30.5%
35-44	46%	36.2%	42%
45-54	39.5%	32.5%	36%
55 and more	22.7%	20%	22%
Total	27%	31.5%	29%

The explanatory phase of epidemic investigation (analytic phase)

It consists of the following four steps:

VII. Develop hypotheses.

VIII. Evaluate hypotheses epidemiologically.

IX. Reconcile epidemiology with laboratory and environmental findings.

X. Conduct additional studies as necessary.

Step VII: Developing the hypothesis/hypotheses

In the context of an outbreak, a hypothesis is an educated guess about an association between an exposure and outcome, and/or about mode of spread. In other words, it is a guess about the cause or causes of this outbreak.

In reality, outbreak investigators do not wait until Step 7 to develop hypotheses. Usually, investigators begin to hypothesize about possible causes from the very beginning of the investigation. Hypothesis should be in a form that allows it to be tested.

To understand hypotheses, we should review exposures and outcomes, because every hypothesis is a guess about the relationship between an exposure and an outcome. By exposure we mean any factor that may cause or influence the likelihood of getting the disease. The factor can be something that someone chooses to do (like smoking cigarettes), or something that happens to someone (like getting bitten by a mosquito). By outcome, we mean the health effect, which can be a health condition, disease, injury, etc. The table below shows examples of exposures and outcomes.

Exposure	**Outcome**
Eat contaminated meat	Develop E. coli infection
Driving without a seat belt	Injured in vehicle accident
Drink alcohol	Become drunk
Drink alcohol everyday x 10 years	Liver damage
Live near mosquito breeding site	Contract malaria
Use insecticide-treated bed nets	Do not contract malaria
Smoke cigarettes	Lung cancer

How investigators develop a hypothesis:

- Subject matter knowledge – known sources, vehicles, transmission modes
- Review descriptive epidemiology – what would account for most?
- Outliers (unique exposure opportunities)
- Talk to case-patients – what do they think?
- What do local health officials think?

Subject-matter knowledge for hypothesis generation:

Knowledge of the disease, its reservoir, its mode of transmission, and other features is probably the most common way of developing hypotheses. For a disease without a confirmed diagnosis yet, ask yourself what kinds of agents can cause this clinical presentation. For a known disease but unknown source or reservoir or vehicle, ask, "What are the agent's usual reservoirs? How is the agent usually transmitted? What are the most common vehicles for transmitting this agent to humans? What are the known risk factors? What are the usual suspects for this disease?" E.g., for cholera – the source is usually water pollution, for measles – the source is contact with a case.

Descriptive epidemiology for hypothesis generation:

One of the reasons descriptive epidemiology is so important is that it provides clues that we can use to develop hypotheses.

Time – Sometimes the shape of the epidemic curve hints at the type of epidemic spread. A point source outbreak usually has a single peak, sometimes with a steeper upslope and more gradual downslope. An outbreak caused by continuing common source rises but stays high. An outbreak caused by an intermittent source has intermittent cases. A propagated outbreak such as measles usually has successive waves of cases.

Place – Why is the attack rate of a disease particularly higher in one area than other areas? What is special for this area that makes it affected more?

Person – Which group(s) by age, sex, occupation, etc. have highest rates?

Example: An outbreak of pneumonia occurred among residents of a nursing home. The nursing home is divided into east and west wings, and the cases were clustered in the east wing only. The hypothesis about the place which might be developed from this distribution of cases is: "Some exposure limited to residents of the east wing has happened."

Suppose the distribution of cases was not only among residents of the east wing, but the west wing residents also affected, then the new hypothesis should be consistent with this distribution of cases. So, the new hypothesis might be, "An exposure unrelated to the location of the residents' room has occurred." This requires us to explore what else these particular residents have in common. For example, if a visitor to the nursing home (such as a family member) developed the same illness as the nursing home residents, what exposure in the nursing home did that visitor have?

Looking at the outliers for hypothesis generation:

- By time: Date of onset of outlier could be early before the start or late after the end of the outbreak. What exposure did that patient (outlier) have in common with others, but perhaps at a different time?
- By place: A non-resident visitor to the nursing home as an example usually has a limited number of exposures, this can narrow the possibilities.
- By person: Again, what exposure did that patient (who is different from the others by age, sex, etc.) have in common with the other patients?

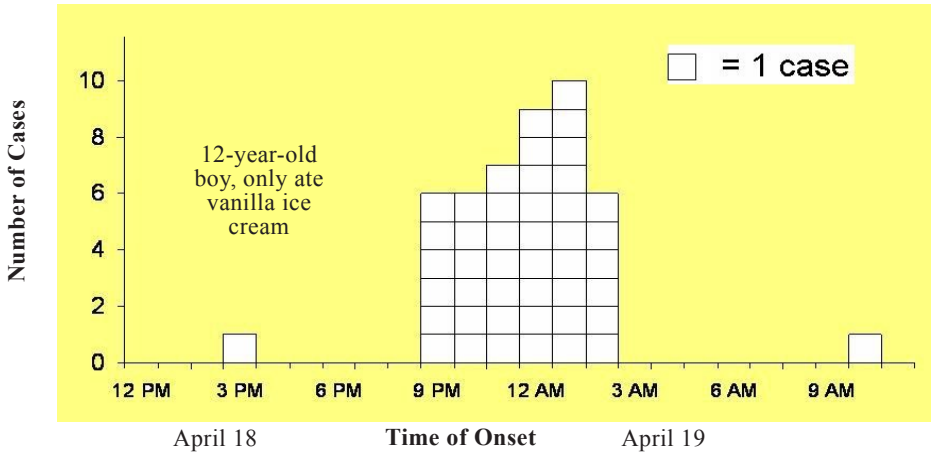
As we have discussed previously, outliers can provide important clues. An outlier may have only one exposure in common with most of the other cases – what is that exposure? For example, if a visitor to the nursing home in the previous example developed the same illness as the nursing home residents, what exposure in the nursing home did that visitor have? In this case, the visitor might be considered the outlier and the index case of this outbreak.

The epidemic curve below is for an outbreak of gastroenteritis caused by Staph aureus. It occurred among a group of about 80 people who had attended a picnic on the evening of April 18. All became sick 4-7 hours after eating.

Notice the early case that occurred at 3 pm. Clearly that case occurred even before the picnic took place. Wouldn't it be interesting to interview that case and see if that person ate any of the foods later served at the picnic, and if so, which one or ones? After interviewing that case, it happened that he was a 12-year-old who ate vanilla ice cream only before the start of the picnic.

However, despite the importance of outliers for hypothesis generation, the outlier could be a totally unrelated case, or a miscoding of the time of onset.

Outbreak of Staph gastroenteritis among a group of people who attended a picnic



Talking with patients for hypothesis generation

Patients usually have plenty of time to think about their own exposures and may have talked with others about what they have in common. So, interviewing a number of patients might be of great value in the formulation of the hypothesis. Different scenarios can be followed during the interview:

- Have open-ended conversations.
- What do they think is the source?
- For foodborne outbreak, ask about foods.
- Sometimes, you may need to look at the kitchen.
- You may bring some patients together to chat and see whether they have any common exposures.

Talking with local authorities for hypothesis generation

Finally, if you are coming from outside the area, you may interview local health officials and other public authorities or community informants about their views regarding the source of the outbreak, they can tell about recent festivals or other recent events or gathering in the area.

Example on developing a hypothesis:

Scenario: Suppose you were notified that several cases of meningococcal meningitis had recently occurred among newborns in Hospital X.

The subject matter knowledge about meningococcal meningitis: meningitis is inflammation of the lining of the brain, characterized by fever, headache, stiff neck, irritability, vomiting, and poor feeding. It is transmitted from person-to-person through droplets of respiratory or throat secretions from carriers. Close and prolonged contact (e.g., kissing, sneezing or coughing, or living in close quarters with an infected person) facilitates the spread of the disease.

What is your hypothesis? The hypothesis should include the outcome and possible exposure. The outcome is meningococcal meningitis, but what is the exposure?

Possible answers include hospital X – yes, but this is not helpful. What about the hospital, or where in the hospital? Nursery? Maternity ward? It may depend on the hospital where newborns stay. Which individual at the hospital was the source of the infection? e.g., staff person, mother, or visitor. It happened that this hospital did not have a newborn nursery. All newborns stayed with their mothers in the maternity ward. Because meningitis is spread from person-to-person, the investigators

developed a series of hypotheses to explore that the babies were exposed to someone in the maternity ward, either a member of the hospital staff such as a doctor, nurse, technician or other staff, or one of the mothers, or a visitor to the maternity ward.

Step VIII: Evaluate (test) hypotheses epidemiologically

Once we have developed some hypotheses, we have to evaluate them.

In some cases, when combining the laboratory, clinical, environmental and/or epidemiological evidence, the evidence becomes strong enough to determine the association without further testing. For example, if the newborns in the meningitis example tested positive for Neisseria meningitides, and staff, mothers, and visitors were tested and only one visitor tested positive for Neisseria meningitides, and that visitor was around all of the babies who tested positive, you may not need to conduct further epidemiological study.

However, if the evidence is not so strong, or if there is some question as to the cause, analytic epidemiology is used to test if an association exists, and if so, how strong it is. The most common types of analytic studies used in outbreak investigations are retrospective cohort study and case-control study.

Cohort study for outbreak investigation

In the context of an outbreak, a cohort study is a good choice if the outbreak occurs among a well-defined group of people, such as a school or wedding or funeral attendees. We can get a complete list of everyone in the group, design a data collection tool such as a questionnaire, and collect information from or about everyone in the group. Then we can calculate attack rates for those who were and were not exposed to various factors. For example, what was the attack rate among those who ate the hypothesized food? What was the attack rate among those who did not eat that food? We compare the attack rates, usually by dividing the attack rate of those who ate food on the attack rate of those who did not eat the food. This fraction is called a relative risk or risk ratio. The type of cohort study used in such situation is a retrospective (historical) cohort study. The steps to be followed are:

- Include everyone in the community (schools, funeral attendees, weddings, etc.) where the outbreak occurred.
- Collect information from or about everyone.
- Calculate attack rates among those exposed and not exposed to various hypothesized factors.
- Calculate the ratio of attack rates (relative risk, risk ratio).

Example: Consider the data shown in the table below. An outbreak occurred among people who attended a picnic. Investigators interviewed almost everyone who attended the picnic. Of the 75 people interviewed, 46 had developed gastroenteritis within a few hours. Among the 54 people who ate a particular item, 43 got sick, for an attack rate of 79.6%. Of the 21 who did not eat that item, only 3 got sick, for an attack rate of 14.3%. The relative risk equals the attack rate among the exposed group divided on the attack rate among the unexposed group, which is 5.6, that means people who ate food A were 5.6 times more likely to get sick as compared to those who did not eat that food item.

Outbreak of Gastroenteritis

	Ill	Well	Total	Attack rate
Ate food A	43	11	54	79.6%
Did not eat food A	3	18	21	14.3%
Total	46	29	75	61.3%

Relative risk =79.6 / 14.3 = 5.6

Case-control study for outbreak investigation

Sometimes, a cohort study is impractical, particularly if we do not have a well-defined group to enroll.

An alternative is a case-control study. In a case-control study, investigators enroll the cases, then find comparable people who did not develop the illness. These are the controls. The investigator collects similar information about exposures from each group. The exposure experience of cases versus controls can be compared with a measure similar to a relative risk that is called an Odds Ratio. The odds ratio is interpreted similar to what we just discussed for relative risks, i.e., an OR > 1 is consistent with a harmful effect, an OR close to 1 indicates no effect, and an OR < 1 is consistent with a potentially beneficial or protective effect.

The steps to be followed are:

- Include people with the disease (cases).
- Enroll comparable group who do not have the disease (controls).
- Collect exposure information equally from both groups.
- Compare exposure experience between cases and controls using odds ratio.
- Interpret odds ratio similar to risk ratio (OR > 1, OR = 1, OR < 1).

Example: **Calculation of OR in Case Control Study**

	Exposed	**Not Exposed**	**Total**
Cases	20 (a)	30 (b)	50
Controls	5 (c)	45 (d)	50
Total	25	75	100

Odds ratio= ad / bc=20*45/ 30*5= 6

The OR of 6 means that cases were at risk to be exposed to the risk factor/s 6 times more than the control group.

Steps IX and X: Reconcile epidemiology with laboratory and environmental findings and conduct additional studies as necessary

While epidemiologists usually have confidence in their epidemiologic findings, laboratory and environmental findings are essential before some authorities are willing to take action. Ideally, the epidemiologic, laboratory, and environmental evidence all point in the same direction. If the epidemiologic and laboratory evidence point in different directions, additional studies and efforts must be made to understand why.

The response phase of the epidemic

It consists of the following three steps:

Step XI: Implement and evaluate prevention and control measures.

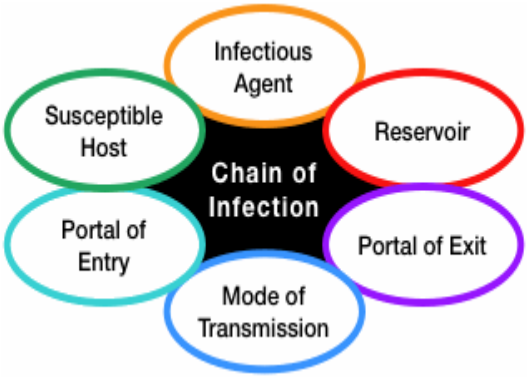
Step XII: Initiate or maintain surveillance.

Step XIII: Communicate findings.

Step XI: Implementing and evaluating prevention and control measures

Remember that the goal of most outbreak investigations is to control and prevent disease transmission. Implementing prevention and control measures helps prevent further exposure and future outbreaks by eliminating or treating the source. Prevention and control measures should be initiated as soon as they are known and available.

The diagram below shows what is called the chain of transmission, which is composed of the reservoir (habitat), where an infectious agent normally lives and multiplies, (such as humans, animals, or the environment), the infectious agent and the susceptible host.



Chain of Transmission of Infectious Disease

The infectious agent must be transmitted in the appropriate way to a susceptible host, and the agent must enter that host through an appropriate portal of entry. In general, control strategies focus on any of the 3 parts of the chain: Control of the reservoir, interrupt transmission and protect the host. Often, control and prevention measures use multiple strategies.

Control strategies for reservoir:

Some control strategies are aimed at the agent where it lives, i.e., its reservoir. Specific strategy will differ depending on the type of reservoir. If human reservoir, treat infected patient (symptomatic patient or asymptomatic carrier) to eliminate infection (Example: Sexually transmitted diseases, pertussis, etc.).

If environmental reservoir, could try to decontaminate or disinfect (Example: Legionella in water system of hospital)

If animal reservoir, could attempt to vaccinate (rabies vaccine for domestic dogs, cats; attempts to vaccinate raccoons in wild)

Cull (eliminate) potentially infected animals (control rat population to reduce risk of plague; millions of chickens potentially infected with avian flu killed in Southeast Asia).

The second target in the chain of transmission is the route of transmission itself. There are many different ways that infections can be transmitted. Usually, these routes are grouped into direct transmission and indirect transmission.

Direct transmission includes:

- Touching, kissing, intercourse
- Droplet (e.g., from coughing, that is expelled by the cougher, then falls a few feet away, so only someone very nearby is affected)
- Trans placental

Strategies to prevent direct transmission include:

- Treatment/isolation of an infected person.
- Barriers to prevent agent from leaving host (bandages, dressings, condoms).

Indirect routes of transmission include:

- Airborne – infectious agents that remain suspended in the air and can be breathed in.
- Vector-borne – transmitted by an arthropod such as a mosquito, tick, louse, mite, etc.
- Vehicle – transmitted by an inanimate object such as food or water, biologic such as a blood transfusion, or fomite such as a towel or surgical instrument.

Control strategies for indirect routes of transmission:

- For airborne transmission, you can isolate the patient in a private room, with negative pressure, door closed, wear N95 masks.
- For vector-borne diseases, eliminate breeding sites, kill the vector by using insecticides.
- For vehicle borne, you can control these by using heat, pasteurization or other chemical processes. A common example of this method of control is to prohibit an infected food handler from working. Chlorination of water supply is another example. For biologic products such as blood, bone marrow, skin grafts, or the like, either don't use at all, or sterilize. Similarly, fomites (inanimate objects such as towels, water bottles, etc.) can be disinfected or sterilized.

Control strategies for the protection of the host:

We can try to prevent entry and protect a potentially susceptible host.

- Messages to practice healthy behaviors are not only targeted to persons with infections, but also those that may be at risk.
- At risk persons can be removed or excluded to avoid exposure. For example, mass gatherings can be cancelled altogether to avoid close contacts and opportunity for spread.
- Long sleeves and long pants are recommended to reduce the risk of being bitten by mosquitoes that can transmit a particular disease.
- Staff and visitors may wear masks to reduce the risk of infections spread by droplet or air (PPEs).
- Vaccination is used to booster the host's immune response, so that, if the body encounters the pathogen, the immune system acts quickly to prevent infection.
- Alternatively, for some organisms such as hepatitis A, exposed persons can be given immune globulin made up of antibodies to combat infection.
- Pre-exposure prophylaxis, that is, taking preventive medication such as antimalarial, is intended to prevent infection from occurring, even if you are exposed to the pathogen.

- Post-exposure prophylaxis, used, for example, for rabies.
- Contact tracing is the process of listing all contacts of a case and finding them. For polio, contacts can be vaccinated. If the contact tracers detect any early signs, the contacts can be immediately isolated and treated in a treatment facility.

Outbreak control measures can be divided into immediate and long-term control measures. For example, what is an immediate measure and a longer-term control measure for a waterborne outbreak? Immediate: Boil water.

Long-term: Chlorinate public water supply.
Immediate control measures primarily involve working to reduce immediate individual risk. Examples of immediate control measures may include recommendations to sterilize drinking water, destroying local mosquito breeding sites, or recalling a food product. It is also important at this stage to have good communication with the public so that individuals are aware of the need to boil water, for example. This can be done even before a specific pathogen is identified.

In immediate control measures, we work with persons at risk, while in long-term control measures, we work with regulators or government.

Each outbreak results from a change – in the environment, in the reservoir, in the host population, or in something else. The question we need to ask at the end of our investigations is: Do these conditions still exist? Could another outbreak occur again? What do I need to do to prevent a future outbreak?

Long-term control measures are more extensive than immediate control measures and may focus on engineering and policy changes. Examples of such measures are recommending different food safety procedures in a restaurant, training staff on sanitation regulations, enhancing ventilation systems, and implementing better disinfection protocols at a local swimming pool. This stage often involves working with government regulators, industry, and health educators.

Step XII: Initiate or maintain surveillance
We have to continue conducting surveillance of the disease if a surveillance system exists or initiate one if the disease is not currently under surveillance.

Continued surveillance and monitoring are essential to determine whether your control measures are working.

Step XIII: Communication of findings
While it is the last step, that does not mean you should wait until the end of the outbreak investigation to communicate any findings. As with prevention and control communication, should be ongoing and should occur whenever there is important information to disseminate. Regular communication among team members is essential while conducting the investigation.

It is also important to keep the public informed. A spokesperson should be designated, and clear concise messages should be developed. The media can be used to convey messages which may

include information to keep people informed, to allay panic, or to provide concrete steps that individuals can take to protect themselves and prevent the spread of disease.

Information should be communicated to health professionals as well. This may include case definitions, requests for reporting, or information on vaccination or treatment protocols. It is also important to keep public health officials and policy makers aware of the situation so that resources and planning decisions can be made.

At the end of the investigation, findings are communicated both through an oral briefing and a written report. The oral briefing usually provides information to local health authorities and the persons responsible for prevention and control measures.

The written report is usually a formal document completed at the end of an investigation. A written report can serve many purposes and is an essential aspect of any investigation. Here are the most common reasons for writing reports:

- A written report recommends actions that can be taken to prevent and control current and future outbreaks.
- Reports share new information or insights about the outbreak, such as a newly discovered transmission mechanism.
- Reports also serve as a record of your performance and document the magnitude of health problems.
- Public health officials may reference past outbreak reports to review the type of investigation, relevant findings, and important lessons learned.
- It is also useful to analyze data from multiple outbreaks to present a summary of outbreaks over time.
- Some health departments use the reports to track how many outbreaks they investigate in a given year, and what types of pathogens were linked to them.
- Reports can also help support research and evaluation activities and the development of recommendations. For example, if your department has been conducting multiple outbreak investigations due to diarrheal disease in preschools, you may decide that school personnel need to be re-educated about the importance of hand-washing practices among themselves and their students.

References

1. Last JM, editor. Dictionary of epidemiology. 4th ed. New York: Oxford University Press; 2001. p. 61.
2. Torrence ME. Understanding Epidemiology. Mosby's Biomedical Science Series. Missouri, Mosby-Yearbook Inc., 1997.
3. Lilienfeld DE, Stolley PD. Foundations of epidemiology, 3rd ed. New York, Oxford University Press, 1994.
4. Rothman KJ. Epidemiology: An Introduction. New York, Oxford University Press, 2002. (ISBN: 0 195 135 547).
5. Morris JN. Uses of epidemiology. Edinburgh: Livingstone; 1957.
6. Tyler CW, Last JM. Epidemiology. In: Last JM, Wallace RB, editors. Maxcy-Rosenau-Last public health and preventive medicine, 14th ed. Norwalk (Connecticut): Appleton & Lange; 1992. p. 11.)
7. Snow J. Snow on cholera. London: Humphrey Milford: Oxford University Press; 1936.
8. Doll R, Hill AB. Smoking and carcinoma of the lung. Brit Med J 1950; 2:739–48.
9. Kannel WB. The Framingham Study: its 50-year legacy and future promise. J Atheroscler Thromb 2000; 6:60–6.
10. Tuberculosis in New York State, 2011 Annual statistical Report, Bureau of tuberculosis control.
11. Standardization, A classic epidemiological method for the comparison of rates, Epidemiological Bulletin, Vol. 23 No. 3, September 2002.
12. Mitali Verma, (n.d.) Fertility of Population Dynamics: Measures of Fertility Analysis. Accessed online from http://www.yourarticlelibrary.com/population-geography/fertility-ofpopulation-dynamics-2-measures-of-fertility-analysis/43246
13. Santosa A, Wall S, Fottrell E, Högberg U, Byass P (2014). "The development and experience of epidemiological transition theory over four decades: a systematic review". Global Health Action. 7: 23574. doi:10.3402/gha.v7.23574. PMC 4038769. PMID 24848657.
14. Omran, A. R. (2005) [1971], "The epidemiological transition: A theory of the epidemiology of population change" (PDF), The Milbank Quarterly, 83 (4): 731–57, doi:10.1111/j.1468-0009.2005.00398.x, PMC 2690264, PMID 16279965, archived from the original (PDF) on 2013-04-12. Reprinted from the Milbank Memorial Fund Quarterly, 49 (4, Pt 1): 509–538, 1971
15. Wahdan, M. H. (1996). "The epidemiological transition". Eastern Mediterranean Health Journal. 2 (1): 2.
16. Corruccini, Robert S. & Kaul, Samvit S (1983), "The epidemiological transition and the anthropology of minor chronic non-infectious diseases", Medical Anthropology, 7 (3): 36–50, doi:10.1080/01459740.1983.9987039.
17. Kelsey, J., Thompson, W., & Evans, A. (1986). Methods in Observational Epidemiology. London: Oxford University Press.
18. Jordan WM. Pulmonary embolism. Lancet. 1961; 278:1146-1147.
19. Pearce NE. The ecologic fallacy strikes back. J Epidemiol Community Health 2000; 54:326-7.

20. Schlesselman, J. (1982). Case-Control Studies Design, Conduct, Analysis. London: Oxford University Press.
21. Elwood, J. (1998). Critical Appraisal of Epidemiological Studies and Clinical Trials. New York, USA: Oxford University Press.
22. 22.Meinert, CL. Clinical trials: design, conduct, and analysis. New York, Oxford University Press, 1986.
23. Dipika Sur, M.D., R. Leon Ochiai, M.H.S., Sujit K. Bhattacharya, M.D., Nirmal K. Ganguly, et al. A Cluster-Randomized Effectiveness Trial of VI Typhoid Vaccine in India, N Engl J Med 2009; 361:335-344 DOI: 10.1056/NEJMoa0807521).
24. Vishnu Renjith, Renjulal Vesodharan, Judith A. Noronha, et al. Qualitative Methods in Health Care Research, Int J Prev Med. 2021; 12: 20].
25. Polit DF, Beck CT. Philadelphia: Lippincott Williams & Wilkins; 2008. Nursing Research: Generating and Assessing Evidence for Nursing Practice. Nieswiadomy RM. 6th ed. Boston: Pearson; 2012. Foundations of Nursing Research.
26. Sorrell JM. Qualitative research in clinical nurse specialist practice. Clin Nurse Spec. 2013; 27:175–8. [PubMed].
27. Munhall PL. 5th ed. Sudbury, MA: Jones and Bartlett Learning; 2012. Nursing Research: A Qualitative Perspective.
28. Butler A, Hall H, Copnell B. A guide to writing a qualitative systematic review protocol to enhance evidence-based practice in nursing and healthcare. Worldviews Evid Based Nurs. 2016; 13:241–9.
29. Karlsson E, Sävenstedt S, Axelsson K, Zingmark K. Stories about life narrated by people with Alzheimer's disease. J Adv Nurs. 2014; 70:2791–9. [PubMed] [Google Scholar].
30. Jyothi Cornelio S, Nayak BS, George A. Experiences of mothers on parenting children with leukemia. Indian J Palliat Care. 2016; 22:168–72. [PMC free article] [PubMed] [Google Scholar].
31. Williams K, King J, Fox JR. Sense of self and anorexia nervosa: A grounded theory. Psychol Psychother. 2016; 89:211–28. [PubMed] [Google Scholar].
32. LeBaron V, Beck SL, Maurer M, Black F, Palat G. An ethnographic study of barriers to cancer pain management and opioid availability in India. Oncologist. 2014; 19:515–22. [PMC free article] [PubMed] [Google Scholar].
33. Lubold AM. Historical-qualitative analysis of breastfeeding trends in three OECD countries. Int Breastfeed J. 2019; 14:36. [PMC free article] [PubMed] [Google Scholar].
34. Beal CC, Millenbruch J. A qualitative case study of poststroke sexuality in a woman of childbearing age. J Obstet Gynecol Neonatal Nurs. 2015; 44:228–35. [PubMed] [Google Scholar].
35. Zachariah R, Harries D, Ishikawa N et al. Operational research in low income countries: what, why and how? Lancet Infect Dis 2009; 9: 711-717.
36. Rajesh Kunwar and V. K. Srivastava, Operational Research in Health-care Settings, Indian J Community Med. 2019 Oct-Dec; 44(4): 295–298. doi: 10.4103/ijcm.IJCM_4_19.
37. Operational Research to Improve Health Services, a guide for proposal development ,2013, Department of Pediatrics and Child Health, Faculty of Medicine and Health Sciences, Stellenbosch University International Union Against Tuberculosis and Lung Disease .
38. John M. Last, The Dictionary of Public Health, first edition.
39. Sackett DL (1979). Bias in analytic research. Journal of Chronic Diseases 32:51-63.

40. Glenn M. Hymel, in Clinical Massage in the Healthcare Setting, 2008, https://www.sciencedirect.com/topics/nursing-and-health-professions/confounding-variable.
41. (Batwala, V., Magnussen, P., Mirembe, J. et al. Timing of malaria messages for target audience on radio airwaves. Malar J 11, 283 (2012). https://doi.org/10.1186/1475-2875-11-283).
42. 42. Mayrent SL, editor. Epidemiology in Medicine. New York: Lippincott Williams & Wilkins; 1987.)
43. Christenfeld NJ, Sloan RP, Carroll D, Greenland S. Risk factors, confounding, and the illusion of statistical control. Psychosom Med. 2004;66:868–75.
44. G. Berry, M. L. Newhouse and P. Antonis British Journal of Industrial Medicine Vol. 42, No. 1 (Jan., 1985), pp. 12-18).
45. Simon Sanderson, Iain D Tatt, Julian PT Higgins International Journal of Epidemiology, Volume 36, Issue 3, June 2007, Pages 666–676, https://doi.org/10.1093/ije/dym018.
46. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health 2005; 95: S144-50.
47. Olsen, J. (2003). What characterizes a useful concept of causation in epidemiology? Journal of Epidemiology and Community Health, 57, 86 - 88.
48. Weed DL. Causal criteria and Popperian refutation. In: Rothman JK, ed. Causal Inference. Massachusetts, Epidemiology Resources Inc., 1988.
49. Hill, Austin Bradford (1965). "The Environment and Disease: Association or Causation?" Proceedings of the Royal Society of Medicine. 58 (5): 295–300. doi: 10.1177/003591576505800503. PMC 1898525. PMID 14283879.
50. Abdel Rahim- Omran, the Clark Omran system of research design in epidemiology, Bahrain medical Bulletin, April 1984, vol. 6, No.1.
51. Health research methodology, a guide for training in research methods, second edition, World Health Organization, regional office for the western pacific Manila, 2001.
52. Commission on chronic illness: Chronic illness in the United States, Vol.1Common Wealth Fund, Harvard University Press, Cambridge, 1957, p.45.
53. Strong K, Wald N, Miller A, Alwan A. Current concepts in screening for noncommunicable disease. J Med Screen 2005; 12:12-9.
54. Mausner and Kramer, Epidemiology, an Introductory Text, 1985.
55. (Wilson, J.M.G, and Jungner, F.: Principles and practice of screening for disease (Public health papers no.34). WHO, Geneva, 1968.
56. Parikh R, Mathai A, Parikh S, Sekhar GC, Thomas R. Understanding and using sensitivity, specificity and predictive values. Indian J Ophthalmol (2008) 56:45–50.10.4103/0301-4738.37595 [PMC free article] [PubMed] [CrossRef] [Google Scholar].
57. Fox, J.P., Hall, C.E., et al.: Epidemiology: Man and Disease. Macmillan Publishing Co., Inc., New York, 1970.
58. https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/epidemic-theory.
59. Diekmann O, Heesterbeek JA, Metz JA. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. J. Math. Biol. 1990; 28:365–381.
60. [Dietz K. The estimation of the basic reproduction number for infectious diseases. Stat. Methods.

Med. Res. 1993; 2:23–41. [PubMed] [Google Scholar].

61. Becker NG. Analysis of Infectious Disease Data. Chapman & Hall; Boca Raton, FL, USA: 1989.
62. Principles of Epidemiology in Public Health Practice , third Edition , an Introduction to Applied Epidemiology and Biostatistics, CDC, Atlanta, 2012.
63. World Health Organization. Report of the technical discussions at the twenty-first World Health Assembly on 'national and global surveillance of communicable diseases.' Geneva: World Health Organization; 18 May 1968, p. A21.
64. Thacker SB, Berkelman RL. Public health surveillance in the United States. Epidemiol Rev 1988; 10:16.
65. Thacker SB, Stroup DF, Parrish RG, Anderson HA. Surveillance in environmental public health: issues, systems, and sources. Am J Public Health 1996; 86:633–8.
66. Vaughan JP, Morrow RH. Manual of epidemiology for district health management. Geneva: World Health Organization; 1989.
67. Leon Gordis, Epidemiology, fifth edition.
68. R. Bonita, R. Beaglehole, T. Kjellstrom, Basic epidemiology, second edition, World Health Organization.
69. Orenstein WA, Bernier RH. Surveillance: information for action. Pediatr Clin North Am 1990; 37:709–34.
70. WHO recommended surveillance standards. Geneva, World Health Organization. 1997.
71. Teutch Sm and Churchill RE, principles and practice of public health surveillance, Oxford university press, 1994.
72. Henning K.J. Overview of syndromic surveillance: what is syndromic surveillance? Morb. Mortal. Wkly. Rep. 2004; 53(Suppl.):5–11. [Google Scholar].
73. Leder K., Torresi J., Libman M.D., Cramer J.P., Castelli F., Schlagenhauf P., Wilder-Smith A., Wilson M.E., Keystone J.S., Schwartz E., Barnett E.D., von Sonnenburg F., Brownstein J.S., Cheng A.C., Sotir M.J., Esposito D.H., Freedman D.O., for the GeoSentinel Surveillance Network GeoSentinel surveillance of illness in returned travelers, 2007–2011. Ann. Intern. Med. 2013; 158(6):456–468.
74. Douglas N. Klaucke, M.D. James W. Buehler, M.D. et al. Guidelines for Evaluating Surveillance Systems, https://www.cdc.gov/mmwr/preview/mmwrhtml/00001769.htm
75. A new public health world order. The Lancet Infectious Diseases, 2004, 4: 4:475.
76. WHO proposals for the revision of the International Health Regulations (http:/ /www.who.int/csr/ihr/revisionprocess/working_paper/en/).
77. Gottlieb MS, Schroff R, Schanker HM, Weisman JD, Fan PT, Wolf RA, et al. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. N Engl J Med 1981; 305:1425-31.
78. Goodman RA, Buehler JW, Koplan JP. The epidemiologic field investigation: science and judgment in public health practice. Am J Epidemiol 1990; 132:9–16.
79. Mathieu E, Sodahlon Y. Epidemic Investigation. International Encyclopedia of Public Health. 2008:401–13. doi: 10.1016/B978-012373960-5.00183-0. Epub 2008 Aug 26. PMCID: PMC7149716.

Relevant websites used and for further reading:

1. https://www.lkouniv.ac.in/site/writereaddata/siteContent/202004032240235577Anupama_Singh_Mortality_and_its_measures.pdf
2. https://www-doh.state.nj.us/doh shad/view/sharedstatic/MaternalMortalityRate.pdf
3. Reference: https://www.britannica.com/science/case-fatality-rate
4. https://en.wikipedia.org/wiki/Life_table
5. https://en.wikipedia.org/wiki/Life_table
6. https://www.britannica.com/topic/epidemiologic-transition
7. https://en.wikipedia.org/wiki/Epidemiological_transition, Omran theory on epidemiological transition
8. https://www.google.com/search?source=univ&tbm=isch&q=epidemiologic+triad+of+disease&fir=GqQjWYo6xWN3pM%252CI36H1_ubYBQ-JM%252C_%253BnJkse5frwTznNM%252CI36H
9. https://www.atrainceu.com/content/2-understanding-chain-infection.
10. https://www.atrainceu.com/content/2-understanding-chain-infection.
11. https://www.google.com/search?q=naturalhistory+of+disease&oq=Natural&aqs=chrome.0.69i59l3j46i275i512j0i512j69i57j0i512l4.8162j1j15&sourceid=chrome&ie=UTF-8
12. https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/epidemic-theory
13. "Herd immunity | immunology". Encyclopedia Britannica. Retrieved 13 June 2021.
14. https://apic.org/monthly_alerts/herd-immunity/
15. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7149515/
16. http://www.flowminder.org/ – Flowminder (last accessed on 17.07.16.)
17. http://www.who.int/ihr/alert_and_response/outbreak-network/en/ – Global Outbreak Alert and Response Network (last accessed on 17.07.16.).
18. https://www.google.org/flutrends/about/ – Google Flu (last accessed on 17.07.16.)
19. http://www.emergency.cdc.gov/han/ – Health Alert Network (last accessed on 17.07.16.
20. https://www.verywellmind.com/what-is-a-cross-sectional-study-2794978, (How Does the Cross-Sectional Research Method Work?
21. https://www.medicalnewstoday.com/articles/281703.
22. https://www.healthknowledge.org.uk/e-learning/epidemiology/practitioners/introduction-study-design-ccs
23. https://www.google.com/search?q=experimental+studies+in+medicine&oq=experimental+studies+in+medicine&aqs=chrome..69i57j0i22i30l3.23244j0j15&sourceid=chrome&ie=UTF-8
24. https://scholar.google.com/scholar?q=cluster+randomized+trial+example&hl=en&as_sdt=0&as_vis=1&oi=scholart
25. https://www.explorepsychology.com/factorial-design-definition-examples.
26. https://geekymedics.com/sensitivity-specificity-ppv-and-npv/
27. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5701930/
28. https://wiki.ecdc.europa.eu/fem/Pages/Outbreak%20investigations%2010%20steps,%2010%20pitfalls.as

Self-Assessment Questions

1. Consideration in the selection of control group for a particular group of cases include all of the following EXCEPT:
   a. Assurance that information on study factors can be obtained from the control group in a manner similar to that by which it was obtained from the cases.
   b. Whether or not deliberately to select (match) the control in such a way as to make them similar to the cases with respect to certain confounding variables.
   c. The desirability that the controls derive from a population generally similar to that which gave rise to the cases.
   d. Practical and economic considerations.
   e. Wide, diverse control group.

2. Advantages of prospective study includes all the following EXCEPT it:
   a. Provides a direct estimate of the risk of developing disease B when A is present, whereas in the retrospective method can only be obtained indirectly.
   b. Decreases the risk of subjective bias, provided the criteria and the procedures are established in advance.
   c. Decreases the likelihood of misclassifying individuals with and without the characteristic.
   d. Is a quick way to collect data
   e. Pertains mainly to incidents of disease.

3. Which of the following do you consider to be the best definition of epidemiology?
   a. The study of epidemics
   b. The study of infectious diseases
   c. The study of chronic diseases epidemics
   d. The study of distribution and determinants of diseases in man
   e. The study of distribution and determinants of diseases and any health relate event in the population in man

4. In a study of 500 cases of a disease and 500 controls, the suspected etiological factor is found in 400 of the cases and 100 of the controls. The absolute risk (incidence) of disease in persons with the factor is:
   a. 80%
   b. 40%
   c. 16%
   d. 20%
   e. Cannot be calculated from the data given.

5. Epidemic refers to:
   a. A disease that has a low rate of occurrence but that is constantly present in a community or region.
   b. An attack rate in excess of 10 per 1000 population.
   c. The occurrence of illnesses of similar nature clearly in excess of the normal expectation for that population at that time.
   d. An attack rate in excess of 10 per 100,000 population.
   e. A disease that is constantly present in a country or region.

6. When a new treatment is developed that prevents death but does not produce recovery from a disease, the following will occur:
   a. Prevalence of disease will decrease.
   b. Incidence of the disease will increase.
   c. Prevalence of the disease will increase.
   d. Incidence of the disease will decrease.
   e. Incidence and prevalence of the disease will decrease.

A community of 100, 000 persons, during a certain year, there were 1000 deaths from all causes. All cases of tuberculosis were found, and they were 300(200 males and 100 females). During the same year, there were 60 deaths from tuberculosis, 50 of them in males. Use the above data for questions (7-11).

7. Crude mortality rate in the community is:
   a. 300 per 100, 000
   b. 60 per 1000
   c. 10 per 1000
   d. 100 per 1000
   e. Cannot be calculated from the given data

8. The proportionate mortality ratio due to tuberculosis is:
   a. 20%
   b. 30%
   c. 6%
   d. 3%
   e. None of the above

9. Case fatality rate for tuberculosis is:
   a. 6%
   b. 20%
   c. 2%
   d. Equal in males and females
   e. None of the above

10. Cause-specific mortality rate for tuberculosis is:
    a. 60 per 100, 000
    b. 300 per 100, 000
    c. 200 per 1000
    d. 20%
    e. None of the above

11. Sex-specific mortality rate for tuberculosis in males is:
    a. 0.5 per 1000
    b. 25%
    c. 5%
    d. 50 per 300
    e. Cannot be calculated from the given data

12. Communities X and Y have equal age-adjusted mortality rates. Community X has a lower crude mortality rate than community Y. One may conclude that:
    a. The two communities have identical age distributions.
    b. X has an older population than Y.
    c. Y has younger population than X.
    d. X has a younger population than Y.
    e. None of the above.

Use the data in the following 2 tables to answer questions 13 and 14.

Number of Persons Who Ate Each Specified Combination of Food Items

	Ate Chicken	**Did Not Eat Chicken**
Ate Fish	100	100
Did Not Eat Caviar	100	100

Number of Sick Persons with Gastroenteritis Who Ate Each Specified Combination of Food Items

	Ate Chicken	**Did Not Eat Chicken**
Ate Fish	50	20
Did Not Eat Caviar	50	25

13. What is the gastroenteritis attack rate in persons who ate both chicken and fish?
    a. 50/50
    b. 50/70
    c. 50/75
    d. 50/100
    e. 50/200

14. Which of the following food items (or combination of food items) is (are) most likely to be the infective item(s):
    a. Chicken only
    b. Fish only
    c. Neither chicken nor fish
    d. Both chicken and fish
    e. Cannot be known from the given data

Results of Screening Test for Diabetes Compared with Confirmatory Test Results

	Positive Confirmatory Test	**Negative Confirmatory Test**	**Total**
Positive Screening Test	34	20	54
Negative Screening Test	116	9830	9946
Total	150	9850	10000

15. Use the data in the above table and select the correct answer.
    a. Sensitivity = 34/150, specificity = 9830/9850, Predictive value positive = 34/54
    b. Sensitivity = 34/54, specificity = 20/9850, Predictive value positive = 34/150
    c. Sensitivity = 34/10000, specificity = 34/150, Predictive value positive = 9830/9946
    d. Sensitivity = 116/150, specificity = 9830/9850, Predictive value positive = 54/10000

16. All of the following are true for epidemiology except one:
    a. Is the basic science of public health
    b. Approaches problems systematically and quantitatively
    c. Uses rigorous techniques and methods to produce valid (accurate) and reliable (precise) findings
    d. Is concerned only with study of distribution of diseases

17. Core epidemiologic functions include:
    a. Public health surveillance
    b. Outbreak investigations
    c. Evaluation of disease control programs
    d. All the above

18. All of the following are true regarding the relation between incidence and prevalence except one:
    a. Prevalence is affected by incidence and duration
    b. If a disease has long duration, Prevalence ~= incidence
    c. If a disease has long duration, in general, Prevalence > incidence
    d. Prevalence tells us about the proportion of cases among the total population at any given time while incidence tells us the probability of a new onset of disease among those at risk for developing the illness

19. The ongoing, usual level of or constant presence of a disease within a given population or a geographic area is referred to as:
    a. Endemicity of disease
    b. Epidemicity of disease
    c. Secular trend of disease
    d. Surveillance of disease

20. All of the following are observational studies except one:
    a. Cohort study
    b. Case-control study
    c. Ecological study
    d. Clinical trials

21. All of the following are true regarding correlation studies except one:
    a. These studies provide a crude way of exploring associations between
    b. factors and disease.
    c. They are considered to be hypothesis generating rather than hypothesis testing.
    d. The group rather than the individual is the unit of comparison.
    e. They are expensive and time consuming to conduct.

22. Lack of temporality in cross-sectional studies means, one answer:
    a. This study design can measure prevalence rather than incidence
    b. It is unable most of the time to test the hypothesis
    c. Inability to identify sequence of events between exposure and disease necessary for drawing causal inferences.
    d. All the above

23. By using the table below, one can calculate the relative risk to be:
    a. 12
    b. 3.8
    c. 2.8
    d. None of the above

Exposure	**Diseased**	**Not Diseased**	**Total**	**Rate**
Yes	150	50	200	79.6%
No	20	80	100	14.3%
Total	170	130	300	61.3%

24. The attributable risk in prospective studies means, one answer:
    a. How much of the disease can we prevent by removing the exposure.
    b. The excess risk due to the exposure.
    c. How much of the disease in that group is “added” because of the exposure.
    d. All the above

25. All of the following are true for cohort studies except one:
    a. Clear temporal sequence of exposure and disease
    b. Well-suited for rare exposures
    c. Can measure incidence of disease
    d. Suitable for studying rare diseases

26. All of the following are true for case-control studies except one:
    a. They are well suited for rare diseases because you can enroll all of the cases.
    b. Because you are collecting information from cases and controls, you can ask about many different types of exposures.
    c. Case-control studies tend to require fewer subjects at entry than cohort studies.
    d. They are well-suited for rare exposures

27. All of the following are true for randomization in clinical trials except one:
    a. Allocation of each participant to the study groups by chance.
    b. Purpose of randomization, is to achieve baseline comparability between groups of both measured and unmeasured characteristics.
    c. Removes investigator bias in assigning patients to groups
    d. Increases validity of measurement by removing the awareness of the subject and the investigator.

28. Investigators in the past found an association between low altitude and cholera which proved to be later on as non-causal association. What was the indirectly associated factor?
    a. The indirect factor was the low altitude.
    b. The indirect factor was the fetid air.
    c. The indirect factor was the polluted water
    d. None of the above

29. A table shell is the (choose the correct answer):
    a. Skeleton of the table without labels or title.
    b. Table with data but without the title and labels.
    c. Table with labels and title but without the data.
    d. None of the above.

30. Which of the following may be useful in generating hypotheses in an outbreak setting? One answer
    a. Review of literature
    b. Look at descriptive epidemiology
    c. Look at the outliers
    d. All of the above

31. The key feature of an analytic study is (select one best answer):
    a. Analysis of data by person, time and place
    b. Calculation of a risk ratio or odds ratio
    c. Presence of a comparison group
    d. It is observational in nature

32. The most appropriate measure of association is (choose one answer):
    a. Attributable risk
    b. Odds ratio
    c. Chi-square
    d. Risk ratio (relative risk)

33. Misclassification of exposure, inappropriate comparison group , differences in participation rates are examples of (select one answer):
    a. Information bias
    b. Selection bias
    c. Confounders
    d. Measurement bias

34. All of the following are true for a confounder except (select one answer):
    a. It is associated with the outcome, independent of the exposure.
    b. It is associated with the exposure.
    c. It can result in overestimation or underestimation of true association.
    d. It works as effect modifier.

35. Using the data in the table below and choose the right answer.

	Disease Present	**Disease Absent**	**Total**
Positive Test	4000	1000	5000
Negative Test	500	9500	10000
Total	4500	10500	15000

    a. Sensitivity =80%, Specificity=95%, predictive value +ve=88.9%, predictive value –ve=90.5%
    b. Sensitivity =88.9%, Specificity=90.5%, predictive value +ve=80%, predictive value –ve=95%
    c. Sensitivity =80%, Specificity=95%, predictive value +ve=80%, predictive value –ve=90.5%
    d. Sensitivity =88.9%, Specificity=90.5%, predictive value +ve=95%, predictive value –ve=90.5%

A screening test for breast cancer was administered to 400 women with biopsy-proven breast cancer and to 400 women without breast cancer. The test results were positive for 100 of the proven cases and 50 of the normal women. Use these data for questions 36 and 37.

36. The sensitivity of the test is:
    a. 87%.
    b. 67%.
    c. 25%.
    d. 33%.
    e. 12%.

37. The specificity of the test is:
    a. 87%.
    b. 67%.
    c. 25%.
    d. 33%.
    e. 12%.

38. In a cross-sectional study of peptic ulcer in a community. Persons meeting the symptomatic criteria for peptic ulcer were found in 80 per 100000 men aged 35-49 and 90 per 100000 women aged 35-49.
    The inference that, in this age group, women are at a greater risk of developing peptic ulcer is:
    a. Correct.
    b. Incorrect due to failure to distinguish between incidence and prevalence.
    c. Incorrect because rates were used to compare males and females.
    d. Incorrect due to failure to recognize a possible cohort effect.
    e. Incorrect because there is no comparison or control group.

39. Epidemiological study of the roles of a suspected factor in the etiology of a disease may be observational or experimental. The essential difference between experimental and observational studies is that in experimental studies:
    a. The study and control groups are equal in size.
    b. The study is prospective.
    c. The study and control groups are always compatible.
    d. The investigator determines who shall be exposed to the suspected factor and who shall not.
    e. Controls are used.

40. Which of the following factors is the most important to the validity of the conclusions drawn from a clinical trial?
    a. Equal numbers of treated and placebo groups.
    b. Follow-up of 100% of the participants.
    c. Effective randomization of participants.
    d. A relatively high incidence of the disease in the population studied.
    e. Inclusion in both groups of individuals of all ages.

41. In a prospective study of a disease, the cohort originally selected consisted of:
    a. Persons who are found to have the disease.
    b. Persons without the disease.
    c. Persons with the factor under investigation.
    d. Persons with a family history of the disease.
    e. Persons without the factor under investigation.

42. A double-blind study of a vaccine is one in which:
    a. The study group receives the vaccine, and the control group receives a placebo.
    b. Neither observers nor subjects know the nature of the placebo.
    c. Neither observers nor subjects know which subjects receives the vaccine and which receives a placebo.
    d. Neither the study group nor the control group knows the identity of the observers.
    e. The control group does not know the identity of the study group.

43. Which of the following is an advantage of a retrospective study?
    a. There is little or no bias in assessment of exposure to the factor.
    b. Multiple disease outcomes following a selected exposure can be readily studied.
    c. Dependence on recall by subjects in the study is minimized.
    d. It is possible to determine the true incidence rate of the disease.
    e. It may be used to study etiology of a rare disease.

44. In 1945, 1000 women were identified wo worked in a factory painting radium dial on watches. The incidence of bone cancer in these women up to 1975 was compared to that of 1000 women who worked as telephone operators in 1945. 20 of the radium dial painters and 4 of the telephone operators developed bone cancer between 1945 and 1975. This study is an example of a:
    a. Prospective study.
    b. Experimental study.
    c. Clinical trial.
    d. Cross-section study.
    e. Retrospective study.

45. The following vaccine trial was performed: 1000 randomly selected children two years of age were given a vaccine against a certain disease and followed for 10 years. Of these, 80% were never afflicted with the disease. Which is the most correct conclusion regarding the efficacy of the vaccine?
    a. The vaccine is an excellent one because of the high rate of efficacy.
    b. No conclusion is possible, since no follow-up was made of non-vaccinated children.
    c. The vaccine is not very effective because it should have produced a higher efficacy rate.
    d. No conclusion is possible, since no test of statistical significance was performed.

A study of age versus prevalence of obesity resulted in the following data:

Age	**Prevalence%**
20-30	19
30-40	25
40-50	15
50-60	5

46. The inference that, as people grow older, they become thinner is:
    a. Correct.
    b. Incorrect because a rate is necessary to support the observation.
    c. Incorrect because no control or comparison group is used.
    d. Incorrect because no such conclusion should be made from cross sectional data.

47. The risk of acquiring a disease is measured by the:
    a. Incidence rate.
    b. Incidence rate times the average duration of the disease.
    c. Incidence rate divided by the prevalence rate.
    d. Prevalence rate.
    e. Prevalence rate times the average duration of the disease.

48. The strength of an association between a factor and a disease is best measured by:
    a. Prevalence over the incidence of the disease.
    b. Incidence of the disease.
    c. Prevalence of the factor.
    d. Attributable risk.
    e. Relative risk.

49. An investigation of an outbreak of diarrhea reveals that the proportion of cases eating in restaurant A was 85%, in restaurant B was 15%, restaurant c was 55%, and the proportion consuming public water was 95%. Which of the following conclusions is valid?
    a. The source is restaurant A because it has the highest proportion of cases among the restaurants.
    b. The source is not restaurant B because it has the lowest rate.
    c. The source is the water supply because it has the highest proportion of cases.
    d. The source could be either restaurant A or restaurant B or the water supply.
    e. No definite conclusion can be made because there is no comparison between those exposed and those not exposed.

50. Case fatality rate for a given disease refers to:
    a. The crude death rate per 1000 population.
    b. Cause-specific death rate due to the disease.
    c. A fatal outcome of any disease.
    d. The percentage of deaths among cases of the disease.
    e. The proportion of deaths due to a disease among all deaths from all causes.

51. An investigator is interested in the etiology of neonatal jaundice. To study this condition, he selected 100 children who were diagnosed with this condition and 100 children born in the same time period and in the same hospital who did not have a diagnosis of neonatal jaundice. He then reviewed the obstetric and delivery records of their mothers to determine various prenatal and perinatal exposures. This is an example of a:
    a. Cross-sectional study.
    b. Retrospective study.
    c. Prospective study.
    d. Clinical trial.
    e. Ecological study.

52. Herd immunity refers to:
    a. Immunity naturally acquired in a population.
    b. Vaccination of domestic animals to prevent disease transmission to human.
    c. Genetic resistance to species-specific disease.
    d. The prevention of disease transmission to susceptible individuals through acquired immunity in others.
    e. The high levels of antibody present in a population following an epidemic.

53. Cases of group A streptococcal disease were reported recently in a certain area. To determine whether or not an outbreak was occurring, one would need to know all of the following EXCEPT:
    a. Incubation period of the disease.
    b. The usual pattern of this disease.
    c. Reporting practices.
    d. Diagnostic accuracy.
    e. A case definition.

54. An official from the ministry of health visits emergency room in a hospital to determine the number of cases of post exposure prophylaxis for rabies. The official's action is an example of:
    a. Case finding.
    b. Secondary prevention.
    c. Active surveillance.
    d. Outbreak investigation.
    e. Screening.

55. Screening program is an example of:
    a. Primary prevention.
    b. Secondary prevention.
    c. Primordial prevention.
    d. Tertiary prevention.

56. The variable that must be associated with both the exposure and the outcome is:
    a. Effect modifier.
    b. Confounder.
    c. Intervening variable.
    d. Necessary cause.

57. The variable that alters the nature of a true relationship between an exposure and an outcome is:
    a. Effect modifier.
    b. Confounder.
    c. Intervening variable.
    d. Necessary cause.

58. The relative risk for lung cancer in smokers is X, and it is Y in asbestos workers, and the risk in those with both exposures is XY. This is an example of:
    a. External validity.
    b. Internal validity.
    c. Synergism.
    d. Virulence.

59. A study population resembles the larger population from which it was drawn, this is called:
    a. Internal validity.
    b. External validity.
    c. Biological plausibility.
    d. Synergism.

60. Study results are obtained in an unbiased manner, this is called:
    a. Internal validity.
    b. External validity.
    c. Biological plausibility.
    d. Synergism.

61. You are interested to estimate the level of measles immunity in your study population and derive data useful for generating vaccination policy. You should:
    a. Conduct a case-control study.
    b. Conduct a randomized trial of measles vaccination.
    c. Conduct a retrospective cohort study of measles vaccination.
    d. Conduct a cross-sectional survey of vaccination status.

62. In a case control study of myocardial infarction, which group of subjects would be a poor choice as controls?
    a. Subjects with similar age distribution as cases.
    b. Subjects with similar sociodemographic characteristics as cases.
    c. Subjects with similar cardiac risk factors as cases.
    d. Subjects admitted to hospital for non-cardiac disease.

63. Performing carotid endarterectomy in a patient with transient ischemic attacks, Is an example of:
    a. Primary prevention.
    b. Secondary prevention.
    c. Tertiary prevention.
    d. Health promotion.

64. Recommending regular physical activity to a person with no known medical problems, is an example of:
    a. Primary prevention.
    b. Secondary prevention.
    c. Tertiary prevention.
    d. Health promotion.

65. A screening program detects lung cancer early. The survival time in those screened is 3 months longer than in those not screened who present with symptoms. This difference is likely due to:
    a. Length bias.
    b. Effect modification.
    c. Lead-time bias.
    d. Observer bias.

	Disease	**No disease**	**Total**
Positive Screening Test	A	b	a+b
Negative Screening Test	C	d	c+d
Total	a+c	b+d	a+b+c+d

66. In the table above, False-positive rate of screening test equal:
    a. b/b+d
    b. c/a+c
    c. b/a+b
    d. c/c+d

67. The probability of disease in a patient with negative screening test in the table above is:
    a. a/a+c
    b. c/a+c
    c. c/c+d
    d. d/c+d

68. Among elderly subjects who are fit, vigorous exercise reduce the risk of heart disease. Among elderly subjects who are unfit, the initiation of vigorous exercise might precipitate a myocardial infarction. So, fitness may be considered:
    a. A risk factor.
    b. An effect modifier.
    c. A confounder.
    d. A necessary cause

69. In a cohort study, the groups are compared on the basis of:
    a. Disease status.
    b. Exposure status.
    c. Inclusion criteria.
    d. Exclusion criteria.

70. Which one of the following examples might represent a retrospective cohort study?
    a. Cases of lung cancer assessed for prior exposures.
    b. Subjects with current angina followed for the development of myocardial infarction.
    c. Subjects with prior radiation exposure followed for the development of lymphoproliferative cancers.
    d. Subjects with skin cancer assessed for life long cumulative sun exposure.

71. In the definition of epidemiology, "determinants" generally includes all of the following EXCEPT:
    a. Agents
    b. Causes
    c. Control measures
    d. Risk factors

72. The hallmark feature of an analytic epidemiologic study is: (Choose one best answer):
    a. Use of an appropriate comparison group.
    b. Laboratory confirmation of the diagnosis
    c. Publication in a peer-reviewed journal
    d. Statistical analysis using logistic regression

73. When analyzing surveillance data by age, which of the following age groups is preferred? (Choose one best answer)
    a. 1-year age groups
    b. 5-year age groups
    c. 10-year age groups
    d. Depends on the disease

74. A study in which children are randomly assigned to receive either a newly formulated vaccine or the currently available vaccine, and are followed to monitor for side effects and effectiveness of each vaccine, is an example of which type of study?
    a. Observational
    b. Cohort
    c. Case-control
    d. Clinical trial

75. A cohort study differs from a case-control study in that:
    a. Subjects are enrolled or categorized on the basis of their exposure status in a cohort study but not in a case-control study.
    b. Subjects are asked about their exposure status in a cohort study but not in a case control study.
    c. Cohort studies require many years to conduct, but case-control studies do not.
    d. Cohort studies are conducted to investigate chronic diseases, case-control studies are used for infectious diseases.

76. The epidemiologic triad of disease causation refers to: (Choose one best answer)
    a. Agent, host, environment.
    b. Time, place, person.
    c. Source, mode of transmission, susceptible host.
    d. Agent, reservoir, mode of transmission.

77. All of the following pertain to cross-sectional study EXCEPT:
    a. It usually provides information on prevalence rather than incidence
    b. It is more useful for descriptive epidemiology than it is for analytic epidemiology.
    c. It is synonymous with survey.
    d. Temporality is preserved.

78. Disease control measures include the following:
    a. Eliminating the reservoir
    b. Eliminating the vector
    c. Interrupting mode of transmission
    d. Reducing host susceptibility
    e. All of the above.

79. A propagated epidemic is usually the result of what type of exposure?
    a. Point source.
    b. Continuous common source.
    c. Intermittent common source.
    d. Person-to-person.

80. Number of women in Country A who died from lung cancer in 2004 estimated NUMBER OF women living in Country A on July 1, 2004. This is an example of
    a. Proportionate mortality ratio.
    b. Cause-specific mortality rate among women.
    c. Crude mortality rate among women.
    d. Sex-specific mortality rate.

Investigators enrolled 100 diabetics without eye disease in a cohort (follow-up) study. The results of the first 3 years were as follows:
Year 1: 0 cases of eye disease detected out of 92; 8 lost to follow-up
Year 2: 2 new cases of eye disease detected out of 80; 2 had died; 10 lost to follow-up.
Year 3: 3 new cases of eye disease detected out of 63; 2 more had died; 13 more lost to follow-up.

81. Based on the above information, the person-time incidence rate is calculated as:
    a. 5/100=50 per 1000 person-years of observation.
    b. 5 / 63=79.4 per 1000 person-years of observation.
    c. 5 / 23=217.4 per 1000 person-years of observation.
    d. 5/250=20 per 1000 person-years of observation.

Within 10 days after attending a wedding, an outbreak of a certain disease occurred among attendees. Of the 83 guests and wedding party members, 79 were interviewed; 54 of the 79 met the case definition. The following two-by-two table shows consumption of wedding cake and illness status.

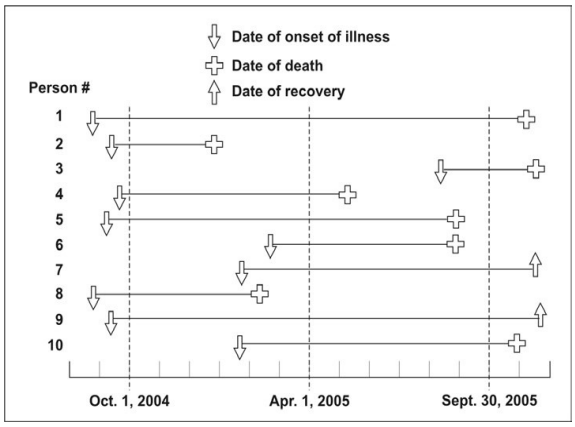
	Ill	**Not Ill**	**Total**
Ate Cake	50	3	53
Did Not Eat Cake	4	22	26
Total	54	25	79

82. From the data in the above table, the fraction 54/79 is a/an:
    a. Attack rate among persons who ate the cake.
    b. Overall attack rate among persons who attended the wedding.
    c. Attack rate among persons who did not eat cake.
    d. Relative risk.

83. What is the following fraction?
    a. Number of deaths due to septicemia among men aged 65–74 years in 2004 Estimated number of men aged 65–74 years alive on July 1, 2004
    b. Age- cause-specific mortality rate.
    c. Age- cause-sex-specific mortality rate.
    d. Cause-specific mortality rate.
    e. Sex-specific mortality rate

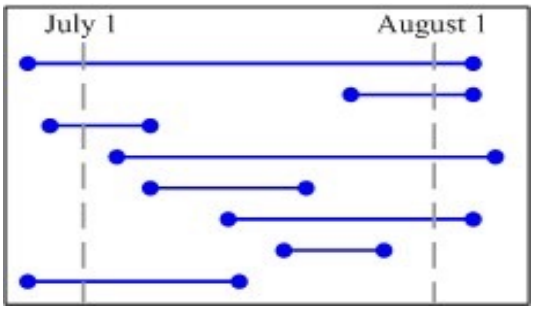
84. Depending on the data in the below diagram, which choice is incorrect from the following:
    a. Incidence rate from October 1, 2004, to September 30, 2005 = (4 / 18) x 100 = 22 new cases per 100 population.
    b. The point prevalence on April 1, 2005 = (7 / 18) x 100 = 38.89%.
    c. The period prevalence from October 1, 2004, to September 30, 2005 = (10 / 18) x 100=55.6 per 100 population.
    d. The period prevalence from October 1, 2004, to September 30, 2005 = (10 / 20) x 100 = 50 per 100 population.

Cases of Illness from October 1, 2004–September 30, 2005



85. In a state that did not require varicella (chickenpox) vaccination, a boarding school experienced a prolonged outbreak of varicella among its students that began in September and continued through December. To calculate the probability or risk of illness among the students, which denominator would you use?
    a. Number of susceptible students at the ending of the period (i.e., June)
    b. Number of susceptible students at the midpoint of the period (late October/early November)
    c. Number of susceptible students at the beginning of the period (i.e., September)
    d. Average number of susceptible students during outbreak

Use the following diagram for Questions86 and 87. Assume that the horizontal lines in the diagram represent duration of illness in 8 different people, out of a community of 700.



86. What is the prevalence of disease during July?
    a. 3/700
    b. 4/700
    c. 5/700
    d. 8/700

87. What is the incidence of disease during July?
    a. 3/700
    b. 4/700
    c. 5/700
    d. 8/700

88. For the formula below, which choice is right?
    $$\frac{\text{Number of children < 365 days of age who died in Country A in 2004}}{\text{Number of live births in Country A in 2004}}$$
    a. Child mortality rate
    b. Post neonatal mortality rate
    c. Neonatal mortality rate
    d. Infant mortality rate

89. Vaccine efficacy measures are:
    a. The proportion of vaccinees who do not get the disease
    b. The attack rate among vaccinees
    c. The proportionate reduction in disease among vaccinees
    d. Disease attributable to the vaccine

90. The best time to create table shells is:
    a. Just before planning a study
    b. As part of planning the study
    c. Just after collecting the data
    d. Just before analyzing the data
    e. As part of analyzing the data

91. Public health surveillance includes which activities?
    a. Data collection
    b. Data analysis
    c. Data interpretation
    d. Data dissemination
    e. All of the above

92. Current public health surveillance targets which of the following?
    a. Chronic disease
    b. Communicable diseases
    c. Health-related behaviors
    d. Presence of viruses in mosquitoes
    e. All of the above

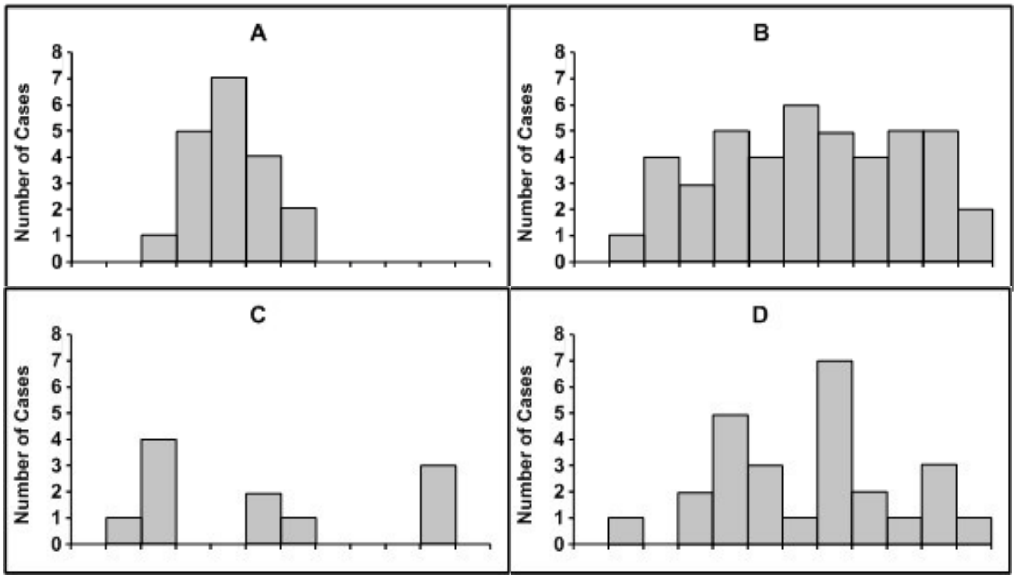
93. Criteria for prioritizing health problems for surveillance include the following except?
    a. Incidence of the problem
    b. Public concern about the problem
    c. Number of previous studies of the problem
    d. Social and economic impact of the problem

94. For an investigation of an outbreak, what is the logical conceptual order of the steps listed below?
    a. 1-2-3-4-5-6-7
    b. 5-6-4-1-2-3-7
    c. 6-5-3-1-2-7-4
    d. 6-5-7-4-1-3-2

- Analyze data by time, place, and person
- Conduct a case-control study
- Generate hypotheses
- Conduct active surveillance for additional cases
- Verify the diagnosis
- Confirm that the number of cases exceeds the expected number
- Talk with laboratorians about specimen collection

95. A case definition during an outbreak investigation should specify all the following Except:
    a. Clinical features
    b. Time
    c. Place
    d. Person
    e. Hypothesized exposure

96. The epidemic curves below represent which of the following choices?



    a. A represents common source epidemic, B represents continuous common source epidemic
    b. A represents common source epidemic, D represents continuous common source epidemic
    c. D represents continuous common source epidemic, C represents intermittent epidemic
    d. D represents propagated

97. All of the following are usually used in generating hypotheses in an outbreak setting EXCEPT?
    a. Review the literature
    b. Look at the descriptive epidemiology
    c. Look at the outliers
    d. Talk with the local health authorities
    e. Conducting a case-control study

98. Possible explanations for a case that occurs substantially later than the other cases in an outbreak include all the following EXCEPT:
    a. Similar but unrelated disease
    b. Secondary case
    c. Case with unusually long incubation period
    d. Time of exposure later than others
    e. Index case

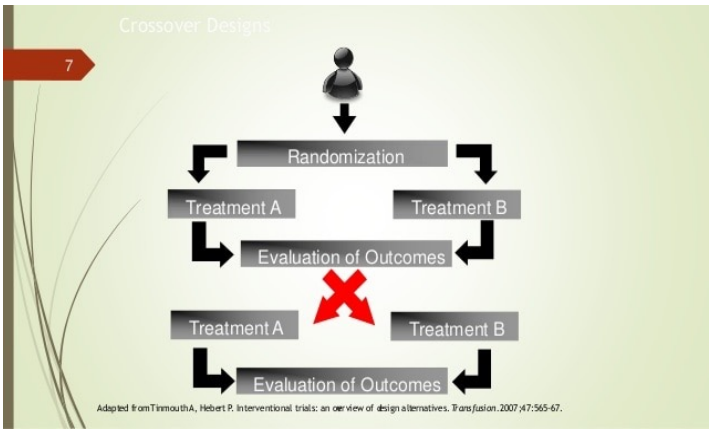
99. All of the following are true for total fertility rate EXCEPT:
    a. It is one of the cohort measures of fertility.
    b. It is a summary measure of age-specific fertility rates.
    c. It refers to the average total number of children a woman will produce during her childbearing age span.
    d. It is calculated by the following formula:
        • Number of live births during a year/Mid-year female population aged 15-49 □ 1000.

100. Regarding epidemiological transition, all the following statements are true EXCEPT:
    a. A theory which "describes changing population patterns in terms of fertility, life expectancy, mortality, and leading causes of death."
    b. Is the process by which the pattern of mortality and disease in a population is transformed from one of high mortality among infants and children and episodic famine and epidemics affecting all age groups to one of degenerative and human-made diseases (such as those attributed to smoking) affecting principally the elderly.
    c. The 20th- and 21st-century declines in mortality in developing countries have been far more rapid than those that occurred in the 19th century in the industrialized countries.
    d. It is the study of distribution and determinants of diseases or health-related problems in the population.

101. The purposes of random allocation in clinical trials are the following EXCEPT:
    a. To achieve baseline comparability between the two groups (two arms of the study) of both measured and unmeasured characteristics, so difference in outcome can be attributed to difference in the intervention.
    b. To remove investigator bias in assigning patients to groups.
    c. To increase validity of statistical tests.
    d. Investigator(s) and/or study participants are kept ignorant of the group to which participants are assigned.

102. Advantages of quasi-experimental design include the following EXCEPT:
    a. Randomization is a must in this study.
    b. They are more frequently used because they are more practical and feasible to conduct.
    c. It is more suitable for real-world natural setting than true experimental research design.
    d. It allows researchers to evaluate the impact of independent variables under naturally occurring conditions.

103. The diagram below represents which design of clinical trials?
   a. Parallel design
   b. Cluster design
   c. Cross-over design
   d. Factorial design



104. All of the following are types of information bias except:
   a. Questionnaire faults
   b. Recall bias
   c. Observer error or interviewer bias
   d. Non-response bias

105. Breastfeeding and diarrhea: lack of breastfeeding is a huge problem ininfants younger than 1 month, less so for infants 1 month or older. So, age is considered here as:
   a. A confounder
   b. An effect modifier
   c. A random error
   d. All of the above

106. The second step in the development of a research proposal is the “Selection and formulation of research problem.” This step includes all of the following EXCEPT:
   a. Choose an appropriate research topic.
   b. Define the nature, extent and significance of the problem.
   c. Frame specific research questions and the possible value of seeking answers to these questions.
   d. State research objectives – immediate and ultimate.
   e. Critical appraisal of existing information.

107. The progress of a disease process in an individual over time, in the absence of intervention refers to:
   a. Chain of infection
   b. Prognosis of the disease
   c. Natural history of the disease
   d. Iceberg phenomenon

108. All of the following refers to herd immunity EXCEPT:
   a. Is a form of indirect protection from infectious disease that can occur with some diseases when a sufficient percentage of a population has become immune to an infection.
   b. Immune individuals are unlikely to contribute to disease transmission, disrupting chains of infection, which stops or slows the spread of disease.
   c. It results only from vaccinating a high proportion of the population against a certain disease.
   d. Herd immunity threshold (HIT) varies depending on the basic reproduction number of the disease.

109. The following are true for basic reproductive number (R0) EXCEPT:
   a. It is affected by the rate of contacts in the host population.
   b. It is affected by the probability of infection being transmitted during contact.
   c. It is affected by the duration of infectiousness.
   d. It is usually smaller than effective reproductive number.
   e. It is used to measure the transmission potential of a disease.

110. Which of the following is true for aggregate surveillance?
   a. It is the surveillance of a disease or health event by collecting summary data on groups of cases but lacks detailed information on specific cases.
   b. It collects detailed information on specific cases.
   c. It involves monitoring cases that meet a clinical case definition for the disease under surveillance, typically without laboratory confirmation.
   d. It can monitor the number of cases and the individual-level data required for specific analyses.

111. Qualitative epidemiological study is characterized by the following except one:
   a. Needs usually small sample size but studied in-depth.
   b. Involves a-priori sample size calculation.
   c. Based usually on purposive sampling methods.
   d. Assumes existence of dynamic and multiple realities.
   e. Doesn’t require predetermined and rigid design.

112. Which of the following is true for operational study?
   a. Operational study is different from other epidemiological studies in that it examines a system or a program rather than focusing on an individual or a group of individuals.
   b. The researcher should spell out well-defined goals and objectives of the health program or the system in question.
   c. The researcher should identify, prioritize constraints and obstacles that prevent the program objectives being achieved.
   d. It requires close interaction between program managers and researcher.
   e. All of the above.

Answers to self-assessment questions

Question number	Answer	Question number	Answer	Question number	Answer
1	e	43	e	81	d
2	d	44	a	82	b
3	e	45	b	83	b
4	e	46	d	84	c
5	c	47	a	85	c
6	c	48	e	86	d
7	c	49	e	87	c
8	c	50	d	88	d
9	b	51	b	89	c
10	a	52	d	90	b
11	e	53	a	91	e
12	d	54	c	92	e
13	d	55	b	93	c
14	a	56	b	94	d
15	a	57	a	95	e
16	d	58	c	96	a
17	d	59	b	97	e
18	b	60	a	98	e
19	a	61	d	99	d
20	d	62	c	100	d
21	d	63	c	101	d
22	c	64	d	102	a
23	b	65	c	103	c
24	d	66	a	104	d
25	d	67	c	105	b
26	d	68	b	106	e
27	d	69	a	107	c
28	b	70	c	108	c
29	c	71	c	109	d
30	d	72	a	110	a
31	c	73	d	111	b
32	d	74	d	112	e
33	b	75	a		
34	d	76	a		
35	b	77	d		
36	c	78	e		
41	b	79	d		
42	c	80	b		

SYNOPSIS IN

GENERAL EPIDEMIOLOGY

First Edition, 2024

SYNOPSIS IN **GENERAL EPIDEMIOLOGY**

This book consists of 10 chapters that cover introduction to epidemiology, epidemiologic transition, various types of epidemiological studies, errors in epidemiological studies, association and causality, design of epidemiological research, screening in public health, general concepts of disease occurrence, public health surveillance, and investigation and control of epidemics.

Examples and exercises were included in each chapter to simplify the texts and make it easily understandable. This little book fits under and post- graduate students studying in various public health fields. It also suits even all types of health workers and health practitioners of different specialties to get oriented and familiar with the general principles of disease occurrence, methods of epidemiological studies, public health surveillance and investigation of epidemics.

In contrast to other introductory books, special and detailed emphasis was devoted to enable various public health workers including the general practitioners to have the essential skills to be efficient members in the epidemiologic investigation team and apply their skills in the conduction of public health surveillance and investigation of epidemics. Figures are used extensively to promote comprehension and retention of the material. Also, it contains at the end, a list of self-assessment questions with their answers to help the reader better understand the different subjects and to give the students an insight about their preparation to enter examinations of epidemiology.

9 789923 012079