
Preface

This book offers a straightforward guide to medical research methods, tailored for undergraduate, postgraduate, PhD scholars, faculty and medical practitioners. It bridges the gap between theoretical knowledge and practical application.

The chapters cover core epidemiological methods, explaining how to study disease trends. Biostatistics is presented with the understanding that complex formulas can be used to make data analysis easy to understand. We emphasize research ethics and practical study design, which are essential for conducting ethical medical research.

We also respect the copyright and wisdom of the authors of these guidelines and research aspects. The similarities between certain words and sentences are to keep the meaning of those important guidelines, methods, and definitions the same as intended in those guidelines and documents. All the important guidelines and resource materials are included in the reference and further reading section. Readers are requested to go through them to further improve skills in epidemiology, research ethics, and biostatistics for becoming responsible and ethical researchers. The publishers, with their extensive experience in ethical and responsible publishing business, have guided us to ensure that copyrights, warranties and obligations related to standard ethical, responsible publications are not violated.

We have also used AI tools as writing assistants to improve the quality of sentences, grammar, and English language. We have checked the correctness of all the content to the best of our abilities, based on the experience acquired over many years in research, research ethics, and publications.

We would also like to thank our teachers, university management, university administration, colleagues, students, participants of various epidemiology workshops, researchers and our national and international collaborators who have shared their knowledge and wisdom over the years. Their contributions have helped us gain the experience and skills that have been translated into this book.

The famous first words of Neil Armstrong upon stepping onto the Moon on July 20, 1969, were “That’s one small step for [a] man, one giant leap for mankind.” This quote reflects this book as an important step for ethical and responsible researchers.

It is an example-oriented book, using scenarios from healthcare settings, published articles, and simulated examples that are valuable for understanding local research contexts.

From formulating research questions to the final steps of publication, this book covers all aspects of the research process. It serves as a one-stop resource for understanding study designs, including clinical trials, and for handling data confidently. We hope this work improves the quality of health research and supports evidence-based, ethical medical practice.

**A Meriton Stanly
Pankaj B Shah
BWC Sathiyasekaran**

Contents

1. Introduction to Epidemiology: As a Basis for Research Involving Human Participants	1
<i>BWC Sathiyasekaran</i>	
2. Measures of Disease Frequency	7
<i>BWC Sathiyasekaran</i>	
3. Epidemiological Studies and Methods in Epidemiology	14
<i>AK Amritha Lekha</i>	
4. Cross-Sectional Studies	22
<i>Vanishree Shriraam</i>	
5. Case Control Studies.....	30
<i>Samya V</i>	
6. Cohort Studies	52
<i>Akila GV</i>	
7. Experimental Studies	66
<i>A Meriton Stanly</i>	
8. Bias in Epidemiological Studies.....	73
<i>Suresh Varadarajan</i>	
9. Sampling Methods.....	83
<i>M Anitha Rani</i>	
10. Sample Size Determination	95
<i>Muthukumar R</i>	
11. Questionnaire-Designing	103
<i>Pragadeesh P</i>	
12. Diagnostic Test Accuracy/Characteristics	117
<i>BWC Sathiyasekaran</i>	

13. Review of Literature.....	127
<i>Susithra R</i>	
14. Writing Research Proposal	135
<i>L Kannan</i>	
15. Test of Statistical Significance	140
<i>T Gayathri</i>	
16. Association, Correlation, and Simple Linear Regression	152
<i>T Gayathri</i>	
17. Normal Distribution and Basis for Statistical Inference Including p-Value and 95% Confidence Interval	159
<i>M Anitha Rani</i>	
18. Artificial Intelligence in Medical Research: Practical Applications and Integration with Digital Research Tools	170
<i>A Aswin</i>	
19. Ethical Guidelines for Biomedical and Health Research Involving Human Participants	180
<i>Pankaj B Shah</i>	

1

Introduction to Epidemiology: A Basis for Research Involving Human Participants

BWC Sathiyasekaran

Dean Research, Sree Balaji Medical College and Hospital, Chennai

Keywords: Epidemiology; Disease distribution; Disease determinants; Public health; Epidemiological approach

INTRODUCTION

Epidemiology is the study of how diseases occur and vary among human populations, and the reasons behind these variations. Almost every disease shows differences in its pattern of occurrence. Such differences may arise due to factors like age, sex, occupation, socioeconomic condition, water supply, environmental surroundings, and socio-cultural practices of the people. Many other related factors can also influence disease occurrence. The primary purpose is to observe variations and analyze the possible reasons. This helps in understanding the causes and mechanisms involved in the development and spread of diseases.

Cholera, HIV/AIDS, and SARS can be quoted as examples of the remote past and recent past in which, even before the actual causative agent was identified, epidemiologists proposed the possible causal mechanisms of the disease by observing the occurrence of the disease and personal characteristics of affected, place of occurrence of the disease, and in terms of time of occurrence of the disease.

The above statement can also be extended to the outcome of the disease. Epidemiology is also the study of variations in the outcome of disease and why such variations should occur.¹ Even when patients with a particular disease are given the same line of treatment, variations occur in the outcome of the disease. Careful observation of these variations and why these variations occur helps in improving diagnostic procedures and better management of the patients (clinical decision making). Clinical epidemiology includes the study of variations in the outcome of disease.

Epidemiology is the scientific foundation for the practice of Public Health.² It is also a basic science for Community Medicine and Clinical Medicine. There is a separate branch in epidemiology called clinical epidemiology. The word epidemiology is derived from ‘*Epidemic*’ which, translated from Greek, means ‘Upon the people’. Epidemiology is concerned more with populations or groups of people rather than individuals. The groups of people may include large populations (the whole district/country) or small groups of people (just 5 to 10 individuals).

Unlike in clinical medicine, the application of epidemiology is not confined to just one individual person or patient.

As mentioned above, the word “epidemiology” means “study of people.” “Epi” stands for “upon” as in epidermis and epididymis. “Demo” stands for “people: as in democracy and demography. “Ology” stands for “study,” as in theology and zoology. Hence, epidemiology is the basis for any study/research involving human participants (earlier known as “study on human subjects,” which to some extent sounds like an unethical term). All studies/research where human beings are involved as study participants, whether observational studies or interventional studies or studies/research on human diseases or any health-related conditions, come under the purview of epidemiology or epidemiological studies. Even social science studies, because it also involves human participants, can be considered under the broad umbrella of epidemiology though social science studies are usually considered as a separate entity by itself.

BRIEF HISTORY OF EPIDEMIOLOGY

The word epidemiology is an ancient word, maybe as old as the term medicine. Epidemiology has roots in the Bible.³ Hippocrates in the fifth century B.C. suggested that human disease might be related to environmental factors.⁴ An important landmark in epidemiology occurred after a span of about 2000 years (1662) when John Graunt in London, for the first time quantified patterns of diseases in a population. It took many years for someone to think on similar lines, and after a span of about 200 years (1839), William Farr applied vital statistical data to evaluate health problems. Another important landmark in epidemiology happened just about 20 years later (1854) when John Snow investigated the cholera epidemic in Broad Street, London. Now, comparing rates of diseases in sub-groups is more common.⁵

DEVELOPMENTS IN MODERN EPIDEMIOLOGY

In 1950, Doll and Hill published one of the early case-control studies of cigarette smoking and lung cancer.⁶ Around the same time, one of the early cohort studies was started in Framingham, “The Framingham Heart Study”.⁷ Again in 1954, a large field trial of the Poliomyelitis vaccine was taken up.⁸ The emergence of non-communicable diseases as major public health problems in developed countries led to the development of new applications in epidemiology. During this period, research methods such as case-control studies and cohort studies became widely used for understanding the causes and patterns of diseases. At the same time, newer approaches were introduced to evaluate the effectiveness of health interventions and preventive measures. These advancements brought significant changes in the field and marked the beginning of modern epidemiology.⁵ Modern epidemiology has further evolved with the addition of many new techniques, methods, and applications in epidemiology, aided by the availability of advanced statistical analysis and high-speed computer facilities.

Epidemiology, which was traditionally concerned with studies of communicable diseases occurring in epidemic proportions, evolved over the years to include all communicable diseases and was further adopted for studies on non-communicable diseases, conditions, interventions, and health programs, and is now applicable in all branches of medicine.

CLINICAL EPIDEMIOLOGY

Clinical epidemiology is considered an important scientific foundation for clinical medicine. It involves the use of epidemiological principles and methods to study and solve problems related

to patient care and clinical practice. Clinical medicine and epidemiology originally developed together, and many early epidemiologists were also practicing clinicians. For instance, John Snow, who is regarded as one of the pioneers of epidemiology, was an anesthetist by profession. However, over the past few decades, the two fields have gradually become separate in their approach and practice. In recent years, there has been growing recognition among clinicians and epidemiologists that both disciplines are closely connected. This understanding has contributed to the development and importance of clinical epidemiology in modern medicine.

The word “clinical” is derived from the Greek word “*Kline*,” which means “bed.” Clinical medicine is often referred to as “bedside medicine.” Though clinicians take care of many patients, the word “clinical” specifically implies care of a patient at the individual level. Though the terms clinical and epidemiology are not compatible with each other (one mostly concerned with care of individual patients and the other mainly concerned with groups of people), the term clinical epidemiology has been in use for more than 50 years and has gained worldwide acceptance. As described earlier, “Clinical Epidemiology” improves patient care (diagnosis and treatment). The scope of clinical epidemiology would include community medicine, clinical decision-making, clinical economics, health services research and related disciplines.

The term “Clinical Epidemiology” was first introduced by John Paul in 1938 as a new basic science for preventive medicine,⁹ emphasizing population rather than individual patient orientation. Over the period, the focus of clinical epidemiology shifted from a community perspective to individual patients. Sackett in 1969 defined clinical epidemiology as “the application, by a physician who provides direct patient care, of epidemiologic and biostatistical methods to the study of diagnostic and therapeutic processes in order to effect an improvement in health”. Later Alvan Feinstein defined clinical epidemiology as follows, “clinical epidemiology represents the way in which classical epidemiology, traditionally oriented toward general strategies in the public health of community groups, has been enlarged to include clinical decisions in personal encounter care for individual patients”. In recent times, clinical epidemiology is increasingly involved in the evaluation of diagnostic tests, predicting patient prognosis, conducting randomized controlled trials, and in evidence-based medicine.

BRANCHES IN EPIDEMIOLOGY

The traditional methods used in epidemiology have also been extended further and applied in different branches of medical research and allied sciences. This resulted in newer methods and techniques in epidemiology, and new branches of epidemiology developed. Each branch is a specialty by itself. Epidemiology has developed into several specialized branches that focus on different aspects of diseases and health conditions. Infectious disease epidemiology studies the spread and control of communicable diseases, while chronic disease epidemiology mainly deals with long-term illnesses such as diabetes, hypertension, and heart disease. Clinical epidemiology applies epidemiological methods in patient care and clinical practice. Environmental epidemiology examines the effects of environmental factors on human health, whereas occupational epidemiology studies diseases and health problems related to the workplace. Genetic epidemiology focuses on role of heredity and genetic factors in disease occurrence. Sero-epidemiology involves study of diseases through blood serum and antibody testing. Cancer epidemiology analyzes the causes and distribution of different types of cancers, while neuro-epidemiology concentrates on disorders related to the nervous system.

The branching out of epidemiology and diversification has resulted in the invention of newer methods, techniques and applications of epidemiology to suit the needs of different specialties, thereby enriching “Epidemiology” overall.

DEFINITIONS OF EPIDEMIOLOGY

There are many definitions for epidemiology, which reflect the evolution of epidemiology over the years. The following are the 3 important definitions made over a period of time:

1. Epidemiology is that branch of medical science that treats Epidemics (Parkin 1873).¹⁰
2. Epidemiology is the study of the distribution and determinants of disease frequency in man (Mac Mahan 1960).¹⁰
3. Epidemiology is the study of distribution and determinants of health-related States or events in specified populations and the application of this study to the control of health problems. (John M Last 1988).¹¹

As already described, epidemiology was initially used in the study of diseases which occurred in epidemic forms, later extended to study of all diseases communicable and non-communicable, acute and chronic, and recently further extended to studies of even health systems, and this evolution of epidemiology is well reflected by different definitions of epidemiology given by different authors over a period of time. The three definitions reflect the evolution of the concepts and applications of epidemiology over the years.

In most definitions of epidemiology, three important terms are commonly used either directly or indirectly. These terms are determinants, distribution, and disease frequency. In the definition given by John M. Last, the words “states” and “events” are used to describe disease frequency. Here, “states” refer to prevalence, while “events” indicate incidence. These concepts are considered basic and essential in epidemiology because they help in understanding how diseases occur, spread, and affect populations.

DETERMINANTS OF DISEASE

Agent factors are the ones that are directly involved in the causation of a disease. It is much easier to identify the agent factor in communicable diseases compared to non-communicable diseases.

Example 1: The cholera bacilli as an agent factor in cholera. In non-communicable diseases, the agent factor is much more difficult to identify and by itself may be made up of many components, for example, as in the case of coronary heart disease.

For a disease to be caused, there needs to be a close interplay between the agent factor, host factors and the environmental factors which is also called as epidemiological triad in communicable diseases.

Example 2: The mere presence of *Vibrio cholerae* as agent factor in water will not result in cholera unless a person drinks this water and even if a person drinks this water, cholera will not occur unless the person is susceptible to the infection (host factor) and no one will drink this contaminated water unless by ignorance, indifference or by other compelling environmental or social factors. A close interaction between the agent factors, host factors, and environmental factors is essential for the causation of any disease, either communicable or non-communicable.

Examples of agent factors are infective agents, nutritional agents, carcinogens, allergens, etc. Age, sex, race, genetic makeup, immunity, physiological state, nutritional state, etc., are some examples of host factors. The environmental factors can be classified as physical environment, biological environment, and socio-economic environment. Examples of the physical environment are air, water, light, sound, heat, ionizing radiation, non-ionizing radiation, etc. Domestic animals, wild animals, other reservoirs, vectors, etc. are examples of the biological environment, and habits, culture, religion, customs, behavior, lifestyle, family size and composition, overcrowding, occupation, income and social status, etc. are examples of the socio-economic environment.

Epidemiological triad is the basis for the multifactorial approach in the causation of disease. It would be useful to have a complete understanding of the complex interactions involved in the causation of disease for identifying the weak link in the process. Striking at the weak link for the control and prevention of disease would be relatively more cost-effective. Even when there is only incomplete knowledge about the agent, host, and environmental factors and their interactions, it is possible to control and prevent the disease; however, it may not be cost-effective.

In the case of poliomyelitis, the best choice of strategy for control/eradication was to improve the host immunity rather than focusing control measures directly on the agent or environmental factors. In the case of leprosy control, the chosen strategy was to mainly focus on the agent (bacteria) by early diagnosis and chemotherapy with a multidrug regimen rather than trying to improve the host immunity or the environmental factors. In the case of guinea worm, the strategy adopted was to focus on environmental factors (removal of step wells) rather than improving host immunity or using chemotherapy against the agent factor (guinea worm).

DISTRIBUTION OF DISEASE

The distribution of disease can be described in terms of personal characteristics of the affected person, place of occurrence of disease (spot map) and time of occurrence of disease (epidemic curve). Describing the occurrence of disease in terms of personal characteristics of the affected persons, place of occurrence of disease, and time of occurrence of disease would also help to a very great extent in understanding the causal mechanisms of the disease.

DISEASE FREQUENCY

Disease frequency is basically measured by incidence and prevalence. Incidence is a measure used to identify the occurrence of new cases of a disease within a specific population during a particular period of time. It helps in understanding the risk of developing a disease and is therefore important in studies related to disease causation and prevention. In contrast, prevalence refers to the total number of existing cases, including both old and newly diagnosed cases, present in a population at a given time. Prevalence mainly indicates the overall burden of illness in the community and provides useful information for health planning, resource allocation, and the organization of healthcare services.

USES AND APPLICATIONS OF EPIDEMIOLOGY

Epidemiology has many uses and applications, and the number of applications keeps increasing. Finding out causes of diseases, predicting illness in individuals and to forecast disease occurrence in populations, investigating epidemics, identifying disease syndromes, finding out the natural history of diseases, predicting prognosis of diseases, evaluation of symptoms and signs, clinical decision making, evaluating health status of populations (community diagnosis) and planning

health care delivery system and their evaluation are the most common uses and applications of epidemiology.

As described earlier, epidemiology is the study of people. Epidemiology is the basis for all studies/research involving human participants.

REFERENCES

1. Weiss NS, Koepsell TD. *Epidemiologic Methods: Studying the Occurrence of Illness*. 2nd ed. New York: Oxford University Press; 2014.
2. Bonita R, Beaglehole R, Kjellström T. *Basic Epidemiology*. 2nd ed. Geneva: World Health Organization; 2006. p. 1.
3. Wallace RB, editor. *Maxcy-Rosenau-Last Public Health and Preventive Medicine*. 14th ed. Stamford (CT): Appleton & Lange; 1998. p. 5.
4. Hippocrates. On Airs, Waters, and Places. *Med Classics*. 1938;3:19–85.
5. Hennekens CH, Buring JE. *Epidemiology in Medicine*. Boston: Little, Brown and Company; 1987. p. 5–6.
6. Doll R, Hill AB. Smoking and carcinoma of the lung. *Br Med J*. 1950;2:739–48.
7. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the Epidemiology of Cardiovascular Disease: A Historical Perspective. *Lancet*. 2014;383(9921):999–1008.
8. Meldrum M. “A calculated risk”: the Salk polio vaccine field trials of 1954. *BMJ*. 1998;317:1233–1236.
9. Paul JR. Clinical epidemiology. *J Clin Invest*. 1938;17:539–41.
10. Park K. *Park’s Textbook of Preventive and Social Medicine*. 25th ed. Jabalpur: Banarsidas Bhanot; 2019. p. 60.
11. Last JM. *A Dictionary of Epidemiology*. 4th ed. New York: Oxford University Press; 2001.