INTRODUCTION

The aim of this article is to give a brief introduction to epidemiology, provide an overview of the main study designs, outline the key ethical issues and its applications to modern day medicine.

WHAT IS EPIDEMIOLOGY?

Epidemiology is defined as “the study of the distribution and determinants of health-related states or events (including disease) in specified populations and the application of this study to control health problems.”[1,2] In Greek, “epi” means upon or among, “demos” means people and “logos” means study. The “distribution” of disease can refer to the geographical location, age, sex, socioeconomic status or occupation of the population studied. The “determinants” or risk-factors are the variables that lead to disease.[3] Epidemiology has several uses. It is used to describe disease patterns in human populations, to identify the causes of diseases and finally, to provide data or information necessary for the development, management and evaluation of health services for the prevention, control and management of disease.[1] The discipline of epidemiology combines medicine, health sciences, social sciences and statistics. In contrast to clinicians who diagnose and manage individual patients, epidemiologists measure disease outcomes in populations and give advice to populations in order to reduce the overall burden of disease.[3] This “population” refers to a group of people who are susceptible to the condition being studied.

TYPES OF EPIDEMIOLOGICAL STUDY DESIGNS

Epidemiological studies involve data collection, analysis and interpretation from a defined population over a specific time period.[3] The two elements of an epidemiological study are exposure and outcome.[4] Risk factors (such as tobacco use, high blood pressure), diagnostic tests or medications are the usual exposures and outcomes are usually measured by the occurrence of disease or death. Studies can be broadly divided into two categories – observational and experimental studies [Figure 1]. In observational studies, researchers observe or describe the health status of the study population. They are called observational studies as they observe the natural path of the disease without any intervention. In experimental (intervention) studies, researchers intervene on the risk factors (by health education, procedures or medications) to alter the course of the disease and observe the changes.[5]
There are two types of observational studies: Ecological and analytical. The main difference in these two approaches is that they differ primarily in the evidence they can provide about a possible causal association. Ecological studies are used to measure incidence and prevalence of disease in a population and also to monitor population health. One of the main uses of ecological studies is to monitor population health and make population level comparisons and observe the relationship between population levels of exposure to risk-factors and the outcome of interest. One of the drawbacks of this study design is termed “ecological fallacy,” which occurs when relationships that exist for population groups are assumed to be true for individual participants of that group. One of the drawbacks of this study design is termed “ecological fallacy,” which occurs when relationships that exist for population groups are assumed to be true for individual participants of that group. For example, a study of people with spectacles showed that they had above average intelligence. To use that information and conclude that someone wearing spectacles has above average intelligence is an ecological fallacy [Table 1].

Cross-sectional studies are conducted to estimate the prevalence of a disease/outcome for a study population at a particular time. In simple terms, it is a snap-shot of the current outcome of interest in the target population at a specific point of time. Since cross-sectional studies are carried out at one-time point, they do not give any indication about the sequence of events. Hence, these studies cannot be used to infer causality. The main use is to collect descriptive information about the target population and determine the prevalence of the outcome of interest. Cross-sectional studies evaluate the prevalent rather than incident outcomes and thus eliminate people who develop the outcome, but die before the study; hence there is a bias in the study sample favoring healthier participants.

In a case–control study (retrospective study), the participants are identified as those with the outcome/condition at the beginning of the study. Controls are defined as those who do not have the outcome, but are sourced from the same population. Information about exposure to risk factors is collected retrospectively. One of the biggest advantages of case–control studies is that it can be used to investigate rare diseases or diseases with a prolonged latency period without having to follow several thousand people over several years because participants are chosen at the beginning of the study and data are collected retrospectively. Hence, these studies are quicker, relatively low-cost and easier to conduct. One of the main challenges of this study design is the selection of the control as the distribution of exposure should be the same among cases and controls. Other disadvantages are recall bias and confounding (a variable which is associated with the exposure and is the cause of the outcome). Since the outcome of interest is known at the beginning of the study, participants with the outcome will have a better understanding and recollection of their exposure compared with the controls. One way to decrease bias is to “match” controls to the cases by certain variables, such as age and sex. An alternative method of overcoming confounding is to collect appropriate data on the confounding variables during the study and adjust for this during analysis.

A landmark case–control study published in 1950 was by Doll and Hill. Doll and Hill recruited 709 lung cancer patients in 20 hospitals in London and asked them a set of smoking related questions. For each patient enrolled, an age and sex matched control was selected and asked the same set of questions. Although there were potential sources of bias (recall and selection), the authors concluded, that the results were unlikely to have been affected by bias and that there was an association between smoking and lung cancer. In order to address some of the concerns about the case–control design of their study, Doll and Hill began a large cohort study called the British Doctor’s Study.

A cohort study follows a group of participants with different levels of exposure to risk factors and evaluates their health over time. It is a longitudinal study which involves repeated measurements of variables overtime (several cross-sectional surveys) in the same population. The main advantage of this design is that
**EXPERIMENTAL STUDIES**

The most common experimental study design is the randomized control trial (RCT). In an RCT, participants are randomly assigned to one of two groups – one that receives the intervention that is being tested (intervention group), and the other that receives the conventional treatment (control group). Randomization of the participants into intervention or control is essential to eliminate systematic bias.

**ETHICS IN EPIDEMIOLOGICAL STUDIES**

In current times, epidemiologists face a number of ethical issues that need careful consideration while carrying out research studies. Whichever epidemiological study one chooses, researchers have ethical and professional obligations to reduce risks and to avoid causing harm to the study participants. They also have the duty to maintain confidentiality and privacy of the study participants, ensure equitable distribution of the benefits of the study to the participants and provide all

---

**Table 1: Advantages and disadvantages of different type of epidemiological studies**

| Study type          | Advantage                                                                 | Disadvantage                                                                 |
|---------------------|---------------------------------------------------------------------------|----------------------------------------------------------------------------|
| Cross-sectional     | Estimate prevalence of outcome of interest                                | Cannot make causal inference                                               |
|                     | Several outcomes and risk factors can be measured                         | Cannot capture outcomes with seasonal variations as this is done in one time point |
|                     | Low-cost and time saving                                                  | Prevalence-incidence bias (neyman bias)*                                   |
|                     | No loss to follow-up                                                      |                                                                            |
|                     | Useful for planning intervention and hypothesis generation                |                                                                            |
| Case-control        | Efficient design for rare conditions                                      | Cannot provide information on incidence                                    |
|                     | Can study multiple exposures                                              | Subject to bias-ecall bias, neyman bias                                    |
|                     | Does not require a huge sample size                                       | Selection of control is challenging                                         |
|                     | Low cost and time saving                                                  | Takes a long-time, very expensive                                          |
| Cohort study        | Demonstrate a temporal sequence of exposure and outcome                   | Loss to follow-up                                                           |
|                     | Allows to calculate incidence rate of disease                             |                                                                            |
|                     | Allows evaluation of multiple outcomes                                    |                                                                            |
|                     | Can be used to study rare or uncommon exposures                           |                                                                            |
| Ecological study    | Easy to conduct                                                           | Subject to ecological fallacy**                                            |
|                     | Examines population level data                                            | Difficult to analyze complicated exposure-outcome relationship            |
| Randomized control  | Effective to determine a cause-effect relation                            | Costly                                                                      |
| study               | Randomization eliminates selection bias and confounding factors           | Loss to follow-up                                                          |
|                     |                                                                           | Lack of generalizability*                                                  |

* Neyman bias: Is a form of selection bias in cross-sectional and case-control studies attributed to selective survival among the prevalent cases (i.e., fatal cases being excluded from the study), hence the sample of cases offers a distorted frequency of the exposure, ** Ecological fallacy: Is a fallacy in the interpretation of data where conclusions about the nature of individuals are derived from inference for the study population to which those individuals belong, *When the study participants who agree to participate in the trial are very different to the general population on whom the results will be applied

since the exposure occurs prior to the outcome, this study design is less subject to bias because exposure is measured before the disease status is known. Since new or incident cases are identified over the course of the study, disease progression or the natural history of the disease can be studied carefully. The main drawbacks of this study design are that it is very expensive and time-consuming and is challenging to follow-up individuals over time. One of the most famous cohort studies is the Framingham study which followed-up 5209 adults (30–62 years) in Framingham, Massachusetts in 1948. The study, which was designed to report on the epidemiology of cardiovascular diseases, has been used study several others conditions and has guided many health policies in the USA.
the necessary information to the study participants so that they can make an informed decision before agreeing to participate in the study. Researchers have a duty to ensure that the potential benefits or harms of the studies are communicated to society in a timely fashion.[18]

The origins of ethics in human research go back to the 19th century when Thomas Percival wrote the first set of guidelines for physicians using new treatments in the Manchester Infirmary in England.[19] In 1835, a Code of Medical Ethics Convention was produced and adopted by the Ohio Medical Convention in the United States of America.[20] At the end of the second world war, a set of guidelines called the Nuremburg Code was written to guide research involving humans. These guidelines were further developed by the World Medical Association (WMA) and a set of ethical principles, called the declaration of Helsinki was produced.[19] These guidelines provide direction to clinicians, researchers and others doing research involving human participants. These guidelines have since become the cornerstone of human research ethics worldwide. Since 1964, these guidelines have been amended 10 times, the most recent update was in October 2013 at the 64th WMA general assembly in Brazil.[21] The declaration includes principles on safeguarding research subjects, taking informed consent, minimizing risk to participants, adhering to an approved research protocol, seeking approval of a research ethics committee, maintaining privacy and confidentiality of participants, registration of research in a publically accessible database and dissemination of results.

THE ROLE OF EPIDEMIOLOGY IN CLINICAL PRACTICE

Clinical epidemiology extends the principles of epidemiology to medical practice. The term clinical epidemiology, coined by Paul in 1938,[22] is a discipline that uses the principles of epidemiology into the medical environment, focusing on patients rather than populations.[23] Clinical epidemiology helps the physician in making a diagnosis, discussing prognosis and choosing treatment pathways. It includes the location, evaluation and application of the best evidence available for the management of a patient. Observational studies that help to describe the population and the determinants of disease are useful in the process of diagnosis and estimating prognosis. The choice of the appropriate therapy is based on the proven effectiveness of interventions using RCTs.[23] Thus, population-based data and epidemiological studies form the basis of clinical practice, including diagnosis, prognosis and selection of therapy.

CONCLUSION

Epidemiology is a diverse field which originates from laboratories researching cell lines to global health using current technologies and responding to the current burden of disease.[24] Since the second world war, there has been rapid progress in the development of epidemiology.[25] The disease focus for epidemiological studies has been evolving overtime, as infectious diseases have been contained; interest has shifted to chronic conditions like cardiovascular diseases and cancer, which continue to be key areas of research. Epidemiological methods are progressing from observational studies to RCTs and now with the emergence of translational epidemiology, which involves translating scientific discoveries and evidence into population health impact.[26] the discipline of epidemiology has come a long way.

REFERENCES

1. World Health Organisation. Epidemiology. Health Topics; 2014. Available from: http://www.who.int/topics/epidemiology/en/.[Last accessed on 2015 Jan 01].
2. Last JM, editor. A Dictionary of Epidemiology. 4th ed. New York: Oxford University Press; 2001.
3. Woodward M. Epidemiology: Study Design and Data Analysis. Washington: Chapman and Hall/CRC; 2004.
4. Jepsen P, Johnsen SP, Gillman MW, Sørensen HT. Interpretation of observational studies. Heart 2004;90:956-60.
5. Pearce N. Classification of epidemiological study designs. Int J Epidemiol 2012;41:393-7.
6. Levin KA. Study design VI-Ecological studies. Evid Based Dent 2006;7:108.
7. Schwartz S. The fallacy of the ecological fallacy: The potential misuse of a concept and the consequences. Am J Public Health 1994;84:819-24.
8. Levin KA. Study design III: Cross-sectional studies. Evid Based Dent 2006;7:24-5.
9. Carlson MD, Morrison RS. Study design, precision, and validity in observational studies. J Palliat Med 2009;12:77-82.
10. Song JW, Chung KC. Observational studies: Cohort and case-control studies. Plast Reconstr Surg 2010;126:2234-42.
11. Doll R, Hill AB. Smoking and carcinoma of the lung: preliminary report. Br Med J 1950;2:739-48.
12. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years’ observations on male British doctors. BMJ 2004;328:1519.
13. Dawber TR, Kannel WB. An epidemiology study of heart disease: The Framingham study Nutr Rev 1958;16:1-4.
14. Sibbald B, Roland M. Understanding controlled trials. Why are randomised controlled trials important? BMJ 1998;316:201.
15. Stolberg HO, Norman G, Trop I. Randomized controlled trials. AJR Am J Roentgenol 2004;183:1539-44.
16. Kendall JM. Designing a research project: Randomised controlled trials and their principles. Emerg Med J 2003;20:164-8.
17. Tuberculosis Trials Committee. Streptomycin Treatment of Pulmonary Tuberculosis: A Medical Research Council Investigation. British Medical Journal 1948;2:769-82.
18. Coughlin SS. Ethical issues in epidemiologic research and public health practice. Emerg Themes Epidemiol 2006;3:16.
19. Waters WE. Ethics and epidemiological research. Int J Epidemiol 1985;14:48-51.
20. Ohio Medical Convention. Code of medical ethics. Boston Med Surg J 1846;35:39-42.
21. World Medical Association. WMA Declaration of Helsinki-Ethical Principles for Medical Research Involving Human Subjects; 2013. Available from: http://www.wma.net/en/30publications/10policies/b3/. [Last accessed on 2015 Jan 01]
22. Anderson GM, Bronskill SE, Mustard CA, Culyer A, Alter DA, Manuel DG. Both clinical epidemiology and population health perspectives can define the role of health care in reducing health disparities. J Clin Epidemiol 2005;58:757-62.
23. Díaz-Vélez C, Soto-Cáceres V, Peña-Sánchez RE, Segura MA, Galán-Rodas E. Clinical Epidemiology and Its Relevance for Public Health in Developing Countries. In: Rodriguez-Morales A, editor. Current Topics in Public Health, 2013. ISBN: 978-953-51-1121-4, InTech, DOI: 10.5772/54901. Available from: http://www.intechopen.com/books/current-topics-in-public-health/clinical-epidemiology-and-its-relevance-for-public-health-in-developing-countries. [Last accessed on 2014 Nov 25].
24. Bhopal R, Macfarlane GJ, Smith WC, West R, Management Executive Committee for the XIX World Congress of Epidemiology. What is the future of epidemiology? Lancet 2011;378:464-5.
25. Hennekens CH, Buring JE. Epidemiology in Medicine. Philadelphia, PA: Lippincott, Williams & Wilkins; 1987. p. 515.
26. Khoury MJ, Gwinn M, Ioannidis JP. The emergence of translational epidemiology: From scientific discovery to population health impact. Am J Epidemiol 2010;172:517-24.
27. Delgado-Rodríguez M, Llorca J. Bias. J Epidemiol Community Health 2004;58:635-41.

Cite this article as: Joshi R, Joshi R. Deconstructing epidemiology. CHRISMED J Health Res 2015;2:3-7.
Source of Support: Nil. Conflict of Interest: No.