Epidemiology

An Introduction for Medical Students

流行病学 医学生入门指南

Edited by
Wang Peishan
主编 汪培山

天津科技翻译出版公司
Tianjin Science & Technology Translation & Publishing Corp.

Epidemiology

An Introduction for Medical Students

Contributed by

Wang, Peishan Professor of epidemiology, Tianjin Medical University, Tianjin, China

West, Roy Professor of epidemiology, Memorial University of Newfoundland, San Johns, Canada

Edlavitch, Stanly A. Professor of Epidemiology, Missouri University Medical School, Kansas City, USA

Wang, Peter P. Professor of epidemiology, Memorial University of Newfoundland, San Johns, Canada

Qi, Xiuying Professor of epidemiology, Tianjin Medical University, Tianjin, China

天津科技翻译出版公司
Tianjin Science & Technology Translation & Publishing Corp.

图书在版编目(CIP)数据

流行病学:英文/汪培山主编. —天津:天津科技翻译出版公司,2011.7
ISBN 978-7-5433-2874-7

Ⅰ. ①流… Ⅱ. ①汪… Ⅲ. ①流行病学 - 英文
Ⅳ. ①R18

中国版本图书馆 CIP 数据核字(2011)第 061816 号

出　　版: 天津科技翻译出版公司
出 版 人: 刘 庆
地　　址: 天津市南开区白堤路 244 号
邮政编码: 300192
电　　话: 022-87894896
传　　真: 022-87895650
网　　址: www. tsttpc. com
印　　刷: 高等教育出版社印刷厂
发　　行: 全国新华书店
版本记录: 889 × 1194　16 开本　11.375 印张　260 千字
2011 年 7 月第 1 版　2011 年 7 月第 1 次印刷
定价:38.00 元

(如发现印装问题,可与出版社调换)

Preface

The main goal of public health is to prevent and control disease in human populations. To reach this goal, we attempt to understand the nature of diseases, explore the causes of diseases, take action (public health programs) and evaluate the effectiveness of the actions. Epidemiology plays a crucial role in realizing this goal and deciding which public health programs to develop and in evaluating those programs.

In recent decades, it appears that epidemiology has increasingly become a very important approach, not only for public health but also for clinical medicine and other areas of medical research and practice. Advances in computer design and availability and the development of advanced statistical methods and readily accessible powerful software have made it feasible to manage and statistically analyze very large data bases. Epidemiological techniques have been applied to many areas of biomedical research and as a result, quite a few branches of epidemiology have emerged. Today, you might find epidemiologists who specialize in chronic disease or infectious disease epidemiology or genetic epidemiology or molecular epidemiology or pharmacoepidemiology or nutritional epidemiology, to name only a few of the specialties.

The aim of this book is to provide an overview of principles of epidemiology for students of clinical medicine. It is also usable for the students of other health professions such as nursing, dentistry, pharmacy, public health service etc. It also can be used as the reference book for other health professionals. For medical students, this book assumes that they have successfully finished a short course of medical statistics.

In this textbook, most of the chapters are concentrated on the basic principles, concepts and methodology of epidemiology (first 10 chapters). The remaining chapters consist of some applications of epidemiology. These may be selected by instructor for lecture or for used by students as self-readings.

It is important to note that this text book just serves as an introduction of epidemiology. Because of the obvious constraints of any single textbook, it does not comprise an exhaustive answer to many questions in epidemiology. I would appreciate any suggestions on what and how the contents of this book can be improved.

Wang Peishan
Tianjin Medical University
Tianjin, China
psw@ tijmu.edu.cn
November 2010

Acknowledgements

I wish to express my indebtedness to Professor Roy West and Professor Stanley Edlavitch, for their generous contributions on this book. I could not have accomplished this work without their dedication. Both are senior professors of epidemiology and experienced on teaching.

The support of the institution, International College, Tianjin Medical University, was essential for this book. I especially want to acknowledge Professor Hao Xishan, the president of the university and Professor Guo Fenglin, the dean of the college for their encouragement and substantive support.

I am grateful to my good fiends, Professor Peter Wang and Professor Qi Xiuying for their wonderful contributions to this book.

I also thank the editors and publisher for helpful advice and support.

I would like to acknowledge many great professors and their books or articles that taught me about epidemiology. I especially recommend the book *Epidemiology* edited by professor Leon Gordis (1st ed. 1996 and 3rd ed. 2004). I have recommended it as the main reference book for my students.

It is an honor to dedicate this book as a memorial to Professor Geng Guanyi, who brought me to the area of epidemiology. Professor Geng, who passed away in 1998, was the distinguished epidemiologist in China and the founder of School of Public Health, Tianjin Medical University.

I appreciate all of the help and comments from my students, including my Master students having worked as my teaching assistants and many international medical students in my classes.

I also very appreciate the contributions of Professor Pat West, Roy's wife, who is also a professor of epidemiology for her zealous work on typing and English editing on part of this book. I would like to thank all authors' spouses for their support.

Finally I am very grateful to my family, particularly my wife, for their encouragement and support of time and patience.

Wang Peishan
Tianjin Medical University
Tianjin, China
psw@ tijmu.edu.cn
November 2010

Contents

Chapter 1

Introduction

What is epidemiology?

Epidemiology is a fundamental medical science that focuses on the distribution and determinants of disease frequency in human populations. If there is disease occurring in the population, we always want to know the cause or the trend of the problem. Who is developing a disease, and where and when, and why are they developing it? We like to be able to compare the disease distributions in different places, different years and in different groups of people. We want to know what is the possible cause of the occurrence of the disease. Based on this knowledge, we may take some measures to control or prevent the disease. Some very common questions are, for example:

— How many new cases with flu occurred last year in this country and how many cases this year.
— If we immunized the population with vaccine, does it prevent the epidemic of flu or not.
— How many cases with diabetes now in this community?
— What is the trend of the breast cancer among women in the urban and rural areas.

Epidemiological methods may help us to answer these questions. Epidemiology is also important to the practice of medicine as increased knowledge of disease occurrence aids in diagnosis and treatment.

The term epidemiology consists of the Greek words επι (epi) = among, δημοζ(demos) = people, and λογοζ (logos) = doctrine. It was translated as "a statement of what is upon the people." Therefore epidemiology is a human science and it concerns the health problems in the human population. (WHO 2002)

Definition of epidemiology

There are many definitions of epidemiology in the literature. Here is the definition from the Dictionary of Epidemiology, edited by Dr. Last, which seems to well describe the work of epidemiology:

"Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations and the application of this study to control of health problems." (Last 2002)

Brief history of epidemiology

In the history of human beings, many epidemic or pandemic diseases were spread in the population such as plague, small pox, cholera, leprosy, syphilis, etc., which caused numerous deaths or serious social problems. It was estimated that there were over 60 million people died from plague in the world. (Oleckno, p14) Up to the early 20th century, more than 600 epidemics had been documented in China. (Geng 1996) The people always attempted to find the reasons why these diseases occurred and were trying to control the spread of the diseases. These efforts contributed to the development of epidemiology.

Early epidemiology

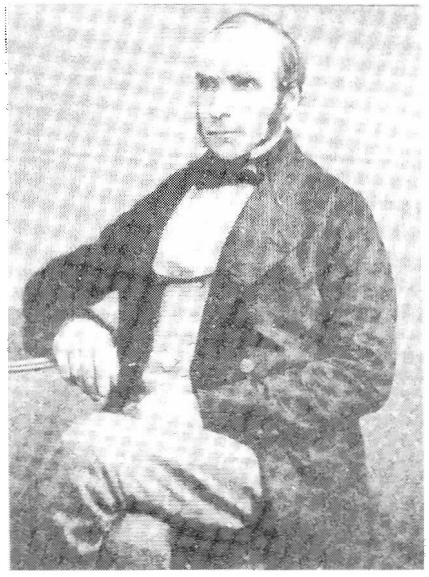
The development of epidemiology occurred over a long stretch of time and numerous doctors and persons contributed to its progress. Many epidemiological observations were made. Today these observations may seem simple or not critical in design, but they provided very useful information towards describing diseases and their control.. Even though microorganisms were not identified until the 19th century, the people found that the diseases might be infectious and spread from person to person or from animal to person. Gradually the people knew that the isolation of the ill people might control the transmission of the disease. In 14th century, in Europe the ships were asked to stay in an isolated area before arrival into the port for 40 days to protect the community against the Black Death (plague), this is the origin of the quarantine. People knew more and more about the possible route of transmission of disease and invented some control methods for disease, such as the vaccine to prevent smallpox (by Edward Jenner, 1749 - 1823). In fact, the process of the fight against disease led to the development of epidemiology.

Hippocrates, the physician who lived about two thousands years ago, was often mentioned as the pioneer of epidemiology. In his books he described the epidemics of disease and recognized associations between environmental and other factors and certain diseases.

For the collection of health data, such as mortality of the population, the UK medical registration of death is a

good example that was introduced in 1801, and in 1838 William Farr (1807 - 1883) introduced a national system of recording causes of death. The analyses of these data involving such techniques as life tables and the standardization of rates gave a picture of the health status of the population of Great Britain.

As we learn the history of epidemiology, **John Snow's work and the Broad Street pump story** is a very common example to show the contribution of epidemiology on the control of the disease. Dr John Snow (1813 - 1858), was a surgeon (actually an anesthesiologist) and pioneer of the science of epidemiology. The work of John Snow is generally quoted as an example of a brilliant analytical investigation which can lead to the identification of a pathogenic agent and its elimination from the environment.



John Snow

The first epidemic of cholera hitting Great Britain was in 1831 - 1832 and caused at least 60,000 deaths. Dr. Snow directly investigated the subsequent major epidemic episodes in London in 1849 and 1854, focusing attention on the role that polluted water might have played in the spread of the disease.

In the 1854 London epidemic, the worst-hit areas first occurred in the districts supplied by the Southwark and Vauxhall Water Company. At the same time, there was a markedly smaller number of deaths in those districts supplied by the Lambeth Company. Because in some areas the water supplies for the two companies happened to be closely intermixed, Dr Snow carried out a door to door inquiry to ascertain to which company supplied the water to their homes. The results clearly demonstrated the association between water supply and later the deaths from cholera. When they calculated death rates from cholera, for the districts supplied by Southwark and Vauxhall it was 4. 4 per thousand population, but for the districts supplied by the Lambeth company it was only 0. 2 per thousand. They was more than a 20 times difference.

There were also many other factors that led Snow to isolate the cause of the cholera to the Broad Street pump which belonged to the Southwark and Vauxhall Company. For instance, of the 530 inmates of the Poland Street workhouse, which was only round the corner, only five people had contracted cholera; but no one from the workhouse drank the pump water, for the building had its own well. Among the 70 workers in a Broad Street brewery, where the men were given an allowance of free beer every day and so never drank water at all, there were no fatalities at all. And an army officer living in St John's Wood had died after dining in Wardour Street, where he too had drunk a glass of water from the Broad Street pump.

Though the authorities were reluctant to believe his conclusion, they

The pump handle (Source: Steven Johnson, Ghost Map, Riverhead Books, New York, 2006)

agreed to remove the pump handle as an experiment. When they did so, the spread of cholera dramatically stopped.

It is very interesting that even though John Snow by 1854 concluded that cholera could be transmitted through contaminated water, the etiologic agent of cholera, Vibrio cholerae, was not identified until almost three decades later. Dr. Snow had proved the cause of the cholera epidemic in Broad Street, London and controlled it by stopping the water supply from the suspected pump was before the development of the theory of disease caused by microorganisms.

Later in the 19th century, some bacteria were recognized by some scientists such as Louis Pasteur in France and Robert Koch in Germany. They demonstrated that microorganisms could cause infectious disease, and developed the criteria to judge the relationship between the microorganism and a specific disease. Along with the progress of microbiology and techniques of statistics, epidemiology began its rapid evolution. There was the first association of epidemiology in London in 1850 and there was even a journal on epidemiology in Russia in 1870. (Geng G.)

Modern epidemiology

For traditional epidemiology, most of the work was concentrated on infectious diseases, such as plague, small pox, cholera, typhoid fever. It was reasonable because epidemics or pandemics of these infectious diseases historically caused millions of deaths. This led to authorities and health workers trying to find the cause of such diseases and to control the spread or eliminate the diseases from the world. Later there were also some individuals who contributed great work on non-infectious diseases such as the identifying of causes and control for pellagra, scurvy, beriberi, etc. In fact James Lind (1716 - 1794) had identified the dietary factors associated with scurvy in 18th century and found the way to prevent and treat the problem, and it eventually led to the development of experimental epidemiology.

Around World War II and later, there were some epidemiological studies on non-communicable chronic diseases such as cancer and cardiovascular disease. A very good story is about smoking and health issue. There were a couple of studies published in the United States and in the United Kingdom in the late 1930's. Doll and Hill in UK had carried out a series studies including case-control study and prospective cohort studies. The air pollution was very serious in London and people thought it could be the main cause of the increase mortality

from lung cancer. However the studies of Dr. Doll and his colleagues suggested a very significant association between cigarette smoking and lung cancer. the Framingham heart study in the US is another classical example, it was started in 1948 when heart disease had become an important cause of death of the population in the United States.

Today, epidemiological methods of investigation have become tools for answering the questions in medicine and public health regarding their biological and social facets. Especially due to the progress of the use of computers, analyses of large databases and complicated calculations have become feasible. Epidemiology has contributed to the understanding of diseases in the population, the studies of etiology, and control of some health problems including preventive or therapeutic actions for a lot of important diseases, e.g., cardiovascular diseases especially ischemic heart disease, asthma, and some cancers. Regarding the identifying of the possible causal risk factors for some emerging diseases such as AIDS, Legionnaires' disease, Lyme disease, and *Helicobacter* infections, epidemiology has played an significant role.

Epidemiological methods have been applied in various areas or jointly with other disciplines and there have been many branches of epidemiology emerge in recent decades such as cancer epidemiology, cardiovascular epidemiology, or sero-epidemiology, pharmacoepidemiology, clinical epidemiology, molecular epidemiology. All of these applications are based on epidemiological methods, the basic principles we will learn in this book.

An overview of methods of epidemiologic studies

To answer the questions about disease distributions, trends of the diseases, etc., we need to collect data from the population and then describe the data or analyze the relationship between the exposure and outcome or cause-and-effect association. **Exposure** is a very common used term in epidemiology, it refers to the causal factors that may associated with the disease. For example, contact with a biologic agent, a harmful materials, or some characteristics such as the age or blood groups which may put an individual at increased risk. Exposure may of course be beneficial rather than harmful, e.g., exposure to proper physical exercise. **Outcome** is the disease or other change in health status. It is the possible result associated with the causes, risk factors or preventive measures.In epidemiology, we collect data about the exposure and outcome and then describe them and/or analyze the association between them.

How to describe the disease in the population? Person, place, and time

Firstly we need to know the distribution of the disease. Primarily it refers to how the morbidity (incidence and prevalence of the disease) or mortality is distributed in a given population or a community. Generally we describe morbidity (illness) or mortality (deaths) in terms of **person, place and time.** We want to know what is the rate of illness in different age groups, whether they have the same death rates, whether there is a difference among various countries, and what are the trends of the pattern of death in the last decades. By this description we may

have the idea of disease in the population and furthermore it is possible to explore or find the potential causes of the disease distribution.

By person

Age, sex, race group, education, occupation, socioeconomic status, marital status, immunization status and some lifestyle or behavior habits are common variables used to describe the distribution of disease. For example, we might describe the distribution of coronary heart disease in one community by sex, age, occupation, and some lifestyle factors such as smoking, drinking and physical exercise. These descriptions may give us a very good idea of this disease. These data will be very useful for the estimate the burden of the disease and policy making. These descriptions may also provide us some thinking of who are in the risk group and what is the potential reason of the disease. More detailed examples of distribution of disease by person will be seen in the chapter of communicable diseases.

By place

Place variables include various geographic characteristics. They might be based on administrative areas such as countries, provinces, cities, towns, or neighborhoods, or might be based on specific geographic characteristics such as urban and rural area, mountainous area and plain district, or in different latitude or altitude. For example, Figure 1. 1 gives us the idea of the liver cancer death rates are varied among countries. And Figure 1.2 is the distribution of esophagus carcinoma in Taihang Mountain area, a central place in China where the death rates of this disease were high. The description of the distribution might be very helpful for the etiological study of this cancer.

By time

"By time" refers to the distribution of given disease by hours, days, weeks, months, years, or even by decades. From the description it will reveal the temporal patterns of the disease. We might know the trends of the disease or some change that is unexpected. Common temporal patterns of morbidity and mortality include short-term fluctuations, cyclic changes and secular trends.

Short-term fluctuations refers to relatively brief, unexpected increases in the frequency of a particular disease. Short-term fluctuations are commonly manifested in epidemics.

Cyclic trends represent periodic increase in frequency of a particular disease morbidity or mortality. It often is predicable.

Seasonal variation is a typical cyclic pattern. The morbidity or mortality of a particular disease might be influenced by weather or temperature. Many diseases have this phenomenon, including infectious disease and non-infectious chronic diseases. The examples for infectious diseases are malaria, influenza, and the examples for non-infectious diseases are some cardiovascular diseases such as stroke.

However, cyclic trends are not only seasonal variations. The curve of time distribution of some illness might have elevation every other year or with an interval of several years. For example measles, before the wide use of vaccine for the immunization, epidemics were common in the community at two or three years intervals. The assumed reason is of

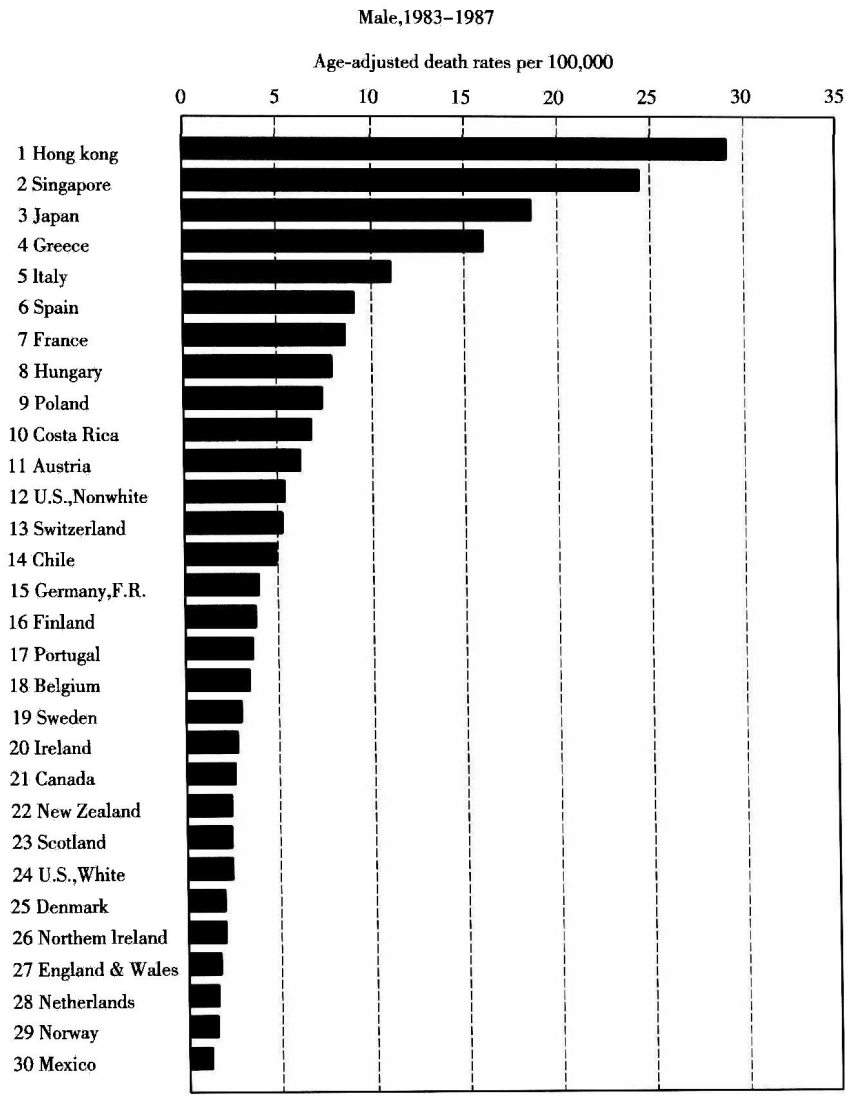


Figure 1. 1 **Age-adjusted death rates of malignant neoplasm of liver and intrahepatic bile ducts (Males) in selected countries and districts (1983 - 1987).** (Source: Tominaga S., Kuroishi T, Aoki K. ed: Cancer mortality statistics in 33 countries 1953 - 1992. UICC 1998 p50)

the accumulation of susceptible children. The same phenomenon may also be found in the distribution of hepatitis A in some places.

Secular trend is also called long-term changes. It represents the pattern of trends for a long time such as many years, decades, or even centuries. From Figure 1. 3 the data shows us that female breast cancer has been increasing in most countries even though the rates are varied. Figure 1. 4 is the trends of incidence and mortality of infectious diseases in China from 1980 to 2000. We can know both incidence and mortality have decreased dramatically.

Table 1. 1 summarizes the commonly used variables of person, place and time. Some more examples also are to be found in chapter 3, communicable diseases.

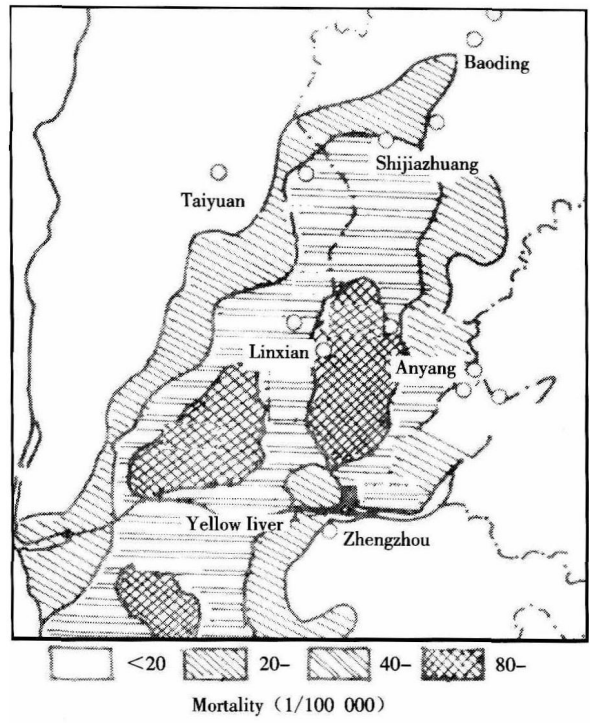


Figure 1.2 **Mortality of esophagus carcinoma in Mountain Taihang area, China, 1979.** (Source: Geng G, 1996)

Table 1.1 Commonly used variables of person, place and time

Person	Place	Time
Age	Administrative boundaries	Short-term fluctuations
Sex	International	Cyclic changes
Social class — income, education, occupation	Provincial	Seasonal variations
Marital status	Natural boundaries	Secular trends
Lifestyle factors — smoking, physical activity	Environmental factors	
Race and/or ethnic group	Residence	
Inherited factors	Birthplace	
Immunization status	Workplace	

Descriptive and analytic epidemiology

Descriptive epidemiology is concerned with the variations of morbidity and mortality in a community. It concentrates on the description of distribution of morbidity or mortality by person, place and time. Several types of studies are in this category, such as case report, case series, ecological study and prevalence survey (or cross-sectional study). The information obtained by these descriptive studies is not only useful in health care planning, but also can provide clues to the potential causes of disease.

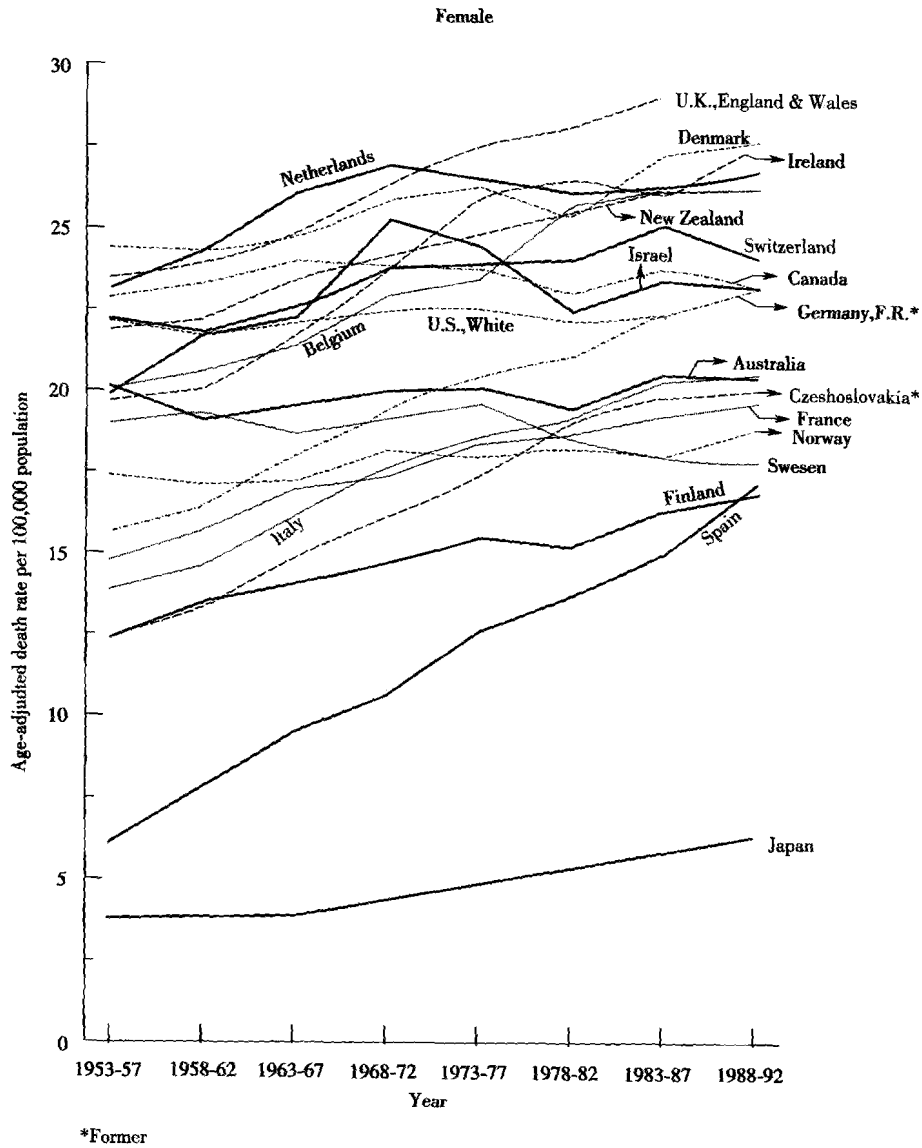


Figure 1.3 **Trends in age-adjusted death rates of malignant neoplasm of breast (Female) in selected countries of the world (from 1953-57 to 1988-92).** (Source: Tominaga S., Kuroishi T, Aoki K. ed: Cancer mortality statistics in 33 countries 1953-1992. UICC 1998 p75.)

Analytic epidemiology, it is more concerned with the relationship between the exposures and outcomes. Common designs are cohort study, case-control study and experimental study. Generally one or more hypotheses are generated base on the information from descriptive studies and then tested to help identify the potential causes of the morbidity or mortality.

Epidemiologic studies should answer the W5 questions: **What** disease? **Who** is diseased? **Where** did it occur? **When** did they become ill? **Why**? Descriptive epidemiology commonly answers the questions of who, where and when, and analytic epidemiology can answer the question why.

However, the distinction between descriptive and analytic studies may not

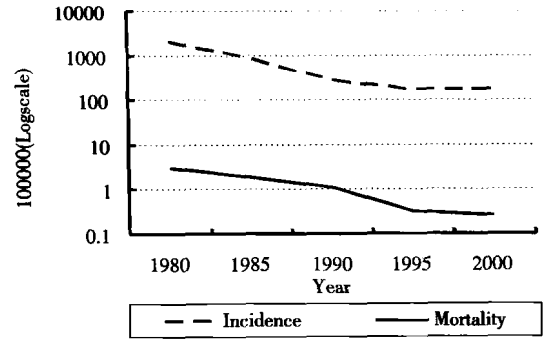


Figure 1.4 **Reported Incidence and Mortality of Infectious Diseases, China, 1985-2000.**

always be clear-cut. A descriptive study, may compare the distribution of the disease in subgroups. It may result in the relationship of two or more variables and provide clues for more analytic studies. In an analytic study, before measuring the association and making an inference between exposure and outcome, it may also need to describe the distribution of variables.

Observational and experimental studies

An **Experimental study** randomizes the subjects into intervention and control groups (exposed and unexposed groups) and follows them over time to compare their rates of disease development. In experimental studies the investigator will control the allocation of subjects to the different groups. Such as in a clinical trial the subjects will be divided into two groups of different treatment or preventive measures. It is scientifically more rigorous. It is good to test the efficacy of a medication or vaccine. It is also good to evaluate the results if some factors associated with the disease are able to be changed, generally we call this an intervention study. However sometimes it is very difficult to carry out the experimental study in epidemiologic research. They are not suitable for the study of rare disease outcomes since it may take a long time to perform, they often present complex ethical problems, or may simply not be feasible.

Most of the time an epidemiologist will just observe and measure the occurrence of the disease or other health related conditions in different groups with various characteristics. In this way, he/she will attempt to identify the causal association between the exposure and outcome. The investigators do not intervene in any way therefore it is called an **observational study**. All cross-sectional studies, case-control studies and cohort studies are in this category.

Prospective and retrospective study

Based on time to collect the information for the study or on the direction of the study, studies can be divided into prospective and retrospective studies. They are sometimes confused. Generally the experimental studies and cohort studies are designed to follow up the outcome, therefore they are called prospective studies. Case-control studies need the investigator to collect the exposure information after the disease has occurred. Therefore they are called retrospective studies. However in cohort studies sometimes we may analyze the association

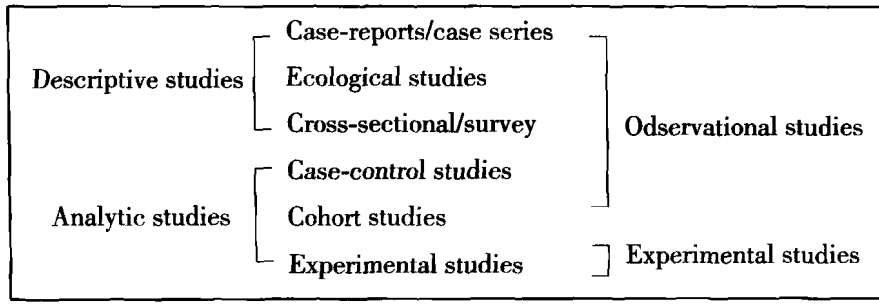


Figure 1.5 **Hierarchy of Epidemiological Study Design**

based on the historical exposure and outcome data, the direction is also from exposure to outcome. The historical cohort study and concurrent cohort study are also called retrospective cohort study and prospective study.

Summary of methods of epidemiology

Now we have a brief idea that epidemiologic methods include a series designs and most of them we will learn later. Figure 1.5 is a hierarchy of epidemiological studies.

Goals of epidemiology

What might we do with epidemiology? The main goal of public health is to identify the health problems within a community, or in the population (a group of the people), then to identify what is causing the problems, and to try to take some control measures and to test if these measures are working to resolve or reduce the problem. Epidemiology is fundamental in providing the data needed to make public health judgments in each of these areas. With descriptions and comparisons of the factors of the population, epidemiological studies may provide substantial information for the public health services and improve the health of the community.

Here are some general goals of epidemiology. In the rest of the book, we will discuss these in more detail with examples.

To describe the distribution of the diseases and the health related states in the population

We may observe and record the occurrence of disease or health status, including some factors related to the disease, and then describe the changes in the health status (by time), or differences between groups of people, or between different geographical areas (by persons and by places). These descriptive data will give us the background needed for action or for further investigation. These data will be also very important for health authorities and planners who need to know the nature and size of the health challenges faced by their communities.

To identify the causes or risk factors of a disease

The purpose of public health professionals is to try to control diseases in the population. Therefore it is very important that we know the reason that caused the disease or caused the disease epidemic. A very important role of epidemiology is to identify the causes or risk factors of a disease. Our final goal is to control the