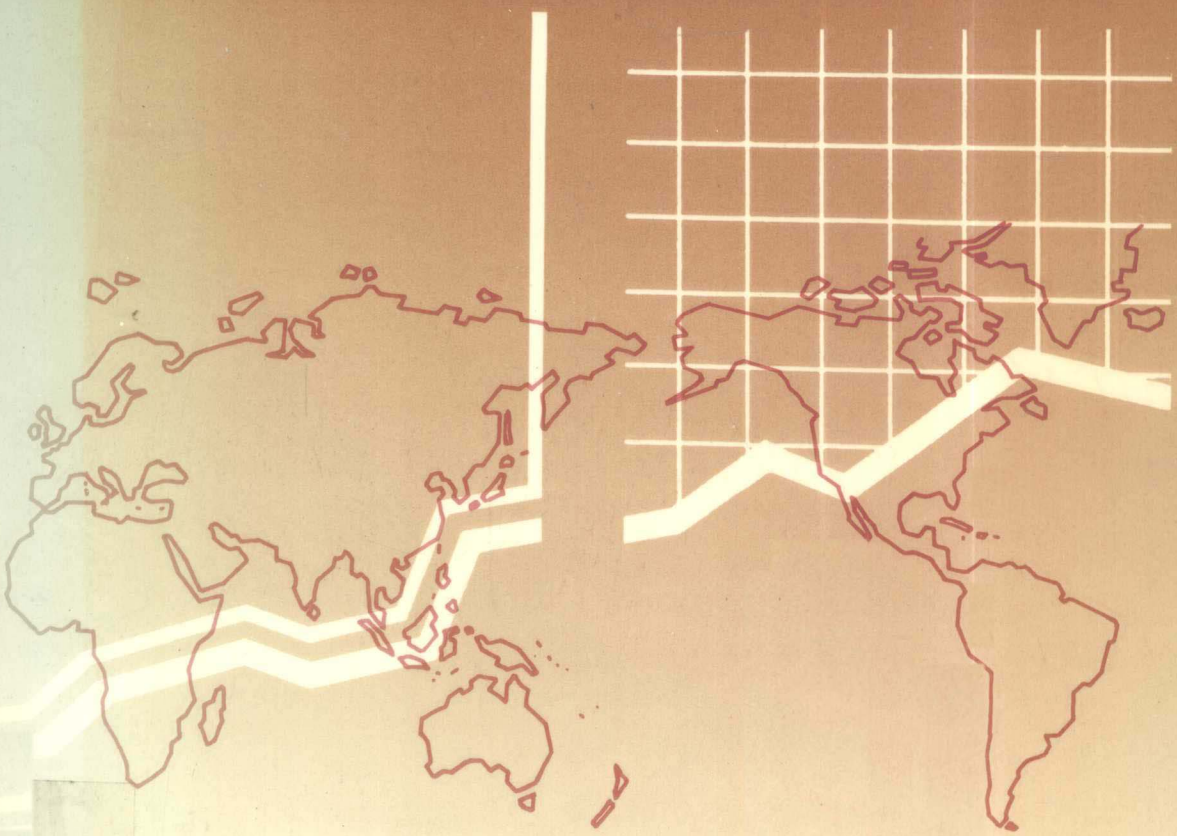
Teaching Health Statistics

Twenty lesson and seminar outlines



Edited by
S. K. Lwanga and Cho-Yook Tye

World Health Organization
Geneva

Teaching Health Statistics

Twenty lesson and seminar outlines

Edited by

S.K. Lwanga and **Cho-Yook Tye**

Epidemiological and Statistical Methodology, World Health Organization, Geneva, Switzerland

Department of Social Medicine and Public Health, National University of Singapore, Singapore

World Health Organization
Geneva
1986

ISBN 92 4 156090 8

TYPESET IN INDIA
PRINTED IN BELGIUM

85/6446—Macmillan/Ceuterick—6000

Preface

The need for a statistical approach is now well recognized in epidemiology and public health since these fields are concerned with communities or populations where the laws of large numbers and random fluctuations clearly apply. Medical teachers and students, however, have been slower to recognize the need for a knowledge of statistics, even though all aspects of medical diagnosis and prognosis are affected by rules of probability.

The extent of the statistical knowledge and skills that medical students need to acquire varies from country to country according to such factors as the common health problems and methods of delivering health care in the country, and the career prospects of the students on graduation. Nevertheless, there is a core of statistical knowledge that all students need to have, irrespective of their country of training.

The present set of outlines has been prepared in response to a recommendation of a WHO Interregional Conference on Teaching Statistics to Medical Students, held in Karachi in 1978, and the topics covered form an internationally acceptable standard basic curriculum for teaching health statistics to medical students. The lessons and seminars highlight fundamental concepts of probability and ways of thinking that are useful to medical students, and are meant for selective use by teachers of statistics in preparing their courses.

The preparation of this publication was conceived some years ago by Dr B. Skrinjar while she was Chief Medical Officer in WHO in charge of the development of health statistical services. The work is a result of the close collaboration of a number of eminent teachers of medical statistics, and has been coordinated and edited by Mr S. K. Lwanga, Statistician, Epidemiological and Statistical Methodology, World Health Organization, with valuable assistance from Professor C.-Y. Tye of the National University of Singapore.

I wish to express my deep gratitude to the following people, who contributed the individual lesson and seminar outlines: Professor O. Ayeni, Medical Statistics Unit, Department of Preventive and Social Medicine, University of Ibadan, Ibadan, Nigeria; Professor C. R. Lowe, Emeritus Professor of Community Medicine, Welsh National School of Medicine, Cardiff, Wales, United Kingdom; Mr S. K. Lwanga, Epidemiological and Statistical Methodology, World Health Organization, Geneva, Switzerland; Dr B. Skrinjar, Formerly Chief, Development of Health Statistical Services, World Health Organization, Geneva, Switzerland; Professor R. N. Srivastava, Department of Social and Preventive Medicine, M.L.B. Medical College, Jhansi, Uttar Pradesh, India; Professor Cho-Yook Tye, Department of Social Medicine and Public Health, National University of

Singapore, Singapore; Mr G. Wakely, Auckland Hospital Board, Auckland, New Zealand.

I also wish to thank the many staff members of the World Health Organization who, together with the following people, reviewed the early drafts of the outlines: Professor T. Colton, School of Medicine, Boston University School of Public Health, Boston, MA, USA; Dr T. Fernandez, Department of Preventive and Social Medicine, Cebu Institute of Medicine, Cebu City, Philippines; Mr T. M. Frazier, Division of Surveillance, Health Hazard Evaluations and Field Studies, National Institute for Occupational Safety and Health, Cincinnati, OH, USA; Dr D. L. Heymann, International Health Program Office, Centers for Disease Control, Atlanta, GA, USA; Dr A. Indrayan, University College of Medical Sciences, New Delhi, India; Ms R. Loewenson, Department of Community Medicine, University of Zimbabwe, Harare, Zimbabwe; Dr G. Luković, Andrija Štampar School of Public Health, University of Zagreb, Zagreb, Yugoslavia; Dr J. Otete, Ministry of Health, Nairobi, Kenya; Professor B. Pirc, Ljubljana, Yugoslavia; Dr T. Watts, Department of Community Health, University of Zambia, Lusaka, Zambia.

Special thanks are also due to all those who helped in the preparation of the manuscript.

K. Uemura
Director, Division of Epidemiological Surveillance and Health Situation and Trend Assessment, World Health Organization

Introduction

Knowledge of, and competence in, the application of statistical principles and methods are necessary, not only for an understanding of the biological and medical sciences, but also for effective practice in any of the health professions. Because of the variability of biological, clinical, and laboratory data, a knowledge of statistics is necessary and central to their understanding and interpretation.

Some of the more important reasons why every student training in the health fields should be exposed to a course in health statistics are given below.

1. A knowledge of statistics is required in order both to understand the rationale on which diagnostic, prognostic and therapeutic decisions are, or should be, based, and to appreciate that medicine is highly dependent on concepts of probability.
2. Health workers need to interpret, within their competence, laboratory tests and bedside observations and measurements in the light of a knowledge of physiological, observer, and instrument variation.
3. Health workers must know and understand the statistical and epidemiological facts about the etiology and prognosis of the diseases they treat, in order to give the best advice to their patients about how to avoid, or limit the effects of, these diseases.
4. Health workers are the primary generators of the data on which health statistics is based. They therefore need to know how data can and should be used, both for the benefit of their own practices, and for the organization and delivery of health care in their countries.
5. Health managers need to know how to interpret and draw inferences from the statistics that describe their country's health problems, and how best to use the resources available to meet them.
6. The study of statistics helps to foster in students the critical and deductive faculties they will need throughout their studies and, after graduation, in their practices.

The course

In recognition of the diversity of teachers and teaching methods, the lesson and seminar outlines given here are deliberately presented in a variety of ways. An attempt has been made, however, to present all of them with clearly stated aims and enabling objectives. (The few enabling

objectives requiring computational skills are marked with a dagger, "†", to remind the teacher not to overemphasize these skills at the expense of the knowledge required to understand and apply the statistical principles.)

No specific textbooks have been recommended for general use in the course, so that teachers who do not have access to particular books will not feel handicapped. Where examples have been extracted from published documents or books, full references are given to allow teachers to refer to the original if they so wish.

The outlines are intended to be a guide for teachers in preparing lessons and seminars, and in deciding on course content. They are *not* intended to be substitutes for fully prepared lessons and seminars. Moreover, they are written neither as self-instructional materials for students, *nor* as a textbook in statistics for teachers lacking in formal statistical training.

The outlines are divided into three parts:

(*a*) Part I (Outlines 1–10) covers statistical principles and methods;

(*b*) Part II (Outlines 11–16) covers health statistics, including demography and vital statistics;

(*c*) Part III (Outlines 17–20) deals with the use of statistics in medicine, including medical records.

For most of the lessons, and some of the seminars, suggested handouts to students are appended. The teacher should judge whether the suggested handouts contain examples that are relevant for the students. For example, the data used in the handout may not be applicable to the country where the teacher is based, and so relevant data for that particular country should be used instead. Additional exercises are presented in the final section, entitled 'Lesson Exercises', which can be found on pp. 217–226. These exercises are graded according to the level (part I, part II or part III) of the course that the student has completed. The teacher may, of course, select any relevant exercises for purposes of revision or illustration of a particular point in a lesson.

There is no fixed number of sessions for each lesson or seminar. Teachers should feel free to design lessons and seminars themselves, on the basis of the outlines, and covering any number of sessions depending on availability of time. Time should also be provided for class exercises.

Teaching of statistics should not be carried out in isolation from the other disciplines in the health curriculum, but whenever possible should be integrated with their teaching. The role of a statistics course in a health curriculum, as part of the training in information support for the health field, should not be forgotten. Statistics should not be taught as an end in itself, but as a means through which other disciplines may be better understood and implemented.

CONTENTS

* The outlines marked with an asterisk are for seminars; the remaining outlines are for lessons.

PART I
STATISTICAL PRINCIPLES AND METHODS

Outline 1 – Lesson

Introduction to the role of statistics in human biology and health care delivery

INTRODUCTION TO THE LESSON

In public health and clinical medicine, many statistical concepts are employed, consciously or subconsciously, in making decisions related to such matters as: clinical diagnoses; predicting likely outcomes of an intervention programme in communities, or the course of a disease in individual patients; selecting appropriate intervention programmes for particular communities or treatments for patients, etc. In laboratory medicine, statistics are constantly used in everyday practice. Knowledge of statistics has also become essential for an understanding and critical appreciation of communications in medical journals. A thorough grasp of statistical principles is thus an absolute necessity for the planning, conduct and analysis of studies to assess health situations and trends, as well as for the conduct of biomedical, clinical and public health research.

Objective of the lesson

The objectives of this lesson are to introduce to the students the role of statistics in the study of human populations, human biology, and medicine, and to create in them an awareness of the need to acquire an understanding of statistical principles and methods.

Enabling objectives

At the end of the lesson the student should be able to:

(*a*) Discuss the role of statistics in health care delivery, and explain the main uses of statistical methods in the broad field of health care.

(*b*) Indicate, through examples, how statistical principles and concepts are relevant in the following situations:

- handling of variation in characteristics (e.g., physiological, chemical) encountered in the field of health care;
- diagnosis of patients' ailments and health problems of communities;
- prediction of likely outcomes of disease intervention programmes in communities or of diseases in individual patients;
- selection of appropriate forms of treatment for individual patients;
- public health administration and planning;
- planning, conducting, analysing, interpreting and reporting of medical research.

Required previous knowledge

The students should:

- have some experience of making and using measurements that have demonstrated biological variation;
- have some knowledge of the concerns of medicine and health systems, and the broad objectives of health services;
- know the meaning of diagnosis, prognosis, and treatment;
- understand the terms science and scientific method.

New terms and concepts

The following is a list of new terms and concepts dealt with in this lesson:

biostatistics; collection, reduction, summary, analysis, and presentation of data; deterministic and probabilistic approaches; exact and inexact science; health statistics; medical statistics; probability; statistical concepts and principles; statistical methods; statistics; uncertainty; variation; vital statistics.

LESSON CONTENT

The lesson should cover the topics outlined below.

Definitions of statistics

The term "statistics" is used in two ways. First, it refers to the everyday use of:

— data
— numerical observations
— quantitative information.

Examples

1. Number of trained community health workers in the different districts of the country.
2. Birth weights of babies.
3. Age (in completed years) of patients seen at an outpatient clinic on a specific day.
4. Prevalence of schistosomiasis, per 1000 of population, in local government areas.
5. Amount of creatinine in mg per litre in a 24-hour urine specimen.

Statistics also refers to the *discipline*, comprising:

— statistical methods
— the study of scientific methods of collecting, processing, reducing, presenting, analysing and interpreting data, and of making inferences and drawing conclusions from numerical data.

Main uses of statistical methods

Three main uses of statistical methods are:

(*a*) *To collect data in the best possible way*

This includes methods of:
— designing forms for data collection
— organizing the collection procedure
— designing and executing research
— conducting surveys in a population.

Examples

1. Collection of data on participants in a disease intervention programme.
2. Systematic collection of data on births and deaths.
3. Collection of data to compare the relative effects of ergometrine + oxytocin and ergometrine alone in the third-stage management of obstetric labour.
4. Collection of data on persons with pulmonary tuberculosis in a defined population.

(*b*) *To describe the charateristics of a group or a situation*

This is accomplished mainly by:
— data reduction
— data summary
— data presentation.

(*c*) *To analyse data and to draw conclusions from such analyses*

This involves the use of various analytical techniques and the use of probability concepts in drawing conclusions.

Use of statistical concepts and principles in health care delivery

The use of statistics is essential in health care delivery, at the levels of both the community and individual patients. Medicine deals with individuals who exhibit differences in various characteristics such as weight, height, blood pressure, cholesterol, immunoglobulin levels, blood sugar, etc. What constitutes a healthy state with respect to each characteristic varies from individual to individual. No two patients or groups of individuals are ever exactly alike, yet decisions affecting patients or the community must be based on experience with other patients or communities of similar biological and social characteristics. It must be recognized that, because of the differences, these decisions cannot be exact: they are always accompanied by some uncertainty. This is the **probabilistic nature of medicine**.

It is thus necessary to be conversant with the proper techniques for coping with such differences and uncertainty.

The application of statistics is also useful in developing a critical thinking faculty, in order to be able to:

— think scientifically, logically and critically about medical problems.
— assess properly available evidence for decision-making.
— be aware of possible risks associated with medical decisions.
— identify decisions and conclusions that lack a scientific and logical basis.

Statistical principles and concepts are applied in various areas in medicine. Some examples are given below.

(*a*) *Handling of variation*

Variation in a characteristic (or factor, or measurement) occurs when its value changes from subject to subject, or from time to time within the same subject. Nearly all characteristics encountered in health care delivery, whether physiological, biochemical or immunological, exhibit variations.

Examples. Age, weight, height, blood pressure, cholesterol level, bilirubin, albumin, immunoglobulin levels, platelet count, glucose load.

Problems arise in trying to: summarize a characteristic for a group of patients, or for a community; decide for a particular characteristic what is the ideal value, the normal value, the average value, etc.; and compare two groups of patients, or two communities, with respect to a particular characteristic. Only when these problems have been clearly defined can appropriate statistical methods for solving them be decided on.

(*b*) *Diagnosis of patients' ailments and community health*

Diagnosis is the process whereby the health status of an individual, or group of individuals, and factors producing it, are identified. The various disease categories, one distinct from the other, based on clustering of signs, symptoms and magnitude of biochemical values, have often been established by procedures employing implicit statistical methods.

In placing an individual or a community's health status in one of these categories there is always some uncertainty. It may happen that the stated signs and symptoms are not exactly the same as those listed for, and defining, that category. Conversely, more than one category may have the same set of signs and symptoms ascribed to it.

Statistical reasoning is often unconsciously employed when a health worker selects a disease category with the best chance of being correct. Explicit statistical methods are available for placing two or more categories in ascending order of probability of being right.

(*c*) *Prediction of likely outcome of a disease intervention programme in a community or of disease in individual patients*

Prognosis is the assessment or prediction of the likely outcome of an intervention programme in a community or of disease in patients in the light of the presenting symptoms, signs and circumstances. The procedure draws on previous experience with similar intervention programmes or patients, and the exercise is, in principle, mainly statistical.

Records must be kept of characteristics observed at initial examination and during treatment, and of the eventual outcome of the disease in the community or in patients previously seen by the clinician. These can then be analysed to find out the detailed results for the different individuals. Determination of the possible outcome of a new intervention programme or patient treatment may be based on the results of such an analysis.

(*d*) *Selection of appropriate intervention for a patient or community*

This is based on the following:

— previous experience with similar patients or communities that had received the intervention.

— reports in the literature of clinical trials or experiments to assess the relative efficacy of different drugs and other methods of treatment.

— objective assessment of the health worker's previous experiences.

The design, execution and analysis of medical experiments and intervention programmes must employ sound statistical principles and methods if the findings and conclusions are to be valid. Otherwise, interventions may unknowingly be ineffective and even harmful.

(*e*) *Public health, health administration, and planning*

The major application here is in the use of data relating to illness in the population in order to make community diagnosis. This requires knowledge of:

— characteristics such as size and age structure of the population;

— the health profile of the population, in terms of disease or risk factor distribution;

— influence of environmental factors;

— use of vital statistics (data on births and deaths).

In health administration and planning, use is also made of data on the distribution of all levels of health care resources (need, availability, utilization etc.).