# Table of Contents

Preface ........................................................................................................... vii
Acknowledgments ......................................................................................... ix
How To Use This Book ................................................................................. xi
  To the Student ............................................................................................ xi
  To the Instructor ........................................................................................ xi
Goals and Objectives ..................................................................................... xiii
  Goals ........................................................................................................... xiii
  Objectives .................................................................................................. xiii

Chapter 1 Investigation of an Epidemic ...................................................... 1
  Objectives Covered .................................................................................. 1
  Study Notes ................................................................................................. 1
  Exercises: An Outbreak of Hepatitis in a Rural Community .................. 6

Chapter 2 Measures of Mortality ............................................................... 17
  Objectives Covered .................................................................................. 17
  Study Notes ................................................................................................. 17
  Exercises .................................................................................................... 21

Chapter 3 Incidence and Prevalence .......................................................... 23
  Objectives Covered .................................................................................. 23
  Study Notes ................................................................................................. 23
  Exercises .................................................................................................... 25

Chapter 4 Measures of Risk ...................................................................... 29
  Objectives Covered .................................................................................. 29
  Study Notes ................................................................................................. 29
  Exercises .................................................................................................... 34

Self-Assessment 1 ....................................................................................... 37
iv  EPIDEMIOLOGY AND BIOSTATISTICS

Chapter 5  Biological Variability ............................................................ 45
Objectives Covered ........................................................................ 45
Study Notes .................................................................................. 45
Exercises ...................................................................................... 49

Chapter 6  Probability ............................................................................. 53
Objectives Covered ........................................................................ 53
Study Notes .................................................................................. 53
Exercises ...................................................................................... 56

Chapter 7  Screening ............................................................................... 59
Objectives Covered ........................................................................ 59
Study Notes .................................................................................. 59
Exercises ...................................................................................... 66

Chapter 8  Sampling................................................................................ 69
Objectives Covered ........................................................................ 69
Study Notes .................................................................................. 69
Exercises ...................................................................................... 70

Chapter 9  Statistical Significance........................................................... 75
Objectives Covered ........................................................................ 75
Study Notes .................................................................................. 75
Exercises ...................................................................................... 81

Chapter 10 Correlation............................................................................. 87
Objectives Covered ........................................................................ 87
Study Notes .................................................................................. 87
Exercises ...................................................................................... 91

Chapter 11 Multiple Regression ............................................................... 95
Objectives Covered ........................................................................ 95
Study Notes .................................................................................. 95
Exercises ..................................................................................... 99

Self-Assessment 2 ................................................................................. 101

Chapter 12 Case-Control (Retrospective) Studies ................................... 109
Objectives Covered ........................................................................ 109
Study Notes .................................................................................. 109
Exercises ..................................................................................... 116
The two disciplines of epidemiology and biostatistics have become central tools for problem solving in health and medical science. In this book we provide coordinated instruction in these subjects. We cover this material at a level that will enable the student to understand and read the health sciences literature critically. The final chapter introduces a systematic procedure to aid the student in the evaluation of such articles.

This book is intended for students of the entire spectrum of the health sciences, including medicine, nursing, dentistry, pharmacy, public and community health, and allied health sciences. It is appropriate for a variety of educational formats including both traditional and problem-based learning approaches. The scope and level are appropriate for students in professional schools, universities, four-year colleges, community colleges, and, in fact, any setting where courses in health sciences are taught. Our text is based on a course given to second-year students at the University of Maryland School of Medicine. Our students have found this book an effective study guide. It has been used successfully for individual, self-paced instruction by selected students.

Our book has a distinct method, based on pedagogic principles. The 46 objectives, expressed in behavioral terms, cite the concepts to be learned and the level at which students are expected to perform. There are three elements to the book, the first being the study notes. These notes may be read as the sole source of input to cover the material, or they may be used to supplement attendance at a lecture series. They are not designed to replace or compete with existing textbooks of epidemiology and biostatistics, which are specifically referenced following each chapter. (Students are encouraged to augment their learning by referring to these books.) The second element is the exercises that accompany each chapter. The exercises encourage students to use their newly acquired knowledge immediately, and improve retention by practice. Detailed feedback is provided in the exercise answers, which are amplified at points where we have learned that some students have difficulty. The third element in the instruction process is the multiple-choice examinations, which have the same scope and are on the same level that students may expect to encounter in professional examinations. Students are challenged both progressively as material is introduced and comprehensively in a final examination.
A guiding principle for us in choosing new material has always been to keep the book no more complicated than necessary to achieve the goal of providing a basic understanding of the biostatistical and epidemiological content found in the health sciences literature. For example, the concept of statistical power is hardly a new idea, having been discussed by statisticians for over 70 years. However, despite its elemental role in statistical inference, it is a topic we have avoided in previous editions. This is because, until recent years, it was infrequently mentioned in the literature likely to be consumed by our readers. Now, almost all clinical trials are required to use a power calculation to justify the number of subjects they enroll. Power information is thus often included in the published reports of these trials. Consequently, we have added a new section in Chapter 9 devoted to the subject of statistical power. Introducing the reader to statistical power also enabled a discussion of two other new topics in Chapter 9, namely, one-tailed $P$ values and study-wise significance levels. Confidence intervals and tests of statistical significance had been discussed separately in previous editions. Our experience has been that students benefit by recognizing the relationship of one to the other. A description of this relationship is now included in Chapter 9.

Other topics have been added to this edition for reasons similar to that described previously for statistical power. A discussion of receiver operating curve (ROC) analysis now appears in Chapter 7. The concept of propensity scoring for the reduction of bias in observational studies is described in Chapter 16. The discussion of meta-analysis in Chapter 17 has been expanded to include current methods for assessing study-to-study variation in results and publication bias.
Acknowledgments

We thank the students of the University of Maryland School of Medicine for providing us with the experience necessary to make this book possible.
How To Use This Book

To the Student

We suggest that as you begin each chapter, you first read the accompanying study notes. If they are not sufficiently detailed, please refer to the standard textbooks listed in the Recommended Readings for that chapter, which are cited in ascending order of difficulty.

Following each chapter are exercises; you should answer the questions, preferably writing down your responses, before consulting the solutions provided. When the first third of the book is completed, you will encounter a multiple-choice self-assessment exam consisting of 25 questions. Complete the exam and score your efforts (answers are given on page 201). You should score at least 60% before proceeding to the next chapters. If you should fail to achieve this level at your first attempt, we suggest that you restudy the sections with which you had difficulty. Similar self-assessment exams are provided following Chapters 11 and 17. A passing grade of 60% is also suggested for these. The final examination, covering all the objectives, may then be attempted. We suggest that 75% constitutes a passing grade on the total 125 questions and that 90% indicates an honor grade.

To the Instructor

This book is versatile. It may be used as a course textbook for a formal lecture series given to large groups. It also has a role in a seminar series, freeing the instructor from didactic teaching, thus enabling more complete discussion of relevant examples. It may be used as a vehicle for an independent study program in which the faculty member assumes the role of a tutor. The exclusive feature of our large question bank provides the instructor with a ready-made assessment instrument to monitor the progress of the students. Specific weaknesses may be identified and focal remediation given.
Goals and Objectives

Goals

After completing this study guide, the student will be able to:

1. Apply epidemiologic methods to critically evaluate the evidence used in medical decision making.
2. Assess data by using epidemiological and biostatistical principles and evaluate conclusions based on study data.

Objectives

1. Interpret the distribution of disease in a population according to time, place, and person.
2. Describe the composition of a rate, in terms of the numerator and denominator, and explain the relationship between them and the importance of time.
3. Explain the use of rates for comparative purposes.
4. Define attack rate, and use it to identify a vehicle of transmission in a common-source outbreak of disease.
5. Define crude mortality rate, specific mortality rate (age, sex, race, and cause), case fatality rate, and proportionate mortality ratio. Cite one example of the correct use of each of these rates, and interpret statements containing them.
6. Explain what is meant by confounding.
7. State the reasons for adjustment of rates, and interpret statements containing adjusted rates.
8. Define incidence, both cumulative incidence and incidence density, and prevalence; state the relationship between them.
9. Name the factors that may cause variation in each measurement. Give the uses of each rate.
10. Define absolute risk, relative risk, and attributable risk, and interpret statements that employ these terms.
11. Explain what is meant by effect modification.
12. State the purpose of a frequency distribution and cumulative frequency distribution in describing a set of biological measurements.
13. Define mean, median, mode, and percentile, and describe the features of a distribution that each characterizes.

14. Contrast the features of a normal (Gaussian) distribution to those of a skewed distribution.

15. Explain why the mean ±2 standard deviations is often used to establish the “normal range” and what practical difficulties might be encountered using this procedure in clinical practice.

16. Determine probabilities and odds from frequency distributions.

17. Explain what is meant by conditional probability.

18. Calculate the probability of complex events by applying the addition and multiplication rules.

19. Describe the application of screening to primary, secondary, and tertiary prevention; list the conditions necessary to justify screening.

20. Define the validity measurements of sensitivity and specificity, explain the purpose of receiver operating characteristic (ROC) analysis; as well as predictive value and reliability of a screening test; compute and use validity and predictive value; and interpret reliability measures.

21. Use the standard error to compute 95% confidence limits for a mean or a proportion, and interpret statements containing confidence limits.

22. Explain sampling bias, and describe how random sampling operates to avoid bias in the process of data collection.

23. Distinguish between the standard deviation and the standard error, and give one example of the use of each.

24. Interpret statements of statistical significance with regard to comparisons of means and frequencies, explain what is meant by a statement such as \( P < 0.05 \) and distinguish between the statistical significance of a result and its importance in clinical application.

25. Explain the following regarding statistical tests of significance: the power of a test; the relationship between significance tests and confidence intervals; one- versus two-tailed tests; and comparison-wise versus study-wise significance levels.

26. Interpret the relationship between two variables as displayed on a scattergram, distinguishing between positive, negative, and zero correlation.

27. Explain the information provided by a regression equation as well as that provided by a correlation coefficient.

28. Interpret statements of statistical significance with regard to the correlation coefficient.

29. Distinguish between a simple regression equation and a multiple regression equation.

30. Interpret the slope coefficients in a multiple regression as well as statements regarding their statistical significance.

31. Explain the information provided by the coefficient of determination.
32. Describe applications of the logistic, Cox proportional hazards, and longitudinal regression models.
33. Distinguish between experimental and observational studies.
34. Describe a case-control (retrospective) study.
35. Describe cohort (prospective) and cross-sectional studies.
36. Define cohort, and recognize a cohort effect when interpreting cross-sectional data.
37. Describe a randomized clinical trial.
38. Describe the circumstances that require prospective studies to control for differences in time under observation.
39. Describe how the following methods control for differences in time under observation, and indicate what assumptions underlie each:
   a. Person-time
   b. Life table, Kaplan-Meier analysis
40. Describe the role of proportional hazards methods in survival analysis.
41. Illustrate with one example the concept of multifactorial causation of disease.
42. Define the following types of association:
   a. Artifactual
   b. Noncausal
   c. Causal
43. Distinguish between association and causation, and list five criteria that support a causal inference.
44. Use the abstract of a journal article from the medical literature to ascertain the study objectives, target population, and especially the study design.
45. Use eight standard questions as a guide for the critical examination of the article paying special attention to those that touch on specific areas of vulnerability of each design type.
46. Use systematic reviews and meta-analyses to weigh evidence derived from several studies.