BIOSTATISTICS

Faculty

Stefanos Bonovas, MD, MSc, PhD
Associate Professor of Medical Statistics

Year/Semester

Year 4 - first semester

Credits: 4

Suggested textbooks

Biostatistics:

- Altman DG. Practical Statistics for Medical Research. Chapman and Hall/CRC.
- Bland M. An Introduction to Medical Statistics. Oxford University Press.

Epidemiology:

- Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Lippincott Williams & Wilkins.
- Bonita R, Beaglehole R, Kjellström T. Basic Epidemiology. World Health Organization. (available free at: http://apps.who.int/iris/bitstream/10665/43541/1/9241547073_eng.pdf)

Overview of the course

Biostatistics is crucial in ensuring that findings and practices in public health and medicine are grounded in reliable evidence. Epidemiology, meanwhile, focuses on the study of the distribution and determinants of health and disease in human populations, as well as the application of methods to improve health outcomes. Together, these two scientific fields form the foundation of health research and evidence-based public health policy and practice.

This course introduces the basic principles and methods of biostatistics and epidemiology, with an emphasis on critical thinking, analytical skills, and application to clinical practice and research. Epidemiological and statistical concepts are taught in parallel to promote integration of knowledge across both disciplines.

The course will cover three domains: (i) Biostatistics: basic concepts; types of data; frequency distributions; methods of summarizing data; central tendency and dispersion; concepts of probability; basic concepts of statistical inference; hypothesis testing; types of statistical errors; parametric and non-parametric methods in medical research; power and sample size calculations; statistical analysis using real data and statistical software with emphasis in concepts and interpretation; (ii) Epidemiology: basic principles; measures of disease occurrence; measures of association; epidemiological study designs; cohort, case-control, cross-sectional and ecological studies; experimental designs and randomized controlled trials; criteria of causality; biases (systematic errors); confounding; synthesis and evaluation of the evidence by means of systematic reviews and meta-analysis; and methodologic approaches for the development of clinical practice guidelines; and (iii) Research methodology: how to prepare a research manuscript for publication in a medical journal; how to submit an article for publication; practical tips for publishing; research ethics and misconduct.

General learning goals

Upon completing this course, the students should be able to describe the basic epidemiological concepts, interpret and critically evaluate literature relevant to public health professionals, select and apply appropriate statistical methods for managing common types of health data, interpret and communicate the results to public health professionals and to the general public.

Assessment

The final exam for this course will consist of 30 multiple-choice questions, with an additional bonus question offering a chance to earn extra credit. Each correct answer earns one point, and you will have 60 minutes to complete the exam. The questions will assess your understanding of the material from lectures, workshops, and statistical software sessions. A thorough grasp of these topics is crucial for success. Only correct answers contribute to your score; unanswered or incorrect responses will not be penalized. To pass, you must achieve a minimum score of 18 out of 30. A perfect score of 31 out of 31, including the bonus question, will earn you the Cum Laude distinction.

Analytical content of the course

1. Introduction to biostatistics: summarizing data (2 hours)

Learning goals: To distinguish the different types of data, explain how to construct a frequency distribution, calculate and interpret the measures of central location: mode, median, and mean, demonstrate how to apply the most appropriate measure of central location for a distribution, and apply and interpret four measures of spread: range, interquartile range, standard deviation, and confidence interval (for mean).

2. Introduction to epidemiology: historical evolution (2 hours)

Learning goals: To define epidemiology and explain its goals, identify historical figures and events that played a decisive role in the evolution of epidemiology and public health, and list the key features and uses of descriptive and analytical epidemiology.

3. Measures of disease occurrence (2 hours)

Learning goals: To understand how to calculate and interpret key measures of disease occurrence, including proportion, ratio, and rate, as well as prevalence, incidence proportion, incidence rate, and attack rate. A brief workshop (exercises and discussion) will help the students apply the lesson content.

4. Measures of association (2 hours)

Learning goals: To learn how to calculate and interpret key measures of association, including risk ratio, rate ratio, and odds ratio, as well as measures of public health impact, such as attributable proportion, population attributable fraction, and efficacy or effectiveness. A brief workshop (exercises and discussion) will help the students apply the lesson content.

5. Experimental study designs: clinical trials (2 hours)

Learning goals: To discuss the key concepts of clinical trials: fundamentals of clinical trial design, strengths and limitations, intention-to-treat and per-protocol analysis.

6. Special issues in clinical trials (2 hours)

Learning goals: To further discuss some concepts in clinical trials: multiple testing, missing data, interim monitoring and stopping rules, as well as quality assessment of clinical trials.

7. Observational study designs: cohort and case-control studies (2 hours)

Learning goals: First, to discuss the features of a cohort study, distinguish between prospective, retrospective, and ambidirectional cohort studies, and explain the strengths and weaknesses of the cohort study design. Second, to define and explain the features of a case-control study, identify the types of questions that can be addressed by case-control studies, learn how to estimate and interpret odds ratios from case-control studies, and discuss their strengths and limitations.

8. Workshop (2 hours)

Learning goals: This workshop, through a series of exercises and discussions, aims to help students apply and reinforce the concepts from previous lessons.

9. Statistical testing in medical research I (2 hours)

Learning goals: To explain the concepts of sample, population, and statistical inference, describe the practical procedures for hypothesis testing, discuss the types of statistical errors, and learn the basic parametric and non-parametric statistical methods used in medical research: one-sample t-test and sign test, two-sample t-test and two-sample Wilcoxon rank-sum (Mann-Whitney) test, paired t-test and matched-pairs Wilcoxon signed-rank test, chi squared test, and correlation analysis.

10. Statistical testing in medical research II (2 hours)

Learning goals: To explain the concepts of sample, population, and statistical inference, describe the practical procedures for hypothesis testing, discuss the types of statistical errors, and learn the basic parametric and non-parametric statistical methods used in medical research: one-sample t-test and sign test, two-sample t-test and two-sample Wilcoxon rank-sum (Mann-Whitney) test, paired t-test and matched-pairs Wilcoxon signed-rank test, chi squared test, and correlation analysis.

11. Bias and confounding, causal inference (2 hours)

Learning goals: To define bias (systematic error) and differentiate between the three main types: selection bias, misclassification/information bias, and confounding. Students will learn to identify common sources of bias, recognize which types of studies are prone to each bias, distinguish between association and causality, and describe and apply Hill's criteria for evaluating causal relationships.

12. Screening and test validity (2 hours)

Learning goals: To discuss which types of diseases are suitable for screening, define, calculate, and interpret the sensitivity, specificity, positive predictive value, and negative predictive value of a screening test, explain how adjusting the cut-off value affects the sensitivity and specificity of the test, and understand how the predictive values are influenced by disease prevalence. A brief workshop (exercises and discussion) will help the students apply the lesson content.

13. Evidence synthesis: systematic reviews and meta-analysis (2 hours)

Learning goals: To discuss the main concepts of systematic reviews and meta-analyses, explain how to interpret their results, and apply to a patient scenario.

14. Evidence-based medicine: GRADE & Oxford systems (2 hours)

Learning goals: To familiarize with current methodological approaches for the guidance of clinical practice, discuss the GRADE (Grading of Recommendations Assessment, Development and Evaluation) system for developing clinical practice guidelines, as well as the OCEBM (Oxford Centre for Evidence-Based Medicine) Levels of Evidence that can be used by clinicians to find the likely best evidence and answer clinical questions very quickly.

15. How to prepare, submit and publish a research article (2 hours)

Learning goals: To discuss how to prepare a research manuscript for publication in a medical journal, how to submit it for publication, and some practical tips for publishing.

16. Research ethics and misconduct (2 hours)

Learning goals: To discuss ethics in research, focus on the moral principles that researchers must follow in clinical research, highlight the importance of ethical publishing, and outline the definitions of misconduct.

17–22. Statistical Software Sessions (six lessons: 12 hours)

Learning goals: To work with Stata software, explore the data, and deepen our understanding of biostatistics. These six sessions will include: Introduction to Stata software, presentation of data using graphical methods, statistical analysis using real datasets, hypothesis testing, comparison of means and proportions, chi square test, correlation analysis, non-parametric tests, as well as sample size and power calculations. Emphasis will be given in concepts and interpretation.

23. Open discussion on the central concepts of the course: Q/A (2 hours)

Learning goals: To perform an overview/synthesis of the course, look back over the topics covered and try to weave them together, and identify common themes, and recurring issues. Students can ask any questions at all, from material in the course to any other topic of interest.

24. Simulation of the final examination (2 hours)

Learning goals: To offer students practice with the types of questions they will encounter on the actual test, enhancing their readiness and confidence.