UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Big Data in Health Care

2526-2-FDS01Q026-FDS01Q029M

Aims

This course aims to provide the basic concepts of clinical epidemiology that are at the basis of a proper methodological approach to a research project in medicine. The student will be able to deal with big data in health care particularly focusing on several aspects including design, data management and analysis. The student will be able to implement optimal design strategies for clinical trials and observational studies. The student will be able to apply regression models for the analysis of time to event data and to use these models for risk prediction and for the assessment of relations between variables. The student will be able to use proper statistical methods for the analysis of omics data.

Knowledge and understanding

This course provides basic knowledge and understanding on:

- clinical study design
- fundamental quantities in survival analysis (survival, cumulative incidence, hazard funcitons)
- statistical methods to deal with time-to-event data boh in standard and high-dimensional contexts
- predictive and utility performance measures of a clinical prediction model

Ability to apply knowledge and understanding

At the end of this course the students will be able to:

- estimate and compare survival functions using non-parametric methods (Kaplan-Meier estimator and Log-Rank test)
- fit a Cox regression model and interpret the coefficients
- use the model to predict profile-specific survival function
- evaluate the prediction and utility performance of a clinical prediction model
- use a penalized model in the context of high dimensional data (omics)

Making judgements

At the end of this course the students will be able to:

- verify the validity of the assumptions required by the Cox regression model
- judge and compare the calibration and discrimination performance and the clinical utility of prediction models

Contents

Clinical epidemiology. Designs of clinical trials and observational studies. Statistical methods for time to event data, competing risks, predictive and utility performance measure for prediction models. Statistical methods for the analysis of omics data.

Detailed program

Basics in clinical epidemiology.

Study design in medicine: cross-sectional, case-control, case-cohort, randomized trials, biomarker driven designs.

Statistical methods for time-to-event data (Kaplan-Meier estimator, Log-Rank test, Cox regression model), competing risks and multistate models, risk prediction models (building and validation).

Penalized regression methods for the analysis of time-to-event data in omics.

Prerequisites

Descriptive and inferential statistics.

Teaching form

Teaching with face-to-face lectures and laboratory activities:

- 3 lectures of 3, 2 and 2 hours (7 hours total) conducted in face-to-face delivery mode;
- 7 lectures of 3 or 2 hours (18 hours total) of computer laboratory activities (through the use of R software) in interactive in-presence mode.

Textbook and teaching resource

Teaching slides.

The following textbooks are suggested:

Machin D., Campbell M.J. (2005). Design of studies for medical research, Chichester: John Wiley & Sons.

E. Marubini, M.G. Valsecchi (1995). Analysing survival data from clinical trials and observational studies, Chichester: John Wiley & Sons.

Vittinghoff, E., Glidden, D. V., Shiboski, S. C., McCulloch, C. E. (2005). Statistics for biology and health. Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models. New York, NY, US: Springer Publishing Co.

Laird N.M., Lange C. (2011). The fundamentals of modern statistical genetics. New York, NY, US: Springer Publishing Co.

Semester

Second semester

Assessment method

The exam will consist of two tests:

- individual project concerning the analysis of big data to test the student's ability to apply the research methodology in clinical studies. The student's ability to develop a complete analysis using the R software and to correctly interpret the results will be assessed. The vote (out of thirty) will weigh 60% on the final judgment.
- questionnaire with multiple choice answers to test the preparation on the theoretical aspects of the overall program. The student's knowledge of some key concepts concerning both the design and the analysis of clinical studies will be evaluated. The vote (out of thirty) will weigh 40% on the final judgment.

No in-course evaluations are planned.

Office hours

On request by e-mail.

Sustainable Development Goals

GOOD HEALTH AND WELL-BEING | LIFE ON LAND