Epidemiology and big data (INFOMEBD)

BLOCK D

General information

Course ID

INFOMEBD

Course type

Course

Credits

7.5 ECTS

Category / Level

Master

Instruction language

English

Offered by

Utrecht University - Faculty of Science - Graduate School of Natural Sciences - - - -

Content

Aims

a. Use and explain the basic concepts used in epidemiology  
b. Understand mechanisms giving rise to missing data  
c. Describe key assumptions and apply imputation methods to deal with missing data  
d. know when to apply a mixed model in practice;  
e. be able to perform mixed model analyses using statistical software (R);  
f. Understand the advantages, limitations and key characteristics of IPD-MA in intervention, diagnostic and prognostic research  
g. Understand the relevance of between-study heterogeneity, and be familiar with statistical methods for investigating and reporting this.  
h. Be familiar with statistical methods for summarizing relative treatment investigating diagnostic test accuracy using IPD from multiple sources  
  
*Assessment*  
Assignments (reflection, evaluation), case study (project).

Content

This course provides insight in the basic principles used in epidemiology, such as bias and confounding and students will learn statistical methods to address missing data and correlated data. Possible mechanisms for data being missing,  
their potential impact in terms of bias, and methods to handle missing data will be discussed.  
Correlated data may occur because response variables are observed more than once per individual, or when there is a hierarchical (multilevel) structure in the data, e.g. patients within hospitals, pupils within classrooms, etcetera.  
Mixed models are one way of analyzing this kind of data. Systematic reviews and meta-analyses are methods to summarize published aggregate data, but it is increasingly common that individual participant data (IPD) are obtained from multiple primary studies leading to big data.

In this course, we therefore also will discuss how a meta-analysis involving IPD can be interventions, to investigate the accuracy of diagnostic tests, to develop clinical prediction models and to externally validate such models.

Students will apply all these methods during computer labs and assignments.  
  
*Pre-requisites*  
Mandatory course Data Wrangling & Data Analysis  
  
*Course form*  
Lectures, tutorials and practicals.  
  
*Study materials*  
• Lecture notes Epidemiology and Big Data  
• Reader  
• Computer labs  
• Data sets for assignments and case study