Bayes Bayesian Biostatistics Leiden, NL 2025

Short Course

# Bayesian and Causal Networks for Clinical and Epidemiological Data: Concepts, Implementation and Interpretation

Marco Scutari
Dalle Molle Institute for Artificial Intelligence and Consultant.

**ABSTRACT:**
Clinical and epidemiological data have become increasingly complex as we try to model the interplay between risk factors and comorbidities with better detail and precision. Bayesian networks have long been a powerful tool to model them, offering flexibility and interpretability over classical statistics and machine learning approaches. Causal networks augment them to express cause-effect relationships, allowing for interventions, counterfactuals, and confounding. However, they present unique challenges that require thoughtful design and careful use of expert information. This course explores the practical challenges in using these models and offers practical solutions using the bnlearn R package (https://www.bnlearn.com/).
The course will first introduce Bayesian networks as the prototypical artificial intelligence model that learns a working representation of a clinical phenomenon and can then answer arbitrary queries. It will then extend them to Causal networks and discuss their interpretation and use for causal reasoning with interventions and counterfactuals.
These concepts will be explained hands-on using bnlearn and two fully reproducible analyses developed with Pfizer (EHR data) and La Roche-Posay/L'Oréal (infodemiology). Participants will have the opportunity to work through all steps, including data preparation, network structure learning, parameter learning, model evaluation/validation and answering clinically relevant queries. They will also explore the impact of different design decisions, gaining insights into the trade-offs of various options.