Practical for "Causal Learning" on the UBRA Data Train – December 2021

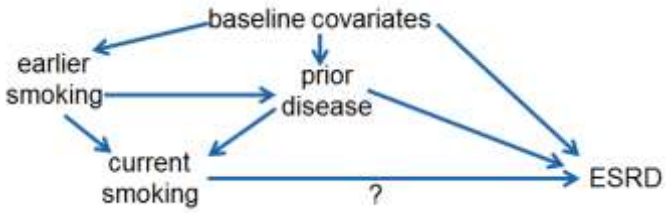
Vanessa Didelez, Christine W Bang, Ronja Foraita

Part 1 (Thursday morning)

(1.1) Consider the causal DAG below (Staplin et al, 2016). ESRD = 'end-stage renal disease' is the binary outcome and all other nodes represent binary covariates with a set of baseline variables not further detailed.
Use the web-based version of `DAGitty` (or if you prefer, there is some example R code in `datatrain1`).

Enter the DAG below in DAGitty and investigate:

- (i) With / (ii) without data on 'early smoking' are there any testable implications of the causal model?
- If we want to identify the causal effect of (i) prior disease, or (ii) current smoking, or (iii) earlier smoking on ESRD, what do we need to adjust for (and for what not), respectively?
- Specifically: can the causal effect of current smoking on ESRD be identified without data on 'early smoking'? Convince yourself that your chosen adjustment set blocks all back-door paths.
- Add a node for a further, unobservable, variable that directly affects prior disease and ESRD. Can the effect of current smoking on ESRD still be identified from the measured data? Explain your conclusion.



(1.2) A study aims at investigating the causal effect of exposure to a particular toxic substance on time-to-pregnancy (TTP) as a measure of fertility. Couples are recruited from clinical records around the time of giving birth with recruitment taking place in period of a few months. They are given a questionnaire and, among others, asked about past exposures to various toxic substances and about the duration from initiation to pregnancy ('initiation' = when a couple stops using contraceptives). TTP ranges between a few months to >10 years.

The study design implies that couples selected into the study who have had a long TTP must have initiated in earlier years than those with short TTP. Note also that due to new legislation, exposure to the toxic substance is much less likely in recent than in former years.

Represent this situation as a causal DAG and investigate whether the effect of exposure on TTP can be estimated via adjustment in such a study.