**Day 4 case study assignment**

**Part 1: Introductory Presentation**

We ask the participant who introduced the case study to present max 3 slides about the background, addressing the following aspects:

* Some brief clinical background to the topic
* What interventions are of interest for the population of interest – e.g. treatments, extra screening, lifestyle interventions.
* What type of data you have in mind – how was it collected / where / in which time period / sample size / type of variables etc. [you do not need to have access to the data – it can be thinking of a future project]

**Part 2: Estimand**

Within the working group you can pretend that the participant who introduced the case study is your ‘clinical collaborator’ who wants the rest of the group to help them develop a great causal prediction model.

The rest of the group starts to elicit information from the ‘clinical collaborator’ that they need to formulate an exact estimand. You can use the following table to denote the answers:

|  |  |  |
| --- | --- | --- |
| **Estimand element** | **Answer** | **Questions that help formulate the estimand element** |
| Target population |  | * To which individuals will the prediction model be applied? * In which health care setting will the prediction model be applied? |
| Moment(s) of intended use |  | * At which moment(s) is the prediction model (re)consulted to inform the intervention decision? |
| Intervention options |  | * Which intervention options are relevant at the moment(s) of making the intervention decision? * For how long should the intervention strategy be fixed? * Should the duration to fix the intervention option be aligned with the time till next moment of prediction? |
| Outcome and prediction horizon |  | * Which outcome(s) are informative for the intervention decision? * What prediction horizon provides important information for the intervention decision: a short-term or long-term horizon? * Should the outcome be defined differently because of the specified intervention option(s)? * Should the prediction horizon be aligned with the time till next moment of prediction? |
| Predictor(s) |  | * Which clinical and/or demographic patient factors are predictive of the outcome of interest? * Which measurements are available at the moment(s) of intended use? * Note predictors need not be the same as confounder |

**Part 3: Assessing the design of the data source**

Now ask the ‘clinical collaborator’ about the type of data they would expect to have available.

1. How were individuals selected for inclusion in the data set? Was the inclusion process in some way related to treatment or outcome?
2. Specify how treatment decisions were made. How was adherence to the treatment?
3. Draw a (first version of) a causal DAG depicting the causal structure of the setting where the data was collected.

**Part 4: Development and evaluation of a causal prediction model**

* 1. a. What strategy would you choose to identify the estimand?
     1. What are the assumptions required for this identification?
     2. Are these assumption plausible?
  2. How do you plan to handle treatment in the analysis? Decide whether to use the treatment as a variable in the model, restrict on it (for baseline treatments only), censor or otherwise.
  3. How will confounding be addressed?
  4. How is time zero, i.e. baseline, defined? How will you align the time of decision making, eligibility, treatment assignment, and start of follow-up?
  5. What data preparation steps do you envisage, e.g. restructuring?