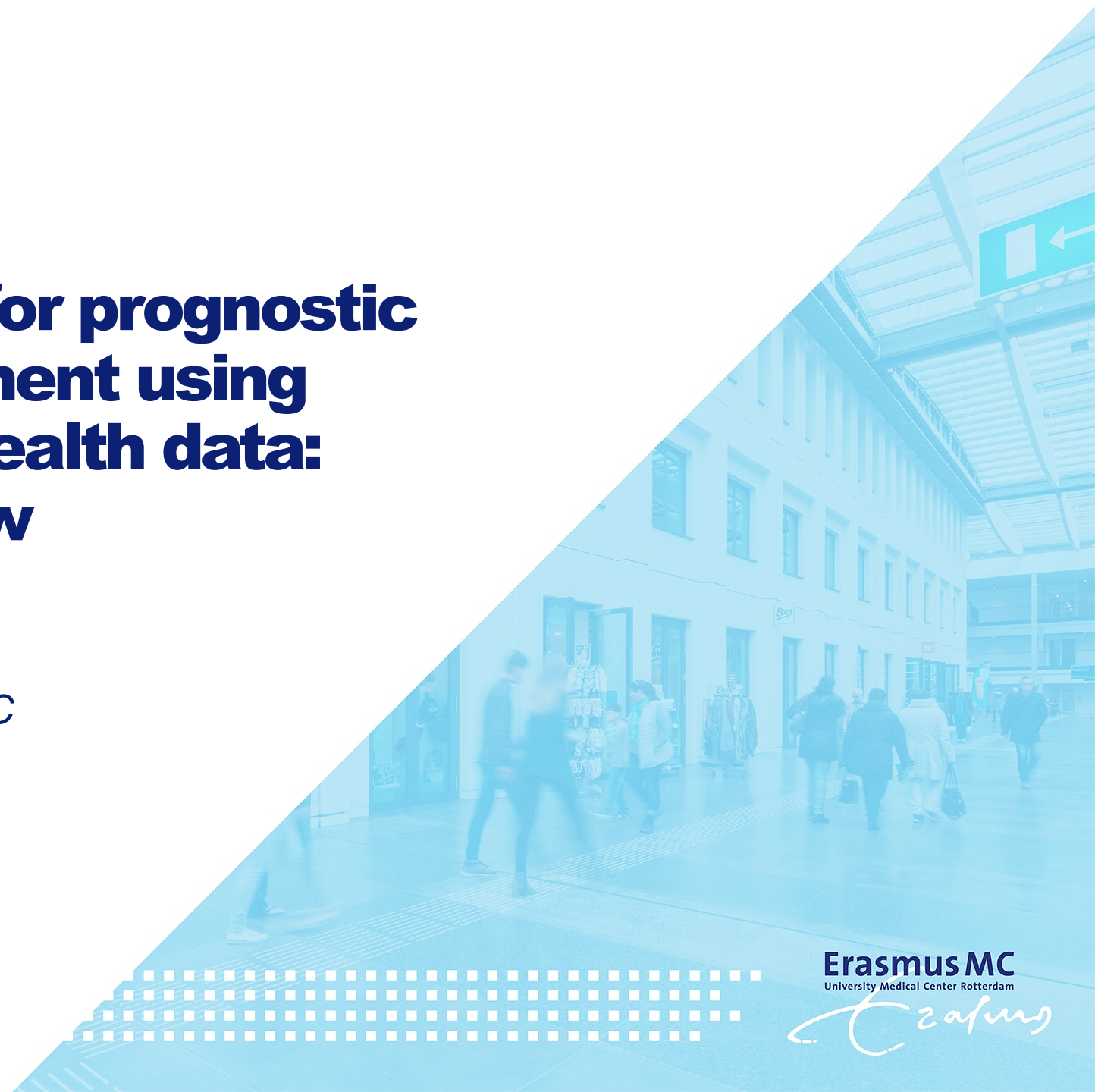
# Best practices for prognostic model development using observational health data: a scoping review

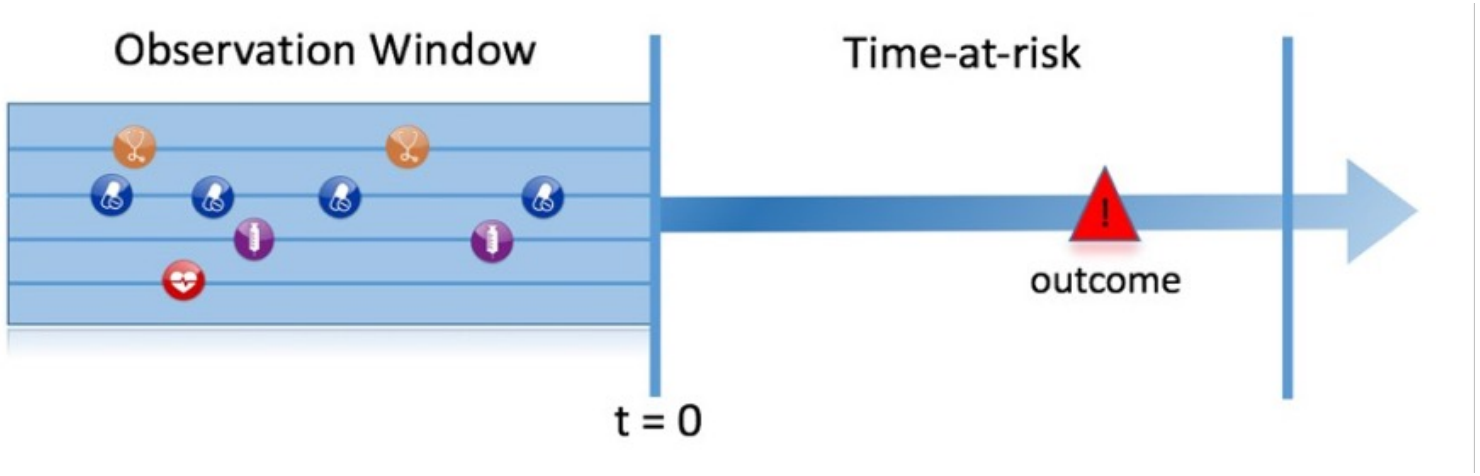
***Cynthia Yang***
*PhD Student at Erasmus MC*



Erasmus MC
University Medical Center Rotterdam

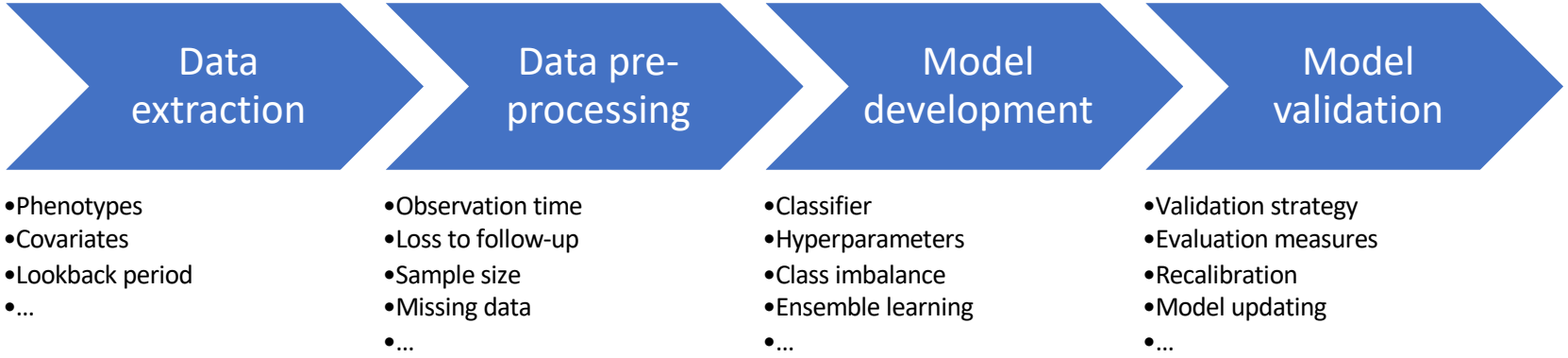
# Prognostic model development

- Among a **target population** of patients at an index date (t = 0), predict which patients will experience some **outcome** during a **time-at-risk** period.



# Study design choices

- Four study design components of prognostic model development:



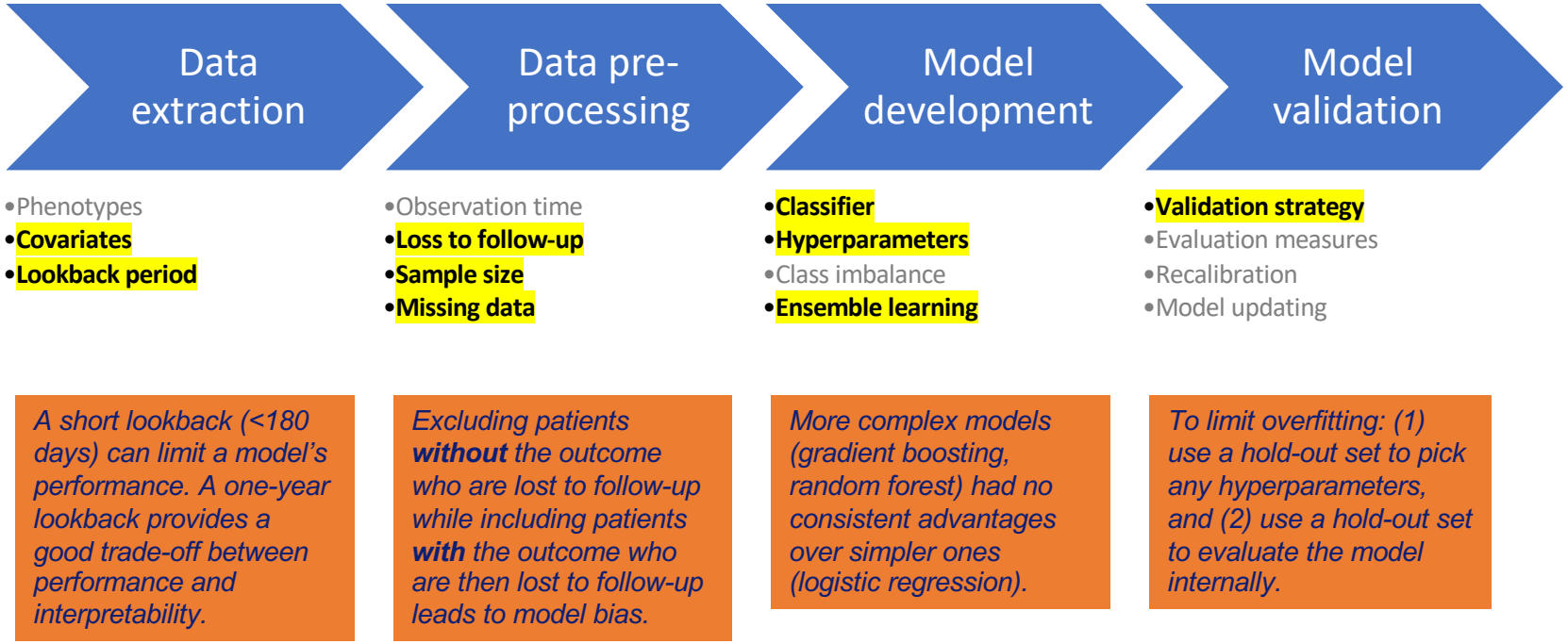
# Scoping review

- Aim: to present an overview of current knowledge and **empirical evidence**.

- Preliminary PubMed search for empirical studies.

1. 6,671 titles inspected of studies from January 2018 to June 2022.
2. 36 potentially relevant studies, 11 papers selected.
3. Data extraction: study design component, recommendations for best practices, whether the OMOP CDM was used, and classifier.

# Preliminary results



# Preliminary results

- 6 out of 11 studies used OMOP CDM data.
- 3 out of 11 studies used MIMIC-III data.
- 2 out of 11 studies used other data.

*MIMIC-III data → OMOP CDM data?*

- 5 out of 11 studies only used logistic regression.
- 6 out of 11 studies also used other classifiers.

*Logistic regression → Other classifiers?*

# Conclusions

- Model development should follow learned best practices.
- A lot of study design choices still require more research.

- Next steps: improve search strategy, more empirical research!

- For more information or questions please visit me at my poster.