UNIVERSITY OF BIRMINGHAM | COLLEGE OF MEDICAL AND DENTAL SCIENCES

NIHR Statistics group: Diagnosis and Prognosis

Dr Sue Mallett, Alice Sitch & Christina Easter

Test Evaluation Research Group

Institute of Applied Health Research, University of Birmingham

NIHR Statistics group: Diagnosis and Prognosis

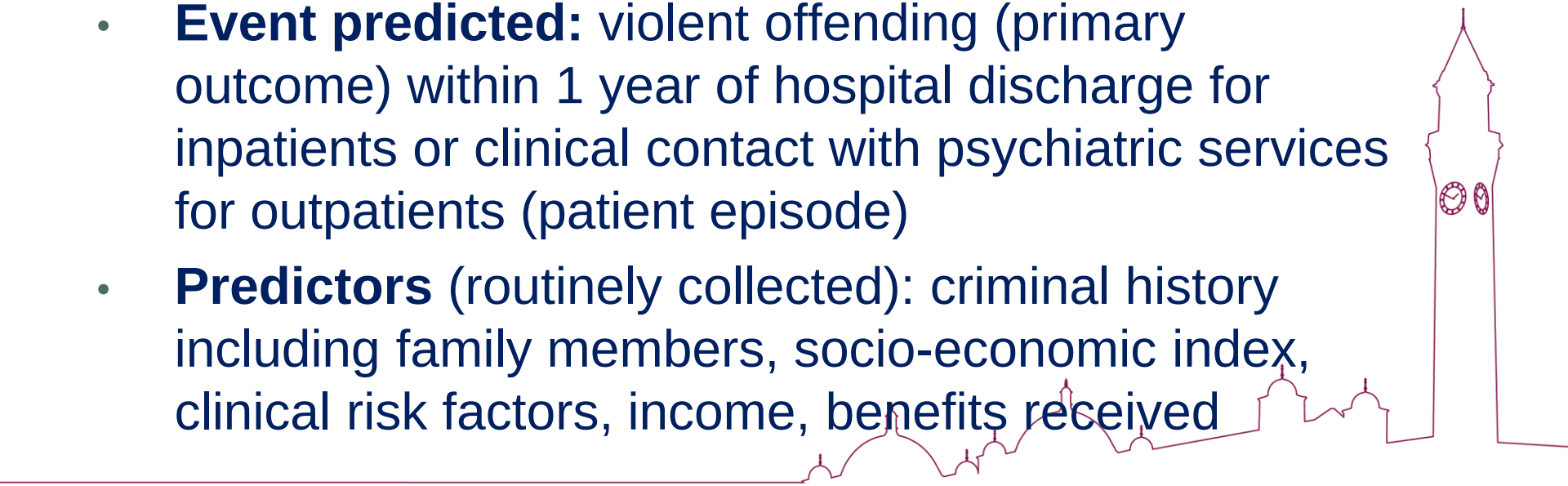
- Welcome to the first meeting of the NIHR Statistics group for Diagnosis and Prognosis
- Current main contact for group: s.mallett@bham.ac.uk
- Aim from meeting
  - Network people working in diagnosis and prognosis
  - Identify people interested in joining a group
  - Identify people interested in helping organise a group

Overview

- Background on ongoing programme of work
  - Example of collaborative project
  - Prof Seena Fazel (Oxford) and Sue Mallett (Birmingham)
- Scenarios for small group discussions

Prediction model: OXMIV

- **Participants:** national cohort of 75,158 Swedish individuals aged 15–65 years
  - with a diagnosis of severe mental illness (schizophrenia spectrum or bipolar disorder)
  - 574 018 patient episodes between Jan 1, 2001, and Dec 31, 2008
- **Event predicted:** violent offending (primary outcome) within 1 year of hospital discharge for inpatients or clinical contact with psychiatric services for outpatients (patient episode)
- **Predictors** (routinely collected): criminal history including family members, socio-economic index, clinical risk factors, income, benefits received



Prediction model: OXMIV

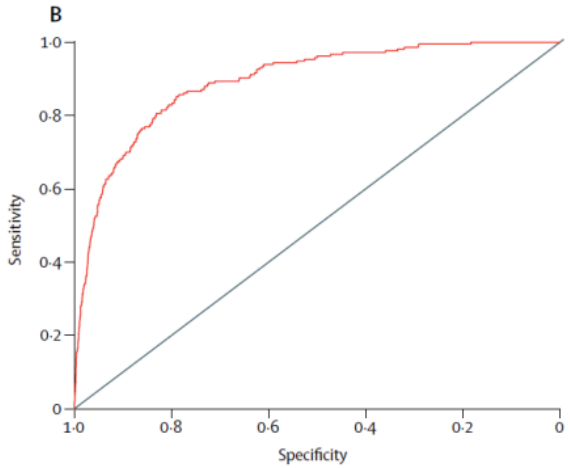
Identification of low risk of violent crime in severe mental illness with a clinical prediction tool (Oxford Mental Illness and Violence tool [OxMIV]): a derivation and validation study

Seena Fazel, Achim Wolf, Henrik Larsson, Paul Lichtenstein, Susan Mallett, Thomas R Fanshawe

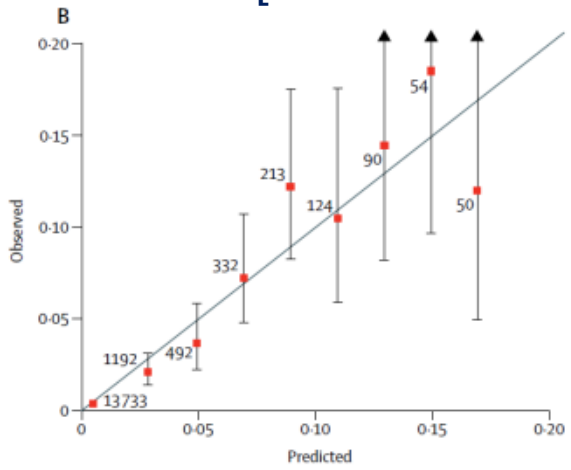
- Derivation dataset: 58,771 with 830 events
- External validation dataset: 16,387 with 220 events
- Data split by geographical region, stratified by urban/rural
- 16 Predictors
- Web calculator tool

Steps in evaluating prediction tools

OXMIV external validation



c-index of 0·89 [95% CI 0·85–0·93]



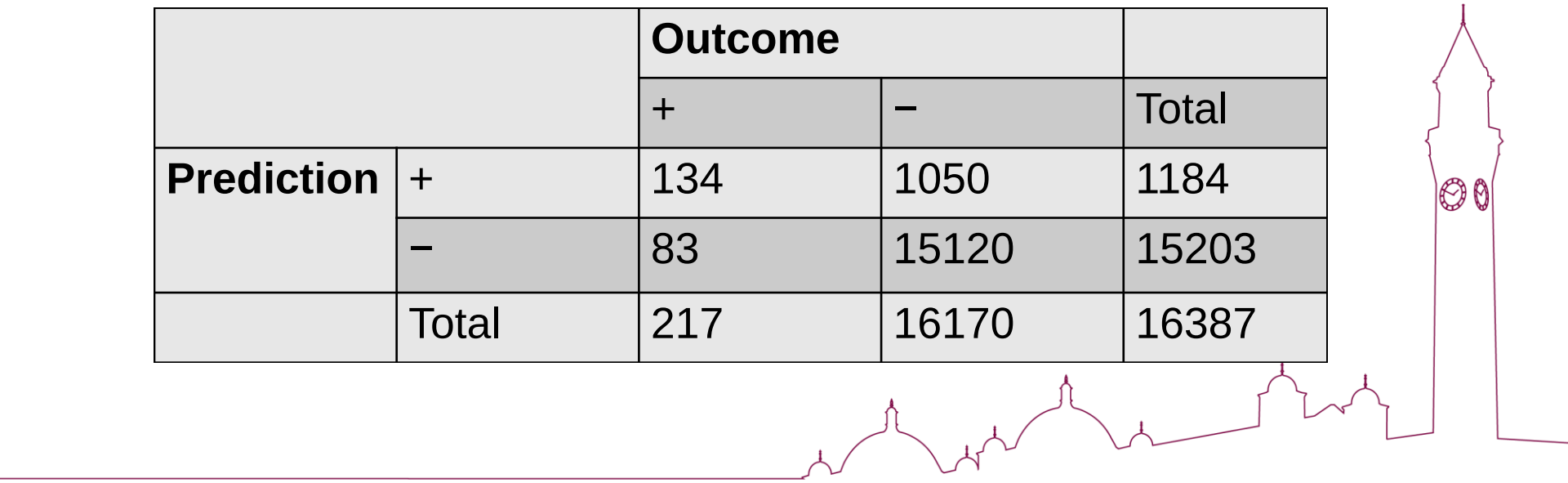
- Most prediction model articles are about developing or validating a model
- Results may be given as
  - Discrimination: c-index (average performance across all thresholds)
  - Calibration: plot of observed vs predicted probabilities
  - Classification or accuracy: performance at a particular threshold e.g. sensitivity and specificity

Steps in evaluating prediction tools

At pre-specified 5% risk cut off for violent crime in 1 year

- sensitivity 62% [95% CI 55–68]
- specificity 94% [93–94]
- Positive predictive value 11%
- Negative predictive value >99%.

		Outcome		
		+	–	Total
Prediction	+	134	1050	1184
	–	83	15120	15203
	Total	217	16170	16387



Choice of pre-specified risk cut off

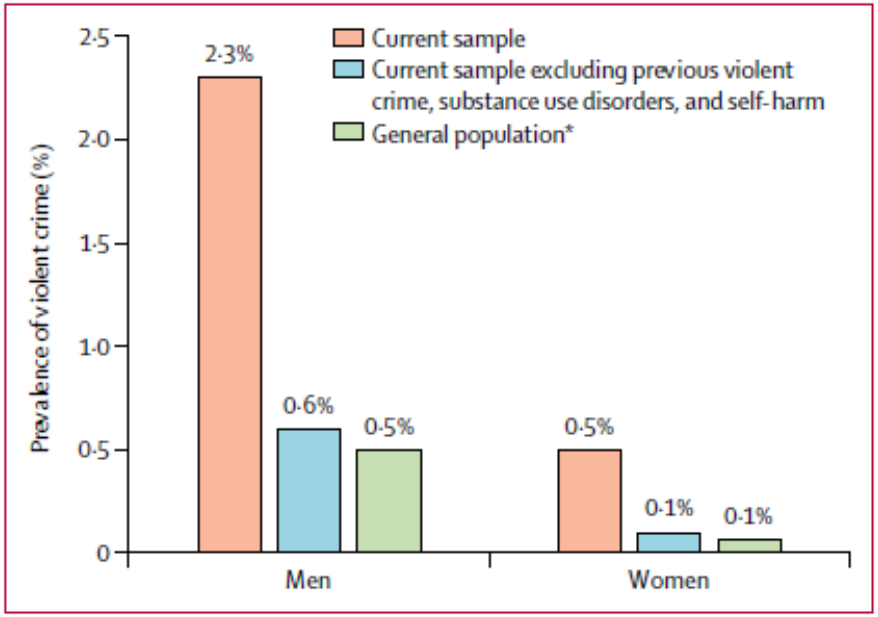


Figure 1: **Violent crime over a 12 month period in different populations by sex**

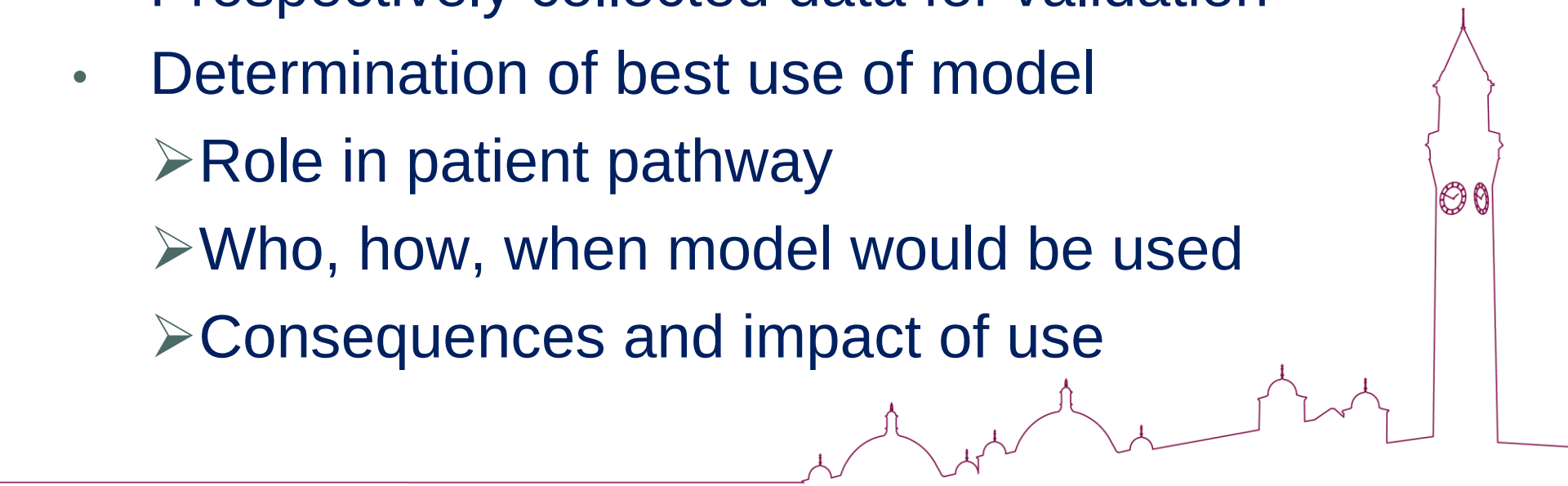
*Data taken from the general population sample in Fazel and colleagues.[1]

- Previous research: incidence of violent crime in schizophrenia spectrum disorders at 1 year of 4%
- Pre-specified 5% cutoff for low-to-high risk of violent offending, as the previous data were based on less severe and younger participants

Overall aim

To determine if model should be used in NHS. Need to conduct studies to evaluate model in NHS setting

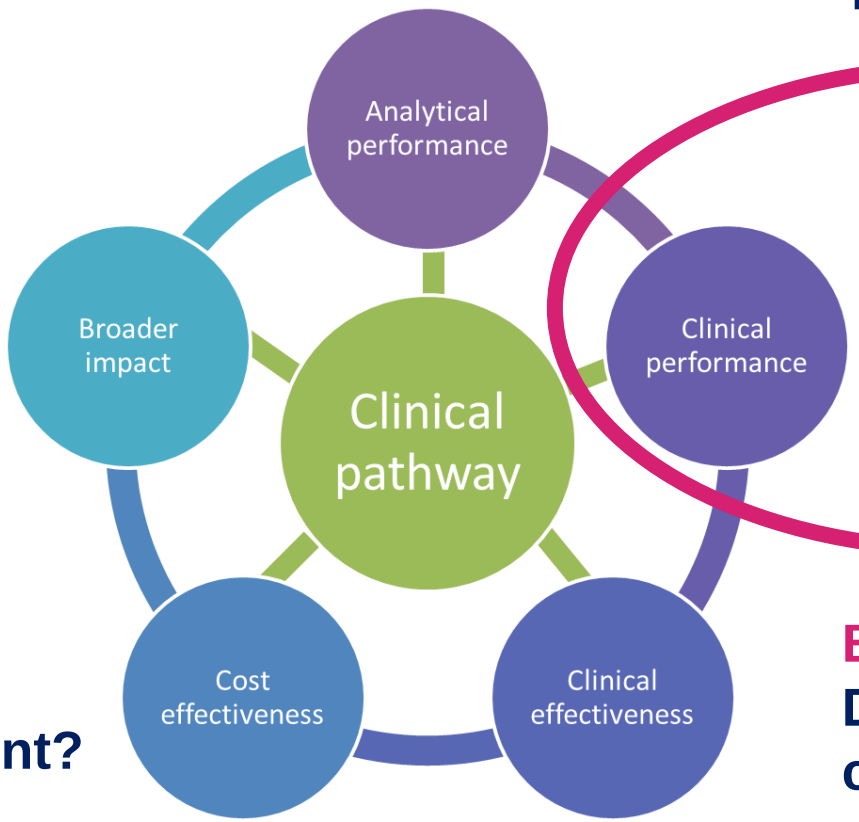
- External validation on NHS data
- Prospectively collected data for validation
- Determination of best use of model
  - Role in patient pathway
  - Who, how, when model would be used
  - Consequences and impact of use



Where does test [prediction tool] accuracy fit?

Is the tool reliable?
Are results repeatable?

Does the tool have social, legal, ethical, societal, etc. consequences?



Accuracy
Does the tool correctly predict outcomes?

Is the tool resource-efficient?

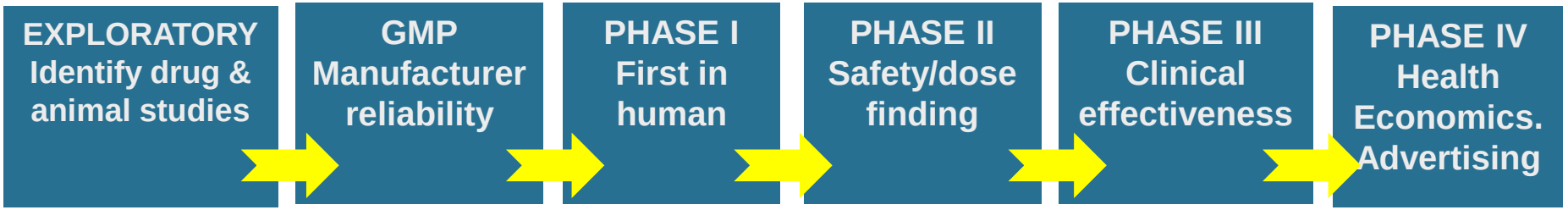
Beyond Accuracy
Does the tool improve clinical outcomes?

UNIVERSITY OF BIRMINGHAM | COLLEGE OF MEDICAL AND DENTAL SCIENCES

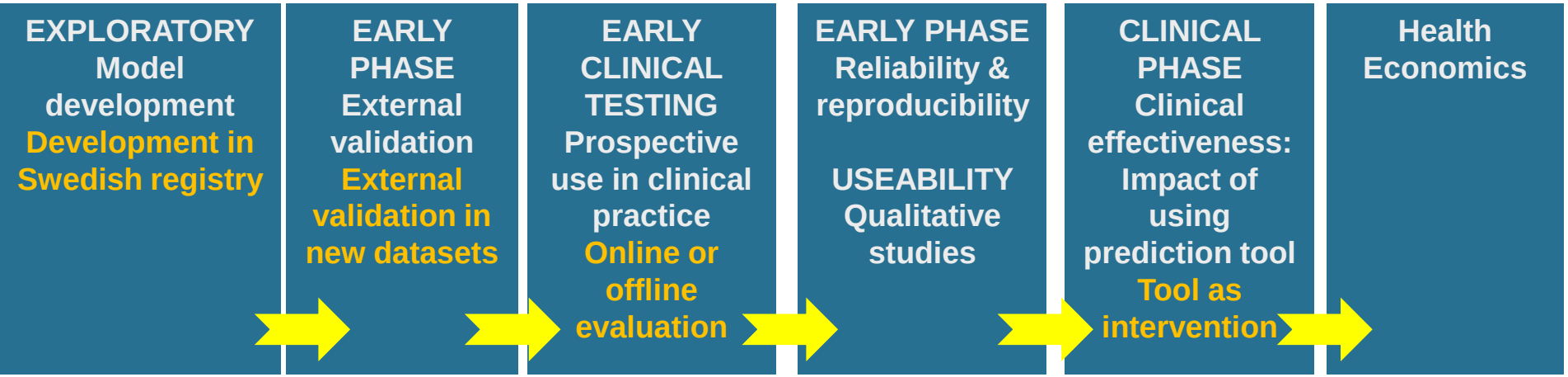
Adapted from similar concepts in Diagnostic Test Accuracy (Horvath AR et al. From biomarkers to medical tests: the changing landscape of test evaluation. Clin Chim Acta. 2014 Jan 1;427:49-57.)

Stages of study design

Intervention studies e.g. developing a new drug



Prediction tool studies



Small group discussions

- Split into small groups according to interest
- Discuss one (or more) scenario from slide handouts

Scenario 1: External validation choice of data source

We want to evaluate model using new participants (external validation) in the NHS

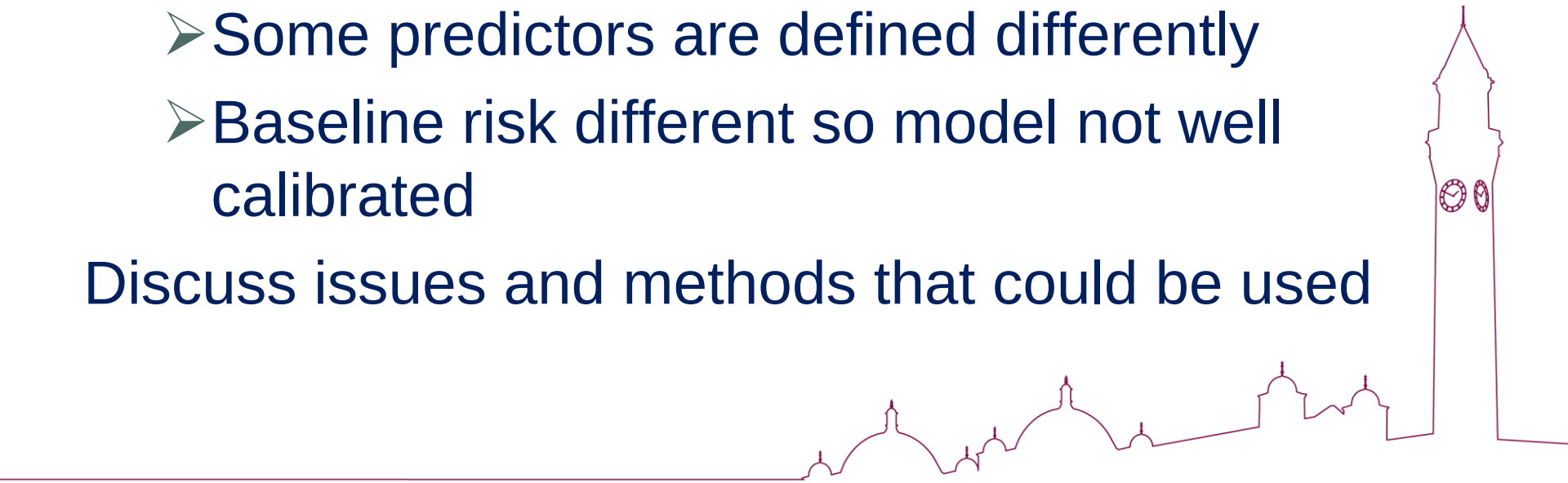
- Discuss different study designs and data sources we could use from NHS patients
- Discuss advantages and disadvantages of different data sources

Scenario 2: External validation and challenges with predictors

We want to evaluate model using new participants (external validation) in a different country setting

- Evaluate in new data collected in NHS however
  - Not all predictors are available
  - Some predictors are defined differently
  - Baseline risk different so model not well calibrated

Discuss issues and methods that could be used



Scenario 3: Is the model as a clinical tool suitable for use in NHS?

We want to evaluate the model as a clinical tool for use in the NHS

- What evidence do we need to know to inform whether the clinical tool should be used in NHS?
- Discuss study designs and methods that could be used

Scenario 4: Is the model as a clinical tool suitable for use in NHS?

We want to know what the **accuracy** of the clinical tool is for determining high risk patients. This would enable targeting extra resources and support to the higher risk population

- Discuss study design and key features of a study to evaluate this

Scenario 5: Is the model as a clinical tool suitable for use in NHS?

We want to know what the impact of the clinical tool will be on patients if this model is used in the NHS

- Discuss issues, study designs and methods that could be used to evaluate this