Regression and Clinical prediction models

Session 2
Steps in planning and conducting CPM research – Part 1

Pedro E A A do Brasil
pedro.brasil@ini.fiocruz.br
2023

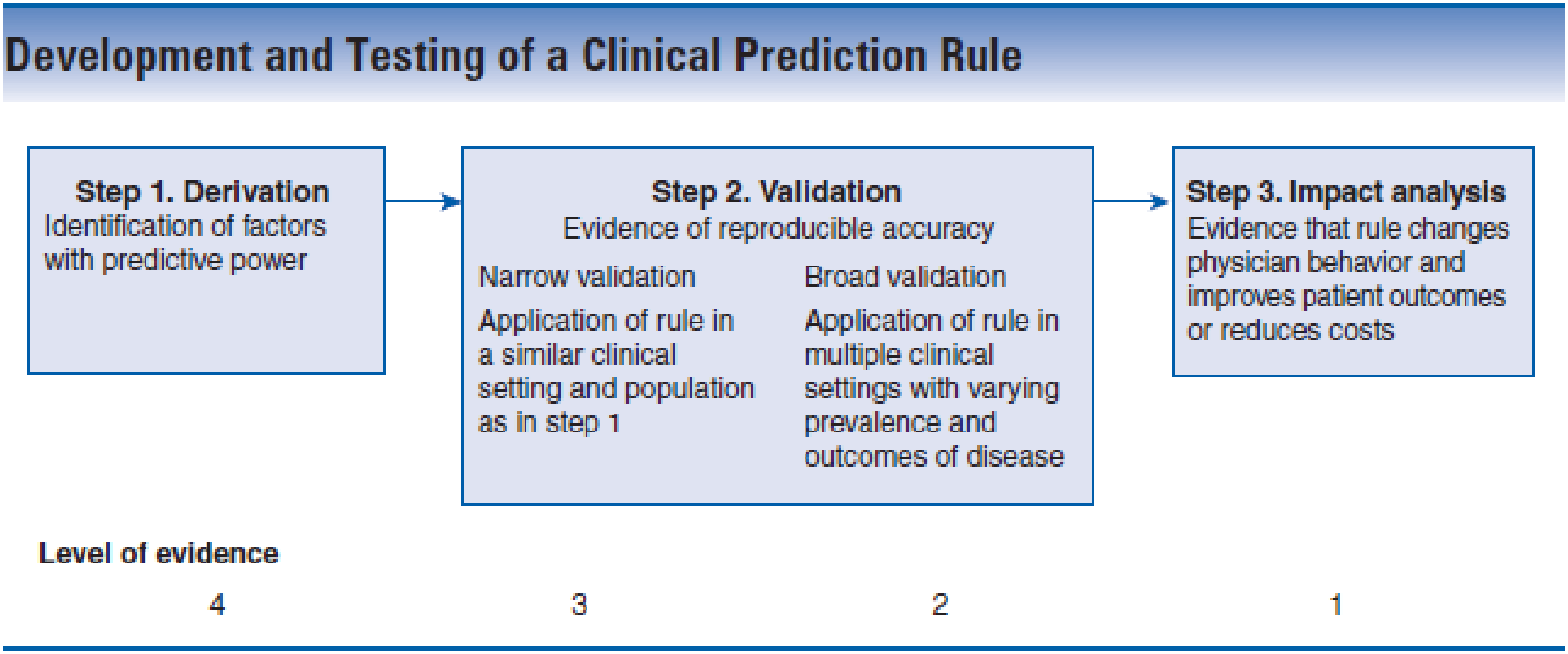
Objectives

- Clinical prediction models have some unique characteristics which make them different from other observational studies.
- In this session, usual steps in planning and conducting CPM research will be introduced and commented.
- One must be well aware of which state of development the research line is, to know what additional evidence is necessary to have a prediction model available.

CPM use

- To use a CPM, one must know what are the strengths and limitations of CPM in general and how the particular CPM may improve the decision making.
- Before planning and conducting a CPM study, besides the familiarity with prediction models and decision making the in topic (e.g. stroke), one must review the recommended and available methods. A good start is editorial recommendations and websites of groups active in the field.
  - http://prognosismethods.cochrane.org/
  - http://www.equator-network.org/reporting-guidelines/tripod-statement/
  - http://www.probast.org/

Derivation & Validation



Guyatt. Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice, 2º ed. 2008.

Derivation & Validation

Users' Guide to Clinical Prediction Rules

Level I

Has the rule undergone at least 1 prospective validation in a population separate from the derivation set, plus 1 impact analysis that demonstrates a change in clinician behavior with beneficial consequences? If yes, clinicians can use the rule in a wide variety of settings with confidence that they can change clinician behavior, facilitate patient decision making, improve patient outcomes, or reduce costs.

Level II

Has the rule shown accuracy either in 1 large prospective multicenter study including a broad spectrum of patients and clinicians or validation in several smaller settings that differ from one another? If so, but if there is no impact analysis, clinicians can use it in various settings with confidence in their accuracy but with no certainty that patient outcomes will improve.

Level III

Has the rule been validated in only 1 narrow prospective sample? If so, clinicians may consider using the CPR with caution and only if patients in the study are similar to those in their clinical setting.

Level IV

Has the rule been derived but not validated or validated only in split samples, large retrospective databases, or through statistical techniques? If so, this is a CPR that needs further validation before it can be applied clinically.

Abbreviation: CPR, clinical prediction rule.

Guyatt. Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice, 2º ed. 2008.

Derivation & Validation

Table 11.2 Methodological standards for Derivation and Validation of a Clinical Prediction Rule[3]

Derivation
1. Were all important predictors included in the derivation process?
2. Were all predictors present in a significant proportion of the study population?
3. Were all the outcome events and predictors clearly defined?
4. Were those assessing the outcome event blinded to the presence of the predictors and those assessing the presence of predictors blinded to the outcome event?
5. Was the sample size adequate (including adequate number of outcome events)?
6. Does the rule make clinical sense?
Validation
1. Were the patients chosen in an unbiased fashion and do they represent a wide spectrum of severity of disease?
2. Was there a blinded assessment of the criterion standard for all patients?
3. Was there an explicit and accurate interpretation of the predictor variables and the actual rule without knowledge of the outcome.
4. Was there 100% follow-up of those enrolled?

Knottnerus & Buntinx .*The Evidence Base of Clinical Diagnosis: Theory and methods of diagnostic research, 2nd Ed* 2009

Usual steps of CPM

- CPM planning
  - Research question
    - Aim: predictors/prediction?
  - Intended application?
    - Clinical practice/research; adjusting for case-mix?
  - Outcome
    - Clinically relevant?
  - Predictors
    - Reliable measurement?
    - Comprehensiveness

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Usual steps of CPM

- Study design
  - Retrospective/prospective?
  - Cohort (data from trials); case–control; cross-sectional
- Statistical model
  - Appropriate for research question and type of outcome?
  - Sample size sufficient for aim?
  - Steps related with data analysis plan will be shown in later sessions.

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Research question

- The research question may vary substantially, and one must make an effort to suit the methods to reach the appropriate answer.
- Types of research questions and applications of prediction models were introduced in previous session.
- There are some comments below on which research questions suits better different research designs.
- At the end, the intended CPM will make clinical sense, and support decision making?

Intended application

- Discussed in previous session.

Outcome

- Outcome clearly defined?
  - "Hard" endpoints are generally preferred (especially mortality) conditions
  - Relevant clinical endpoints are usually those which need a change in the course of action.
  - When cause-specific mortality is considered, a reliable assessment of the cause of death is required.
  - Composite end points have the advantage of increasing the effective sample size and hence the power for statistical analyses (eg: death or rehospitalization)
  - Outcome should be determined with similar rigor as in an etiologic study or randomized clinical trial

Outcome

- Types of outcomes
  - Non-fatal events (e.g. disease recurrence)
  - Patient centered outcomes (e.g. scores on quality of life questionnaires)
  - Indicators of burden of disease (e.g. absence from work, days with mechanical ventilation)
  - Between binary, ordered and continuous outcomes, the latter are preferred from a statistical perspective, since they provide more power in the analysis.
- Outcome was blinded?
  - Information bias

Predictors

- Important predictors included in the process?
  - Strength of association; previously identified
- Predictors had a significant prevalence in the study population?
  - OR = 2 & P = 50% vs OR = 3 & P = 1%
  - Continuous predictors must cover a clinical relevant range
- Predictors clearly defined?
  - Concepts; reproducible; data quality; no missing values; biological variability; “regression dilution bias”

Predictors

- Type of predictors
  - Demographics (e.g. age, sex, race, socio-economic status)
  - Type and severity of disease (e.g. principal diagnosis, presenting characteristics, severity scores)
  - History characteristics (e.g. previous disease episodes, risk factors, history of past exposures)
  - Comorbidity (concomitant diseases)
  - Physical functional status (e.g. Karnofsky score, WHO performance score)
  - Subjective health status and quality of life (psychological, cognitive, psychosocial, functioning)
  - Diagnostic tests or biomarkers (e.g laboratory data; physical examination data)

Predictors

- Characteristics of good predictors
  - Definitions and scorings that are in line with daily practice (pragmatic research)
  - Are quite readily available
  - Not too costly to obtain (cost & burden)
  - Can be measured with reasonable precision

Study design

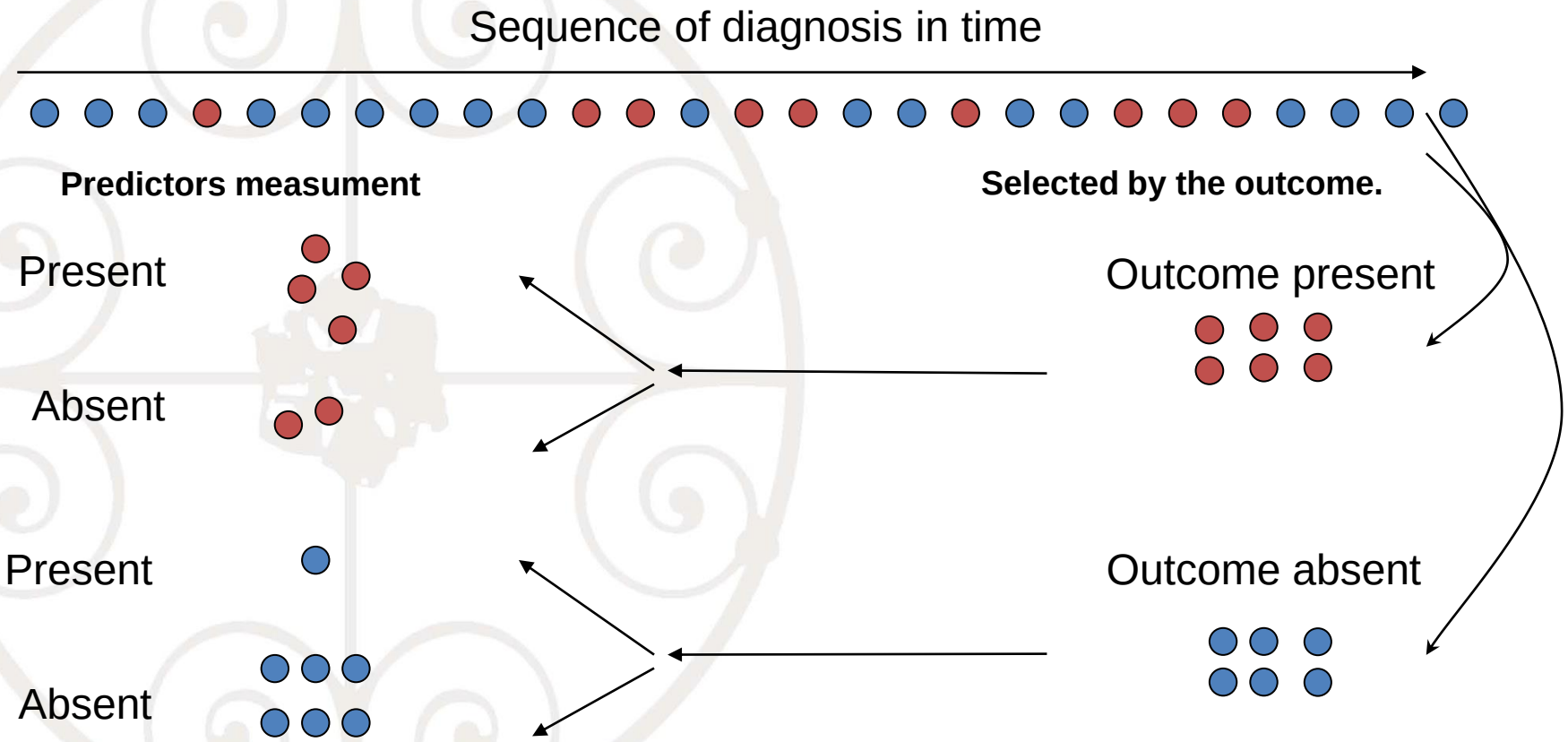
- Study design must suit the research aim
  - Diagnosis
    - What does my patient have?
    - Should I order additional tests for this patient?
    - What are the pre-tests probabilities of this condition for this patient?
    - Cross-sectional; Case-control
  - Prognosis
    - Is my patient going to die/get better from this condition?
    - How long will the patient live with this condition?
    - Follow-up (Cohort); Case-control

Study design

- Study design must suit the research aim
  - Treatment
    - Should I give my patient any treatment?
    - Will my patient improve more with this treatment?
    - Which patients are likely to benefit from this treatment?
    - Will additional treatment be necessary due to side effects?
    - Follow-up (Cohort); RCT

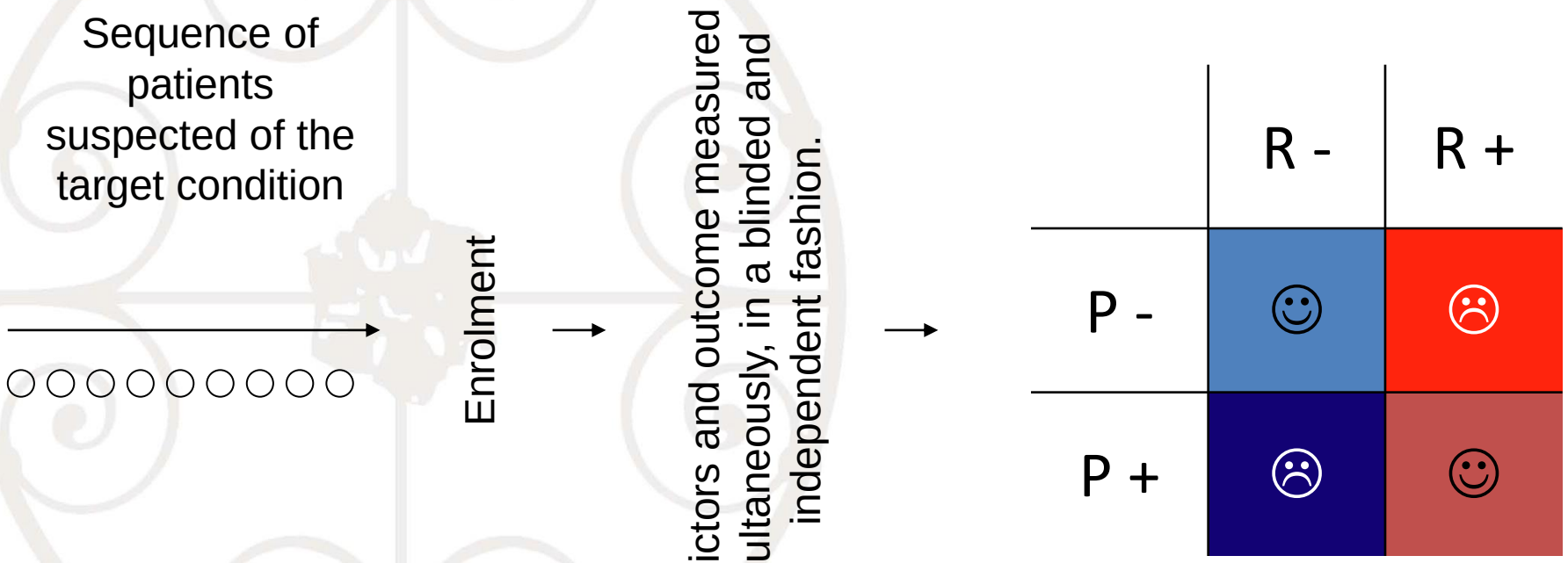
Study design

- Diagnosis – Case-control



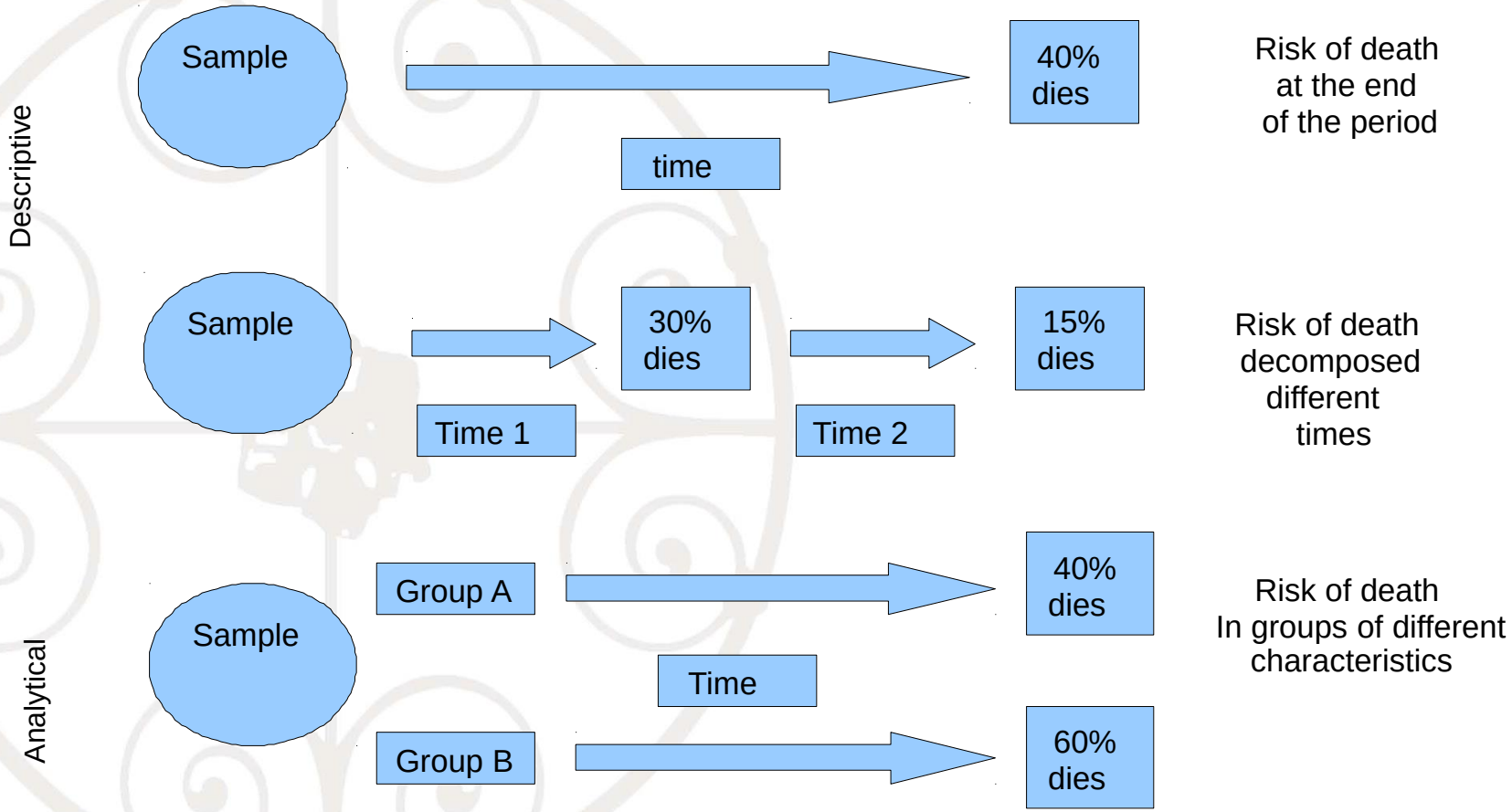
Study design

- Diagnosis – Cross-sectional



Study design

- Follow-up studies (cohort).



Study design

- Follow-up with case-control
  - Every case-control is nested in a population/cohort (real or imaginary)
  - Controls should represent the exposure experience from the population/cohort.
  - The only way to do that is to be sure that controls are selected from the same population/cohort as cases (that is why case-controls are always nested).

Study design

- Follow-up with case-control
  - One serious problem of case-controls is the inherent definition of time which may mass up the interpretation of the risk of the outcome.
  - Look 4 slides above
    - Eg: Update prediction model with an expensive biomarker
      - Collect blood from everyone and follow them for **a year**
      - Select cases and randomly select controls from the remaining cohort with **a year** of follow-up after biomarker is collected.
      - The answer will be the probability of the outcome **at day 365 after biomarker measurements**.
      - The time period must be defined at design phase to allow results inference.

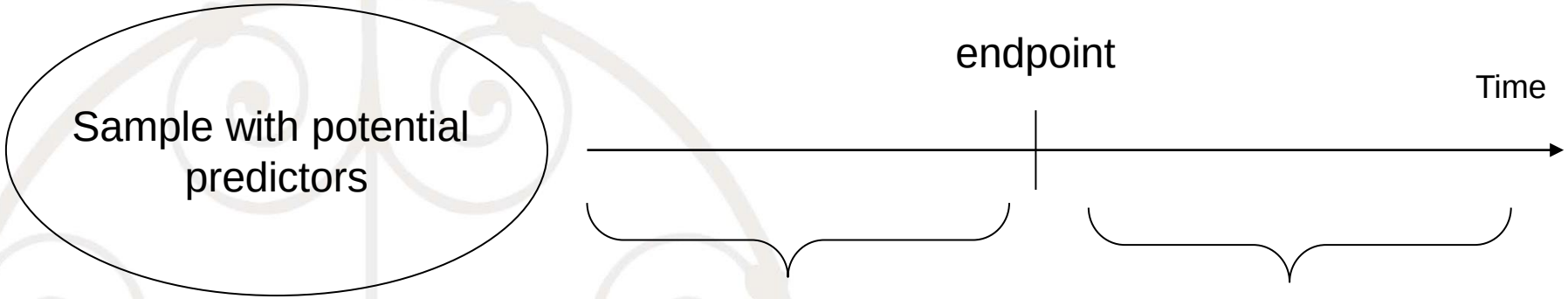
Study design

Table 3.1 Study designs for prognostic studies

Design	Characteristics	Strengths	Limitations
Retrospective	Often single-centre studies	Simple, low costs	Selection of patients Definitions and completeness of predictors Outcome assessment not by protocol
Prospective	Often multicentre RCT	Well-defined selection of patients Prospective recording of predictors Prospective assessment of outcome according to protocol	Poor generalizability because of stringent in- and exclusion criteria
Registry	Complete coverage of an area/participants covered by insurance	Simple, low costs Prospective recording of predictors	Outcome assessment not by protocol
Case–control	Efficient when outcome relatively rare	Simple, low costs	Selection of controls critical Definitions and completeness of predictors Outcome assessment not by protocol

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Study design



The fact that predictors of interest were registered somewhere before research were planed, and only later, after research began, they were established is not related with the study classification according with the time, rather the outcome is important for this classification.

Study design

- Retrospective studies
  - Strengths: simplicity and feasibility; relatively low costs
  - Limitations: identification of patients has to be done in retrospect; missing information; incorrectly recorded information; reliability of information (predictors and outcome)

Study design

- Prospective studies
  - Strengths: better check specific inclusion and exclusion criteria; clear and consistent definitions of predictors and outcome; assessment of patient outcomes at pre-defined time points.
  - Limitations: more complex and costs more (when compared to retrospective); if from (randomized) trials, stringent selection of patients may limit the generalizability of a model (specially from single center RCT).

fim

Session 2
Steps in planning and conducting CPM research – Part 1

Pedro E A A do Brasil
pedro.brasil@ini.fiocruz.br
2023