FORUM PBC
Session II: Scientific Validity and Statistical Challenges

When the Going Gets Tough:
Emulating a Control Arm in PBC
When a Randomized Trial is no Longer Feasible

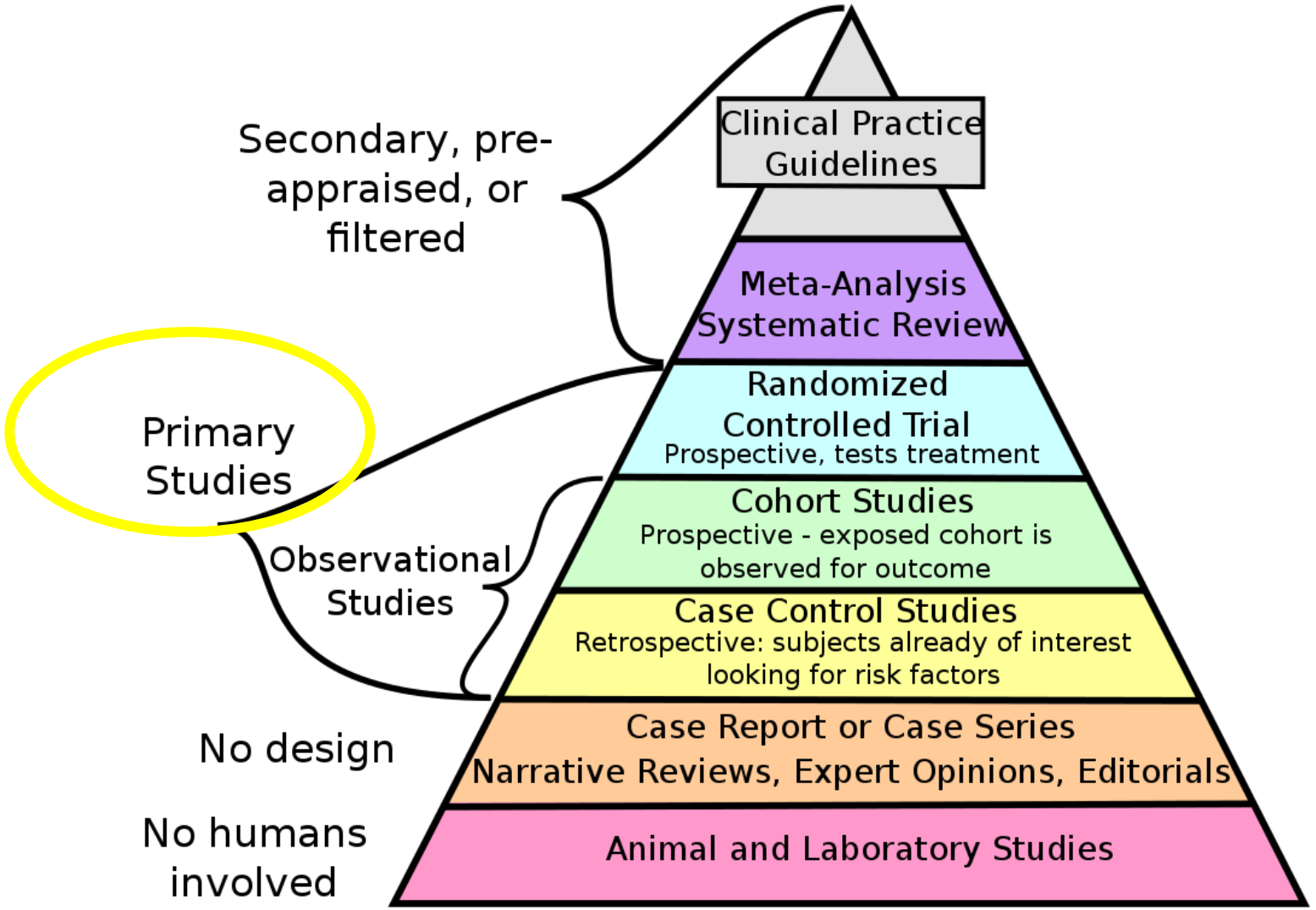
Edinburgh April 7, 2025

Bettina E Hansen

Erasmus MC, Rotterdam, the Netherlands

IHPME, University of Toronto, Canada

Hierarchy of Scientific Evidence



❖https://commons.wikimedia.org/wiki/File: Research design and evidence.svg#/media/File:Research design and evidence.svg

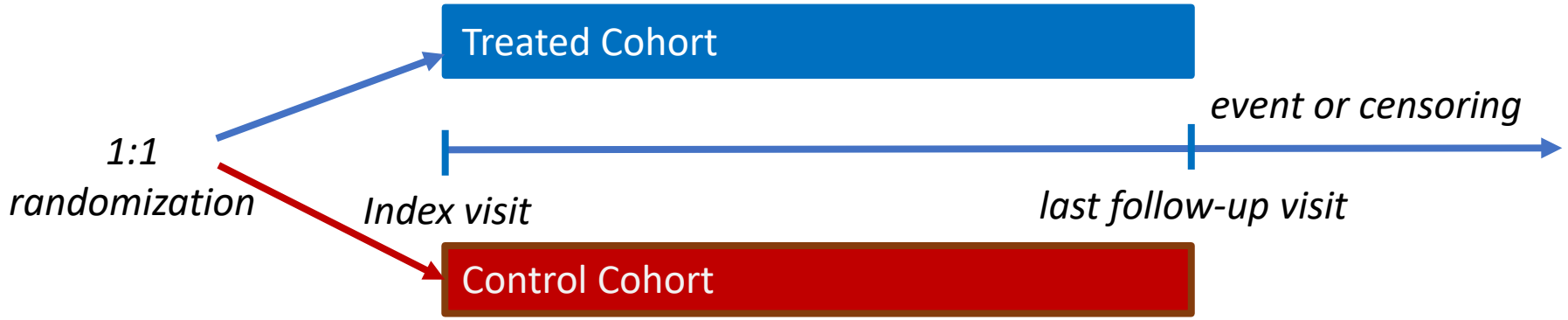
The ideal phase 3 confirmatory study in PBC

Ideal

- Clinical hard endpoint
- Inclusion insufficient response to UDCA population
- Follow-up ?

In reality 2016 (before any 2nd line Rx)

- Clinical hard endpoint
- Inclusion advanced population
- Follow-up ~ 7-8 years



In 2016 the first 2nd line Rx OCA gets accelerated approval

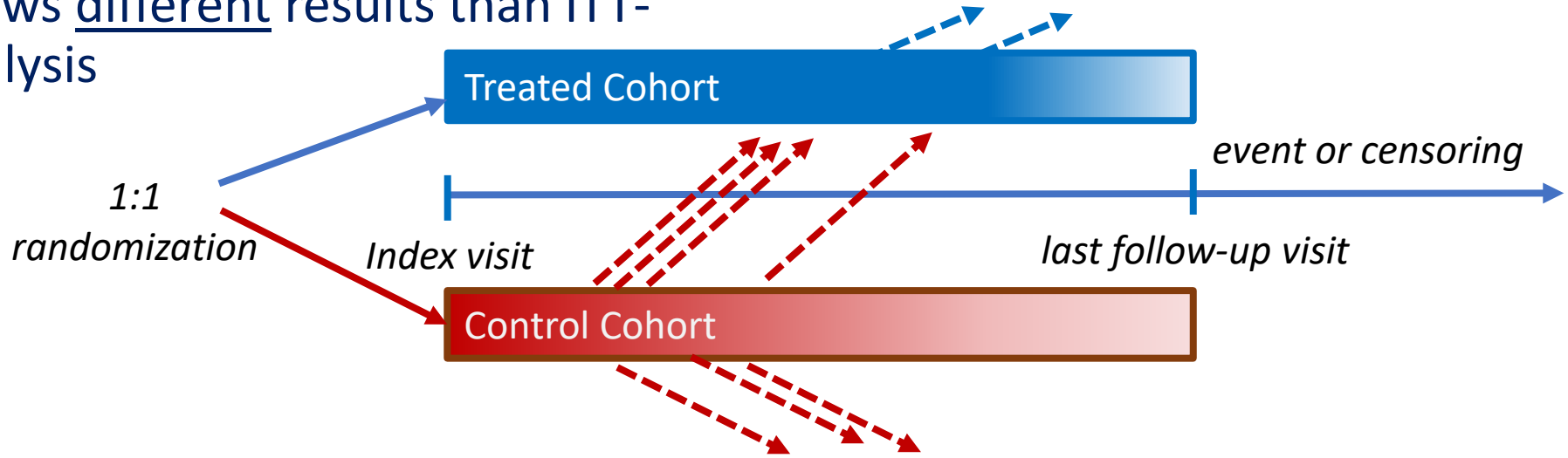
When the Going Gets Tough

What went wrong

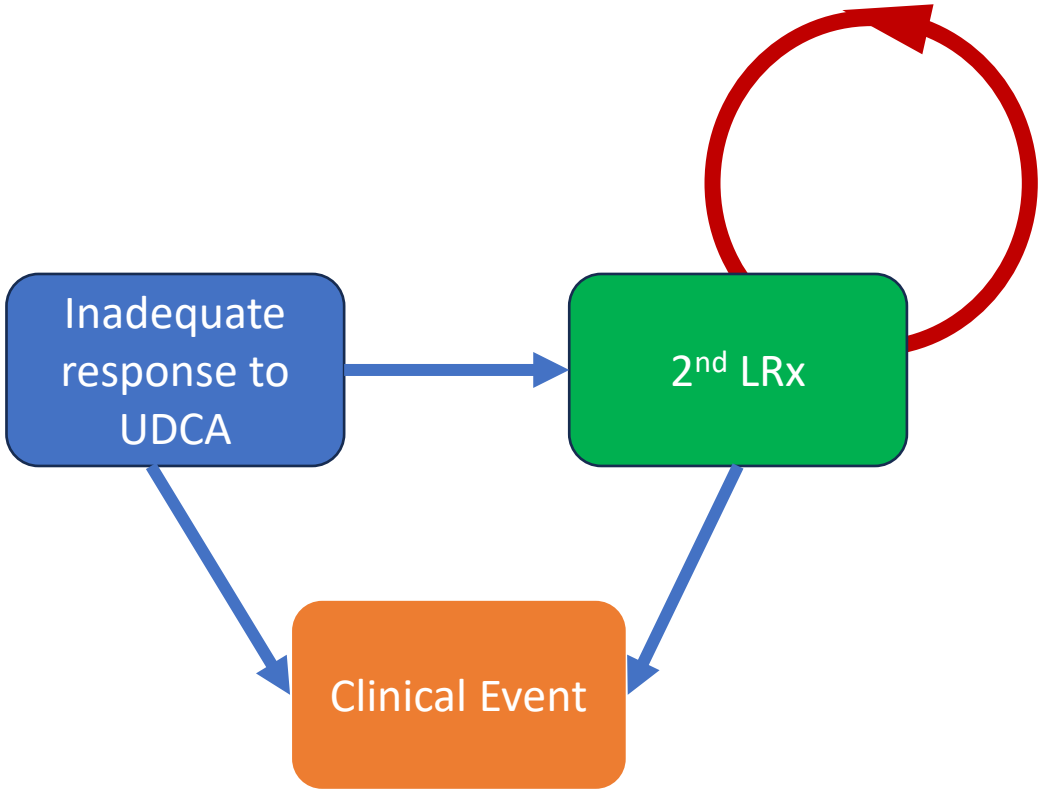
- Patients left the confirmatory study in an unbalanced way
- Trial is no longer balanced or powered to show effects
- Statistical repair of patients switching shows different results than ITT-analysis

In 2025 (multiple 2nd line Rx options)

- Confirmatory study faces same retention problems and inclusion has become even more challenging
- And … *(next slide)*



Added complication in the control cohort *(and treated arm)*



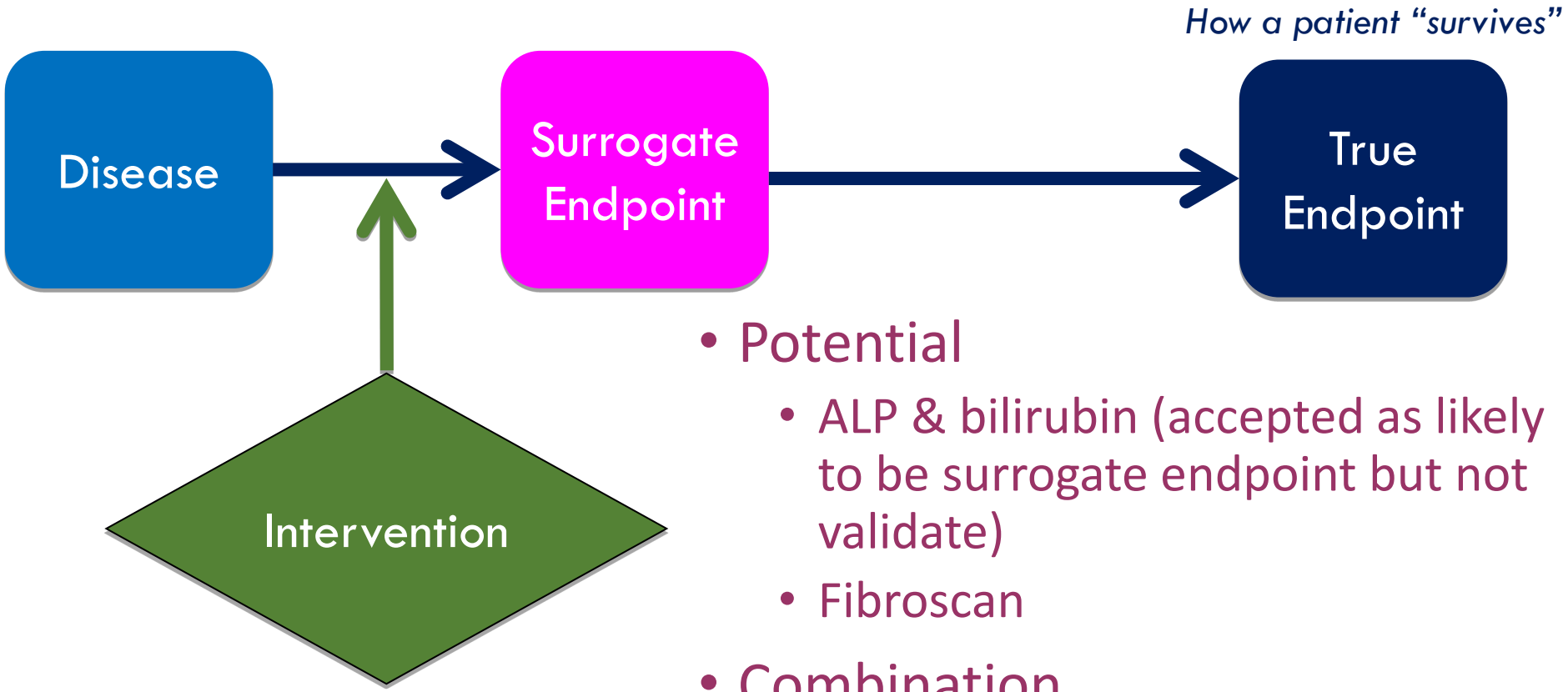
Intercurrent events

- In case of inadequate response to UDCA subjects are *after 2016* eligible for 2nd line treatment (2nd LRx)
- In case of continued inadequate response another 2nd LRx/dose may be offered

The tough get going

- Here and now
- May 6 in Amsterdam

Option 1: A validated surrogate endpoint

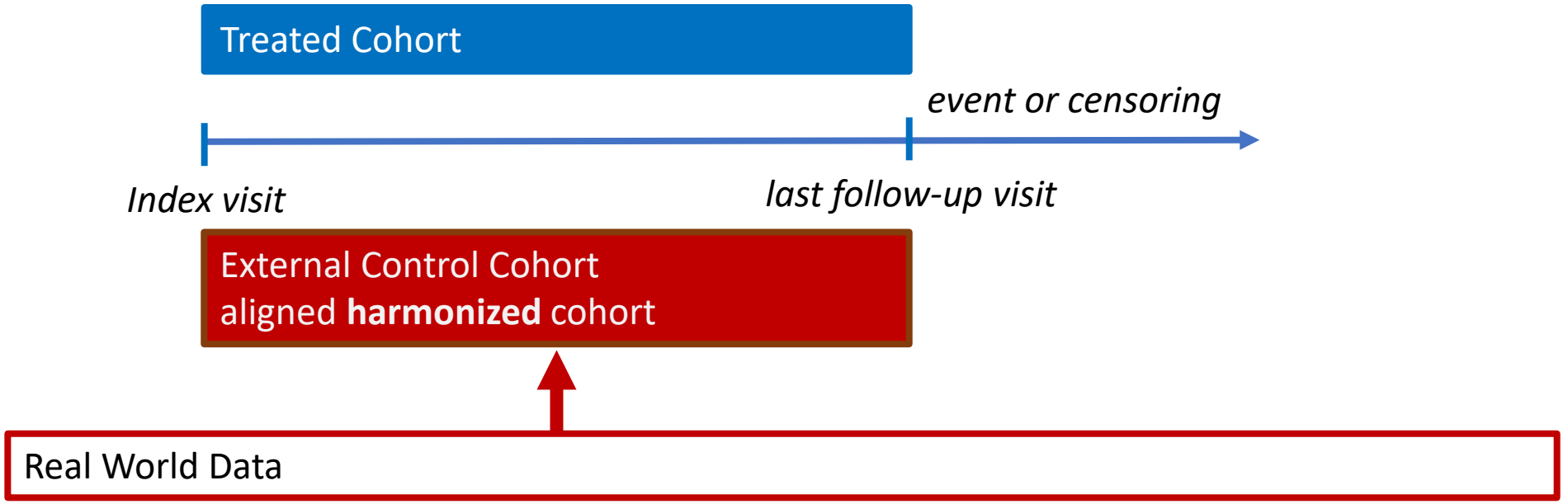


- Potential
  - ALP & bilirubin (accepted as likely to be surrogate endpoint but not validate)
  - Fibroscan
- Combination
- Proof and validation needed

Option 2: use RWD to create RWE

Aim

To compare time to clinical event in treated patients from an open label study(treated only cohort) with external controls



Option 2: use RWD to create RWE

Different phases can be applied

- Phase 2 studies with open label extension
- To replace the confirmatory study:
  - Phase 3 study:
  - Duration 1-2 years
  - Endpoint: ALP & bilirubin response (/fibroscan endpoint), with safety
  - Followed with an open label extension (placebo-arm offered active treatment)
  - Continued collection of safety
- Phase 4 study, active treatment

What kind of RWD

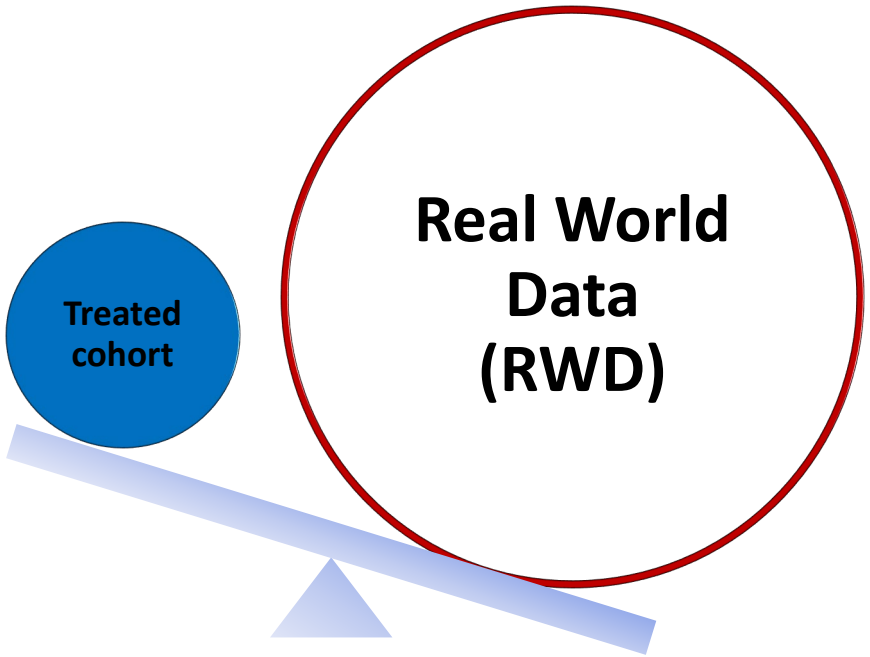
Requirements versus achievable

Keep in mind PBC is a chronic rare disease

- Retrospective versus prospective
- Shared placebo between pharma
- Ownership
  - patients? academia? pharma?
  - identified problem:

  pharma has the commitment to show regulators results, patients and academia do not have these obligations

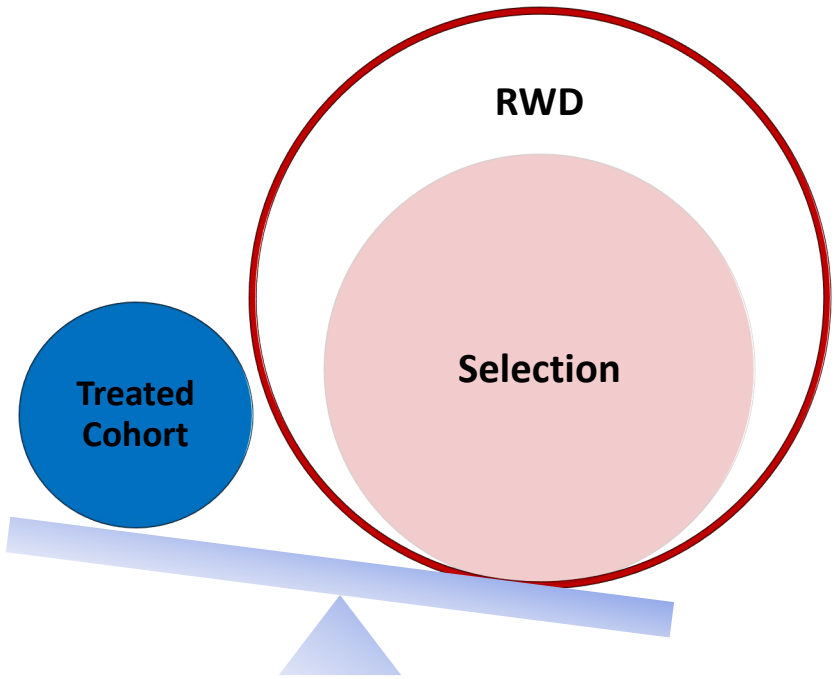
Harmonize Design with the guidelines from EMA (FDA)



Is the RWD Fit for Purpose?

- Quality of RWD
  - storage and governance
  - stability of data collection over time
  - completeness
- Diagnose defined
- Outcome(s) – same definition(s), adjudication
- Coverage - time span and countries
- Lab-values – different labs, ULN, unit
- Patient and disease characteristics
- Treatment
- Confounders
- Power analysis

Identification of patients & visits

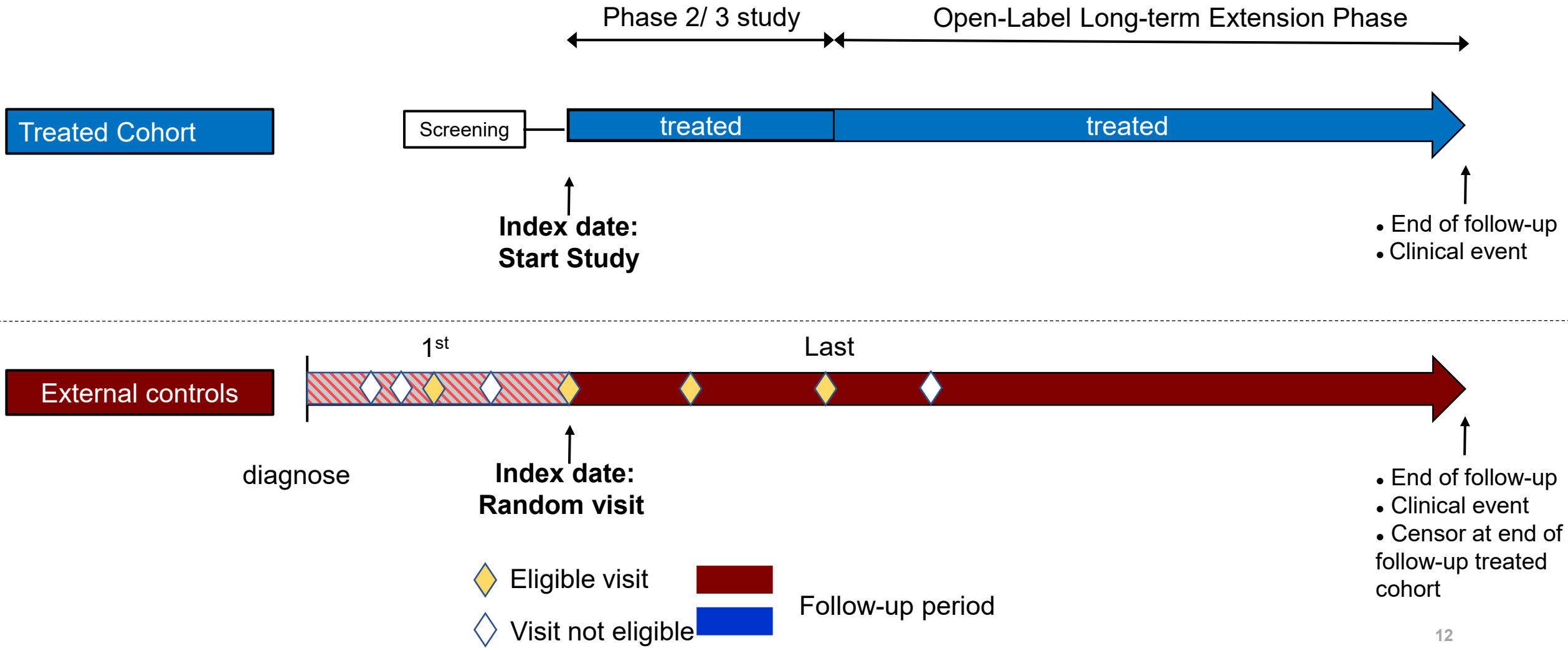


Selection process

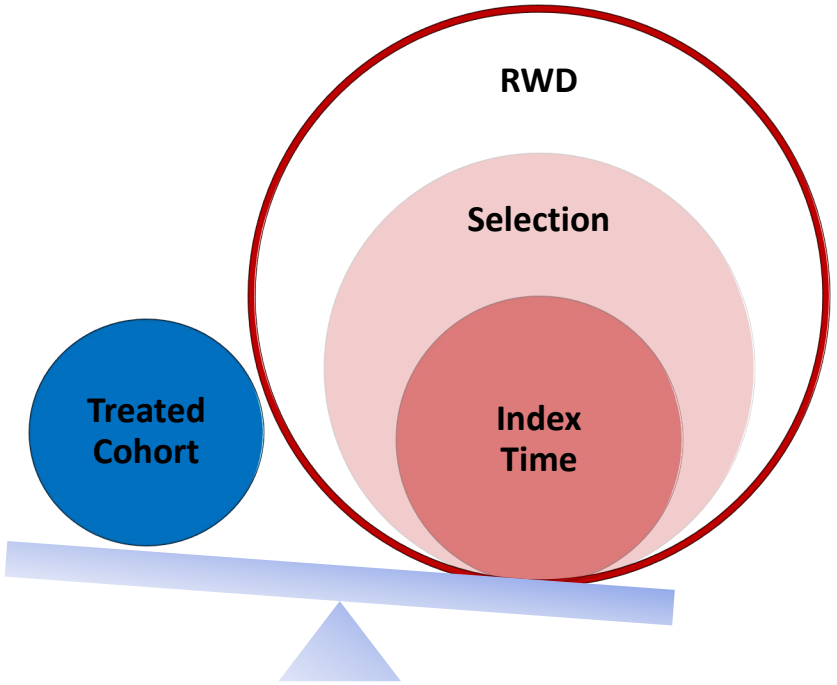
- Apply aligned inclusion/exclusion criteria
- Overlay sites / regions
- Overlay calendar time / SOC treatment

RWD = Real-World Data

Selection of index date (example open label extension)



Index visit

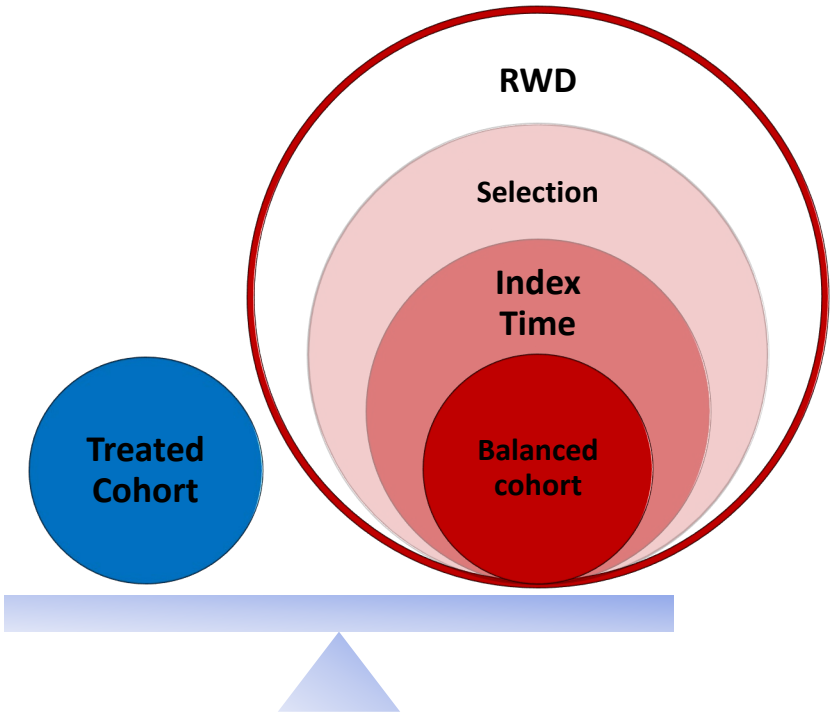


Choice of Index Time = start of follow-up

- First visit
- Random visit(s)
- Last visit
- Other methods:
  - Confirmatory visit
  - Multiple visits
  - ML-method

RWD = Real-World Data

Balanced design using weights



Assessment of balance

- pre-specified check and tests
- Estimate weights
  - Propensity scores
  - IPTW
  - ATT weights

RWD = Real-World Data

Harmonize Design while blinded for outcome to mimic RCT

Firewall: blinded for outcome

Fit for purpose

- Define outcome, confounders
- Quality of Lab-values, patient and disease factors, missingness
- Power analysis

Selection

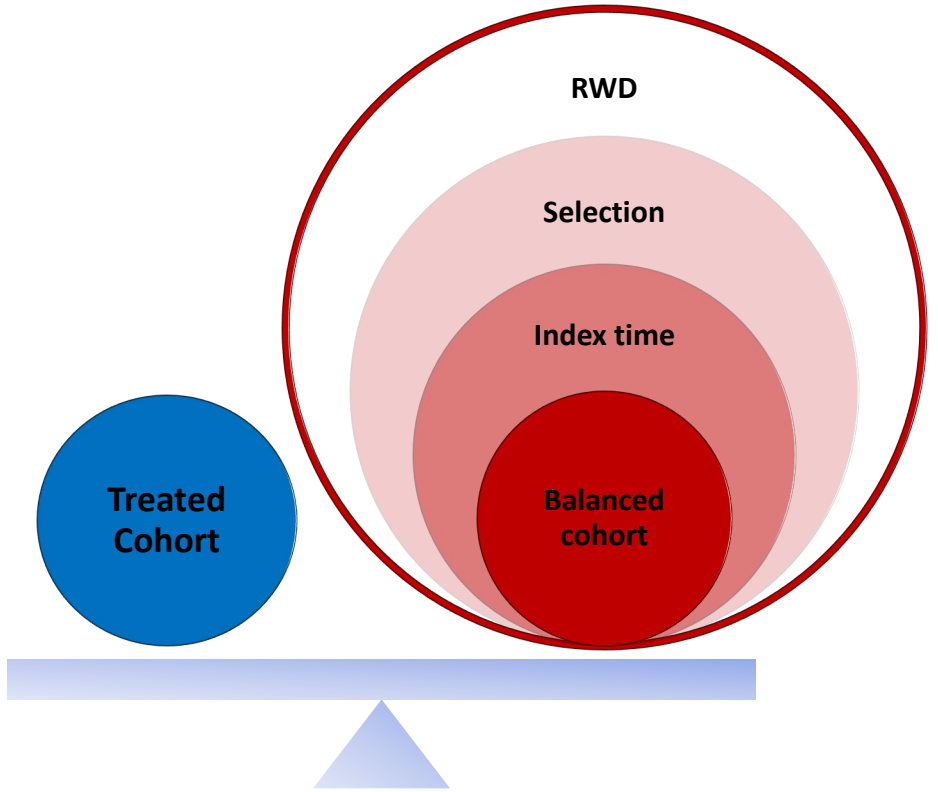
- Apply aligned inclusion/exclusion criteria
- Overlay sites / regions / calendar time

Index Time

- First visit, confirmatory visit, random visit(s), last visit, other methods

Assessment of balance

- pre-specified check and test
- weights: propensity scores, IPTW, ATT, …



RWD = Real-World Data

Analysis of time to event, unblinded for outcome

Firewall off
un-blinded for outcome

Rx arm

- Check for informative censoring

Composite endpoint

- Characterize type of events over time in both Rx arm and RWD-selection (especially for composite endpoints)

Analysis of endpoint

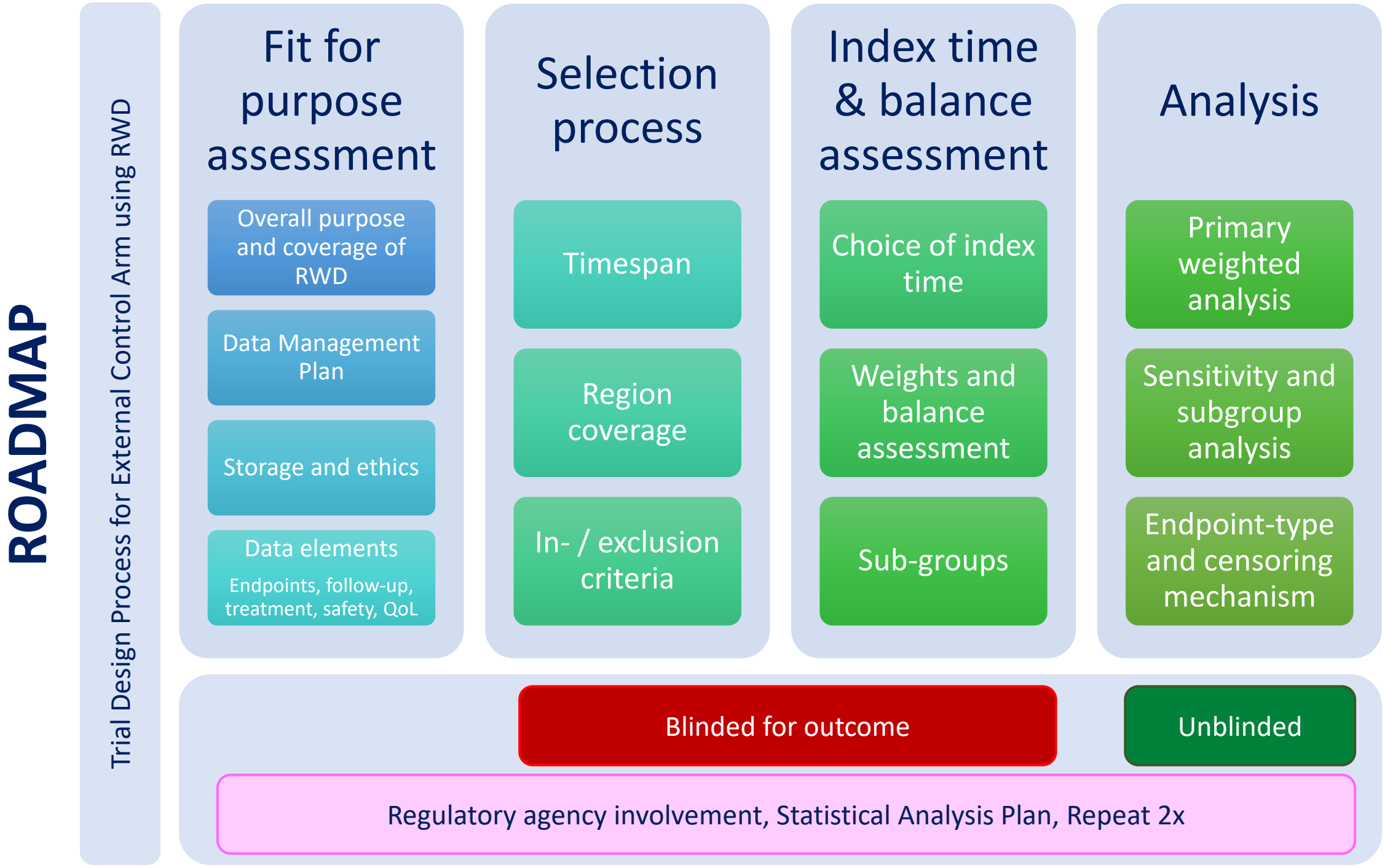
- Kaplan-Meier and Cox regression methods
- Primary: Weighted
- Secondary: Crude effect and adjusted for confounders

Sensitivity analyses

- Range of selection of index time
- Pruning of time to avoid immortal time bias

Subgroup analysis

- Concurrent calendar time
- Same region/ sites/or different sites


ROADMAP
Trial Design Process for External Control Arm using RWD
Fit for purpose assessment
Overall purpose and coverage of RWD
Data Management Plan
Storage and ethics
Data elements
Endpoints, follow-up, treatment, safety, QoL
Selection process
Timespan
Region coverage
In- / exclusion criteria
Index time & balance assessment
Choice of index time
Weights and balance assessment
Sub-groups
Analysis
Primary weighted analysis
Sensitivity and subgroup analysis
Endpoint-type and censoring mechanism
Blinded for outcome
Unblinded
Regulatory agency involvement, Statistical Analysis Plan, Repeat 2x

RWD as external control arm

S Trial: all participants gets treated, enthusiasm for inclusion, less dropout, reduction of time to proof of benefit, harm

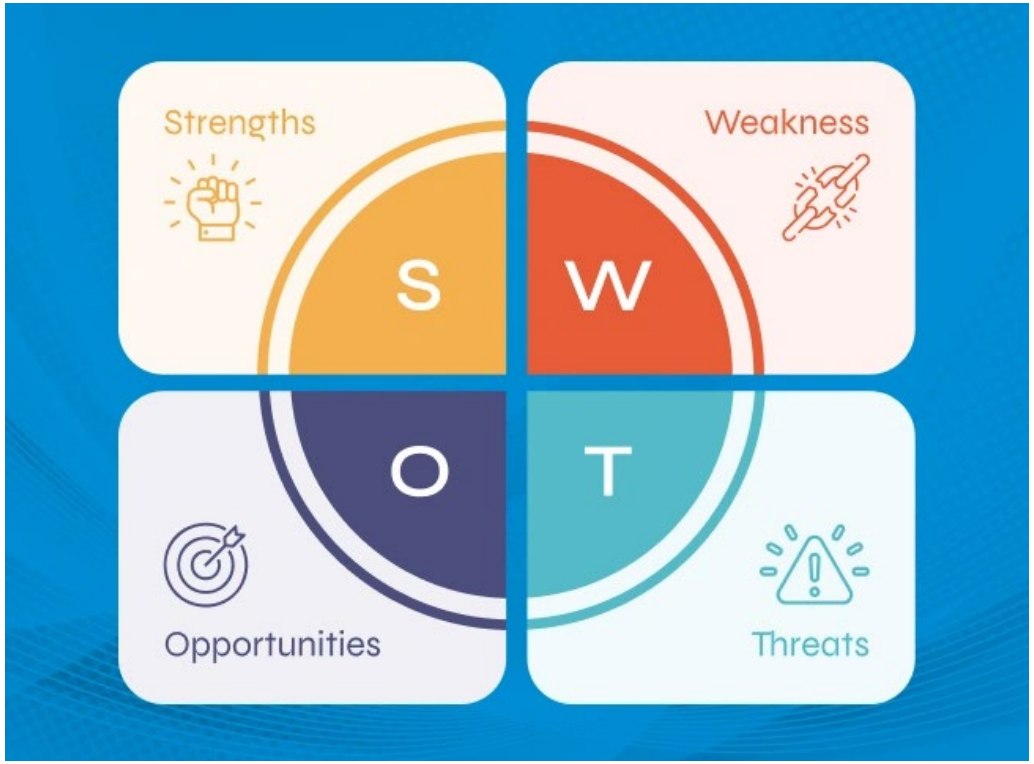
external control: all participates

W RWD different labs,

soft safety data lacking

O Collaboration

T Approval of method



Discussion

- Option 1: surrogate endpoint validation
- Option 2: use of RWE

- Extra’s

Thank you