Correction methods for unmeasured confounding in non-interventional studies

Luca Grassano
Supervisor: Aurélie Bardet

Introduction

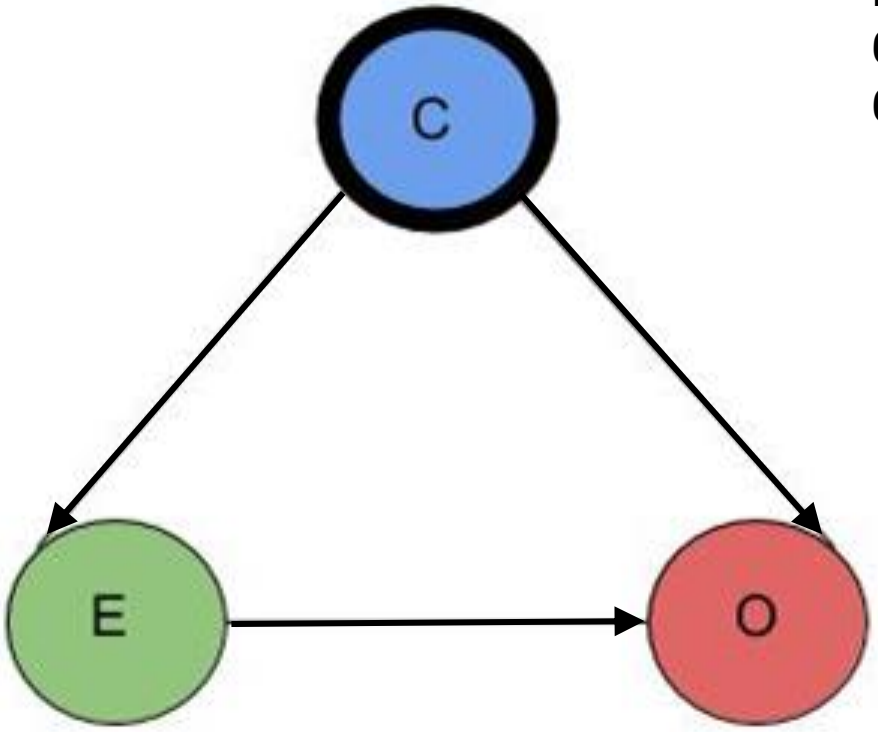
Propensity Score Calibration

Monte Carlo simulation

Conclusions

Confounding

E = Exposure
O = Outcome
C = Confounder



E = Treatment admin
O = Death
C = Age

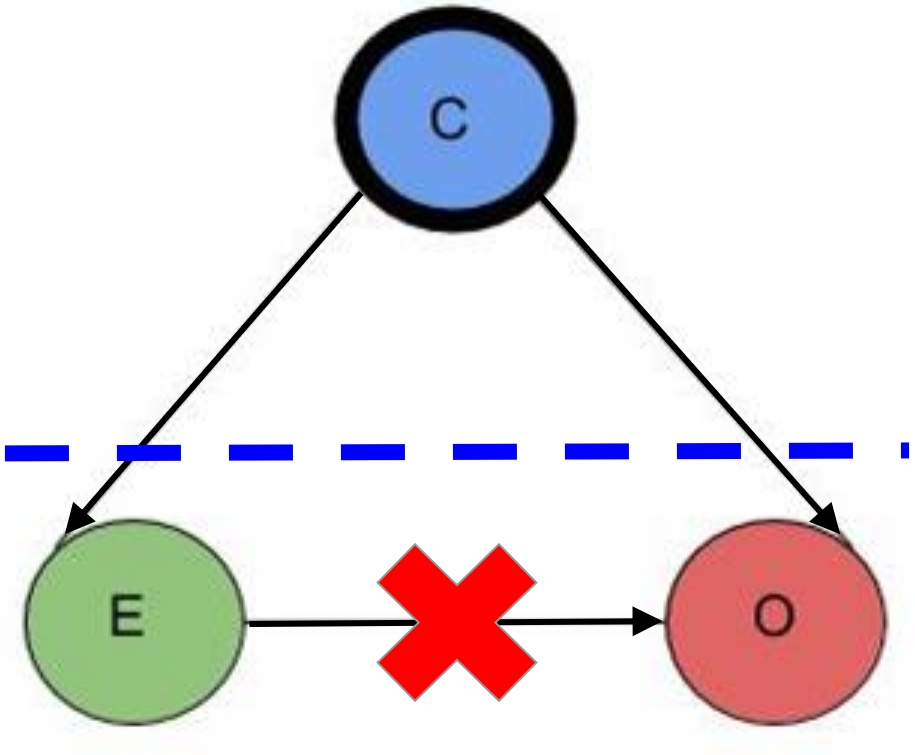
Confounding definitions

"Confounding may be considered a confusion of effects. […] A variable must be associated with *both the exposure under study and the disease* under study to be a **confounder**. [...] A confounding factor must be associated with disease occurrence *apart* from its association with exposure."

Rothman, K.J., Greenland, S. and Lash T.L. (2008) *Modern Epidemiology*, Lippincott Williams & Wilkins

Confounding

E = Exposure
O = Outcome
C = Confounder



When the confounder is **not** taken into **account** in the analysis, the estimation of the treatment effect is affected by **bias**.

Confounding in NIS

RANDOMIZED CLINICAL TRIALS

Randomization of treatment assignment allows to avoid systematic associations between exposure and patient characteristics, hence confounding.

The estimator of treatment effect is considered UNBIASED (even if…)

NON-INTERVENTIONAL STUDIES

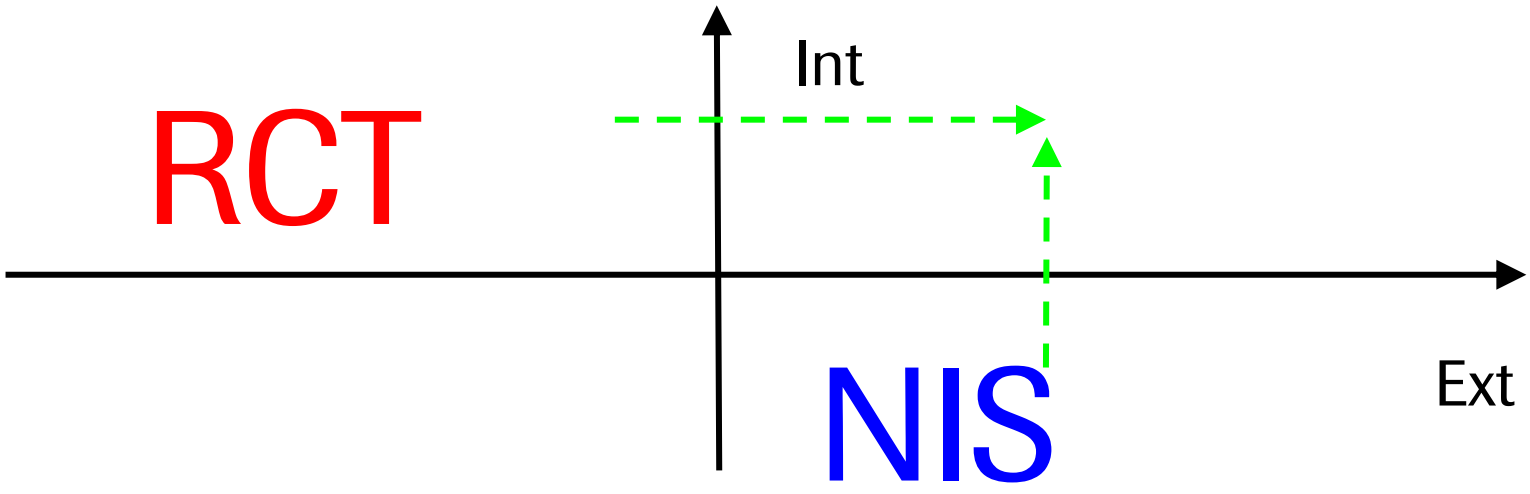
It is often the case that patient characteristics or other factors are **determinants** of treatment assignment and also of the outcome of interest (e.g. confounding by indication).

The estimator of treatment effect is considered BIASED.

External vs Internal Validity

Internal : the validity of the inferences drawn as they pertain to the members of the source population of the study. → **Reliable estimates**

External : measure of how much the previous results are **generalizable** to a target population or to a more general set of circumstances.



RCT = Randomized Clinical Trials NIS= Non-Interventional Studies

Unmeasured confounding

Bias can be avoided accounting for the confounder in the analysis, hence removing the effect of **measured confounding**.

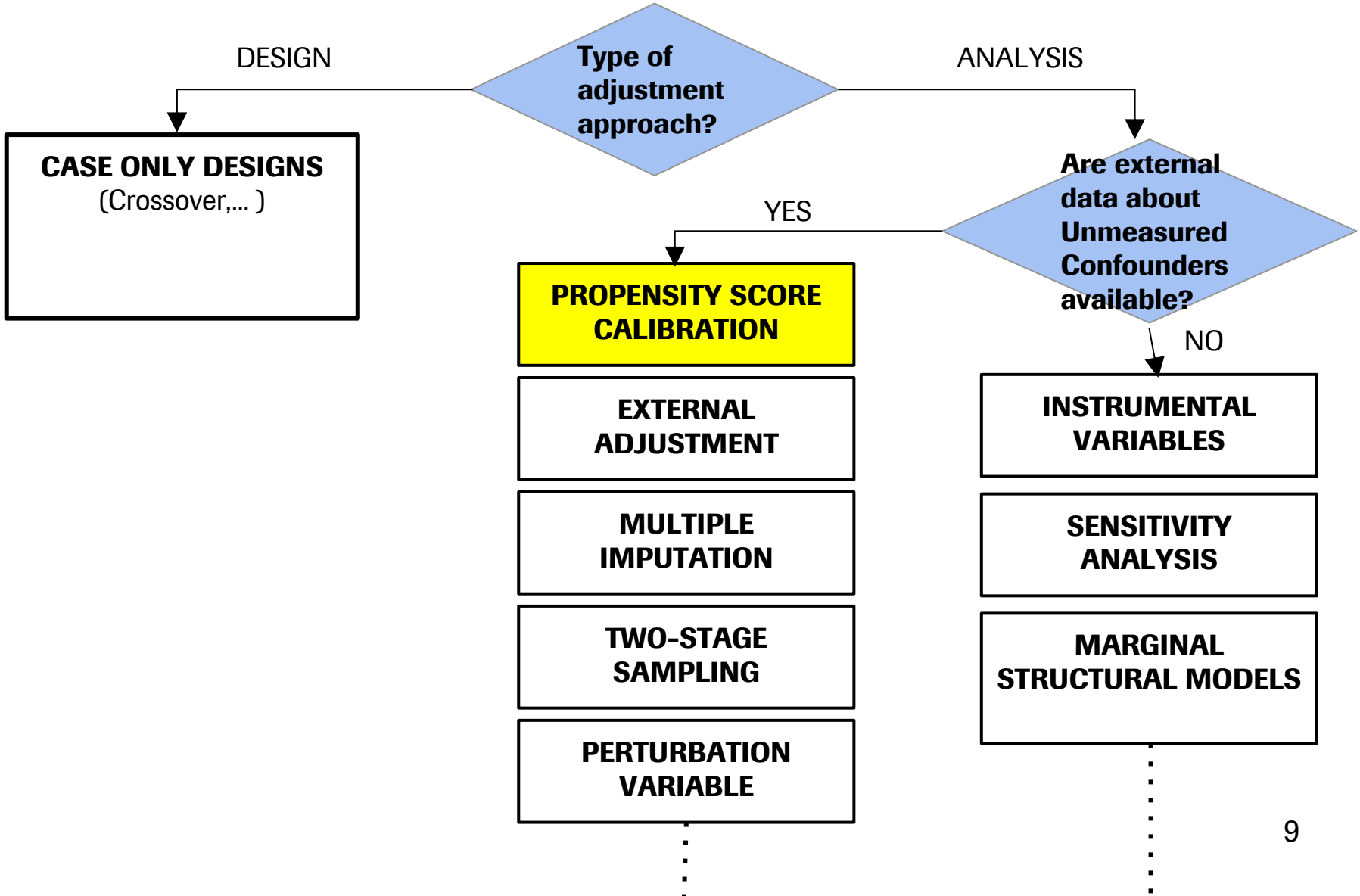
In several contexts this is not possible:

- Measurement errors
- Unfeasible
- Missing
- …

Unmeasured confounding

Both measured and unmeasured confounding can be present in a study.

Unmeasured confounding: methods



Introduction

Propensity Score Calibration

Monte Carlo simulation

Conclusions

Propensity Score

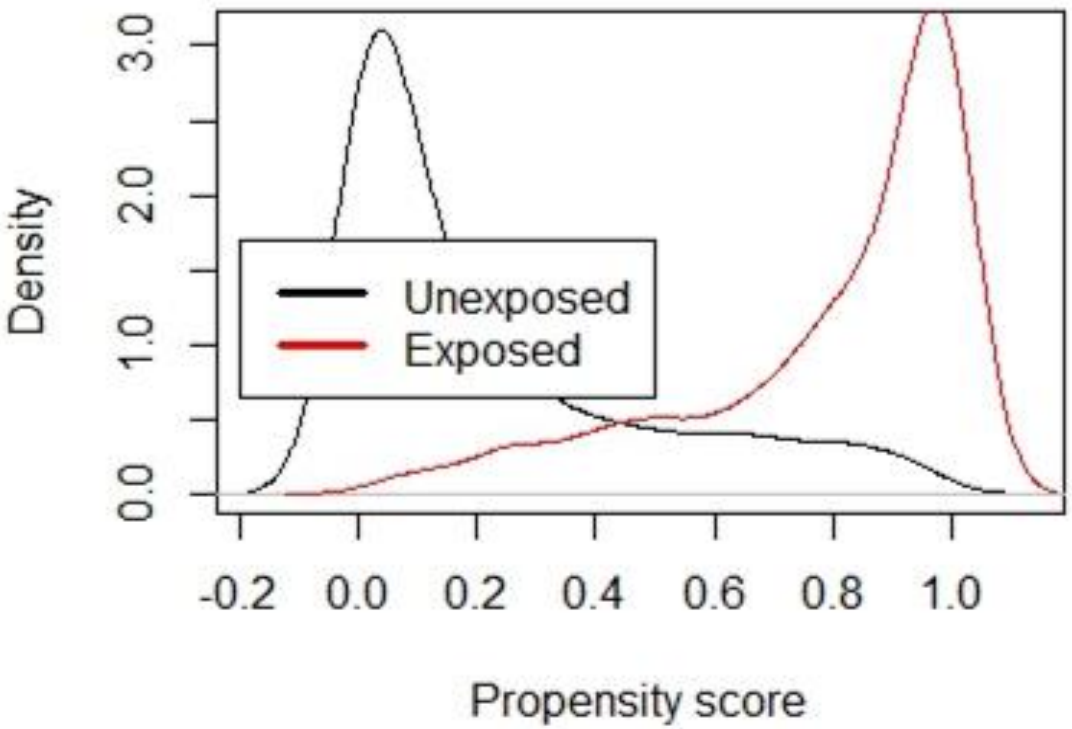
The method aims at **summarizing** covariates through a unique variable (the PS) such that treatment and control groups are **comparable** according to the distribution of covariates within the same PS level and any effect can be attributed to treatment.

$$\mathbf{PS = e(x) = P(treatment \mid X=x)}$$

where X is the vector of covariates.

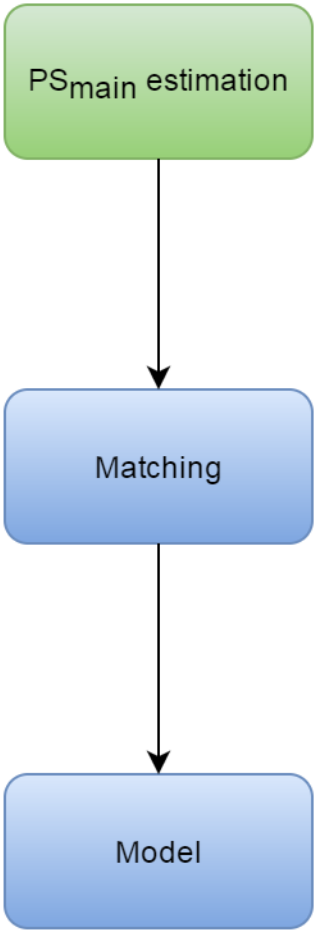
This can be achieved by **logistic regression**.

Propensity Score

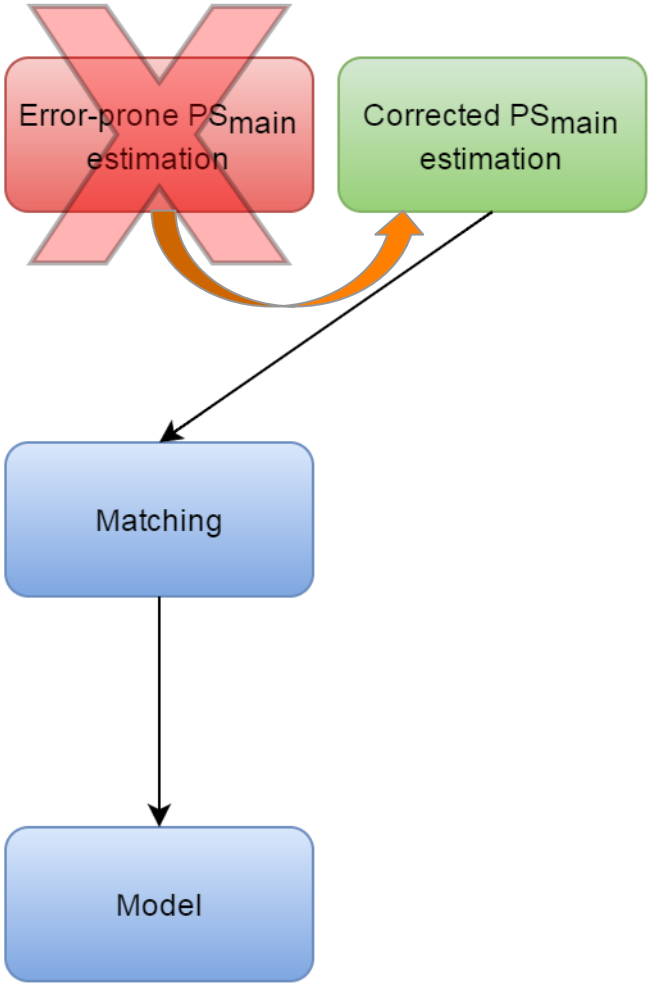


Measured vs Unmeasured Confounding using Propensity Score (PS)

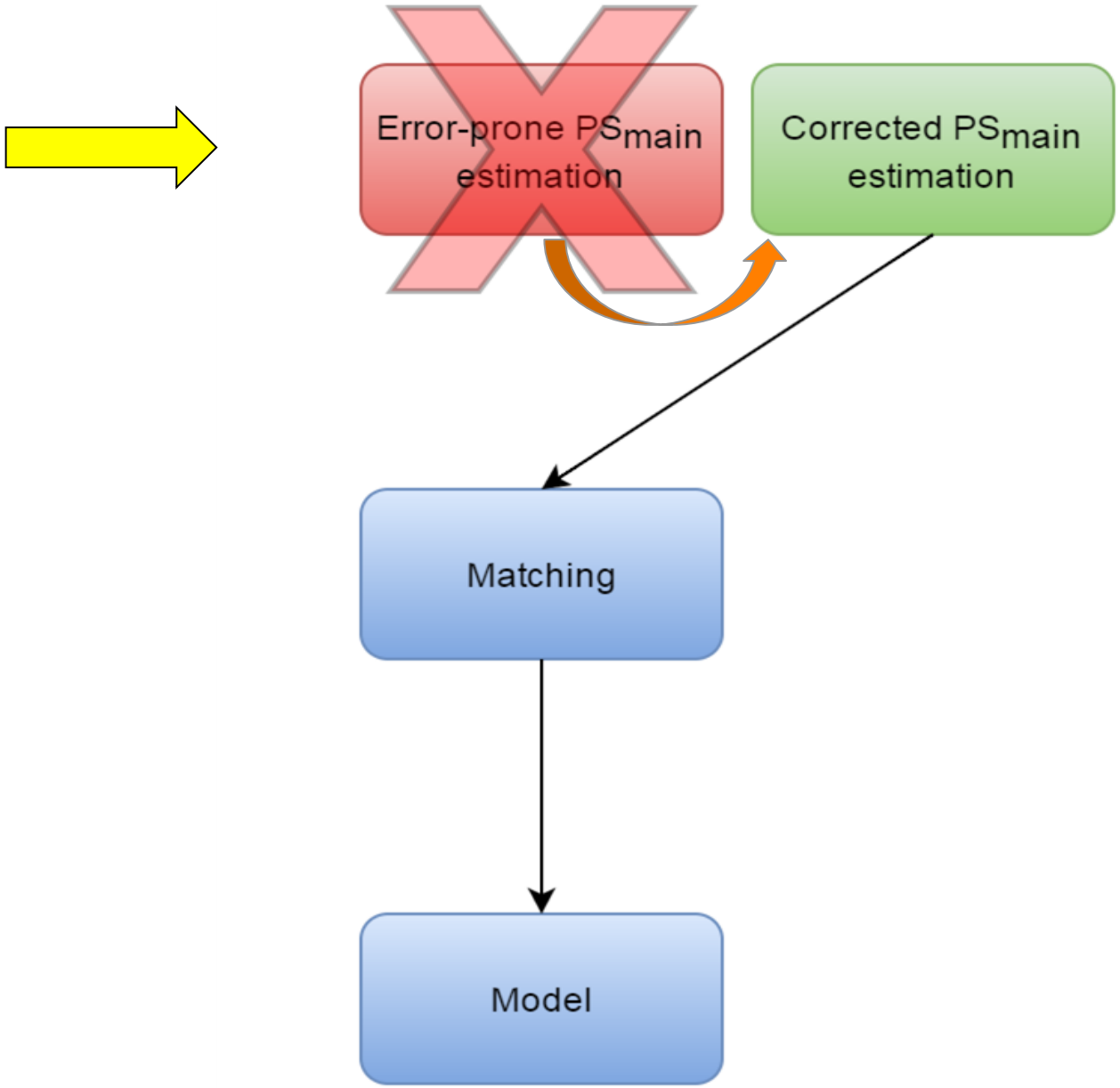
Measured



Unmeasured

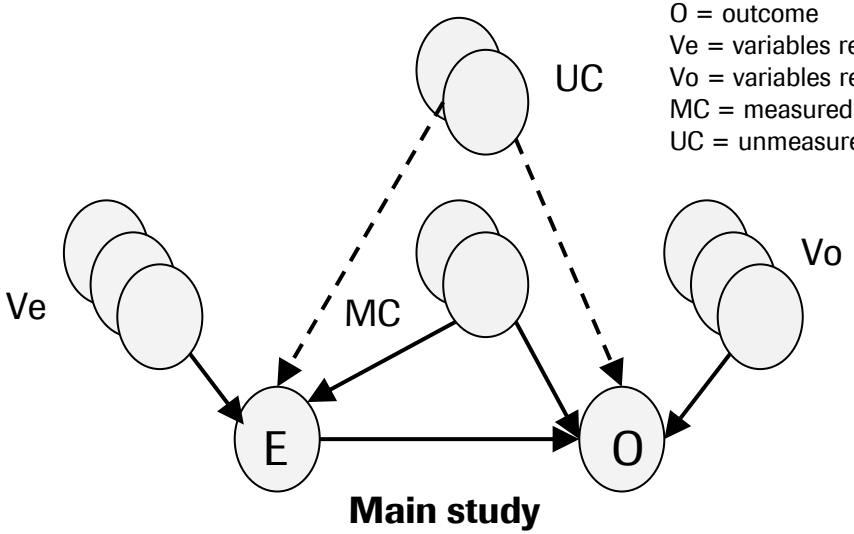


Unmeasured Confounding using PS

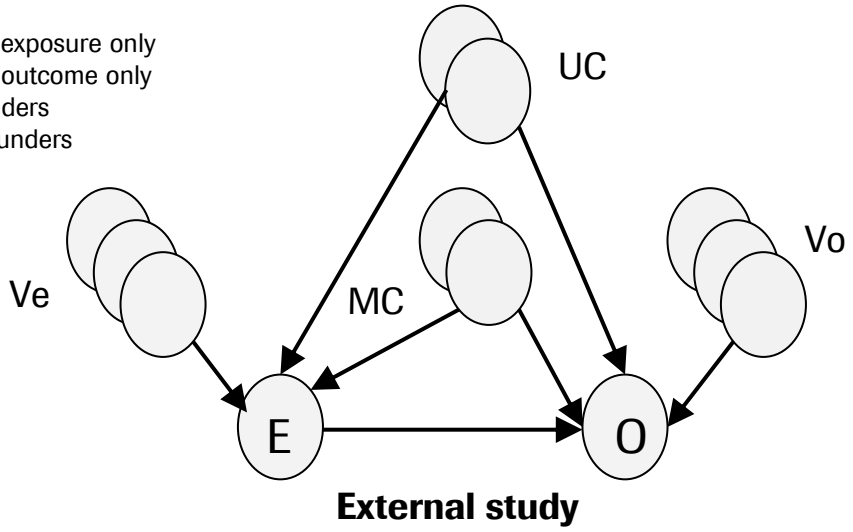


Frequentist Propensity Score Calibration

E= exposure
O = outcome
Ve = variables related to exposure only
Vo = variables related to outcome only
MC = measured confounders
UC = unmeasured confounders



Main study



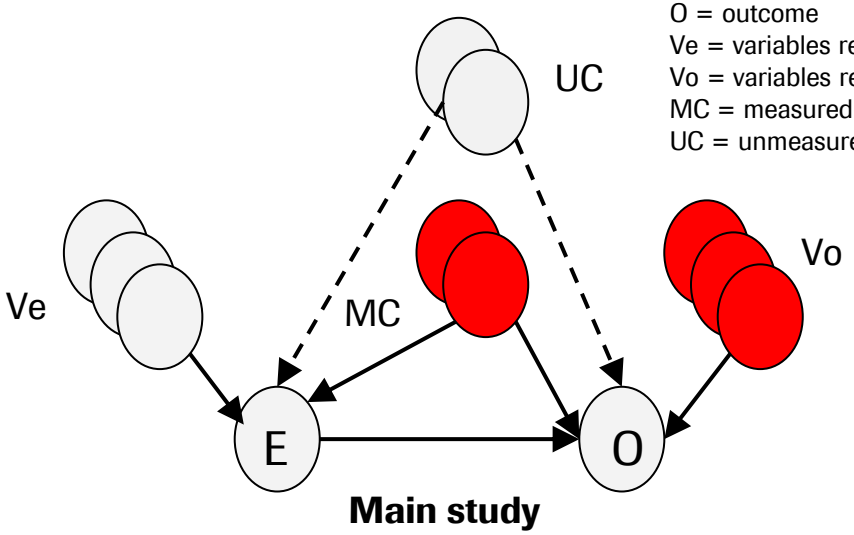
External study

E	O	Ve	Vo	MC	UC
x	x	x	x	x	?
x	x	x	x	x	?
x	x	x	x	x	?

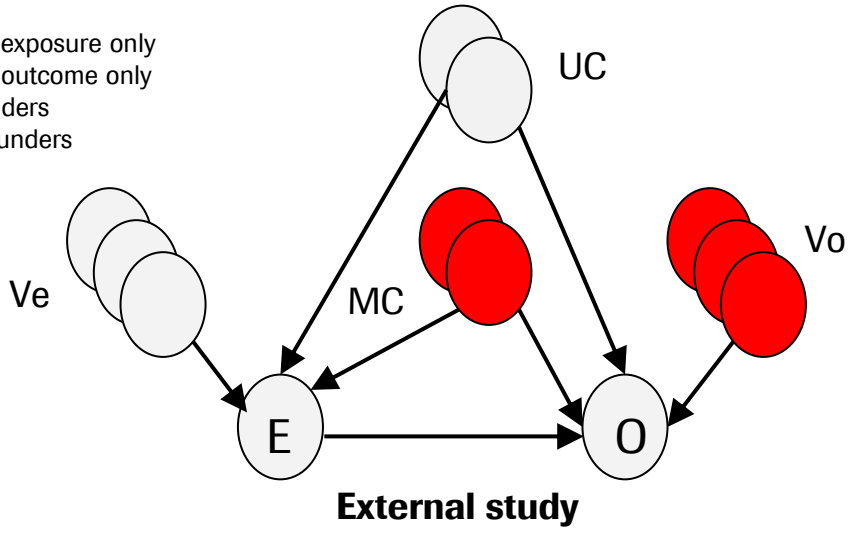
E	O	Ve	Vo	MC	UC
x	(x)	x	x	x	x
x	(x)	x	x	x	x
x	(x)	x	x	x	x

Frequentist Propensity Score Calibration

E= exposure
O = outcome
Ve = variables related to exposure only
Vo = variables related to outcome only
MC = measured confounders
UC = unmeasured confounders



Main study



External study

Error-prone PS_{main}

Logistic regression

E ~ MC + Vo

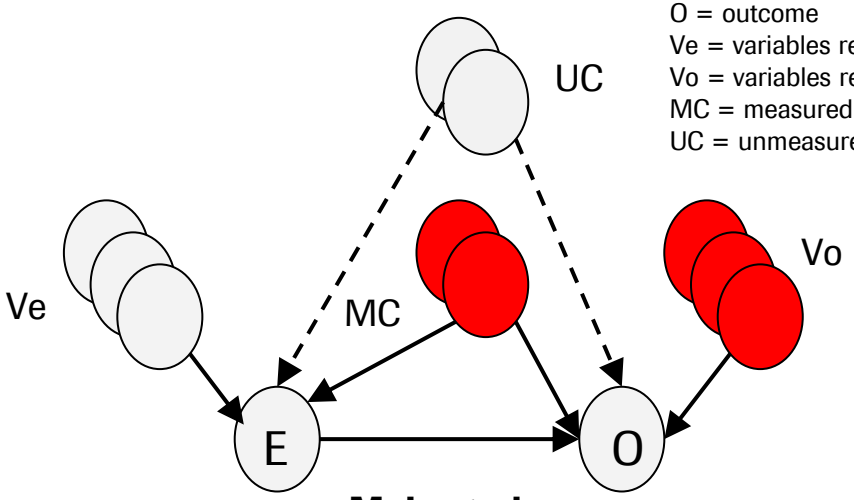
Error-prone $PS_{external}$

Logistic regression

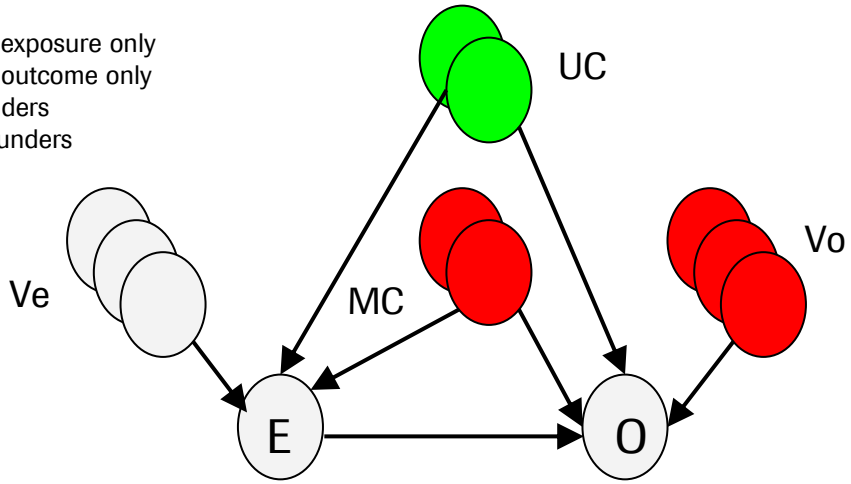
E ~ MC + Vo

Frequentist Propensity Score Calibration

E= exposure
O = outcome
Ve = variables related to exposure only
Vo = variables related to outcome only
MC = measured confounders
UC = unmeasured confounders



Main study



External study

Error-prone PS_{main}

Logistic regression

E ~ MC + Vo

Error-prone $PS_{external}$

Logistic regression

E ~ MC + Vo

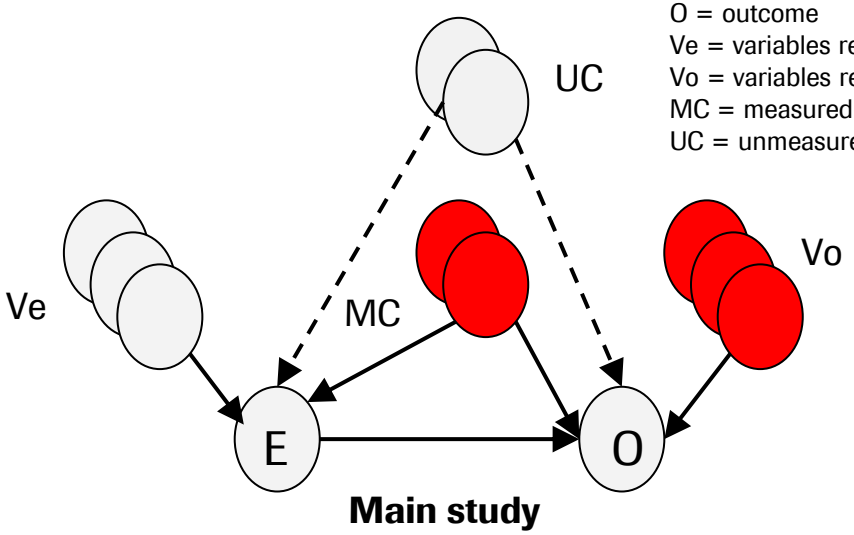
Gold standard $PS_{external}$

Logistic regression

E ~ MC + UC + Vo

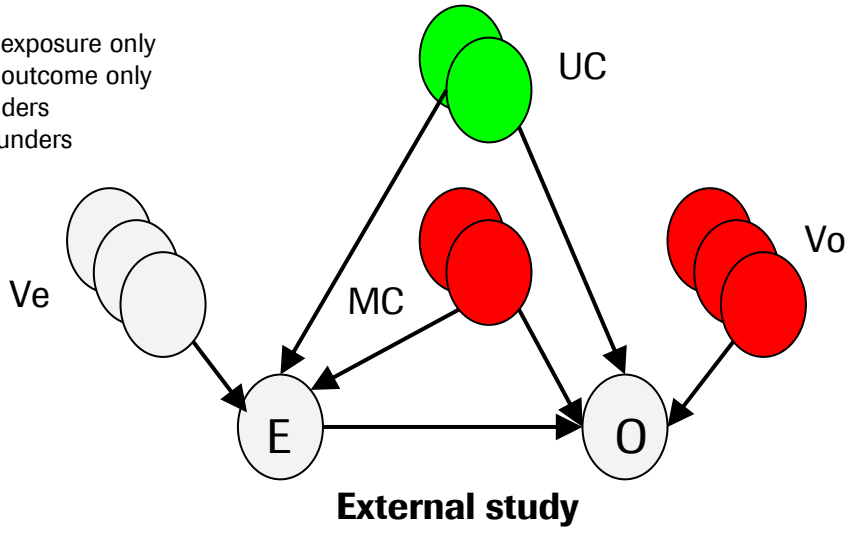
Frequentist Propensity Score Calibration

E= exposure
O = outcome
Ve = variables related to exposure only
Vo = variables related to outcome only
MC = measured confounders
UC = unmeasured confounders



Main study

Error-prone PS_{main}



External study

Error-prone $PS_{external}$

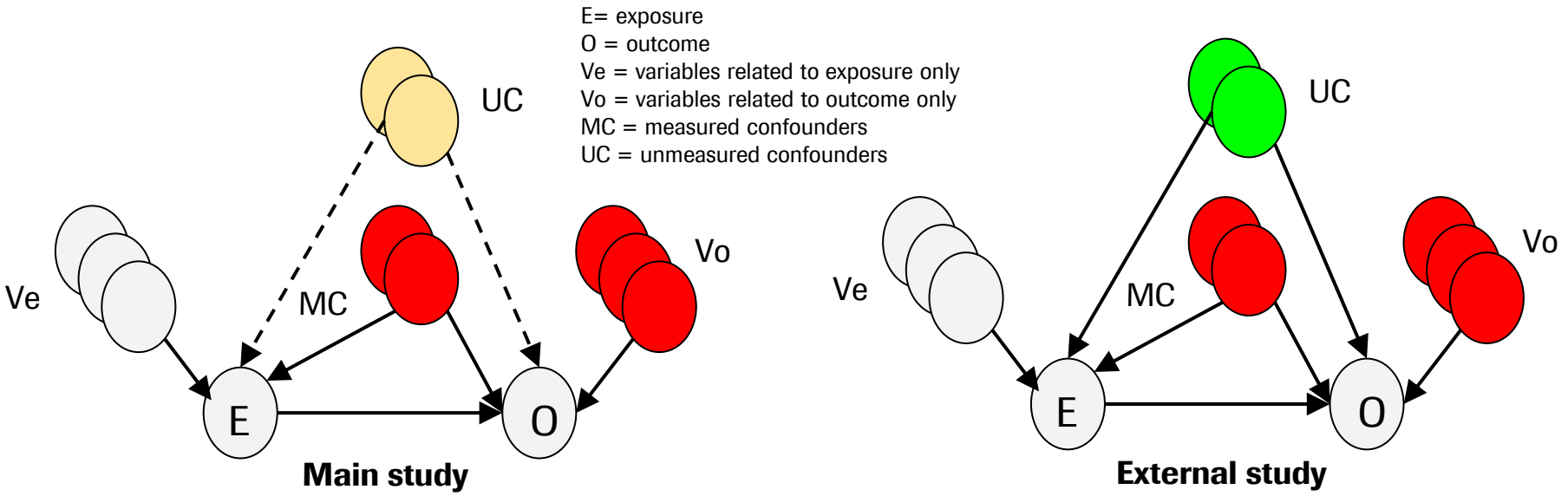
Linear regression

$PS_{GS} \sim E + PS_{EP}$

CALIBRATION

Gold standard $PS_{external}$

Frequentist Propensity Score Calibration



Error-prone PS_{main}

Error-prone $PS_{external}$

CALIBRATION

Gold standard $PS_{external}$

Corrected/adjusted PS_{main}

Surrogacy hypothesis

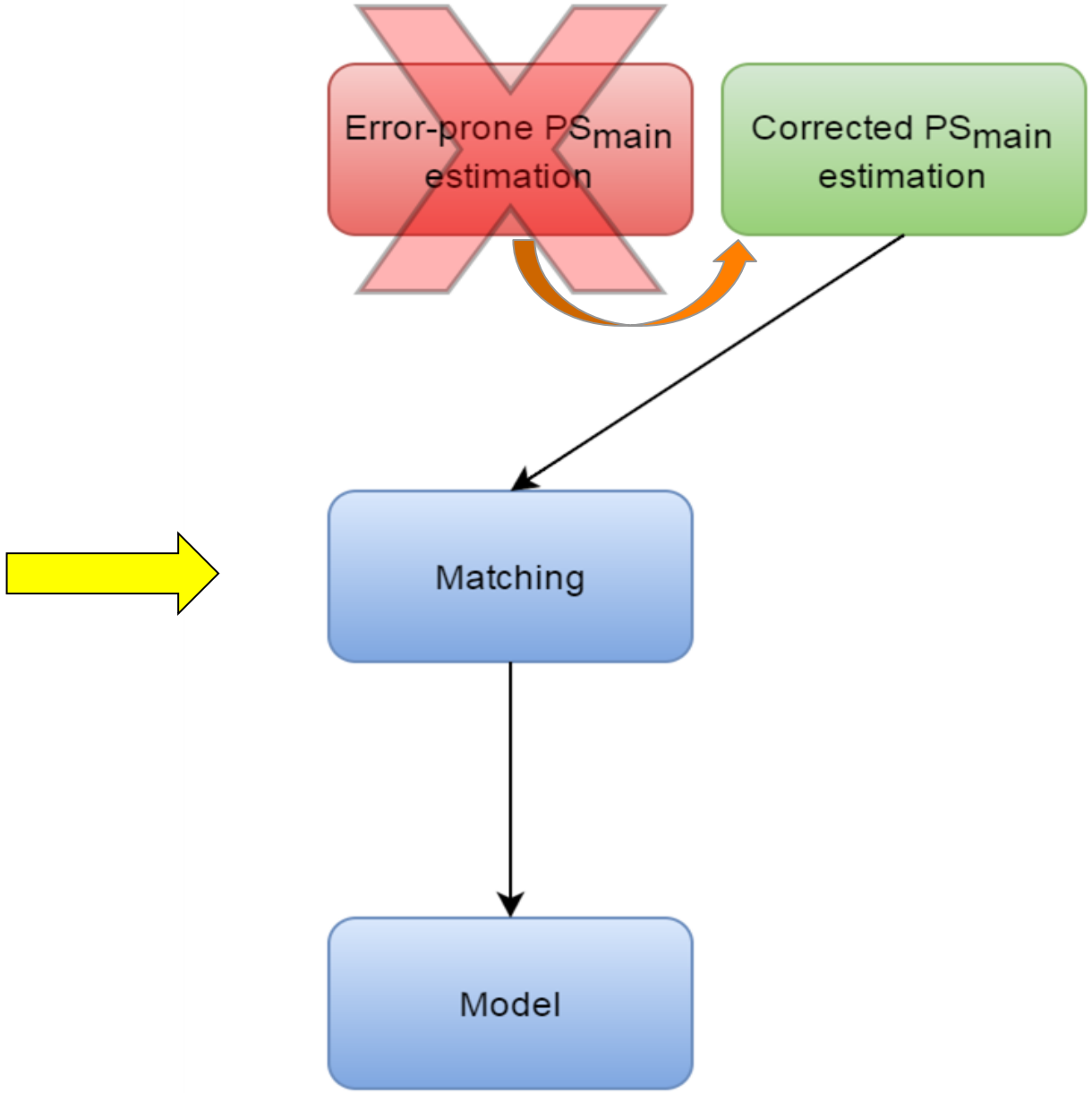
The effectiveness of the Propensity Score Calibration (PSC) strongly relies on the **surrogacy hypothesis,** requiring that

"the error-prone propensity score is independent of the outcome of interest given the gold standard PS and the exposure".
(i.e. the error-prone PS is a **surrogate** of the gold standard PS).

O independent $\mathbf{PS_{EP} \mid PS_{GS}}$ **, E**

Stürmer et al., (2005), Adjusting Effect Estimates for Unmeasured Confounding with Validation Data using Propensity Score Calibration, American Journal of Epidemiology, 162 (3): 279–289

Unmeasured Confounding using PS

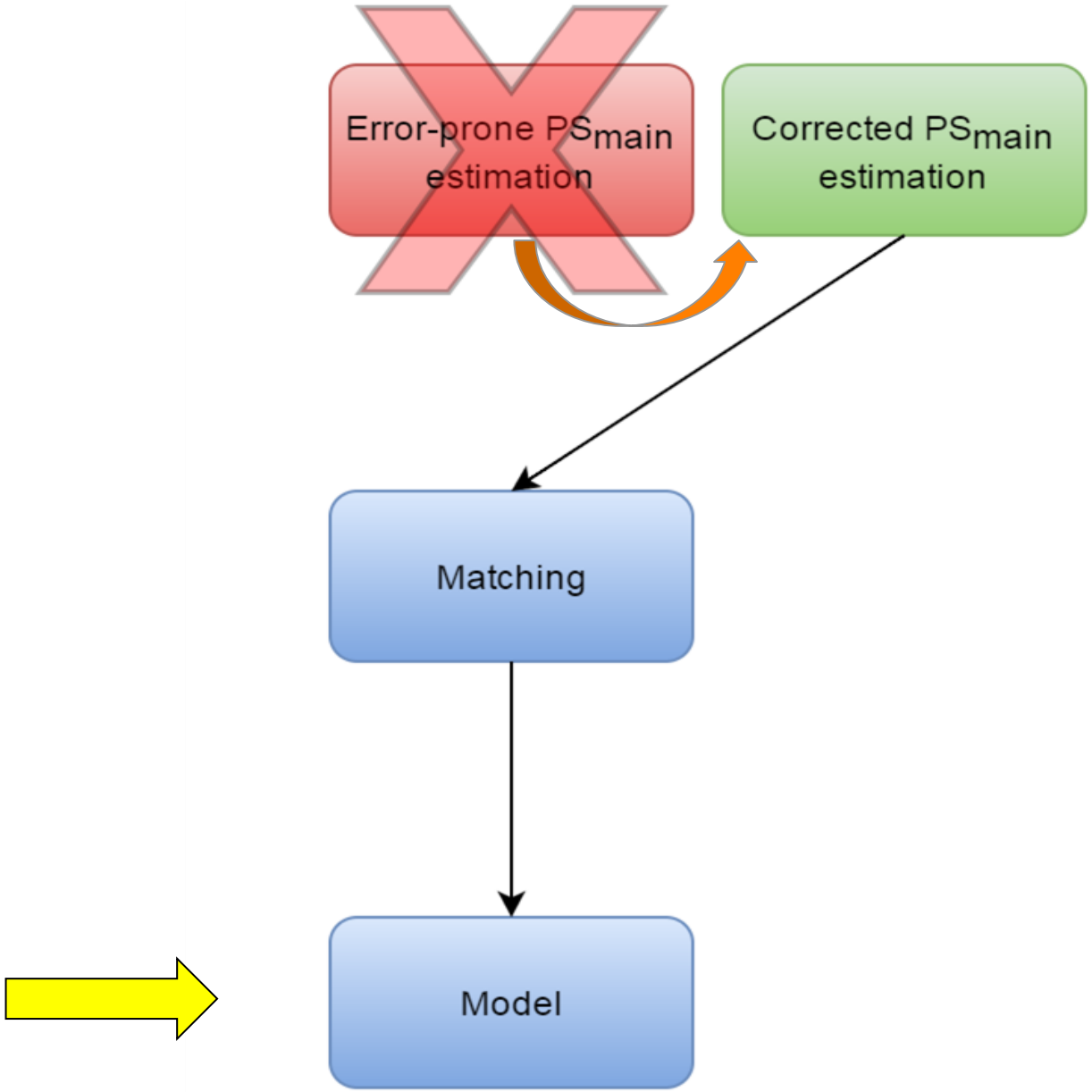


- Use of external data
- Calibration

Matching

1:1 with no replacement within a fixed caliper
by greedy algorithm

Unmeasured Confounding using PS



- Use of external data
- Calibration

- Matching in the main study population

Model

Marginal treatment effect estimated by the appropriate model on the matched population (e.g. Cox model)

O ~ E

O= Outcome E=Exposure

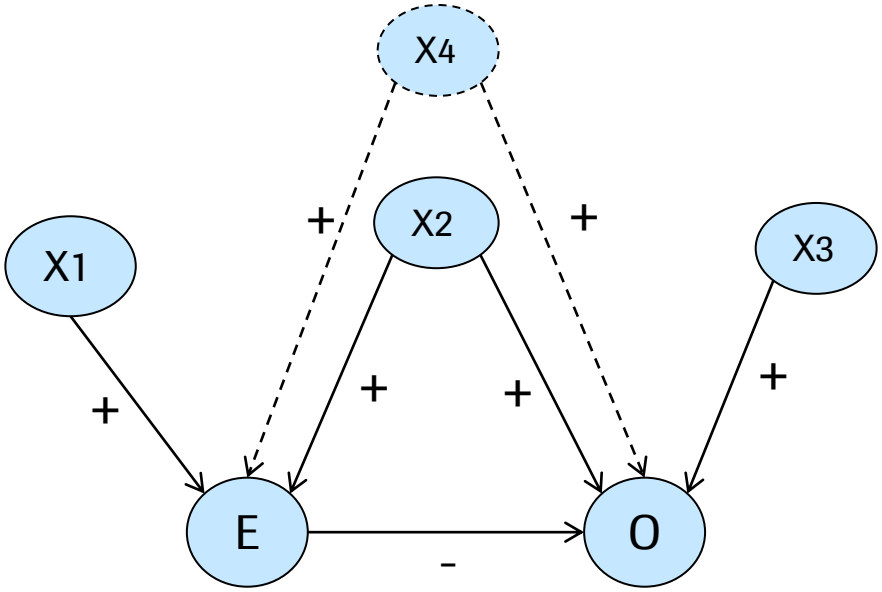
Introduction

Propensity Score Calibration

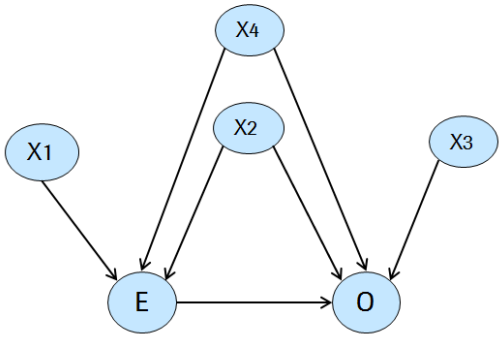
Monte Carlo simulation

Conclusions

Simulation for Unmeasured Confounding



Main study
1000 subjects



External study
250 subjects

X1, X2, X3, X4 ~ Bernoulli(0.5)
logit(p) = 1.5*X_1+1.5*X_2+1.5*X_4+logit(0.05)
E ~ Bernoulli(p)
log(r)= log(0.76)*E+log(2)*X_2+log(2)*X_4+log(2)*X_3-3.5
O ~ exp (r)
No censoring

Simulation for Unmeasured Confounding

Example of PS models and calibration

log(OR)	Error Prone PS	Gold standard PS
Intercept	-1.17 ***	-1.82 ***
X2	1.49 ***	1.47 ***
X3	-0.26	-0.21
X4	NA	1.08 ***

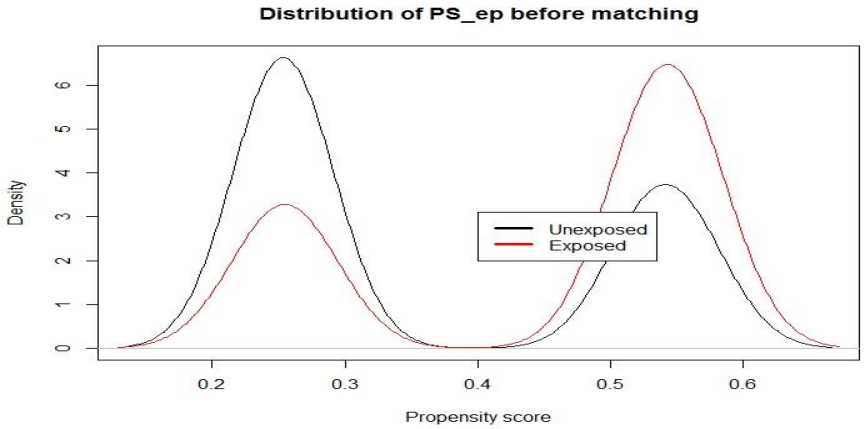
p<= 0.001 ***
0.001 < p <= 0.01 **
0.01 < p <= 0.05 *
0.05 < p <= 0.1 .
p>0.1

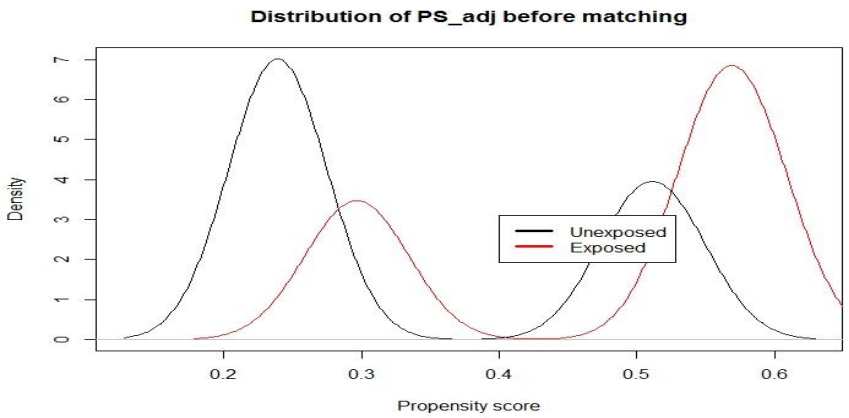
	Calibration
Intercept	0.00
E	0.06 ***
PS ep	0.94 ***

Simulation for Unmeasured Confounding

Example of PS models and calibration

	Calibration
Intercept	0.00
E	0.06 ***
PS ep	0.94 ***





Simulation for Unmeasured Confounding

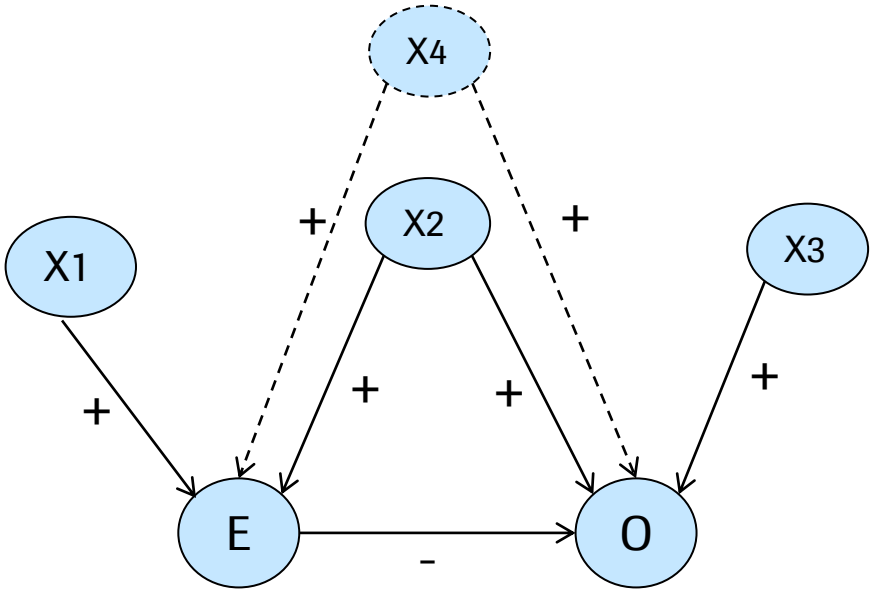
Results for 1000 repetitions

True = 0.80	Point est (HR)	95% confint	Relative Bias	Matched subjects
Crude	1.12	[0.99, 1.29]	40%	NA
Matching (EP)	0.96	[0.84, 1.10]	20%	716
Matching (adj)	0.90	[0.58, 1.18]	18%	657

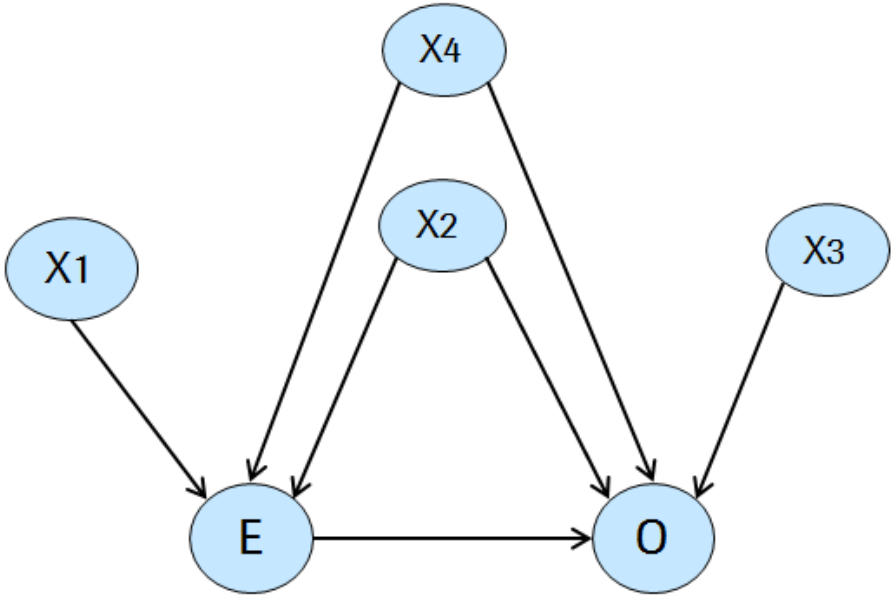
Legend

- Point est: mean Hazard Ratio (HR) across 1000 repetitions
- 95% confint: 95% empirical confidence interval using 2.5% and 97.5% quantiles across 1000 repetitions
- Relative Bias: mean relative absolute bias across 1000 repetitions
- Matched subjects: mean number of matched patients across 1000 repetitions

Other scenarios: larger external study



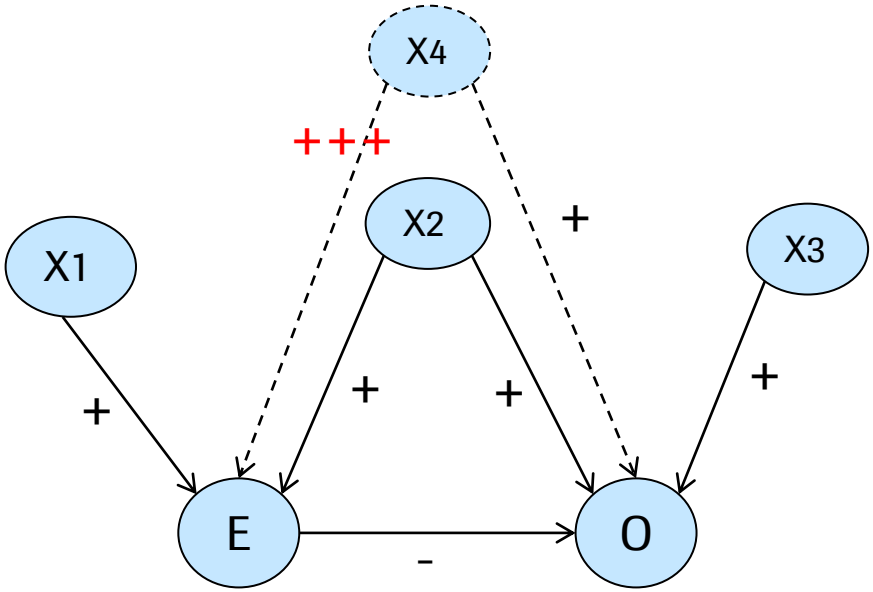
Main study
1000 subjects



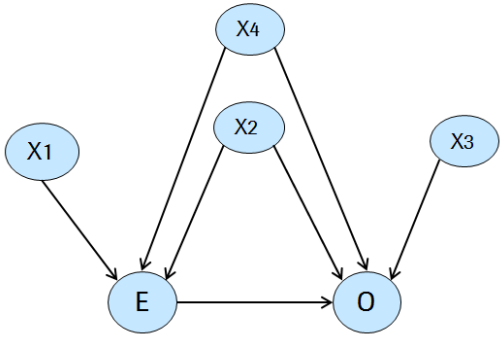
External study
1000 subjects

Variance reduction for the PSC estimate

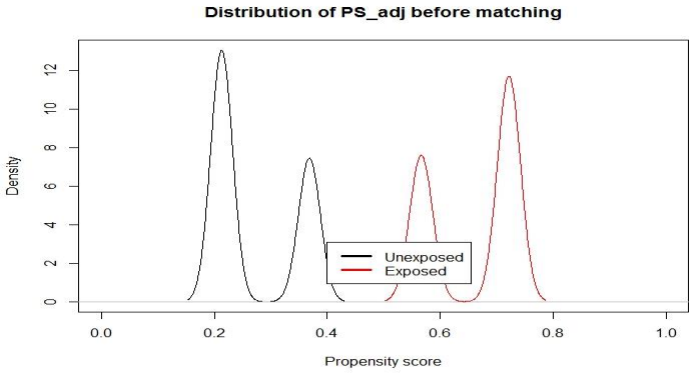
Other scenarios: predominant UC



Main study
1000 subjects

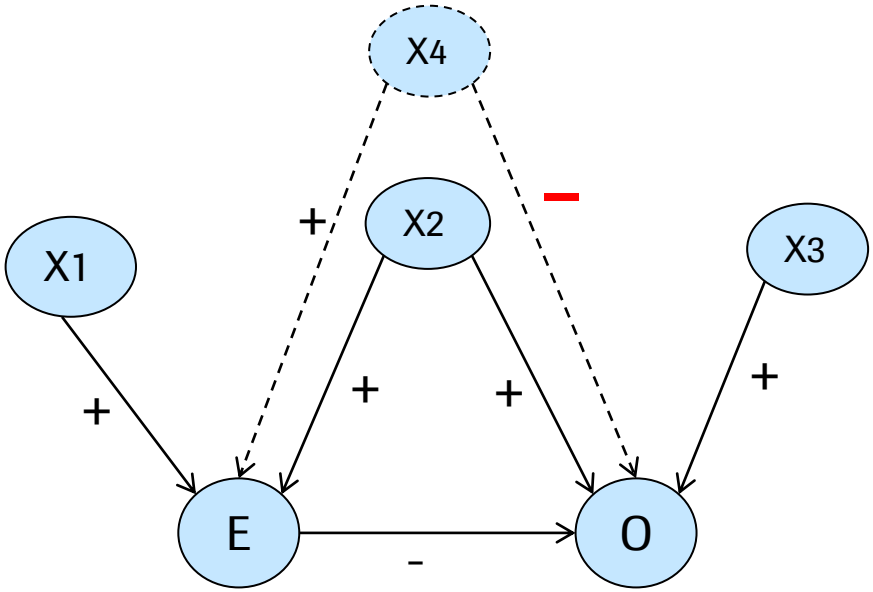


External study
250 subjects

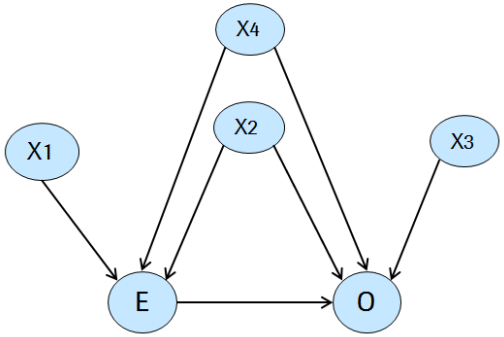


- **No overlapping → matching problem**
- **Balance of confounders is not achieved**

Other scenarios: surrogacy violation



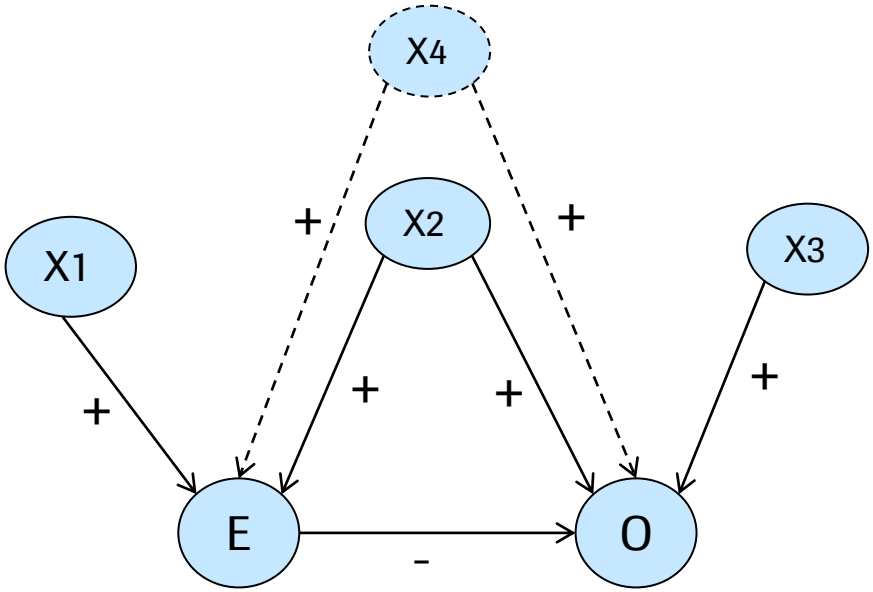
Main study
1000 subjects



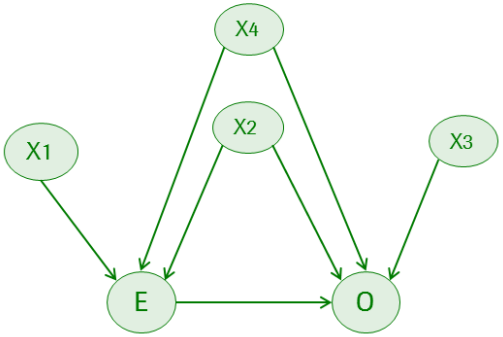
External study
250 subjects

Biased estimate

Other scenarios: transportability



Main study
1000 subjects



External study
250 subjects

Perturbations in:

- **Exposure prevalence**
- **Exposure Causal Parameters**

PSC correction may not be significant or reliable

Introduction

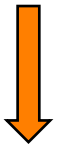
Propensity Score Calibration

Monte Carlo simulation

Conclusions

Conclusions

- The **surrogacy test** enables to define a context where PSC **reduces** confounding bias
- This allows to use PSC to perform sensitivity analyses, once the related practical elements have been properly defined



- The **matching algorithm** and the **caliper** may significantly impact on the performance of the method: different options may be undertaken
- The choice of the **external database** (and its size) plays an important role, even if this point is often related to practical feasibility constraints and it is difficult to check
- Correlations and different strengths of covariates effects on treatment and outcome require further assessment

Doing now what patients need next