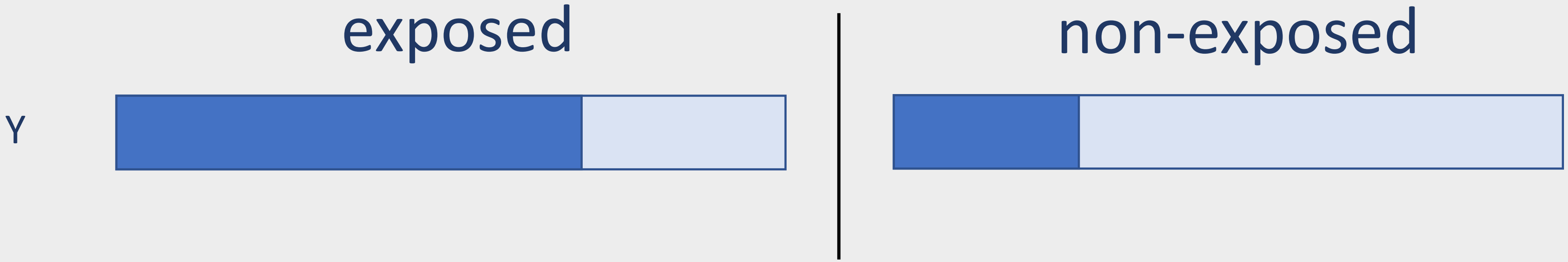
Agenda

- Data Simulation
- Paper
- E. Schnitzer & R. Platt

Agenda

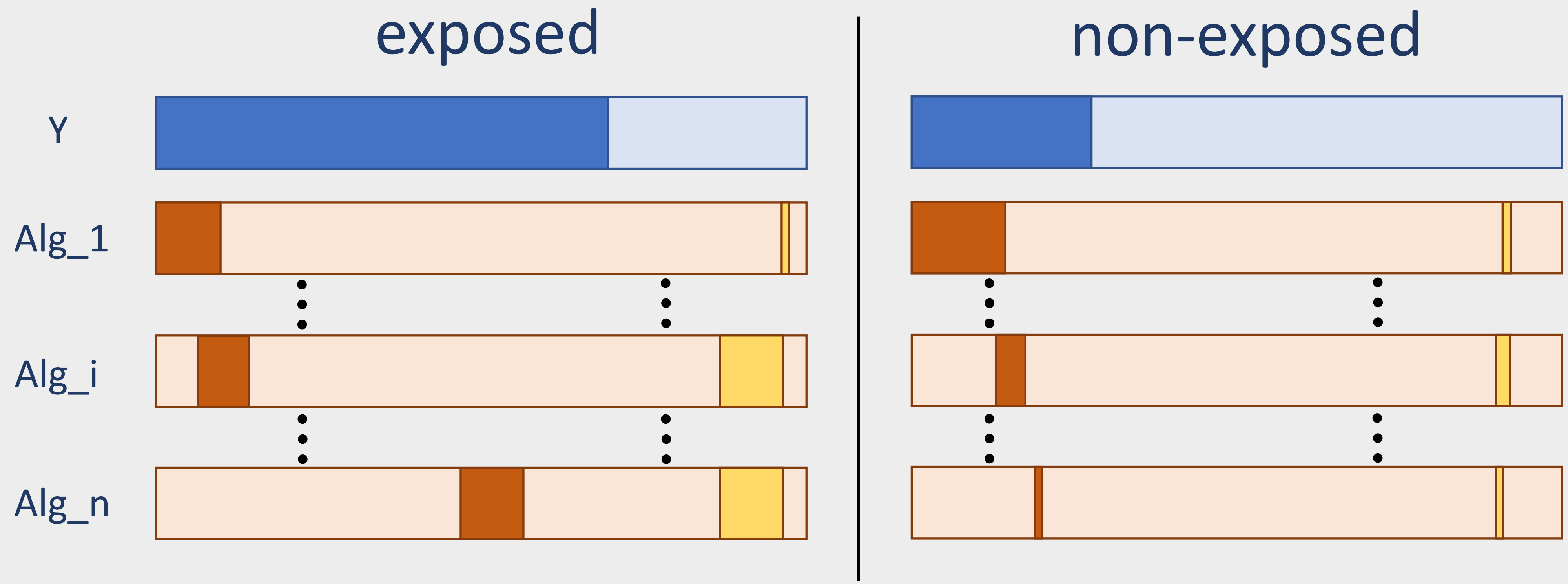
- Data Simulation
- Paper
- E. Schnitzer & R. Platt



Y = 0

Y = 1

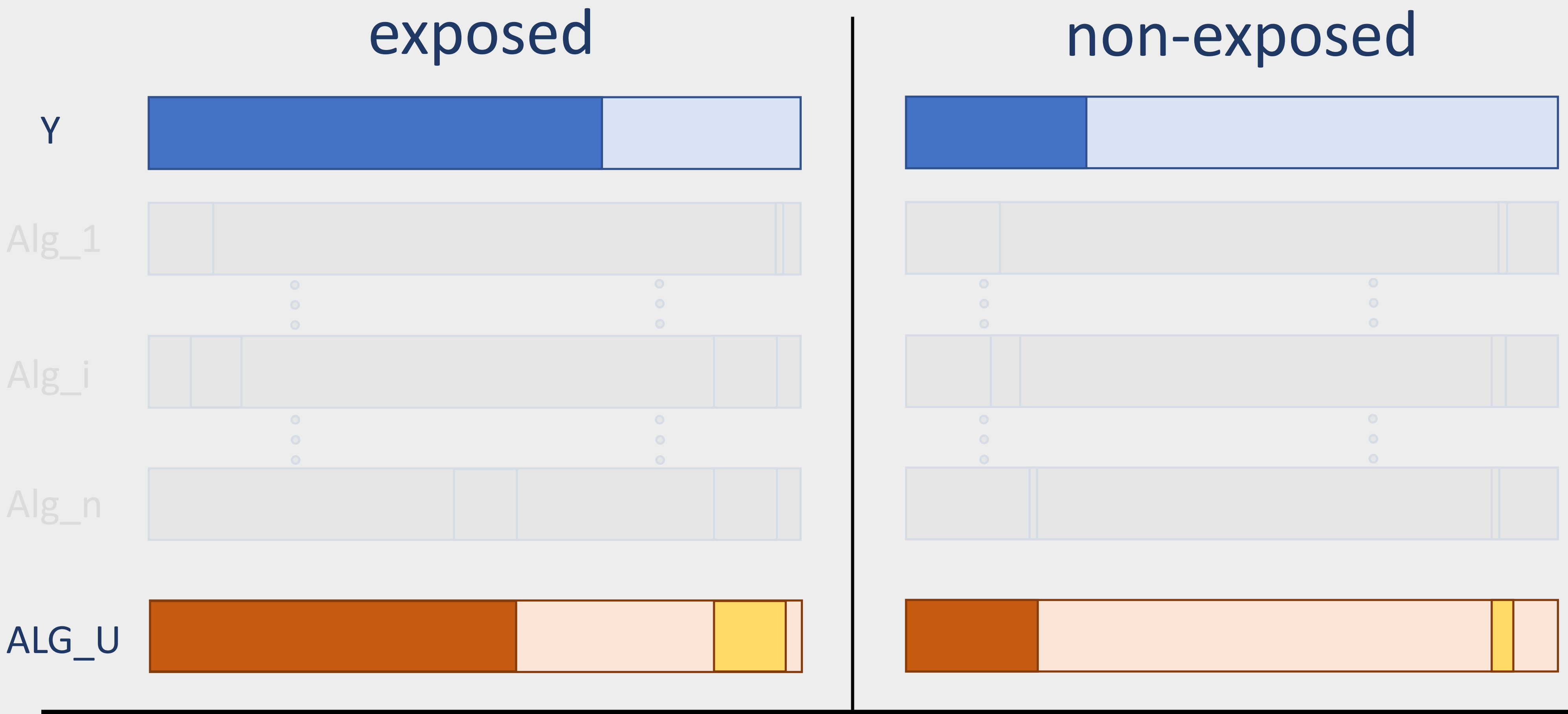
- Y could be different between groups



Y = 0
Y = 1

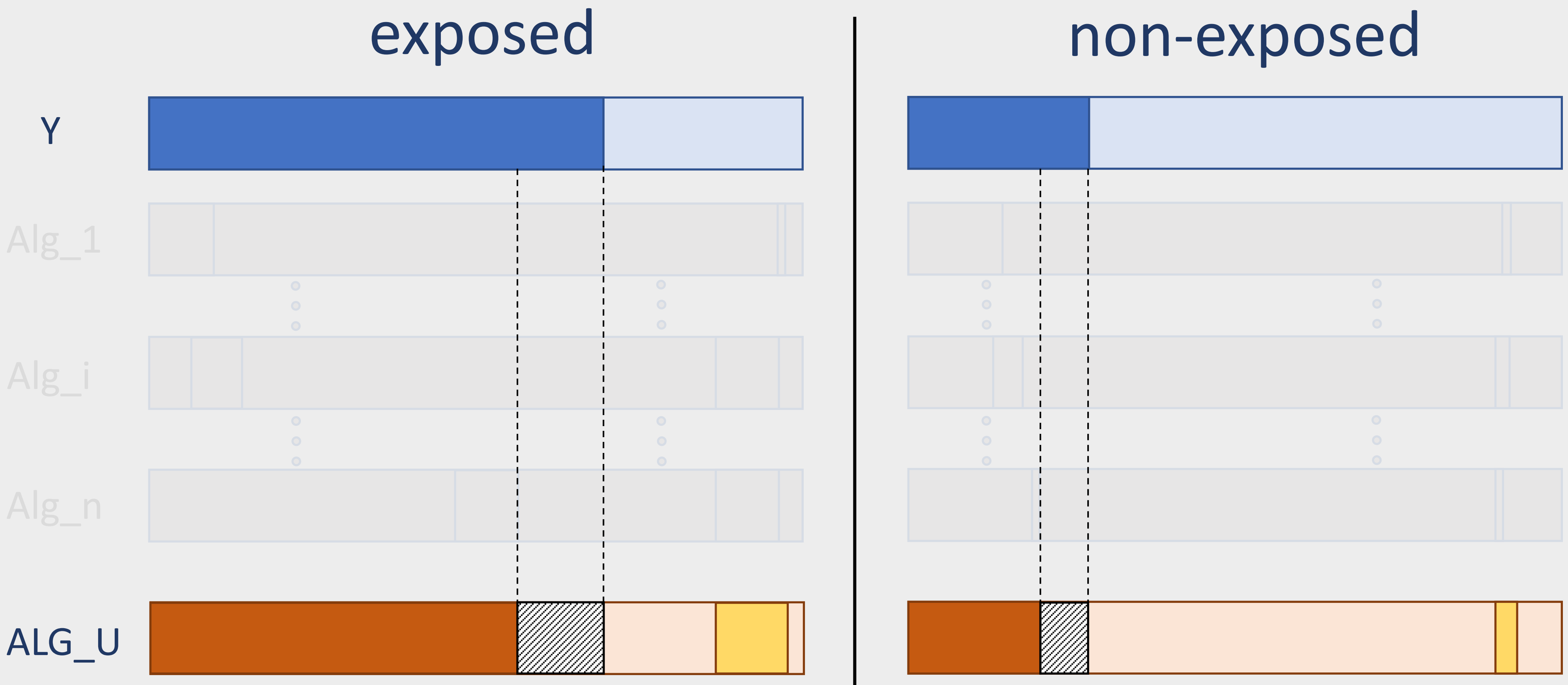
FP
TP
Alg = 0

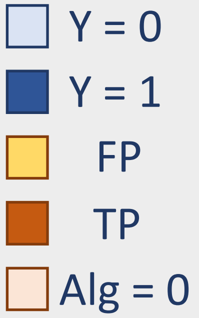
- Many algorithms can be combined, each one finding a part of the positives
- The validation indices of the individual algorithms can be non-differential



Y = 0
Y = 1
FP
TP
Alg = 0

- ALG_U corresponds to Alg_1 U…U Alg_n U…U Alg_n





FN

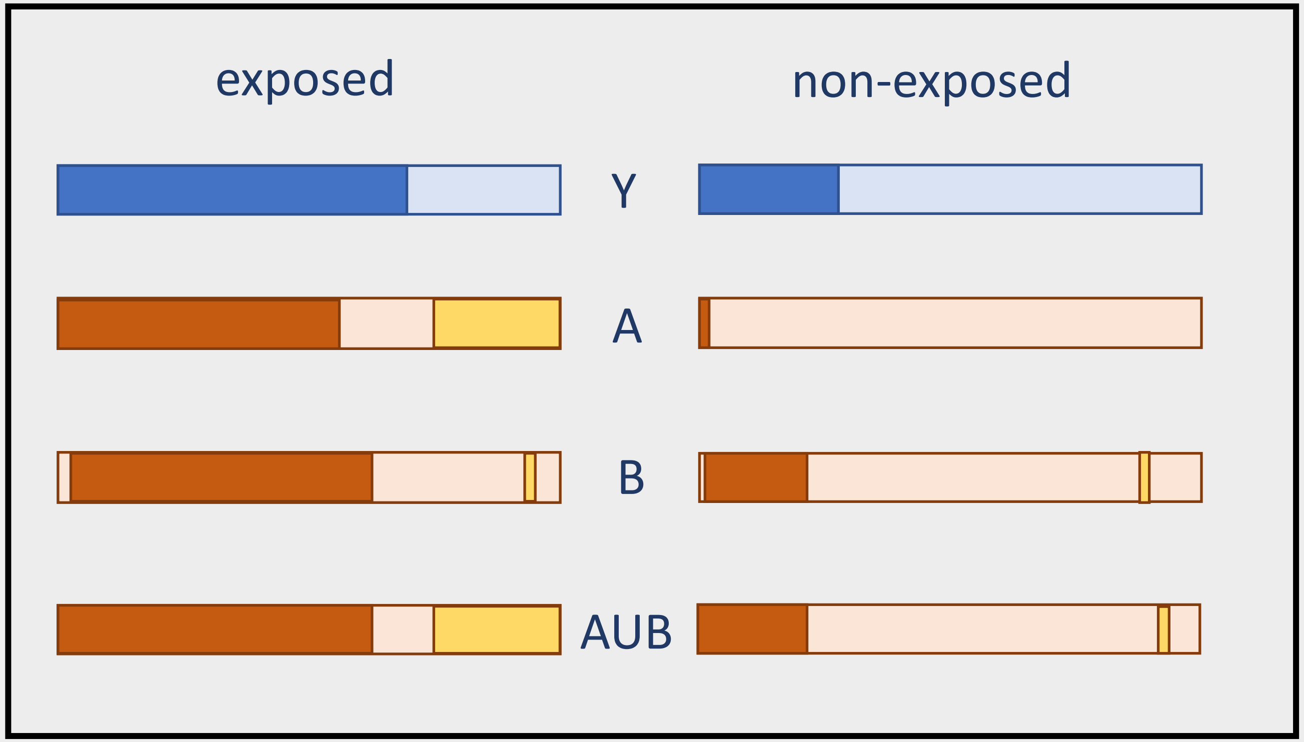
The proportion of false negative of ALG_U (so the sensitivity), is assumed to be the equal in both groups, and it must be due to 'MAJOR' cause:

- Non-sistematic errors
- Private medicine
- …

All the others cases must detected in al least one algorithm.

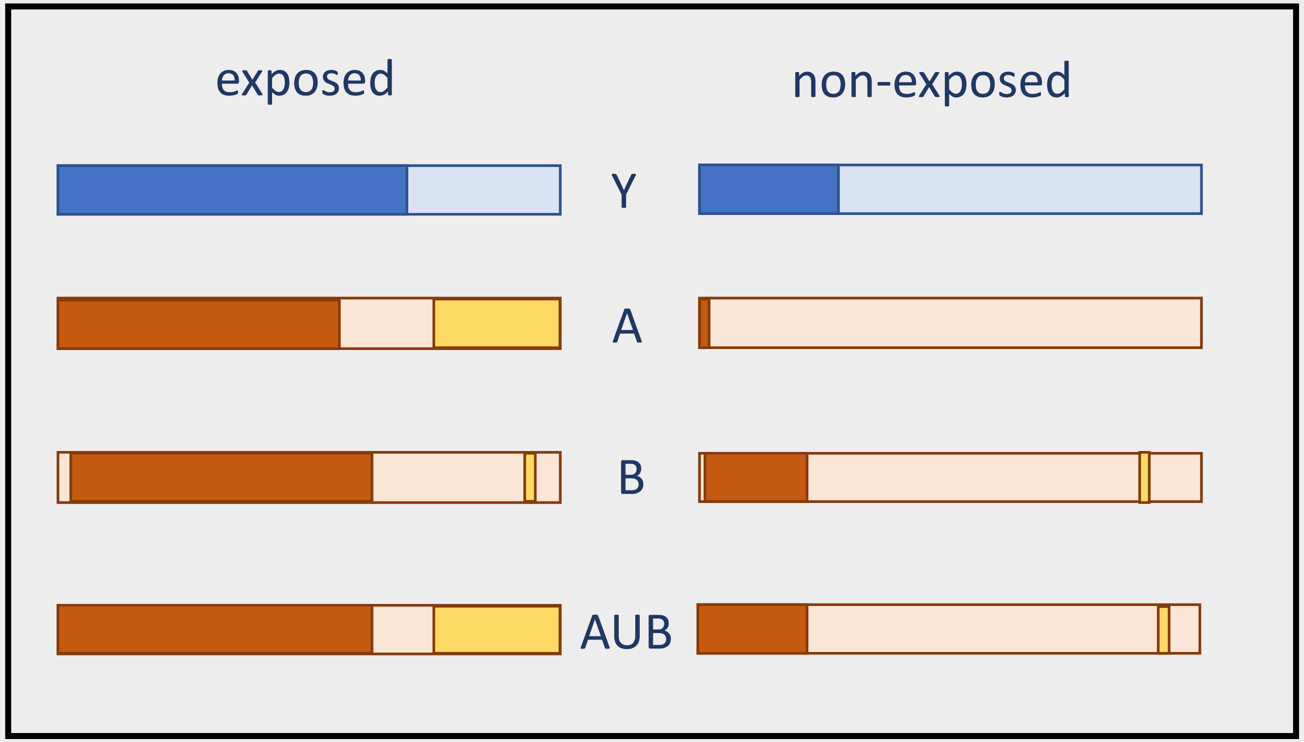
Example

Some events are certainly found in a patient's medical history (e.g. lung cancer), unless there are ' MAJOR ' causes. We can think of an exposure positively correlated with the event (e.g. smoking), and a first algorithm A including all subjective diagnoses and a second algorithm B including all objective diagnoses.



Example

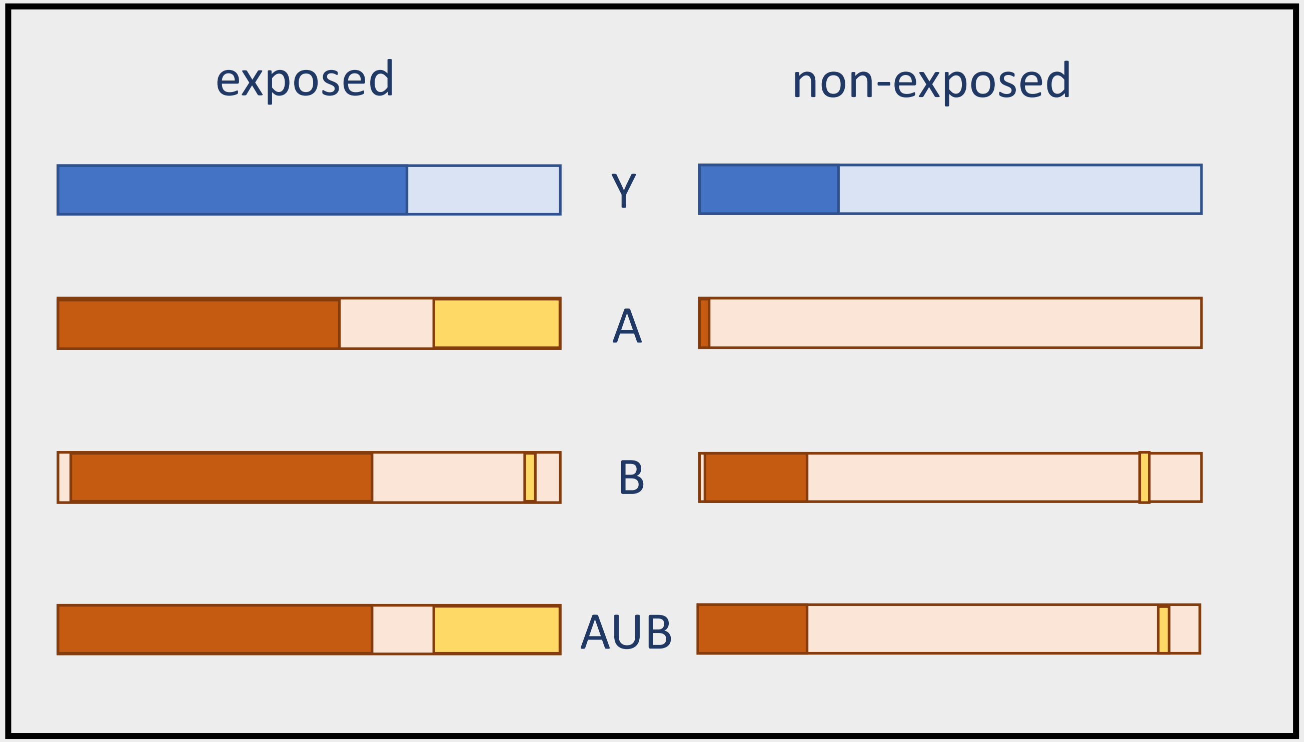
Some events are certainly found in a patient's medical history (e.g. lung cancer), unless there are ' MAJOR ' causes. We can think of an exposure positively correlated with the event (e.g. smoking), and a first algorithm A including all subjective diagnoses and a second algorithm B including all objective diagnoses.



- events are higher among the exposed group

Example

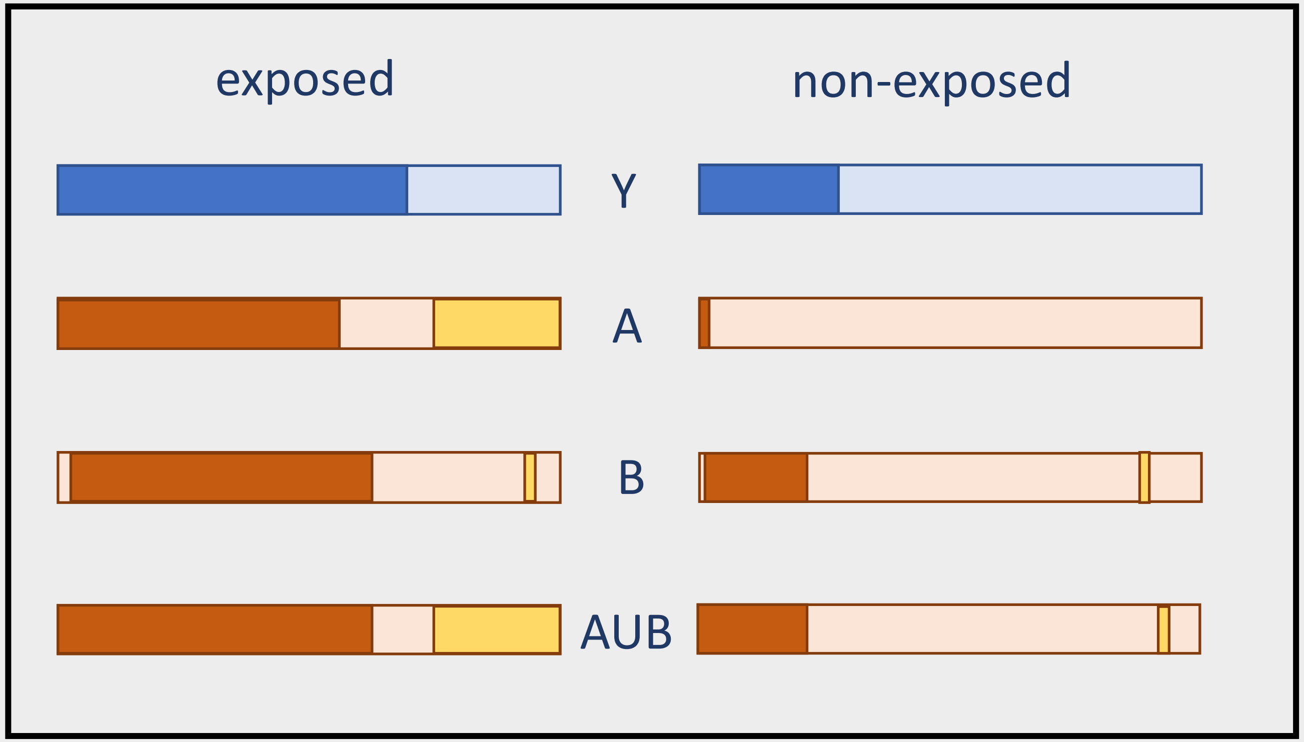
Some events are certainly found in a patient's medical history (e.g. lung cancer), unless there are ' MAJOR ' causes.
We can think of an exposure positively correlated with the event (e.g. smoking), and a first algorithm A including all subjective diagnoses and a second algorithm B including all objective diagnoses.



- events are higher among the exposed group
- a subjective diagnosis would result in A having a higher SE in the exposed group, but also a lower SP, due to the influence of exposure
- A may have a higher SP in the non-exposed group, but also a lower SE

Example

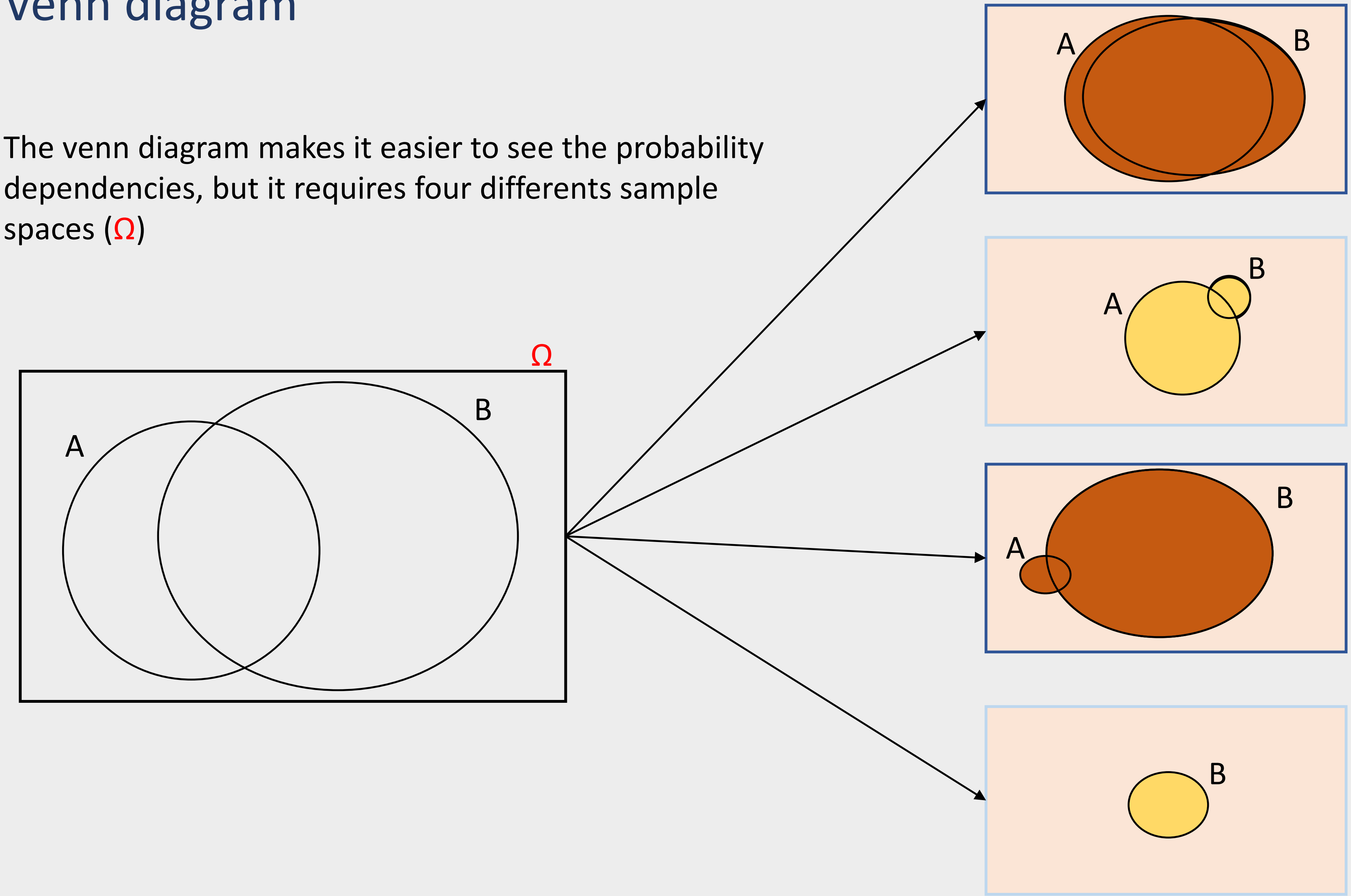
Some events are certainly found in a patient's medical history (e.g. lung cancer), unless there are ' MAJOR ' causes.
We can think of an exposure positively correlated with the event (e.g. smoking), and a first algorithm A including all subjective diagnoses and a second algorithm B including all objective diagnoses.



- events are higher among the exposed group
- a subjective diagnosis would result in A having a higher SE in the exposed group, but also a lower SP, due to the influence of exposure
- A may have a higher SP in the non-exposed group, but also a lower SE
- the proportion of FP in AUB should not depend on the exposure group (MAJOR causes ⊥ E)

Venn diagram

The venn diagram makes it easier to see the probability dependencies, but it requires four differents sample spaces (Ω)



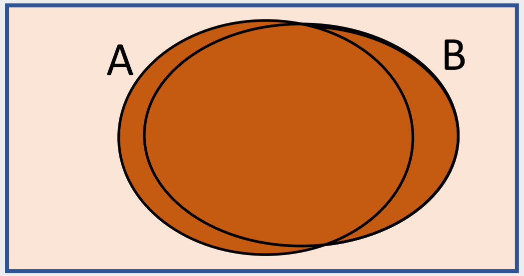
Ω: Y == 1 & E == 1
P(A) = SE_A^e
P(B) = SE_B^e

Ω: Y == 0 & E == 1
P(A) = 1 - SP_A^e
P(B) = 1 - SP_B^e

Ω: Y == 1 & E == 0
P(A) = $SE_A^{\bar{e}}$
P(B) =$SE_B^{\bar{e}}$

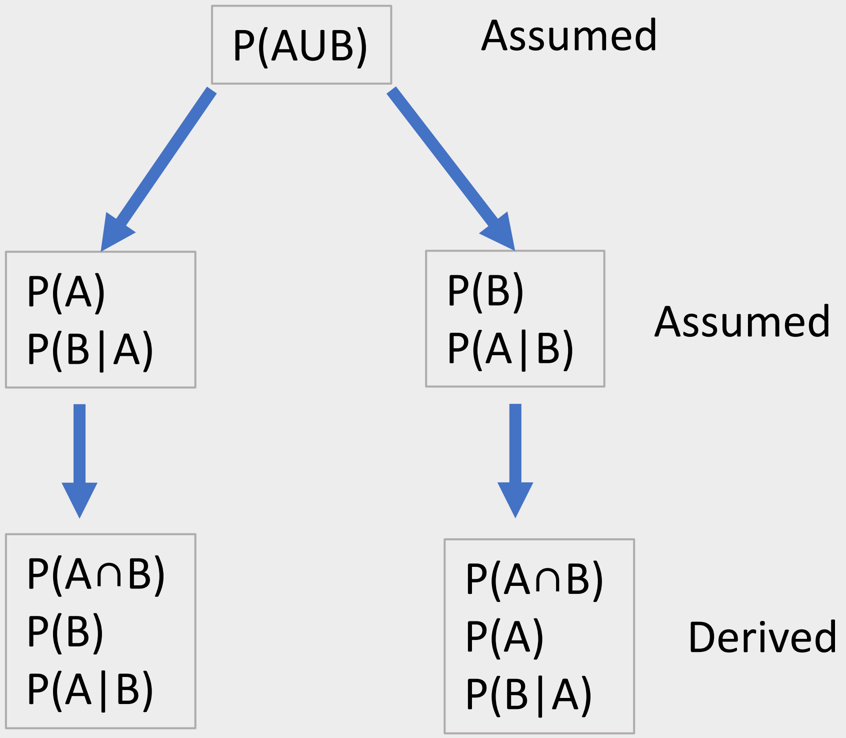
Ω: Y == 0 & E == 0
P(A) = 1 - $SP_A^{\bar{e}}$
P(B) = 1 - $SP_B^{\bar{e}}$

Venn diagram

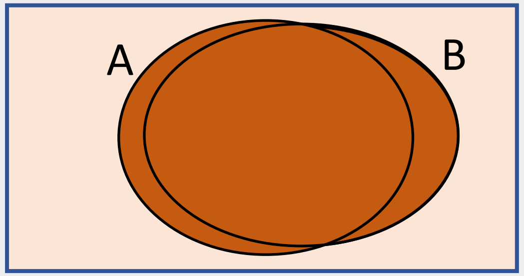


For each Ω six different probabilities exist:

- P(A)
- P(B)
- P(AUB)
- P(A∩B)
- P(A|B)
- P(B|A)

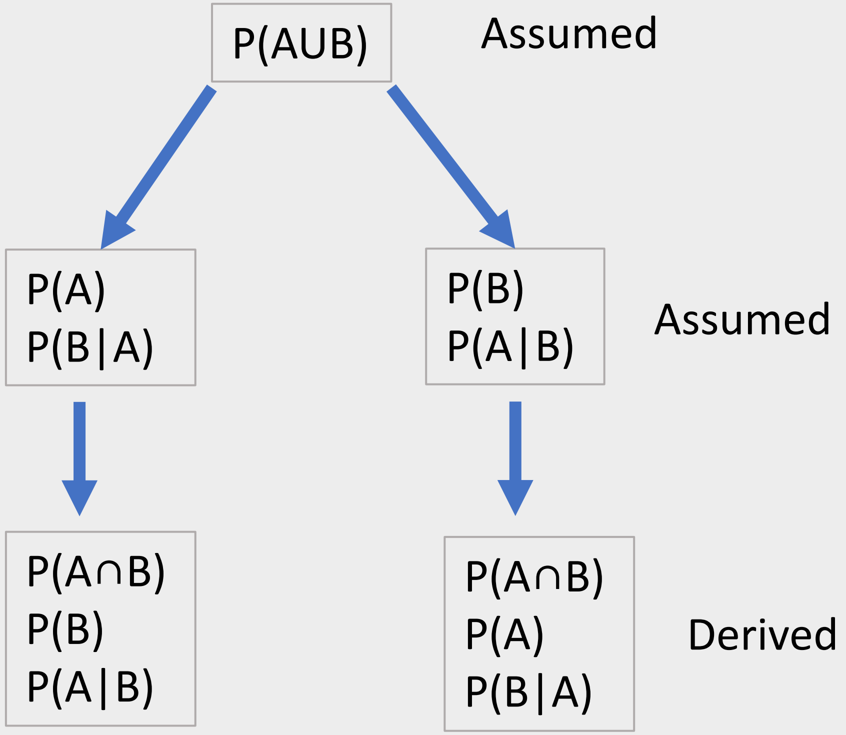


Venn diagram



For each Ω six different probabilities exist:

- P(A)
- P(B)
- P(AUB)
- P(A∩B)
- P(A|B)
- P(B|A)



P(AUB) = 0.80
P(A) = 0.7
P(B|A) = 0.9

P(A∩B) = P(A) * P(B|A) → P(A∩B) = 0.7 * 0.9 = 0.63
P(B) = P(AUB) – P(A) + P(A∩B) → P(B) = 0.8 - 0.7 + 0.63 = 0.73
P(A|B) = P(A∩B) / P(B) → P(A|B) = 0.63 / 0.73 = 0.86

Agenda

- Data Simulation
- Paper
- E. Schnitzer & R. Platt

Agenda

- Data Simulation
- Paper
- E. Schnitzer & R. Platt

E. Schnitzer & R. Platt

	E. Schnitzer & R. Platt	Pregnancy Algorithm	Non-differentility test
Data	• Defined cohort of pregnancies	• Cohort of pregnancies defined by multiple components	• Simulated data
Exposure	• Gestational age-specific exposure		• Binary exposure

Causal Inference	• Target trials (ITT):
	  • IPW
	  • G-computation
	  • TMLE
			• Differential misclassification
			• Risk Ratio

Possible developement		• Can missed pregnancies affect causal effects?	