G-computation and propensity score weighting to estimate a marginal effect

yohann.foucher@univ-poitiers.fr

Best practices and recent advances in causal analyses, October 14, 2025

Université de Poitiers

CHU Poitiers

Inserm
La science pour la santé
From science to health

Plan

Introduction

Inverse probability weighting (IPW)

G-computation (GC)

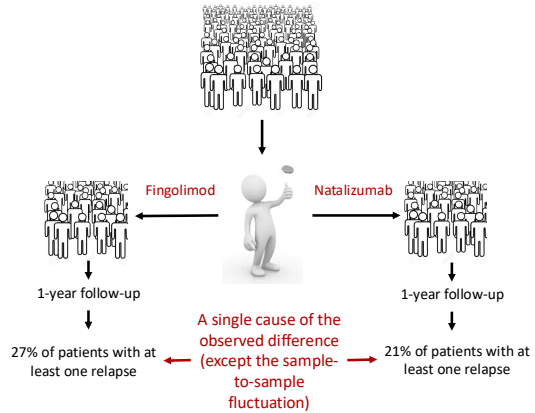
Assumptions of causal inference

References

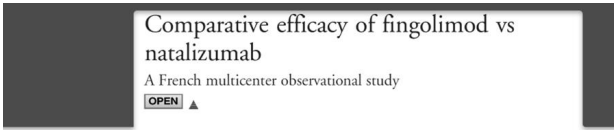
A use case in multiple sclerosis (MS)

- A MS relapse is the occurrence of new MS symptoms or the worsening of old ones.
- Fingolimod and natalizumab share the same indication for second-line treatment.
- Which treatment is most effective in preventing MS relapse?

Randomization allows us to demonstrate a causal relationship



Confounders in observational studies: data from Barbin et al. [1]



Comparative efficacy of fingolimod vs natalizumab

A French multicenter observational study

OPEN

Laetitia Barbin, PhD*
Chloe Rousseau, MSc*
Natacha Jousset, BSc
Romain Casey, PhD
Marc Debouverie, MD, PhD

ABSTRACT

Objective: To compare natalizumab and fingolimod on both clinical and MRI outcomes in patients with relapsing-remitting multiple sclerosis (RRMS) from 27 multiple sclerosis centers participating in the French follow-up cohort Observatoire of Multiple Sclerosis.

Baseline characteristics	All patients	Fingolimod	Natalizumab
No. of patients considered for analysis	629	303	326
Female, n (%)	479 (76.2)	225 (74.3)	254 (77.9)
Age at treatment initiation, y, mean (SD)	37.0 (9.6)	37.2 (9.2)	36.8 (9.9)
Disease duration at treatment initiation, y, mean (SD)	8.5 (6.4)	9.0 (6.8)	8.0 (6.1)
EDSS score, mean (SD)	2.6 (1.3)	2.4 (1.3)	2.8 (1.3)
EDSS score between 3.0 and 5.5, n (%)	288 (45.8)	122 (40.3)	166 (50.9)
Previous immunomodulatory treatment,* n (%)	556 (88.4)	263 (86.8)	293 (89.2)
Relapse in the preceding year, n (%)			
≥1 relapse	526 (83.6)	233 (76.9)	293 (89.9)
≥2 relapses	264 (42.0)	92 (30.4)	172 (52.8)

[1] Barbin et al. Neurology (2016)

Confounders in observational studies

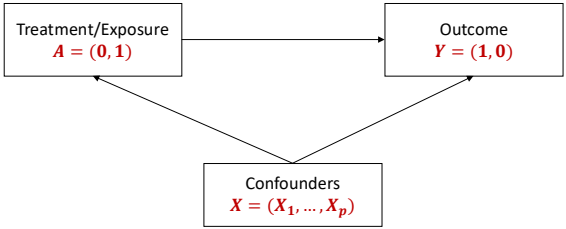
Baseline characteristics	All patients	Fingolimod	Natalizumab
No. of patients considered for analysis	629	303	326
Female, n (%)	479 (76.2)	225 (74.3)	254 (77.9)
Age at treatment initiation, y, mean (SD)	37.0 (9.6)	37.2 (9.2)	36.8 (9.9)
Disease duration at treatment initiation, y, mean (SD)	8.5 (6.4)	9.0 (6.8)	8.0 (6.1)
EDSS score, mean (SD)	2.6 (1.3)	2.4 (1.3)	2.8 (1.3)
EDSS score between 3.0 and 5.5, n (%)	288 (45.8)	122 (40.3)	166 (50.9)
Previous immunomodulatory treatment,[a] n (%)	556 (88.4)	263 (86.8)	293 (89.2)
Relapse in the preceding year, n (%)			
≥1 relapse	526 (83.6)	233 (76.9)	293 (89.9)
≥2 relapses	264 (42.0)	92 (30.4)	172 (52.8)

Endpoints	All patients	Fingolimod	Natalizumab
Clinical			
No. of patients considered for relapse analysis	629	303	326
At least one relapse, n (%)	151 (24.0)	82 (27.1)	69 (21.2)

The greater efficacy of the Natalizumab may be underestimated

Context of this session

- Binary exposure and binary outcome.



- From the observations (y_i, a_i, x_i), with $i = 1, ..., N.$, we aim to estimate the average treatment effect (no mediation):

$$\text{ATE} = \mathbb{E}_X[\Pr(Y = 1 \mid A = 1)] - \mathbb{E}_X[\Pr(Y = 1 \mid A = 0)]$$

The ATE in our use case

Baseline characteristics	All patients	Fingolimod	Natalizumab
No. of patients considered for analysis	629	303	326
Female, n (%)	479 (76.2)	479 (76.2)	479 (76.2)
Age at treatment initiation, y, mean (SD)	37.0 (9.6)	37.0 (9.6)	37.0 (9.6)
Disease duration at treatment initiation, y, mean (SD)	8.5 (6.4)	8.5 (6.4)	8.5 (6.4)
EDSS score, mean (SD)	2.6 (1.3)	2.6 (1.3)	2.6 (1.3)
EDSS score between 3.0 and 5.5, n (%)	288 (45.8)	288 (45.8)	288 (45.8)
Previous immunomodulatory treatment,[a] n (%)	556 (88.4)	556 (88.4)	556 (88.4)
Relapse in the preceding year, n (%)			
≥1 relapse	526 (83.6)	526 (83.6)	526 (83.6)
≥2 relapses	264 (42.0)	264 (42.0)	264 (42.0)

Plan

Introduction

Inverse probability weighting (IPW)

G-computation (GC)

Assumptions of causal inference

References

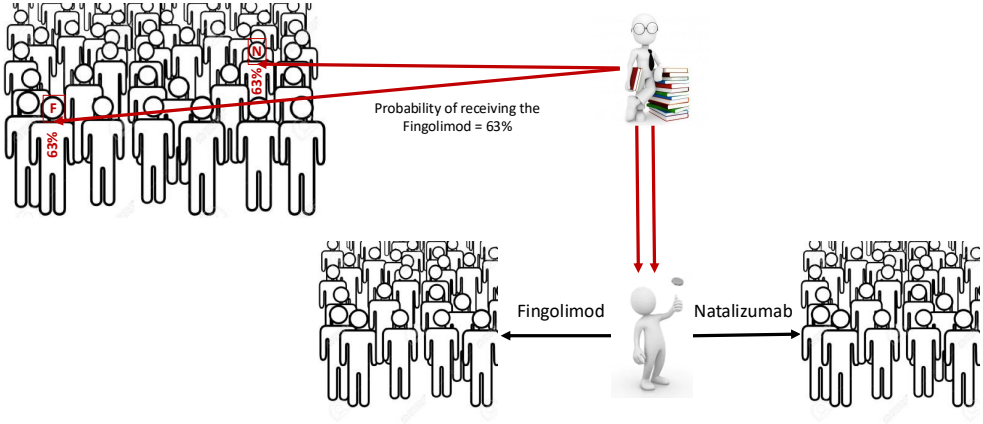
The definition of a propensity score (PS)

- The PS is the individual probability to be exposed:

$$\pi_i = \Pr(A = 1 \mid X_i)$$

- The related **exposure model** can be obtained from a logistic regression.
- Once estimated, two main methods are used:
  - Matching on the PS.
  - Weighting on the Inverse of the Treatment Probability (IPW).

The principle of matching on the PS



Three main limitations of matching on the PS[2]

- The targeted population cannot be defined a priori.
- The sample size is reduced due to non-matched individuals.
- The matched samples vary due to the random matching process.

[2] Chatton et al. Scientific report (2020)

The principle of weighting on the PS

Fingolimod ($A = 1$)				Natalizumab ($A = 0$)			
relapse +1y (Y)	relapse -1y (X)	relapse +1y (Y)	relapse -1y (X)	relapse +1y (Y)	relapse -1y (X)	relapse +1y (Y)	relapse -1y (X)
0	1			0	0		
0	1			1	0		
1	1			0	1		
0	0			0	0		
0	1			1	1		
0	0			0	0		
1	1			0	0		
0	1			0	0		
1	0			0	0		
0	1			0	0		
30%	70%			20%	20%		
Real world		Conterfactual world		Real world		Conterfactual world	

The principle of weighting on the PS

The PS represents the individual probability of receiving Fingolimod.

$$P(A = 1|X)$$

Fingolimod				Natalizumab			
relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y
0	1			0	0		
0	1			1	0		
1	1			0	1		
0	0			0	0		
0	1			1	1		
0	0			0	0		
1	1			0	0		
0	1			0	0		
1	0			0	0		
0	1			0	0		

The principle of weighting on the PS

The PS represents the individual probability of receiving Fingolimod:

$$P(A = 1|X)$$

For patients with a relapse before the prescription, 7/9 received Fingolimod:

$$\text{PS} = P(A = 1|X = 1) = \mathbf{78\%}.$$

Fingolimod				Natalizumab			
relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y
0	**1**			0	0		
0	**1**			1	0		
1	**1**			0	**1**		
0	0			0	0		
0	**1**			1	**1**		
0	0			0	0		
1	**1**			0	0		
0	**1**			0	0		
1	0			0	0		
0	**1**			0	0		
	7				**2**		

The principle of weighting on the PS

The PS represents the individual probability of receiving Fingolimod:

$$P(A = 1|X)$$

For patients with a relapse before the prescription, 7/9 received Fingolimod:

$$\text{PS} = P(A = 1|X = 1) = 78\%.$$

For patients without relapse before the prescription, 3/11 received Fingolimod:

$$\text{PS} = P(A = 1|X = 0) = 27\%.$$

Fingolimod				Natalizumab			
relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y
0	1			0	**0**		
0	1			1	**0**		
1	1			0	1		
0	**0**			0	**0**		
0	1			1	1		
0	**0**			0	**0**		
1	1			0	**0**		
0	1			0	**0**		
1	**0**			0	**0**		
0	1			0	**0**		
	3				**8**		

The principle of weighting on the PS

For patients with a relapse before the prescription, 7/9 received Fingolimod:

$$\mathrm{PS} = P(A = 1|X = 1) = 78\%.$$

We aim to give less weight to patients who have had a relapse prior to the prescription of Fingolimod.

	Weight	Fingolimod probability
Real world	1	78%
Contrefactual world	W	50% (10/20)

Fingolimod				Natalizumab			
relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y
0	1		1*0,641	0	0		
0	1		1*0,641	1	0		
1	1		1*0,641	0	1		
0	0			0	0		
0	1		1*0,641	1	1		
0	0			0	0		
1	1		1*0,641	0	0		
0	1		1*0,641	0	0		
1	0			0	0		
0	1		1*0,641	0	0		

W = 1 * 0,5 / 0,78 = 0,641

The principle of weighting on the PS

For patients without relapse before the prescription, 3/11 received Fingolimod:

$$\mathrm{PS} = P(A = 1|X = 0) = 27\%.$$

We aim to give higher weight to patients who have no relapse prior to the prescription of Fingolimod.

	Weight	Fingolimod probability
Real world	1	27%
Contrefactual world	W	50% (10/20)

W = 1 * 0,5 / 0,27 = 1,852

Fingolimod				Natalizumab			
relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y
0	1		1*0,641	0	0		
0	1		1*0,641	1	0		
1	1		1*0,641	0	1		
0	0		0*1,852	0	0		
0	1		1*0,641	1	1		
0	0		0*1,852	0	0		
1	1		1*0,641	0	0		
0	1		1*0,641	0	0		
1	0		0*1,852	0	0		
0	1		1*0,641	0	0		

The principle of weighting on the PS

For patients with a relapse before the prescription, 2/9 received Natalizumab:

$$\mathrm{PS} = P(A = 0|X = 1) = 22\%.$$

We aim to give higher weight to patients who have a relapse prior to the prescription of Natalizumab.

	Weight	Natalizumab probability
Real world	1	22%
Contrefactual world	W	50% (10/20)

W = 1 * 0,5 / 0,22 = 2,273

Fingolimod				Natalizumab			
relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y
0	1		1*0,641	0	0		
0	1		1*0,641	1	0		
1	1		1*0,641	0	1		1*2,273
0	0		0*1,852	0	0		
0	1		1*0,641	1	1		1*2,273
0	0		0*1,852	0	0		
1	1		1*0,641	0	0		
0	1		1*0,641	0	0		
1	0		0*1,852	0	0		
0	1		1*0,641	0	0		

The principle of weighting on the PS

For patients without relapse before the prescription, 8/11 received Natalizumab:

$$\text{PS} = P(A = 0|X = 0) = 73\%$$

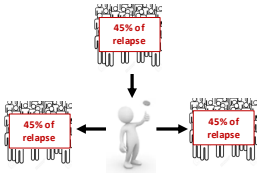
We aim to give lower weight to patients who have a relapse prior to the prescription of Natalizumab.

	Weight	Natalizumab probability
Real world	1	73%
Contrefactual world	W	50% (10/20)

W = 1 * 0,5 / 0,73 = 0,689

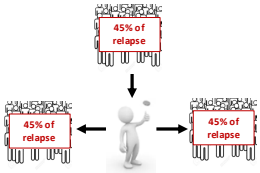
Fingolimod				Natalizumab			
relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y
0	1		1*0.641	0	0		0*0.689
0	1		1*0.641	1	0		0*0.689
1	1		1*0.641	0	1		1*2.273
0	0		0*1.852	0	0		0*0.689
0	1		1*0.641	1	1		1*2.273
0	0		0*1.852	0	0		0*0.689
1	1		1*0.641	0	0		0*0.689
0	1		1*0.641	0	0		0*0.689
1	0		0*1.852	0	0		0*0.689
0	1		1*0.641	0	0		0*0.689

The principle of weighting on the PS



Fingolimod				Natalizumab			
relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y
0	1		1*0,641	0	0		0*0,689
0	1		1*0,641	1	0		0*0,689
1	1		1*0,641	0	1		1*2,273
0	0		0*1,852	0	0		0*0,689
0	1		1*0,641	1	1		1*2,273
0	0		0*1,852	0	0		0*0,689
1	1		1*0,641	0	0		0*0,689
0	1		1*0,641	0	0		0*0,689
1	0		0*1,852	0	0		0*0,689
0	1		1*0,641	0	0		0*0,689
	70%		44,7%		20%		45,3%

The principle of weighting on the PS



Fingolimod				Natalizumab			
relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y	relapse +1y	relapse -1y
0	1	0*0,641	1*0,641	0	0	0*0,689	0*0,689
0	1	0*0,641	1*0,641	1	0	1*0,689	0*0,689
1	1	1*0,641	1*0,641	0	1	0*2,273	1*2,273
0	0	0*1,852	0*1,852	0	0	0*0,689	0*0,689
0	1	0*0,641	1*0,641	1	1	1*2,273	1*2,273
0	0	0*1,852	0*1,852	0	0	0*0,689	0*0,689
1	1	1*0,641	1*0,641	0	0	0*0,689	0*0,689
0	1	0*0,641	1*0,641	0	0	0*0,689	0*0,689
1	0	1*1,852	0*1,852	0	0	0*0,689	0*0,689
0	1	0*0,641	1*0,641	0	0	0*0,689	0*0,689
30%	70%	31,2%	44,7%	20%	20%	29,5%	45,3%

$$\text{ATE} = 31.2 - 29.5 = 1.7\%$$

Main steps to estimate ATE by using IPW (1)

- Construct the exposure model, for instance a logistic regression:

$$\text{logit}(Pr(A = 1 \mid X)) = \beta_0 + \beta_1 X_1 + ... + \beta_p X_p$$

- Predict the probability of being exposed according to the individual characteristics:

$$\hat{\pi}_i = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + ... + \hat{\beta}_p X_{pi})}{(1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + ... + \hat{\beta}_p X_{pi}))}$$

- Transform the predictions into individual (stabilized) weights:[3]

$$\hat{\omega}_i = \mathbb{1}\{A_i = 1\} \frac{\Pr(A_i = 1)}{\hat{\pi}_i} + \mathbb{1}\{A_i = 0\} \frac{1 - \Pr(A_i = 1)}{1 - \hat{\pi}_i}$$

[3] Austin and Stuart. Statistics in medicine (2015)

Main steps to estimate ATE by using IPW (2)

- Estimate the ATE in the weighted sample, for instance an univariate logistic regression obtained by maximising the weighted log-likelihood:

$$\text{logit}(Pr(Y = 1 \mid A)) = \gamma_0 + \gamma_1 A$$

- Compute the ATE:

$$\text{ATE} = \mathbb{E}_X[\Pr(Y = 1 \mid A = 1)] - \mathbb{E}_X[\Pr(Y = 1 \mid A = 0)]$$

$$\widehat{ATE} = \frac{\exp(\hat{\gamma}_0 + \hat{\gamma}_1)}{(1 + \exp(\hat{\gamma}_0 + \hat{\gamma}_1))} - \frac{\exp(\hat{\gamma}_0)}{(1 + \exp(\hat{\gamma}_0)}$$

- Bootstrap all the previous steps to obtain the 95% confidence intervals.

ATE by using IPW and R software

```
# The exposure model
glm1 <- glm(A~X1+... +Xp, family=binomial(link="logit"))

# Individual predicted probabilities
pi <- glm1$fitted.values

# Weights
pA <- mean(A)
w <- (A==1)*pA/pi + (A==0)*(1-pA)/(1-pi)

# ATE estimation
glm2 <- glm(Y~A,  weights = w, family=binomial(link = "logit"))
p0 <- exp(glm2[1,1])/(1+exp(glm2[1,1]))
p1 <- exp(glm2[1,1]+ glm2[2,1])/(1+exp(glm2[1,1]+glm2[2,1]))
ATE <- p1 - p0
```

The average treatment effect on the treated (ATT) as an alternative

Baseline characteristics	All patients	Fingolimod	Natalizumab
No. of patients considered for analysis	629	303	326
Female, n (%)	479 (76.2)	225 (74.3)	254 (77.9)
Age at treatment initiation, y, mean (SD)	37.0 (9.6)	37.2 (9.2)	36.8 (9.9)
Disease duration at treatment initiation, y, mean (SD)	8.5 (6.4)	9.0 (6.8)	8.0 (6.1)
EDSS score, mean (SD)	2.6 (1.3)	2.4 (1.3)	2.8 (1.3)
EDSS score between 3.0 and 5.5, n (%)	288 (45.8)	122 (40.3)	166 (50.9)
Previous immunomodulatory treatment,[a] n (%)	556 (88.4)	263 (86.8)	293 (89.2)
Relapse in the preceding year, n (%)			
≥1 relapse	526 (83.6)	233 (76.9)	293 (89.9)
≥2 relapses	264 (42.0)	92 (30.4)	172 (52.8)

ATE

Baseline characteristics	All patients	Fingolimod	Natalizumab
No. of patients considered for analysis	629	303	326
Female, n (%)	479 (76.2)	479 (76.2)	479 (76.2)
Age at treatment initiation, y, mean (SD)	37.0 (9.6)	37.0 (9.6)	37.0 (9.6)
Disease duration at treatment initiation, y, mean (SD)	8.5 (6.4)	8.5 (6.4)	8.5 (6.4)
EDSS score, mean (SD)	2.6 (1.3)	2.6 (1.3)	2.6 (1.3)
EDSS score between 3.0 and 5.5, n (%)	288 (45.8)	288 (45.8)	288 (45.8)
Previous immunomodulatory treatment,[a] n (%)	556 (88.4)	556 (88.4)	556 (88.4)
Relapse in the preceding year, n (%)			
≥1 relapse	526 (83.6)	526 (83.6)	526 (83.6)
≥2 relapses	264 (42.0)	264 (42.0)	264 (42.0)

ATT

Baseline characteristics	All patients	Fingolimod	Natalizumab
No. of patients considered for analysis	629	303	326
Female, n (%)	225 (74.3)	225 (74.3)	225 (74.3)
Age at treatment initiation, y, mean (SD)	37.2 (9.2)	37.2 (9.2)	37.2 (9.2)
Disease duration at treatment initiation, y, mean (SD)	9.0 (6.8)	9.0 (6.8)	9.0 (6.8)
EDSS score, mean (SD)	2.4 (1.3)	2.4 (1.3)	2.4 (1.3)
EDSS score between 3.0 and 5.5, n (%)	122 (40.3)	122 (40.3)	122 (40.3)
Previous immunomodulatory treatment,[a] n (%)	263 (86.8)	263 (86.8)	263 (86.8)
Relapse in the preceding year, n (%)			
≥1 relapse	233 (76.9)	233 (76.9)	233 (76.9)
≥2 relapses	92 (30.4)	92 (30.4)	92 (30.4)

Plan

Introduction

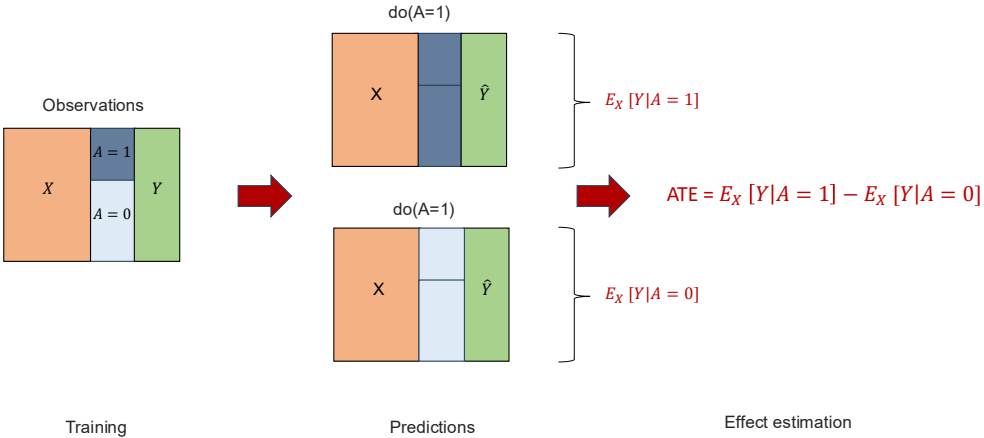
Inverse probability weighting (IPW)

G-computation (GC)

Assumptions of causal inference

References

The principle of the G-computation (GC) for ATE estimation



Main steps to estimate ATE by using GC (1)

- Construct the outcome model, for instance a logistic regression:

$$\text{logit}(Pr(Y = 1 \mid X, A)) = \beta_0 + \beta_1 X_1 + ... + \beta_p X_p + \beta_{p+1} A$$

- Predict the individual probabilities of event if all the subject are exposed:

$$\hat{\pi}_{1i} = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + ... + \hat{\beta}_p X_{pi} + \hat{\beta}_{p+1})}{(1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + ... + \hat{\beta}_p X_{pi} + \hat{\beta}_{p+1}))}$$

- Predict the individual probabilities of event if all the subject are unexposed:

$$\hat{\pi}_{0i} = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + ... + \hat{\beta}_p X_{pi})}{(1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + ... + \hat{\beta}_p X_{pi}))}$$

Main steps to estimate ATE by using GC (2)

- Compute the ATE:

$$\widehat{ATE} = n^{-1} \sum_i \hat{\pi}_{1i} - n^{-1} \sum_i \hat{\pi}_{0i}$$

- Bootstrap all the previous steps to obtain the 95% confidence intervals.

ATE by using GC and R software

```
# The outcome model
glm1 <- glm(Y~X1+...+Xp+A, family=binomial(link="logit"), data=mydata)

# The conterfactual prediction
mydata0 <- mydata1 <- mydata
mydata0$A <- 0
mydata1$A <- 1
p0i <- predict(glm1,  newdata=mydata0,  type="response"))
p1i <- predict(glm1,  newdata=mydata1,  type="response"))

# The ATE estimation
p0 <-mean(p0i)
p1 <- mean(p1i)
ATE <- p1 - p0
```

Plan

Introduction

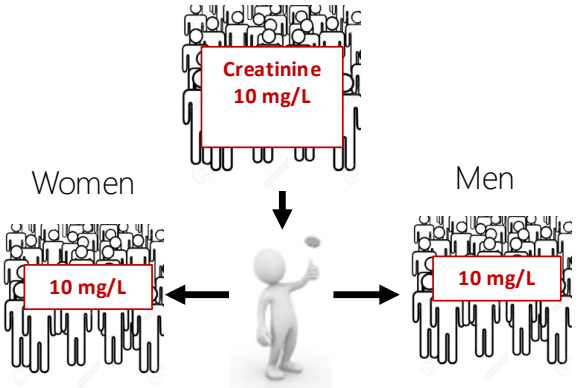
Inverse probability weighting (IPW)

G-computation (GC)

Assumptions of causal inference

References

The exposure should be modifiable



The correct specification of the working models

- Do not categorize continuous explanatory variables X.
- Respect the assumptions of the model (log-linearity in logistic regression).
- Consider the interaction when necessary.
- AUC maximization does not represent an objective.
- Standardized mean differences should not exceed 10%. [4]
  - For categorical covariate, SMD is defined by:

$$SMD = \left(\hat{p}_{T=1} - \hat{p}_{A=0}\right) \Big/ \left(\sqrt{(\hat{p}_{A=1}(1-\hat{p}_{A=1}) + \hat{p}_{A=0}(1-\hat{p}_{A=0}))/2}\right)$$

  - For continuous covariates, SMD is defined by:

$$SMD = \left(\bar{X}_{A=1} - \bar{X}_{A=0}\right) \Big/ \left(\sqrt{(s_{A=1}(X)^2 + s_{A=0}(X)^2)/2}\right)$$

[4] Austin PC. Communications in Statistics - Simulation and Computation (2009)

No missing confounders

- Draw a directed acyclic graph (DAG) before analyses.
- Sensitivity analysis to evaluate the potential impact of missing confounders (ex: E-value).

Positivity assumption

- Draw the histograms of the propensity scores in the two groups.
- Distinguish theoretical violation and near violation (i.e. by chance). [5]
  - Inclusion criteria should be updated in case of theoretical violation.
  - GC is more robust to near-violation.

[5] Léger et al. Biometrical Journal (2022)

Consistency and no interference

- Consistency: well-defined expositions/treatment.
- No interference: whether one individual receives treatment (or not) has no effect on the potential outcomes of any other individual.

Plan

Introduction

Inverse probability weighting (IPW)

G-computation (GC)

Assumptions of causal inference

References

References

1. Barbin et al. Comparative efficacy of fingolimod vs natalizumab: A French multicenter observational study. Neurology. 2016 Feb 23;86(8):771-8.
2. Chatton et al. G-computation, propensity score-based methods, and targeted maximum likelihood estimator for causal inference with different covariates sets: a comparative simulation study. Sci Rep. 2020 Jun 8;10(1):9219.
3. Austin and Stuart. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med. 2015 Dec 10;34(28):3661-79.
4. Austin,. Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research. Communications in Statistics - Simulation and Computation. 2009. 38(6), 1228–1234.
5. Léger et al. Causal inference in case of near-violation of positivity: comparison of methods. Biom J. 2022 Dec;64(8):1389-1403.