



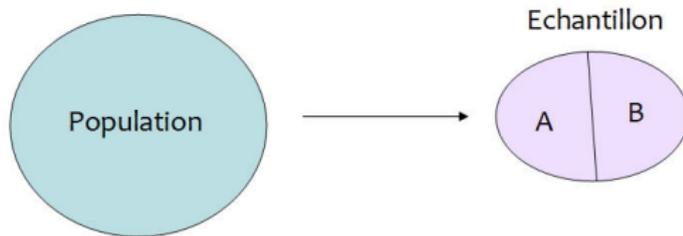
Les scores de propension



Clémence Leyrat

INSERM CIC 202 - CHRU de Tours

Objectif : comparer l'effet d'une intervention A à une intervention B sur un échantillon de patients



Les groupes de l'étude doivent être **comparables** pour que l'effet estimé du traitement ou de l'intervention étudié soit sans biais :

Les groupes ne diffèrent-ils que par l'intervention reçue ?

Essais randomisés vs. études observationnelles

Scores de
propension

La **randomisation** permet d'équilibrer les variables observées et non observées entre les groupes

Contexte
Inférence et
biais
Ajustement

Le score de
propension

Analyse par
SP

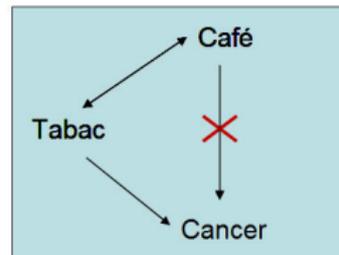
Conclusion

Cependant, d'autres sources d'informations que celles issues d'essais randomisés peuvent être utiles :

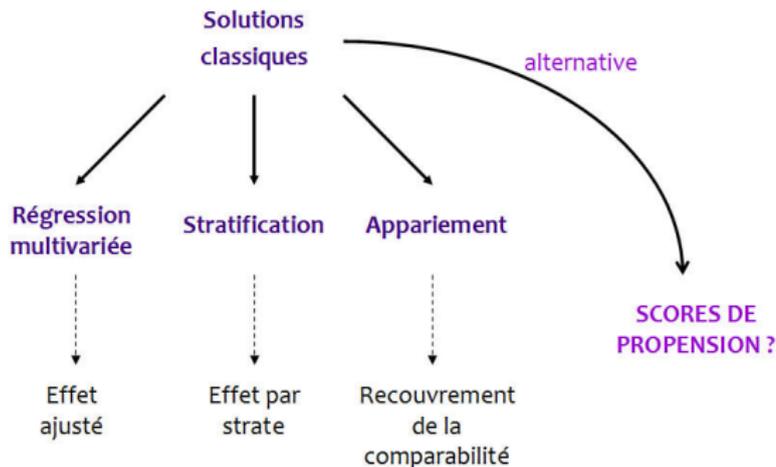
- données de registres
- études observationnelles ...

Dans ces cas, les groupes que l'on veut comparer **peuvent être différents** (biais de confusion)

Exemple : cohorte pour étudier le lien entre la consommation de café et le risque de cancer



Si le déséquilibre n'est pas pris en compte \Rightarrow **estimation biaisée de l'effet de intervention**
(comparaison de moyennes, de proportions brutes)



Les scores de propension

Scores de propension

Contexte

Le score de propension

Définition

Estimation

Exemple

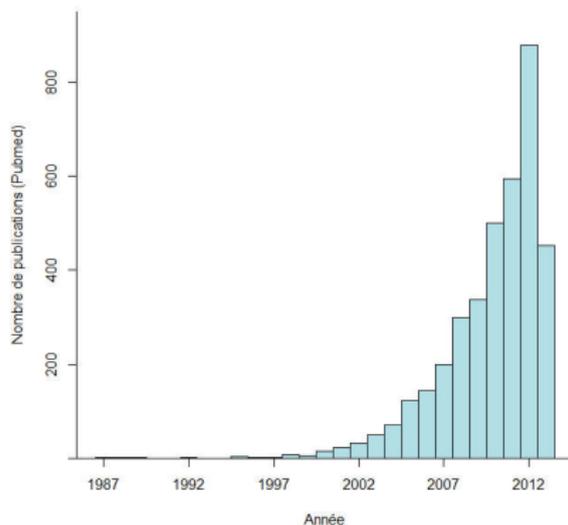
Analyse par SP

Conclusion

Méthode proposée en 1983 par Rosenbaum et Rubin pour pallier les **biais de sélection dans les études observationnelles**

But : **mimer un essai randomisé** en recouvrant l'équilibre entre les groupes

Méthode en pleine expansion depuis 20 ans (Pubmed)



Qu'est-ce que le score de propension ?

Définition

Le score de propension (SP) est la probabilité qu'a un patient de recevoir un traitement plutôt qu'un autre conditionnellement à ses caractéristiques :

$$e(x) = P(Z = 1|\mathbb{X}),$$

où Z désigne le traitement et \mathbb{X} le vecteur de covariables

Le score de propension est estimé puis utilisé dans l'analyse pour corriger le biais : **méthode en deux temps**



La vraie valeur du SP est **inconnue** mais le SP peut être estimé :

- ↔ si l'assignement au traitement dépend des variables mesurées
- ↔ le plus souvent par un modèle logistique

Variable dépendante : intervention (traitement vs. placebo par ex.) ou exposition

Variables explicatives : variables supposées associées à l'allocation au traitement ou à l'exposition

Les SP individuels sont les **prédictions** issues de ce modèle

9104 patients ayant subi un infarctus du myocarde : prescription ou non de statine à la sortie d'hospitalisation.

Estimation du SP à être traité
ou non selon 27 covariables
(et 229 interactions)

⇒ Biais : groupes non comparables

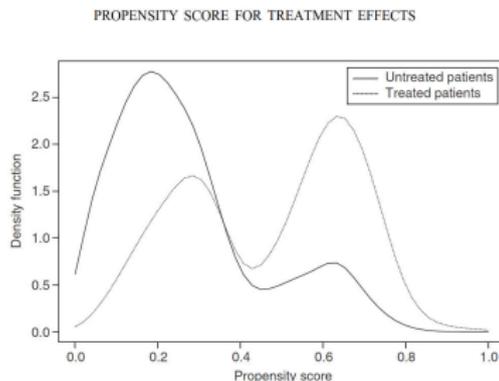


Figure 1. Distribution of propensity score in treated/untreated patients.

Austin, P, Mamdani, M. A Comparison of Propensity Score Methods : A Case-Study Estimating the Effectiveness of Post-AMI Statin Use. **Statistics in Medicine**. 2006. ; 25 :2084-2106.

Le SP permet de mettre en évidence le biais mais également de **corriger le biais** au moment de l'analyse

Il permet de **résumer** une quantité importante d'information en un seul score :

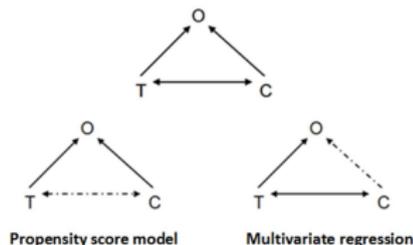
- intérêt quand beaucoup de variables et peu d'événements
- effectif trop peu important pour ajuster

Approche différente de la régression ajustée classique :

C : covariable

O : critère de jugement (outcome)

T : traitement (exposition)



Ajustement direct sur le SP :

- le SP est utilisé comme une variable d'ajustement dans un modèle de régression (linéaire, logistique, de Cox, mixte...)
- simple à mettre en oeuvre et interpréter
- résultats proches d'une régression multivariée

Pondération par l'inverse du SP :

- les patients sont pondérés pour retrouver une représentativité de l'échantillon
- Poids = $\frac{1}{e(x)}$ chez les traités (exposés), $\frac{1}{1-e(x)}$ chez les non traités
- les patients ayant un poids de 0 ne sont pas inclus

Stratification selon les quintiles du SP :

- création de 5 strates selon les quintiles du SP
- une analyse par strate puis combinaison des résultats
- utilise l'ensemble de l'information mais peut garder un biais résiduel

Appariement sur le SP :

- les patients ayant des valeurs proches du SP sont appariés
- bonne capacité à réduire le biais
- tous les patients ne peuvent pas être appariés : perte d'effectif

Critère de jugement : mortalité à 3 ans

Après vérification que le SP permette de recouvrer l'équilibre entre les groupes :

Table V. Treatment effect for Statin at discharge—adjusted estimates.

Method	Odds ratio	95 per cent confidence interval	P-Value
Crude	0.49	(0.44, 0.55)	<0.0001
<i>Propensity score methods</i>			
Stratifying on PS quintiles	0.77	(0.66, 0.89)	0.0003
Stratifying on PS quintiles—within stratum regression adjustment	0.76	(0.65, 0.89)	0.0007
Covariate adjustment—linear term for PS	0.84	(0.74, 0.96)	0.0099
Covariate adjustment—quadratic terms for PS	0.81	(0.70, 0.93)	0.0033
Covariate adjustment—cubic spline for PS	0.81	(0.70, 0.93)	0.0028
Covariate adjustment—linear term for PS with adjustment for additional confounders	0.80	(0.69, 0.93)	0.0038
PS Matching	0.85	(0.72, 0.99)	0.0372
PS weighting—simple model	0.77	(0.64, 0.92)	0.0041
PS weighting—complex model	0.76	(0.63, 0.90)	0.0022
<i>Direct regression adjustment</i>			
Regression adjustment—simple predictive model	0.75	(0.65, 0.85)	<0.0001
Regression adjustment—backwards selection method	0.73	(0.64, 0.84)	<0.0001
Regression adjustment—complex predictive model	0.78	(0.67, 0.91)	0.0014

Scores de
propension

Contexte

Le score de
propensionAnalyse par
SPIntérêt
Méthodes
Exemple
Limites

Conclusion

Méthodes **conservatrices** : la variabilité liée à l'estimation du SP n'est pas prise en compte dans l'analyse

Problème de la **sélection des variables** : facteurs de confusion uniquement, mais pas toujours identifiables *a priori*

Problème de **convergence** lors de l'estimation du SP s'il y a trop de variables par rapport à l'effectif

Les SP permettent de détecter le biais et sont une alternative à la régression classique pour l'analyse

Régression classique : estimation conditionnelle

Scores de propension : estimation conditionnelle ou causale dans certains cas

⇒ Choix de la méthode **guidée par la problématique**