

# Le score de propension, une technique de contrôle de biais dans les études observationnelles

Bachirou Taddé, Biostatistician

28/09/2020

# Section 1

## Rappel

# Les différents types de covariables

- **Variable dépendante** (événement) : variable d'intérêt (Y)
- **Variable indépendante** (exposition) : variable explicative de la variable dépendante (X)
- **Variable médiatrice** (intermédiaire) : variable situé sur le chemin causal X vers Y
- **Facteur de confusion** : Variable liée à la fois à la variable indépendante et à la variable dépendante
- **Variable instrumentale**
- **Variable liée uniquement à l'événement** (Moulis and Lapeyre-Mestre 2018)

# Les différents biais

Un **biais** est une **Erreur de raisonnement ou de procédure** qui amène à une représentation fausse de la réalité

- **Biais de sélection** : erreur de procédure faite lors de la sélection des sujets de l'étude
- **Biais d'information** : erreur de procédure due à une mauvaise mesure des covariables
- **Biais d'indication** (ou biais d'évaluation, erreur de raisonnement): une erreur systématique due au fait que la mesure du critère de jugement n'est pas réalisée de la même manière dans les groupes comparés
- **Biais d'interprétation** : une erreur systématique d'interprétation des résultats d'analyse

**NB** : Importance d'éviter ou de contrôler les biais.

## Section 2

### Differentes types d'études

# Différentes types d'étude(1/2)

- **Etude expérimentale (essai randomisée)**

- étude contrôlée qui permet de répondre à la question : « est-ce que l'exposition a un effet sur la survenue de l'événement en conditions expérimentales ? »
- méthodologie de référence la moins biaisée pour faire de l'inférence causale

- **Forces & faiblesses**

- tient compte des variables observables et non-observables
- conditions expérimentales loin de la vie réelle
- impossible à mettre en œuvre dans certaines situations (chirurgie ou maladie rare, etc.)

## Différentes types d'étude(2/2)

- **Etude observationnelle** : permet de répondre à la question : « quel est l'effet de l'exposition sur la survenue de l'événement en conditions réelles ? »
- **Forces & faiblesses**
  - Inférence causale dans des conditions complexes de vie réelle
  - plusieurs sources de biais
  - L'allocation du traitement non aléatoire peut entraîner un biais d'indication
  - nécessite une méthode supplémentaire de contrôle de biais
  - la non-prise en compte des facteurs de confusion non observables peut entraîner des biais résiduels

# Méthode de contrôle de biais dans une étude observationnelle

- **Modèle d'analyse multivarié**

- approche classique et facile à mettre en œuvre
- requiert des hypothèses de modélisation et des règles strictes pour une interprétation causale (DAG, (Pearl and others 2000))
- limitation du nombre de covariables sinon problème de sur-ajustement

- **Méthode du Score de propension** (Rosenbaum and Rubin 1984)

- maîtriser par une seule covariable (le score de propension) le biais de confusion
- exclus les variables non-observables (ou non-observée)
- Possible biais résiduel du fait de variables non-observables



## Section 3

### Score de propension

# Méthode du Score de propension

le score de propension de chaque individu est la probabilité d'être exposé conditionnellement à des covariables de l'étude :

$$sp = P(X = 1|S = s) \quad (1)$$

avec:

- **Y** : la variable « Evènement », la variable d'intérêt ou variable dépendante.
- **X** : la variable « Exposition », variable binaire 1 = exposé, 0 = non-exposé
- **C** : l'ensemble des autres covariables de l'étude
- **S** : un sous ensemble de covariables :  $S \subset C$

le jeu de données est complété par les SP des individus calculés.

# Création et validation du score de propension

## ● Méthode de création du SP

- Régression logistique
- Autre méthode de scoring normalisée (LDA, QDA, etc. ???)
- Choix inclusive ou exclusive des variables à inclure dans la création du score

## ● Validation du SP

- Le SP permet de restaurer une équiprobabilité entre les groupes comparés
- le SP est en soit une variable dérivée
- conduire à la construction d'un pseudo jeu de données équilibré
- vérifier que les groupes sont comparables au regard de la SP
- Vérifier que le SP restaure l'équiprobabilité d'exposition entre les groupes (représentations graphiques, distribution des scores dans chaque groupe).
- Vérifier que les caractéristiques des groupes sont les mêmes pour toutes les covariables dans le pseudo-jeu de données.

# Utilisation du Score de propension

Le SP peut être utilisé de plusieurs manières :

- par Ajustement dans un modèle de régression
- par appariement
- par pondération IPTW)
- par stratification De toutes ces méthodes, quelle est la meilleure?

# Conclusion

- le score de propension est une méthode utilisée dans les études observationnelle pour contrôler les biais pouvant entacher les conclusions causales.
- Facile à construire
- Important de le valider avant l'utilisation
- Plusieurs manières d'utiliser le score de propension
- Possible biais résiduel contrôlable par
  - Utilisation de la propensity score calibration (PSC)
  - Utilisation de témoins négatifs

## Références

Moulis, G, and M Lapeyre-Mestre. 2018. "Score de Propension: Intérêts, Utilisation et Limites. Un Guide Pratique Pour Le Clinicien." *La Revue de Médecine Interne* 39 (10): 805–12.

Pearl, Judea, and others. 2000. "Models, Reasoning and Inference." *Cambridge, UK: CambridgeUniversityPress*.

Rosenbaum, Paul R, and Donald B Rubin. 1984. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79 (387): 516–24.