

SFES Lyon, le 19 Avril 2012

*Enjeux et Méthodologie  
de L'évaluation Comparative des Politiques  
Sanitaires Publiques*

Pr Robert Launois

28, rue d'Assas

75006 Paris – France

Tel . 01 44 39 16 90 – Fax 01 44 39 16 92

E-mail : reesfrance@wanadoo.fr - Web : www.rees-france.com

# Plan de l'intervention

- ✦ La preuve de l'effet causal
- ✦ Les biais : un risque des études observationnelles
- ✦ Gérer les interférences parasites
- ✦ Faire parler les banques médico-administratives
- ✦ Choisir le bon schéma d'étude

# Les Enjeux

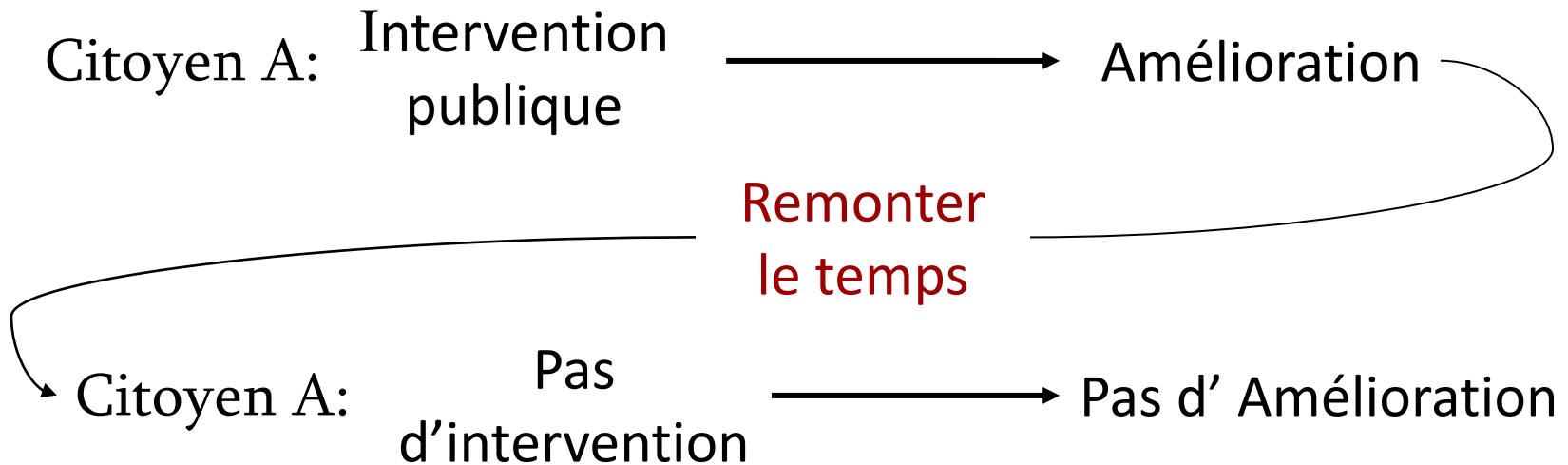
- ✦ **Produire des relations causales** non biaisées (validité interne) **entre intervention et résultats** est au cœur de toute démarche évaluative.
- ✦ le **programme a-t-il les effets attendus** sur les individus ou la population ciblée, les ménages, les institutions, les bénéficiaires du programme
- ✦ Comment **isoler les effets produits par un programme** de santé, des **effets produits par d'autres facteurs** concomitants ?
- ✦ On mène ces recherches, pour fournir aux décideurs politiques et aux acteurs sociaux **des arguments qui reposent sur des études empiriques**
- ✦ Ce qui **exige des données et un modèle statistique** capable de fournir un résumé : par exemple une différence de moyennes. Ce sont ces résumés que l'on interprète comme des résultats

# LA PREUVE DE L'EFFET CAUSAL

# Cause de l'Effet Ou Effet de la Cause ?

- ✦ **L'Approche structurelle** fait un lien précis **ex ante** entre le modèle formel où les comportements fondamentaux des agents sont modélisés et le modèle statistique estimé pour prédire l'impact de programme en projet
- ✦ **L'Approche en forme réduite** étudie **ex post** l'effet de deux variables d'intérêt (« *Les résultats potentiels* ») pour un individu celui où il bénéficie du TrT et celui où il n'en bénéficie pas, sans détailler les mécanismes intermédiaires, ni estimer les paramètres fondamentaux

# Principe du Raisonnement *Contrefactuel*



# Effet Causal Individuel du Traitement

- ✦ **Objectif** : évaluer pour chaque individu  $i$  le lien de cause à effet
  - Entre l'Intervention  $T_i$  (**traitement** , programme)
  - Et
  - Le résultat  $Y_i$  ( indice thérapeutique d'efficacité)
- ✦ Pour un bénéficiaire du programme doté de caractéristiques observables  $X_i$  l'effet causal ( $\Delta_i$ ) du traitement sur le résultat égale

$$\Delta_i = (Y_{i1} | X_i, T_i=1) - (Y_{i0} | X_i, T_i=0)$$

- On observe le niveau de santé atteint après intervention (noté  $T=1$ ) ( $Y_{i1} | X_i, T_i=1$ )
- Mais on observe pas le niveau de santé qui eut été atteint sans l'intervention (noté  $T=0$ ). ( $Y_{i0} | X_i, T_i=0$ )
- ✦  $Y_{i0}$  correspond au résultat qui aurait été obtenu si le sujet n'avait pas pas participé au programme, c'est un résultat contrefactuel parce que non observable. L'estimation de l'effet causal **individuel est impossible** : une seule des deux variables est observée

**le même sujet, ne peut pas au même instant, être bénéficiaire et non bénéficiaire d'un traitement ou d'une politique publique**

# Effet Causal Moyen Populationnel

- ✦ La situation contrefactuelle ne peut être estimée qu'en calculant la différence  $\Delta^{TT}$  entre les moyennes de la variable **Y** entre deux populations : celle des bénéficiaires et celle des non bénéficiaires
- ✦ Les différences entre bénéficiaires et non bénéficiaires ont deux sources
  - l'impact réel des politiques ou des traitements : ( $\Delta E^{TT}$ )
  - le biais de sélection : ( $B^{TT}$ ). Ce biais correspond à la différence qui serait enregistrée en l'absence de TrT entre ceux qui en bénéficient et les autres, **compte tenu des caractéristiques initiales**  $X_i$  des deux groupes
- ✦  $\Delta E^{TT} = E(Y_{i1} | X_i, T_i=1) - E(Y_{i0} | X_i, T_i=0)$   
 $\Delta E^{TT} = \underbrace{E(Y_{i1} | X_i, T_i=1) - E(Y_{i0} | X_i, T_i=0)}_{\Delta^{TT}} + \underbrace{E(Y_{i0} | X_i, T_i=0) - E(Y_{i0} | X_i, T_i=0)}_{B^{TT}}$   
 $\Delta E^{TT} = \Delta^{TT} - B^{TT}$
- ✦ Les mesures qui oublie le biais de sélection prennent à tort la mesure de la quantité  $\Delta^{TT}$  pour la mesure de  $\Delta E^{TT}$  Quant on oublie de soustraire le biais de sélection, on surestime « la quantité d'effet »



# Quantité d'Effet vs Association

- ✦ Les populations des groupes ainsi constitués sont elles comparables ?
- ✦ OUI par randomisation → Quantité d'effet mesurée par La réduction des risques absolus (RRA), La réduction risques relatifs (RRR), odds ratio (OR) etc...
- ✦ NON si le groupe de test et le groupe de contrôle ne sont pas au départ des substituts parfaits. **La situation moyenne des individus qui ont reçu le traitement n'aurait pas été la même en l'absence de traitement que celle des individus n'ayant pas reçu le traitement → association**

# Portée des Evaluations Expérimentales

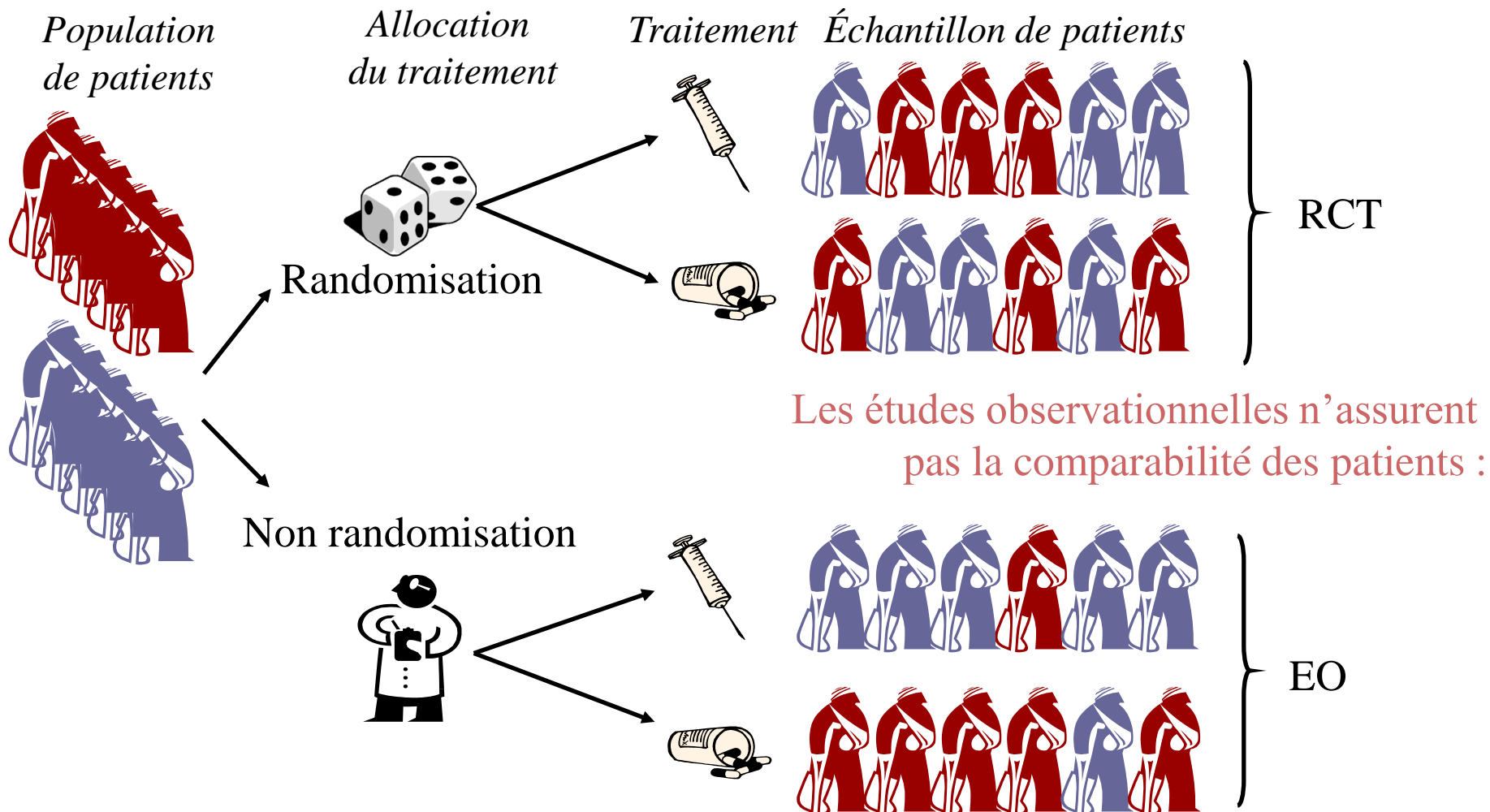
- ✦ Dans les essais randomisés en double insu et avec groupe contrôle, trois conditions permettent d'établir une relation causale:
  - **L'affectation aléatoire garantit la comparabilité des populations** entre les deux groupes, le groupe verum et le groupe placebo, en neutralisant les caractéristiques personnelles observables et non observables des patients inclus dans chacun d'entre eux,
  - **Le placebo garantit la comparabilité des effets** des traitements en évitant les interférences qui pourraient être associées à d'administration d'un autre médicament .
  - **le double insu garantit la comparabilité des comportements** : les effets blouse blanche, les effets patient, les effets structure sont neutralisés.
- ✦ Dans ces conditions les différences qui apparaissent entre les deux groupes, « toutes choses égales d'ailleurs », sont exclusivement dues au traitement. **Le rapport de cotes ne mesure pas une association mais exprime réellement une quantité d'effet**, le lien de cause à effet est démontré.

# LES BIAIS : UN RISQUE INHÉRENT À LA NATURE MÊME DES ETUDES OBSERVATIONNELLES

# Evaluations Non expérimentales, Biais et Facteurs de confusion

- ✦ Un biais, est une erreur systématique qui s'introduit dans une étude et qui tend à produire une estimation (incidence, prévalence, RR...) différant systématiquement de la vraie valeur et conduisant une interprétation erronée du résultat.
- ✦ Un facteur de confusion est un facteur parasite tiers agissant de façon concomitante qui biaise les résultats de l'étude parce que son association avec l'exposition (i.e le programme, le traitement, le fait générateur) et avec la survenue de l'événement fait faussement croire que le facteur d'exposition est associée à la survenue de l'événement.

# Le Biais de Recrutement



Biais de recrutement : les patients sont sélectionnés selon certaines caractéristiques corrélées à leur devenir.

# Le Biais d'Indication

- ✦ Les différences observées dans la survenue des événements peuvent ne pas être liées au choix thérapeutique **mais aux raisons pour lesquelles ce choix a été fait**
- ✦ Les traitements ne sont pas prescrits au hasard
  - Ils sont explicitement laissés au choix du médecin: l'indication est posée en raison des caractéristiques des patients ou sur la base de son jugement clinique
  - Ou implicitement sélectionnés sur la base d'habitudes de prescriptions ou du désir d'apporter une réponse à une souffrance psychologique ou sociale
- ✦ → Biais d'indication: on ne peut pas distinguer l'effet des caractéristiques propres des sujets et l'effet du traitement sur le critère de jugement

# Un Impératif : Travailler sur des Populations Comparables

- ✦ On peut légitimement imaginer dans le cadre d'une étude observationnelle que ceux qui bénéficient d'une prise en charge spécifique
  - présentent des caractéristiques particulières
  - qui sont autant de facteurs de risque,
  - Et que ces caractéristiques auraient joué en leur défaveur,
  - même s'ils n'avaient pas reçu le traitement.
- ✦ Si ce traitement s'avère néfaste, on doit s'attendre à ce qu'ils subissent de son fait un préjudice **encore plus grand** que celui qui eut été le leur si ce traitement ne leur avait pas été prescrit → Le **VRAI** préjudice **de santé est surestimé**.
- ✦ La tâche de l'évaluateur est donc de reconstituer un groupe contrôle à partir du groupe socialement constitué, en **sélectionnant des individus ressemblants autant que possible** à ceux du groupe auxquels le traitement a été administré.

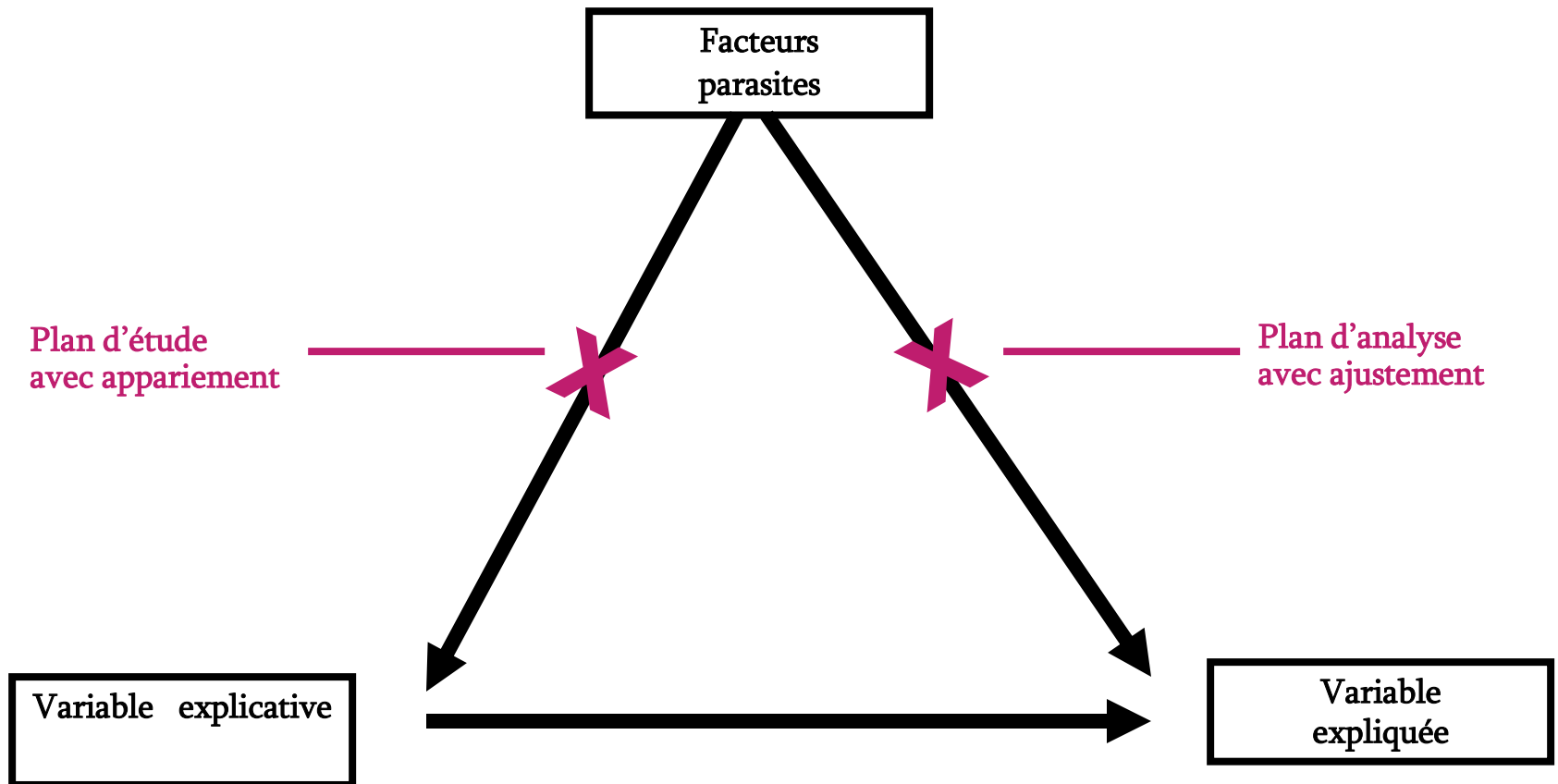
# GERER LES INTERFERENCES PARASITAIRES



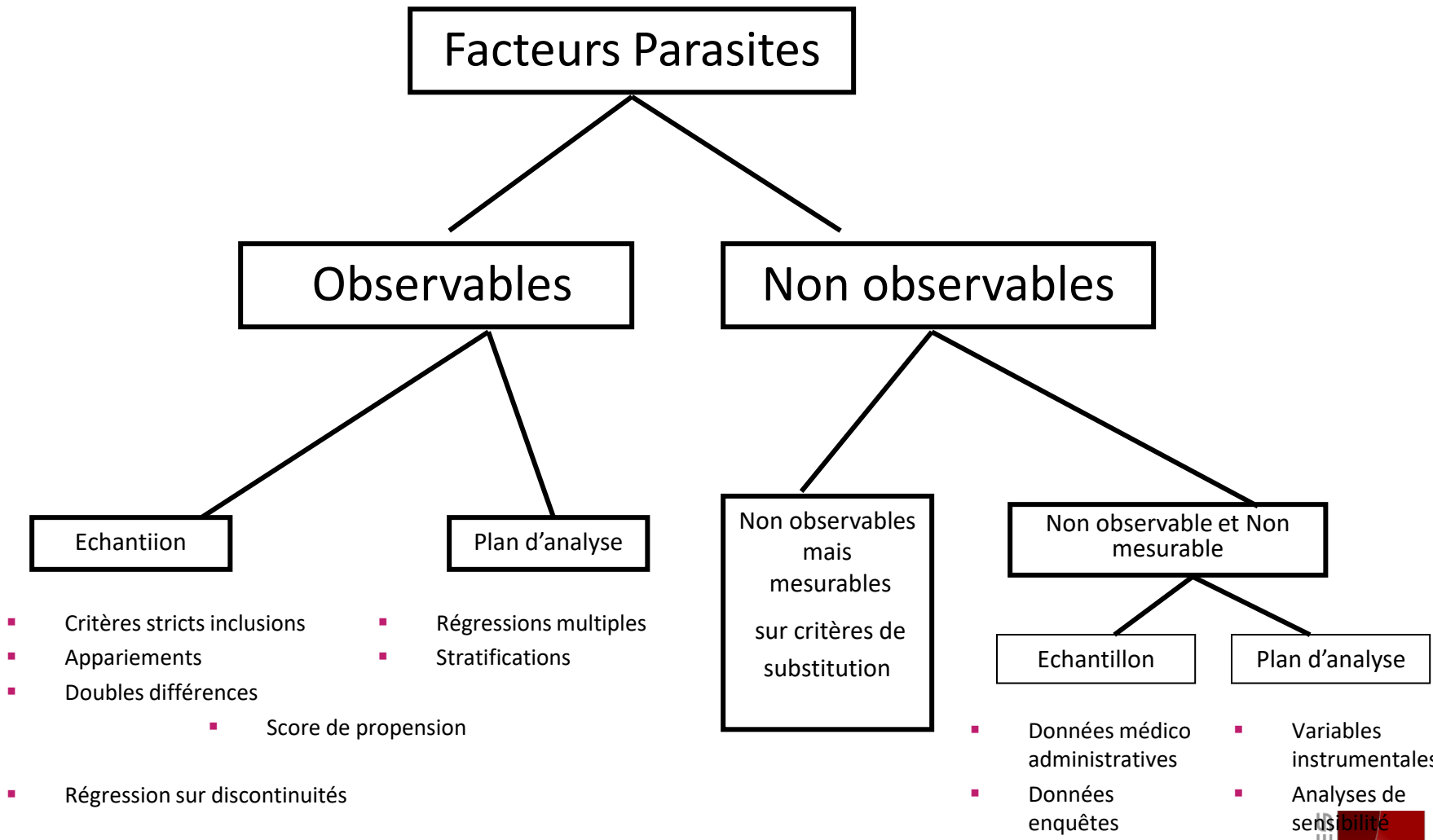
# Deux Approches Pour Circonscrire et Exclure « les Autres Causes Possibles du Dommage »(les Biais)

- ✦ Modéliser *a priori* la relation entre les caractéristiques observables du patient et l'exposition à un traitement (score de propension)
- ✦ Modéliser *a posteriori* la relation entre les caractéristiques observables du patient et la survenue observée des événements (analyse multivariée)

# Appariement *a priori* ou Ajustement *a posteriori*



# Une Grande Diversité d'Outils Disponibles



# L'Appariement Exhaustif Sur Les Caractéristiques Observables Est Impossible

- ✦ Pour les caractéristiques de chaque patient traité,
- ✦ Il faudrait trouver un patient non traité avec des caractéristiques identiques
- ✦ Si l'on suppose :
  - 10 tranches d'âge,
  - Sexe : masculin/féminin,
  - Antécédents diagnostiques: 2 modalités (oui/non), 5 antécédents,
  - Antécédents thérapeutiques: 2 modalités (oui/non), 5 antécédents
  - 5 niveaux de coût
- 102 400 appariements possibles

# Appariement sur la Base d'un Résumé Unidimensionnel des Caractéristiques : le Score de Propension

## ■ Définition :

- SP : Probabilité pour un sujet de bénéficier du traitement étudié, quel que soit le groupe dont il relève, en fonction des caractéristiques initiales qu'il présente =  $P(\text{traitement} \mid \text{caractéristiques du patient})$
- Si le score  $SP_{\text{Traitement}} = \text{le score } SP_{\text{Contrôle}} \rightarrow$  Les sujets traités et les contrôles auront en moyenne les mêmes caractéristiques initiales

## ■ Usage :

- **Échantillon** avec **appariement** sur le SP
    - Contrôle du biais de recrutement
    - Moins de patients si SP trop dissemblables  $\rightarrow$  perte de puissance
  - **Analyse** avec **ajustement** sur le SP
    - Moindre impact sur la taille de l'échantillon
    - Une seule variable d'ajustement
  - **Analyse** avec **stratification** sur le SP
- Le SP **simule un essai randomisé** en assurant la comparabilité des patients sur les variables **observées**.

# Opérationnalisation de la Méthode SP

1. Rassembler le plus grand nombre d'informations possibles sur les caractéristiques cliniques initiales des patients;
2. Modéliser la probabilité que les malades soient amenés à recevoir le traitement en fonction des caractéristiques initiales qu'ils présentent à l'aide du score de Propension (PS);

– PS = P(Traitement | Caractéristiques initiales)

– A l'aide d'une fonction logistique :

$$\ln\left(\frac{PS}{1-PS}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

3. Les groupes de patients dont les scores de propension sont similaires sont considérés comparables.
- ➔ Le SP **simule un essai randomisé** en assurant la comparabilité des patients sur les variables **observées**.

# Procédure d'Appariement



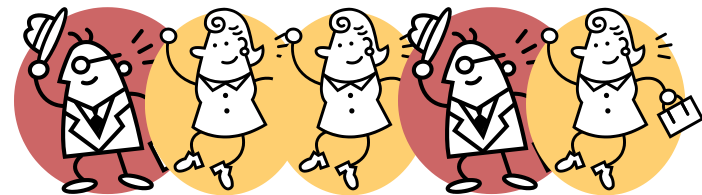
$$SP_H = P(T|Homme) = 4/6 = 0.66$$

$$SP_F = P(T|Femme) = 1/4 = 0.25$$

	T	C	
H	4	2	6
F	1	3	4

Groupe test : T

Groupe contrôle : C



SP = 0.25

SP = 0.66

SP = 0.66

**Appariement sur le SP**



# Mesurer la “Quantité” de Biais Avant et Après Appariement par le SP

- ✦ Inconvénients de la p-value:
  - Ne renseigne pas sur l’importance du déséquilibre
  - Dépend du modèle mis en oeuvre

- ✦ Différences standardisées :

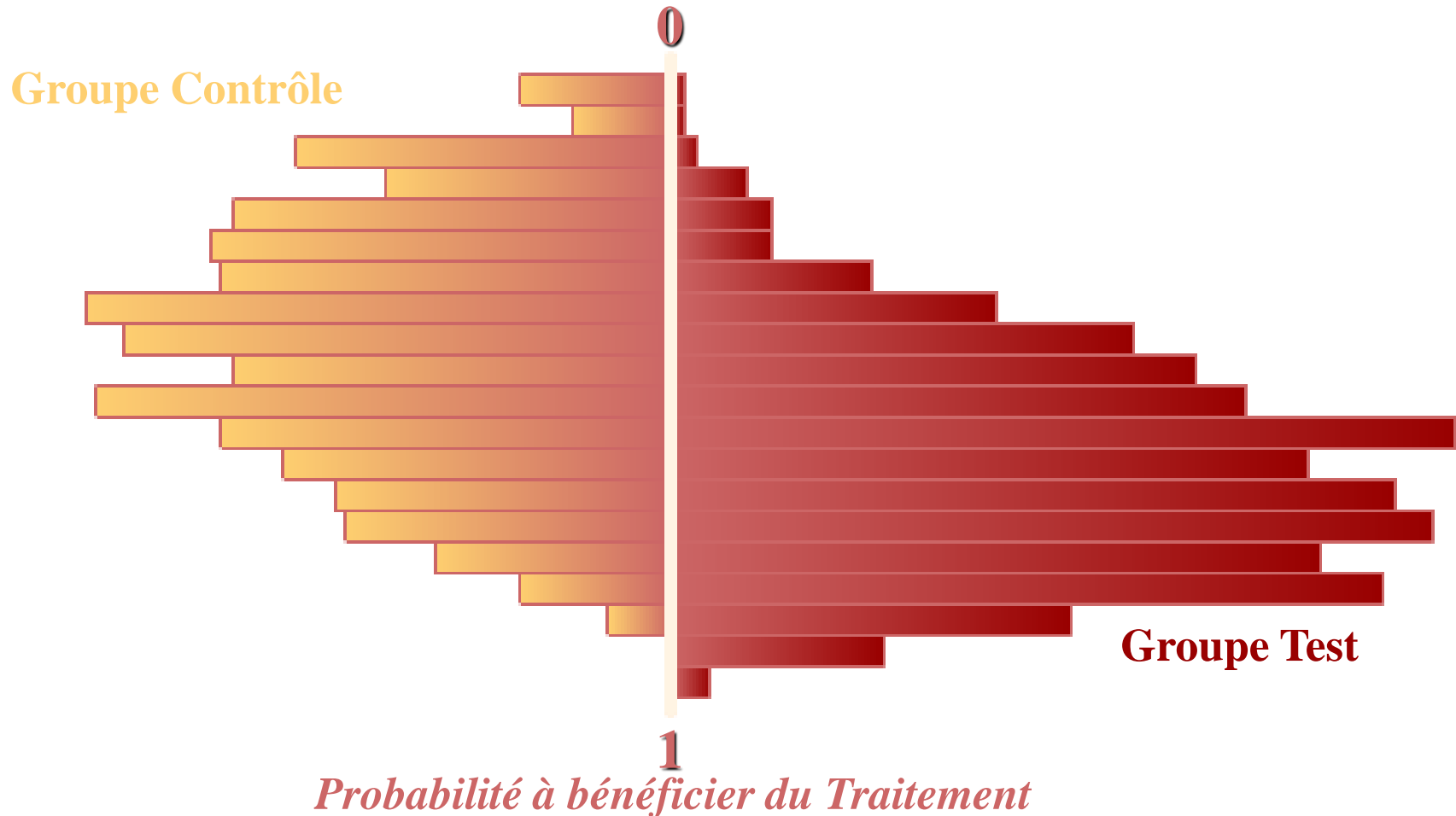
$$d = \frac{(\overline{x_{traitement}} - \overline{x_{contrôle}})}{\sqrt{\frac{s_{traitement}^2 + s_{contrôle}^2}{2}}}$$

- ✦  $|d| > 10\%$  → déséquilibre entre les deux bras



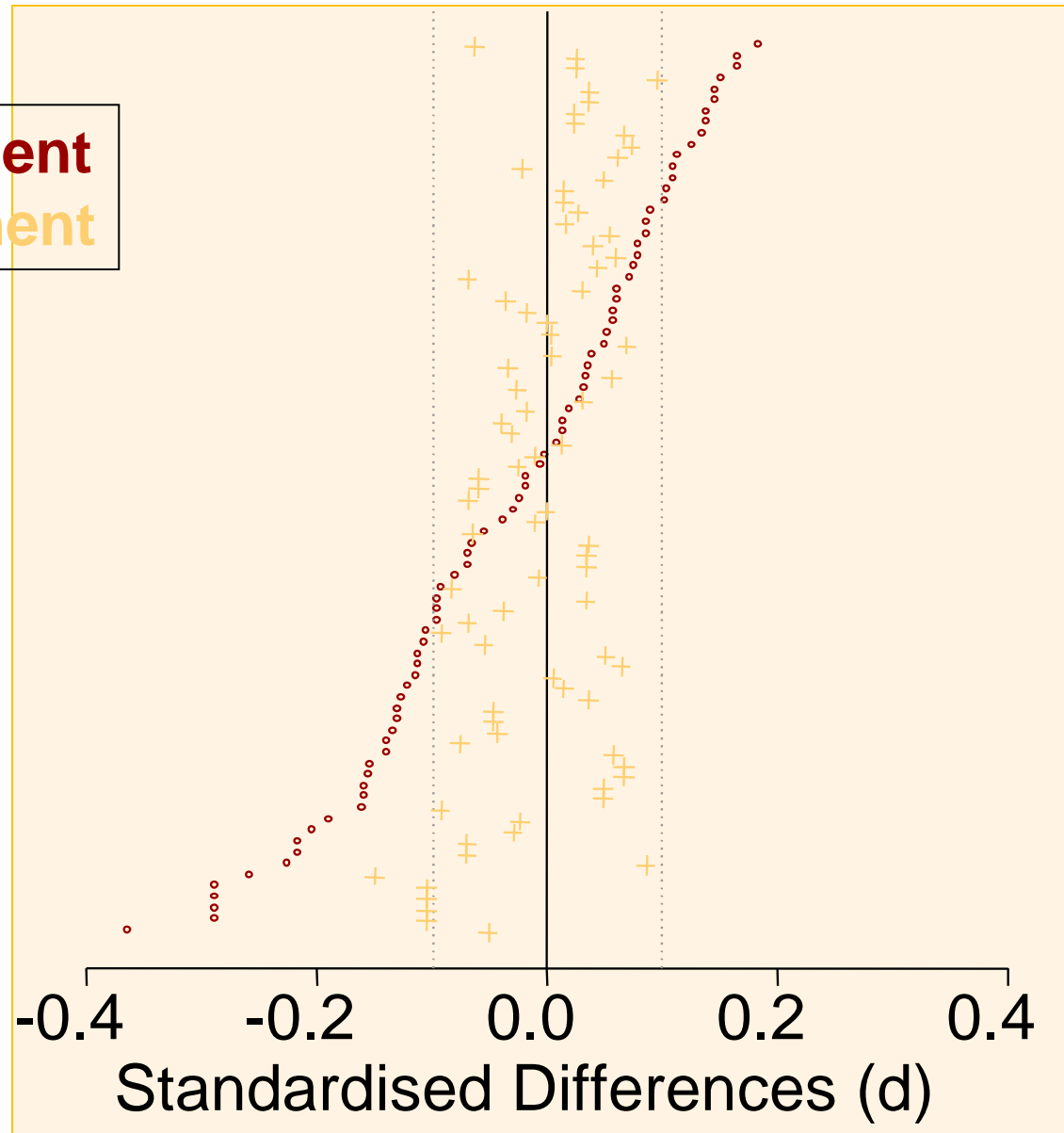
# Biais de Recrutement

Distribution du SP dans le groupe contrôle et le groupe test



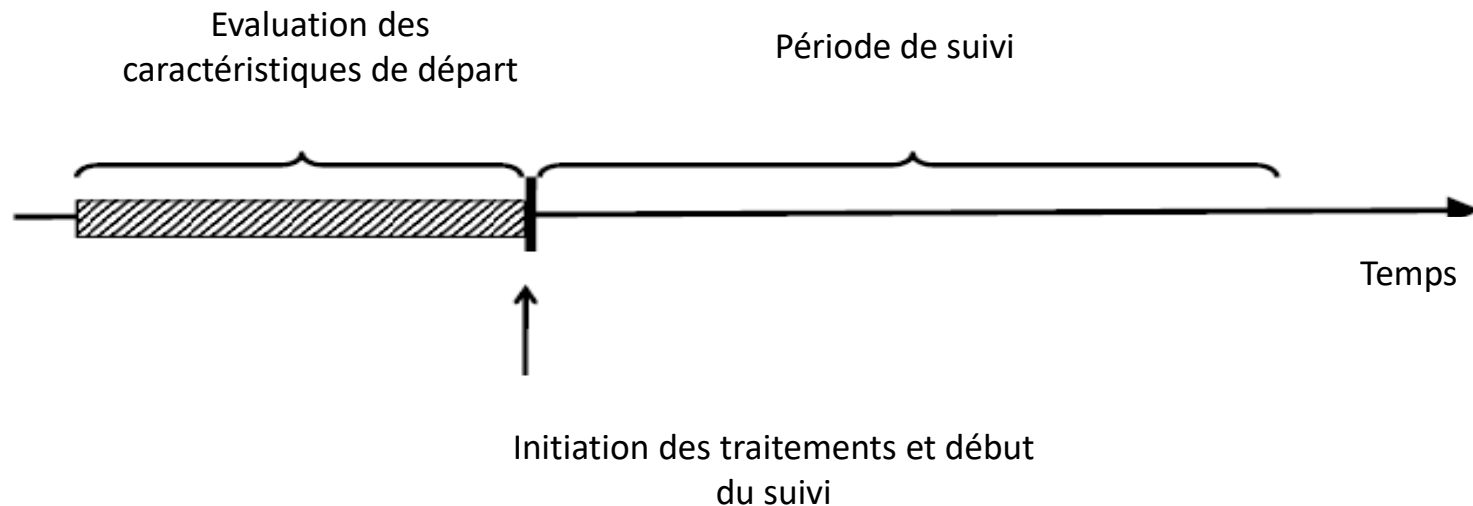
# L'Appariement sur le PS Réduit les Biais

o Avant appariement  
+ Après appariement



# FAIRE PARLER LES BANQUES DE DONNEES MEDICO ADMINISTRATIVES

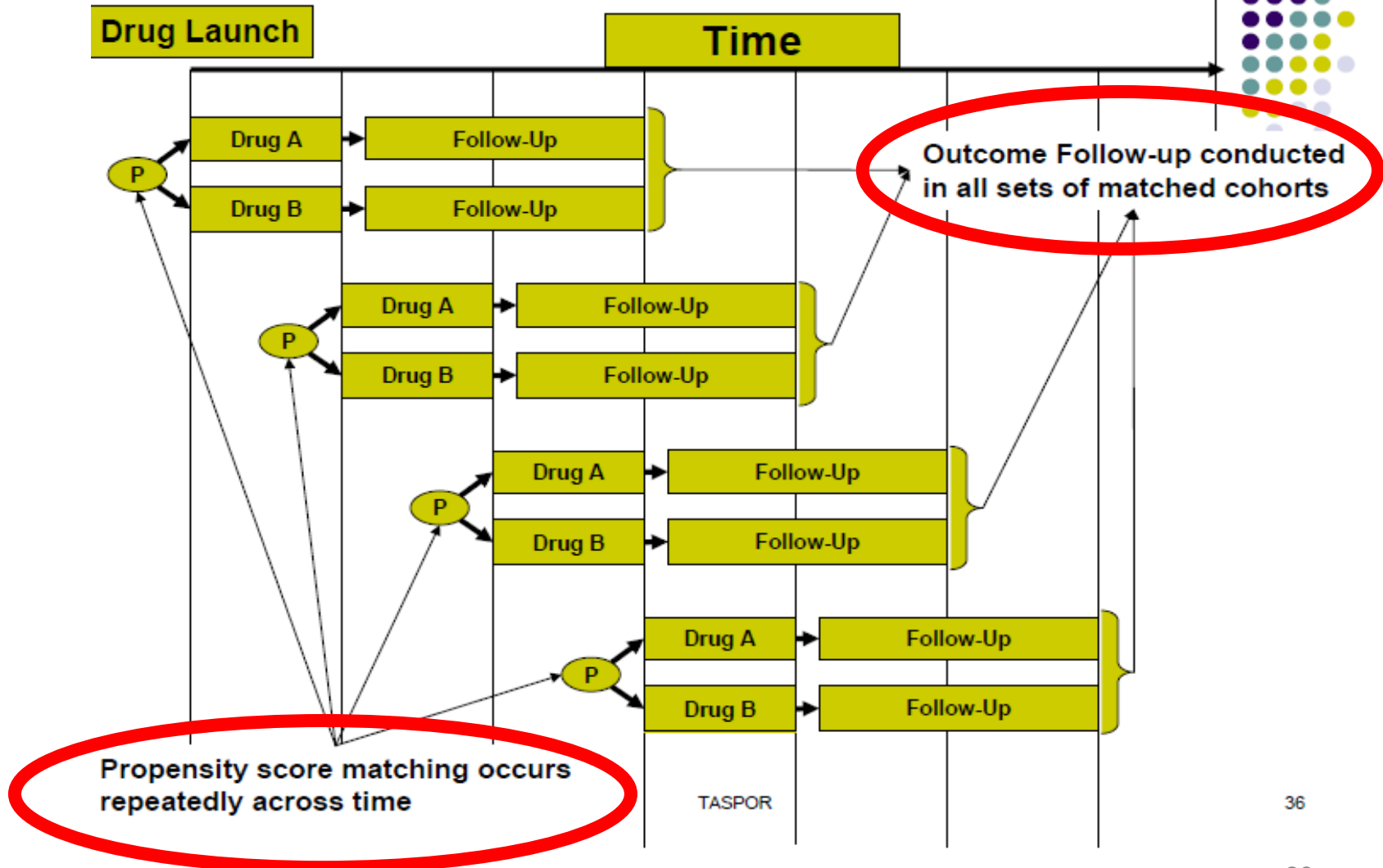
# Schémas de Cohorte Simplifié pour l'Analyse Longitudinale d'une BDMA\*



\*Banque de Données Médico-Administrative

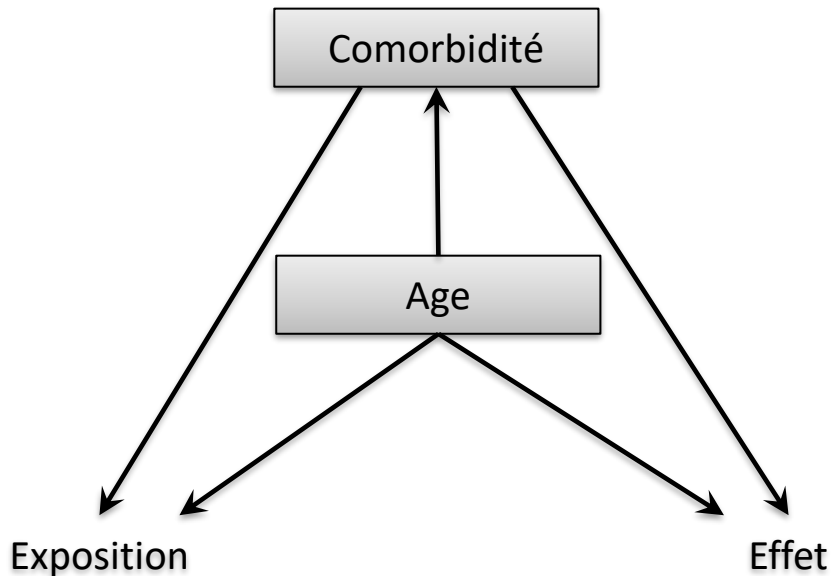
# Technique Longitudinale d'Appariement

Matching within calendar time period



# Identification des Critères de Substitution dans les BDMA

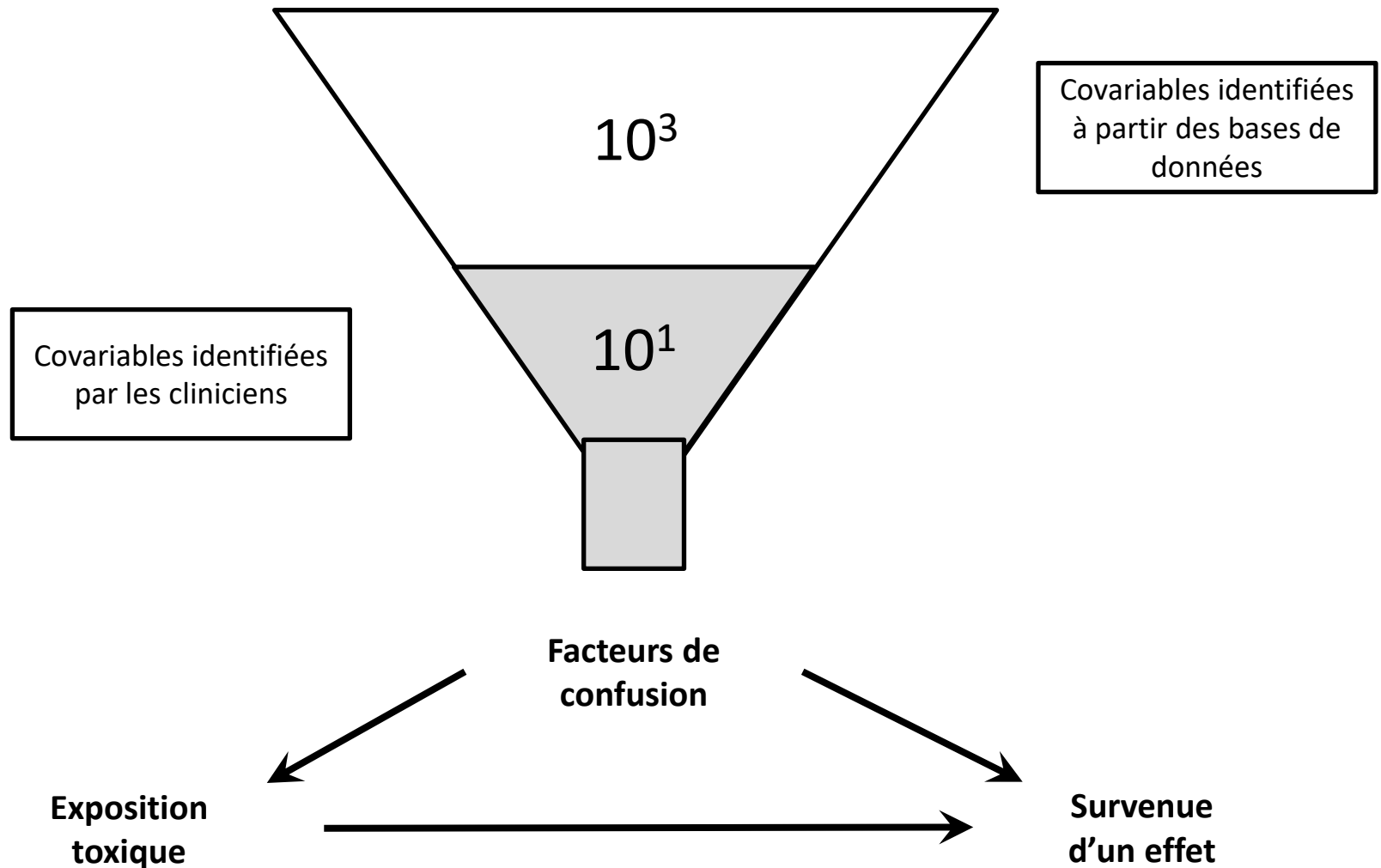
- Les facteurs de confusion identifiés à travers les dépenses reconnues et remboursées sont autant de critères d'évaluation de l'état de santé



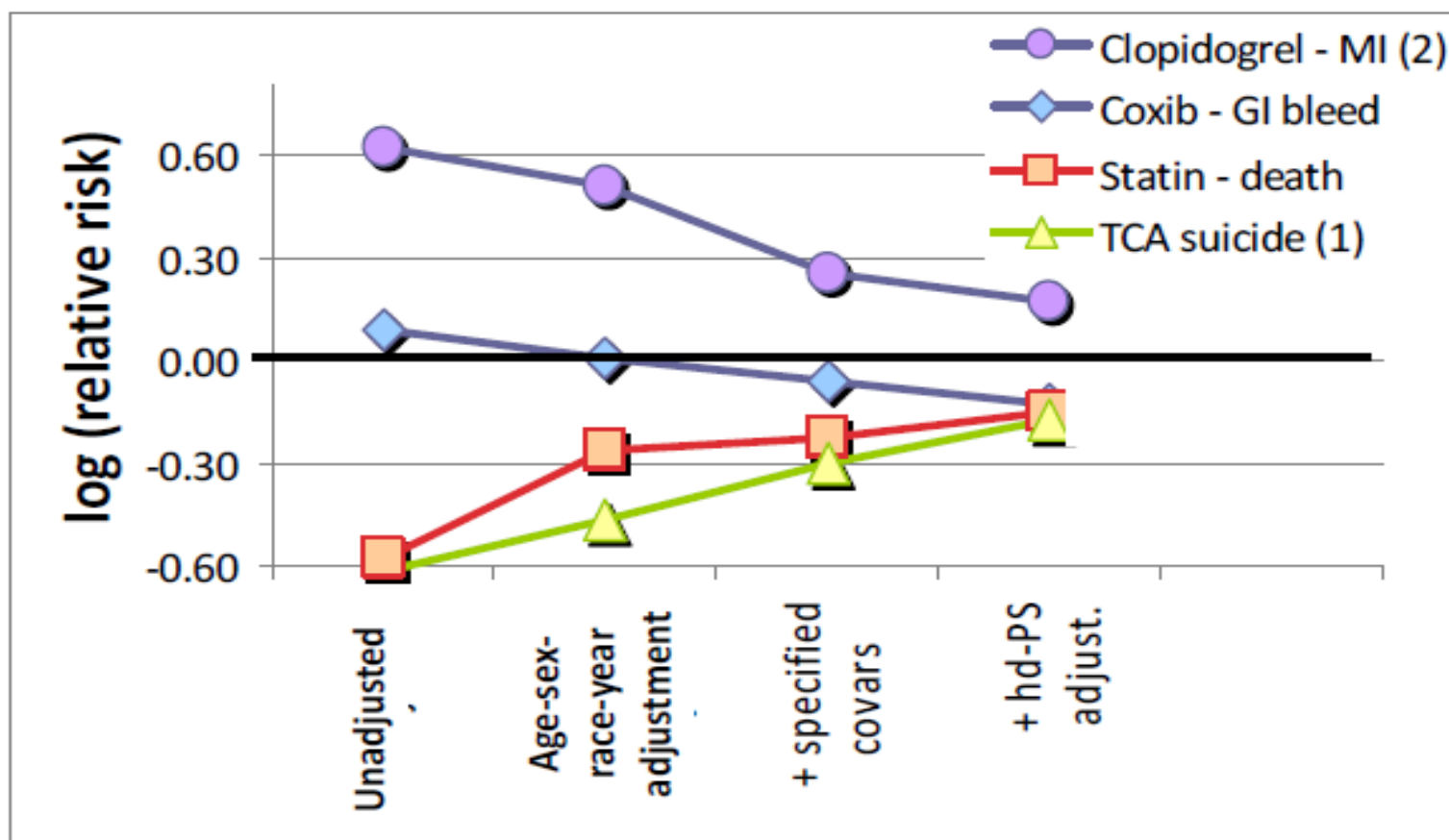
Plus il y a de critères de substitution mieux c'est

Critère de substitution	Etat de santé
Utilisation de bouteilles d'oxygène	Santé très fragile
Diagnostic d'hypertension artérielle durant un séjour hospitalier	Pas trop mauvaise santé
Check-up annuel et coloscopie	Très observant
1 <sup>ère</sup> statine à 70 ans	Plutôt en bonne santé
Polymédication, nombreuses consultations	Plutôt en mauvaise santé

# Identification des Covariables



## Performance of different adjustment procedures, including hd-PS adjustment





## Coronary Heart Disease

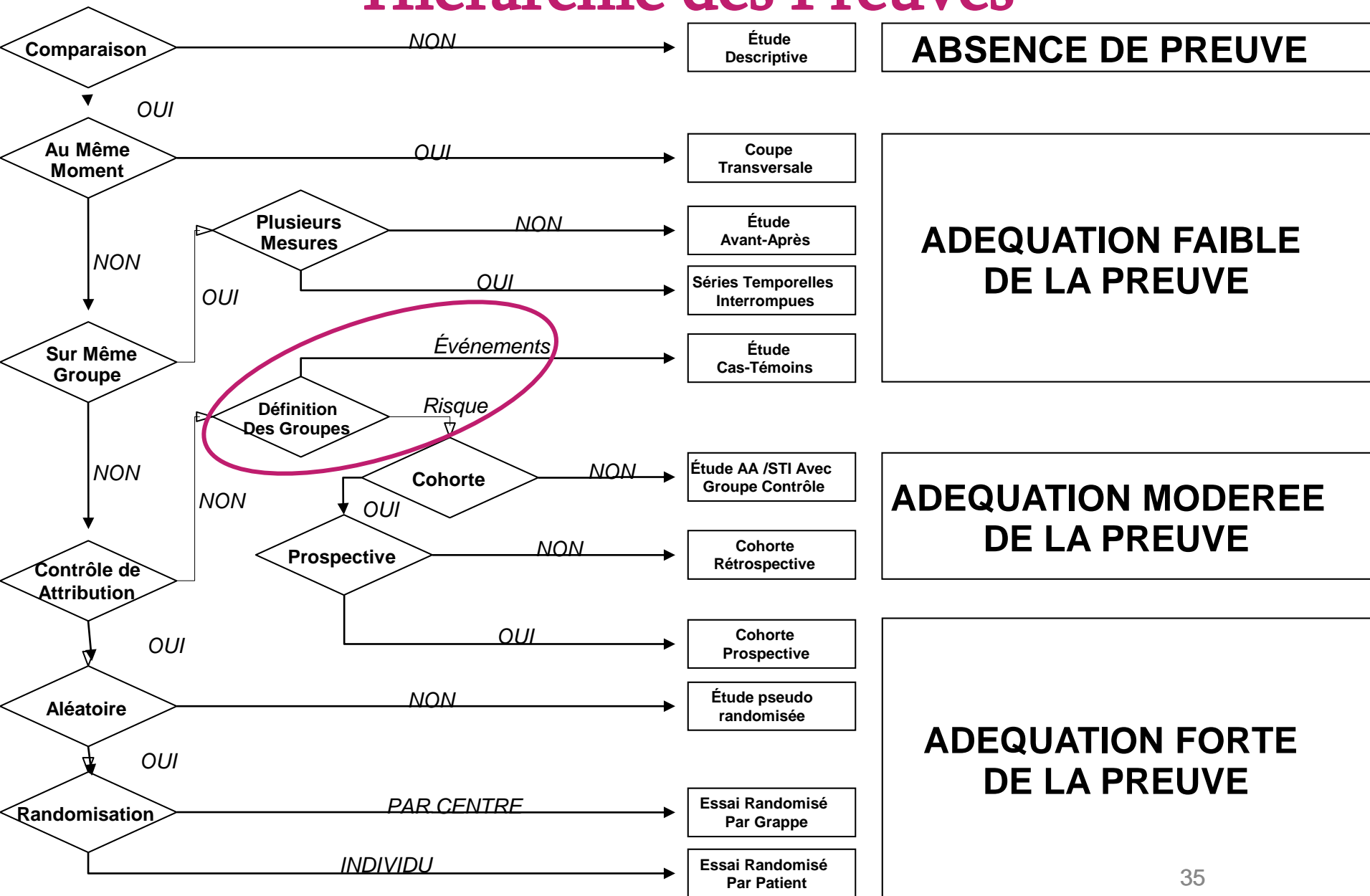
### Cardiovascular Outcomes and Mortality in Patients Using Clopidogrel With Proton Pump Inhibitors After Percutaneous Coronary Intervention or Acute Coronary Syndrome

Jeremy A. Rassen, ScD; Nitesh K. Choudhry, MD, PhD;  
Jerry Avorn, MD; Sebastian Schneeweiss, MD, ScD

Outcome	British Columbia (n=10 391)	RR (95% CI)			P for Heterogeneity
		Pennsylvania (n=4176)	New Jersey (n=3998)	Pooled* (n=18 565)	
Primary end point					
MI hospitalization or death					
Unadjusted	1.90 (1.47–2.46)	1.96 (1.34–2.88)	1.26 (0.85–1.88)	1.74 (1.44–2.10)	0.19
Age/sex-adjusted	1.69 (1.30–2.20)	2.01 (1.37–2.96)	1.18 (0.79–1.75)	1.62 (1.34–1.96)	0.15
Multivariate-adjusted	1.33 (1.01–1.75)	1.71 (1.14–2.56)	0.97 (0.64–1.47)	1.32 (1.08–1.61)	0.16
PS-adjusted†	1.22 (0.92–1.61)	1.69 (1.09–2.63)	1.08 (0.70–1.68)	1.22 (0.99–1.51)	0.31

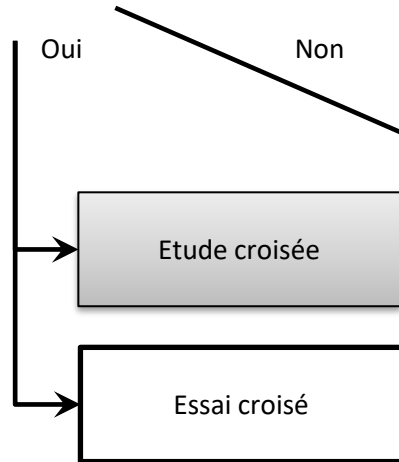
# COMMENT CHOISIR LE BON SCHEMA D'ETUDE ?

# Choix du Schéma d'Etude en Fonction de la Hiérarchie des Preuves

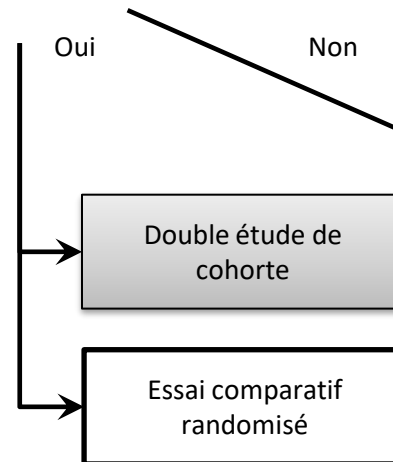


# Choix du Schéma d'Etude En Fonction des Facteurs de Variabilité

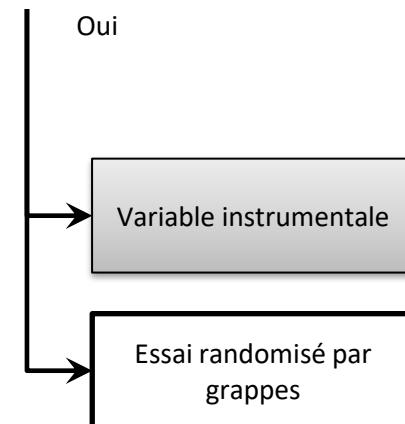
**Variabilité Intra-individuelle**



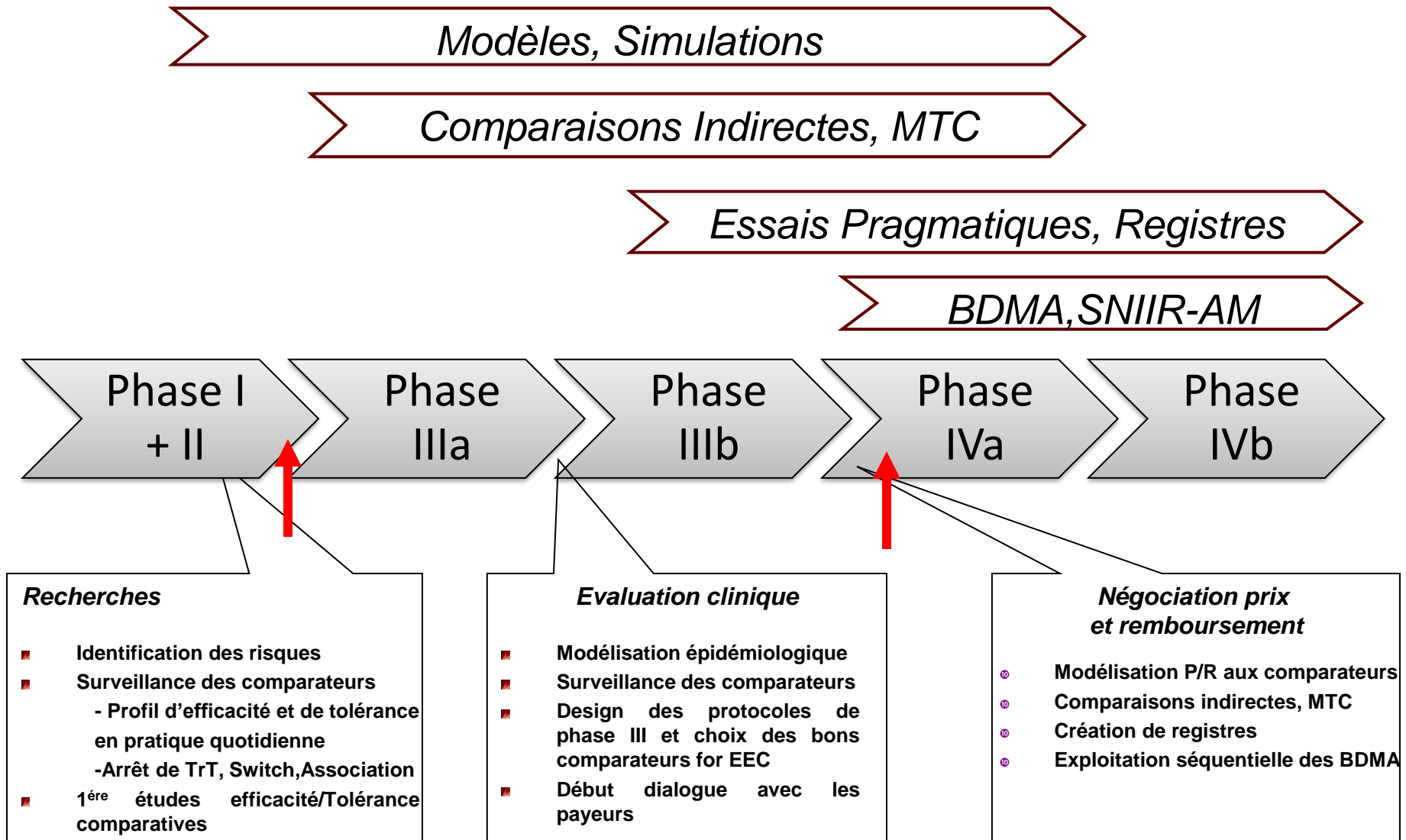
**Variabilité Inter-individuelle**



**Variabilité Inter-structure**



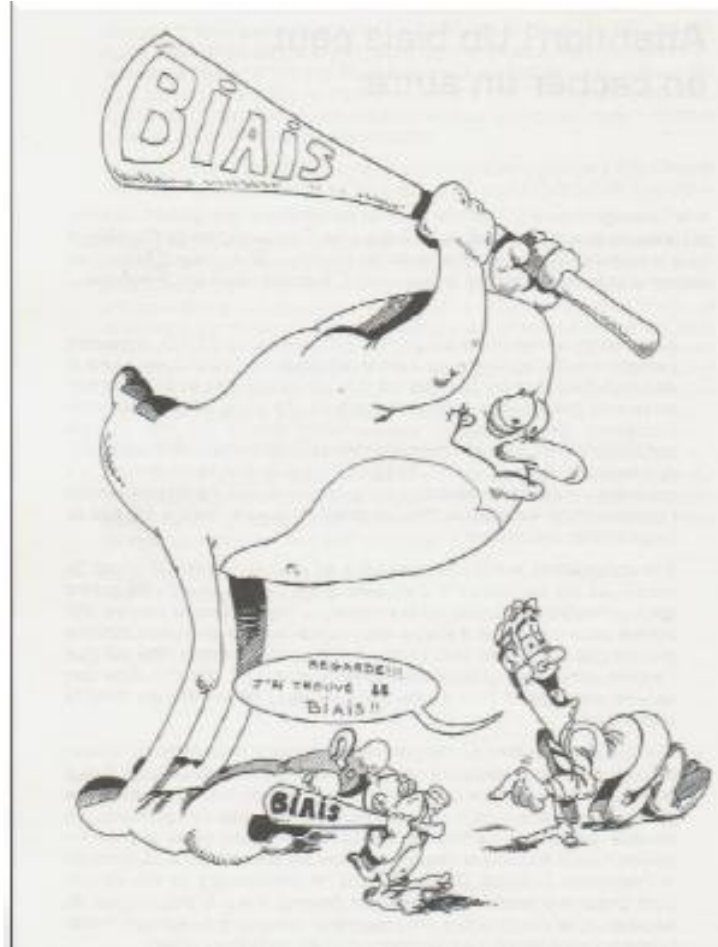
# Choix du Schéma d'Etude Selon la Phase de Développement dans Laquelle se Trouve le Produit



# CONCLUSIONS

- ✦ L'évaluation des politiques publiques constitue une « valeur montante » en France
  - Création d'indicateurs dans le cadre de la LOFT (2001)
  - La possibilité d'expérimentation est inscrite dans la Constitution depuis la loi du 28 mars 2003
  - Lancement d'expérimentations-évaluations RSA
- ✦ Les méthodes d'expérimentation sociale et d'évaluation aléatoire porte une ambition forte puisque leur objectif est « *d'identifier les politiques efficaces et celles qui ne le sont pas* » *Esther Dufflot*

# La Chasse au Biais



# Sélection Bibliographique

- Austin P. An introduction to propensity-score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research*. 2011;46(3):399-424.
- Brodsky T, Crépon B, Fougère D. Les méthodes micro-économétriques d'évaluation et leurs applications aux politiques actives de l'emploi. *Economie et prévision*. 2007;177(1):93-118.
- Cling J, Razafindrakoto M, Roubaud F. L'évaluation d'impact des politiques publiques : enjeux, méthodes, résultats. Paper presented at: Les Journées de Tam Dao2008.
- Duflo E. L'approche expérimentale en économie du développement. *Revue d'économie politique*. 2009:691-726.
- Fougère D. Les méthodes économétriques d'évaluation. *Revue Française des affaires sociales*. 2010;1-2:105-128,
- Givord P. Méthodes économétriques pour l'évaluation de politiques publiques. Paris: INSEE;2010.
- Heckman J. Building bridges between structural and program evaluation approaches to evaluating policy. Cambridge: National Bureau of economic research;2010.
- Kurth T, Sonis J. Assessment and control of confounding in trauma research. *J Trauma Stress*. Oct 2007;20(5):807-820.
- Kurth T, Walker AM, Glynn RJ, et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *Am J Epidemiol*. Feb 1 2006;163(3):262-270.
- Little RJ, Rubin DB. Causal effect in clinical and epidemiological studies via potential outcomes : concepts and analytical approaches. *Annu. Rev. Public Health*. 2000;21:121-145.
- Rassen JA, Brookhart MA, Glynn RJ, Mittleman MA, Schneeweiss S. Instrumental variables I: instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships. *J Clin Epidemiol*. Dec 2009;62(12):1226-1232.
- Rassen JA, Choudhry NK, Avorn J, Schneeweiss S. Cardiovascular outcomes and mortality in patients using clopidogrel with proton pump inhibitors after percutaneous coronary intervention or acute coronary syndrome. *Circulation*. Dec 8 2009;120(23):2322-2329.
- Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. *Am J Epidemiol*. Nov 1 2001;154(9):854-864.
- Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. Apr 2005;58(4):323-337.
- Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf*. May 2006;15(5):291-303.
- Schneeweiss S, Solomon DH, Wang PS, Rassen J, Brookhart MA. Simultaneous assessment of short-term gastrointestinal benefits and cardiovascular risks of selective cyclooxygenase 2 inhibitors and nonselective nonsteroidal antiinflammatory drugs: an instrumental variable analysis. *Arthritis Rheum*. Nov 2006;54(11):3390-3398.
- Schneeweiss S. Developments in post-marketing comparative effectiveness research. *Clin Pharmacol Ther*. Aug 2007;82(2):143-156.
- Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology*. Jul 2009;20(4):512-522.
- Schneeweiss S, Gagne JJ, Glynn RJ, Ruhl M, Rassen JA. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. *Clin Pharmacol Ther*. Dec 2011;90(6):777-790.
- Seeger JD, Kurth T, Walker AM. Use of propensity score technique to account for exposure-related covariates: an example and lesson. *Med Care*. Oct 2007;45(10 Supl 2):S143-148.
- Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci*. Feb 1 2010;25(1):1-21.
- Stuart EA, Marcus SM, Horvitz-Lennon MV, Gibbons RD, Normand SL. Using Non-experimental Data to Estimate Treatment Effects. *Psychiatr Ann*. Jul 1 2009;39(7):41451.\*





# BACK UP