

*CONSTRUIRE UN MODÈLE
POUR DISPOSER D'UNE
PAILLASSE VIRTUELLE*

« S'il vous plaît, Dessine Moi »...un Modèle

- **Un modèle : un objet modélisé? ou un objet modélisant?**
 - . Une représentation stylisée ou un instrument d'expérimentation?
→ adopter la 2^{ème} définition
 - « Pour un observateur B, un objet A* est un modèle d'un objet A,
« dans la mesure où B peut utiliser A*
« pour répondre à des questions qui l'intéressent au sujet de A ». (Legay 1963)
- Tout modèle comporte :
 - une **structure**, sous la forme d'un arbre logique de survenue des événements, suite à une prise de décision,
 - **Un ensemble de lois** c.a.d. des relations déterministes ou probabilistes liant les **entrées** du système et ses **sorties** en termes de conséquences dommageables ou bénéfiques,
 - des **paramètres** c.a.d. des constantes ou des variables pouvant prendre des valeurs numériques, qui permettent de spécifier ces relations,
- Loin d'être une « *boite noire* », c'est un « *outil de facilitation* » qui permet de **décrire** une situation complexe **dans des termes compréhensibles et mesurables** (Varennnes 2009)

A Quoi le Modèle Peut il Servir?

- **Outil prédictif +++ : Élaborer des variantes prévisionnelles** pour éclairer les choix en fonction des scénarios retenus. Objectif : « *passer d'une médecine de la réactivité à une médecine d'anticipation* ». Le choix final reste entre les mains du politique (**vision HAS/CEPS**).
- **Outil normatif ++: Contribuer à la fixation de règles de gestion** (Seuils pour fixer l'effort financier socialement acceptable) . Attention la beauté de la carrosserie peut cacher la faiblesse du moteur !!!
- **Outil descriptif +/- : « mimer » le génie évolutif** de la maladie en formalisant les connaissances et les hypothèses dans un cadre cohérent .
→ *Construire le modèle en fonction des besoins de ceux dont il doit éclairer la prise de décision*

Que Mesure le Paramètre ?

■ Modèles « individus centrés » :

- C'est la distribution du paramètre T (la taille par ex) **dans l'échantillon** qui nous intéresse pour pouvoir les utiliser comme estimateurs des paramètres liés à la population.
- Loi Normale : $T \sim \text{Norm}(\bar{X}, s^2)$, avec :
 - \bar{X} = moyenne de l'échantillon
 - s^2 = variance estimée à partir de l'échantillon

■ Modèles à l'échelle d'une population : Majorité des modèles en évaluation économique.

- C'est la distribution **de l'espérance du paramètre** qui nous intéresse. $E[T]$ est la valeur centrale d'une distribution de probabilité qui est une représentation de l'incertitude de la moyenne dans la population, alors que la moyenne caractérise un échantillon]
- Loi normale : $E[T] \sim \text{Norm}(\mu, \sigma^2)$,
 - $\sigma^2 = s^2/n$ est une représentation de l'incertitude sur la variabilité de la moyenne dans la population

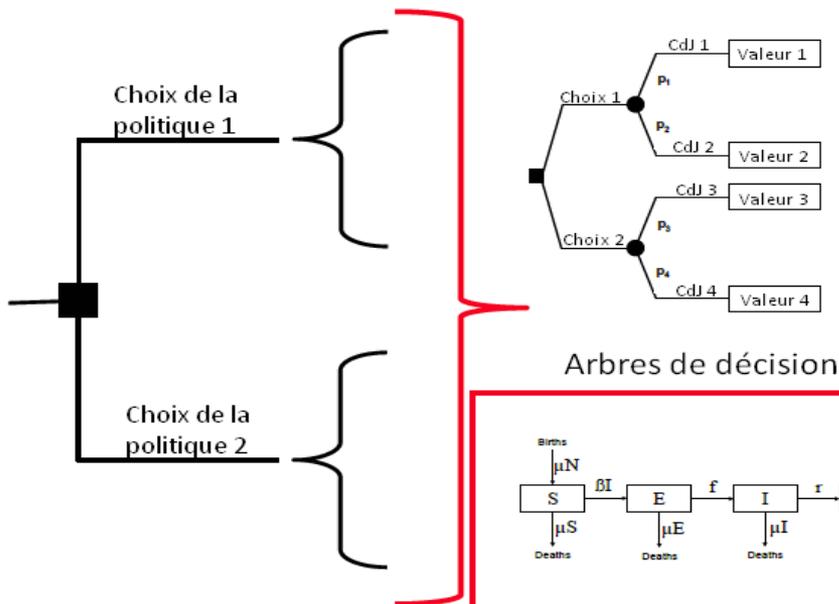
Typologie des Modèles

			Modèles de cohorte / agrégés		Modèles Individu-centrés	
			Déterministes	<i>probabilistes</i>	<i>probabilistes</i>	
Prise en compte du temps	Implicite		Sans Interactions	Arbre de décision	Arbre de décision (simulation Monte Carlo 2 nd ordre)	Arbre de décision (simulation Monte Carlo 1 ^{er} et 2 nd ordres)
	Explicite	Discret		Modèle de Markov	Modèle de Markov (simulation Monte Carlo 2 nd ordre)	Modèle de Markov (simulation Monte Carlo 1 ^{er} et 2 nd ordres)
		Continu	Avec Interactions	Modèles dynamiques (systèmes d'équations différentielles)	Chaîne de Markov à temps discret (CMTD)	Micro simulations
				Chaîne de Markov à temps continu (CMTC)	Modèles à événements discrets	

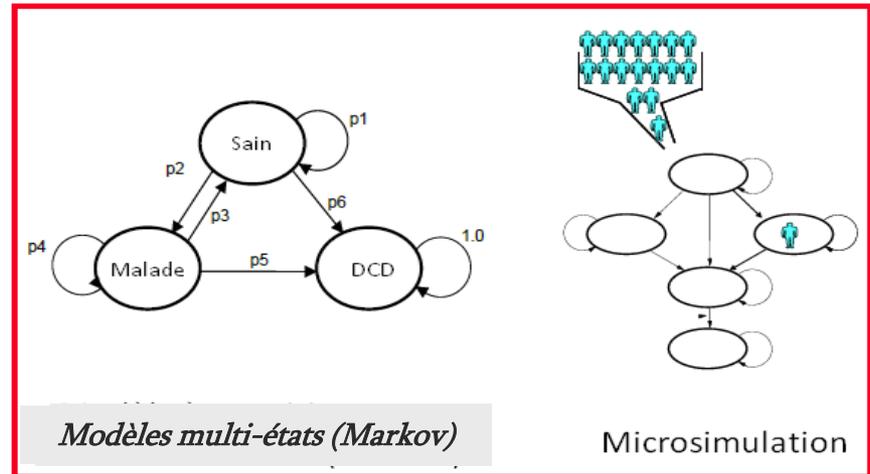
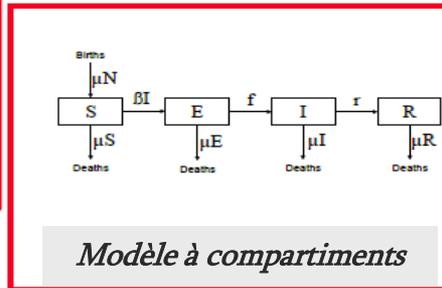
Source : HAS, d'après Brennan et al. (2006)

Référence : Brennan A, Chick SE, Davies R. A taxonomy of models structures for economic evaluation of health technologies. Health Economics 2006; 15: 1295-1310.

Une Grande Variété d'Options Possibles :



Arbres de décision



$$v(h, l) = \text{Max}\{r(h, l), l + \lambda \sum_{h'} \sum_{l'} P(h', l' | h, l) v(h', l')\}$$

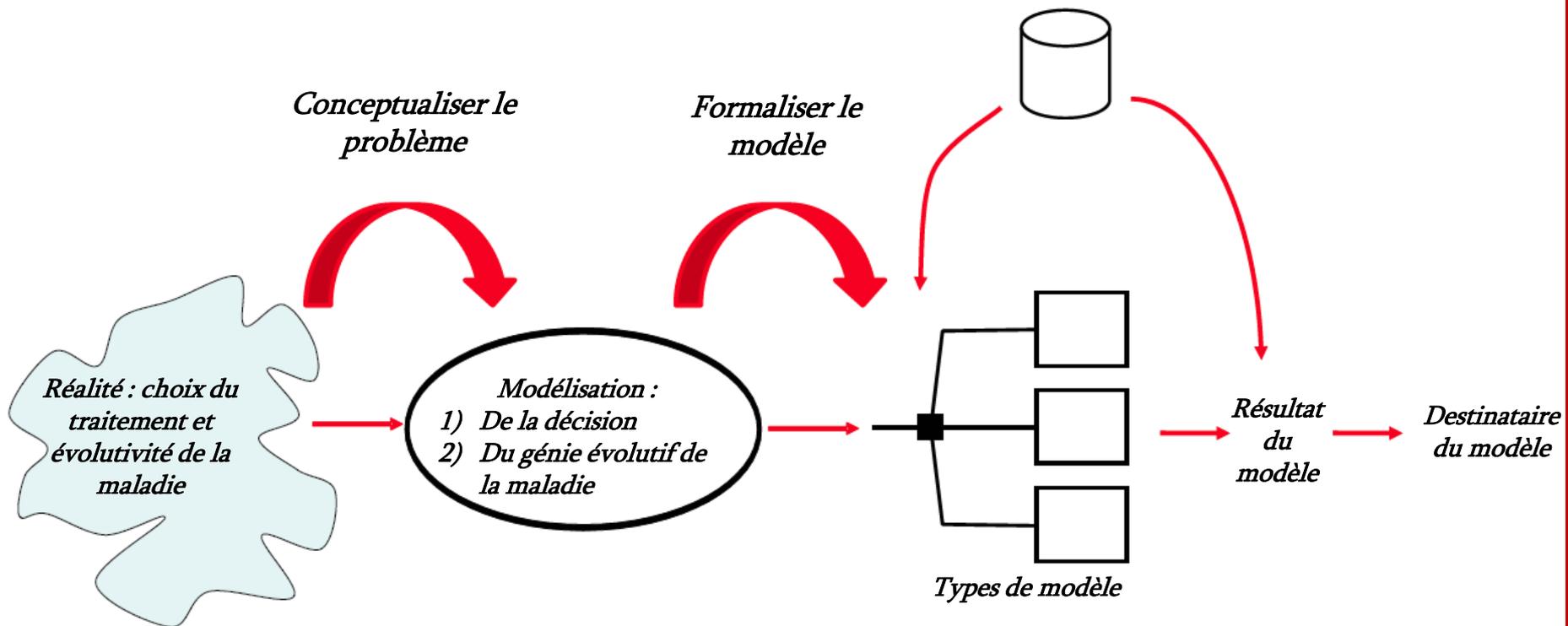
Optimisation mathématique

Comment Choisir ?

- **l'unité de modélisation** doit elle être
 - **agrégée** sur un groupe de sujets présentant des caractéristiques communes (*modèle de cohorte/ agrégé*)? ou
 - **Individu centrée** pour refléter la spécificité des antécédents médicaux de chacun (*modèle individu centré /microsimulation/SED*) ?
- **L'incertitude doit elle intégrée sous la forme**
 - d'une **probabilité moyenne** de survenue (*modèle déterministe*) ? ou
 - par une **plage de valeurs** -une distribution de probabilité- qui se substituerait à tout facteur d'incertitude (*modèle probabiliste*) ?
- **le temps doit il être pris en compte** de façon **implicite** ou **explicite**? Doit il être mesuré de façon **continu** ou à intervalles discrets?
- **les individus interagissent ils entre eux?**
 - Ils n'ont pas d'interaction entre eux : → *modèle multi-états/ de Markov ou de microsimulation*
 - Ils interagissent-ils entre eux → *modèle dynamique /à compartiments*

Réfléchir Avant de Programmer

- La modélisation est une abstraction de la réalité
 - Le modélisateur fait d'abord des hypothèses sur ce qui caractérise au mieux le tableau clinique dans l'indication,
 - Avant de sélectionner un type de modèle et de formaliser pour en faire un bon instrument d'expérimentation

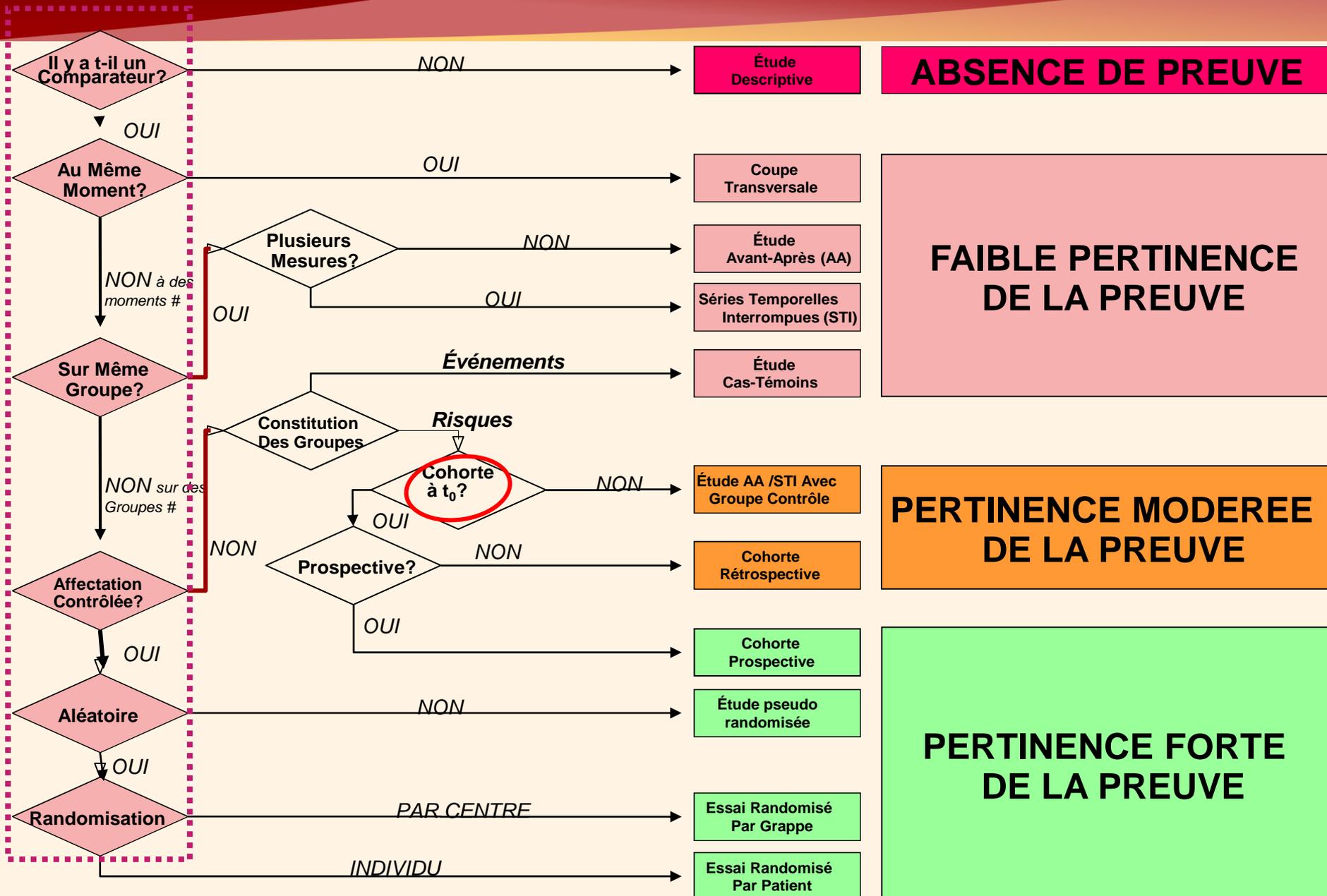


Construire Une Population Virtuelle Réaliste (PVR[†])

- **Recueil de données réelles** sur coupe transversale reflétant la pratique médicale quotidienne française,
- **Modélisation** intégrant les connaissances disponibles dans une reproduction informatique de la population à prendre en charge (Multicohorte)
- **Simulation numérique** par le biais d'une série « d'expérimentations sur le modèle ». Stockage du résultat de ces calculs de Monte Carlo d'ordre 2 dans une base numérique
- **Exploration de différents scénario alternatifs dans le cadre** du clone de la population réelle ainsi reconstituée in silico « la PVR[†] »
- **Vérification expérimentale** : Si les résultats numériques et expérimentaux sont proches, la simulation est déclarée valide

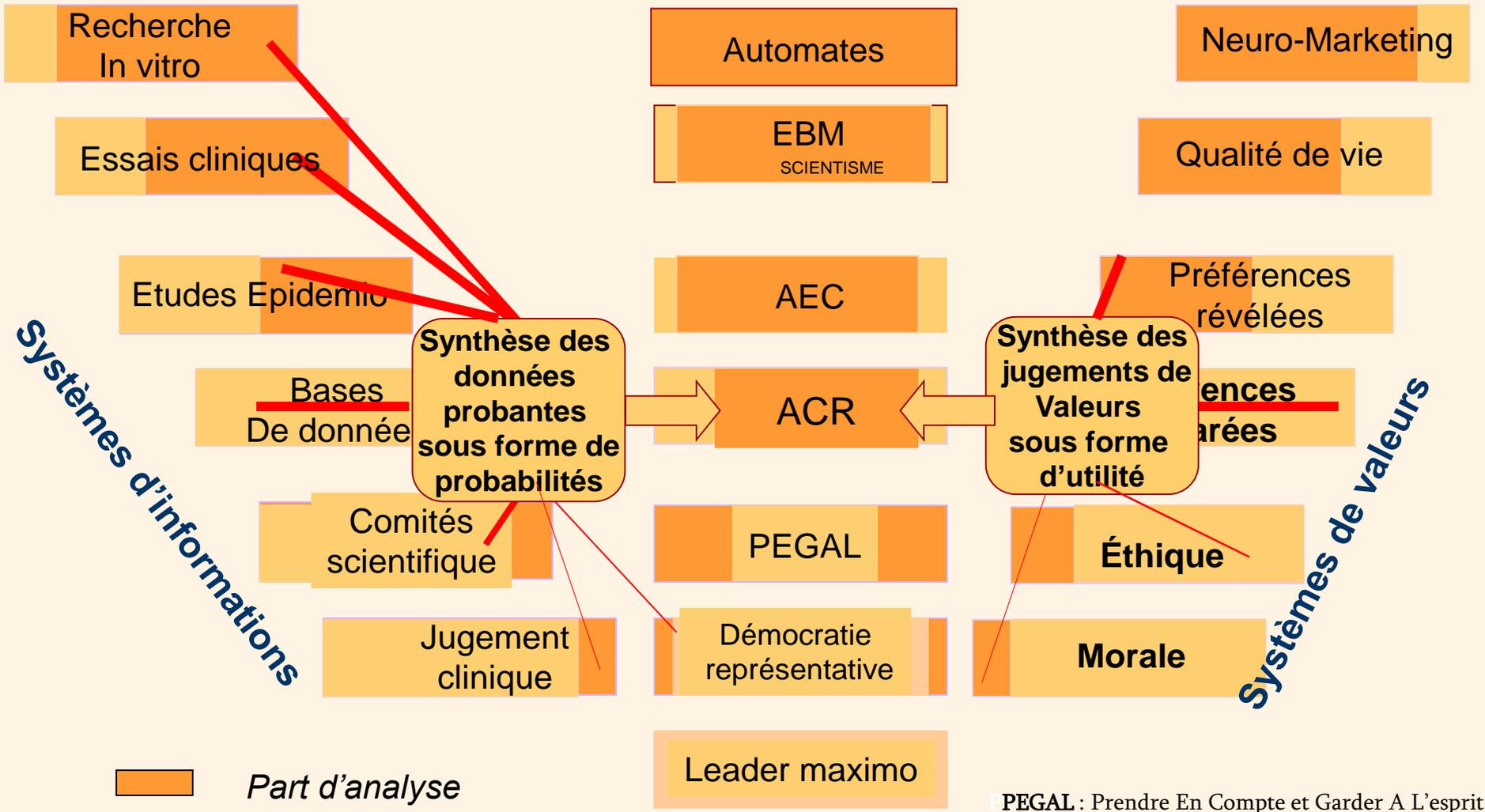
[†] F. Gueyffier, Brun Strang C, Berdeaux G. Apport de la modélisation et des populations virtuelles pour transposer les résultats des essais cliniques à la vie réelle et éclairer la décision publique .xxviii Atelier de Giens. Thérapie 2012 Juillet-Août; 67 (4): 359–366

En Rassemblant des Données du Monde Réel

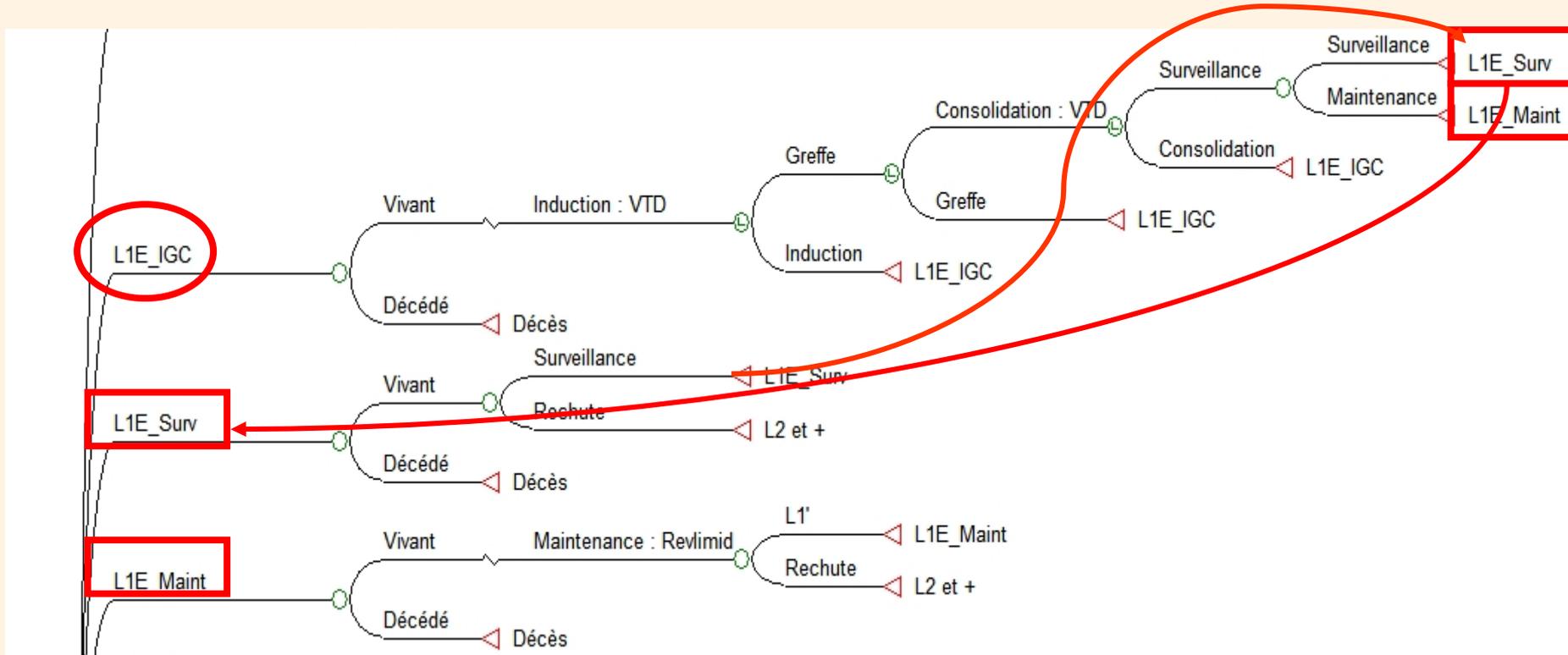


Dans le Cadre d'un Modèle de Décision Quantitatif et Explicite

Systemes de gouvernance



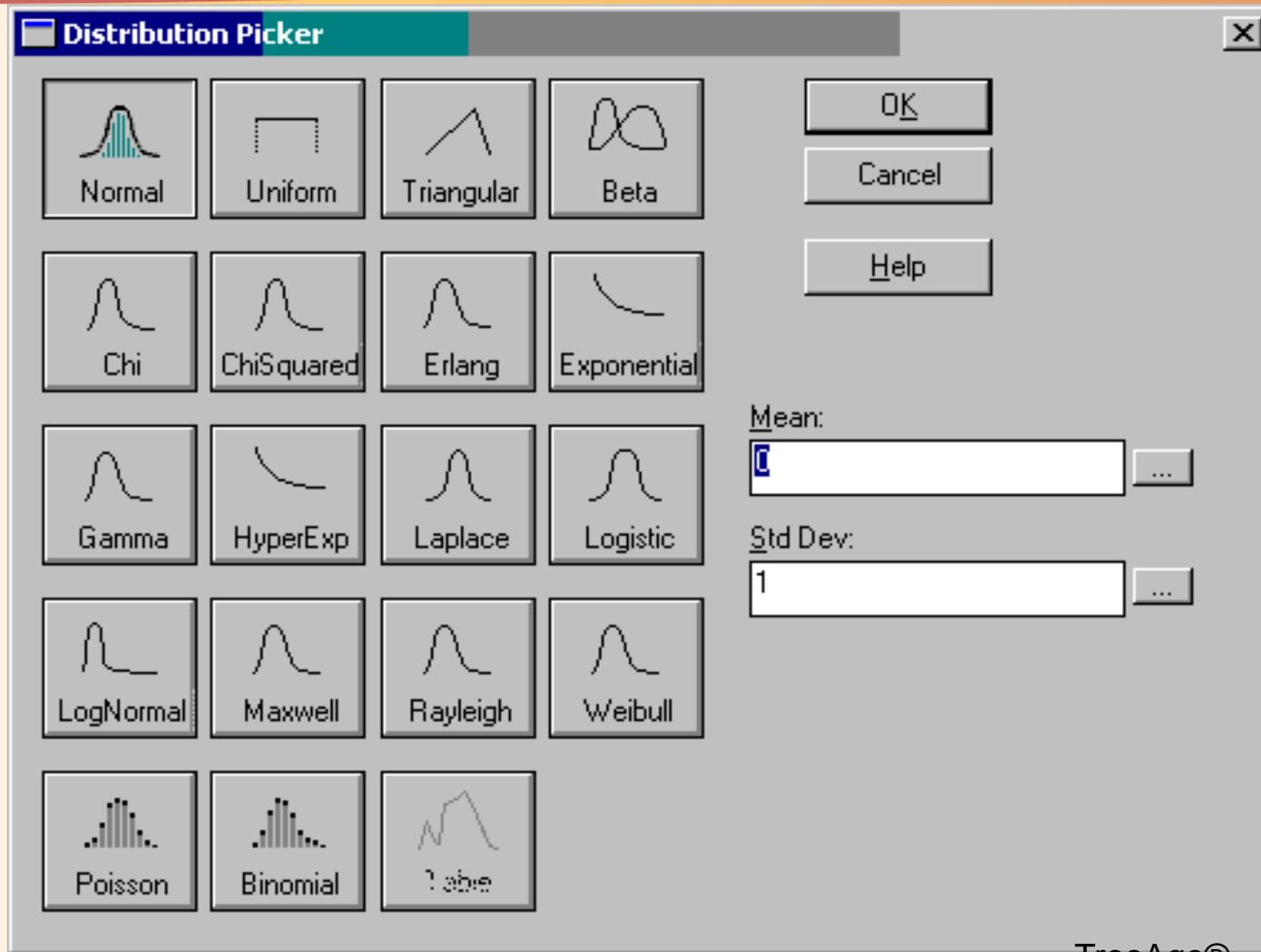
Programmation TreeAge® : Les Sorties d'un Cycle Sont les Entrées du Cycle Suivant



Traduction TreeAge module 1 : Patients Eligibles, 1^{ère} Ligne

On file REES

Après y Avoir Introduit des Lois de Probabilité Pour Passer de «l'Expérience de Pensée» à la Simulation Numérique†:



† *Simuler c'est faire des expérimentations sur un modèle. Pour que les équations mathématiques du modèle puissent être comprises par l'ordinateur, il faut les traduire en algorithmes numériques, les seules formulations qu'un ordinateur soit capable de traiter. Les résultats des calculs sont stockés dans une base de données numériques*

Finalité: Contribuer à L'Établissement Des Conditions d'Efficiency

- **En synthétisant** les critères de jugement clinique à l'aide d'un **indicateur de résultats de morbi-mortalité**
- **En réintroduisant les préférences** des patients ou des citoyens dans les choix individuels et collectifs
- **En chaînant les coûts** et en reconstituant les trajectoires de prise en charge
- **En permettant d'extrapoler les résultats à d'autres populations** ou à d'autres types de pratiques,

La Modélisation: L'Instrument Incontournable des Etudes de Primo Inscription

- **Pour généraliser les résultats des essais** sur la Population Virtuelle Réaliste que l'on peut construire à partir des études observationnelles,
- **Pour estimer quantitativement**, et non plus de manière intuitive, la fréquence de survenue des événements évolutifs,
- **Pour isoler les variables-clés** et **caractériser l'incertitude** qui les entoure



***INCERTITUDES ET
SIMULATIONS NUMÉRIQUES***

Incertitudes : Définition et Causes

➤ *Incertitude = incapacité à donner une valeur unique*

■ Quatre catégories d'incertitude

➤ *Incertitude par essence* due à la *variabilité intrinsèque* de la grandeur concernée pour cause de fluctuations d'échantillonnage

➤ *Incertitude par ignorance* due au *manque d'information* sur le phénomène étudié

➤ *Incertitude due à l'hétérogénéité naturelle des tableaux cliniques*

➤ *Incertitude sur le bien fondé du construit*

■ La différence porte sur **la capacité à réduire l'incertitude**:

➤ Les incertitudes par essence *sont irréductibles*

➤ Les incertitudes par ignorance sont *réductibles* par apport de données nouvelles.

■ **Tout ce qui est incertain est probabilisé**, pas de paramètres fixes; à la sortie on n'a pas un valeur mais une distribution de valeur

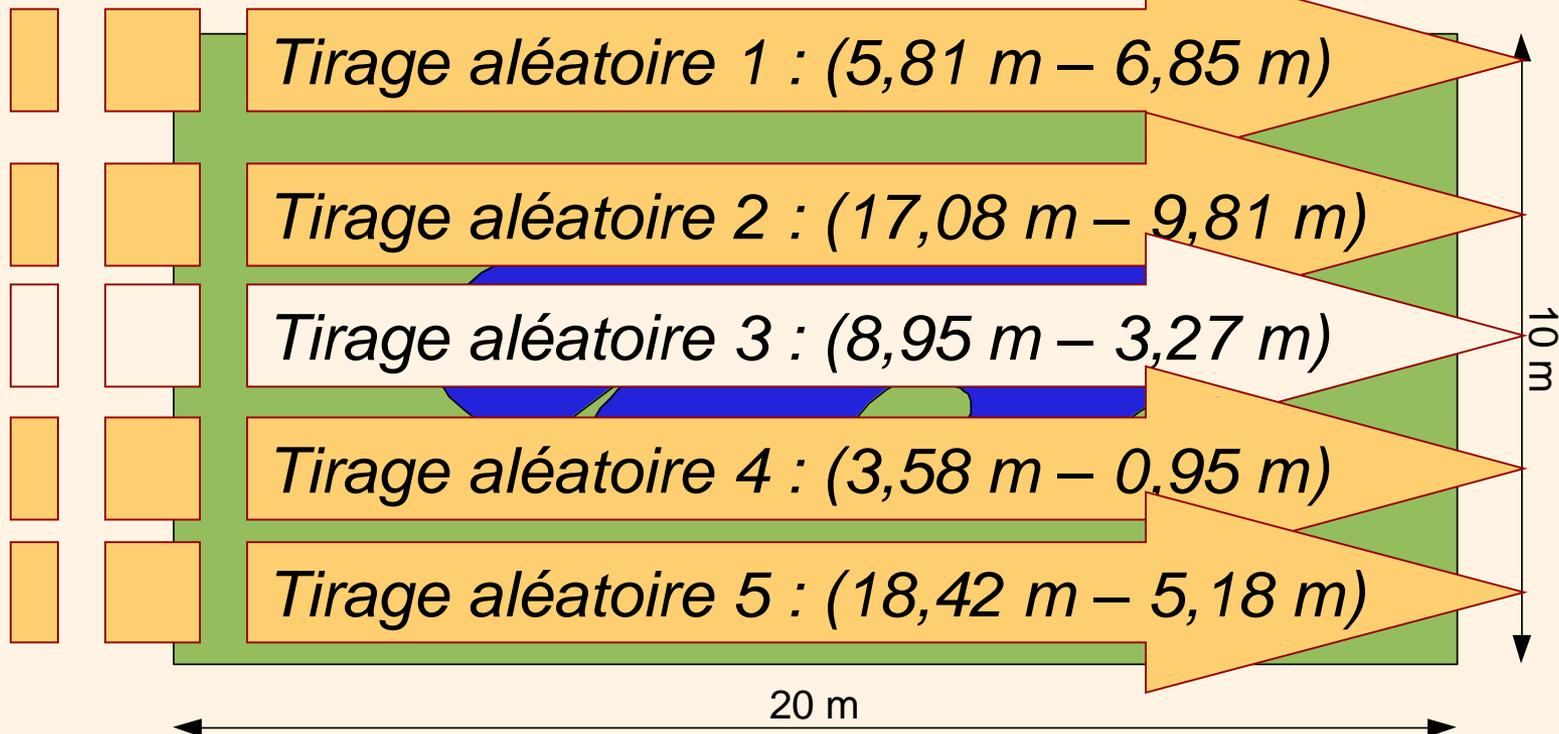
Variabilité - Incertitude - Hétérogénéité

Sources des incertitudes	Nature de l'incertitude	Autre terminologie	Analogie avec la régression
<i>Variabilité stochastique</i>	Incertaine par essence : variation purement aléatoire entre les mesures successives d'un même phénomène, → irréductible	Fluctuation d'échantillonnage. Erreur de Monté Carlo. Incertaine d'ordre 1 ou à une dimension 1D	Erreur aléatoire
<i>Incertaine sur les paramètres</i>	Incertaine par ignorance, → réductible par introduction de nouvelles données	Incertaine du second ordre ou à deux dimensions 2D	Erreur standard de l'estimateur
<i>Hétérogénéité des populations</i>	Variabilité interindividuelle attribuable à la diversité des tableaux cliniques, → réductible par analyse en sous-groupes	Variabilité. Hétérogénéité observée ou expliquée	Coefficients beta (la valeur de la variable dépendante change selon les patients)
<i>Incertaine structurelle</i>	Bien fondé des hypothèses sur la structure du modèle de décision → réductible par hypothèses alternatives	Incertaine sur le construit de la modélisation	Forme du modèle de régression (linéaire, log linéaire).

Simulation : Définition

- ✦ Toute méthode calculant une valeur numérique à partir de procédés aléatoires. [les statistiques sont accumulées durant la période des itérations pour évaluer les grandeurs d'intérêt à la fin de l'exécution du programme]

- ✦ Calcul de la surface d'un lac :



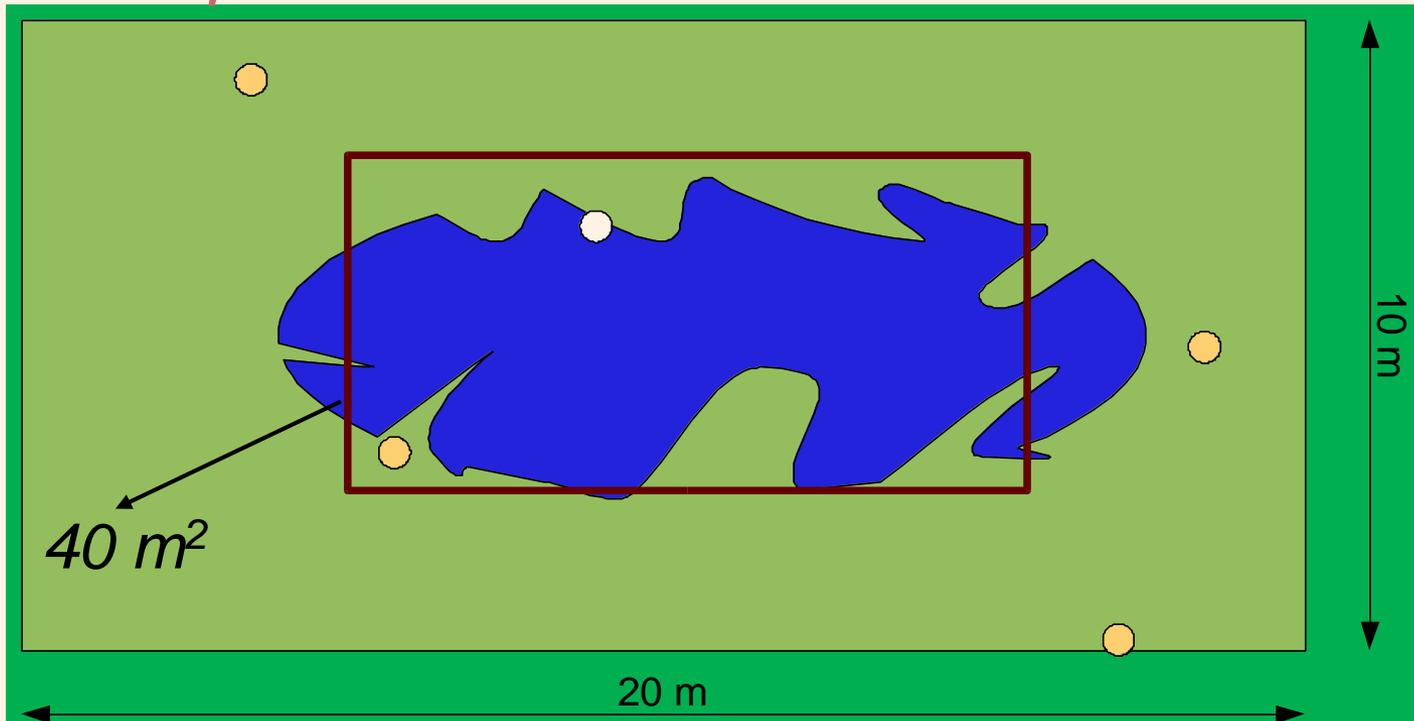
Simulation : Définition

- Estimation : chercher à **donner la valeur la plus proche possible de la valeur « vraie »** du paramètre inconnu
- Simulation : Toute méthode permettant d'obtenir une valeur numérique à partir de procédés aléatoires en l'absence de solution analytique connue pour la loi de probabilité des sorties du modèle
- La procédure consiste à accumuler les statistiques durant la période des itérations pour évaluer les grandeurs d'intérêt à la fin de l'exécution du programme

Ex d'Estimation par Simulation de Monte Carlo

- ✦ *Calcul de la surface d'un lac; surface du terrain $10 \times 20 = 200 \text{m}^2$. 5 tirs effectués. 1/5^e des tirs touchent le lac \rightarrow surface du lac = $200 \text{m}^2 / 5 = 40 \text{m}^2$*

Attention à procéder à un nombre suffisant de tirages !



Types De Monte-Carlo

- **Simulation de premier ordre** : La simulation de premier ordre consiste à tirer au hasard le devenir d'un patient dans l'arborescence:
 - On compare la valeur (R) prise par une variable aléatoire tirée d'une loi de *distribution uniforme* comprise entre 0 et 1 à *la probabilité fixe de survenue* (p) de chacun des événements qui se trouvent sur son chemin.
 - si $R < p$ → pas de changement d'état,
 - si $R > p$ → changement d'état
 - Sur un grand nombre d'itinéraires, on retrouve la probabilité de survenue de chaque conséquence possible. Seule la variabilité est étudiée; l'incertitude ne l'est pas car *la stratégie optimale reste toujours la même*
- **Simulation de second ordre** : tirage au sort de chaque variable aléatoire en fonction de sa loi de probabilité, puis calcul des résultats attendus
 - Prise en compte de l'incertitude sur les variables
 - Selon les valeurs obtenues, le choix de la stratégie optimale pourra différer
- **Simulation de troisième ordre** : combinaison des deux simulations précédentes

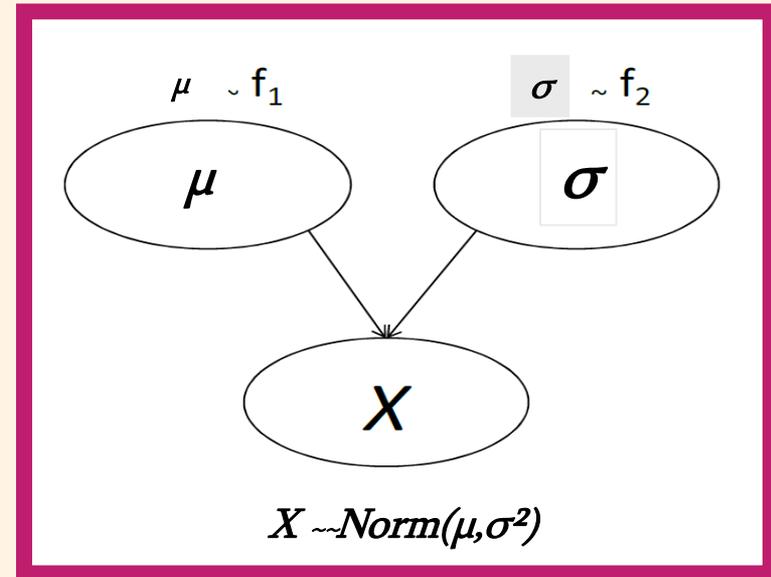
→TreeAge® permet les trois variantes

Simulation de Monte Carlo du 1^{er} ordre

- Les sujets sont envoyés un par un dans le modèle
- Au franchissement du premier embranchement de l'arborescence, on simule une des réalisations possibles de la loi Uniforme $[0,1]$:
 - Si sa valeur est comprise $[0, 0.1]$, le sujet décède
 - Si sa valeur est comprise $[0.1, 0.3]$, le sujet tombe malade
 - Si sa valeur est comprise $[0.3, 1]$, il reste en bonne santé
 - Lorsque l'expérience est répétée sur un grand nombre d'individus, les effectifs relevés dans les différents état de santé sont les mêmes que ceux obtenus dans un modèle de simulation agrégé de cohorte
- Le modèle permet de retracer l'histoire de la maladie
 - Une variable binaire permet de distinguer le premier épisode de celle ci et les rechutes
 - La valeur des probabilités de transition seront fonction de la valeur de la variable binaire

Simulation de Monte Carlo du 2^{ème} ordre

- Dans le cadre d'un MC d'ordre 2, la variabilité et l'incertitude sont appréhendées dans deux dimensions différentes
- La variabilité d'une statistique d'un échantillon suit une loi normale de moyenne x_1 et d'écart type= y_1 , **conditionnellement** aux valeurs que peuvent prendre ces valeurs en fonction de **la forme des lois de probabilité qui régissent leur réalisation** qui est caractérisée par leur moyenne: μ et leur erreur standard: σ . μ et σ sont appelés des hyperparamètres.



Une modélisation hiérarchique

Incertitude et Médecine

- L'incertitude est **consubstantielle** à l'activité médicale. Quand un médecin prend une décision c'est toujours dans l'angoisse de prendre la mauvaise
- Les analyses économiques sont entachées de la même caractéristique à la fois sur l'estimation de l'efficacité, des coûts et des conséquences de la pathologie
- Le fait qu'on soit dans l'incertitude doit inciter à se poser des questions : « **What if ?** » : **Qu'est ce qu'il se passerait si ?**.
- L'analyse de sensibilité a pour but d'évaluer la stabilité des conclusions d'une analyse par rapport aux hypothèses émises
- Il y **plusieurs manières de la faire**

Analyse De Sensibilité Classique

■ Modalités

- Analyse uni,bi ou tri-dimensionnelles
- Jeu limité de scénarios
- Chaîne de corrélations entre variables
- Recherche des seuils de renversement des choix

■ Limites :

- On suppose que les autres paramètres demeurent constants
- Ne permet pas d'explorer les variations conjointes de tous les paramètres.(pas plus de 3 paramètres en même temps)
- Certaines valeurs pour une variable **sont plus probables** que d'autres

→ **Le traitement de l'incertitude exige un regard nouveau**

Analyse de Sensibilité Probabiliste : 3 Méthodes

- **Monte Carlo** : obtenir la distribution d'échantillonnage à partir de la distribution *a priori* des paramètres du modèle: analyse de **sensibilité paramétrique**.
- **Bootstrap** : obtenir la distribution d'échantillonnage à partir de la distribution empirique des observations par tirage au sort avec remise: analyse de **sensibilité non paramétrique**
- **Distribution a posteriori Bayésienne** . On choisit une distribution a priori des paramètres pour représenter l'incertitude de la moyenne dans la population et sa variabilité et on actualise l'information, par apport de données nouvelles. À la sortie on a une distribution de valeur.

Analyse de Sensibilité Paramétrique : *« Réunir les Incertitudes dans de Grands Sacs d'Ignorance »*

Bernier J., Parent E., Boreux JJ. Paris 2000

- ✦ *Faire une typologie de ces sacs,*
- ✦ *définir leur forme à partir d'un petit nombre de paramètres*
- ✦ *simuler des tirages au hasard issus de ces lois*

Analyse de Sensibilité Paramétrique

- ✦ A chaque variable aléatoire, on associe non pas une *probabilité moyenne* mais une *distribution de probabilité* qui décrit leur fréquence d'apparition anticipée
- ✦ Pour une famille arrêtée de courbes, on *caractérise la valeur de ses paramètres* théoriques non observables qui simule le mieux la réalité observée
- ✦ Une fois caractérisée la loi de distribution, *on tire au sort* la réalisation de chaque variable,
- ✦ l'incertitude *est incorporée et se propage* dans tout le modèle.
- ✦ le résultat d'une analyse quantitative stochastique des risques est une *distribution de probabilité*.
- ✦ Sur un grand nombre de tirages, *la moyenne des sorties du modèle approche leur espérance*

*CHOIX DES LOIS DE PROBABILITÉ ET
DÉDUCTION DU PARAMÉTRAGE
À PARTIR DES STATISTIQUES
DESCRIPTIVES*

Lois de Probabilité Usuelles

- **Uniforme : Unif(Min = a ; Max = b)**
 - Loi **bornée sur [a;b]**
 - Utile lorsque l'on n'a aucun a priori sur la probabilité des valeurs prises par un paramètre, mais que l'on connaît son intervalle de variation
- **Normale : $N(\mu; \sigma^2)$ avec $\sigma^2 = s^2/n$**
 - **Le support de la distribution normale est l'ensemble des réels de moins l'infini à plus l'infini $]-\infty ; +\infty[$,**
 - symétrique, comporte des valeurs négatives, à utiliser avec césure
- **Log-Normal: LogNorm($\alpha ; \beta$)**
 - Définie sur $[0 ; +\infty[$
 - Recommandée pour les distributions asymétriques
- **Beta : Beta(a ; b)**
 - **Définie sur [0 ;1]**
 - Distribution polymorphe
- **Gamma: Gamma($\alpha ; \beta$)**
 - Définie sur $[0 ; +\infty[$
 - Distribution polymorphe

Formes Fonctionnelles Recommandées

- Probabilités $[0 ; 1]$: loi Beta
- Risques relatifs $[0 ; \infty[$: log normale ou loi Gamma
- Utilité $]-\infty ; 1]$: loi Beta ou Normale tronquée à 1
- Coûts $[0 ; \infty[$ loi Gamma ou Lognormale
- Valeurs inconnues : Loi Uniforme non informative

Caractérisation d'une Proportion

- Bornée sur $[0 ; 1]$ → Loi Beta

- Méthode des moments :

Retrouver α et β à partir des caractéristiques de la distribution : moyenne, variance , médiane, bornes d'un Intervalle de Confiance...

- Méthode Pragmatique :

- α = Nombre de succès

- β = Nombre d'échecs

Loi Beta

✦ *Beta*($\alpha ; \beta$)

✦ Continue sur $[0 ; 1]$

✦ Densité :

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$$

✦ Espérance :

$$E[X] = \frac{\alpha}{\alpha + \beta}$$

✦ Variance :

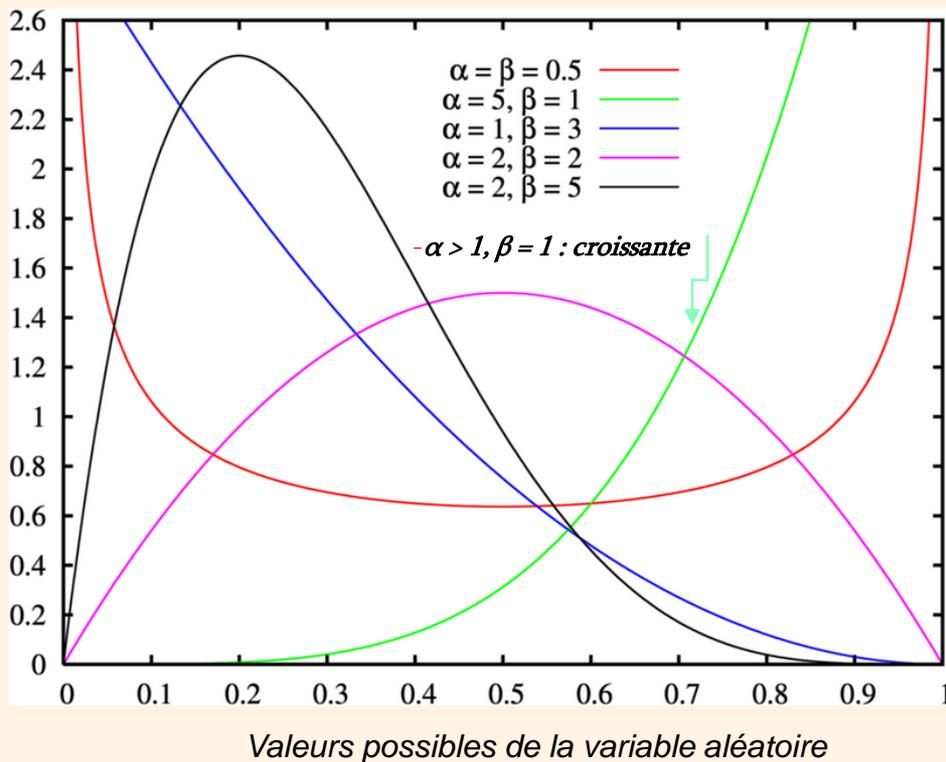
$$V[X] = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

✦ *Caractéristiques de forme* :

- $\alpha = \beta = 1$: Unif(0 ; 1) ; $\alpha = \beta$: symétrique
- $\alpha < 1, \beta < 1$: Courbe en U
- $\alpha = 1, \beta > 1$: décroissante ; $\alpha > 1, \beta = 1$: croissante

Distributions de Probabilité de la loi Beta

Densité de probabilité (pdf)

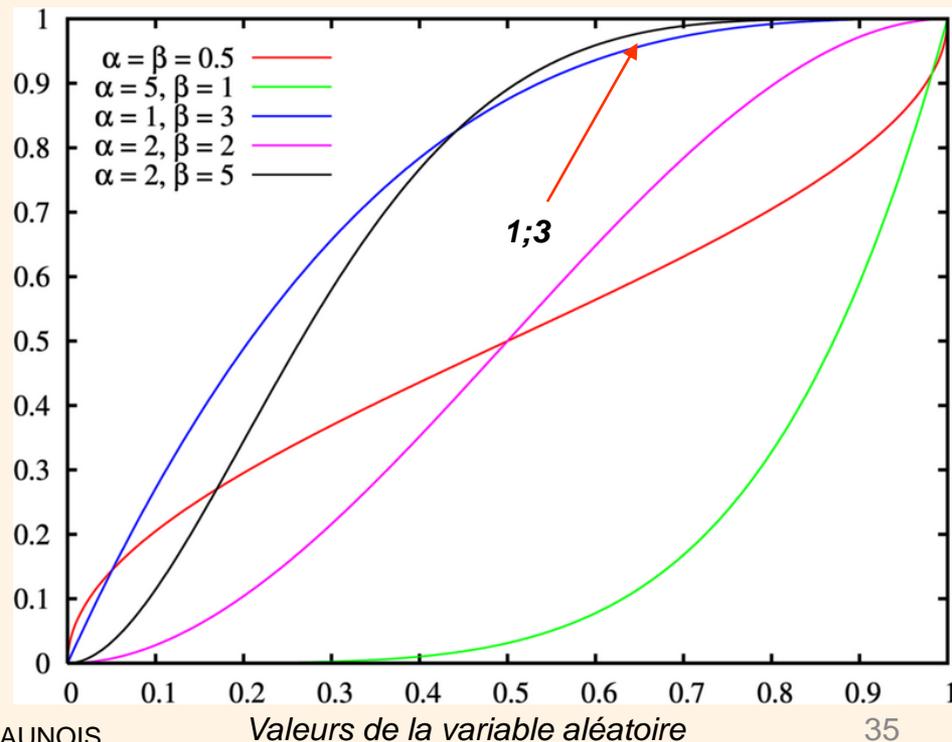


Définitions

- ✦ "Les distributions de probabilité sont associées à des variables numériques continues ou discrètes. Elles précisent comment se répartissent les probabilités associées aux différentes valeurs que peut prendre une variable aléatoire"
- ✦ "Il y a deux manières équivalentes de représenter les distributions : soit par leur densité, soit par leur fonction de répartition ou probabilité cumulée"

J B Denis 2005

Probabilité cumulée (cdf)



Caractérisation des Lois à Partir des Statistiques Descriptives : [Méthode des Moments]

- **les moments** : la moyenne et la variance sont les plus connus (le premier moment = la moyenne d'une distribution, le deuxième moment = la variance). Les moments d'ordre plus élevés sont utilisés pour *caractériser d'autres aspects de la distribution*. Le troisième moment est par exemple lié à l'asymétrie ou la dissymétrie
- **la méthode des moments** consiste à **égaliser les moments (connus) de l'échantillon et les moments correspondants (inconnus) de la distribution** : on estime les moments théoriques par les moments empiriques et on résout le système d'équations.

Calcul des Moments de Beta

■ \bar{x} est la moyenne empirique et s l'écart type de l'échantillon

■ Egalons les premiers moments $\bar{x} = \alpha / (\alpha + \beta)$

■ Puis les seconds $s^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$

■ Et résolvons α p/r aux variables observables et β p/r à α

$$\alpha = \bar{x}(\bar{x}(1 - \bar{x}) / (s^2 - 1))$$

$$\beta = \alpha(1 - \bar{x}) / \bar{x}$$

■ Calcul simplifié: beta est caractérisée par deux paramètres $(r; n-r)$ où r décompte le nombre d'événements, et alpha son complément c.a.d le nombre de non événements $n-r$: beta $(r; n-r)$.

Briggs A, Smdm 2002 ;22:298-308

$$\alpha = \bar{x}(\alpha + \beta); s^2 = \frac{\bar{x}(\alpha + \beta) \times (\alpha + \beta) - \bar{x}(\alpha + \beta)^2}{n^2 \times (n + 1)} = \frac{\bar{x} n^2 \times (1 - \bar{x})}{n^2 \times (n + 1)}; n + 1 = \frac{\bar{x} \times (1 - \bar{x})}{s^2}$$

Variantes

- $p_1 = 48\%$, estimé sur 50 patients

- Méthode des moments: utilisation de la loi normale
 - Si on connaît \bar{x} et son $IC_{95\%} \rightarrow IC = \bar{x} \pm 1.96*s$
 - $EV = 7,9$ ans. $IC_{95\%} = [4,9 ; 10,9]$ ans :
 - $IC = 7,9 \pm 3,0$
 - $IC = 7,9 \pm 1,96*s$ On mesure l'étendue par rapport à une des bornes
 - $\sigma = 3/1,96 = 1,53$ ans
 - $EV \sim \text{Norm}(7,9 ; 1,53^2)$

- Méthode pragmatique : utilisation d'une loi beta
 - $\alpha =$ Nombre de « succès » = 24
 - $\beta =$ Nombre d'« échecs » = 26
 - $p_1 \sim \text{Beta}(24 ; 26)$

Calcul des Moments de Gamma

Soit une loi gamma avec un premier paramètre de forme α et un second paramètre de forme β , la moyenne dans la population est égale à $\alpha\beta$, la variance à $\alpha\beta^2$; \bar{x} est la moyenne empirique et s l'écart type de l'échantillon

$$\theta \sim \text{Gamma}(\alpha, \beta)$$

✦ *Espérance :* $E[\theta] = \alpha\beta$
(Version normale)

✦ *Variance :* $V[\theta] = \alpha\beta^2$
(Version normale)

✦ *Egalisation des moments* $\bar{x} = \alpha\beta$ $s^2 = \alpha\beta^2$

✦ *Résolution p/r variables observées*

$$\alpha = \frac{\bar{x}^2}{s^2} \quad \beta = \frac{s^2}{\bar{x}}$$



***MODÈLE DE DÉCISION INTÉGRÉ
COMBINANT MTC ET ACE***

Choix Structurants du Modèle de Simulation

[Check list de l'évaluation de l'efficience HAS 2013]

- **Type d'analyse économique** : *analyse cout utilité*
- **Choix du critère d'évaluation du résultat** : *RDCR, CAS*
- **Choix de la perspective** : *Systeme de soins*
- **Horizon temporel** : *5 ans :début 2011 à fin 2015*
- **Actualisation** : *4%*
- **Population d'analyse** : *cas prévalents de PR ayant eu une réponse inadéquate à un traitement de fond classique (DMARD)*
- **Stratégies à comparer** :
 - *5 anti-TNF α : adalimumab (Humira[®]), certolizumab pegol (Cimzia[®]), etanercept (Enbrel[®]), golimumab (Simponi[®]), et infliximab (Remicade[®])*
 - *1 anticorps contre le récepteur de (IL6) : tocilizumab (RoActemra[®])*
- **Méthode d'analyse de sensibilité**: *Distribution bayésienne a posteriori*

Modélisation

[Check list de l'évaluation de l'efficience HAS 2013]

- **Type de modèle** : *MCMC sur Multicohorte ouverte*
- **Critères de jugement cliniques** : ACR 50, abandons de traitement, infections
- **Identification des états de santé,**
- **Explicitation des hypothèses structurelles**
- **Durée des cycles** : semaines,
- **Population simulée et sous groupes de population**
- **Probabilités**: méthodes d'estimation et sources,
Paramétrage des distributions de probabilité
- **Explicitation des hypothèses d'extrapolation**
- **Discussion de la validité du modèle**

■ **Méthode de valorisation des état de Santé**

- Description des conséquences en termes de santé y compris les EI
- Estimation de l'espérance de vie
- Estimation des qaly's (source et méthode)
- Résultats présentés en espérance de vie et en utilité

■ **Méthode de valorisation des coûts**

- Liste des catégories de coût prise en compte
- Dénombrement des ressources utilisées
- Présentation des couts unitaires
- Valorisation
- Résultats présentés par poste de coût

Analyse de Sensibilité

[Check list de l'évaluation de l'efficience HAS 2013]

- Type d'analyse de sensibilité: DSA, PSA
- Choix des bornes ou des distributions
- Représentation de trois courbes d'acceptabilité pour trois niveaux de prix différents du produit
- Présentation des résultats de ces analyses

Présentation des Résultats

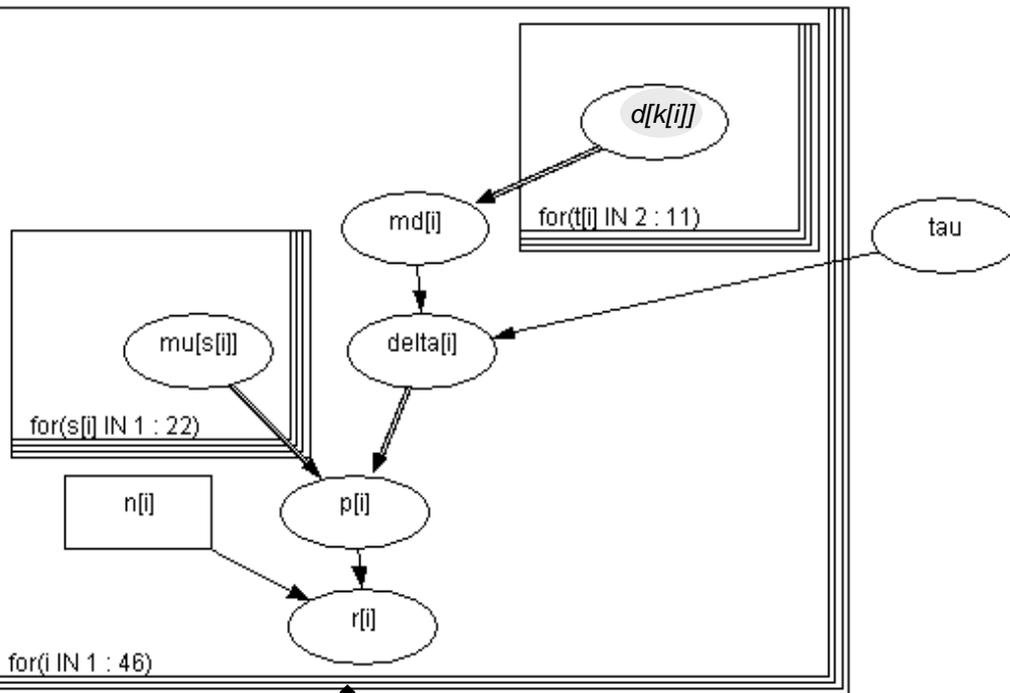
[Check list de l'évaluation de l'efficience HAS 2013]

- représentation graphique de la Frontière d'efficience
- caractérisation du degré d'incertitude associé aux résultats quantitatifs
- discussion de l'impact de l'incertitude sur les résultats

Modèle MTC Nombre de Réponse ACR50

Graphe Orienté Acyclique [DAG directed acyclic graph]

- Le raisonnement inductif de l'Inférence Statistique, consiste à parcourir le graphe dans le sens inverse des flèches. Il s'agit de partir des données et de faire remonter l'information vers les paramètres
- Le raisonnement déductif de la Modélisation parcourt le graphe dans le sens des flèches, Il s'agit de descendre des paramètres vers les données
- Les variables aléatoires sont représentées sous la forme d'ellipses. une distribution de probabilité leur est associée
- Les données apparaissent dans un carré ou un rectangle.
- flèches simples : relations stochastiques conditionnelles ;
- flèches doubles : relations logiques déterministes.
- L'empilement des feuilles symbolise le nombre d'itérations.
- Les variables indépendantes (les paramètres) sont à l'extérieur du cadre



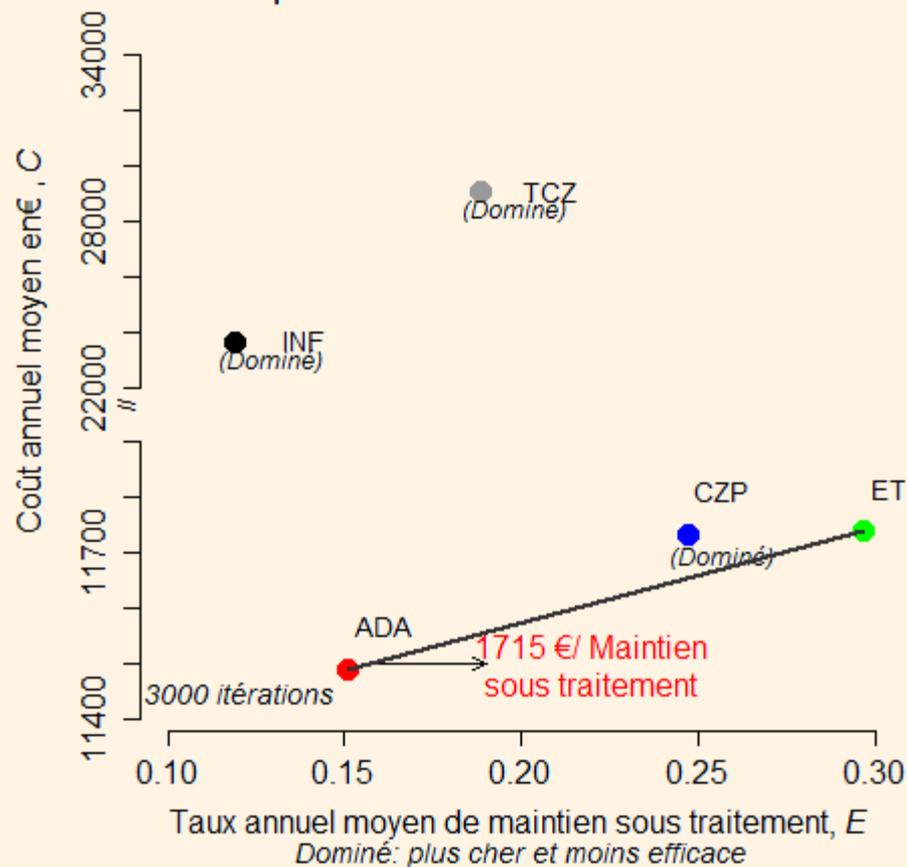
Indices

- * i : bras 46, $s[i]$: études 22, $k[i]$: traitements 7+3+Pb
- * $r[i]$: nombre de répondeurs ;
- * $p[i]$: probabilité de réponse ;
- * $n[i]$: effectif ;
- * $\mu[s[i]]$: le log de la **cote** (logOdds) pour chaque étude;
- * $\delta[i]$: log du **rapport de cotes** (logOddsRatio);
- * $md[i]$: moyenne « vraie » de la distribution dans la population de l'effet relatif du ttx k du bras $[i]$ par rapport au ttx de référence, mesuré en log du rapport de cotes (logOR)
- * $d[k[i]]$: distribution dans la population de l'effet relatif du ttx k du bras $[i]$ par rapport a un ttx de référence mesuré en log du rapport de cotes (logOR) = une différence de log de cote,
- * τ : la précision du logOR.

Plan CE et Frontière d'Efficiéce

- Adalimumab et etanercept forment la frontiére d'efficiéce
- Certolizumab pegol, infliximab et tocilizumab sont forteméce dominés
- Le coût additionnel annuel par rapport à adalimumab d'un maintien en rémission plus long sous etanercept est de 1715€ en moyenne

Coût-Efficacité des Traitements anti-TNF en 1ère Ligne après Echec d'un Traitement de Fond



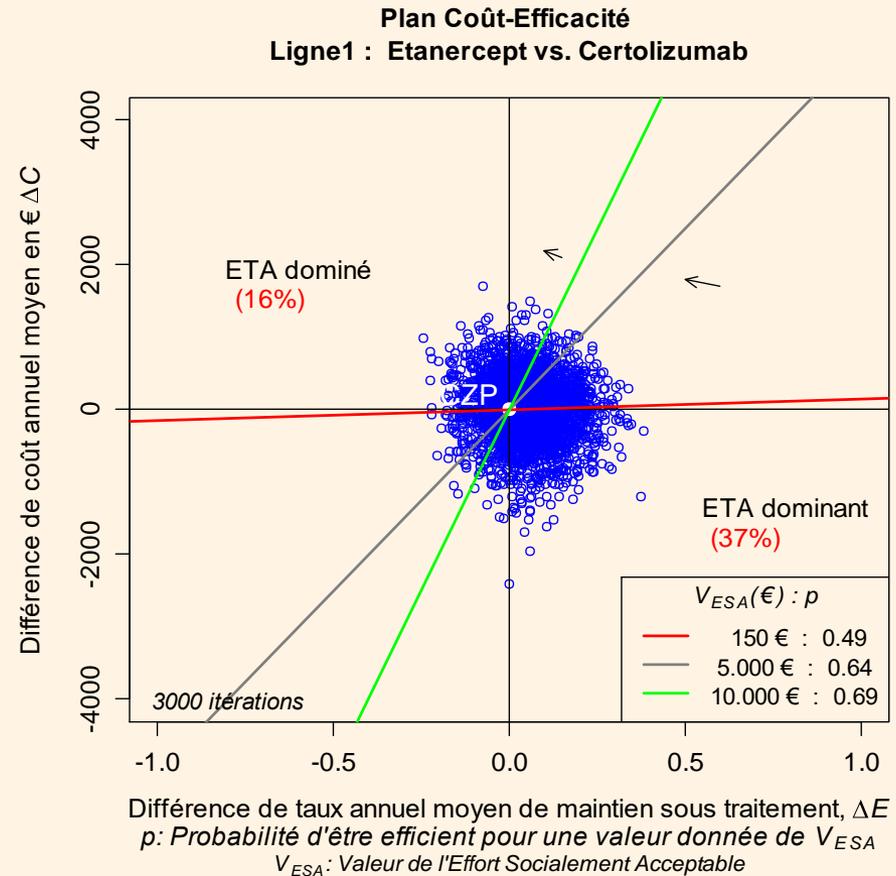
CZP, INF, TCZ Sont Plus Onéreux et Moins Efficaces



Plan CE et Volonté de Payer

Etanercept (ETA) vs. Certolizimab (CZP)

- Dans 37% des simulations, etanercept domine certolizumab pegol. Il est dominé dans 16% des cas.
- Pour une valeur donnée de l'effort socialement acceptable (V_{ESA}) égale à 150€, la probabilité que etanercept soit plus efficace que certolizumab pegol est de 0,5.
- Lorsque que la V_{ESA} est estimée à 5000€, cette probabilité est de 0,64.



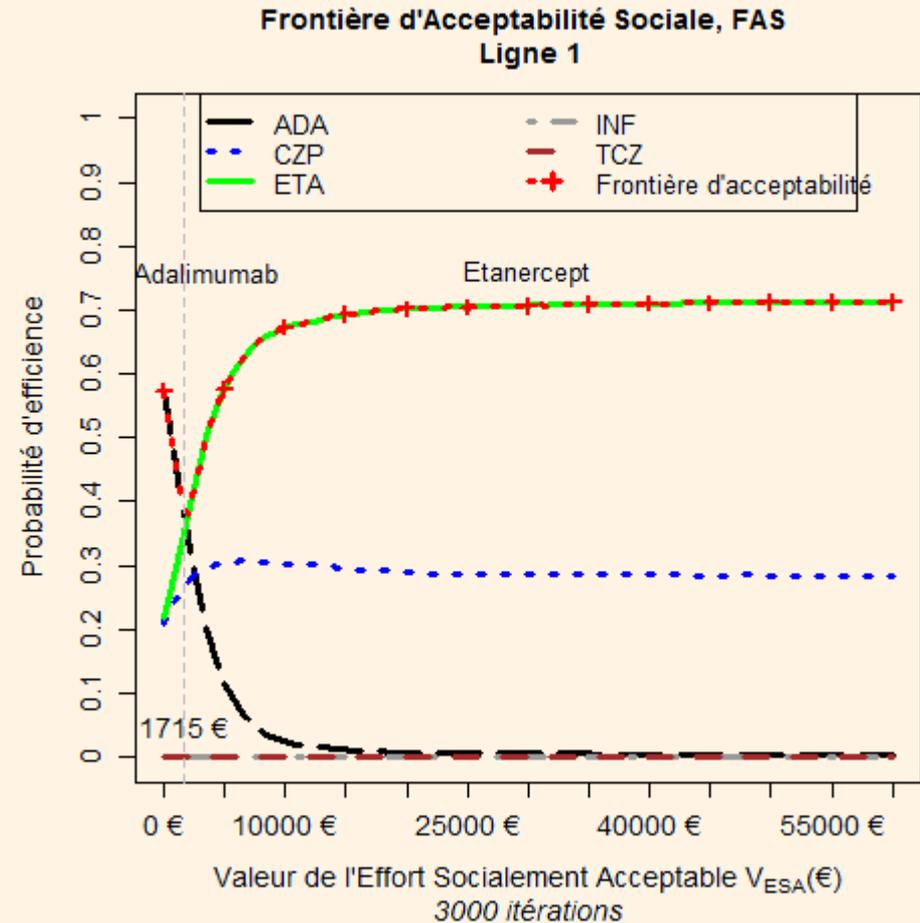
Launois R, Le Moine JG, Huynh TM ; ISPOR Washington 2012



Frontière d'Acceptabilité Sociale

Comparaison Multitraitements

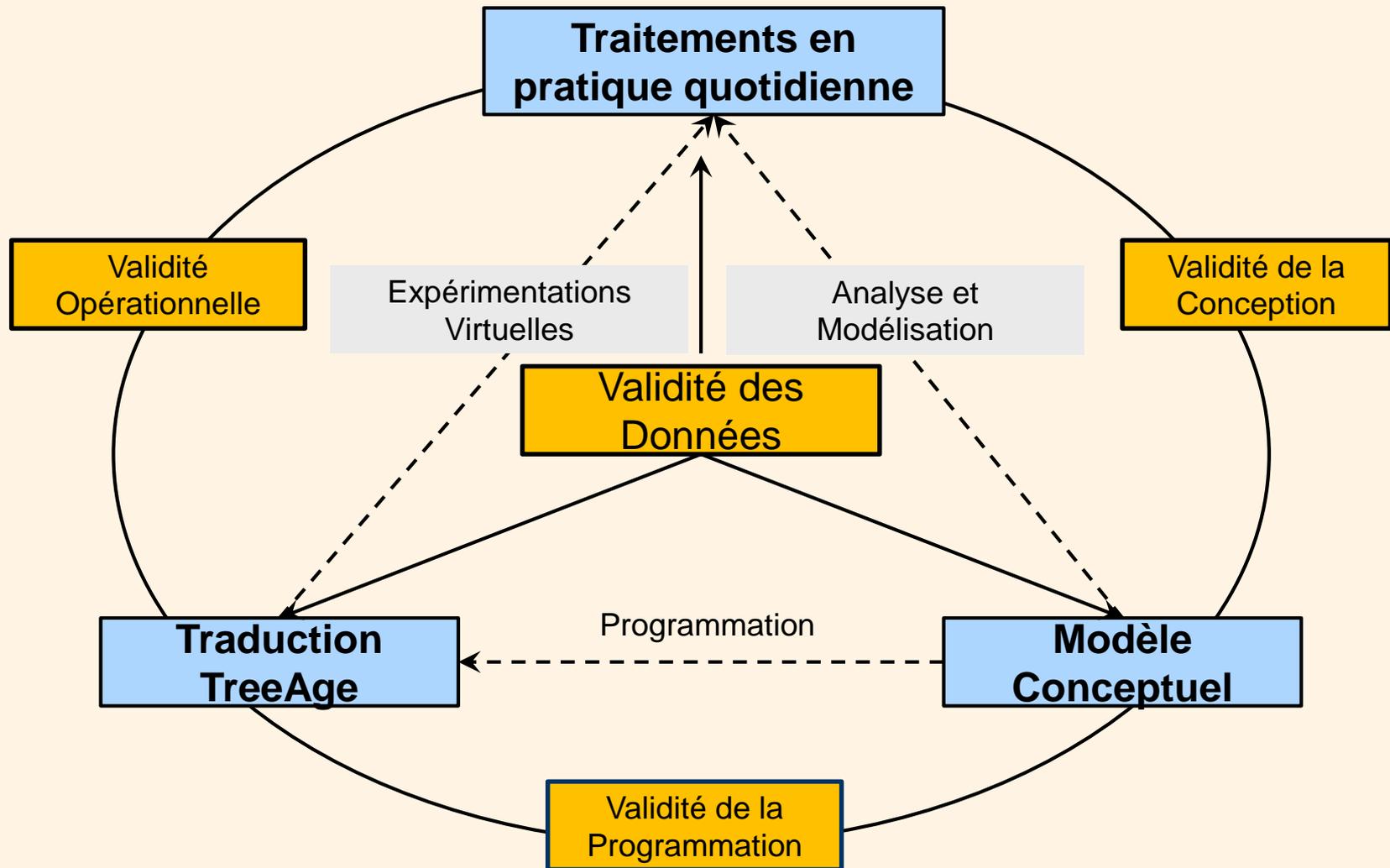
- Adalimumab et etanercept sont situés sur la frontière d'acceptabilité sociale.
- Adalimumab représente le traitement le moins couteux, etanercept représente le traitement le plus efficace.



Launois R, Le Moine JG, Huynh TM ; ISPOR Washington 2012



Valider la Simulation Numérique Produite Par le Logiciel

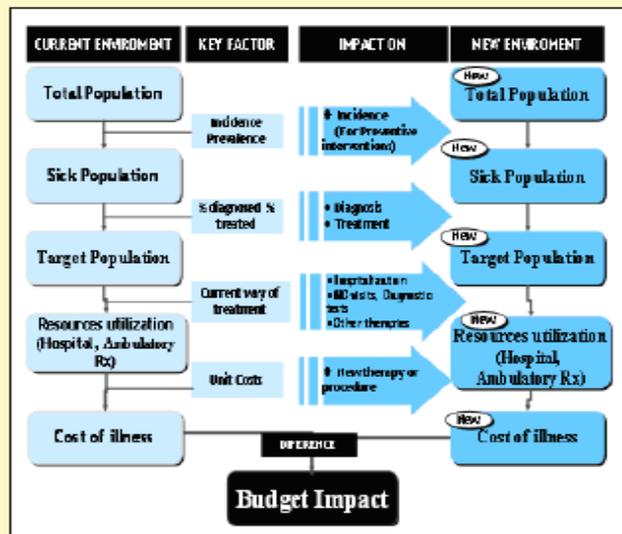


3) POUR LE CEPS: BASES DE DONNÉES MÉDICO ADMINISTRATIVES & EVALUATION DE L'IMPACT BUDGETAIRE



Principles of Good Practice for Budget Impact Analysis

Task Force lead by Jo Mauskopf / Sean Sullivan / Dan Mullins



RECOMMENDATIONS FOR INPUTS AND DATA SOURCES

Six key elements requiring inputs for BIA framework:

- Size and characteristics of affected population
- Current intervention mix without the new intervention
- Costs of current intervention mix
- New intervention mix with the new intervention
- Cost of the new intervention mix
- Use and cost of health condition and treatment-related health care services

Les Magasins du SNIIR-AM

Système National d'Informations Inter-Régimes de l'Assurance Maladie

Un accès organisé par grands domaines :

- **Architecture :**



Datamarts

- **Les outils :**

Outils requêteurs (SPSS, SQL+, SAS), mais aussi analyses statistiques complexes (*multidimensionnel, datamining*)

Données Anonymes et Individuelles

L'entrepôt SNIIR-AM

■ Composantes

- *L'entrepôt inter-Régimes*
- *L'entrepôt national Régime général (ERASME national)*
- *L'échantillon de bénéficiaires au 1/97 : (500 000 personnes; actuellement uniquement Régime général; à terme échantillon inter-Régimes); depuis 2003 (échantillon pérenne)*
- *Les données de consommation inter-régimes : DCIR*

■ Contenu

- *Données de remboursement*
- *Codes CIM 10 pour les patients en ALD et en invalidité*
- *Chaînage avec PMSI (RSA)*
- *Codages médicament (CIP), biologie, actes (CCAM), dispositifs (LPP)*
- *Anonymes (patients)*
- *Nominatives (professionnels de santé et établissements)*

■ Les accès autorisés à ce jour

- *L'entrepôt inter-Régimes : pas d'accès CNAMTS en routine*
- *L'entrepôt national Régime général (ERASME national) : accès CNAMTS en routine (demande CNIL pour l'Etat, les agences, les chercheurs)*
- *L'échantillon au 1/97 : accès autorisé y compris pour l'Etat, les agences, les chercheurs*

Source Philippe Ricordeau

Données Anonymes et Agrégées

les Magasins de l'entrepôt SNIIR-AM

■ Contenu

Le magasin (Datamart) « offre de soins », inclut par exemple...

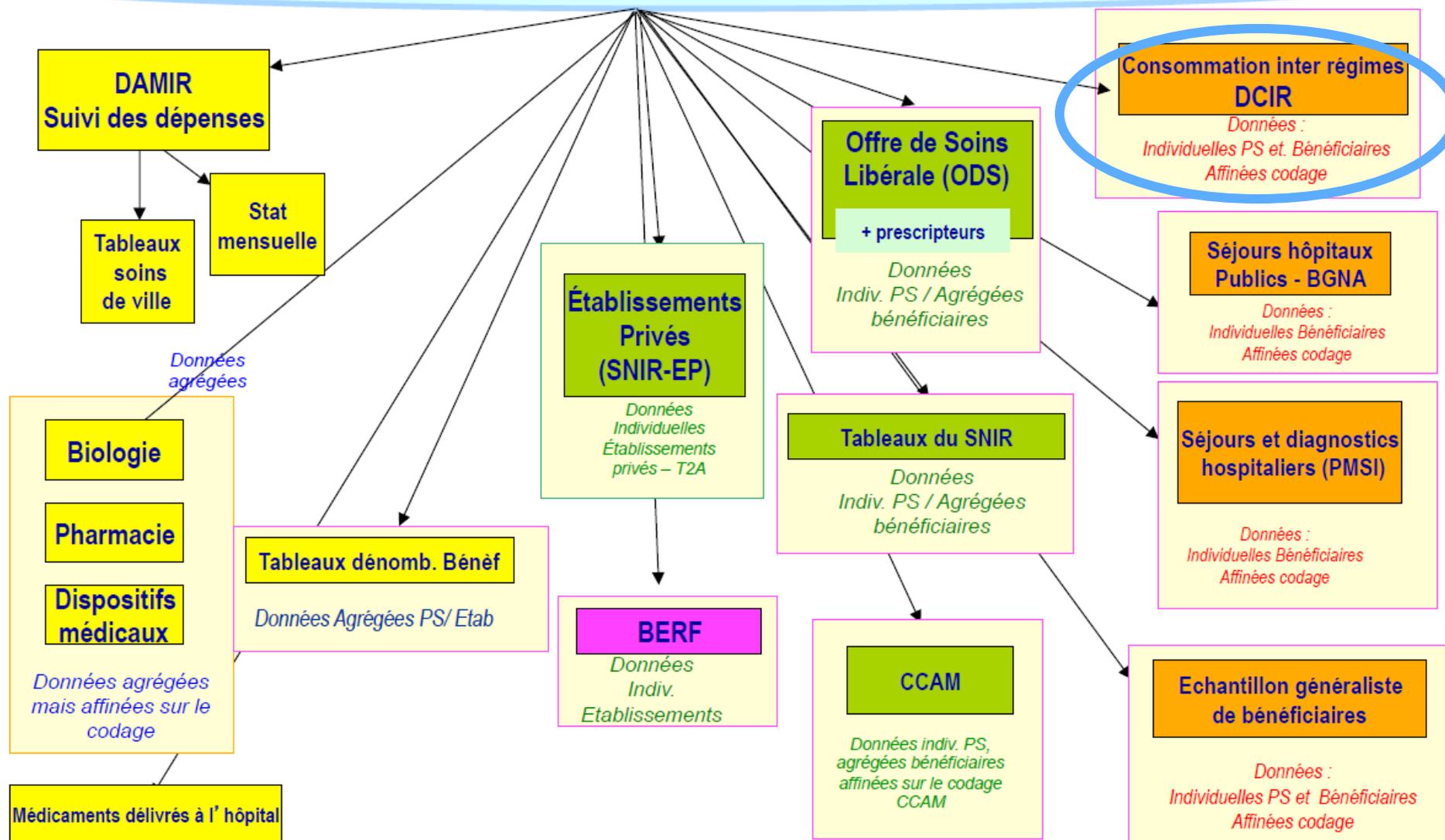
- *Tableaux de démographie, d'activité et d'honoraires des professionnels de santé libéraux*
- *Tableaux de bord du codage de la pharmacie,*
- *Tableaux de bord sur le codage de la biologie,*
- *Bases de données individuelles par professionnel (en cours de mise en production)*

■ Les accès autorisés à ce jour

- *Organismes d'assurance maladie*
- *Etat et services déconcentrés (Ministères, DRESS, ARH)*
- *Les agences ou équivalents (InVS, AFSSAPS, HAS, HCAM, INCA)*
- *Les professionnels de santé et les URML*
- *Certains organismes de recherche (INSERM, IRDES, CNRS, CETAF)*

Source Philippe Ricordeau

Données disponibles par le portail SNIIR-AM



Données agrégées bénéficiaires et d'offres de soins
Pas de limite de conservation

Données agrégées bénéficiaires et individualisées offre de soins
Conservation : 10 ans

Référentiels

Données individualisées bénéficiaires et offre de soins
Conservation 3 années + année en cours si PMSI 10 ans
20 ans pour échantillon EGB

RECAP : Conditions D'Accès SNIIR-AM

Arrêté de juillet 2013

		Données exhaustives individuelles anonymisées	Extractions (échantillons spécifiques)	Echantillon Généraliste de Bénéficiaires	Données agrégées
CNAMTS ⁶ , CCMSA ⁶ , RSI ⁶		Accès possible	Autorisation CNIL	Accès possible	Accès possible
Organisme poursuivant un but non lucratif	CNSA, médecins salariés des ARS (ex URCAM et ARH) ³ , INVS ⁵ , ANSM, HAS	Accès possible	Approbation IDS et autorisation CNIL	Accès possible	Accès possible
	Ministères, agences (ATIH, HCAAM...), grands organismes de recherche (CNRS, INSERM, IRDES...)	Pas d'accès		Accès possible	Accès possible
	Autres organismes (exemple : CHU, ORS Université, etc...)	Pas d'accès		Approbation IDS	Approbation IDS
Les membres de l'IDS		Pas d'accès	Pas possible ? / Approbation IDS et autorisation CNIL ?	Accès possible	Accès possible
Structures adhérant aux membres de l'IDS ou les constituant		Pas d'accès		Pas d'accès ⁵	Accès possible
Unions Régionales de Professionnels de Santé (toute profession de santé)		Pas d'accès		Pas d'accès	Accès possible ³
17 organismes complémentaires d'assurance maladie participant au projet MONACO		Pas d'accès		Pas d'accès	Accès possible
Organisme poursuivant un but lucratif (exemple : industrie du médicament et des produits de santé)		Pas d'accès		Pas d'accès	Pas d'accès

Quelle Procédure Pour les Industriels?

- **Pour les données individuelles** : Demande d'accord
 - Relève de l'IDS (Institut des Données de Santé), au titre du chapitre X de loi du 6 janvier 1978, L'IDS définit son propre dossier type d'avis pour les demandes d'accès.
 - Le comité technique de l'IDS a déjà retenu les critères suivants : La charge de travail liée à l'accès aux données du SNIIRAM côté CNAMTS (extraction....). La méthodologie, la qualité technique et scientifique de l'étude, le respect d'une charte de déontologie
- **Pour les données agrégées** : Avis du COPIIR (Comité de pilotage Inter régimes) qui décide
 - Extraction par les services
 - Accès temporaire limité

ERASME Régional : Une éventuelle Solution de Dépannage Après Anonymisation

Données Nominatives Individuelles

- *Réparties dans les entrepôts régionaux,*
- *Identifiant en clair du bénéficiaire*
- *Données du seul Régime général (hors sections locales mutualistes comme la MGEN, les fonctionnaires, etc)*
- *Pas de chaînage PMSI,*
- *Pas de chaînage en routine avec les ALD,*
- *Historique de 24 mois glissants*
- *Accès réservé aux **utilisateurs de l'assurance maladie***

Source Philippe Ricordeau

Analyse Coût Efficacité vs Impact Budgétaire

Coût-efficacité

Impact budgétaire

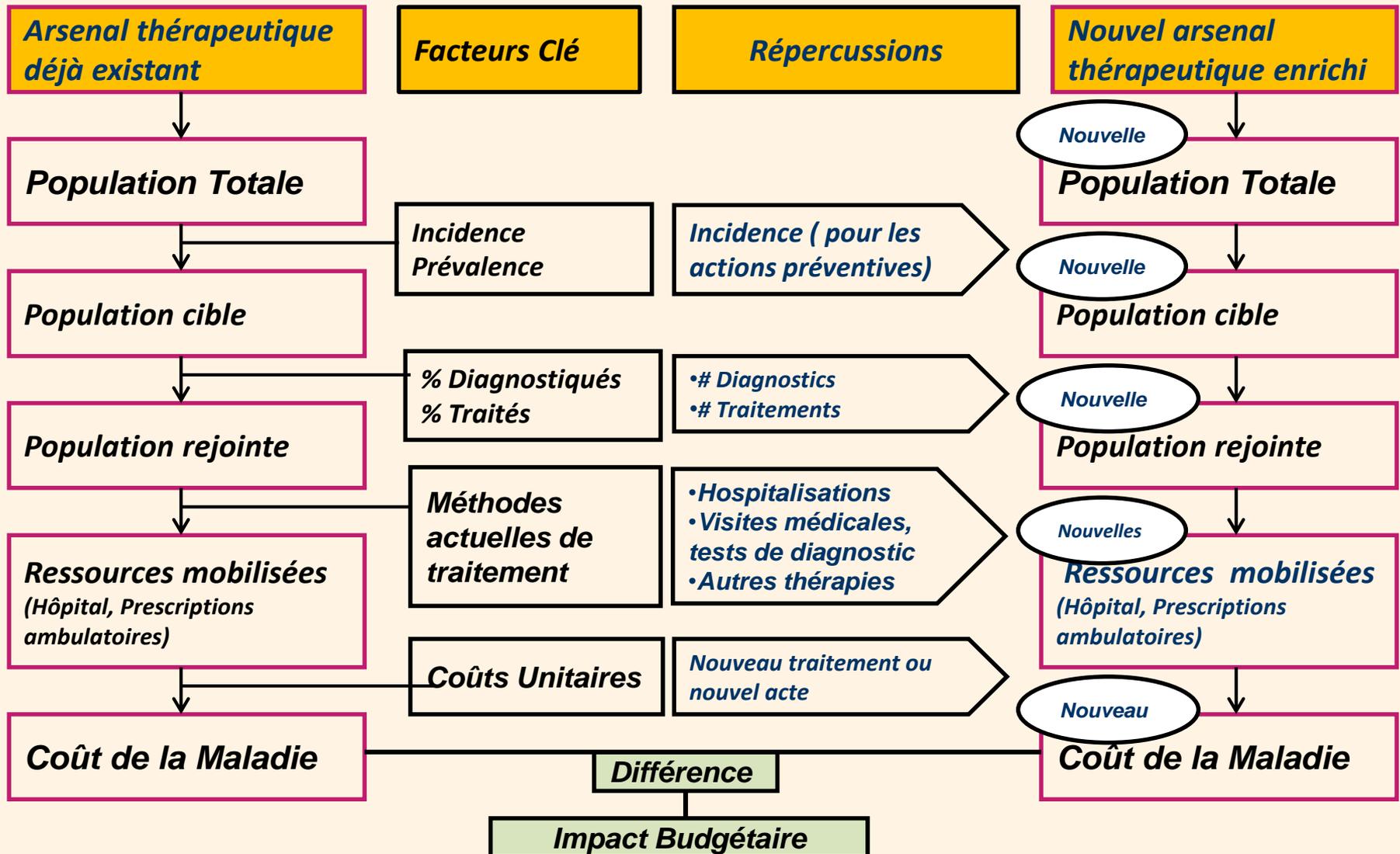
C vs N

(O + C) vs (O + C + N)

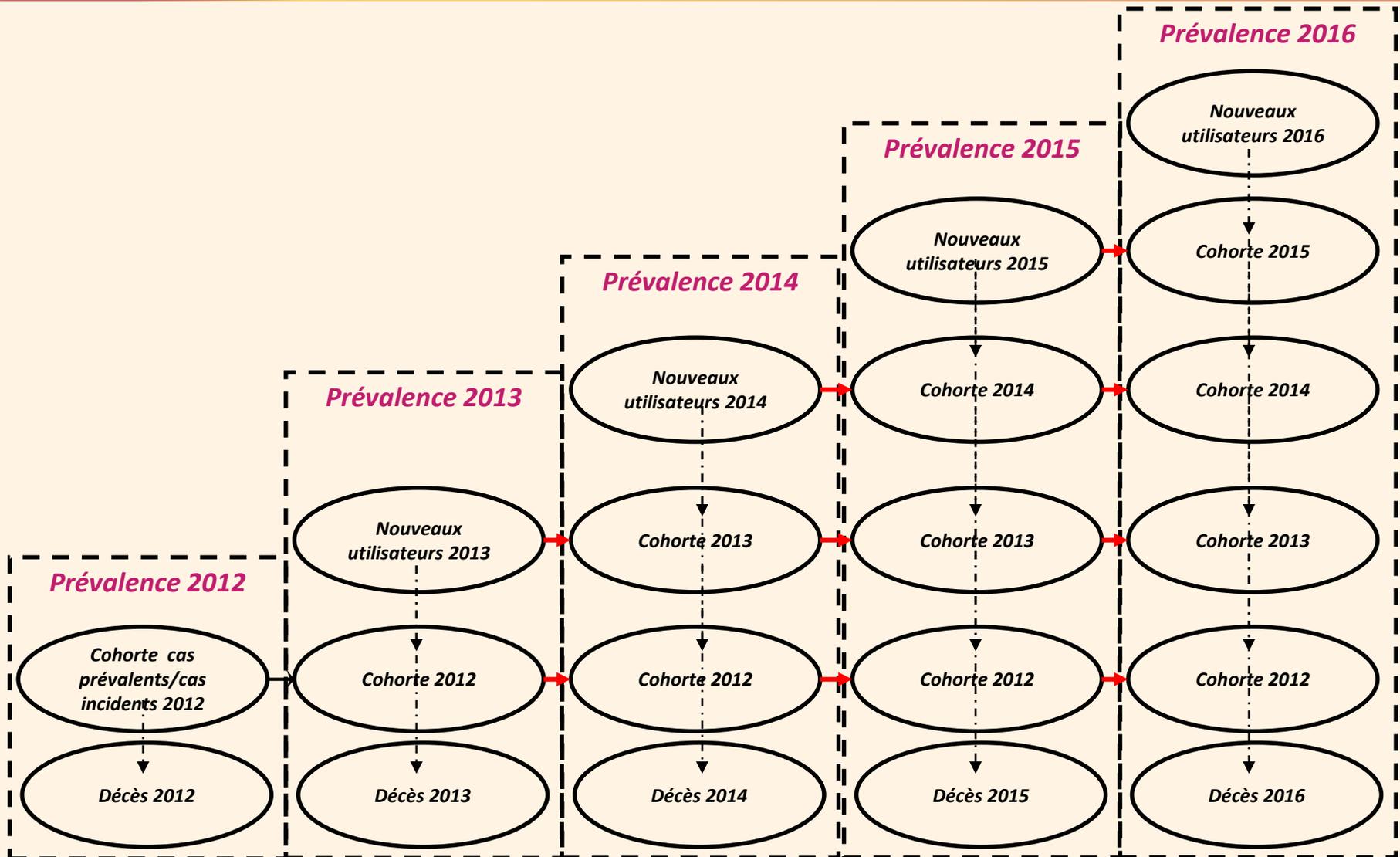
Légende O: malades non traités; C: traitement conventionnel; N : traitement nouveau; O → N : élargissement du du marché ; C → N: substitution et accroissement des parts de marché

Adapté de Ethgen. Cohort, Multi-cohort or Population Model? Why, When and How? ISPOR 2011

Une Comparaison Toutes Stratégies Confondues

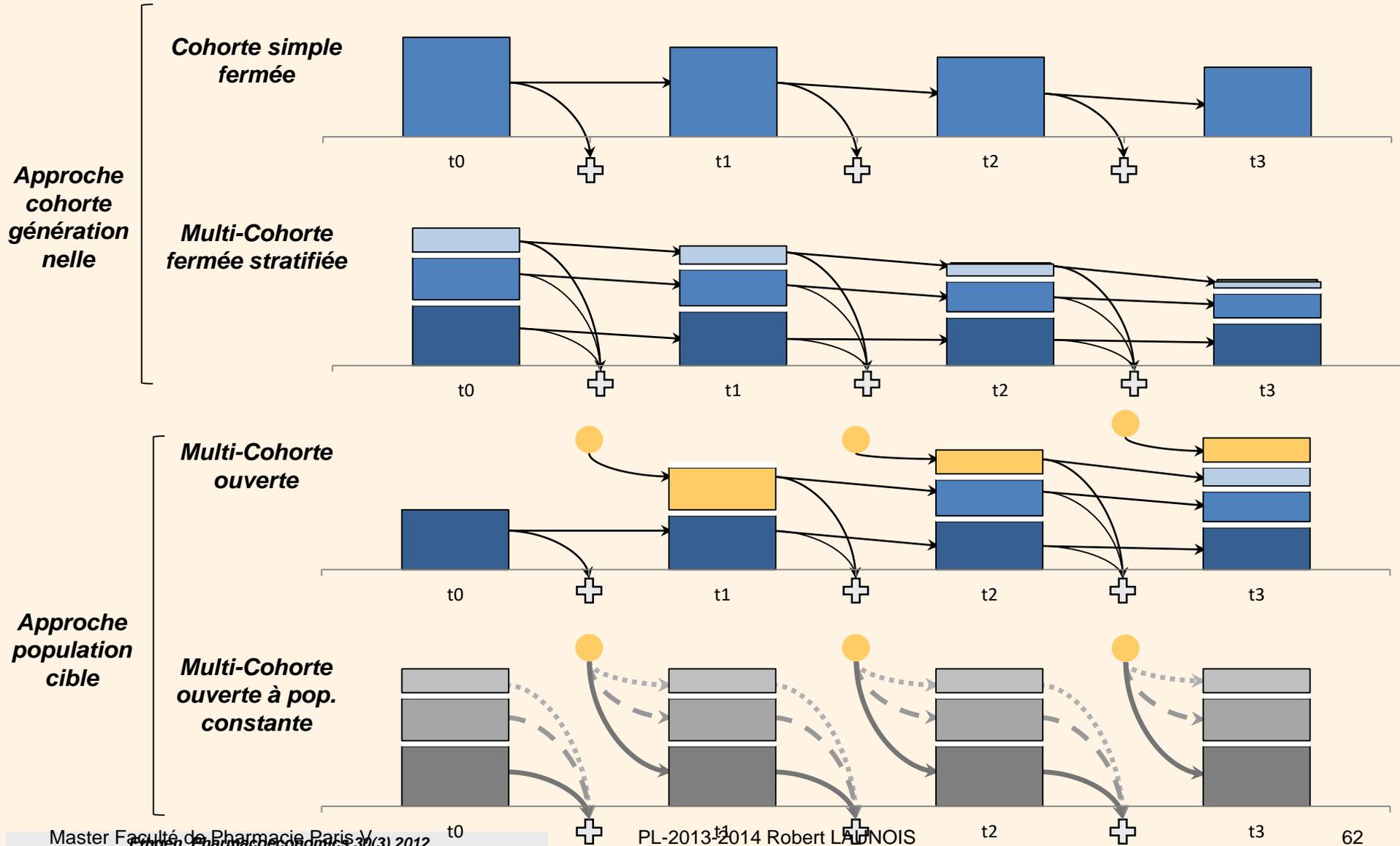


Principe de l'Analyse Multi-Cohortes



Source : R Launois et al. « L'évaluation des technologies médicales : les protocoles de deuxième génération » 61
JEM 2006

Cohorte Simple vs Multi-Cohortes



Des Approches Radicalement Différentes

	AIB- Risque sharing	Analyse coût efficacité
Perpectives	Acheteur	Assurances/Société
Rationnel	Risque financier	RDCE < λ
Épidémiologie	Prévalence	Incidence
Modélisation	Multicohorte fermée/ouverte	Cohorte simple
Fenêtre temporelle	Court terme	Vie entière
Dynamique populationnelle	Rôle clé (volume)	Variable omise
Dynamique du marché	Substitution /addition	Variable omise
Dynamique maladie	RCT, BDMA	RCT, BDMA
Etats de santé	NON	QALYs
Coût	Prix*volume*CPR[†]	Coût communautaire
Actualisation	NON	4%

[†] CPR/Contrats de Partage de Risque

Sélection Bibliographique

- ++ Annemans L. Budget impact analysis. *Pharmaco-economics : from theory to practice* 2009.
- CES. Guide méthodologique pour la mise en place d'une analyse d'impact budgétaire. Paris: Collège des économistes de la santé;2008.
- CES. Guide méthodologique pour l'évaluation économique des stratégies de santé - Recommandations méthodologiques. Paris: Collège des économistes de la santé;2003.
- Cohen JP, Stolk E, Niezen M. Role of budget impact in drug reimbursement decisions. *J Health Polit Policy Law*. Apr 2008;33(2):225-247.
- ++ **HIQA. Guidelines for the Budget Impact Analysis of Health Technologies in Ireland. Dublin: Health Information and Quality Authority;2010.**
- HAS (Haute Autorité de Santé). Polyarthrite Rhumatoïde : Synthèse des recommandations. Recommandations professionnelles, septembre 2007
- Huang E, Esraïlian E, Spiegel BM. The cost-effectiveness and budget impact of competing therapies in hepatic encephalopathy - a decision analysis. *Aliment Pharmacol Ther*. Oct 15 2007;26(8):1147-1161.
- ++ **Launois R., Avouac B., Berenbaum F., Blin O., Bru I., Fautrel B., Joubert J.-M., Sibilia J., Combe B., Comparison of Certolizumab Pegol with other anti-cytokine agents for the treatment of rheumatoid arthritis : a multiple-treatment bayesian meta-analysis, The Journal of Rheumatology,2011 May;38(5):835-45**
- ++ **Launois R, Payet S, Saidenberg-Kermanac'h N, Francesconi C, Franca LR, Boissier MC. Budget impact model of rituximab after failure of one or more TNFalpha inhibitor therapies in the treatment of rheumatoid arthritis. Joint Bone Spine. Dec 2008;75(6):688-695.**
- ++ Launois R, Payet S, Riou Franca L, et al. L'évaluation des technologies de santé : les protocoles de deuxième génération. *Journal d'Économie Médicale*. 2006;24(5).
- Launois R, Payet S. Évaluation pharmacoéconomique des épisodes dépressifs majeurs et des récides. *Neuropsychy news*. 2003;2(5):209-221.
- ++ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. Oct 30 2004;23(20):3105-3124.
- Maravic M « Economic impact of rheumatoid arthritis (RA) biotherapies in France » *Joint, bone, spine : revue du rhumatisme* 2010;77(6):546-51.
- Mauskopf JA, Sullivan SD, Annemans L, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on good research practices--budget impact analysis. *Value Health*. Sep-Oct 2007;10(5):336-347.
- Mauskopf JA, Earnshaw S, Mullins CD. Budget impact analysis: review of the state of the art. *Expert Rev Pharmacoecon Outcomes Res*. Feb 2005;5(1):65-79.
- Mauskopf J. Meeting the NICE requirements: a Markov model approach. *Value Health*. Jul-Aug 2000;3(4):287-293.
- ++ **Mauskopf J. Prevalence-based economic evaluation. Value Health. Nov 1998;1(4):251-259.**
- NICE. Developing costing tools - Methods guide. London: National Institute for Health and Clinical Excellence;2008.
- Niezen MG, de Bont A, Busschbach JJ, Cohen JP, Stolk EA. Finding legitimacy for the role of budget impact in drug reimbursement decisions. *Int J Technol Assess Health Care*. Jan 2009;25(1):49-55.
- ++ **Orlewska E, Gulacsi L. Budget-impact analyses: a critical review of published studies. Pharmacoeconomics. 2009;27(10):807-827.**
- Orlewska E, Mierzejewski P. Proposal of Polish guidelines for conducting financial analysis and their comparison to existing guidance on budget impact in other countries. *Value Health*. Jan-Feb 2004;7(1):1-10.
- Shaya T, Mullins CD, Wong WC. Clinical, Epidemiological, and Economic Dimensions of Formulary Decision Making: PART 2 of 2. *ISPOR connections*. 2003;9(1):1.
- Shaya T, Mullins CD, Wong WC. Clinical, Epidemiological, and Economic Dimensions of Formulary Decision Making: PART I of 2. *ISPOR connections*. 2002;8(5):1.

4) POUR LA RÉÉVALUATION : EFFECTIVITÉ ET EFFICIENCE COMPARATIVES DANS LE MONDE REEL



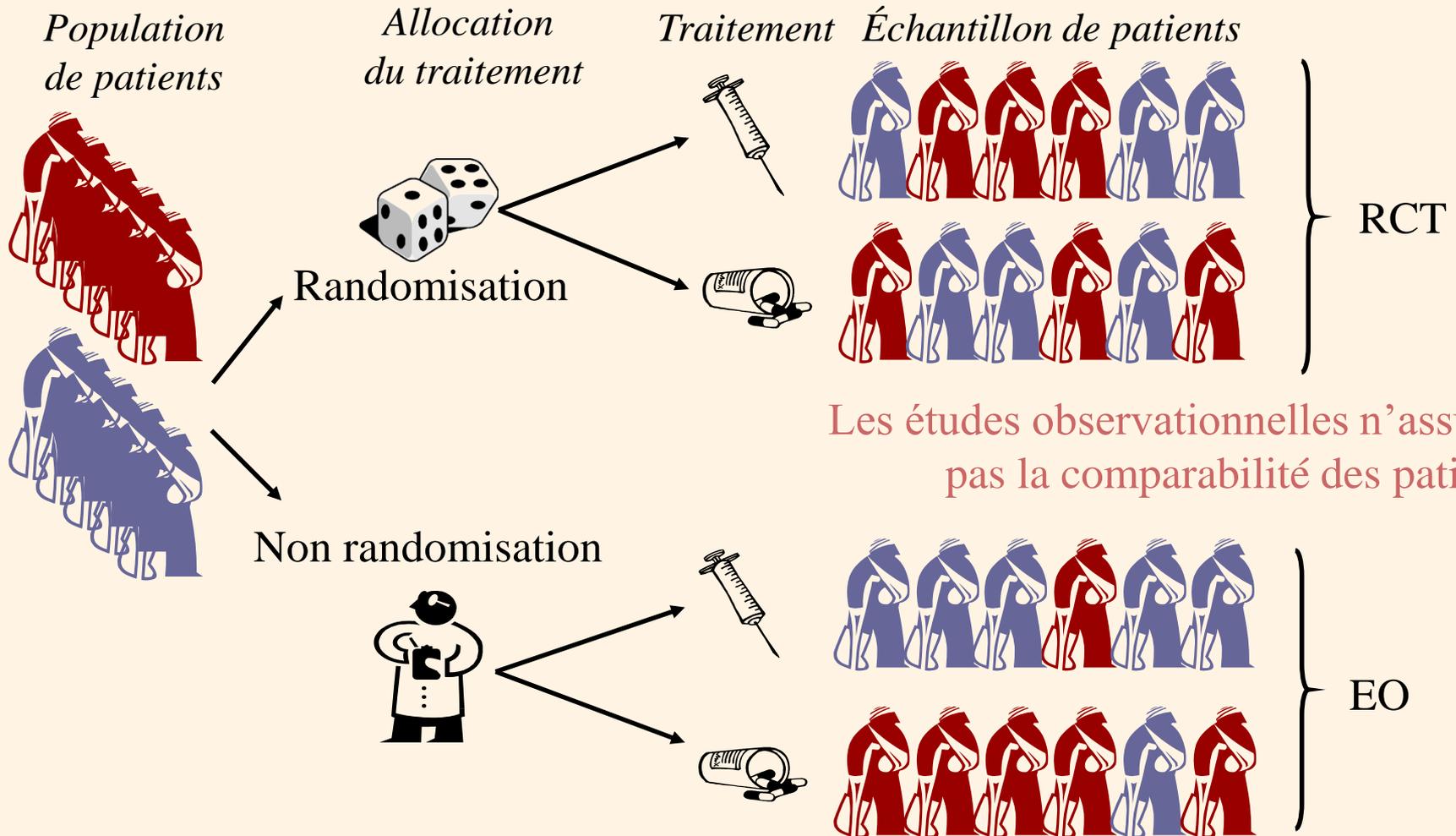


*LES FAITS NE SONT PAS
TOUJOURS
DES PREUVES*

Evaluations Non expérimentales, Biais et Facteurs de Confusion

- Un biais, c'est une *erreur systématique* qui introduit un écart entre la valeur estimée des variables d'intérêt et leur vraie valeur, ce qui entraîne une interprétation erronée du résultat.
- Un facteur de confusion est un *mélange d'effets* (facteur tiers et facteur d'exposition) qui biaise l'interprétation des résultats de l'étude, parce que *l'association simultanée du facteur tiers*
 - avec l'exposition (i.e le programme, le traitement)
 - et avec la survenue de l'événementfait croire faussement que le facteur d'exposition est le fait générateur de la survenue de l'événement. Ex cheveux gris /infarctus /age

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
 - l'indication du médecin est posée en raison des caractéristiques **observables** cliniques, fonctionnelles ou comportementales des patients. Or ces paramètres sont autant de facteurs pronostiques de l'évolutivité de la maladie
 - l'indication est sélectionnée par le médecin sur la base de facteurs de risque subtils (par ex le désir d'apporter une réponse à une souffrance psychologique ou sociale) qui sont autant de facteurs parasites **inobservables**
 - Biais d'indication: il convient de **distinguer et d'isoler l'effet des caractéristiques observables ou inobservables des sujets de l'effet du traitement sur le critère de jugement**



LA CHASSE AU BIAIS

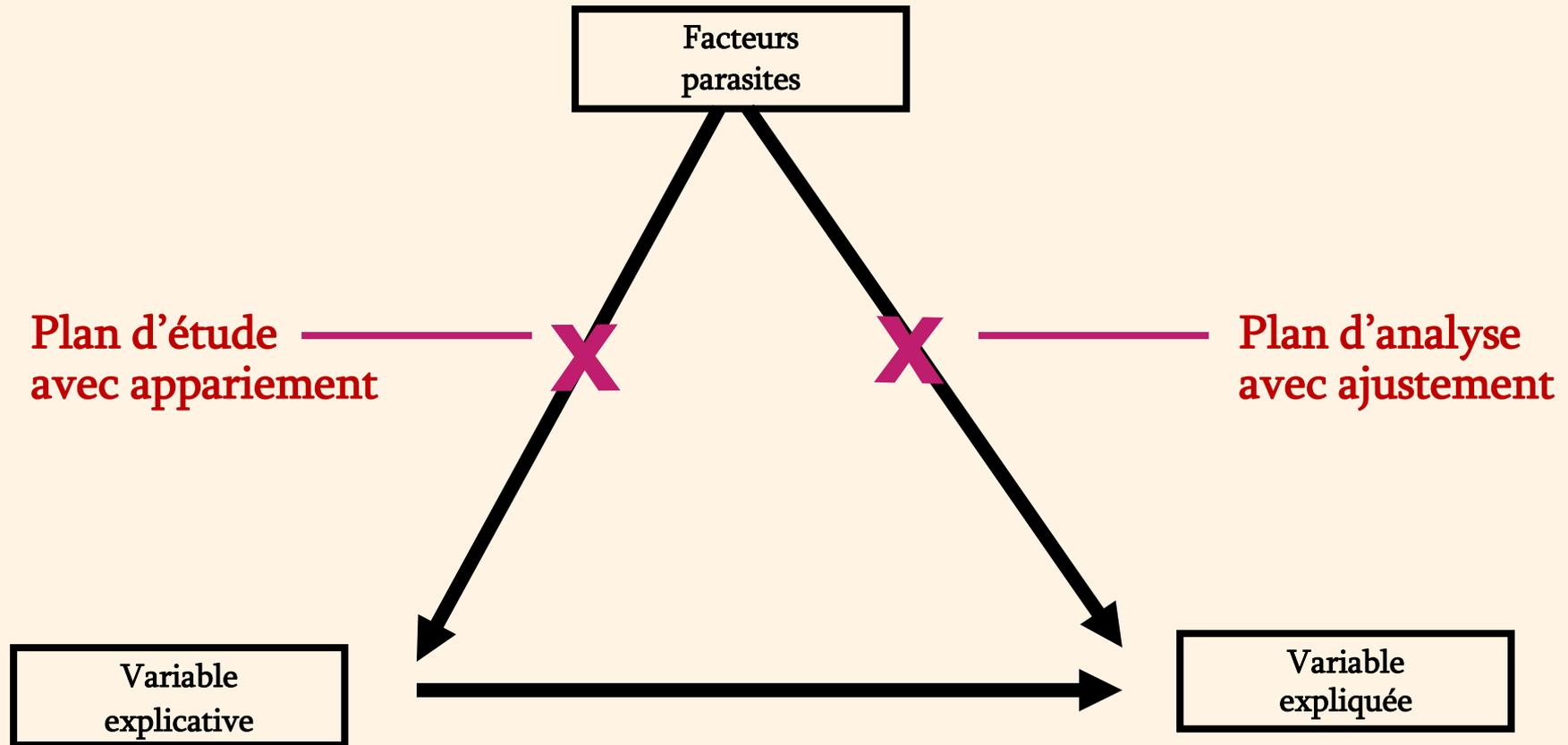
Un Impératif : Travailler sur des Populations Comparables

- 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 caractéristiques favorables,
 - Et que ces caractéristiques auraient joué en leur faveur,
 - même s'ils n'avaient pas reçu le traitement.
- Si ce traitement s'avère bénéfique, on doit s'attendre à ce qu'ils retirent de son fait un avantage **encore plus grand** que celui qui eut été le leur si ce traitement ne leur avait pas été prescrit □ Le **VRAI** bénéfice **du traitement est surestimé.**
- La tâche de l'évaluateur est donc de reconstituer un groupe contrôle à partir du groupe qui a été défini sur des critères d'appartenance prédéfinis et non aléatoires (territoriaux, socio-professionnels), en **sélectionnant des individus ressemblants autant que possible** à ceux du groupe auxquels le traitement a été administré.

- **Au moment du choix du schéma d'étude**
 1. concevoir le design de l'étude indépendamment de la variable d'intérêt, en modélisant *a priori* la relation entre les caractéristiques observables du patient et le modèle d'exposition au traitement (score de propension)
 2. construire ensuite le modèle d'effet causal : exposition au ttx → résultats

- **Au moment de l'analyse**, modéliser *a posteriori* la relation entre les caractéristiques observables du patient et la survenue de l'événement d'intérêt (analyse de régression)

Appariement *a priori* ou Ajustement *a posteriori*





CONSTRUCTION DU SCORE DE PROPENSION

L'Appariement Direct de Toutes Les Facteurs de Confusion 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
- $10 * 2 * 2^5 * 2^5 * 5 = 102\ 400$ appariements possibles

Le Score de Propension Permet d'Exprimer l'Ensemble des Facteurs de Confusion par un Scalaire

■ Définition :

- Score de Propension SP : Modélise le processus selon lequel le médecin a une propension à choisir le ttx en fonction des caractéristiques initiales du patient (les covariables ou les facteurs de confusion)
- = Probabilité pour un sujet de recevoir le traitement étudié, quel que soit le groupe dont il relève, compte tenu des paramètres cliniques ou biologiques qu'il présente initialement = $P(\text{traitement} \mid \text{caractéristiques du patient})$
- Si le score $SP_{\text{Traitement}} = \text{le score } SP_{\text{Contrôle}} \rightarrow$ Les sujets traités et les sujets non traités auront, aux fluctuations d'échantillonnage près, les mêmes caractéristiques initiales.

La randomisation est assurée par les covariables plutôt que par un processus aléatoire d'affectation

Prerequisite : Imputations Multiples des Données Manquantes

- **Méthode utilisée** : Monte Carlo par Chaîne de Markov («MCMC») (méthode recommandée lorsque les données manquantes ne sont pas monotones).
- **Stratégie de sélection des variables** pour le modèle d'imputation : Pour chacun des items documentés on dispose d'une distribution par niveau de réponse. A partir de cette distribution, des tirages seront effectués à 5 reprises pour calculer la moyenne des réponses dont on se servira pour imputer les valeurs manquantes de non répondants.
- **Règle** : Pour remplacer les données manquantes à l'aide de cette méthode, 50% des items de la dimension concernée doivent avoir été documentés.
- **Analyse** : Faite sur la base des données finales après intégration de la moyenne de 5 imputations précédentes.

Spécification du Modèle

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 (SP);
 - $SP = P(\text{Traitement} \mid \text{Caractéristiques initiales})$
 - 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..

Choix des Variables Explicatives du SP

■ Plusieurs stratégies

- Toutes les variables liées au choix du ttx : déconseillé, utiliser dans ce cas les variables instrumentales,
- Toutes les variables corrélées au choix du traitement et au résultat c.a.d les *facteurs de confusion vrais* déconseillé, augmente la variance,
- Toutes les *facteurs de confusion potentiels* liés aux résultats mais non au choix du traitement et vice versa: variables cliniques initiales et variables pronostiques du résultat,

■ Ne pas utiliser les critères classiques d'adéquation des fonctions logistiques:

- GOF (*Goodness-of-Fit*) ou Test de Homer-Lemeshow
- Statistique C qui mesure l'aire sous la courbe

→ Ces tests permettent de prédire les choix thérapeutiques en fonction des caractéristiques individuelles mais ne permettent pas d'identifier des facteurs de confusion importants qui auraient pu être oubliés



*IDENTIFICATION DES DÉSÉQUILIBRES
ET VÉRIFICATION DE LEUR
CORRECTION*

Mesure des Déséquilibres entre les Covariables Initiales de Chacun des 2 Groupes

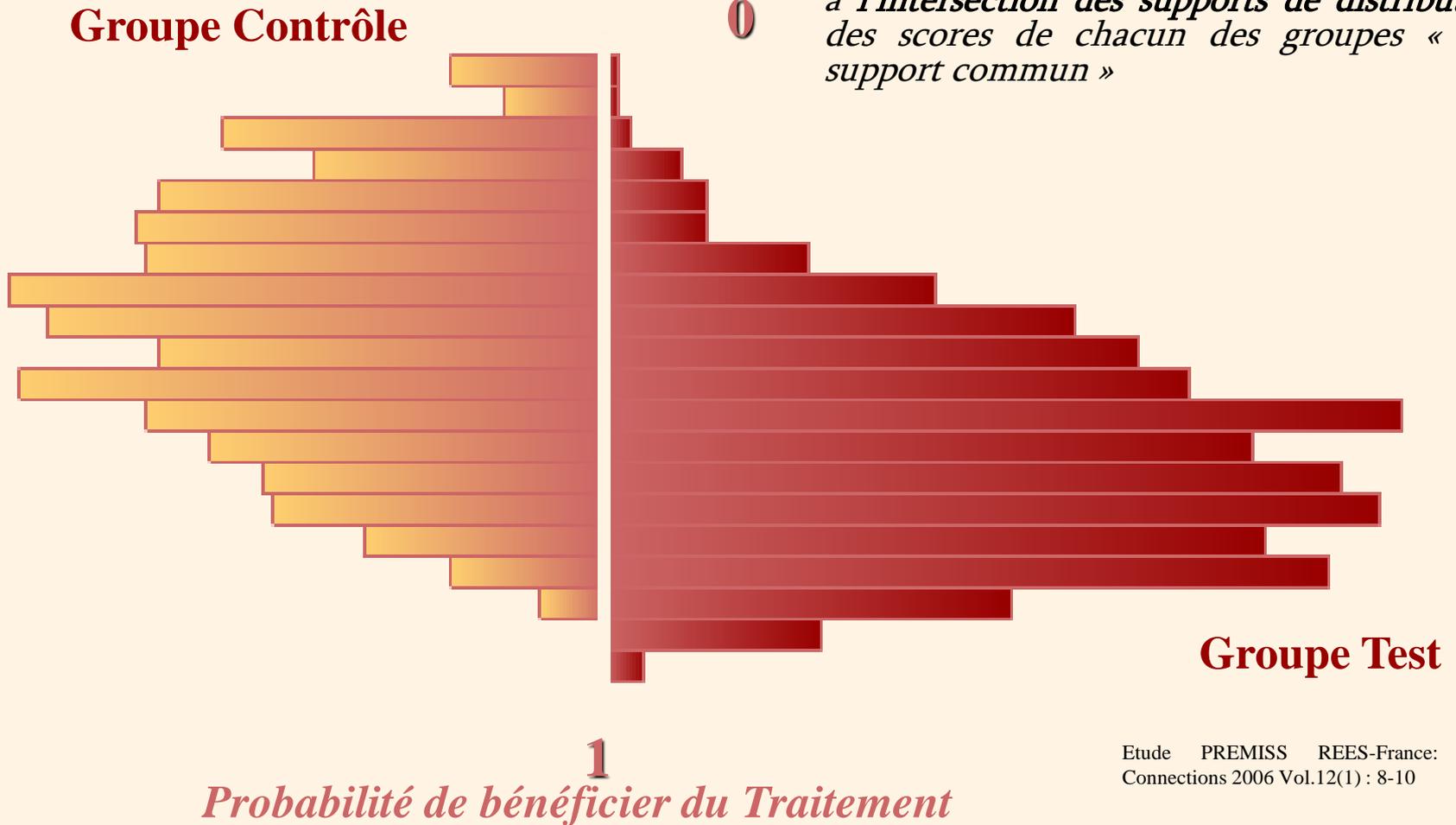
- Calcul des p-value *Chi* tests (pour des variables dichotomiques) où *t* tests pour les variables continues **à proscrire** :
 - Majore l'importance du déséquilibre en fonction de la taille de l'échantillon. La diff. peut être statistiquement significative avec grd échantillon sans l'être cliniquement
- → recours aux **différences standardisées** : différence de moyennes ou de proportions, mesurée en unités **d'écart-type moyen** des deux groupes [racine carrée de la variance moyenne]

$$d = \frac{(\overline{x}_{\text{traitement}} - \overline{x}_{\text{contrôle}})}{\sqrt{\frac{S_{\text{traitement}}^2 + S_{\text{contrôle}}^2}{2}}}$$

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

Distribution du Score de Propension sur le Groupe Test et le Groupe Contrôle

On ne peut construire de contre factuel que pour les sujets dont les scores appartiennent à l'intersection des supports de distribution des scores de chacun des groupes « le support commun »



Etude PREMISS REES-France: ISPOR
Connections 2006 Vol.12(1): 8-10

Algorithme d'Appariement [Méthode du plus proche voisin]



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



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

Groupe test : T

	T	C	
H	4	2	6
F	1	3	4

Groupe contrôle : C



SP = 0.25

SP = 0.66

SP = 0.66



Choix de la Métrique d'Appariement

- Appariement avec le plus proche voisin (*Nearest Neighbor Matching*), appariement de chaque sujet du groupe contrôle à un sujet traité en sélectionnant celui qui a le score de propension le plus proche du sien
- Appariement sous réserve du respect d'une valeur seuil donnée (*Caliper Matching*), impose une distance maximale entre les scores de propension de sujets traités et non traités au-delà de laquelle l'appariement est interdit
- Appariement dans le cadre d'un périmètre défini par une valeur donnée de son rayon (*Radius Matching*), se réfère non seulement au voisin le plus proche par rapport à une distance fixée a priori, mais à tous les individus non traités qui se trouvent à l'intérieur d'un périmètre dont le rayon est donné
- Appariement par noyaux (*Kernel Matching*), des coefficients de pondération qui diminuent le poids des observations distantes

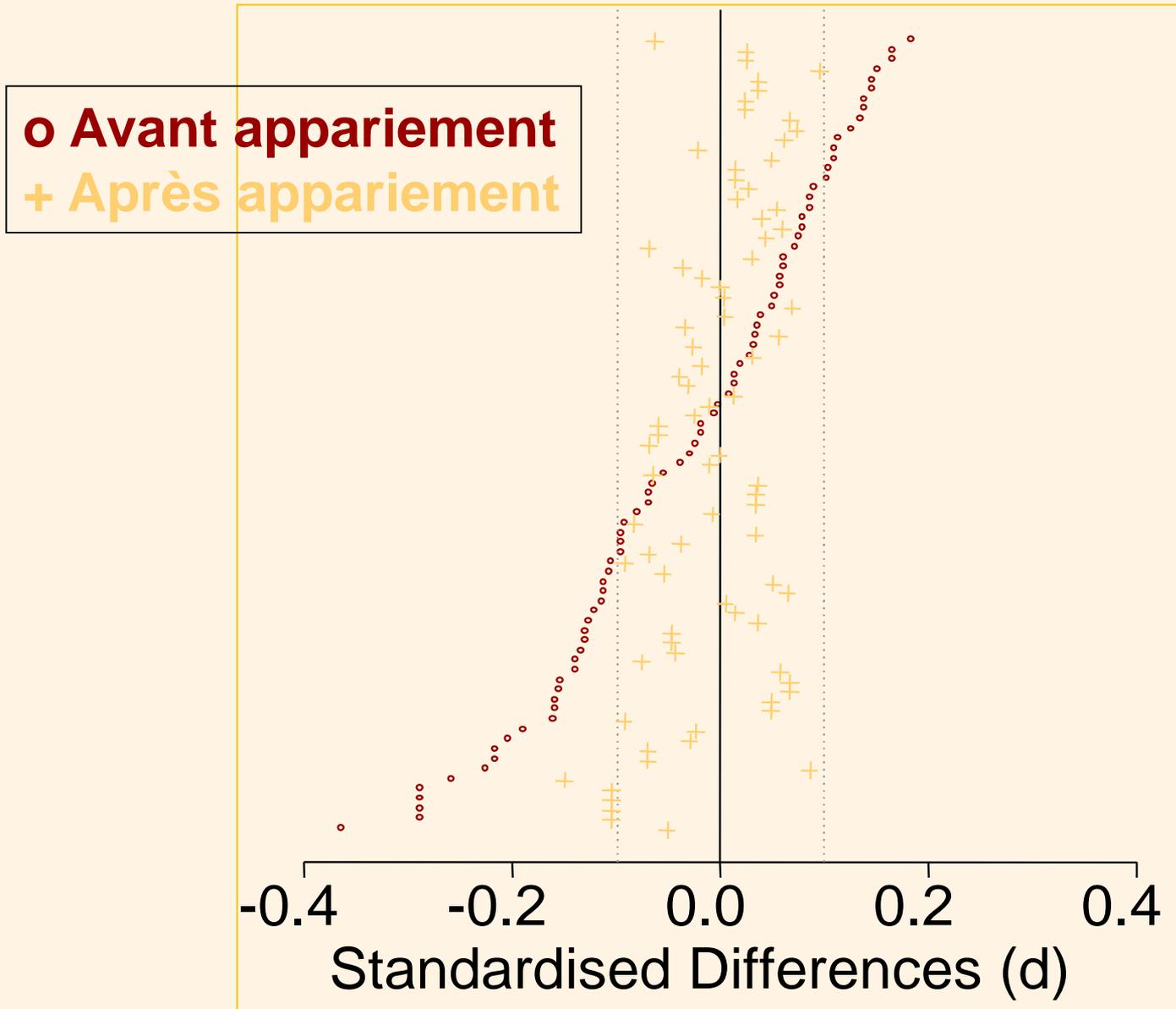
Vérifier Après Appariement que les Différences Standardisées sont Moindres qu' Auparavant

■ → Différences standardisées :

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

■ $|d| < 10\%$ → équilibre entre les deux bras

Différences Standardisées Avant et Après Appariement





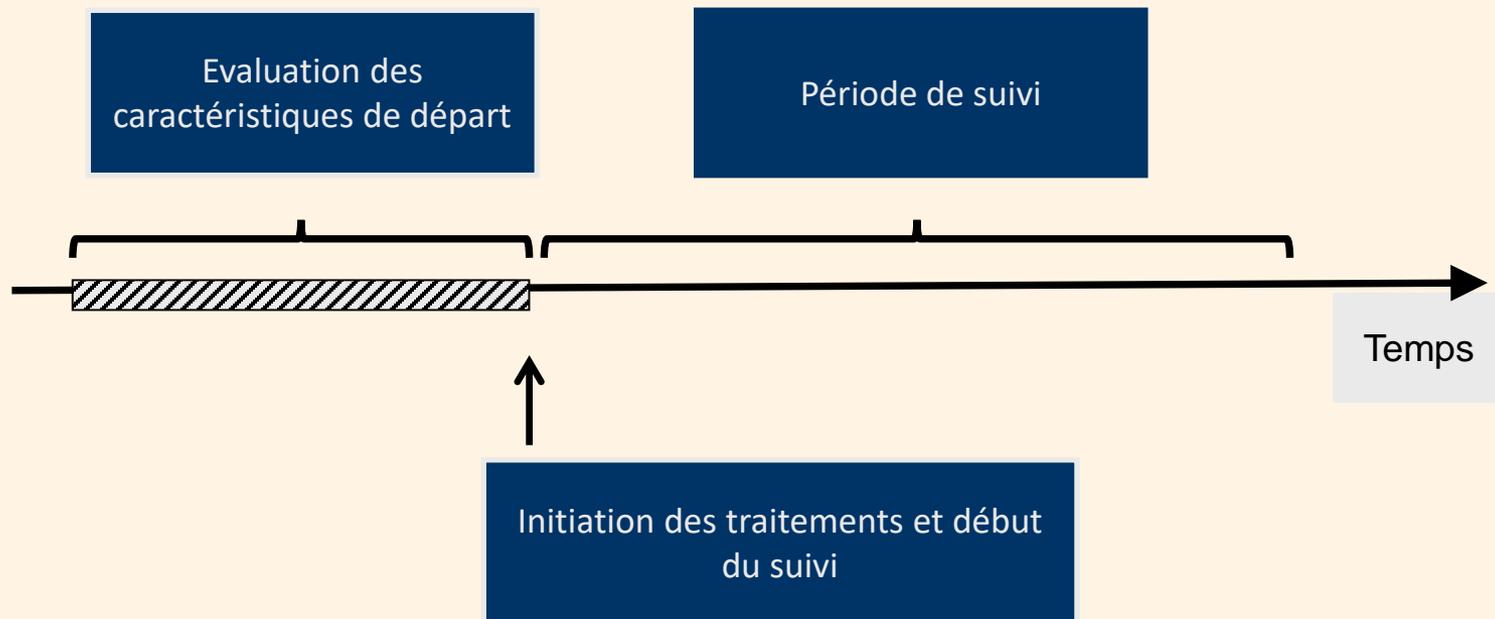
EXPLOITATION DU SCORE DE PROPENSION

Modalités de Mise en Oeuvre Possibles

- Protocole d'étude avec **appariement** sur le SP
 - Contrôle du biais de recrutement
 - Moins de patients si SP trop dissemblables → 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 , cinq strates permettent d'éliminer 90% des biais.

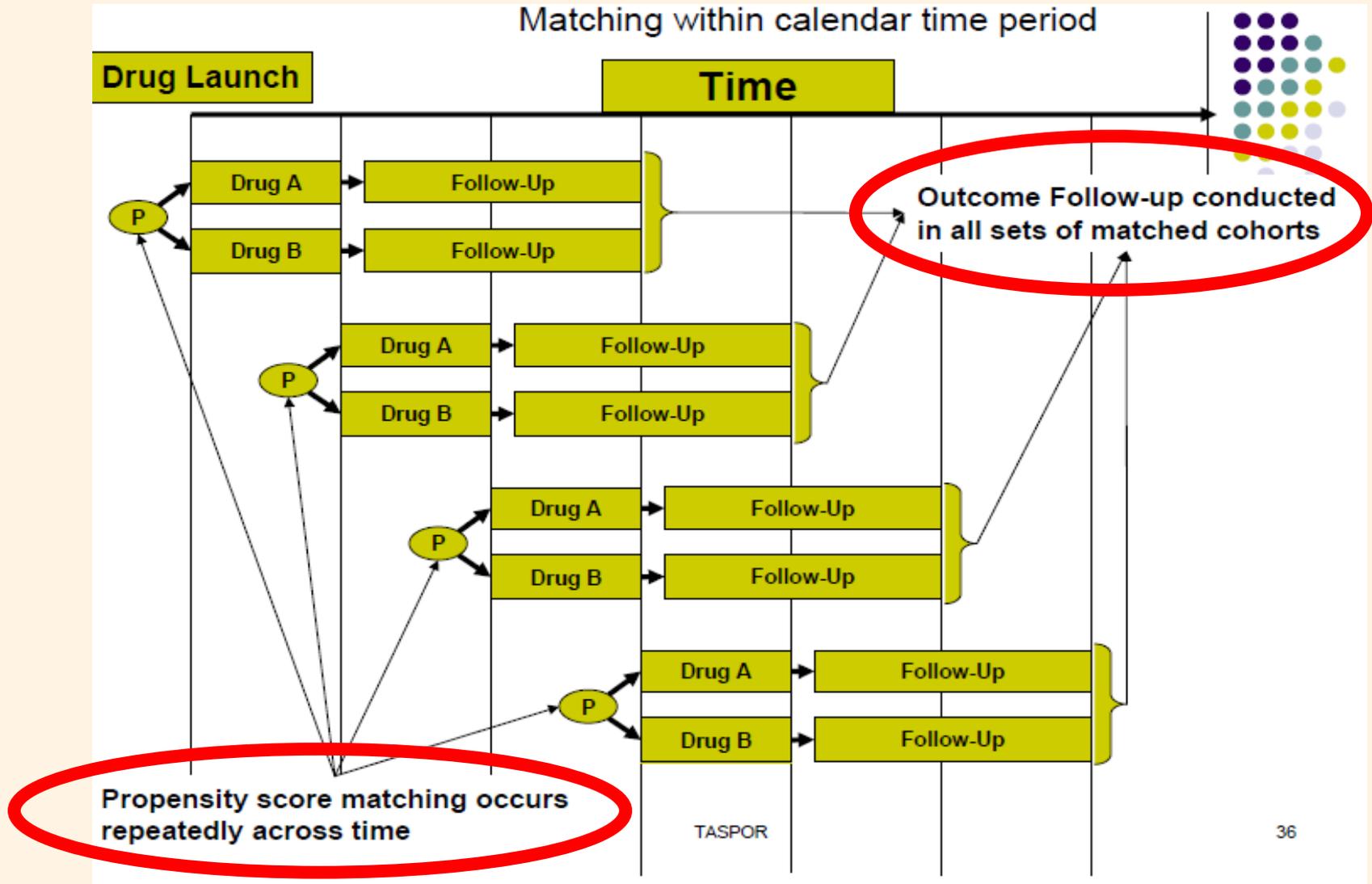
Dans les trois cas, le SP **simule un essai randomisé** en assurant la comparabilité des patients sur les variables **observées**.

Mettre en Place des Schémas Nouveaux Utilisateurs



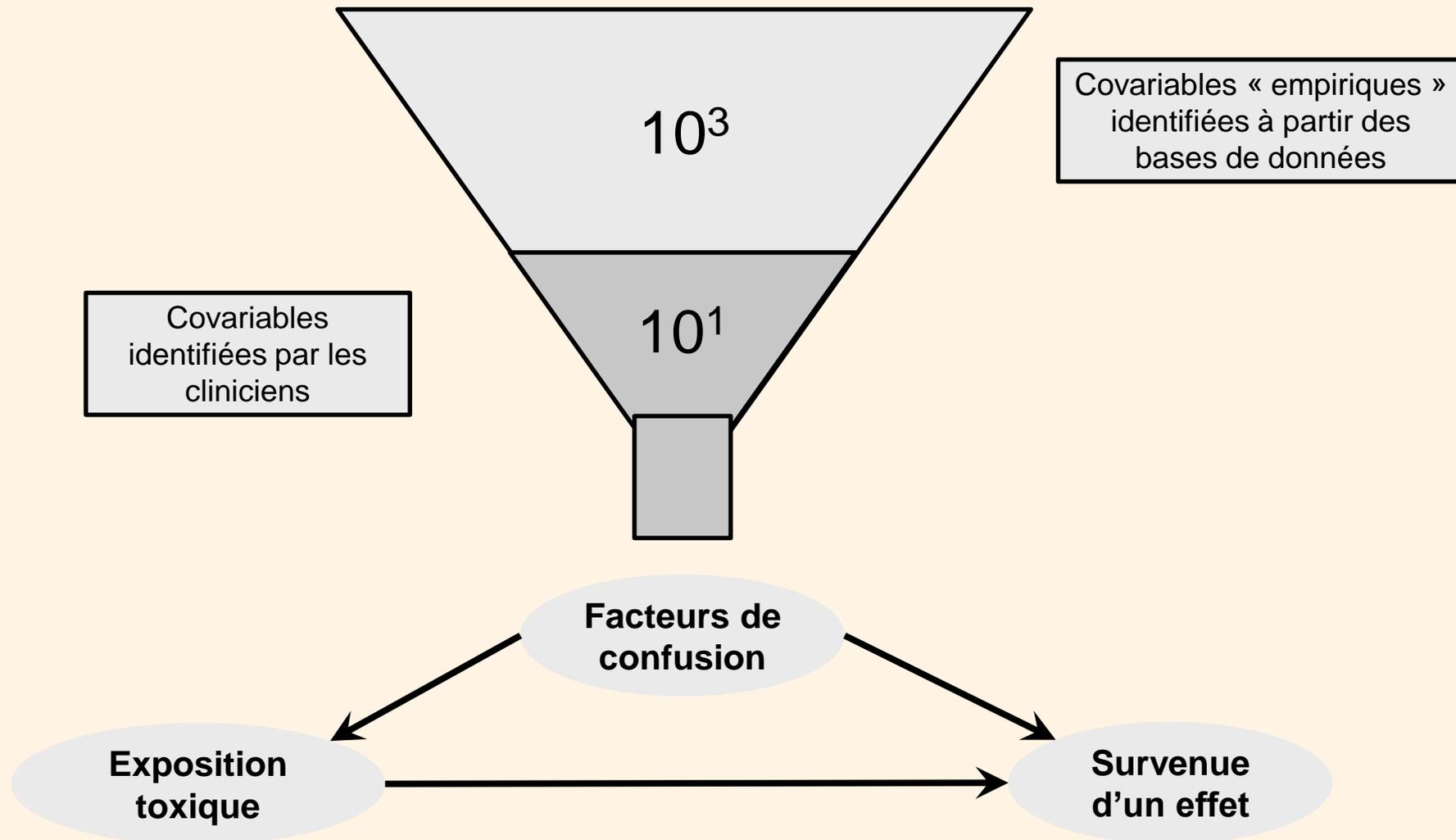
*Banque de Données Médico-Administrative

Procéder à des Appariements Séquentiels



Neutraliser les Biais dus Aux Facteurs Inobservables

Score de propension enrichi : hd-PS



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; Niteesh 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