

Consommation médicamenteuse et risque d'accident de la route :

*Exploration par simulation de schémas d'études
applicables à partir de données médico-
administratives*

Née Mélanie
Equipe PPCT
Inserm U897

Maître de stage :

Marta Avalos (MCF biostatistique)

Co-encadrant :

Ludivine Orriols (Post-doctorante)

Emmanuel Lagarde (Directeur de recherche)

Contexte (1)

- Utilisation croissante des bases de données médico-administratives
 - Large échantillon (évènement et/ou exposition rare, puissance...)
 - Réduction coûts...
- Exemple : étude de l'impact des médicaments sur le risque d'accident de la route
 - Canada, Norvège, France (étude CESIR)...

Contexte (2)

- Particularités de ces bases de données :
 - Facteurs de risque ou de confusion absents
 - Pas de témoins (conducteurs non accidentés)...→ Choix des schémas d'étude optimaux complexe
- Approche : choix du/des schéma(s) d'étude « optimal/aux », c'est-à-dire, permettant de :
 - Eviter l'introduction de biais
 - Conserver une bonne puissance

via des simulations

Objectif

Etudier, via des simulations, différents schémas d'études

- Cas-témoin
- Cas-témoin appariés
- Cas-croisé (ou *case-crossover*)
- Série de cas

permettant d'analyser le lien entre exposition aux

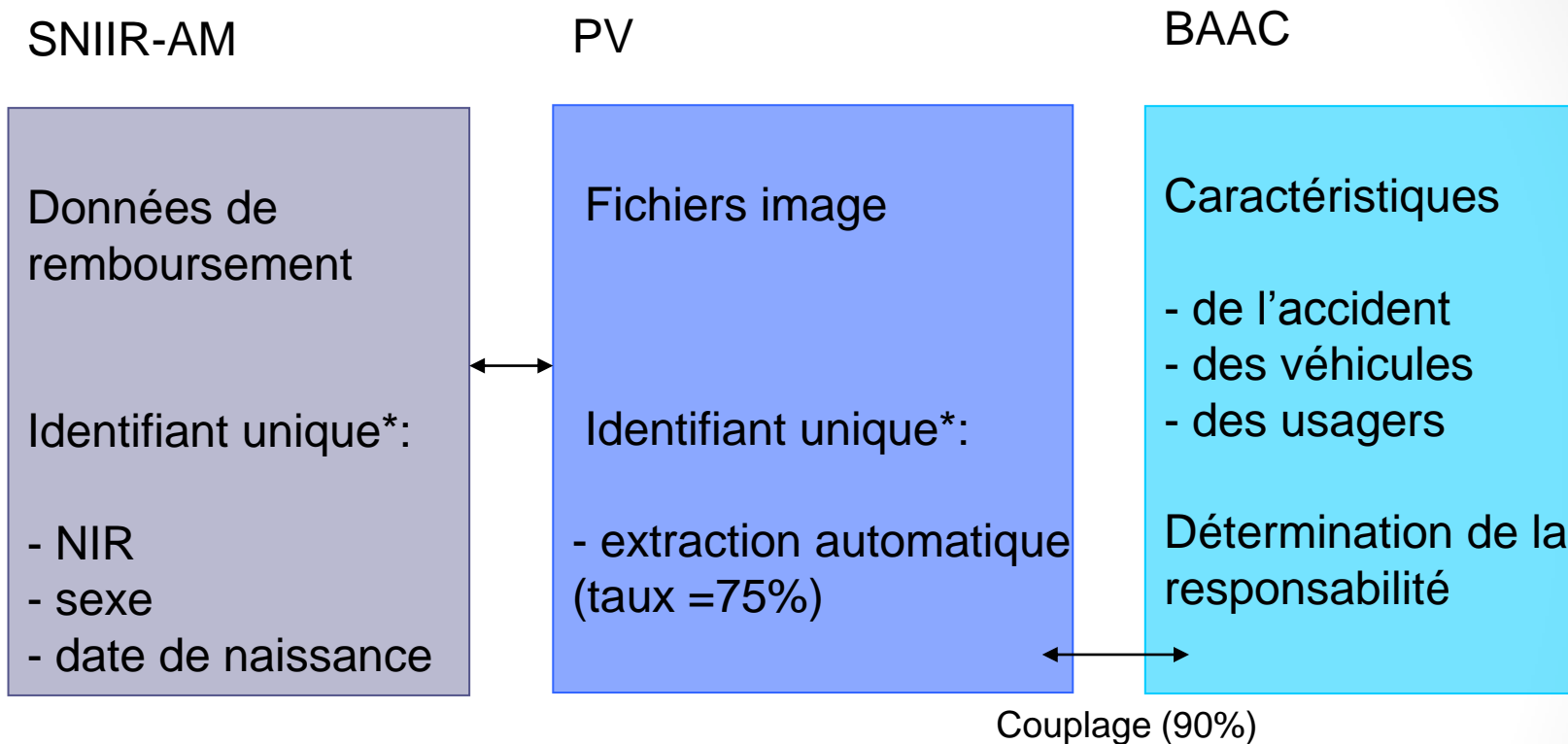
- Benzodiazépines (effets sur la conduite sont bien établis)

et le risque d'être :

- Responsable d'un accident
- Impliqué dans un accident

en utilisant comme cadre les données CESIR

Méthode : CESIR, population de référence (1)



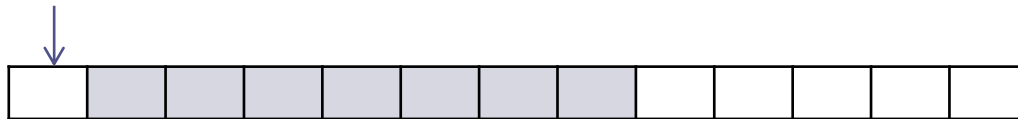
- 72 685 conducteurs impliqués dans un accident corporels entre juillet 2005 et mai 2008
- données sur la consommation de soins 6 mois avant et 2 mois après accident

* Identifiant anonymisé par la fonction FOIN pour les échanges CNAMTS / INSERM / IFSTTAR

Méthode : CESIR, population de référence (1)

- Variables disponibles dans CESIR
 - Variables d'ajustement caractérisant :
 - Les conducteurs : âge, sexe, alcoolémie, ALD...
 - L'accident : date de l'accident, heure de l'accident, localisation...
 - Exposition médicamenteuse

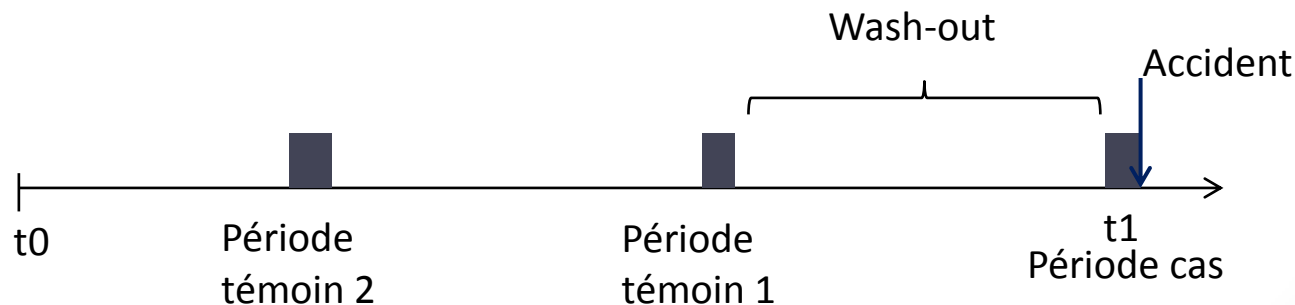
Date de
délivrance



Exposition :
enquête EPPM, par classe
ATC (niveau 5)

Méthode : schéma cas-croisé (1)

- Etudier l'effet d'une exposition ponctuelle sur la survenue d'un évènement (Maclure, 1991)
- Chaque cas est son propre témoin → contrôle des facteurs fixes
- Exposition période cas comparée à celle durant période(s) témoin(s)
- Choix des périodes témoins : unidirectionnelle, bidirectionnelle...
- $RC = \frac{\text{Nombre de cas exposés uniquement durant fen\^etre cas}}{\text{Nombre de cas exposés uniquement durant fen\^etre témoin}}$



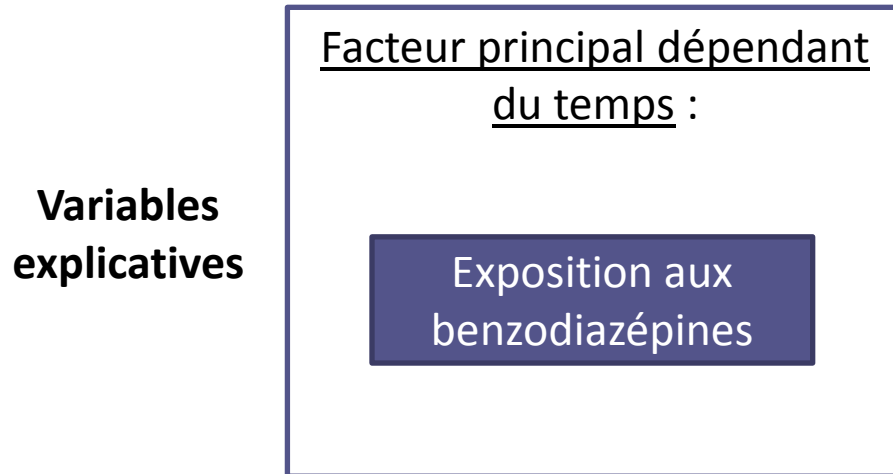
Méthode : schéma cas-croisé (2)

- Applicabilité aux données de CESIR
 - Cas-croisé unidirectionnel
 - Evènement modifie exposition future (Clarivet, 2013, rapport de stage M2 Epidémiologie)
 - Expositions ponctuelles à effets immédiats
 - Pas de tendance dans l'exposition
 - Données 6 mois avant l'accident
 - Période cas et témoins courtes
 - Wash-out : 30 jours

Méthode : questions des simulations

- a) L'augmentation du nombre de témoins par cas permet-elle d'augmenter significativement la puissance des résultats ?
- b) Quel impact a le fait de ne pas savoir si la personne conduisait le jour témoin ?

Méthode : génération des données (1)



Variable à expliquer

Méthode : génération des données (1)

**Variables
explicatives**

Facteur principal dépendant
du temps :

Exposition aux
benzodiazépines

Facteurs secondaires :

- Indépendants du temps :

Age

Alcool

Sexe

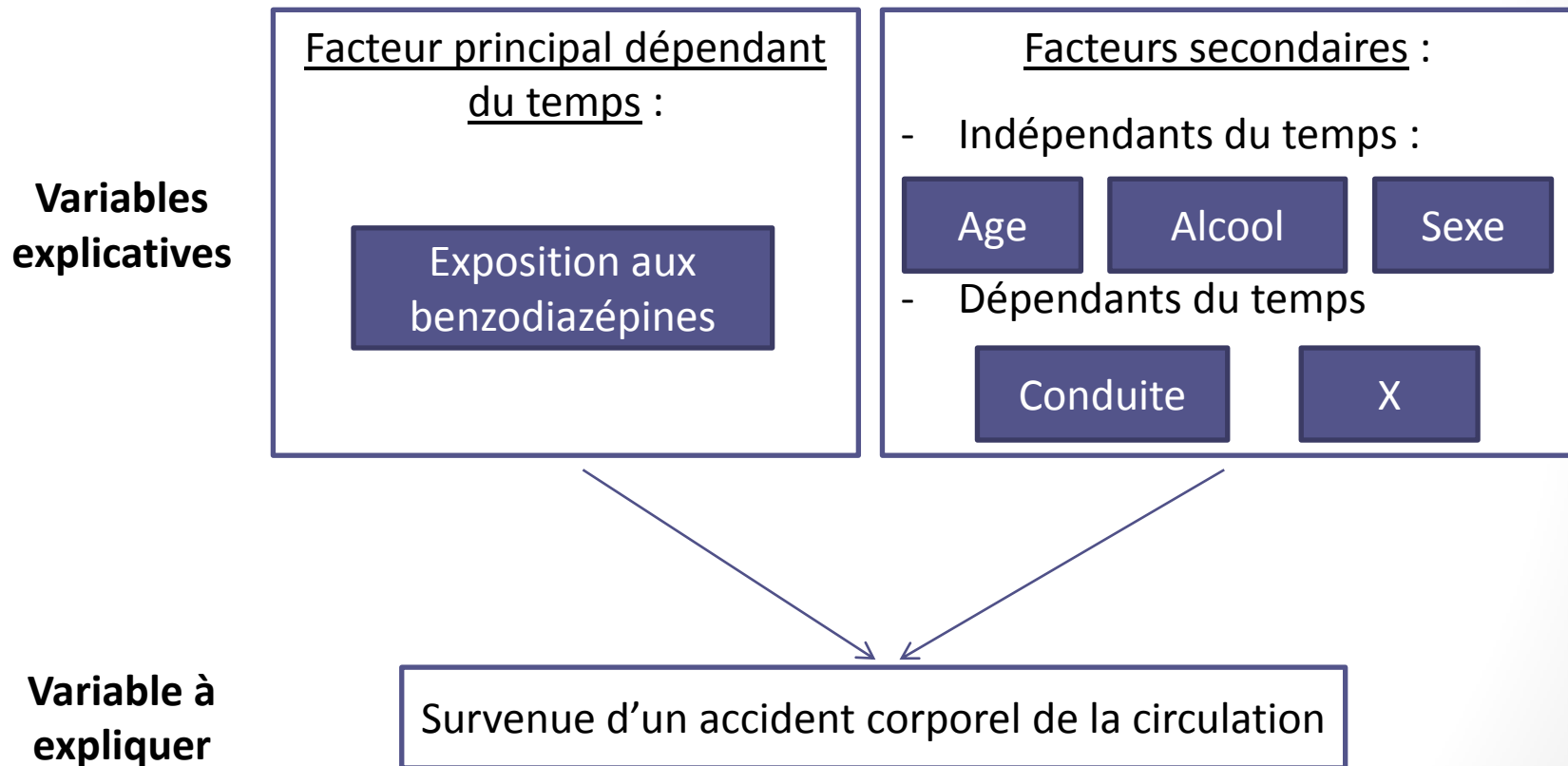
- Dépendants du temps

Conduite

X

**Variable à
expliquer**

Méthode : génération des données (1)



Méthode : génération des données (2)

- Variables d'ajustement fixes :
 - Sexe : variable binaire
 - Âge : variable catégorielle à 4 modalités
 - Consommation d'alcool : variable catégorielle à 5 modalités

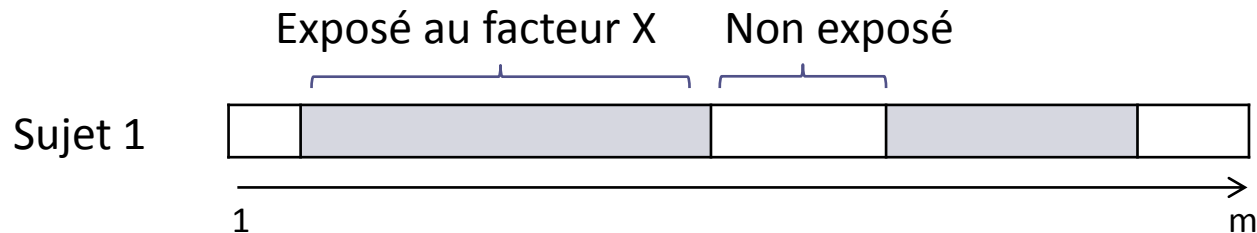
	Loi	CESIR (N=72 685)
Femmes (%)	Binomiale (N,p)	31,5
Age (%)	Multinomiale (N, pi)	
<25 ans		24,6
[24;45[44,2
]45;65[24,5
≥ 65 ans		6,7
Alcool (le jour de l'accident)	Multinomiale (N, pi)	
non renseigné		13,6
< 0,5		80,8
[0,5 ; 1,2[1,9
[1,2 ; 2,0[1,9
>2		1,8

Méthode : génération des données (3)

- Variable d'ajustement dépendante du temps : facteur X
 - Etape 1 : probabilité d'être exposé : $P(X_{\text{temp}}=1)=15\%$
 - Etape 2 : chez les $P(X_{\text{temp}}=1)$, on génère alternativement :
 - Une date de début d'exposition selon une loi uniforme $U[m]$
 - Une durée selon l'expression $14 + 7 \times \mathcal{P}(2)$

Puis jusqu'à ce que l'exposition d'un sujet couvre la période de m jours, on génère alternativement :

- Périodes « non exposé » selon l'expression $14 + 7 \times \mathcal{P}(3)$
- Périodes « exposé » selon l'expression $14 + 7 \times \mathcal{P}(2)$



Méthode : génération des données (4)

- Facteur principal : exposition médicamenteuse
 - Etape 1 : fait d'être exposé au moins une fois (E_{temp})

$$\text{logit}(P(E_{\text{temp}i}=1)) = \alpha_0 + \alpha_1 \text{Sexe}_i + \alpha_2 \text{Age}_{]24:44]i} + \alpha_3 \text{Age}_{]44:64]i} + \alpha_4 \text{Age}_{\geq 64i}$$

- Etape 2 : historique de l'exposition
 - Profils d'exposition dans CESIR
 - Tirage aléatoire sans remise

Méthode : génération des données (5)

- Conduite
 - Scénario 1 : la probabilité de conduire est la même chez les exposés et les non exposés \rightarrow Conduite $\sim B(N, p=0,75)$
 - Scénario 2 : la probabilité varie selon l'exposition (exposé/non exposé):
 - si sujet non exposé : 75% \rightarrow Conduite $\sim B(N_1, p=0,75)$ où N_1 le nombre de sujets non exposés
 - si sujet exposé : 45% \rightarrow Conduite $\sim B(N_2, p=0,45)$ où N_2 le nombre de sujets non exposés

Méthode : génération des données (6)

- Variable d'intérêt : adaptation de l'algorithme de MacKenzie et Abrahamowicz (J Am Statist Assoc, 1996)

Etape 1 : Génération des N temps de survie $T_i \sim U(183, m-61)$

Etape 2 : Génération des N temps de censure $C_i \sim U(184, (m-61) \times 3/2)$

Etape 3 : Pour chaque sujet i ($i=1, \dots, N$)

- $t_i^* = \min(T_i, C_i)$ (temps d'observation)
- $\delta_i = I\{T_i \leq C_i\}$ (indicateur de non censure), $i=1, \dots, N$

(t_i^*, δ_i) sont triés
tel que $t_i^* \leq t_{i+1}^*$.

Méthode : génération des données (7)

Etape 4 : Partant du temps observé le plus petit t_1^* :

a) Pour chaque sujet, détermination au temps t^* :

- de l'exposition aux benzodiazépines
 - du statut vis-à-vis de la variable X
 - de la conduite
- } un vecteur de valeurs de covariables $X_s(t_i^*)$ pour chaque sujet s

b) Assignment aléatoire de chacun des N couples (t_i^*, δ_i^*) a un vecteur de covariables $X_s(t_i^*)$

- Si $\delta_i^* = 0$, tirage aléatoire simple dans l'échantillon à risque R_i avec une équiprobabilité $p_i = 1/\text{taille}(R_i)$
- Si $\delta_i^* = 1$, probabilité basée sur la vraisemblance partielle de Cox :

$$P_{s,t_i^*} = \frac{\exp [\beta' X_s(t_i^*)]}{\sum_{j \in R_i} \exp [\beta' X_j(t_i^*)]} \quad \text{où } R_i = \text{l'ensemble des individus à risque au temps } t_i^*$$

Méthode : génération des données (8)

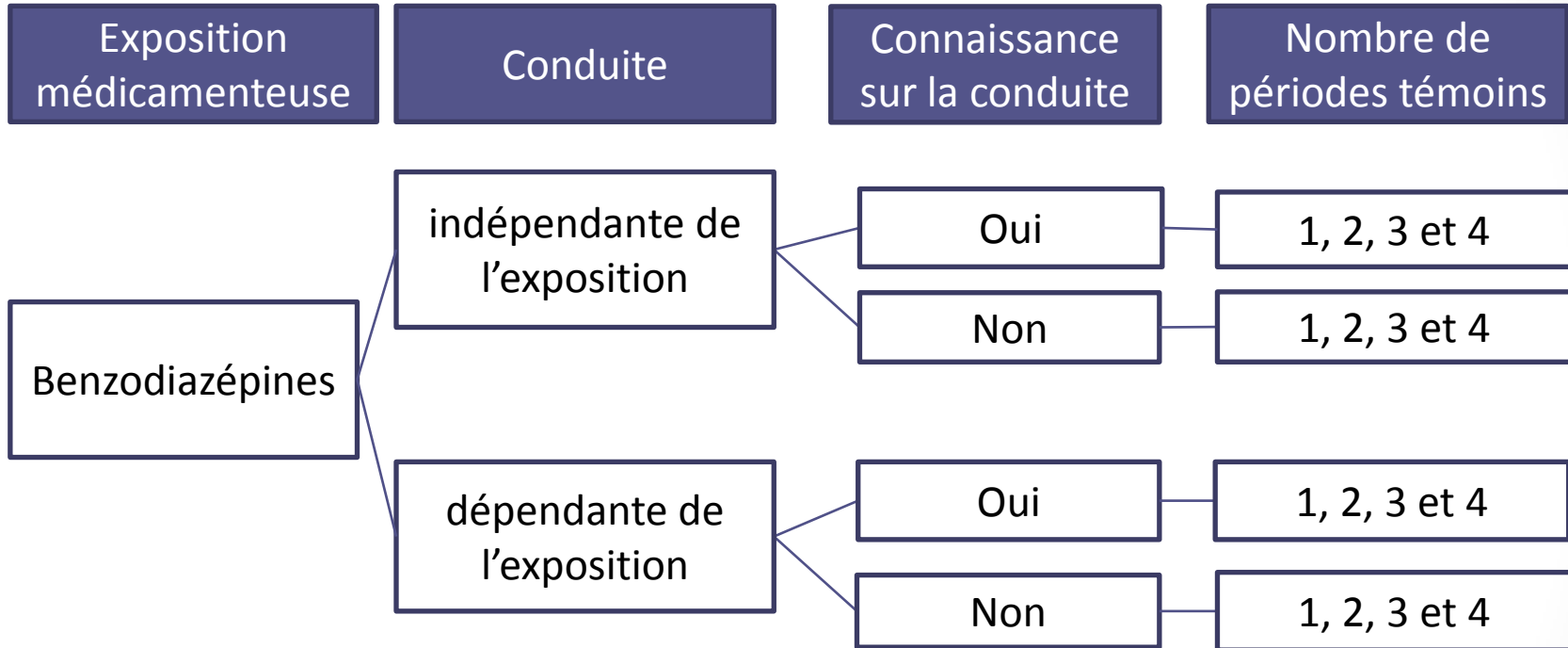
- Application de l'algorithme
 - N=40000 sujets et m=730 jours
 - probabilité d'avoir un accident au temps t :

$$h(t|X,E(t))=h_0(t) \exp(\theta_1 \text{ Sexe} + \theta_2 \text{ Age}_{]24-44]} + \theta_3 \text{ Age}_{]24-64]} + \theta_4 \text{ Age}_{\geq 64} + \theta_5 \text{ Al}_{<0,5} + \theta_6 \text{ Al}_{[0,5-1,2[} + \theta_7 \text{ Al}_{[1,2-2,0[} + \theta_8 \text{ Al}_{\geq 2} + \theta_9 X(t) + \theta_{10} E(t))$$

où Al désigne la consommation d'alcool et E l'exposition.

- On conserve :
 - Sujets non censurés
 - Les données sur la période 6 mois avant et 2 mois après accident

Analyse : scénarios testés



Analyse : critères d'évaluation

- Pour chacun des scénarios ont été relevés :
 - la moyenne des RC sur l'ensemble des 1000 jeux de données
 - écart-type des RC sur l'ensemble des 1000 jeux de données
 - biais relatif a également été calculé selon la formule :

$$\frac{\bar{\beta} - \beta_{th}}{\beta_{th}}$$

où β_{th} désigne le log(RC) qui a généré les données et $\bar{\beta}$ désigne la moyenne des 1000 RC estimés.

- proportion de p-valeurs < 5% au test de Wald d'hypothèse nulle de non association entre exposition et risque d'accident de la route

Résultats (1) : description

- Échantillon sujets non censuré :
 - 28165 en moyenne dont 408 consommateurs de benzodiazépines
- Distribution des variables d'ajustement (âge, sexe et consommation d'alcool) et exposition équivalente à CESIR

Résultats (2) :

- Question: l'augmentation du nombre de périodes témoins permet-elle d'augmenter significativement la puissance des résultats ?

Nombre périodes témoins	Benzodiazépines ($\beta_{th}=0,329$)				$\pi_{p<5\%}$
	Conduite connue	$\bar{\beta}$	$\sigma_{\bar{\beta}}$	Biais relatif (%)	
1	Oui	0,336	0,20	2,1	32,9
2	Oui	0,331	0,20	0,6	33,3
3	Oui	0,330	0,22	0,3	31,6
4	Oui	0,331	0,24	0,6	26,6

Résultats (2) :

- Question: l'augmentation du nombre de périodes témoins permet-elle d'augmenter significativement la puissance des résultats ?

Nombre périodes témoins	Benzodiazépines ($\beta_{th}=0,329$)				$\pi_{p<5\%}$
	Conduite connue	$\bar{\beta}$	$\sigma \bar{\beta}$	Biais relatif (%)	
1	Oui	0,336	0,20	2,1	32,9
2	Oui	0,331	0,20	0,6	33,3
3	Oui	0,330	0,22	0,3	31,6
4	Oui	0,331	0,24	0,6	26,6
1	Non	0,341	0,17	3,6	43,7
2	Non	0,339	0,15	3,0	56,8
3	Non	0,338	0,14	2,7	61,5
4	Non	0,339	0,13	3,0	64,0

Résultats (3) :

- Question : quel est l'impact du fait de ne pas savoir si la personne conduisait le jour témoin ?
 - Benzodiazépines
 - $\beta_{th}=0,329$

Connaissance sur la conduite	P(Cond=1)=0,75			
	$\bar{\beta}$	$\sigma \bar{\beta}$	Biais relatif (%)	$\pi_p < 5\%$
Oui	0,336	0,20	2,1	32,9
Non	0,341	0,17	3,6	43,7

Résultats (3) :

- Question : quel est l'impact du fait de ne pas savoir si la personne conduisait le jour témoin ?
 - Benzodiazépines
 - $\beta_{th}=0,329$

Connaissance sur la conduite	P(Cond=1)=0,75				P(Cond=1 E=1)=0,75 et P(Cond=1 E=0)=0,45			
	$\bar{\beta}$	$\sigma_{\bar{\beta}}$	Biais relatif (%)	$\pi_p < 5\%$	$\bar{\beta}$	$\sigma_{\bar{\beta}}$	Biais relatif (%)	$\pi_p < 5\%$
Oui	0,336	0,20	2,1	32,9	0,320	0,24	-2,7	30,1
Non	0,341	0,17	3,6	43,7	-0,315	0,18	-194,8	39,9

Discussion

- Augmentation du nombre de périodes témoins :
 - Résultats en accord avec ceux de Mittleman et al. (Am J Epidemiol, 1995): peu d'effets sur les estimations mais réduction des intervalles de confiance
- Impact du fait de ne pas savoir si les sujets conduisaient :
 - Impact important (sous-estimation de l'effet) si probabilité de conduire dépend fortement de l'exposition → en accord avec Young (Epidemiology, 2012)
 - Méthodes pour corriger les estimations (Redelmeier et Tibshirani, N Engl J Med, 1997) en l'absence de connaissance sur la conduite discutées

Bibliographique

- Quantin C, Goldberg M et al. Bases de données médico-administratives et épidémiologie : intérêt, limites. *Rev Épidémiologie Santé Publique*. 2008;56(5):S254.
- Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol*. 1991;133:144–53.
- MacKenzie T, Abrahamowicz M. Time-dependant hazard ratio: modeling and hypothesis testing with application in lupus nephritis. *J Am Statist Assoc*. 1996;91:1432-1439.
- Richardson DB. Occupational exposures and lung cancer: adjustment for unmeasured confounding by smoking. *Epidemiology*. mars 2010;21(2):181-186.
- Groenwold R. Bias sensitivity analysis of unmeasured confounding (invited session New methods to control for unmeasured confounding). In: 35th Annual Conference of the International Society for Clinical Biostatistics. Vienna (Austria); 2014.
- McCandless L. A Bayesian perspective on unmeasured confounding in large administrative databases. (invited session New methods to control for unmeasured confounding). In: 35th Annual Conference of the International Society for Clinical Biostatistics. Vienna (Austria); 2014.
- Burne R, Abrahamowicz M. New statistical methods for using validation subsamples to adjust for unmeasured confounders in survival analysis (invited session New methods to control for unmeasured confounding). In: 35th Annual Conference of the International Society for Clinical Biostatistics. Vienna (Austria); 2014.
- Mittleman M, Maclure M et al. Control Sampling Strategies for Case-Crossover Studies: An Assessment of Relative Efficiency. *Am. J. Epidemiol*. 1995;142,91-98.
- Young RA. Cell Phone Use and Crash Risk: Evidence for Positive Bias, *Epidemiology*. 2012;23(1):116-18.
- Redelmeier DA, Tibshirani RJ. Association between cellular-telephone calls and motor vehicle collisions. *N Engl J Med*. 13 févr 1997;336(7):453-458.