

# Appariement des bases de données médico-administratives

A propos d'un cas d'usage

« CANARI »

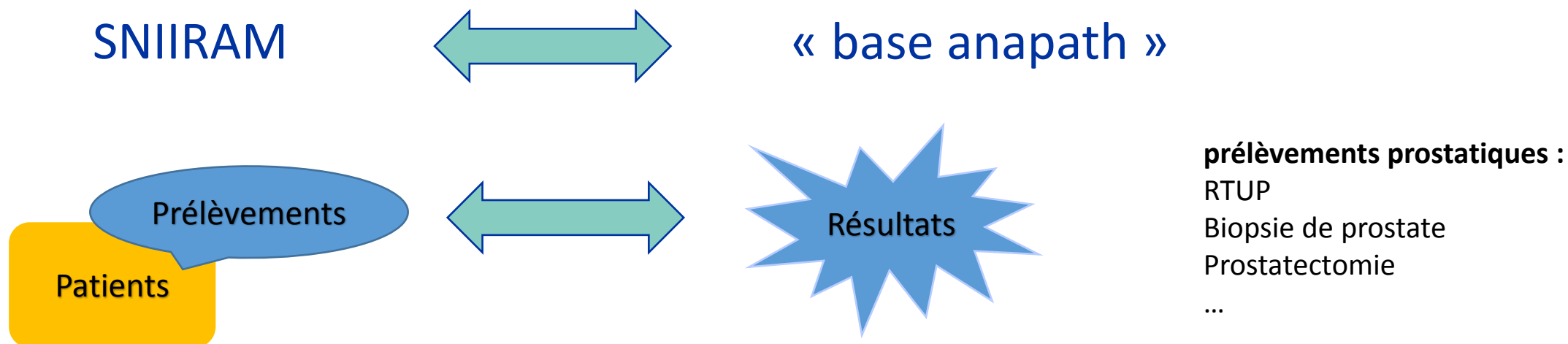
---

L-M. Scailteux<sup>1,2</sup>, Frédéric Balusson<sup>2</sup>, Sébastien Vincendeau<sup>3,4</sup>,  
Nathalie Rioux-Leclercq<sup>5,6</sup>, Emmanuel Nowak<sup>7</sup>

1. Centre Régional de Pharmacovigilance, de Pharmaco-épidémiologie et d'information sur le médicament, CHU Rennes
2. UPRES-EA 7449 « REPERES », Université de Rennes 1 et EHESP
3. Service d'Urologie, CHU Rennes
4. INSERM CIC 1414, CHU de Rennes
5. Service d'Anatomie et Cytologie Pathologiques, CHU Rennes, Université de Rennes 1
6. IMR 1085 – IRSET, Université Rennes 1
7. CHU de Brest et Université de Brest, INSERM CIC 1412, IFR 148, Brest



- D** Objectif : Etudier l'association entre cancer de la prostate de haut grade et la prise d'un traitement par 5ARI pour une hypertrophie bénigne de la prostate
  - ▷ stade du cancer de la prostate ▶ score Gleason ▶ absent du SNIIRAM
  - ▷ collecte locale ▶ laboratoires d'anatomo-pathologie de Bretagne





### **D SNIIRAM:** système national d'information inter-régime de l'assurance maladie

- base de données individuelle, anonyme
- 99 % de la population française
- données de remboursement des prescriptions médicales et des actes biologiques et médicaux

☹ Pas de résultat biologique

▷ Appariement une autre source de données pour enrichir la base

➤ Appariement « exact » dans l'idéal

☹ Pas d'identifiant unique permettant de relier les bases de données entre elles



# Problématique



Faisabilité ?

Quelle méthode ?



Est-ce que cela fonctionne ?



### **D Base de données médico-administratives ➤ CNIL**

Données anonymisées

Chainage de bases de données anonymes

▷ Appariement « semi-déterministe »

▷ Clés d'appariement

▷ Vérification de la cohérence



### **D Cohorte**

patients bretons traités pour une supposée hypertrophie bénigne de la prostate (HBP)

▷ inhibiteurs de la 5  $\alpha$ -réductase (5ARI),  $\alpha$ -bloquants, phytothérapie en 2010-2011

### **D Cas :** patients avec code diagnostic cancer de prostate CIM-10 C61 ou D40.0 (hospitalisation ou ALD) entre 2012-2013 et naïf de cancer de prostate < 2012

▷ Identification des actes de prélèvement prostatique et chainage des 2 bases

➤ date index = date du 1<sup>er</sup> score Gleason identifié

➤ cancer de bas grade : Gleason < 8

➤ cancer de haut grade : Gleason  $\geq$  8

### **D Témoins :** patients naïf de cancer de prostate 2010-2013 : pas de code C61 ou D40.0 ni d'ALD, ni de remboursement d'hormonothérapie, ni d'orchidectomie.



### D Appariement « semi-déterministe »

#### Clés d'appariement

- ▷ mois et année de naissance
- ▷ date du prélèvement prostatique dans le SNIIRAM et date de réception du prélèvement pour analyse dans la base anapath (autorisation de 3 voire 4 jours à partir de la date de prélèvement)

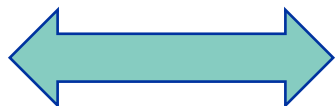
#### Cohérence

- ▷ localisation géographique : département lieu prélèvement / lieu analyse



# CANARI – bases à appairer

SNIIRAM

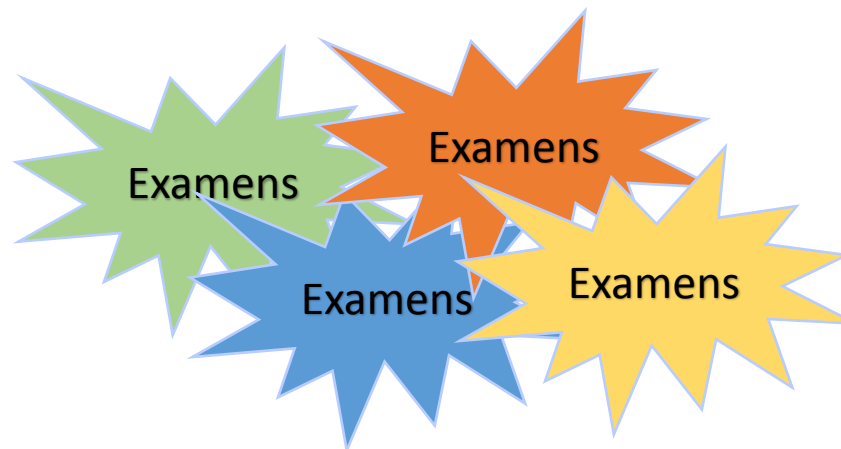
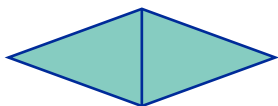


« base anapath »

9 laboratoires bretons,  
privés et publics

≥ 1  
Prélèvement  
prostatique /  
patient

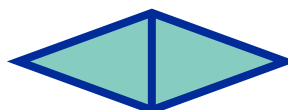
Patients  
Ttt indiqué HBP  
Diag CaP (CIM-10 C61,  
D40.0)



Examens de  
prélèvements  
prostatiques,  
**toutes indications  
confondues**

≥ 1  
Prélèvement  
prostatique /  
patient

Patients  
Ttt indiqué HBP  
Diag CaP (CIM-10 C61,  
D40.0)



**Score Gleason**





## D Données disponibles

Laboratoire d'anapath	Numéro examen	Mois-Année de naissance	Date de réception dans le laboratoire d'anapath	Date de prélèvement	Etablissement préleveur	Score Gleason
-----------------------	---------------	-------------------------	---	---------------------	-------------------------	---------------



**SNIIRAM**

**74 596 patients** avec au moins 1 délivrance de médicament pour une HBP  
Entre 01/01/2010 et 31/12/2011

**Base anapath**

**13 796 prélèvements prostatiques** (toutes indications et résultats confondus)  
Entre 01/01/2012 et 31/12/2013

**Patients avec cancer de prostate**  
Entre 2012 et 2013  
N = 1 055

**Actes de prélèvements de prostate dans SNIIRAM**  
Entre 2012 et 2013  
N = 1 623

**Patients sans cancer de prostate**  
Jusqu'au 31/12/2013  
N = 67 730

**Exclusion** : prélèvements non appariés (n=348); appariements inconsistants (n=30); pas de Gleason (n=46)

Après appariement, conservation du 1<sup>er</sup> prélèvement positif (n=340)

**859 patients**  
// 859 scores Gleason



- D** Cohorte SNIIRAM ▷ 859 cas de cancer de prostate
  - ▶ challenge d'apparier 2 bases de données anonymes
    - ≠ Autres études publiées disposant d'un n° unique d'identification entre les bases ou bases déjà reliées

### **D** Limites, difficultés

- Gestion des données longue : collecte et uniformisation des fichiers reçus des labo d'anapath
- Collecte des données locales et non pas nationales
- Certains prélèvements de la base d'anapath non appariés liés à l'absence de correspondance SNIIRAM



## Conclusion / appariement



Faisabilité ?

Est-ce que cela fonctionne ?

Fiabilité ?

Oui

