



Améliorer la santé de demain



# La cohorte CONSTANCES

## Appariements avec le SNIIR-AM et le SNGC-Cnav

Marie Zins, Marcel Goldberg  
Équipe Constances et les CES de l'Assurance Maladie

*Plateforme de recherches « Cohortes épidémiologiques en population »  
Inserm U1018-CESP Villejuif - Université Versailles St-Quentin*



# Une cohorte en population pour la recherche et la santé publique

Recherche en épidémiologie, notamment

- Vieillissement
- Risques professionnels
- Inégalités sociales de santé
- Santé des femmes

*Maladies chroniques +++*

*Facteurs génétiques et interactions gène-environnement*

Santé publique : décrire

- La santé de la population adulte
- Les parcours de prévention et de soins

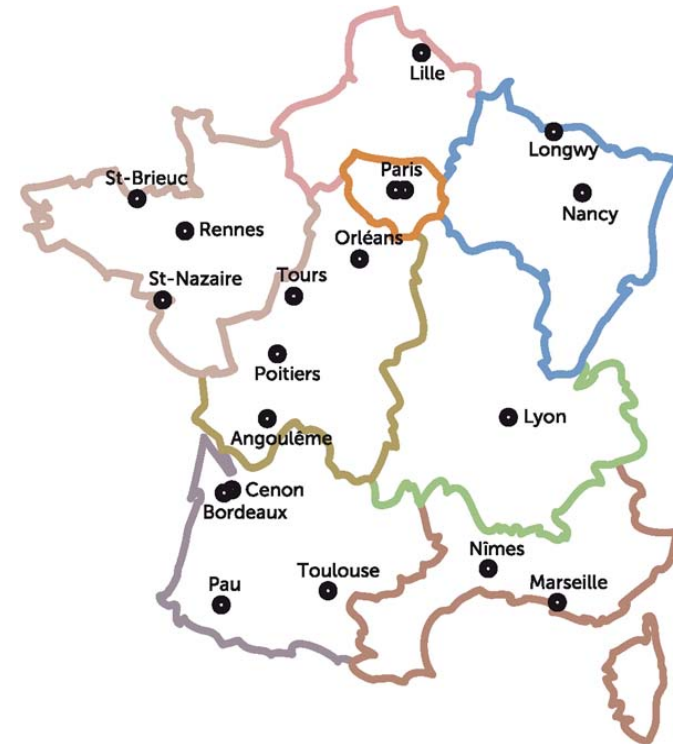
# Schéma de l'étude

## Échantillon

- Représentatif de la population affiliée au Régime général de la Sécurité sociale, âgés de 18 à 69 ans à l'inclusion
- Effectif : 200 000 personnes

## Inclusion et suivi

- Dans 17/110 Centres d'examens de santé (CES), dans 16 départements
- Inclusion graduelle sur 5 ans
- Questionnaires annuel, examen tous les 5 ans
- Appariement aux bases médico administratives nationales



# Apport des bases médico administratives

1- Tirage au sort

2- L'inclusion et le suivi

- Des données de sources diverses
- Des difficultés et des limites
  - a) Maitrise de la base
  - b) Repérage et validation des diagnostic

3- Etude des effets de sélection

# 1- Le tirage au sort

Objectif : échantillon représentatif et proportionnel à la population

- Tirage au sort à probabilités inégales
  - Caisse AM, Sexe, DDN, PCS,
  - Et le NIR pour suivi dans les Bases

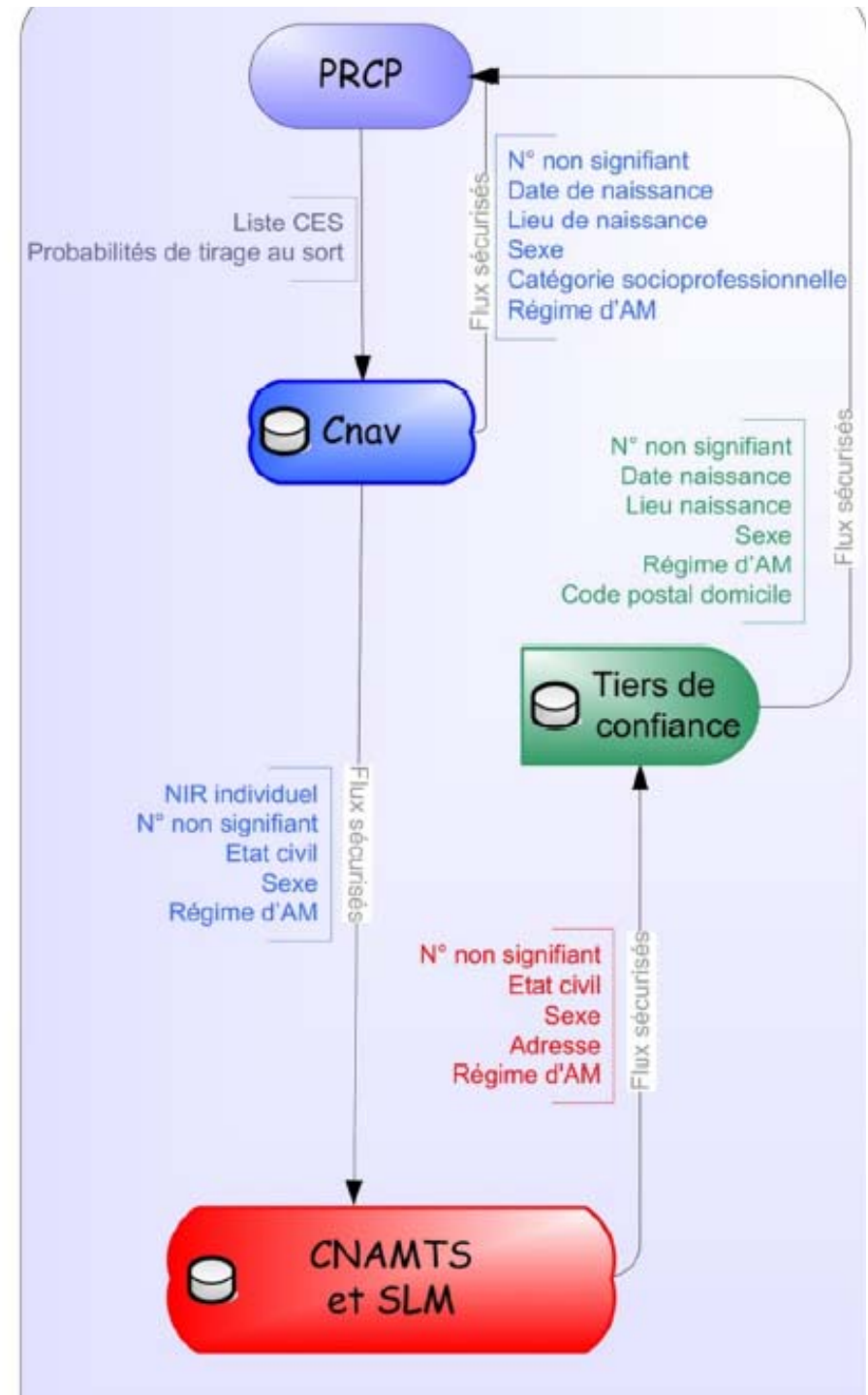
Intervenant : la CNAV

Les bases

- RNIAM (NIR, Caisse AM, sexe, DDN)
- SNGC (PCS)

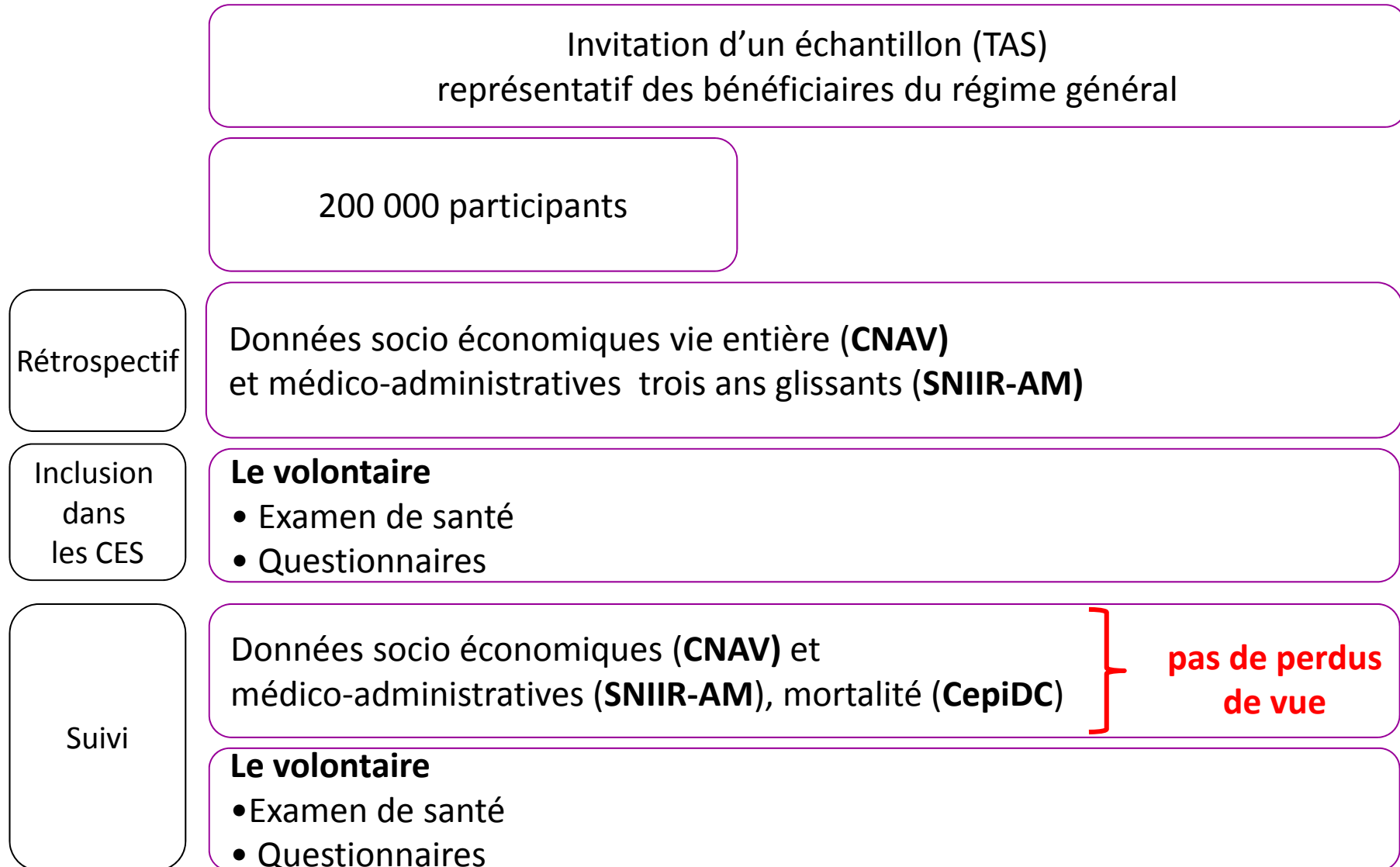
La CNAV procède au TAS

- Crée un Numéro non identifiant
- Adresse le fichier à la CNAMTS et SLM pour l'enrichir des adresses
- Avec le NIR pour le SNIIRAM



## 2- Inclusion et suivi :

Une grande variété de données issues de sources diverses



## 2- Inclusion et suivi

Une grande variété de données issues de sources diverses

### **Les données recueillies au niveau individuel**

- Caractéristiques sociodémographiques, statut et situation sociale
- Comportements
- Facteurs professionnels
- Environnement (géocodage des adresses )
- Données de santé
  - Poids, taille, TA, ECG, fréquence cardiaque, vision, audition, spirométrie, biologie
  - 45 ans et plus : tests de fonctionnement cognitif et physique
  - Échelles de santé, qualité de vie, pathologies (liste de pathologies déclarées, handicaps, limitations, incapacités) santé de femmes
- Biobanque (sang, urine)

### **Les données des bases (Cnav, SNIIRAM)**

## 2 a - Une difficulté majeure : la complexité des bases de données

Objectif : rendre « utilisables » les données

- Constitution d'une équipe « *Bases de données médico-administratives* » au sein de la plateforme
- Mise en place et animation de groupes de travail
  - « *Mains dans le cambouis* » comprendre les données brutes
  - « *Indicateurs épidémiologiques* » construction de variables synthétiques à partir des données brutes
- Catalogue de données en cours de finalisation : tables et variables associées, règles d'utilisation, glossaire, documents et textes de référence, tables de valeurs ; historisé



## 2 b - Une limite : exhaustivité et précision des diagnostics

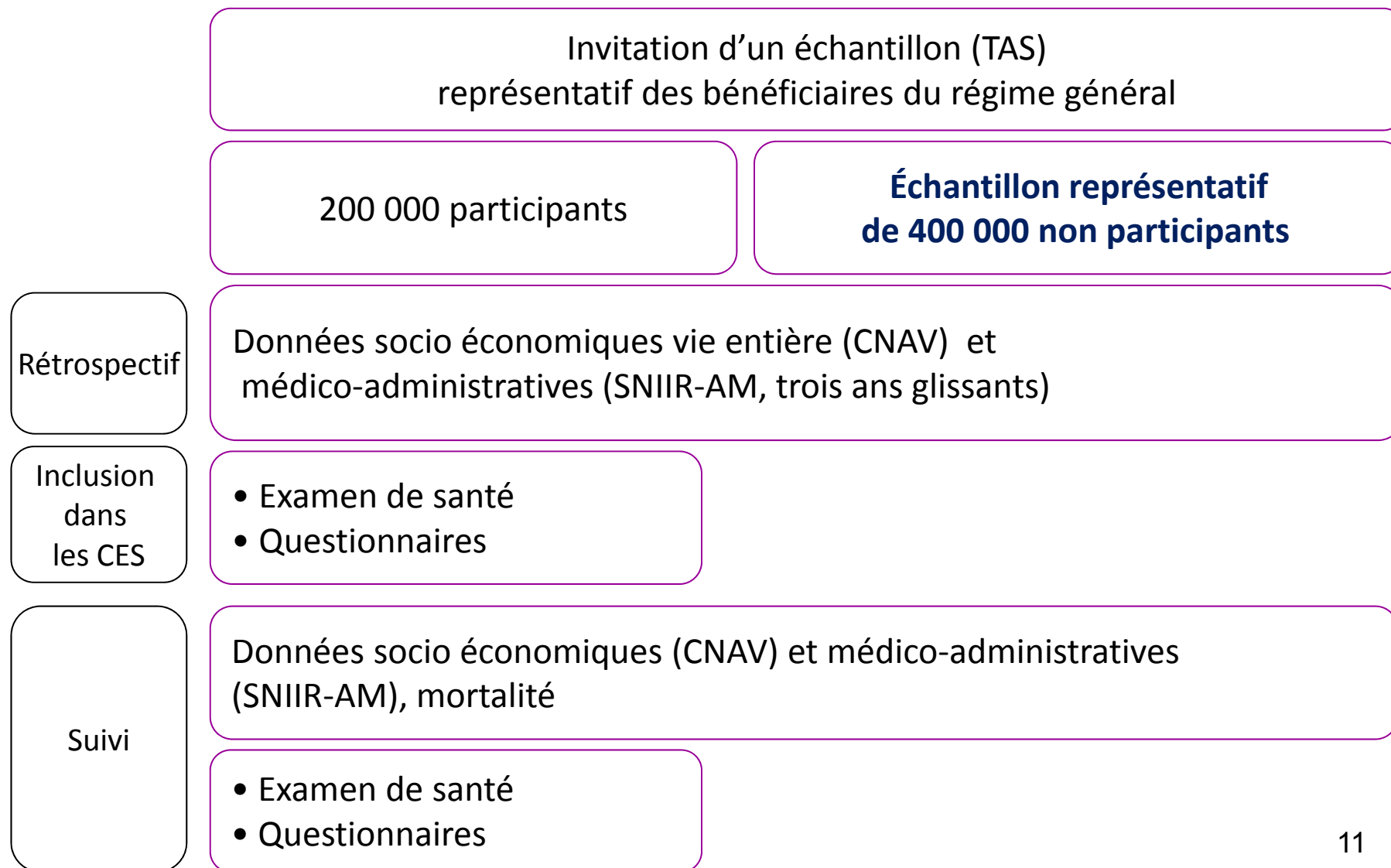
### Repérage et validation de pathologies incidentes

- Repérer la pathologie
  - « Multisources » : données des bases SNIR-AM et PMSI, AQ
- Valider
  - Retour aux dossiers médicaux (hôpitaux, médecin traitant, volontaires) et validation par des Comités de validation externes
    - Focus actuel : cardiopathies ischémiques, cancers, maladies neurodégénératives
- Algorithmes de repérage et validation
  - Participation à un réseau en cours de constitution (CNAMTS, InVS, DIM)

# 3- Prise en compte des biais de sélection

- Le volontariat induit des effets de sélection
  - À l'inclusion
  - Pendant le suivi (attrition)
- Il existe des méthodes statistiques pour prendre en compte les effets de sélection
  - Imputation
  - Pondération
- Elles reposent sur des données auxiliaires
  - Associées à la participation et aux phénomènes d'intérêt de l'étude
  - **Disponibles pour les répondants et non répondants**

# Etudier les effets de sélection dans Constances



# Premiers résultats du pilote (SNGC, CNAV)

Risque de participer ajusté sur âge

	Odds Ratio	[95% Interval	Conf.]
<b>Sexe</b>			
Hommes	1		
Femmes	1,16	1,03	1,30
<b>Nb employeurs/ an</b>			
0	1		
1	2,07	1,34	3,21
2 ou plus	2,15	1,37	3,37
<b>Catégorie sociale</b>			
Cadre	1		
Pi	0,90	0,74	1,08
Employé	0,58	0,48	0,69
Ouvrier	0,39	0,33	0,47
Manquant	0,47	0,38	0,59
<b>Brut mensuel plafonné</b>			
0-499	1		
500-999	1,21	1,02	1,43
1000-1499	1,46	1,21	1,77
1500-1999	1,85	1,48	2,32
3000 et plus	1,95	1,28	2,98
Missing	0,61	0,21	1,81

# Premiers résultats du pilote (SNIIR-AM)

## Caractéristiques des patients

Remboursement 2007/2009	Participant	Non participant	Total
Au moins un soin de ville remboursé	99,1%	96,2%	96,5%
Cout moyen remboursé honoraires médicaux en €	269.6 (388.4)	242.2 (398.8)	244.4 (398)
Cout remboursé moyen en soin de ville en €	694,3 (1329.4)	742,3(1953.2)	738.5 (1911.3)
Nb consultations dentaires en 2008			
% bénéficiaires	50,6%	38,2%	39,1%
Nb Moyen(écart-type)	2,7 (2,2)	3,0 (2,6)	2,9 (2,6)

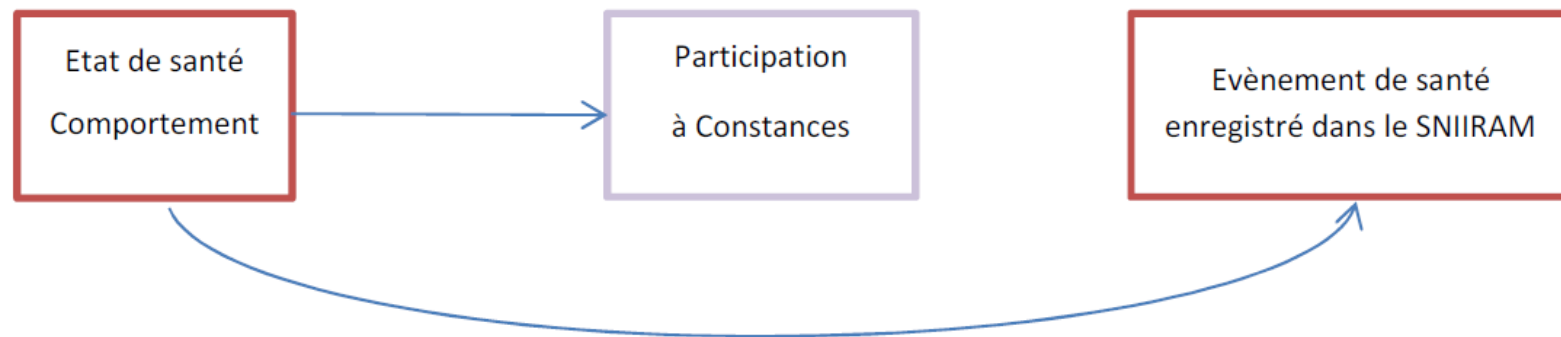
# Premiers résultats du pilote (SNIIR-AM)

## Données médicalisées (ALD)

Remboursement 2007/2009	Participant	Non participant	Total
Au moins un remboursement au titre d'une ALD	10,1%	12,8 %	12,6%
ALD affection psychiatrique	1.4%	3.3%	3.2%
ALD diabète	1.3%	2.3%	2.2%
ALD cancer	3.0%	2.5%	2.5%
ALD cardiovasculaire	2.1%	2,7%	2,7%
Hospitalisations	37,2%	32,7%	33.1%

# Prise en compte des effets de sélection à l'inclusion

- Modélisation de la probabilité de participation et pondération
- Possibilité de mise à jour des probabilités de participation au cours du suivi



Label de qualité statistique du CNIS

Guéguen A. et al. Ottawa, Canada. 27-30 Octobre 2009

Sitta R. et al. ISCB Bergen, Norvège. 19-23 aout 2012

# Synthèse de l'apport des BDMA dans la mise en place de la cohorte Constances

- Constitution de la cohorte
- Quantification des effets de sélection et pondérations
- Suivi de la cohorte sans perdus de vue
- Enrichissement des données dans des domaines cruciaux : santé, situation et trajectoire socioprofessionnelle
- Repérage et validation des événements de santé

**En conclusion : Constances n'aurait pas été possible  
à cette échelle sans l'appariement avec les BDMA**

**Merci aux équipes de la Cnav et de la Cnamts !**