



# Prise en compte de la non-participation dans Constances Premières explorations



# Contexte

- > Constances : cohorte épidémiologique généraliste, construite pour :
  - > Constituer une infrastructure de recherche
  - > Permettre des études de santé publique et de surveillance épidémiologique
- > Néanmoins, pas dans les standards des enquêtes descriptives à visée représentative :
  - > Contraintes logistiques
    - Plan de sondage différent selon les années et les CES
    - Poids de sondage dispersés pouvant entraîner des estimations de variance élevées
  - > Déplacement des personnes aux CES (recueil de données de santé)
    - Taux de participation attendu ( ~ 7%)
    - Potentiels biais de non-réponse importants
  - > Dans le cas d'événements rares, combinaison de plusieurs années d'inclusion pour obtenir des effectifs importants
    - Conditions nécessaires de stabilité de la population et de la mesure de la variable d'intérêt





# Objectifs

- > Présenter les problèmes rencontrés pour estimer des prévalences dans Constances et les premiers traitements effectués

Population d'intérêt : personnes âgées de 30 à 69 ans en 2013, vivantes au 31 janvier 2014, résidant dans un département Constances et affiliées au Régime Général, à la Camieg, la MFPs ou la MGEN.

*Exclusion des moins de 30 ans (taux de participation très faible)*





# Matériel et méthodes





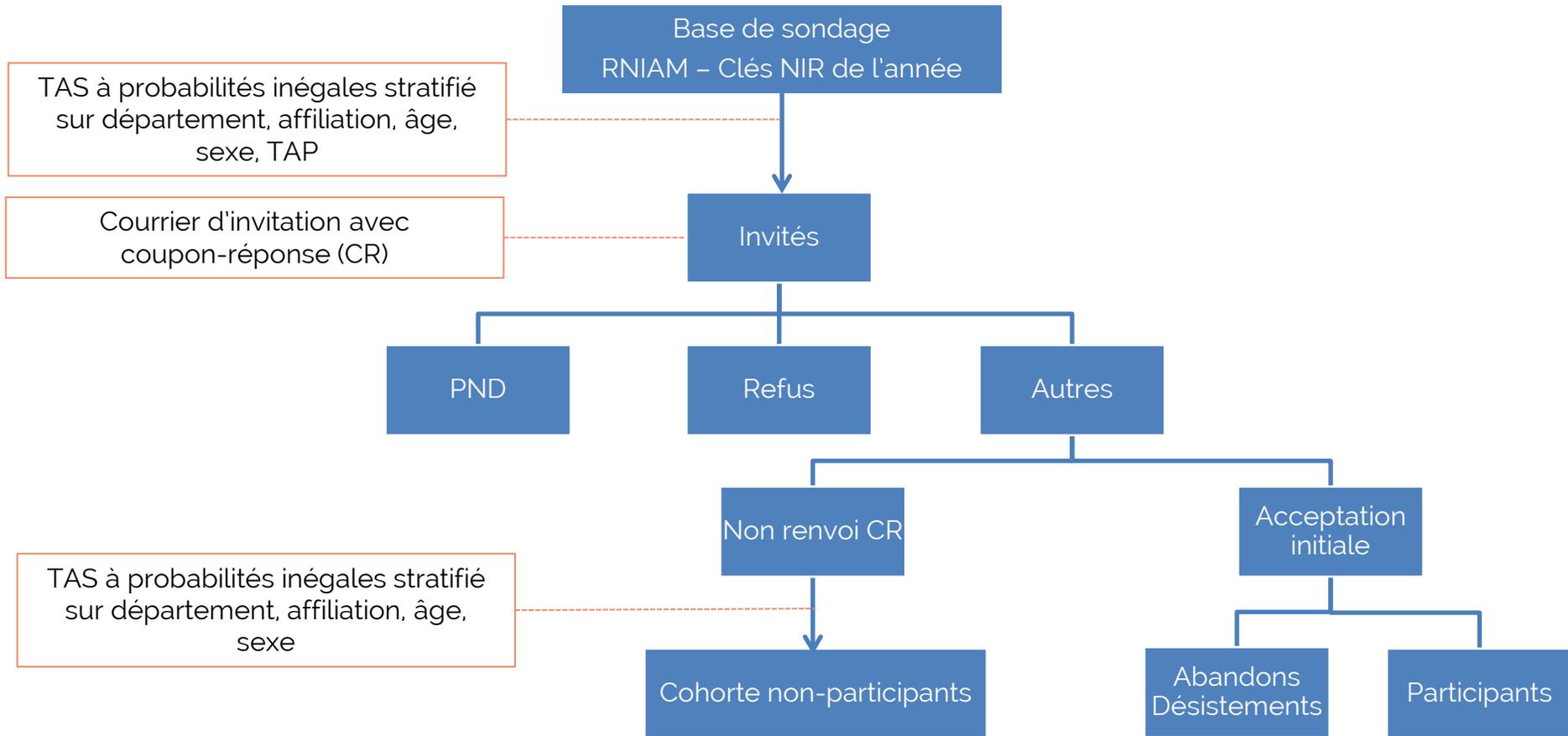
# Plan de sondage général

- > **Population cible** : Personnes résidant dans un département Constances, affiliées au RGSS (sens large) âgées de 18 à 69 ans au moment de leur invitation
- > **Base de sondage** : le RNIAM
  - > Personnes âgées de 18 à 69 ans, appartenant au RGSS (sens strict), CAMIEG, LMDE, MFPs, MGEN, affiliées à l'une des CPAM couvrant un département Constances et résidant dans l'un d'entre eux
    - Informations directement disponibles : âge, sexe, affiliation, département de résidence
    - Appariement avec le SNGC : dernière PCS connue et activité/inactivité = typologie d'activité professionnelle (TAP)
  - > Inclusion sur 6 ans (2012 à 2017) : partitionnement du RNIAM en 6 groupes distincts selon les clés NIR des personnes
    - Chaque année : une base de sondage = un groupe





# Plan de sondage pour une année donnée



# Estimation des prévalences à l'inclusion en 2013

## > **Population d'intérêt (domaine) :**

Personnes âgées de 30 à 69 ans en 2013

## > **Echantillon**

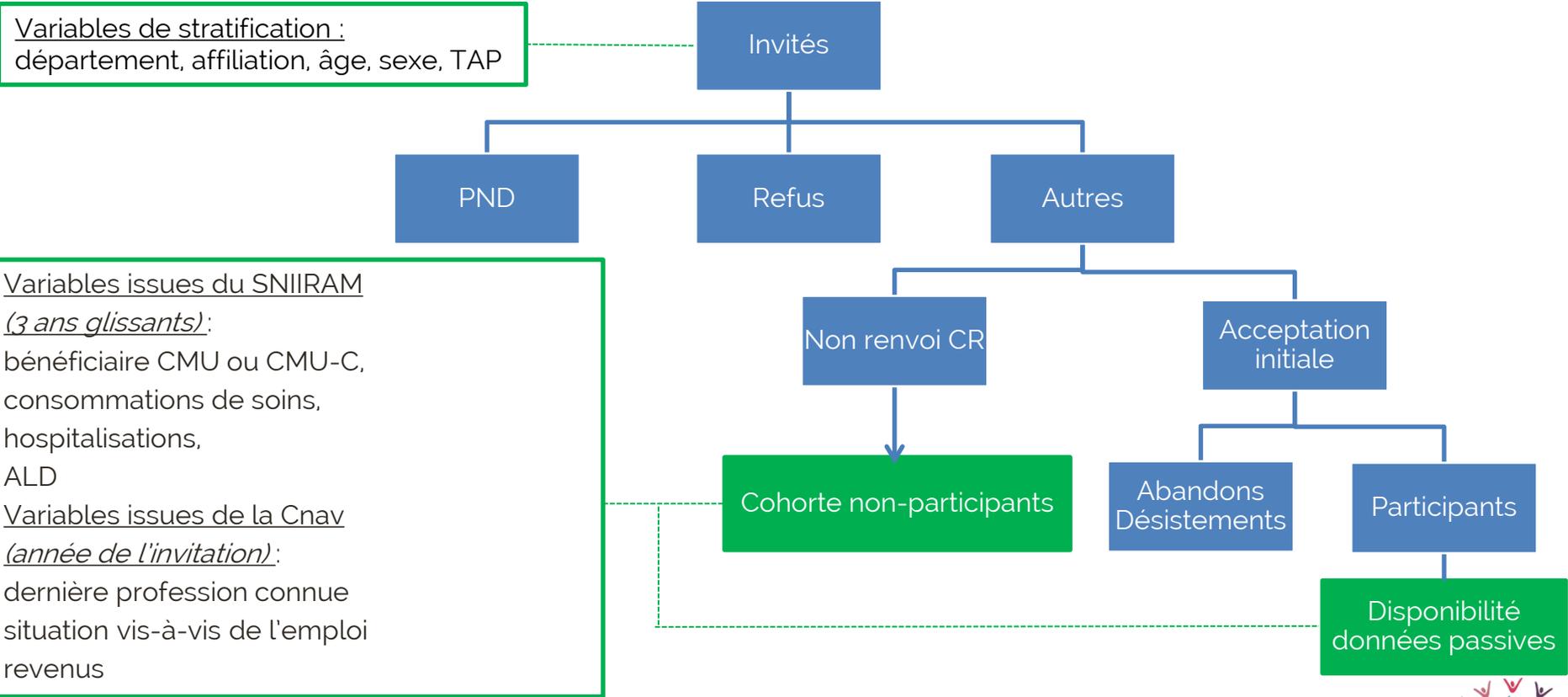
- > Approche classique : utiliser les données des invités de 2013
- > Approche moins classique : ajouter aux données des invités de 2013 les données des invités de 2012
  - Combinaison de plusieurs années déjà proposée pour estimer des prévalences d'événements rares
  - Démarche acceptable si les deux populations sources sont disjointes et correspondent en grande partie à la même population cible
    - OK pour les invités de 2012 et 2013



# Estimation des prévalences à l'inclusion en 2013

## Pondérations

- > **Poids de sondage** : Inverse de la probabilité d'inclusion
- > **Facteur correctif de la non-participation** : données disponibles





# Estimation des prévalences à l'inclusion en 2013

## Pondérations

### > **Facteur correctif de la non-participation : estimation en deux étapes**

- > Prise en compte des « non renvois de coupon-réponse »
- Etude de la participation des personnes avec données du SNIIRAM et CNAV
  - Probabilité de participation selon les variables de stratification, du SNIIRAM et de la CNAV estimée par régression logistique pondérée
  - Construction de 20 groupes homogènes vis-à-vis de la participation (GHR) par la méthode des scores
    - corr1 : inverse du taux de participation estimé dans chaque GHR
- > Prise en compte des courriers non distribuables et des refus
  - Calage de corr1 sur les variables de stratification
    - corr\_fin

### > **Poids final : produit du poids de sondage et du facteur correctif de la non-participation**

$$\text{poids\_cor} = \text{poids\_sond} * \text{corr\_fin}$$



# Estimation des prévalences à l'inclusion en 2013

## > Variables d'intérêt

- > Mesurées au CES : LDL-Cholestérol, obésité
- > Questionnaire : symptomatologie dépressive (CES-D), statut tabagique, respect mérité au travail pour les personnes en activité professionnelle

## > Prévalences estimées

- > Prévalences pondérées pour les hommes et les femmes :
  - Sans correction de la non-participation (MCAR) :  $\hat{y}_{\text{poids\_sond}}$
  - Avec correction de la non-participation (MAR) :  $\hat{y}_{\text{poids\_cor}}$

## > Changements relatifs liés à la correction de la non-participation

$$CR = \frac{\hat{y}_{\text{poids\_cor}} - \hat{y}_{\text{poids\_sond}}}{\hat{y}_{\text{poids\_cor}}} * 100$$





# Résultats





# Taux de participation

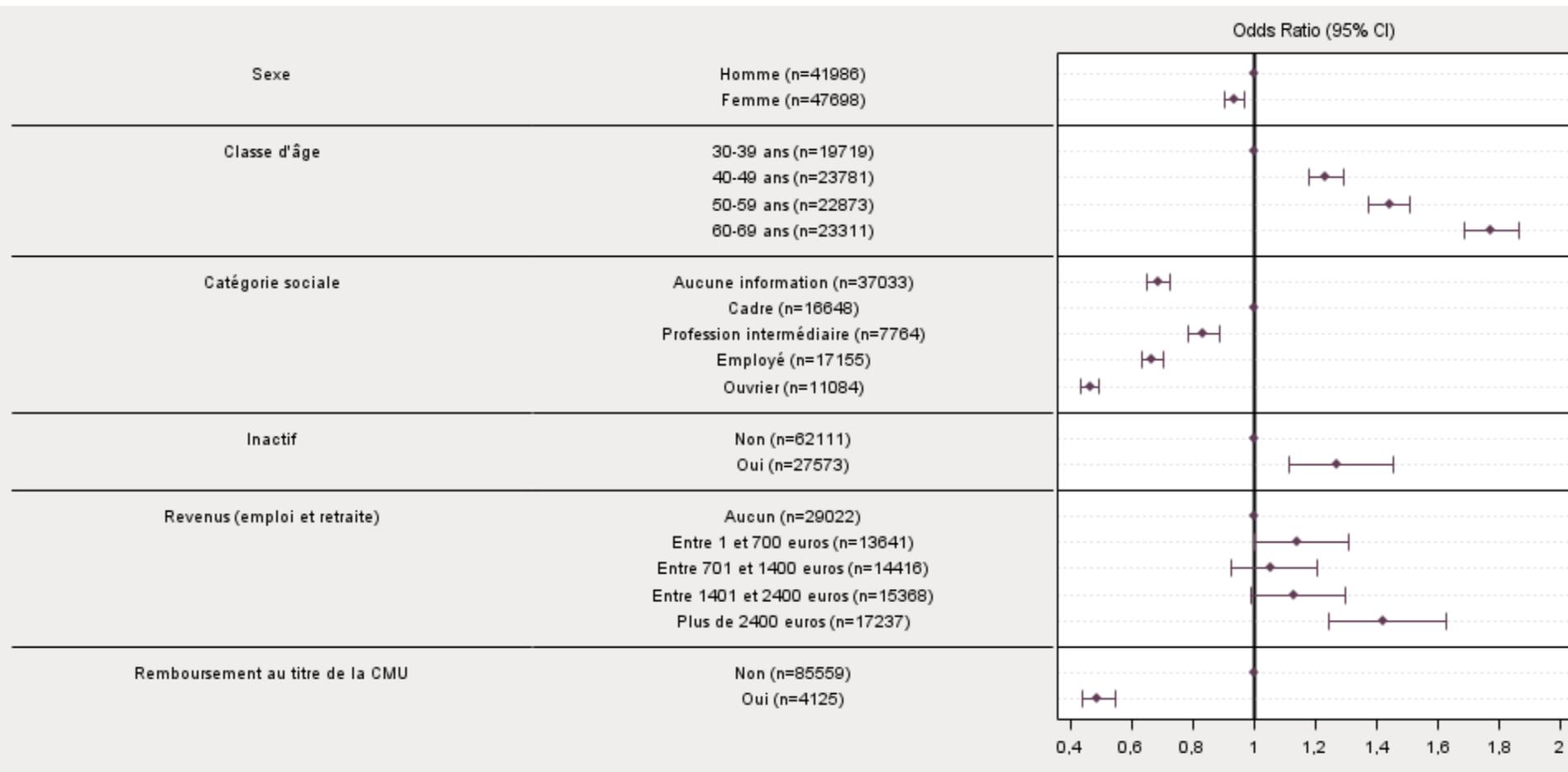
- > 471 152 invités en 2012 et 2013 âgés d'au moins 30 ans au 1<sup>er</sup> janvier 2013
- > 7,5 % de taux de Plis Non Distribuables parmi les invités
  - > 1,5% de taux de refus parmi les invités ayant reçu un courrier d'invitation
    - > 7,4% de taux de participation (n=31 642) parmi les invités ayant reçu un courrier d'invitation et n'ayant pas refusé de participer
- > Données du SNIIRAM et de la CNAV pour 28 914 participants





# Facteurs associés à la participation

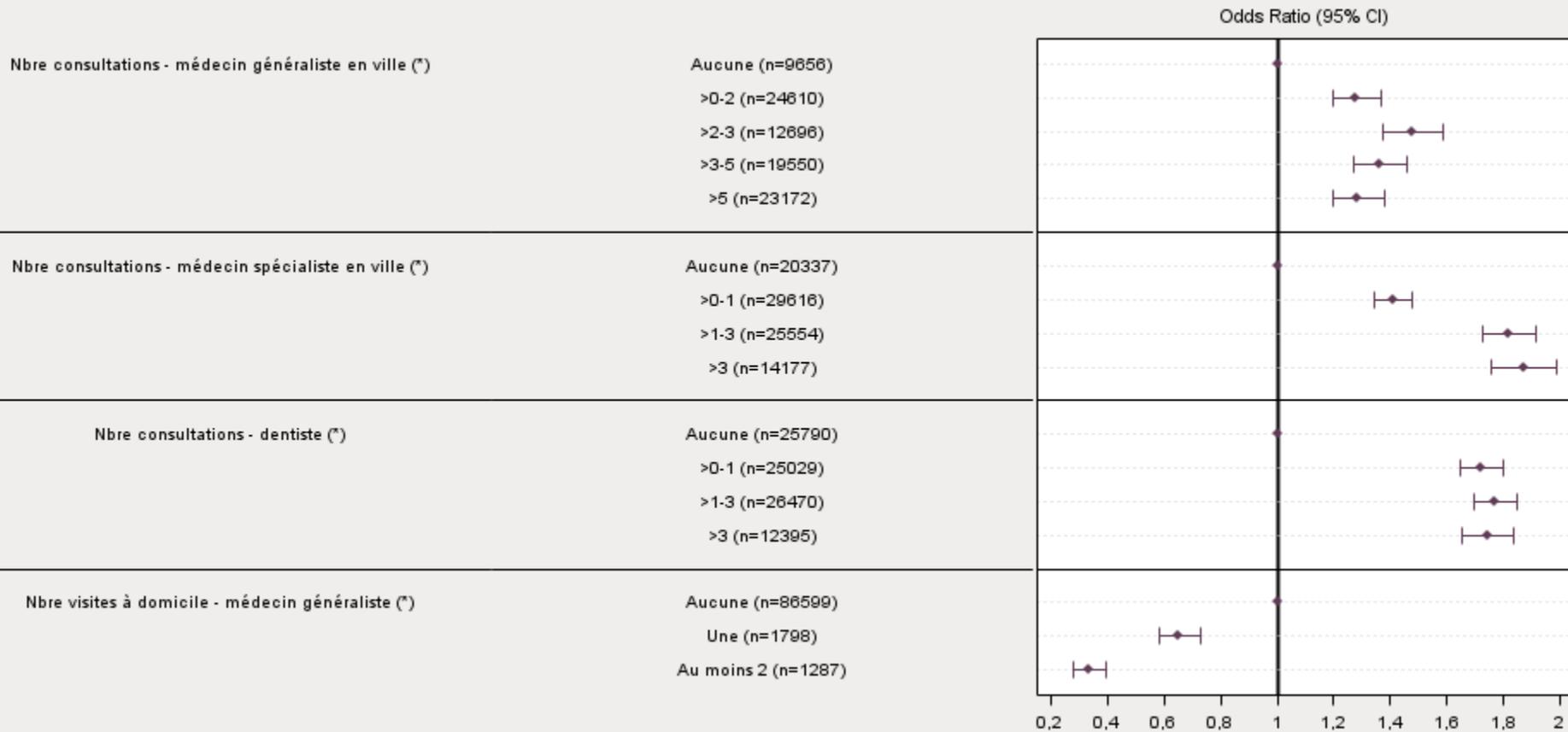
## Modèle final (1/3)





# Facteurs associés à la participation

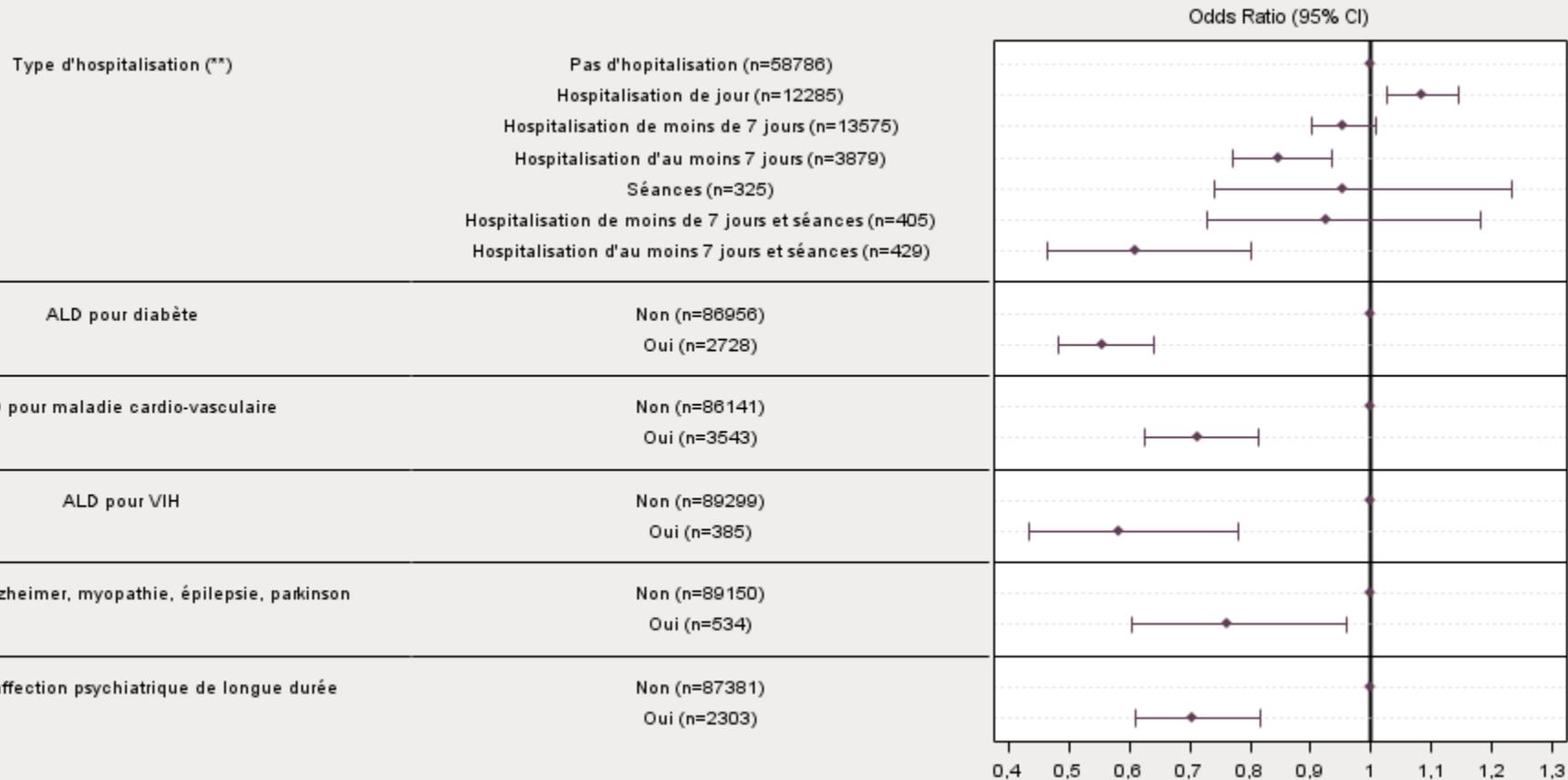
## Modèle final (2/3)





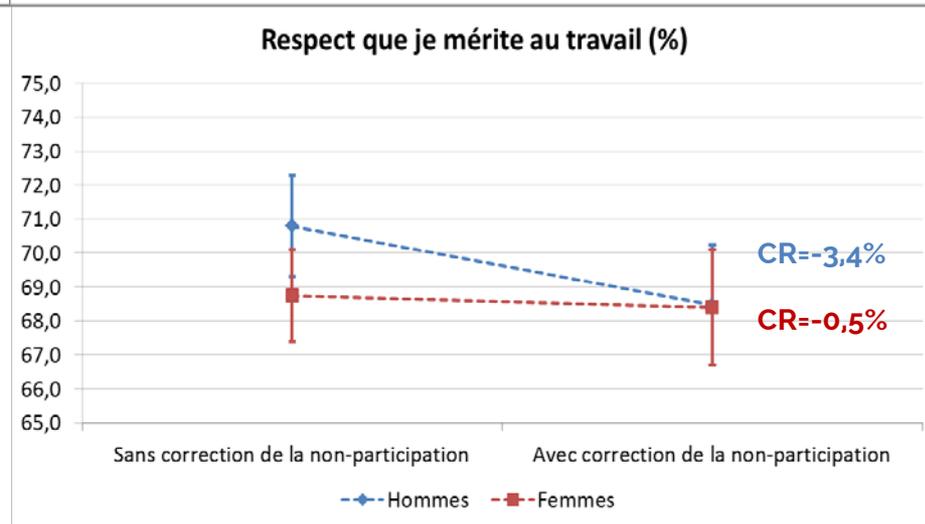
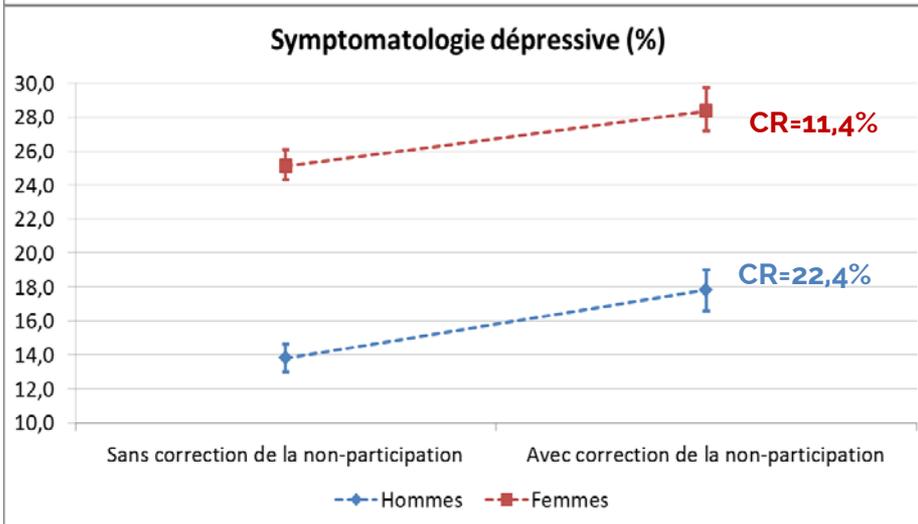
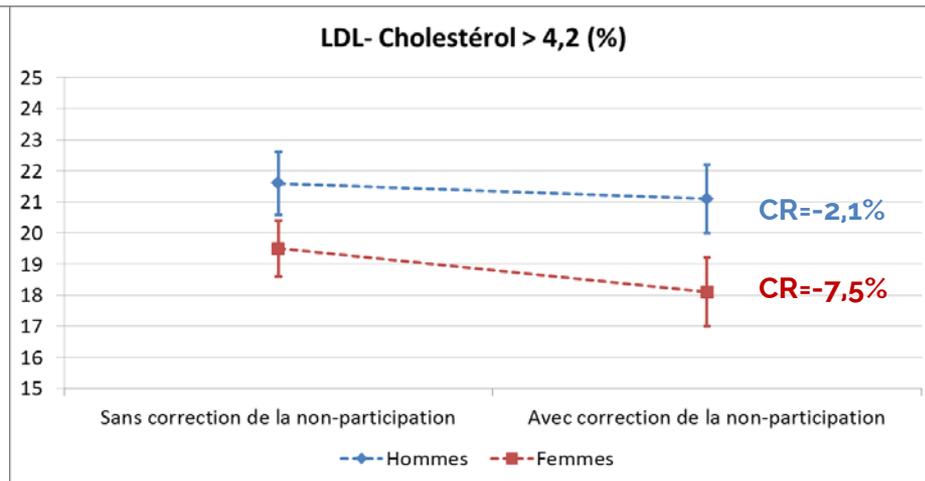
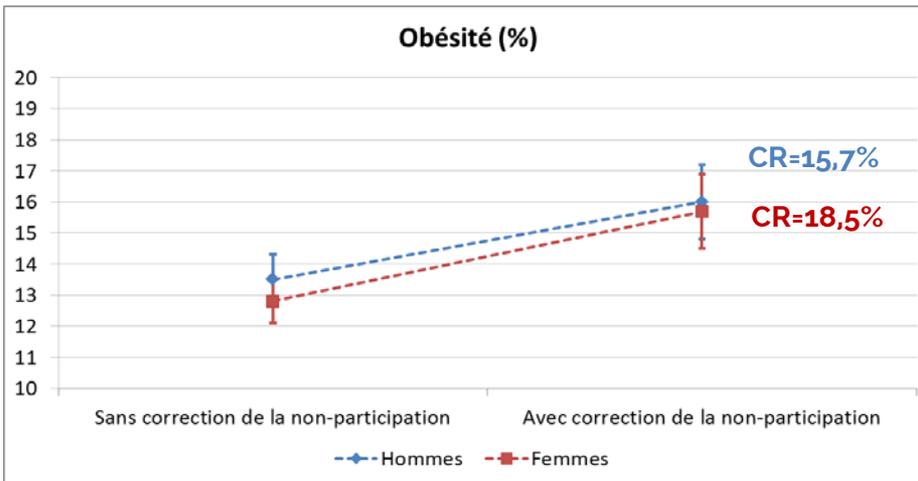
# Facteurs associés à la participation

## Modèle final (3/3)





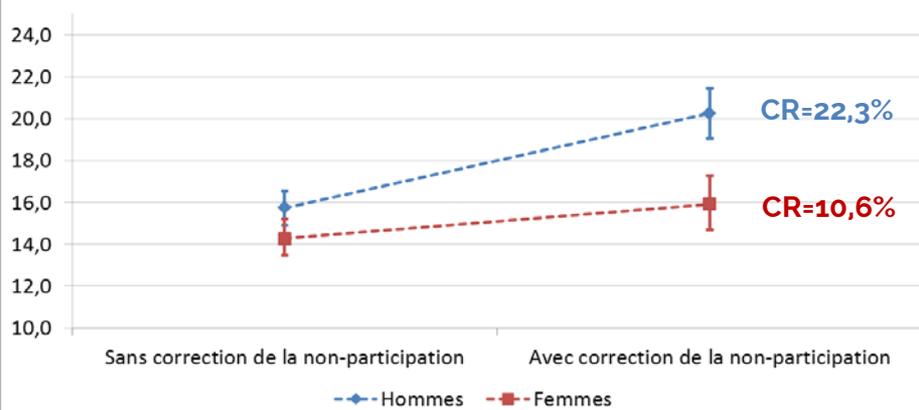
# Prévalences estimées sans et avec correction de la non-participation



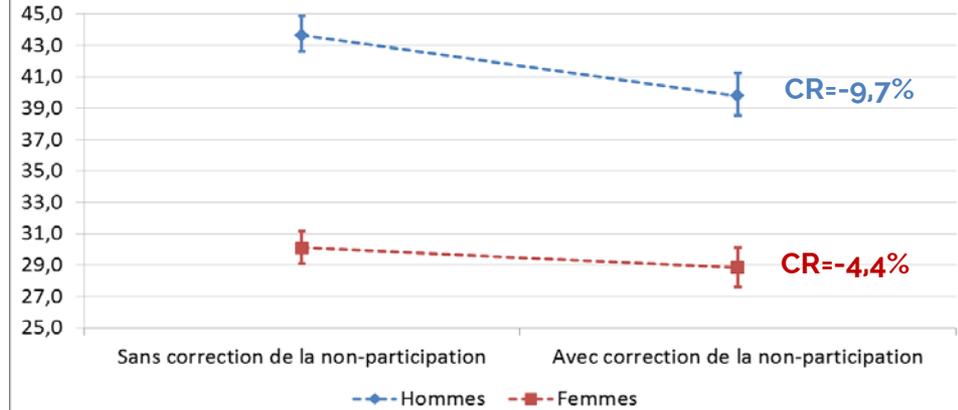


# Prévalences estimées sans et avec correction de la non-participation

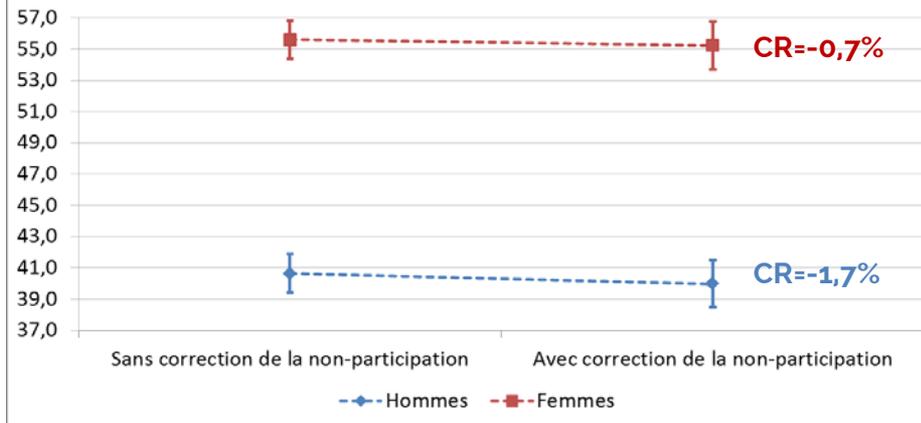
Statut tabagique - Fumeur (%)



Statut tabagique - Ex fumeur (%)



Statut tabagique - Non fumeur (%)





# Discussion - Perspectives





# La participation

- > Participation associée à des variables socio-économiques, professionnelles et relatives à la santé
  - Concordant avec la littérature épidémiologique
  
- > Prise en compte de la participation dans l'estimation des prévalences ayant un effet faible à important selon les variables d'intérêt
  - > Prévalences augmentées pour la symptomatologie dépressive, l'obésité et le statut tabagique « fumeur »
  - > Peu de variations pour « respect que je mérite au travail »
    - Concordant avec la littérature épidémiologique
  
- > Cohérence des résultats mais prise en compte de tous les facteurs liés à la participation non garantie





## La combinaison de deux années d'invitation

- > Validité de cette démarche qui repose sur les hypothèses suivantes :
  - > Population source de 2012 et 2013 ayant la même structure et changements de strate d'une année à l'autre négligeables
  - > Pas d'effet période affectant la variable d'intérêt
- > Combinaison conduisant à des effectifs importants :
  - > Avantage lorsqu'on étudie une pathologie ou une exposition rares
  - > Incertitude estimée par les intervalles de confiance réduite





# Apport de Constances pour la santé publique

- > Label d'intérêt général et de qualité statistique du CNIS
- > Cohorte qui permettra de compléter les informations de santé publique déjà disponibles comme dans l'Enquête ESPS et le Baromètre Santé
- > Faible taux de participation (7%) mais :
  - > Disponibilité de données mesurées dans les CES
  - > Appariement avec des bases de données médico-administratives





# Perspectives

- > **Nombreuses !**
- > Etude plus fine des pondérations :
  - > Modèles de non-participation plus complexes
  - > Analyses de robustesse (prise en compte des poids dispersés)
  - > Calage
  - > Analyses de sensibilité
- > Traitement de la non-réponse partielle
- > Pour une pathologie ou une population d'intérêt particulière, construction de modèles spécifiques
- > Calage supplémentaire pour extrapoler les résultats à l'ensemble des affiliés du Régime général, voire de la population française possible par les chercheurs s'ils considèrent que cette extrapolation supplémentaire engendre des biais négligeables





Merci pour votre attention

