

**L'utilisation des données génétiques
pour mieux comprendre l'étiologie des maladies complexes
dans les études épidémiologiques en population générale**

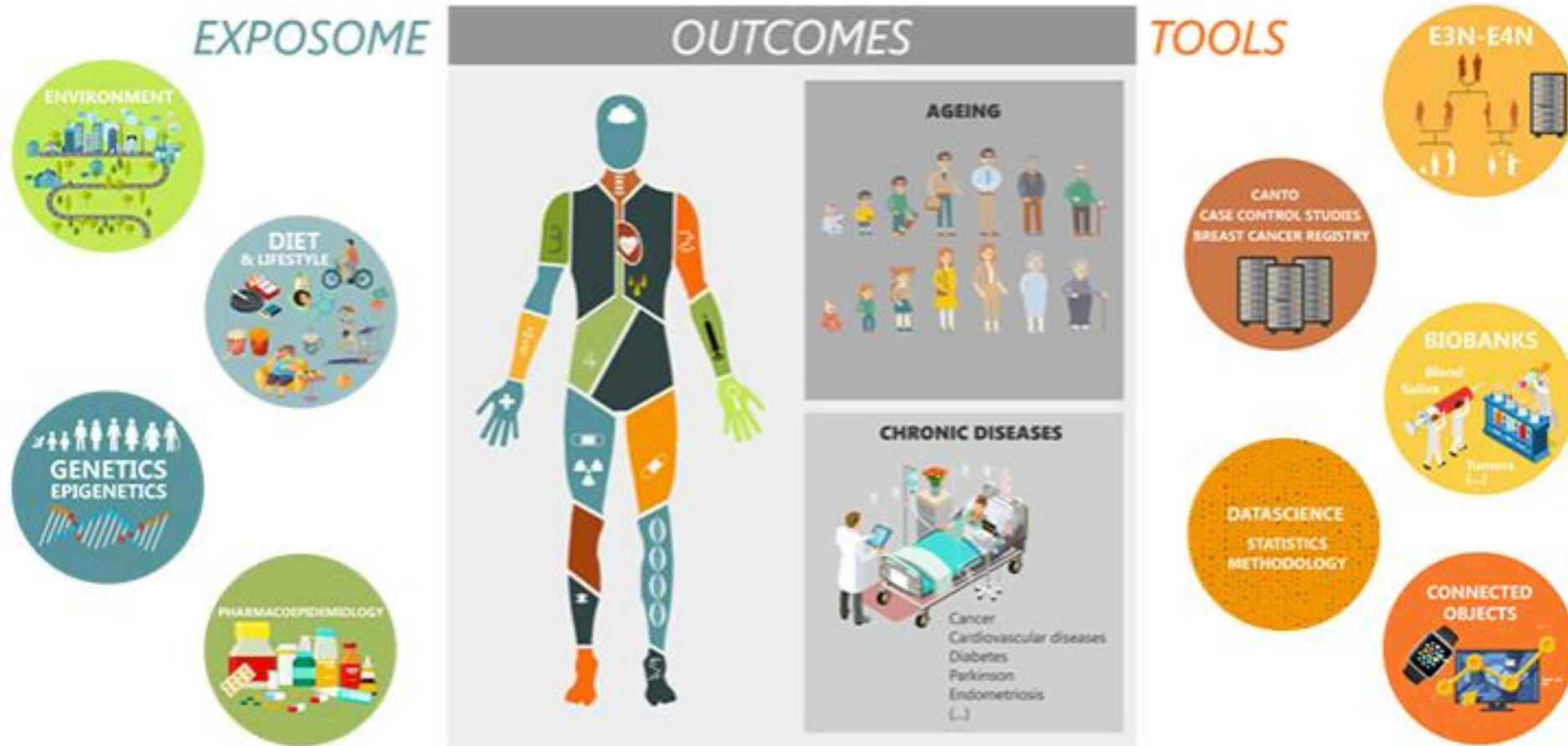
Thérèse Truong

INSERM U1018 - CESP

Centre de recherche en Epidémiologie et Santé des Populations

Equipe Exposome et Hérité

Equipe Exposome et Hérité (dir: Gianluca Severi)



- Exemples de 2 études épidémiologiques :
- Etude cas/témoin de Nouvelle-Calédonie
 - Cohorte E3N-Génération

Contexte en Nouvelle-Calédonie

Incidence des cancers de la thyroïde dans la période ~1985-1990

- + fréquent chez les femmes
- Plus élevée dans les Iles du Pacifique : ++ en NC
- Différence d'incidence par ethnie : ++ mélanésiens de NC
- Seul facteur de risque connu = radiations ionisantes
Mais pas d'essai nucléaire en NC

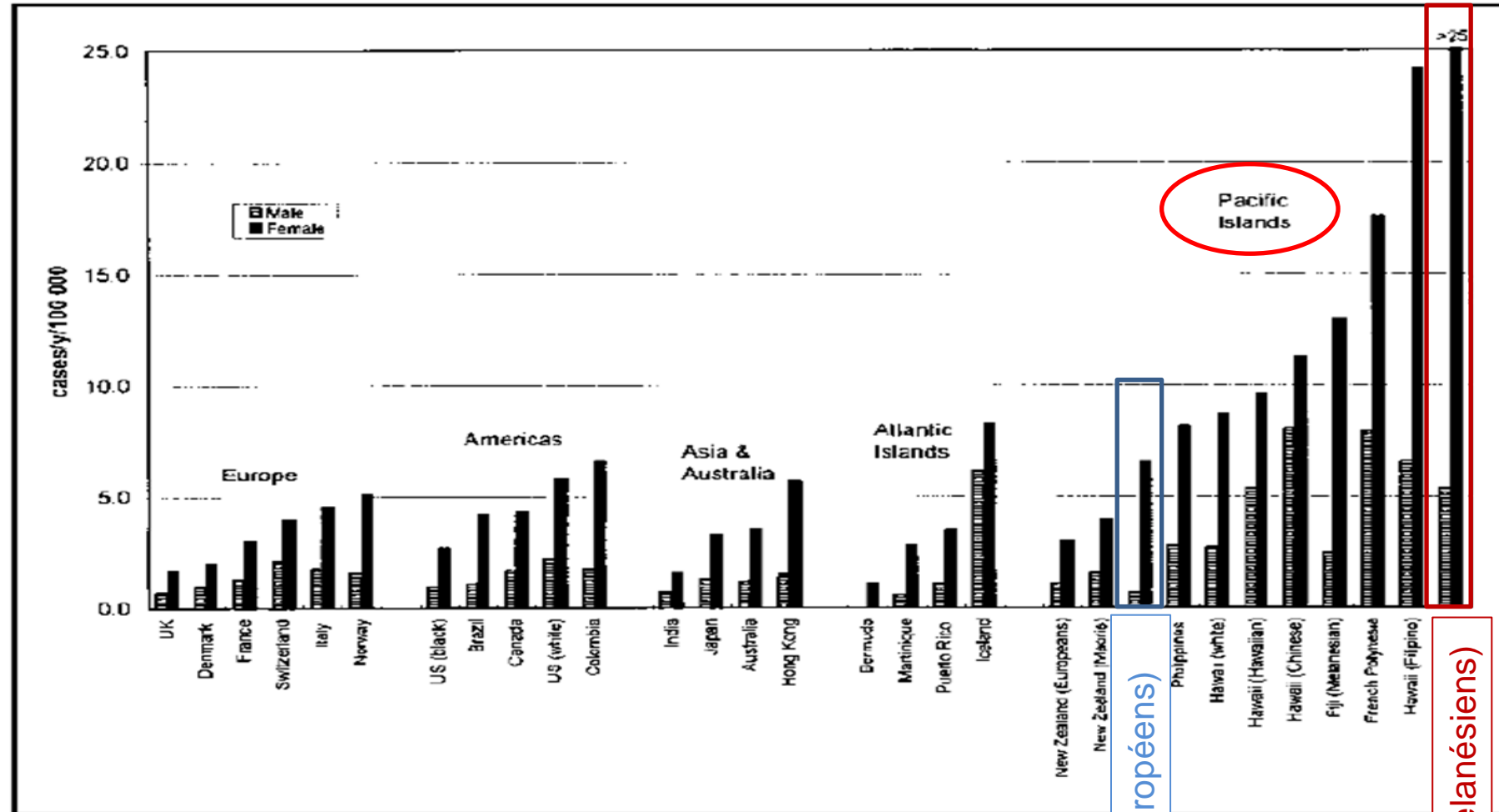


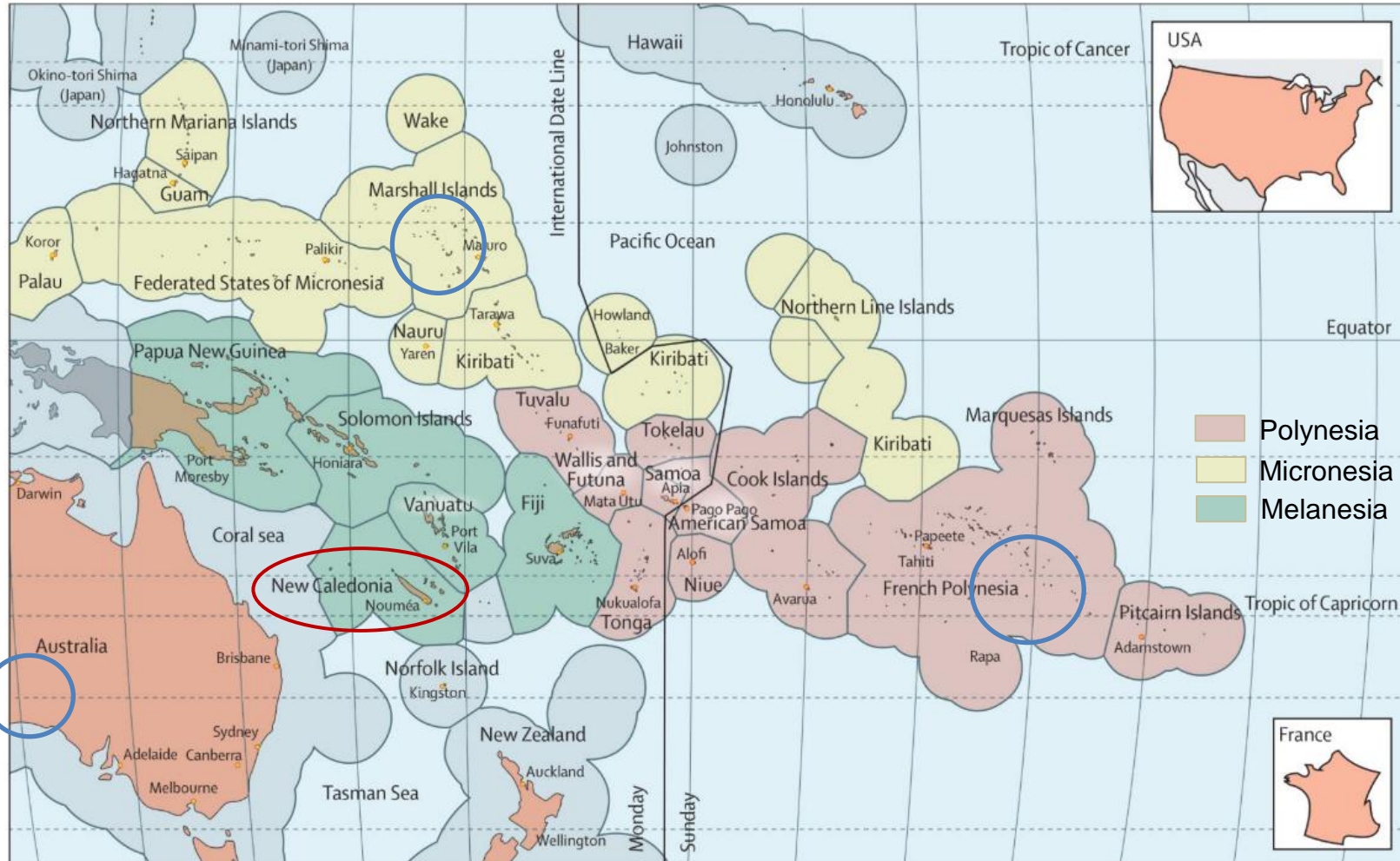
Fig. 1. Rates for selected countries and territories. The data were derived in most part from cancer registries reporting for the mid-1980s to the International Agency for Research on Cancer program of cancer incidence in five continents (1), as well as from updated rates from cancer registries operating in the South Pacific (2). For several countries with multiple cancer registries (e.g., the U.K. and France) the within-country rates across the within-country registries are shown.

Nouvelle-Calédonie



Map from Sarfati et al. Lancet Oncology 2019 –
Cancer control in the Pacific: big challenges facing small islands states

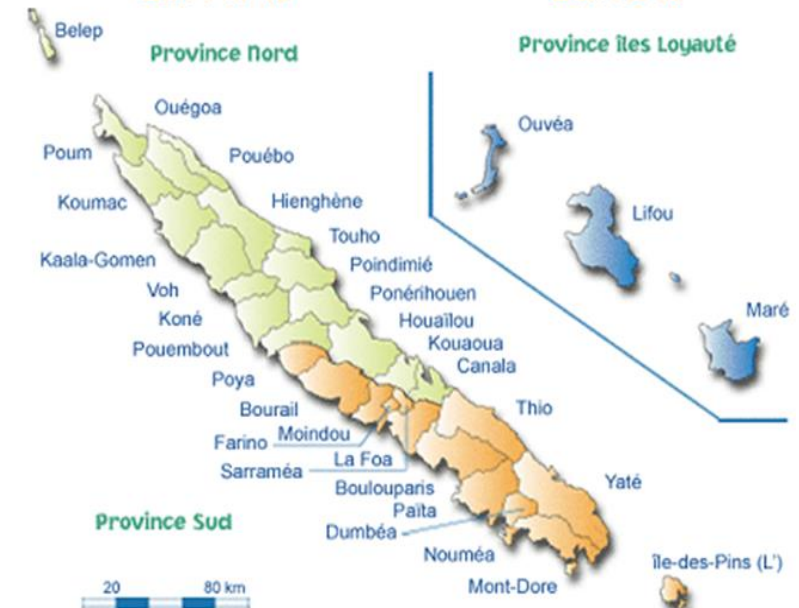
Nouvelle-Calédonie



200 000 habitants en 2000
(45 % Mélanésiens, 35% Européens)

21 % de la pop. totale
Mél. : 79 %
Eur. : 17 %

11 % de la pop. totale
Mél. : 97 %
Eur. : 2 %



68 % de la pop. totale
Mél. : 26 %
Eur. : 44 %

Map from Sarfati et al. Lancet Oncology 2019 –
Cancer control in the Pacific: big challenges facing small islands states

Etude cas-témoins en Nouvelle-Calédonie

Etude conduite en 1998-2002

- 332 cas de cancer de la thyroïde diagnostiqués entre 1993-1999
- 412 témoins appariés sur âge et sexe
- Recueil des données via un questionnaire détaillé (facteurs socio-démo, alimentation, taille-poids, facteurs hormonaux, etc)
- **Association nette avec la parité et l'obésité**
- **Pourraient expliquer jusqu'à environ 20% de l'excès de risque chez les Mélanésiens**

*Truong et al. Am J Epidemiol 2005,
Guignard et al. Am J Epidemiol 2007,
Truong et al. Cancer Causes & Control 2010
Leux et al. Cancer Causes & Control 2012
Paquet et al. Am J Epidemiol 2020*

Etude ancillaire THYROGENE en 2008

- Prélèvements salivaires (participation ~70%) et nouveau consentement pour l'analyse des données génétiques

Etude EPITHYR en 2016

- Génotypage pangénomique par la puce OncoArray (Illumina) dans le cadre d'un consortium
- **Pas de nouveau signal identifié chez les Mélanésiens/Polynésiens**

*Tcheandjieu et al. Int J Cancer 2016
Truong et al. Int J Cancer 2021*

Difficultés et questionnements

■ Difficultés d'analyse des données sur la pop mélanésienne

- Manque de puissance statistique
- Qualité de l'imputation des données génétiques faible
- Pas de population de réplication

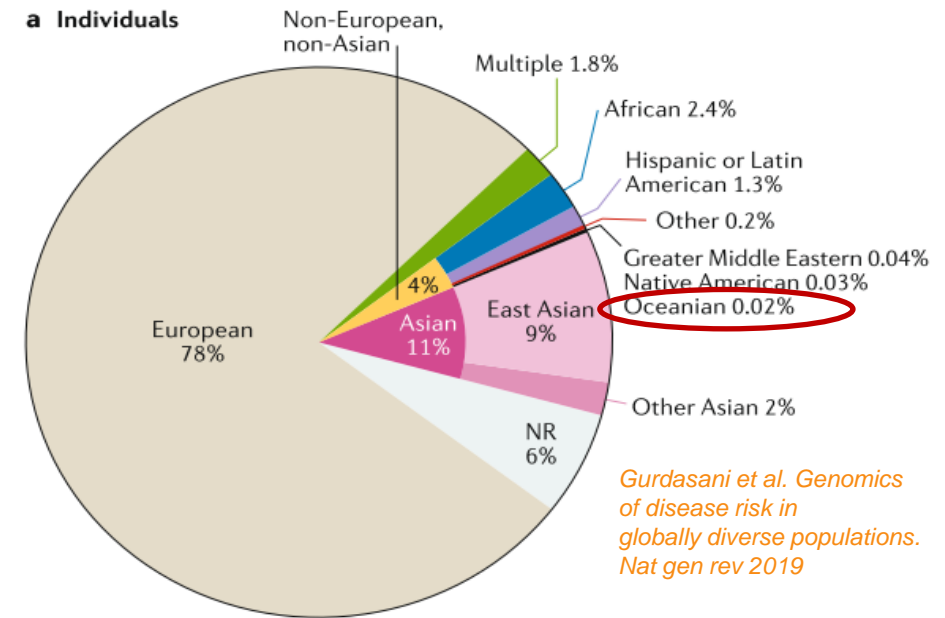
■ Perspectives

- Séquençage des participants Mélanésiens et Polynésiens (Collab : F. de Vathaire, CNRGH)
- Analyse des segments HBD (variants récessifs rares) (Collab : AL. Leutenegger, H. Perdy)

■ En 2017: Demande d'une équipe Institut Pasteur pour une analyse de génétique des populations sur les océaniens

- Consentement individuel spécifique pour l'étude cancer de la thyroïde
- Demande de l'institut Pasteur au Sénat Coutumier → refus
 - selon le droit coutumier: « la génétique des Mélanésiens n'appartient pas à l'individu mais à son clan »
 - contexte politique tendu (référendum indépendance NC)

→ Mise en conformité RGPD de l'étude, questionnements éthiques (discussions en cours avec CNIL et DPD Inserm)



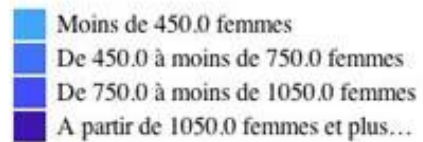
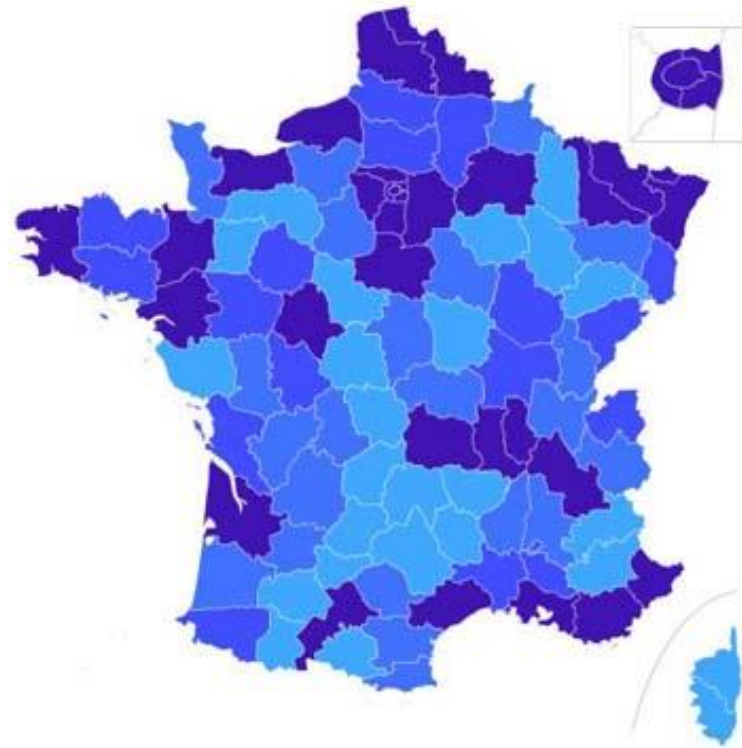
Cohorte E3N

Femmes E3N



98 998 femmes nées
entre 1925 et 1950
affiliées à la MGEN
(Mutuelle Générale de
l'Education Nationale)

Répartition des femmes à Q1

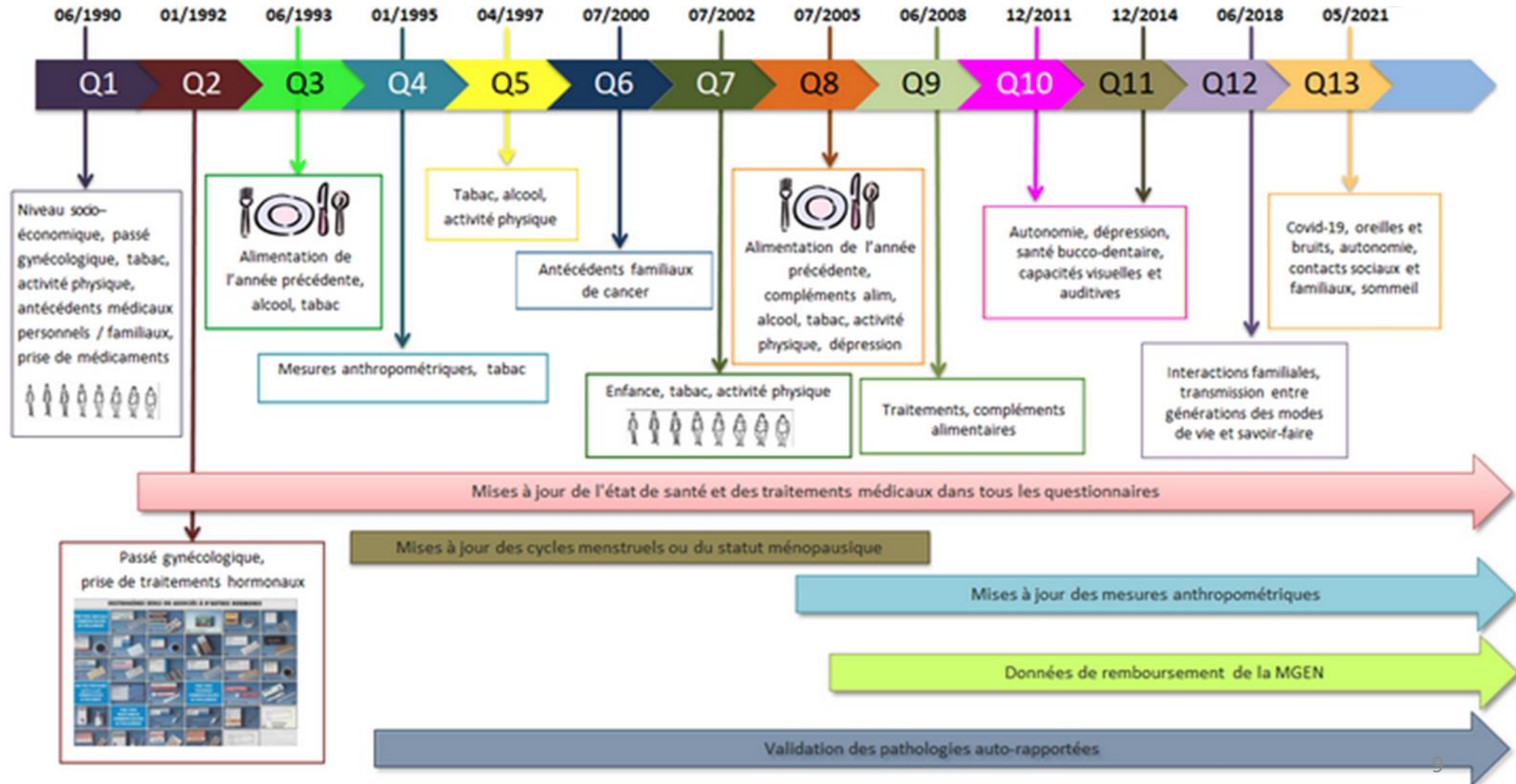


Cohorte E3N

Femmes E3N



98 998 femmes nées entre 1925 et 1950 affiliées à la MGEN (Mutuelle Générale de l'Education Nationale)



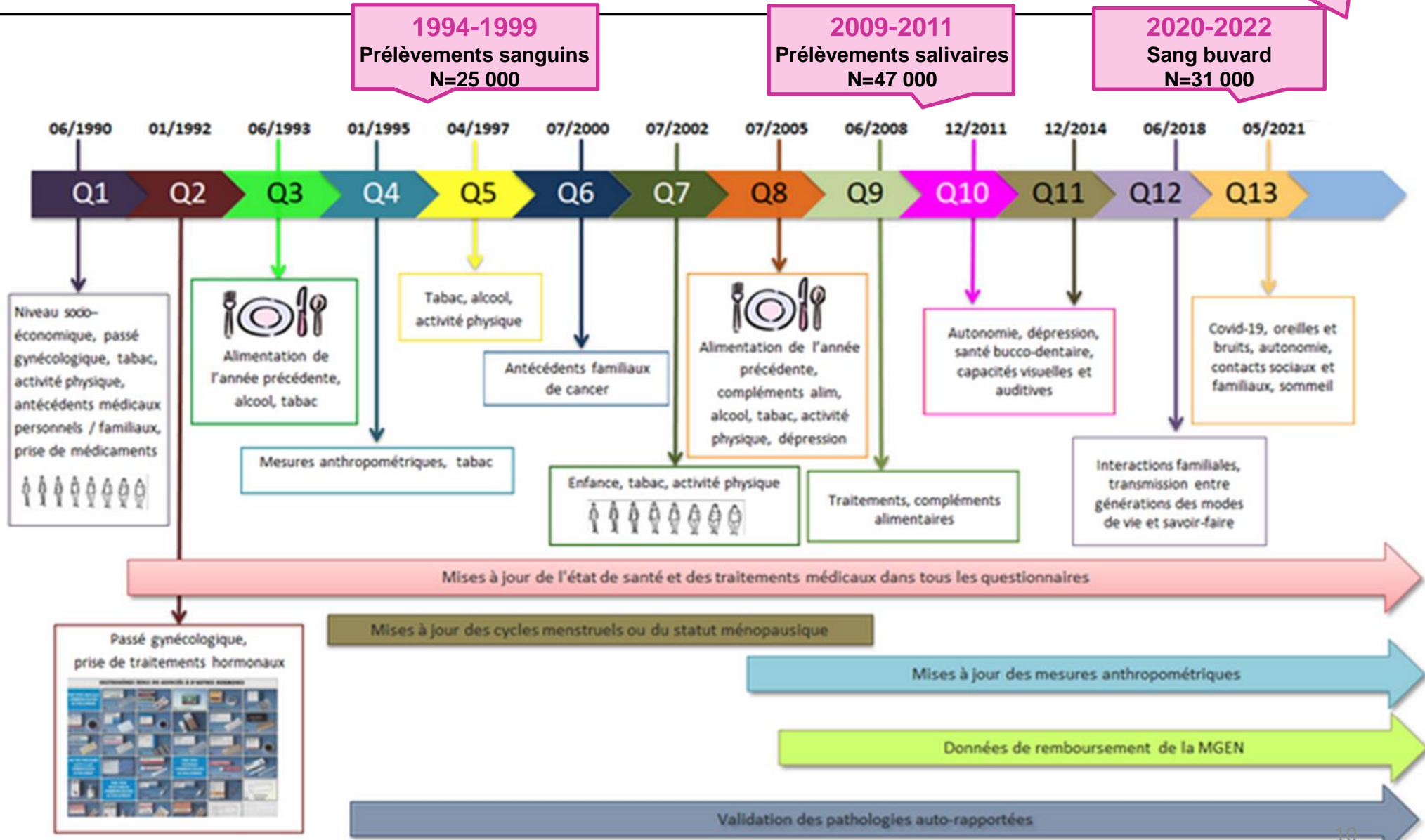
Cohorte E3N

En cours
Tumeurs

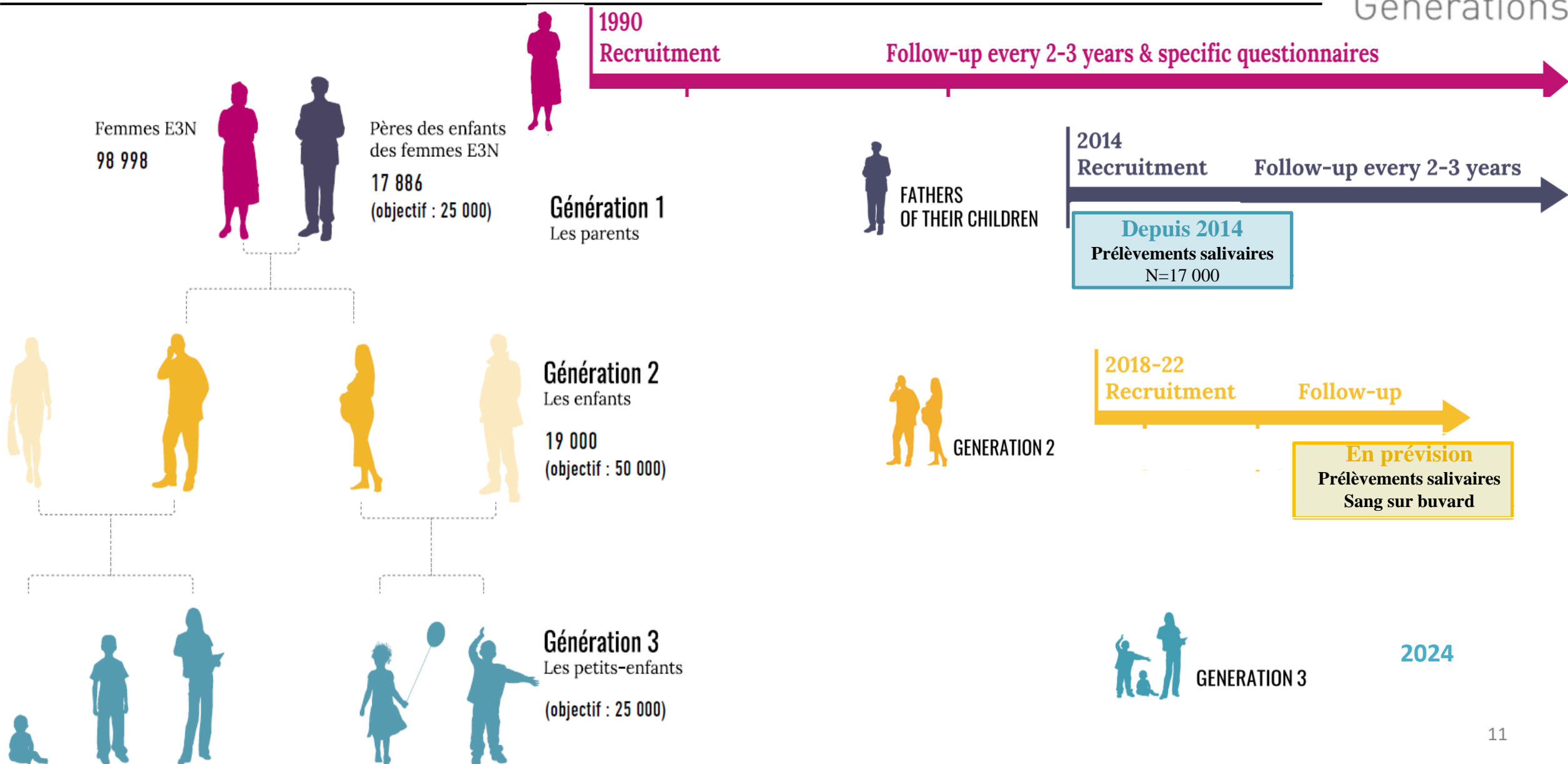
Femmes E3N



98 998 femmes nées entre 1925 et 1950 affiliées à la MGEN (Mutuelle Générale de l'Education Nationale)



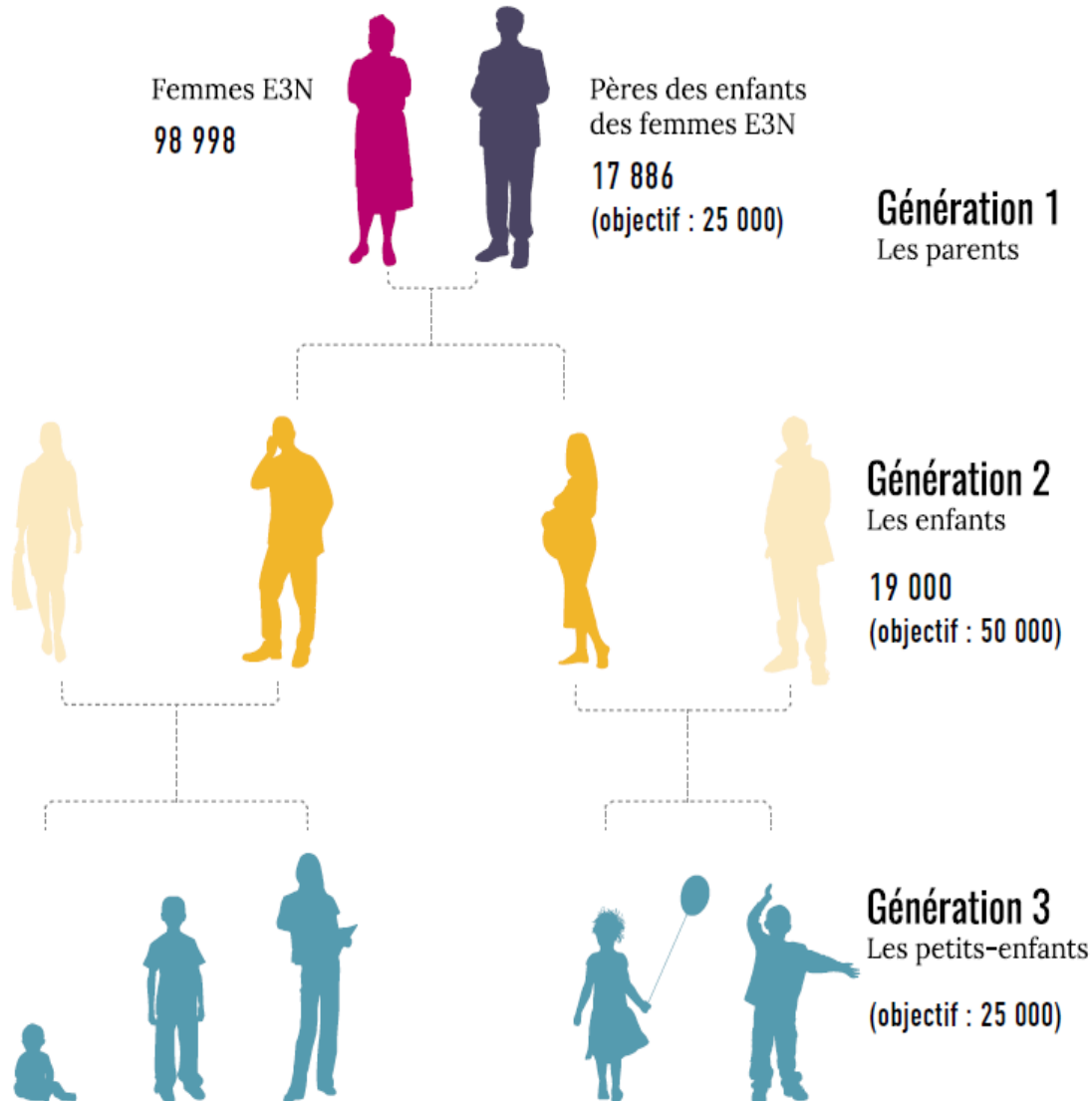
Cohorte E3N-Génération



Cohorte E3N-Génération



~ 136 000 participants (objectif : 200 000)



Biobanque centralisée au CEPH

- Génération 1
 - 65 000 prélèvements salivaires
 - 25 000 prélèvements sanguins
 - 31 000 prélèvements sanguins sur buvard
- Génération 2 (en prévision)

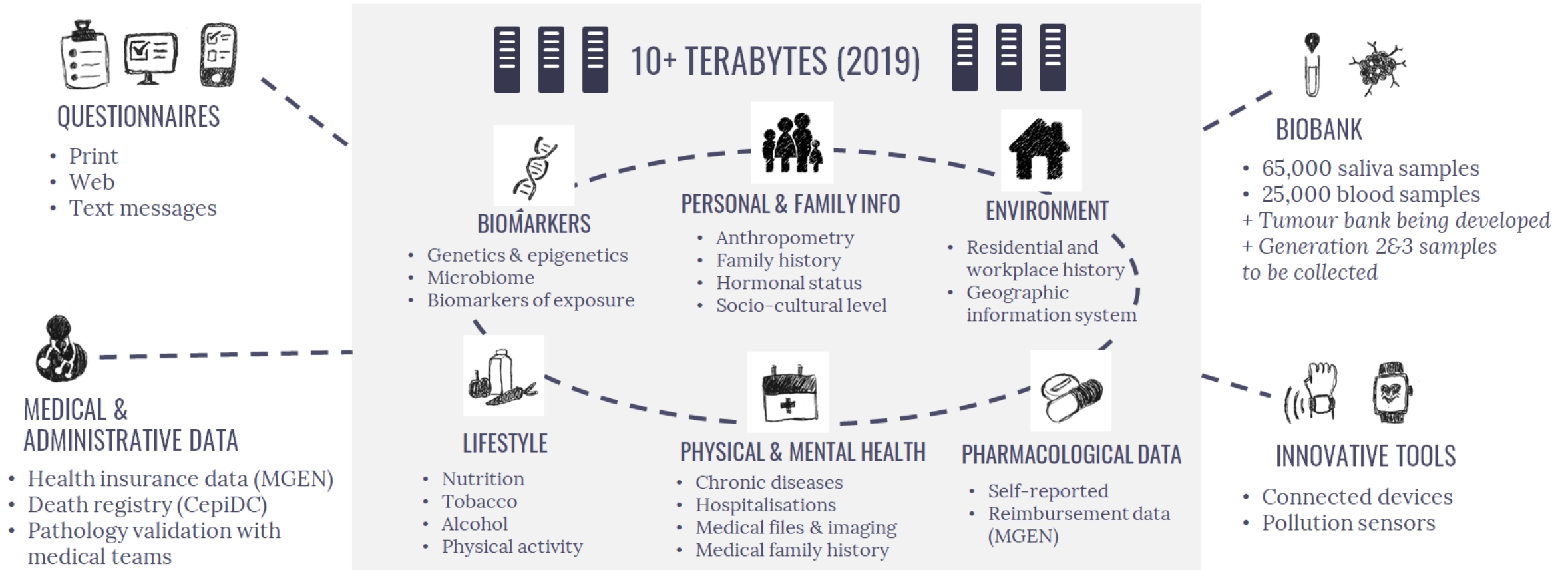
Données génétiques

- Génotypage de G1 prévu fin 2023-2024 (projet BioCF)

Autres données –omiques (divers projets)

- Épigénétique
- Métabolomique
- Microbiome
- Signatures mutationnelles

Cohorte E3N-Génération



Apport des données génétiques

- **Interaction GxE**

Améliorer la prédiction de risque / Mécanisme biologique sous-jacent

- **Randomisation Mendélienne**

Inférence causale entre un facteur de risque modifiable et risque de maladie

- **Pléiotropie**

Analyse des facteurs de risque génétique communs à différents phénotypes pour identifier mécanismes biologiques

- **Signature mutationnelle**

Identification des processus mutationnels ou expositions mutagènes à partir de l'ADN tumoral

Interaction GxE

Exemple : Cancer du poumon

- GWAS ont identifié locus 15q25 (comprend les gènes *CHRNA3*, *CHRNA5* et *CHRNA4* encodant des sous-unités du récepteur nicotinique)
- rs16969968 dans *CHRNA5* également associé à l'addiction au tabac

→ Pas d'association avec cancer du poumon chez les non fumeurs

CHRNA5_rs16969968

	Ca	Co	OR	95%CI
Co-dominant	6888	10040	1.26	1.20-1.32

By histology (p-heterogeneity=0.5763)

Adenocarcinomas	2826	10040	1.21	1.14-1.29
Squamous	1640	10040	1.30	1.20-1.40
Large cell	327	10040	1.27	1.08-1.49
Small cell	967	10040	1.22	1.10-1.34

By smoking status (p-heterogeneity=0.0001)

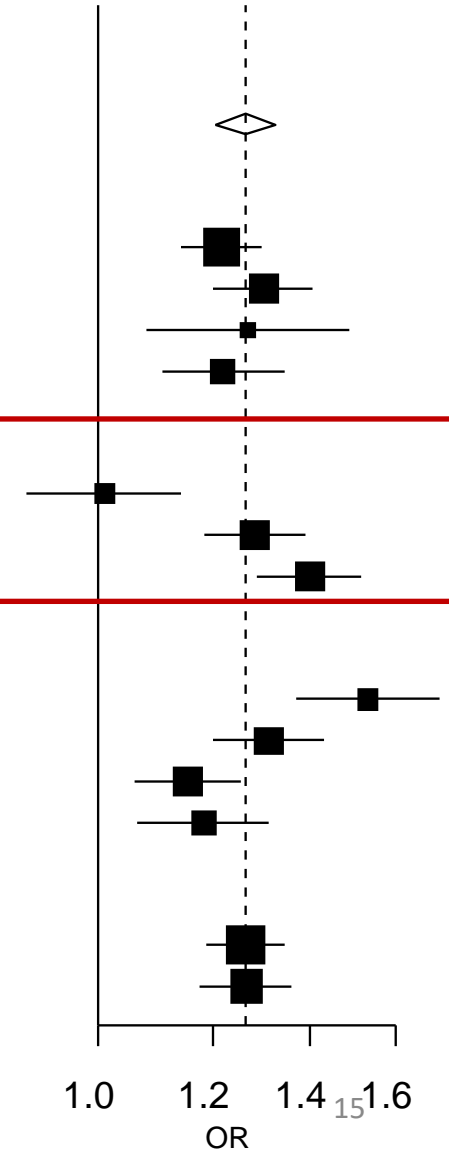
Never smokers	754	3531	1.01	0.89-1.14
Former smokers	2659	3648	1.28	1.18-1.39
Current smokers	3290	2404	1.40	1.28-1.52

By age (p-heterogeneity=0.0004)

<50	1025	2099	1.53	1.37-1.72
50-60	1907	2861	1.31	1.20-1.43
60-70	2317	3131	1.15	1.06-1.25
>=70	1639	1949	1.18	1.06-1.31

By gender (p-heterogeneity=0.9591)

Men	4002	5812	1.26	1.18-1.34
Women	2886	4228	1.26	1.17-1.36



Interaction GxE

Exemple : Cancer du poumon

- GWAS ont identifié locus 15q25 (comprend les gènes *CHRNA3*, *CHRNA5* et *CHRNA4* encodant des sous-unités du récepteur nicotinique)
- rs16969968 dans *CHRNA5* également associé à l'addiction au tabac

→ Pas d'association avec cancer du poumon chez les non fumeurs

→ Association avec le nombre de cigarettes/jour

Nombre de cigarette/jour selon génotype

	All			Controls		
	n	mean	CI 95%	n	mean	CI 95%
rs16969968 (<i>CHRNA5</i>) GG	5425	20.74	20.36-21.12	2610	17.99	17.45-18.53
GA	6597	21.85	21.49-22.20	2701	19.20	18.67-19.78
AA	2039	23.48	22.92-24.04	746	20.56	19.68-21.44
<i>p-trend</i>			7×10^{-19}			6×10^{-9}

Means are adjusted for age, sex, study and case/control status when appropriate

Interaction GxE

Exemple : Cancer du poumon

- GWAS ont identifié locus 15q25 (comprend les gènes *CHRNA3*, *CHRNA5* et *CHRNA4* encodant des sous-unités du récepteur nicotinique)
- rs16969968 dans *CHRNA5* également associé à l'addiction au tabac

→ Pas d'association avec cancer du poumon chez les non fumeurs

→ Association avec le nombre de cigarettes/jour

Nombre de cigarette/jour selon génotype

	All			Controls		
	n	mean	CI 95%	n	mean	CI 95%
rs16969968 (<i>CHRNA5</i>) GG	5425	20.74	20.36-21.12	2610	17.99	17.45-18.53
GA	6597	21.85	21.49-22.20	2701	19.20	18.67-19.78
AA	2039	23.48	22.92-24.04	746	20.56	19.68-21.44
<i>p-trend</i>			7×10^{-19}			6×10^{-9}

Means are adjusted for age, sex, study and case/control status when appropriate

Prise en compte de E et de G

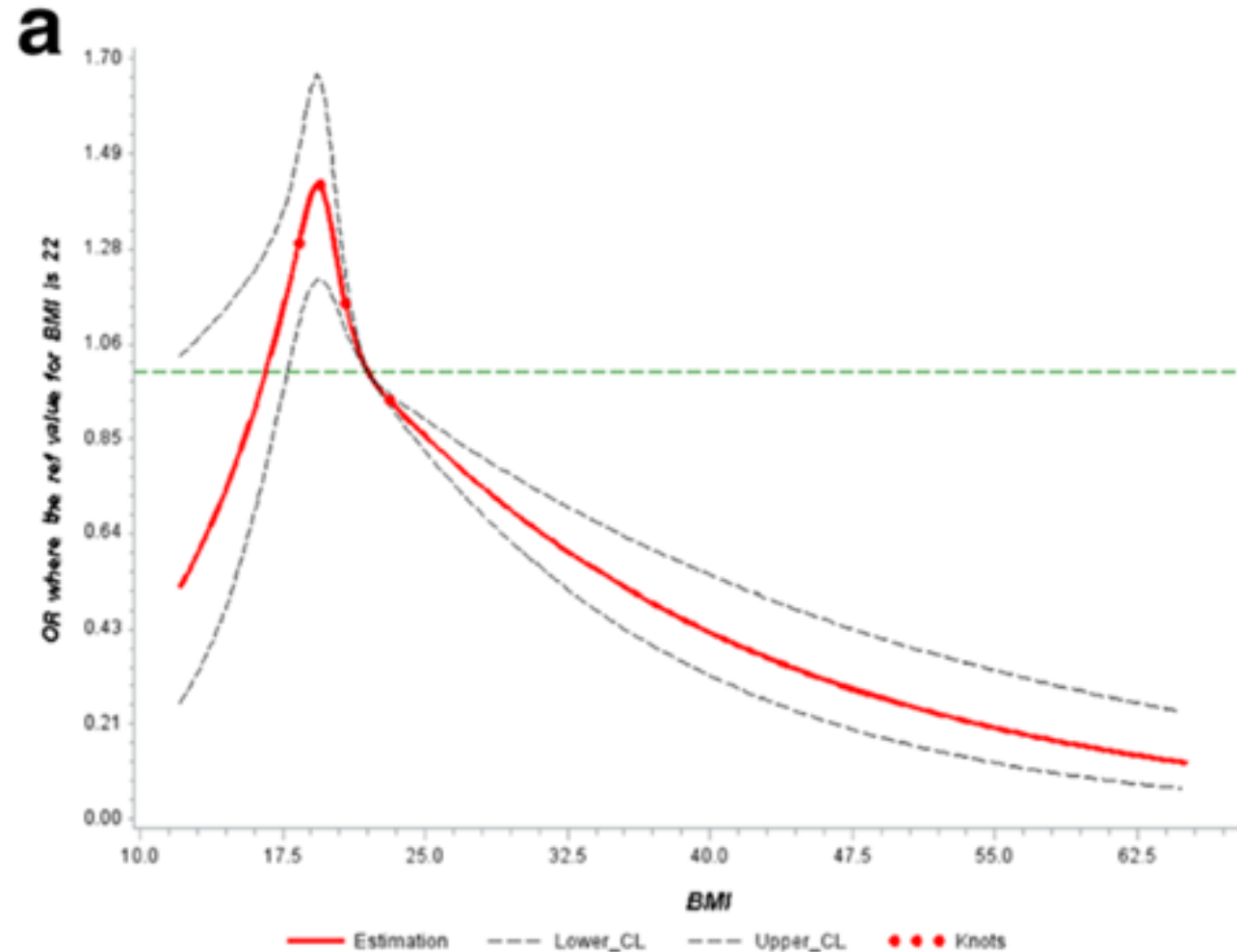
- Hypothèse sur le mécanisme biologique sous-jacent
- Potentiellement importante pour la prédiction de risque

Randomisation mendélienne

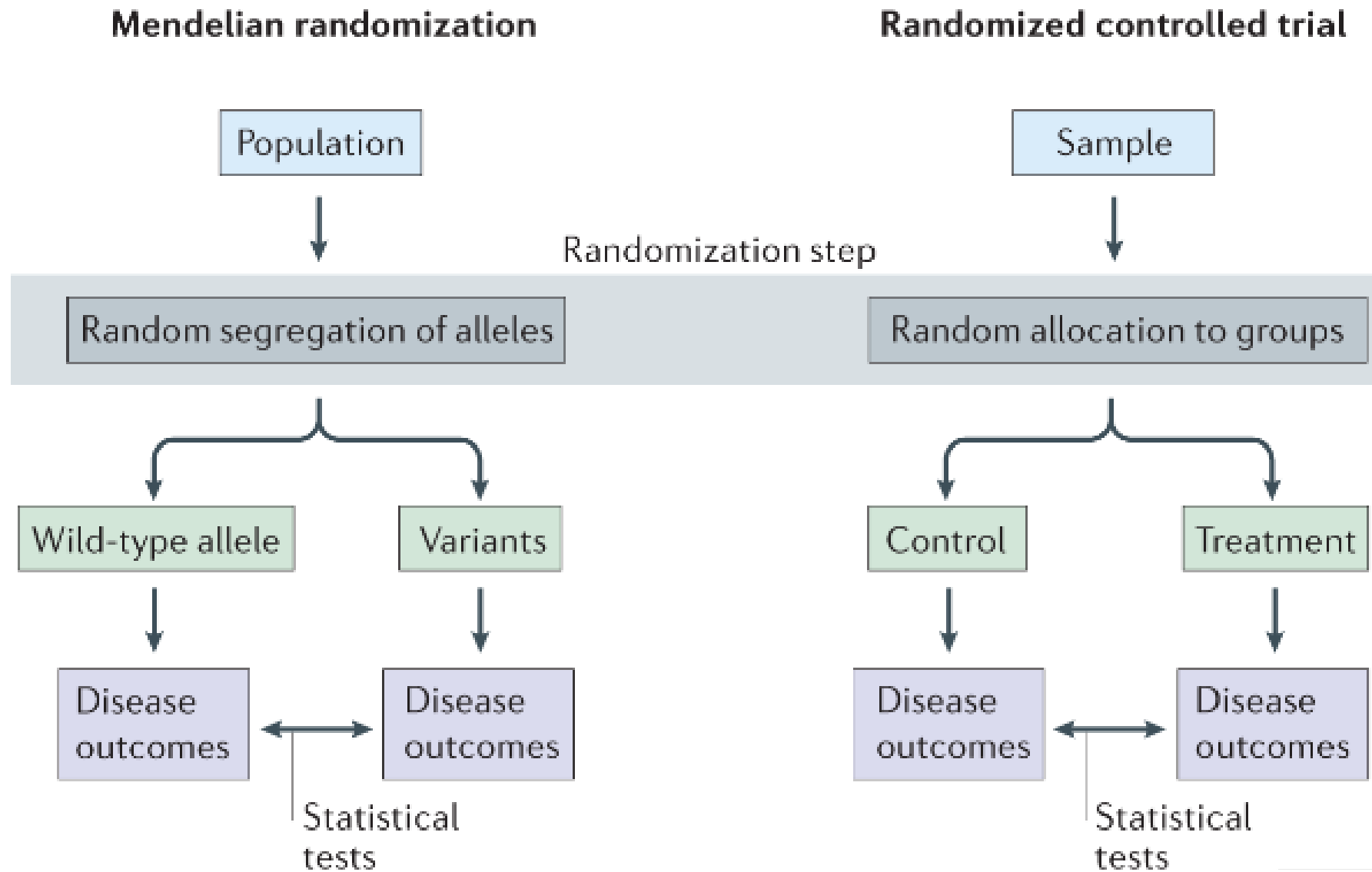
Exemple :

Obésité associée négativement au risque de cancer du poumon

- Causalité inverse ?
- Confusion dû à l'association entre consommation de tabac et poids ?

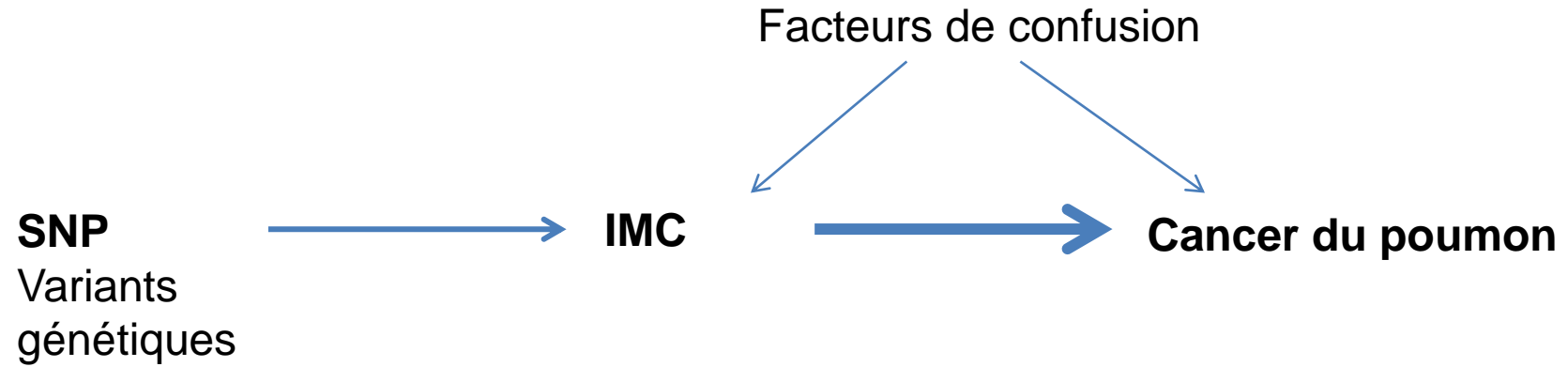


Randomisation mendélienne



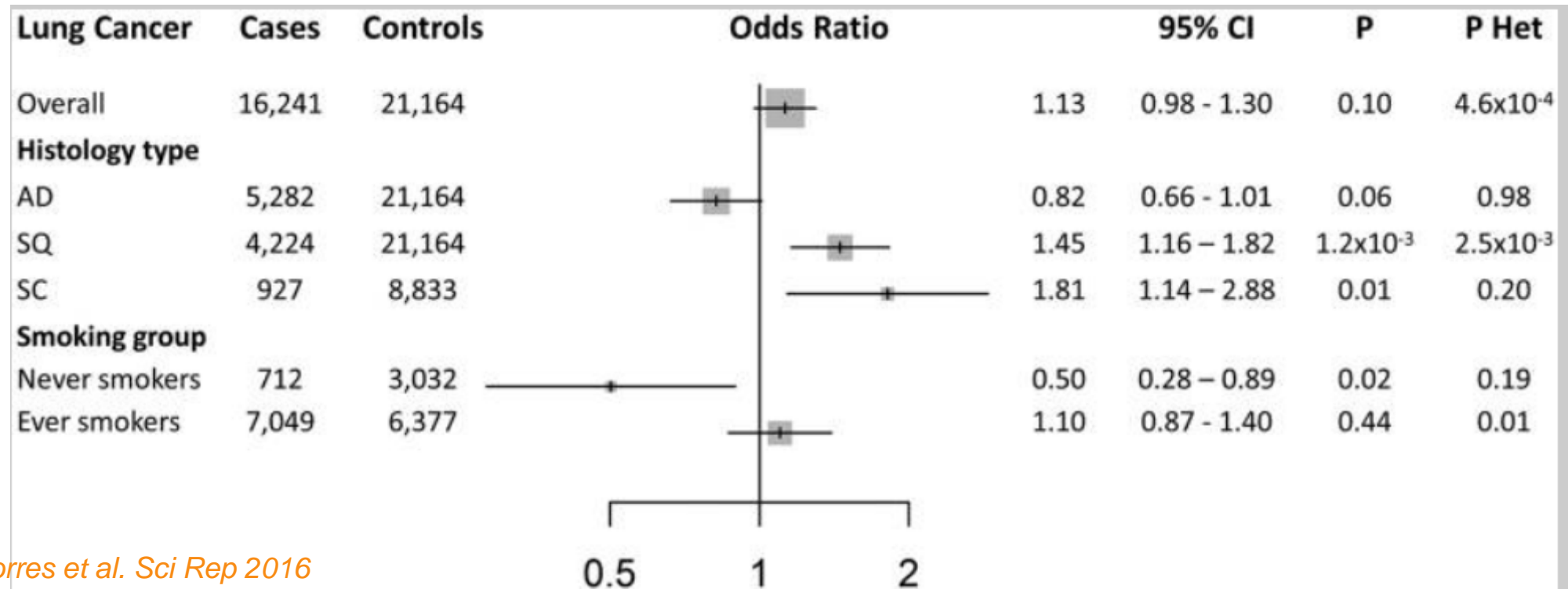
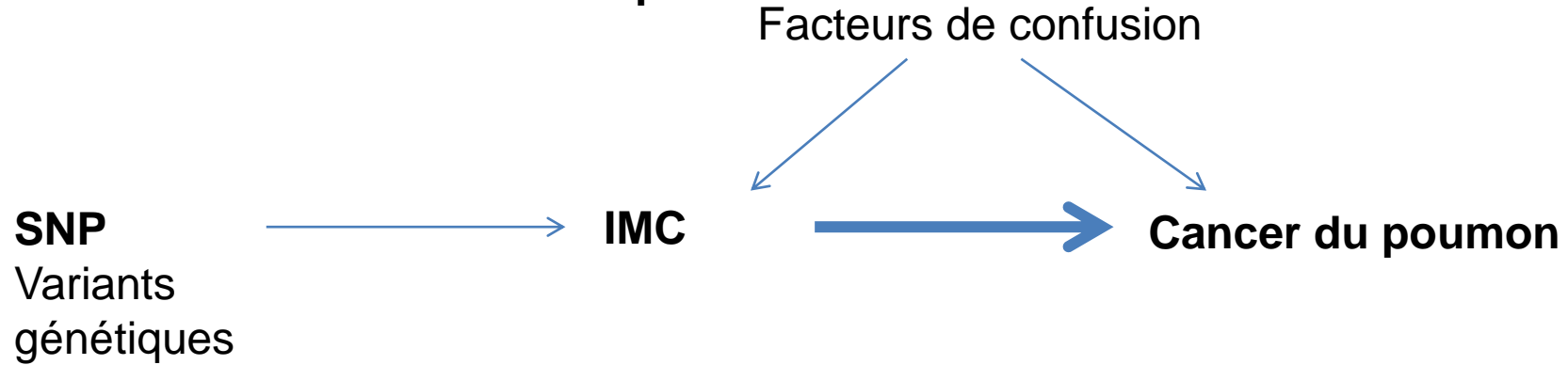
Randomisation mendélienne

Association causale entre IMC et cancer du poumon



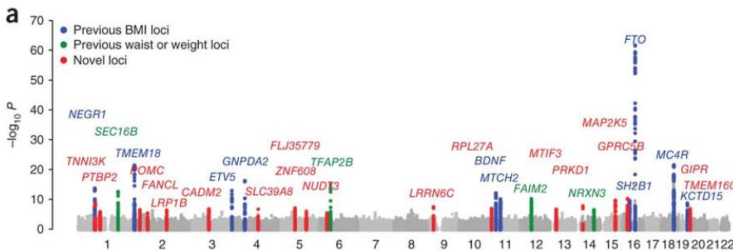
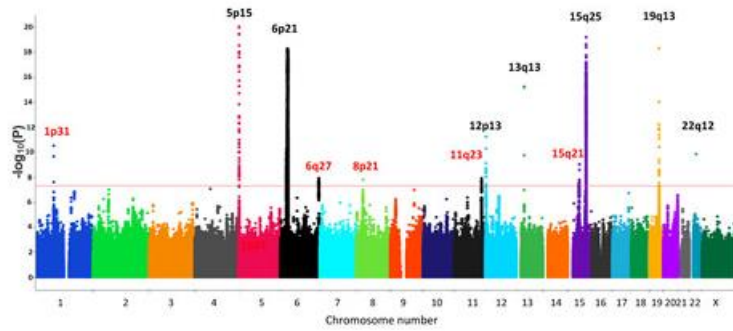
Randomisation mendélienne

Association causale entre IMC et cancer du poumon



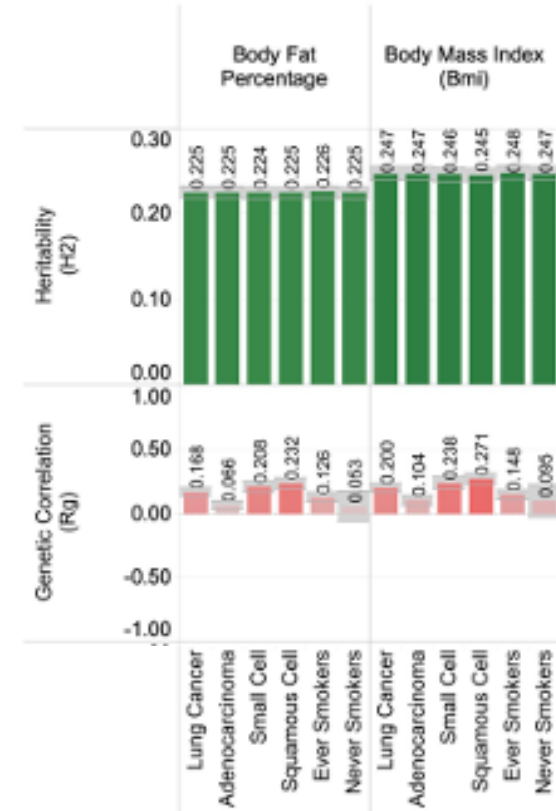
Pléiotropie

GWAS cancer du poumon



GWAS BMI

Corrélation génétique



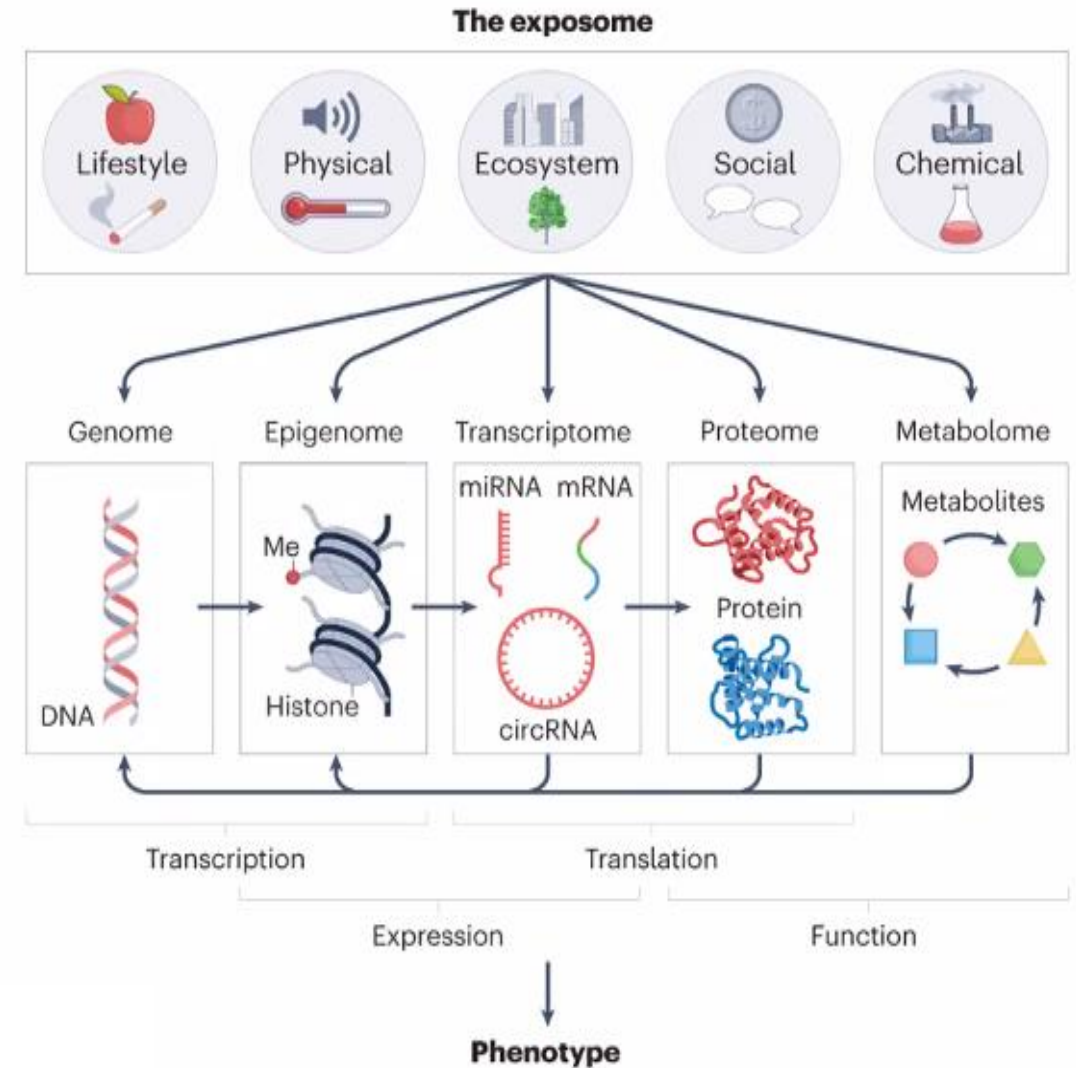
→ Analyse des variants ou gènes associés aux 2 phénotypes pour identifier des mécanismes biologiques communs

Pettit et al. The shared genetic architecture between epidemiological and behavioral traits with lung cancer. Scientific Report 2021

Opportunités

Constitution de bases de données épidémiologiques de plus en plus complexes

- Approche génétique en complément des approches épidémiologiques classiques
- Design familial pour les analyses d'association pangénomique
- Intégration multi-omiques pour mieux comprendre le lien entre exposome et maladies complexes



Challenges

- **Manque de diversité dans les études génétiques**
 - 80% des participants de GWAS sont Européens alors qu'ils ne représentent que 16% de la population (Duggal et al. AJE 2019)
- **Partage des données avec les collaborateurs tout en garantissant la protection des données ?**
 - Mise à jour des conventions de collaboration en conformité avec le RGPD dans les consortia internationaux
 - Données massives : gestion informatique ?
 - Quel référentiel de sécurité pour les études de cohorte en population générale (HDS, EDS, SNDS, ...) ?
- **Questionnement éthique**
 - ADN de l'individu versus ADN du clan?
 - Mise en conformité des anciennes études cas/témoins initiée avant le RGPD
 - Ré-information des personnes ?
 - Durée de conservation des échantillons vs délai pour la valorisation des données ?

Remerciements

Equipe Exposome et Hérité



