

# Les cohortes en France

Depuis 2009, des cohortes sont mises en place, sélectionnées et financées dans le cadre des Très Grandes Infrastructures de recherche (TGIR) et du Grand Emprunt. Elles sont destinées à améliorer la connaissance de certaines pathologies ou certains groupes de population.

## Les cohortes « historiques » en France

**Annick Alpérovitch**  
Inserm U708

Pour toute une génération d'épidémiologistes, celle du dernier tiers du xx<sup>e</sup> siècle, le mot « cohorte » est indissociable de l'étude de Framingham. « *In the epidemiological imagination, the Framingham Heart Study has attained iconic status, both as the prototype of the cohort study and as a result of its scientific success... In constructing their investigation, Framingham's initiators had to invent new approaches to epidemiological research.* »

Avec des outils très simples – pression artérielle au cabinet du médecin généraliste, cholestérol total, indice de masse corporelle, consommation de tabac –, l'étude de Framingham a défriché le champ de l'épidémiologie cardio-vasculaire. Les tout premiers articles issus de l'étude de Framingham ont donné des bases – toujours solides – à la prévention des maladies cardio-vasculaires, à un moment où, dans les pays développés, les maladies chroniques étaient reconnues comme un problème de santé publique prioritaire [30]. L'étude de Framingham est devenue rapidement un modèle que des équipes à travers le monde se sont approprié. Autour de Daniel Schwartz, au milieu des années 1960, la première étude de cohorte française – l'Étude prospective parisienne (EPP) – a été lancée [39]. Pendant plus de vingt ans, l'Étude prospective parisienne fut la seule grande cohorte française. Mais la décennie 1990-2000

a vu le lancement de grandes études de cohorte, encore actives pour la plupart – Gazel, E3N, Paquid, EVA, 3C –, qui vont être brièvement présentées.

### L'Étude prospective parisienne

Cette cohorte est constituée par 7 746 hommes âgés de 42 à 55 ans, nés en France métropolitaine, travaillant à la préfecture de police de Paris, inclus dans l'étude entre 1967 et 1972 à l'occasion d'un examen de médecine du travail, comprenant un bilan biologique standard. Les données de suivi ont été recueillies lors des visites médicales ultérieures, par questionnaire postal, ou par connaissance du décès. La mortalité de cette cohorte, en relation avec les paramètres initiaux, continue de faire l'objet d'analyses. Un article sur l'association entre le rythme cardiaque au repos et la mortalité par cancer a ainsi été publié en 2011 [29]. L'un des coauteurs de ce papier, Pierre Ducimetière, faisait partie du groupe qui, vingt-cinq ans auparavant, fut à l'origine de l'EPP. Cette étude a fourni les premières estimations de l'incidence des maladies vasculaires en France : cardiopathie ischémique, accident vasculaire cérébral, artérite oblitérante des membres inférieurs. Ce faisant, elle montra que le risque de maladie vasculaire était plus faible en France que dans d'autres

Les références entre crochets renvoient à la Bibliographie générale p. 51.



## Apport des cohortes à la connaissance de la santé

pays développés (Amérique du Nord, Grande-Bretagne et Europe du Nord). Ce résultat reste globalement vrai et a nourri trois décennies de débats sur l'origine de cette protection.

### Gazel

Gazel aussi est une cohorte recrutée en milieu professionnel. En 1989, à l'initiative de Marcel Goldberg qui travaillait alors dans l'unité Inserm 88 dirigée par François Grémy, 15 010 hommes âgés de 40 à 50 ans et 5 614 femmes âgées de 35 à 50 ans, appartenant au personnel d'Électricité-Gaz de France (d'où le nom de la cohorte) ont été volontaires pour aider la recherche en participant à une étude prospective de longue durée [47]. Gazel, à la différence de l'étude de Framingham et de l'EPP, est au départ une étude généraliste, ne ciblant aucune pathologie, ni aucun facteur de risque précis. Les informations sur les participants ont été principalement recueillies par questionnaire, le suivi étant surtout réalisé par questionnaire postal (des prélèvements biologiques ont été effectués sur environ 4 600 participants en 2000). Le projet s'est régulièrement enrichi de nouvelles thématiques au cours du suivi. Le fait que le personnel de cette entreprise bénéficie d'un système de couverture médicale particulier a permis à l'équipe de recherche de disposer, avec l'avis favorable de la Commission nationale informatique et liberté (Cnil), de données individuelles sur les événements de santé ayant entraîné un arrêt de travail. Si Gazel est à l'origine de nombreuses publications scientifiques, sa contribution est particulièrement significative dans le domaine des relations entre travail, environnement social et santé physique ou psychologique. Enfin, un atout unique de Gazel est de pouvoir étudier prospectivement l'évolution de la santé au moment du départ en retraite, plus de 90 % des hommes et 87 % des femmes étant retraités fin 2010. Gazel a ainsi pu montrer que, dans cette cohorte, le départ à la retraite s'accompagne d'une amélioration de la santé subjective globale et de différents troubles (sommeil, céphalées, fatigue et symptômes dépressifs). Gazel se voulait une plateforme ouverte à d'autres équipes et d'autres idées de recherche, et a atteint cet objectif : une trentaine de groupes de recherche, en France et hors de France, ont eu accès aux données de Gazel. Ces collaborations ont permis, entre autres, des études comparatives particulièrement intéressantes dans le domaine des inégalités sociales de santé.

### L'étude E3N

Lancée pratiquement en même temps que Gazel, l'étude E3N (pour Étude épidémiologique auprès des femmes de l'Éducation nationale) est une étude de très grande taille – la plus grande des cohortes françaises — en milieu professionnel. L'objectif principal d'E3N, projet de l'équipe de Françoise Clavel-Chapelon, était d'étudier les facteurs de risque (nutrition, contraceptifs et traitement hormonal substitutif) de cancer dans une cohorte de

100 000 femmes, âgées de 40 à 65 ans à l'inclusion [9]. La cohorte a été régulièrement suivie par questionnaire postal, permettant d'actualiser les expositions et d'enregistrer la survenue des événements de santé majeurs. Une banque d'ADN a été constituée au cours du suivi, du matériel génétique étant disponible pour environ 25 % de la cohorte E3N. Tout comme Gazel, du fait d'un système de couverture sociale particulier aux enseignants, E3N dispose de données précises sur la consommation de médicaments, qui ont permis des études très intéressantes sur le lien entre la consommation de traitements hormonaux et le risque de cancer. E3N est l'une des cohortes de l'European Prospective Investigation into Cancer and Nutrition, et est partie prenante dans la plupart des travaux de ce groupe.

### Paquid

L'étude Paquid, lancée en 1988 par l'équipe bordelaise de Jean-François Dartigues, fait date à double titre : par sa thématique et par sa méthodologie. Pour ce qui est de la thématique, Paquid est le premier programme de recherche épidémiologique sur le vieillissement, en particulier les troubles cognitifs liés à l'âge, répondant au besoin de connaissances solides pour affronter les problèmes de santé publique résultant de l'évolution démographique. Pour ce qui est de la méthodologie, Paquid est la première grande cohorte recrutée en population générale (tirage au sort sur les listes électorales de 75 communes de Gironde et Dordogne) et suivie par des examens répétés (en moyenne, tous les deux ans) des 3 777 participants âgés de 65 ans et plus. Paquid a fourni les toutes premières estimations de la prévalence et de l'incidence de la démence dans la population française, et les premières données sur les facteurs associés au risque de démence, au premier rang desquels l'éducation. Paquid a permis à la France d'être présente dans le premier programme européen de recherche épidémiologique sur la démence, EuroDem, qui a fourni des estimations de la prévalence des démences en fonction de l'âge qui restent pertinentes vingt ans après. Paquid incluait aussi un important volet sur les incapacités, développé par Pascale Barberger-Gateau. Du fait de l'âge et du nombre de participants, Paquid est une cohorte en voie d'extinction. Au dernier point (vingt-deux ans de suivi), seulement 498 participants étaient encore en vie ; 10 étaient centenaires. Une proportion importante des travaux les plus récents de Paquid porte sur le lien entre troubles cognitifs, incapacités et mortalité.


### Les études EVA et 3C

Une nouvelle étape dans la conception des cohortes a été franchie avec l'étude EVA (pour Épidémiologie du vieillissement artériel, mais qui portait aussi sur le déclin cognitif), qui peut être considérée comme un prototype pour l'étude 3C (Étude des 3 cités). En effet, EVA, coordonnée par Pierre Ducimetière et Annick Alperovitch, a été, en 1991, la première cohorte française en popu-

lation générale (1 400 personnes résidant à Nantes) à utiliser des investigations habituellement réservées aux études cliniques, telles que l'échographie carotidienne, le doppler transcrânien ou l'imagerie par résonance magnétique (IRM). EVA a permis de décrire l'évolution de la paroi carotidienne au cours du vieillissement, et de démontrer l'association entre hypertension artérielle, déclin cognitif et lésions vasculaires cérébrales silencieuses en IRM. L'étude 3C a bénéficié de l'expérience acquise dans EVA pour étudier, dans une cohorte de près de 10 000 personnes de 65 ans et plus recrutées à Bordeaux, Dijon et Montpellier, la relation entre facteurs vasculaires et risque de démence [42]. Mais alors qu'environ 1 200 échographies et 800 IRM avaient été effectuées dans EVA, on dispose dans 3C de plus de 6 000 échographies et de plusieurs milliers d'IRM. 3C est un projet de recherche collaboratif associant, à l'origine, 6 équipes de recherche fondatrices (voir <http://www.three-city-study.com> pour la liste des responsables de ces équipes) ayant des compétences, des moyens techniques et des domaines d'intérêt complémentaires. Bien que l'étude ait été conçue dans l'objectif de tester une hypothèse précise, la liste des publications montre la variabilité des thèmes qui ont été abordés à partir des données recueillies dans 3C. Après plus de dix ans de suivi des participants, 3C est devenue, au moins en partie, une cohorte généraliste, ouverte à de nombreuses équipes extérieures. Un atout récent de 3C est que, dans le cadre du volet recherche du Plan

Alzheimer, le génome de l'ensemble de la cohorte a pu être séquencé. Il en résulte des développements qui n'avaient pas été anticipés, l'étude 3C étant très sollicitée non seulement pour participer à des études GWA (Genome Wide Association) concernant des mesures disponibles dans 3C (mesures IRM ou carotidiennes en particulier), mais aussi pour fournir des groupes témoins dans des études GWA dans diverses pathologies. Un investissement très important est ainsi rentabilisé par des travaux scientifiques nombreux et variés.

### En conclusion

De ce bref tour d'horizon des cohortes « historiques », on peut retenir quelques points. Les études de cohorte sont un domaine où les équipes de recherche épidémiologique françaises ont été depuis cinquante ans, et sont toujours aujourd'hui, très compétitives. Cet investissement très lourd, tant par les financements, qu'il faut trouver, que par l'énergie demandée aux investigateurs, n'est scientifiquement productif que plusieurs années après la mise en route de l'étude, mais le reste pendant des années. Pour prendre l'exemple de 3C, seulement 10 articles ont été publiés au cours des cinq ans qui ont suivi le recrutement de la cohorte (1999-2000), alors que 46 ont paru dans la seule année 2011. Enfin, quel que soit le modèle de la cohorte (avec ou sans un objectif précis prédéfini), la richesse des données permet d'aborder de nombreux sujets, offrant à de nombreuses équipes un matériel de recherche unique. 

## Les grandes cohortes en santé 2008-2011

Une cohorte est la réunion d'un groupe de sujets partageant un certain nombre de caractéristiques, suivis longitudinalement à l'échelle individuelle, selon un protocole préétabli. À des fins de recherche en santé publique, elle repose sur la collecte d'informations concernant des caractéristiques et des expositions des sujets suivis, qui sont recueillies avant le moment d'apparition des effets analysés, susceptibles de modifier le risque de survenue de phénomènes de santé divers, en particulier de pathologies. Les informations collectées peuvent provenir à la fois de questionnaires sociodémographiques, comportementaux et cliniques, de mesures de caractéristiques physiques, biomédicales, psychométriques ou/et d'attitudes, d'exams cliniques et paracliniques, et bien sûr de prélèvements biologiques. La taille des cohortes s'échelonne habituellement de quelques milliers à plusieurs centaines de milliers de sujets inclus, et elles nécessitent la mobilisation d'infrastructures et de personnels nombreux, ainsi qu'un

grand professionnalisme dans le suivi logistique puis l'analyse. L'intérêt des cohortes est important pour la surveillance, la connaissance, le suivi et l'évaluation des actions dans le domaine de la santé.

Une Très Grande Infrastructure de recherche (TGIR) est un outil établi en vue de mener une recherche propre d'importance et pouvant assurer une mission de service pour une ou plusieurs communautés scientifiques de grande taille (cf. feuille de route Predecob<sup>1</sup>, <http://www.roadmaptgi.fr>). Son coût de construction et d'exploitation est tel qu'il justifie un processus de décision et de financement concerté au niveau national, et éventuellement européen ou international, ainsi qu'une programmation pluriannuelle.

**1.** Plate-forme de recherche et développement des grandes cohortes biomédicales : le projet Predecob a pour objectif l'organisation de la coordination des grandes études de cohorte (groupe d'une population suivie chronologiquement) biomédicales.

### Alfred Spira

Université Paris Sud/  
Institut de recherche  
en santé publique  
(IReSP)