

PARIS, le 6 juillet 2007 - N° 155 /D130

## Interformation STATISTIQUES SOCIALES

---

Réunion du 6 juin 2007

COMPTE RENDU DE LA RÉUNION .....	2
PROJET D'AVIS .....	27
LISTE DES DOCUMENTS PRÉPARATOIRES .....	28

---

COMPTE RENDU DE LA RÉUNION DE  
L'INTERFORMATION STATISTIQUES SOCIALES - 6 juin 2007 -

---

Président : Mme Claire BAZY MALAURIE, Rapporteur général et Président de Chambre à la Cour des Comptes, Présidente de la formation « Éducation, formation » du Cnis

Rapporteur : M. Stéfan LOLLIVIER, Directeur des statistiques démographiques et sociales (Insee), Rapporteur de la formation « Démographie, conditions de vie »

Responsable de la formation : Marie-Christine CHAMBET-ROSSET (01 41 17 52 65)

*RAPPEL DE L'ORDRE DU JOUR*

Introduction .....	4
I. Présentation du projet d'étude longitudinale française depuis l'enfance – Elfe.....	5
II. Les projets longitudinaux en Grande-Bretagne : une expérience de soixante ans.....	9
III. Place de l'épidémiologie dans le projet ELFE .....	10
IV. Débat .....	12
V. Dimensions déontologiques et éthiques du projet Elfe .....	18
VI. Débat .....	21
VII. Projet d'avis .....	26

## LISTE DES PARTICIPANTS

Jean-Paul AMANN	COLLÈGE DE FRANCE
Sophie BAUDET-MICHEL	UNIVERSITÉ PARIS DIDEROT
Claire BAZY MALAURIE	COUR DES COMPTES
Sébastien BERTHE	CNIDFF
Pierre BOISSELOT	DREES MIN SANTÉ SOLIDARITÉ
Jeanne BOSSI	CNIL
Jean-Claude BOURQUIN	UFC QUE CHOISIR
Caroline BRAY	OSII
Chantal CASES	IRDES
Marie-Christine CHAMBET-ROSSET	CNIS
Philippe CHOFFEL	DÉLÉGATION INTERMINISTÉRIELLE VILLE
Olivier COHEN	HC FORUM
Christiane CRÉPIN	CNAF
Philippe CUNEO	Secrétaire général du CNIS
Jean-Philippe DAMAIS	
Jacques DESFONTAINE	UNION FRANÇAISE SANTÉ BUCCO DENTAIRE
Béatrice DUCOT	INSERM
Denis DURAND	CGT BANQUE DE FRANCE
Françoise DUSSERT	Secrétaire générale adjointe du CNIS
Elsa FERRI	UNIVERSITÉ DE LONDRES
Mélanie FLEURANT	IAAT
Bernard GUIBERT	INSEE
Jean-Baptiste HERBET	DREES MIN SANTÉ SOLIDARITÉ
Evelyne HOUDOIN	DREES MIN SANTÉ SOLIDARITÉ
Aude JOANNON	LJCOM
Stéphane JUGNOT	INSEE
Jean-Claude LAMBERT	INSERM
Jacques LAVERTU	INSEE
Jean-Pierre LE GLÉAU	INSEE
Sylvie LEDOUX	INSERM
Nadine LEGENDRE	DREES MIN SANTE SOLIDARITE
Henri LERIDON	INED
Frédéque LESAULNIER	CNIL
Stéfan LOLLIVIER	INSEE
Christian LUBINEAU	MIN TRANSPORTS ÉQUIPEMENT
Efi MARKOU	INED
Philippe MONIER	INSEE
Alain MOREAU	SGAR AQUITAINE
Daniel NAULLEAU	UNIVERSITÉ PARIS 6
Françoise OEUVRARD	DEPP MIN ÉDUCATION ENSGT SUP RECHERCHE
Alain PELC	CNSA
Claudine PIRUS	INED
Ruxandra POPA	INED
Laurent RAPHAEL	INED
Benoît RIANDEY	INED
Benoît ROBIN	CGT FO
Fabienne ROSENWALD	DEPP MIN ÉDUCATION ENSGT SUP RECHERCHE
Georges SALINES	INSTITUT NAL DE VEILLE SANITAIRE

Bruno SUCHAUT  
 Bernard SUJOBERT  
 Jean-Paul TEGLAS  
 Gilles TROUessin  
 Stéphanie VANDENTORREN  
 Matie ZINS

UNIVERSITÉ DE BOURGOGNE CNRS  
 INSEE  
 INSERM  
 OPPIDA SUD  
 INSTITUT NATIONAL DE VEILLE SANITAIRE  
 INSERM

*Absents excusés :*

Isabelle CHAPTAL  
 Marcel GOLDBERG  
 Michel ISNARD  
 Catherine QUANTIN  
 Robert PIERRON  
 Nicolas PROKOVAS

GROUPE DES INDUSTRIES MÉTALLURGIQUES  
 INSERM  
 INSEE  
 CHRU DIJON  
 CONSEIL RÉG AQUITAINE  
 CGT ANPE

*La séance est ouverte à 14h35 sous la présidence de Claire BAZY MALAURIE.*

## Introduction

### **Claire BAZY MALAURIE, Présidente**

Je suis Rapporteur général et Président de Chambre à la Cour des Comptes, et par ailleurs Présidente de la formation « Éducation, formation » du Cnis. C'est la première fois que je préside l'interformation « Statistiques sociales », qui sera aujourd'hui un peu particulière. Le seul sujet à l'ordre du jour est le projet Elfe, Étude longitudinale française dès l'enfance. Nous nous réunissons à la demande du Conseil national de l'information statistique.

Le projet Elfe a déjà été présenté à l'interformation le 6 octobre 2006, lors d'une séance présidée par Jean-Baptiste de Foucauld. Cette première présentation et les premières enquêtes ont recueilli un avis d'opportunité favorable. Sur le principe, Elfe répond à une demande importante que le Cnis a pensé devoir relayer, qui a comme objectif la connaissance plus précise des trajectoires à multiples dimensions des enfants. Ce type d'études longitudinales a déjà été lancé dans de nombreux autres pays, et nous évoquerons tout à l'heure l'expérience britannique avec Elsa Ferri.

Le projet Elfe, compte tenu de sa nouveauté et de sa complexité, a été évoqué à l'Assemblée plénière du Cnis en décembre 2006. Le Conseil a demandé à être associé aux étapes successives de son élaboration, en particulier en termes de garantie de la confidentialité des données et d'accès aux fichiers pour la recherche. Il a également souhaité l'organisation de cette réunion sur les enjeux sociétaux de l'opération et les conditions de sa mise en œuvre.

Elfe est un projet très ambitieux, qui associe et intéresse des partenaires d'horizons très différents, organisés dans une structure de groupement d'intérêt scientifique présidé par Henri Leridon. Il implique la collecte de très nombreuses informations sur les enfants et les parents, allant de la santé, avec les questions épidémiologiques, à l'éducation, la socialisation, les trajectoires sociales des parents ou l'environnement. Cette collecte sera effectuée par le biais d'enquêtes statistiques couvertes par la loi de 1951 et d'enquêtes épidémiologiques répondant à d'autres règles, Monsieur Salines nous en reparlera. Ces informations seront croisées avec des données issues des sources statistiques et d'autres enquêtes.

Le projet Elfe permettra de suivre des enfants depuis leur naissance jusqu'à l'âge adulte, dans des

composantes très personnelles et sensibles appelant des précautions particulières. L'intérêt scientifique de ce projet est évident, mais nous avons une obligation de respect de principes éthiques. Les questions relatives à ces principes doivent être soulevées pour pouvoir être traitées.

Les deux maîtres mots de cette opération seront la rigueur, scientifique bien sûr mais aussi concernant la protection des données, et la transparence. Les étapes de l'évolution du projet doivent faire l'objet d'informations qui ne prêtent pas à confusion dans l'esprit des partenaires, des producteurs de données, et des utilisateurs. Nous essaierons aujourd'hui de recueillir toutes les questions qu'un tel projet peut susciter, et nous espérons y apporter quelques réponses grâce aux interventions de personnes connaissant bien le dossier. Nous avons aujourd'hui besoin d'une interaction sereine et franche pour permettre au projet Elfe de progresser.

La présentation de ce projet par Monsieur Leridon sera suivie de l'intervention d'Elsa Ferri, *Senior Research Fellow* au Centre pour les études longitudinales à l'Institut d'éducation de l'Université de Londres. Ce centre a eu en charge plusieurs grandes cohortes britanniques, notamment des cohortes d'enfants. Madame Ferri a travaillé sur deux études : *Parenting in the 90s* (La fonction parentale dans les années 90), et *Changes in Lone Parenthood* (Évolutions de la monoparentalité) en 2003. Georges Salines, Directeur du département Santé-environnement de l'Institut national de veille sanitaire, évoquera ensuite les aspects épidémiologiques du programme. Nous nous arrêterons pour un débat avant qu'Henri Leridon n'aborde la protection des données et des personnes dans le projet Elfe. Enfin nous terminerons par un second temps d'échange.

Je voudrais remercier Jean-Paul Amann de sa présence aujourd'hui. Il est Assistant au Collège de France, et appartient au groupe de travail Éthique et philosophie des sciences dans le cadre de la chaire d'Anne Fagot-Largeault, qui s'intéresse tout particulièrement aux recherches médicales impliquant des enfants. Monsieur Amann interviendra en tant que de besoin, comme il l'a souhaité, au cours de cette réunion. Je remercie également Elsa Ferri de s'être jointe à nous.

## **I. Présentation du projet d'étude longitudinale française depuis l'enfance – Elfe**

### **Henri LERIDON, Directeur du GIS Elfe**

Je me réjouis de cette rencontre, souhaitée en octobre dernier. Le projet Elfe est atypique à plusieurs titres, et porte une grande ambition. Il doit répondre à des critères de qualité, de rigueur, et de transparence. Je suis convaincu que, même si nous en avons eu l'idée, monter une telle opération dans la clandestinité serait impossible. L'information sur ce projet doit donc être la plus claire possible, y compris en vue de l'obtention du consentement des familles.

Les partenaires engagés dans le groupement d'intérêt scientifique Elfe sont des organismes de recherche et d'études comme l'Ined, l'Inserm et l'Insee, une agence de santé, l'InVS, trois directions dépendant de divers ministères - la Drees, la Depp et la DGS- et la Caisse d'Allocations Familiales.

### **1. Origines du projet Elfe**

Le projet Elfe a été précédé de demandes assez explicites d'institutions françaises et de nombreuses expériences à l'étranger qui ont montré l'intérêt de ce type de démarche.

Des préoccupations ont été exprimées par la Drees, le Commissariat au Plan et le CERC, insistant sur le fait que pour comprendre certaines situations et les mécanismes qui les sous-tendaient en matière de pauvreté et de transferts sociaux, et en pensant particulièrement aux enfants, il fallait s'appuyer sur des études longitudinales qui n'existaient pas encore. A l'Ined, on s'intéressait également à ces questions, en particulier par le biais de la transformation des structures familiales, très importante dans les dernières décennies, et qui a rendu caducs un certain nombre de modes d'observation classiques et a conduit à envisager des observations de suivi. Enfin, le rapport de Stéfan Lollivier et Mylène Chaleix (Insee) de 2004 effectué dans le cadre du Cnis préconisait de s'orienter davantage vers des observations de panels dans divers domaines.

Du côté de la santé, un rapport du Haut Comité de la santé publique en 1998 évoquait déjà des études longitudinales dès la petite enfance, visant à dépister les troubles essentiels, les accidents, les violences, analyser les comportements à risques, et évaluer l'impact des interventions entreprises. Le Haut Comité insistait dès cette époque sur l'importance de recherches pluridisciplinaires associant sciences sociales, cliniques et épidémiologiques si l'on voulait être efficace. La loi de Santé publique de 2004 fait quant à elle des recommandations visant à combler des lacunes dans un certain nombre de domaines, et qui peuvent également passer par des études longitudinales.

Enfin, la France a adopté en 2004 un plan national Santé et environnement qui propose parmi une dizaine d'actions prioritaires la réalisation d'une étude épidémiologique « enfants », en liaison avec l'étude américaine *National Children's Study*, afin de :

- suivre la croissance et l'état de santé des enfants - nous manquons d'informations sur la continuité du développement de l'enfant ;
- mesurer les expositions aux principaux polluants - c'est extrêmement important car nous avons peu d'informations sur le sujet mais de nombreuses raisons de soupçonner que ces expositions sont encore plus graves pour les enfants que pour les adultes ;
- analyser les relations entre ces expositions et l'état de santé.

Ce que nous envisageons aujourd'hui existe déjà dans un certain nombre de pays. La Grande-Bretagne et les États-Unis ont déjà organisé des cohortes d'enfants depuis quelques décennies, en commençant à la naissance, parfois avant, parfois après, et le plus souvent sur une base nationale, mais parfois régionale. Les thématiques sont regroupées autour de l'idée du développement de l'enfant, avec un très large éventail de dimensions. Il s'agit de reconstituer les histoires entremêlées du contexte familial, économique, éducatif, ou de la socialisation de l'enfant. On a parfois des éléments biographiques rétrospectifs, mais on a rarement la possibilité d'étudier rétrospectivement les interactions entre ces différentes composantes. La dimension santé est aussi prise en compte. Elle interfère avec plusieurs aspects du développement de l'enfant, y compris en termes d'éducation ou de ressources. Cette dimension ouvre la porte à de nombreux modes d'observation. On n'utilise pas systématiquement des examens cliniques ou des recueils biologiques. Des tests peuvent être effectués par des professionnels ou par les parents. Si l'étude commence pendant la grossesse ou à la naissance, des études de périnatalité sont menées.

En Grande Bretagne, la première opération de suivi longitudinal d'enfants a commencé en 1946 et est toujours en cours. La moitié des personnes suivies au commencement le sont encore aujourd'hui, soixante ans après. De nouveaux échantillons ont été constitués en 1958, 1970, et 2000 (date du lancement de la *Millennium Cohort Study*). Ces échantillons comptent de 15 000 à 20 000 personnes. Je mentionnerai également l'étude Alspac, grande enquête de référence pour les épidémiologistes. Le mode de constitution, de fonctionnement et les problématiques de cette étude régionale ressemblent à ce qui nous intéresse aujourd'hui.

Au Canada, l'étude nationale Elnej est partie d'enfants de 0 à 11 ans, pour les suivre jusqu'à 20 ans, et au Québec, Eldeq a étudié un échantillon de femmes enceintes jusqu'aux huit ans de leur enfant.

Aux États-Unis, un vaste projet est en cours depuis 2002, la *National Children's Study*. Il permettrait de suivre 100 000 enfants, en mettant l'accent sur la santé et l'environnement, jusqu'à l'âge adulte. Les crédits ont été momentanément coupés par l'administration Bush, puis rétablis par le Congrès. L'enquête devrait donc démarrer cette année.

L'outil cohorte est de plus en plus utilisé en France, en particulier dans le domaine de la santé. Un recensement réalisé par l'Inserm a pointé une soixantaine de cohortes existantes, et autant de projets pour les années à venir. Mais seulement cinq ou six d'entre elles, essentiellement orientées vers des questions épidémiologiques, concernent des enfants et elles ne sont pas toujours représentatives au plan national ou local. Par ailleurs, nous n'avons pas, en France, d'approche interdisciplinaire telle que celle que le projet Elfe ambitionne d'avoir.

Dans certains domaines, la cohorte paraît indispensable. Un groupe de travail de la Fondation européenne pour la science s'est intéressé aux variations de l'espérance de vie en santé en Europe, en relation avec les inégalités sociales. Le constat montre que les inégalités sociales de santé existent et tendent à s'aggraver dans la majorité des pays d'Europe. Il est logique de s'en préoccuper. Et pour

comprendre ce qui se passe, nous avons besoin de données longitudinales, parce que ces inégalités résultent d'un cumul d'expositions au fil du temps. Une enquête transversale, même un peu rétrospective, ne permettrait pas d'appréhender convenablement ces petits épisodes ou les diverses dimensions d'exposition accumulés au cours de la vie, et notamment au cours de l'enfance. Ce groupe de travail a également insisté sur l'interaction avec les données économiques et sociales. Pour espérer améliorer l'interprétation causale des divers événements susceptibles d'intervenir, les études longitudinales fournissent au moins une donnée temporelle et chronologique. Il est beaucoup plus difficile de démêler un ensemble de facteurs transversaux et de comprendre lequel peut être la cause de l'autre.

La démographie a par ailleurs vu se développer l'approche « *Life Course* » (« l'histoire de vie » ou « analyse biographique »). De nombreux textes montrent l'intérêt de la démarche et développent des outils concrets pour la rendre efficace.

## 2. Le projet Elfe

Notre projet a démarré en 2006, associant les partenaires déjà cités, auxquels d'autres pourront se joindre. Plusieurs institutions sont déjà demandeuses. Nous avons d'abord constitué une équipe pluridisciplinaire, puis lancé un appel à proposition très large auprès des équipes de recherche pour définir des thèmes possibles, et les construire ensuite au sein du groupe de projet.

Le thème général est celui du développement de l'enfant :

- en interaction avec la vie familiale, sociale et les éléments économiques, l'environnement géographique du cadre de vie de la famille ;
- d'un point de vue intellectuel et psychomoteur, son parcours éducatif et plus tard, peut-être, son insertion professionnelle ;
- du point de vue de la santé, traitant au départ de la périnatalité, puis des événements pouvant affecter l'enfant, en relation en particulier avec les expositions environnementales.

Ces questions peuvent intéresser une très grande diversité de spécialistes dans différents domaines.

Le contenu scientifique du projet Elfe n'était pas préétabli. Ceci fait partie des spécificités et des difficultés de ce projet. Nous avons défini un cadre général, en demandant aux équipes de proposer des projets pour lesquels le suivi longitudinal est essentiel et qui ont un aspect pluridisciplinaire, car on ambitionne de collecter de l'information qui doit pouvoir être partagée. Une soixantaine d'équipes ont répondu, autant en sciences sociales qu'en santé, représentant au total plus de quatre-vingt projets. Nous avons constitué une petite quinzaine de groupes thématiques qui regroupent les équipes autour de leur thématique privilégiée. Certaines équipes participent à plusieurs groupes, car elles sont par nature déjà transversales. Dans chaque groupe, un coordonnateur est chargé de rapporter au groupe de projet - qui réunit ces coordonnateurs et l'équipe permanente - les propositions de questions et de modalités d'observation, tout ce que les équipes souhaitent faire après de premiers arbitrages entre elles. L'équipe de projet, en liaison à la fois avec le groupe et le Conseil scientifique, essaie de définir les étapes, les modalités d'observation concrètes et les questionnaires pour aboutir à un ensemble cohérent et faire en sorte que l'on ne sollicite pas les familles. Nous nous sommes fixés des formats de questionnaires conformes aux normes les plus classiques, et le rythme d'observation ne doit pas être une charge pour les familles, nous y reviendrons.

Nous souhaitons par ailleurs que le projet Elfe serve à d'autres enquêtes, qui auraient été menées de toute façon. C'est un souci d'économie générale, mais c'est aussi l'occasion d'augmenter l'efficacité de ces interventions ponctuelles. Par exemple, la première étape d'Elfe sera la prochaine enquête de périnatalité française, commanditée comme les précédentes par la Direction générale de la santé, la Drees et réalisée avec une équipe de recherche de l'Inserm. Cette enquête a déjà été réalisée trois ou quatre fois et porte sur l'ensemble des naissances observées en France pendant une période donnée.

Pourquoi est-ce possible et pourquoi avons-nous convergé assez vite sur cette proposition ? Nous cherchions à étudier un échantillon représentatif de l'ensemble des naissances d'une année. L'Insee nous a proposé de le prendre à l'intérieur du nouvel échantillon démographique permanent (EDP) de l'Insee, et de suivre des enfants nés pendant les seize jours de l'EDP, mais uniquement en 2009. Nous pourrions donc enrichir nos données avec les données classiques de l'EDP, le recensement essentiellement. Nos

enfants seront nés à certaines dates précises, ce qui permettra de les retrouver plus facilement par ailleurs.

Nous voudrions également être en articulation avec des enquêtes ayant lieu en milieu scolaire : les échantillons de ces enquêtes seraient les enfants Elfe, ou les enfants éligibles pour Elfe, et non plus des échantillons indépendants. Pour chacune de ces opérations, l'intérêt longitudinal sera évident. L'enquête périnatalité s'arrête à la naissance. Il sera très intéressant pour les chercheurs ayant mené cette enquête de pouvoir suivre l'évolution de la santé de ces enfants, par exemple au regard d'une hypotrophie ou d'une prématurité observée à la naissance. Lorsque les enfants seront dans le système scolaire, nous voudrions que notre cohorte devienne un pseudo panel pour l'Éducation nationale. Cela nous permettrait de suivre les parcours scolaires sans avoir à interroger les familles. Enfin, nous pensons rapprocher les données de diverses sources externes, comme celles des CAF ou de l'Assurance maladie.

Au total, sur la durée de l'observation que nous espérons mener jusqu'à l'âge adulte, nous aurons :

- des relevés dans les dossiers de maternité ;
- des questionnaires aux parents, et plus tard aux enfants, à partir de 14 ans environ, qui seront effectués en face à face, comme les enquêtes de l'Insee, ou par téléphone ;
- des auto-questionnaires par exemple sur les comportements alimentaires ;
- des tests psychologiques et psychomoteurs dans des versions simplifiées pouvant être passées par des non-professionnels ;
- des données provenant d'autres sources (du système scolaire par exemple) ;
- du point de vue médical, peut-être des examens de santé scolaire à six et onze ans ; des données de la Sécurité sociale, des registres de maladie (pour les enfants souffrant de cancers par exemple, au sujet desquels nous monterons un projet à dimension internationale) ;
- des variables de localisation, en particulier pour l'environnement ; nous géo-localiserons les logements pour mettre en relation avec des facteurs d'exposition, par exemple avec des capteurs de pollution atmosphérique ou des tests sur la qualité de l'eau. Certaines mesures d'exposition seront peut-être effectuées au domicile des enfants, comme la mesure très simple d'exposition au Radon, encore peu connue en France, à l'aide d'un petit capteur laissé quelques mois dans le logement ;
- à la naissance, des prélèvements biologiques à la maternité : le sang du cordon (prélèvement non invasif) et de l'urine et des cheveux de la mère, avec son accord. Il n'est pas prévu de prélèvements ultérieurement.

Nous partirons donc des naissances extraites de l'EDP en 2009 et commencerons avec l'enquête périnatale en maternité, c'est-à-dire un entretien de vingt minutes avec la mère. Nous lui demanderons alors son accord pour une interview à domicile six à huit semaines plus tard. Nous aurons ensuite des interviews vers trois ans et six ans. Dans l'intervalle, nous utiliserons probablement des interviews téléphoniques plus légères, et les examens en milieu scolaire à 6 et 11 ans (et peut-être 3 ans).

La structure globale du projet Elfe est aujourd'hui en place. La convention de GIS est signée, l'équipe de projet est en place. Les groupes de travail thématiques fonctionnent, et nous encourageons les équipes à un travail transversal en créant des groupes d'interface entre les spécialités. Les thématiques constituées sont :

- Démographie famille ;
- Économie précarité ;
- Socialisation éducation, que nous séparerons peut-être lorsque les enfants seront scolarisés ;
- Alimentation nutrition - cette structure est par nature très transversale ; l'obésité, le surpoids deviennent un problème important en France, y compris chez les enfants, et pour comprendre ce phénomène nous avons besoin d'observations de natures diverses. La mauvaise alimentation est liée aux pratiques sociales et aux ressources familiales. Mais il peut exister des facteurs biologiques ou génétiques de prédisposition, aggravant ou amortissant les effets des comportements. Nous avons de très bons spécialistes dans tous ces domaines, et Elfe devrait nous permettre de mieux comprendre les facteurs d'obésité et de surpoids, qui ont eux-mêmes des conséquences sociales ;
- Développement psychomoteur, en axant les recherches sur la motricité ou le langage. Nous n'irons pas très loin dans le domaine de la santé mentale, en partie hors de portée de notre projet tel qu'il est construit ;
- Périnatalité ;

- Recours au soin, question primordiale en termes d'inégalités sociales ;
- Accidents et traumatismes, un facteur de risque très important pour les enfants ;
- Contamination des milieux, en particulier l'air et l'eau ;
- Expositions chimiques. Des listes sont établies au niveau européen sur les risques liés aux produits chimiques. 100 000 sont classés à risque je crois, et 30 000 dans la « liste courte ». Ce sont des produits que nous utilisons dans la vie quotidienne. Les enfants sont particulièrement exposés. Nous voudrions déterminer quelles mesures sont envisageables à l'encontre de certains de ces produits, et formuler des hypothèses quant à leurs effets éventuels sur la santé ;
- Expositions physiques : le Radon ou la radiologie peuvent être mauvais pour les enfants, et nous manquons d'informations sur la dose de radiations que peut subir un enfant au cours de sa jeunesse ;
- Asthme et allergies, qui progressent beaucoup ces dernières années ;
- Cancer, pathologie plus rare chez les enfants, au sujet duquel nous nous associerons sans doute à un consortium international qui rapprochera les données de plusieurs cohortes.

Nous avons par ailleurs constitué des groupes transversaux, plus techniques, de méthodologie de la collecte, observation en maternité, informatique et protection des données, éthique et communication.

Notre Conseil scientifique compte une quinzaine de membres représentant une grande diversité de compétences. Il est présidé par la généticienne Anne Cambon-Thomsen, spécialisée en éthique ; Chantal Cases en est la vice-présidente. Cinq de ses membres sont des personnalités étrangères très compétentes dans le domaine des cohortes en Angleterre, au Danemark, en Suisse et en Espagne.

En matière d'éthique, nous mettons en place une structure, également animée par Anne Cambon-Thomsen, avec des spécialistes de l'éthique qui n'ont aucun rapport avec notre projet, Jean-Paul Amann, Nicolas Lechopier de la Faculté de Médecine de Lyon, Joëlle Le Marec, professeur à l'École normale supérieure de Lettres de Lyon, et Emmanuelle Rial-Sebbag, de la Faculté de Médecine et dans une unité Inserm de Toulouse. En France, nous avons une bonne expérience de l'éthique dans le domaine de la santé, mais nous allons construire dans le domaine des sciences sociales.

## **II. Les projets longitudinaux en Grande-Bretagne : une expérience de soixante ans**

**Elsa FERRI, Former Senior Research Fellow au Centre for Longitudinal Studies, University of London**

Je voudrais dire d'abord combien je suis heureuse de votre invitation. Je vais vous présenter un bref historique des études longitudinales menées en Grande-Bretagne.

La première remonte à plus de soixante ans. Trois des quatre études nationales en cours ont suivi la vie et le développement de tous les individus nés en Grande-Bretagne pendant une semaine respectivement des années 1946, 1958 et 1970. Chaque échantillon comptait environ 18 000 enfants. Ces études sont toutes pluridisciplinaires, et s'adressent à tous les domaines majeurs de la vie de l'enfant puis de l'adulte : la santé physique et mentale, l'éducation, la famille, l'emploi et les revenus, le logement, la communauté et l'environnement, et les attitudes sociales. De plus un tiers des enfants des personnes appartenant à la cohorte de 1958 ont aussi été étudiés.

Nous avons donc des données sur trois générations, ce qui nous a permis d'analyser les continuités et les discontinuités des caractéristiques et des circonstances, et de traiter de la question vitale de la transmission du désavantage. Le pourcentage de participants a toujours été satisfaisant au cours des années, à un taux moyen de 70 %. Les échantillons sont restés suffisamment représentatifs. L'existence de trois cohortes comptant plus de 30 000 sujets et séparées par 12 ans permet aussi l'investigation des effets respectifs de l'âge et du contexte social et économique dans une période déterminée. Par exemple, quelles sont les similarités ou les différences entre les individus ayant atteint leurs trente ans en l'an 2000 ou en 1978 ? Nous menons aussi des études spéciales, plus qualitatives que quantitatives, sur des sous-populations des cohortes, comme les familles monoparentales ou recomposées, les individus avec problème spécifique physique ou mental, les désavantagés, etc.

Les enquêtes générales ne permettent jamais d'inclure des questions suffisamment détaillées pour éclairer les circonstances et les expériences des groupes spécifiques. On doit donc planifier des études spéciales soit dès le début, soit rechercher des financements quand émergent des résultats intéressants ou, comme souvent, imprévus.

L'étude longitudinale britannique la plus récente, la 4<sup>ème</sup>, est la *Millenium Cohort Study* ou MCS. Elle compte plus de 18 000 enfants nés en 2000 et 2001, sélectionnés dans des circonscriptions électorales stratifiées pour que les groupes désavantagés et les minorités ethniques soient surreprésentés. Nous savions des cohortes précédentes que le taux d'attrition de ces groupes était supérieur à la normale. Nous voulions donc, grâce à un échantillon plus important, retenir un nombre suffisant d'entre eux pour pouvoir les étudier dans le futur. La MCS, elle aussi pluridisciplinaire, a déjà conduit trois enquêtes à neuf mois, trois ans et cinq ans. C'est l'étude la plus comparable au projet français actuel.

L'intérêt scientifique ou politique des études longitudinales n'est plus à démontrer. L'un de leurs résultats les plus importants et renommés concerne l'effet nocif de la cigarette pendant une grossesse sur le poids de l'enfant à la naissance, ainsi que sur son développement physique et pédagogique jusqu'à l'âge de 16 ans. La cohorte de 1946 a montré que l'effet de faible poids était toujours évident à l'âge de 43 ans. Toutes les informations sur les caractéristiques, les événements, les circonstances, enregistrées par ordre chronologique sont sans pareilles pour analyser et chercher des relations causales, pour orienter le développement de la politique et des services pour la famille et les enfants. Plus ces études seront menées dans d'autres pays, plus nous pourrons faire des comparaisons internationales enrichissantes.

L'histoire de ces études en Grande-Bretagne a été longue et positive, mais non sans problèmes. Le premier d'entre eux est la difficulté de maintenir une entreprise sociale si monumentale sur une période prolongée, des décennies. La question du financement à long terme, qui assure l'infrastructure du projet et inclut le personnel de l'équipe centrale, ne peut pas être sous-estimée. La tension est endémique, dans ces études longitudinales, pour maintenir la cohérence de l'enquête à travers les années tout en abordant les thèmes d'intérêt ou d'importance plus ou moins immédiat souvent favorisés par les financeurs gouvernementaux, ce qui est bien compréhensible. Des tensions permanentes existent également entre le nombre de questions que l'on voudrait poser et la longueur possible d'un entretien. Nous avons établi que le maximum absolu d'interview est une heure et demie, mais il est difficile de s'y tenir lorsque l'on aborde tous les domaines de la vie. La somme des questions indispensables à chaque équipe aboutirait à un entretien de cinq ou six heures ! Très souvent, on laisse quelques questions pour l'étape suivante, ce qui ne fait que déplacer et amplifier le problème. Enfin, nous avons continuellement été confrontés au problème de l'attrition de l'échantillon. La situation en France est peut-être un peu différente, mais en Grande-Bretagne, les sujets des études longitudinales n'ont aucune obligation de participer. Il faut donc à chaque fois susciter leur intérêt. Mais je ne voudrais pas vous faire peur et je vais m'arrêter là.

#### **Claire BAZY MALAURIE**

Je vous remercie infiniment, Madame, pour cet exposé complet et très concret, qui touche au cœur du problème.

### **III. Place de l'épidémiologie dans le projet ELFE**

#### **Georges SALINES, InVS**

Je dirige le département Santé-environnement de l'Institut de Veille Sanitaire. Je ne suis plus impliqué au jour le jour dans le projet Elfe, mais Stéphanie Vandentorren, coordinatrice de la thématique Santé, environnement dans l'équipe projet pourra m'aider si vous avez des questions très précises. Je vais vous parler de la place de l'épidémiologie dans le projet Elfe. L'épidémiologie étudie la répartition et les déterminants des phénomènes de santé dans les populations et sert à :

- la recherche, pour étudier les facteurs de risques ou les causes des maladies ;
- la surveillance, avec l'enregistrement continu, le suivi de certains phénomènes dans le temps pour en mesurer les évolutions ;

- l'alerte, lors de l'émergence d'une nouvelle maladie par exemple ;
- l'investigation, pour déterminer les causes d'une maladie non pas en général mais dans un cas particulier, comme une épidémie de légionellose par exemple ;
- l'évaluation des programmes de santé ou des méthodes thérapeutiques.

Nous employons donc une multitude de méthodes différentes, notamment des enquêtes transversales, ou des enquêtes longitudinales ce qui permet de mieux distinguer les causes des effets, en utilisant soit des études cas témoins soit des études de cohortes. Nous utilisons toujours des méthodes d'observation de la vie réelle, et non de laboratoire. Dans les études cas témoins, qui sont toujours rétrospectives, on interroge un groupe témoin de personnes non malades et un groupe de personnes malades en essayant de reconstituer ce qui, dans leur passé, peut expliquer l'apparition de la maladie. Puis on compare les deux groupes. Dans les études de cohorte, qui sont en général prospectives, on commence à enregistrer des données alors que personne n'est malade et on regarde les facteurs de risque, puis on suit les gens dans le temps et on enregistre l'apparition de maladies. C'est la meilleure méthode pour élaborer un raisonnement causal.

L'InVS est l'un des huit partenaires de la cohorte Elfe. Ce projet répond à deux de nos quatre missions définies dans le Code de la santé publique : la surveillance et l'observation permanente de l'état de santé de la population, et également la veille et la vigilance sanitaires. Nous devons donc détecter des facteurs de risque et pour chaque type de risque, étudier et répertorier les populations les plus fragiles ou les plus menacées. C'est au titre de ces deux missions que nous nous sommes impliqués dans le projet Elfe. Les enfants sont une population prioritaire, notamment et principalement en Santé-environnement car :

- Leurs organes, leur immunité et leur mécanisme d'élimination des toxiques sont immatures ; ils se défendent moins bien contre les agressions de l'environnement ;
- Leurs comportements les exposent davantage à certains polluants, en particulier lorsqu'ils portent les objets à la bouche pour les goûter ;
- Proportionnellement, ils mangent, boivent et respirent plus que les adultes ;
- Ils sont en croissance. Leurs cellules se divisent donc plus vite que les nôtres, et leur ADN est plus facilement soumis aux agressions des polluants ou aux radiations ;
- Ils sont plus près du sol et de la poussière que les adultes ;
- Leur surface corporelle est, proportionnellement à leur poids, supérieure à la nôtre : ils ont donc une plus grande surface d'échange ;
- On constate depuis quelques années une plus grande fréquence de certaines maladies chez les enfants, comme l'asthme, les allergies, et l'on s'interroge sur les cancers, les perturbations endocriniennes (la puberté semble de plus en plus précoce), l'obésité et les maladies neurologiques dégénératives.

Nous ne sommes pas les seuls au sein de l'InVS, à considérer les enfants comme une population prioritaire : le département des maladies infectieuses s'intéresse à la transmission mère-enfant, et les unités qui s'occupent des accidents et traumatismes et de la nutrition sont également concernés.

Cette préoccupation du lien entre environnement et santé des enfants n'est pas propre à la France. En Europe, l'OMS a organisé une grande conférence à Budapest en 2004, qui a adopté un plan 2004-2010 mettant l'accent en particulier sur les problèmes liés à la pollution de l'eau, aux accidents et traumatismes, à l'obésité, à l'asthme en lien en particulier avec la pollution de l'air, aux cancers et aux problèmes de neurotoxicité, d'immunotoxicité, de génotoxicité, de perturbations endocriniennes et d'allergies liées aux substances chimiques.

La France, de son côté, a adopté le Plan national Santé-environnement dont tout un chapitre est consacré au renforcement de la protection des enfants et des femmes enceintes. L'action 26, qui fait partie des douze actions prioritaires du Plan, demandait la réalisation d'une « étude épidémiologique enfants en lien avec l'étude américaine *National Children's Study* devant faire l'objet d'un partenariat national construit autour d'un axe InVS-Inserm ». Nous sommes en fait aujourd'hui dans un consortium beaucoup plus large.

En complément des interventions d'Henri Leridon et Elsa Ferri, je voudrais citer les cohortes étrangères ayant un volet Santé-environnement et avec lesquelles nous avons des relations : la *National Children's*

*Study*, la ICCCC (le consortium international étudiant les cancers chez l'enfant, et dont nous faisons partie), l'Alspac en Grande-Bretagne (cohorte régionale ayant une très large composante Santé-environnement), l'INMA en Espagne, et de grandes cohortes dans les pays scandinaves. La démarche est donc mondiale, et nous essayons de nous y insérer.

Henri Leridon a déjà évoqué les différentes étapes du projet Elfe. J'aimerais simplement ajouter quelques éléments sur les dimensions santé et épidémiologie qui posent des problèmes particuliers. A la naissance, l'enquête périnatale sera complétée par des données prises dans le dossier médical. On travaille sur des pistes, comme l'agrégation des données issues des bilans de santé obligatoires et l'utilisation des données de l'assurance maladie sur les recours aux soins. On s'interroge sur la manière de recueillir en continu les données sur la croissance et l'âge d'apparition de la puberté, et sur la possibilité d'avoir un examen médical à l'âge de 5-6 ans.

La toute première étape, le jour de la naissance, sera le prélèvement du sang du cordon, pour doser le plomb, certains nutriments et les OMICS. « OMICS » est le suffixe de termes anglo-saxons tels que *toxicogenomics* ou *metabonomics*, qui sont des marqueurs d'effets. Il s'agit de mesurer des quantités biologiques qui traduisent l'impact de l'environnement sur le génome et les métabolismes. Entre le deuxième et le quatrième jour, nous doserons les phtalates (composés du plastique, soupçonnés d'être des perturbateurs endocriniens), les pesticides, la cotinine (composé nicotinique), les alkyl-phénols (solvants), et les organo-étains (produit utilisé pour tuer les vers du bois) dans les urines de la mère. Dans ses cheveux, nous doserons le mercure. Au bout de six à huit semaines, nous essaierons d'obtenir du lait maternel pour doser les indicateurs de PCB (polluants organiques persistants).

Tous ces prélèvements biologiques posent des problèmes particuliers d'acceptabilité, nous travaillons à des pilotes. Comme nous n'effectuerons aucune prise de sang, nous avons bon espoir que les mères acceptent. Les prélèvements posent également des problèmes de logistique pour les transporter, les conserver et les analyser – ce qui coûte aussi très cher. Par ailleurs, comme pour les questionnaires se posera un problème de volume, puisque les quantités prélevées seront limitées. Toutes les analyses espérées par les chercheurs ne pourront donc pas être effectuées, d'autant que nous souhaitons qu'une partie de chaque prélèvement soit conservée dans une biobanque pour pouvoir y revenir ultérieurement en fonction des découvertes scientifiques.

Concernant les données environnementales, comme Henri Leridon l'a expliqué, nous déposerons des capteurs de Radon, gaz radioactif posant des problèmes dans les régions à sol granitique comme la Bretagne ou le Corse, et des capteurs de poussière pour compter les moisissures sources d'allergies. Nous travaillerons aussi à partir de l'adresse de l'enfant, puisque de nombreuses données environnementales sont géo-référencées. Nous rapprocherons donc le lieu de résidence de l'enfant de données sur la pollution de l'air ou la qualité de l'eau, par exemple. C'est une manière économique de travailler, et sans déranger les familles.

#### **IV. Débat**

##### **Denis DURAND, CGT**

Je me réjouis de la tenue de cette réunion. Elle répond à une attente des membres du CNIS, et la démarche est très utile à ce stade du projet. J'aimerais attirer votre attention sur plusieurs points, sans exprimer de réserves sur le principe même du projet. Mais son existence soulève d'importantes questions, du point de vue de la CGT, que je représente aujourd'hui.

D'abord, dans l'état d'avancement du projet, il me semble qu'il y a un contraste entre certains points qui sont très avancés et précis - comme ce que vient d'expliquer Georges Salines ou le projet de brochure mis à notre disposition aujourd'hui, qui préfigure sans doute celle qui sera distribuée aux mères - et d'autres qui sont encore mal définis. Par exemple, pendant combien de temps la cohorte sera-t-elle suivie ? Il y aura des interrogations sur l'aspect scolaire à 6 et 11 ans. Est-il envisagé d'aller plus loin, que se passe-t-il lorsque les enfants deviennent adultes ? Ce flou est peut-être dans la nature de ce type d'enquête.

Comme l'a montré Elsa Ferri, les études longitudinales portent en elles une tension permanente entre l'aubaine que représente le lancement d'une grande opération, avec la tentation de profiter de son cadre pour lancer des études à moindre coût, et les limites du projet. Ceci nous renvoie à des aspects de politique statistique et de coût-avantage. Quand devient-il intéressant d'entrer dans le projet, ou de mener une enquête spécifique ? Ceci a des conséquences directes sur les politiques des organismes commanditaires, donc de l'Insee.

Par ailleurs, nous devons affronter les questions fondamentales que soulève le projet Elfe. Le problème de l'obtention du consentement des unités enquêtées est ici particulièrement compliqué, en raison de la durée du suivi et parce que les enfants, ne pouvant répondre eux-mêmes, seront représentés par leur famille. On voit qu'un organisme de représentation des familles est prévu dans le dispositif de l'enquête, un comité d'éthique a été constitué. Mais il reste la question du consentement de la mère. J'ai moi-même mené une enquête statistique non scientifique auprès d'un échantillon totalement biaisée de mères de famille auquel j'ai exposé les principes de Elfe, et j'ai été confronté à un mouvement de recul spontané, presque un mouvement d'horreur de la part de ces mères de famille. On sent que chacun est conscient que les précautions à prendre en matière de protection des données, protection des individus sont cruciales dans cette affaire. Mais j'attire quand même l'attention sur le fait que la portée de l'interrogation va peut-être au-delà des préoccupations directement liées à la nécessité de réaliser l'enquête.

Je me pose une autre question, un peu radicale et épistémologique : ce type d'interrogation longitudinale, ce type de relation créé entre les données, cette priorité donnée à la trajectoire individuelle sont-ils, pour chaque domaine concerné par le projet, ceux qui répondent le mieux aux problèmes ? On a vu que dans certains cas la réponse est affirmative, mais est-on capable de répondre dans tous les cas étant donné la multiplicité des sujets traités ?

Ceci nous renvoie à une interrogation plus générale et fondamentale sur les précautions à prendre lorsque l'on développe de très grands panels. On a eu l'occasion d'exprimer des réserves sur certains aspects du projet EDP++. Je profite de l'occasion pour dire que ce n'est pas une mise en cause des panels en tant que tels, mais mon intervention vise à attirer l'attention sur le fait que cette façon d'aborder les problèmes renvoie à des questions qui sont peut-être encore plus générales et fondamentales qu'on ne le pense au moment où on lance l'enquête. Peut-être faudra-t-il imaginer d'autres formes de réflexion, associant d'autres partenaires qui ne sont pas nécessairement directement associés à la sphère de la statistique publique, compte tenu de la grande généralité des questions posées à l'occasion de cette enquête.

#### **Alain PELC, Caisse nationale de solidarité pour l'autonomie (CNSA)**

Le projet comporte une part sanitaire importante, mais je n'ai vu nulle part dans le projet les mots « handicap », « incapacité » ou « déficience ». Est-ce volontaire ou s'agit-il d'un oubli qui pourrait être corrigé dans le futur ? La prévalence du handicap est relativement faible, il y aura peu d'enfants handicapés dans le panel. Mais c'est un sujet de préoccupation, 2 % des enfants sont handicapés. Je m'interroge aussi sur la représentativité et l'acceptation des familles. Les enfants handicapés pourraient être, selon la réaction de leurs parents, sous-représentés ou surreprésentés.

#### **Docteur Jacques DESFONTAINE, Union Française pour la Santé Bucco-Dentaire**

Suite à l'appel à propositions, pourrions-nous prendre connaissance de la liste des soixante équipes retenues sur les quatre-vingts projets suite à l'appel à proposition ?

#### **Claire BAZY MALAURIE**

Monsieur Leridon va commencer à répondre, et j'invite chacune des personnes participant au projet à apporter sa contribution à la réponse.

#### **Henri LERIDON**

Sur la dernière question, nous travaillons avec soixante équipes représentant quatre-vingts projets. Nous

n'avons pas fait de sélection, hormis pour quelques cas clairement « hors des clous » de la proposition de recherche. La sélection sera effectuée en fonction des projets présentés concrètement et de leur intégration logique et scientifiquement justifiée au sein de l'ensemble du projet Elfe.

Concernant l'intervention de Denis Durand, l'utilisation abusive de l'outil cohorte est un souci majeur pour nous. L'une des contraintes de l'appel à proposition était le caractère indispensable de l'observation de cohorte dans les projets présentés. Toutes les équipes n'ont pas suivi cette recommandation, et dans une opération de ce genre, il existe un risque de dérive « attrape-tout ». Compte tenu de la multiplicité des demandes de la part des équipes, nous aurons des choix à faire. Nous leur avons récemment rappelé que la sélection à venir intégrerait spécifiquement le critère de la nécessité de l'observation de cohorte, et le fait que cette cohorte d'enfants paraît être la bonne. Il est de la responsabilité de l'équipe de projet, avec l'appui du Conseil scientifique et des autres instances qui souhaiteraient intervenir dans la sélection, d'être cohérents sur ce point.

La question de savoir s'il est opportun d'absorber certaines enquêtes qui auraient lieu de toute façon sous forme transversale mérite en effet d'être débattue. Il ne s'agit pas seulement d'un souci d'économie au sens strict du terme, car l'enquête, de périnatalité par exemple, aurait eu lieu en 2008-2009 en dehors de la cohorte. L'enjeu est surtout scientifique. Les concepteurs de l'enquête peuvent penser qu'il est utile de savoir ce qui s'est passé avant et ce qui se passera après leur observation. Il reste encore de nombreux points d'interrogation au regard des problèmes pratiques ; nous déterminons au fur et à mesure ce qui est réalisable et pertinent, tout en vérifiant que l'économie globale du projet est cohérente. Pour ma part, je pense que la cohorte sera un mode de travail économe, autant financièrement qu'en investissements méthodologiques. Nous n'avons pas réinventé l'enquête de périnatalité, nous avons repris celle qui existait.

Par ailleurs, la question du consentement est en effet absolument majeure. Nous avons tous réalisé des « enquêtes sauvages » comme celle de Denis Durand et obtenu le même type de réactions de recul. Il nous est pourtant permis d'être plus optimistes grâce à l'enquête pilote que nous avons menée en avril en maternité, dans des conditions loin d'être optimales car nous n'étions pas complètement prêts. Nous savons comment améliorer les choses à l'avenir. Nous avons donc proposé à des femmes qui avaient répondu à l'enquête de périnatalité de participer à l'enquête Elfe. L'immense majorité d'entre elles n'avait jamais entendu parler du projet, ce qui n'est pas très bon. Les enquêtrices, ici les sages femmes, devaient les convaincre rapidement, le temps passé en maternité étant de plus en plus court, et les visites aux accouchées très nombreuses. Malgré cela, le taux d'acceptation est de l'ordre de 60 %. Nous plaçons l'objectif plus haut, mais ce n'est pas un mauvais résultat au vu des conditions de réalisation de l'enquête. Nous avons essuyé des refus, liés souvent aux maris, parfois même alors que les femmes avaient accepté. L'annonce de l'enquête au domicile deux mois après la naissance semble inquiéter. Elle paraît parfois vécue comme une forme de contrôle. Nous devons mieux former nos enquêtrices sur les informations à donner sur le contenu du questionnaire à six ou huit semaines, qui est pour une grande part une enquête « classique » de l'Insee sur nos thématiques. Elle comporte de nombreuses descriptions de la situation socio économique car c'est le premier contact avec la famille. Il faut bien montrer aux familles que l'enquêteur n'est pas un enquêteur social, mais un intervieweur sur les thématiques de l'enquête. Nous devons également être capables de désamorcer les craintes fondamentales à l'encontre d'un système de type *Big Brother*, en expliquant que personne n'aura accès à toutes les informations - j'y reviendrai dans mon second exposé - et que l'enquête n'a pas pour vocation de contrôler. Nous établirons un lien interactif entre Elfe et les familles, avec un retour d'information à leur intention.

Enfin, pour répondre à Alain Pelc, nous avons bien entendu des questions sur le handicap, dans la mesure où nous travaillons sur la santé des enfants. Il existe en effet un risque que les familles d'enfants handicapés soient moins (ou plus) volontaires que les autres et faussent la représentativité. Nous essaierons également de ne pas passer à côté des familles aux situations les plus précaires, et sommes conscients que nous rencontrerons des problèmes linguistiques, peut-être les plus irrémédiables. Nous n'avons pas encore d'hypothèses sur une possible sur ou sous représentation de certaines populations. Les deux phénomènes sont possibles. Cependant, notre modalité de sélection initiale montrant clairement que nous nous intéressons à tous les enfants nés pendant quatre fois quatre jours, sans exception, évitera peut-être les déviations que nous aurions rencontrées dans un système basé sur le volontariat.

Néanmoins, il est très difficile de savoir comment se forge l'intime conviction des personnes de répondre ou pas. L'intérêt de l'échantillon totalement représentatif est qu'on peut essayer de vérifier s'il l'est au regard de critères que nous n'avons pas envisagés au départ mais pour lesquels on dispose d'autres sources.

### **Stéfan LOLLIVIER, Insee, rapporteur de l'interformation**

Au sujet de la durabilité de l'opération, nous pouvons potentiellement suivre très longtemps les enfants de la cohorte dans les sources administratives, même jusqu'au décès, et faire le lien entre les causes de ce décès et ce que nous savons d'eux depuis leur naissance. Cela pose un problème de bonne pratique en matière de valorisation conjointe de données de collecte et de données administratives. Cette bonne pratique n'a pas encore vraiment émergé en France, mais nous y travaillons grâce à la cohorte. Par exemple, grâce aux fichiers de paie, quand les sujets commenceront à travailler nous pourrions étudier la discrimination dans l'accès à l'emploi en lien avec la petite enfance. D'après les travaux de la DEPP, les différenciations sociales s'opèrent apparemment avant l'école primaire : une étude longitudinale nous permettra de comprendre sous quelles conditions se créent les inégalités.

### **Henri LERIDON**

L'horizon du projet fixé officiellement aujourd'hui est l'âge adulte. Mais bien entendu, le financement du projet n'est pas assuré sur vingt-cinq ans : il ne s'agit donc pas d'un engagement mais d'un objectif. Aurons-nous la constance de nos collègues britanniques ? Le problème sera en tout cas continuellement reposé aux familles, puis aux enfants directement.

### **Daniel NAULLEAU, Université Paris VI**

Je suis enseignant et je m'occupe également des problèmes d'informatique et libertés. J'aimerais poser une question technique. Quel pourcentage d'acceptation devrez-vous obtenir pour que l'enquête soit valide ? Je pense au refus massif des parents d'élèves parisiens lors d'une enquête de la MGEN. N'y a-t-il pas un risque de biais si une partie des gens, ceux qui sont sensibilisés à la problématique Informatique et libertés, refusent de répondre ?

### **Henri LERIDON**

Il n'y a pas de réponse absolue. J'espère que nous atteindrons un taux d'acceptation largement supérieur à 60 %. Dans la pratique, les enquêtes publiques ont des taux d'acceptation de l'ordre de 80 à 85 % pour celles qui fonctionnent bien. C'est peut-être même plutôt le record. Je me réfère aux enquêtes que j'ai menées dans le passé, avec l'Ined ou l'Inserm, pour lesquelles les taux d'acceptation étaient élevés. Concernant Elfe, nous commencerions à nous poser des questions sur les biais si le taux était inférieur à 60 %. Mais nous aurons toujours des biais, même à 85 %. Nous voudrions nous donner les moyens de les contrôler. L'un des intérêts pour nous de l'enquête périnatale est qu'elle comprend quelques questions de description socio-démographique et qu'elle atteint un taux de 95 % d'acceptation. Nous pourrions donc comparer dès la naissance les personnes qui auront répondu à l'enquête périnatale et celles qui auront accepté de participer à l'opération Elfe.

### **Stéphanie VANDENTORREN, InVS**

Je coordonne les aspects Santé-environnement de la cohorte. Je voudrais apporter un complément technique sur la question du handicap. Le problème du déficit neurosensoriel, par le biais des expositions chimiques, nous tient particulièrement à cœur. Nous avons également un groupe sur les accidents et traumatismes qui s'intéressera aux séquelles.

### **Chantal CASES, Directrice de l'Irdes**

Pour compléter les réponses aux questions de Denis Durand, je voudrais ajouter qu'une grande partie de la dernière séance du Conseil scientifique d'Elfe a été consacrée à la question du choix qui sera vraisemblablement nécessaire entre les projets qui se sont manifestés. Afin d'aider l'équipe Elfe à sélectionner certains projets, nous avons élaboré des critères établissant la nécessité de la dimension longitudinale et de la pluridisciplinarité des études, ou à tout le moins le fait qu'elles se préoccupent de

questions de santé et de questions sociales. Ce point est essentiel. Je représente avec d'autres la composante « sciences sociales » au sein du Conseil scientifique. Je tiens à souligner combien ce type de données est utile et novateur, en particulier dans les problématiques des inégalités sociales de santé que nous traitons beaucoup à l'Irdes. Cette problématique est particulièrement sensible en France. Les indicateurs de la mortalité prématurée nous placent par exemple en tête de l'Europe de l'Ouest en matière d'inégalités sociales de santé. Des travaux ont été réalisés pour tenter de comprendre la formation de ces inégalités. Le corpus d'hypothèses est très varié : on évoque la génétique, le social ou le comportemental. On sait que ces inégalités sociales de santé existent dès la naissance, elles se renforcent pendant l'enfance et sont très prégnantes à l'âge adulte. Le projet Elfe est une occasion de vérifier ou d'infirmier un certain nombre d'hypothèses. Nous avons besoin d'une collecte de données de grande ampleur, parce que la santé est multifactorielle et qu'étudier l'exposition à des polluants ne suffit pas. Nous avons besoin de croiser les facteurs environnementaux, comportementaux et sociaux.

### **Sophie BAUDET-MICHEL, Université Paris-Diderot**

Je suis géographe, enseignante et chercheur. Je travaille sur des questions de santé et d'environnement. J'aimerais savoir si une attention sera portée aux différents types d'espaces auxquels « appartiendront » ces enfants, par exemple à la taille de leur ville ou à la région qu'ils habitent. On sait que les inégalités sociales et de santé ne sont pas distribuées de manière aléatoire dans l'espace, et qu'il y a des concentrations d'inégalité.

Y aura-t-il des questions relatives aux lieux fréquemment fréquentés par les enfants (lieu de garde, de scolarisation, ...), qui ne sont pas obligatoirement identiques, en termes d'expositions, au lieu de résidence ? Les questions de mobilité, la manière dont ils se déplacent, sont également importantes.

Enfin, j'ai une interrogation sur la manière dont vous sélectionnez les équipes qui auront accès aux résultats de ces études. Changeront-elles au fur et à mesure, et pourra-t-on postuler dans le futur pour accéder aux données ?

### **Henri LERIDON**

Étant donné le mode de constitution de l'échantillon, il sera géographiquement représentatif de la France entière. Nous utiliserons des modalités de description du lieu de résidence classiques, et le géocodage de l'adresse, qui permettra de mener des études sur les caractéristiques de l'environnement, elles-mêmes référencées. Concernant les autres lieux de passage de l'enfant, par exemple s'il est gardé ou scolarisé ailleurs, Stéphanie Vandentorren saura peut-être nous éclairer sur le stade actuel du projet.

### **Stéphanie VANDENTORREN**

Nous avons un module intitulé « modes de garde » pour savoir comment l'enfant est gardé. Nous avons proposé de recueillir également l'adresse de ce lieu de garde pour pouvoir en étudier l'environnement comme au domicile. Mais il sera plus facile d'obtenir l'adresse d'une crèche ou d'une assistante maternelle que celle, par exemple, des grands-parents. Pour les modes de garde usuels, ce sera vraisemblablement possible.

### **Henri LERIDON**

Ceci supposerait d'obtenir chaque fois une adresse précise ; nous ne pouvons pas multiplier les adresses individuelles, surtout pour des tiers.

### **Sophie BAUDET-MICHEL**

Vous répondez sur la petite enfance, mais il me semble, pour faire écho aux questions d'actualité sur la carte scolaire, que le lieu de scolarisation peut être très différent du lieu de résidence. Les expositions seront donc différentes.

### **Henri LERIDON**

Nous travaillerons sur cette question, qui est pertinente. Concernant la sélection des équipes de

recherche, le projet est actuellement élaboré par un ensemble d'équipes de recherche qui se sont manifestées lors de nos appels à proposition. Si nous constatons un manque, nous pourrions éventuellement faire un nouvel appel, mais le Conseil scientifique ne nous encourage pas à trop élargir. En revanche, dans le futur, nous devrons certainement rouvrir l'appel à proposition lors des différentes étapes du projet. Et un accès aux données sera assuré à toutes les équipes de recherche. Cette accessibilité nous est demandée par diverses instances et financeurs.

### **Caroline BRAY, OSII**

Certaines parties de la population seront-elles surreprésentées dans votre échantillon, notamment la catégorie des enfants d'immigrés ?

### **Henri LERIDON**

La question de la surreprésentation *a priori* a été réglée par le choix de notre mode de sélection qui est de prendre tous les enfants qui naissent certains jours donnés en France. En revanche, il serait envisageable que des enfants qui émigrent en France à trois, quatre ou cinq ans, et qui sont nés dans la période déterminée pour Elfe, soient éligibles pour l'enquête. Il serait difficile de les retrouver avant l'école, mais à partir de leur scolarisation, ils seront éligibles pour le projet.

### **Jean-Pierre LE GLÉAU, Insee**

J'aimerais demander une précision à Elsa Ferri. Vous avez parlé uniquement « d'enquêtes » dans les études britanniques. Les données de vos cohortes ont-elles également été enrichies de données administratives ?

### **Elsa FERRI**

Dans toutes les études de cohortes, nous avons utilisé les données des recensements décennaux, qui appartiennent au domaine public. Plus récemment, nous avons utilisé des fichiers médicaux. Pour la MCS, nous avons eu accès à des fichiers de naissance des maternités, puis aux fichiers d'hospitalisation des enfants avec les motifs d'hospitalisation. A l'âge de cinq ans, nous avons obtenu les données scolaires. Il existe en Angleterre des profils pédagogiques pour chaque enfant, centralisés dans le département de l'éducation. En Écosse, en Irlande du Nord et au Pays de Galles, il a fallu s'adresser directement aux écoles pour demander un accès aux données. Le taux d'acceptation a été de 70 % en Écosse et de 50 % au Pays de Galles. Nous ne savons pas à quoi est due cette différence.

Dans tous ces cas d'utilisation des fichiers médicaux ou pédagogiques, le consentement explicite des parents a été une exigence.

### **Henri LERIDON**

Le consentement des parents a-t-il été obtenu facilement ?

### **Elsa FERRI**

Je ne connais aucun cas de refus. Ils ont accepté dans 100 % des cas.

### **Georges SALINES**

Dans les cohortes danoise et norvégienne, comptant chacune 100 000 couples mère-enfant et s'intéressant aux effets à long terme des expositions pendant la grossesse, les femmes sont recrutées pendant leur grossesse et le recueil de données spécifiques se fait essentiellement pendant la grossesse et à la naissance, avec des prélèvements biologiques. Mais tout le suivi, à l'exception de quelques rares enquêtes téléphoniques, est fait par agrégation de données diverses, notamment issues des registres de cancers, de malformations, des données d'hospitalisations, etc. Ce sont des systèmes statistiques complètement différents des nôtres.

### **Claire BAZY MALAURIE**

Voilà qui fait la transition avec la question de l'utilisation des données et les problèmes éthiques.

## V. Dimensions déontologiques et éthiques du projet Elfe

### Henri LERIDON

Il est absolument essentiel que nous soyons capables de protéger les données que nous recueillerons dans le cadre du projet Elfe. Ces données seront particulièrement sensibles, pour différentes raisons. D'une part, elles seront recueillies sur le long terme et on envisage de les croiser avec des données externes, ce qui représente potentiellement une masse d'information considérable. D'autre part, elles concerneront des enfants représentés par leurs parents, ce qui pose des problèmes spécifiques. Nous devons vérifier, lorsque les enfants auront grandi, qu'ils consentent eux-mêmes au projet, mais il sera trop tard pour revenir totalement en arrière.

#### 1. Les instances d'arbitrage internes

Certaines de nos instances nous servent de guides et de garde-fou :

- Le *Comité d'orientation stratégique* est l'instance dirigeante du projet. Il associe de plein droit les huit institutions membres du projet Elfe. C'est l'assurance d'une diversité de sensibilité et de responsabilité. Il n'intervient pas au quotidien mais prend les grandes décisions et est en charge d'entériner nos choix. Au cours des dernières réunions, les directeurs de l'Ined, de l'Inserm et de l'InVS étaient généralement présents en personne.
- Le *Conseil scientifique* nous aide à prendre des décisions par rapport aux choix, par exemple des projets de recherche. Il a également été saisi de plusieurs problèmes éthiques, et nous a encouragés à constituer un groupe spécifique à ce sujet.
- Pour constituer le *Groupe éthique*, nous avons eu des difficultés à trouver des spécialistes axés sur les sciences sociales. Le problème éthique principal du projet Elfe est en effet que l'on disposera d'un grand nombre d'informations sur une même personne et que cette personne sera, pendant plusieurs années, « représentée » par ses parents. Ces questions, à ma connaissance, n'ont pas encore été abordées en France.
- Nous voulons constituer *un groupe représentant la société civile*, qui pourrait peut-être compter des participants à cette enquête ou à d'autres cohortes. Il pourra donner un avis très direct sur ce qui est acceptable et ce qui ne l'est pas dans notre démarche. Trouver des représentants qui s'engagent n'est pas facile. Nous n'étudions pas une population malade : nous ne pouvons donc pas nous appuyer sur des organisations de patients. Nous sommes prêts à étudier toute proposition pour constituer ce groupe.

#### 2. Les instances d'arbitrage externes

Cinq instances interviennent pour vérifier que nous travaillons dans des conditions convenables :

- La *Cnil* : nous avons eu plusieurs rencontres informelles depuis 2006. Un dossier a été déposé pour la première enquête pilote ; un deuxième le sera pour la prochaine, puis pour chaque enquête pilote suivantes, et bien sûr pour l'enquête en 2009.
- Le *Cnis* : la prochaine étape sera la présentation au Comité du label en septembre.
- Le *CCTIRS (Comité Consultatif pour le Traitement Informatique dans les Recherches sur la Santé)* : il traite de tous les projets d'enquête ou de recherche impliquant des données personnelles sur la santé. Nous lui avons remis un dossier pour le premier pilote, qui a reçu un avis favorable, et devons lui en remettre pour toutes les enquêtes suivantes. Ce comité vérifie que les questions que nous posons ou les prélèvements que nous envisageons sont pertinents par rapport à la recherche proposée et que le traitement des données est convenable. Il travaille en articulation avec la Cnil.
- Le *CPP (Comité de Protection des Personnes)* : c'était autrefois le CCPPRB (comité de protection des personnes dans les projets de recherche biomédicale). Il rentre plus dans le détail des aspects concrets de la démarche de prélèvement, vérifie que les présupposés scientifiques sont convenables et que les conditions techniques et éthiques de recueil sont assurées. Nous lui transmettrons un dossier pour le prochain pilote, le premier n'ayant pas nécessité de prélèvements.
- La *DGS (Direction Générale de la Santé)* : en parallèle au dossier du CPP il est prévu désormais de faire un dossier auprès de la DGS.

### 3. Les autres contacts

Nous essayons de développer d'autres contacts, par exemple avec les professionnels concernés par la santé, et en particulier la petite enfance. A travers leurs sociétés, leurs conseils nationaux, nous avons approché les pédiatres, les gynécologues obstétriciens, les sages-femmes et les spécialistes de périnatalité ou de néonatalogie. Il est très important qu'ils soient tous informés de notre projet et qu'ils aient la possibilité d'émettre leurs remarques et d'éventuelles réserves dès maintenant. Nous n'avons pas eu de réactions *a priori* négatives et avons reçu des conseils pratiques importants pour le succès de l'opération. Nous faisons également des présentations dans des congrès scientifiques en France et à l'étranger, ce qui nous permet d'avoir des réactions, de comparer notre projet avec d'autres, déjà menés ou en construction.

Enfin, nous ne négligeons pas de tester nos questionnaires afin de recueillir les réactions des personnes enquêtées et des enquêteurs professionnels eux-mêmes. Il est très enrichissant pour nous de connaître les parties du questionnaire qui fonctionnent bien ou moins bien, pour aménager la présentation, changer l'ordre des questions, voire exclure certaines parties. Nous avons l'expérience de l'enquête périnatale et des expériences plus réduites du questionnaire six à huit semaines qui fera prochainement l'objet d'une véritable enquête pilote.

Mon expérience en termes d'enquêtes m'a appris que si l'on se donne la peine de travailler en profondeur un questionnaire, et en interaction avec les personnes qui y répondent, on peut obtenir des réponses que je crois non contraintes et sincères pour une grande majorité, et allant beaucoup plus loin que ce que l'on imagine *a priori*. J'ai l'expérience des enquêtes sur les comportements sexuels en France, qui ne sont pas les plus faciles, qui ont été réalisées par une équipe multidisciplinaire, non sans tensions internes parfois importantes. Les questionnaires n'ont finalement provoqué aucun drame. Il a parfois été difficile pour les enquêteurs de poser certaines questions, mais la plupart des enquêtés ont semblé contents de pouvoir se confier sur un sujet qu'ils avaient rarement abordé, voire jamais. Un questionnaire doit être construit, justifié scientifiquement, et testé du point de vue de son acceptabilité. Par ailleurs, il faut faire attention à ne pas interpréter une question hors de son contexte, et ce n'est pas facile à faire comprendre à des non professionnels.

### 4. Procédures d'information des participants et consentement

La participation sera, bien sûr, volontaire. Elle sera proposée aux mères pendant leur séjour en maternité après le questionnaire de l'enquête périnatale. Une information préalable aussi large que possible aura été faite. Il est actuellement envisagé que les CAF, qui nous l'ont proposé, envoient une information avec le courrier qui est adressé au septième mois de grossesse à toutes les femmes ayant déclaré leur grossesse. Elle couvrira environ 90 % des femmes enceintes. Des dépliants et des affiches seront disposés dans les consultations des maternités, pour que les femmes qui consultent au cours du troisième trimestre de grossesse dans le service où elles vont accoucher puissent encore prendre connaissance du projet.

Nous utiliserons aussi les médias. Nos contacts avec les médias locaux lors de notre enquête pilote ont été très positifs. L'utilisation des médias, bien sûr, est à double tranchant, mais nous devons jouer l'information et la transparence. Nous ferons une information au plan national le moment venu.

Tous les professionnels susceptibles d'intervenir doivent être bien informés, notamment dans les maternités. Si tout le monde est au courant dans la maternité, le contact avec la mère est grandement facilité.

Une lettre d'information précise destinée à la mère, et éventuellement au père, expliquera ce que nous attendons d'eux à court et moyen termes. Le consentement se concrétisera avec la communication de l'adresse du domicile.

Pour les prélèvements biologiques, nous aurons besoin d'un recueil écrit de consentement, distinct du consentement à la participation au projet Elfe. Les personnes pourront refuser les prélèvements ou l'analyse des prélèvements qui auront été faits et en même temps accepter de participer à Elfe.

Pour chaque nouvelle étape, nous redemanderons un consentement et réexpliquerons le processus. L'acceptation initiale ne vaut pas engagement pour vingt ans. Les personnes peuvent décider à tout moment de ne plus participer à l'enquête.

Enfin, nous informerons les familles par l'envoi régulier d'un journal d'information sur le déroulement du projet et des résultats non susceptibles d'influencer leur comportement. On envisage également l'accès à un site web réservé où elles pourront poser des questions et obtenir plus d'informations.

## 5. La protection des données

Depuis plusieurs mois, nous avons mis en place un groupe spécifique intitulé « Informatique, protection des données », composé de spécialistes de l'Ined, l'Inserm, l'Insee et l'InVS. Ils s'informent auprès d'experts en anonymisation et cryptage des données. Leur mission est d'aider à choisir les méthodes qui assureront la protection totale de l'anonymat des participants, la sécurité des fichiers de données qui doivent être protégés contre tout accès illicite – tout en ayant la possibilité de suivre les enfants nominativement dans le temps –, le contrôle des accès aux données pour les analyses scientifiques, et tout ceci dans la durée. Comme le disait Elsa Ferri, il n'est pas facile d'assurer le fonctionnement d'un système en continu sur plusieurs décennies.

Les principes que nous avons pour l'instant retenus sont les suivants :

- Les données de chaque collecte donneront lieu à un fichier particulier et anonyme, comportant un identifiant spécifique, propre à cette collecte. Par exemple, l'enquête périnatale constituera un fichier, l'enquête de six à huit semaines un autre fichier, et une collecte de données extérieures un troisième.
- Le seul fichier nominatif sera celui comportant les adresses des familles et tous les éléments permettant d'assurer la continuité de l'observation (les téléphones des personnes relais si elles ont accepté de nous les donner, la réponse ou pas à telle étape).
- La liaison entre ces différents fichiers serait assurée par un tiers de confiance externe, seul détenteur des clefs de passage entre les différents fichiers. Il délivrerait les identifiants propres à une étape donnée avec les adresses, mais n'aura pas lui-même les divers fichiers de données.
- Les prélèvements biologiques seront conservés en bibliothèques, protégés, avec des conditions d'accès au produit pour des recherches scientifiques elles mêmes réglementées.

Il n'y aura donc aucune mégabase Elfe. Aucun fichier ne comportera jamais la totalité des informations du projet. Quelqu'un qui accéderait aux fichiers de deux étapes différentes ne serait pas en mesure de rapprocher les individus, puisque les identifiants seraient différents.

Nous prendrons par ailleurs toutes les précautions nécessaires de duplication pour assurer la survie des données en cas de problème technique.

En revanche, nous abandonnons l'option du hachage irréversible de l'identifiant, qui aurait supposé la constitution d'une méga base et aurait empêché tout retour en arrière. Ceci aurait pu poser problème dans une étude de cohorte suivie, dont on ne peut prévoir toutes les questions qu'elle soulèvera à l'avenir.

Toute équipe de recherche du secteur public – celles travaillant actuellement sur le projet comme les autres – et désirant utiliser les données Elfe devra soumettre une demande d'accès à une instance *ad hoc*. Le rôle de cette instance sera de vérifier l'intérêt scientifique du projet et la pertinence des données demandées au regard de ce projet. L'intérêt du projet Elfe réside dans le fait que les observations sont successives et pluridisciplinaires. Nous ne pouvons donc bien entendu pas cloisonner l'accès aux données. Les chercheurs doivent avoir le droit d'accéder à toutes les bases, mais en justifiant la pertinence de leur demande.

La difficulté majeure sera de vérifier que l'association de ces bases ou parties de bases n'augmente pas le risque d'identification indirecte. L'instance devra donc réfléchir assez tôt aux variables potentiellement identifiantes et aux modalités de communication qui peuvent être envisagées. On peut penser qu'un chercheur ayant besoin, dans le cadre de sa recherche, d'une localisation géographique fine n'aura pas accès à des données descriptives du statut économique détaillées. Nous devons avoir plusieurs échelles

de modalités possibles ; dans chaque cas, compte tenu de la nature de la recherche, nous donnerons accès à tel ou tel type d'informations qui n'auront pas toujours le niveau de détail maximum.

Enfin, c'est encore une hypothèse, si l'on veut éviter que trop de bases ne circulent, nous pourrions imaginer que le fichier ouvert à un chercheur ne lui soit pas directement transmis, mais soit mis à sa disposition sur un serveur dédié. C'est techniquement possible, mais lourd à gérer. Nous cherchons des centres pouvant assurer une accessibilité vingt-quatre heures sur vingt-quatre et une sécurité d'accès et physique maximales. Il faudrait par ailleurs que chaque équipe de recherche trouve sur ce serveur tous les outils d'analyse qui lui sont nécessaires. Mais nous travaillons dans la perspective de 2010 ou 2011. Les modalités d'accès à des données à distance, les sécurités et les temps d'accès auront sans doute été améliorés d'ici là, et nous pourrions alors envisager cette option.

Nous travaillons énormément sur ces aspects de protection des données, avec des personnes compétentes et qui vont chercher les compétences qu'elles n'ont pas elles-mêmes là où il faut. Nous sommes demandeurs de tout complément d'information sur des expériences utiles. Nous nous orientons en tout cas vers un système de protection et d'accès très construit et correspondant à l'état actuel et futur de la science en France pour des observations de ce genre.

Bien entendu, le choix des procédures de protection et gestion des données s'effectuera sous le contrôle des instances d'arbitrage que j'évoquais plus haut, et qui ratifieront les choix.

## **VI. Débat**

### **Bernard SUJOBERT, Insee**

Je suis statisticien. Pour avoir assisté à la réunion d'octobre dernier, je mesure combien les choses ont avancé et combien l'effort d'information aujourd'hui est louable. Qu'en est-il de la transparence des réflexions menées par les instances que vous avez évoquées (Conseil scientifique, etc.) ? Jusqu'à présent, nous ignorions quelle était la nature des questions et des choix auxquels elles étaient confrontées. Il serait intéressant que nous puissions avoir une certaine visibilité de la progression de leurs réflexions. Cela pourrait répondre à des questions qui ont été évoquées aujourd'hui.

### **Jean-Philippe DAMAIS, Géographe**

Je suis géographe. Ma question porte sur l'après projet. Vous avez dit que l'accès aux données Elfe sera réservé en principe aux chercheurs du secteur public. Cela diminuerait mes craintes. Mais il existe un risque d'aboutir à une typologie de parcours, donc un risque d'utilisation stigmatisante de ces différents types de parcours, se rapportant à des populations elles-mêmes typées. C'est une question que je me pose, peut-être à tort.

### **Claire BAZY MALAURIE**

Même s'il est évident qu'il s'agit là d'éviter des appropriations à objectif commercial ou autre, la question du secteur public posera problème, car ce secteur est très étroit et ne représente pas la recherche dans son entier. Nous reviendrons sur la question je pense.

### **Benoît ROBIN, CGT-FO**

Le projet Elfe est très ambitieux, et les réalisations ayant été menées à l'étranger montrent bien son intérêt scientifique. Malgré tout, nous avons un certain nombre d'interrogations, notamment parce que les données sont sensibles. Les conditions dans lesquelles semble se dérouler l'étude concourent à donner pleinement satisfaction sur les précautions prises pour mener ces enquêtes, mais reste la question de l'accès aux données. La création d'une instance *ad hoc* donnant les autorisations d'accès a été évoquée. Mais le projet, parce qu'ambitieux, sera attractif. Les éléments d'une enquête s'intéressant aux inégalités ne sont pas tous strictement objectifs. Il faudra donc veiller à ce que l'on ne puisse pas agréger n'importe quelle information aux données existantes, avec la tentation de caricaturer ou de créer des pseudos ghettos.

Par ailleurs, la « France entière » couverte par Elfe comprendra-t-elle les DOM-TOM ?

### **Jeanne BOSSI, Cnil**

Je voudrais rappeler l'état de la réflexion de la Cnil sur les principes et les modalités de mise en œuvre du projet Elfe. Nous sommes en contact avec Henri Leridon et son équipe depuis plus d'un an pour discuter de la mise en place de l'enquête et de ses principes. Nous avons ce genre de contacts avec l'ensemble des chercheurs préparant un projet de longue haleine ou la mise en place de cohortes. Je rappelle que la protection des données personnelles ne s'oppose pas à la mise en place de telles cohortes. Il est important de le rappeler. On ne peut donc pas se servir de l'informatique et des libertés pour s'opposer à ce genre de projets.

Les modalités de mise en œuvre sont essentielles. Au regard de la loi Informatique et libertés, la Cnil se concentre précisément sur la finalité du projet et la nature des données collectées, sur les modalités d'information et sur la sécurité. Ici, la finalité pluridisciplinaire pose des questions particulières sur la protection des données personnelles, qui s'organise normalement autour d'une notion de finalité centrale. Il est très difficile, dans une enquête de ce type qui collecte des informations sociologiques, sociales, médicales, de trouver une finalité qui les rassemble et qui puisse permettre de dire que toutes ces données sont pertinentes au regard de la finalité affichée de l'enquête. C'est la raison pour laquelle nous avons beaucoup travaillé sur la nature de ces données, et nous continuons à le faire. Et nous nous interrogeons sur le libellé de certaines questions ou sur le fait que des questions que nous considérons comme plus sensibles et plus intrusives que d'autres devraient davantage être posées sous forme d'auto-questionnaires.

Il est important que la cohorte se construise dans le temps et que la Cnil soit tenue au courant et associée à chacune des étapes de cette construction. L'utilisation des données Elfe au bénéfice d'autres enquêtes épidémiologiques sera elle aussi soumise à la Cnil, et nous ne pouvons pas nous prononcer dès aujourd'hui sur ces enquêtes. La Cnil les appréciera donc au cas par cas.

Par ailleurs, le retentissement médiatique de deux enquêtes récentes de la MGEN et de l'Inserm (cette dernière ayant été autorisée par la Cnil) centrées sur la prédisposition des enfants ou de leur santé a montré combien l'information préalable de la population est capitale. Il faut informer à travers les médias, laisser le temps aux personnes d'enregistrer le message, de l'analyser, de poser des questions, de comprendre, et ce temps n'a pas été suffisant dans les deux exemples. Cet enseignement est très important pour la cohorte Elfe, qui touche elle aussi des enfants, sujet sensible. Tous les moyens doivent être mis en œuvre pour que les femmes aient entendu parler de l'enquête avant d'être interrogées. La collégialité des commissaires de la Cnil y attachera énormément d'importance.

Enfin, la Cnil apprécie les mesures de sécurité permettant d'assurer la confidentialité des données. Le système d'anonymisation retenu fait penser à d'autres systèmes retenus dans des projets de cohorte que la Cnil a pu autoriser, qui sont assez solides. La Cnil pourrait être plus directement associée au groupe sur l'informatique et la protection des données qui réfléchit en ce moment sur le dispositif d'anonymisation et sur le choix du tiers de confiance. Nous gagnerions du temps, et la Cnil pourrait certainement apporter dans ce domaine.

Quant aux conditions d'accès aux données, la distinction entre secteurs public et privé me paraît aujourd'hui dépassée. La question porte plus sur les modalités d'accès pratiques et le type de fichiers auxquels on pourra accéder. Des organismes, dont l'Insee, se posent ces questions aujourd'hui. Il n'est pas question de fournir des données identifiantes. Alors, pourquoi bloquer en amont l'accès à des données anonymisées aux équipes privées ou publiques dès lors que les conditions qui seront posées pour cet accès seront respectées ? Là encore, je considère qu'on est dans une phase d'amont et que ces conditions d'accès devront être définies au fil du temps et seront nécessairement amenées à évoluer. Voilà les quelques éléments que je souhaitais porter à votre connaissance.

## **Henri LERIDON**

Concernant la transparence des processus de réflexion des instances d'arbitrage, nous pourrions demander leur accord pour mettre en ligne au fur et à mesure les comptes rendus des réunions faisant état des décisions prises. Je soumettrai cette idée, notamment au Conseil scientifique.

Sur l'instance à créer pour contrôler l'accès aux données, nous avons déjà eu des expériences similaires en France. Certains organismes (la Drees, l'Irdes, l'Insee pour son enquête de santé, ...) lancent des appels d'offre pour exploitation de données existantes. Le comité scientifique de l'appel d'offre décide de celles qui méritent d'être soutenues. Il est possible d'évaluer la pertinence des projets proposés.

Sur la distinction public/privé, c'est moi qui l'ai proposée. L'esprit global de notre opération appartient en effet au secteur public, pour les équipes qui y travaillent, les organismes partenaires et le financement. Le COS (Comité d'Oriente) ne m'a pas encouragé à chercher des fonds privés. Mais cette position n'est pas définitive ; nous en débattons le moment venu avec les instances compétentes.

Je partage les remarques de Jeanne Bossi, en particulier sur l'importance de l'information par exemple. Je comprends les difficultés d'interprétation des finalités. C'est difficile pour nous aussi. Notre terrain n'est pas balisé, et nous devons à la fois éviter de donner l'impression que nous répondrons à tout et de donner l'impression que nous ne savons pas ce que nous allons faire. Pour le moment, notre projet global est une somme de projets. Nous l'avons présenté pour la première fois dans son entier à la dernière réunion du Conseil scientifique, qui commence donc à en avoir une idée concrète. Il faut considérer la logique du projet plutôt que les intentions initiales des différents projets. Nous associerons très volontiers la Cnil à nos réflexions. Nous ne l'avons pas fait plus tôt par souci de séparation des rôles.

Enfin, nous avons prévu d'intégrer les DOM aux premières étapes de nos recherches (nous verrons si le suivi sur le plus long terme est possible), mais pas les COM.

## **Jean-Paul AMANN, Collège de France**

Une partie des dernières questions posées relève d'une interrogation éthique. Des précautions légales sont requises, et seront prises. L'analyse éthique de ce genre d'enquêtes, dont les exemples sont plus nombreux à l'étranger, se mène en termes de risques et de bénéfices escomptés. On distingue ensuite pour qui sont les risques, et pour qui les bénéfices.

Ici, les bénéfices escomptés pour la collectivité sont considérables, escomptés bien entendu. Pour citer un exemple bien connu, il y a aujourd'hui des grandes campagnes internationales sur le tabac, parce que l'accord scientifique est assez général sur les méfaits du tabagisme, en termes d'incidence sur le cancer broncho pulmonaire. Ceci a été obtenu grâce à des études épidémiologiques bien connues, études de cohortes datant des années 46-50. Pour Elfe, les bénéfices attendus pour la collectivité sont bien entendu considérables. Les bénéfices pour les individus qui participeront seront en revanche mineurs, on ne les soignera pas, ils ne recevront pas de gratification. Ce sont des volontaires, il est important qu'ils aient conscience qu'ils vont participer à une entreprise qui comporte à la fois des risques et des bénéfices.

Dans une discussion comme celle d'aujourd'hui, l'objet est de mettre au jour les risques. Le risque principal pour les individus serait la rupture de l'anonymisation. C'est en partie une question technique, sur laquelle le projet doit être irréprochable, et qui doit s'entourer des meilleurs spécialistes de cryptage. Un autre risque est celui de la stigmatisation de certains comportements ou de certains groupes. Si l'on s'apercevait rétrospectivement qu'une corrélation statistique existe entre le fait d'être jumeau et une pathologie sociale, qu'advierait-il ? Ce genre de risques de stigmatisation nécessite, pour être anticipé, à la fois les meilleures garanties scientifiques, pour que les données ne soient pas exploitées de façon indue, et une profonde réflexion sociale sur la place de la connaissance scientifique, et ses limites à l'intérieur des comportements sociaux. C'est pourquoi la suggestion qui a été faite, et je pense acceptée, de rendre transparente la réflexion éthique interne et sans doute aussi externe qui sera menée sur les finalités et sur les moyens me semble extrêmement importante. Il faut sans doute un large consensus social sur la nécessité d'aborder ce genre de questions, faute de quoi on ira au devant de difficultés qui risquent de remettre en question des efforts et de les rendre peut-être inutiles.

### **Denis DURAND**

Je me félicite à mon tour de la qualité du premier débat qui s'est déroulé aujourd'hui. La sécurité des données Elfe reposera sur des systèmes informatiques, et je voudrais attirer l'attention sur le fait que tous les systèmes, même ceux qui sont parfaits, sont faillibles. Nous en avons tous fait l'expérience. Je craignais d'ailleurs un peu que les exposés d'aujourd'hui ne nous fournissent une masse d'informations techniques prouvant que le système serait parfaitement construit. Je sais gré à Henri Leridon de ne pas l'avoir fait, à la fois parce que nous ne sommes pas en mesure de juger de la qualité technique d'un système informatique, et parce que les systèmes parfaits n'existent pas. Ils doivent être insérés dans une organisation du travail et un contexte social. Plus un système informatique est grand et compliqué et plus le bug est assuré. Il me semble que le projet Elfe présente un risque maximal, étant donné le volume des données collectées, le nombre des parties prenantes, le nombre d'équipes utilisatrices, l'hétérogénéité des données, leur sensibilité, la possibilité de croiser ces données avec des données administratives, la durée de conservation des données, et je ne suis pas sûr d'être exhaustif. Par ailleurs, les chercheurs ne sont pas une population particulièrement préparée à considérer la sécurité des données comme une priorité. Peut-être faut-il réfléchir à des mesures exceptionnelles pour tester la capacité du dispositif dans son ensemble à protéger la sécurité des informations. Est-il envisageable de faire réaliser des audits d'organismes indépendants quand le projet se mettra en place ?

### **Jean-Pierre LE GLÉAU**

La participation au projet Elfe étant volontaire, quelles seraient les conséquences si une personne décidait de tout arrêter ? On pourrait soit arrêter de l'interroger mais continuer à collecter des données administratives, soit arrêter toute nouvelle collecte d'information, soit effacer toute trace des informations collectées dans le passé. Les trois possibilités ne sont pas équivalentes du point de vue de l'acceptabilité du projet.

Je voudrais par ailleurs faire une remarque sur le respect du secret statistique. Il est à distinguer de manière très nette des recommandations pouvant être faites par la Cnil et faisant référence à la loi Informatique et libertés de 1978. Le secret statistique, lui, est établi par la loi du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques. Parmi les informations Elfe, certaines seront collectées dans le cadre d'enquêtes statistiques ayant reçu un visa à la suite de leur passage devant diverses instances du Conseil national de l'information statistique - Formation et Comité du label - puis ayant été reconnues enquêtes statistiques par une publication au Journal officiel. La loi dit, à propos des données recueillies sur les personnes physiques dans ce cadre, que les informations individuelles ne peuvent faire l'objet d'aucune communication de la part du service dépositaire avant une période de cent ans. L'interprétation du texte considère que les informations individuelles en question sont celles qui permettent d'identifier au moins une personne. Cela signifie que personne ne doit pouvoir identifier des individus par les informations qu'elle possède et les renseignements individuels qui lui seront fournis. Il s'agira donc de protéger l'anonymat des enquêtés à la fois vis à vis des demandeurs extérieurs, mais aussi vis à vis des organismes participant au projet. Si l'information a été collectée par la Drees, l'Insee ne doit pas pouvoir identifier personne. Des informations tirées d'une enquête statistique ne pourront pas être mises à disposition de l'InVS si l'InVS détient par ailleurs des données qui lui permettraient d'identifier les personnes. Les huit partenaires du projet se heurteront donc à des problèmes pour partager les données collectées par le canal de la loi de 1951. Mais les moyens pour y parer existent sans doute, et ont peut-être déjà été envisagés.

### **Daniel NAULLEAU**

Concernant la distinction privé/public, je pense qu'elle n'a plus beaucoup de sens, l'autonomie des universités étant à l'ordre du jour. Il n'existe plus beaucoup d'équipes de recherche totalement publiques, sauf peut-être en sciences sociales. On peut imaginer que certaines entreprises privées auront intérêt à avoir accès à ce type de données, pas nominatives, pour avoir une meilleure connaissance de certaines populations.

Par ailleurs, j'ai compris que vous transmettez aux chercheurs des morceaux de fichiers. Je ne comprends pas bien comment les équipes pourront mener des études longitudinales avec de tels fichiers. Les identifiants devront leur permettre de recoller les morceaux entre eux.

Enfin, combien de temps comptez-vous conserver les fichiers de données ? Les fichiers déclarés à la Cnil par exemple doivent être détruits dans un laps de temps fonction de la finalité. Ce qui est acceptable et accepté aujourd'hui le sera-t-il encore dans dix ou vingt ans ? Nous ne savons pas quel type de pouvoir sera en place. Je rappelle que le fichier des empreintes génétiques (FNAEG) qui concernait uniquement les criminels sexuels à l'origine a été étendu à l'ensemble de la population fréquentant les prisons, et dans certains cas les commissariats. La création d'un fichier comme Elfe n'est-elle pas une prise de risque pour l'avenir, même s'il ne concerne qu'une petite partie de la population ?

### **Jeanne BOSSI**

Les données sont en effet toujours collectées pour une durée déterminée, à l'issue de laquelle elles sont détruites ou totalement anonymisées. Il existe dans certains fichiers rassemblant des millions de personnes en France des dispositions qui prévoient la destruction physique du fichier en cas d'atteinte au territoire ou de survenue d'un régime qui ne respecterait pas les libertés individuelles. Nous avons tendance à ne plus inscrire ce genre de dispositions, parce que la destruction est facilement réalisable.

Il convient au départ, comme l'expliquait Jean-Paul Amann, d'évaluer les risques. Une cohorte en comporte plusieurs, mais ces risques doivent être mis en balance avec les bénéfices attendus pour la santé publique et la population en général. L'exemple du FNAEG est typique du développement de fichiers construits à l'origine pour des finalités très encadrées et particulières, et dont le champ d'application est élargi au fil du temps par la loi. On revient rarement en arrière par la suite. Ceci relève d'un choix de société ; il convient de sensibiliser la population, donc le législateur, à ce type de dérive possible.

### **Denis DURAND**

Je souhaite ajouter que la CGT n'est pas favorable à l'introduction dans le projet Elfe de financements privés, ni à l'utilisation des données à des fins privées.

### **Henri LERIDON**

Je conviens que tous les systèmes informatiques sont faillibles, mais nous n'avons pas d'autre choix. Certaines données sont même directement collectées de façon informatisée, ce qui représente un risque supplémentaire par rapport au questionnaire papier. Nous devons trouver des systèmes de protection et de duplication qui minimisent ce risque. Concernant les bibliothèques, nous avons l'intention d'en choisir deux pour minimiser les risques de perte des prélèvements. Je retiens en tout cas l'idée d'une évaluation du système informatique, et nous passerons peut-être par des audits s'il le faut.

Sur la question des personnes se retirant en cours de route, elles peuvent le faire sans nous en informer, en cessant simplement de répondre, ou au contraire exprimer clairement leur volonté de se retirer du projet. Doit-on alors supprimer les données collectées antérieurement ? Le point a été discuté récemment en France dans un colloque international réunissant des spécialistes de grandes cohortes. Dans la mesure où le consentement avait été donné pour les étapes précédentes, et où la suppression de données fausserait les chiffres sur lesquels auraient déjà travaillé des chercheurs, nous sommes plutôt portés à penser que nous n'effacerons pas les données déjà collectées les concernant. Quant aux apports de données externes, la Cnil considérera sans doute que tout dépendra des conditions initiales de consentement. Peut-être cela dépendra-t-il également des conditions de retrait du consentement, délibéré ou par perte de vue. Ces points méritent d'être encore débattus.

Sur l'objection évoquant la loi de 1951, nous avons me semble-t-il déjà répondu en expliquant que tout fichier, quelle que soit la source, sera transmis avec un identifiant ne signifiant rien pour personne, et différent de celui des autres fichiers. Mais tout ceci doit être examiné dans des discussions plus approfondies.

## VII. **Projet d'avis**

La formation propose le [projet d'avis](#) ci-après, qui sera soumis à la prochaine Assemblée plénière du Cnis.

### **Claire BAZY MALAURIE**

Je vous remercie pour votre présence, votre attention et la qualité des débats que nous avons eus. Je pense que les promoteurs du projet Elfe sont eux aussi satisfaits, car ils savent qu'ils peuvent continuer à avancer.

*La séance est levée à 17h55.*

---

### **Adoption du compte rendu**

Le présent compte rendu sera modifié en fonction des observations que vous voudrez bien faire parvenir au Secrétariat général du Cnis avant le 1<sup>er</sup> septembre 2007 ( [cnis@insee.fr](mailto:cnis@insee.fr) - 01 41 17 52 62). Ces modifications seront prises en compte sur le site du Cnis.

---

PROJET D'AVIS

---

proposé par l'interformation Statistiques sociales le 6 juin 2007  
*sera soumis à l'approbation de l'Assemblée plénière du Cnis en fin d'année 2007*

Le Conseil se félicite de la présentation détaillée de **l'état d'avancement du projet d'étude longitudinale française depuis l'enfance (Elfe)** devant l'inter formation « statistiques sociales » du Cnis. Il confirme son soutien à ce projet qui permettra d'étudier les interactions entre les parcours familiaux, sociaux, scolaires, sanitaires, le contexte environnemental et le devenir des enfants. Il confirme l'avis d'opportunité des deux premières enquêtes : l'enquête en maternité et l'enquête 6-8 semaines à réaliser en 2009.

Le Conseil a noté que la loi informatique et liberté ne s'oppose pas par principe à une cohorte de ce type. Il prend acte de la demande de la Cnil d'être associée à l'évolution des réflexions tout particulièrement en ce qui concerne la confidentialité et la sécurité des données ainsi que l'accès aux fichiers.

Conscient de l'importance de ce projet et des enjeux sociétaux qu'il recouvre, le Conseil demande à être informé du déroulement du projet en mettant l'accent sur les solutions proposées sur chacun des points qualifiés d'essentiels comme l'horizon temporel du panel, les conditions de respect des droits des personnes et des informations les concernant, les droits d'accès aux données.

La présentation a mis en évidence les impératifs qu'Elfe doit respecter notamment en matière de rigueur et de transparence. Le Conseil recommande que tous les moyens soient mis en œuvre pour informer largement la population de ce projet très en amont de sa réalisation. Il insiste sur l'importance de sélectionner en priorité les études pour lesquelles la dimension longitudinale est indispensable.

---

DOCUMENT PRÉPARATOIRE

---

- [Note de présentation du projet](http://www.cnis.fr/Agenda/DPR/DPR_0539.PDF) disponible sur le site cnis.fr  
([http://www.cnis.fr/Agenda/DPR/DPR\\_0539.PDF](http://www.cnis.fr/Agenda/DPR/DPR_0539.PDF))