

Direction de la recherche, des études,
de l'évaluation et des statistiques
DREES

SERIE
STATISTIQUES

**DOCUMENT
DE
TRAVAIL**

Échantillon interrégimes de cotisants (EIC)
Tome 1 :
Procédures statistiques de constitution de l'EIC

Laurent Caillot, Hélène Chaput,
Christel Colin, Najat El Mekkaoui-de Freitas, Hélène Michaudon.

n° 50 – mars 2003

MINISTERE DES AFFAIRES SOCIALES, DU TRAVAIL ET DE LA SOLIDARITE
MINISTERE DE LA SANTE, DE LA FAMILLE ET DES PERSONNES HANDICAPEES

SOMMAIRE

PRÉAMBULE	5
1. LES OBJECTIFS DE L'EIC	7
2. CALENDRIER ET MÉTHODES MIS EN ŒUVRE POUR LA RÉALISATION DE L'EIC	12
3. MOYENS MIS EN ŒUVRE POUR LA RÉALISATION DE L'EIC	15
4. LE MODE DE COLLECTE	16
4.1 LES PRINCIPES D'ORGANISATION DE L'ENQUÊTE (TECHNIQUE DU DOUBLE AVEUGLE)	16
4.2 SÉLECTION DE L'ÉCHANTILLON.....	19
4.3 CHAMP ET PÉRIODE DE RÉFÉRENCE	22
4.4 CIRCUITS DE COLLECTE.....	23
4.4.1 ... pour le régime général, les régimes complémentaires obligatoires, les régimes alignés, les régimes des professions libérales, les régimes spéciaux de salariés	23
4.4.2 ...pour la Fonction publique.....	26
4.4.3 Appariement complémentaire : les fichiers de l'UNEDIC	27
5. INFORMATIONS SOLLICITÉES	27
5.1 NATURE DE L'INFORMATION	27
5.2 CARACTÉRISTIQUES GÉNÉRALES DES FICHIERS TRANSMIS PAR LES RÉGIMES	29
6. PERSPECTIVES	29
7. ANNEXES	30
7.1 GLOSSAIRE.....	30
7.2 LISTE DES ORGANISMES GESTIONNAIRES DE RÉGIMES DE RETRAITE OBLIGATOIRE PARTICIPANTS À L'OPÉRATION EIC	31
7.3 AUTRES ORGANISMES PARTICIPANTS À L'OPÉRATION EIC :	32
7.4 COMPTES-RENDUS DES RÉUNIONS BILATÉRALES AVEC LES DIFFÉRENTS PARTENAIRES : ERREUR! SIGNET NON DÉFINI.	
<i>Compte-rendu de la réunion bilatérale DREES-Ministère de la Défense</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-régime des Mines.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-CAVIMAC.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-CRPNAC.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-SEITA (ALTADIS).....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-CRPCEN.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-ENIM</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES- IEG Pensions</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion DREES-UNEDIC.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-CNAVPL.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-CNAVTS.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion DREES- Agents des collectivités locales et des hôpitaux</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-Banque de France.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-ORGANIC.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-CCMSA.....</i>	<i>Erreur! Signet non défini.</i>
<i>Compte-rendu de la réunion bilatérale DREES-CANCAVA</i>	<i>Erreur! Signet non défini.</i>

Préambule

Le système statistique français souffre, à l'heure actuelle, de l'absence d'un outil permettant de connaître précisément les droits acquis en matière de retraite pour les générations d'âge actif, de simuler les retraites futures et d'analyser leurs effets sur l'équilibre financier du système de retraite. C'est la vocation de l'Échantillon interrégimes de cotisants (EIC) d'apporter ces éclairages.

La publication de trois documents de travail sur l'EIC sont prévus. Le premier, présenté ici, porte sur les procédures statistiques de constitution de l'échantillon. Le second traitera du cadre juridique de l'opération. Le dernier sera un guide méthodologique pour l'utilisation de l'EIC.

La collecte de l'EIC fait suite à une phase de préparation technique incluant des réunions du comité de pilotage et une série de réunions bilatérales avec l'ensemble des régimes de retraites qui participeront à la première édition de l'EIC prévue en 2003, et les autres organismes partenaires de l'opération¹.

Pour cette phase, la DREES a bénéficié du concours de Monsieur Yves Ullmo, conseiller-maître honoraire à la Cour des Comptes.

L'élaboration du contenu de l'EIC s'est donc faite en étroite concertation avec les différents partenaires que nous remercions vivement pour leur collaboration.

Les réunions avec les différents régimes partenaires de l'opération ont donné lieu à des comptes rendus (voir annexe 7.4).

¹ Voir en annexe 7.2 la liste des organismes participants à l'opération.

1. Les objectifs de l'EIC

Afin d'étudier les retraites futures des personnes d'âge actif, la DREES met en œuvre un outil statistique permettant de reconstituer les droits à pension acquis en matière de retraite des cotisants aux différents régimes. La création d'un tel outil (échantillon interrégimes de cotisants) figure dans la loi de financement de la Sécurité sociale pour 2001 (article 27 de la loi n° 2000-1257 du 23 décembre 2000 de financement de la Sécurité sociale pour 2001) modifiée par l'article 62 de la loi n° 2002-1487 du 20 décembre 2002 de financement de la Sécurité sociale pour 2003.

Actuellement, les rares analyses sur les retraites futures, ou sur l'âge auquel les actifs d'aujourd'hui pourront liquider leurs retraites, se basent essentiellement sur des données d'enquêtes statistiques, éventuellement prolongées par des techniques de simulation. On peut, à cet égard, citer l'exemple de l'étude « Parcours professionnels et retraite : à quel âge partiront les actifs d'aujourd'hui ? » (*France, portrait social*, 1999) menée à partir de l'enquête Patrimoine (1998) de l'INSEE, ou des études réalisées à partir du modèle de microsimulation dynamique Destinie de l'INSEE (« Évaluation de trois réformes du régime général à l'aide du modèle de microsimulation Destinie », document de travail de la DESE G2002/07).

Or les données utilisées ne répondent qu'imparfaitement aux besoins : d'une part, les calendriers d'activité reconstitués font appel à la mémoire des personnes interrogées et distinguent mal les années travaillées et les années validées ; d'autre part, ces données ne donnent aucune indication sur les revenus perçus au cours de la carrière et donc sur les droits acquis. En ce sens, la précision des données administratives fait défaut. Des données administratives, issues des fichiers de salaires portés aux comptes de la Caisse nationale d'assurance vieillesse, ont, quant à elles, été utilisées pour un exercice de projection des retraites du régime général à l'horizon 2015, mais celui-ci se limite aux retraites versées par la CNAV (« La retraite des salariés du secteur privé en 2015 », *Retraite et société*, n° 20, 1997).

Dans une optique d'éclairage des choix publics, la combinaison de l'exigence de précision des données administratives et d'une vision inter-régimes conduit à mettre en place un Échantillon interrégimes de cotisants (EIC), donnant, avec une périodicité régulière, une photographie de la situation des droits acquis par un échantillon d'actifs de différentes générations, dans l'ensemble des régimes de retraite où ils ont cotisé. Cet échantillon sera le symétrique de l'Échantillon interrégimes de retraités (EIR) construit par la DREES tous les quatre ans (encadré 1), et qui est le seul outil permettant une connaissance globale des pensions des individus à la retraite à une date donnée, la pension d'un individu étant très souvent la somme de plusieurs retraites versées par différents régimes.

L'échantillon interrégimes de cotisants servira alors de base à des projections sur les niveaux de retraites futures et les âges de liquidation des droits, selon différentes hypothèses sur les fins de carrière, dépendant par exemple des taux d'activité, du taux de chômage, du taux de croissance des salaires, du développement du travail à temps partiel, etc.

De manière plus précise, l'EIC doit permettre :

- d'estimer les montants des pensions (y compris les minima de pension comme le minimum contributif) à servir par les différents régimes de retraite, qu'ils s'adressent aux

salariés du secteur privé, de l'agriculture, ou des régimes spéciaux, aux fonctionnaires, aux professions libérales, aux commerçants, artisans ou exploitants agricoles, d'où la nécessité de prendre en compte dans l'échantillon le maximum de régimes de retraite de base et complémentaires obligatoires ;

- d'estimer l'âge futur de départ à la retraite (au sens de liquidation des droits).

Or les pensions de retraite sont souvent la somme de différents éléments. Dans l'idéal, afin d'avoir une vision complète des retraites futures, il faudrait pouvoir estimer :

- le montant des avantages principaux de droit direct (avantages acquis en contrepartie de la carrière, pour les années travaillées et les années validées) ;
- le montant des avantages accessoires, complémentaires à l'avantage principal : bonification pour trois enfants et plus, majoration pour conjoint à charge, majoration pour aide constante d'une tierce personne, allocation supplémentaire du minimum vieillesse, etc. ;
- le montant des pensions de réversion.

Les informations à mobiliser pour ces calculs sont toutefois de différents niveaux. Ainsi, pour déterminer si une personne aurait droit ou non au minimum vieillesse, des informations au niveau du ménage (ensemble des revenus, situation familiale) sont nécessaires. De même, le calcul des pensions de réversion nécessite par exemple de connaître les droits acquis par le conjoint, ou la survenance d'un décès au sein du couple.

Les rencontres bilatérales avec les différents régimes ont montré que seuls quelques régimes, essentiellement des régimes spéciaux qui gèrent parfois également les prestations familiales de leurs salariés, disposent d'une information sur le nombre d'enfants et sur la situation matrimoniale (parfois sur une base déclarative). Sauf exception, les régimes ne disposent pas, par ailleurs, de données sur les revenus du conjoint qui pourraient servir de base à des estimations des pensions de réversion.

En conséquence, dans le cadre de l'EIC, sera collectée de l'information sur les droits à la retraite acquis à titre individuel. En ce qui concerne les avantages familiaux, les données nécessaires à leur calcul seront recueillies auprès des quelques régimes qui en disposent. Pour les autres, le calcul des avantages familiaux se fera sur la base d'imputations du nombre d'enfants en fonction des caractéristiques individuelles et des données démographiques disponibles par ailleurs, par exemple dans les enquêtes de l'INSEE.

Pour connaître les droits à la retraite acquis par les différentes générations, une bonne connaissance tant des durées validées dans les différents régimes que des rémunérations perçues est nécessaire. Sur ce plan, l'EIC apportera des informations quantitatives originales sur les périodes validées pour la retraite, y compris celles dites « assimilées », validées lors de périodes où les individus n'ont pas travaillé. Ainsi, grâce aux échanges d'informations entre organismes d'assurance vieillesse et organismes d'assurance maladie, d'assurance chômage et caisse nationale d'allocations familiales, l'EIC doit en particulier permettre de quantifier pour les différentes générations les périodes validées au titre :

- du chômage et de la préretraite (ainsi que des périodes de formation ou reconversion),
- de la maladie, de la maternité, de l'invalidité, des accidents du travail,

- du service militaire, des périodes de guerre (en partie du moins, ces informations ne sont parfois disponibles qu'à la liquidation des droits),
- de l'assurance vieillesse des parents au foyer.

La connaissance des carrières en sera améliorée dans la mesure où, par rapport aux panels de déclarations de salaires existants, on pourra, dans de nombreux cas, déterminer la nature des périodes où les individus n'ont pas été rémunérés, et si ces périodes ouvrent des droits à la retraite. Des zones d'ombre subsisteront néanmoins. Ainsi, les périodes de chômage non indemnisé ne seront que partiellement repérées, celles de perception du Revenu minimum d'insertion (qui n'ouvre pas de droits à la retraite) ne pourront être qualifiées, ainsi que, par exemple, certaines périodes de travail à l'étranger.

Enfin, par rapport aux projections disponibles jusqu'à présent sur les retraites, celles réalisées à partir de l'EIC seront plus précises, dans la mesure où elles pourront tenir compte de la diversité des situations individuelles et pas seulement des évolutions moyennes. Un des principaux intérêts de l'EIC est, en effet, de pouvoir, partant d'informations individuelles, connaître les carrières cohorte par cohorte, en distinguant selon le sexe, selon le secteur d'emploi, etc. Ceci corrigerait les faiblesses soulignées, en 1999, par la Cour des Comptes dans son rapport annuel sur la Sécurité sociale : « la multiplicité des régimes et la profusion des règles constituent des limites aux exercices globaux de projection, qui ne peuvent pas rendre compte des disparités entre les retraités. Les réformes nécessaires ne peuvent pourtant pas faire l'impasse sur cette structure et cette diversité ». Ces projections permettront un meilleur pilotage des régimes, y compris par eux-mêmes.

La constitution de cet instrument statistique nécessite donc la collecte de nombreuses informations sur les carrières auprès de l'ensemble des caisses de retraite, pour un échantillon représentatif des personnes d'âge actif au sein des différents secteurs socio-professionnels. L'objectif est de disposer ainsi d'un système d'information aussi complet que possible sur les cotisants.

Il faut souligner qu'à l'issue de la phase de collecte proprement dite, seules les informations brutes sur les droits acquis fin 2001 par différentes générations seront disponibles. L'analyse des futurs droits à la retraite et des âges probables de liquidation des droits n'interviendra qu'ensuite, et nécessitera une phase d'études et de simulations à partir des données de l'EIC.

Encadré 1 - L'Échantillon interrégimes de retraités (EIR)

L'Échantillon interrégimes de retraités (EIR) a pour objectif de reconstituer le montant de la retraite globale des individus, ainsi que ses éléments constitutifs. L'opération consiste à interroger les organismes de retraites sur les caractéristiques d'un échantillon de retraités : nature et montant des prestations versées, conditions de liquidation des droits à la retraite (âge de liquidation, taux de liquidation, durée de carrière validée, par exemple). Le rapprochement, individu par individu, des montants en provenance des différents régimes permet notamment de reconstituer la pension globale de chaque retraité.

La quasi-totalité¹ des organismes de retraites obligatoires (régimes de base, régimes complémentaires obligatoires et régimes spéciaux) sont interrogés. Les retraites issues de régimes complémentaires facultatifs sont exclues du champ de l'enquête², ainsi que les revenus provenant de l'épargne individuelle volontaire.

L'EIR 2001 est la quatrième enquête menée : la première a eu lieu en 1988 ; l'expérience a ensuite été reconduite en 1993 et en 1997.

L'Échantillon interrégimes se présente sous la forme d'un panel. Les personnes appartenant à l'échantillon initial sont sélectionnées à nouveau à chaque enquête (exceptés les individus décédés entre deux vagues). Disposer d'un panel permet notamment d'étudier de manière détaillée l'évolution des retraites entre deux vagues : il est ainsi possible de distinguer ce qui dans cette évolution relève d'une variation nette des pensions (revalorisations affectant les montants perçus par les retraités déjà présents dans l'échantillon), et ce qui correspond à un effet « noria », de renouvellement de la population. Une partie de l'évolution globale des pensions est en effet due à l'arrivée à l'âge de la retraite de nouvelles générations et au décès des individus appartenant aux générations les plus anciennes. Actuellement, l'effet de génération joue positivement sur le montant des pensions : plus la génération est récente, plus les individus ont effectué leurs carrières dans des contextes favorables, et ont bénéficié de conditions de liquidation meilleures que leurs aînés.

L'échantillon a été enrichi à chaque vague, pour mieux prendre en compte la population des retraités et répondre aux besoins d'information des acteurs sociaux. Les échantillons de 1988 et 1993 ont été conçus de manière à observer la population des retraités de 65 ans et plus nés en France métropolitaine. En 1988, lors de la première opération, seules quatre générations de retraités ont été sélectionnées (1906, 1912, 1918, 1922). En 1993, une génération a été ajoutée, celle de 1926, suite au vieillissement de l'échantillon.

L'échantillon de 1997 marque une rupture par comparaison aux deux enquêtes précédentes : il a été étendu à l'ensemble des retraités de 55 ans et plus et concerne désormais douze générations, toujours nées en France métropolitaine. En 1997, les six générations les plus anciennes représentent les personnes âgées de plus de 65 ans, ce qui correspond au champ des EIR 1988 et 1993 (la génération 1930 a simplement été ajoutée aux générations précédemment enquêtées). Les six générations les plus jeunes (1932, 1934, 1938, 1940, 1942) représentent les retraités de 55 à 64 ans. Ce rajeunissement avait notamment pour objectif de mieux connaître la situation des « jeunes retraités », qui liquident leurs droits plus tôt que les autres, ainsi que d'étudier le cumul emploi-retraite à partir de 55 ans. Le nombre de jours de tirage dans les générations déjà présentes et dans les nouvelles générations a en outre été augmenté, afin d'avoir un échantillon suffisamment large pour réaliser des études régionales. Cette optique a été poursuivie lors de la constitution de l'EIR 2001 : l'échantillon est conçu pour représenter les personnes âgées de 55 ans et plus.

¹ Le régime complémentaire de l'ORGANIC, facultatif, constitue une exception à ce principe. Il est en effet inclus dans l'EIR.

² Seuls quelques régimes ne sont pas pris en compte, ayant peu d'adhérents : notamment la Caisse nationale des barreaux français (avocats), la Chambre de commerce et d'industrie de Paris, l'Opéra de Paris et le Port autonome de Strasbourg.

2. Calendrier et méthodes mis en œuvre pour la réalisation de l'EIC

L'idée initiale évoquée en juin 1999, de l'élaboration d'un outil statistique sur les cotisants, qui serait le pendant de l'échantillon interrégimes de retraités pour les générations encore d'âge actif était d'expérimenter la constitution d'un tel échantillon en commençant par la CNAV, l'ARRCO et l'AGIRC. Finalement, les réflexions ultérieures ont plutôt conduit à inclure le maximum de régimes, afin d'avoir la vision la plus complète possible des droits acquis à la retraite.

Dès février 2000, dans la perspective d'une possible loi portant diverses dispositions d'ordre social, une première version d'un article de loi « créant un système d'informations sur les droits acquis à la retraite, basé sur l'exploitation d'un échantillon statistique anonyme et représentatif de personnes ayant acquis des droits à la retraite » a été rédigée. Compte tenu du calendrier législatif, c'est finalement dans la loi de financement de la Sécurité sociale pour 2001, votée en décembre 2000, que l'article législatif créant l'EIC trouvera sa place, ainsi qu'un article créant le répertoire national des retraites et des pensions.

Une réunion de lancement rassemblant l'ensemble des partenaires potentiels (fournisseurs de données et utilisateurs) a eu lieu en avril 2001. Un comité de pilotage, comprenant l'INSEE, la Cour des Comptes, la Direction de la prévision, le COR², la CNAV, l'ARRCO, l'AGIRC, la CDC, la Banque de France, la CCMSA, la CANCAVA, la SNCF et la DSS a été monté. Il s'est réuni pour la première fois en septembre 2001 : la DREES a présenté les premières conclusions du groupe de travail « Fonction publique » (cf. infra), et les premières réunions bilatérales avec les régimes (cf. infra), et le groupe s'est exprimé sur les objectifs du projet. Le comité de pilotage s'est ensuite réuni en janvier 2002 : l'état d'avancement du projet a été présenté, ainsi qu'un calendrier prévisionnel. Enfin, en septembre 2002, une réunion de lancement de la phase de constitution de l'EIC a regroupé l'ensemble des partenaires et des membres du comité de pilotage : la DREES a présenté le cahier des charges et le volet juridique de l'opération.

Dans l'intervalle se sont tenues vingt réunions bilatérales (encadré 2), entre août 2001 et juin 2002, pour lesquelles la DREES a bénéficié du concours de Monsieur Yves Ullmo, conseiller maître honoraire à la Cour des Comptes. L'objectif initial de ces réunions était de présenter le projet aux partenaires et de faire le point sur les données disponibles, sur la base d'une liste indicative élaborée par la DREES en mai 2001. Il est vite apparu que ces réunions étaient également l'occasion de réunir un matériau très riche sur les spécificités des régimes et leur fonctionnement, ainsi qu'une documentation éditée par les caisses. La méthode de travail adoptée a été la rédaction rapide d'un compte-rendu après chaque réunion, et sa validation par les participants avant diffusion (annexe 6.4), de manière à disposer d'un support à l'élaboration ultérieure du cahier des charges et des dessins de fichiers. Ces comptes rendus ont servi à la rédaction, en janvier 2002, d'une note intitulée « synthèse relative à la disponibilité des données pour l'Échantillon interrégimes de cotisants », présentée lors de la deuxième réunion du comité de pilotage.

² Pour obtenir la signification de ces sigles, voir le glossaire en annexe 7.1.

Encadré 2 - Calendrier de l'ensemble des réunions

- Réunion de lancement : 25 avril 2001
- GT Fonction publique : 25 juin 2001
- CCMSA : 29 août 2001
- ORGANIC : 5 septembre 2001
- Comité de pilotage n°1 : 10 septembre 2001
- CNAV : 19 septembre 2001
- CANCAVA : 24 septembre 2001
- AGIRC : 26 septembre 2001
- BDF : 26 septembre 2001
- ARRCO : 3 octobre 2001
- CNRACL : 8 octobre 2001
- IRCANTEC : 7 novembre 2001
- CNAVPL : 6 novembre 2001
- SNCF : 12 novembre 2001
- DGCL : 22 novembre 2001
- CANSSM : 26 novembre 2001
- DHOS : 29 novembre 2001
- RATP : 10 décembre 2001
- INSEE : 10 décembre 2001
- CIG petite couronne : 17 décembre 2001
- IEG-Pensions : 20 décembre 2001
- UNEDIC : 21 janvier 2002
- Comité de pilotage n°2 : 28 janvier 2002
- CRPCEN : 5 juin 2002
- ENIM : 10 juin 2002
- CAVIMAC : 13 juin 2002
- CRPN : 14 juin 2002
- SEITA : 28 juin 2002
- INSEE (aspects juridiques) : 2 juillet 2002
- ARRCO-AGIRC (projet de dessin de fichier) : 11 juillet 2002
- Réunion plénière de lancement : 12 septembre 2002
- CICSS : 12 septembre 2002
- DGCP : 26 septembre 2002
- OSD : 11 octobre 2002
- INSEE et DGCP : 16 octobre 2002
- CNIL : 8 novembre 2002
- OSD : 14 janvier 2003

Parallèlement aux réunions bilatérales, un groupe de travail spécifique au champ de la Fonction publique a été monté. En effet, il n'existe pas de données sur les carrières au sein des organismes qui gèrent la retraite des fonctionnaires, tant au niveau de la Fonction publique d'État que de la Fonction publique territoriale. La mission du groupe de travail « Fonction publique » était de voir dans quelle mesure le recours à d'autres producteurs de données pouvait permettre de remédier à cette absence. Ce groupe comprenait, outre la DREES, la Cour des Comptes, la CDC – Branche retraite, l'IRCANTEC, l'INSEE, le Service des pensions de l'État, la DGAFF, la DGCP, le COR. Il s'est réuni dans sa forme complète en juillet 2001. À l'issue de cette réunion, différentes pistes de réflexion ont été proposées ; elles

ont été explorées, au cours des réunions bilatérales avec la CNRACL, la DGCL, la DHOS, et le CIG Petite Couronne et elle ont été présentées, en janvier 2002, au comité de pilotage, à partir d'une note intitulée « Échantillon interrégimes de cotisants - point sur le recueil des informations concernant les agents de la Fonction publique ». Il a été établi que le recours à la base des cotisants actuels et passés de la CNRACL, aux fichiers de paye de la Comptabilité publique, au panel des Déclarations annuelles de données sociales (DADS) et aux données du ministère de la Défense permettait de réunir une grande partie des informations souhaitées. Demeurent manquantes la reconstitution des débuts de carrière dans certains cas, et les périodes dues à certains types de « trous de carrière » (congé parental, départ à l'étranger, congé formation, création d'entreprise...). Il est également apparu que l'interrogation directe des organismes employeurs (ministères, collectivités locales, hôpitaux) sur ces données manquantes serait lourde et coûteuse. Il a donc été décidé de privilégier les techniques habituelles de « modélisation sous hypothèses » et d'imputation, en se calant sur des données provenant d'autres enquêtes. Une enquête complémentaire auprès des cotisants eux-mêmes, qui présenterait l'avantage d'apporter des informations sur d'autres thèmes (situation familiale et patrimoniale par exemple), reste néanmoins envisageable ; cette possibilité est d'ailleurs mentionnée dans le dossier CNIL.

Le volet juridique de l'opération a comporté plusieurs étapes :

- la sécurisation de la base légale : un article législatif a été passé dans la Loi de financement de la Sécurité sociale (LFSS) pour 2003 afin de compléter la disposition créant l'EIC dans la LFSS pour 2001 en incluant l'UNEDIC dans le champ des organismes partenaires ;
- la préparation du décret en Conseil d'État relatif à la fois à l'EIC et à l'EIR et contresigné par quatre membres du gouvernement (le ministre des affaires sociales, du travail et de la solidarité, le ministre de l'économie, des finances et de l'industrie, le ministre délégué au budget et la ministre de la défense). Le décret s'est inspiré du décret n°2002-240 du 20 février 2002 relatif à l'Échantillon national interrégimes d'allocataires de minima sociaux (ENIAMS), élaboré par le bureau « lutte contre l'exclusion » à la DREES. Cette architecture juridique commune pour l'EIC et l'EIR a permis de « régulariser » la base juridique de l'EIR, auparavant régi par un arrêté, et d'articuler les deux échantillons en un même « traitement automatisé de données individuelles anonymisées à des fins statistiques en matière de retraite ». Le décret a été présenté au Comité interministériel de coordination en matière de Sécurité sociale (CICSS), animé par la Direction de la Sécurité sociale (DSS), puis soumis à l'avis de la CNIL et enfin à celui du Conseil d'État ;
- la préparation de l'arrêté relatif à l'EIC, qui a également bénéficié du précédent de l'ENIAMS. Le projet d'arrêté a été présenté devant le CICSS, soumis à l'avis de la CNIL puis signé par les quatre ministres co-signataires du décret commun à l'EIC et à l'EIR.

Par ailleurs, deux protocoles bilatéraux de coopération ont été préparés, d'une part avec l'INSEE, d'autre part avec la DGCP.

3. Moyens mis en œuvre pour la réalisation de l'EIC

Comme toute opération statistique de consolidation de données issues de nombreux fournisseurs, la réalisation de l'EIC mobilise des ressources humaines importantes et des compétences variées, de nature statistique, mais aussi économique et juridique.

Tableau 1 - Récapitulatif des moyens mis en œuvre

Étape	Moyens
Préparation statistique et juridique de l'EIC	
1. Genèse de l'EIC (1999-2000)	Chef du bureau BPHVD, conseillère scientifique
2. Concertation avec les partenaires (2001-2002)	Chef du bureau BPVHD, adjointe, conseillère scientifique, avec le concours d'un conseiller-maître honoraire de la Cour des Comptes
3. Élaboration du cahier des charges (2002)	Chef du bureau BPVHD, attachée INSEE, conseillère scientifique, avec le concours du chargé de mission auprès du sous-directeur et de l'adjointe
4. Textes réglementaires et conventions (2002-2003)	Chargé de mission auprès du sous-directeur avec concours de la chef du bureau BPVHD, de l'adjointe, d'une attaché INSEE, de la conseillère scientifique et de la mission coordination des programmes
Constitution technique et exploitations	
5. Constitution et apurement de l'EIC 2003 (2003)	Création d'un poste d'attaché INSEE et redéploiement d'un demi-poste pour gérer et exploiter le traitement
6. Édition de résultats statistiques (2004-...)	À définir
7. Modélisation (2004-...)	À définir

Compte tenu de la mise en place de l'EIC, qui a vocation à être actualisé tous les 4 ans en alternance et en articulation avec l'EIR, l'investissement initial, c'est-à-dire le coût fixe de la création du panel, a été particulièrement lourd, d'autant plus que, sur le plan juridique, il y a également eu consolidation du régime de l'EIR qui forme désormais, conjointement avec l'EIC, un « traitement automatisé permanent de données individuelles à des fins statistiques en matière de retraite ».

Cette phase préparatoire a nécessité la constitution d'une équipe pluridisciplinaire de 5 à 6 personnes de catégorie A en charge de l'EIC au sein de la sous-direction de l'observation de la solidarité : pour le bureau « Politique de la vieillesse, du handicap et de la dépendance » (PVHD), la chef de bureau et son adjointe, une puis deux attachés de l'INSEE (le second poste étant créé à partir de 2003) et une conseillère scientifique ; également, pour la partie juridico-administrative, le chargé de mission auprès du sous-directeur (administrateur civil), qui a notamment rédigé les textes juridiques, le secrétariat du bureau PVHD dont le rôle est central étant donné le nombre de partenaires (annexe) et de réunions (encadré 2), des personnes «ressources» de catégorie A : un conseiller-maître honoraire à la Cour des Comptes, le sous-directeur de l'observation de la solidarité et la Mission « Coordination des programmes » en les personnes de la chef de mission et d'un attaché.

4. Le mode de collecte

4.1 Les principes d'organisation de l'enquête (technique du double aveugle)

Dans son principe, l'EIC consiste à recueillir auprès de l'ensemble des régimes de retraite³ ainsi que de l'INSEE, la Comptabilité Publique, l'Unédic et le ministère de la Défense des données permettant le calcul des droits à pension acquis pour un échantillon de cotisants. Le rapprochement, individu par individu, des informations provenant des différentes sources permettra de reconstituer l'ensemble des droits acquis par chaque cotisant.

Pour respecter les obligations de la loi « Informatique et libertés », la collecte des informations comporte trois étapes et utilise la « technique du double aveugle », qui permet de garantir l'anonymat des données collectées (schéma 1).

Première étape :

La DREES détermine les règles de sélection de l'échantillon (voir point 2.2). Cette sélection est effectuée sur la base des jours et années de naissance. L'INSEE tire dans le Répertoire national d'identification des personnes physiques⁴ (RNIPP) un échantillon de personnes correspondant aux dates de naissance désignées et affecte à chaque individu un numéro d'ordre anonyme. La liste des personnes établie par l'INSEE comprend, en sus de ce numéro d'ordre anonyme, le Numéro d'identification au répertoire (NIR), le nom patronyme, le (ou les) prénom(s), le sexe, le jour, le mois, l'année, la commune et le pays de naissance. Cette liste est envoyée par l'INSEE aux différents organismes d'assurance vieillesse et aux autres partenaires concernés.

Deuxième étape :

À l'aide des éléments d'identification contenus dans la liste de l'INSEE, dont le NIR, les organismes d'assurance vieillesse isolent les individus de l'échantillon relevant de leur régime en tant que cotisant actuel ou passé. Pour les personnes qui cotisent ou ont cotisé à leur régime, ils recueillent les données correspondant au dessin de fichier demandé, puis les transmettent à la DREES en les accompagnant uniquement du numéro d'ordre. Les autres organismes effectuent une extraction de leurs fichiers sur le champ de l'échantillon et procèdent de la même manière. Les éléments d'identification, en particulier le NIR, sont préalablement effacés pour rendre anonymes les données.

Troisième étape :

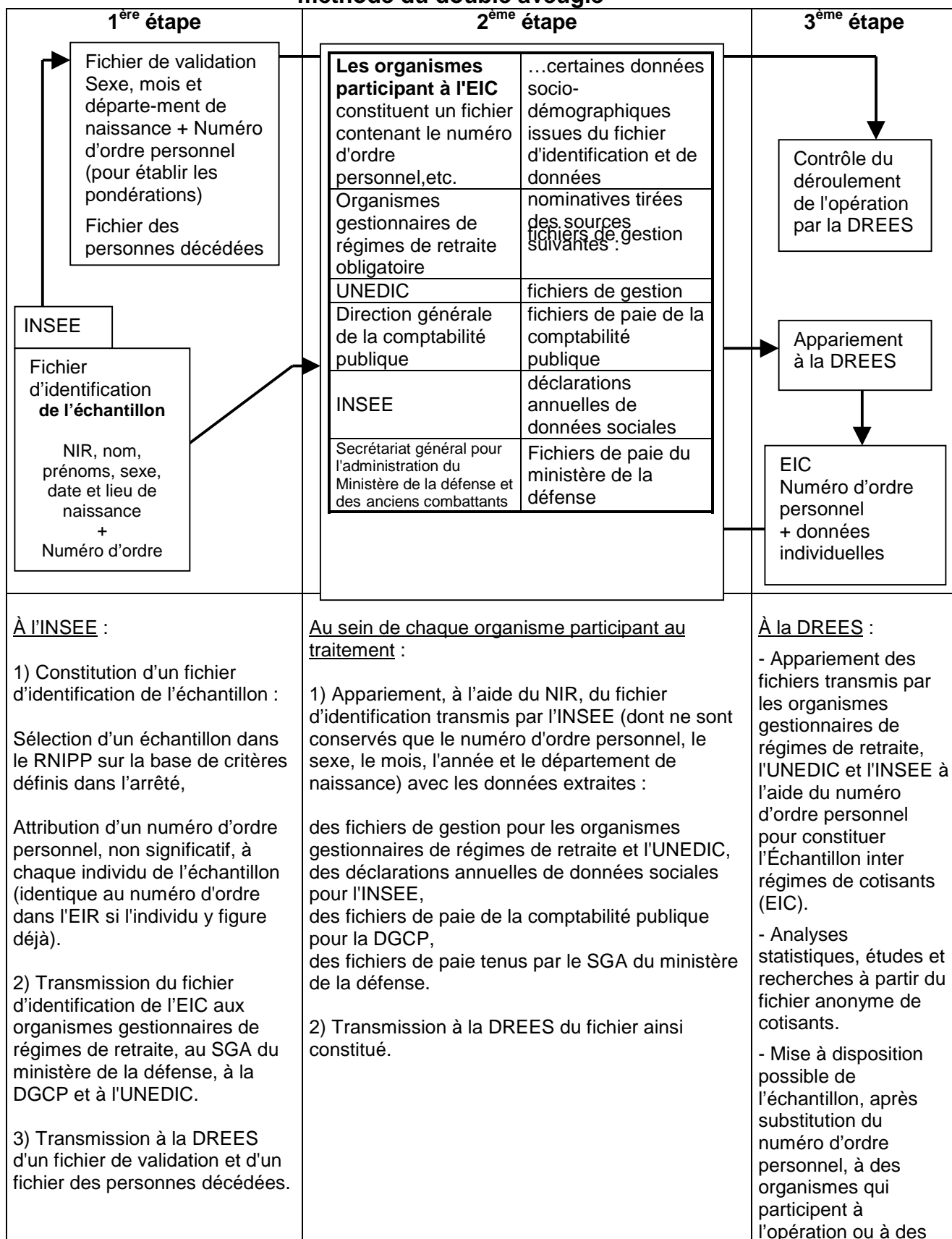
La DREES fusionne, sur le numéro d'ordre anonyme, les informations provenant des différents organismes, puis réalise des exploitations statistiques.

³ Voir en annexe 7.2 la liste des organismes gestionnaires de régimes de retraite obligatoire participants à l'opération.

⁴Le RNIPP comprend toutes les personnes ayant acquis des avantages sociaux sur le territoire français, qu'elles soient nées en métropole, dans les DOM ou à l'étranger. L'INSEE procédera ainsi au tirage sur l'ensemble du RNIPP, y compris la « Section hors métropole », gérée par la CNAV pour le compte de l'INSEE.

La procédure utilisée garantit l'anonymat des informations recueillies. Aucun des partenaires de l'opération ne possède la totalité des informations (NIR, numéro d'ordre, ensemble des données sur les droits acquis). La correspondance avec les identifiants du Répertoire national est détenue isolément par chaque organisme fournisseur de données, qui ne peut en aucun cas la transmettre à la DREES ou à un autre organisme.

Schéma 1 - Constitution de l'échantillon interrégimes de cotisants suivant la méthode du double aveugle



<p><u>À l'INSEE :</u></p> <p>1) Constitution d'un fichier d'identification de l'échantillon :</p> <p>Sélection d'un échantillon dans le RNIPP sur la base de critères définis dans l'arrêté,</p> <p>Attribution d'un numéro d'ordre personnel, non significatif, à chaque individu de l'échantillon (identique au numéro d'ordre dans l'EIR si l'individu y figure déjà).</p> <p>2) Transmission du fichier d'identification de l'EIC aux organismes gestionnaires de régimes de retraite, au SGA du ministère de la défense, à la DGCP et à l'UNEDIC.</p> <p>3) Transmission à la DREES d'un fichier de validation et d'un fichier des personnes décédées.</p>	<p><u>Au sein de chaque organisme participant au traitement :</u></p> <p>1) Appariement, à l'aide du NIR, du fichier d'identification transmis par l'INSEE (dont ne sont conservés que le numéro d'ordre personnel, le sexe, le mois, l'année et le département de naissance) avec les données extraites :</p> <p>des fichiers de gestion pour les organismes gestionnaires de régimes de retraite et l'UNEDIC, des déclarations annuelles de données sociales pour l'INSEE, des fichiers de paie de la comptabilité publique pour la DGCP, des fichiers de paie tenus par le SGA du ministère de la défense.</p> <p>2) Transmission à la DREES du fichier ainsi constitué.</p>	<p><u>À la DREES :</u></p> <ul style="list-style-type: none"> - Appariement des fichiers transmis par les organismes gestionnaires de régimes de retraite, l'UNEDIC et l'INSEE à l'aide du numéro d'ordre personnel pour constituer l'Échantillon inter régimes de cotisants (EIC). - Analyses statistiques, études et recherches à partir du fichier anonyme de cotisants. - Mise à disposition possible de l'échantillon, après substitution du numéro d'ordre personnel, à des organismes qui participent à l'opération ou à des organismes extérieurs, à des fins d'étude statistique et de recherche en matière de retraite.
---	---	--

4.2 Sélection de l'échantillon

L'échantillon EIC 2003 est conçu pour représenter les personnes ayant entre 31 ans et 67 ans au 31 décembre 2001 (entre 30 ans et 66 ans au 1^{er} janvier 2001). Il sera tiré dans le RNIPP (y compris dans la section hors métropole), dans sa version de septembre 2002.

L'EIC sera constitué de personnes nées en octobre d'années paires. En effet, une partie des informations collectées pour l'EIC sera extraite du panel DADS qui couvre ce champ, et de plus, le choix de ces dates est également imposé par l'articulation souhaitée entre EIC et EIR (voir infra).

Le taux de sondage pour l'EIC sera de 1 génération sur 4 et 2,67 % par génération tirée (10 jours sur 30 du mois d'octobre qui compte pour 1/12,5 des naissances d'une année)⁵. Chaque individu de l'échantillon sera ensuite affecté d'une pondération établie sur la base des données démographiques permettant d'extrapoler à l'ensemble des individus de 31 à 67 ans

⁵ Exception faite de la génération 1934 pour l'EIC 2003, voir infra.

au 31 décembre 2001. Les analyses par génération seront toutefois privilégiées dans l'exploitation de l'EIC.

L'EIC sera un panel : les individus tirés dans un EIC seront suivis dans les EIC suivants, afin de pouvoir observer l'évolution des droits acquis par les individus de l'échantillon entre deux EIC successifs (acquisition de nouveaux droits ou mise à jour d'informations au sein des régimes). Les individus qui décéderont entre deux EIC ne seront pas remplacés, les écarts de structure qui pourront en résulter seront corrigés par les pondérations.

D'autre part, la DREES souhaite une articulation entre l'EIC et l'EIR, afin de pouvoir comparer, dès lors qu'un individu passe du statut de cotisant à celui de retraité, les droits estimés à partir du dernier EIC aux droits versés effectivement par chaque régime de retraite. Ceci permettrait d'avoir une quantification de l'ampleur des droits que les personnes font reconnaître au moment de la liquidation de leur pension et qui n'étaient pas recensés dans les bases des régimes.

L'articulation EIC-EIR prendra 2 formes :

- les 4 générations les plus anciennes de l'EIC 2003 (1934, 1938, 1942, 1946) sont déjà représentées dans l'EIR 2001 et, à ce titre, un échantillon d'individus de ces générations a déjà été tiré pour l'EIR. Ces individus seront inclus dans l'EIC 12003, avec un éventuel tirage complémentaire (tableau 2). Les décès survenus au cours de l'année 2001 et qui auront été enregistrés pour mettre à jour le RNIPP pourront néanmoins entraîner des différences d'échantillon entre l'EIR et l'EIC, y compris pour les générations communes (l'EIR 2001 a été tiré dans la situation du RNIPP de novembre 2000).

- un certain nombre de personnes identifiées comme des cotisants dans l'EIC 2003 seront retraitées lors de la mise en place de l'EIR 2005. Pour la génération 1950 qui entrera dans l'EIR 2005 et qui aura fait l'objet d'un tirage pour l'EIC 2003, seront sélectionnés les individus faisant partie de l'EIC 2003. La même procédure se répètera pour les EIR ultérieurs.

L'articulation entre l'EIC et l'EIR aura donc des incidences sur le tirage des prochains EIR, mais pas sur les informations demandées.

La liste des personnes de l'échantillon est finalement composée :

- pour la génération 1934, des individus faisant partie de l'EIR 2001,
- pour les générations 1938 à 1946, des individus faisant partie de l'EIR 2001, et d'un complément d'individus sélectionnés par l'INSEE dans le RNIPP sur la base des dates de naissance désignées dans le tableau 2,
- pour les générations 1950 à 1970, d'individus sélectionnés par l'INSEE de manière aléatoire dans le RNIPP sur la base des dates de naissance désignées dans le tableau 2.

Le nombre de personnes sélectionnées par l'INSEE devrait être de l'ordre de 200 000, compte tenu des critères de tirage indiqués dans le tableau 2.

Tableau 2 - Critères de tirage de l'échantillon de l'EIC 2003

Année de naissance	Génération représentée dans l'EIR 2001 et jours de naissance	Jours de naissance (en complément de 1934 à 1946)
1934 (1)	oui : 1 ^{er} au 12 octobre	
1938	oui : 1 ^{er} au 6 octobre	7 au 10 octobre
1942	oui : 1 ^{er} au 6 octobre	7 au 10 octobre
1946	oui : 1 ^{er} au 6 octobre	7 au 10 octobre
1950	Non	1 ^{er} au 10 octobre
1954	Non	1 ^{er} au 10 octobre
1958	Non	1 ^{er} au 10 octobre
1962	Non	1 ^{er} au 10 octobre
1966	Non	1 ^{er} au 10 octobre
1970	Non	1 ^{er} au 10 octobre

(1) Sur-représentée dans l'EIR 2001 pour le calcul des taux de remplacement (1^{er} au 6 octobre + 7 au 12 octobre).

Le tableau 3 donne une simulation de la taille de l'EIC, sur la base des effectifs par génération au recensement de la population (RP) de 1999 et non du RNIPP fin 2001.

Tableau 3 - Simulation de l'échantillon de l'EIC 2003

Année de naissance	Population RP 1999	Effectif tiré pour les jours de naissance indiqués dans le tableau 2
Génération 1934	558 816	17 882
Génération 1938	544 294	14 515
Génération 1942	548 499	14 627
Génération 1946	803 502	21 427
Génération 1950	864 722	23 059
Génération 1954	846 527	22 574
Génération 1958	844 114	22 510
Génération 1962	852 382	22 730
Génération 1966	865 182	23 072
Génération 1970	831 808	22 182

Le tableau 4 indique, pour les générations 1934, 1938, 1942 et 1946 communes à l'EIR 2001 et l'EIC 2003, le nombre d'individus sélectionné par l'INSEE pour l'EIR 2001 et le nombre de ceux déjà identifiés comme retraités dans l'EIR 2001.

Tableau 4 - Échantillon de l'EIR 2001 pour les générations communes à l'EIR 2001 et l'EIC 2003

Génération	Nombre d'individus tirés par l'INSEE	Nombre d'individus identifiés comme retraités
1934	15 904	14 425
1938	7 911	5 924
1942	8 297	1 178
1946	11 532	568

Source : DREES, Échantillon interrégimes de retraités 2001

Les informations transmises par les caisses concernant les individus des générations 1934 à 1946 qui auront déjà été identifiés comme retraités dans l'EIR n'auront pas à être réactualisées pour les EIC ultérieurs, puisqu'elles seront a priori définitives (au moins dans certains régimes). La situation peut varier d'un régime à l'autre. Ces individus feront néanmoins partie des échantillons ultérieurs tirés par l'INSEE pour l'EIC, sauf pour la génération 1934 qui sera sortante dans l'EIC 2007 et pour les décédés entre deux EIC. Les régimes devront être en mesure de renvoyer à la DREES les informations concernant les retraités pour tous les EIC ultérieurs : ils ont le choix entre conserver l'information transmise à la DREES entre deux EIC ou refaire pour chaque EIC la recherche sur l'ensemble des individus.

4.3 Champ et période de référence

Le champ de l'EIC comprend tous les individus de l'échantillon qui cotisent actuellement à un régime ou y ont cotisé à un moment ou à un autre. « Cotiser » s'entend ici au sens large d'être affilié et d'acquérir des droits (au cours d'une période de travail ou au cours d'une période assimilée). Autrement dit, il ne suffit pas d'identifier uniquement les cotisants actuels, mais aussi de repérer tous les anciens cotisants, même radiés, et de collecter les données pour chaque type de population. L'appariement peut se faire sur le NIR, ou à défaut sur les autres variables d'identification (nom, prénom, date de naissance).

La période de référence des informations sur les droits accumulés va jusqu'au 31 décembre 2001. Les personnes dont la date d'affiliation à un régime est postérieure au 31 décembre 2001 ne doivent pas être considérées comme cotisant à ce régime.

Parmi les générations les plus anciennes, les régimes repèreront à la fois des cotisants et des retraités ; le même type d'informations est demandé pour les deux populations. Le fichier comprendra une variable décrivant le statut (statut actuel, renseignée par chaque régime) qui permettra de les distinguer. Par ailleurs, dans certains régimes, il est possible d'être à la fois retraité et salarié (exemple : à l'IRCANTEC, un cotisant peut percevoir une retraite parce qu'il a exercé un mandat d'élu local indemnisé). Ces deux situations, différenciées par la

variable relative au statut, seront décrites séparément (un code caisse différent sera attribué à l'IRCANTEC-élus et à l'IRCANTEC).

L'enregistrement des décès dans le RNIPP est effectué avec des retards plus ou moins importants selon la commune qui enregistre le décès. De plus, des décès peuvent intervenir entre la date de tirage de l'échantillon et la collecte des données par les régimes. Il est donc demandé aux régimes de repérer les personnes figurant dans l'échantillon, mais décédées avant le 31 décembre 2001.

4.4 Circuits de collecte...

4.4.1 ... pour le Régime général, les régimes complémentaires obligatoires, les régimes alignés, les régimes des professions libérales, les régimes spéciaux de salariés

Le fichier comportant la liste des personnes sélectionnées pour l'échantillon sera transmis par l'INSEE aux caisses en mars-avril 2003. La nécessité d'identifier non seulement les cotisants actuels mais également les anciens cotisants impose à chaque caisse de traiter l'intégralité du fichier. Toutefois, l'existence des différences en matière de gestion de données nécessitera des délais supplémentaires pour certains régimes (tableau 5).

En effet, pour un régime donné, les individus concernés par l'EIC sont tous ceux qui y ont cotisé à un moment donné de leur carrière. L'information demandée aux régimes sera donc relative à l'ensemble des cotisants, qu'ils soient actifs, préretraités ou radiés. Or, les données rétrospectives ne sont pas toujours conservées par les régimes. Des recherches sur des « dossiers-papiers » pour compléter les données de l'EIC sont à prévoir. La reconstitution d'historique et la recherche manuelle dans certains cas va donc nécessiter des délais supplémentaires. L'exploitation des dossiers papiers est à prévoir à la CNAVPL, l'ENIM, la CRPCEN, l'ARRCO et la RATP (tableau 5).

L'information n'est, par ailleurs, pas disponible au même niveau dans tous les régimes. Ainsi les régimes complémentaires, tels que l'AGIRC et l'ARRCO⁶, et certains régimes spéciaux, tels que celui de la SNCF ou d'IEG-Pensions ne centralisent pas les données sur les assurés, ce qui, dans la perspective de l'EIC, pose le problème de la collecte de l'information à des niveaux très décentralisés. La nécessité pour ces régimes de procéder à une interrogation de chacune de leurs caisses implique une charge de travail plus lourde et des délais supplémentaires.

Le NIR est largement utilisé comme identifiant par les régimes de retraite. Toutefois, une difficulté particulière se pose pour quelques régimes qui utilisent plutôt un matricule interne pour les opérations de gestion courante, par exemple pour le régime de la RATP. Pour constituer l'EIC, l'utilisation de certaines applications des régimes concernés, telle que l'application paye à la RATP, permettra de repérer au sein de l'échantillon de NIR tiré par l'INSEE, les agents actuellement et anciennement présents à la RATP et donnera pour ceux-ci

⁶ Toutefois, l'AGIRC et l'ARRCO procèdent aujourd'hui à un rapprochement de leurs bases de données. L'acquisition de nouvelles informations et le rapprochement de leurs bases vont permettre une meilleure gestion des carrières des affiliés, et ainsi une transmission plus complète des données nécessaire à l'élaboration de l'EIC.

la correspondance « numéro NIR - matricule RATP ». Cette démarche va néanmoins nécessiter des délais supplémentaires de collecte de données.

La procédure globale est donc, dans un premier temps, d'identifier les personnes de l'échantillon qui cotisent ou ont cotisé dans le régime, à l'aide du n° NIR ou à défaut du nom, du prénom et de la date de naissance puis de recueillir pour chacune de ces personnes l'information détaillée requise (tableau récapitulatif en annexe).

Les modalités de recueil des données, résumées dans le tableau 5 ci-après, diffèrent selon le système d'information de chaque régime : appariement sur le numéro NIR avec une base de données centralisée, utilisation d'autres bases en particulier pour obtenir la correspondance numéro NIR - numéro de gestion interne, recherche de tout ou partie de l'information dans des dossiers papier, transmission de la demande à des caisses ou établissements à un échelon décentralisé. La dernière étape consiste en l'envoi à la DREES d'un fichier contenant les données demandées pour les individus cotisants actuels ou anciens cotisants. Ce fichier ne contient pas les éléments d'identification, mais uniquement le numéro d'ordre.

Tableau 5 - Les différents modes de recueil de l'information selon le régime

Régime	Recours à ...			Modalités particulières d'organisation de la collecte
	... des données centralisées	... des données décentralisées	... des dossiers papier	
CNAVTS	oui	non	non	
ARRCO	non	oui	oui	- 76 caisses ARRCO - transmission de l'ensemble des n°NIR à chacune de ces caisses - points gratuits pour motif de maladie connus sur une base déclarative, donc sous forme papier
AGIRC	non (à terme projet BACCARA)	oui	non	- 34 caisses - transmission de l'ensemble des n°NIR à chacune de ces caisses - à terme utilisation de BACCARA au niveau central
IRCANTEC	oui	non	non	
CANCAVA	oui	non	non	
ORGANIC	oui	non	non	
CCMSA	oui	non	non	- 2 régimes, celui des exploitants et celui des salariés : le code caisse les différencie.
CNAVPL	non	oui	oui	- 12 sections professionnelles : information à rechercher au niveau de chacune d'entre elles - recherche papier pour l'historique nécessaire dans certaines sections
IEG Pensions	oui	oui	non	- retour aux établissements pour certaines informations (comme les primes)
CANSSM	oui	non	non	
SNCF	oui	oui	non	- gestion des cotisations faite par l'établissement qui emploie l'agent - pour identifier les agents et leur établissement de rattachement, appariement des n°NIR avec la liste du personnel au niveau central, puis interrogation des établissements
RATP	oui	non	oui	- le n°NIR n'est pas un identifiant dans les fichiers de la RATP : appariement des n°NIR avec l'application paye pour le repérage des agents et la correspondance n°NIR-matricule RATP - traitement mensuel antérieur aux 24 derniers mois, montant des cotisations et distinction primes/traitement à recherche sur microfiches au service de la paie
Banque de France	oui	non	non	
CRPCEN	oui	non	oui	- NIR non certifiés, souvent absents pour les radiés avant 1990 - dossiers papier pour données antérieures à 1990
CAVIMAC	oui	oui	non	- dans certains cas, interrogation directe des affiliés
CRPNAC	oui	non	non	
SEITA	oui	non	non	
ENIM	oui	non	oui	- dossiers papier pour l'information antérieure à 1972

4.4.2 ... pour la Fonction publique

En l'absence de données sur les carrières des fonctionnaires au sein des organismes qui gèrent leur retraite, quatre sources seront mobilisées : le panel DADS, la base CNRACL-FSPOEIE, les fichiers de paie de la Comptabilité Publique et les fichiers du ministère de la défense.

L'INSEE procédera à une extraction du panel DADS pour l'ensemble des NIR sélectionnés dans l'EIC de manière à recueillir des informations sur les carrières des individus ayant été rémunérés par une collectivité locale depuis 1989 ou par un hôpital depuis 1976 (qu'ils soient ou non fonctionnaires).

Par appariement avec la base de cotisants actuels et passés de la CNRACL et du FSPOEIE à l'aide du NIR, on pourra :

- distinguer les fonctionnaires des non-fonctionnaires ;
- repérer les individus ayant été fonctionnaires dans une collectivité locale avant 1989 ou dans un hôpital avant 1976 ou ouvriers d'État avant 1978 ;
- récupérer la date d'affiliation de manière à pouvoir calculer la durée validée à la CNRACL ou au FSPOEIE.

Les titulaires de la Fonction publique locale à temps non complet ne sont pas affiliés à la CNRACL, mais à l'IRCANTEC. Cette procédure permettra donc bien de retrouver les affiliés à la CNRACL, cette notion ne recouvrant pas exactement celle de titulaire/non titulaire.

Par extraction des fichiers de paie annuels de la Comptabilité publique pour l'ensemble des NIR sélectionnés dans l'EIC, on recueillera des informations sur les carrières des individus ayant été agent civil de l'État ou d'un établissement public national depuis 1978.

Les fichiers de paie annuels de la Comptabilité publique ne comprennent pas les NIR des militaires et les personnels civils du ministère de la défense ne figurent que partiellement dans ces fichiers. C'est pourquoi on mobilisera sur ce champ les données du ministère de la défense. La constitution du fichier sera effectuée par l'Observatoire social de la défense (OSD), en liaison avec l'Observatoire économique de la défense (OED).

Toutefois, malgré la mobilisation de ces différentes sources de données, les carrières ainsi reconstituées ne seront pas complètes. Les carrières des fonctionnaires d'État ne seront pas connues avant 1978 et devront être simulées avant cette date afin d'en déduire, en particulier, un âge d'entrée dans la Fonction publique d'État, et ce, en fonction du sexe, de la génération, de la catégorie statutaire de l'individu.

En ce qui concerne les agents des collectivités locales et des hôpitaux, la base des cotisants permettra de mieux connaître leur durée d'assurance mais sans précision, par exemple, sur les rémunérations perçues avant 1989 ou 1976.

4.4.3 Appariement complémentaire : les fichiers de l'UNEDIC

Afin de couvrir les situations de chômage indemnisé, de chômage non indemnisé et de préretraite, il est prévu de recueillir des données extraites des fichiers UNEDIC de manière à pouvoir qualifier certains des trous de carrière subsistant après l'appariement des fichiers de cotisants fournis par les différentes caisses. L'extraction ne pourra pas être effectuée à partir du seul Fichier national des Assedic (nouveau FNA). En effet, ce fichier n'est considéré comme complet que pour les individus qui sont passés par une période de chômage ou entrés en préretraite depuis 1993 (ou dont la période de préretraite inclut l'année 1993). Pour les autres, le Fichier national des allocataires (ancien FNA), considéré comme complet depuis 1984-1985 et comprenant un historique allant jusqu'au 31 décembre 1997, sera mobilisé. Ce fichier ne couvre pas le champ des chômeurs non indemnisés.

Le fichier de l'UNEDIC couvre tous les individus de l'échantillon INSEE ayant connu une période de chômage indemnisé ou non indemnisé ou de préretraite sur la période 1993-2001 et ceux ayant connu au moins une période de chômage indemnisé ou de préretraite sur les années 1984-1992.

5. Informations sollicitées

5.1 Nature de l'information

L'EIC doit permettre d'estimer les montants des pensions de retraite et les âges de liquidation. L'évaluation des droits acquis nécessite donc de connaître les rémunérations perçues par les individus ainsi que les différentes périodes ouvrant des droits à la retraite (ie : périodes validées et assimilées).

Le calcul des avantages de droits directs nécessite, en conséquence, la connaissance du nombre de trimestres validés (cotisés ou non), des périodes assimilées, des rémunérations, du montant des cotisations et du nombre de points (cotisés ou non).

Ainsi, pour les régimes de base et pour chaque année de carrière, la rémunération et ses composantes principales, le montant des cotisations, le nombre de trimestres durant lesquels il y a eu cotisation dans le régime, le nombre de trimestres assimilés ainsi que les motifs correspondants sont sollicités.

Pour les régimes complémentaires, le nombre de points acquis avec ou sans contrepartie de cotisations ainsi que les motifs correspondants sont demandés (tableau 6). L'année de début de cotisation dans le régime, ainsi que celle de fin, s'il y a eu radiation, sont également requises.

Pour les régimes qui procèdent à des estimations des droits acquis, soit dans le cadre de reconstitutions de carrières soit par un suivi en temps réel, le taux de liquidation déjà acquis, la durée totale de carrière sont, par exemple, demandés.

À des fins d'études, d'autres types d'informations sont sollicitées telles que le statut, la situation d'activité et la condition d'emploi (temps complet, temps partiel). Les aléas de

carrière pourront, par exemple, être évalués à partir des périodes non cotisées. Les informations sur les périodes d'interruption d'activité ainsi que les causes, celles relatives au chômage (indemnisé ou non), à la maladie (validée ou non) et à la maternité, par exemple, sont dès lors requises auprès des régimes.

Les informations sur la situation personnelle et familiale de l'individu seront fournies par les régimes qui en disposent. Ces données permettent de savoir si l'individu a le droit à des avantages accessoires tels que la bonification par enfant ou la majoration pour conjoints à charge, par exemple.

Sur le champ de la Fonction publique, les caractéristiques professionnelles (catégorie statutaire, catégorie socio-professionnelle, conditions d'emploi), la rémunération détaillée, le domaine d'emploi, le secteur d'activité, la catégorie juridique de l'organisation et le nombre d'enfants sont requis, à la fois pour le calcul de la pension et à des fins d'études.

Tableau 6 - Tableau simplifié des données sollicitées auprès des différents organismes participants à l'opération

Catégories de données par organisme et source statistique				
Organismes gestionnaires de régimes de retraite obligatoire	UNEDIC	DGCP	INSEE	OSD
Fichiers de gestion	Fichiers de gestion	Fichiers de paie	Déclarations annuelles de données sociales	Fichiers de gestion
1. Caractéristiques individuelles permanentes du cotisant (notamment, données relatives à la dernière année ou dernière période d'affiliation), à la fin 2001 (année de référence de la première version de l'échantillon)				
<ul style="list-style-type: none"> -situation administrative vis-à-vis de l'organisme, -état matrimonial légal, -catégorie socio-professionnelle, -situation d'activité et conditions d'emploi, -dernier département de résidence, -durée totale de carrière -durée totale d'affiliation -cumul des trimestres et points validés dans le régime, -taux de liquidation acquis, -année probable de liquidation, -nombre d'enfants, -droits acquis pour raisons familiales. 	<ul style="list-style-type: none"> -dernier département de résidence, -caractéristiques du dernier emploi exercé. 	<ul style="list-style-type: none"> -dernier département de résidence. 	<ul style="list-style-type: none"> -dernier département de résidence. 	<ul style="list-style-type: none"> -situation administrative vis-à-vis de l'organisme, -état matrimonial légal, -catégorie socio-professionnelle, -dernier département de résidence, -cumul des trimestres et points validés dans le régime, -nombre d'enfants, -droits acquis pour raisons familiales.
2. Caractéristiques individuelles du cotisant pour chaque année ou période d'affiliation, depuis 1945 et jusqu'à 2001 compris				
Pour chaque année d'affiliation	Pour chaque période d'affiliation à allocation constante	Pour chaque année de présence	Pour chaque année de présence où l'organisme employeur fait partie du panel des DADS	Pour chaque année d'affiliation
<ul style="list-style-type: none"> -situation de l'individu vis-à-vis du régime au 31 décembre de chaque année, -nombre total de trimestres validés, cotisés, rachetés et gratuits, -nombre total de points validés, cotisés, rachetés et gratuits, -rémunération et ses composantes, -montant total des cotisations d'assurances vieillesse, -conditions d'emploi, -nombre d'heures et de jours travaillés. 	<ul style="list-style-type: none"> -statut d'indemnisation, -types de droits versés, -motif de radiation. 	<ul style="list-style-type: none"> -catégorie statutaire, -conditions d'emploi, -rémunération détaillée, -catégorie d'organisme employeur. -nombre d'enfants. 	<ul style="list-style-type: none"> -caractéristiques de la période travaillée -conditions d'emploi, -rémunération, -domaine d'emploi et secteur d'activité, -catégorie juridique de l'organisme, type de budget 	<ul style="list-style-type: none"> -situation de l'individu vis-à-vis du régime au 31/12 de chaque année, -nombre total de trimestres validés, -rémunération et ses composantes, -montant total des cotisations d'assurances vieillesse -condition d'emploi, -nombre d'heures et de jours travaillés.

5.2 Caractéristiques générales des fichiers transmis par les régimes

Les fichiers sont hiérarchisés ; ils comprennent trois niveaux : un niveau « individu », un niveau « année de présence » et un niveau « période ».

Le niveau « individu » concerne les caractéristiques individuelles de l'affilié (date de naissance, sexe, etc) et les informations sur la dernière année de cotisation, données qui ne changent pas au cours du temps. Le niveau « année de présence » décrit chaque année d'affiliation. Enfin, il existe un troisième niveau, dit « période », destiné à permettre à certains régimes de faire figurer des cumuls de points ne pouvant, pour des raisons diverses, être ventilés année par année.

6. Perspectives

L'échantillon devrait servir de support à deux types de travaux. À partir de l'EIC stricto sensu, il sera possible d'établir des **typologies de carrières**, par sexe et par génération : les trajectoires entre les régimes, les profils de rémunération ainsi que les changements de situation (temps partiel, chômage ou inactivité) pourront être étudiés. Les **droits acquis** à la retraite, en termes de durée d'assurance, pourront être calculés pour chacune des générations, en distinguant les motifs de validations des périodes (travail, chômage, maladie...).

Des **cas-types pertinents et représentatifs**, dont on pourra mesurer la fréquence, pourront être construits. Pour les individus communs à l'EIC et à l'EIR, on pourra mesurer **l'écart entre la ou les pensions estimées à partir de l'EIC et les pensions réellement perçues**, données par l'EIR.

À partir de l'EIC enrichi de simulations, il sera possible de **prolonger les carrières**, et **d'estimer le montant des pensions et l'âge de départ** : ces travaux seront menés d'abord sur les générations les plus anciennes, dont la carrière sera connue dans sa quasi-totalité, puis élargis aux générations les plus jeunes.

7. ANNEXES

7.1 Glossaire

AGIRC : Association générale des institutions de retraites des cadres
ARRCO : Association des régimes de retraites complémentaires
AVTS : Allocation aux vieux travailleurs salariés
AVPF : Assurance vieillesse des parents au foyer
CANCAVA : Caisse autonome nationale de compensation de l'assurance vieillesse artisanale
CANSSM : Caisse autonome nationale de la Sécurité sociale dans les mines
CARMF : Caisse autonome de retraite des médecins de France
CARPIMKO : Caisse autonome de retraite et de prévoyance des infirmiers, masseurs-kinésithérapeutes, pédicures-podologues, orthophonistes et orthoptistes
CAVIMAC : Caisse d'assurance vieillesse, invalidité et maladie des cultes
CNAV-TS : Caisse nationale d'assurance vieillesse des travailleurs salariés
CNAVPL : Caisse nationale d'assurance vieillesse des professions libérales
CNRACL : Caisse nationale de retraite des agents des collectivités locales
CDC : Caisse des dépôts et consignations
CGP : Commissariat général au plan
CIG : Centre interdépartemental de gestion
CNIL : Commission nationale de l'informatique et des libertés
COR : Conseil d'orientation des retraites
DADS : Déclaration annuelle des données sociales
FSPOEIE : Fonds spécial des pensions des ouvriers des établissements industriels de l'État
IEG : Industries électriques et gazières
IRCANTEC : Institut de retraite complémentaire des agents non titulaires de l'État et des collectivités publiques
MSA : Mutualité sociale agricole
NIR : Numéro d'identification au répertoire
ORGANIC : Organisation autonome nationale de l'industrie et du commerce
OSD : Observatoire social de la défense
PCS : Profession et catégorie socioprofessionnelle
RATP : Régie autonome des transports parisiens
RNIPP : Répertoire national d'identification des personnes physiques
SAM : Salaire annuel moyen
SEITA : Société nationale d'exploitation industrielle des tabacs et allumettes
SGA : Secrétariat général pour l'administration
SMIC : Salaire minimum interprofessionnel de croissance
SNCF : Société nationale des chemins de fer français
TOM : Territoires d'Outre-Mer

7.2 Liste des organismes gestionnaires de régimes de retraite obligatoire participants à l'opération EIC

- la Caisse nationale d'assurance vieillesse des travailleurs salariés ;
- la caisse centrale de la mutualité sociale agricole – salariés et exploitants agricoles ;
- l'Association des régimes de retraites complémentaires ;
- l'Association générale des institutions de retraite des cadres ;
- l'Institution de retraite complémentaire des agents non titulaires de l'État et des collectivités publiques régime général, régime des élus et régime des médecins ;
- la Caisse de compensation de l'organisation autonome nationale de l'industrie et du commerce – régime de base et régime complémentaire ;
- la Caisse autonome nationale de compensation de l'assurance vieillesse artisanale – régime de base et régime complémentaire ;
- la Caisse nationale de retraite des agents des collectivités locales ;
- la Caisse de retraite des notaires (régime de base et régime complémentaire) ;
- la Caisse autonome de retraite des médecins de France (régime de base, régime complémentaire et régime des praticiens conventionnés) ;
- la Caisse autonome de retraite des chirurgiens dentistes (régime de base, régime complémentaire et régime des praticiens conventionnés) ;
- la Caisse d'assurance vieillesse des pharmaciens (régime de base, régime complémentaire et régime des praticiens conventionnés) ;
- la Caisse autonome de retraite des sages-femmes françaises (régime de base et régime des praticiens conventionnés) ;
- la Caisse autonome de retraite et de prévoyance des infirmiers, masseurs-kinésithérapeutes, pédicures-podologues, orthophonistes et orthoptistes (régime de base, régime complémentaire et régime des praticiens conventionnés) ;
- la Caisse autonome de retraite et de prévoyance des vétérinaires (régime de base et régime complémentaire) ;
- la Caisse de retraite de l'enseignement, des arts appliqués, du sport et du tourisme (régime de base et régime complémentaire) ;
- la Caisse d'allocation vieillesse des agents généraux et des mandataires non salariés de l'assurance et de la capitalisation (régime de base et régime complémentaire) ;
- la Caisse interprofessionnelle de prévoyance et d'assurance vieillesse (régime de base et régime complémentaire) ;
- la Caisse d'allocation vieillesse des experts-comptables et des commissaires aux comptes (régime de base et régime complémentaire) ;
- la Caisse autonome nationale de la sécurité sociale dans les mines (service de Paris) ;
- la Caisse de retraite de la société nationale des chemins de fer français ;
- le Département des prestations invalidité-vieillesse-décès d'électricité de France-gaz de France ;
- l'Établissement national des invalides de la marine ;
- la Caisse de retraite et de prévoyance des clercs et employés de notaire ;
- la division pensions retraites de la régie autonome des transports parisiens ;
- la Caisse d'assurance vieillesse invalidité et maladie des cultes ;
- la Caisse de retraite du personnel navigant professionnel de l'aéronautique civile ;
- la Caisse de réserve des employés de la Banque de France ;
- le Service d'exploitation industrielle des tabacs et allumettes ;
- le Fonds spécial des pensions des ouvriers des établissements industriels de l'État.

7.3 Autres organismes participants à l'opération EIC :

- L'Union nationale pour l'emploi dans l'industrie et le commerce ;
- L'Institut national de la statistique et des études économiques ;
- La Direction générale de la comptabilité publique du ministère de l'économie, des finances et de l'industrie ;
- Le Secrétariat général pour l'administration du ministère de la défense.