
Une méthodologie nouvelle pour l'enquête Famille 2025, descendante d'une longue lignée

Lucile RICHET-MASTAIN, Elisabeth ALGAVA ()*

() Insee, Direction des statistiques démographiques et sociales,
Division Enquêtes et études démographiques*

`lucile.richet-mastain@insee.fr, elisabeth.algava@insee.fr`

Mots-clés : Collecte, Multimode, Recensement de la population, Famille, Démographie

Domaines Collecte, Multimode

Résumé

Cet article présente le nouveau mode de collecte retenu pour la future enquête Famille 2025. Cette enquête recueille des informations détaillées sur la composition des familles, les événements familiaux, la fécondité, l'origine sociale et géographique. Elle permet ainsi des analyses sur la fécondité et les familles qui croisent différentes dimensions : les générations, les origines sociales, les trajectoires géographiques sur de nombreux thèmes tels que la descendance finale et le calendrier des naissances, la vie familiale des enfants de parents séparés, les parcours conjugaux et familiaux. La taille de son échantillon permet une analyse de populations rares : les couples de même sexe, les familles homoparentales, les pères isolés ou ayant la garde exclusive de leurs enfants, l'orphelinage, les familles adoptives, etc.

Depuis 1954, l'enquête Famille est « associée » au recensement de la population : la collecte a toujours été réalisée en même temps que le recensement, auprès d'un échantillon de ménages recensés. La dernière édition de l'enquête Famille, l'enquête Famille et Logements (EFL), a eu lieu en 2011 et a été collectée entièrement sous format papier (questionnaire de 4 pages auto-administré). C'était la première édition à ne plus être associée à un recensement exhaustif mais à une collecte annuelle de recensement, ce qui a nécessité des ajustements méthodologiques et du protocole¹.

Depuis 2011, la collecte du recensement a évolué, tout particulièrement par l'introduction, à côté de la collecte papier, d'une collecte par internet d'ampleur croissante. Cela a conduit à retenir une collecte multimode de l'enquête Famille 2025, avec des protocoles très distincts cependant entre la collecte internet et la collecte papier. L'instruction des différents protocoles envisagés a montré qu'il n'était pas opportun d'ajouter un module « enquête Famille » à la collecte internet du recensement et qu'il était préférable de dissocier les deux collectes d'un point de vue informatique. Ce protocole place ainsi l'enquête dans deux filières, celle du recensement et

1. Trois présentations à partir de l'Enquête Famille et Logement 2011 ont été réalisées aux JMS 2015 (exemple : <http://jms-insee.fr/jms2015s21-1/>)

celle de la filière des enquêtes ménages à l’Insee (Métallica), qui ont chacune leurs équipes, leurs outils, leurs particularités et leurs calendriers. Cela constitue un enjeu fort pour cette nouvelle enquête : les différences devront être *in fine* aplanies afin que les fichiers issus des deux collectes soient réconciliés en un seul, pour produire des résultats de qualité, représentatifs et comparables à ceux de l’EFL 2011.

La taille de l’échantillon de l’enquête Famille est importante avec 400 000 personnes interrogées. L’échantillon sera représentatif au niveau régional et tiré selon des grappes de districts dans les petites communes et d’Iris² dans les grandes communes. La répartition sera de deux tiers de femmes et d’un tiers d’hommes interrogés, l’enquête ayant pour but privilégié à l’origine d’étudier la fécondité des femmes.

La collecte envisagée pour l’édition de l’enquête Famille 2025 sera menée selon le protocole suivant :

- elle sera réalisée sous format papier pour les personnes qui répondent au recensement en format papier et cette collecte sera réalisée en même temps que celle du recensement. Cette collecte sera proche de celle menée en 2011 qui avait été la première enquête adossée au recensement, et dont l’expérience s’est avérée concluante. Cependant, certaines évolutions seront à prendre en compte telles que le suivi de la collecte désormais réalisé *via* l’application Omer ou la mise en place en 2022 de la collecte séquentielle (dépôt des documents dans la boîte aux lettres pour les habitants de maisons individuelles) ;
- elle sera réalisée par internet pour les personnes qui répondent au recensement par internet. La collecte internet de l’enquête sera réalisée dans le cadre de la filière Métallica, sans attendre la fin de la collecte du recensement. Elle s’appuiera sur l’utilisation, à des fins d’enquête statistique, du mail donné par les répondants au recensement. Les personnes recensées par internet et ayant communiqué une adresse mail recevront, quelques jours après avoir répondu au recensement, un mail les invitant à répondre à l’enquête Famille. Le délai de quelques jours entre la réponse au recensement et l’envoi du mail sera mis à profit pour constituer, à intervalles réguliers, des échantillons grâce aux données brutes collectées au recensement, qui seront entrés en *input* du questionnaire internet de l’enquête. Chaque fichier échantillon contiendra à la fois les informations de contact et des données sur la composition du ménage. La récupération des réponses au recensement permettra d’élaborer un questionnaire plus adapté aux situations, sans reposer les questions de composition du ménage par exemple, ce qui ne serait pas compris par les répondants internet au recensement³. Cette amélioration de la qualité de la collecte ne concernera que les répondants internet, mais ils seront majoritaires, puisque l’on vise 70 % de réponse sur internet.

Grâce à l’effet d’entraînement du recensement, *via* une collecte concomitante sur papier et rapprochée sur internet, un taux de réponse élevé est attendu pour l’enquête Famille 2025. Ceci devrait limiter les besoins de relance, notamment téléphonique, et donc le coût de collecte de l’enquête. L’objectif de répondants est de 365 000, effectif similaire à celui de l’EFL 2011. Après la collecte, les données de l’enquête seront complétées avec les réponses complètes au recensement et un appariement avec les données fiscales et sociales est également envisagé. L’enquête Famille 2025 constituera ainsi une source de données extrêmement riches.

2. Îlots Regroupés pour l’Information Statistique.

3. Dans l’enquête de 2011, collectée sur papier uniquement, comme l’EAR de 2011, il était indispensable par exemple de relister les enfants du logement pour compléter les questions les concernant. Pour la future enquête Famille sur internet, redemander la liste de ces enfants ne sera pas nécessaire si on a déjà cette information récupérée par l’EAR internet. La « répétition » restera en revanche nécessaire pour la collecte papier de l’enquête Famille.

Abstract

This article describes the new methodology of the future Family survey which will take place in 2025. This survey collects detailed information on the composition of families, fertility, social and demographic background. It is the descendant of a long line since the first edition took place in 1954. Data collection has always been linked to the census survey. Until the last edition in 2011, the survey was collected using a paper questionnaire.

The collection method for the 2025 edition is innovative : for people who answer to the census of the population in paper format, it will be conducted in paper format and for people who answer to the census of the population on the internet, it will be conducted in internet format. The reminder that can be implemented in case of low response rate will be conducted by telephone. After describing the survey, the article presents the main advantages and disadvantages of these collection methods, depicts the richness of the survey data and reviews the main risks and challenges of this project.

Introduction

Cet article présente le nouveau mode de collecte qui sera mis en place pour la future enquête Famille 2025. La première partie porte sur une présentation de l'enquête : ses thématiques, son champ, son échantillon, son questionnaire ainsi que l'exploitation et la valorisation des données. La deuxième partie présente le nouveau protocole de collecte de l'enquête Famille : collecte papier et *via* internet, éventuellement par téléphone pour la relance, ainsi que les principaux avantages et inconvénients de ces modes de recueil de données. La troisième partie décrit la richesse des données attendues de l'enquête et la dernière partie synthétise les principaux défis et risques de ce projet.

1 Présentation de l'enquête Famille

1.1 Historique de l'enquête Famille

L'enquête Famille 2025 (EF) sera la descendante d'une longue lignée. De 1954 à 1999, une enquête Famille a été adossée à chaque recensement exhaustif de la population (RP). La collecte a toujours été réalisée en même temps que celle du recensement, auprès d'un échantillon de ménages recensés. En 1999, l'enquête Famille nommée « Étude de l'histoire Familiale » s'est appuyée sur le dernier recensement exhaustif. La dernière édition en date de l'enquête, l'enquête Famille et Logements (EFL), a eu lieu en 2011 et a été collectée entièrement sous format papier. C'était la première édition à ne plus être associée à un recensement exhaustif mais à une enquête annuelle de recensement (EAR), ce qui a nécessité des ajustements méthodologiques et du protocole [1].

Jusqu'à l'édition de 1999, l'enquête Famille était adressée uniquement aux femmes. Elle a d'abord porté sur les femmes mariées ayant achevé leur vie féconde (entre 45 et 54 ans) pour élargir ensuite son champ aux femmes mariées, veuves ou divorcées de moins de 65 ans et par la suite à l'ensemble des femmes qu'elles aient ou non été mariées. Ce processus d'élargissement s'est poursuivi en 1999 avec l'extension de l'interrogation aux hommes. En plus de thématiques récurrentes, chaque édition de l'enquête consacre une série de questions à un thème spécifique, comme l'activité féminine en 1982 ou la garde des jeunes enfants en 1990. L'édition de l'enquête Famille de 1999 a porté sur la transmission familiale des langues tandis que celle de 2011 a porté sur la garde des enfants de moins de 4 ans et la multirésidence.

Depuis 1954, les enquêtes Famille ont donc évolué, mais certains traits distinctifs perdurent (**figure 1**) et l'enquête Famille 2025 se situera dans la continuité.

Figure 1 : L'enquête Famille de 1962, un questionnaire portant sur les enfants et les unions, collectée par les agents recenseurs, auprès des femmes dans les districts tirés au sort

RECENSEMENT DE 1962 Institut National de la Statistique et des Études Économiques

BULLETIN SPÉCIAL POUR L'ÉTUDE DES FAMILLES (Imprimé n° 2 F)

Complément à l'instruction aux Agents recenseurs (Imprimé 2 F)

Objet de l'enquête. Votre district (ou vos districts) de recensement a été désigné pour être soumis à une enquête par sondage dont l'objet est de réunir une documentation d'intérêt scientifique sur la structure des familles.

Personnes à interroger. L'enquête consiste à remplir, en plus du bulletin individuel (imprimé n° 2) un *bulletin spécial pour l'étude des familles* (imprimé n° 2 F) pour toute femme mariée, veuve ou divorcée âgée de moins de 70 ans le 1^{er} janvier 1962, c'est-à-dire née en 1892 et après.

Consignes générales. Contrairement aux autres informations qui sont rassemblées au cours du recensement général de la population, les données relatives aux familles ne sont recueillies qu'auprès d'une fraction de la population : celle qui réside dans les districts qui, comme le (ou les) vôtre ont été désignés par le sort.

1.2 Objectifs de l'enquête Famille

L'enquête Famille recueille des informations détaillées sur la composition des familles, les événements familiaux, la fécondité, l'origine sociale et géographique. Elle permet ainsi des analyses sur la fécondité et les familles qui croisent différentes dimensions : les générations, les origines sociales, les trajectoires géographiques sur de nombreux thèmes tels que la descendance finale et le calendrier des naissances ou la vie familiale des enfants de parents séparés.

L'enquête Famille fait l'objet d'une très forte demande sociale. Le recensement de la population apporte une photographie de la population à un moment donné, qu'il convient de compléter par l'histoire familiale et conjugale des personnes. Cela est rendu possible grâce à la datation des principales étapes biographiques (unions, ruptures, naissance, décès) disponible dans les enquêtes Famille. Comme depuis sa création, l'EF demeure ainsi la seule enquête à permettre d'étudier la descendance finale et le calendrier des naissances des femmes et des hommes selon les générations. Elle offre également une description de la vie familiale des enfants de parents séparés plus riche que celle dorénavant permise par le questionnaire rénové du recensement [2], en reconstituant les parcours de vie de ces enfants (datation des événements, rang de ces enfants dans les fratries en tenant compte des (demi-)frères ou sœurs qui vivent ailleurs, etc.).

De plus, la taille importante de l'échantillon de l'enquête Famille permet une analyse de populations rares : les couples de même sexe, les familles homoparentales, les pères isolés ou ayant la garde exclusive de leurs enfants, l'orphelinage, les familles adoptives, etc. Elle est également représentative au niveau des régions.

1.3 Utilisation des résultats de l'enquête Famille par une large palette d'acteurs

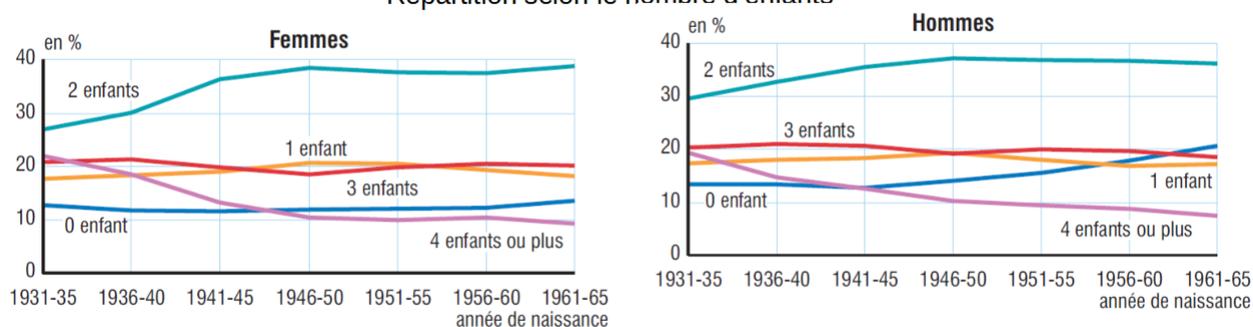
L'enquête Famille est très attendue par les partenaires de l'Insee ainsi que par l'Insee tant au niveau national qu'au niveau régional : elle répond à une demande sociale de connaissances

actualisées ou nouvelles sur les différents sujets liés à la famille qu'elle permet d'aborder.

L'enquête Famille 2025 fait partie des huit enquêtes de l'équipement en sciences sociales Lifeobs, observatoire français des parcours de vie, porté par l'Ined. Il s'agit d'une coopération nationale entre institutions clés : l'Ined, porteur du projet, certaines universités, l'infrastructure de recherche Progedo et l'Insee. Lifeobs est une infrastructure de recherche qui vise à développer un programme d'enquêtes longitudinales et innovantes sur les comportements familiaux. Se mettent en place des synergies, des échanges avec d'autres organismes porteurs d'enquêtes sur les familles, la mutualisation d'investissements méthodologiques novateurs sur la collecte multimode et les appariements avec des données administratives et une réflexion sur la diffusion des données.

Plus de 70 publications ont été réalisées à partir de l'enquête Famille et Logements de 2011. Elles portent sur des thématiques variées étudiées au niveau national (par exemple les couples et les familles [3] (**figure 2a**),

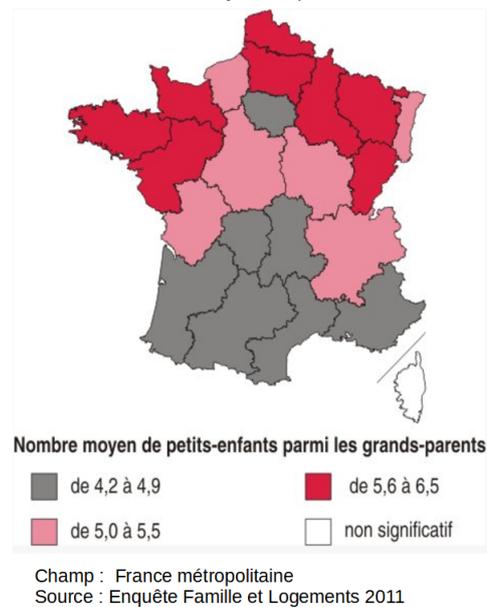
Figure 2a : Exemple de résultats de l'enquête Famille et Logements 2011
Répartition selon le nombre d'enfants



Champ : France métropolitaine
Source : Enquête Famille et Logements 2011

la famille à distance [4], les familles monoparentales [5]), sur des thématiques spécifiques (par exemple les familles d'immigrés [6], les familles des agents de la fonction publique [7], les couples de même sexe [8]) ainsi que des thématiques exploitées au niveau régional (par exemple les grands-parents [9] (**figure 2b**)).

Figure 2b : Exemple de résultats de l'enquête Famille et Logements 2011
Nombre moyen de petits-enfants



L'enquête de 2011 a également été mobilisée pour des travaux méthodologiques, en comparant ses résultats à ceux du recensement [10] ou encore de l'enquête Emploi (sur les structures familiales par exemple [11]) pour comparer notamment les tailles de ménages décrites dans ces différentes sources.

1.4 Questionnaire de l'enquête Famille

Le questionnaire de l'enquête Famille est court : il tient sur un format de 4 pages A4. Il est auto-administré et doit donc contenir des questions simples et facilement compréhensibles, sans l'aide d'un enquêteur. Le questionnaire contient également le moins de filtres possibles. L'édition de l'EFL de 2011 contenait deux tableaux relativement importants : le premier décrivant les enfants de la personne interrogée et de son éventuel conjoint vivant dans le logement et le second décrivant les enfants de la personne interrogée qui ne vivent plus dans le logement ou qui sont décédés.

Il existe un questionnaire pour les femmes et un autre pour les hommes (pour faciliter la collecte, réalisée selon les endroits uniquement auprès de hommes ou auprès de femmes – voir 1.6) mais ils diffèrent uniquement par la féminisation de certains noms et les accords grammaticaux.

Un comité scientifique qui réunit les principaux partenaires de l'enquête (Institut national d'études démographiques - Ined, Caisse nationale des allocations familiales - Cnaf, Service statistique ministériel du Ministère des Solidarités et de la santé – Drees) se réunit régulièrement depuis fin 2021, notamment pour élaborer la nouvelle version du questionnaire de l'EF 2025. Le thème spécifique pour cette édition est celui des solidarités familiales (contacts, en distinguant si possible échanges financiers et non financiers, proximité du lieu de résidence).

Une approche omnimode est mise en œuvre pour l'enquête Famille 2025 : ainsi, le questionnaire sera conçu indépendamment du mode de collecte, aucune question ne sera *a priori* posée dans un mode de collecte et non pas dans l'autre. Le questionnaire papier sera le premier à être élaboré. Les adaptations nécessaires dans chacun des modes de collecte seront réalisées dans un

second temps afin de limiter au maximum les effets de mode. Les deux tableaux importants notamment seront remplacés par une série de questions dans le questionnaire en ligne [12] et des filtres en fonction de certaines réponses au recensement pourraient être introduits pour la collecte internet.

1.5 Champ de l'enquête Famille

Les dernières éditions de l'enquête Famille ont été menées uniquement en France métropolitaine. Dans les départements d'outre-mer (DROM), une enquête Migrations, Famille et Vieillesse est menée de façon apériodique⁴ dans laquelle sont abordés des thèmes relativement proches de ceux de l'EF. La nouveauté de l'enquête Famille 2025 est d'élargir son champ géographique afin d'assurer une représentativité au niveau de la France, nationale comme régionale, mais aussi au niveau de chacun des DROM : Guadeloupe, Martinique, Guyane, La Réunion et Mayotte. Les enquêtes annuelles de recensement (EAR) sont désormais réalisées dans les DROM au même titre qu'en Métropole, y compris à Mayotte, depuis l'EAR 2021, ce qui rend possible cet élargissement. Comme il est observé dans les collectes de recensement, on envisage un taux de réponse sous format papier plus élevé dans les DROM qu'en métropole.

L'enquête Famille est menée auprès des personnes vivant en ménages ordinaires. Comme en 2011, les personnes vivant en collectivité ne font pas partie du champ de l'enquête⁵.

1.6 Échantillon de l'enquête Famille

L'échantillon de l'enquête Famille est tiré dans les enquêtes annuelles de recensement ; il est représentatif au niveau régional et tiré selon des grappes de districts dans les petites communes et d'Iris dans les grandes communes. La répartition des personnes interrogées dans l'échantillon est de deux tiers de femmes et d'un tiers d'hommes, l'enquête ayant pour thème privilégié depuis son l'origine l'étude de la fécondité des femmes.

Il existera des zones de tirage « Femmes » où toutes les femmes de 18 ans ou plus du ménage seront interrogées et des zones de tirage « Hommes » où tous les hommes de 18 ans ou plus du ménage seront interrogés.

Les échantillons des enquêtes Famille sont en effet, depuis leur origine, constitués de zones correspondant au travail des agents recenseurs. Cela facilite la collecte puisque certains agents recenseurs font l'enquête, dans toute leur zone, et d'autres ne sont pas du tout concernés. Depuis 1999, certains agents recenseurs enquêtent toutes les femmes de leur zone, d'autres tous les hommes. Toutefois, en 2011, maintenir cette simplicité de collecte tout en ayant un échantillon représentatif, au niveau régional qui plus est, a constitué un véritable défi. En effet, le nouveau contexte d'un recensement par enquête annuelle a conduit à complexifier le plan de sondage qui « s'appuie sur une intersection de deux échantillons de logements : d'une part, l'échantillon des logements de l'EAR de 2011, d'autre part, un échantillon de logements spécifiques à l'enquête EFL résultant d'un plan à deux degrés. Ce dernier est constitué par un tirage stratifié de communes (recensées en 2011) dans lesquelles on sélectionne au hasard (par tirage aléatoire simple, donc à probabilités égales) des grappes prédéfinies de logements, lesquelles sont en fait des regroupements de districts contigus, voire d'Iris contigus lorsqu'il s'agit de grandes communes. » [1].

4. La dernière édition date de 2020-2021.

5. L'enquête de 1999 avait été étendue aux communautés et notamment aux prisons (avec un échantillon de 6 600 personnes vivant en communauté, dont 1 700 détenus de sexe masculin).

Ensuite, les deux tiers de ces grappes sont celles dans lesquelles les femmes seront enquêtées et le tiers restant celles pour les hommes. Faire coïncider les grappes ainsi tirées avec les zones de collecte effectives de l'enquête n'est pas simple. Cela suppose d'obtenir le consentement de chacune des communes dans lesquelles des grappes sont sélectionnées et de signer une convention avec elles, puisque les communes sont désormais responsables de la collecte du recensement. Cela suppose aussi de faire coïncider les grappes avec les zones de travail des agents recenseurs, zones définies par leur employeur, à savoir la commune. Toutes ces étapes sont très complexes : elles seront à reproduire pour l'enquête Famille 2025 et mobiliseront également à l'Insee la division Sondages, la maîtrise d'ouvrage et les correspondants de l'enquête Famille en régions.

Néanmoins, l'enquête Famille et Logements de 2011 constitue un précédent sur lequel nous appuierons très largement. Quelques différences envisagées néanmoins : le recensement par internet et le recensement séquentiel (incitation à répondre par internet, puis, en cas de non-réponse, proposition de la collecte papier) réduisent le temps de collecte du recensement et incitent à agrandir les zones de collecte des agents recenseurs, ce qui sera à prendre en compte dans le tirage des grappes. L'opération repose sur une collaboration avec les communes : les convaincre de participer à l'EF 2025 sera donc un point important, même si en 2011 cela n'a pas *in fine* constitué une difficulté. Faute de moyens pour le mettre en œuvre et d'indications quant à son efficacité, nous n'utiliserons pas l'argument d'une contrepartie sous la forme de résultats détaillés pour la commune elle-même comme en 2011 [13].

La taille de l'échantillon visé de l'enquête 2025 sera proche de celle de l'EFL de 2011, soit 365 000 répondants.

Pour anticiper un taux de réponse à l'EF 2025 éventuellement plus faible que lors des éditions précédentes de l'enquête du fait du décalage temporel de la collecte internet (voir 1.7), le tirage d'un échantillon de réserve sera prévu dans les communes appartenant à l'échantillon principal ; ce dernier sera débloqué si nécessaire au cours de la collecte de l'EF 2025 pour pouvoir atteindre un taux de réponse suffisant.

1.7 Taux de réponse à l'enquête Famille

Jusqu'à l'édition de 2011, l'enquête Famille n'avait pas le caractère obligatoire de réponse. Cependant, en 2011 le taux de réponse a été très élevé (84 %) puisque l'EF a bénéficié de l'effet d'entraînement du recensement de la population, la collecte ayant lieu au même moment que le recensement. Elle a bénéficié également de l'ensemble du dispositif de grande ampleur mis en oeuvre pour le RP : la communication sur le recensement auprès de la population est très importante, ainsi que la relation entre les agents recenseurs, qui collectaient aussi l'enquête Famille, et les enquêtés. En 2025, cet effet d'entraînement devrait jouer pour favoriser le taux de réponse de la collecte papier. Il existe plus d'incertitudes quant à la collecte sur internet, décalée dans le temps, ce qui conduit à se poser la question de demander l'obligation de réponse.

L'objectif pour l'EF 2025 est d'atteindre 365 000 répondants, soit un effectif similaire à celui de l'EFL de 2011. En dehors du recensement, l'enquête Famille est donc l'enquête la plus importante de l'Insee menée auprès des ménages, en termes de taille d'échantillon. Cette taille d'échantillon est possible parce que l'enquête est auto-administrée.

La première édition de l'enquête Famille menée en 1954 concernait un échantillon d'environ 50 000 femmes non célibataires âgées de 45 à 54 ans. Dès l'édition de 1962, l'échantillon s'est

considérablement agrandi avec environ 250 000 femmes non célibataires mais âgées de moins de 70 ans. Ensuite, il est passé à 300 000 répondants, puis davantage lorsque le questionnaire a été étendu aux hommes (**figure 3**).

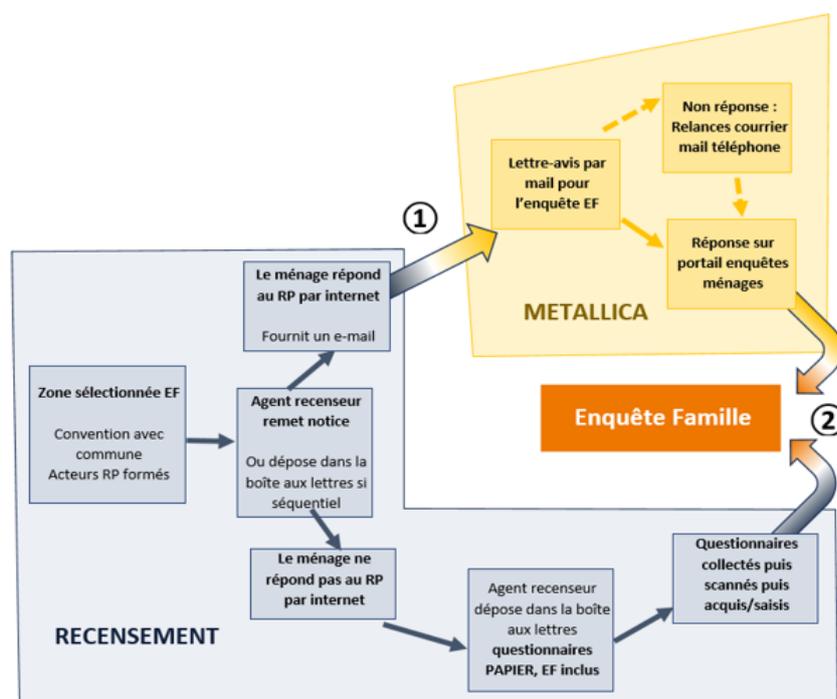
Figure 3 : Nombre de répondants aux différentes enquêtes Famille

Année de collecte	Nombre de répondants
1954	50 000
1962	250 000
1975	250 000
1982	310 000
1990	340 000
1999	380 000
2011	365 000

2 Un protocole de collecte innovant pour l'enquête Famille 2025

Depuis la dernière édition de l'enquête Famille en 2011, la collecte du recensement a évolué, tout particulièrement avec l'introduction depuis 2015, en plus de la collecte papier, d'une collecte *via* internet d'ampleur croissante. Cela conduit à retenir une collecte multimode de l'enquête Famille 2025, avec des protocoles très distincts entre la collecte papier et la collecte internet. L'instruction des différents protocoles envisagés a en effet montré qu'il n'était pas opportun d'ajouter un module « enquête Famille » à la collecte internet du RP et qu'il était préférable de dissocier les deux collectes d'un point de vue informatique. Ce protocole place ainsi l'enquête dans deux filières, celle du recensement et celle des enquêtes ménages de l'Insee (Métallica), qui ont chacune leurs équipes, leurs outils, leurs particularités et leurs calendriers. Cela constitue un enjeu fort pour cette nouvelle édition de l'enquête Famille (**figure 4**).

Figure 4 : Une enquête, deux filières



Note : les numéros correspondent aux points de jonction entre les deux filières.

Ainsi, pour la première fois, la collecte de l'enquête Famille 2025 sera multimode : en version papier pour les personnes qui répondent au recensement en version papier et en version internet pour les personnes qui répondent au recensement *via* internet.

En appliquant la répartition des modes de collecte observée dans le recensement, on part sur l'hypothèse d'arriver à environ 30 % des répondants à l'enquête Famille sous format papier et 70 % des répondants *via* internet⁶.

En fonction du taux de réponse, une relance des non-répondants à l'enquête est également envisagée sous la forme de courriers postaux, de relances par mails mais aussi d'une passation complète du questionnaire par téléphone, ce qui constituera un troisième mode de collecte.

2.1 La collecte papier

La partie collecte sous format papier de l'EF 2025 sera entièrement adossée à la collecte papier du recensement de la population, comme pour les éditions précédentes. Elle concernera les personnes qui répondent au recensement en format papier et cette collecte sera réalisée en même temps que celle de l'enquête annuelle de recensement (EAR)⁷. Une information sur l'enquête

6. En 2020, 62 % des bulletins individuels de recensement ont été collectés *via* internet.

7. La collecte de l'EAR débute le 3^{ème} jeudi de janvier pour une durée de 5 semaines et demi. En 2022 par exemple, la collecte du recensement a lieu du jeudi 20 janvier au samedi 26 février en France

Famille sera apposée sur les documents de l'EAR, qu'ils soient adressés aux maires ou aux habitants. Si les ménages souhaitent répondre à l'EAR sous format papier, un questionnaire papier de l'EF sera fourni par les agents recenseurs pour chacun des membres du ménage concernés (pour toutes les femmes de 18 ans ou plus si le ménage est dans une zone « Femmes » et pour tous les hommes de 18 ans ou plus si le ménage est dans une zone « Hommes »).

Ce mode de collecte sera proche de celui mené en 2011 : cela avait été la première enquête adossée aux enquêtes annuelles de recensement, et l'expérience s'est avérée concluante. Cependant, certaines évolutions seront à prendre en compte telles que le suivi de la collecte des bulletins de recensement désormais réalisé via l'application Omer. Il s'agira également de s'appuyer sur la méthode de collecte séquentielle du RP mise en place en 2022 : les agents recenseurs déposent les documents du recensement directement dans la boîte aux lettres des ménages si ces derniers résident en maison individuelle. Au bout de quelques jours, s'ils n'ont pas répondu *via* internet, l'agent recenseur les rencontrera. Dans les logements collectifs, les agents recenseurs continueront, comme pour les éditions antérieures des EAR, à rencontrer les ménages pour leur remettre les documents en mains propres. Ce nouveau mode de dépôt des bulletins du recensement limite au maximum les contacts entre les agents recenseurs et les enquêtés.

Comme lors des éditions précédentes de l'enquête, aucune relance spécifique ne sera réalisée auprès des non-répondants pour la partie collecte papier. On considère que les agents recenseurs qui relancent les non-répondants au RP le feront en même temps pour l'enquête Famille 2025.

Cette partie de la collecte est très fortement liée aux travaux des équipes du RP de l'Insee, depuis l'éditique, les contacts avec les communes et les formations des acteurs (superviseurs, coordonnateurs communaux, agents recenseurs) jusqu'au suivi de la collecte et la saisie des questionnaires. On prévoit que des correspondants enquête Famille seront présents dans les équipes RP des services statistiques des établissements régionaux de l'Insee, comme pour l'enquête de 2011, pour gérer les relations avec les communes concernées, suivre la collecte, etc. D'autre part, l'impression, le transport et la saisie des questionnaires papier de l'EF seront réalisés par les prestataires, *a priori*, comme en 2011, ceux qui prennent en charge les documents papier du recensement.

Avec l'hypothèse de 30 % de réponses à l'EF sous format papier, la collecte internet devrait concerner environ 110 000 questionnaires individuels.

2.2 La collecte *via* internet

La collecte de l'enquête Famille 2025 sera réalisée *via* internet pour les personnes qui répondent à l'enquête annuelle de recensement *via* internet. Cette partie de la collecte sera mise en œuvre dans le cadre de la filière Métallica des enquêtes auprès des ménages de l'Insee. Elle débutera quelques jours après le début de la collecte de l'EAR et sera un peu plus longue que la durée de collecte du RP, notamment du fait des relances auprès des non-répondants.

La partie collecte *via* internet de l'enquête Famille sera entièrement déconnectée de celle de l'EAR. Elle s'appuiera sur l'utilisation, à des fins d'enquête statistique, du mail fourni par les répondants à l'EAR. Après avoir renseigné *via* internet leur feuille de logement et leur(s) bulletin(s) individuel(s) du recensement, les ménages reçoivent un mail d'accusé de réception de la part de l'équipe du recensement de l'Insee. Dans les zones d'adresses concernées par l'enquête Famille, il est actuellement envisagé (*en cours d'instruction*) que ce mail soit spécifique et indique

métropolitaine, Antilles et Guyane et du jeudi 3 février au samedi 12 mars à La Réunion et à Mayotte.

aux ménages qu'ils vont être contactés pour répondre à l'EF. Ensuite, quelques jours après avoir validé leur réponse au RP, les ménages recevraient un mail les invitant à répondre à l'enquête Famille avec leur identifiant/mot de passe et le lien de connexion pour y répondre.

Le délai de quelques jours entre la validation de la réponse à l'EAR et l'envoi du mail pour répondre à l'EF aura en effet été mis à profit pour constituer un échantillon composé des ménages ayant validé leur réponse à l'EAR *via* internet et vivant dans des zones EF. Des échantillons seront ainsi constitués à intervalles réguliers, grâce aux données brutes collectées à l'EAR, et entrés (sélection des données utiles à préciser) en *input* de Métallica. Chaque fichier échantillon intégré à Métallica contiendra à la fois les informations de contact et des données sur la composition du ménage ayant répondu à l'EAR. La récupération des réponses fournies à l'EAR permettra ainsi de cibler les personnes devant répondre à l'enquête Famille (selon la zone, toutes les femmes ou tous les hommes de 18 ans ou plus du ménage). Cela permettra aussi d'élaborer un questionnaire de l'enquête plus adapté aux situations, sans reposer les questions de composition du ménage par exemple. Les ménages concernés ont probablement une tolérance moindre quant à l'absence de filtres et d'adaptation à leur situation que ceux qui répondent sur un questionnaire papier. Cette étape constituera donc une amélioration de la qualité de la collecte de l'enquête Famille par rapport aux éditions précédentes.

Ce point de jonction entre les deux filières est crucial pour l'enquête : la réussite technique (rapidité et fluidité de la circulation des informations entre les deux filières) est une condition indispensable à l'obtention d'un taux de réponse élevé à l'enquête Famille sur internet. Les tests nous diront si la rapidité de transition entre la réponse au recensement et la réception du mail-avis permettent à l'enquête internet de bénéficier d'un effet d'entraînement comme c'est le cas sur papier.

Cette partie de la collecte de l'EF est très fortement liée aux travaux de la filière Métallica des enquêtes auprès des ménages de l'Insee. L'offre standard de la filière Métallica sera mobilisée pour élaborer le questionnaire internet (avec filtres et contrôles), effectuer les envois de mails et de courriers, mettre en place le protocole de relances (par courrier, par mail voire passation du questionnaire par téléphone) et récupérer les données collectées en *output* de la plateforme de collecte *via* internet. L'EF bénéficiera également de l'offre de codification automatique proposée dans Métallica pour les trois nomenclatures qu'elle mobilise dans le questionnaire complémentaire à celui du RP (communes, pays et catégories socioprofessionnelles (PCS)).

De plus, en 2025, le questionnaire internet de l'enquête Famille sera *responsive design*, c'est-à-dire permettra aux enquêtés un accès, une lisibilité et une possibilité de répondre de manière homogène quel que soit le matériel dont ils disposent, y compris un smartphone ou une tablette.

Avec l'hypothèse de 70 % de réponses à l'EF sous format internet, cette partie de la collecte devrait concerner environ 260 000 questionnaires individuels, soit beaucoup plus qu'une enquête auprès des ménages classique menée à l'Insee.

Ce sont bien les enquêtés qui vont choisir sous quelle forme ils vont répondre à l'enquête Famille, sachant que ce mode de réponse sera identique à celui de l'EAR. Les deux modes de collecte (papier et internet) seront en concurrence puisque la possibilité de répondre sur internet sera offerte en même temps (ou avec un décalage de quelques jours) que la phase de collecte des questionnaires papier. On se placera donc sous un protocole de réponse **mixte**[14].

2.3 La relance et l'introduction d'un mode de collecte par téléphone

Si le taux de réponse au volet internet de l'enquête est faible, une relance par courrier postal puis par mail sera mise en place. Néanmoins, cette relance concernera uniquement les non-répondants à l'enquête Famille qui ont répondu à l'EAR *via* internet puisqu'il faudra connaître leurs nom-prénom et adresse mail⁸. Une relance par téléphone et la passation du questionnaire par téléphone sont également envisagées, si les relances mises en place ne suffisent pas à relever le taux de réponse à l'enquête Famille 2025. Elle serait menée sur un échantillon de non-répondants internet qu'il est prévu de sélectionner selon certains critères : par exemple, il contiendrait davantage d'adresses de certaines régions qui n'ont pas un taux de réponse suffisant. Puisque la relance s'adresse aux répondants à l'EAR, on pourra également, à partir des informations déclarées au recensement, sélectionner pour la relance téléphonique certaines catégories de populations rares qu'on aimerait étudier et dont le taux de réponse serait plus faible (couples de même sexe, familles recomposées, personnes en couple non co-résident, etc.). On se placera ici dans un mode de collecte séquentiel, la proposition d'une alternative n'étant faite qu'aux non-répondants de la phase antérieure [14].

Il reste à instruire comment récupérer le numéro de téléphone des personnes répondantes à l'EAR qu'on souhaite relancer à l'enquête Famille. Deux pistes sont actuellement envisagées : soit un appariement des nom-prénom-adresse avec le fichier démographique d'origine fiscale sur les logements et les personnes (Fideli) ; soit le recours à une société extérieure (de type La Poste ou Orange qui pourrait fournir l'information sur le numéro de téléphone). Cette relance téléphonique consisterait en un passage complet du questionnaire par téléphone par un prestataire extérieur à l'Insee (voir 2.4).

Actuellement, le nombre de non-répondants à relancer, que ce soit ou non par téléphone, n'est pas connu mais l'enquête test de 2023 (voir 2.4.2) devrait fournir une première estimation du taux de réponse à l'enquête. Le coût des relances par courrier postal ou par mail est relativement faible. En revanche, la taille de l'échantillon de non-répondants à relancer par téléphone sera liée au budget final dont on disposera pour l'enquête.

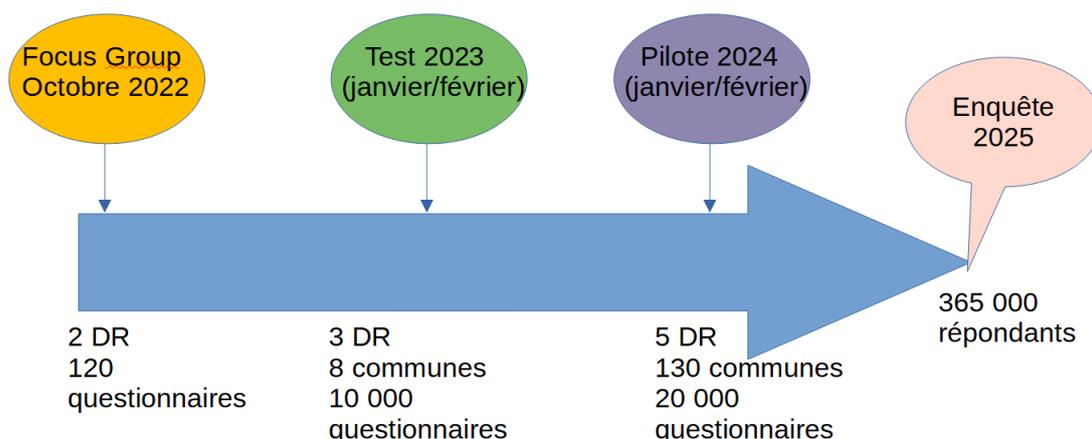
Le protocole consistant à mobiliser l'information disponible pendant la collecte (données recueillies par questionnaire ou grâce aux parodonnées) afin d'ajuster la stratégie de collecte pour le reste de l'échantillon est celui du *Responsive design* [14]. Il sera ainsi décidé, en cours de collecte de l'EF 2025, de porter l'effort sur un profil particulier en le ciblant directement ou indirectement en recourant à un mode de collecte particulier (le téléphone dans le cas présent). Ce type de protocole est innovant actuellement dans le cadre des enquêtes de la statistique publique française.

2.4 Les tests menés dans le cadre de ce nouveau protocole

Différents tests vont être menés pour anticiper la collecte de l'enquête Famille adossée à l'EAR 2025 (**figure 5**).

8. On rappelle que les nom-prénom des personnes ne sont pas connus au moment du tirage de l'échantillon de l'enquête Famille puisque ce sont des grappes de zones d'adresses qui sont tirées (voir 1.6). En théorie, cela génère un biais de relancer uniquement les répondants à l'EAR mais le taux de réponse à l'EAR en 2020 étant de 96 %, ce biais sera très limité.

Figure 5 : Calendrier des tests et collecte de l'enquête Famille



En lien avec les équipes du recensement, des enquêtes ménages et du programme Métallica, il a été construit un programme de tests et un calendrier qui s'échelonne de la façon suivante :

2.4.1 Focus group (automne 2022)

L'établissement régional de Rouen et la direction régionale de Nantes de l'Insee ont été choisis pour mener le Focus group de l'enquête Famille en octobre 2022. Une dizaine d'enquêteurs ménages feront remplir à des personnes de leur connaissance les questionnaires papier du recensement puis de l'enquête Famille et recueilleront leurs impressions et commentaires lors d'un entretien de bilan. Ce test sera donc réalisé uniquement sur papier. 120 questionnaires sont attendus.

En amont, un comité scientifique associant les principaux partenaires de l'enquête (Ined, Cnaf, Drees) se sera réuni plusieurs fois pour élaborer la version du questionnaire papier à tester lors de ce Focus group (voir 1.4).

Il est également envisagé que l'Ined, principal partenaire de l'enquête, organise un test en amont du Focus group : par exemple, deux versions du questionnaire pourraient être testées et plusieurs sous-échantillons spécifiques (par exemple des couples de même sexe, des familles monoparentales) seraient enquêtés, avec pour objectif de voir quelle version du questionnaire est la mieux adaptée pour eux. Au moment de la rédaction de cet article, il n'est pas encore décidé du calendrier ni de la taille de l'échantillon de ce test.

2.4.2 Premier test (T1 2023)

Ce premier test sera réalisé au moment de l'EAR 2023 auprès de trois directions régionales de l'Insee (deux en métropole et celle de La Réunion) dans 8 communes. Pour les enquêtés papier, ce test pourrait se dérouler comme l'enquête en vraie grandeur. Il est nécessaire de limiter le nombre de communes concernées à moins d'une dizaine pour ne pas perturber la collecte du recensement et gérer de façon très individualisée les relations avec ces communes. En revanche, le nombre de ménages contactés pourrait être important, autour de 10 000, notamment sur internet afin de pouvoir estimer le taux de réponse. Ce test permettra :

- de valider les deux versions des questionnaires (papier et internet) : il servira à vérifier que les ménages acceptent l'enquête, à voir si les questions sont comprises et si le déroulement est fluide. A l'issue de ce test, les questionnaires devraient être pour l'essentiel stabilisés ;
- d'avoir une première évaluation du rôle des différents acteurs lors de la collecte du recensement, leur charge de travail, leur formation et l'impact de l'enquête Famille sur la collecte du recensement ;
- de valider la procédure d'extraction et de transmission des échantillons en *input* du site de collecte en ligne ;
- d'estimer le taux de réponse sur internet : ce sera la première occasion de dérouler la collecte internet, de vérifier la façon dont les différentes étapes se succèdent (recensement sur internet, contact par mail pour répondre à l'EF quelques jours après puis saisie des réponses sur le site de collecte) ;
- si possible, de tester différents calendriers de relance par courrier postal et mails, pour voir quelles relances sont les plus efficaces (il n'y aura pas de relance téléphonique à ce test) ;
- de tester les étapes de maquettage, d'impression et de saisie pour la collecte papier : contractualisation, spécifications, échanges.

Pour valider le questionnaire, différentes informations seront disponibles. Il y aura d'abord les retours des différents acteurs de la collecte lors des bilans. Ensuite, l'analyse des réponses permettra d'apporter des éléments : taux d'abandon, taux de non réponse partielle aux différentes questions, modalités de réponse qui concentrent trop ou trop peu de répondants, cohérence des informations avec celles collectées au recensement (nombre et âge des enfants, situation conjugale, etc.) et différences de réponse entre les deux modes (même si c'est à considérer avec précaution car les profils des répondants sont différents).

Enfin, la bonne compréhension du questionnaire sur internet sera évaluée au moyen des paradonnées de l'enquête, c'est-à-dire les données d'environnement de la collecte. L'enjeu est d'identifier, en l'absence d'enquêteur, les passages du questionnaire mal compris ou nécessitant un effort cognitif trop important. Il s'agira pour cela d'analyser la non-réponse partielle et les temps de passation des différentes parties du questionnaire. On examinera aussi le recours à l'aide en ligne et on tâchera d'identifier si certaines questions ou parties du questionnaire concentrent des abandons. Enfin, il s'agira d'instruire la façon d'identifier les questionnaires de mauvaise qualité sur la base d'indicateurs fondés sur les paradonnées.

2.4.3 Enquête pilote (T1 2024)

Compte tenu des innovations importantes, il est indispensable de réaliser une enquête pilote de grande ampleur, adossée à l'EAR 2024, année précédant l'enquête. Cette enquête pilote sera réalisée dans 130 communes sélectionnées dans cinq directions régionales de l'Insee (au moins une dans les DROM) dans 130 communes. Elle pourrait porter sur environ 20 000 personnes enquêtées. Elle permettra :

- de valider la méthode d'échantillonnage sur le terrain en allant de la sélection des grappes à leur traitement par les agents recenseurs, en passant par les ajustements nécessaires pour adapter les grappes aux zones des agents recenseurs ;
- de tester le protocole de collecte de l'enquête Famille à grande échelle : vérifier la bonne insertion de l'enquête associée dans les processus de production du recensement, avoir une première mesure du taux d'acceptation de l'enquête par les communes et du bon déroulement de la procédure de conventionnement et d'améliorer le dispositif pour l'enquête en 2025 ;
- en fonction du taux de réponse, à nouveau évalué dans le cadre de cette enquête pilote,

de tester le protocole de relance téléphonique pris en charge par un prestataire extérieur, sur un sous-échantillon de non-répondants.

Cette enquête pilote de 2024 permettra d'apporter également des enseignements importants concernant les traitements aval (saisie des questionnaires papier, appariements, codification des professions, etc.).

2.5 Les principaux avantages et inconvénients de chaque mode de collecte

Les principaux avantages et inconvénients des différents modes de collecte de l'enquête Famille 2025 sont présentés ci-dessous :

2.5.1 Pour la collecte papier

Avantages : L'ensemble du protocole de collecte est maîtrisé. Du fait de l'effet d'entraînement apporté par le recensement, le taux de réponse devrait être proche de celui de l'édition de l'enquête de 2011 (plus de 80 %). La durée de la collecte est la même que celle de l'enquête annuelle de recensement et il n'y a pas de relances spécifiques à effectuer.

Inconvénients : Le questionnaire est moins « souple » : on ne peut pas introduire de filtres et les enquêtés doivent comprendre, en fonction de certaines de leurs réponses, par quelle question ils doivent enchaîner (on leur indique « aller à la question n°x »). Aucun contrôle des réponses fournies ne peut être mis en place pour analyser la cohérence entre les réponses. Les informations déclarées à l'EAR ne peuvent pas être récupérées en amont du remplissage de l'EF : il est donc indispensable par exemple, comme en 2011, de demander aux enquêtés de lister à nouveau les membres de leur famille pour compléter les questions les concernant, alors qu'ils l'ont déjà fait dans leur réponse au recensement. Il faut également saisir les données collectées ce qui occasionne un coût, un délai supplémentaire et parfois des erreurs de saisie.

2.5.2 Pour la collecte *via* internet

Avantages : La possibilité pour les enquêtés de pouvoir répondre *via* internet contribue à améliorer l'image de l'Insee : il n'est pas possible de demander à des personnes qui viennent de répondre à l'EAR par internet de répondre à l'enquête Famille par un seul mode, le papier. L'introduction de la collecte *via* internet va permettre d'obtenir des données de meilleure qualité. Dans la version internet du questionnaire de l'EF, les personnes de la famille, déclarées comme habitants du logement dans la feuille de logement de l'EAR, seront « affichées » dans le questionnaire et l'enquêté aura simplement à confirmer leur présence. De plus, des filtres et des contrôles seront intégrés lors de réponses aberrantes par exemple, et des listes de réponses pourront être affichées (pour les communes, les pays de naissance ou la nomenclature PCS) ce qui contribuera à l'amélioration de la qualité des données recueillies. Les enquêtés devraient apprécier ces facilités de remplissage, impossibles à implémenter dans la version papier.

Le multimode va permettre d'alléger la charge des agents recenseurs : pour tous les répondants au recensement *via* internet, la collecte de l'enquête Famille 2025 sera relativement transparente pour eux. Les correspondants de l'EF situés dans les établissements régionaux de l'Insee auront également une charge de travail moins importante puisqu'ils seront uniquement responsables du suivi de la collecte papier. Enfin, il n'y a aucun coût de saisie des questionnaires ce qui allège la

charge financière et réduit le délai de récupération des données.

Inconvénients : La présentation du questionnaire doit être complètement revue. Sous format papier, traditionnellement, le questionnaire était composé de tableaux de taille importante. En effet, pour décrire par exemple le nombre d'enfants vivant dans le logement (enfants de l'enquêté et de son conjoint), dans l'EFL de 2011, il y avait un tableau comportant 6 lignes (1 ligne par enfant) et 8 colonnes (1 colonne par caractéristique demandée). De même pour le nombre d'enfants de l'enquêté vivant hors du logement ou décédés, le tableau comportait 7 lignes et 8 colonnes. Pour le questionnaire en ligne, les tableaux sont fortement déconseillés. Il faut envisager de tout décomposer en posant une question différente pour chacune des caractéristiques attendues (les « colonnes ») pour chacun des enfants (les « lignes »). De plus, pour qu'il soit *responsive design*, le questionnaire en ligne doit contenir uniquement une question par page afin que le questionnaire puisse être renseigné sur smartphone et non pas sur ordinateur exclusivement. Ces changements dans la présentation du questionnaire pourraient entraîner une compréhension et un mode de remplissage peut-être différents de la part des enquêtés, voire des ruptures de séries par rapport aux éditions précédentes de l'enquête.

Enfin, le taux de réponse à l'EF 2025 est inconnu. En 2021, le taux de réponse aux enquêtes de l'Insee *via* internet tourne autour de 30 %. Les tests de l'EF menés pendant les EAR de 2023 et 2024 vont permettre d'obtenir une estimation du taux de réponse final dans le cadre du protocole adopté.

2.5.3 Pour la relance par téléphone

Les deux modes de collecte « principaux » que sont le papier et internet sont des modes de collecte auto-administrés, qui ont donc de nombreux points communs. En revanche, une relance par téléphone nécessitera la présence d'une personne qui pose les questions de l'enquête Famille directement aux enquêtés. Or, il existe deux biais principaux induits par la présence ou l'absence d'un enquêteur [14] :

- le biais de *satisficing* désigne le fait que certains enquêtés renoncent à fournir l'effort nécessaire pour renseigner la réponse exacte. Ce phénomène survient plus fréquemment avec des questionnaires auto-administrés, difficiles, longs ou répétitifs pour lesquels le niveau de concentration et d'engagement sont moindres que face à un enquêteur. Le biais devrait donc être plus fort sur papier et sur internet que par téléphone. Le questionnaire de l'enquête Famille est court mais il peut sembler répétitif puisque les mêmes caractéristiques sont demandées sur les différents membres de la famille (exemple, en 2011, les questions portant sur la mère et sur le père de l'enquêté sont strictement identiques). Il peut se manifester par des réponses approximatives, par une tendance à choisir plutôt les premières modalités proposées (*primacy effect*) ou les dernières (*recency effect*), à ne pas recourir à des documents, à choisir des modalités médianes ou encore par de la non-réponse partielle, voire un abandon du questionnaire ;
- la *désirabilité sociale* désigne le fait que certains enquêtés peuvent être amenés à fournir une réponse donnant une image valorisante d'eux-mêmes ou dont ils croient qu'elle satisfait les attentes normatives. Elle survient plus fréquemment lors d'une interrogation menée par un enquêteur et pour des questions d'opinion ou des questions qui peuvent renvoyer à une certaine image que l'on donne de soi et de ses relations sociales (par exemple la profession des parents, la fréquence des contacts, etc.). Ce biais devrait donc être plus marqué lors des relances téléphoniques.

Avantages : La présence d'un enquêteur pourra permettre d'améliorer la compréhension des questions et des attendus ; le fait que l'enquêteur déroule le questionnaire en ligne par téléphone permettra de bénéficier des filtres et contrôles mis en place et améliorera la qualité du questionnement et des réponses.

Inconvénients : Les enquêteurs, qui plus est, extérieurs à l'Insee, vont interroger une population particulièrement réfractaire à répondre à l'enquête puisque la relance téléphonique fera suite à différentes relances antérieures. La relance téléphonique va allonger la durée de collecte de l'enquête Famille. De plus, ce troisième mode de collecte va nécessairement complexifier ensuite la réconciliation et l'exploitation des données. Enfin, la prise en charge d'un volet téléphonique par un prestataire extérieur sera totalement inédite pour une enquête auprès des ménages menée par l'Insee pour ses deux volets principaux.

3 La richesse des données de l'enquête Famille 2025

Les différences de mode de collecte de l'EF 2025 devront être *in fine* aplanies afin que les fichiers issus des deux (voire trois) modes de collecte soient réconciliés en un seul, pour produire des résultats de qualité, représentatifs et comparables à ceux de l'EFL de 2011. Les données collectées seront ensuite enrichies des données du recensement (avec de nouvelles variables disponibles, en comparaison avec le RP de 2011) et de celles disponibles dans certains fichiers administratifs (par appariements *ex-post*). Les données permettront ainsi d'aborder de nombreux sujets et de produire des études très complètes et diversifiées : publications scientifiques, ouvrages et chapitres d'ouvrage, thèses, etc.

3.1 La réconciliation des données de l'enquête

La refonte du protocole de collecte de l'EF 2025 va induire une réingénierie des traitements statistiques. L'introduction du multimode proposé aux enquêtés va générer différents modes de réponse possibles (papier, internet voire téléphone en cas de relances) pour les enquêtés et devra être pris en compte dans le traitement de la non-réponse. Les modèles de non-réponse pourraient cependant gagner en pouvoir explicatif à s'appuyer sur les paradonnées retraçant la difficulté à joindre les ménages et la chronique des différentes tentatives pour pouvoir les contacter. Une attention particulière devra être portée aux éventuels effets de mode, et à la comparabilité des données entre l'enquête Famille 2025 et les enquêtes précédentes. En 2026, de nombreux dispositifs l'auront précédé dans ce passage au multimode, ce qui permettra d'avoir de supports d'appui et de comparaison (les enquêtes de victimation, SRCV, l'enquête Emploi, les enquêtes Logement).

Les opérations statistiques d'apurement, de codification, de pondération, de documentation et de diffusion seront comme en 2011, prises en charge par la maîtrise d'ouvrage de l'Insee, avec un appui éventuel (toujours comme en 2011) de l'Ined et de la Drees. Pour toutes ces étapes, les outils en cours de développement à l'Insee, notamment issus de la filière Métallica, devront être mobilisés.

3.2 Les nouvelles variables issues du RP

L'appariement des données de l'EF 2025 avec les données du recensement permettra de récupérer l'ensemble des informations collectées au recensement (feuille de logement, bulletins individuels de tous les habitants du logement et traitements post-collecte) pour les personnes

ayant répondu à l'enquête Famille sur papier ou *via* internet (ou téléphone). Cette opération va être complexifiée par rapport aux éditions précédentes de l'enquête puisqu'il va falloir gérer les identifiants différents entre la filière RP et la filière Métallica. Lors des éditions précédentes, le taux d'appariement entre l'enquête Famille et les données du RP était de 99 %. La réconciliation des deux filières, que ce soit du point de vue collecte ou du point de vue appariement et restitution des données va induire des travaux du côté des équipes, en charge de Métallica comme de celles du recensement, que la maîtrise d'ouvrage va devoir piloter.

La reprise et la saisie des listes ABC⁹ ainsi que l'extension de l'analyse ménages-familles¹⁰ ont été des opérations très coûteuses pour l'édition 2011 de l'enquête Famille : elles ne seront plus nécessaires pour l'enquête Famille 2025 grâce à la refonte en 2018 de la feuille de logement et de l'analyse ménages-familles du RP, désormais réalisée pour tous les ménages. La refonte de la feuille de logement à partir de l'EAR 2018 a ainsi permis d'améliorer les informations collectées, concernant la description des liens familiaux et de la multirésidence (qui était le thème spécifique de l'EFL de 2011).

Des évolutions concernant la codification des professions dans le recensement sont également prévues avant l'EF 2025, notamment avec l'introduction de la PCS 2020¹¹. Le coût des traitements des PCS pour l'enquête Famille 2025 par rapport à celle de 2011 devrait ainsi être revu à la baisse.

Enfin, l'introduction de la PCS2020 permettant de réduire le questionnement, il pourrait être envisagé d'introduire trois nouvelles variables dans le bulletin individuel du recensement à partir de 2024¹²) : concernant le télétravail, l'indicateur GALI de limitation d'activités fonctionnelles, et les départements et pays de naissance des parents (qui étaient posées dans le questionnaire de l'enquête Famille). Un tel enrichissement permettrait de supprimer ces deux questions sur les parents dans le questionnaire de l'enquête Famille 2025 et de gagner de la place pour pouvoir introduire des questions supplémentaires. L'introduction du GALI serait également utile pour l'EF 2025, pour la connaissance par exemple des situations familiales des personnes ayant des limitations fonctionnelles. La décision d'inclure ou non ces variables dans le bulletin individuel du RP pourrait avoir lieu en 2022 (suite à la présentation de résultats de tests menés par l'équipe du recensement au Cnis en juin 2022).

3.3 Les appariements avec d'autres sources

Plusieurs pistes d'appariements des données de l'EF 2025 avec d'autres sources de données administratives sont envisagées et en cours d'instruction.

9. Il s'agissait des tableaux des habitants permanents et non permanents, avec des liens familiaux indiqués en clair, qui n'étaient pas saisis dans le cadre des traitements du recensement, et avaient été saisis pour les répondants à l'enquête Famille en 2011. Il n'y a plus de liens en clair sur la feuille de logement depuis l'EAR 2018, ce qui fait disparaître les coûts de saisie des listes ABC pour la nouvelle enquête Famille.

10. L'Analyse Ménage Famille (AMF) était réalisée sur l'exploitation complémentaire du recensement, car elle nécessitait dans certains cas des traitements manuels. Depuis la refonte de l'EAR de 2018, les traitements sont automatiques et intégrés à l'exploitation principale. Il n'y a plus de surcoût à réaliser l'AMF sur tous les ménages concernés par l'enquête Famille.

11. Réunion de la commission démographie et questions sociales du 2 décembre 2021, consacrée au projet de refonte du questionnaire individuel du recensement, <https://www.cnis.fr/evenements/demographie-et-questions-sociales-2021-2e-reunion/?category=1092>

12. Réunion de la commission démographie et questions sociales du 2 décembre 2021, consacrée au projet de refonte du questionnaire individuel du recensement, <https://www.cnis.fr/evenements/demographie-et-questions-sociales-2021-2e-reunion/?category=1092>

En premier lieu, il est envisagé de mettre en place un suivi passif tous les ans, pendant 5 ans par exemple, des personnes de l'enquête Famille présentes dans les fichiers fiscaux et sociaux (actuellement le fichier démographique d'origine fiscale sur les logements et les personnes - Fideli), afin de suivre les ruptures d'unions ou les remises en couple et les naissances. Cela permettra également d'obtenir toutes les informations fiscales et sociales présentes dans le fichier Fideli (évolution du niveau de vie, type de ressources disponibles dans le ménage dont prestations sociales, enfants à charge).

En deuxième lieu, sera instruite la possibilité d'appariement avec le fichier des séparations dans le cadre de l'ARIPA (agence de recouvrement des impayés de pensions alimentaires) de la CNAF, qui sera disponible à partir de 2025. Il permettra de disposer d'éléments sur les séparations et les pensions alimentaires versées.

En troisième lieu, il est envisagé d'apparier les données de l'EF avec les données de l'état civil pour le suivi passif de la mortalité, par exemple à 5 ans puis à 10 ans (vérification dans le RNIPP à partir du NIR), comme cela a été le cas pour l'enquête Famille de 1999.

Pour réaliser ces appariements, on souhaite s'appuyer autant que possible sur le CSNS (code statistique non significatif). Si cela s'avérait nécessaire, les traits d'identité seront mobilisés, comme pour l'Échantillon Démographique Permanent, pour retrouver les personnes enquêtées à l'enquête Famille, à partir de leur sexe, nom-prénom¹³, date et lieu de naissance, afin de retrouver leur NIR au RNIPP et de traiter ensuite les appariements (dans un cadre conforme aux réglementations en vigueur).

4 Les défis et les risques de ce nouveau protocole de collecte

La réalisation de l'enquête Famille 2025, enquête très attendue notamment par les partenaires de l'Insee tels que l'Ined, la CNAF et la Drees s'avère donc être un défi important compte tenu de toutes les nouveautés envisagées, notamment concernant le protocole de collecte des données.

À l'Insee, le séminaire de lancement d'une enquête a justement pour objectif de décrire les défis auxquels les équipes reponsables d'un projet vont devoir faire face, et d'analyser les risques afin de trouver des solutions pour les limiter. Ce séminaire doit également permettre de réunir et fédérer tous les acteurs de l'enquête, que ce soit du point de vue de la maîtrise d'ouvrage statistique, de la maîtrise d'œuvre informatique et des partenaires. Au moment de la rédaction de cet article, le séminaire de lancement n'a pas encore pu être organisé pour la nouvelle édition de l'enquête Famille. Les éléments décrits ci-dessous sont donc partiels et à confirmer ultérieurement.

4.1 Les principaux défis

Le premier défi est l'obligation de revoir profondément l'organisation de la collecte de l'enquête. Auparavant entièrement adossée au recensement de la population, l'enquête Famille bénéficiait de la compétence des équipes du recensement, de son obligation légale et de son effet d'entraînement concernant le taux de réponse. Le passage en deux filières (« système recensement » pour la version papier et « filière Métallica » pour la version internet) va complexifier le mode de collecte : il va falloir gérer les interactions entre ces deux systèmes, ce qui n'a jamais été

13. dont la qualité de saisie sera pour les enquêtés de l'EAR appartenant à l'enquête Famille et répondants par papier plus exigeante (dite de niveau « qualité EDP »).

réalisé jusqu'à présent. Compte-tenu des charges pesant sur le recensement, il faut avoir le moins de modifications possibles dans ses chaînes de traitement. La filière Métallica est pour sa part en cours de construction à l'Insee et sa montée en charge progressive ne pourra pas permettre d'offrir toutes les fonctionnalités attendues dès les tests réalisés en amont de l'enquête Famille par exemple.

Le deuxième défi sera la relance téléphonique. Dans un premier temps, il faut instruire comment récupérer une information fiable sur les numéros de téléphone et dans un délai relativement court. Ensuite, la relance téléphonique sera un défi d'une part, par la sélection de l'échantillon de non-répondants à contacter par téléphone et d'autre part par le fait que cette relance sera prise en charge par un prestataire extérieur. Ces deux éléments sont à l'heure actuelle tout à fait inédits à l'Insee.

Enfin, le troisième défi pour les équipes informatiques de l'Insee va être de pouvoir absorber une charge importante de connexions simultanées sur le site de collecte internet de l'enquête. Les ménages vont être incités à se connecter au site pour répondre à l'enquête dès la réception du mail contenant les informations de connexion. Les ménages ayant tendance à réagir rapidement aux sollicitations, il est probable que le premier week-end suivant le début du lancement de l'enquête Famille soit un moment propice à beaucoup de connexions. Or la taille de l'échantillon de l'EF est beaucoup plus importante que celle des autres enquêtes auprès des ménages menées par l'Insee.

Il faudra également être vigilant à la collecte de l'enquête réalisée pour la première fois dans les DROM avec sans doute une communication et des formations à adapter pour que la collecte de l'enquête Famille 2025, adossée à celle du recensement, se déroule au mieux dans ces départements.

4.2 Les principaux risques

Les principaux risques identifiés à ce stade sont :

- le risque informatique de pouvoir faire se réconcilier les deux systèmes RP et Métallica : les fichiers issus des deux (voire trois) modes de collectes devront être réconciliés en un seul, pour produire des résultats de qualité, représentatifs et comparables à ceux de l'enquête de 2011 ;
- les risques liés aux nombreux acteurs impliqués dans le projet : un élément clé de la réussite de l'enquête Famille est également la coordination des différents acteurs, des rôles et calendriers des opérations attendues par chacun. La complexification du protocole par rapport à l'édition de 2011 engendre aussi une complexification de l'organisation en interne à l'Insee. Un suivi et des échanges réguliers sont en œuvre, et le séminaire de lancement sera un moment de formalisation des échanges important, notamment en termes de comitologie.
- le risque de ne pas contacter « la bonne personne » : l'adresse mail qui sera utilisée pour envoyer les liens de connexion au site internet de l'enquête Famille est associée à la personne qui a répondu au recensement. Or, il se peut que la personne destinataire de ce mail ne soit pas la personne concernée par le questionnaire de l'enquête Famille : si on est dans une zone de collecte « Hommes » et que les bulletins de recensement (Feuille de logement et Bulletin individuel) ont été renseignés par une femme du logement, ce

sera elle le point de contact pour le recensement et donc la destinataire du mail-avis de l'EF. Dans quelle mesure cette femme va-t-elle transmettre les liens de connexion à tous les hommes majeurs de son logement ? Va-t-elle remplir les questionnaires destinés aux hommes adultes de son logement, avec le risque d'avoir beaucoup de réponses inexactes ou non renseignées ?

Les proportions de non-réponse ou de réponses par *proxy* liées à cette difficulté pourront être évaluées lors des tests de l'EF et notamment dès le test de 2023. Plusieurs éléments peuvent être envisagés afin de limiter le phénomène : lister dans le mail-avis les prénoms des personnes du logement qui doivent répondre à l'enquête (information qui pourra être récupérée à partir des bulletins de l'EAR), avoir des consignes plus ou moins strictes concernant les *proxy*, etc.

Il convient de noter que ce risque était déjà présent lors des enquêtes Famille précédentes et le sera aussi dans le volet papier de l'enquête : si une seule personne remplit pour le logement la feuille de logement du recensement, les bulletins individuels, les bulletins enquête Famille sont sans doute remplis également par cette personne, qui n'est pas forcément celle directement concernée par l'EF.

- le risque de baisse du taux de réponse : le délai de quelques jours entre la réponse à l'EAR et l'envoi du mail-avis de l'EF 2025 va limiter l'effet d'entraînement du recensement qui aboutissait, lors des éditions précédentes de l'enquête Famille, à des taux de réponse très élevés. Ce risque pourra être anticipé lors des tests de l'EF et notamment dès le test de 2023. Il s'agira ensuite de prendre des mesures adéquates telles que les relances par mail, par courrier, voire par téléphone. Il est aussi envisagé de demander le caractère obligatoire de réponse pour l'EF 2025. Il est estimé que le caractère obligatoire d'une enquête peut entraîner une hausse d'environ 10 points du taux de réponse pour les enquêtes ménages à l'Insee menées en face à face.

Conclusion

Le dispositif de collecte de l'enquête Famille 2025 est en cours de construction à l'heure des JMS 2022. Sa genèse et sa conception en font un dispositif d'enquête tout à fait innovant, entre collecte papier adossée à l'enquête de recensement de la population et questionnaire auto-administré sous internet *via* la filière Métallica des enquêtes ménages de l'Insee. La collecte de l'enquête est donc soumise à des risques et des défis que l'Insee et ses partenaires vont devoir relever. En règle générale, l'évolution vers le multimode constitue une occasion d'améliorer les enquêtes sur de nombreux aspects, mais elle représente aussi un coût organisationnel et un défi pour la gestion des effets de mode : ceci nous invite à apporter le plus grand soin dans les choix méthodologiques envisagés au moment de la conception de l'enquête Famille 2025. L'enquête est en effet très attendue au sein de l'Insee mais aussi par ses principaux partenaires¹⁴ et il est absolument essentiel que, malgré le protocole de collecte innovant, on puisse aboutir, sur les nombreux sujets démographiques abordés dans l'enquête, à des résultats fiables et comparables à ceux de l'édition précédente de l'enquête en 2011.

14. 13 organismes et une dizaine de chercheurs étaient présents lors de la réunion de recueil des besoins sur le questionnaire de l'enquête Famille 2025, organisée le 7 janvier 2022.

Bibliographie

- [1] Pascale BREUIL-GENIER et al. “Enquête Famille et logements adossée au recensement de 2011 : comment s’adapter à la nouvelle méthodologie des enquêtes annuelles et quels apports?” In : *Economie et statistique* 483.1 (2016), p. 205-226. URL : <https://www.insee.fr/fr/statistiques/fichier/2017652/ES483J.pdf>.
- [2] Élisabeth ALGAVA. “La nouvelle feuille de logement et la refonte de l’analyse Ménages-Familles en 2018 : Quels apports pour l’étude des structures familiales?” In : *Document de travail F2021-01 INSEE* (2021). URL : <https://www.insee.fr/fr/statistiques/5351799>.
- [3] Marceline BODIER et al. “Couples et familles”. In : *INSEE Références* (2015). URL : <https://www.insee.fr/fr/statistiques/2017528>.
- [4] Christophe IMBERT, Éva LELIÈVRE, David LESSAULT et al. “La famille à distance : Mobilités, territoires et liens familiaux”. In : *collections Questions de population, Ined* (2018). URL : <https://www.ined.fr/en/publications/editions/questions-populations/la-famille-a-distance/>.
- [5] Marie ACS, Bertrand LHOMMEAU et Émilie RAYNAUD. “Les familles monoparentales depuis 1990”. In : *Dossiers Solidarité Santé n°67, DREES* (2015). URL : <https://drees.solidarites-sante.gouv.fr/publications/dossiers-solidarite-et-sante-1998-2016/les-familles-monoparentales-depuis-1990>.
- [6] Florent DOMERGUE. “Familles d’immigrés et de descendants : entre spécificités et convergence vers le pays d’accueil”. In : *Infos migrations n°86* (2016). URL : <https://www.immigration.interieur.gouv.fr/Info-ressources/Actualites/Focus/Familles-d-immigres-et-de-descendants-entre-specificites-et-convergence-vers-le-pays-d-accueil>.
- [7] Emma DAVIE. “Les agents de la fonction publique et leur famille en 2011. Plus d’un quart des agents de la fonction publique en couple ont un conjoint lui aussi dans la fonction publique”. In : *Point Stat, Statistiques et recherche sur la fonction publique, DGAFP* (2014). URL : http://www.fonction-publique.gouv.fr/files/files/statistiques/point_stat/Point_Stat_famille_v7.pdf.
- [8] Wilfried RAULT. “Les mobilités sociales et géographiques des gays et des lesbiennes. Une approche à partir des femmes et des hommes en couple”. In : *Sociologie, 7(4), pp 337-360* (2016). URL : <https://www.cairn.info/revue-sociologie-2016-4-page-337.htm>.
- [9] Martial MAILLARD et Patrick CATTELIN. “Les grands-parents picards ont six petits-enfants en moyenne”. In : *INSEE Picardie Flash n°11* (2013). URL : <https://www.insee.fr/fr/statistiques/1560133>.
- [10] Loïc TRABUT et al. “Famille et recensement font-ils bon ménage?” In : *Population-F n°70 (3), pp 603-630, Ined* (2015). URL : <https://www.cairn.info/revue-population-2015-3-page-637.htm>.
- [11] Guillemette BUISSON et Aude LAPINTE. “Les structures familiales en France : comparaison entre le recensement, l’enquête Famille et logements et l’enquête emploi”. In : *Document de travail DSDS F1703, INSEE* (2017). URL : <https://www.insee.fr/fr/statistiques/2847257>.

- [12] Heidi KOUMARIANOS et Amandine SCHREIBER. “Conception de questionnaires auto-administrés”. In : *INSEE, Document de travail n°2021-03* (2021). URL : <https://www.insee.fr/fr/statistiques/6010788>.
- [13] Pascal ARDILLY. “Estimations communales exploitant les données de l’enquête Famille et logements 2011 et du recensement : une opération à hauts risques”. In : *Économie et Statistique* 483.1 (2016), p. 227-258. URL : <https://www.insee.fr/fr/statistiques/fichier/2017654/ES483K.pdf>.
- [14] François BECK et al. “Le multimode dans les enquêtes auprès des ménages : une collecte modernisée, un processus complexifié”. In : *Courrier des statistiques n°7* (2022). URL : <https://www.insee.fr/fr/information/6035950>.