

DREES MÉTHODES

N° 7 • novembre 2022

Enquête santé européenne (EHIS) 2019 : Bilan méthodologique

France métropolitaine et DROM

Aude Leduc (DREES-Insee)

Enquête santé européenne (EHIS) 2019 : Bilan méthodologique

France métropolitaine et DROM

Aude Leduc (DREES-Insee)

Remerciements : Philippe Raynaud pour ses précieux conseils tout au long de ce projet ;
Benoît Ourliac pour sa relecture attentive du document.

Retrouvez toutes nos publications sur : drees.solidarites-sante.gouv.fr

Retrouvez toutes nos données sur : data.drees.solidarites-sante.gouv.fr



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*



DREES MÉTHODES

N° 7 • novembre 2022

Synthèse

Enquête santé européenne (EHIS) 2019 :

Bilan méthodologique

France métropolitaine et DROM

Aude Leduc (DREES-Insee)

Retrouvez toutes nos publications sur : drees.solidarites-sante.gouv.fr

Retrouvez toutes nos données sur : data.drees.solidarites-sante.gouv.fr



STATISTIQUE
PUBLIQUE

La DREES fait partie
du Service statistique
public piloté par l'Insee.

SYNTHÈSE

À la suite de la publication des premiers résultats de l'enquête (Leduc *et al.*, 2021), ce document dresse un bilan méthodologique du projet EHIS (European Health Interview Survey) en France métropolitaine et dans les cinq DROM. La partie France métropolitaine de cette enquête répond au règlement européen 1338/2008 de la Commission européenne et a vocation à être rééditée tous les 6 ans.

Il s'agit à la fois de documenter l'enquête 2019 et de mettre en avant ce qui pourrait être amélioré pour une future édition de l'enquête. Ce document est divisé en deux parties portant sur la première sur l'enquête en France métropolitaine et pour la seconde sur l'enquête dans les DROM. En effet, l'enquête en France métropolitaine a été réalisée en collaboration avec l'Irdes et le terrain a été porté par un prestataire extérieur (Kantar) tandis que l'enquête dans les DROM a été faite en collaboration avec l'Insee qui en a assuré sa maîtrise d'œuvre.

Chaque partie commence par un rappel du calendrier détaillé de toutes les étapes de l'enquête et un bilan du questionnaire. Concernant le questionnaire, les points d'attention quant à sa bonne mise en conformité par rapport au règlement européen sont mis en avant et pour les DROM, un bilan est dressé de l'ajout des questions spécifiques aux DROM qui se sont avérées pour beaucoup ne concerner qu'un nombre restreint de répondants.

Les plans de sondage concernant la France métropolitaine et les DROM sont décrits. La comparaison des résultats de l'enquête DROM avec ceux de l'enquête SRCV 2018 tirée dans le recensement ne met pas en avant des biais particuliers à avoir échantillonné l'enquête pour les quatre DROM historiques dans Fidéli. Les méthodes de calcul de la pondération sont précisées dans deux notes complémentaires, l'une pour la France métropolitaine, l'autre pour les DROM, disponibles sur la page¹ consacrée à EHIS sur le site Internet de la DREES. La méthodologie des calculs de variance menés en France métropolitaine conformément aux attentes du rapport qualité européen est également explicitée dans une note complémentaire. Dans les DROM, les calculs de variance prenant en compte le plan de sondage n'ont pas été menés, faute de temps.

L'enquête EHIS a été une des premières enquêtes à avoir été tirée dans Fidéli. Par ailleurs, son volet France métropolitaine avait une collecte multimode (téléphone/face-à-face). Aussi, il était intéressant d'analyser quels retours la collecte de EHIS pouvait apporter à la connaissance de Fidéli comme base de sondage pour les enquêtes « ménages ». Ainsi, les parties « protocole » et « bilan de collecte » mettent en avant l'usage qui a été fait de Fidéli et notamment des nombreux numéros de téléphone à disposition pour la collecte et leur rentabilité en termes de collecte. Ce bilan est permis par l'analyse des données décrivant la collecte de l'enquête mises à disposition par le prestataire, notamment l'historique des appels.

Dans le protocole de l'enquête France métropolitaine, un échantillon méthodologique avait été introduit pour évaluer les effets de mode, plus précisément les effets de mesure. Celui-ci a été exploité. La partie « effet de mode » du document donne les résultats concernant les effets de mesure relatifs à un questionnaire santé entre le téléphone et le face-à-face ainsi qu'entre le téléphone et l'auto-administré. En effet, lorsque les enquêteurs étaient en face-à-face, ils devaient à la fin du questionnaire prêter leur tablette à l'enquêté pour que celui-ci renseigne par lui-même le module « sensible » (portant sur la santé mentale, la consommation de tabac et d'alcool) directement sur la tablette. La collecte a pu donc être faite en autoadministré sur une partie du questionnaire et de l'échantillon.

Dans les DROM, cette partie de questionnaire a eu lieu sous casque. Pour la passation en Guyane et à Mayotte, la partie sous-casque avait été traduite en différentes langues, trois pour la Guyane et deux pour Mayotte. Le document dresse un bilan de l'intérêt de ce travail de traduction mettant en avant que, mis à part le shimaore à Mayotte, les autres langues ont été très peu utilisées pendant la collecte.

Enfin, que ce soit en France métropolitaine ou dans les DROM, le document met bien en avant la qualité de l'appariement de l'enquête avec le SNDS. La méthode de collecte du NIR a été différente selon le territoire (indirecte en France métropolitaine et déclarative dans les DROM) mais l'appariement donne des résultats très satisfaisants avec un taux d'appariement proche de 95 % en France métropolitaine, aux Antilles, à La Réunion, de l'ordre de 83 % en Guyane et de 40 % à Mayotte. Les données seront également bientôt appariées avec les fichiers fiscaux et sociaux.

Ce document accompagné de la documentation présente sur la page Internet du site de la DREES dédiée à l'enquête fournit une documentation très complète pour les utilisateurs de l'enquête et plus généralement pour ceux qui s'intéressent aux enquêtes portant sur la santé et au développement du multimode dans les enquêtes ménages.

¹ <https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/enquete-sante-europeenne-ehis>

SOMMAIRE

■ INTRODUCTION	3
Contexte européen	3
Extension aux départements et régions d'outre-mer	4
Partenariats et coût de la troisième édition	4
■ ENQUÊTE SANTÉ EUROPÉENNE : FRANCE MÉTROPOLITAINE	5
Calendrier	5
Questionnaire	10
Variables européennes	10
Variables françaises	13
Durée du questionnaire	13
Bilan enquêteurs sur le questionnaire	15
Documents de collecte	15
Plan de sondage	16
Protocole	19
Description	19
Bilan de l'enrichissement téléphonique	20
Caractéristiques sociales des échantillons du terrain téléphone et face-à-face	21
Type de numéro de téléphone disponible selon les caractéristiques sociales dans l'échantillon du terrain téléphone	24
Bilan de collecte	29
Taux de réponse	29
Déroutement du terrain téléphone : très différent selon la vague d'enquête	32
Déroutement du terrain face-à-face : relativement similaire selon la vague d'enquête	40
Correction de la non-réponse partielle	44
Étude des effets de mode : exploitation de l'échantillon méthodologique	45
Contexte et design expérimental	45
Méthodes d'analyse	46
Résultats et interprétations	47
Appariements	57
Système national des données de santé (SNDS)	57
Fichiers fiscaux et sociaux	60
■ ENQUÊTE SANTÉ DOM	61
Calendrier	61
Questionnaire	66
Questions spécifiques aux DROM	66
Durée du questionnaire	72
Documents de collecte	73
Plan de sondage	74
Description du plan de sondage dans les quatre DROM historiques	74
Mise en perspective du tirage dans Fidéli pour les quatre DROM historiques	76
Description du plan de sondage à Mayotte	79
Protocole	80
Bilan de collecte	81
Taux de réponse	81
Taux de proxy	82
Matériel à Mayotte (balance et télémètre)	82
Difficultés avec les deux Kish	83
Bilan du casque	84
Traductions	85
Non-réponse partielle et module alcool	88
Correction de la non-réponse partielle	89

Appariements.....	90
Système national des données de santé (SNDS)	90
Données fiscales et sociales	94
■ CONCLUSION : MISE À DISPOSITION DES RÉSULTATS, DES DONNÉES ET DE LA DOCUMENTATION	95
■ POUR EN SAVOIR PLUS.....	96
Annexe 1. Effets marginaux moyens des régressions logistiques relatives au fait de ne disposer d'aucun numéro de téléphone dans le fichier et d'un numéro de mobile pour la personne échantillonnée dans l'échantillon téléphone.....	97
Annexe 2. Nombre de non-réponse partielle par question santé (France métropolitaine et DROM) 99	
Annexe 3. Distribution du score de propension entre échantillon méthodologique (face-à-face) et échantillon téléphone	106

■ INTRODUCTION

Contexte européen

L'*European Health Interview Survey* (EHIS) est régie par le règlement cadre 1338/2008. L'enquête santé européenne 2019 correspond à la troisième édition de EHIS qui est rendue obligatoire dans les pays de l'Union européenne par le règlement d'application 2018/255. L'enquête a été conduite dans tous les pays de l'Union européenne ainsi qu'en Islande, en Norvège, en Serbie et en Turquie. Il s'agit d'une enquête généraliste sur la santé, le recours aux soins et les habitudes de vie en lien avec la santé.

La première édition de l'enquête n'était pas obligatoire (*gentlemen's agreement*) mais la France y a participé au travers de l'enquête Handicap Santé ménages en 2008-2009 réalisée en collaboration entre la DREES et l'Insee. La seconde édition de l'enquête, cette fois rendue obligatoire par le règlement 141/2013, a également été portée en France par une autre enquête, à savoir l'enquête Santé et protection sociale, un panel historiquement conçu par l'Irdes. Elle a ainsi fait l'objet d'une collaboration entre la DREES et l'Irdes. L'édition 2019 est la première qui fait de EHIS une enquête indépendante alors que l'enquête Santé et protection sociale de l'Irdes cesse. De ce fait, cette édition a été l'occasion de revoir la méthodologie de l'enquête.

Une enquête sous règlement européen est contrainte. Pour la troisième édition de EHIS, le règlement impose :

- l'année de collecte (2019) ;
- la durée de collecte (au moins trois mois dont un mois d'automne) ;
- le champ (les personnes âgées de 15 ans ou plus, les départements et collectivités d'outre-mer sont exclus) ;
- une liste d'indicateurs qui doivent être renseignés grâce aux réponses à l'enquête ;
- un niveau de précision à atteindre sur la part de personnes en situation de handicap (fortement limitées) au sens du GALLI.

On voit que le mode de collecte est laissé au libre choix des pays. Il y a une certaine marge de manœuvre sur le questionnaire dans la manière de formuler les questions qui permettront de renseigner les indicateurs demandés par Eurostat. Des questions type sont proposées en anglais dans le manuel méthodologique afférent à l'enquête (Eurostat, 2018). Rien n'est imposé quant au plan de sondage tant que le niveau de précision est atteint. Il est en effet très dépendant du type de données dans lesquelles les enquêtes sont échantillonnées qui peut varier selon les pays.

Il est possible de demander des dérogations au règlement européen. Par exemple, la Belgique a réalisé la troisième édition de l'enquête en 2018 car celle-ci est portée par une enquête nationale réalisée à rythme régulier (règlement d'application relatif aux dérogations 257/2018). Des dérogations portent également sur les indicateurs à renseigner. Par exemple, pour la seconde édition, la France avait demandé à ne pas transmettre les variables relatives à la consommation d'alcool pour poser à la place le module AUDIT-C historiquement présent dans l'enquête Santé et Protection sociale de l'Irdes.

Pour la troisième édition de l'enquête, alors que la conception de celle-ci était totalement revue en France, il a été décidé de se rapprocher le plus possible du règlement européen de manière à favoriser les comparaisons internationales. Ainsi, aucune dérogation n'a été demandée.

En France, l'enquête dont le questionnaire se rapproche le plus de celui de EHIS est le Baromètre Santé de Santé publique France². Il s'agit d'une enquête sur la santé initiée en 1992 et réalisée très régulièrement (tous les ans ces dernières années) qui est plutôt généraliste mais éclaire selon les années des thématiques spécifiques. À la date de 2019, la méthodologie de cette enquête est très différente de celle de EHIS puisqu'il s'agit d'une enquête téléphonique basée sur la génération aléatoire de numéros de téléphone.

La quatrième édition de EHIS sera réalisée dans le contexte du nouveau règlement cadre relatif aux statistiques sociales européennes (*Integrated european social statistics – IESS*) qui homogénéise les attentes sur l'ensemble des enquêtes ménages européennes. Certains aspects de ce nouveau règlement étaient anticipés dans le règlement relatif à la troisième édition, par exemple les attentes en matière de précision. Mais des modifications sont à attendre de l'inscription de EHIS dans le contexte de IESS. Ces aspects pourront être investigués avec l'Insee qui a réfléchi au basculement de ses enquêtes européennes dans IESS (par exemple aux règles de rattachement des personnes au ménage dans le cadre de la refonte du tronc commun des ménages). En 2019, l'enquête santé est la seule enquête ménages européenne sous règlement obligatoire à ne pas être sous maîtrise d'ouvrage Insee.

² Voir : <https://www.santepubliquefrance.fr/etudes-et-enquetes/barometres-de-sante-publique-france#block-65430>

Extension aux départements et régions d'outre-mer

Même si le règlement ne l'obligeait pas, l'enquête a été étendue aux cinq DROM. Cela faisait suite au constat fait par la Cour des comptes du manque de données relatives à la santé dans les DROM publié en juin 2014 dans un rapport intitulé « La Santé dans les outre-mer : une responsabilité de la République ». On pouvait y lire : « Les dispositifs de recueil des données socio-sanitaires doivent être rapidement confortés afin de prendre pleinement la mesure des écarts et retards les plus préoccupants et d'adapter en conséquence les actions correctrices à court et moyen terme. À cette fin, l'établissement d'un schéma directeur des données de santé pour les outre-mer doit constituer une priorité par le ministère des affaires sociales et de la santé ».

Le Baromètre santé DOM 2014 de Santé publique France était la dernière enquête en date, généraliste sur la santé, réalisée dans les quatre DROM historiques, selon la même méthodologie qu'en France métropolitaine (par téléphone avec génération aléatoire de numéros de téléphone).

Aucune enquête généraliste sur la santé n'avait eu lieu auparavant à Mayotte. Mais EHIS s'est montée parallèlement à l'enquête Unono wa maore³ de Santé publique France qui s'est déroulée quasiment en même temps que EHIS à Mayotte (à partir d'avril 2019 pour l'enquête Unono wa Maore). Des échanges ont eu lieu entre les équipes en charge de ces enquêtes. La DREES et l'Insee ont proposé à Santé publique France de fusionner les deux enquêtes tout en grossissant un peu l'échantillon initialement prévu pour EHIS. Une première visite aurait pu être faite par un enquêteur de l'Insee pour recueillir le questionnaire commun aux deux enquêtes et une seconde visite aurait permis de recueillir le questionnaire propre à l'enquête Unono wa maore et de mener les prélèvements (sang, urines, ...) et examens de santé prévus dans le protocole de l'enquête Unono wa maore. Malgré cette proposition, appuyée par le directeur de la DREES dans une lettre datée du 21 février 2018, le directeur général de Santé publique France n'a pas souhaité reculer le terrain de l'enquête Unono wa maore, initialement prévue fin 2018. Ainsi, ces deux enquêtes se sont déroulées dans des délais très rapprochés à Mayotte, avec des questionnaires en partie redondants. Les échantillons ont été disjointes par l'Insee conformément aux bonnes pratiques de la statistique publique en matière d'enquêtes. Toutefois, les objectifs des deux enquêtes étaient assez différents. L'inscription de EHIS dans un contexte national rendait nécessaire un questionnaire commun en France métropolitaine et dans les cinq DROM pour mener les comparaisons. L'enquête Unono wa maore était spécifique à Mayotte avec un questionnaire plus adapté au territoire et accompagné de prélèvements/examens de santé.

Partenariats et coût de la troisième édition

L'enquête en France métropolitaine a été l'objet d'une collaboration entre la DREES et l'Irdes. Le prestataire retenu pour réaliser le terrain de l'enquête était Kantar public. Le coût externe du terrain de l'enquête s'est élevé à 1,850 millions d'euros. Un total de 14 192 questionnaires a été obtenu.

L'enquête santé DOM est le fruit d'un partenariat entre la DREES et l'Insee. C'est le réseau d'enquêteurs de l'Insee qui a mené l'enquête. Un total de 9 764 questionnaires a été obtenu. Le coût direct de l'opération s'élève à 1,3 millions d'euros et le coût complet qui, en plus, réaffecte les coûts des fonctions support aux opérations, s'élève à 2,2 millions d'euros supportés à hauteur de 760 000 euros par la DREES, de 240 000 euros par la Direction générale de l'Outre-mer et de 50 000 euros par l'Agence régionale de santé de la Martinique. Le reste du coût était à la charge de l'Insee.

L'enquête EHIS en France métropolitaine et dans les DROM s'est construite de manière parallèle mais très distincte du fait de partenariats différents. La suite du document présente le bilan de ces deux enquêtes de manière séparée.

³ Voir : <https://www.santepubliquefrance.fr/docs/enquete-de-sante-a-mayotte-unono-wa-maore-questionnaire>

■ ENQUÊTE SANTÉ EUROPÉENNE : FRANCE MÉTROPOLITAINE

Calendrier

La préparation de l'enquête a duré 2 ans et demi à partir de fin 2016. Le terrain de l'enquête a duré 8 mois. Il s'est déroulé en deux vagues, chacune sur une partie aléatoire de l'échantillon total, séparée par le mois d'août : du 30 avril 2019 au 1^{er} août pour la vague 1 et du 27 août 2019 au 1^{er} février 2020 pour la vague 2. Les traitements de l'aval statistique ont pris un peu de retard du fait de l'engagement des équipes de EHIS à la DREES dans la production de l'enquête EpiCov à partir du premier confinement de mars 2020. La correction de la non-réponse partielle sur les variables santé et le calcul de la pondération se sont déroulés entre août 2020 et décembre 2020. La validation des données par Eurostat s'est fait en mars 2021. Les premiers résultats de l'enquête sont parus en avril 2021 (Leduc et al., 2021). Il faut donc compter presque 5 ans entre le début du projet et la publication des premiers résultats. Le rapport qualité de l'enquête a été validé par Eurostat en janvier 2022.

Du point de vue du calendrier et des grandes étapes de l'organisation du projet, les aspects suivants auraient pu être améliorés :

- Le questionnaire de l'enquête aurait sans doute gagné à la mise en place d'un groupe de travail sur le questionnaire regroupant plusieurs institutions. C'est une remarque qui avait été formulée par le Comité du label de la statistique publique. Le questionnaire est très contraint par le règlement européen. Ceci dit, que ce soit du point de vue de la formulation des questions européennes en français que de l'ajout de questions spécifiquement françaises, une consultation plus large aurait pu être mise en place avec des réunions régulières. La collaboration avec l'Irdes ancrerait les questions spécifiquement françaises sur le thème de l'économie de la santé et notamment le sujet de la complémentaire santé. Mais la prochaine édition gagnerait à une consultation plus large autour du questionnaire. Il serait souhaitable d'associer notamment Santé publique France en vue d'une meilleure articulation avec le Baromètre santé qui, d'ici la prochaine édition de EHIS, pourrait s'être engagé dans une démarche de rapprochement des standards de qualité des enquêtes de la statistique publique.
- On a essayé de s'y prendre le plus tôt possible pour les démarches CNIL d'autorisation de l'enquête mais le CESREES demandait à disposer du relevé de décision du Comité du label avant de se prononcer sur le projet. De ce fait, le dossier a été transmis le 22 novembre 2018 au CESREES qui s'est prononcé le 13 décembre 2018. La CNIL a mis un peu plus de quatre mois à rendre sa décision. L'autorisation a été obtenue le 29 avril 2019. Le terrain a été lancé le 30 avril 2019. Depuis la modification de la loi informatique et liberté de juillet 2019, il n'est maintenant plus nécessaire pour le service statistique public de demander une autorisation de la CNIL pour la réalisation d'enquêtes portant sur la santé. Une autorisation restera nécessaire pour l'appariement avec le SNDS mais cela contraint moins le calendrier opérationnel de collecte.
- La durée du terrain de l'enquête a été un peu courte. Il n'est pas évident de prévoir la bonne durée du terrain et cela est très lié aux moyens enquêteurs dévolus au projet. Ceci dit, le protocole de l'enquête était en multimode séquentiel puisqu'une partie des non-répondants par téléphone étaient basculés en face-à-face, ce qui demandait une certaine durée de collecte pour pouvoir procéder à la collecte selon les deux modes. La vague 1, qui en plus comportait 60 % de l'échantillon (hors réserves), soit 13 680 personnes, a eu une durée de collecte trop courte. Initialement, elle devait se terminer début juillet et 2 mois et demi de collecte n'ont pas suffi. C'est aussi lié au fait que le début du terrain a été retardé du fait de difficultés de programmation du questionnaire chez Kantar. La vague 1 a alors été prolongée jusqu'à fin juillet et une partie des non-répondants par téléphone de la vague 1 ont finalement été interrogés en face-à-face lors de la vague 2. On aurait gagné à prolonger encore plus la vague 1 si on n'avait pas buté sur le mois d'août. La vague 2 a été mieux calibrée. Elle comportait 40 % de l'échantillon initial et les 20 % de l'échantillon de réserve et a eu une durée de collecte de 5 mois, notamment grâce à la prolongation d'un mois jusqu'à fin janvier 2020. Il aurait été préférable de prévoir une durée totale de collecte de 10 mois, 5 mois pour la vague 1 et 5 mois pour la vague 2 et donc d'avancer le début de collecte (et donc des préparatifs) puisque la collecte devait se dérouler en 2019. On y aurait gagné également à répartir l'échantillon initial à raison de 50 % pour la vague 1 et 50 % pour la vague 2 car la vague 1 était moins rôdée que la vague 2 chez le prestataire si bien qu'a posteriori cela semble audacieux de confier une grosse partie de l'échantillon au début. Eurostat avait validé en 2019 un prolongement de la collecte en janvier 2020.
- Comme précisé plus haut, les traitements de l'aval statistique ont pris du retard du fait de la pandémie de Covid 19 et en particulier du projet EpiCov qui mobilisait les mêmes équipes que EHIS à la DREES. Ainsi, une dérogation a été demandée à Eurostat pour la date de livraison des données et du rapport qualité. Le règlement indique que les données doivent être livrées 9 mois après la fin de collecte (établie à fin 2019, le mois de collecte de janvier 2020 ne comptait pas), ce qui menait à septembre 2020 et le rapport qualité doit être livré trois mois après la livraison des

données. Les données ont été validées en mars 2021 et le rapport qualité intégrant les calculs de variance en janvier 2022. La France n'était pas le seul pays à être en retard dans la livraison. Sans le projet EpiCov, les délais auraient peut-être été tenus mais ces délais demandent une mobilisation importante des équipes statistiques dès la réception des données pour tenir l'échéance. Les délais restent les mêmes avec IESS. Les travaux orientés recherche de prise en compte du SNDS dans les traitements d'aval statistique de l'enquête et d'étude des effets de mode n'ont pas pu être intégrés dans la livraison des fichiers Eurostat.

- Un groupe d'exploitation de l'enquête n'a pas pu être mis en place du fait de la place qu'occupait EpiCov au sein du bureau « État de santé de la population » à la DREES, mais il serait intéressant d'en mettre un en place dans la continuité du groupe de travail sur le questionnaire pour la prochaine édition.

Le tableau 1 ci-dessous présente le calendrier détaillé du projet.

Tableau 1 • Calendrier du projet en France métropolitaine

Date	Étape
13-14 octobre 2016	Réunion de la <i>Task Force</i> EHIS (questionnaire, modules complémentaires envisagés pour test basé sur le volontariat : santé des enfants, santé mentale, handicap, expérience du patient, nutrition)
7 novembre 2016	Réunion avec la division Sondages sur le choix de la base de sondage (recensement, fichiers fiscaux ou bases de l'Assurance maladie)
18 janvier 2017	Réunion avec la division Sondages notamment sur les conséquences d'un effet de grappe en matière de précision en cas d'interrogation de plusieurs personnes par ménage
31 janvier 2017	Réunion avec la division Recueil et traitement de l'information sur le choix de la base de sondage, l'unité d'échantillonnage, le TCM et le protocole de collecte
9-10 février 2017	Réunion du <i>Technical Group</i> EHIS
23 février 2017	Réunion avec l'équipe en charge de l' <i>European social survey</i> à Sciences-po sur leur protocole de collecte
23 février 2017	1 ^{er} comité de pilotage portant sur le choix de la base de sondage, les protocoles envisagés pour maîtriser le coût de l'enquête, le calendrier et le questionnaire. Il est acté de faire une demande à la DGFIP pour tirer l'échantillon dans les fichiers fiscaux.
8 mai 2017	Réunion avec l'équipe en charge de l'enquête INCA à l'ANSES sur leur protocole de collecte
17 mai 2017	Réunion avec l'équipe en charge du Baromètre Santé de Santé publique France sur leur protocole de collecte
29 juin 2017	Réunion avec la division Sondages sur les tirage des zones d'action enquêteurs (ZAE) pour l'enquête qui comportera nécessairement du face-à-face
6-7 juillet 2017	Réunion de la <i>Task Force</i> EHIS (au sujet d'IESS et du questionnaire)
11 juillet 2017	Déclaration auprès de la CNIL des tests de l'enquête. La CNIL a accepté pour ces tests une simple déclaration.
25 juillet 2017	Réunion avec la division Sondages. Validation du nombre de ZAE (zones d'action enquêteur) à échantillonner.

6 septembre 2017	2 ^e comité de pilotage : validation des protocoles de collecte qui seront soumis dans l'appel d'offres aux prestataires de collecte (1 : tirages d'individus, collecte téléphone et face-à-face/2 : tirage de logements, enquête de tout le ménage dans le champ, collecte face-à-face et autoadministrée/3 : idem 3 en version individus, sans enquêter tout le ménage pour éviter l'effet de grappe) et arbitrages sur le questionnaire
4 octobre 2017	Séance de la Commission services publics et services aux publics du Cnis – Obtention de l'avis d'opportunité Validation de la demande d'accès aux fichiers fiscaux au titre de l'article 7bis pour le tirage de l'enquête
10 novembre 2017	Signature de la convention DREES/Insee/GIP-Irdes
29 novembre 2017	Publication de l'appel d'offres public pour la réalisation du terrain de l'enquête. Il est demandé de chiffrer les trois scénarios envisagés par la 2 ^e comité de pilotage.
7 décembre 2017	Déclaration de sous-traitance au Comité du secret pour l'Irdes
29 janvier 2018	Ouverture des plis de l'appel offres
13 février 2018	Réunion avec la division Sondages. Description de Fidéli pour le tirage.
19 février 2018	Notification du marché pour la réalisation du terrain de l'enquête – la société Kantar Public est retenue
29-30 mai 2018	Réunion de la <i>Task Force</i> EHIS (diffusion de EHIS 2 ^e édition et questionnaire des modules proposés pour test sur la base du volontariat)
26 juin 2018	Déclaration de sous-traitance au Comité du secret pour Kantar Public
27 juin 2018	Lettre d'acceptation d'Eurostat pour une bourse de 25 000 euros pour tester le WHO5 dans EHIS (bourse qui sera finalement déclinée par la suite par faute de temps au sein de l'équipe conceptrice pour mener un bilan pour Eurostat sur ce module de questions)
19 juillet 2018	3 ^e comité de pilotage. Information sur le prestataire retenu. Fixation des objectifs du pilote. Discussion sur les partenariats (CGET et OQAI). Arbitrage sur la profondeur de l'appariement avec le SNDS.
14 septembre 2018	Réunion avec les équipes de la BRPP au sujet de la reconstitution des NIR avec Fidéli
18 septembre – 9 octobre 2018	Test du questionnaire par téléphone avec Kantar Public
16 octobre 2018	Réunion avec la CNAM pour l'appariement de l'enquête avec le SNDS
12 novembre 2018	Envoi du dossier de l'enquête au Comité du label de la statistique publique
22 novembre 2018	Envoi du dossier relatif à l'autorisation de la CNIL à l'Institut national des données de santé
12 décembre 2018	Séance du Comité du label. Obtention du label d'intérêt général et de qualité statistique avec le caractère obligatoire.
13 décembre 2018	Avis favorable du CEREES en vue d'obtenir l'autorisation de la CNIL

19 décembre 2018	Signature de la convention entre l'Irdes et le CSTB valant liaison entre l'enquête de l'OQAI et EHIS contre participation financière à EHIS
Janvier 2019	Signature de la convention entre l'Irdes et le CGET pour l'extension quartiers prioritaires de la politique de la ville (QPV) de l'échantillon contre participation financière à EHIS
9 janvier 2019	Dépôt du dossier complet à la CNIL
11 janvier 2019	Avis de la CNIL d'une prolongation du délai d'examen du dossier de deux mois en plus des deux mois légaux
4 mars 2019	Envoi de l'échantillon à Kantar
5 mars – 18 mars 2019	Enrichissement de l'échantillon avec les numéros annuaire
19 mars 2019	Tirage des 2 000 individus de l'échantillon méthodologique par la division Sondages
21 mars 2019	Réunion du <i>Technical Group</i> EHIS (IESS, rapport qualité, expériences nationales de EHIS 3 ^e édition, module santé de SILC)
15 février-27 avril 2019	Programmation du questionnaire
4 avril 2019	Mise en ligne du site internet de l'enquête
8 avril-7 juin 2019	Formation des enquêteurs de la vague 1
29 avril 2019	Autorisation de la CNIL. Numéro : 919006. Autorisation de mener l'enquête et d'apparier avec les données 2015-2024 du SNDS.
30 Avril -1 ^{er} août 2019	Vague 1 du terrain de l'enquête
2-3 juillet 2019	Réunion de la <i>Task Force</i> EHIS (modules proposés pour test, extension de EHIS aux personnes vivant en institution, module santé de SILC)
22 août - 2 octobre 2019	Formation des enquêteurs de la vague 2
27 août 2019 – 1 ^{er} février 2020	Vague 2 du terrain de l'enquête
13 novembre 2019	4 ^{ème} comité de pilotage. Discussion autour des difficultés relatives à la collecte. Réunion nécessaire avec la CNAM pour mettre en place l'appariement avec le SNDS.
19 novembre 2019	Signature de la convention avec la CNAM pour l'appariement avec le SNDS
3 décembre 2019	Signature du contrat de sous-traitance de l'Irdes
28 janvier 2020	Signature de la convention avec l'Insee pour la reconstitution des NIR (BRPP)
Février 2020	Envoi de l'état civil de Fidéli par la division Sondages à la BRPP pour identification sur NIR
26-27 février 2020	Réunion de la <i>Task Force</i> EHIS (articulation avec le module santé de SILC et extension de EHIS aux personnes vivant en institutions)

6 février 2020	Signature de la convention de tirage de l'échantillon avec l'Insee et la DGFiP
17 février 2020	Signature de l'avenant à la convention tripartite DREES-Insee-GIP Irdes
18 mars 2020	5 ^e comité de pilotage. Bilan de la collecte. Calendrier des traitements aval statistiques.
7 avril 2020	Dépôt par l'Insee des NIR sur la plateforme Safe de la CNAM (entre février et avril difficultés pour mettre le fichier au format attendu par Safe et pour la transmission du fichier à la CNAM)
24 juillet 2020	Signature de la convention avec l'Insee pour la codification Sicore
Août 2020	Réception de l'appariement SNDS 2015-2018
Septembre 2020	Livraison de la codification Sicore par l'Insee
Août-décembre 2020	Correction de la non-réponse partielle pour les variables santé et calcul de la pondération
24 mars 2021	Validation des données par Eurostat (premier envoi en décembre 2020)
9 avril 2021	Publication du dossier de la DREES de premiers résultats (n°78)
17 mai 2021	Mise en ligne des indicateurs de santé en open data sur le site de la DREES ⁴
24-25 Novembre 2021	Réunion du <i>Working Group on Public Health</i> dont une journée était consacrée à EHIS
21 janvier 2022	Validation du rapport qualité par Eurostat (premier envoi en octobre 2021)
3-4 mars 2022	Réunion de la <i>Task Force</i> EHIS avec présentation des rapports qualité ⁵
Mars 2022	Mise à disposition de la base Eurostat à la recherche via le CASD
Fin 2022	Livraison de l'appariement socio-fiscal (millésime 2019) par l'Insee
Jusqu'en 2024	Appariements chaque année avec le SNDS, mise à disposition de ces données appariées.

⁴ Voir : <https://data.drees.solidarites-sante.gouv.fr/explore/dataset/indicateurs-ehis/information/>

⁵ Le rapport qualité de la France est disponible sur la page Internet consacrée à EHIS sur le site Internet de la DREES : <https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/enquete-sante-europeenne-ehis>

Questionnaire

Variables européennes

Pour l'édition 2014, comme EHIS s'inscrivait dans le contexte de l'enquête Santé et protection sociale de l'Irdes, des dérogations avaient été demandées pour poursuivre des séries de l'enquête Santé et protection sociale. Les questions relatives à l'aide informelle (IC1 à IC3) n'étaient pas posées ; pour l'alcool, le module AUDIT-C était préféré, ainsi les questions pour cerner la consommation d'alcool ne correspondaient pas au règlement européen. Les questions sur le renoncement aux soins, même si elles ne figuraient pas dans la liste des dérogations, n'étaient pas posées de la même façon que ce qui était demandé par le règlement européen (besoins de soins non satisfaits). D'ailleurs, ces variables ne sont pas publiées pour la France sur le site internet d'Eurostat.

En même temps que EHIS devenait une enquête indépendante en 2019, il a été décidé de faire correspondre au maximum le questionnaire aux indicateurs européens demandés par Eurostat. Toutes les questions de 2019 ont pour but de cerner l'ensemble des notions du manuel européen des indicateurs. Rappelons que ce qui est demandé par Eurostat, c'est une liste d'indicateurs. Le manuel des indicateurs précise la signification de l'indicateur et conseille une question formulée en anglais pour recueillir les informations nécessaires à la construction de l'indicateur.

Pour un certain nombre de questions, la liste des modalités proposées dans le questionnaire 2019 est plus détaillée que ce qui est demandé dans l'indicateur européen. Il s'agit des variables AC2, PA2, PA3, PA4, PA5, PA6, PA7, PA8, IC1.

Le tableau 2 ci-dessous présente les principales remarques quant à la comparaison du questionnaire de 2019⁶ avec le manuel méthodologique d'Eurostat (Eurostat, 2018). La plupart de ces remarques portent sur des légers écarts de concepts entre le questionnaire et les indicateurs européens. Beaucoup de consignes ont été ajoutées dans le questionnaire 2019 par rapport au questionnaire 2014 pour essayer de coller au plus près au manuel européen. Parfois, nous n'avons pas trouvé de solution pleinement satisfaisante. Les questions relatives à la liste de maladies (questions CD1) ont fait l'objet d'une traduction que l'on a voulu la plus proche possible du manuel européen en collaboration avec un médecin de l'Irdes.

Tableau 2 • Mise en perspective du questionnaire français par rapport au manuel méthodologique européen

Nom de la variable	Signification	Mise en perspective par rapport au manuel des indicateurs européens (3 ^e édition)
HS1	État de santé général perçu ; 1 ^{ère} question du mini-module européen sur la santé	La troisième modalité de la question-type proposée, « fair » en anglais, doit être traduite par une notion neutre d'après le manuel, du type « ni bon, ni mauvais ». Ce n'est pas totalement cohérent car « fair » en anglais a une connotation positive. En français, elle est historiquement traduite par « assez bon ». La question de l'évolution de la formulation du mini-module européen s'est posée au début du projet EHIS 2019. On a envisagé de changer les formulations de HS1 et HS3 mais il apparaît qu'il faut conserver les formulations historiques pour maintenir les séries longues, d'autant que cela concerne aussi SILC, LFS, etc. qui comportent le module (même l'ensemble des enquêtes Ménages européennes avec IESS).
HS2	Présence de maladies chroniques ; 2 ^e question du mini-module européen sur la santé	On a ajouté, conformément au manuel, une consigne pour définir le terme « maladie chronique » qui est complexe (« une maladie chronique est une maladie qui a duré ou peut durer pendant six mois au moins »).
HS3	GALI ; 3 ^e question du mini-module européen sur la santé	Le manuel recommande de poser deux questions. La première relative à la présence de limitations et la seconde relative à leur durée (« est-ce que cela dure depuis au moins 6 mois ? ») car Eurostat considère que la formulation en une question est complexe avec la notion de limitations et la notion de la durée depuis

⁶ Le questionnaire est disponible sur la page internet de l'enquête EHIS de la DREES : <https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/enquete-sante-europeenne-ehis>

		laquelle ces limitations sont présentes. La question est historiquement posée en une unique question en France. Un dossier de la DREES (Dauphin et Eideliman, 2021) a récemment souligné la pertinence de cette question pour mesurer le handicap et la stabilité des niveaux observés dans les différentes enquêtes rappelant l'importance de ne pas changer la formulation. La question issue de SILC est par ailleurs utilisée pour mesurer l'espérance de vie sans incapacité.
CD2	Santé perçue bucco-dentaire	Même remarque que HS1 sur la traduction de la modalité « fair ».
PL5	Difficulté à entendre dans une pièce où il y a du bruit	La question n'est pas très bien comprise : « s'il y a du bruit, je n'entends pas... ». Eurostat nous avait conseillé de donner l'exemple d'une pièce où plusieurs personnes parlent en même temps. Cela a été ajouté en consigne.
PL9	Difficulté à croquer dans des aliments durs comme une pomme ferme	Eurostat impose cette question pour les 55 ans ou plus. Nous l'avons élargie à l'ensemble du champ de l'enquête. Par ailleurs, la réponse ne doit pas tenir compte de l'usage éventuel d'un dentier d'après le manuel. Cela nous a semblé incohérent avec les autres incapacités (par exemple, les lunettes doivent être prises en compte pour les difficultés de vue). Nous avons indiqué de tenir compte du dentier si la personne avait répondu en avoir un.
PA1	Vaccination contre la grippe	Normalement, Eurostat veut le mois et l'année de la dernière vaccination contre la grippe si celle-ci a eu lieu depuis l'année précédente à la collecte, c'est-à-dire après le 31 décembre 2017 (donc entre 2018 et la date de l'enquête). En France métropolitaine, la campagne de vaccination s'étend d'octobre à mars. La période de référence mélange donc potentiellement les campagnes de vaccination 2017/2018 et 2018/2019. Nous n'avons pas trouvé de façon simple de collecter cette information. Aussi, nous avons simplement demandé si la personne avait été vaccinée au cours des 12 derniers mois. Comme la collecte a démarré en mai 2019, on exclut des gens qui aurait été vacciné entre janvier et mars 2018 pour la campagne 2017/2018 et qui ne l'aurait pas été pour la campagne 2018/2019.
Questions UN	Besoins de soins non satisfaits	Le concept recueilli est le même que celui demandé par Eurostat. Par rapport à la question proposée dans le manuel, la modalité « non concerné » a été placée en premier dans le questionnaire pour bien faire la distinction avec le fait de ne pas avoir eu des besoins de soins non satisfaits. Il y a eu beaucoup de discussions sur la formulation en français de ces questions. Ce travail a été l'occasion de rendre proches les formulations de SRCV et de EHIS dans le cadre de la refonte 2020 de SRCV. En effet, entre 2015 et 2019, concernant le renoncement aux soins, SRCV utilisait une question filtre pour savoir si la personne avait eu besoin de soins et dans le cas d'une réponse positive uniquement, la question sur le renoncement était posée. Il a été craint que la question filtre ne réduise artificiellement le taux de soins non satisfaits. C'est pour cette raison que les trois modalités « non concerné », « besoin de soin non satisfait », « pas de besoin de soins non satisfait » ont été proposées en réponse à une unique question. Cependant, nous avons eu des retours enquêteurs disant que la formulation était trop compliquée (cf. partie « bilan enquêteurs »). De ce fait la formulation de SRCV 2020, lors de la refonte, a été un peu revue par rapport à celle de EHIS en utilisant notamment le mot « renoncer ». De plus, dans les données de EHIS, on

		observe beaucoup d'incohérences entre les réponses des personnes aux différentes questions sur les besoins de soins non satisfaits. Eurostat a indiqué que des incohérences étaient également présentes mais dans une moindre ampleur dans les données des autres pays.
PE2, PE3, PE4, PE5	Temps passé à faire des trajets d'au moins 10 minutes à pied ou à vélo	Le manuel européen indique d'exclure les trajets réalisés dans le cadre du travail. Cette précision n'a pas été indiquée dans le questionnaire français. On s'est demandé comment exclure ces trajets et pourquoi avoir choisi de les exclure. Pourquoi inclure un trajet pour se rendre à son entreprise (<i>commuting</i>) et pas les trajets réalisés par un livreur à vélo ?
PE6, PE7, PE8	Temps passé à faire du sport dans le cadre du loisir	Le manuel européen indique de préciser que ces questions excluent les activités de trajet à pied ou à vélo déclarées dans les questions précédentes. Nous n'avons pas mis cette précision dans le questionnaire qui nous semblait rendre le questionnaire complexe.
DH1, DH2, DH3 et DH4	Consommation de fruits et de légumes	On a essayé de coller au plus près du manuel mais certaines consignes, que l'on a pourtant mises dans le questionnaire, ne nous ont pas semblé pertinentes. Pourquoi exclure toutes les préparations à base de fruits et de légumes surtout si elles sont faites maison ? Par exemple, les soupes sont exclues alors qu'elles peuvent n'être composées que de légumes et représenter une forte consommation de légumes l'hiver en France.
DH5	Consommation de jus maison et 100 % pur jus	Le manuel Eurostat évoque les jus de fruits et de légumes. Seuls les jus de fruits ont été considérés dans le questionnaire.
HHINCOME	Niveau de vie du ménage (quintiles)	Calculé à partir des données socio-fiscales et non déclaré dans l'enquête (de ce fait, la question du revenu n'a pas été posée dans l'enquête). Millésime 2017 dans la première livraison à Eurostat. Le millésime 2019 sera livré dans un second temps.

L'ordre du questionnaire relatif à la santé n'a pas été exactement celui recommandé par Eurostat du fait que les questions relatives au PHQ9, au tabac et à l'alcool étaient posées sous casque dans les DROM. De ce fait, en France métropolitaine également, les questions du PHQ9 ont été regroupées avec les questions sur le tabac et l'alcool à la fin du questionnaire. En dehors de cela, l'ordre préconisé par Eurostat a été respecté.

La question s'est posée de faire passer les questions socio-démographiques à la fin de l'enquête pour débiter l'enquête par des questions sur la santé, comme annoncé avec la thématique de l'enquête. On ne l'a pas fait mais cela pourrait être expertisé pour une prochaine édition de l'enquête. Certains enquêteurs ont souligné dans le « bilan enquêteurs » la difficulté d'avoir une première partie assez longue, personnelle et administrative sans lien avec la santé : « Sur le début du questionnaire, les gens n'arrivent pas à comprendre qu'on leur pose toutes ces questions sur leur vie privée, avant de poser les questions médicales. »

Lorsqu'un mineur était échantillonné, la partie socio-démographique était posée à un adulte et les questions santé au mineur. Ce n'était pas forcément évident, notamment par téléphone, il fallait alors avoir deux personnes pour l'enquête. L'analyse des taux de réponse à l'enquête (cf. partie « taux de réponse ») semble confirmer une certaine difficulté à ce propos : le taux de réponse est plus faible sur cette classe d'âge par téléphone.

Le proxy était autorisé conformément au règlement européen. Au total, 284 personnes ont répondu via proxy. Cela représente 2,0 % des répondants. C'est le niveau à ne pas dépasser tel que c'est recommandé dans le manuel méthodologique européen. Le proxy est plus fréquent en face-à-face que par téléphone où le recueil par proxy doit être plus complexe. Il y a 2,8 % (172 questionnaires) des enquêtes réalisées en face-à-face qui l'ont été via proxy et 1,4 % (112) de celles réalisées par téléphone. Le retour des enquêteurs sur la passation via proxy a été positif si ce n'est un bémol sur le fait que les tournures du questionnaire n'étaient pas adaptées spécifiquement à ce cas (cela aurait signifié un questionnaire à la troisième personne du singulier en plus du vouvoiement du cas sans proxy, ce qui est coûteux à développer).

Variables françaises

Des questions françaises, en plus du questionnaire européen, ont été ajoutées au questionnaire. Il s'agit de :

- La PCS pour la personne de référence (au sens du TCM) du ménage (et pas seulement celle du répondant demandée par Eurostat)
- Une question sur la crainte de perdre son emploi
- Un module sur la complémentaire santé
- Un module sur les caractéristiques du logement
- Un module sur les origines (plus complet que les variables européennes)
- Deux questions sur l'opération de la cataracte
- Cinq questions sur la littératie en santé (dimension 9 du *Health Literacy Questionnaire* (HLQ))
- Deux questions sur la dernière personne vue pour un problème de santé
- Deux questions sur les besoins de soins non satisfaits pour l'optique et les prothèses auditives
- Deux questions d'échelle (note entre 0 et 10) quant au goût pour le risque et à la satisfaction par rapport à la vie (échelle Cantril)
- La neuvième question du PHQ9

Durée du questionnaire

Le questionnaire a une durée moyenne de 44,0 minutes et une durée médiane de 41,4 minutes (tableau 3). Le quart des personnes qui a passé le moins de temps sur l'enquête a passé moins de 35,3 minutes à remplir le questionnaire. Les 10 % des personnes qui ont passé le plus de temps sur l'enquête ont passé plus de 58,7 minutes à remplir le questionnaire. La durée de passation par téléphone et en face-à-face est quasi similaire.

La durée de passation varie un peu selon les caractéristiques sociales des personnes enquêtées (tableau 4) Les plus âgés, les plus modestes, ceux qui habitent dans un quartier prioritaire de la politique de la ville (QPV) et surtout ceux qui ont un état de santé dégradé passent plus de temps à répondre à l'enquête. Pourtant, le questionnaire de l'enquête EHIS est très peu filtré si bien que tout le monde passe dans la quasi intégralité des questions. En particulier, les personnes en mauvais état de santé n'ont pas plus de questions. Il y a toutefois un module de questions spécifiques pour les plus de 55 ans sur les limitations fonctionnelles. Les écarts de durée s'interprètent plutôt comme une difficulté à renseigner le questionnaire ou comme la nécessité de réfléchir plus à chaque question parce que l'on est plus concerné par les questions quand on a des problèmes de santé que quand, par exemple, on n'a presque aucun contact avec le système de soins.

Tableau 3 • Durée du questionnaire par mode de collecte (en minutes)

	Q25	Médiane	Moyenne	Q75	Q90
Ensemble	35,3	41,4	44,0	49,4	58,7
Téléphone	36,8	42,4	44,4	49,7	57,8
Face-à-face	33,1	39,7	43,6	48,8	60,6

Lecture > La durée médiane de passation de l'enquête est 41,4 minutes.

Champ > France métropolitaine

Source > DREES-Irdes, EHIS 2019 (France métropolitaine)

Tableau 4 • Durée moyenne du questionnaire selon les caractéristiques des répondants et le mode de collecte

	Téléphone	Face-à-face
Ensemble	44,4	43,6
Femmes	45,0	44,1
Hommes	43,6	43,0
15-17 ans	43,5	42,5
18-24 ans	43,2	40,4
25-34 ans	41,8	41,7
35-44 ans	42,7	41,6
45-54 ans	43,7	43,0
55-64 ans	44,8	44,8
65-74 ans	46,5	45,6
75-84 ans	48,7	47,5
85 ans ou plus	52,8	48,8
d1 : les 10 % les plus modestes (niveau de vie)	46,2	44,8
d5 : médiane	44,3	43,0
d10 : les 10 % les plus aisés	43,2	43,7
Quartier prioritaire de la politique de la ville (QPV)	47,8	46,6
Hors QPV	44,0	43,0
État de santé général : très bon	42,1	40,8
État de santé général : très mauvais	50,2	48,3
Maladie chronique : oui	45,9	45,5
Pas de maladie chronique	43,4	42,3
GALI : handicapé (fortement limité)	48,9	47,4
GALI : pas limité	43,3	42,1
Pas proxy	44,3	43,5
Proxy	48,7	44,4

Lecture > La durée moyenne de passation de l'enquête est 44,4 minutes par téléphone, 45,0 minutes pour les femmes et 43,6 minutes pour les hommes.

Champ > France métropolitaine

Source > DREES-Irdes, EHIS 2019 (France métropolitaine)

Bilan enquêteurs sur le questionnaire

Le « bilan enquêteurs » met en évidence des différences d'appréciation selon qu'il s'agissait d'enquêteurs téléphone ou face-à-face, en lien avec les spécificités de ces deux modes de passation.

Concernant la durée du questionnaire, les enquêteurs du téléphone soulignent que c'était assez long : « Le questionnaire est long et c'est fatiguant, tant pour l'enquêteur que pour la personne interrogée. Mais peu s'en sont plaints et ça se passait bien en général. Cela pose surtout problème pour les personnes âgées ou fatiguées relativement à la fiabilité des réponses. ». En revanche, les enquêteurs face-à-face n'ont pas exprimé de difficulté quant à la longueur du questionnaire : « Je n'ai eu aucune remarque sur la durée du questionnaire ; ce qui prouve que les gens ont répondu comme il le fallait, l'étude est intéressante et bien préparée. Alors la durée de temps n'existe plus, enfin moins ... ».

Les questions qui ont appelé le plus de remarques dans le « bilan enquêteurs » sont celles relatives aux besoins de soins non satisfaits⁷ dont on avait changé la formulation pour respecter l'indicateur européen. La formulation n'est pas jugée satisfaisante. Par exemple, on peut lire dans le « bilan enquêteur » : « Les questions concernant les difficultés ou non de pouvoir se payer des soins ou « non concernées » sont très mal formulées. J'ai dû expliquer à plusieurs reprises car certaines personnes ne comprenaient pas de suite. » ; ou encore : « La question sur le besoin de soins et la capacité de l'interviewé à se le permettre financièrement ou non était assez « indigeste » et sujette à interprétation d'autant qu'elle arrive en fin de questionnaire et que la vigilance de l'interviewé comme celle de l'enquêteur baisse. ». Finalement, certains enquêteurs disent avoir découpé la question en deux pour savoir si la personne avait eu besoin de soins et ensuite, le cas échéant, si elle y avait renoncé : « Les questions sur le report de rendez-vous à cause des délais de transport étaient complexes et longues. Il fallait à chaque fois ajouter avez-vous eu des consultations ou eu besoin de lentilles ou lunettes, etc.. Il fallait vérifier si non concerné. ». C'est justement ce que l'on souhaitait éviter avec la nouvelle formulation afin qu'une question filtre ne vienne pas abaisser le taux de renoncement aux soins, comme cela avait pu être observé dans les données de SRCV dans les années antérieures au moment d'un changement de formulation.

Le PHQ9 a été mis en évidence dans le « bilan enquêteurs » comme un module difficile par certains enquêteurs, notamment sans doute la neuvième question ajoutée en France en plus du PHQ8 : « Les questions sur les difficultés de relations sociales, c'était un sujet sensible sur lequel parfois les personnes semblaient gênées de répondre. La question sur l'envie de mourir a aussi été difficile. Une dame a fondu en larmes à ce moment-là, difficile de savoir quoi dire ».

Les questions relatives au nombre de portions consommées par jour de fruits et de légumes semblent sujettes à caution pour certains enquêteurs. C'est d'autant plus vrai que, par téléphone, les enquêteurs ne pouvaient pas montrer les cartes-codes relatives à la mesure d'une portion⁸. On peut lire dans le « bilan enquêteurs ». « Les gens ne comptabilisaient pas clairement les unités de légumes et de fruits, je pense. Ils en oubliaient sûrement. ». Même en face-à-face, alors qu'il y a eu un gros travail sur les cartes-codes, certains enquêteurs le relèvent : « Bien que les portions soient bien imagées elles ne semblaient pas être facilement convertibles spontanément et les réponses récoltées semblaient bien approximatives. ».

Documents de collecte

Les documents de collecte comportaient les lettres-avis, la plaquette de présentation de l'enquête et le livret des cartes code pour le face-à-face. Ces documents figure dans la rubrique « documents de collecte » de la page Internet consacrée à EHS 2019 sur le site de la DREES⁹. Le prestataire avait plastifié une lettre avis qu'il a donnée à tous les enquêteurs face-à-face pour faciliter le contact. Il a dit que cela avait aidé les enquêteurs.

Le kit de communication de l'enquête (affiche, plaquette, vidéo, communiqué de presse) figure sur la page internet de l'enquête sur le site internet de la DREES dans la rubrique « Communication ».

Avant la collecte, le site internet relatif à l'enquête a été mis en ligne. Un mail a été envoyé aux mairies des communes où des personnes avaient été échantillonnées pour informer sur l'enquête.

⁷ Voir le questionnaire sur la page internet consacré à EHS sur le site de la DREES : questions UN1A à UN2E2.

⁸ Voir le livret des cartes codes sur la page internet consacrée à l'enquête EHS sur le site de la DREES dans la rubrique « Documents de collecte ».

⁹ <https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/enquete-sante-europeenne-ehis>

Plan de sondage

L'échantillon total a été réparti aléatoirement en deux vagues de collecte (vague 1, collecte au printemps et vague 2, collecte à l'automne et en hiver) et comportait quatre premiers échantillons de réserve de taille, chacun, de 5 % de l'échantillon initial et quatre autres échantillons de réserve spécifiques au suréchantillon QPV de taille, chacun, de 10 % du suréchantillon QPV. Lors de la collecte, du fait d'un taux de réponse plus faible qu'anticipé en vague 1, l'ensemble des échantillons de réserve ont été mis en collecte lors de la vague 2. On obtient donc la répartition de l'échantillon présentée dans le tableau 5.

Tableau 5 • Répartition de l'échantillon par strate de tirage et affectation aux vagues de collecte

	Échantillon principal...	... dont QPV	Suréchantillon QPV	Total
Taille initiale	21 600	1 728	1 200	22 800
Réserves	4 X 1 080 = 4 320	4 X 86 = 344	4 X 120 = 480	4 800
Affectés à la vague 1	60 % X 21 600 = 12 960	60 % X 1 728 = 1 037	60 % X 1 200 = 720	13 680
Affectés à la vague 2	40 % X 21 600 + 4 320 = 12 960	40 % X 1 728 + 344 = 1 035	40 % X 1 200 + 480 = 960	13 920

Lecture > La taille de l'échantillon initial est 21 600 personnes.

Champ > France métropolitaine

Source > DREES-Irdes, EHIS 2019 (France métropolitaine)

L'échantillon est issu d'un tirage en deux étapes (dit « à deux degrés »). La première étape a consisté à tirer des zones de collecte et la seconde étape était le tirage des personnes à enquêter au sein de ces zones de collecte. La première étape fait nécessairement perdre de la précision aux estimations issues de l'enquête mais elle avait pour but de faciliter la collecte face-à-face. À ce titre, la liste de zones échantillonnées avait été publiée lors de l'appel d'offres pour recruter le prestataire de collecte. L'objectif était de permettre au prestataire de s'ajuster au mieux aux zones tirées en recrutant éventuellement des enquêteurs dans les zones qui auraient été difficilement accessibles des enquêteurs en poste. C'est la division Sondages de l'Insee qui a effectué ces tirages.

Méthode de tirage du premier degré de tirage :

Ce tirage s'appuie sur la construction d'unités primaires dans le cadre du projet Nautille de l'Insee. Chaque unité primaire est une agrégation de communes respectant le nombre minimal de 2 500 résidences principales. Les unités primaires ont été constituées de façon à être les plus compactes possibles afin de faciliter le travail des enquêteurs, en se basant sur une résolution du problème du voyageur de commerce qui a conduit à sélectionner un découpage minimisant les temps de trajet des enquêteurs en respectant les frontières départementales (Chevalier et al., 2022). Ce découpage a été effectué sur la base de la géographie légale au 1^{er} janvier 2016. Au total, 35 772 communes sont dénombrées constituant 5 155 unités primaires.

Étant donnée la taille de l'échantillon d'EHIS visé et la nécessité d'être représentatif de la France métropolitaine, il a été décidé de tirer environ 250 unités primaires. Le tirage est stratifié selon les 13 régions de la France métropolitaine au 1^{er} janvier 2016, Corse incluse. Les probabilités d'inclusion des unités primaires sont calculées selon une méthode itérative :

1. Le nombre d'unités primaires à sélectionner dans chaque région est calculé proportionnellement au nombre de résidences principales pour chaque région.

$$n_{\text{région}} = 250 \times \frac{(\text{total des résidences principales dans la région})}{(\text{total des résidences principales en métropole})}$$

2. La probabilité d'inclusion d'une unité primaire k (p_k) dans la région dépend du nombre de résidences principales de la zone dans la région.

$$p_k = n_{\text{région}} \times \frac{(\text{total des résidences principales dans la zone } k)}{(\text{total des résidences principale dans la région})}$$

3. Les zones pour lesquelles la probabilité est supérieure à 1 sont considérés comme exhaustives : elles seront de toutes façons sélectionnées dans l'échantillon. Leur probabilité d'inclusion est fixée à 1. Le nombre de zones à échantillonner est ensuite réduit de 1 et l'algorithme repart de l'étape 1. Il s'arrête lorsqu'il n'y a plus de zone exhaustive.

En sommant les probabilités d'inclusion, on obtient pour des raisons d'arrondis un total de 249 zones à tirer (ou 291 en comptant séparément les arrondissements de Paris, Lyon et Marseille ; chacune de ces trois villes étant considérée comme une unique commune pour le calcul des probabilités d'inclusion et représentant une zone exhaustive).

Le tirage est donc stratifié par régions et également équilibré selon 13 variables. L'équilibrage a pour but de réduire la variance des estimations issues de l'enquête. Il est fait sur des variables possiblement corrélées au sujet de l'enquête. Cela signifie que la répartition de ces variables au sein de la base de sondage est respectée par le tirage. Comme ces variables sont corrélées au sujet de l'enquête, cela améliore la précision des estimations issues de l'enquête. Les variables retenues pour l'équilibrage ont été la probabilité d'inclusion (pour obtenir le bon nombre de zones), le revenu fiscal en 2014, le montant des pensions en 2014, la pyramide des âges à partir de 15 ans par tranches d'âges de 15 ans, le nombre de résidences principales en pôle urbain, en couronne d'un pôle urbain, en zone multi-polarisée (périurbain), en zone hors influence des pôles (rural).

La méthode de tirage repose sur le cube local (échantillonnage spatialement équilibré). L'échantillon obtenu est constitué de 249 zones (291 en comptant séparément les arrondissements de Paris, Lyon, Marseille), dont 10 étaient exhaustives (52 en comptant séparément les arrondissements de Paris, Lyon, Marseille). Au total, 1 231 communes appartiennent à l'échantillon (1 273 avec les arrondissements de Paris, Lyon, Marseille). Les allocations initiales sont respectées par le tirage. La répartition des zones par région est présentée dans le tableau 6.

Tableau 6 • Répartition des zones de collecte par région

Région	Nombre de zones enquêteurs
Ile-de-France	57
Centre-Val de Loire	11
Bourgogne-Franche-Comté	12
Normandie	14
Nord-Pas-de-Calais-Picardie	24
Alsace-Champagne-Ardenne-Lorraine	23
Pays de Loire	15
Bretagne	14
Aquitaine-Limousin-Poitou-Charentes	25
Languedoc-Roussillon-Midi-Pyrénées	23
Auvergne-Rhône-Alpes	39
Provence-Alpes-Côte d'Azur	33
Corse	1
Total	291

Méthode de tirage du second degré de tirage :

L'échantillon de EHIS a été tiré dans Fidéli 2018. C'était les débuts des tirages dans Fidéli, surtout en ce qui concerne les tirages d'individus (et non de logements). Plus précisément, la division Sondages de l'Insee avait deux fichiers à disposition : un fichier des logements ordinaires au sens des enquêtes « ménages » qui a été filtré sur les résidences principales et un fichier d'individus connus des services fiscaux (à l'impôt sur le revenu, à la taxe d'habitation ou aux deux). Ce fichier est filtré sur les 15 ans ou plus à début avril 2019 (date prévue pour le lancement de l'enquête). C'est l'intersection de ces deux fichiers

qui est conservée et restreintes aux communes tirées au sein des 249 zones de collecte. Ainsi la base de sondage contient 10 146 114 personnes.

La taille de l'échantillon a été fixée à 27 600 personnes à l'aide d'hypothèses sur les taux de réponse. L'échantillon a été réparti en un échantillon principal (25 920 personnes) et un sur-échantillon QPV (1 680 personnes). Comme l'information sur les QPV est présente dans la base de sondage, le tirage se fait au sein de deux strates : hors QPV (au total 23 848 personnes, cf. tableau 5) et QPV (3 752 personnes issues de l'échantillon principal et du sur-échantillon).

Dans chacune des strates de tirage, l'échantillon est réparti entre une partie initiale et des réserves (cf. tableau 5). Ensuite l'échantillon est réparti entre les deux vagues de collecte : 60 % de l'échantillon principal est attribué à la vague 1 et les 40 % restants ainsi que les réserves sont attribués à la vague 2 (cf. tableau 5).

Pour finir un échantillon méthodologique est sélectionné parmi l'échantillon principal hors réserves. Il a pour but l'étude des effets de mode (cf. parties « protocole de collecte » et « effets de mode »). Cet échantillon est constitué de 2 000 personnes tirées de façon à avoir 1 200 personnes en vague 1 (60 %) et 800 en vague 2 (40 %). Ces individus sont tirés parmi les personnes ayant un numéro de téléphone à disposition après enrichissement annuaire et seront interrogés directement en face-à-face.

Le tirage est réalisé au sein de chacune des strates : hors QPV et QPV. La méthode de tirage consiste à tirer un nombre fixe d'individus par zone de collecte (à l'exception des 52 zones exhaustives, pour lesquelles le nombre d'individus à enquêter dépend de la taille de la zone). Comme les zones de collecte ont été tirées proportionnellement au nombre de logements dans la zone, tirer ensuite un nombre fixe d'individus par zone conduit à des probabilités d'inclusion des individus très proches. En effet, une zone avec peu de logements a une plus faible probabilité d'être sélectionnée qu'une zone avec beaucoup de logements dans le premier degré de tirage mais dans le second degré de tirage, un tirage de n individus représentera une plus forte proportion des individus d'une zone avec peu de logements que d'une zone avec beaucoup de logements. Comme la probabilité de tirage de l'individu est le produit de ces deux probabilités, celle-ci va être proche pour tous les individus. Le sondage est dit auto-pondéré.

La méthode de tirage est un tirage systématique à probabilités inégales (le poids de tirage de l'unité primaire est pris en compte pour permettre d'avoir le même nombre d'individus tirés par zone) trié sur l'identifiant de l'unité primaire, puis la présence ou non d'un numéro de téléphone dans la base de sondage, l'âge, le niveau de vie. Le tirage systématique consiste à sélectionner un individu avec un « pas » de tirage régulier (par exemple, un individu tous les 10 000), où le pas de tirage est défini de manière à avoir le bon nombre d'individus souhaités à l'issue du tirage. La base de sondage est auparavant triée par des variables liées au sujet de l'enquête pour améliorer la précision des estimations (selon un principe de stratification implicite).

La méthode qui a été utilisée a permis de s'assurer qu'un seul individu par logement est tiré. Pour y arriver, un individu dans chaque logement a été tiré au préalable par sondage aléatoire simple et il a été dupliqué autant de fois qu'il y a de personnes dans le champ de l'enquête dans le logement. Ces individus ayant les mêmes caractéristiques, ils sont à côté à l'issue du tri de la base de sondage et le tirage systématique n'en sélectionnera qu'au plus un.

Enfin, une indicatrice valant 1 pour une moitié aléatoire de l'échantillon et 0 pour l'autre moitié est créée. Cela indique les personnes éligibles à la « bascule » du téléphone vers le face-à-face en cas de situation « impossible à joindre » par téléphone (cf. partie « protocole »).

Protocole

Description

L'enquête était multimode : téléphone et face-à-face. Les personnes pour lesquelles un numéro de téléphone (Fidéli ou enrichissement annuaire) (cf. partie « enrichissement annuaire ») était disponible étaient sollicitées par téléphone. Les personnes pour lesquelles aucun numéro de téléphone n'était disponible étaient sollicitées en face-à-face.

Par ailleurs, pour essayer de maximiser le taux de réponse, un impossible à joindre¹⁰ (IAJ) sur deux par téléphone a été sollicité à la suite en face-à-face (« bascule »). Le fait d'en choisir un sur deux et de se concentrer sur les IAJ plutôt que l'ensemble des non-répondants par téléphone a été lié à des contraintes budgétaires. En effet, on anticipait un meilleur taux de réponse en face-à-face parmi les IAJ du téléphone que parmi les refus du téléphone. À ce titre, le taux de réponse parmi les IAJ du téléphone a d'ailleurs été proche de celui des personnes sans numéro de téléphone (53 % versus 58 %, cf. partie « taux de réponse »).

Enfin, l'enquête comportait un échantillon méthodologique. Il s'agit de personnes qui avaient un numéro de téléphone (via Fidéli ou enrichissement annuaire) et qui ont été interrogées directement en face-à-face. Le but de cet échantillon est d'étudier l'effet du mode de collecte (téléphone ou face-à-face) sur les réponses données à l'enquête. L'effet de mode se décompose en un effet de sélection (ce sont des gens différents qui répondent selon les différents modes de collecte) et un effet de mesure (une même personne peut donner une réponse différente selon le mode sur lequel elle est interrogée). On a tiré l'échantillon méthodologique aléatoirement parmi les personnes qui disposent d'un numéro de téléphone et on l'a interrogé en face-à-face directement. Le but est de réduire au maximum l'effet de sélection entre l'échantillon méthodologique et les personnes directement interrogées par téléphone car elles appartiennent au même ensemble des personnes avec numéros de téléphone, pour pouvoir ainsi identifier un effet de mesure. Les personnes de l'échantillon méthodologique ont été tirées par la division Sondages de l'Insee une fois l'enrichissement annuaire mené par le prestataire de collecte. L'étude des effets de mode grâce à cet échantillon méthodologique est présentée dans la partie « effets de mode ».

En fonction des effectifs échantillonnés (cf. partie « plan de sondage ») et de l'enrichissement par numéros de téléphone annuaire (cf. partie « enrichissement annuaire »), le tableau 7 donne les effectifs de personnes sollicitées selon les différents modes de collecte.

Tableau 7 • Répartition de l'échantillon selon les modes de sollicitation

Terrain	CATI et CAPI	CATI et CAPI	CATI	CAPI	CAPI	CATI et CAPI	CAPI
Échantillon	Tiré	Mis en collecte	Par téléphone	En face-à-face...	... dont sans numéro de téléphone	... dont bascules dont échantillon méthodologique
Taille totale (sollicités)	27 600	27 595	19 354	10 457	6 242	2 216	1 999
Vague 1 (d'origine)	13 680	13 676	9 411	5 087	3 066	8 22	1 199
Vague 2 (d'origine)	13 920	13 919	9 943	5 370	3 176	1 394	8 00
Numéro de téléphone	Certains	Certains	Tous	Certains	Aucun	Tous	Tous

Note > La différence entre l'échantillon tiré et mis en collecte correspond à des hors-champs. Par ailleurs, certains IAJ de la vague 1 ont été basculés en face-à-face lors de la vague 2 faute de temps sur la vague 1. Ils sont comptabilisés dans leur vague d'origine (vague 1)

Lecture > La taille totale de l'échantillon tiré est 27 600 personnes.

Champ > France métropolitaine

Source > DREES-Irdes, EHIS 2019 (France métropolitaine).

¹⁰ On sépare classiquement l'échantillon en différentes populations : les répondants, les non-répondants et les hors-champs. Parmi les non-répondants, on distingue les « refus » et les IAJ. Les IAJ sont ceux que l'on n'a jamais réussi à joindre.

Lorsque la collecte était réalisée en face-à-face, il était proposé à l'enquêté de renseigner les questions les plus sensibles de l'enquête (PHQ9, tabac, alcool) directement par lui-même sur la tablette de l'enquêteur. Ces questions étaient regroupées en fin de questionnaire. Ce choix a été fait pour assurer la meilleure comparabilité possible des réponses avec l'enquête dans les DROM où ces modules étaient posés sous casque.

Enfin, une difficulté particulière de collecte tient au fait que dans Fidéli, pour les plus jeunes, on ne dispose pas du prénom/nom/sexe/jour et mois de naissance. Ainsi, dans ce cas, les enquêteurs doivent trouver la personne à enquêter dans le ménage uniquement à partir de son année de naissance. Les enquêteurs ont souligné dans le « bilan enquêteurs » que cela était une gêne dans leur travail, notamment dans la collecte par téléphone : « L'absence de nom a en effet été gênante pour les mineurs car ça fait perdre en crédibilité lorsque vous avez le parent : méfiance au moment de la prise de contact. Parfois vous avez le jeune en ligne mais il ne se sent pas concerné et passe un autre interlocuteur. Ça a généré des rappels alors que nous avions déjà la cible au bout du fil. ». En face-à-face, les retours n'étaient pas très négatifs : « Non, pas de problème à partir du moment où les parents ont bien reçu le courrier. Mais avec le prénom ça serait un plus. ». D'ailleurs, le taux de réponse de cette classe d'âge est relativement plus faible par téléphone mais pas en face-à-face (cf. partie « taux de réponse »).

Bilan de l'enrichissement téléphonique

La base Fidéli est très riche en numéros de téléphone. Il a été décidé de garder les mobiles des conjoints pour tenter d'enquêter la personne échantillonnée. Cela nous semblait être un moyen intéressant de toucher la personne échantillonnée. Une recherche annuaire faite par le prestataire est venue compléter le fichier Fidéli. La recherche a porté sur les mobiles et les fixes dont les 09. Les éléments quantitatifs sont résumés dans le tableau 8. Seuls les numéros différents de ceux de Fidéli ont été enregistrés comme « numéros annuaire ». On remarque que la recherche annuaire apporte des numéros de fixe géographique (01 à 05) et des 09 mais quasiment aucun numéro de mobile et fait passer la part de personne sans numéro de téléphone de 31 % à 20 %.

Tableau 8 • Qualification de l'échantillon avec des numéros de téléphone

Part de l'échantillon avec au moins un numéro...	... de mobile pour la personne échantillonnée	... de mobile pour le conjoint	... de fixe pour le ménage	... de box pour le ménage	... de téléphone
Fidéli	31 %	17 %	33 %	3 %	69 %
Recherche annuaire (hors numéros déjà présents dans Fidéli)	1 %	2 %	13 %	8 %	22 %
Total	32 %	19 %	46 %	11 %	80 % (77 % quand on enlève les 09 annuaire)

Lecture : La part de l'échantillon avec au moins un numéro de mobile rattaché à la personne échantillonnée est 31 %.

Champ : France métropolitaine

Source : DREES-Irdes, EHS 2019 (France métropolitaine).

Comme on disposait de beaucoup de numéros de téléphone, on a fait le choix d'exclure les 09 annuaire de la collecte en faisant l'hypothèse qu'ils seraient moins rentables sur le terrain que d'autres numéros. Le but était de resserrer l'algorithme de rotation des numéros de téléphone sur ceux qui avaient le plus de chance d'aboutir à une interview. En effet, la programmation de cet algorithme a été complexe pour le prestataire et n'a pas donné entière satisfaction. On avait peur qu'en laissant trop de numéros par personne, les numéros ayant une bonne chance d'obtenir un « décroché » ne soient pas assez souvent appelés, ou pas le samedi par exemple où la plage horaire des appels était plus restreinte. La répartition du nombre de numéros disponibles au sein de l'échantillon (09 annuaire exclues) fait ressortir un nombre raisonnable de numéros par personne (tableau 9). Seules 1 % des personnes ont plus de 2 numéros. La partie « succès des appels » montre que les numéros de fixe géographique annuaire sont moins rentables que les numéros de fixe Fidéli pour la collecte. La rentabilité des 09 annuaire aurait sans doute été encore moindre. La rentabilité des 09 annuaire a, en effet, été étudiée avec les données de la vague 1 de EpiCov et confirme cette moins bonne rentabilité que les numéros de fixe annuaire, eux-mêmes moins rentables que les numéros de fixe Fidéli (Charrance et al., 2022).

Tableau 9 • Répartition du nombre de numéros de téléphone après enrichissement (hors 09 annuaire)

Nombre de numéros	0	1	2	3	4 ou 5
Part de l'échantillon	23 %	57 %	19 %	1 %	0 %

Lecture > 57 % de l'échantillon a un unique numéro de téléphone dans le fichier après enrichissement annuaire.

Champ > France métropolitaine.

Source > DREES-Irdes, EHS 2019 (France métropolitaine).

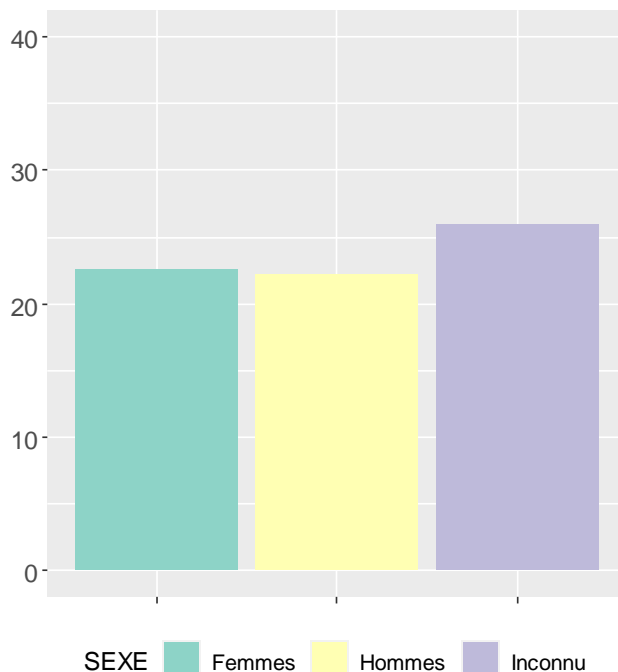
La présence de variables sociales dans Fidéli permet de décliner la probabilité d'avoir ou non un numéro de téléphone à l'issue de l'enrichissement annuaire selon les caractéristiques des personnes. Les différences de probabilités induisent des différences dans les caractéristiques sociales des échantillons mis en collecte par téléphone ou en face-à-face.

Caractéristiques sociales des échantillons du terrain téléphone et face-à-face

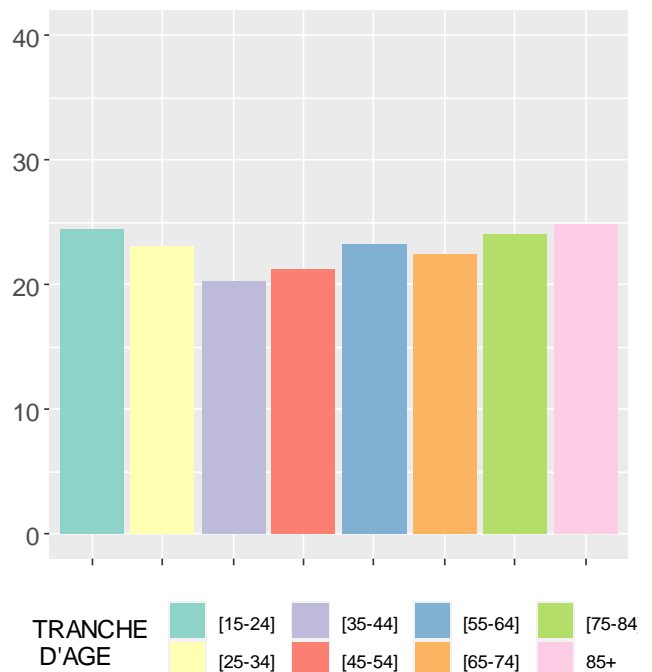
Le graphique 1 présente la probabilité de n'avoir aucun numéro de téléphone suite à l'enrichissement (23 % de l'échantillon, 09 annuaire exclues) selon les caractéristiques sociales (sexe¹¹, tranches d'âge, vivre dans un quartier prioritaire de la politique de la ville (QPV), décile de niveau de vie et situation familiale). On constate que cette probabilité est plus faible parmi les 35-54 ans¹², plus élevée dans un QPV et pour les personnes vivant seules, moins élevée pour les couples avec enfant. On observe également un gradient social : plus le niveau de vie est faible, plus la probabilité de n'avoir aucun numéro de téléphone disponible dans le fichier est élevée.

Graphique 1 • Probabilité de n'avoir aucun numéro de téléphone après enrichissement annuaire selon les caractéristiques sociales (en %)

**A AUCUN numéro
Ménage
15 ans +**



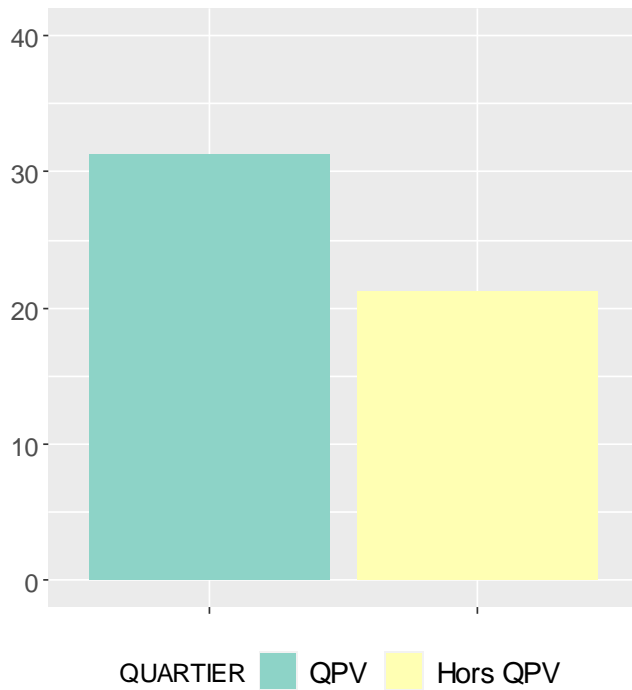
**B AUCUN numéro
Ménage
15 ans +**



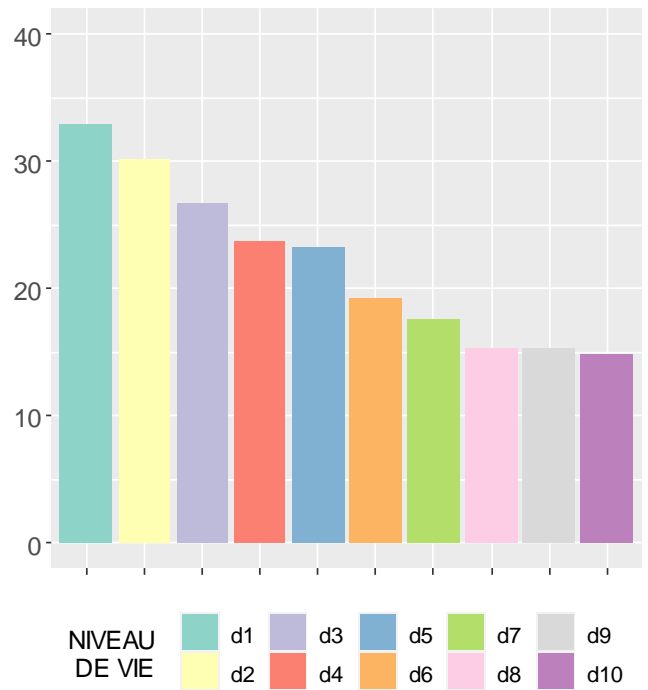
¹¹ Le sexe est manquant pour les plus jeunes dans la base de données, d'où la modalité « Inconnu ».

¹² Pour les plus jeunes de l'échantillon, les numéros récupérés dans Fidéli sont ceux du référent fiscal du ménage ou de son conjoint.

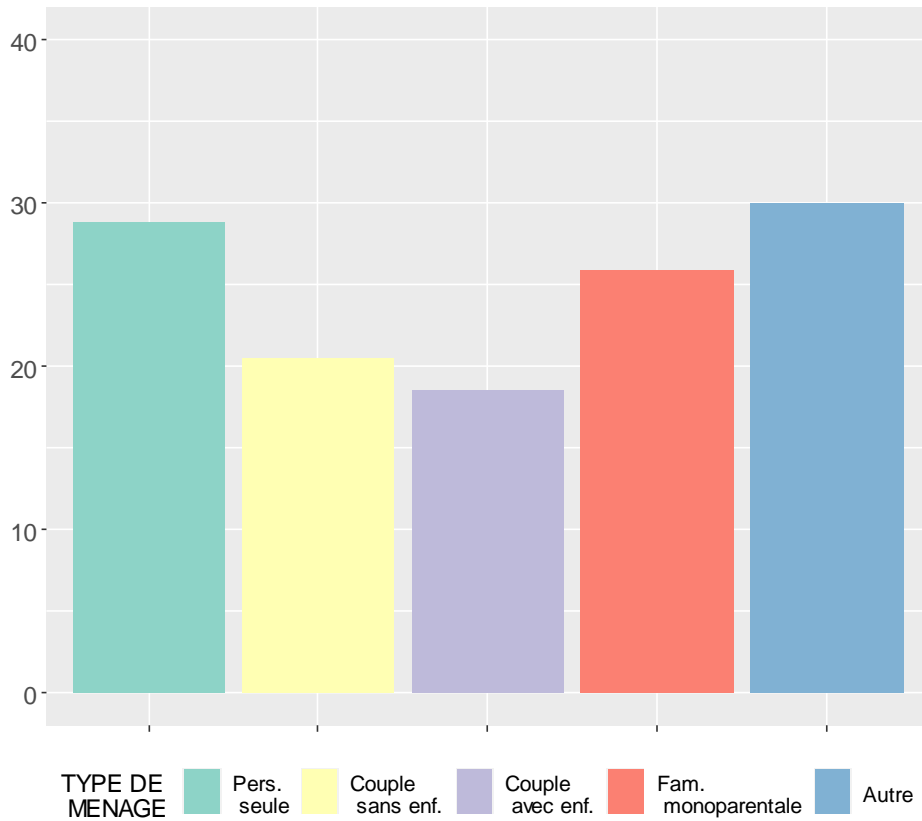
A AUCUN numéro
Ménage
15 ans +



B AUCUN numéro
Ménage
15 ans +



AUCUN numéro
Ménage
15 ans +



Note > Statistiques brutes sur l'échantillon, pas pondérées par le poids de tirage.

Les résultats de la régression de la probabilité de n'avoir aucun numéro de téléphone en fonction de ces caractéristiques sociales sont présentées en annexe 1. On peut résumer ainsi l'effet des caractéristiques sociales (tableau 10) :

Tableau 10 • Effet des caractéristiques sociales sur la probabilité de n'avoir aucun numéro de téléphone à disposition pour l'enquête (cf. annexe 1)

	Hommes	Jeunes	Âgés	QPV	Les plus modestes	Les plus aisés	Pers. seule	Couple avec enfant	Famille monoparentale
Probabilité de n'avoir aucun numéro de téléphone	ns	-	+	+	+	-	+	-	ns

Lecture > La probabilité de n'avoir aucun numéro de téléphone à disposition dans le fichier est plus élevée pour les plus âgés.

Champ > France métropolitaine.

Source > DREES-Irdes, EHS 2019 (France métropolitaine).

De ce fait, les échantillons du terrain téléphone et face-à-face ont des caractéristiques sociales différentes (tableau 11). L'écart est le plus important entre l'échantillon téléphone et l'échantillon face-à-face sans numéro de téléphone. L'échantillon face-à-face sans numéro de téléphone possède relativement légèrement plus de personnes âgées, plus de personnes vivant en QPV, plus de personnes modestes, plus de personnes seules, moins de couples avec enfant. L'échantillon méthodologique, aussi mis en collecte en face-à-face, a été tiré parmi les personnes pour lesquelles un numéro de téléphone est disponible. Il a donc une structure sociale très proche de celle de l'échantillon téléphone. L'échantillon « bascule » est déformé par rapport à l'échantillon téléphone montrant que les plus jeunes, les personnes vivant en QPV, les plus modestes, les couples avec enfant ont été plus fréquemment « IAJ » par téléphone.

Tableau 11 • Caractéristiques sociales des échantillons selon le mode de collecte

	Échantillon téléphone	Échantillon face-à-face sans numéro de téléphone (1)	Échantillon face-à-face méthodologique (2)	Échantillon face-à-face bascule (3)	Échantillon face-à-face (1)+(2)+(3)	Échantillon total
Hommes	46 %	45 %	48 %	44 %	46 %	46 %
Femmes	49 %	49 %	48 %	49 %	49 %	49 %
Sexe inconnu	5 %	6 %	4 %	7 %	5 %	5 %
15-24 ans	14 %	15 %	13 %	18 %	15 %	14 %
25-34 ans	14 %	14 %	12 %	13 %	13 %	14 %
35-44 ans	15 %	13 %	17 %	17 %	14 %	15 %
45-54 ans	17 %	15 %	16 %	16 %	15 %	16 %
55-64 ans	15 %	16 %	16 %	16 %	16 %	16 %
65-74 ans	14 %	13 %	14 %	11 %	13 %	14 %
75-84 ans	7 %	8 %	8 %	5 %	8 %	8 %

85 ans +	4 %	5 %	5 %	4 %	5 %	5 %
QPV	13 %	19 %	6 %	17 %	16 %	14 %
Hors QPV	87 %	81 %	94 %	83 %	84 %	86 %
décile 1 : les 10 % les plus modestes	10 %	16 %	9 %	14 %	14 %	11 %
décile 5	10 %	10 %	9 %	10 %	10 %	10 %
décile 10 : les 10 % les plus aisés	10 %	6 %	12 %	9 %	7 %	9 %
Pers. seule	17 %	24 %	18 %	19 %	22 %	19 %
Couple sans enf.	25 %	23 %	27 %	20 %	24 %	25 %
Couple avec enf.	40 %	31 %	38 %	42 %	33 %	38 %
Famille monoparentale	10 %	12 %	10 %	11 %	11 %	10 %
Autre situation familiale	8 %	11 %	7 %	9 %	10 %	8 %

Note > Statistiques brutes sur l'échantillon, non pondérées par le poids de tirage. Il y a moins de personnes en QPV dans l'échantillon méthodologique car il a été tiré hors suréchantillon QPV. Les déciles de niveau de vie ont été découpés selon les bornes au sein de la population totale.

Lecture > 17 % de l'échantillon sollicité par téléphone vit seul, C'est le cas de 24 % de l'échantillon face-à-face sans numéro de téléphone.

Champ > France métropolitaine.

Source > DREES-Irdes, EHS 2019 (France métropolitaine).

Pour les personnes pour lesquelles on dispose d'un numéro de téléphone pour l'enquête, il est également possible de s'intéresser plus en détails au type de numéro disponible à l'issue de l'enrichissement annuaire. La part de l'échantillon avec un numéro de mobile pour la personne à enquêter est 32 %. Il s'agit d'un mode de contact privilégié pour enquêter la personne. On voit, en effet dans la partie « succès des appels » qu'un numéro de portable augmente la probabilité d'obtenir un « décroché ».

Type de numéro de téléphone disponible selon les caractéristiques sociales dans l'échantillon du terrain téléphone

Le graphique 2 présente la probabilité d'avoir un certain type de numéro de téléphone (mobile Fidéli pour la personne échantillonnée, mobile Fidéli pour une autre personne du logement, fixe géographique Fidéli, 09 Fidéli ou fixe géographique annuaire) selon les caractéristiques sociales parmi l'échantillon du terrain téléphone. Les mobiles annuaire ne sont pas considérés comme ils sont très peu fréquents, ni les 09 annuaire, comme ces numéros n'ont pas été utilisés pour le terrain.

On constate que le fait d'avoir un numéro de mobile est lié au sexe : les hommes sont plus concernés. Il est possible que ce soit expliqué par les déclarations d'impôts communes qui seraient plus fréquemment remplies par les hommes qui laisseraient alors leurs propres coordonnées. Cet écart entre les sexes est corrigé au niveau du ménage par l'indication du mobile d'une autre personne du logement, le plus souvent une femme, en plus de celui du référent fiscal. Cependant la partie « succès des appels » a montré que le mobile d'une autre personne du logement est moins efficace que le mobile de la personne échantillonnée pour avoir une interview. Les personnes de sexe inconnu dans Fidéli sont le plus souvent des mineurs. Il n'y a pas de numéro de mobile disponible dans Fidéli pour eux car ce ne sont pas eux qui font la déclaration fiscale mais une autre personne du logement (le plus souvent un des parents).

La probabilité d'avoir un numéro de mobile à disposition dans le fichier décroît fortement avec l'âge mais à l'inverse, celle d'avoir un numéro de fixe (Fidéli ou annuaire) augmente fortement avec l'âge. Ces effets se compensent si bien que la probabilité d'avoir à disposition un numéro de téléphone est assez peu différente selon l'âge, mais la répartition entre mobile et fixe est très différente selon l'âge.

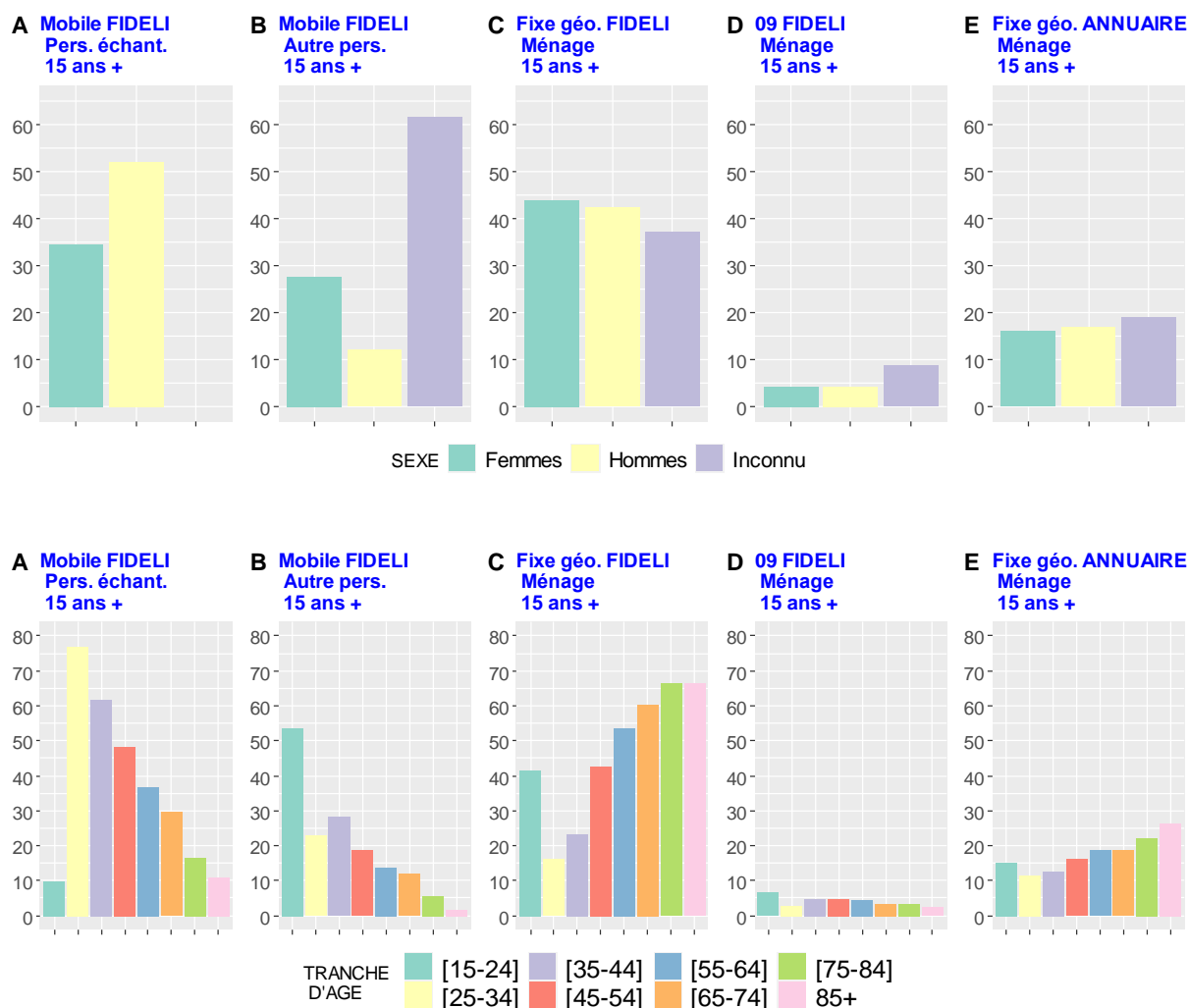
La probabilité d'avoir, pour l'échantillon du terrain téléphone, un numéro de mobile est plus élevée dans les QPV que hors des QPV. À l'inverse, le fait d'avoir dans le fichier un numéro de fixe est moins fréquent dans les QPV que hors des QPV. S'il est moins fréquent d'avoir à disposition un numéro de téléphone pour l'enquête dans les QPV, on constate donc, que, quand un numéro est disponible, c'est plus fréquemment celui d'un mobile que celui d'un fixe. Hors des QPV, c'est, à l'inverse, plus

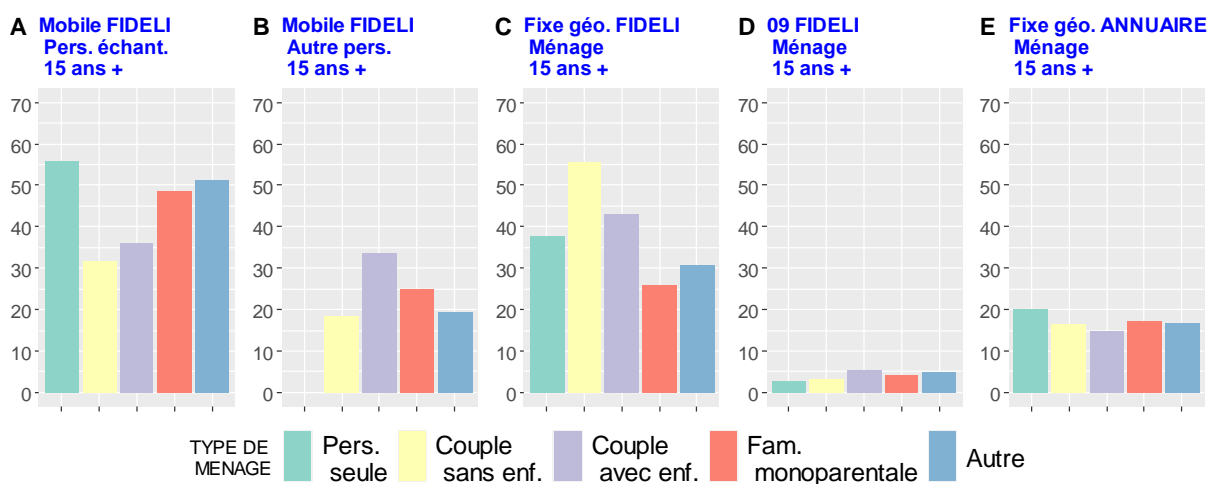
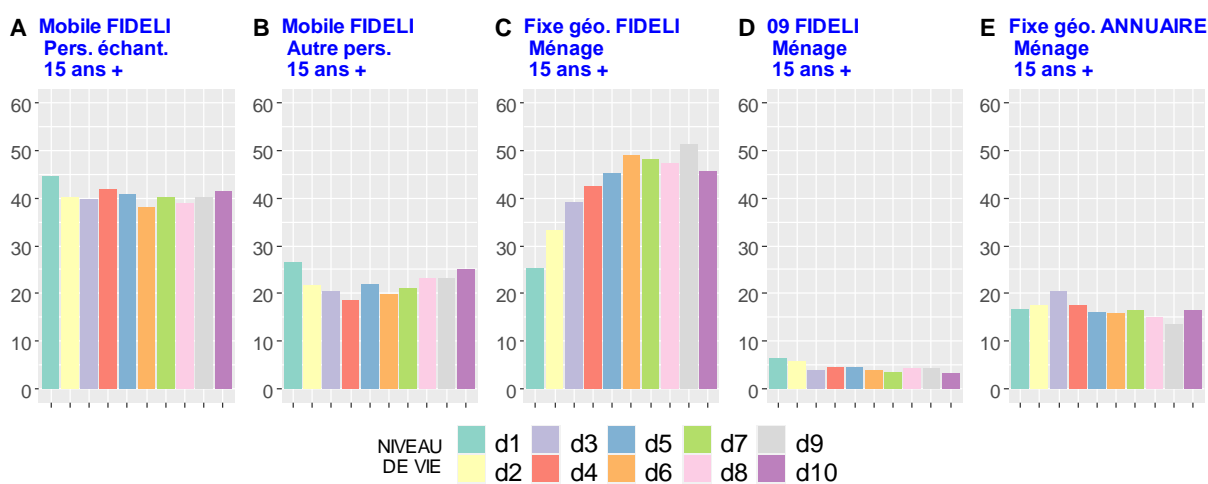
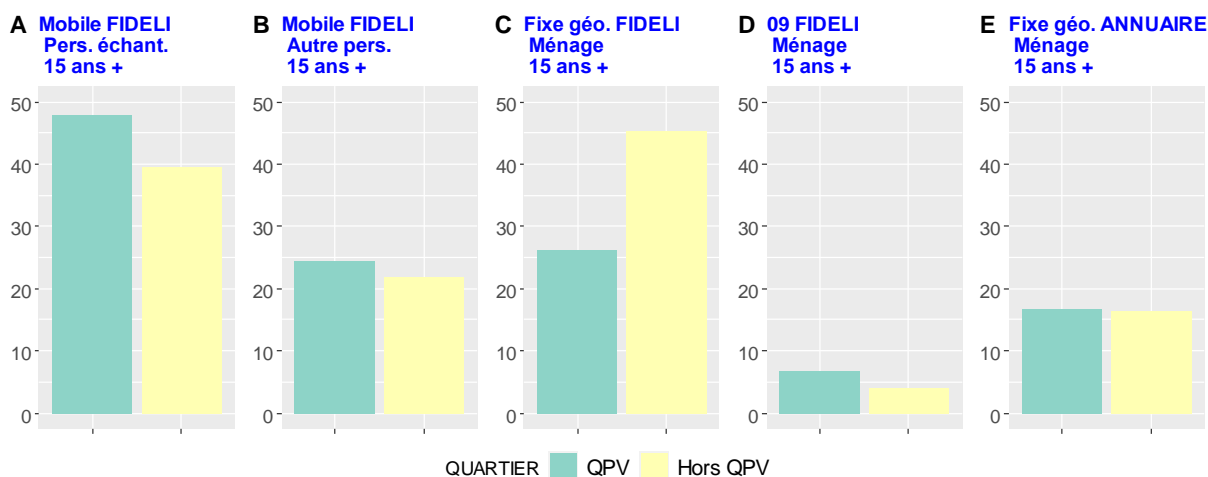
fréquemment celui d'un fixe que celui d'un mobile. Cela peut être lié à la structure par âge ainsi qu'à la structure sociale de la population de ces quartiers. En effet, plus le décile de niveau de vie est élevé, plus la probabilité de disposer d'un fixe Fidéli pour l'enquête est élevée. Pour les numéros de mobile, la probabilité est assez stable selon le niveau de vie.

Pour l'échantillon pour lequel on dispose d'au moins un numéro de téléphone, la probabilité d'avoir un numéro de mobile dans Fidéli est assez stable selon le décile de niveau de vie. En revanche, la probabilité d'avoir un numéro de fixe géographique augmente avec le décile de niveau de vie. C'est sans doute corrélé à l'âge.

Les couples sans enfant ont le moins fréquemment, parmi les numéros à disposition pour l'enquête, un mobile Fidéli et le plus fréquemment un numéro de fixe Fidéli. Il y a un effet d'âge derrière ce résultat. En revanche, il reste un écart entre les couples avec enfant et les familles monoparentales en termes de type de numéro disponible. La probabilité d'avoir un numéro de mobile est plus élevée pour les familles monoparentales que pour les couples avec enfant ; c'est l'inverse pour les fixes géographiques Fidéli. Les personnes seules qui ont un mobile Fidéli sont en moyenne plus jeunes que celles qui ont un fixe Fidéli.

Graphique 2 • Probabilité d'avoir un certain type de numéro de téléphone selon les caractéristiques sociales dans l'échantillon du terrain téléphone (en %)





Note > Statistiques brutes sur l'échantillon, non pondérées par le poids de tirage.

Lecture> Parmi l'échantillon mis en collecte sur le terrain téléphone, 36 % des femmes ont au moins un numéro de mobile les concernant en propre contre 56 % des hommes.

Les résultats de la régression de la probabilité d'avoir pour les personnes de l'échantillon téléphone un numéro de mobile en fonction de ces caractéristiques sociales sont présentées en annexe 1. On peut résumer ainsi l'effet des caractéristiques sociales (tableau 12) :

Tableau 12 • Effet des caractéristiques sociales sur la probabilité d'avoir un numéro de mobile à disposition pour l'enquête (champ des personnes du terrain téléphone, cf. annexe 2)

	Hommes	Jeunes	Vieux	QPV	Les plus modestes	Les plus aisés	Pers. seule	Couple avec enfants	Famille monoparentale
Probabilité d'avoir un numéro de mobile	+	+	-	+	ns	+	+	-	+

Lecture > Parmi les personnes pour lesquelles un numéro de téléphone est disponible dans le fichier, la probabilité d'avoir un numéro de mobile est plus faibles pour les plus âgés.

Champ > France métropolitaine.

Source > DREES-Irdes, EHS 2019 (France métropolitaine).

L'échantillon disposant d'un mobile pour la personne échantillonnée (ce qui est un atout pour la collecte, cf. partie « succès des appels ») est donc déformé par rapport à l'ensemble de l'échantillon téléphone. Il est relativement plus masculin, plus jeune (sauf les très jeunes) et est caractérisé par plus de personnes seules et de familles monoparentales que l'échantillon téléphone (tableau 13).

Tableau 13 • Caractéristiques sociales de l'échantillon téléphone et téléphone avec numéro de mobile de la personne échantillonnée

	Échantillon téléphone	Échantillon mobile personne échantillonnée
Hommes	46 %	58 %
Femmes	49 %	42 %
Sexe inconnu	5 %	0 %
15-24 ans	14 %	3 %
25-34 ans	14 %	25 %
35-44 ans	15 %	23 %
45-54 ans	17 %	20 %
55-64 ans	15 %	14 %
65-74 ans	14 %	10 %
75-84 ans	7 %	3 %
85 ans +	4 %	1 %
QPV	13 %	14 %
Hors QPV	87 %	86 %
Décile 1 : les 10 % les plus modestes	10 %	11 %

Décile 5	10 %	10 %
Décile 10 : les 10 % les plus aisés	10 %	10 %
Personne seule	17 %	23 %
Couple sans enfant	25 %	20 %
Couple avec enfant	40 %	36 %
Famille monoparentale	10 %	12 %
Autre situation familiale	8 %	9 %

Note > Statistiques brutes sur l'échantillon, non pondérées par le poids de tirage.

Lecture > Les 25-34 ans représentent 14 % de l'échantillon téléphone et 25 % de l'échantillon téléphone avec un numéro de mobile pour la personne échantillonnée

Champ > France métropolitaine.

Source > DREES-Irdes, EHS 2019 (France métropolitaine).

Les différences de caractéristiques sociales des différents échantillons mis en collecte peuvent avoir, comme le mode de collecte utilisé, une incidence sur le taux de réponse. La partie consacrée au bilan de collecte se pose la question de savoir si cette partition de l'échantillon par mode de collecte selon la disponibilité d'un numéro de téléphone semble être efficace pour la collecte au regard des taux de réponse attendus selon l'effet conjugué des caractéristiques sociales et du mode de collecte.

Taux de réponse

Le taux de réponse à l'enquête est 55 % (tableau 14). Il est un peu plus élevé parmi les femmes que parmi les hommes (57 % versus 55 %). Il augmente avec l'âge jusqu'à la tranche 65-74 ans ; au-delà il baisse. Il est le plus bas aux âges extrême de l'échantillon : environ 50 % parmi les 15-34 ans et 41 % parmi les 85 ans ou plus. Le taux de réponse est plus faible de 7 points dans les QPV comparé aux autres quartiers. Il augmente graduellement avec le niveau de vie : de 48 % pour le premier décile de niveau de vie jusqu'à 63 % pour le huitième décile avant de redescendre un peu parmi les deux déciles le plus aisés (60 % pour le dernier décile). Les personnes seules et les familles monoparentales ont des taux de réponse plus faibles que les couples avec ou sans enfants.

Les taux de réponse varient fortement selon les échantillons de collecte. Le taux de réponse global est de 45 % pour l'échantillon téléphone contre 63 % pour l'échantillon face-à-face.

Le meilleur taux de réponse, de loin, est obtenu pour l'échantillon méthodologique du terrain face-à-face (68 %). On peut comparer les taux de réponse par caractéristiques sociales de cet échantillon à ceux de l'échantillon du téléphone comme l'échantillon méthodologique a été tiré aléatoirement parmi ceux qui disposent dans le fichier d'un numéro de téléphone. L'écart de taux de réponse est de 23 points, mettant clairement en évidence les meilleurs résultats de la collecte face-à-face par rapport à la collecte téléphone. L'écart de taux de réponse entre les deux échantillons est d'autant plus fort que l'âge s'élève pour atteindre 30 point parmi les 74-85 ans et 47 points parmi les 85 ans ou plus. Il est également fort pour les 15-24 ans (30 points) sans doute du fait de la difficulté à enquêter par téléphone les plus jeunes sans disposer du prénom. L'écart est stable et du niveau moyen (22-25 point) dans et hors des QPV. Il est le plus fort lorsque le niveau de vie est le plus faible. Ainsi, il est de 31 points pour le premier décile contre 16 points pour le dernier décile. La collecte face-à-face est donc particulièrement rentable parmi les ménages les plus modestes. L'écart est plus faible que la moyenne pour les personnes vivant seules (15 points) et plus élevé pour les couples avec ou sans enfant (25-26 points). Donc, si l'idéal serait une collecte en face-à-face pour tout le monde, en termes d'optimisation de la collecte à moindre coût, on aurait intérêt à enquêter plutôt les jeunes dont on ne dispose pas du prénom, les plus âgés, les plus modeste et les couples avec enfants en face-à-face et de privilégier le téléphone pour les autres car l'écart de taux de réponse est plus faible.

On peut rapprocher ces résultats de la déformation en termes de caractéristiques sociales de l'échantillon face-à-face sans numéro de téléphone. La partie précédente concluait que ne disposer d'aucun numéro de téléphone, et donc se retrouver sur le terrain face-à-face, était lié positivement à un âge élevé, au fait de résider en QPV, d'être dans les premiers déciles de niveau de vie et de vivre seul. Or, on constate que le face-à-face est particulièrement rentable par rapport au téléphone pour les plus âgés et les personnes d'un niveau de vie modeste. Donc, une partie de la bonne rentabilité du terrain de l'échantillon face-à-face sans numéro de téléphone comparée à celle du terrain téléphone doit tenir dans la déformation des caractéristiques sociales de cet échantillon. Pourtant, le taux de réponse de l'échantillon face-à-face sans numéro de téléphone (63 %) reste en deçà de celui de l'échantillon méthodologique (68 %). C'est notamment le cas pour les plus âgés (environ -10 points pour les 65-84 ans et -17 points pour les 85 ans ou plus) et les plus aisés (-13 points parmi le dernier décile). Cela tend à prouver qu'il y a un effet de sélection du fait de ne pas disposer de numéro de téléphone dans le fichier qui rend la collecte plus complexe même en face-à-face sur les plus âgés qui répondent généralement particulièrement mieux en face-à-face. En revanche, pour les personnes de niveau de vie modeste, le taux de réponse de l'échantillon sans numéro de téléphone est très proche de celui de l'échantillon méthodologique. Il semble donc ne pas y avoir d'effet de sélection à ne pas disposer de numéro de téléphone dans le fichier. Par ailleurs, le taux de réponse face-à-face est plus élevé que le taux de réponse de l'échantillon téléphone particulièrement pour cette population (26 points de plus pour le premier décile contre 18 points d'écart moyen de taux de réponse). Cela tend à mettre en évidence que la partition induite par le fait de ne pas disposer de numéro de téléphone qui conduit à surreprésenter les plus modestes sur le terrain du face-à-face est plutôt rentable pour la collecte de l'enquête. Pour les personnes résidant en QPV, il semble aussi ne pas y avoir d'effet de sélection car le taux de réponse de l'échantillon sans numéro de téléphone et de l'échantillon méthodologique sont égaux (60 %). Par ailleurs, l'écart de taux de réponse entre la collecte téléphone et face-à-face sur cette population est plus élevé que l'écart moyen sur l'échantillon (24 points de plus contre 18 points). Il y a donc une rentabilité particulière à enquêter également cette population en face-à-face par rapport à une autre population. En outre, comme son taux de réponse est très faible par téléphone, il est préférable d'enquêter cette population en face-à-face. Pour les personnes qui vivent seules, on voit que la rentabilité du face-à-face est plutôt faible (+15 points de collecte entre l'échantillon méthodologique et l'échantillon téléphone contre 23 points en moyenne) et qu'elle est d'autant plus faible dans l'échantillon sans numéro de téléphone (taux de réponse de 55 % contre 59 % pour l'échantillon méthodologique). Il s'agit peut-être d'un groupe qui répondrait finalement pour partie bien par internet si on pouvait réduire la durée du questionnaire ou le proposer en plusieurs fois et qu'on disposait pour eux d'une adresse mél. Sinon, il faudrait parvenir à augmenter l'inscription d'un numéro de téléphone dans les déclarations fiscales dans ce groupe. En tous cas, si le face-à-face sur les personnes vivant seules sans numéro de téléphone était dans le contexte de la collecte 2019 incontournable, il ne s'avère pas très rentable sur cette population.

L'échantillon « bascule » a le plus faible taux de réponse du face-à-face (58 %). Ceci dit, c'est un taux de réponse plus élevé que celui de l'échantillon téléphone alors qu'il s'agit d'IAJ du téléphone. Cela met bien en évidence l'efficacité du multimode séquentiel : commencer par le téléphone pour réduire les coûts, puis interroger en face-à-face avec encore une bonne rentabilité sur les non-répondants du téléphone. Il serait intéressant d'avoir des informations à partir d'autres enquêtes sur cette efficacité sur les refus du téléphone. Si l'on compare les résultats de l'échantillon « bascule » à ceux de l'échantillon méthodologique, comme ils disposent tous deux de personnes avec un numéro de téléphone, on se rend compte que les taux de réponse de l'échantillon « bascule » sont particulièrement faibles passé 65 ans. D'une certaine manière, si on n'arrive pas à joindre par téléphone cette population pour l'enquête, c'est mauvais signe pour le face-à-face derrière. En revanche, l'écart entre l'échantillon « bascule » et l'échantillon méthodologique est plus faible (- 7 points) que la moyenne (- 10 points) sur les 15-24 ans mettant en évidence à nouveau des difficultés de collecte par téléphone sur les plus jeunes, peut-être du fait de barrages des parents et donc l'efficacité relative du face-à-face sur cette population. Il en est de même pour les personnes appartenant au deuxième décile, vivant en QPV et celles appartenant à un couple avec enfant, une famille monoparentale ou à une autre situation familiale. L'écart de taux de réponse entre l'échantillon « bascule » et l'échantillon méthodologique est plus faible que la moyenne montrant une efficacité relative plus importante de la bascule parmi ces personnes. On voit que la déformation en termes de caractéristiques sociales de l'échantillon « bascule » (plus jeunes, plus de personnes vivant en QPV, plus de personnes modestes, plus de couples avec enfants) s'accompagne d'une rentabilité relative meilleure du terrain face-à-face sur ces personnes spécifiquement. La déformation de l'échantillon bascule serait alors bien liée à une difficulté à joindre ces personnes en particulier par téléphone plutôt qu'à un refus de l'enquête déguisé par une impossibilité à joindre. La partie « succès des appels » montre, en effet, que ce sont des groupes où il est plus difficile d'avoir la personne échantillonnée au téléphone du fait d'un taux plus faible de « décroché » (manque de disponibilité ou filtrage des appels) ou du fait de barrages de proches alors qu'une fois qu'on a la personne échantillonnée au téléphone, elle n'accepte l'interview pas moins que les autres. Cela vient renforcer l'intérêt d'une approche séquentielle en termes de mode de collecte.

En ce qui concerne l'efficacité du terrain téléphone, le taux de réponse est plus important lorsqu'un numéro de mobile de la personne échantillonnée est disponible pour la collecte. Au total, le taux de réponse est supérieur de 5 points à celui de l'échantillon téléphone global. La hausse est plus forte passé 64 ans. Cela traduit sans doute un effet de sélection, les personnes âgées ayant un mobile étant sans doute celles en meilleure santé et donc les plus aisées à enquêter quel que soit le mode de collecte. Sinon, la hausse est répartie de manière assez égalitaire quelles que soient les caractéristiques sociales et conduit à améliorer les taux de réponse des groupes surreprésentés parmi ceux pour lesquels un numéro de mobile est disponible. En revanche, l'échantillon qui dispose d'un numéro de mobile pour le conjoint mais pas pour la personne échantillonnée a un taux de réponse légèrement plus faible que l'échantillon téléphone global (- 1 point). Il est difficile à ce stade d'en tirer des conclusions fermes sur l'efficacité des mobiles pour la collecte car ces résultats dépendent aussi du nombre d'appels sur ces téléphones en particulier, des moments où les appels sont passés, etc. La partie « bilan de collecte téléphone » investigate ce sujet en considérant l'historique des appels et conclut à une plus forte efficacité des mobiles de la personne échantillonnée dans le collecte mais pas de celle des mobiles du conjoint qui est due à un taux de contact avec la personne échantillonnée meilleur avec les numéros de mobiles de la personne échantillonnée.

Tableau 14 • Taux de réponse selon les caractéristiques et les échantillons de collecte

	Échantillon téléphone	Échantillon 06 personne échantillonnée	Échantillon 06 autre pers. du ménage	Échantillon face-à-face	Échantillon sans numéro de téléphone	Échantillon méthodologique	Échantillon bascule	Échantillon total
Hommes	45 %	50 %	41 %	62 %	62 %	67 %	57 %	55 %
Femmes	47 %	51 %	47 %	63 %	63 %	70 %	59 %	57 %
Sexe inconnu	31 %		33 %	61 %	60 %	65 %	59 %	47 %
15-24 ans	37 %	43 %	37 %	61 %	59 %	67 %	60 %	51 %
25-34 ans	44 %	48 %	37 %	49 %	50 %	51 %	43 %	50 %

35-44 ans	46 %	49 %	45 %	60 %	62 %	65 %	52 %	55 %
45-54 ans	46 %	50 %	47 %	65 %	64 %	68 %	65 %	57 %
55-64 ans	49 %	56 %	55 %	69 %	70 %	72 %	65 %	61 %
65-74 ans	52 %	57 %	59 %	71 %	69 %	79 %	65 %	62 %
75-84 ans	48 %	59 %	67 %	67 %	66 %	77 %	57 %	57 %
85 ans +	27 %	37 %	57 %	60 %	56 %	73 %	54 %	41 %
QPV	34 %	39 %	30 %	59 %	60 %	60 %	54 %	49 %
Hors QPV	47 %	53 %	46 %	63 %	63 %	69 %	59 %	56 %
Décile 1 : les 10 % les plus modestes	31 %	37 %	27 %	58 %	59 %	62 %	51 %	48 %
Décile 2	36 %	39 %	33 %	63 %	64 %	63 %	60 %	51 %
Décile 3	37 %	43 %	36 %	62 %	62 %	68 %	58 %	51 %
Décile 4	43 %	49 %	40 %	63 %	67 %	67 %	50 %	54 %
Décile 5	44 %	49 %	49 %	64 %	65 %	66 %	62 %	55 %
Décile 6	47 %	55 %	51 %	65 %	65 %	70 %	60 %	57 %
Décile 7	48 %	50 %	49 %	65 %	65 %	70 %	60 %	57 %
Décile 8	54 %	60 %	51 %	68 %	65 %	75 %	65 %	63 %
Décile 9	55 %	64 %	52 %	67 %	68 %	71 %	58 %	62 %
Décile 10 : les 10 % les plus aisés	53 %	60 %	52 %	62 %	58 %	71 %	62 %	60 %
Personne seule	44 %	50 %		55 %	55 %	59 %	51 %	52 %
Couple sans enfant	50 %	56 %	55 %	68 %	68 %	77 %	59 %	60 %
Couple avec enfant	45 %	51 %	44 %	66 %	66 %	70 %	62 %	56 %
Famille monoparentale	38 %	47 %	31 %	61 %	63 %	62 %	54 %	51 %
Autre situation familiale	39 %	43 %	36 %	57 %	58 %	58 %	55 %	50 %
Total	45 %	51 %	44 %	63 %	63 %	68 %	58 %	55 %

Lecture > Le taux de réponse à l'enquête des hommes est 45 % dans l'échantillon téléphone et 50 % dans l'échantillon téléphone avec un numéro de mobile pour la personne échantillonnée.

Champ > France métropolitaine.

Source > DREES-Irdes, EHS 2019 (France métropolitaine).

Déroulement du terrain téléphone : très différent selon la vague d'enquête

Le terrain téléphonique s'est déroulée en deux vagues. La première vague a eu lieu du 30 avril au 27 juillet 2019 et la seconde vague du 27 août au 28 décembre 2019.

Un total de 3 738 questionnaires ont été réalisés en vague 1 et 4 353 en vague 2. Le taux de réponse¹³ de la vague 2 est un peu plus faible que celui de la vague 1 (44 % versus 46 %).

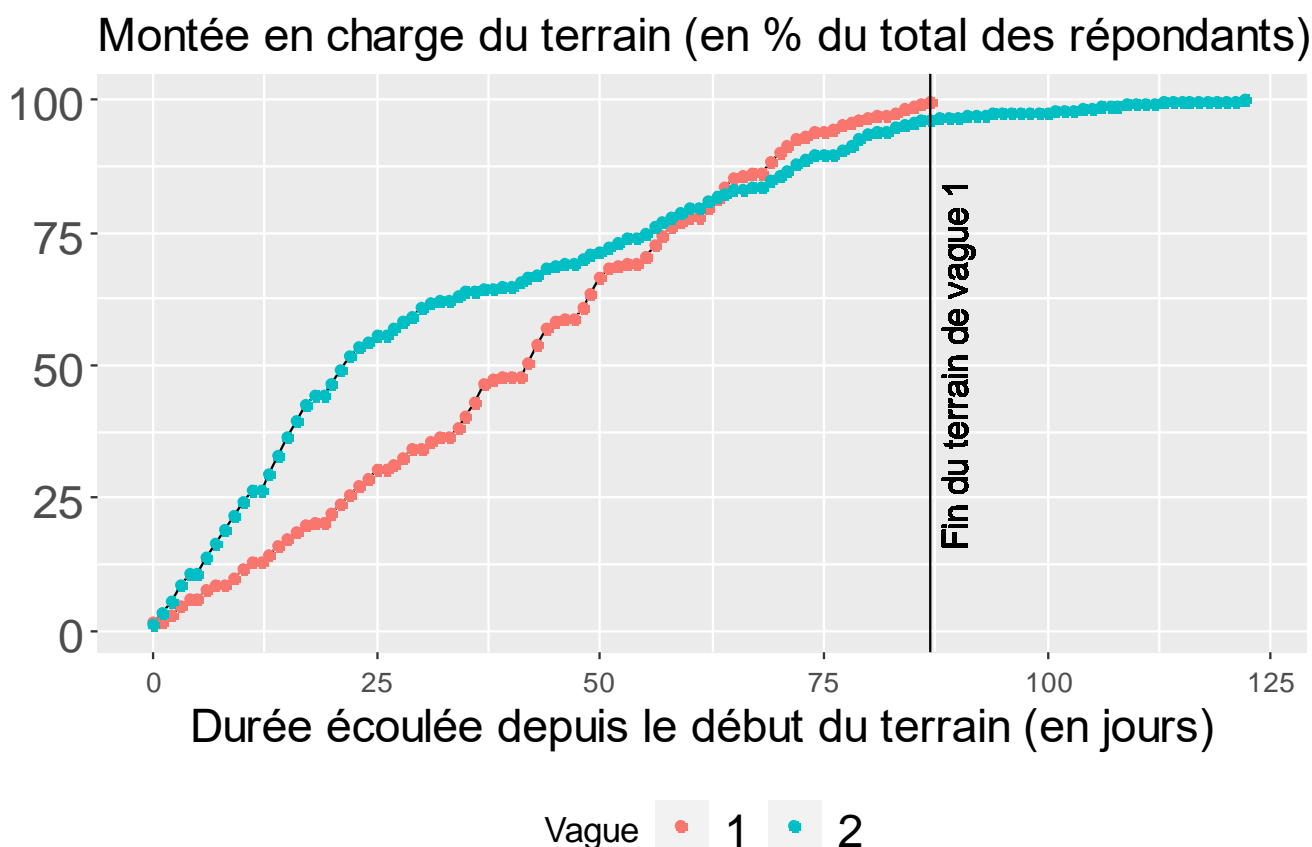
Description des interviews

La dynamique de montée en charge des réponses est très différente entre la vague 1 et la vague 2 (graphique 3). Le début du terrain de la vague 1 a été marqué par un manque important d'enquêteurs disponibles, ce qui est très visible dans le nombre d'interview réalisées. Par ailleurs le terrain téléphonique de la vague 1 a duré un mois de moins que celui de la vague 2 si bien qu'on n'observe pas vraiment de ralentissement dans la montée en charge des réponses, la courbe est assez linéaire. Au contraire en vague 2, on constate un tassement des réponses à la fin de l'enquête. À l'observation des courbes, il semble possible que le nombre d'interviews eût augmenté si la vague 1 n'avait pas buté sur le mois d'août où le terrain a été arrêté.

Le prestataire a confirmé que peu d'enquêteurs étaient présents en vague 1 en fin de semaine. On observe clairement moins d'interviews en fin de semaine en vague 1 comparé à la vague 2 (graphique 4). Il est probable que plus d'enquêteurs qu'en vague 1 étaient affectés à l'enquête en matinée la semaine lors de la vague 2. En effet, on observe plus d'interviews le matin en semaine en vague 2 qu'en vague 1. Le samedi, il y a peu d'écart entre les vagues.

La comparaison des deux vagues montre l'impact de l'organisation du terrain sur la collecte.

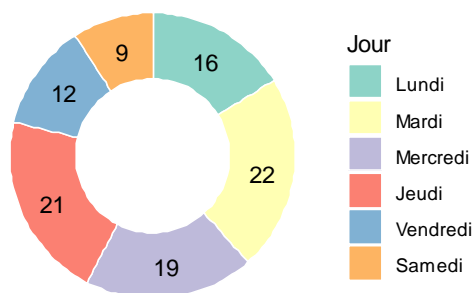
Graphique 3 • Montée en charge du terrain : des dynamiques très différentes entre la vague 1 et la vague 2



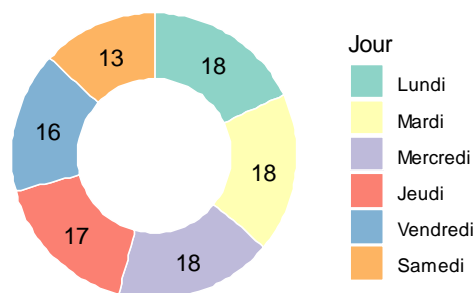
¹³ Pour le calcul de ces taux de réponse, les personnes IAJ par téléphone puis basculées en face-à-face et répondantes à ce moment-là sont incluses dans le calcul et considérées comme non-répondants sur le terrain du téléphone.

Graphique 4 • Jours et heures des interviews : des écarts entre la vague 1 et la vague 2

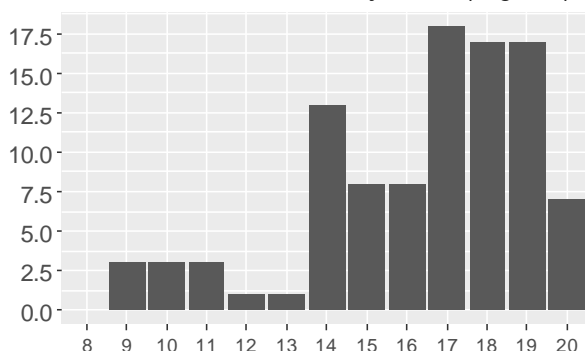
A CATI : % / jour de la semaine (vague 1)



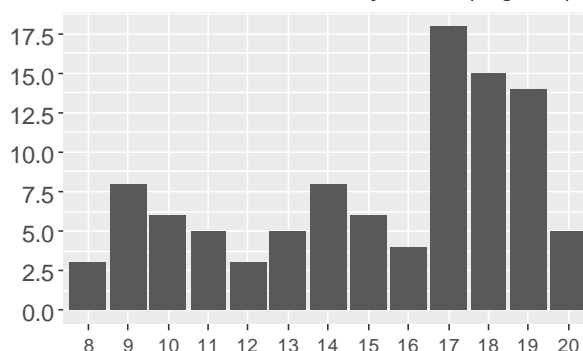
D CATI : % / jour de la semaine (vague 2)



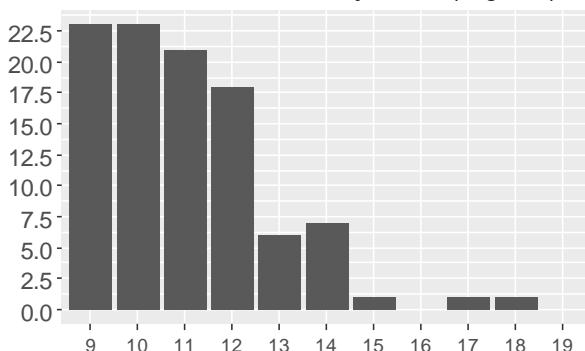
B Semaine - % / heure de la journée (vague 1)



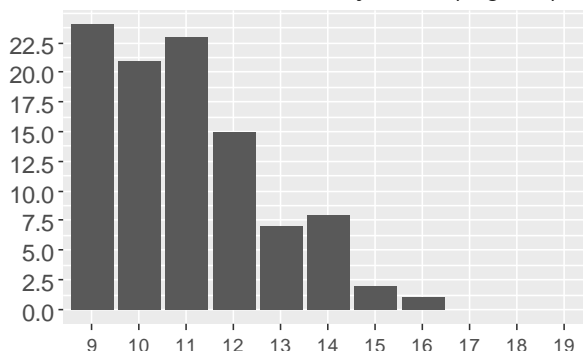
E Semaine - % / heure de la journée (vague 2)



C Samedi - % / heure de la journée (vague 1)



F Samedi - % / heure de la journée (vague 2)



Un protocole d'appels avait été défini de manière à exploiter au mieux tous les numéros de téléphone disponibles. Ce protocole comportait quatre critères que chaque fiche adresse devait respecter avant que son exploitation ne cesse. Les critères étaient les suivants :

- 25 appels par numéro dans la mesure où celui-ci appartient à la personne échantillonnée ; 10 par numéro s'il s'agit d'un numéro du conjoint ; un numéro de fixe est toujours considéré comme appartenant à la personne échantillonnée ; quand un unique numéro est disponible le nombre d'appels est toujours 25 ;
- 5 appels au moins après 18h par personne à enquêter ;
- 5 appels au moins le samedi par personne à enquêter ;
- Une fiche adresse est exploitée pendant au moins 18 jours ouvrés.

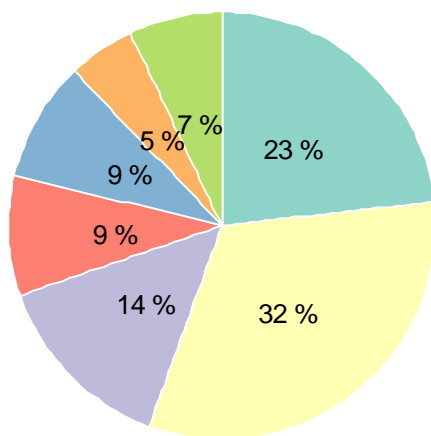
De ce protocole d'appels a découlé la programmation par le prestataire d'un algorithme de rotation des numéros de téléphone. Cet algorithme n'a pas bien fonctionné si bien qu'environ 30 % des fiches adresse en vague 1 et 20 % en vague 2 n'ont pas respecté au moins un de ces quatre critères à l'issue du terrain. L'algorithme qui a mal fonctionné en vague 1 a été reprogrammé pour la vague 2. Cela s'est conjugué avec une plus forte quotité enquêteurs en début de vague 2 par rapport au début de vague 1. Ainsi les numéros ont fortement tourné en début de vague 2 et le nombre maximal d'appels a été atteint très vite et a alors bloqué les numéros alors que les autres critères (soirée, samedi) n'étaient pas atteints. Le prestataire a finalement pu débloquer certains numéros pour les relancer sur le terrain, si bien que la part des interviews à l'issue d'un grand nombre d'appels a été plus important en vague 2 qu'en vague 1 (graphique 5).

Même si le nombre de fiches adresse qui ne respectaient pas les quatre critères a été plus bas en vague 2 qu'en vague 1, le taux de réponse a été plus faible en vague 2. On pense que l'exploitation très rapide des numéros en début de terrain a finalement été défavorable au terrain. On observe aussi à l'issue de la vague 2 un nombre beaucoup plus important qu'en vague 1 d'IAJ qui vont être en partie basculés sur le terrain face-à-face : 1394 (10 % du terrain téléphone de la vague 2) contre 822 (6 %) en vague 1. C'est également le signe que le terrain téléphone a été moins rentable. La partie « succès des appels » en détaille les causes possibles.

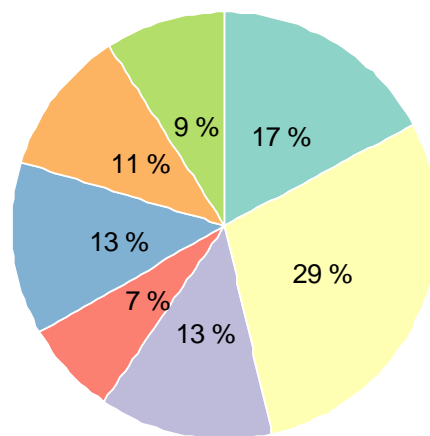
En revanche, les numéros sur lesquels ont eu lieu les interviews sont en proportion les mêmes (graphique 6).

Graphique 5 • Nombre d'appels pour avoir une interview : plus d'interviews après un grand nombre d'appels en vague 2

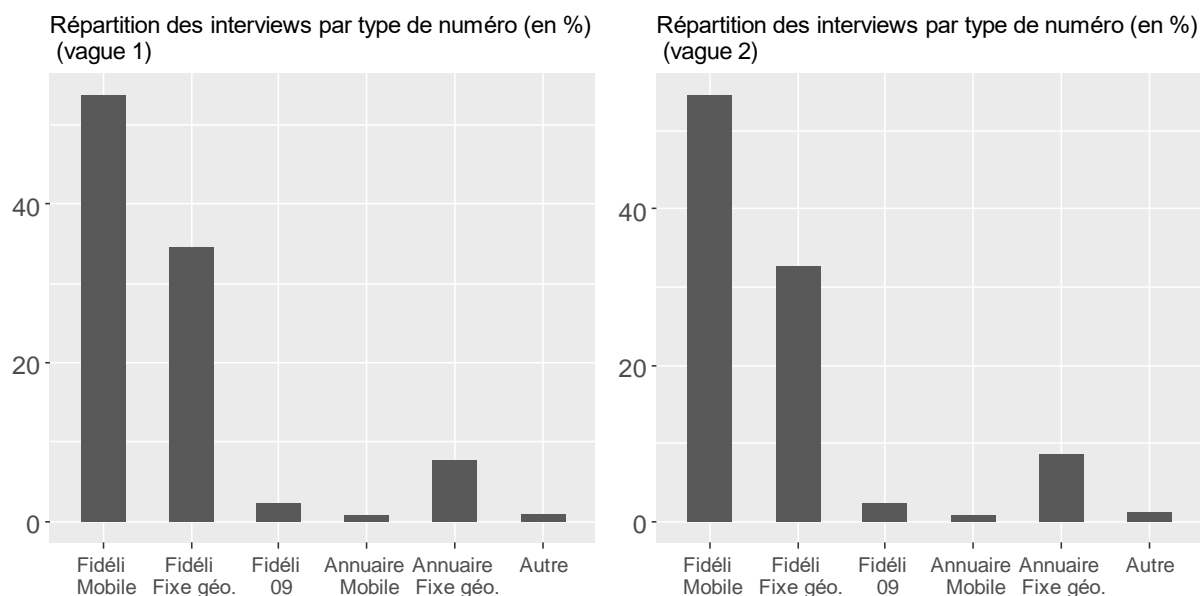
Nombre d'appels pour avoir une interview (vague 1)



Nombre d'appels pour avoir une interview (vague 2)



Graphique 6 • Répartition des interviews par type de numéros de téléphone : identique selon les vagues



Succès des appels

Cela pose la question du succès des appels. Il est intéressant d'examiner quels sont les appels les plus rentables dans le protocole de l'enquête du point de vue des jours, horaires, numéros appelés en fonction de la vague d'enquête et des caractéristiques sociales de la personne à enquêter. La rentabilité a été étudiée comme le fait d'avoir un « décroché », un contact avec la personne échantillonnée et une interview.

Cette analyse a été faite grâce à l'historique des appels de l'enquête qui enregistre pour chaque personne dont on connaît les caractéristiques sociales dans Fidéli l'ensemble des appels réalisés avec leur jour et moment et les numéros appelés. Au total, l'historique des appels contient ces informations à propos des 467 948 appels passés pour l'enquête.

Quatre types de modèle ont été estimés. Les variables à expliquer sont les suivantes : la probabilité de décrocher, la probabilité d'avoir un contact avec la personne échantillonnée sachant qu'il y a eu un « décroché », la probabilité d'avoir une interview sachant qu'il y a eu un contact avec la personne échantillonnée et la probabilité d'avoir une interview. Les trois premiers modèles décomposent les mécanismes qui mènent au quatrième modèle. Les variables explicatives sont les caractéristiques socio-économiques des personnes, les moments des appels, le type de numéro utilisé et le fait d'avoir été basculé sur un « serveur refus » (en cas de refus, les fiches adresse étaient sollicitées à nouveau par des enquêteurs expérimentés pour essayer d'obtenir tout de même une interview). Les modèles ont été estimés séparément pour la vague 1 et la vague 2 (tableau 15). La constante représente la probabilité correspondant à chaque variable à expliquer pour la situation de référence. Chaque coefficient exprime de combien cette probabilité baisse ou augmente en points de pourcentage lorsque la situation dévie de la situation de référence pour la modalité considérée.

Concernant l'effet des caractéristiques sociales, on constate que la plus faible probabilité d'avoir une interview quand la personne échantillonnée est un homme est liée à la fois à une plus grande difficulté à avoir un contact avec la personne échantillonnée alors que le taux de décroché est le même et à une plus faible acceptation des interviews parmi les hommes lorsque la personne échantillonnée a pu être contactée. Plus l'âge augmente, plus il est facile d'avoir un décroché au sein du ménage mais le contact avec la personne échantillonnée est plus faible, du fait peut-être de comportements de barrage des proches plus fréquents. Il est également moins fréquent d'avoir une interview quand on a la personne échantillonnée au bout du fil. Il y a sans doute relativement plus de refus exprimés parmi cette population, peut-être en lien avec la durée de l'enquête qui est longue ou le grand âge. Le fait d'habiter en QPV n'a pas d'impact sur la probabilité de décrocher et aussi en vague 2 sur la probabilité de joindre la personne échantillonnée. En revanche, en vague 1, on observe un peu moins d'interviews une fois la personne au bout du fil. La probabilité d'avoir une interview est plus élevée parmi les plus aisés. On voit que c'est lié à une plus grande probabilité d'avoir une interview lorsque la personne échantillonnée est au bout du fil. Pour le dernier décile, cela semble lié à une fréquence relativement plus élevée d'avoir un décroché. La probabilité d'avoir une interview est plus faible parmi les personnes seules, c'est lié à un taux de « décroché » plus faible.

En ce qui concerne l'effet des moments d'appels, on constate des écarts importants entre les deux vagues de collecte mettant en évidence que les terrains téléphone de la vague 1 et 2 ont été très différents. En vague 1 comme en vague 2, on observe une probabilité de « décroché » beaucoup plus importante le matin qu'aux autres moments de la journée, quel que soit le jour de la semaine. En fait, il y a eu très peu d'appels en semaine le matin (9 % du total des appels ont été fait en semaine le matin). Le marché avec le prestataire prévoyait des appels entre 13h00 et 20h30/21h en semaine (et entre 10h et 15h le samedi). Donc, les matins en semaine ont surtout été consacrés à des rendez-vous pris avec les enquêtés, c'est pour cette raison qu'il y a peu d'appels et relativement plus de « décrochés » et de contacts avec la personne échantillonnée. En vague 2, l'effet des rendez-vous sur les interviews est celui attendu : la probabilité d'avoir une interview lorsque l'on a la personne échantillonnée au bout du fil est plus importante. En revanche, en vague 1, c'est très différent. La probabilité d'avoir une interview lorsque l'on a la personne au bout du fil est plus faible. On peut imaginer que cet effet est dû aux grandes difficultés rencontrées par le prestataire à honorer les rendez-vous en vague 1. Il est possible que les personnes à enquêter aient décroché (il y a un effet de sélection : ce sont des personnes avec lesquelles un rendez-vous a déjà été convenu) mais aient refusé l'interview agacées de ne pas avoir été appelées au moment convenu pour le rendez-vous. Au total, l'effet des matins sur la probabilité d'avoir une interview est positif mais il est plutôt plus faible en vague 1 qu'en vague 2 sans doute du fait des rendez-vous non honorés.

Si l'on s'intéresse maintenant à la rentabilité relative des différents types de numéros de téléphone, on constate que les numéros de téléphone mobile Fidéli permettent d'avoir plus fréquemment un « décroché » que les numéros de téléphone fixe Fidéli. En revanche, une fois que l'on a la personne échantillonnée au bout du fil, ils permettent d'obtenir moins fréquemment une interview qu'un fixe Fidéli. C'est un résultat intéressant. Cela signifie que les personnes sont peut-être moins souvent disponibles pour une longue interview lorsqu'on les joint sur leur mobile plutôt que sur un téléphone fixe. Au total, la probabilité d'avoir une interview quand on dispose d'un numéro de mobile de personne échantillonnée est plus importante mais c'est surtout en lien avec des « décrochés » plus fréquents. Un numéro de mobile du conjoint permet d'obtenir relativement plus de « décroché » mais pas plus (ni moins) de contact avec la personne échantillonnée que quand il y a eu un décroché sur un fixe Fidéli. La rentabilité en matière d'interview du mobile du conjoint est plutôt faible. La probabilité d'avoir une interview sur le numéro de mobile du conjoint est beaucoup plus faible que sur un numéro de fixe de Fidéli même quand on a la personne échantillonnée au bout du fil. Les numéros de fixe annuaire ont une moins bonne rentabilité que les numéros de fixe Fidéli à tout point de vue (« décroché », contact avec la personne échantillonnée et interview). Les « autres numéros » ont une excellente rentabilité pour les interviews : en effet, ce sont des numéros récupérés en cours de collecte via le site internet de l'enquête ou les enquêteurs pour réaliser les interviews. À nouveau, on constate que cette rentabilité est meilleure en vague 2 qu'en vague 1, sans doute en lien avec les rendez-vous qui n'ont pas été honorés en vague 1. Un travail récent mené sur l'historique des appels de la vague 1 de l'enquête EpiCov aboutit à des résultats similaires quant à la rentabilité des différents types de numéros de téléphone, et cela dans un contexte un peu différent car marqué par le premier confinement de l'épidémie de Covid 19 en France (Charrance et al., 2022).

Les enquêtés qui avaient refusé l'enquête ont été basculés vers un serveur refus où des enquêteurs expérimentés devaient les convaincre de participer à l'enquête. On voit, dans les données, la difficulté à convaincre ces personnes. Si la probabilité d'avoir un décroché est plus importante (il y a un effet de sélection : ce sont des personnes qui ont déjà décroché comme elles ont exprimé un refus), la probabilité d'avoir un contact avec la personne échantillonnée en vague 1 et la probabilité d'avoir une interview quand on a la personne échantillonnée au bout du fil en vague 2 restent inférieures à celle des personnes qui ne sont pas sur le serveur refus. Mais on a pu tout de même récupérer par ce biais 323 interviews (18 % des personnes passées sur le serveur refus). Il y a, en vague 1, un petit effet positif sur les interviews du serveur refus.

Enfin, quand on estime le même modèle sur les données des deux vagues en ajoutant une variable « vague de collecte », on constate que les effets en point de pourcentage associés au fait d'être en vague 2 plutôt qu'en vague 1 sont les suivants : « décroché » : -1,2*** points (probabilité pour la situation de référence : 5,4 %) ; « contact avec la personne échantillonnée sachant « décroché » » : +5,2*** points (probabilité pour la situation de référence : 68,2 %) ; « interview sachant contact avec la personne échantillonnée » : +10,2*** points (probabilité pour la situation de référence : 25,5 %) ; « interview » : +0,5*** points (probabilité pour la situation de référence : 0,8 %). On a montré précédemment que le taux de réponse avait été plus faible en vague 2 qu'en vague 1. En fait, cela ne semble pas lié à une moins bonne rentabilité des appels comme les coefficients relatifs aux interviews sont positifs. Au contraire, une fois la personne au bout du fil, il semble qu'il ait été plus facile d'avoir une interview en vague 2 qu'en vague 1. Cela peut être un effet « enquêteurs ». Nous n'avons en effet pas eu les mêmes enquêteurs en vague 2 qu'en vague 1 (sous-traitant différent). Le taux de réponse plus faible de la vague 2 semble alors tenir dans un nombre plus faible d'appels (187 150 contre 211 475 en vague 1) pour un taille d'échantillon mis en collecte téléphone un peu plus élevée en vague 2 qu'en vague 1 (9 943 contre 9 411), un plus faible taux d'appel sur les portables de la personne échantillonnée (35 % des appels contre 39 % des appels en vague 1) dont on a vu qu'il était rentable, notamment en termes de « décroché », et une plus faible disposition de numéros récupérés en cours de collecte (10 161 autres numéros en vague 1 contre 3 575 en vague 2) sans que l'on sache trop pourquoi. Il pourrait également être lié au fait qu'il y a plus de plis non distribués en vague 2, certainement du fait de l'ancienneté plus grande de l'échantillon, et que cette absence de lettre-avis ait réduit le nombre de « décrochés ». Par ailleurs, en vague 1, le numéro de téléphone affiché par le prestataire était régionalisé en fonction du lieu d'habitation de la personne échantillonnée, ce qui n'était pas le cas en vague 2. Cela a pu jouer sur le taux de « décrochés » à la faveur de la vague 1.

Par la suite, les régressions du tableau 15 ont été estimées en enlevant de l'historique d'appels tous les appels qui semblaient correspondre à des rendez-vous. Pour ce faire, un appel a été supprimé s'il correspondait à un rendez-vous pris sur le portail Internet de l'enquête ou s'il suivait directement un appel dont le code résultat correspondait à une prise de rendez-vous. En effet, dans ce cas, normalement la personne cessait d'être appelée jusqu'au moment du rendez-vous pris. On a ainsi supprimé 14 222 appels de l'historique (comprenant 398 625 appels). L'objet de ce test partait de l'idée que si tous les rendez-vous étaient pris sur un même type de téléphone, cela pouvait biaiser l'estimation de la rentabilité des types de numéros car un rendez-vous a une probabilité plus élevée d'aboutir à une interview. Cependant, les résultats des régressions restent quasiment inchangés. Les interprétations sont les mêmes. Il est noté que la rentabilité des matins reste meilleure même une fois les rendez-vous supprimés. Il n'est toutefois pas certains que tous les rendez-vous aient pu être repérés. On sait que des données concernant les rendez-vous pris sur le portail internet ont été perdues pendant la vague 1. Par ailleurs, des rendez-vous n'ont pas été honorés conduisant l'appel après un code résultat « rendez-vous » à finalement ne pas être un rendez-vous.

Tableau 15 • Régressions logistiques à partir de l'historique des appels (présentation des effets marginaux moyen)

	Décroché v1	Contact pers. échant./ décroché v1	Interview/contact pers. échant. v1	Interview v1	Décroché v2	Contact pers. échant./ décroché v2	Interview/contact pers. échant. v2	Interview v2
Constante	5,5 %	67,1 %	23,1 %	0,8 %	5,8 %	63,8 %	42,8 %	1,5 %
Femmes	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Hommes	0,2	-2,9***	-2,4***	-0,2**	0,2	-5,4***	-2,6**	-0,2***
Sexe inconnu	1,6*	0,3	-0,8	0,2	0,7	-3,3	2,2	0,1
[15-24]	-1,3**	-0,6	0,0	-0,2	-0,8	-2,8	0,7	-0,4*
[25-34]	-1,3***	-0,1	0,8	-0,1	-1,2**	1,7	-4,4**	-0,4***
[35-44]	-0,6	2,2	-0,2	0,0	-1,3***	2,6*	-2,0	-0,3**
[45-54]	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
[55-64]	0,7	0,1	2,2	0,3*	0,6	-1,2	-0,2	0,1
[65-74]	3,2***	-2,9*	2,3	0,7***	2,7***	-1,5	-3,1*	0,4**
[75-84]	6,0***	-4,5**	-2,2	0,6**	7,6***	-4,0*	-9,7***	0,7***
85+	10,1***	-14,1***	-9,3***	-0,1	6,0***	-15,8***	-20,6***	-1,0***
QPV	0,3	-2,1*	-3,4**	-0,3*	-0,7	0,0	-1,7	-0,3**
Hors QPV	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
d1 : les plus modestes	0,3	-4,7**	-2,5	-0,2	-1,8***	-10,5***	-1,0	-0,6***
d2	0,0	-2,4	-1,0	-0,1	-0,9	-4,9**	-0,8	-0,4*
d3	0,7	-4,2**	-1,4	0,0	-1,0	-4,3*	-0,2	-0,3*
d4	0,3	-2,8	0,9	0,1	0,1	-2,1	1,0	0,1
d5	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.

d6	0,0	-1,7	2,2	0,2	0,3	-2,1	2,9	0,2
d7	0,4	0,7	-0,1	0,1	0,6	0,5	2,7	0,3
d8	0,8	1,9	4,8***	0,6***	0,7	2,5	5,0**	0,5***
d9	1,0	-0,5	3,7**	0,5***	1,5**	1,6	4,5**	0,7***
d10 : les plus aisés	1,6**	-0,5	1,1	0,5**	1,7**	1,3	0,6	0,4*
Personne seule	-2,3***	3,5***	2,1*	-0,1	-2,2***	6,9***	-0,3	-0,3*
Couple sans enfant	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Couple avec enfant	0,6	-2,5*	-1,0	0,1	0,5	-4,0***	-5,6***	-0,3**
Fam. monoparentale	-1,3**	-1,3	-0,8	-0,2	-1,4**	0,5	-5,5***	-0,4**
Autre situation familiale	-0,4	-3,2*	-1,0	-0,1	-1,8***	-4,7**	-6,8***	-0,8***
Pas de déménagement	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Déménagement (même commune)	1,0	3,6	-2,0	0,0	-1,0	-4,3	-7,4**	-0,7***
Déménagement (commune différente)	-0,7	1,8	-1,2	-0,1	-0,8	-1,5	-2,9	-0,4*
Lundi matin	15,5***	15,4***	-8,3***	1,4***	7,4***	8,6***	12,6***	3,4***
Lundi après-midi	0,6*	1,3	6,0**	0,4***	0,0	1,9	2,4	0,2
Lundi soir	3,6***	2,6	3,4*	0,8***	3,9***	4,5*	-1,2	0,8***
Mardi matin	15,4***	21,9***	-9,8***	1,2***	3,1***	8,7***	13,3***	2,1***
Mardi après-midi	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Mardi soir	2,7***	1,5	3,1*	0,6***	1,1**	0,7	-3,0	0,0
Mercredi matin	16,1***	19,8***	-7,5***	1,8***	2,0***	7,9**	13,9***	1,8***
Mercredi après-midi	-0,4	-4,2*	3,7*	0,1	0,6	1,7	0,8	0,2
Mercredi soir	0,7**	1,7	1,1	0,2*	0,7	2,3	-3,3	0,0
Jeudi matin	18,4***	13,3***	-9,6***	1,3***	1,8***	8,7***	10,5***	1,5***
Jeudi après-midi	1,1***	3,8*	0,3	0,3**	-2,3***	-3,6	0,3	-0,6***
Jeudi soir	1,5***	2,7	4,0*	0,5***	-1,0**	-0,7	-6,5**	-0,6***
Vendredi matin	19,7***	15,2***	-7,0***	2,2***	0,2	11,1***	8,2**	1,0***

Vendredi après-midi	12,4***	9,5***	2,9	2,5***	-1,9***	6,0*	2,5	-0,1
Vendredi soir	8,2***	5,9***	-0,3	1,3***	2,0***	1,5	-5,2*	0,1
Samedi matin	6,7***	4,5**	5,1**	1,6***	2,6***	3,1	-1,5	0,5**
Samedi après-midi	8,0***	-4,2	9,4*	2,0***	5,7***	9,6*	17,7***	3,5***
Fidéli N° fixe	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Fidéli N° mobile pers. échant.	1,4***	17,2***	-1,5	0,5***	5,0***	17,4***	-3,6**	1,4***
Fidéli : N° mobile autre pers.	2,3***	-1,6	-8,2***	-0,3***	7,5***	1,6	-15,3***	0,3**
Fidéli : N° 09	-0,6	0,5	-1,4	-0,2	-0,6	1,9	-1,0	-0,1
Annuaire N° fixe	-1,5***	-6,3***	-7,1***	-0,7***	-1,3***	-5,1***	-8,5***	-0,7***
Annuaire N° mobile	-0,7	6,9	-9,4**	-0,6*	-1,5	1,2	-8,2	-0,5*
Autre numéro : collecte	2,1**	19,4***	11,7***	1,9***	18,9***	30,4***	17,5***	10,7***
Serveur refus	4,8***	-5,4***	-1,0	0,4**	4,9***	7,9***	-11,3***	-0,1
N'a pas exprimé de refus	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.

Note > Matin : avant 14h ; Après-midi : entre 14h et 17h ; Soir : après 17h. *** : p-value<0,001 ; ** : p-value<0,01 ; * : p-value<0,05

Lecture > En vague 1, un appel aboutit dans 5,5 % des cas à un décroché pour la situation de référence. Appeler sur un mobile Fidéli de la personne échantillonnée par rapport à un fixe Fidéli augmente cette probabilité de 1,4 points.

Champ > France métropolitaine.

Source > DREES-Irdes, EHIS 2019 (France métropolitaine).

Déroulement du terrain face-à-face : relativement similaire selon la vague d'enquête

Le terrain face-à-face s'est déroulé en deux vagues comme le terrain téléphone. La vague 1 a duré du 30 avril au 1^{er} août 2019 et la vague 2 du 27 août 2019 au 1^{er} février 2020. En cours de collecte, le terrain face-à-face intégrait les fiches adresse des IAJ du téléphone. Le terrain a été trop court en vague 1 pour que cela soit possible, ainsi les IAJ téléphone de la vague 1 ont été basculés en vague 2 en face-à-face¹⁴.

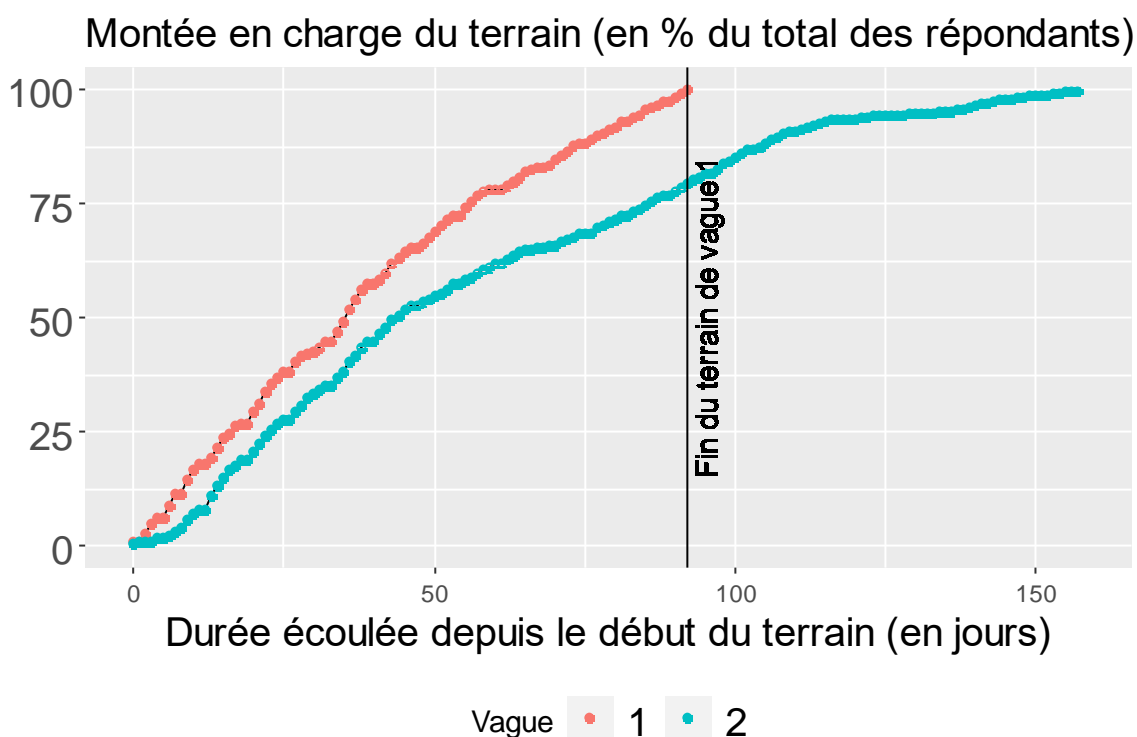
Au total, 2 201 questionnaire ont été réalisés en vague 1 et 3 900 en vague 2¹⁵. Le taux de réponse est à nouveau un peu meilleur en vague 1 qu'en vague 2 (65 % versus 62 %). En vague 2, l'échantillon était plus ancien. Il y a eu plus de plis non distribués, signal d'un nombre plus important de déménagements. Il est possible que cela ait joué à la baisse sur le taux de réponse.

Description des interviews

La montée en charge des réponses (graphique 7) a une dynamique similaire dans les deux vagues mais elle est plus rapide en vague 1, sans doute en lien avec une durée de terrain anticipé plus courte.

La répartition des interviews par jour de la semaine est assez similaire selon la vague (graphique 8). Il y a toutefois un peu plus d'interviews le samedi en vague 2 qu'en vague 1 (16 % versus 11 %).

Graphique 7 • Montée en charge du terrain : même dynamique en vague 1 et 2, quoique plus rapide en vague 1



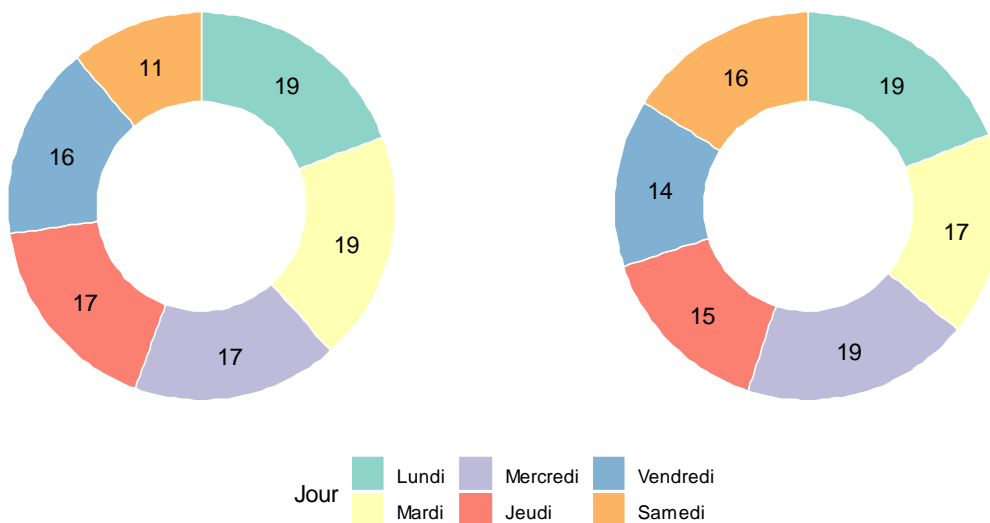
¹⁴ Pour certaines fiches adresse l'exploitation face-à-face a commencé en vague 1 mais s'est poursuivie en vague 2.

¹⁵ On considère dans ces calculs la vague définitive de collecte. C'est-à-dire les IAJ du téléphone en vague 1 collectés en face-à-face lors de la vague 2 sont considérés comme des répondants face-à-face de la vague 2. La même règle est appliquée pour le calcul des taux de réponse.

Graphique 8 • Jours des interviews : un peu plus d'interviews le samedi en vague 2

CAPI : % / jour de la semaine (vague 1)

CAPI : % / jour de la semaine (vague 2)

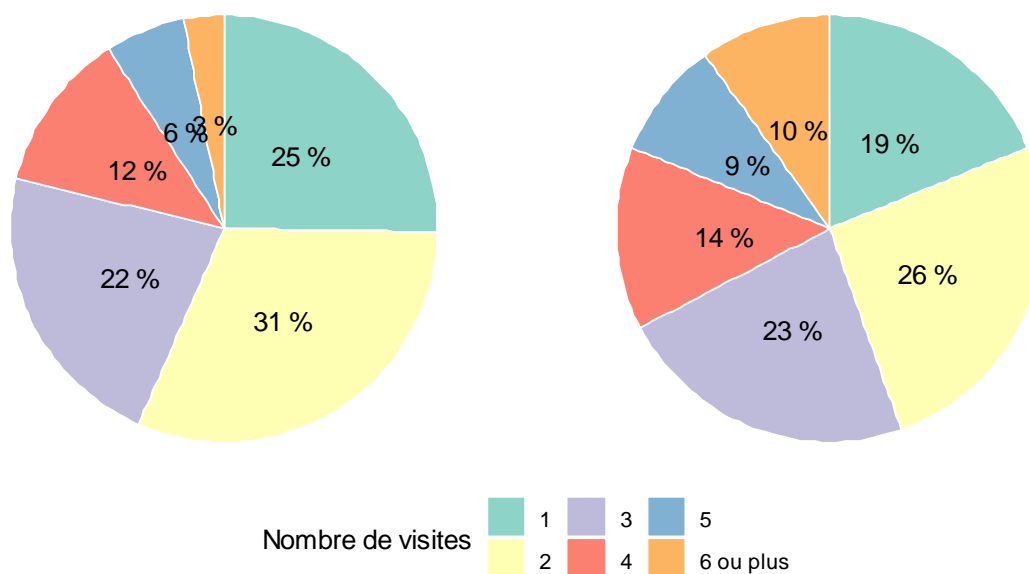


Les enquêteurs devaient noter leurs tentatives de contacts avec les personnes à enquêter. Le protocole prévoyait cinq visites maximum au domicile de la personne à enquêter. Les enquêteurs ont parfois été au-delà. On constate que même 10 % des interviews en vague 2 ont été réalisées au-delà de 6 visites (graphique 9), ce qui laisse penser une certaine rentabilité des visites au-delà de cinq. La différence à ce propos entre la vague 1 et la vague 2 s'explique en partie par les personnes IAJ du téléphone de vague 1 basculées en vague 2 en face-à-face. Ces adresses ont commencé à être exploitées l'été en fin de vague 1 et l'ont été aussi en vague 2 si bien que le nombre de visites de cinq maximum a souvent été dépassé.

Graphique 9 • Nombre de visites par interview : plus d'interviews après plus de 5 visites en vague 2

Nombre de visites pour avoir une interview (Vague 1)

Nombre de visites pour avoir une interview (Vague 2)



Bilan de l'utilisation de la tablette pour les questions les plus sensibles

Quand l'enquête se déroulait en face-à-face, la tablette était proposée à la personne enquêtée en fin de questionnaire pour qu'elle puisse répondre par elle-même aux questions jugées les plus sensibles de l'enquête (PHQ9, consommation de tabac et d'alcool). Le but était d'assurer la meilleure comparabilité des réponses avec l'enquête DROM où ce module était passé sous casque.

Le « bilan enquêteur » fait état d'un protocole qui a bien fonctionné : « Les questions auto-administrées se sont très bien déroulées. Seules deux personnes âgées ont demandé que je saisisse la réponse. ».

Le questionnaire comportait une question pour savoir si le module « sensible » avait été renseigné directement par l'enquêté avec la tablette. Cela permet de calculer la part des enquêtés qui n'ont pas utilisé par eux-mêmes la tablette selon les caractéristiques sociales et l'état de santé (tableau 16). Au total 4 706 personnes ont répondu par elles-mêmes avec la tablette, donc, 77 % des enquêtés face-à-face. On constate que la part de ceux qui n'ont pas utilisé la tablette est constante selon le sexe, et augmente, comme attendu, avec l'âge pour atteindre 73 % des 85 ans ou plus. Cette part est plus forte parmi les personnes des ménages modestes que celles des ménages aisés et plus forte dans les QPV. On remarque également que cette part est très liée au sujet de l'enquête. Moins bonne est la santé, plus la part des enquêtés qui ne remplissent pas par eux-mêmes le questionnaire avec la tablette est importante. Cela peut altérer l'intérêt du dispositif. L'hypothèse était que la tablette favoriserait les réponses sincères à l'enquête mais on voit que le biais éventuel de déclaration lié au mode de remplissage du module est lié au sujet d'intérêt de l'enquête. Les réponses seraient, suivant cette hypothèse, moins sincères pour ceux qui sont en plus mauvaise santé. La partie « effet de mode » étudie l'effet de l'interrogation en auto-administré avec tablette par rapport au téléphone sur la réponse à l'enquête.

En ce qui concerne la non-réponse partielle, elle est un peu plus élevée parmi les enquêtés qui ont rempli avec la tablette par rapport à ceux qui ont répondu avec l'enquêteur (tableau 17).

Tableau 16 • Part des enquêtés qui n'ont pas renseigné le module « sensible » avec la tablette

	Part qui n'a pas utilisé la tablette parmi les répondants face-à-face	Effectif total de répondants face-à-face
Femmes	23 %	3212
Hommes	23 %	2889
15-17 ans	8 %	296
18-24 ans	7 %	644
25-34 ans	11 %	662
35-44 ans	13 %	873
45-54 ans	17 %	1040
55-64 ans	27 %	1079
65-74 ans	35 %	889
75-84 ans	54 %	427
85 ans ou plus	73 %	191
d1 : les 10 % les plus modestes	32 %	829
d5 : Médiane	24 %	589
d10 : les 10 % les plus aisés	12 %	498

QPV	34 %	890
Hors QPV	21 %	5211
HS1 : Très bon	11 %	1674
HS1 : Très mauvais	58 %	121
HS2 : Maladie chronique	32 %	2425
HS2 : Pas de maladie chronique	17 %	3676
HS3 : Fortement limité	47 %	743
HS3 : Pas limité	17 %	4389
Proxy	41 %	172

Note > Statistiques non pondérées.

Lecture > Il y a eu un effectif total de 3 212 femmes qui ont répondu à l'enquête. Parmi elles, 23 % ont répondu au module sensible sans la tablette.

Champ > France métropolitaine.

Source > DREES-Irdes, EHS 2019 (France métropolitaine).

Tableau 17 • Répartition de la population selon le nombre de questions non renseignées parmi les questions non filtrées du module « sensible »

	0	1	2	3	4	Plus de 4
Enquêteur	96,7 %	1,5 %	0,8 %	0,2 %	0,1 %	0,6 %
Tablette	92,5 %	3,2 %	1,7 %	0,9 %	0,4 %	1,3 %

Note > Cet indicateur est construit à partir des 14 questions non filtrées du module sur tablette : les 9 questions du PHQ9, 3 questions sur le tabac et 2 sur l'alcool. Statistiques non pondérées.

Lecture > 1,3 % des personnes ayant renseigné le module sensible avec la tablette ont laissé plus de quatre questions sans réponse. C'est le cas de 0,6 % des personnes qui ont renseigné ce module avec enquêteur.

Champ > France métropolitaine.

Source > DREES-Irdes, EHS 2019 (France métropolitaine).

Correction de la non-réponse partielle

Le nombre de non-réponses partielles par question est présenté en annexe 2. Les questions pour lesquelles la non-réponse partielle dépasse 1,0 % des répondants concernent la première question de littératie en santé (LITER1) portant sur la capacité à remplir des formulaires médicaux, la date de la dernière mesure par un professionnel de santé du taux de cholestérol et de la glycémie (PA3 et PA4), la taille et le poids¹⁶ (BM1 et BM2), la durée des trajets à pied (PE3), deux questions de l'échelle d'Oslo : le fait que l'entourage montre ou non de l'intérêt pour ce que l'on fait (SS2) et la difficulté avec laquelle on peut obtenir de l'aide des voisins (SS3) ainsi que trois questions du PHQ9 (avoir peu d'intérêt ou de plaisir à faire les choses, avoir une mauvaise opinion de soi-même, bouger ou parler lentement ou être au contraire très agité).

La correction de la non-réponse partielle concernant les variables santé a été effectuée de la même manière pour l'ensemble des variables. Elle a été imputée via un *hotdeck* séquentiel à l'aide du package VIM de R. Les classes considérées pour mener le *hotdeck* ont été définies en fonction du sexe et les observations ont été triées selon l'âge en tranches quinquennales puis le niveau de vie en continu. Toutes les variables manquantes pour une même personne ont été imputées à l'aide d'un même donneur pour respecter les corrélations entre les variables.

Des indicatrices indiquant qu'une imputation a été menée figurent dans les données si bien que les imputations peuvent être repérées et refaites pour les besoins d'études particulières.

¹⁶ En fait, pour ces deux variables, on a considéré que lorsque l'une des deux était manquante, on imputait à la fois la taille et le poids. Sur les 165 valeurs imputées pour chacun de ces deux variables, il manquait le poids et la taille dans 19 cas, la taille uniquement dans 31 cas et le poids uniquement dans 115 cas.

Étude des effets de mode : exploitation de l'échantillon méthodologique

Contexte et design expérimental

Cette partie s'intéresse à l'analyse des effets de mesure entre les différents modes de collecte. Pour cela, elle se base sur l'exploitation de l'échantillon méthodologique de l'enquête.

Pour rappel, cet échantillon a été tiré parmi les personnes possédant dans le fichier un numéro de téléphone (après enrichissement annuel et hors suréchantillon QPV) et a été interrogé en face-à-face. La majeure partie du questionnaire a été posée en face-à-face et les questions les plus sensibles ont été posées de façon auto-administrée : l'enquêteur a prêté sa tablette à l'enquêté. Cet échantillon permet ainsi d'étudier les effets de mesure entre le téléphone et le face-à-face pour les questions posées en face-à-face et entre le téléphone et l'auto-administration via la tablette pour les questions les plus sensibles.

Quand on étudie les effets de mode, l'enjeu est toujours d'essayer de bien distinguer les effets de sélection des effets de mesure. Les effets de sélection recouvrent le fait que ce sont souvent des populations différentes qui répondent aux différents modes de collecte du fait de la spécificité de chacun des modes. Les effets de mesure qualifient le fait qu'une même personne peut répondre différemment à un même questionnaire selon le mode sur lequel elle est interrogée.

Le design du protocole rend les effets de sélection les plus réduits possibles. Cependant ceux-ci peuvent exister du fait de la non-réponse. À ce titre, les taux de réponse sont assez différents entre le téléphone (45 %, cf. partie « bilan de collecte ») et l'échantillon méthodologique (68 %). La comparaison des répondants par téléphone hors suréchantillon QPV et hors proxys (7 643 personnes), des répondants en face-à-face (1 265 personnes) et des répondants via tablette (1 037 personnes) met en évidence des profils socio-démographiques toutefois très proches entre les répondants (tableau 18). Il y a néanmoins plus de personnes âgées dans l'échantillon qui a répondu en face-à-face (échantillon méthodologique) par rapport à l'échantillon téléphone. On constate que cet effet se réduit pour la tablette puisque moins de personnes âgées ont répondu via la tablette (cf. tableau 16). Il semble donc y avoir moins d'écart entre l'échantillon téléphone et l'échantillon tablette. Les méthodes d'estimation des effets de mesure proposées ci-dessous ont pour objectif de rapprocher au maximum les caractéristiques de ces échantillons pour contrôler les effets de sélection. Une attention est portée à bien rapprocher ces échantillons selon la répartition par âge et le handicap.

Tableau 18 • Caractéristiques sociodémographiques des répondants selon les modes

Mode de passation	Téléphone	Face-à-face	Tablette
Âge moyen	49,4	51,4	48,5
<25 ans	13 %	14 %	16 %
80 ans ou +	3 %	7 %	3 %
Femmes	53 %	51 %	51 %
Niveau de vie médian	1 986	1 886	1 969
Vit en QPV	5 %	5 %	4 %
Né à l'étranger	10 %	11 %	9 %
Aucun diplôme	13 %	16 %	12 %
Bac+3 ou plus	28 %	22 %	25 %
En emploi	54 %	47 %	52 %
Chômage	5 %	5 %	6 %
Vit seul	19 %	17 %	15 %

État de santé général : bon ou très bon	74 %	72 %	75 %
État de santé général : mauvais ou très mauvais	6 %	8 %	6 %
Maladie chronique	38 %	39 %	37 %
GALI : Handicapé	6 %	10 %	7 %

Note > Statistiques non pondérées.

Lecture > Parmi l'échantillon téléphone (sans proxy et sans suréchantillon QPV), il y a 6 % de personnes en situation de handicap ; parmi l'échantillon méthodologique face-à-face il y en a 10 % et il y en a 7 % parmi ceux qui parmi l'échantillon méthodologique ont utilisé la tablette pour répondre au module sensible.

Champ > France métropolitaine.

Source > DREES-Irdes, EHIS 2019 (France métropolitaine).

Méthodes d'analyse

Deux méthodes ont été testées pour estimer les effets de mesure. Ces méthodes ont été appliquées à un très grand nombre d'indicateurs issus de l'enquête (tableaux 19, 20 et 21) sur l'échantillon composé des répondants par téléphone et des répondants issus de l'échantillon méthodologique.

Première méthode : appariement sur score de propension

La première méthode utilise un appariement sur score de propension pour détecter les effets de mesure. C'est une méthode très répandue pour ce type d'analyses (Rosenzweig, 2022, Dabet, 2022) qui provient du champ de l'évaluation des politiques publiques. Les codes permettant d'implémenter cette méthode sont repris de document de travail Insee de Simon Quantin sur l'estimation du score de propension en R (Quantin, 2018).

On considère deux populations : les « traités », c'est à dire les répondants en face-à-face et les « non traités » encore appelés « échantillon de contrôle », c'est-à-dire ceux qui ont répondu par téléphone. La méthode consiste à estimer l'effet causal du traitement, c'est-à-dire l'effet de répondre en face-à-face plutôt que par téléphone. Les résultats présentent l'effet moyen du traitement sur les traités (ATT : *average treatment effect on the treated*) et l'effet moyen du traitement (ATE : *average treatment effect*), tous deux corrigés du biais estimés selon la méthode de Abadie et Imbens (2011). Dans notre cas, la population des « traités » et la population de l'échantillon total (échantillon téléphone et échantillon méthodologique) représentent la même population, à savoir ceux qui ont un numéro de téléphone (issu de Fidéli ou de l'enrichissement annuel). Donc il est normal que l'ATT et l'ATE soient très proches.

On estime pour chaque personne de l'échantillon considéré un score de propension : il s'agit de la probabilité d'être traité, donc d'être interrogé en face-à-face estimée par une modélisation logit. Les variables explicatives doivent être corrélées au sujet de l'enquête et expliquer le mode de passation. Elles ne doivent pas être influencées par des effets de mesure pour ne pas fausser l'estimation finale. Il est donc usuel de prendre des variables de la base de sondage ou des variables très factuelles de l'enquête. On teste les variables suivantes : sexe (enquête), niveau d'étude (enquête), catégorie socio-professionnelle (enquête), indicatrice d'habiter en QPV (base de sondage), indicatrice d'avoir effectué une mobilité résidentielle en 2017 (base de sondage), tranches d'âge (enquête), décile de niveau de vie (base de sondage), indicatrice de pauvreté monétaire (base de sondage), situation familiale (base de sondage), tranche de taille d'unité urbaine (collecte), situation vis-à-vis de l'emploi (enquête), le fait d'être né à l'étranger (enquête). On retient toutes les variables qui sont significatives au seuil de 10 % et on ajoute le sexe car il est très lié au sujet de la santé même si la p-value vaut 0,16. Cela conduit à retenir toutes les variables sauf la situation vis-à-vis de l'emploi, l'indicatrice de pauvreté monétaire, l'indicatrice d'habiter dans un QPV et le fait d'être né à l'étranger.

On apparie à toute personne de l'échantillon face-à-face un jumeau (en termes de score de propension) issu de l'échantillon téléphone. La différence de prévalence de la variable d'intérêt entre ces deux groupes forme l'ATT. L'effet est estimé sur un échantillon de la taille du groupe traité, l'échantillon méthodologique face-à-face. On peut également appairier à toute personne du téléphone un jumeau issu de l'échantillon face-à-face (en dupliquant forcément des observations du face-à-face comme l'échantillon est plus petit). La différence forme alors l'ATE. L'effet est alors estimé sur un échantillon de la taille du groupe de contrôle, l'échantillon téléphone. On calcule l'ATT (résultats des colonnes ATT1, ATT2 et ATT3) et l'ATE (résultats de la colonne ATE1). On peut exiger que l'appariement soit exact sur certaines variables mais cela dégrade la qualité de l'appariement sur le score puisque les scores jumeaux ne sont cherchés que dans des sous-groupes formés par les variables pour lesquelles on veut un appariement exact. On teste deux variantes : d'une part, un appariement sur score de propension et exact sur le sexe (résultats de la colonne ATT3) et d'autre part, un appariement sur score de propension

et exact sur le sexe et les tranches d'âge¹⁷ (résultats des colonnes ATT1, ATT2 et ATE1) puisque le sexe et l'âge sont des variables importantes en matière de santé.

En calculant l'ATT issu de cette méthode sur le mini-module européen sur la santé, on constate un effet positif d'être interrogé en face-à-face sur les variables du GALI (plus de personnes en situation de handicap en face-à-face) et d'état de santé général (plus de personnes qui se disent en très bon état de santé et moins qui se disent en bon état de santé) (résultats de la colonne ATT1). Il nous semble que cela s'apparente plus à des effets de sélection mal contrôlés par le score de propension qu'à des effets de mesure. Par exemple, on constate dans les statistiques descriptives (tableau 18) une part plus importante de personnes en situation de handicap interrogées en face-à-face par rapport au téléphone. On choisit donc d'ajouter ces variables dans le calcul du score de propension (résultats des colonnes ATT2, ATT3, ATE1). La colonne ATT1, pour comparer, présente les résultats sans ces deux variables dans le calcul du score de propension.

La distribution du score de propension avant et après appariement montre un support commun important en matière de score de propension pour mener l'appariement. Par ailleurs, on voit que même avant appariement les groupes de traités (face-à-face) et non traités (téléphone) sont très proches sur les variables qui composent le score de propension (annexe 3).

Deuxième méthode : régression

La seconde méthode testée est une régression sur l'ensemble de l'échantillon (téléphone et échantillon méthodologique). L'effet d'être interrogé en face-à-face plutôt que par téléphone est alors le coefficient de l'indicatrice du face-à-face. On teste deux méthodes d'estimation : un modèle logit (résultats de la colonne « Logit ») et un modèle à probabilité linéaire (les résultats ne sont pas présentés mais sont tout à fait similaires aux autres méthodes). Les variables de contrôle introduites dans le modèle sont les variables du modèle d'estimation du score de propension dans sa version définitive (donc avec les deux variables santé : GALI et état de santé général).

Résultats et interprétations

Les raisons des effets de mesure peuvent être classifiées en plusieurs catégories :

- La *désirabilité sociale* : l'enquêté peut avoir tendance à vouloir donner une image de lui en conformité avec des normes sociales valorisantes. On s'attend à ce que cet effet soit d'autant plus fort que l'enquêté est à proximité de l'enquêteur. L'effet serait donc plus fort en face-à-face que par téléphone et plus fort par téléphone qu'en auto-administré.
- La *gêne de parler de sujets intimes* : certaines questions en lien avec la santé peuvent créer une gêne chez l'enquêté, ce qui peut les conduire à ne pas répondre la vérité en présence d'un enquêteur. Au contraire, en auto-administré, on s'attend à ce que les réponses soient plus conformes à la réalité. Cet effet peut se rapprocher de la désirabilité sociale.
- Le *satisficing* : l'enquêté peut se contenter d'une réponse qu'il juge satisfaisante sans chercher à donner la réponse exacte. Il est attendu que cet effet soit le plus fort lorsque l'enquêté remplit le questionnaire en auto-administré sans personne pour le motiver à bien répondre. Il peut être fort éventuellement par téléphone en lien avec les conditions de passation où l'enquêteur va très vite et où il y a moins de place qu'en face-à-face pour un échange avec l'enquêteur afin de bien comprendre les questions. Une conséquence du *satisficing* peut être de choisir des modalités extrêmes sans réfléchir à une réponse plus nuancée qui aurait pu être plus juste.
- Les *conditions de passation* : cela rejoint le point précédent. En face-à-face, les enquêteurs disposent du livret des cartes codes dans lequel sont écrites les modalités des questions complexes. Cela peut aider les enquêtés à répondre à ces questions complexes. Par ailleurs, en face-à-face, des proches peuvent écouter l'échange entre l'enquêté et l'enquêteur, ce qui peut avoir un effet sur les réponses données (sans doute plutôt en faveur d'une réponse plus honnête de manière générale sauf cas particuliers : violences dans le ménage ou autre). Par téléphone, l'échange est rapide et il peut y avoir plus de difficultés de compréhension du questionnaire qu'en face-à-face. En face-à-face, les enquêteurs prennent sans doute plus de temps pour expliquer les questions ou aider les enquêtés à se remémorer des éléments utiles aux réponses.

¹⁷ Les tranches d'âge considérées sont : 15-19 ; 20-29 ; 30-39, 40-49 ; 50-59 ; 60-69 ; 70-79 ; 80 ans ou plus.

Effets de mesure liés au fait d'être interrogé en face-à-face plutôt que par téléphone

Les résultats sont présentés dans les tableaux 19 (santé perçue), 20 (recours aux soins) et 21 (habitudes de vie en lien avec la santé). Les coefficients significatifs sont indiqués par des étoiles. Chaque coefficient s'interprète alors comme l'augmentation ou la baisse de la prévalence¹⁸ en points de pourcentage induite par une interrogation en face-à-face plutôt que par téléphone. La validité de cette estimation porte sur la population ayant un numéro de téléphone (77 % de l'échantillon total, partie « protocole »). La prévalence non pondérée estimée sur l'échantillon interrogé par téléphone est présentée en dernière colonne pour pouvoir mettre en regard le niveau des effets de mesure d'être interrogé en face-à-face plutôt que par téléphone et le niveau de la prévalence par téléphone.

De manière générale, on constate un nombre important de coefficients significatifs traduisant la présence d'effets de mesure.

Au compte de la désirabilité sociale, il nous semble y avoir l'effet sur les difficultés auditives et sur les difficultés à se souvenir que l'on déclare environ un quart de moins en face-à-face que par téléphone (tableau 19). La note de satisfaction vis-à-vis de la vie (échelle Cantril) est également plus élevée en face-à-face que par téléphone (environ +0,3 sur la note moyenne de 6,9 par téléphone tableau 19). La note relative au goût pour le risque semble aussi un peu plus élevée en face-à-face que par téléphone (environ +0,2 sur la note moyenne de 3,7 par téléphone tableau 19). Il est possible que le fait d'aimer prendre des risques corresponde à un trait de caractère valorisé dans certains milieux. Il apparaît également que les enquêtés déclarent plus fréquemment que les autres montrent beaucoup d'intérêt pour ce qu'ils font en face-à-face que par téléphone et également qu'il est très facile d'avoir de l'aide de leurs voisins (tableau 21). Les questions sur la littératie en santé (tableau 20) porte sur des notions de compréhension qui peuvent sembler basiques et pour lesquelles il peut être difficile d'avouer que l'on a des difficultés (« comprendre ce que le professionnel de santé vous demande de faire », « lire et comprendre des informations écrites sur la santé », « lire et comprendre toutes les instructions sur la prise de médicaments ») si bien qu'on observe des effets de mesure très importants sur ces questions. Les enquêtés répondent beaucoup plus fréquemment que c'est « très facile » en face-à-face que par téléphone (+ 20 points environ sur une prévalence par téléphone de 40 % environ). Les vases communicants entre modalités se font principalement avec la modalité juste en dessous (« généralement facile ») mais aussi avec la modalité encore en dessous (« parfois difficile ») qui est environ moitié moins dite en face-à-face que par téléphone.

Certains effets peuvent s'interpréter comme de la gêne à parler de sujets intimes. C'est le cas de la moins forte prévalence d'incontinence urinaire observée en face-à-face par rapport au téléphone (tableau 19) : l'effet porte sur plus d'un tiers de la prévalence observée par téléphone. Il semble y avoir un petit effet sur le fait de déclarer aller voir un psychologue ou un psychiatre mais les effectifs sont très faibles. On n'observe pas d'effet sur le fait de déclarer une dépression. Concernant les tests de dépistage (recherche de sang dans les selles, coloscopie, mammographie, frottis, tableau 20), les effectifs sont aussi très faibles si bien qu'il est difficile d'avoir un effet significatif au seuil de 5 %. Cela semble toutefois être le cas pour les frottis moins déclarés en face-à-face que par téléphone. C'est aussi le cas pour la coloscopie au seuil de 10 %. Cela peut être lié à la gêne d'aborder ces sujets.

Un bon nombre d'effets sont sans doute liés aux différences de conditions de passation. Par exemple, l'effet très significatif du fait d'être interrogé en face-à-face plutôt que par téléphone concernant la prise de médicaments non prescrits (+ 8 points environ pour une prévalence de 28 % par téléphone) pourrait tenir au fait que les enquêteurs en face-à-face prennent plus le temps de citer des médicaments non prescrits (le paracétamol par exemple) et qu'ainsi les enquêtés se rappellent en avoir pris dans les deux dernières semaines. Les questions liées au renoncement aux soins (ou aux besoins de soins non satisfaits) est un autre très bon exemple. Ces questions étaient très longues et complexes. Le bilan des enquêteurs l'a bien montré (partie « questionnaire »). La première modalité citée était « Vous n'êtes pas concerné : vous n'avez pas eu besoin de soins ou d'examen médicaux » et les deux autres modalités étaient « Oui, ça vous est arrivé : vous n'avez pas pu payer les soins dont vous aviez besoin » et « Non, vous avez pu payer les soins dont vous aviez besoin ». On constate (tableau 20) qu'il y a beaucoup moins de personnes qui déclarent la première modalité « non concerné » en face-à-face et à l'inverse, plus de personnes qui déclarent ne pas avoir renoncé. Il est fort probable que les questions aient été mieux comprises en face-à-face, du fait notamment de l'usage du livret des cartes codes. Par téléphone, il est possible que des personnes aient répondu la première modalité pour aller vite assimilant le « non concerné » au fait de ne pas avoir renoncé. L'effet est très fort, il peut aller jusqu'à baisser de moitié le taux de « non concernés » ce qui joue de façon importante sur le taux de renoncement aux soins puisque ceux-ci deviennent des personnes qui n'ont pas renoncées et font ainsi baisser le taux de besoins de soins non satisfaits. Cela met en évidence le fait que la formulation de la question ne convient pas même si beaucoup de travail avait été effectué dessus. Les résultats des estimations mettent également en avant une part de la population plus élevée qui déclare ne pas faire de trajet à pied d'au moins 10 minutes d'affilée dans la semaine en face-à-face que par téléphone (tableau 21). Le résultat peut paraître étonnant, il va à l'encontre de la désirabilité sociale. En fait, quand on regarde les données plus en détail, on constate qu'il y a des vases communicants entre 0 jours par semaine et 7 jours par semaine. Une interprétation peut être que la partie de la question qui porte sur « pendant au moins 10 minutes d'affilée » est interprétée plus littéralement en face-à-face que par téléphone du fait des conditions de passation (plus d'interactions avec l'enquêteur, plus de temps pour

¹⁸ Pour deux variables uniquement, il s'agit d'une moyenne et non d'une prévalence car ce sont des variables continues : note entre 0 et 10 relative à la satisfaction par rapport à la vie (échelle Cantril) et note entre 0 et 10 relative à l'aversion au risque.

réfléchir). Ainsi par téléphone, les gens surestimeraient le nombre de trajets par semaine parce qu'ils feraient moins attention à la partie de la question consacrée à la durée d'un trajet (au moins 10 minutes d'affilée). Les enquêtés déclarent plus fréquemment en face-à-face manger des fruits ou des légumes tous les jours et quand ils le déclarent, ils précisent plus fréquemment manger plus d'une portion par jour (tableau 21). Cet effet peut à nouveau être interprété comme un effet du livret des cartes codes. Un gros effort avait été fait sur l'illustration de ces questions avec des dessins de portions de fruits et de légumes (voir la rubrique « documents de collecte » de la page Internet consacrée à EHIS sur le site de la DREES). Ainsi en voyant ces dessins, les enquêtés se sont très certainement rappelés plus de fruits et de légumes et les ont déclarés plus fréquemment et en quantité plus nombreuse. L'effet est très fort : par exemple +14 points sur la part de la population qui mange des légumes tous les jours (estimée à 60 % par téléphone).

Il reste des questions sans effet de mesure. Ce sont pour beaucoup des questions très factuelles (liste de maladies, hospitalisation, recours au généraliste, au spécialiste). Il y a aussi certains effets que l'on explique mal : prévalence plus élevée d'allergies et d'asthme en face-à-face, du fait de ne pas faire de sport, du fait de consommer de boissons sucrées tous les jours. Pour la première, on a émis l'hypothèse que cela pourrait être lié à la période d'enquête comme l'échantillon méthodologique a été enquêté pendant une période plus resserrée que l'échantillon du téléphone. Cependant, il semble que les deux échantillons ont eu une part équivalente de questionnaires renseignés pendant les périodes où les allergies sont répandues (printemps, début de l'automne). Sur les deux derniers aspects, cela peut être lié au fait que les réponses seraient plus sincères en face-à-face qu'au téléphone parce que l'enquêteur est au domicile, dans l'environnement de l'enquêté ou parce que des proches de l'enquêté assistent au questionnaire.

Tableau 19 • Effet d'être interrogé en face-à-face plutôt que par téléphone (santé perçue)

Variables et modalités	ATT 1	ATT 2	ATT 3	ATE 1	Logit	Prévalence (téléphone)
État de santé général : Très mauvais	-0,3	-0,4	0,0	0,0	X	1 %
État de santé général : Mauvais	1,8	-0,2	1,2	0,3	X	5 %
État de santé général : Assez bon	-0,4	0,9	-1,6	1,6	X	20 %
État de santé général : Bon	3,0	-0,8	-0,2	-0,1	X	41 %
État de santé général : Très bon	-4,0*	0,5	0,6	-1,7	X	34 %
Maladie chronique : oui	1,1	0,1	-1,6	-0,4	-1,4	38 %
GALI : Handicapé	3,0**	0,9	1,4	0,8	X	6 %
GALI : Limité pas fortement	-2,4	-1	-0,7	-0,9	X	16 %
GALI : Pas limité	-0,7	0,2	-0,7	0,2	X	78 %
Dents : Très mauvais	0,5	1,7*	1,4*	1,5*	0,8*	2 %
Dents : Mauvais	0,4	0,5	0,9	-0,2	0,7	8 %
Dents : Assez bon	1,1	-1,7	-1,2	-0,5	-1,2	24 %
Dents : Bon	1,9	2,0	0,0	3,0	1,3	42 %
Dents : Très bon	-3,9*	-2,5	-1,1	-3,9*	-2,6*	25 %
Douleurs : Nulles	-1,0	0,7	1,5	1,6	1,2	41 %

Douleurs : Très faibles	-1,2	-0,5	-0,5	-2,4	-1,2	18 %
Douleurs : Faibles	-0,6	-2,4	-2,8*	-1,1	-2,0	16 %
Douleurs : Modérées	-0,1	-0,3	-0,8	0,1	-0,2	17 %
Douleurs : Fortes	2,6*	2,3*	2,4*	1,4	1,5*	6 %
Douleurs : Très fortes	0,3	0,1	0,2	0,4	0,1	2 %
Absence travail du fait de la santé	0,5	-1,6	-4,0	0,3	-1,1	32 %
Asthme : oui	-2,4*	-2,1*	-2,5**	-2,5**	-2,5**	8 %
Hypertension : oui	1,0	2,5	0,4	1,9	0,4	16 %
Mal au dos : oui	-2,6	-1,0	-2,1	-0,6	-2,0	34 %
Arthrose : oui	-1,8	-3,2*	-2,8	-3,1*	-2,5*	22 %
Mal à la nuque : oui	-0,4	1,3	0,8	-0,1	0,1	21 %
Diabète : oui	-0,7	0,5	0,4	0,0	-0,4	6 %
Allergies : oui	-3,3*	-5,5***	-5,3***	-5,3***	-5,7***	26 %
Incontinence urinaire : oui	-1,5	-4,6***	-3,3***	-3,5***	-3,5***	8 %
Dépression : oui	0,2	-1,7	-0,9	-1,7	-0,7	7 %
Cholestérol : oui	-3,8**	-2,6	-2,3	-1,3	-2,2*	14 %
BPCO : oui	-0,5	-1,2	-0,3	-1,4	-1,3	6 %
Crise cardiaque : oui	0,0	-0,5	-0,5	-0,1	-0,3	1 %
Maladie des artères coronaires : oui	0,1	-0,6	-0,3	-0,2	-0,6	2 %
AVC : oui	-0,1	-0,9*	-1,0*	-0,2	-0,4	1 %
Cirrhose du foie : oui	-0,1	-0,1	-0,2	-0,1	-0,2	0 %
Maladie des reins : oui	2,5***	1,5	1,4	1,5*	1,0*	3 %
Accident circulation	0,4	0,4	0,6	0,3	0,5	1 %
Accident au domicile	0,4	0,2	0,7	0,3	0,6	4 %
Accident dans les loisirs	2,1*	0,8	1,5*	1,4	1,2*	5 %
Lunettes : oui	-0,1	0,9	0,1	0,2	0,3	74 %
Vue : difficultés	0,6	1,0	-0,6	-1,4	0,0	18 %
Appareil auditif : oui	1,1	0,2	0,3	0,7	0,3	3 %
Audition pièce silencieuse : difficultés	-1,9	-4,0***	-3,0**	-3,1**	-3,3***	9 %

Audition pièce bruyante : difficultés	-5,3**	-6,8***	-7,0***	-7,4***	-7,2***	31 %
Marche : difficultés	2,4*	0,9	2,6**	1,4	0,7	6 %
Escalier : difficultés	2,7*	0,7	2,0	1,4	0,8	7 %
Souvenir : difficultés	-5,6***	-6,0***	-6,2***	-5,9***	-6,5***	24 %
Croquer : difficultés	-0,6	1,7	-0,6	-0,4	-0,8	12 %
CANTRIL : note moyenne	0,3***	0,4***	0,3***	0,4***	0,4***	6,9
GOUT DU RISQUE : note moyenne	0,2	0,3***	0,2*	0,2*	0,2**	3,7

Note > ATT1 : les variables santé (état de santé perçue et GALI) ne sont pas introduites dans le modèle de calcul du score de propension et l'appariement se fait sur score de propension ainsi que de manière exacte sur sexe et tranches d'âge. **ATT2** : les variables santé sont introduites dans le modèle de calcul du score de propension et l'appariement se fait sur score de propension et de manière exacte sur sexe et tranches d'âge. **ATT3** : les variables santé sont introduites dans le modèle de calcul du score de propension et l'appariement se fait sur score de propension et de manière exacte sur le sexe. **ATE1** : les variables santé sont introduites dans le modèle de calcul du score de propension et l'appariement se fait sur score de propension et de manière exacte sur sexe et tranches d'âge. **Logit** : résultats de la régression : effets marginaux moyens du logit et coefficients de la régression linéaire pour les deux variables continues (note de satisfaction par rapport à la vie : échelle Cantril et note de goût pour le risque). **Prévalence** : prévalence sur l'échantillon (interrogé par téléphone) non pondérée.

« difficultés » s'entend comme avoir « quelques difficultés », « beaucoup de difficultés » ou « ne pas pouvoir du tout » pour la variable considérée.

*** : p-value < 0,001 ; ** : p-value < 0,01 ; * : p-value < 0,05

Lecture > Lorsque l'on interroge les personnes en face-à-face, la prévalence de déclarer des difficultés à se souvenir est en moyenne inférieure de 5,6 à 6,5 points (selon les estimations) par rapport à lorsque l'on les interroge par téléphone. Pour comparer, cette prévalence a un niveau de 24 % dans l'échantillon interrogé par téléphone (non pondéré). L'effet est significatif au seuil de 0,1 %.

Champ > France métropolitaine.

Source > EHIS 2019, France métropolitaine, DREES-Irdes.

Tableau 20 • Effet d'être interrogé en face-à-face plutôt que par téléphone (recours aux soins)

Variables et modalités	ATT 1	ATT 2	ATT 3	ATE 1	Logit	Prévalence (téléphone)
Littératie 1 : très facile	19,0***	23,1***	21,2***	20,9***	21,3***	46 %
Littératie 2 : très facile	21,3***	22,6***	23,4***	21,8***	23,5***	49 %
Littératie 3 : très facile	24,3***	25,8***	26,4***	25,6***	25,2***	41 %
Littératie 4 : très facile	23,8***	23,2***	22,7***	23,6***	23,7***	51 %
Littératie 5 : très facile	23,0***	24,2***	23,4***	23,4***	24,6***	53 %
Littératie 1 : généralement facile	-18,7***	-20,6***	-19,7***	-18,4***	-20,4***	40 %
Littératie 2 : généralement facile	-17,4***	-17,9***	-18,7***	-18,1***	-19,6***	40 %
Littératie 3 : généralement facile	-18,7***	-19,0***	-20,0***	-19,4***	-20,3***	41 %
Littératie 4 : généralement facile	-21,4***	-19,5***	-19,2***	-19,5***	-21,6***	39 %
Littératie 5 : généralement facile	-22,3***	-20,7***	-21,0***	-20,8***	-22,8***	41 %
Littératie 1 : parfois difficile	-1,4	-3,2**	-2,6*	-2,6*	-3,2**	10 %
Littératie 2 : parfois difficile	-3,3***	-3,8***	-4,1***	-2,9**	-4,3***	8 %

Littératie 3 : parfois difficile	-5,7***	-6,4***	-6,0***	-5,5***	-7,2***	14 %
Littératie 4 : parfois difficile	-2,7**	-4,0***	-4,3***	-4,3***	-4,4***	8 %
Littératie 5 : parfois difficile	-0,8	-3,0***	-2,1**	-2,3**	-2,1**	5 %
Littératie 1 : généralement difficile	0,5	0,9	0,3	-0,5	0	2 %
Littératie 2 : généralement difficile	-0,8	-1,2*	-0,9	-0,8	-0,9*	2 %
Littératie 3 : généralement difficile	-1,2*	-1,8**	-1,5**	-1,6**	-1,5**	3 %
Littératie 4 : généralement difficile	-0,6	-0,5	0,1	-0,5	-0,1	2 %
Littératie 5 : généralement difficile	0	-0,5	-0,3	-0,3	-0,4	1 %
Littératie 1 : toujours difficile	0,7	-0,1	0,8	0,5	0,4	1 %
Littératie 2 : toujours difficile	0,2	0,3	0,4	-0,1	0,0	1 %
Littératie 3 : toujours difficile	1,3**	1,3**	1,2*	0,9*	0,6**	1 %
Littératie 4 : toujours difficile	1,0*	0,8	0,7	0,6	0,3	1 %
Littératie 5 : toujours difficile	0,1	0,0	0,1	0,1	-0,1	0 %
Hospitalisation de nuit : oui	0,1	-0,7	-1,7	0,2	-0,5	10 %
Hospitalisation de jour : oui	-0,1	-2,3	-1,4	-1,4	-0,8	18 %
Dentiste (12 derniers mois)	1,3	-0,2	-0,2	1,5	0,6	61 %
Généraliste (12 derniers mois)	0,7	-0,7	0,0	0,0	0,5	85 %
Spécialiste (12 derniers mois)	4,6*	1,8	1,6	1,2	2,1	49 %
Kiné (12 derniers mois)	0,8	-1,3	-0,9	-1,3	-0,7	36 %
Psy (12 derniers mois)	0,0	-2,7**	-0,8	-2,1*	-1,1	8 %
Médicaments prescrits : oui	2,6	2,3	2,8	3,4*	2,7*	48 %
Médicaments non prescrits : oui	8,8***	8,3***	8,2***	8,6***	8,2***	28 %
Vaccin grippe : oui	-1,1	-2,2	-1,8	-0,8	-1,0	21 %
Tension (12 derniers mois)	3,8*	2,4	3,2*	2,8	4,2**	78 %
Test cholestérol (12 derniers mois)	0,7	1,0	-0,1	-0,1	0,6	51 %
Test glycémie (12 derniers mois)	-1,8	-1,3	-2,1	-1,9	-1,2	54 %
Test de sang dans les selles (vie)	-1,7	-2,5	-2,6	-2,0	-1,8	32 %
Coloscopie (vie)	-1,4	-2,9	-2,6	0,1	-2,0	26 %
Mammographie (vie)	0,1	-2,0	-4,1	-0,4	-0,7	66 %

Frottis (vie)	-4,2**	-2,9	-4,0*	-1,6	-2,2*	86 %
Renoncement aux soins...						
... du fait du délais pour un rdv : non concerné	-7,5***	-4,4**	-5,0***	-4,6**	-6,7***	20 %
... du fait du délais pour un rdv : oui	-3,9**	-5,1***	-5,6***	-4,5**	-5,3***	23 %
... du fait du délais pour un rdv : non	11,4***	9,5***	10,5***	9,1***	11,4***	58 %
... du fait des transports : non concerné	-11,2***	-7,7***	-8,9***	-9,3***	-11,9***	23 %
... du fait des transports : oui	1,0	0,0	0,6	0,4	0,4	3 %
... du fait des transports : non	10,3***	7,7***	8,3***	8,9***	10,8***	74 %
Besoins de soins non satisfaits (raisons financières)...						
... consultations : non concerné	-11,4***	-7,4***	-8,7***	-10,0***	-11,7***	23 %
... consultations : oui	-0,9	-0,8	-0,5	-0,9	-1,0	4 %
... consultations : non	12,2***	8,1***	9,3***	10,9***	12,3***	72 %
... soins dentaires : non concerné	-13,1***	-12,9***	-12,7***	-13,4***	-14,4***	36 %
... soins dentaires : oui	-0,4	0,9	0,9	0,2	0,3	7 %
... soins dentaires : non	13,5***	12,0***	11,8***	13,2***	13,5***	57 %
... médicaments : non concerné	-12,7***	-11,0***	-10,4***	-11,2***	-13,2***	24 %
... médicaments : oui	-0,1	0,4	0,1	-0,2	0,1	1 %
... médicaments : non	12,8***	10,6***	10,3***	11,4***	12,6***	74 %
... soins en santé mentale : non concerné	-12,5***	-10,2***	-11,0***	-11,6***	-9,8***	83 %
... soins en santé mentale : oui	-0,3	-0,5	-0,4	-0,8	-0,3	2 %
... soins en santé mentale : non	12,7***	10,7***	11,5***	12,4***	9,9***	15 %
... lunettes : non concerné	-11,2***	-12,5***	-11,1***	-10,8***	-12,2***	47 %
... lunettes : oui	0,3	0,8	0,2	0,4	0,3	3 %
... lunettes : non	10,9***	11,7***	11,0***	10,4***	11,7***	50 %
... prothèses auditives : non concerné	-10,0***	-10,6***	-9,7***	-9,5***	-7,7***	90 %
... prothèses auditives : oui	0,0	0,3	0,4	0,0	0,2	1 %
... prothèses auditives : non	10,0***	10,2***	9,3***	9,4***	7,4***	9 %

Note > Même définition des colonnes que dans le tableau 19.

Littératie 1, 2, 3, 4, 5 renvoie au nom des questions portant sur la littératie en santé dans le questionnaire : LITER1, 2, 3, 4, 5.

*** : p-value<0,001 ; ** : p-value<0,01 ; * : p-value<0,05

Lecture > Lorsque l'on interroge les personnes en face-à-face, la prévalence de se déclarer « non concerné » par le besoin de soins lié aux prothèses auditives est en moyenne inférieure de 7,7 à 10,0 points (selon les estimations) par rapport à lorsque l'on les interroge par téléphone. Pour comparer, cette prévalence a un niveau de 90 % dans l'échantillon interrogé par téléphone (non pondéré). L'effet est significatif au seuil de 0,1 %.

Champ > France métropolitaine.

Source > EHIS 2019, France métropolitaine, DREES-Irdes.

Tableau 21 • Effet d'être interrogé en face-à-face plutôt que par téléphone (habitudes de vie en lien avec la santé)

Variables et modalités	ATT 1	ATT 2	ATT 3	ATE 1	Logit	Prévalence (téléphone)
Pas de trajet à pied dans la semaine (10 min.)	6,4***	5,9***	7,0***	6,0***	4,9***	14 %
Pas de vélo dans la semaine (10 min.)	-0,2	0,7	-1,1	-0,4	-0,4	84 %
Pas de sport dans la semaine (10 min.)	9,6***	6,4***	8,1***	8,6***	7,9***	46 %
Pas de musculation dans la semaine	5,5***	5,5***	5,2***	5,7***	5,9***	72 %
Fruits tous les jours	4,8**	5,6**	5,1**	5,5**	5,1***	59 %
Jamais de fruits	0,1	-0,2	0,0	0,5	0,2	2 %
Fruits : consommation intermédiaire	-4,9**	-5,4**	-5,0**	-6,0***	-5,3***	38 %
Une portion de fruits par jour (vs pls portions)	-9,0***	-5,9**	-6,2**	-4,8*	-7,0***	30 %
Légumes tous les jours	14,3***	13,8***	13,9***	12,8***	14,5***	60 %
Jamais de légumes	0,6	0,6	0,5	0,4	0,4	1 %
Légumes : consommation intermédiaire	-14,8***	-14,4***	-14,4***	-13,2***	-15,2***	40 %
Une portion de légumes par jour (vs pls portions)	-6,2**	-8,5***	-7,2***	-7,8***	-8,5***	27 %
Jus de fruits tous les jours	-0,9	0,6	0,5	1,4	1,2	24 %
Jamais de jus de fruits	4,3*	3,3	3,6*	3,4	2,7	34 %
Jus de fruits : consommation intermédiaire	-3,3	-3,9*	-4,0*	-4,8**	-4,1**	42 %
Boissons sucrées tous les jours	2,6*	2,8**	1,4	1,2	1,8*	7 %
Jamais de boissons sucrées	-0,1	2,2	2,7	2,5	2,1	50 %
Boissons sucrées : consommation intermédiaire	-2,5	-5,0**	-4,2*	-3,7*	-4,2**	43 %
Personnes sur qui on peut compter : aucune	-0,3	-0,7	-0,5	-0,4	-0,5	1 %
Intérêt des autres : aucun	-0,1	-1,5	-0,5	-1,3	-1,1	5 %
Intérêt des autres : peu	-0,7	-1,4	-1,3	-2,7*	-2,1	15 %
Intérêt des autres : un certain intérêt	-2,8	-2,1	-2,3	-3,5	-3,3*	43 %

Intérêt des autres : beaucoup d'intérêt	3,5	3,6	3,6*	6,3***	5,3***	43 %
Aide des voisins : très facile	6,6***	8,3***	7,6***	6,2***	6,6***	24 %
Aide des voisins : très difficile	1,9	2,5*	2,9**	2,7**	2,4***	6 %
Aide des voisins : intermédiaire	-8,5***	-10,7***	-10,5***	-8,9***	-9,3***	70 %
Soin à une autre personne : oui	-2,2	-3,5*	-2,9	0,3	-1,5	27 %

Note > Même définition des colonnes que dans le tableau 19.

*** : p-value<0,001 ; ** : p-value<0,01 ; * : p-value<0,05

Lecture > Lorsque l'on interroge les personnes en face-à-face, la prévalence de déclarer ne pas faire de trajet à pied d'au moins 10 minutes est en moyenne supérieure de 4,9 à 7,0 points (selon les estimations) par rapport à lorsque l'on les interroge par téléphone. Pour comparer, cette prévalence a un niveau de 14 % dans l'échantillon interrogé par téléphone (non pondéré). L'effet est significatif au seuil de 0,1 %.

Champ > France métropolitaine.

Source > EHS 2019, France métropolitaine, DREES-Irdes.

Effets de mesure liés au fait d'être interrogé de manière auto-administrée (tablette) plutôt que par téléphone

Le tableau 22 présente les résultats sous la même forme que les tableaux précédents. Les coefficients significatifs s'interprètent comme l'effet en points de pourcentage en plus ou en moins sur la prévalence du fait d'être interrogé en auto-administré via la tablette plutôt que par téléphone. On constate un effet de désirabilité sociale. Le fait de boire de l'alcool tous les jours est déclaré plus fréquemment en auto-administré (via la tablette) que par téléphone. L'effet est de l'ordre d'un quart de la prévalence. Les enquêtés déclarent également moitié plus souvent penser qu'il vaudrait mieux être mort en auto-administré plutôt que par téléphone. Les enquêtés déclarent en revanche moins fréquemment des problèmes de sommeil en auto-administré que par téléphone, c'est peut-être parce qu'il s'agit d'un sujet facile à aborder dans une discussion.

Tableau 22 • Effet d'être interrogé de manière auto-administrée (tablette) plutôt que par téléphone (module le plus sensible)

Variables et modalités	ATT 1	ATT 2	ATT 3	ATE 1	Logit	Prévalence (téléphone)
Fume tous les jours	-1,3	-0,1	0,3	1,2	0,8	16 %
Fume occasionnellement ou a fumé	0,4	-0,1	-1,0	-1,6	-0,1	38 %
Boit de l'alcool tous les jours	2,2	3,0*	2,6*	3,0*	2,4**	10 %
Boit de l'alcool toutes les semaines	1,6	2,2	0,6	3,5	1,7	38 %
Peu d'intérêt à faire les choses	-0,9	-0,7	-0,7	-1,0	-1,2	10 %
Être triste	-0,8	0,1	0,2	-0,1	-0,1	6 %
Difficultés à s'endormir ou dormir trop	-2,4	-4,9**	-2,3	-4,0**	-3,1*	19 %
Se sentir fatigué	-2,9	-1,2	-2,6	-2,1	-2,6*	17 %
Peu d'appétit ou manger trop	-2,8*	-1,0	-1,2	-1,2	-1,8	11 %
Mauvaise opinion de soi-même	-0,1	1,1	-0,2	0,5	0,0	6 %
Difficultés de concentration	-0,9	-0,2	-0,8	0,8	-0,4	5 %
Bouger lentement ou au contraire	-1,4*	-2,0**	-1,9**	-1,7**	-2,2**	3 %

Penser qu'il vaudrait mieux mourir	1,9*	1,7	2,3**	2,3**	1,8***	4 %
Présence de symptômes dépressif	-1,8	-1,9	-1,1	-1,8	-1,5	9 %
Présence de symptômes dépressif sévères	-1,0	0,0	-0,4	-0,9	-0,8	3 %

Note > Même définition des colonnes que dans le tableau 19.

*** : p-value<0,001 ; ** : p-value<0,01 ; * : p-value<0,05

Lecture > Lorsque l'on interroge les personnes de manière auto-administrée (avec la tablette), la prévalence de déclarer qu'on pense qu'il vaudrait mieux mourir est en moyenne supérieure de 1,7 à 2,3 points (selon les estimations) par rapport à lorsque l'on les interroge par téléphone. Pour comparer, cette prévalence a un niveau de 4 % dans l'échantillon interrogé par téléphone (non pondéré). Dans la plupart des modèles, l'effet est significatif au moins au seuil de 5 %.

Champ > France métropolitaine.

Source > EHIS 2019, France métropolitaine, DREES-Irdes.

Appariements

Système national des données de santé (SNDS)

L'appariement de l'enquête avec le SNDS a été autorisé par la CNIL en même temps que la collecte de l'enquête. Il porte sur une profondeur de 10 ans entre 2015 et 2024. Cette profondeur avait été choisie pour faire la jonction entre deux éditions de l'enquête EHIS. Comme il s'agissait d'un tirage d'individus, la méthode choisie a été une collecte indirecte des informations d'état civil dans la base de sondage (Fidéli) : prénom, nom, sexe, date de naissance, commune de naissance. Les NIR ont alors été reconstitués par interrogation de la BRPP (base du répertoire des personnes physiques) à l'Insee puis envoyés à la CNAM via la procédure Safe pour extraction des consommations de soins. Du point de vue du circuit de données, c'est la division Sondages de l'Insee qui a préparé les fichiers avec l'état civil des personnes échantillonnées, l'a transmis au service de la BRPP qui a transmis les NIR à la CNAM. Une table de passage a été créée par la division Sondages entre l'identifiant anonyme de l'enquête et l'identifiant anonyme transmis à la CNAM.

La transmission des NIR a porté sur l'ensemble de l'échantillon, donc à la fois les personnes répondantes et non répondantes. L'appariement avait été annoncé à l'ensemble des personnes de l'échantillon dans la lettre avis. Les personnes pouvaient s'opposer à l'appariement. Un total de 8 personnes se sont opposées à l'appariement de leurs données avec le SNDS. Elles ont donc été retirées du fichier transmis à la BRPP par la division Sondages. L'appariement des non-répondants avait pour but d'aider au travail de correction de la non-réponse totale à l'enquête en permettant d'enrichir les modèles de correction de la non-réponse de variables issues du SNDS, très corrélées au sujet de l'enquête et possiblement corrélées au comportement de réponse¹⁹.

Pour un certain nombre de personnes dans Fidéli, on ne dispose pas de l'état civil complet mais uniquement de la date de naissance. C'est le cas des personnes à charge au sens fiscal. De ce fait, la reconstitution du NIR n'a pas pu être menée (tableau 23). Cela concerne notamment les personnes de 16 ans ou moins à la date de validité de la base de sondage, le 1^{er} janvier 2018 (tableau 24).

Tableau 23 • Effectifs en vue de l'appariement avec le SNDS aux différentes étapes du circuit de données

Étapes	Nombre de personnes dans le fichier
Échantillon EHIS France métropolitaine	27 600
Envoyé par la division Sondages à la BRPP (la division Sondages a supprimé les personnes pour lesquelles on n'avait ni prénom, ni nom)	26 207
Nombre de NIR retrouvés via la BRPP (il restait dans le fichier des personnes avec un état civil incomplet)	24 535
Nombre de NIR arrivés à la CNAM	24 516
Nombre d'identifiants uniques dans le référentiel des bénéficiaires (table IR_BEN_R) du SNDS transmis par la CNAM à la DREES	24 035

Champ > France métropolitaine.

Source > EHIS 2019, France métropolitaine, DREES-Irdes.

L'appariement des répondants à l'enquête, utile pour les études, a été analysé plus finement. Il y a au total 12 552 personnes à la fois dans l'enquête et dans le SNDS, soit un taux d'appariement de 88 %.

¹⁹ À la date de publication de ce document, la modélisation de la non-réponse dans le calcul des poids mis à disposition n'inclut pas les variables du SNDS, faute de temps pour investiguer ce sujet.

Tableau 24 • Taux d'appariement des répondants par tranches d'âge

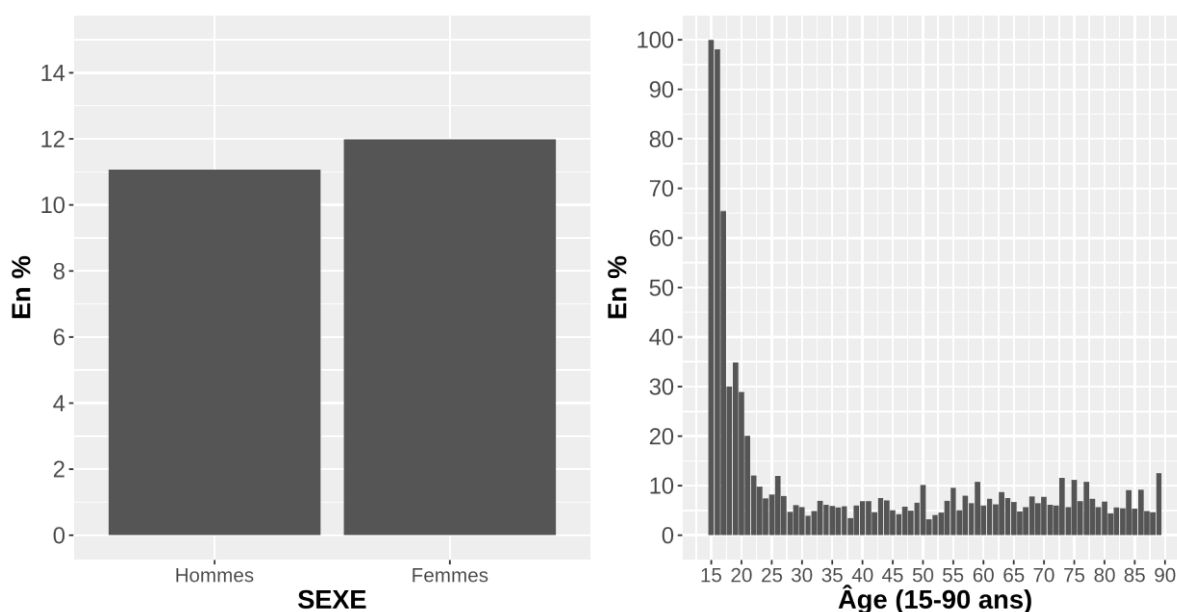
Tranche d'âge	15 ans ou plus	16 ans ou plus	17 ans ou plus	18 ans ou plus	19 ans ou plus	20 ans ou plus	30 ans ou plus
Taux d'appariement	88 %	90 %	91 %	92 %	93 %	93 %	93 %
Taux d'appariement de l'âge minimal de la tranche	0 %	2 %	35 %	70 %	65 %	71 %	94 %

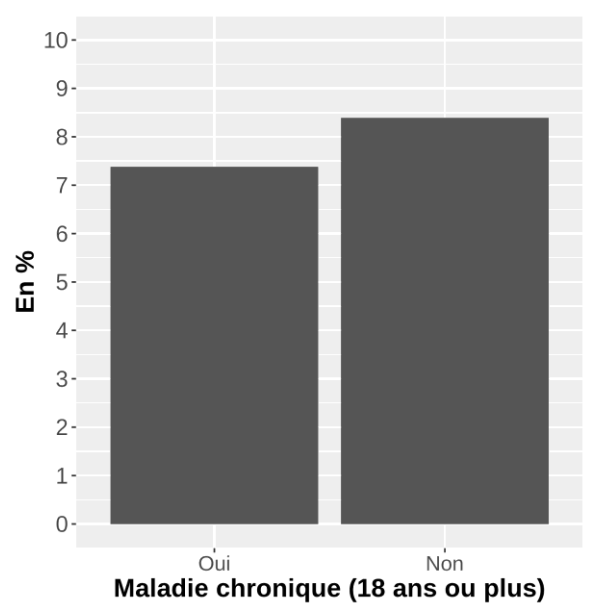
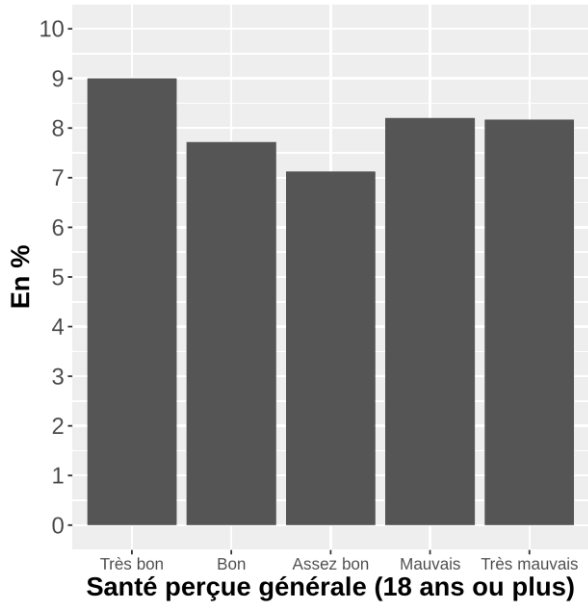
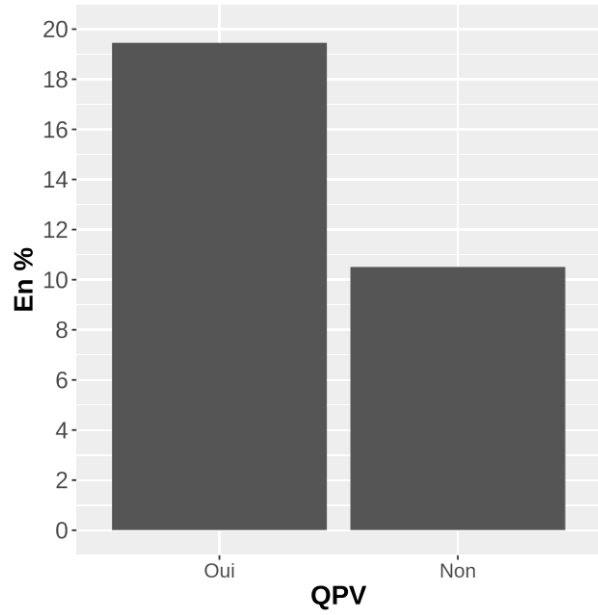
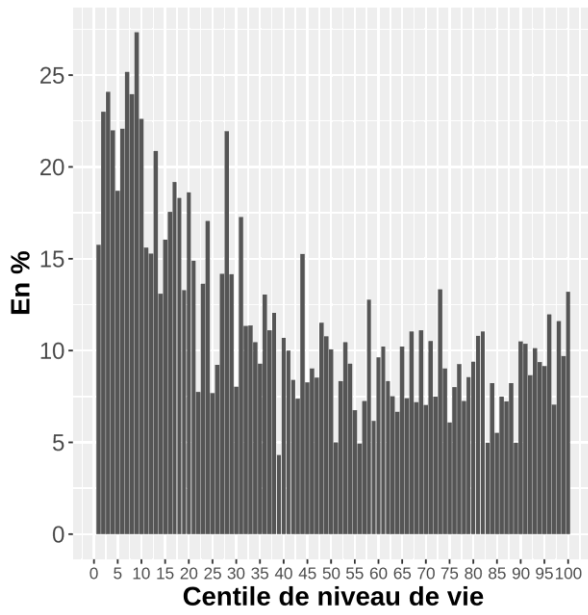
Champ > France métropolitaine.

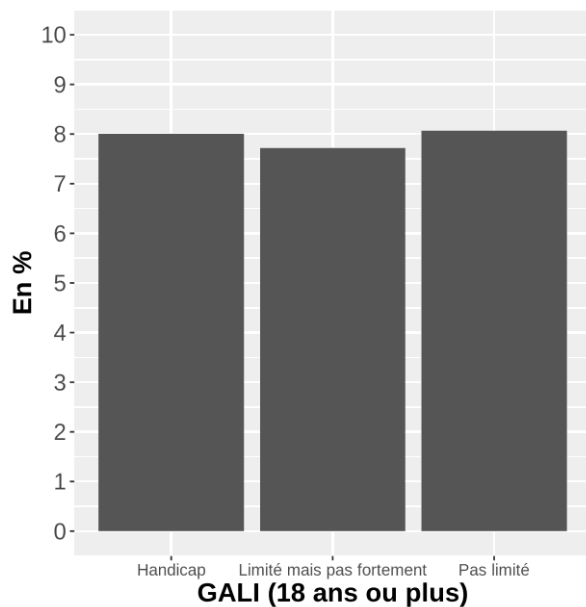
Source > EHIS 2019, France métropolitaine, DREES-Irdes.

Le taux d'appariement des répondants à l'enquête est identique selon le sexe : 88 % pour les femmes et 89 % pour les hommes. Pour vérifier l'appariement, on a comparé le sexe et l'année de naissance des personnes dans l'enquête et dans le SNDS. Parmi les 12 552 personnes appariées, il y en a 16 qui ont un sexe différent dans l'enquête et le SNDS et 59 une année de naissance différente.

La part de personnes non appariées avec le SNDS parmi les répondants selon le sexe, l'âge, le centile de niveau de vie, l'indicatrice du fait d'habiter dans un quartier prioritaire de la politique de la ville (QPV) et le mini module européen sur la santé est présenté dans le graphique 10. Hormis sur la question de l'âge mentionnée précédemment, il est difficile de distinguer ce qui tient au fait que le NIR n'a pas été retrouvé à partir de Fidéli ou qu'il n'a pas été retrouvé dans le SNDS car on ne dispose pas du détail des identifiants anonymes envoyés à la CNAM. La corrélation entre le non-appariement et le niveau de vie laisse plutôt penser que cela est dû à Fidéli. En tous cas, une fois les plus jeunes retirés, on n'observe pas de corrélation importante entre le fait de ne pas avoir été apparié et les réponses au mini module européen sur la santé.

Graphique 10 • Taux de non-appariés selon l'âge, le centile de niveau de vie et les réponses au mini module européen sur la santé





Fichiers fiscaux et sociaux

Il est prévu que l'enquête soit appariée avec les fichiers fiscaux et sociaux dans le courant de l'année 2022 pour obtenir des variables de revenus et transferts sociaux plus complètes et relatives à l'année 2019. En l'état, les informations de revenus et niveaux de vie sont issues de Fidéli et relatives à l'année 2017. La variable déclarative de revenu avait été supprimée du questionnaire car on allait récupérer des informations plus fiables à partir de la base de sondage et de l'appariement socio-fiscal. À l'Insee, dans les enquêtes où le niveau de vie est central, il est conseillé de laisser un niveau déclaratif de revenu ou une liste de questions relatives au fait de percevoir des revenus par types de source possibles (salaire, retraite, RSA, etc.) pour aider à l'imputation de ceux qui n'auront pas pu être appariés.

■ ENQUÊTE SANTÉ DOM

Calendrier

L'enquête santé DOM s'est déroulée en parallèle de l'enquête santé en France métropolitaine. Comme la maîtrise d'œuvre de cette enquête était réalisée par l'Insee, celle-ci a suivi le calendrier classique de la mise en place d'une nouvelle enquête à l'Insee, marqué par de nombreux tests. Le séminaire de lancement de l'enquête a eu lieu le 15 juin 2017 marquant le début de la préparation de l'enquête, deux ans avant la collecte. Cependant, cette préparation avait débuté doucement plusieurs mois auparavant en même temps que le travail portant sur l'enquête France métropolitaine. Par exemple, les réunions méthodologiques avec l'Insee étaient mutualisées sur les deux enquêtes. La comitologie suivait les règles de l'Insee : elle était ponctuée par des réunions des comités de suivi et de pilotage. Le terrain de l'enquête a eu lieu à peu près au même moment que celui de la seconde vague de l'enquête France métropolitaine. Il a eu lieu du 8 juillet au 31 décembre 2019. Il a ainsi été prolongé de quinze jours par rapport à la date de fin initialement prévue, du fait d'un début de collecte assez lent en lien avec les vacances scolaires. Le calendrier du projet est présenté dans le tableau 25.

Tableau 25 • Calendrier du projet dans les DROM

Date	Étape
7 novembre 2016	Réunion avec la division Sondages comportant des éléments de cadrage sur les enquêtes dans les DROM (champ, bases de sondage, méthode Kish)
13 septembre 2016	Réunion de mise en place de l'organisation autour de l'enquête entre l'Insee et la DREES
13 décembre 2016	Réunion avec les concepteurs de l'enquête « Adult education survey » à l'Insee sur le protocole de tirage de 2 individus « Kish » au sein d'un ménage
Début 2017	Décision de tirer 2 individus « Kish » par ménage
7 juin 2017	Réunion avec les directions régionales de l'Insee dans les DROM et les ARS des DROM sur l'adaptation du questionnaire
15 juin 2017	Séminaire de lancement Insee de l'enquête
29 juin 2017	Réunion avec l'équipe en charge du Baromètre Santé DOM de Santé publique France sur leur expérience de collecte d'enquêtes Santé dans les DROM et sur la coordination avec l'enquête Unono wa Maore à Mayotte
29 juin 2017	Réunion avec la division Sondages sur les modalités de tirage et la taille des échantillons
11 juillet 2017	Déclaration auprès de la CNIL des tests de l'enquête. La CNIL a accepté pour ces tests une simple déclaration.
25 juillet 2017	Réunion avec la division Sondages : évocation des bases de sondage pour les tirages, notamment de la comparaison Fidéli/recensement en Guyane. Discussion autour du calendrier de tirage
4 octobre 2017	Séance de la Commission services publics et services aux publics – Obtention de l'avis d'opportunité Validation de la demande d'accès aux fichiers fiscaux au titre de l'article 7bis pour le tirage de l'enquête
6-25 novembre 2017	Test papier à La Réunion et à Mayotte

10 novembre 2017	Signature de la convention DREES/Insee/GIP-Irdes
10 janvier 2018	1 ^{ère} réunion du Comité de suivi (enquête de Santé publique France à Mayotte, bilan du focus group, préparation du test CAPI, module sous casque, matériel pour peser et mesurer à Mayotte)
16 janvier 2018	Visioconférence spécifique à la préparation de l'enquête à Mayotte
13 février 2018	Réunion avec la division Sondages et le Criem. Description de la méthodologie de tirage pour les quatre DROM historiques et pour Mayotte. Discussion autour de la base de sondage pour la Guyane.
14 février 2018	1 ^{ère} réunion du Comité de pilotage (positionnement par rapport à l'enquête Unono wa Maore de Santé publique France, évolutions du questionnaire, échantillonnage, validation de l'utilisation du casque, calendrier)
21 février 2018	Courrier du Directeur de la DREES au Directeur de Santé publique France actant que Santé publique France n'a pas souhaité prolonger les discussions en vue d'une fusion de l'enquête Unono wa maore à Mayotte avec l'enquête santé DOM
14 mars 2018	Visioconférence spécifique à la préparation de l'enquête à Mayotte
29 mars 2018	2 ^e réunion du Comité de suivi (organisation du codage du questionnaire sous Blaise, partie sous casque, préparation du test CAPI)
Avril-août 2018	Codage du questionnaire en Blaise en vue du test CAPI
6 avril 2018	Présentation de l'enquête dans le cadre d'un séminaire des chefs de division enquêtes Ménages à l'Insee
12 avril 2018	Décision de la DREES de tirer l'échantillon de la Guyane dans Fidéli (et non dans le recensement)
18 mai 2018	Demande de devis en vue des traductions pour le casque en Guyane
25 mai-16 juin 2018	Traduction de la partie sous casque pour la passation en Guyane par la société Filogis (bushinenge, créole guyanais, portugais-brésilien)
12 juin 2018	3 ^e réunion du Comité de suivi (préparation du test CAPI, ajout d'un test CAPI à Mayotte en janvier 2019, préparation du dossier pour le Label)
28 juin 2018	2 ^e réunion du Comité de pilotage (validation de l'augmentation de la taille des échantillons du fait de l'effet de grappe induit par les 2 individus « Kish », demande de matériel pour le poids et la taille à Mayotte, validation de l'ajout d'un test CAPI à Mayotte pour tester le casque et le matériel spécifique : balances, télémètres)
12 juillet 2018	Signature de la convention de participation financière entre la DREES et l'ARS Martinique (financement à hauteur de 50 000 euros par l'ARS Martinique)
3 septembre-6 octobre 2018	Test CAPI à La Réunion
10 septembre 2018	Lancement des travaux sur le communication autour de l'enquête
14 septembre 2018	Réunion avec les équipes de la BRPP au sujet de la reconstitution des NIR avec Fidéli

17 septembre-13 octobre 2018	Test CAPI en Guyane
26 septembre 2018	Réunion afin que les enquêteurs de Mayotte testent le matériel (balances, télémètres)
16 octobre 2018	Réunion avec la CNAM pour l'appariement de l'enquête avec le SNDS
17 octobre 2018	CHSCT spécial enquêteurs qui voit l'acceptation de l'utilisation du matériel (balances et télémètres) pour l'enquête à Mayotte
22-30 octobre 2018	Traduction de la partie sous casque pour Mayotte par la société Filogis (shimaore et kibushi)
25 octobre 2018	4 ^e réunion du Comité de suivi (bilan du test CAPI, préparation du test CAPI supplémentaire à Mayotte)
5 novembre 2018	3 ^e réunion du Comité de pilotage (bilan du test CAPI, circuit des données, préparation du test CAPI à Mayotte)
12 novembre 2018	Envoi du dossier de l'enquête au Comité du label
22 novembre 2018	Envoi du dossier relatif à l'autorisation de la CNIL à l'Institut national des données de santé
12 décembre 2018	Séance du Comité du label. Obtention du label d'intérêt général et de qualité statistique avec le caractère obligatoire.
13 décembre 2018	Avis favorable du CEREES en vue d'obtenir l'autorisation de la CNIL
9 janvier 2019	Dépôt du dossier complet à la CNIL
11 janvier 2019	Avis de la CNIL d'une prolongation du délai d'examen du dossier de deux mois en plus des deux mois légaux
11-21 janvier 2019	Test CAPI supplémentaire à Mayotte (avec notamment test du casque et du matériel : balances et télémètres)
5 février 2019	5 ^e réunion du Comité de suivi (bilan du test CAPI supplémentaire à Mayotte, organisation de la répétition générale, préparation de la collecte)
27 février 2019	4 ^e réunion du Comité de pilotage (bilan du test CAPI à Mayotte, information sur les démarches CNIL en cours, calendrier)
Mars 2019	Répétition générale de l'enquête
Avril-mai 2019	Tirage des échantillons dans les cinq DROM par le Criem
9 avril 2019	6 ^e réunion du Comité de suivi (bilan de la répétition générale, corrections des traductions pour le casque en lien avec les Directions régionales de l'Insee, préparation des documents de collecte, de la communication et des formations pour la collecte)
13 mai 2019	7 ^e réunion du Comité de suivi (dernières corrections du datamodel, calendrier pour la collecte)
19 juin 2019	8 ^e réunion du Comité de suivi : préparation de la collecte

1 ^{er} juillet 2019	Autorisation de la CNIL. Numéro : 919210. Autorisation de mener l'enquête et d'apparier avec les données 2015-2024 du SNDS.
8 juillet- 31 décembre 2019	Collecte de l'enquête
12 septembre 2019	Réunion d'information en vue des publications avec les Services Études et Diffusion des DROM à l'Insee
23 septembre 2019	9 ^e réunion du Comité de suivi : avancement de la collecte et projets de publications
11 octobre 2019	Comité de pilotage du Criem : le Criem prendra en charge le calcul de la pondération dans les cinq DROM ; livraison de la pondération prévue en mai 2020
19 novembre 2019	Signature de la convention avec la CNAM pour l'appariement avec le SNDS
27 novembre 2019	10 ^e réunion du Comité de suivi : déroulement de la collecte, décalage de la fin de la collecte du 14 décembre au 31 décembre 2019
20 janvier 2020	Réunion de bilan de collecte avec les divisions enquêtes ménages de l'Insee
28 janvier 2020	Signature de la convention avec l'Insee pour la reconstitution des NIR (BRPP)
Février 2020	Envoi de l'état civil de Fidéli par le Criem à la BRPP pour identification sur NIR (4 DROM historiques, ménages des logements échantillonnés à la date de tirage)
3 février 2020	Présentation de l'enquête aux journées ARS
5 février 2020	5 ^e réunion du Comité de pilotage (bilan de collecte, redressements, publications)
6 février 2020	Signature du contrat de sous-traitance entre la DREES et l'Insee
15 février 2020	11 ^e réunion du Comité de suivi : bilan de la collecte et appariements
17 février 2020	Signature de l'avenant à la convention tripartite DREES-Insee-GIP Irdes
7 avril 2020	Dépôt par l'Insee des NIR sur la plateforme Safe de la CNAM (entre février et avril difficultés pour mettre le fichier au format attendu par Safe et pour la transmission du fichier à la CNAM)
Mai 2020	Livraison par le service informatique de Lille de l'Insee des NIR et/ou de l'état-civil des répondants à l'enquête pour identification par la BRPP
Août 2020	Réception de l'appariement SNDS 2015-2018 pour les membres des ménages habitant à la date de tirage dans les logements échantillonnés
Septembre 2020	Livraison de la codification Sicore par l'Insee
octobre 2020	Livraison de la pondération définitive par le Criem (le décalage par rapport à la date de mai 2020 est dû à l'implication du Criem dans les redressements de l'enquête EpiCov vague 1)
Novembre 2020	Dépôt par l'Insee des NIR des répondants à l'enquête sur la plateforme Safe de la CNAM

Fin 2020	Signature de la convention entre la DREES, l'Insee et la DGOM portant participation de la DGOM au financement de l'enquête à hauteur d'environ 240 000 euros
9 avril 2021	Publication du dossier de la DREES de premiers résultats (n°78)
17 mai 2021	Mise en ligne des indicateurs de santé en <i>open data</i> sur le site de la DREES
Août 2021	Réception de l'appariement SNDS 2015-2019 pour les répondants à l'enquête (identification par validation du NIR déclaré ou reconstitution du NIR via l'état civil déclaré)
Été 2022	Mise à disposition de la base à la recherche via le CASD
Fin 2022	Livraison de l'appariement socio-fiscal (millésime 2019) par l'Insee
Jusqu'en 2024	Appariements chaque année avec le SNDS, mise à disposition de ces données appariées.

Questionnaire

Questions spécifiques aux DROM

L'adaptation du questionnaire aux DROM a fait l'objet d'une réunion avec les Agences régionales de santé concernées et les équipes de l'Insee dans les DROM le 25 juillet 2017. Par la suite, le questionnaire a été amélioré au fil des différents tests de l'enquête : le test papier à La Réunion et Mayotte permettant d'obtenir une cinquantaine de réponses par DROM en novembre 2017, le test CAPI à La Réunion et en Guyane en septembre-octobre 2018 permettant d'obtenir 80 questionnaires à La Réunion et 140 en Guyane et le test CAPI supplémentaire à Mayotte en janvier 2019 avec 220 questionnaires. Ces tests ont été à chaque fois accompagnés de déplacements des concepteurs et « chef de projet en organisation statistique », ce qui a été l'occasion de préparer au mieux la formation à l'enquête, d'estimer la durée de l'enquête, de faire du suivi d'enquêteurs et d'échanger avec les enquêteurs et les divisions « Enquêtes Ménages » à propos du questionnaire et de l'enquête dont le sujet était considéré comme nouveau pour l'Insee. Ces missions ont été très importantes pour tirer profit de ces tests et se rendre compte de la réalité du terrain dans les DROM, d'autant que les concepteurs ont manqué de temps pour exploiter les bases de données issues de ces tests. À Mayotte, la durée de passation du questionnaire avait été estimée à un niveau assez haut (beaucoup plus haut que la moyenne effectivement observée à l'issue de l'enquête), si bien qu'il y a une demande appuyée du Service territorial de l'Insee à Mayotte pour alléger le questionnaire. Les passations à Mayotte sont généralement plus longues que dans les autres DROM du fait de la nécessité de traduire de façon quasi systématique en shimaore ou kibushi. La formation à l'enquête initialement prévue sur une journée a été jugée trop courte à l'issue du test papier. Elle a été augmentée à deux jours, ce qui n'a pas été possible en France métropolitaine du fait du coût d'ajouter une journée de formation supplémentaire pour l'ensemble des enquêteurs.

De manière générale, ce qui a guidé l'élaboration de ce questionnaire a été la volonté de reprendre au maximum le questionnaire France métropolitaine de manière à permettre les comparaisons tout en ajoutant quelques questions spécifiques pour tenter de capter des enjeux particuliers aux DROM. Les ajouts n'ont cependant pas beaucoup modifié le questionnaire. Dans l'analyse, pour certains de ces ajouts, le point de comparaison « France métropolitaine » a manqué.

Le questionnaire débute par le tronc commun des ménages (TCM) de l'Insee qui doit être renseigné par un adulte du ménage. L'usage du TCM de l'Insee a permis de collecter plus d'informations socio-démographiques que ce qu'il a été possible de faire en France métropolitaine. Le module « logement » est celui dans lequel des adaptations ont été menées pour les DROM. Pour Mayotte, deux questions ont été ajoutées sur la nature du sol et la manière de s'approvisionner en eau qui sont des marqueurs sociaux à Mayotte. Elles sont issues du recensement. Pour tous les DROM, des questions ont été ajoutées portant sur la chaleur dans le logement, la présence d'un ventilateur, de la climatisation et de rats autour du logement. Chaque modalité de ces variables possède des effectifs relativement importants (tableau 26). Pour Mayotte et la Guyane, un bloc de questions a été ajouté sur la présence de l'eau courante, de l'eau chaude, de WC à l'intérieur du logement et d'une baignoire ou d'une douche. Ce type d'information est demandé dans le recensement dans les DROM. À Mayotte, les effectifs sont importants dans toutes les modalités. Toutefois, en Guyane, il n'y a que 77 logements (sur un total de 1 344 logements enquêtés) qui n'ont pas l'eau courante et parmi ceux qui l'ont, seuls 25 dans l'enquête n'ont pas de WC à l'intérieur du logement et 24 pas de douche ou baignoire dans le logement. En revanche 431 logements n'ont pas l'eau chaude dans l'enquête en Guyane.

Une fois le TCM et le module logement passés, l'enquête devient « individuelle ». Elle s'adresse à la ou les personne(s) tirées selon la méthode Kish. Il s'agit d'une personne du ménage âgée de 15 ans ou plus conformément au champ de l'enquête EHIS. Un module NIR a été ajouté pour recueillir ou pouvoir reconstituer le NIR du répondant. En effet, contrairement à la France métropolitaine, ce sont des logements qui ont été échantillonnés (pour faciliter une collecte en face-à-face) si bien qu'il n'est pas possible de reconstituer le NIR du répondant avec Fidéli. Il s'agit de recueillir le NIR ainsi que toutes les informations permettant de le reconstituer (cf. partie « appariements »). Pour Mayotte et la Guyane, une question préalable a été ajoutée pour savoir si les personnes sont affiliées à l'Assurance maladie (dite « Sécurité sociale » dans le questionnaire). Cette question a été exploitée et il est finalement dommage de ne pas en avoir disposer pour l'ensemble des DROM. Pour la Guyane, une question a été ajoutée sur le fait de disposer de l'Aide médicale d'État. La question n'a été ajoutée qu'en Guyane car le dispositif n'existe pas à Mayotte et qu'il a semblé que le recours était faible dans les autres DROM. Ceci étant, c'est une information qui est sans doute connue de manière beaucoup plus fiable dans les registres. On voit déjà que, sur le taux de bénéficiaires de la CMU-C, il y a des écarts importants entre la donnée d'enquête et les chiffres du Fonds CMU. Par ailleurs, l'AME est un sujet sensible, donc il est complexe de communiquer dessus si les chiffres déclaratifs sont jugés de moins bonne qualité que des chiffres de registre. À Mayotte, le questionnaire sur la complémentaire santé est allégé. Le taux de couverture par une complémentaire santé est bas (12 % d'après l'enquête). Les questions sur le coût de la complémentaire santé et la satisfaction relative à la complémentaire santé ont été supprimées. Les tests ont mis en évidence que beaucoup de personnes à Mayotte ne savaient pas ce qu'était une complémentaire santé. Cette modalité a été ajoutée pour l'ensemble des DROM dans les raisons de la non couverture par une complémentaire santé. C'est la modalité la plus utilisée parmi ceux qui à Mayotte déclarent être affiliés à l'Assurance maladie mais ne pas avoir de complémentaire santé (373 personnes sur 2026 répondants au total à Mayotte). Dans les autres DROM, les effectifs concernés par la question relative aux raisons de ne pas avoir de complémentaire santé sont faibles de manière générale et seules entre 3 et 6 personnes déclarent ne pas savoir ce qu'est

une complémentaire santé. Suite aux tests à La Réunion, il a paru important pour le bon déroulement du questionnaire d'ajouter une question sur le chèque santé du département. Il s'agit d'une aide accordée aux personnes de 60 ans ou plus sous conditions de ressources. Les enquêteurs ne savaient pas bien s'il devait inclure ou non ce chèque santé dans le coût de la complémentaire santé. Il a été choisi de demander combien le ménage payait en plus du chèque santé pour la complémentaire santé, ce chèque étant déduit du coût de la complémentaire santé demandé au ménage et d'un montant connu (entre 300 et 360 euros).

La partie du questionnaire consacrée au recueil du diplôme a été sortie du TCM de manière à ne la poser que pour le répondant et pas pour l'ensemble du ménage. Cela visait à alléger la charge de collecte alors qu'en France métropolitaine le diplôme n'était demandé que pour le répondant.

Concernant le questionnaire santé, quatre maladies ont été ajoutées dans la liste des maladies (sous la forme de maladie ou de séquelles de cette maladie) : il s'agit du paludisme, de la dengue, du chikungunya et du zika. Ces questions devaient être enlevées avant la collecte s'il n'y avait pas d'épidémie. Elles ont été laissées. Dans le bilan enquêteurs, l'agacement de certains enquêtés dans un DROM face à ces questions alors qu'aucune épidémie de ce type n'était présente a été notée. À Mayotte et en Guyane, presque personne ne déclare ces maladies et de manière générale, le paludisme et zika sont très rarement évoqués. Concernant la liste de maladies, le bilan enquêteurs fait état du souhait d'avoir une catégorie « Autres maladies, précisez » face à des enquêtés qui n'ont pas compris ne pas pouvoir évoquer leur maladie (« prostate », « cancer »). Ce point avait été évoqué lors des tests de l'enquête où la catégorie était présente mais, outre le fait que ce type de questions ouvertes n'est pas facilement exploitable, certains enquêteurs s'étaient retrouvés dans des situations difficiles lorsque les enquêtés avaient évoqué leurs « autres » maladies. Lors des tests, il avait également été discuté de demander pour chaque maladie si un diagnostic avait été posé pour essayer d'améliorer la fiabilité du recueil déclaratif alors que l'on s'attendait à ce que l'appariement avec le SNDS soit de moins bonne qualité qu'en France métropolitaine. Mais cette idée a été abandonnée. Cela rallongeait la passation de cette liste pour une plus-value discutable. Forcément, une enquête avec examen de santé permet un meilleur recueil des maladies.

Les deux questions sur l'opération de la cataracte n'ont pas été posées à Mayotte pour alléger le questionnaire et car il a été jugé à partir du SNDS que la prévalence serait trop faible. Cependant, c'était finalement assez gênant pour l'exploitation de ne pas disposer du chiffre pour Mayotte. Cette remarque reste valable pour les questions sur les limitations fonctionnelles (ADL et IADL²⁰). La population est jeune à Mayotte si bien qu'il a été pensé que ces questions relatives aux limitations fonctionnelles ne seraient pas exploitables pour disposer de statistiques sur Mayotte et qu'elles ne serviraient qu'à avoir des statistiques France entière où le poids de Mayotte est de toutes façons très faible. Il avait été envisagé de surreprésenter dans l'échantillon les personnes âgées pour disposer d'un effectif assez important mais ce n'était pas possible avec la base de sondage (les données cartographiques). Finalement, toujours dans le souci d'alléger le questionnaire à Mayotte, ces questions ont été posées partout sauf à Mayotte. L'exploitation des données (Leduc et al., 2021) a montré une très forte prévalence du handicap à Mayotte parmi la population âgée donc il est finalement dommage que ces questions n'aient pas été administrées à Mayotte, d'autant que le champ de l'enquête Unono wa Maore de Santé publique France s'arrêtait aux personnes âgées de 69 ans.

La question AM10A portant sur le type de personnes vue lors du dernier problème de santé avait été introduite principalement dans le but de capter le recours aux médecines « parallèles » dans les DROM. Elle a été mise pour comparaison en France métropolitaine également. L'exemple du « fundi²¹ » était donné pour Mayotte et celui du shaman ou « médecin feuille » pour la Guyane. Ces exemples étaient issus de discussions avec les enquêteurs lors des formations aux tests. Cependant, de manière générale, la question AM10A n'a pas été facile à formuler et la dernière modalité (« une autre personne ») a été très peu utilisée (tableau 26).

Quelques questions ont été introduites pour capter la spécificité du recours aux soins à Mayotte : deux questions sur le recours au dispensaire²² et une question sur la manière de se procurer les médicaments. Lors des tests, des enquêteurs ont expliqué que les médicaments pouvaient être obtenus à la pharmacie ou gratuitement au dispensaire et qu'une partie de la population y recourait. Les effectifs sont importants dans chacune des modalités de ces deux variables. Lors des tests également, il a été constaté qu'une partie de la population semblait ne pas connaître les tests de dépistage à Mayotte. Une question a donc été ajoutée pour savoir si une prise de sang avait déjà été faite à laquelle il est plus facile de répondre que sur le type d'analyse menée à partir du sang.

²⁰ ADL : *activities of daily life* ; IADL : *instrumental activities of daily life*.

²¹ Le « fundi » est celui qui a le savoir. Il y a un fundi pour réparer sa voiture par exemple. Pour des problèmes de santé, on peut aussi recourir à un fundi (mais ce n'est pas le même que celui pour la voiture).

²² Concernant le recours au dispensaire, le service d'étude de l'Insee en charge de la Guyane a regretté que la question n'ait pas été posée en Guyane où le recours au dispensaire semble être possible.

Pour l'ensemble des DROM, une question a été ajoutée pour savoir si les personnes s'étaient déjà faites soigner en dehors de leur département de résidence et si elle avait renoncé à des soins faute du spécialiste nécessaire dans le département.

Concernant le poids et la taille, les questions ont été remplacées par un recueil mesuré à Mayotte car les tests ont montré que beaucoup de personnes ne connaissaient pas leur poids et leur taille à Mayotte (cf. partie « protocole »). Si la taille figure sur la carte d'identité française et comorienne, il est apparu, lors du test, que le chiffre indiqué n'était pas toujours fiable. Il aurait été intéressant de garder un recueil déclaratif pour étudier l'écart entre le déclaratif et le mesuré. C'est ce qu'a fait Santé publique France dans l'enquête Unono wa maore. Par ailleurs, la non-réponse a été assez importante en Guyane notamment sur le poids et la taille. Lors de la préparation de l'enquête la question s'était posée de savoir si on étendait le recueil mesuré à d'autres DROM que Mayotte. Au vu des résultats, il aurait sans doute été intéressant de l'étendre à la Guyane même si le recueil mesuré n'a pas donné pleinement satisfaction à Mayotte (cf. partie « protocole »).

Le recueil concernant la question sur la fréquence de la consommation de jus de fruits maison ou 100 % pur jus n'est probablement pas de bonne qualité dans les DROM. Outre le fait qu'il faille être au courant que le jus acheté est « 100 % pur jus », il y a beaucoup de jus maison dans les DROM et il est d'usage d'ajouter du sucre dans les jus maison. La consigne avait été donnée aux enquêteurs de ne pas compter ces jus dans lesquels les gens ajoutaient du sucre car l'objet de cette question Eurostat était plutôt de recueillir la part de fruits qui est consommée sous forme de jus et non de s'intéresser au caractère sucré de la consommation de jus. C'est ainsi que c'est formulé dans le manuel européen et il y avait eu des discussions à Eurostat sur le sujet. Mais finalement pour les DROM, exclure le sucre ajouté conduisait très certainement à exclure la plupart des jus. Par ailleurs, on ne sait pas bien comment les enquêteurs ont *in fine* pu suivre cette consigne d'exclure les jus maison s'il y avait du sucre ajouté (ont-ils par exemple systématiquement demandé aux gens s'ils ajoutaient du sucre, comme c'était demandé en formation ?). De ce fait, cette question n'a pas été exploitée dans la publication des premiers résultats de l'enquête.

Deux questions ont été introduites juste avant la collecte dans l'ensemble des DROM relatives à l'inquiétude des répercussions pour la santé de la pollution présente dans les aliments ou l'air. Cela a fait suite à la formation « enquêteurs » en Martinique où des critiques ont porté sur le fait que l'enquête Santé n'aborde pas la problématique du chlordécone. Sur ce point précisément, il semblait que l'appariement SNDS (ou l'exploitation de l'exhaustivité du SNDS) était plus à même de répondre à cette problématique avec par exemple l'étude de la survenue de cancers. Deux questions ont toutefois été ajoutées sur le ressenti par rapport à la pollution. Elles ont été exploitées par les SED de l'Insee dans leurs premières publications sur la Martinique, la Guadeloupe et la Guyane. Il aurait été intéressant d'en disposer également en France métropolitaine pour pouvoir comparer les niveaux.

La question ayant pour but la mesure de l'aversion au risque (note donnée entre 0 et 10) a été supprimée dans les DROM car au vu des tests, elle semblait très mal comprise. L'échelle Cantril a toutefois été conservée car elle paraissait mieux comprise. L'exploitation de cette question s'est révélée intéressante dans les premiers résultats montrant une note particulièrement faible à Mayotte. Il était initialement prévu que cette question soit posée sous casque mais elle a été sortie du casque car il apparaissait que le travail de l'enquêteur pour expliquer la question était nécessaire.

Concernant le questionnaire sous casque (santé mentale, alcool, tabac, drogues), deux questions ont été ajoutées pour cerner la consommation de tabac à Mayotte. Il s'agit d'une sorte de poudre de tabac qui est avalée. Au total 167 personnes déclarent en consommer dans l'échantillon de 2026 personnes. La question sur la quantité consommée est sans doute peu intéressante, d'autant que trouver une unité de mesure (la cuillère) n'était pas simple. Le module sur la consommation de drogues n'a pas été bien conçu même si des experts du sujet avaient été rencontrés. La consommation de cocaïne est déclarée au cours de la vie par entre 9 et 44 personnes selon le DROM. Donc la question sur la consommation au cours des douze derniers mois ne semble pas vraiment exploitable car les effectifs absolus sont très faibles (entre 3 et 13 personnes). La consommation de médicament du type « Artane, Rivotril, Rohypnol » au cours de la vie est déclarée par entre 4 et 18 personnes. Du fait de ces effectifs absolus très faibles, cette question relative au douze dernier mois ne semble pas vraiment exploitable, ni celle pour savoir si la consommation de ces médicaments s'est faite avec de l'alcool ou du cannabis. Par ailleurs, cette question avait été reprise du Baromètre Santé de Santé publique France mais nous n'étions pas bien sûrs de l'actualité de la liste de ces médicaments. Enfin, ces questions n'ont pas pu être posées en France métropolitaine faute de place dans le questionnaire, donc il n'est pas possible de mettre en perspective les résultats avec ceux de la France métropolitaine. Les questions sur la consommation de cannabis (vie et douze derniers mois) posée en France métropolitaine et dans les DROM auraient peut-être été suffisantes. Les questions sur la consommation de la « chimique », une drogue, à Mayotte (vie et douze derniers mois) ne sont pas très intéressantes car c'est déclaré par 15 personnes (vie). La pertinence d'avoir posé ce module sous casque est questionnée dans la partie « bilan du casque ».

Certaines idées issues des tests ont été abandonnées. Par exemple, il a été envisagé après le test papier de novembre 2017 d'équiper les enquêteurs de Mayotte de douchettes pour scanner les boîtes de médicaments afin d'en inférer les pathologies des personnes.

La force du dispositif a bien été de disposer de questions communes à la France métropolitaine et aux DROM et de manière générale il n'est pas simple d'exploiter une question qui n'a pas son pendant en France métropolitaine. L'ajout de questions spécifiques aux DROM aurait sans doute profité d'un groupe de travail dédié à ce sujet. Le temps ne l'a pas vraiment permis

et du reste, le nombre de questions concernées était faible. Ceci étant, la durée de passation du questionnaire dans les DROM avait été surestimée par l'Insee.

Tableau 26 • Tris à plat des questions spécifiques aux DROM et de quelques questions des modules français concernant le logement et la complémentaire santé

	Guade- loupe	Marti- nique	Guyane	La Réunion	Mayotte
Nombre de logements (table "ménages")	1300	1195	1341	1077	1218
Sol : terre battue					387
Sol : béton					264
Sol : carrelage ou autre					570
Humidité salle de bains : non	728	576	796	656	864
Humidité salle de bains : oui, à quelques en- droits	391	390	341	276	233
Humidité salle de bains : oui, à beaucoup d'en- droits	181	229	204	145	121
Bruits : pas gênant	815	662	781	683	576
Bruits : un peu gênant	284	331	303	236	430
Bruits : très gênant	157	122	178	119	173
Bruits : insupportable	43	80	81	40	43
Souffrir chaleur dans le logement : oui	715	703	798	488	933
Climatisation dans le logement : oui	710	394	718	312	302
Ventilateur dans le logement : oui	1105	959	1184	627	730
Eau courante : non			77		493
Eau chaude : non			431		543
WC à l'intérieur : non			25		166
Baignoire/douche : non			24		207
Rats autour du logement ou dedans : oui	642	669	547	442	876
Approvisionnement en eau (si pas eau cou- rante) :					
Dans la cour					284
Chez un parent					21
Chez un tiers					106

	Guade- loupe	Marti- nique	Guyane	La Réunion	Mayotte
Borne, fontaine					48
Puit, citerne					13
Autre					21
Nombre de répondants (table "individus")					
	1982	1817	2176	1763	2026
Complémentaire santé					
Assurance maladie ("Sécurité sociale") : non			136		742
Aide médicale d'état : oui			53		
CMU-C : oui	387	308	680	479	
Complémentaire santé (si pas CMU-C) : oui	1348	1310	981	1131	183
Raisons de ne pas avoir de complémentaire santé :					
Pas les moyens	45	26	39	11	212
Pris en charge à 100 % par la Sécurité sociale	26	23	14	5	199
Ne sait pas ce qu'est une mutuelle	4	6	3	0	373
Pas besoin	9	14	14	6	77
Pas le temps	20	18	32	6	82
Vous ne savez pas comment faire	11	4	17	2	55
En cours adhésion	39	30	47	19	23
Questionnaire Santé					
Liste de maladies					
Paludisme : oui	2	3	25	4	17
Dengue : oui	100	45	50	108	1
Chikungunya : oui	285	131	41	114	14
Zika : oui	78	48	14	1	0
Recours aux soins					
Dernière visite pour un problème de santé :					
Médecin	1785	1474	1948	1450	1776

	Guade- loupe	Marti- nique	Guyane	La Réunion	Mayotte
Autre professionnel de santé	183	330	191	297	160
Autre personne	7	9	22	14	73
Dispensaire : visite il y a moins de 12 mois					722
Dispensaire : visite il y a 12 mois ou plus					696
Dispensaire : aucune visite					599
A eu besoin de se faire soigner à l'extérieur du DROM	133	102	205	23	184
-> Cela a été pris en charge par l'Assurance maladie	111	82	120	17	78
-> Cela a été pris en charge par vous	21	17	77	5	82
-> Cela a été pris en charge par une autre personne	1	2	8	1	24
Médicaments provenant du dispensaire					456
Médicaments provenant de la pharmacie					1475
Ne prend jamais de médicaments					94
Parmi ceux qui déclarent n'avoir jamais fait de test de glycémie ou de test de cholestérol (603 personnes), nombre qui déclarent avoir déjà eu une prise de sang					366
Renoncement aux soins car pas le soin dans le DROM	57	55	148	16	185
Habitudes de vie en lien avec la santé					
Alimentation : très inquiet par la pollution	887	826	688	686	584
Alimentation : un peu inquiet par la pollution	586	556	700	638	609
Alimentation : pas inquiet par la pollution	498	426	779	425	815
Environnement : très inquiet par la pollution	985	1036	686	643	603
Environnement : un peu inquiet par la pollution	539	477	699	635	553
Environnement : pas inquiet par la pollution	447	295	782	472	851
Tibacu : tous les jours					94
Tibacu : occasionnellement					24
Tibacu : par le passé					49

	Guade- loupe	Marti- nique	Guyane	La Réunion	Mayotte
Tibacu : jamais					1835
Cannabis au cours de la vie : oui	205	228	331	314	96
Cannabis au cours des 12 derniers mois : oui	67	77	117	117	36
Cocaïne au cours de la vie : oui	14	25	44	33	9
Cocaïne au cours des 12 derniers mois : oui	3	4	13	8	4
Chimique au cours de la vie : oui					15
Chimique au cours des 12 derniers mois : oui					6
Autre drogue au cours de la vie : oui	32	37	37	45	24
Autre drogue au cours des 12 derniers mois : oui	12	8	13	13	7
Médicaments type Artane, rivotril, rohypnol :					
Au cours de la vie	20	22	19	32	11
Au cours des 12 derniers mois	8	6	6	8	4
Avec alcool ou cannabis au cours de la vie	4	8	6	18	5
Avec alcool ou cannabis au cours des 12 der- niers mois	2	3	2	5	1

Lecture > Dans l'enquête en Guadeloupe, 2 personnes ont déclaré avoir pris des médicaments type Artane, rivotril ou rohypnol avec de l'alcool ou du cannabis au cours des 12 derniers mois.

Champ > DROM

Source > EHS 2019, DROM, DREES-Insee.

Durée du questionnaire

Conformément au plan de sondage, l'échantillon était séparé en deux sous-échantillons : le numéro 1 pour lequel on anticipait qu'il n'y aurait qu'une personne Kish dans l'enquête (donc une seule personne de 15 ans ou plus dans le ménage) et le numéro 2 pour lequel on anticipait qu'il y aurait 2 Kish. Il s'agissait d'une prédiction à partir des informations de composition familiale de la base de sondage (Fidéli). Lors de la collecte, certains ménages du sous-échantillon 1 avaient finalement deux personnes de 15 ans ou plus (donc 2 Kish) et d'autres du sous-échantillon 2 n'avaient finalement qu'une personne de 15 ans ou plus dans le ménage parce que la situation familiale avait évolué depuis le millésime de la base de sondage. La durée moyenne observée pour le sous-échantillon 1 est de 54 minutes, celle pour le sous-échantillon 2 est de 1 heure 12 minutes. À Mayotte, il n'y avait pas de sous-échantillon comme il n'était pas possible d'anticiper la composition du ménage avec les données cartographiques. La durée moyenne de l'enquête a été de 1 heure 06 minutes. Finalement, la durée moyenne de l'enquête a été plus courte que celle qui avait été anticipée à partir des données des tests CAPI de l'enquête (et donc utilisée pour la prévision de la quotité « enquêteurs ») : 10 minutes de moins en moyenne pour le sous-échantillon 1, 7 minutes pour le sous-échantillon 2 et surtout 28 minutes de moins pour Mayotte, ce qui est vraiment beaucoup. Il aurait été possible de laisser les questions sur ADL et IADL facilement et même d'ajouter des questions spécifiques à Mayotte.

Documents de collecte

Les documents de communication et de collecte France métropolitaine et DROM étaient très proches. Ils figurent sur le page internet de l'enquête EHIS sur le site internet de la DREES dans la rubrique « documents de collecte »²³. Un effort important a porté sur les cartes codes des fruits et des légumes. Ces deux cartes codes ont été adaptées pour chaque DROM avec des exemples de « portions » de fruits et de légumes propres à chaque territoire. Le graphiste de la DREES a été sollicité pour réaliser ces illustrations. En plus des traductions nécessaires pour le casque (cf. partie « bilan du casque »), une traduction concertée a été menée pour la liste des maladies et les tests de dépistage en créole guyanais, bushinenge, portugais brésilien pour l'enquête en Guyane et en shimaore et kibushi pour l'enquête à Mayotte. Les traductions ont été demandées à la société Filogis et ont été retravaillées avec les enquêteurs de l'Insee lors des tests. Les enquêteurs ont dit dans le bilan que ces traductions avaient été très utiles. Elles permettent, en plus de les aider dans leur travail, d'améliorer l'homogénéité des informations collectées. Ces traductions sont disponibles sur la page consacrée à l'enquête sur le site internet de la DREES. Une carte avec des numéros d'aide en lien avec la santé était laissée aux enquêtés sur le modèle de ce qui était fait pour l'enquête Cadre de vie et sécurité de l'Insee.

²³ <https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/enquete-sante-europeenne-ehis>

Plan de sondage

L'échantillon de l'enquête santé DOM a été tiré par le Criem de l'Insee.

Description du plan de sondage dans les quatre DROM historiques

Le champ de l'enquête est l'ensemble des logements ordinaires qui constituent une résidence principale dans les quatre DROM historiques : Guadeloupe, Martinique, Guyane et La Réunion.

Champ géographique :

- Guadeloupe : Guadeloupe continentale et Marie-Galante (la Désirade et les Saintes sont exclues)
- Martinique : Aucune exclusion
- Guyane : Les 10 communes suivantes sont exclues : Régina, Iracoubo, Maripasoula, Papaïchton, Grand-Santi, Saint Elie, Ouanary, Camopi, Saül et Awala-Yalimapo. Il s'agit d'un échantillon sur la zone dite « Guyane côtière élargie ». Par ailleurs, il y a des exclusions d'îlots pour commodité d'enquête au sein des communes dans le champ de l'enquête, de l'ordre de 12 % des logements.
- La Réunion : La Réunion entière privée de la commune de Cilaos. Des îlots difficiles d'accès sont exclus également de la base de tirage de La Réunion, notamment le cirque de Mafate.

Critère d'âge : l'ensemble des ménages vivant en résidence ordinaire avec au moins une personne de 15 ans ou plus sont dans le champ de l'enquête.

Unité de tirage : le logement. Par la suite, lorsque l'enquêteur est au sein du ménage et a décrit les personnes le composant, un tirage Kish a lieu : on tire une personne à enquêter dans le ménage lorsqu'il y a dans celui-ci une unique personne de 15 ans ou plus ; on en tire deux lorsque le ménage est composé d'au moins deux personnes de 15 ans ou plus au sein du ménage.

Base de tirage : Fidéli 2018. Il a été décidé de tirer l'échantillon dans Fidéli plutôt que dans le recensement car l'appariement avec le SNDS des personnes occupant les logements dans le millésime de la base de tirage devait fournir des variables auxiliaires intéressantes pour les modèles de correction de la non-réponse totale (voir partie « appariements »). Il n'a finalement pas été possible d'exploiter cet appariement pour la correction de la non-réponse totale faute de temps.

Trois types de logements ont été considérés : les logements connus comme résidences principales à la taxe d'habitation, des logements secondaires ou vacants soupçonnés d'être des résidences principales car un ménage déclare des revenus à l'adresse du logement (reclassement) et des logements non présents à la taxe d'habitation qui sont probablement des logements plus précaires et informels mais à l'adresse desquels un ménage déclare ses revenus (création de logement) (tableau 27).

Tableau 27 • Répartition des logements considérés pour le tirage de l'enquête

	Base initiale	Reclassement	Création de logement	Total
Guadeloupe	153 782	173	10 941	164 896
Martinique	160 754	57	3 443	164 254
Guyane	51 465	196	8 857	60 518
La Réunion	322 413	193	11 540	334 146

Taille de l'échantillon :

L'échantillon devait initialement avoir une taille de 1 500 répondants par DROM. Par la suite, il a été décidé de tirer deux Kish par logement lorsqu'il y avait au moins 2 personnes âgées de 15 ans ou plus dans le logement, ce qui induit un effet de grappe et réduit donc la précision des estimateurs. Il a donc été décidé d'augmenter la taille de l'échantillon pour pallier cet effet de grappe. Le Criem a estimé un coefficient de corrélation entre les réponses des personnes d'un même ménage au mini-module européen sur la santé à partir des données de l'enquête Emploi où tous les adultes d'un ménage sont interrogés. Cela a permis d'en déduire de combien il fallait augmenter la taille de l'échantillon pour obtenir le même niveau de précision, en interrogeant quand c'est possible 2 personnes par ménage qu'avec 1 500 répondants à raison d'un répondant par ménage. L'échantillon a été découpé en deux sous-échantillons : le sous-échantillon 1 dans lequel on anticipe à partir de la base de sondage qu'il n'y aura qu'un Kish tiré, c'est-à-dire qu'il n'y a qu'une personne de 15 ans ou plus dans le ménage et le sous-échantillon 2 dans lequel on anticipe qu'il y aura deux Kish tirés, c'est-à-dire qu'il y a au moins deux personnes de 15 ans ou plus dans le ménage (tableau 28).

Tableau 28 • Répartition du nombre de logements à tirer par sous-échantillon

Nombre de logements tirés	Sous-échantillon 1	Sous-échantillon 2	Total
Guadeloupe	556	1 154	1 710
Martinique	545	1 123	1 668
Guyane	288	1 738	2 026
La Réunion	331	1 244	1 575

Tirage : Il s'agit d'un tirage systématique de logements à probabilités égales dans chaque strate, les strates étant définies comme le croisement entre le sous-échantillon et des strates géographiques classiquement utilisées par l'Insee dans les tirages d'échantillon dans les DROM. Le tirage est donc proportionnel au nombre de logements dans chaque strate. Les variables de tri pour le tirage systématique sont les suivantes :

- Lieu de naissance des adultes occupants : nombre d'adultes nés dans un DROM, en France métropolitaine, à l'étranger ;
- Âge moyen des habitants du logement ;
- Revenu fiscal du ménage par unité de consommation.

Poids de tirage :

Pour chaque sous-échantillon d'une strate, les poids initiaux des logements sont égaux :

Poids du logement du sous-ech. i dans la strate j = (nombre total de logement du sous-ech. i de la strate j) divisé par (nombre de logements du sous-ech. i de la strate j à tirer d'après le plan de sondage).

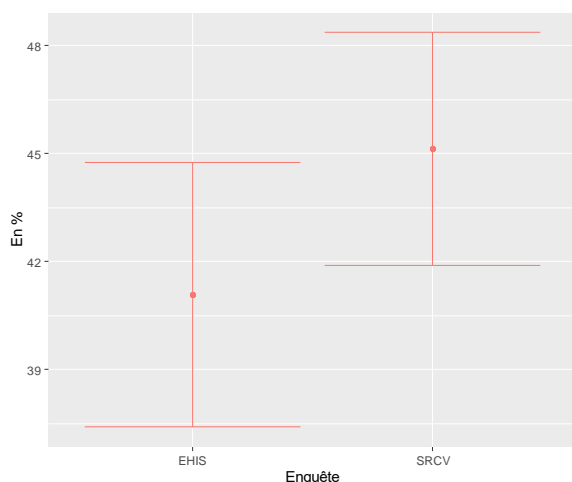
Mise en perspective du tirage dans Fidéli pour les quatre DROM historiques

Une question qui peut se poser est de savoir si le tirage dans Fidéli pour les quatre DROM historiques a généré un biais dans les statistiques. On considère en général que le taux de couverture de Fidéli est de l'ordre de 75 % du recensement en Guyane, 80 % dans les Antilles et 96 % à La Réunion. Pour étudier cela, on compare les résultats de EHIS 2019 au volet DROM de SRCV 2018 dont l'échantillon a été tiré dans le recensement. La comparaison n'est pas parfaite, les taux de réponse peuvent jouer mais il s'agit de deux enquêtes très rapprochées toutes deux réalisées par le réseau d'enquêteurs de l'Insee. Le graphique 11 présente la part de logements sans eau chaude, sans douche ni baignoire et sans WC à l'intérieur en Guyane. Ces questions n'ont été posées qu'en Guyane et à Mayotte dans EHIS 2019. On ne constate que de très faibles écarts entre EHIS 2019 et SRCV 2018. Le graphique 12 s'intéresse au mini-module européen sur la santé en comparaison entre les deux enquêtes dans les quatre DROM. On voit peu d'écart entre les deux enquêtes sur la question d'état de santé général et le GALI. En revanche, sur la présence de maladies chroniques, il y a un écart significatif partout sauf à La Réunion. Il est possible de l'interpréter comme une conséquence du sujet de l'enquête : les enquêtés pourraient déclarer plus fréquemment des maladies chroniques quand on leur annonce une enquête sur la santé (disons qu'ils se concentrent sur ce sujet) tandis qu'ils en déclareraient moins dans une enquête sur les ressources et les conditions de vie faisant moins d'efforts sur les questions relatives à la santé.

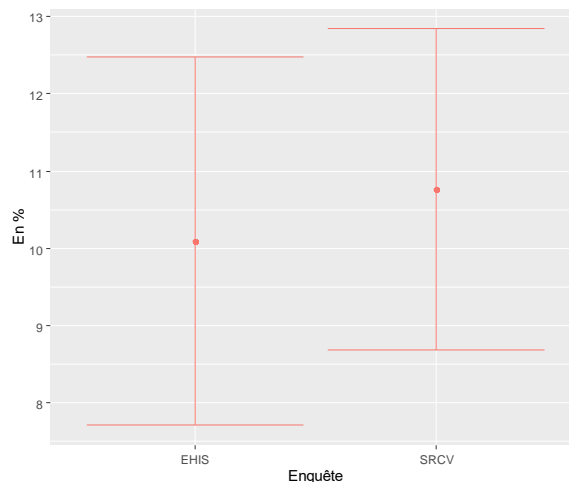
Ces graphiques ne permettent pas d'étayer l'hypothèse d'un fort biais induit par le tirage dans Fidéli plutôt que dans le recensement.

Graphique 11 • Comparaison de EHIS 2019 et SRCV 2018 sur les caractéristiques des logements en Guyane

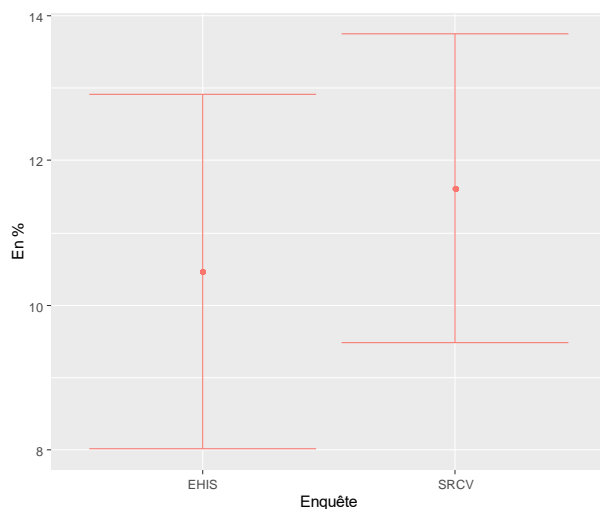
a) Logement sans eau chaude



b) Logement sans douche ni baignoire



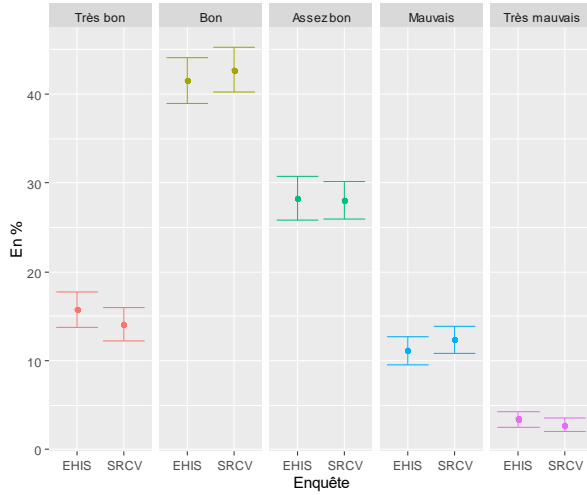
c) Logement sans WC à l'intérieur



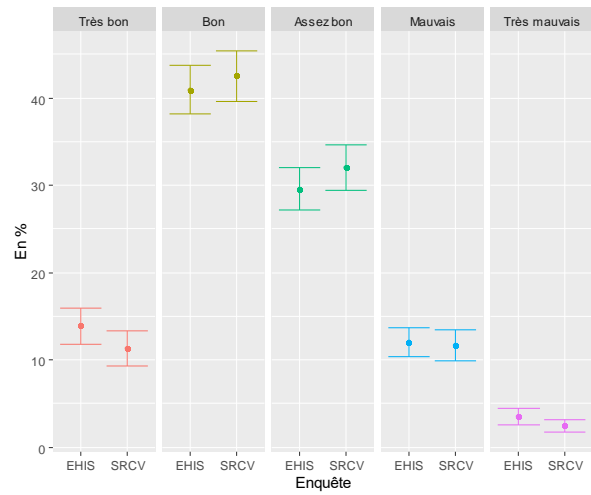
Graphique 12 • Comparaison de EHIS 2019 et SRCV 2018 sur le mini-module européen sur la santé dans les quatre DROM historiques

a) Guadeloupe

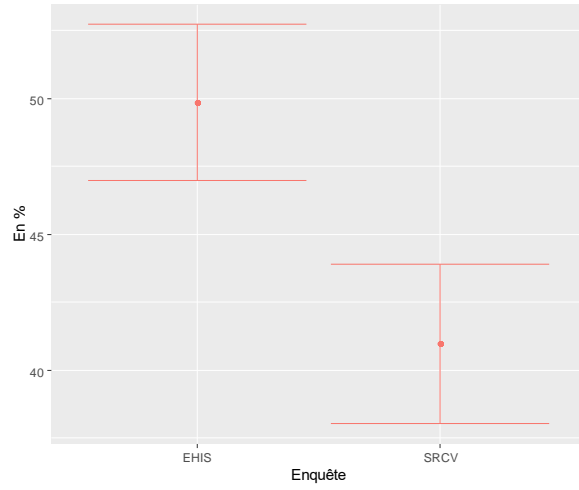
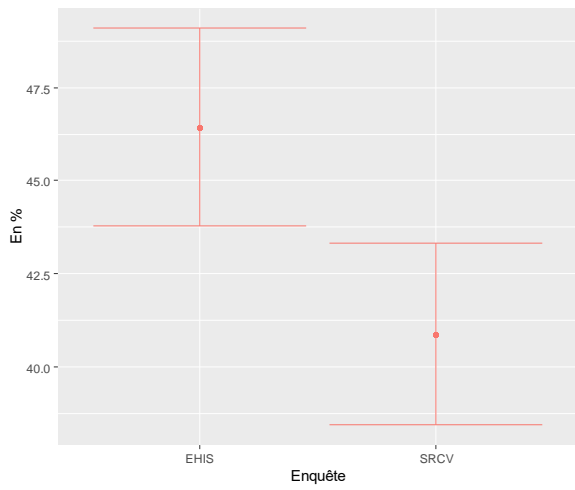
État de santé général



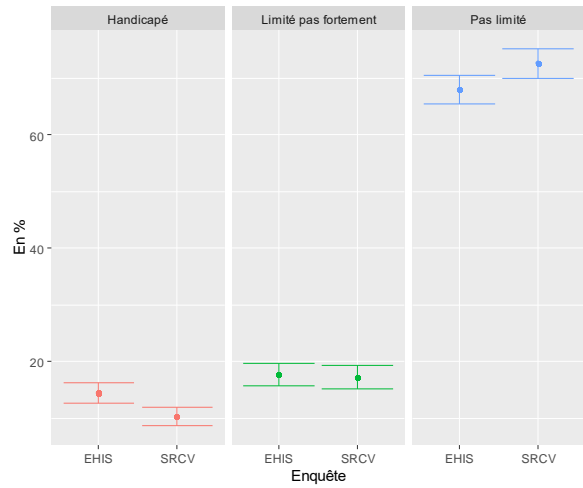
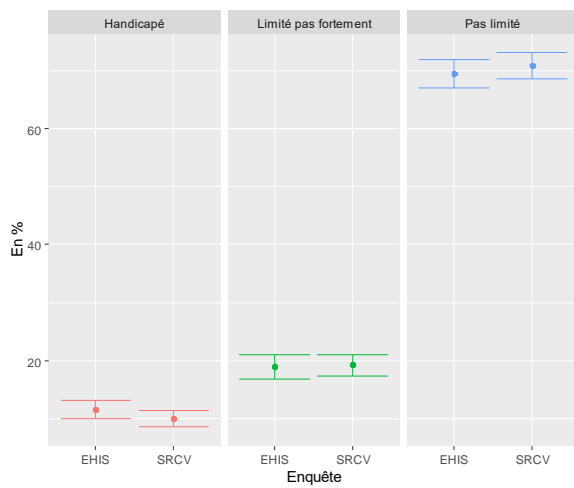
b) Martinique



Présence d'une maladie chronique

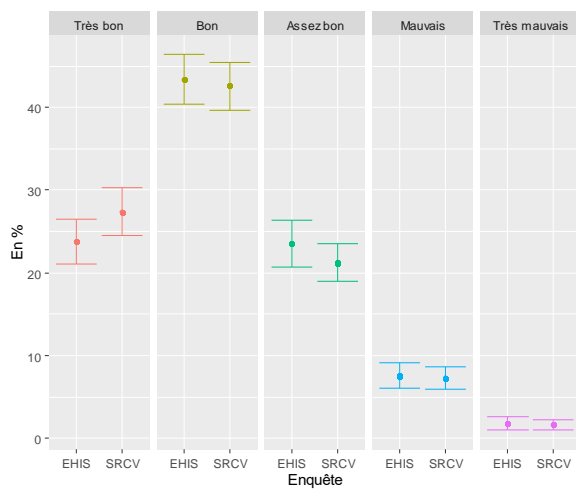


GALI (limitations d'activité)

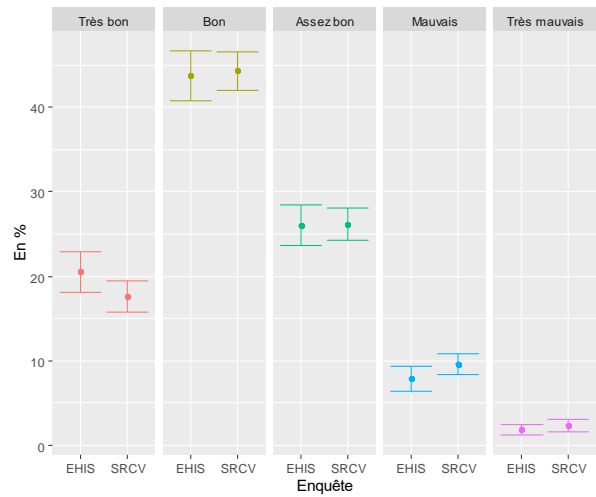


b) Guyane

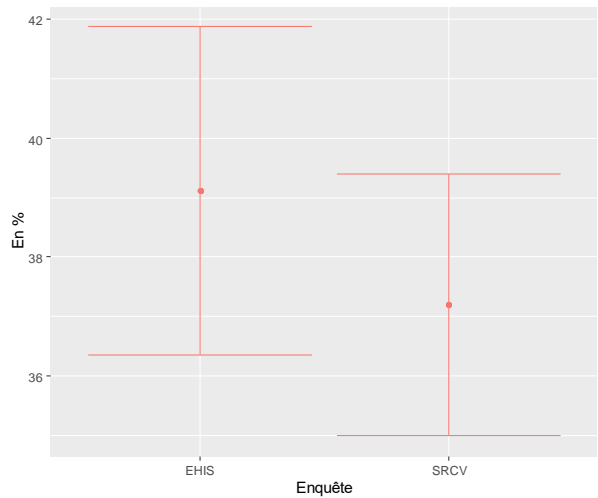
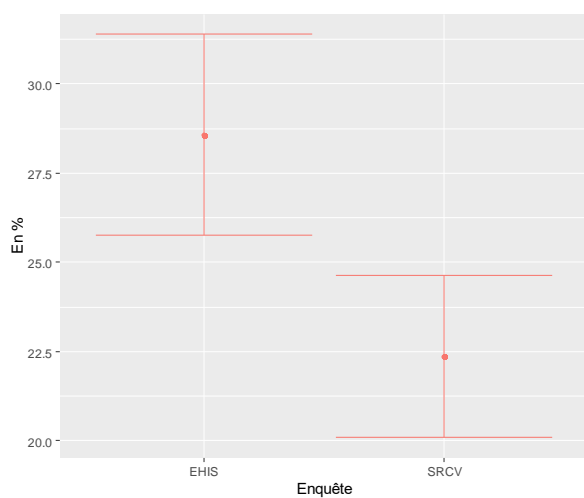
État de santé général



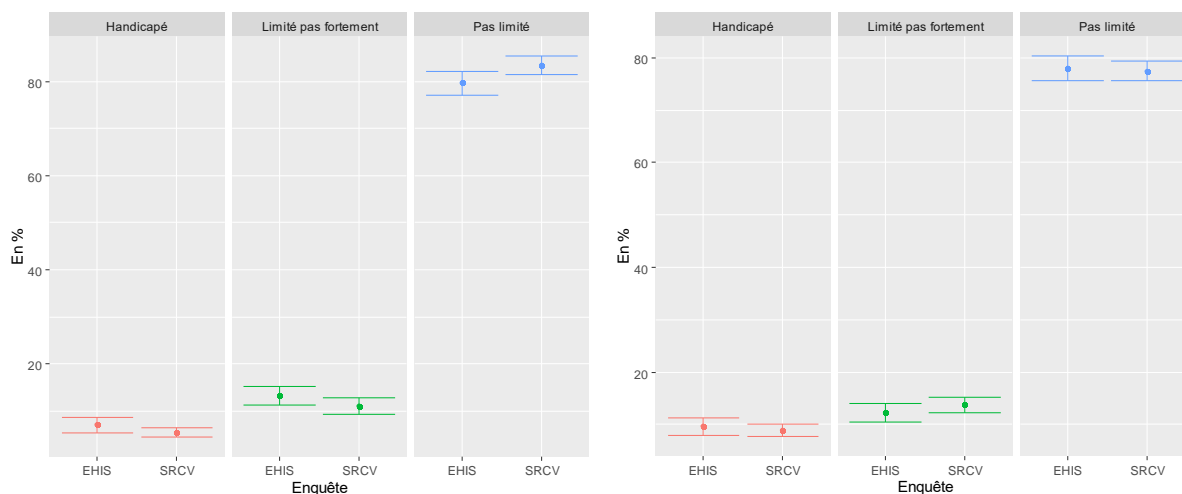
d) La Réunion



Présence d'une maladie chronique



GALI (limitations d'activité)



Note > Statistiques pondérées. Les intervalles de confiance sont estimés sous l'hypothèse de sondage aléatoire simple avec en plus des clusters dans le cas de l'interrogation de deux Kish par ménage.

Description du plan de sondage à Mayotte

Champ géographique : il n'y a aucune exclusion.

Taille de l'échantillon : 1 720 logements.

Base de sondage : données cartographique 2018, groupe de rotation 1 : mis à jour le plus récemment²⁴. Cette base contient 11 173 adresses et 13 368 logements.

La base est découpée en trois groupes : les mono-logements, les petites adresses, les grandes adresses (tableau 29).

Tableau 29 • Répartition des adresses des données cartographique (groupe de rotation 1)

Mono-logements	Petites adresses	Grandes adresses	Total
9 821	1 337	15	11 173
87,9 %	12,0 %	0,1	100 %

Méthode de tirage : le tirage est équilibré selon quatre variables. Il est effectué avec la macro cube. Les quatre variables sont :

- Le type d'adresse (mono-logement : un logement par adresse ; petite adresse : 2 à 10 logement par adresse ; grande adresse : plus de 10 logement par adresse²⁵)
- Le type de bâti (dur ou autre)
- Une variable de regroupement géographique
- La typologie sociale de village de l'Insee à partir du recensement 2012

De ce tirage résulte une quasi équipondération des logements. Tout logement tiré a une probabilité 0,0257 d'être tiré (soit un logement sur 39). La probabilité de tirage des logements est convertie en probabilité de tirage d'adresses.

Échantillon : Le nombre d'adresses et de logements tirés sont présentés dans le tableau 30.

²⁴ Pour disposer d'une base de sondage récente, la base cartographique est actualisée tous les ans sur un cinquième de la base, par groupe de rotation.

²⁵ Dans ce groupe, on effectue un tirage d'adresse proportionnellement à leur nombre, puis à un second degré de tirage, selon un sondage aléatoire simple, on tire 10 logements. Dans les petites adresses, on prend tous les logements.

Tableau 30 • Nombre d'adresses et de logements tirés par type d'adresse

Nombre...	Mono-logement	Petite adresse	Grande adresse
... d'adresses tirées	1 264	172	4
... de logements tirés	1 264	410	40

Poids de sondage :

Chaque logement est équipondéré. Chaque poids de tirage vaut : $1/0,0257=39$.

Protocole

L'enquête était collectée en face-à-face. Une fois entré dans le logement échantillonné l'enquêteur commençait l'enquête avec un adulte du ménage. Un tirage Kish déterminait la ou les personnes devant répondre au questionnaire individuel (dont le questionnaire santé). Lorsqu'il y a une personne dans le champ de l'enquête (15 ans ou plus), cette personne était tirée pour l'enquête ; quand il y avait au moins deux personnes dans le champ, deux personnes au plus étaient tirées pour le questionnaire individuel. Ce protocole avait été repris de l'enquête *Adult Education Survey* de l'Insee. Le tirage Kish était basé sur la date de naissance pour les quatre DROM historiques (les personnes ayant la date de naissance la plus proche du 1^{er} juillet mais située après le 1^{er} juillet ; on a fait l'hypothèse que les indicateurs de santé n'étaient pas liés au mois de naissance). La date du 1^{er} juillet a été choisie en référence avec le début de l'enquête qui commençait en juillet. Cette solution a été préférée à un Kish purement aléatoire afin de permettre aux enquêteurs d'anticiper éventuellement qui serait tiré dans le ménage. Si le premier contact avait lieu par téléphone, il était possible pour l'enquêteur de demander les dates de naissance de tout le ménage afin d'anticiper le tirage Kish et de fixer avec le ménage un rendez-vous au domicile à un moment où les deux personnes Kish étaient présentes de manière à éviter à l'enquêteur d'avoir à retourner deux fois sur place pour mener les deux interviews. Le tirage Kish était purement aléatoire à Mayotte. Certaines dates de naissance identiques étaient données à Mayotte aux personnes qui ne connaissaient pas leur date de naissance, donc un Kish sur la date de naissance était déconseillé. Il en était de même pour un tirage sur le prénom comme par exemple beaucoup de prénoms féminins commencent par la lettre « z ».

En début de questionnaire « individuel », le NIR était recueilli ainsi que les informations permettant de le reconstituer pour pouvoir apparier les données de l'enquête avec le Système national de données de santé. C'était nécessaire pour l'enquête DROM au contraire de l'enquête France métropolitaine où le NIR était reconstitué à partir de Fidéli car l'unité échantillonnée était le logement et non l'individu (cf. partie « appariements »).

Les questions les plus sensibles (santé mentale, alcool, tabac, drogues) étaient posées sous casque. Il s'agissait de la première utilisation du casque par les enquêteurs Insee de Mayotte (cf. partie « bilan du casque »). Contrairement à l'enquête Cadre de vie et Sécurité, pour l'enquête Santé, les enquêteurs avaient pour instructions de mener tout de même le recueil en face-à-face si le casque venait à être refusé par l'enquêté de manière à récupérer le questionnaire. Le casque était laissé au ménage comme cadeau.

Comme les tests avaient mis en avant la difficulté à recueillir de manière déclarative le poids et la taille à Mayotte, une balance et un télémètre ont été mis à disposition des enquêteurs de l'Insee afin qu'ils pèsent et mesurent les enquêtés. Des exemplaires de ce matériel sont parvenus à l'équipe Insee de Mayotte. Une réunion a été organisée avec l'équipe du service territorial de Mayotte dont des enquêteurs et une ergonome de l'Insee afin d'évaluer si ce matériel pouvait être utilisé et dans quelles conditions par les enquêteurs de l'Insee. Les enquêteurs ont par exemple demandé des lingettes pour pouvoir nettoyer le matériel. La demande d'utilisation du matériel pour cette enquête a été acceptée lors de la réunion du CHSCT spécial enquêteurs du 17 octobre 2018 suite à la rédaction d'une note de la direction régionale La Réunion-Mayotte justifiant cette demande de matériel et d'échanges entre la DREES, les équipes de l'Insee en charge du pilotage de cette enquête, une ergonome et le service territorial de Mayotte (dont des enquêteurs).

L'utilisation de ce matériel et du casque ont été à l'origine de la demande par le service territorial de Mayotte de l'ajout d'un test en janvier 2019 qui a été acceptée.

Bilan de collecte

La collecte a duré du 8 juillet au 31 décembre 2019. Initialement, elle devait se terminer le 14 décembre mais elle a été prolongée pour améliorer les taux de réponse alors que la montée en charge des réponses pendant le début de collecte, qui se déroulait pendant les vacances d'été, avait été assez lent.

Taux de réponse

Les taux de réponse ont été très bons, notamment à Mayotte et en Guadeloupe. Le taux de réponse à La Réunion a cependant été un peu décevant par rapport à d'autres enquêtes. Cela semble dû à des problèmes organisationnels à l'Insee La Réunion pendant la collecte.

Le taux de réponse « ménage » présenté dans le tableau 31 considère comme « répondant » un ménage dans lequel un seul des deux Kish aurait répondu. L'idéal serait de disposer d'un taux de réponse individuel pour cette enquête dont l'exploitation se fait à un niveau individuel. Cependant, ce calcul n'est pas aisé dans la mesure où l'on ne sait pas si un ménage « non-répondant » comportait un ou deux Kish. Le tableau 33 donne une approximation de ce taux en multipliant le taux de réponse « ménage » par le taux de réponse des « Kish » au sein d'un ménage considéré comme « répondant », car au moins un Kish a répondu (tableau 32).

Tableau 31 • Taux de réponse « ménage » par DROM

Nombre de « ménages »	Guadeloupe	Martinique	Guyane	La Réunion	Mayotte
Répondants	1303	1201	1344	1081	1224
Non répondants	177	305	403	403	160
Hors champ	230	162	288	91	378
Total	1710	1669	2035	1575	1762
Taux de réponse « ménage »	88,0 %	79,7 %	76,9 %	72,8 %	88,4 %

Lecture > En Guadeloupe 88,0 % des ménages sollicités ont eu au moins un répondant.

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee.

Tableau 32 • Taux de réponse Kish par DROM

Nombre de « Kish »	Guadeloupe	Martinique	Guyane	La Réunion	Mayotte
Répondants	1982	1817	2176	1763	2026
Non-répondants	67	56	94	84	34
Total	2049	1873	2270	1847	2060
Taux de réponse « Kish »	96,7 %	97,0 %	95,9 %	95,5 %	98,3 %

Lecture > Au sein des ménages considérés comme « répondants » car au moins un Kish a répondu à l'enquête, 96,7 % des Kish ont répondu en Guadeloupe.

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee.

Tableau 33 • Estimation du taux de réponse individu par DROM

Approximation du taux de réponse « individuel »	Guadeloupe	Martinique	Guyane	La Réunion	Mayotte
(taux de réponse « ménage » X taux de réponse « Kish »)	85,1 %	77,3 %	73,7 %	69,5 %	86,9 %

Lecture > Le taux de réponse « individu » estimé en Guadeloupe est 85,1 %, il est estimé comme le taux de réponse « ménages » fois le taux de réponse « Kish ».

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee.

Taux de proxy

On a voulu autoriser le recours au proxy de manière à pouvoir disposer des réponses de personnes très âgées ou à l'état de santé dégradé dans l'enquête afin d'améliorer la représentativité. Cependant, les enquêteurs de l'Insee ont l'habitude de recourir assez fortement au proxy lorsque c'est autorisé. C'est par exemple le cas dans l'enquête Emploi de l'Insee. De ce fait certainement, le taux de proxy s'est révélé très élevé en début de collecte (tableau 34), ce qui était gênant pour une thématique assez intime comme peut l'être la santé. Le taux de proxy étant une variable de suivi de collecte, des consignes ont été repassées rapidement après le début de collecte de manière à réduire ce taux de proxy. En formation, nous avons bien précisé que le proxy n'était autorisé que lorsque la personne n'était pas en mesure de répondre par elle-même à l'enquête, notamment en cas d'état de santé dégradé ou si la personne est très âgée. La borne de 2 % maximum de proxy avait été fixée aux divisions « Enquêtes Ménages » en formation. Le taux de proxy a donc pu être abaissé ; il est cependant resté élevé en Martinique.

Tableau 34 • Taux de proxy par DROM au fil de la collecte

Taux de proxy	Guadeloupe	Martinique	Guyane	La Réunion	Mayotte
À un mois de collecte	10,3 %	8,6 %	9,8 %	7,5 %	4,1 %
A deux mois de collecte	8,7 %	5,5 %	6,3 %	5,9 %	3,9 %
A quatre mois de collecte	3,7 %	6,7 %	4,2 %	3,7 %	3,4 %
En fin de collecte	2,9 %	7,4 %	3,0 %	2,0 %	2,5 %

Lecture > Le taux de proxy était de 10,3 % à un mois de collecte. Il est passé à 2,9 % en fin de collecte.

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee.

Matériel à Mayotte (balance et télémètre)

Le test papier à Mayotte mené en novembre 2017 avait mis en évidence, lors des accompagnements « enquêteurs » que beaucoup de Mahorais ne connaissaient pas leur taille et leur poids. L'usage d'une balance et d'un télémètre à Mayotte a donc été très important pour recueillir un poids et une taille à Mayotte. Les consignes données aux enquêteurs consistaient à leur dire qu'ils pouvaient recueillir une taille et un poids déclaratifs mais que l'usage du matériel était préféré. Au cas où les enquêtés ne connaissaient pas leurs mesures, alors les enquêteurs devaient utiliser le matériel de mesure.

Tableau 35 • Nombre de valeurs manquantes à l'issue de l'enquête concernant le poids et la taille

Nombre de valeurs manquantes	Guadeloupe	Martinique	Guyane	La Réunion	Mayotte
Taille	34 (1,7 %)	42 (2,3 %)	165 (7,6 %)	19 (1,1 %)	36 (1,8 %)
Poids	90 (4,5 %)	73 (4,0 %)	178 (8,2 %)	47 (2,7 %)	42 (2,1 %)

Lecture > A l'issue de la collecte, il y avait 34 valeurs manquantes pour la taille en Guadeloupe et 165 en Guyane.

Champ > DROM.

Source > EHIS 2019, DROM, DREES-Insee.

On constate avec le tableau 35 que le taux de non-réponse est élevé en Guyane, ce qui pourrait justifier l'usage de matériel de mesure. Par ailleurs, il y a beaucoup de valeurs arrondies (à la dizaine ou à 5) dans le recueil déclaratif qui aurait peut-être pu être précisées si les mesures avaient été prises.

Cependant, à Mayotte également, l'annexe 9 du dossier de la DREES de premiers résultats (Leduc et al., 2021) a mis en évidence que les valeurs étaient arrondies avec des pics dans les distributions de tailles et poids aux valeurs à la dizaine ou à 5. Cela peut être dû à plusieurs raisons : certains enquêteurs ont pu privilégier un recueil déclaratif, il n'était pas très simple d'utiliser le matériel (la balance devait être posée sur une surface plane, ce qui n'est pas forcément simple quand l'enquête se fait dehors ; les enquêteurs d'une taille plus petite que les enquêtés avaient logiquement des difficultés à utiliser le télémètre, ...). Cela a pu être redressé grâce aux données de l'enquête Unono wa maore de Santé publique France. Dans cette enquête les mesures avaient été recueillies par du personnel médical lors d'un examen de santé.

Ce travail de redressement aurait été certainement bien amélioré si on avait disposé d'une variable renseignant si la taille et le poids avaient été mesurés ou déclarés dans l'enquête.

Ce matériel semble avoir été très utile car sans lui, il est fort probable que l'on aurait manqué de beaucoup de données rendant difficile l'exploitation. Cependant, le constat des arrondis montre aussi qu'il n'est pas simple pour des enquêteurs d'utiliser ce matériel. Le poids et la taille étaient mesurés par du personnel médical lors de l'enquête Unono wa maore à Mayotte ce qui conduisait à des distributions bien plus lisses de tailles et poids. Avant redressements, la répartition par type de corpulences dans EHIS était plus éloignée de celle de l'enquête Unono wa Maore, ce qui peut s'expliquer par différentes raisons (le taux de réponse, les redressements, etc.) mais aussi par la qualité de l'information collectée.

Difficultés avec les deux Kish

Le plus souvent lorsqu'une enquête recourt à un tirage Kish, elle n'échantillonne qu'une personne par ménage avec cette méthode. L'idée de tirer deux Kish venait de l'enquête *Adult Education Survey* de l'Insee. L'objectif était d'économiser sur une partie des temps de déplacements enquêteurs. En compensation, pour pallier l'effet de grappe inhérent à la méthode, la taille d'échantillon a été augmentée par l'Insee d'environ 1 500 fiches adresse à 2 000 fiches adresse par DROM.

On n'a pas vraiment pu faire le bilan de savoir l'économie sur les déplacements et les durées d'enquête (il n'y avait qu'un tronc commun des ménages pour potentiellement deux enquêtés) compensait les 500 fiches adresse ajoutées.

Cependant, cette méthode a été critiquée par les enquêteurs qui disaient devoir souvent venir deux fois dans le ménage pour pouvoir obtenir les questionnaires des deux Kish. Lors de la refonte de l'enquête *Adult Education Survey* qui est devenue l'enquête Formation tout au long de la vie, cette stratégie de tirer deux Kish a été abandonnée suite à un bilan négatif. Le tirage porte sur un Kish par ménage. Par ailleurs, avoir un nombre différent de Kish par logement introduit de la dispersion des poids.

Bilan du casque

La partie de l'enquête passée sous casque correspondait aux modules considérés comme les plus sensibles qui étaient regroupés en fin de questionnaire : santé mentale (PHQ9), consommation d'alcool, de tabac et de drogues. Introduire le casque dans le recueil répondait à l'idée de favoriser des réponses sincères et donc d'améliorer la qualité du recueil sur des sujets qu'il pouvait être difficile d'aborder avec l'enquêteur ou devant son entourage. Le tableau 36 met en avant que beaucoup d'enquêtés ne sont pas isolés lorsqu'ils répondent à l'enquête. Même s'il s'agit d'une consigne donnée aux enquêteurs lors de la formation, d'autant plus que la santé peut être un sujet intime ; en pratique, ce n'est pas forcément facile à réaliser. Le casque rendait également plus acceptable auprès des enquêteurs l'introduction de certaines questions comme les questions relatives aux drogues ou la neuvième question du PHQ9 (« Au cours des deux dernières semaines, à quelle fréquence avez-vous été gêné(e) par les problèmes suivants ? Penser qu'il vaudrait mieux mourir ou vous faire mal d'une manière ou d'une autre. »). Cette question a en effet suscité des remarques des enquêteurs dans le bilan. Certains ont noté avoir dû faire face à des situations qu'ils n'auraient pas dû rencontrer.

Les consignes passées aux enquêteurs étaient de privilégier le casque donc de le proposer systématiquement et d'appuyer son utilisation. Dans ce cas, l'enquêteur branchait le casque à son ordinateur et passait l'ordinateur à l'enquêté en lui ayant présenté les touches qui pouvaient lui être utiles pour répondre, valider une question, passer une question, etc. En plus des bandes sons, le texte des questions s'affichait à l'écran dans la langue choisie (ce qui n'est pas toujours le cas lorsque les questions sont très sensibles comme pour l'enquête Cadre de vie et sécurité) dans le but d'améliorer la compréhension. Le casque était ensuite laissé en cadeau au ménage enquêté. En cas de refus d'utilisation du casque, l'enquête devait être terminée avec l'enquêteur en face-à-face comme c'était fait pour le reste de l'enquête.

Dans l'enquête, deux questions adressées à l'enquêteur permettent de connaître les conditions de la passation : la première juste avant la partie sous casque demande s'il y a d'autres personnes dans la pièce et la seconde après le module sous casque demande si la partie a été passée en utilisant ou non le casque. Il y a certainement eu des situations intermédiaires où l'enquêteur a aidé l'enquêté dans l'utilisation du casque ou de l'ordinateur ou a fini la partie commencée sous casque en face-à-face à la demande de l'enquêté, mais nous n'avons pas d'informations sur ces cas. Le tableau 36 met en avant qu'il est difficile d'isoler l'enquêté avec entre 23 % et 36 % des interviews qui se passent en présence d'autres personnes. Le casque a été largement diffusé, le taux d'utilisation le plus élevé étant atteint pour Mayotte où l'on s'inquiétait de la bonne acceptabilité du protocole comme il s'agissait de la première utilisation du casque par l'Insee à Mayotte. Le fait de ne pas savoir si le module s'est déroulé sous casque ou non concerne au maximum 3 % des interviews (La Réunion). En fait ce sont des personnes qui n'ont pas rempli du tout le module sous casque mais que l'on a gardé comme répondants à l'enquête car ils avaient bien renseigné le reste de l'enquête. Il est possible que sans l'interruption « casque » ils auraient fini l'enquête.

Tableau 36 • Description de la passation du module sensible

Part de la population...	Guadeloupe	Martinique	Guyane	La Réunion	Mayotte
... avec au moins une autre personne dans la pièce au moment de débiter le casque	23 %	34 %	32 %	36 %	34 %
... qui a répondu avec le casque	79 %	75 %	75 %	72 %	82 %

Note > Statistiques non pondérées.

Lecture > En Guadeloupe, pour 23 % des enquêtés il y avait au moins une autre personne dans la pièce en plus de l'enquêté et de l'enquêteur au moment d'aborder le module sous casque. Un total de 79 % des enquêtés en Guadeloupe ont répondu au module « sensible » sous casque.

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee.

Les personnes qui ont répondu avec le casque sont relativement plus jeunes, avec un niveau de vie plus élevé (notamment en Guadeloupe en Martinique et à Mayotte) et un niveau de diplôme plus élevé (tableau 37). Il y a peu d'écart entre les sexes si ce n'est à Mayotte où les femmes sont relativement plus nombreuses parmi les personnes qui ont répondu sans le casque, peut-être en lien avec le niveau d'étude. En lien avec l'âge et le niveau de vie, les personnes ayant répondu avec le casque

sont également en meilleure santé et moins fréquemment handicapées. On constate ainsi qu'il était important de laisser la possibilité de répondre en face-à-face avec l'enquêteur. Sans cela, une population en meilleure santé aurait été sélectionnée.

Tableau 37 • Caractéristiques comparées de ceux qui ont utilisé le casque et ceux qui ne l'ont pas utilisé

Casque ...	Guadeloupe		Martinique		Guyane		La Réunion		Mayotte	
	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non
Âge moyen	50	65	50	67	41	51	45	58	36	51
<25 ans	14 %	2 %	12 %	3 %	22 %	9 %	16 %	5 %	27 %	8 %
80 ans ou +	3 %	16 %	2 %	17 %	1 %	4 %	1 %	8 %	0 %	2 %
Femmes	59 %	59 %	58 %	54 %	56 %	59 %	54 %	56 %	59 %	68 %
Niveau de vie moyen (euros par mois)	1 659	1 151	1 767	1 156	1 565	1 493	1 614	1 517	682	366
1 ^{er} décile	8 %	17 %	8 %	15 %	4 %	8 %	9 %	9 %	9 %	15 %
10 ^e décile	11 %	6 %	13 %	4 %	12 %	12 %	11 %	7 %	10 %	3 %
Sans diplôme	35 %	70 %	27 %	73 %	34 %	58 %	33 %	63 %	63 %	94 %
Plus que le bac	39 %	13 %	43 %	11 %	36 %	24 %	38 %	17 %	23 %	3 %
Né à l'étranger	5 %	12 %	2 %	4 %	25 %	41 %	3 %	2 %	52 %	51 %
État de santé général : bon ou très bon	60 %	34 %	58 %	26 %	68 %	61 %	66 %	45 %	66 %	40 %
État de santé général : mauvais ou très mauvais	11 %	34 %	11 %	36 %	9 %	15 %	7 %	21 %	8 %	31 %
Maladie chronique	44 %	66 %	49 %	70 %	28 %	40 %	38 %	54 %	19 %	40 %
Handicapé	9 %	24 %	9 %	28 %	6 %	13 %	6 %	19 %	3 %	13 %

Note > Les statistiques ne sont pas pondérées.

Lecture > En Guadeloupe, parmi les personnes qui ont répondu avec le casque, 9 % étaient handicapées ; parmi celles qui ont répondu sans le casque 24 % étaient handicapées.

Champ > DROM.

Source > EHIS 2019, DROM, DREES-Insee

Traductions

Pour une passation par casque la plus large possible, les questions concernées ont été traduites en différentes langues. C'est la société Filogis qui s'est chargée de la traduction et de l'enregistrement des bandes sons. Les équipes de l'Insee dans les DROM (dont des enquêteurs) ont fait des remarques sur ces traductions et des corrections ont été apportées à la suite de cette relecture. Les traductions pour les langues concernant la Guyane se sont déroulées sans difficultés et étaient de qualité. En revanche, pour Mayotte, le prestataire a eu plus de mal à obtenir un résultat satisfaisant. Pour l'enquête utilisant le casque qui a suivi, l'enquête « Cadre de vie et sécurité », la traduction a été prise en charge directement par le service territorial de l'Insee à Mayotte. Pour l'enquête santé DOM, les traductions pour Mayotte ont été effectuées en shimaore et en kibushi. Pour « Cadre de vie et sécurité », seul le shimaore a été utilisé. Par ailleurs, dans le cadre de l'enquête santé DOM, le casque était utilisé pour la première fois par l'équipe Insee à Mayotte. En Guyane, les traductions concernaient le créole guyanais, le portugais brésilien et le bushinenge (une langue du fleuve). Des discussions ont eu lieu avec l'équipe Insee en Guyane sur

l'étendue des traductions. Certains enquêteurs ont souligné qu'il fallait ajouter d'autres langues (anglais, créole haïtien, espagnol, etc.). Il a finalement été arbitré de s'arrêter à trois langues en plus du français. Les textes des traductions sont disponibles sur la page internet de EHS sur le site de la DREES. Dans les autres DROM, aucune traduction n'a été proposée. Les enquêtes peuvent avoir lieu en créole mais les personnes parlent aussi français.

Les données permettent de tirer un bilan de l'usage de ces langues. En effet, la langue choisie a été conservée dans les données. En Guyane, 13,7 % des questionnaires qui se sont déroulés sous casque l'ont été via une traduction : 3,1 % en créole guyanais, 3,9 % en portugais brésilien et 6,6 % en bushinenge (tableau 38). À Mayotte, c'est le cas de 46,8 % des questionnaires. En très grande majorité c'est la traduction en shimaore qui a été utilisée (43,1 % des questionnaires). Le kibushi n'a concerné que 3,8 % des questionnaires renseignés sous casque.

Tableau 38 • Nombre de répondants selon la langue utilisée

Nombre de répondants...	Guyane				Mayotte		
	Créole Guyanais	Portugais brésilien	Bushinenge	Français	Shimaore	Kibushi	Français
... utilisant le casque	51	63	108	1403	717	63	885
... n'utilisant pas le casque	9	9	10	504	163	9	189

Note > Le module sous casque n'est pas renseigné pour 19 personnes ; elles ne sont pas incluses dans ce tableau.

Lecture > En Guyane, 51 questionnaires ont été réalisés avec le casque en utilisant les bandes son de la traduction en créole guyanais.

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee

Les personnes qui ont répondu dans une autre langue que le français ont des caractéristiques très différentes de celles qui ont répondu en français (tableau 39). Elles ont un niveau de vie beaucoup plus faible en moyenne et sont moins diplômées. Elles ont plus fréquemment un état de santé mauvais ou très mauvais. Les personnes ayant répondu en créole guyanais ont plus fréquemment une maladie chronique en lien avec leur âge en moyenne plus élevée. Les personnes n'ayant pas répondu en français se déclarent plus fréquemment handicapées que celles qui ont répondu en français sauf celles ayant répondu en bushinenge. En ce qui concerne les questions de la partie sous-casque, en Guyane, les personnes qui ne répondent pas en français sont beaucoup plus fréquemment repérées comme dépressives avec le PHQ9 que les personnes qui ont répondu en français. L'écart est plus faible à Mayotte sauf pour la dépression sévère parmi celles qui ont répondu en kibushi. La consommation de cannabis au cours de la vie est déclarée moins fréquemment par les personnes qui ont répondu au questionnaire avec le casque dans une autre langue que le français.

Tableau 39 • Caractéristiques socio-démographiques et de santé des personnes ayant répondu avec le casque selon la langue utilisée (Guyane et Mayotte)

Langue :	Guyane				Mayotte		
	Français	Créole guyanais	Portugais brésilien	Bus-hinenge	Français	Shimaore	Kibushi
Âge moyen	41	54	45	40	32	39	49
Femmes	56 %	48 %	70 %	60 %	53 %	65 %	70 %
Niveau de vie moyen (euros par mois)	1681	987	863	727	1013	293	488
1er décile	3 %	10 %	8 %	15 %	6 %	12 %	6 %
10ème décile	13 %	6 %	2 %	1 %	17 %	1 %	3 %

Aucun diplôme	27 %	80 %	82 %	81 %	36 %	95 %	92 %
Plus que le bac	41 %	10 %	10 %	3 %	41 %	2 %	2 %
État de santé général : mauvais ou très mauvais	7 %	30 %	17 %	14 %	4 %	12 %	14 %
Maladie chronique	28 %	50 %	27 %	22 %	16 %	22 %	22 %
Handicapé	6 %	12 %	12 %	5 %	2 %	4 %	5 %
Symptômes dépressifs	15 %	34 %	33 %	30 %	20 %	19 %	22 %
Symptômes dépressifs sévères	5 %	18 %	13 %	11 %	7 %	4 %	14 %
Fume tous les jours : oui	11 %	4 %	10 %	10 %	11 %	11 %	13 %
Cannabis (vie) : oui	18 %	12 %	8 %	16 %	8 %	2 %	5 %

Note > Les statistiques ne sont pas pondérées.

Lecture > En Guyane, 18 % des personnes ayant répondu avec le casque en français ont consommé du cannabis au cours de leur vie. C'est le cas de 12 % des personnes ayant répondu avec le casque en créole guyanais.

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee

Comme ces personnes qui ont utilisé les traductions du module sous casque pour répondre à cette partie du questionnaire sont particulières, on peut se demander quel aurait été l'impact de ne pas avoir traduit le module dans ces autres langues. Le tableau 40 présente les prévalences concernant les questions du module sous casque parmi l'ensemble des répondants, parmi les répondants qui ont répondu avec le casque et enfin parmi les répondants qui ont répondu avec le casque en langue française. Les statistiques sont pondérées et on se place dans la situation où aucune imputation ne serait venue compenser l'absence d'informations des personnes n'ayant pas souhaité ou pu utiliser le casque ainsi que des personnes n'ayant pas pu l'utiliser en français. On constate que les écarts sur les prévalences sont très faibles même à Mayotte où près de la moitié de la population a répondu en shimaore, ce qui peut remettre en question l'intérêt d'avoir procédé à ces traductions pour les langues rares d'autant que ça a représenté beaucoup de travail pour mettre en œuvre la prestation et la faire valider par les équipes d'enquêteurs de l'Insee parlant ces langues.

Tableau 40 • Prévalence du module sous casque sur les populations suivantes : population totale, population qui a répondu avec le casque et population qui a répondu avec le casque en français

	Guyane			Mayotte		
	Total	Casque	Casque en français	Total	Casque	Casque en français
Symptômes dépressifs	19 %	19 %	17 %	20 %	19 %	20 %
Symptômes dépressifs sévères	7 %	7 %	6 %	6 %	6 %	7 %
Fume tous les jours	10 %	12 %	12 %	11 %	12 %	11 %
Cannabis (vie)	16 %	18 %	18 %	6 %	7 %	9 %
Cocaïne (vie)	2 %	2 %	2 %	1 %	1 %	1 %

Note > Les statistiques sont pondérées, aucune imputation n'est faite.

Lecture > En Guyane, la part de la population qui présente des symptômes dépressifs est 19 %. Si on ne considère que les personnes qui ont répondu avec le casque, elle est de 19 % également. Si on ne considère que ceux qui ont répondu au casque en français, elle est de 17 %.

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee

Non-réponse partielle et module alcool

La non-réponse est un peu plus élevée avec le casque que sans le casque (tableau 41). Si l'on ajoute aux statistiques de non-réponse partielle relatives au casque les 126 personnes pour lesquelles il n'y a quasiment aucune réponse au module casque (on ne sait pas si le casque a été utilisé ou non car le module n'a pas été renseigné), le taux de non-réponse partielle d'un nombre de 4 questions ou plus passe de 1,2 % à 2,8 % avec le casque contre 0,3 % sans le casque.

Tableau 41 • Nombre de non-réponse partielle parmi les questions non filtrées du casque (hors module alcool) selon que le casque a été utilisé ou non

	0	1	2	3	4	4 ou plus
Casque	93,5 %	3,5 %	1,0 %	0,6 %	0,3 %	1,2 %
Sans casque	95,7 %	3,0 %	0,7 %	0,2 %	0,2 %	0,3 %

Note > Le comptage se fait sur les questions non filtrées du casque hors module alcool : 9 questions du PHQ9, 3 questions sur le tabac et une question sur la consommation de cannabis au cours de la vie. Les variables de consommation de cocaïne et de consommation de médicaments pour une recherche d'effet ne sont pas considérées dans cet indicateur car elles concernent très peu de personnes. Statistiques non pondérées.

Lecture > Parmi les personnes ayant répondu avec le casque, 93,5 % n'ont laissé aucune non-réponse partielle. C'est le cas de 95,7 % des personnes ayant répondu sans le casque.

Champ > DROM.

Source > EHIS 2019, DROM, DREES-Insee

Les variables relatives à l'alcool ont été considérées comme inexploitable par la DREES. Aucune statistique n'a été diffusée et les données ne seront pas mises à disposition. En effet, les réponses se sont avérées totalement incohérentes au sein du module et avec des prévalences très éloignées de celles dont on dispose par ailleurs sur la consommation de l'alcool dans les DROM. Les modalités relatives à la fréquence de consommation d'alcool avaient été séparées en autant de questions qu'il y a de modalités car la liste des modalités était trop longue pour être mémorisée par les enquêtés sans intervention de l'enquêteur, dans le cadre de l'usage du casque. Ainsi, on demandait d'abord à l'enquêté s'il buvait tous les jours et en cas de réponse « non », on lui demandait s'il buvait 5 à 6 fois par semaine et ainsi de suite jusqu'à la réponse « oui » à une fréquence énoncée. Cependant, à aucun moment on ne lui listait l'ensemble des modalités possibles en amont du questionnement. Cela a pu générer un mauvais positionnement des enquêtés parmi les modalités.

De ce fait, l'usage du casque dans le cadre de l'enquête Santé s'avère très mitigé. Le questionnement de ce module était sûrement trop complexe pour être passé sous casque. Il vaut mieux privilégier des questions avec une liste de modalités très restreinte pour l'usage du casque. Il n'a pas vraiment été possible d'analyser l'effet de mesure lié à l'usage du casque sur les réponses données puisque la grande majorité des répondants ont utilisé le casque et que ceux qui ne l'ont pas utilisé sont très sélectionnés du point de vue des caractéristiques socio-démographiques et de santé. Cependant, l'analyse des effets de mesure avec les données de l'enquête en France métropolitaine (cf. partie « effets de mode ») laisse penser que le recours à l'auto-administré favorise la déclaration du fait de boire tous les jours ou de penser qu'il vaudrait mieux mourir. Il est probable que l'on aurait observé le même effet dans les DROM avec le casque si l'on avait disposé d'un échantillon de contrôle.

Les enquêteurs de l'Insee n'ont pas relevé de difficultés particulières à l'usage du casque dans le bilan qu'ils ont dressé de l'enquête.

Correction de la non-réponse partielle

Le nombre de non-réponses partielles par question est présenté en annexe 2. Les questions pour lesquelles la non-réponse partielle dépasse 1,0 % des répondants concernent la première question de littératie en santé (LITER1) portant sur la capacité à remplir des formulaires médicaux, la date de la dernière mesure par un professionnel de santé du taux de cholestérol et de la glycémie (PA3 et PA4), la taille et le poids²⁶ (BM1 et BM2), la durée des trajets à pied (PE3), le temps passé assis ou allongé sans dormir (PE9), deux questions de l'échelle d'Oslo : le fait que l'entourage montre ou non de l'intérêt pour ce que l'on fait (SS2) et la difficulté avec laquelle on peut obtenir de l'aide des voisins (SS3) ainsi que le module sous casque du fait de l'usage du casque. Il est intéressant de constater que, mis à part le module sous casque et la question PE9, il s'agit des mêmes questions qu'en France métropolitaine.

La correction de la non-réponse partielle concernant les variables santé a été effectuée de la même manière pour l'ensemble des variables et comme pour la France métropolitaine. Elle a été imputée via un *hotdeck* séquentiel à l'aide du package VIM de R. Les classes considérées pour mener le *hotdeck* ont été définies en fonction du sexe et les observations ont été triées selon l'âge en tranches quinquennales puis le niveau de vie en continu. Toutes les variables manquantes pour une même personne ont été imputées à l'aide d'un même donneur pour respecter les corrélations entre les variables.

Des indicatrices indiquant qu'une imputation a été menée figurent dans les données si bien que les imputations peuvent être repérées et refaites pour les besoins d'études particulières.

²⁶ En fait, pour ces deux variables, on a considéré que lorsque l'une des deux était manquante, on imputait à la fois la taille et le poids. Sur les 582 valeurs imputées pour chacun de ces deux variables, il manquait le poids et la taille dans 153 cas, la taille uniquement dans 152 cas et le poids uniquement dans 277 cas. Par ailleurs, des redressements ont été effectués concernant le poids et la taille à Mayotte (cf.annexe 9 de Leduc et al., 2021).

Appariements

Système national des données de santé (SNDS)

Comme l'échantillonnage a porté sur des logements dans les DROM, il n'était pas possible de recourir exactement à la même méthode qu'en France métropolitaine pour réaliser l'appariement avec le SNDS. Plusieurs stratégies ont été mises en œuvre pour mener l'appariement.

D'une part, dans le but d'améliorer le modèle de correction de la non-réponse totale nécessaire au calcul de la pondération, les consommations de soins de toutes les personnes du logement pour lesquelles il était possible d'effectuer une reconstitution du NIR à partir de l'état civil issu de Fidéli ont été extraites. Il s'agissait des personnes présentes dans le logement au 1^{er} janvier 2018, millésime de la base de sondage. L'idée était de calculer des indicateurs de santé au niveau du ménage et de les inclure dans le modèle de correction de la non-réponse au niveau du logement. L'extraction a donc concerné les quatre DROM historiques dont les échantillons ont été tirés dans Fidéli (pas Mayotte). Les personnes concernées en étaient informées par la lettre avis et un droit d'opposition pouvait s'exercer. Personne ne s'est opposé à cet appariement. Le circuit de données a consisté en une sélection par le Criem des états civils issus de Fidéli des personnes habitants dans les logements échantillonnés, une transmission au service de la BRPP qui a reconstitué les NIR puis une transmission à la CNAM via la procédure Safe. Une table de passage a été créée entre l'identifiant anonyme de l'enquête et l'identifiant anonyme envoyé à la CNAM. Faute de temps, les travaux de repondération n'ont finalement pas inclus ces informations dans les modélisations permettant de calculer les pondérations mises à disposition dans le fichier DROM à la date de publication de ce document.

D'autre part, comme il y avait une incertitude à la fois sur le fait que les habitants du logement au 1^{er} janvier 2018 était toujours ceux à la date de l'enquête et éventuellement sur le fait de savoir lequel des habitants du logement était le répondant, le NIR et l'état civil ont été demandés dans l'enquête aux répondants. L'objectif d'appariement avec les données de l'Assurance maladie était rappelé dans le questionnaire au moment de récupérer ces informations. Ces variables de l'enquête ont été transmises par le service informatique de l'Insee au service de la BRPP. Au cas où un NIR était renseigné dans l'enquête, c'est le NIR qui a été transmis à la BRPP pour vérification. Dans le cas contraire, c'est l'état civil qui a été transmis pour reconstitution des NIR. Ceux-ci ont ensuite été envoyés à la CNAM pour extraction du SNDS. De même que pour le fichier issu de Fidéli, une table de passage entre identifiants anonymes a été créée.

Enfin, pour les répondants pour lesquels l'appariement n'a pas pu être mené faute de NIR ou d'état civil complet dans l'enquête, il a été recherché parmi l'extraction SNDS des habitants du logement au 1^{er} janvier 2018 s'il se trouvait une personne dont le sexe et l'année de naissance dans le SNDS étaient les mêmes que ceux recueillis dans l'enquête. Dans ce cas, on a considéré que cette personne qui habitait le logement au 1^{er} janvier 2018 et avait le même sexe et la même année de naissance que le répondant était le répondant. L'appariement a ainsi été complété pour 421 personnes de cette manière (tableau 42). On a vérifié que pour ces personnes, il n'y avait aucune autre personne dans le même logement au 1^{er} janvier 2018 avec le même sexe et la même année de naissance (dans le cas de couples homosexuels nés la même année par exemple).

Tableau 42 • Description de l'appariement avec le SNDS

Table	Effectifs	Commentaires
Nombre de NIR issus de la collecte du NIR dans le questionnaire envoyés pour vérification à la BRPP par l'informatique de l'Insee	7 000	On ne connaît pas le nombre de NIR validés par la BRPP
Nombre d'identifiants uniques dans la table reçue de la CNAM (dans IR_BEN_R)	6 554 (en comparant le sexe dans l'enquête et dans le SNDS, on constate un écart pour 24 personnes ; en faisant de même pour l'année de naissance, on constate un écart pour 73 personnes dont pour 30 une année d'écart entre les deux sources)	Il y en a en fait 6 563 mais trois identifiants correspondent à deux répondants de l'enquête en même temps (possiblement des jumeaux), ils sont supprimés et 6 ont en fait été considérés comme non répondants à l'enquête du fait d'une forte non-réponse partielle
Nombre d'états civils collectés dans l'enquête et transmis à la BRPP	2 822	Fichier complémentaire à celui des NIR collectés dans l'enquête : ce sont des personnes différentes

Nombre d'identifiants uniques dans la table reçue de la CNAM (dans IR_BEN_R)	934 (6 n'ont pas le même sexe dans l'enquête et dans le SNDS ; toutes les années de naissance sont similaires)	Il y en a en fait 941 mais deux identifiants correspondent chacun à deux répondants de l'enquête. Ils ont été supprimés. Par ailleurs, 5 sont considérés comme non-répondants à l'enquête
Total de personnes appariés à partir de la collecte de l'enquête (NIR ou état civil)	6 554 + 934 = 7 488 (donc 9 764-7 588=2 276 non appariés)	
Nombre d'identifiants uniques récupérés par l'appariement via les occupants au 1 ^{er} janvier 2018 du logement d'après Fidéli	421	Non présents dans le fichier des 7 488
Total de personnes appariés	7 488+421=7 909 pour 9 764 répondants (taux d'appariement 81 %)	

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee.

Les taux d'appariement diffèrent selon les DROM. Il est finalement plus élevé qu'en France métropolitaine dans les Antilles et à La Réunion, pas loin de 100 %. Le taux d'appariement de la Guyane est 83,3 % et 39,4 % pour Mayotte (tableau 43).

Tableau 43 • Taux d'appariement de l'enquête avec le SNDS par DROM

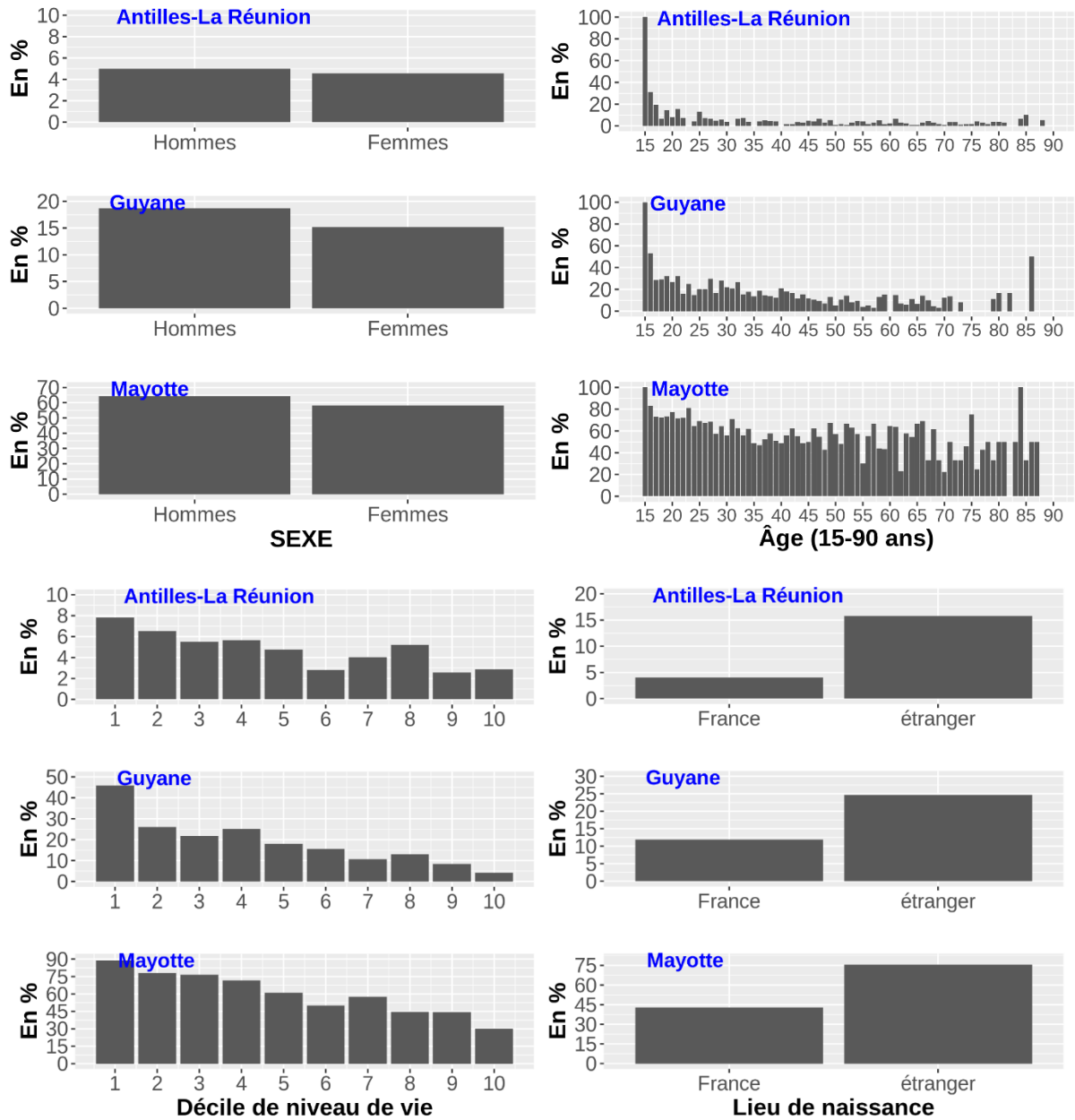
	Guadeloupe	Martinique	Guyane	La Réunion	Mayotte
Nombre d'appariés via collecte du NIR	1 600	1 500	1 379	1 290	785
Nombre d'appariés via collecte de l'état civil	177	187	233	324	13
Nombre d'appariés via les habitants du logement au 1 ^{er} janvier 2018 (Fidéli)	80	63	201	77	0 (la base de sondage n'était pas Fidéli)
Non appariés	125	67	363	72	1 228
Taux d'appariement	93,7 %	96,3 %	83,3 %	95,9 %	39,4 %

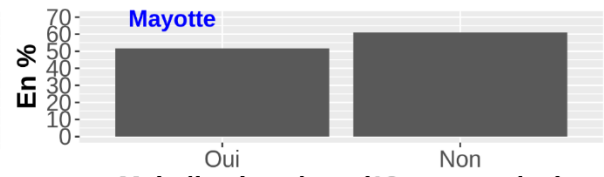
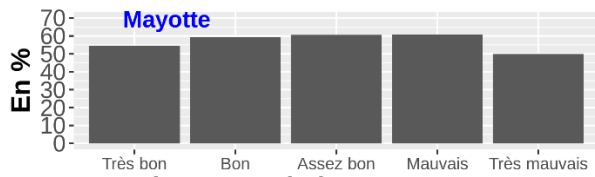
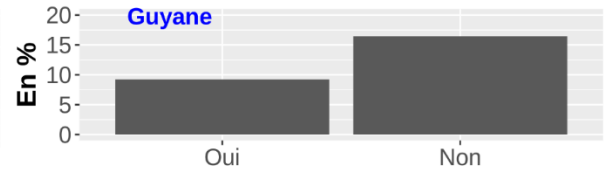
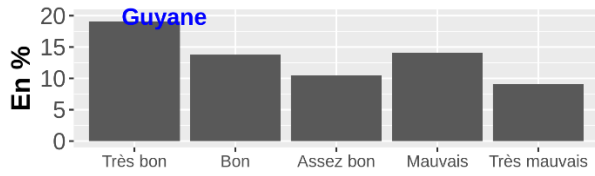
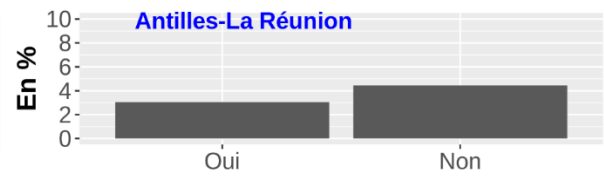
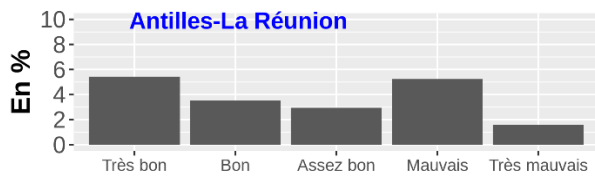
Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee.

La part de personnes non appariées par âge, niveau de vie et selon le mini module européen sur la santé sont précisés dans le graphique 13. On voit que quand le taux d'appariement est plus faible, des écarts socio-démographiques et en matière d'état de santé entre les appariés et les non appariés apparaissent. On a l'impression que le taux d'appariement est plus faible parmi ceux qui sont en meilleure santé sans doute en lien avec le taux d'appariement par âge. Les écarts de taux d'appariement par niveau de vie et par lieu de naissance sont très importants.

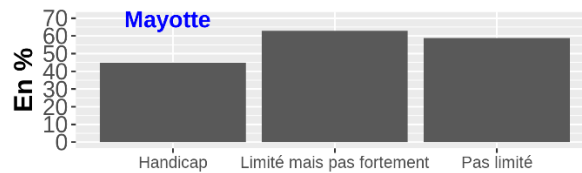
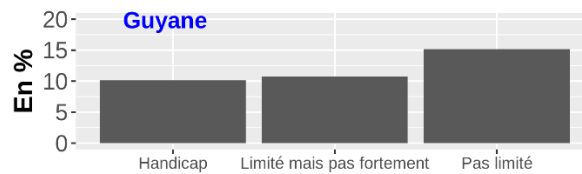
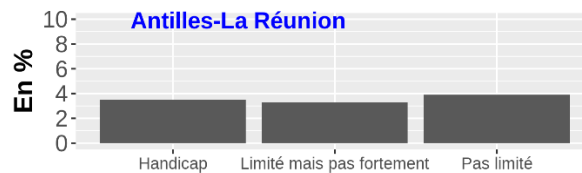
Graphique 13 • Taux de non-appariés selon le sexe, l'âge, le décile de niveau de vie, le lieu de naissance et les réponses au mini module européen sur la santé





Santé perçue générale (18 ans ou plus)

Maladie chronique (18 ans ou plus)



GALI (18 ans ou plus)

Lors du test de l'enquête CAPI de l'enquête qui s'est déroulée en Guyane (249 répondants) et à La Réunion (147 répondants), il était demandé aux enquêtés s'ils accepteraient de donner leur NIR, leur nom de naissance et leur commune de naissance. On n'avait pas d'autorisation pour récupérer ces informations lors du test mais on souhaitait évaluer l'acceptabilité de donner ces informations en vue d'anticiper la qualité de l'appariement avec le SNDS pour la vraie enquête. Il est intéressant de constater que lors de la vraie enquête, il y a une proportion, un peu plus importante de personnes qui ont effectivement donné ces informations comparativement à la proportion de celles qui avaient déclaré qu'elles les donneraient si on leur demandait (tableau 44).

Tableau 44 • Taux d'acceptation de donner ces informations si elles étaient demandées (test CAPI) et taux d'informations effectivement déclarées (enquête)

Information demandée	Guyane – test CAPI	Guyane - enquête	La Réunion – test CAPI	La Réunion - enquête
NIR	170 (68 %)	1 519 (70 %)	96 (65 %)	1 336 (76 %)
NIR ou nom	212 (85 %)	2 029 (93 %)	128 (87 %)	1 757 (100 %)
Commune de naissance	113 (45 %)	1 992 (92 %)	135 (92 %)	1 744 (99 %)
Effectifs totaux	249 (100 %)	2 176 (100 %)	147 (100 %)	1 763 (100 %)

Champ > DROM.

Source > EHS 2019, DROM, DREES-Insee.

Données fiscales et sociales

Les données des DROM seront appariées avec les sources socio-fiscales fin 2022 pour récupérer le revenu et les prestations sociales en 2019. Dans l'enquête DROM, contrairement à l'enquête France métropolitaine, le montant du revenu a été demandé dans l'enquête, comme il s'agissait d'une question préprogrammée du tronc commun des ménages.

■ CONCLUSION : MISE À DISPOSITION DES RÉSULTATS, DES DONNÉES ET DE LA DOCUMENTATION

Les premiers résultats de l'enquête sont parus dans le Dossier de la DREES n°78 (Leduc et al., 2021). Une mise à disposition des indicateurs Eurostat est disponible pour la France métropolitaine et chacun des DROM à cette adresse : <https://data.drees.solidarites-sante.gouv.fr/explore/dataset/indicateurs-ehis/information/>

Le site Internet de la DREES a une page dédiée²⁷ à l'enquête présentant la documentation européenne, les documents de collecte, les questionnaires, les traductions pour la Guyane et Mayotte, le kit de communication, les publications avec les données de l'enquête. Les notes complémentaires à ce document portant sur le calcul de la pondération en France métropolitaine, dans les DROM et sur la méthodologie de calcul de variance en France métropolitaine figurent également sur la page consacrée à EHIS sur le site Internet de la DREES. La fonction R de calcul de la variance pour la France métropolitaine est disponible sur demande auprès du CASD.

Les données sont disponibles au CASD²⁸ avec les questionnaires et dictionnaires des codes. Pour y accéder, il est nécessaire d'y être autorisé par les Archives de France après passage devant le Comité du secret²⁹. Comme il s'agit de données de santé, une autorisation auprès de la CNIL doit également être obtenue ou bien la méthodologie de référence MR004 peut être utilisée par les équipes de recherche.

Les données de l'appariement avec le SNDS seront mises à disposition au CASD.

²⁷ <https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/enquete-sante-europeenne-ehis>

²⁸ <https://www.casd.eu/source/european-health-interview-survey-enquete-sante-en-metropole-fichier-au-format-eurostat/>

²⁹ <https://www.comite-du-secret.fr/>

■ POUR EN SAVOIR PLUS

Abadie, A. et Imbens, G. W. (2011). Bias-Corrected Matching Estimators for Average Treatment Effects. *Journal of Educational and Behavioral Statistics*, 29.

Charrance, G., Cochet, P., Leduc, A., Bondon, M., Merly Alpa, T. (2022, mars). [Apprendre des paradonnées pour améliorer les protocoles de collecte : l'exemple d'EpiCov](#), Actes des Journées de méthodologie statistique.

Chevalier, M. et al., (2022, mai). [Le renouvellement de l'échantillon-maître des enquêtes auprès des ménages et de l'échantillon de l'enquête Emploi de l'Insee](#), Insee Méthodes, 141.

Cour des comptes (2014, juin). [La santé dans les outre-mer. une responsabilité de la République](#)

Dabet, G., Mazari, Z., Oujia, I. (2022, mars). [Estimation et décomposition de l'effet de mode dans les enquêtes multimodes \(Internet/téléphone\)](#), Actes des Journées de méthodologie statistique.

Dauphin, L et Eideliman J.-S. (2021, février). [Élargir les sources d'études quantitative de la population handicapée : Que vaut l'indicateur « GALI » ?](#). DREES, coll. Les Dossiers de la DREES, 74.

Eurostat (2018, mars). [European Health Interview Survey \(EHIS wave 3\). Methodological Manual](#).

Leduc, A., Deroyon, T., Rochereau, T., Renaud, A. (2021, avril). [Premiers résultats de l'enquête santé européenne \(EHIS\) 2019 – Métropole, Guadeloupe, Martinique, Guyane, La Réunion, Mayotte](#). DREES, coll. Les Dossiers de la DREES, 78.

Quantin, S. (2018). [Estimation avec le score de propension sous R](#). Insee. Document de travail, méthodologie statistique. 2018/01

Rosenzweig, A., Razafindranovona, T., Guedj, H. Quiviger, A., Zilloniz, S. (2022, mars). [Effets de mode relatifs aux victimations et aux perceptions en matière de sécurité : estimation à partir de l'enquête Genre et Sécurité \(GENESE\)](#), Actes des Journées de méthodologie statistique.

Annexe 1. Effets marginaux moyens des régressions logistiques relatives au fait de ne disposer d'aucun numéro de téléphone dans le fichier et d'un numéro de mobile pour la personne échantillonnée dans l'échantillon téléphone

	Probabilité de n'avoir aucun n° de tél. à disposition	Probabilité d'avoir un n° de mobile pour la pers. échant. (au sein de l'échantillon téléphone)
Constante	19,3 %	31,7 %
Femmes	Réf.	Réf.
Hommes	0.3	17.7***
Sexe inconnu	1.7	-33.8***
[15-24]	0.1	-32.6***
[25-34]	-2.3**	23.4***
[35-44]	-1.3	13.7***
[45-54]	Réf.	Réf.
[55-64]	3.2***	-13.3***
[65-74]	2.9**	-20.1***
[75-84]	4.3***	-32.3***
85+	3.2*	-39.5***
Pers. seule	3.7***	20.9***
Couple sans enf.	Réf.	Réf.
Couple avec enf.	-1.7*	-7.0***
Fam. monoparentale	1.3	14.3***
Autre situation familiale	6.6***	8.7***
décile de niveau de vie : d1	6.2***	-0.5
d2	4.4***	-2.3
d3	1.8	-1.8
d4	-0.4	-1.0

d5	Réf.	Réf.
d6	-3.9***	-1.1
d7	-5.3***	1.0
d8	-7.6***	0.3
d9	-7.5***	3.0*
d10	-8.1***	5.0***
Allocation chômage : oui	-0.2	5.3***
Allocation chômage : non	Réf.	Réf.
Minima sociaux : oui	3.6***	4.2***
Minima sociaux : non	Réf.	Réf.
QPV	3.8***	3.7***
Pas QPV	Réf.	Réf.
Pas de déménagement	Réf.	Réf.
Déménagement (même commune)	0.4	12.3***
Déménagement (commune différente)	25.2***	12.6***

Lecture > Pour la situation de référence, la probabilité de n'avoir aucun numéro de téléphone à disposition après enrichissement annuaire est 19,3 %. Avoir entre 75 et 84 ans plutôt qu'entre 45 et 54 ans augmente cette probabilité de 4,3 points. L'effet est significatif au seuil de 0,01 %.

*** : p-value < 0,01 % ; ** : p-value < 0,1 % ; * : p-value < 0,5 %.

Champ > France métropolitaine

Source > EHIS 2019, France métropolitaine, DREES-Irdes.

Annexe 2. Nombre de non-réponse partielle par question santé (France métropolitaine et DROM)

	France métropolitaine	DROM
Nombre d'observations :	14 192	9 764
Nombre de valeurs imputées pour la variable :		
HS1	11	11
HS2	17	15
HS3	17	5
CD2	36	12
CD1A	9	1
CD1B	5	7
CD1C	2	2
CD1D	14	11
CD1E	21	19
CD1F	1	3
CD1G	18	12
CD1H	7	9
CD1I	2	4
CD1J	12	23
CD1K	15	11
CD1L	12	14
CD1M	6	7
CD1N	12	16
CD1O	6	8
CD1P	55	52
CD1S1	Question non présente	2
CD1S2	Question non présente	6

CD1S3	Question non présente	12
CD1S4	Question non présente	4
AC1A	0	1
AC1B	0	1
AC1C	1	0
AC2	3	2
AW1	8	9
AW2	28	28
PL1	1	0
PL2	11	5
CATAR1	6	2
CATAR2	8	2
PL3	0	0
PL4	5	2
PL5	29	15
PL6	3	13
PL7	5	5
PL8	8	9
PL9A	4	6
PL9B	16	6
PC1A	6	1
PC1B	8	0
PC1C	6	1
PC1D	6	1
PC1E	9	1
PC2	11	3
PC3	15	5
HA1A	6	2
HA1B	7	0

HA1C	9	1
HA1D	10	1
HA1E	7	1
HA1F	7	2
HA1G	8	1
HA2	13	4
HA3	17	12
PN1	19	14
PN2	34	17
LITER1	248	151
LITER2	93	44
LITER3	106	23
LITER4	76	23
LITER5	46	24
AM10A	36	45
AM10B	Imputations non menées	Imputations non menées
HO1A	1	4
HO1B	11	10
HO2A	6	7
HO2B	34	30
AM1	12	26
AM2	9	19
AM3	30	36
AM4	16	22
AM5	23	20
DISP1	Question non présente	9
DISP2	Question non présente	14
EXT1	Question non présente	2
EXT2	Question non présente	3

AM6A	11	7
AM6B	4	4
AM7	0	1
MD1	2	2
MD2	5	3
MD3	Question non présente	1
PA1A	21	Question non présente
PA1B_M	97	Question non présente
PA1B_A	97	Question non présente
PA2	73	65
PA3	318	261
PA4	351	173
PA4_BIS	Question non présente	4
PA5A	40	51
PA5B	84	60
PA6A	24	29
PA6B	36	39
PA7A	34	11
PA7B	53	26
PA8A	50	26
PA8B	68	48
UN1A	24	10
UN1B	17	8
UN1C	Question non présente	Imputations non menées
UN2A	15	7
UN2B	39	19
UN2C	21	7
UN2D	18	12
UN2E1	29	20

UN2E2	15	18
BM1	165	582
BM2	165	582
PE1	47	16
PE2	78	94
PE3	166	125
PE4	43	24
PE5	47	27
PE6	33	24
PE7	66	25
PE8	31	12
PE9	96	118
DH1	21	36
DH2	63	79
DH3	29	20
DH4	115	52
DH5	56	25
DH6	31	10
POL1	Question non présente	61
POL2	Question non présente	61
SS1	62	32
SS2	169	162
SS3	301	260
IC1	52	8
IC2	54	8
IC3	94	42
GOULDURISQUE	121	Question non présente
CANTRIL	165	0
MH1A	191	339

MH1B	106	276
MH1C	108	234
MH1D	104	244
MH1E	129	248
MH1F	150	294
MH1G	113	268
MH1H	148	297
MH1I	134	340
SK1	57	185
SK2A	59	189
SK2B	61	245
SK3	58	193
SK4	122	250
SK4A (metro = SK4)	122	Question non présente
SK4B	Imputations non menées	Question non présente
SK4C	Imputations non menées	Question non présente
SK5	75	194
SK6	59	183
SK7A	Question non présente	24
SK7B	Question non présente	29
AL1_FR	91	Module inexploitable
AL1_BIS	92	Module inexploitable
AL2	120	Module inexploitable
AL3	134	Module inexploitable
AL4	116	Module inexploitable
AL5	134	Module inexploitable
AL6_FR	124	Module inexploitable
AL6_BIS	141	Module inexploitable
DRO1A	Question non présente	166

DRO1A_BIS	Question non présente	170
DRO1B	Question non présente	161
DRO1B_BIS	Question non présente	162
DRO1C	Question non présente	14
DRO1C_BIS	Question non présente	14
DRO1D	Question non présente	166
DRO1D_BIS	Question non présente	166
DRO2	Question non présente	170
DRO3	Question non présente	171
DRO4	Question non présente	171
DRO5	Question non présente	171

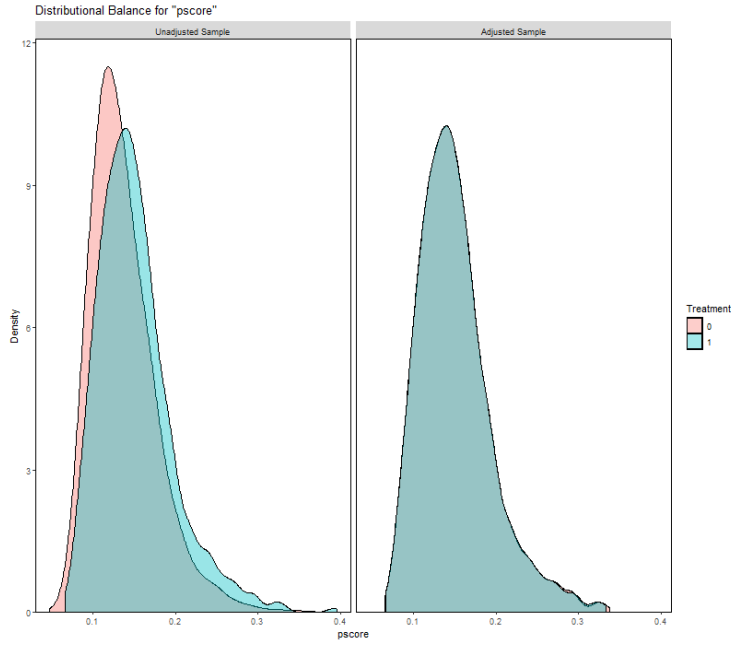
Lecture > A la question HS2, il y a 17 non-réponses en France métropolitaine et 15 dans les DROM.

Champ > France métropolitaine et DROM

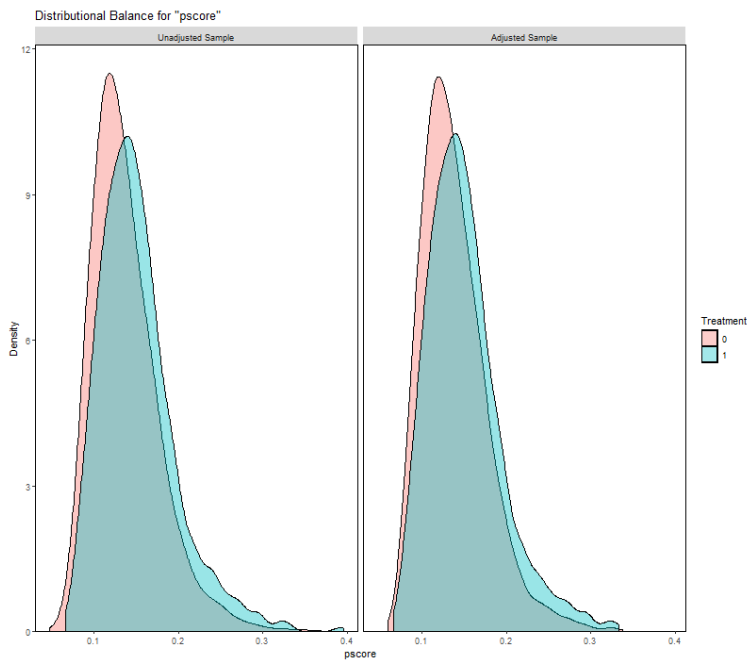
Source > EHIS 2019, France métropolitaine et DROM, DREES-Irdes-Insee.

Annexe 3. Distribution du score de propension entre échantillon méthodologique (face-à-face) et échantillon téléphone

Average treatment effect on the treated (ATT) :



Average treatment effect (ATE) :



DREES MÉTHODES

N° 7 • novembre 2022

Enquête santé européenne (EHIS) 2019 :
Bilan méthodologique

Directeur de la publication
Fabrice Lenglard

Responsable d'édition
Valérie Bauer-Eubriet

ISSN
2495-120X

Ministère des Solidarités et de la Santé
Direction de la recherche, des études, de l'évaluation et des statistiques (DREES)

14 avenue Duquesne - 75 350 Paris 07 SP
Retrouvez toutes nos publications sur drees.solidarites-sante.gouv.fr et nos données sur data.drees.solidarites-sante.gouv.fr
