

DREES MÉTHODES

N° 18 • septembre 2024

Le modèle de microsimulation Ines- Omar 2019

Note méthodologique détaillée

Pierre Poulon (DREES)

Le modèle de microsimulation Ines- Omar 2019

Note méthodologique détaillée

Pierre Poulon (DREES)

Remerciements : Catherine Pollak, Rosalinda Solotareff, Khadija Jabri, Aurore Desjonquères, Pierre-Louis Bithorel (DREES)

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

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

SOMMAIRE

■ INTRODUCTION	2
Omar 2019	2
Ines-Omar 2019	3
Évolutions entre les millésimes du modèle Ines-Omar	4
La prise en compte des DROM dans Ines-Omar	6
■ LES DÉPENSES DE SANTÉ	7
■ LES RETRAITEMENTS DE LA COMPLÉMENTAIRE SANTÉ.....	12
Attribution d'une couverture dans SRCV 2017	12
Redressement des déclarations de la CMU-C.....	13
EHIS 2014	13
SRCV 2017	14
EHIS 2019	14
Contrats d'entreprise ou individuels.....	16
Attribution de l'ACS.....	17
■ PSEUDO-APPARIEMENT EHIS 2014–SRCV 2017	18
■ ATTRIBUTION D'UN CONTRAT DE COMPLÉMENTAIRE SANTÉ	19
Retraitement du type de contrat dans l'enquête OC.....	19
Imputation des contrats	20
■ CALCUL DES COTISATIONS DE COMPLÉMENTAIRE SANTÉ	23
Tarification des contrats dans l'enquête OC	23
Retraitements des informations de cotisation dans l'enquête OC	23
Simulation fine des cotisations	24
Vérifications	25
■ ESTIMATION DES REMBOURSEMENTS DE COMPLÉMENTAIRE SANTÉ.....	28
Modes d'expression des garanties	28
Niveau de détail des simulations.....	29
Remboursements de la CMU-C et de l'ACS.....	29
Remboursement des équipements dentaires et optiques par la CMU-C.....	29
■ LE FINANCEMENT DE L'ASSURANCE MALADIE OBLIGATOIRE	32
Pseudo-appariement au modèle Ines 2019.....	32
Calages macro-économiques	34
Distribuer le financement de l'AMO	36
■ POUR EN SAVOIR PLUS	39
Annexe 1. Questionnaire couverture complémentaire d'EHIS 2019	40
Annexe 2. Modifications par rapport à Ines-Omar 2017.....	42
Annexe 3. Extraits du questionnaire de l'enquête OC 2019 sur les primes.....	44
Annexe 4. Comparaison avec l'enquête Budget de Famille 2017	46
Annexe 5. Pseudo-appariement par forêt aléatoire.....	50
Annexe 6. Pseudo-appariement par hot-deck.....	52

■ INTRODUCTION

Connaître la structure de financement des dépenses de santé au niveau microéconomique, et plus particulièrement les restes à charge supportés par les individus selon leur profil socio-démographique, permet de piloter plus finement les politiques de santé. Pour répondre à ce besoin de connaissances, la DREES a construit un *outil de microsimulation pour l'analyse des restes à charge* (Omar), qui simule le partage de la dépense de santé entre ses trois financeurs : l'assurance maladie publique, les complémentaires santé privées et les ménages. Recourir à la microsimulation est nécessaire car il n'existe pas de source administrative exhaustive concernant les restes à charge des ménages après remboursement des complémentaires santé^{1,2}. Le modèle de microsimulation Omar est rapproché du modèle Ines, qui simule la législation sociale et fiscale française, ce qui permet d'analyser la redistribution opérée par l'assurance maladie obligatoire (AMO) à la fois sous l'effet de ses financements et de ses prestations.

Ce document présente la méthode du millésime 2019 du modèle Ines-Omar, qui fait suite au millésime 2017 (Fouquet, 2021) et à celui de 2012 (Barlet *et al.*, 2019). Le modèle Ines-Omar est construit en deux temps ; tout d'abord la construction du modèle Omar, puis son appariement au modèle Ines pour constituer le modèle final Ines-Omar. Le schéma 1 résume ses différents éléments constitutifs et leur articulation.

Omar 2019

Le modèle Omar permet d'étudier au niveau individuel la répartition de la dépense de santé entre l'assurance maladie publique (y compris les dispositifs publics favorisant l'accès à la complémentaire santé des ménages modestes), les complémentaires santé privées et les ménages. Il fournit également des informations sur les cotisations dont doivent s'acquitter les souscripteurs à une complémentaire. Ces données peuvent être déclinées sur diverses sous-populations selon le type de couverture complémentaire, le niveau de vie, l'âge, le fait ou non de bénéficier du dispositif d'affectation de longue durée (ALD), l'état de santé déclaré, le statut d'activité, etc.

Le millésime 2019 du modèle Omar est issu du regroupement de plusieurs sources de données.

Le cœur de cette édition du modèle Omar est l'enquête santé européenne (EHIS) 2019³, appariée au système national des données de santé (SNDS). L'appariement EHIS 2019 – SNDS permet de disposer d'informations sur les caractéristiques sociodémographiques et la couverture complémentaire santé (type de contrat, cotisation déclarée) pour environ 12 700 individus de 15 ans ou plus, ainsi que de données médico-administratives exhaustives et détaillées sur la consommation de soins et les remboursements par l'AMO. Le type d'un contrat de complémentaire santé déclaré dans l'enquête subit des retraitements, pour qu'il soit cohérent avec les autres informations de l'enquête et avec le contenu du SNDS (voir partie III).

Néanmoins, EHIS 2019 ne fournit pas la couverture complémentaire d'un échantillon représentatif d'individus de moins de 15 ans. En effet, les seuls enfants dont la couverture complémentaire est connue sont ceux couverts par le même contrat que l'adulte interrogé par l'enquête. De ce fait, l'échantillon d'individus de moins de 15 ans du modèle Omar 2019 est constitué des individus de moins de 15 ans de l'enquête statistiques sur les ressources et les conditions de vie (SRCV) 2017 produite par l'Institut national de la statistique et des études économiques (Insee). Les deux enquêtes sources du modèle Omar 2019 (SRCV 2017 et EHIS 2019) sont traitées parallèlement et indépendamment jusqu'à l'appariement d'Omar avec le modèle Ines.

Certains retraitements spécifiques aux individus de moins de 15 ans sont nécessaires. Pour l'essentiel, les traitements appliqués à SRCV 2017 lors de la conception du modèle Omar 2019 sont identiques à ceux effectués pour le modèle Omar 2017 (voir Fouquet, 2021). Il s'agit notamment de retraiter le statut de couverture de l'ensemble des individus (voir partie III), et d'imputer des dépenses de santé fines aux individus de moins de 15 ans de SRCV 2017, comme l'enquête SRCV 2017 n'est pas appariée au SNDS. L'appariement EHIS-ESPS 2014 – SNDS, produit conjointement par l'Irdes et la DREES, fournit pour plus de 3 500 individus de moins de 15 ans des données individuelles identiques à celles présentes dans SRCV 2017 (variables socio-démographiques, état de santé perçu, couverture complémentaire), associées à des données exhaustives et très détaillées sur les dépenses de santé et les remboursements perçus au titre de l'AMO (voir partie II). Les individus de moins de 15 ans de SRCV 2017 sont

¹ Il existe une enquête auprès des ménages, l'enquête Budget de Famille 2017 de l'Insee, sur leurs restes à charge déclarés en santé. La comparaison de ces sources est discutée dans l'annexe 4.

² L'article L1461-1 du code de la santé publique - Légifrance ([legifrance.gouv.fr](https://www.legifrance.gouv.fr)) prévoit que le système national des données de santé (SNDS) puisse rassembler « un échantillon représentatif des données de remboursement par bénéficiaire transmises par des organismes d'assurance maladie complémentaire », mais sa mise en œuvre se heurte jusqu'à présent à l'impossibilité, de la part des organismes complémentaires, de mettre à disposition un tel échantillon.

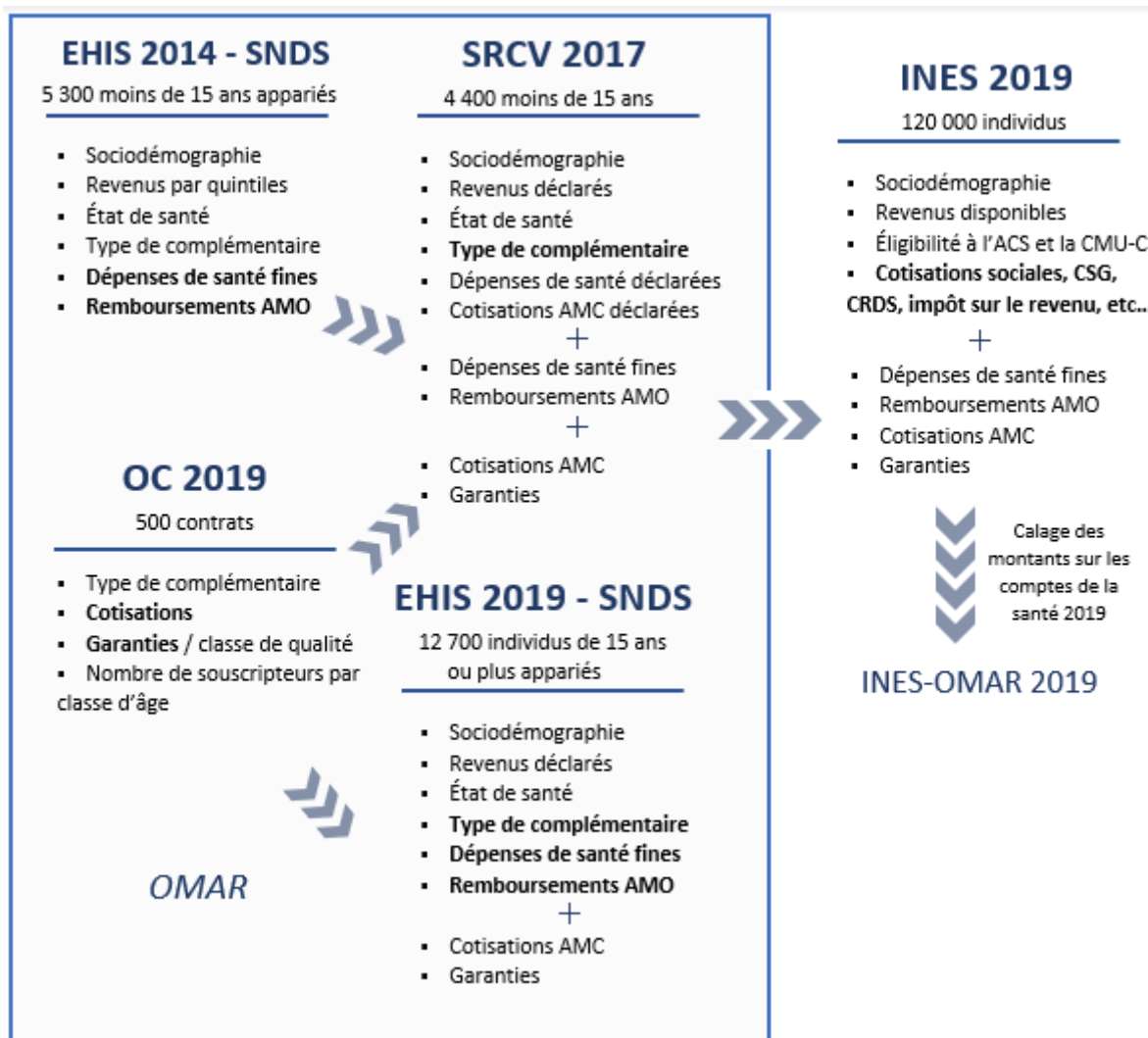
³ L'enquête santé européenne a remplacé l'enquête santé et protection sociale (ESPS), menée par l'Irdes, à partir de l'édition 2014.

ainsi pseudo-appariés à des individus de moins de 15 ans d'EHIS 2014⁴ afin que des données de santé leur soient imputées (voir partie IV).

L'enquête de la DREES auprès des organismes complémentaires offrant une couverture complémentaire santé (enquête OC) 2019 donne de nombreux éléments sur les contrats individuels et collectifs : nombre de souscripteurs par classe d'âge ; cotisations en fonction de l'âge, du nombre de bénéficiaires, du revenu et du lieu de résidence ; niveaux de garantie pour chaque poste de soin. Un contrat tiré de l'enquête OC 2019 est attribué à chaque ouvrant droit de contrat présent dans EHIS 2019 ou dans SRCV 2017 (voir partie V) et sa cotisation est recalculée finement (voir partie VI), ainsi que les remboursements complémentaires perçus par lui et ses ayants droit (voir partie VII).

De l'ajout des individus de moins de 15 ans de SRCV 2017 aux individus de 15 ans ou plus de EHIS 2019 résulte le modèle Omar 2019, qui contient 17 000 individus représentatifs de l'ensemble des ménages ordinaires de France métropolitaine.

Schéma 1 Schéma des éléments constitutifs du modèle Ines-Omar 2019



Ines-Omar 2019

Le modèle Ines, développé conjointement par l'Insee, la DREES et la Caisse nationale des allocations familiales (CNAF), applique la législation sociofiscale à un échantillon représentatif de la population vivant en France métropolitaine en logement ordinaire (ne couvrant donc pas les personnes vivant en institution, comme les étudiants vivant en résidence étudiante ou les personnes âgées ou dépendantes vivant en établissement) (Fredon, Sicsic 2020). Il utilise l'enquête revenus fiscaux et sociaux (ERFS) pour connaître précisément l'ensemble des revenus perçus par près de 120 000 individus et simule les prélèvements sociaux et fiscaux dont ils doivent s'acquitter. Un

⁴ Les individus de moins de 15 ans d'EHIS 2019 ne sont pas appariés au SNDS.

module complémentaire simule les formes de taxation indirecte⁵ : TVA, taxes sur l'alcool et le tabac (André, et al., 2016).

À chaque individu du modèle Ines en France métropolitaine sont attribuées les données d'un individu d'Omar aux caractéristiques sociales et démographiques proches (voir partie VIII). Le modèle résultant de ce pseudo-appariement est dénommé Ines-Omar. Étant donné que la construction du modèle consiste en l'imputation de données issues d'Omar aux individus d'Ines, et non l'inverse, les principales caractéristiques socio-démographiques des individus du modèle Ines-Omar 2019 sont celles du modèle Ines, notamment l'âge, le sexe, l'occupation et les revenus.

Puis les montants de dépenses de santé sont calés sur le niveau macro-économique de dépense établi par l'édition 2022 des comptes nationaux de la santé (CNS) pour l'année 2019, en se restreignant aux dépenses individualisables et remboursables par l'AMO en France entière⁶, hors dépenses exécutées en milieu médico-social. Il n'inclut pas les indemnités journalières ni les rentes pour accidents du travail et maladies professionnelles. Le calage ne modifie pas la répartition de la dépense entre le reste à charge des ménages, les parts remboursées par l'AMO et par l'assurance maladie complémentaire (AMC), estimée par le modèle (voir partie VII).

Par ailleurs, la connaissance des prélèvements obligatoires au niveau individuel permet d'attribuer aux individus du modèle Ines une contribution au financement de l'AMO (voir partie VIII). Le financement total de l'AMO par type de prélèvement est calé sur les montants des Comptes de la Sécurité sociale (Direction de la Sécurité sociale [DSS], 2020).

À la suite de ces calages, il devient possible de calculer le poids de l'ensemble des dépenses de santé, y compris les contributions au financement de l'AMO, dans le revenu des ménages. En effet, comme des ménages entiers sont observés dans le modèle Ines, il est possible de mener des analyses au niveau du ménage avec le modèle Ines-Omar, et non plus seulement à l'échelle de l'individu comme avec le modèle Omar.

In fine, le modèle Ines-Omar 2019 comprend près de 120 000 individus, auxquels ont été attribuées les dépenses de santé des un peu plus de 17 000 individus d'Omar 2019 (tableau 1). Comme le nombre d'individus dont proviennent les dépenses de santé est relativement modeste, il convient de ne pas produire de statistiques sur les dépenses de santé sur des sous-populations trop spécifiques d'Ines-Omar 2019 (par exemple pour les personnes âgées du premier dixième de niveau de vie, qui sont relativement peu nombreuses).

Tableau 1 Tableau récapitulatif de la taille des échantillons utilisés dans le modèle Ines-Omar 2019

	Sources utilisées pour le modèle Omar pour les moins de 15 ans		Sources utilisées pour le modèle Omar pour les 15 ans ou plus	Source Ines sur laquelle est basé le modèle Ines-Omar
	EHIS-ESPS 2014	SRCV 2017	EHIS 2019	Ines 2019
Taille initiale	26 514 (dont 5 304 de moins de 15 ans)	25 321 (dont 4 442 de moins de 15 ans)	14 192 (de 15 ans ou plus)	118 419
Taille après appariement au SNDS	16 757	/	12 745	/
Taille après appariement au SNDS et retraitements	16 496 (dont 3 552 de moins de 15 ans)	24 280 (dont 4 386 de moins de 15 ans)	12 685	118 408

Lecture > L'enquête EHIS 2019 comprend 14 192 individus répondants, tous de 15 ans ou plus, dont 12 685 sont appariés au SNDS et sont conservés à l'issue des retraitements effectués.

Champ > Enquêtes menées auprès de la population vivant en ménage ordinaire ; seule la population du volet France métropolitaine de ces enquêtes est donnée ici.

Évolutions entre les millésimes du modèle Ines-Omar

Des changements méthodologiques importants sont mis en œuvre entre chaque édition, grâce à l'enrichissement des sources à disposition et à la sophistication des méthodes d'imputation. Ces changements méthodologiques

⁵ Le module taxation indirecte reposait sur l'enquête « Budget de Famille » (BdF) 2011 lors de sa création en 2016. Il a depuis été mis à jour à l'aide de l'enquête BdF 2017.

⁶ À ce jour, le modèle Ines n'est pas construit sur les DROM (voir partie ci-après), menant à ne pouvoir élargir Ines-Omar au champ France entière qu'au moment du calage. Cela repose sur l'hypothèse simplificatrice que les dépenses de santé, remboursements et restes à charge moyens sont semblables dans les DROM et en France métropolitaine. Du fait de cette hypothèse forte, le modèle ne peut pas être utilisé pour des analyses uniquement sur les DROM.

rendent impossibles des études en évolution temporelle ainsi que la comparaison des résultats entre publications basées sur des millésimes différents du modèle Ines-Omar.

Par exemple, depuis le millésime 2017 du modèle, les dépenses hospitalières contiennent non plus seulement les dépenses en médecine, chirurgie, obstétrique et odontologie (MCO), mais également la psychiatrie (PSY), les soins de suite et de réadaptation (SSR) et l'hospitalisation à domicile (HAD) [Fouquet, 2021]. Le lien entre le niveau de vie et les garanties et primes des contrats est également plus fiable depuis 2017, du fait de l'imputation d'un contrat (et donc de ses garanties et primes) en tenant compte des primes déclarées dans l'enquête, ainsi que du type de contrat (individuel ou d'entreprise, contrat spécifique ou non à certaines catégories professionnelles).

Le millésime 2019 présente également plusieurs améliorations méthodologiques par rapport au millésime 2017 du modèle Ines-Omar, essentiellement en lien avec les sources utilisées (voir également annexe 2) :

- Le modèle Omar 2019 est de meilleure qualité que le précédent sur les dépenses de santé des adultes, qui sont les dépenses réelles (issues du SNDS) des individus de 15 ans et plus enquêtés d'EHIS 2019, et non plus imputées indirectement. Pour les moins de 15 ans, la qualité ne change pas, puisqu'on utilise toujours l'enquête SRCV 2017 pseudo-appariée à EHIS 2014 afin que des données de santé leur soient imputées. Il n'en demeure pas moins que les dépenses de santé présentes dans le modèle Ines-Omar, après appariement d'Ines avec Omar, sont toujours imputées.
- La cohérence entre années des sources est également améliorée : Ines-Omar 2019 utilise les millésimes 2019 de l'enquête OC et du modèle Ines, tandis qu'Ines-Omar 2017 utilisait l'enquête OC de 2016 et Ines 2017, ainsi que la structure de la consommation de soins et des dépenses de santé de l'année 2014.
- La représentativité des contrats de complémentaire santé est également meilleure : à la suite d'un changement de méthodologie lors du millésime 2019 de l'enquête OC, cette enquête recueille les caractéristiques d'un échantillon représentatif des dix contrats les plus fréquemment souscrits de chaque type (individuel / collectif) des organismes enquêtés, au lieu des trois plus fréquents auparavant.
- Le repérage des bénéficiaires de la CMU-C, de l'ACS et des patients en ALD s'est amélioré. L'appariement d'EHIS 2019 avec le SNDS permet de repérer dans l'enquête les patients en ALD, les bénéficiaires de l'ACS et une part importante des bénéficiaires de la CMU-C. Ainsi, il est possible de redresser les déclarations de CMU-C dans EHIS 2019⁷, et il n'est plus nécessaire d'y tirer aléatoirement les bénéficiaires de l'ACS, comme cela était fait dans SRCV 2017 pour le modèle Ines-Omar 2017.
- Une nouvelle méthode d'appariement entre Omar et Ines a été introduite. La méthode utilisée dans Ines-Omar 2017 comme 2019 est un *hot deck* ; elle consiste à stratifier sur la même manière les deux bases appariées (Ines et Omar) et à associer chaque individu d'Ines à un individu appartenant à la strate correspondante dans Omar. Dans un *hot-deck*, les strates doivent avoir une taille minimale, si bien que les strates de taille jugée insuffisante sont fusionnées. Les méthodes utilisées pour Ines-Omar 2017 et 2019 diffèrent par la manière dont les strates de taille insuffisante sont fusionnées (voir annexe 6).

En revanche, le modèle Ines-Omar 2019 n'intègre pas les changements réglementaires suivants intervenus en 2019 :

- Ines-Omar 2019 comporte le bénéfice de la CMU-C ou de l'ACS plutôt que celui de la complémentaire santé solidaire (CSS) gratuite ou contributive, qui a été introduite en remplacement de la CMU-C et de l'ACS le 1^{er} novembre 2019. En effet, l'enquête EHIS 2019 interroge sur le bénéfice de la CMU-C, mais pas sur celui de la CSS, et le SNDS 2019 comporte encore une information sur le bénéfice de la CMU-C et sur l'ACS. Ce choix méthodologique ne biaise pas les résultats du modèle en ce qui concerne les bénéficiaires de la CMU-C, tous passés à la CSS au 1^{er} novembre 2019, comme les garanties associées à la CSS lors de son entrée en vigueur sont très proches de celles auparavant associées à la CMU-C. En revanche, le modèle comporte un biais sur certains bénéficiaires de l'ACS qui ont choisi de passer à la CSS contributive : en effet, les classes B et C de contrats ACS étaient plus couvrantes que la CMU-C et la CSS, si bien que le modèle surestime les remboursements supplémentaires que perçoivent leurs bénéficiaires après leur passage de l'ACS à la CSS contributive.
- Ines-Omar 2019 n'intègre pas les effets de la réforme du 100 % santé en optique, prothèses dentaires et auditives. En effet, en 2019, seule une partie de la réforme était entrée en vigueur (modification du prix limite de vente et de la base de remboursement des aides auditives par exemple), la mise en place de la réforme s'étant étalée jusqu'en 2021 (voir par exemple Lapinte *et al.*, 2024). Cependant, plusieurs paniers d'équipements avec un reste à charge nul et réservés aux bénéficiaires de la CMU-C ont été mis en place puis enrichis dans le courant de 2019, dont les caractéristiques ont été appliquées pour calculer le remboursement des équipements achetés par les bénéficiaires de la CMU-C dans Ines-Omar 2019.

⁷ Dans SRCV 2017 comme dans EHIS 2019 un biais de sur-déclaration du bénéfice de la CMU-C parmi les personnes de plus de 60 ans est observé : ce biais a été redressé dans EHIS 2019 pour le modèle Ines-Omar 2019, alors qu'il ne l'était pas dans SRCV 2017 pour Ines-Omar 2017.

La prise en compte des DROM dans Ines-Omar

Le modèle Ines-Omar 2019, comme les millésimes précédents, repose sur des sources de France métropolitaine qui ne couvrent pas les DROM, et l'élargissement du champ à la France entière se fait seulement au moment du calage des dépenses de santé. En effet, il n'a pas été possible de constituer un échantillon représentatif de la population vivant en outre-mer dont la complémentaire santé ainsi que les dépenses de santé seraient connues.

Pourtant, l'enquête EHIS 2019 a également été menée dans les DROM, Mayotte comprise, et l'échantillon ultramarin a été apparié au SNDS (Leduc *et al.*, 2021 ; Leduc, 2022), mais il ne couvre pas les individus de moins de 15 ans. Il aurait fallu donc s'appuyer sur d'autres sources pour connaître à la fois leur couverture complémentaire et leurs dépenses de santé, sachant que l'enquête SRCV 2017 n'a pas été menée en outre-mer. L'enquête Budget de Famille (BdF) 2017, qui comporte un module de questions sur la complémentaire santé, a quant à elle été menée en outre-mer, mais n'a pas été appariée au SNDS. L'enquête EHIS-ESPS 2014, appariée au SNDS, n'a pas été menée en outre-mer, et ne pouvait donc pas servir à l'imputation de dépenses de santé aux individus de moins de 15 ans de l'enquête BdF 2017.

Enfin, le modèle Ines ne couvre à ce stade que le champ de la France métropolitaine.

L'extension aux DROM (hors Mayotte) à compter de l'année 2022 de l'enquête SRCV, sur laquelle pourrait être construite un prochain millésime du modèle Ines-Omar, ouvre toutefois des perspectives d'amélioration de la couverture des territoires ultra-marins à l'avenir.

■ LES DÉPENSES DE SANTÉ

Les dépenses de santé du modèle Ines-Omar 2019 sont issues de l'appariement d'EHIS 2019 au SNDS pour les 15 ans ou plus, et de l'appariement d'EHIS 2014 au SNDS pour les moins de 15 ans. Ces deux appariements couvrent le même champ : les dépenses de santé présentées au remboursement, individualisables et remboursables en ville et à l'hôpital, hors médico-social. L'appariement EHIS – SNDS 2019 utilise plus d'information que l'appariement EHIS – SNDS 2014 quant à la reconnaissance d'affections de longue durée (ALD) des individus et à leur bénéfice de l'ACS.

Le SNDS contient des informations sur les soins remboursés en ville ou en établissements de santé. Il est possible d'y calculer le montant des dépenses de santé, la part remboursée par l'AMO, la part de liberté tarifaire ou encore les participations forfaitaires acquittées. L'appariement exhaustif des données sur les dépenses de santé en ville et en hôpital privé, issues du DCIR (datamart de consommation inter-régime) et des dépenses de santé en hôpital public, issues du PMSI (programme de médicalisation des systèmes d'information), dans le cadre de la base RAC (reste à charge) a permis de fiabiliser le calcul de ces indicateurs depuis le millésime 2017 d'Ines-Omar (Adjerad, Courtejoie, 2021). Le champ retenu pour les postes de soins est le plus exhaustif possible. Il couvre les dépenses remboursables et présentées au remboursement (qu'il y ait eu avance de frais ou non), réalisées en ville, au cours d'un séjour à l'hôpital public ou en établissement de santé privé, en médecine, chirurgie et obstétrique (MCO), en soins de suite et de réadaptation (SSR), en psychiatrie (PSY) et en hospitalisation à domicile (HAD).

Comme le précise l'encadré méthodologique décrivant la base RAC dans Adjerad, Courtejoie (2021), « *les dépenses non directement individualisables sont exclues de l'analyse. Il s'agit notamment de celles associées à la rémunération des médecins libéraux sur objectifs de santé publique, qui ne peuvent pas être directement affectées à un bénéficiaire spécifique. Pour les patients du champ médico-social, qui incluent notamment les personnes handicapées ou âgées en établissements spécialisés, les dépenses incluses dans les forfaits de soins de ces établissements médico-sociaux ne sont pas prises en compte, car elles ne sont que partiellement renseignées dans le SNDS.* » Enfin, on ne tient pas compte des dépenses non prises en charge par l'AMO, parce qu'elles ne figurent pas de manière exhaustive dans le SNDS, telles que les suppléments pour confort personnel. Cela conduit ainsi, dans la base RAC comme dans les versions 2017 et 2019 du modèle Ines-Omar, au champ des dépenses de santé présentées au remboursement, individualisables et remboursables en ville et à l'hôpital, hors médico-social.

Par ailleurs, pour les séjours hospitaliers, il est nécessaire de recalculer certains montants à partir des informations disponibles, comme précisé dans les fiches de la documentation collaborative du SNDS⁸. Ainsi, le taux de remboursement est recalculé en fonction de la nature de l'assurance (régime de base, régime Alsace-Moselle, maternité, accident du travail, prévention maladie) et de l'existence éventuelle d'une exonération, tandis que le forfait journalier est recalculé selon la durée du séjour, la nature complète ou partielle de l'hospitalisation, le mode de sortie et les éventuelles exonérations. Un travail a été effectué afin d'intégrer les dépenses associées aux médicaments et dispositifs de la liste en sus, ainsi que les actes et consultations externes.

Enfin, le regroupement des prestations de soins en postes de soins plus généraux nécessite un travail important de croisement des différentes tables du SNDS et d'appariement aux nomenclatures idoines. Seul ce travail fin permet de distinguer les différentes formes de pharmacie, de séparer les prothèses dentaires de l'orthodontie et des actes de conservation dentaire, ou encore l'optique du matériel médical et des audioprothèses.

L'appariement EHIS-SNDS 2014

Le modèle Ines-Omar 2019 utilise les individus de moins de 15 ans de l'appariement entre EHIS 2014 et le SNDS. La jointure EHIS-SNDS 2014 est possible pour 16 757 individus de l'enquête sur 26 514, dont 3 595 de moins de 15 ans sur 5 304. Un recalage des poids individuels selon l'âge, le sexe et le niveau de vie est effectué pour que le sous-échantillon des moins de 15 ans appariés demeure représentatif de la population de France métropolitaine de moins de 15 ans.

Les programmes développés dans le cadre de la base RAC 2017 sont repris et exécutés ici sur l'échantillon EHIS-SNDS 2014. Une documentation plus extensive, ainsi que les programmes de la base RAC sont disponibles sur le Gitlab public de la DREES⁹.

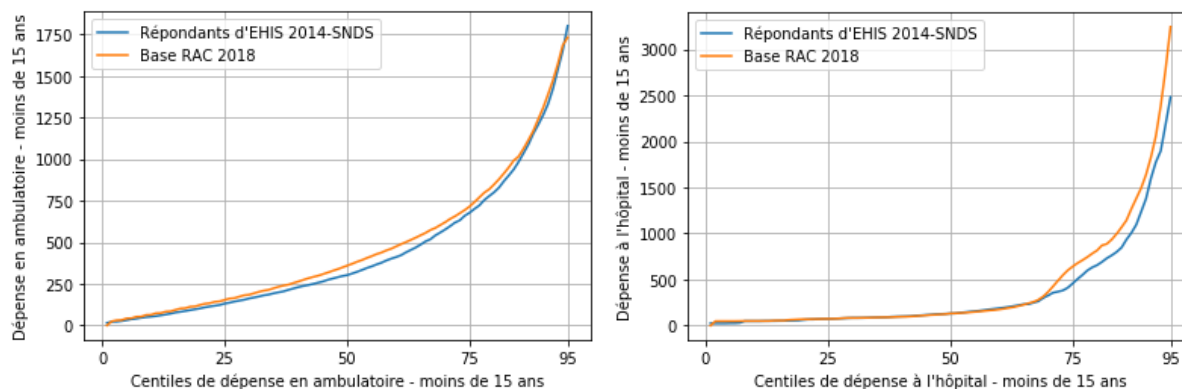
Pour les individus de moins de 15 ans, la distribution de la dépense de santé totale par centiles a un profil très proche de celle obtenue sur le SNDS exhaustif, en particulier en ambulatoire, malgré la faible dimension de la base EHIS – SNDS 2014 (*graphique 1*). Néanmoins, la hausse des dépenses de santé entre 2014 et 2017 ne suffit pas à expliquer un écart observé sur l'hôpital entre les deux distributions, qui apparaît dès le 70^{ème} centile des deux distributions des dépenses à l'hôpital. Une explication possible serait que les personnes ayant les plus fortes

⁸ Ces fiches sont disponibles ici : <https://documentation-snds.health-data-hub.fr>.

⁹ L'adresse du Gitlab est https://gitlab.com/DREES_code.

dépenses de santé soient sous-représentées dans l'enquête EHIS 2014, cette enquête n'ayant pas pu se faire auprès de personnes hospitalisées vivant seules, et ne couvrant pas non plus les personnes vivant en institution.

Graphique 1 Distribution par centiles du montant total individuel de dépenses de santé à l'hôpital ainsi qu'en ambulatoire, pour les individus de moins de 15 ans. Comparaison entre la base RAC 2018 et l'échantillon répondant d'EHIS 2014 apparié au SNDS



Note > Les centiles de dépenses sont affichés jusqu'au 95^e inclus pour ne pas écraser la courbe.

Lecture > Le troisième quartile (75^e centile sur les figures) des dépenses des individus de moins de 15 ans en ambulatoire s'élève à 1 622 euros en 2017 d'après la base RAC, et à 1 709 euros en 2014 d'après l'échantillon répondant EHIS-SNDS.

Champ > Personnes de moins de 15 ans d'un ménage répondant à l'enquête EHIS 2014, menée en France métropolitaine uniquement, et présentes dans l'appariement EHIS 2014-SNDS. Toutes dépenses en hôpital public, y compris SSR/PSY/HAD et dépenses en sus du groupe homogène de séjour. Toutes dépenses en ambulatoire y compris les actes et consultations externes, à l'exclusion du médico-social.

Source > DREES, Base RAC 2018 ; DREES-IRDES, EHIS-SNDS 2014. Traitements DREES.

L'appariement EHIS-SNDS 2019

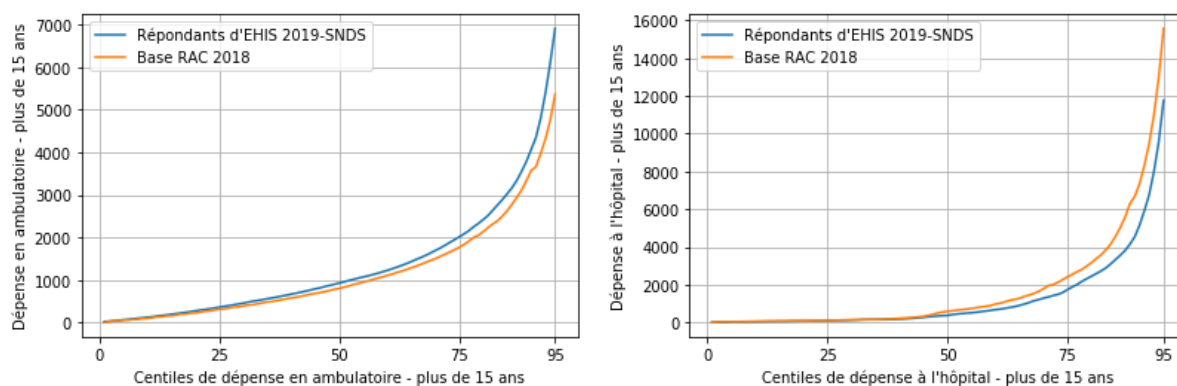
L'appariement entre EHIS 2019 et le SNDS, quant à lui, ne concerne que des individus de 15 ans ou plus. La jointure EHIS-SNDS 2019 est possible pour 12 745 individus, sur 14 192 individus de 15 ans ou plus conservés à l'issue du retraitement d'EHIS 2019. Seuls les individus d'EHIS 2019 de 15 ans ou plus et appariés au SNDS sont utilisés dans Ines-Omar 2019. Un recalage des poids individuels selon l'âge, le sexe et le niveau de vie est effectué pour que ce sous-échantillon demeure représentatif de la population de France métropolitaine de 15 ans ou plus.

Les programmes développés dans le cadre de la base RAC 2018 sont adaptés aux données disponibles pour l'année 2019, et exécutés ici sur l'échantillon répondant d'EHIS 2019 apparié au SNDS.

La distribution par centiles de la dépense de santé totale en ambulatoire a un profil très proche de celle obtenue sur la base exhaustive des dépenses en 2018¹⁰, malgré la plus faible dimension de la base des répondants d'EHIS-SNDS 2019 (*graphique 2*). En revanche, un écart significatif est observé sur l'hôpital dès le 70^{ème} centile de la distribution des dépenses, en dépit de la hausse des dépenses de santé entre 2018 et 2019. Cela pourrait s'expliquer par deux phénomènes. D'une part, l'enquête EHIS est une enquête auprès des ménages ordinaires, qui exclut les personnes âgées vivant en institution, dont celles en toute fin de vie recevant d'importants soins hospitaliers. D'autre part, il est probable que les personnes vivant en ménage ordinaire et ayant connu des périodes d'hospitalisation prolongées, ou plus généralement étant en mauvaise santé, soient sous-représentées dans l'enquête EHIS en raison d'un moindre taux de réponse.

¹⁰ La base RAC 2019 n'ayant pas été produite par la DREES, la comparaison se fait avec la base RAC 2018.

Graphique 2 Distribution par centiles du montant total individuel de dépenses de santé à l'hôpital ainsi qu'en ambulatoire, pour les individus de 15 ans et plus. Comparaison entre la base RAC 2018 et l'échantillon répondant d'EHIS 2019 en France métropolitaine apparié au SNDS



Note > Les centiles de dépenses sont affichés jusqu'au 95^e inclus pour ne pas écraser la courbe.

Lecture > Le troisième quartile des dépenses des individus de 15 ans et plus en ambulatoire s'élève à 1 768 euros en 2018 d'après la base RAC, et à 2 014 euros en 2019 d'après l'échantillon répondant EHIS-SNDS.

Champ > Ensemble de la population ayant eu au moins une dépense de santé remboursable et individualisable dans l'année 2018 (champ de la base RAC 2018) en France métropolitaine ; échantillon représentatif de la population de 15 ans ou plus vivant en France métropolitaine en ménage ordinaire – personnes répondantes à EHIS 2019 et retrouvées dans l'appariement avec le Système national des données de santé (SNDS). Toutes dépenses en hôpital public, y compris SSR/PSY/HAD et dépenses en sus du groupe homogène de séjour. Toutes dépenses en ambulatoire y compris les actes et consultations externes, à l'exclusion du médico-social.

Source > DREES, base RAC 2018, EHIS-SNDS 2019. Traitements DREES.

L'appariement EHIS – SNDS 2019 est étoffé de plusieurs tables absentes de l'appariement EHIS – SNDS 2014, qui permettent de mieux repérer dans Ines-Omar 2019 que dans Ines-Omar 2017 les ALD et les bénéficiaires de la CMU-C ou de l'ACS :

- Les titulaires d'une ALD sont repérés pour la plupart dans le DCIR ; ce sont les personnes ayant consommé des soins enregistrés dans le DCIR en 2019, et qui sont enregistrés dans le DCIR comme titulaires d'une ALD au moins une journée dans l'année. La période de reconnaissance de chaque ALD est enregistrée dans la table du référentiel médicalisé du DCIR (IR_IMB_R). Parmi les personnes qui n'ont pas consommé de soins enregistrés dans le DCIR, les titulaires d'une ALD sont repérés au moyen de leurs dépenses en lien avec une ALD dans le PMSI. En fin de compte, il est possible d'identifier dans Ines-Omar 2019 les remboursements supplémentaires faits par l'AMO au titre de l'ALD pour les individus de 15 ans ou plus, à l'hôpital comme en soins de ville.
- Les bénéficiaires de l'ACS sont repérés dans le SNDS grâce à la table des affiliations à un organisme complémentaire (IR_ORC_C). Par ailleurs les remboursements supplémentaires des organismes complémentaires d'assurance maladie (OCAM) au titre des contrats ACS sont parfois observés dans le DCIR via la table des autres remboursements que le régime obligatoire (ER_ARO_F). Tous les individus pour lesquels des remboursements supplémentaires des OCAM au titre des contrats ACS figurent dans le SNDS sont bénéficiaires de l'ACS, d'après la table d'affiliation à un organisme complémentaire.
- Le bénéfice de la CMU-C est également repéré dans le PMSI et dans le DCIR via la table des affiliations à un organisme complémentaire (IR_ORC_C). De même que pour l'ACS, les remboursements dus à la CMU-C dans le DCIR sont disponibles dans la table des autres remboursements que le régime obligatoire (ER_ARO_F), et cohérents avec la table d'affiliation à un organisme complémentaire.

L'appariement EHIS – SNDS 2014 ne comprend pas la table des affiliations à un organisme complémentaire (IR_ORC_C), mais on peut y identifier les bénéficiaires de la CMU-C grâce aux remboursements effectués par la CMU-C, dans le PMSI comme dans le DCIR.

Décomposition de la dépense de santé

Les programmes de la base RAC, appliqués aux appariements EHIS 2014 – SNDS et EHIS 2019 – SNDS, extraient du SNDS, pour chaque soin :

- Le montant de la dépense (DEP)
- Les remboursements de base de l'AMO (REMB), dont ceux perçus au titre d'un motif d'exonération comme l'ALD.
- Les remboursements supplémentaires enregistrés dans le SNDS (REMB_SUP), dont ceux perçus au titre d'un régime spécial comme celui de l'Alsace-Moselle, de la CMU-C ou de la part des OCAM au titre des contrats ACS.
- Le reste à charge après remboursement de base et remboursements supplémentaires mesurés par le SNDS (RAC_AMO_SUP).

Ces variables s'articulent ainsi : $DEP = REMB + REMB_SUP + RAC_AMO_SUP$.

Les programmes de la base RAC extraient également les composantes du reste à charge après remboursement de base et remboursements supplémentaires mesurés par le SNDS. Il s'agit du montant restant à charge du ménage, et non pas du montant initial de ses composantes avant remboursements supplémentaires.

- Le montant des dépassements (DPA) pour les soins de ville.
- Les participations forfaitaires et franchises¹¹ (PF_FR), non opposables au remboursement des complémentaires santé, pesant par exemple sur les consultations de médecins ou les boîtes de médicament.
- La participation forfaitaire aux actes lourds à l'hôpital, de 24 euros en 2019 (PA24).
- Le forfait journalier à l'hôpital (FJ).

Le Ticket modérateur (TM) est calculé par différence, de sorte à vérifier : $RAC_AMO_SUP = DPA + TM + PA24 + PF_FR + FJ$.

Si cette formule est vraie en toute généralité à l'hôpital comme en soins de ville, elle peut être simplifiée dans la mesure où les variables DPA et PF_FR ne sont positives qu'à la ville, tandis que PA24 et FJ ne sont positives qu'à l'hôpital.

Tableau 2 Composantes du reste à charge pour les soins de ville et les soins à l'hôpital

	Ville	Hôpital
DPA	X	
PF_FR	X	
PA24		X
FJ		X
Définition du TM	$RAC_AMO_SUP = TM + DPA + PF_FR$	$RAC_AMO_SUP = TM + PA24 + FJ$
DPA	X	

Lecture > Le reste à charge après remboursement de l'assurance maladie obligatoire sur les soins de ville se compose d'un ticket modérateur, d'un dépassement et d'un forfait.

Enfin, la dépense de santé peut se décomposer entre la base de remboursement de la sécurité sociale (BASE_REMB), qui sert au calcul du remboursement de base par l'AMO, les dépassements en ville (DPA) et le forfait journalier à l'hôpital (FJ) : $DEP = BASE_REMB + DPA + FJ$ ¹².

La décomposition du reste à charge après remboursements de base et supplémentaires pris en compte dans le SNDS servira ultérieurement à calculer finement les remboursements des complémentaires santé privées (voir partie VII).

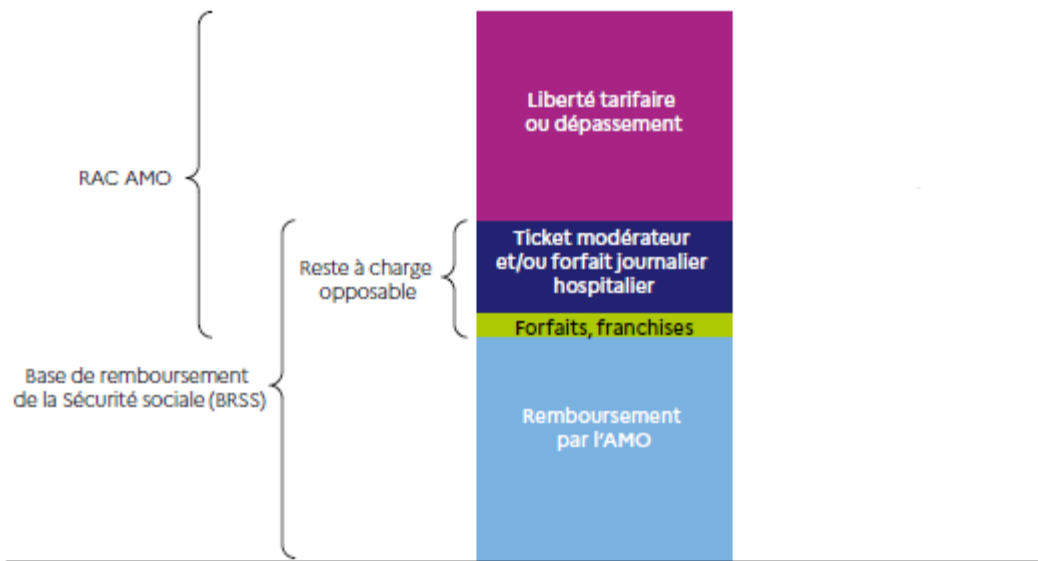
Une partie du reste à charge dite opposable est composée du ticket modérateur et des forfaits et franchises. Cette part est dans sa majorité remboursée par les contrats de complémentaire santé dits responsables, qui ont l'obligation de rembourser le Ticket Modérateur à leurs assurés, et l'interdiction de couvrir les participations forfaitaires et franchises.

La décomposition de la dépense de santé dans les soins de ville peut être résumée par ce schéma :

¹¹ À la suite d'une sous-estimation des participations et franchises par la base RAC, celles-ci sont aussi sous-estimées dans le modèle Ines-Omar 2019 et ne doivent pas être étudiées en tant que telles avec cette source.

¹² En toute rigueur, dans cette présentation, est ajoutée aux dépassements restant à régler après les remboursements supplémentaires, variable extraite du SNDS et notée ici DPA, la part des dépassements faisant déjà l'objet d'un remboursement supplémentaire (par exemple la Caisse d'Assurance Maladie des Industries Électriques et Gazières rembourse les dépassements sur les actes paramédicaux jusqu'à 20 % de la BRSS). Ainsi, comme c'est bien le cas dans le modèle, $DEP = BASE_REMB + DPA + FJ + \text{dépassement pris en charge par les remboursements supplémentaires}$.

Schéma 2 Le reste à charge sur une dépense en ville du panier de soins remboursables par l'assurance maladie obligatoire



AMO : assurance maladie obligatoire ; RAC AMO : reste à charge après remboursement par l'assurance maladie obligatoire.

■ LES RETRAITEMENTS DE LA COMPLÉMENTAIRE SANTÉ

Les trois enquêtes auprès des ménages utilisées dans Ines-Omar 2019 (EHIS 2014, SRCV 2017, EHIS 2019) requièrent des retraitements de leur module complémentaire santé selon l'information disponible dans l'enquête elle-même, et dans le SNDS si celle-ci lui est appariée (tableau 3). Il s'agit, selon les enquêtes, soit de corriger des incohérences entre le statut professionnel déclaré et le type de couverture complémentaire, ou un biais de déclaration de la CMU-C, soit de compléter l'information disponible sur la couverture complémentaire au sein des ménages.

Tableau 3 Informations sur la couverture complémentaire disponibles dans les sources d'Ines-Omar 2019

	EHIS 2019	SRCV 2017	EHIS-ESPS 2014
Couverture des moins de 15 ans	Absents de l'enquête	Non	Oui
Bénéfice de la CMU-C dans le SNDS	Oui	Non appariée	Oui
Bénéfice de l'ACS dans le SNDS	Oui	Non appariée	Non
Bénéfice de l'ACS dans l'enquête	Non	Non	Oui

Les quatre étapes de retraitement de la complémentaire santé, détaillées dans les parties ci-dessous, sont :

- Attribuer une couverture complémentaire à tous les membres des ménages dans SRCV 2017. En effet, seules les personnes de 16 ans ou plus sont interrogées sur leur couverture complémentaire santé dans SRCV 2017, et l'identité de leurs éventuels ayants-droits n'est pas établie explicitement. Ainsi, le statut de couverture des enfants de moins de 16 ans y est inconnu. De plus, un certain nombre d'ayants droit adultes n'ont pas répondu au questionnaire, ou seulement partiellement.
- Redresser les déclarations de CMU-C. Les informations sur le bénéfice de la CMU-C provenant du SNDS peuvent servir à vérifier les déclarations faites dans EHIS-ESPS 2014 et EHIS 2019, comme ces enquêtes sont appariées au SNDS. Par ailleurs, un modèle de prédiction du bénéfice de la CMU-C entraîné sur EHIS-ESPS 2014 est utilisé dans EHIS 2019 et SRCV 2017.
- Vérifier la nature du contrat (contrat d'entreprise ou individuel), car un certain nombre de personnes ne connaissent pas ou mal le type de contrat dont elles bénéficient, conduisant à des situations incohérentes. Ce retraitement est nécessaire pour utiliser EHIS 2014, EHIS 2019 et SRCV 2017.
- Attribuer le bénéfice de l'ACS. Faute d'information dans SRCV 2017, les bénéficiaires de l'ACS parmi les individus de moins de 15 ans de SRCV 2017 sont tirés de sorte à respecter la répartition des bénéficiaires de l'ACS par tranche d'âge.

L'échantillon d'EHIS 2019 et celui d'EHIS-ESPS 2014 sont restreints aux seuls individus appariés avec le SNDS avant de procéder à ces retraitements.

Attribution d'une couverture dans SRCV 2017

Il est nécessaire d'établir pour chaque individu d'un ménage son état de couverture précis : type de contrat, ayant droit ou ouvrant droit, personnes du ménage rattachées au même contrat, cotisation versée pour ce contrat. Cependant, la couverture complémentaire au sein des ménages est mal connue :

- Les enfants ne font pas partie du questionnaire. Ils sont donc seulement comptabilisés indirectement par leur ouvrant droit *via* la variable qui recense son nombre d'ayants droit (variable CombPers) ;
- Un certain nombre d'ayants droit adultes n'ont pas répondu au questionnaire, ou seulement partiellement ;
- Le lien explicite entre l'ouvrant droit et ses ayants droit n'est pas établi par le questionnaire. Il doit donc être établi au mieux en utilisant toutes les informations connues sur les contrats.

Ainsi, sans retraitements, seulement 77 % de la population serait considérée comme couverte par une complémentaire santé et, pour beaucoup, il manquerait la connaissance des caractéristiques du contrat. Après retraitement, plus de 96 % de la population est estimée être couverte en 2017 (Fouquet, 2020).

L'attribution de la couverture maladie universelle complémentaire (CMU-C)

Lorsqu'au moins une personne du ménage s'est déclarée bénéficiaire de la CMU-C, et si cette personne est considérée comme un adulte du ménage, la CMU-C est affectée aux autres membres du ménage car elle est en principe attribuée à l'ensemble des membres d'un ménage. L'ouvrant droit retenu est la personne la plus âgée ayant déclaré bénéficier de la CMU-C. Sont considérés comme n'étant pas bénéficiaires : les personnes d'une génération antérieure à celle de l'ouvrant droit et les enfants de l'ouvrant droit ayant plus de 25 ans, sauf s'ils ont

déclaré explicitement bénéficiaire de la CMU-C. Tous les autres membres du foyer sont considérés bénéficiaires de la CMU-C.

Logique de couverture

L'une des principales variables utilisées pour le retraitement est CombPers, qui recense pour tout ouvrant droit le nombre de personnes rattachées à son contrat dans le ménage, y compris lui-même. Des erreurs ont pu être commises par les répondants à l'enquête ; elles sont corrigées selon la méthode du modèle Ines-Omar 2017 (Fouquet, 2021). Ces erreurs portent notamment sur la variable déclarée par les ouvrants droit (CombPers) qui indique le nombre total de personnes assurées par leur contrat de complémentaire santé au sein de leur ménage : il arrive par exemple que les ouvrants droit ne s'incluent pas dans le décompte des assurés.

Une fois la variable CombPers corrigée, l'enjeu est tout d'abord d'identifier tous les membres du ménage couverts, même s'ils sont enfants ou non-répondants ; ensuite de rattacher chaque ayant droit à son ouvrant droit en fonction des informations disponibles sur son contrat. L'algorithme est le même que celui mis en œuvre dans le modèle Ines-Omar 2017 (voir Fouquet, 2021).

Après mise en œuvre de l'algorithme, les seuls ménages supprimés sont ceux composés uniquement de non-répondants au module complémentaire santé. On passe ainsi de 25 401 individus à 24 280 personnes dans la base finale de SRCV 2017.

Redressement des déclarations de la CMU-C

Il est nécessaire de redresser certaines déclarations de bénéfice de la CMU-C dans les enquêtes utilisées par le modèle Ines-Omar. En effet, on observe par exemple une sur-déclaration de la CMU-C chez les plus de 60 ans dans EHS 2019, et une proportion insuffisante de bénéficiaires de la CMU-C chez les moins de 15 ans de SRCV 2017, après attribution de la couverture. Le taux de souscription par tranche d'âge en France métropolitaine donné chaque année par le Fonds CMU-C (disparu en 2020 avec le remplacement de la CMU-C par la CSS) fait référence.

Ces biais semblent induits par le module sur la complémentaire santé dans SRCV 2017 et EHS 2019 (voir annexe 1). Celui-ci s'ouvre par une question demandant si la personne interrogée « bénéficie actuellement de la couverture médicale universelle complémentaire, appelée plus couramment CMU. » Si la personne interrogée répond affirmativement, alors l'enquêteur lui demande directement qui a souscrit la CMU-C et qui en sont les bénéficiaires au sein du ménage, sans plus l'interroger sur sa couverture complémentaire. Ainsi, une personne qui aurait déclaré hâtivement bénéficiaire de la CMU-C n'a plus l'occasion de se dédire.

Il est possible de redresser le bénéfice de la CMU-C déclaré dans EHS 2019 à l'aide des informations provenant du SNDS, auquel est appariée l'enquête. Cela est opportun car les informations contenues dans le SNDS peuvent être considérées plus fiables que les réponses données à l'enquête.

En revanche EHS-ESPS 2014 ne présente pas de biais de déclaration de la CMU-C. De plus, l'information sur la CMU-C ou l'ACS contenue dans le SNDS 2014 permet de redresser la couverture déclarée dans l'enquête. Ainsi, il est possible d'entraîner sur EHS-ESPS 2014 un modèle *logit* de prédiction du bénéfice de la CMU-C qui ne soit pas biaisé. Ce modèle est ensuite utilisé pour corriger les biais de SRCV 2017 et EHS 2019.

Lorsque le SNDS et l'enquête EHS-ESPS 2014 ou EHS 2019 sont croisés, l'indicatrice de bénéfice de la CMU-C provenant du SNDS contredit souvent les déclarations de l'enquête (voir *tableaux 4 et 6*). Cela n'indique pas nécessairement que l'enquête soit de mauvaise qualité ; par exemple, un adulte enquêté en mars et alors couvert par un contrat individuel peut souscrire la CMU-C en octobre, ce qui le fera compter comme tel dans le SNDS. Cependant, il importe de trancher entre l'information du SNDS et celle de l'enquête. Dans EHS-ESPS 2014 comme dans EHS 2019, la priorité est accordée aux informations du SNDS. En effet, le SNDS contient des montants de remboursements de la CMU-C sur les soins de ville, et il est jugé préférable de modifier la couverture complémentaire de certains individus dans l'enquête, plutôt que de forcer à 0 les remboursements de la CMU-C enregistrés à leur profit dans le SNDS.

EHS 2014

Parmi les 26 514 individus d'EHS-ESPS 2014, 16 757 sont appariés aux SNDS. Après avoir retiré les individus ne déclarant pas de couverture complémentaire dans EHS-ESPS 2014, restent 16 496 individus qui seront utilisés par la suite dans la construction du modèle. Les redressements de la CMU-C s'appliquent à eux.

L'appariement EHS – SNDS 2014 fournit une information sur le bénéfice de la CMU-C pour 88 % d'entre eux. Pour ces individus, plutôt que de modifier les informations provenant du SNDS pour s'aligner sur EHS 2014, il est décidé de conformer les déclarations de l'enquête aux informations du SNDS. Ainsi :

- La CMU-C est attribuée à 1,2 % des individus d'EHS-ESPS 2014 appariés aux SNDS, qui ne se déclarent pas bénéficiaires de la CMU-C dans l'enquête mais sont désignés comme tels dans le SNDS.

- La CMU-C est retirée à 0,8 % d'individus d'EHIS-ESPS 2014 appariés aux SNDS, qui se déclarent bénéficiaires de la CMU-C dans l'enquête, mais que le SNDS n'enregistre pas comme tels. Ils sont alors considérés comme couverts par un contrat individuel.

À la suite de ces deux premières corrections, la part de bénéficiaires de la CMU-C dans la population est réévaluée de 6,6 % à 7,0 %. Le fonds CMU-C indique que le taux de bénéfice de la CMU-C en 2014 en France métropolitaine était de 6,4 %. Cette légère surestimation ne perturbe pas le fonctionnement du modèle Ines-Omar 2019. En effet, la répartition des bénéficiaires de la CMU-C par catégorie d'âge observée dans EHIS 2014 reste très proche de celle déclarée par le Fonds CMU-C :

Tableau 4 Répartition des bénéficiaires de la CMU-C par tranche d'âge en 2014

	0 – 19 ans	20 – 39 ans	40 – 59 ans	Plus de 60 ans
EHIS-ESPS 2014 (France métropolitaine)	43,7 %	29,9 %	21,3 %	4,9 %
Fonds CMU-C (France entière)	44,2 %	28,4 %	22,3 %	5,1 %

Champ > EHIS 2014 : population française métropolitaine résidant en ménage ordinaire ; fonds CMU-C : population de France entière.
Source > Fonds CMU-C ; DREES-IRDES, EHIS-ESPS 2014.

Il est donc possible d'entraîner sur EHIS 2014 un modèle de prédiction du bénéfice de la CMU-C. Il est fait le choix d'un modèle *logit* prenant comme variables l'âge, le sexe, le dixième de niveau de vie, l'occupation professionnelle, l'état de santé et le type de ménage. Ce modèle, entraîné sur l'échantillon des individus couverts par une complémentaire, privée ou publique, présente un pseudo-R2 de 0,50.

SRCV 2017

L'identification des bénéficiaires de la CMU-C par le questionnaire complémentaire de SRCV 2017 est assez limitée, n'incorporant aucun contrôle de cohérence pour s'assurer que l'enquêté ne s'est pas trompé. Il est donc probable que des personnes bénéficiaires de la CMU-C aient déclaré être couvertes par un contrat individuel privé et soient par conséquent considérées comme telles malgré les retraitements effectués.

Sans redresser les déclarations de CMU-C faites dans SRCV 2017, le taux de couverture par la CMU-C serait estimé à 6,6 % après attribution d'une couverture dans SRCV. Or le Fonds CMU-C estimait le taux de couverture par la CMU-C en France métropolitaine à 7,4 % en 2017. En particulier, le taux de souscription parmi les moins de 15 ans estimé dans SRCV 2017 apparaît bien inférieur au taux donné par le fonds CMU-C¹³. Comme seuls les individus de moins de 15 ans de SRCV seront conservés pour la construction du modèle, il importe avant tout de corriger leur couverture.

Pour pallier ce biais, le modèle de prédiction¹⁴ du bénéfice de la CMU-C, entraîné sur EHIS 2014, est appliqué aux individus de 15 ans ou plus de SRCV 2017. 105 individus de 25 à 59 ans précédemment couverts par un contrat individuel, dont le modèle indique une probabilité de bénéfice de la CMU-C supérieure à 40 %, se voient attribuer la CMU-C. Cette correction est modique, puisqu'elle ne concerne que 1,2 % des personnes de 25 à 59 ans de l'enquête.

Puis l'algorithme d'attribution de la couverture aux adultes non-répondants et aux moins de 16 ans est appliqué à nouveau. Cet algorithme va répercuter la correction effectuée sur les 105 individus de 25 à 59 ans sur les enfants vivant dans leur foyer.

Ainsi le taux de couverture par la CMU-C est rehaussé à 14 % pour les 0-9 ans et 13 % pour les 10-14 ans, soit des taux très proches de ce qu'indique le Fonds CMU-C.

EHIS 2019

Parmi les 14 192 individus d'EHIS 2019, 12 745 sont appariés au SNDS, dont 11 753 ont consommé des soins en 2019 d'après le SNDS. L'appariement EHIS - SNDS 2019 traité par les programmes de la base RAC permet d'observer des indicatrices de bénéfice de l'ACS et de la CMU-C pour les individus consommateurs. L'échantillon des 12 745 individus est repondéré, de sorte à conserver les pondérations totales observées dans l'échantillon initial par croisement de sexe, tranche d'âge et décile de niveau de vie.

¹³ D'après SRCV 2017 retraité, 11 % des 0-9 ans et 10 % des 10-14 ans bénéficieraient de la CMU-C en métropole. Or à partir de la répartition des bénéficiaires de la CMU-C donnée par le Fonds CMU-C, on peut estimer que 15 % des 0 – 9 ans et 14 % des 10 – 14 ans en bénéficient en France entière.

¹⁴ Modèle *logit* prédisant le bénéfice de la CMU-C parmi les individus assurés par une complémentaire santé, entraîné sur EHIS – ESPS 2014, prenant comme variables l'âge, l'âge mis au carré, le sexe, l'occupation professionnel, le type de ménage, le niveau déclaré de santé, le décile de niveau de vie.

Les deux indicatrices se contredisent pour 44 individus. Cependant, un individu est compté comme bénéficiaire de la CMU-C dès lors qu'il l'a été au moins un mois dans l'année. Il est par conséquent possible qu'une personne ait successivement souscrit à la CMU-C et l'ACS dans la même année, d'où la contradiction apparente. Les deux indicatrices sont départagées en utilisant les déclarations faites dans l'enquête. 26 individus sur 44 s'y déclarent bénéficiaires de la CMU-C et voient leur indicatrice de bénéfice de l'ACS forcée à 0, les 18 autres conservent le bénéfice de l'ACS aux dépens de la CMU-C.

L'indicatrice de bénéfice de la CMU-C provenant du SNDS contredit souvent les déclarations dans EHIS 2019 (voir *tableau 5*). L'information provenant de l'enquête EHIS 2019 est donc mise en cohérence avec celle venant du SNDS, conservée en l'état, de sorte à corriger le biais de sur-déclaration de la CMU-C parmi les 60 ans ou plus observé dans EHIS 2019. Avant correction, le taux de couverture par la CMU-C des plus de 60 ans est de 4,1 %, contre 1,9 % selon les données du Fonds CMU-C (voir *tableau 6*). Selon le Fonds CMU-C, 8,1 % de la population générale en France métropolitaine souscrivait à la CMU-C en octobre 2019, et autour de 6,3 % de la population de 15 ans ou plus.

Tableau 5 Croisement de l'information sur le bénéfice de la CMU-C dans le SNDS et EHIS 2019

Bénéfice de la CMU-C dans EHIS 2019	Non consommant	Consommant – sans la CMU-C dans le SNDS	Consommant – avec la CMU-C dans le SNDS
Ne sait pas	6	6	0
Oui	73	288	465
Non	913	10 913	81
Total	992	11 207	546

Lecture > Parmi les 12 745 individus d'EHIS 2019 appariés au SNDS, 992 n'ont pas consommé de soins en 2019, si bien que les programmes de la base RAC ne remontent pas l'information sur le bénéfice de la CMU-C ou de l'ACS. 546 individus qui ont consommé des soins en 2019 bénéficient de la CMU-C d'après le SNDS, dont 465 en bénéficient également d'après EHIS 2019.

Champ > Individus d'EHIS 2019 appariés au SNDS et déclarant leur couverture complémentaire dans l'enquête.

Source > DREES, EHIS 2019 - SNDS.

Tableau 6 Taux de couverture par la CMU-C par tranche d'âge en France métropolitaine

	15 ans ou plus	15-19 ans	20-39 ans	40-59 ans	60 +
EHIS 2019 restreint aux appariés avec le SNDS, avant retraitements	6,7 %	8,8 %	8,9 %	6,5 %	4,3 %
Fonds CMU-C en 2019 (estimation)	6,3 %	10,1 %	10,1 %	6,8 %	1,9 %
EHIS 2019 - SNDS après retraitements	6,2 %	8,4 %	9,5 %	6,9 %	2,1 %

Champ > EHIS 2019 : individus vivant en ménage ordinaire en France métropolitaine ; Fonds CMU-C : population de France métropolitaine.

Source > DREES, EHIS 2019 – SNDS ; Fonds CMU-C, calculs DREES.

Le redressement des déclarations d'EHIS 2019 procède en deux temps :

- La CMU-C est attribuée aux 81 individus signalés comme tels dans le SNDS. Cette correction porte le taux de bénéficiaires de la CMU-C à 7,4 % parmi les plus de 15 ans, ce qui est supérieur au taux indiqué par le Fonds CMU-C.
- Le modèle *logit* de prédiction du bénéfice de la CMU-C est appliqué.
 - La CMU-C est retirée à 90 individus de plus de 60 ans avec une probabilité prédite de bénéfice de la CMU-C inférieure à 20 %.
 - La CMU-C est retirée à 77 individus de moins de 60 ans avec une probabilité prédite de bénéfice de la CMU-C inférieure à 0,3 %.
 - La CMU-C est attribuée à 23 individus de moins de 60 ans avec une probabilité prédite de bénéfice de la CMU-C supérieure à 85 %.
 - Les individus auxquels la CMU-C est retirée se voient attribuer un contrat individuel ou collectif en fonction du statut professionnel de leur ouvrant droit.

Ces seuils sont choisis de sorte à obtenir un taux de souscription de la CMU-C par tranche d'âge proche de la cible du fonds CMU-C.

In fine, après suppression des individus d'EHIS 2019 dont la couverture complémentaire reste inconnue, le taux de souscription de la CMU-C par tranche d'âge dans EHIS 2019 est proche de celui estimé à partir des données du fonds CMU-C (voir *tableau 6*).

Contrats d'entreprise ou individuels

Les questions adressées aux ménages dans EHS 2014, SRCV 2017, et EHS 2019 permettent d'identifier si la complémentaire santé est un contrat d'entreprise. Les contrats collectifs qui ne sont pas des contrats d'entreprise (contrats collectifs dans des associations, ...) ne peuvent cependant pas être distingués des contrats individuels ; ils sont donc considérés comme des contrats individuels dans le modèle.

Il est fréquent que des personnes bénéficiant de contrats individuels, notamment les fonctionnaires, pensent être couvertes par une complémentaire d'entreprise car elles les ont souscrits auprès d'une mutuelle spécialisée dans la couverture santé des agents d'un ministère spécifique. Ces erreurs sont corrigées à l'aide du statut d'activité et de la catégorie sociale de l'ouvrant droit, pour lui ainsi que pour tous ses ayants droit. Bien que les variables soient différentes entre enquêtes, la logique suivie est identique : parmi les personnes couvertes par une complémentaire privée, sont considérés comme bénéficiaires de contrats individuels les fonctionnaires, les indépendants, les salariés d'un particulier, les personnes aidant un membre de leur famille, les personnes sans emploi depuis plus d'un an, les inactifs et les retraités.

Dans EHS 2014 et SRCV 2017, toutes les personnes de 16 ans ou plus sont interrogées, si bien que l'ouvrant droit de chaque complémentaire santé répond lui-même à l'enquête (sauf dans le cas marginal où des personnes interrogées sont couvertes par un contrat souscrit par un tiers extérieur à leur ménage). Son statut professionnel peut être considéré comme plus fiable que le type de complémentaire déclaré. La logique de correction du type de contrat est donc toujours suivie dans EHS 2014 et SRCV 2017.

En revanche dans EHS 2019, l'ouvrant droit d'une complémentaire santé ne répond pas toujours lui-même à l'enquête, puisqu'un seul adulte par ménage est interrogé. Ce dernier déclare le statut professionnel des autres membres de son ménage, avec un risque plus fort de faire erreur. Ainsi, la logique de correction dans EHS 2019 n'est appliquée pleinement que lorsque l'adulte interrogé est lui-même l'ouvrant droit de son contrat. Lorsque l'adulte interrogé n'est qu'ayant droit, le type de complémentaire n'est corrigé que si l'ouvrant droit est inactif ou retraité. Ce statut professionnel est susceptible de susciter en effet moins de confusion que les autres statuts professionnels possibles.

Il faut ajouter à cela que le type de complémentaire est imputé dans EHS 2019 pour 81 individus qui ont déclaré être couverts mais ont omis leur type de complémentaire, en utilisant le statut professionnel de leur ouvrant droit.

Le tableau 7¹⁵ récapitule les effets des différents retraitements sur la situation de couverture complémentaire des individus de 15 ans ou plus dans EHS 2019.

Tableau 7 Répartition des situations de couverture des individus de 15 ans ou plus d'EHS 2019 répondants et appariés au SNDS, avant et après retraitements

	Avant tous retraitements	Après correction de la CMU-C	Après tous retraitements (Omar 2019)
Complémentaires d'entreprise	45,0	44,8	36,7
Contrats individuels	44,7	45,6	53,8
CMU-C	6,7	6,3	6,2
Non couverts	3,6	3,3	3,2

Lecture > Avant retraitements, 45,0 % des individus de l'enquête étaient considérés couverts par une complémentaire d'entreprise. Ils étaient 44,8 % après redressement de la CMU-C, et plus que 36,7 % à l'issue de l'ensemble des retraitements.

Champ > Population française métropolitaine résidant en ménage ordinaire.

Source > DREES, EHS 2019. Traitements DREES.

Le redressement des déclarations de couverture complémentaire santé dans Ines-Omar s'écarte légèrement de celui opéré par Pierre (2022), qui exploite également l'enquête EHS 2019. La principale différence est que dans le modèle Ines-Omar, les salariés de la fonction publique ouvrants droit de leur contrat sont considérés comme ayant souscrit un contrat de type individuel, tandis que dans Pierre (2022), les réponses des salariés de la fonction publique qui se déclarent ouvrant droit d'un contrat collectif ne sont pas considérées comme erronées. In fine, Pierre (2022) estime que la part de la population bénéficiaire d'un contrat collectif est un peu plus élevée, et celle bénéficiaire d'un contrat individuel un peu plus faible (de l'ordre de 1 %) que l'estime Ines-Omar 2019.

Par ailleurs, le bénéfice de la CMU-C dans EHS 2019 est également redressé par Pierre (2022) au moyen d'un modèle *logit* de prédiction du bénéfice de la CMU-C entraîné sur l'enquête EHS 2014. Cette méthode a ensuite été reprise dans le modèle Ines-Omar 2019.

¹⁵ Les taux de couverture diffèrent légèrement de ceux publiés dans Pierre (2022) qui ne redresse pas les déclarations de CMU-C à l'aide du SNDS et procède à des corrections du type de complémentaire en fonction du statut professionnel légèrement différentes.

Attribution de l'ACS

L'ACS complète le dispositif CMU-C en soutenant les ménages légèrement au-dessus du seuil d'éligibilité à la CMU-C dans le financement de leur complémentaire santé.

Or le module complémentaire de SRCV 2017 ne permet pas de savoir qui est couvert par l'ACS. Le recours à ce dispositif est donc attribué à des individus, tirés dans SRCV 2017 conformément au nombre de bénéficiaires de l'ACS selon le fonds CMU-C¹⁶ et au pourcentage de personnes éligibles à l'ACS par vingtième de niveau de vie selon le modèle Ines 2019, même si aucun bénéficiaire de l'ACS n'est tiré au sort dans SRCV 2017 au-delà du neuvième vingtième de niveau de vie. Le tirage respecte la répartition des bénéficiaires de l'ACS par tranche d'âge observée dans la base RAC 2018¹⁷, construite à partir du SNDS exhaustif.

Tableau 8 Répartition des bénéficiaires de l'ACS entre tranches d'âge, d'après le SNDS 2018

Bénéficiaires de l'ACS fin octobre 2019 (fonds CMU-C)	Moins de 16 ans	16-29 ans	30-39 ans	40- 49 ans	50- 59 ans	60-69 ans	Plus de 70 ans
1 304 451	22,6 %	12,5 %	10,7 %	11,4 %	12,1 %	15,0 %	15,6 %

Champ > France entière.

Source > Fonds CMU-C pour le nombre de bénéficiaires de l'ACS en 2019 ; répartition des bénéficiaires entre tranches d'âge issue de la base RAC 2018, construite à partir du SNDS 2018.

Le module complémentaire d'EHIS 2019 ne permet pas non plus de savoir qui est couvert par l'ACS. En revanche cette information est fournie par l'appariement EHIS–SNDS 2019. Les bénéficiaires de l'ACS repérés dans le SNDS étant moins nombreux que la cible donnée par le fonds CMU-C pour 2019, quelques bénéficiaires supplémentaires sont tirés au sort parmi les bénéficiaires de contrat individuel (environ trois quarts des bénéficiaires de l'ACS sont repérés dans le SNDS, ce tirage concerne donc un quart supplémentaire). De nouveau, le tirage respecte la répartition des bénéficiaires de l'ACS par tranche d'âge observée dans le SNDS en 2018.

¹⁶ Le Fonds CMU-C donne le nombre de bénéficiaires de l'ACS pour la France entière, alors que le modèle Ines-Omar 2019 utilise uniquement l'échantillon métropolitain de l'enquête SRCV 2017 et de l'enquête EHIS 2019. Cela n'est pas préjudiciable, dans la mesure où le modèle Ines-Omar final a vocation à être représentatif de la France entière.

¹⁷ On utilise SRCV 2017 pour décrire la couverture complémentaire des individus de moins de 15 ans en 2019, sous l'hypothèse que celle-ci a peu évolué entre 2017 et 2019. Il est donc cohérent d'utiliser le nombre de bénéficiaires de l'ACS de 2019 ; on peut supposer par ailleurs que leur répartition entre tranches d'âge n'a que peu évolué entre 2018 et 2019.

■ PSEUDO-APPARIEMENT EHS 2014–SRCV 2017

Les données de santé des individus de moins de 15 ans présents dans EHS-ESPS 2014 sont imputées aux individus de moins de 15 ans de SRCV 2017 par pseudo-appariement. Ce pseudo-appariement pallie l'absence de dépenses de santé connues pour les individus de SRCV 2017. Il pourrait sembler plus aisé d'utiliser EHS-ESPS 2014 directement plutôt que SRCV 2017 pour décrire la couverture complémentaire et les dépenses de santé des individus de moins de 15 ans ; cependant EHS 2014 décrit leur couverture complémentaire avant l'entrée en vigueur en 2016 de l'accord national interprofessionnel (ANI) de 2013, qui a généralisé la couverture collective en entreprise. Il est préférable d'utiliser une source sur la couverture complémentaire des moins de 15 ans postérieure à 2016, et SRCV 2017 est la seule source disponible.

Comme dans Ines-Omar 2017, ce pseudo-appariement a été réalisé au moyen de la méthode des forêts aléatoires (voir Fouquet, 2021), décrite en détail dans l'annexe 5.

La mise en pratique du pseudo-appariement

À cause du faible nombre d'individus de moins de 15 ans dans l'appariement EHS 2014 – SNDS (3 552), il n'a pas été possible de pseudo-apparier séparément les enfants selon la nature de leur couverture complémentaire : individuelle, collective, CMU-C, non couvert. Les bénéficiaires de la CMU-C ont été appariés d'une part, et tous autres les enfants d'autre part, dont les non-couverts ou ceux bénéficiant de l'ACS. EHS 2014 contient 815 enfants bénéficiaires de la CMU-C, et 2 737 enfants avec une autre couverture.

Très peu de variables concernant l'état de santé des moins de 15 ans sont disponibles à la fois dans EHS 2014 et SRCV 2017. Notamment la reconnaissance d'une ALD est absente des deux bases pour les enfants. Pour choisir la liste de variables proposées aux arbres de régression, il faut chercher les variables qui permettent d'obtenir la forêt aléatoire prédisant la dépense de santé totale des moins de 15 ans la plus performante.

Parmi les dix variables retenues pour l'appariement des enfants autres que les bénéficiaires de la CMU-C :

- Certaines concernent l'individu apparié : son âge, son sexe, son état de santé déclaré, le fait de souffrir d'une maladie chronique, le type de couverture complémentaire.
- D'autres concernent son ménage : l'indice de masse corporelle (IMC) moyen des plus de 15 ans, la consommation journalière de cigarettes moyenne des 15 ans ou plus, le diplôme obtenu le plus élevé au sein du ménage, le nombre de personnes dans le ménage, la zone d'études et d'aménagement du territoire (ZEAT) de résidence.

Seules sept variables sont retenues pour le pseudo-appariement des bénéficiaires de la CMU-C, en raison du nombre plus réduit d'observations : âge, sexe, état de santé déclaré, ZEAT du ménage, plus haut diplôme du ménage, IMC moyen et consommation moyenne de cigarettes des individus de 15 ans ou plus du ménage.

Les autres paramètres de la forêt aléatoire sont également choisis de sorte à obtenir la forêt aléatoire qui prédise le mieux la dépense de santé totale des enfants. Les deux forêts aléatoires réalisées comptent 100 arbres (ce nombre étant choisi arbitrairement). Chacun de ces arbres utilise une sélection de quatre variables explicatives tirées parmi la liste indiquée ci-dessus, et dont la probabilité de sélection dépend du pouvoir explicatif de la variable. Enfin, chaque strate de la population des moins de 15 ans d'EHS-ESPS 2014 définie par ces arbres doit compter au moins 20 individus.

■ ATTRIBUTION D'UN CONTRAT DE COMPLÉMENTAIRE SANTÉ

EHIS 2019 et SRCV 2017 contiennent des éléments sur les contrats de complémentaire santé souscrits par les enquêtés : la nature du contrat (contrat individuel / d'entreprise / CMU-C / ou non couvert), la satisfaction vis-à-vis des remboursements et la prime d'assurance mensuelle pour l'ensemble des assurés. Mais ces informations sont insuffisantes pour estimer finement les remboursements par l'organisme complémentaire d'assurance maladie (OCAM), et il n'est pas certain que les primes d'assurance déclarées par les enquêtés soient d'une fiabilité suffisante. Pour ces deux raisons, il est fait appel à l'enquête auprès des organismes offrant une couverture complémentaire santé (enquête OC) 2019 de la DREES, qui contient de nombreuses informations sur les garanties liées aux contrats les plus souscrits ainsi que sur le mode de tarification. L'enjeu est alors d'attribuer un contrat aussi plausible que possible à chaque assuré. La méthode d'attribution d'un contrat de complémentaire santé présentée ici est pratiquement identique à celle utilisée dans la version 2017 d'Ines-Omar.

Cependant la représentativité des contrats de complémentaire santé par l'échantillon de l'enquête OC 2019 s'est améliorée par rapport à l'enquête OC 2016. En effet, l'échantillon de contrats de l'enquête OC 2019 est tiré parmi les dix contrats de chaque type les plus souscrits auprès de chaque organisme complémentaire, et non plus parmi les trois contrats les plus souscrits. Selon Guilloton et Lapinte (2023), « en 2019, la constitution d'une base de contrats intégrant les dix plus gros contrats collectifs et individuels rend l'enquête représentative de respectivement 37 % et 67 % des bénéficiaires. Cela représente un gain de 14 points de couverture pour les contrats collectifs, et de 20 points pour les contrats individuels par rapport aux éditions précédentes de l'enquête, où seuls les trois plus gros contrats de chaque type étaient considérés ». La représentativité des contrats collectif est plus faible que celle des contrats individuels car le marché collectif est plus morcelé.

Retraitement du type de contrat dans l'enquête OC

Dans EHIS 2019 et SRCV 2017, les contrats de complémentaire santé sont divisés entre contrats individuels, d'entreprise et CMU-C. Les contrats collectifs qui ne sont pas des contrats d'entreprise sont donc comptés parmi les contrats individuels. Dans l'enquête OC, la distinction se fait entre individuel et collectif. Il est donc nécessaire de reclasser les contrats collectifs non d'entreprise en contrats individuels. Ceci est possible car les organismes complémentaires déclarent si les contrats sont réservés à une certaine catégorie de population.

Le tableau 9 présente la répartition des contrats collectifs selon la population à laquelle ils sont réservés. Le score de contrat mentionné dans ce tableau correspond à une méthode d'évaluation du niveau de garantie mise au point par la DREES. Il s'interprète comme un niveau de prise en charge théorique, par les contrats de complémentaire santé, du reste à charge après intervention de l'AMO. Le score varie entre 0 (le contrat n'offre aucune prise en charge du reste à charge sur les prestations de soins considérées) et 1 (le contrat offre une prise en charge complète des restes à charge pour les prestations de soins considérées) [Loiseau, 2020 ; Lapinte *et al.*, 2024].

Tableau 9 Statistiques descriptives sur les contrats collectifs selon la population visée

	Retraités, anciens salariés	Fonctionnaires	Salariés du privé	Branche professionnelle	Indépendants / agriculteurs	Autres	Pas d'exclusivité déclarée
Nombre de contrats	13	17	244	59	26	15	89
Part des assurés collectifs	2,4 %	0,8 %	47,9 %	16,7 %	2,8 %	2,4 %	27,0 %
Score moyen	0,82	0,69	0,77	0,69	0,60	0,56	0,68
Prime de référence moyenne	62 €	54 €	76 €	34 €	65 €	46 €	83 €

Lecture > 244 contrats de l'enquête OC 2019 sont déclarés comme étant destinés aux salariés du secteur privé. Ces contrats couvrent 47,9 % des bénéficiaires de contrats collectifs d'après l'enquête OC 2019. Le score moyen pour ces contrats s'élève à 0,77 et la prime moyenne pour un individu « de référence » défini dans le questionnaire de l'enquête est de 76 euros mensuels.

Champ > Contrats collectifs non surcomplémentaires de l'enquête OC.

Source > DREES, enquête OC 2019. Traitements DREES.

Étant donné leur niveau de garanties et leur tarification, les contrats collectifs de l'enquête OC n'ayant pas déclaré d'exclusivité spécifique semblent être des contrats d'entreprise classiques, au sens d'EHIS 2019 et de SRCV 2017, ayant jugé que la question ne les concernait pas. Les catégories « retraités », « fonctionnaires », « indépendants » et « autres » sont considérées comme ne relevant pas de contrats d'entreprise et sont donc classées comme relevant de contrats individuels au sens d'EHIS 2019 et SRCV 2017. À l'inverse, les quelques contrats individuels de l'enquête OC réservés à une branche professionnelle sont reclassés comme des contrats d'entreprise au sens d'EHIS 2019 et SRCV 2017.

Imputation des contrats

L'imputation de contrats de complémentaire santé aux individus interrogés par EHS 2019 et aux ouvrants droit d'un contrat de complémentaire dans SRCV 2017¹⁸ se fait sur la base des éléments permettant de distinguer entre eux les contrats dans l'enquête OC 2019 :

- Type de contrat : individuel ou d'entreprise ;
- Score de qualité de contrat estimé sur la base des garanties offertes, et classé en trois grandes catégories de niveau de couverture ;
- Exclusivité pour une certaine catégorie de population : retraités, étudiants, salariés du privé ou fonctionnaires ;
- Nombre d'assurés par classe d'âge ;
- Nombre de bénéficiaires par assuré.

L'imputation d'un contrat de complémentaire santé à chaque individu de SRCV 2017 et d'EHS 2019 suit la même logique que le pseudo-appariement des enfants dans EHS 2014 et SRCV 2017. D'une part, il s'agit de stratifier la base des donneurs (ici l'enquête OC 2019) et la base des receveurs (ici les individus assurés de SRCV 2017 et d'EHS 2019) de la même manière. Ici, on utilise comme variables de stratification le type de contrat, son niveau de couverture (qui permet d'approximer sa qualité¹⁹) et l'exclusivité du bénéfice du contrat pour une catégorie professionnelle. Ces deux dernières variables ne sont pas disponibles dans EHS 2019 ou SRCV 2017 et doivent être imputées. D'autre part, plusieurs probabilités de tirage différentes peuvent être calculées pour chaque contrat de l'enquête OC 2019 ; ici la probabilité de tirage est différenciée selon l'âge de l'individu receveur et selon que celui-ci a des ayants-droits ou non.

Tableau 10 Disponibilité des variables de stratification et de pondération, dans l'enquête OC 2019 et dans EHS 2019 ou SRCV 2017, avant toute imputation

		OC 2019	EHS 2019 ou SRCV 2017
Stratification	Type de contrat	X	X
	Qualité du contrat	X	
	Contrat au bénéfice exclusif d'une catégorie professionnelle	X	
Pondération	Répartition des ouvrants droit par âge	X	X
	Proportion des ouvrants droit avec des assurés secondaires	X	X

Lecture > La qualité de chaque contrat de l'enquête OC est approchée par le biais d'un score calculé par la DREES, qui synthétise les garanties déclarées dans l'enquête. En revanche on n'observe pas dans EHS 2019 ni SRCV 2017 le niveau de garanties ou la qualité des contrats qui couvrent les individus enquêtés.

Source > DREES, enquête OC 2019 ; INSEE, SRCV 2017 ; DREES, EHS 2019.

Stratification de l'enquête OC et de la base des ouvrants droit

Les strates sont définies ici par le croisement du type de contrat, de la qualité du contrat mesurée par le score, et de l'exclusivité pour certaines catégories de la population.

Les contrats sont regroupés en trois groupes selon leur qualité : contrats les moins couvrants, contrats de milieu de gamme, et contrats les plus couvrants. Les trois catégories rassemblent chacune un tiers des contrats collectifs et un tiers des contrats individuels. Les seuils appliqués au score, qui délimitent ces trois catégories, sont donc différenciés pour les contrats individuels et collectifs²⁰. En revanche, la qualité du contrat souscrit par un ouvrant droit n'est pas observée dans EHS 2019 ni dans SRCV 2017. Il faut donc procéder à son imputation.

Grâce à l'enquête OC, la proportion d'assurés bénéficiant de contrats les moins couvrants, de milieu de gamme, et les plus couvrants est connue pour chaque tranche d'âge et chaque type de contrat souscrit. Par ailleurs les

¹⁸ L'échantillon de SRCV 2017 comprend tous les membres de chaque ménage, donc pour chaque contrat souscrit son ouvrant droit et ses ayants droit. Il suffit donc d'imputer un contrat aux ouvrants droit dans SRCV 2017. En revanche on n'observe qu'un individu par ménage dans EHS 2019, à qui l'on impute un contrat, qu'il soit ouvrant droit ou ayant-droit. Par la suite, on parlera de l'imputation de contrats « aux individus d'EHS 2019 et de SRCV 2017 » sans préciser davantage.

¹⁹ Voir méthode dans Lapinte *et al.*, 2024.

²⁰ Ces seuils sont 0,48 et 0,63 pour les contrats individuels, contre 0,67 et 0,86 pour les contrats collectifs. Si les mêmes seuils avaient été utilisés pour les contrats collectifs et individuels, comme dans Ines-Omar 2017, de sorte à les répartir en trois groupes égaux, les seuils auraient été 0,54 et 0,74.

individus de SRCV 2017 et EHS 2019 d'une même classe d'âge et de même type de contrat sont classés par ordre croissant de cotisations, et le niveau de qualité de leur contrat leur est attribué, conformément aux proportions par tranche d'âge et type de contrat observées dans l'enquête OC. Au préalable, les cotisations déclarées par les individus dans EHS 2019 et SRCV 2017 sont individualisées en prenant en compte le nombre et le type de bénéficiaires, afin d'obtenir le meilleur estimateur possible de la qualité du contrat. Le graphique 2 montre la répartition des souscripteurs entre trois gammes de qualité de contrat, définies grâce à la méthode de classification des contrats de la DREES, selon leur classe d'âge et leur niveau de vie. L'attribution des contrats s'appuie donc sur l'hypothèse que, à âge et type de contrat donné, les différences de qualité de contrat sont liées aux différences de cotisations versées.

EHS 2019 et SRCV 2017 n'indiquent pas non plus si les individus sont couverts par un contrat réservé à leur catégorie professionnelle. Il faut donc également imputer l'exclusivité ou non des contrats dont bénéficient les individus d'EHS 2019 et SRCV. Un calcul est fait pour savoir quelle part des retraités (par exemple) bénéficie de contrats exclusivement réservés aux retraités : pour chaque croisement type de contrat x qualité, la proportion de retraités dans SRCV ou EHS 2019 est mise en regard de la proportion d'assurés de ce croisement ayant un contrat réservé aux retraités dans l'enquête OC. Par exemple, si deux tiers des individus du croisement contrat individuel x qualité intermédiaire dans SRCV ou EHS 2019 sont des retraités, et qu'un tiers des bénéficiaires de contrats individuels et de qualité intermédiaire dans l'enquête OC 2019 bénéficient d'un contrat réservé aux retraités, cela signifie que la moitié des retraités de ce croisement dans SRCV ou EHS 2019 bénéficient d'un contrat qui leur est réservé. Cette probabilité est alors respectée lors de l'attribution des contrats de complémentaire santé : la moitié des retraités du croisement vont recevoir un contrat réservé aux retraités, tandis que l'autre moitié recevra un contrat ouvert à tous.

Probabilité de tirage d'un contrat

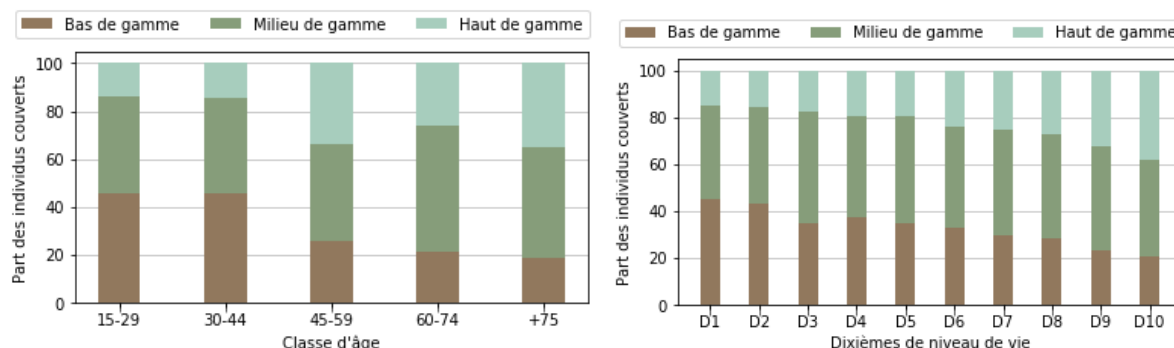
À l'issue de l'étape de stratification, la base des contrats et la base des individus d'EHS 2019 et SRCV 2017 sont toutes deux stratifiées selon le type, la qualité, et l'exclusivité du contrat. Au sein de chaque strate de la base des contrats, la probabilité de tirer un contrat donné est modulée selon le type d'individu receveur potentiel : huit probabilités différentes de tirage sont calculées pour chaque contrat selon l'âge du receveur, catégorisé en quatre classes d'âge, et selon que le receveur partage son contrat avec d'autres assurés. En effet, les contrats qui sont financièrement les plus avantageux pour les familles comptent en moyenne davantage de bénéficiaires par ouvrant droit, et inversement. Ne pas prendre en compte cet effet conduirait à surestimer les primes en attribuant à de nombreuses familles des contrats peu avantageux pour elles.

La probabilité de tirage est modulée selon l'âge en utilisant le nombre d'assurés principaux par tranche d'âge (0 – 24 ans, 25 – 59 ans, 60 – 64 ans, plus de 65 ans) déclaré dans l'enquête OC 2019. Pour chaque contrat de l'enquête OC la proportion d'assurés sans ayant droit peut être estimée grâce au nombre moyen de bénéficiaires par assuré. Cette proportion est ensuite utilisée pour moduler les probabilités d'attribution des contrats selon si l'assuré a ou non des ayants droit.

Chaque individu d'EHS 2019 et de SRCV 2017 ayant été placé dans la strate qui lui correspond, l'imputation à celui-ci d'un contrat donné parmi ceux de la strate correspondante dans l'enquête OC se fait au tirage au sort, en utilisant la pondération des contrats correspondant à sa classe d'âge et au fait qu'il partage ou non son contrat avec d'autres assurés.

Graphique 3 Répartition du niveau de couverture de contrat selon les revenus et l'âge du bénéficiaire dans EHS 2019, après imputation

En %



Lecture > 46 % des personnes couvertes par un contrat de complémentaire santé privé et âgées d'entre 60 et 74 ans ont un contrat dont le niveau de couverture est classifié comme « Milieu de gamme » (score entre 0,48 et 0,63 pour les contrats individuels, et entre 0,67 et 0,86 pour les contrats collectifs). Ce chiffre est de 48 % pour les personnes couvertes du troisième dixième de niveau de vie.

Champ > Ensemble des personnes de 15 ans ou plus couvertes par un contrat de complémentaire santé privée en France métropolitaine.

Source > DREES, EHS 2019 ; DREES, enquête OC 2019. Traitements DREES.

Dans le modèle Ines-Omar 2017, la part de bénéficiaires de contrats haut de gamme diminuait fortement après 60 ans, ce qui n'est plus le cas dans Ines-Omar 2019. Cela s'explique par la définition différente des classes de niveau de couverture des contrats dans les deux modèles : dans Ines-Omar 2017 les mêmes seuils de qualité étaient appliqués à tous les contrats, or les contrats collectifs ont des scores bien plus élevés en moyenne que les contrats individuels. Ainsi dans Ines-Omar 2017, pratiquement tous les contrats les moins couvrants étaient individuels, et presque tous les contrats les plus couvrants étaient collectifs. Comme la part de bénéficiaires de contrats collectifs chute après 60 ans, vers le passage à la retraite, la qualité moyenne des contrats imputés dans Ines-Omar 2017 chutait après 60 ans.

Au contraire Ines-Omar 2019 classe séparément les contrats individuels et collectifs, de sorte que les catégories haut, milieu et bas de gamme rassemblent chacune un tiers des contrats individuels et un tiers des contrats collectifs. Ceci évite un effet de composition et permet de comparer la qualité des contrats à différents âges. Il apparaît que plus un individu assuré est âgé, meilleure est la qualité du contrat qu'il souscrit en moyenne. Il y a une certaine logique à cela, comme les dépenses de santé augmentent avec l'âge.

■ CALCUL DES COTISATIONS DE COMPLÉMENTAIRE SANTÉ

À présent chaque individu d'EHIS 2019 et de SRCV 2017 s'est vu attribuer un contrat issu de l'enquête OC 2019. Les primes d'assurance (cotisations) déclarées dans SRCV 2017 et EHIS 2019 servent à estimer la qualité du contrat, permettant ainsi une imputation plus fine. Néanmoins, elles ne correspondent pas aux contrats attribués. Les informations de tarification disponibles dans l'enquête OC 2019 sont donc mobilisées pour simuler la prime associée au contrat. Ainsi les primes utilisées dans le modèle Omar sont plus directement corrélées au niveau de garanties imputées aux individus d'EHIS 2019 et SRCV 2017 à partir de l'enquête OC 2019. Le choix de s'appuyer sur ces informations semble validé par les bonnes propriétés des distributions de prime obtenues (voir *infra*). Pour les bénéficiaires d'un contrat collectif d'entreprise, la participation de l'employeur à la prime de complémentaire est enfin estimée au moyen du taux de participation moyen par secteur d'activité issu de l'enquête sur la protection sociale complémentaire en entreprise (PSCE) 2017.

Tarification des contrats dans l'enquête OC

L'enquête OC détaille le mode de tarification de chaque contrat et les montants de prime pour un cas-type d'assuré de 40 ans, en faisant ensuite varier ces caractéristiques une à une pour cet individu « de référence ». L'annexe 3 présente la partie du questionnaire de l'enquête OC qui y est relative. Les informations issues de l'enquête OC relatives au mode de tarification, qui seront utilisées dans le modèle Ines-Omar, sont :

- La prime de l'individu « de référence » : prime d'une personne de 40 ans vivant seule et sans enfant, rattachée au régime général de la Sécurité sociale, travaillant à temps plein, percevant un revenu correspondant à la tranche de revenus la plus faible de la grille salariale définie par l'organisme complémentaire, et résidant dans la zone géographique où les cotisations sont les plus élevées.
- La prime de l'individu de référence s'il avait non pas 40 ans, mais 20, 60, 75 ou 85 ans.
- La prime de l'individu de référence s'il habitait non pas la région où les cotisations sont les plus élevées, mais celle où les cotisations sont les plus faibles.
- La prime de l'individu de référence s'il avait non pas un revenu égal au salaire minimal, mais un revenu égal à deux fois le plafond mensuel de la Sécurité sociale de 2018 (6 622 € brut).
- La tarification associée aux enfants : pas de prise en charge des enfants, supplément de primes associé à chaque enfant, gratuité à partir d'un nombre d'enfants, ou forfait familial indépendant du nombre d'enfants.
- La prime de l'individu de référence s'il était marié à une femme de 40 ans, et vivait sans enfant, avec un enfant, deux enfants, ou trois enfants.

Retraitements des informations de cotisation dans l'enquête OC

Les données de l'enquête OC utilisées sont relativement brutes. Étant donné la complexité des différentes modalités de calcul des cotisations, un nombre important de retraitements préalables sont nécessaires.

- Pour 57 contrats, aucune information de cotisation n'est disponible, ni pour l'individu de référence, ni pour aucun autre type d'assuré. Ces contrats ne sont donc attribués à personne.
- Concernant les contrats individuels, il se peut que les cotisations à certains âges ne soient pas déclarées car le contrat est en théorie interdit aux personnes de cet âge. C'est notamment fréquemment le cas aux âges les plus élevés, ou inversement pour la cotisation à 20 ans. Étant donné que les classes d'âge utilisées lors de l'attribution des contrats dans EHIS 2019 ou SRCV 2017 ont des seuils différents des âges auxquels sont indiqués les primes de référence dans l'enquête OC 2019, il ne peut pas être évité que des contrats de ce type soient attribués à des individus auxquels ils ne sont pas destinés. Il est donc nécessaire de simuler une prime à tout âge. Cela est fait en s'appuyant sur la cotisation la plus proche déclarée, par exemple celle à 75 ans pour remplir celle à 85 ans, à laquelle est appliqué le ratio moyen observé sur les contrats individuels pour lesquels les deux variables sont remplies. Néanmoins, étant donné que l'attribution des contrats se fait en tenant compte du nombre d'assurés du contrat dans chaque classe d'âge, lequel doit par définition être très faible si le contrat n'est pas accessible à une classe d'âge spécifique, ce cas de figure est peu fréquent.
- Quelques contrats d'entreprise déclarant moduler leur tarification selon le revenu n'ont pas renseigné la cotisation pour l'individu de référence, tout en ayant renseigné une cotisation pour la tranche de revenus la plus élevée. Nous supposons dans ce cas que l'absence de réponse est liée à la non-pertinence de la question pour ce contrat, c'est-à-dire au fait que le contrat serait gratuit pour la tranche de revenus la plus basse. On renseigne alors une cotisation nulle pour l'individu de référence. Ce cas de figure ne concerne que cinq contrats de complémentaire santé issus de l'enquête OC, ce qui est très anecdotique et n'est donc pas de nature à modifier significativement les résultats issus du modèle.

Simulation fine des cotisations

La cotisation de complémentaire santé d'un ouvrant droit est obtenue en modulant la prime de référence du contrat en fonction des caractéristiques individuelles. Ces caractéristiques à prendre en compte sont :

- L'âge de l'ouvrant droit ;
- La localisation de l'ouvrant droit ;
- Les revenus de l'ouvrant droit ;
- Le nombre de bénéficiaires (adultes et enfants) et leur âge.

Âge • Entre deux âges auxquels les OCAM ont déclaré la cotisation de référence (20, 40, 60, 65, 75, 85 ans), l'hypothèse est faite que les primes évoluent de façon linéaire. En dessous de 20 ans et après 85 ans il est supposé que les primes n'évoluent pas. Ces hypothèses semblent vraisemblables par rapport à ce qui est pratiqué par les OCAM.

Localisation • Pour les OCAM qui modulent leurs primes en fonction de la région de résidence, la prime de leurs ouvrants droit est modulée selon leur lieu de résidence, en l'occurrence selon la zone d'études et d'aménagement du territoire (ZEAT) où ils habitent²¹. Il est fait l'hypothèse que les primes dépendent du reste à charge après remboursement de l'assurance maladie obligatoire (RAC AMO) opposable²² moyen dans la région (reflet du risque à assurer par les OCAM ; voir Lapinte *et al.*, 2024). Les données utilisées pour la modulation géographique sont les RAC AMO opposables moyens pour la tranche d'âge 40-50 ans, issus du SNDS 2018 : les ZEAT où les restes à charge opposables sont les plus élevés (le Nord et le pourtour méditerranéen) se voient attribuer des primes de complémentaire santé plus importantes et inversement (l'Est est la ZEAT avec les restes à charge opposables les plus faibles). Le niveau moyen des primes par ZEAT est estimé à l'aide de l'enquête EHIS 2019. L'hypothèse est faite que les primes moyennes par ZEAT évoluent de façon linéaire entre l'Est et le Nord selon le RAC AMO opposable moyen de la ZEAT.

Revenus • Certains OCAM modulent leurs primes selon les revenus bruts de l'ouvrant droit, notamment pour les contrats d'entreprise. La prime de référence correspond alors à celle pour un salarié à temps plein de la tranche de revenus la plus faible (au smic donc). Pour ces contrats, il est également indiqué la prime pour une personne rémunérée 6 622 euros brut par mois²³ (ou, à défaut, appartenant à la tranche de revenus la plus élevée). La hausse de la prime entre ces deux extrêmes est considérée linéaire. En-dessous du smic et au-delà de 6 622 euros brut par mois, il est considéré que la prime n'évolue plus.

Bénéficiaires • La plupart des contrats comportent un surcoût en cas d'extension du contrat à des ayants droit. Dans ce cas, il est indiqué quelle serait la prime de l'individu de référence s'il couvrirait son conjoint, son conjoint et un enfant, et deux enfants, et trois enfants (en faisant l'hypothèse qu'au-delà de trois enfants le coût du contrat ne change pas). Le surcoût individuel du conjoint et de chaque enfant peut être déduit de ces différents éléments. Chaque situation individuelle, même celle des familles monoparentales, peut ainsi être simulée en cumulant les surcoûts liés à chaque ayant droit à la prime de l'ouvrant droit. Il est fait comme hypothèses que le surcoût lié à la présence de tout autre adulte que l'ouvrant droit est constant, et identique au surcoût du conjoint de l'ouvrant droit.

La simulation de la prime se déroule de manière suivante. Dans un premier temps est calculée la contribution de l'ouvrant droit à la prime : la prime obtenue en fonction de l'âge fait office de prime de base, à laquelle les modulations de la prime selon le revenu ou le lieu de résidence sont appliquées sous forme de coefficients multiplicateurs.

$$\text{contrib. ouvrant droit} = \text{Prime de référence à l'âge de l'ouvrant droit} * \text{Correctif revenu} * \text{Correctif ZEAT de résidence}$$

Puis, dans un second temps sont calculées les contributions de chaque assuré à la prime totale simulée. Pour les adultes bénéficiaires, leur âge est pris en compte pour la modulation du surplus de prime, et non pas l'âge de l'ouvrant droit lui-même, généralement plus élevé, sans quoi les cotisations seraient surestimées. La localisation et les revenus de l'ouvrant droit sont toutefois pris en compte sous forme de multiplicateurs appliqués au surplus de prime.

Si le contrat propose un forfait famille et si un autre adulte est couvert :

$$\text{contrib. famille} = \text{Coût du forfait famille} * \text{Correctif revenu} * \text{Correctif ZEAT} * \text{Correctif âge du 1er adulte bénéficiaire}$$

²¹ Les ZEAT sont des regroupements de départements tels que « Bassin parisien », « Sud-Ouest » ou « Méditerranée ».

²² Le niveau régional moyen des primes déclarées dans EHIS 2019 est plus fortement corrélé avec le RAC AMO opposable moyen entre 40 et 50 ans qu'avec le RAC AMO global moyen entre 40 et 50 ans.

²³ Soit deux fois le plafond mensuel de la Sécurité sociale de 2018.

Si le contrat ne propose pas de forfait famille, la contribution des éventuels autres adultes bénéficiaires est calculée de la manière suivante :

- S'il y a un 1^{er} adulte bénéficiaire,

$$\text{contrib. 1}^{\text{er}} \text{ adulte} = \text{Surcoût adulte} * \text{Corr. revenu} * \text{Corr. ZEAT} * \text{Corr. âge 1}^{\text{er}} \text{ adulte}$$

- S'il y a un 2^e adulte bénéficiaire,

$$\text{contrib. 2}^{\text{e}} \text{ adulte} = \text{Surcoût adulte} * \text{Corr. revenu} * \text{Corr. ZEAT} * \text{Corr. âge 2}^{\text{e}} \text{ adulte}$$

- S'il y a un 3^e adulte bénéficiaire,

$$\text{contrib. 3}^{\text{e}} \text{ adulte} = \text{Surcoût adulte} * \text{Corr. revenu} * \text{Corr. ZEAT} * \text{Corr. âge 3}^{\text{e}} \text{ adulte}$$

Puis le surcoût lié à chaque enfant bénéficiaire :

- S'il y a un 1^{er} enfant bénéficiaire,

$$\text{contrib. 1}^{\text{er}} \text{ enfant} = \text{Surcoût du 1}^{\text{er}} \text{ enfant} * \text{Corr. revenu} * \text{Corr. ZEAT}$$

- S'il y a un 2^e enfant bénéficiaire,

$$\text{contrib. 2}^{\text{e}} \text{ enfant} = \text{Surcoût du 2}^{\text{e}} \text{ enfant} * \text{Corr. revenu} * \text{Corr. ZEAT}$$

- S'il y a au moins 3 enfants bénéficiaires,

$$\text{contrib. 3}^{\text{e}} \text{ enfant} = (\text{Nombre d'enfants} - 2) * \text{Surcoût du 3}^{\text{e}} \text{ enfant} * \text{Corr. revenu} * \text{Corr. ZEAT}$$

La prime totale est la somme de la contribution due à l'ouvrant droit et du forfait famille le cas échéant :

$$\text{prime} = \text{contrib. ouvrant droit} + \text{contrib. famille}$$

Ou bien la somme des contributions dues à l'ouvrant droit, aux autres adultes et aux enfants bénéficiaires :

$$\text{prime} = \text{contrib. ouvrant droit} + \text{contrib. 1}^{\text{er}} \text{ adulte} + \dots + \text{contrib. 1}^{\text{er}} \text{ enfant} + \dots$$

En fin de compte, un calage des primes est opéré selon la ZEAT de résidence. En effet, la majorité des OCAM déclarent ne pas moduler les primes selon la localisation, mais cela peut induire en erreur car de nombreux petits OCAM n'exercent en fait que dans une région précise. Ainsi, un OCAM n'exerçant qu'en Île-de-France pourrait avoir des cotisations bien supérieures à celles d'un OCAM n'exerçant qu'en Alsace, tandis que les deux OCAM déclareraient ne pas moduler leurs primes selon le lieu de résidence. Or la méthode d'attribution des contrats ne permet pas d'éviter qu'un contrat francilien ne soit attribué à un ouvrant droit Alsacien.

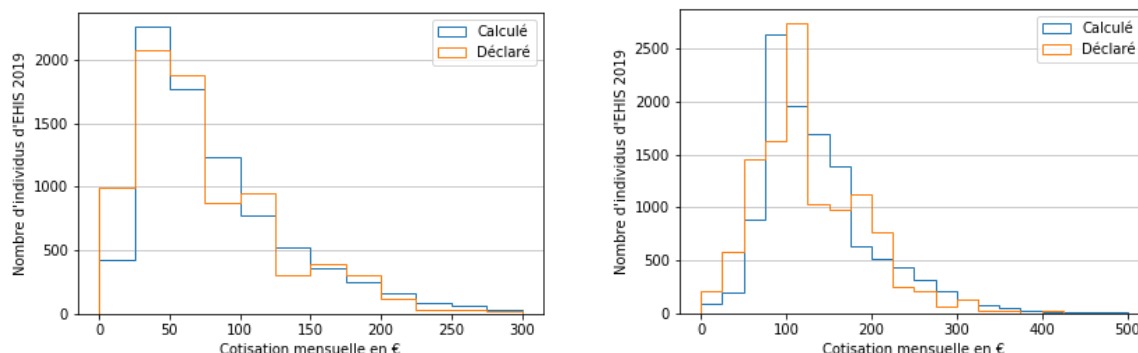
Pour éviter d'introduire un biais, il est fait comme si cette logique était suivie par tous les organismes, même s'ils ont déclaré ne pas moduler leurs cotisations selon le lieu de résidence. Un calage est effectué de sorte à conserver le même niveau global des primes simulées mais à restituer le niveau relatif des primes observé dans l'enquête EHS 2019. Ce calage est effectué séparément pour les contrats collectifs et individuels.

Vérifications

La distribution des primes calculées à partir des contrats imputés peut être comparée à celle des primes déclarées dans SRCV 2017 ou EHS 2019. Dans l'ensemble, la distribution des primes simulées est très proche de celle des primes déclarées, dans EHS 2019 comme dans SRCV 2017, et ce même en se restreignant à des sous-populations précises. Cette cohérence valide à la fois la fiabilité des primes déclarées dans SRCV et EHS 2019, et la méthode employée pour simuler des primes à partir de l'enquête OC. Le choix est fait de conserver la prime simulée à partir de l'enquête OC afin qu'il y ait cohérence entre la prime payée par un ménage et le niveau de garanties dont il bénéficie, lui aussi issu de l'enquête OC.

Le graphique 4 compare les distributions des primes déclarées dans EHS 2019 et de celles recalculées finement. Le champ est restreint aux contrats individuels. À gauche, pour les ouvriers-droits de 15 à 50 ans, à droite pour les ouvriers droit de plus de 65 ans. Pour ces deux populations, les distributions tout comme les moyennes sont très semblables. Il en va de même pour les autres sous-populations analysées : assurés n'ayant aucun ayant droit ; assurés ayant au moins un ayant droit ; personnes des trois premiers dixièmes de niveau de vie ou des trois derniers dixièmes.

Graphique 4 Distributions comparées des primes déclarées dans EHS 2019 et recalculées pour les détenteurs de contrats individuels de 15 à 50 ans (gauche) et de plus de 65 ans (droite)



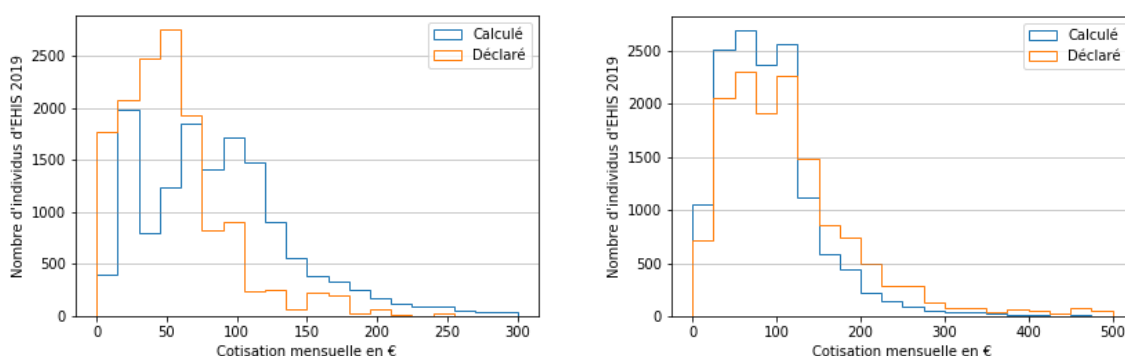
Lecture > Parmi les souscripteurs de contrats individuels âgés de 15 à 50 ans, 1 870 personnes enquêtées dans EHS 2019 ont déclaré une cotisation de complémentaire santé située entre 50 et 75 euros par mois. Elles seraient 1 239 à se situer dans cette tranche de cotisation d'après les calculs fins effectués dans le cadre du modèle.

Champ > Ensemble des bénéficiaires de contrats individuels de 15 à 50 ans (gauche) et de plus de 65 ans (droite).

Source > DREES, EHS 2019 ; DREES, enquête OC 2019. Traitements DREES.

Concernant les contrats d'entreprise, il existe en revanche un important décalage entre les distributions des primes déclarées dans EHS 2019 et des primes calculées (*graphique 5 gauche*). Les enquêtés ont déclaré combien ils paient pour leur complémentaire santé et non pas quel est son coût réel. Pour pouvoir effectuer la comparaison, il est impératif d'ajouter aux primes déclarées la part prise en charge par l'employeur. Les résultats de l'enquête Protection sociale complémentaire d'entreprise (PSCE) 2017 sont mobilisés pour affecter une part moyenne de prise en charge par l'employeur dépendante du secteur d'activité déclaré dans SRCV 2017 et EHS 2019 (Lapinte, *et al.*, 2024 ; Perronnin, 2019). À la suite de cette correction, les deux distributions sont nettement plus cohérentes entre elles (*graphique 5 droite*). Néanmoins, des divergences existent, avec notamment un certain nombre d'ouvrants droit dont la cotisation déclarée corrigée dépasse la cotisation simulée.

Graphique 5 Distributions comparées des primes déclarées dans EHS 2019 et recalculées pour les détenteurs de contrats d'entreprise avant et après prise en compte de la participation employeur



Note > À gauche, les primes déclarées sont directement issues d'EHS 2019, à droite les primes déclarées ont été gonflées de la participation employeur.

Lecture > Parmi les détenteurs de contrat d'entreprise, 1 123 enquêtés d'EHS 2019 ont déclaré une cotisation de complémentaire santé située entre 75 et 100 euros par mois (*graphique de gauche*). Les simulations effectuées dans le cadre du modèle, et qui incluent la part financée par l'employeur, comptabilisent 2 370 ouvrants droit dans cette tranche de cotisations.

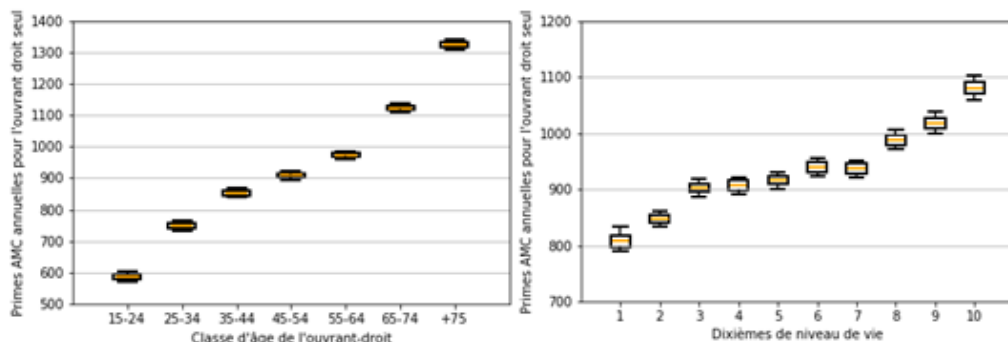
Champ > Ensemble des bénéficiaires de contrats d'entreprise.

Source > DREES, EHS 2019 ; DREES, enquête OC 2019. Traitements DREES.

La méthode employée pour l'imputation des contrats comporte une part d'aléa : pour chaque ouvrant droit, un contrat de sa strate est tiré au sort. Les garanties et le montant des cotisations retenus pour ce dernier sont alors ceux du contrat tiré. Il est légitime de s'inquiéter qu'un autre tirage au sort conduise à avoir des résultats différents. Afin d'étudier la variabilité des résultats, l'opération d'attribution des contrats et de simulation des cotisations a été effectuée 100 fois pour le modèle Omar 2017 (Fouquet, 2021), ayant permis d'estimer 100 fois les moyennes de cotisation par dixième de niveau de vie et par classe d'âge (*graphique 6*). Ce graphique permet de montrer que la variabilité par classe d'âge est très faible et qu'il n'y a pas de risque d'obtenir des moyennes erronées. Concernant

les cotisations par dixième de niveau de vie, la variabilité est également assez faible, l'écart interquartiles étant de l'ordre de quelques dizaines d'euros par an. Néanmoins, quelques valeurs peuvent s'éloigner fortement au-delà des premier et neuvième déciles affichés dans le graphique, ce qui fait courir un léger risque que la prime simulée moyenne par dixième de niveau de vie qui sera retenue *in fine* dans le modèle Ines-Omar soit écartée jusqu'à hauteur d'une centaine d'euros de la valeur médiane des simulations.

Graphique 6 Boîtes à moustaches de répartition des 100 itérations, moyennes de cotisation par classe d'âge (gauche) et par dixième de niveau de vie (droite) dans le modèle Omar 2017



Lecture > Sur 100 itérations de l'algorithme d'attribution des contrats de complémentaire santé de l'enquête OC 2016 aux individus de SRCV 2017 dans le modèle Omar 2017, concernant la cotisation moyenne des individus du 6^e dixième de niveau de vie, le premier décile se situe à 923 euros par an, le premier quartile à 930 euros par an, la médiane à 939 euros, le troisième quartile à 947 euros et le neuvième décile à 954 euros.

Champ > Ensemble des personnes en ménage ordinaire de France métropolitaine.

Source > INSEE, SRCV 2017 ; DREES, enquête OC 2016. Traitements DREES.

ESTIMATION DES REMBOURSEMENTS DE COMPLÉMENTAIRE SANTÉ

À présent, les individus des enquêtes EHS 2019 et SRCV 2017 se sont vu imputer des contrats de complémentaires santé, dont l'enquête OC recueille les garanties détaillées par poste de soin. Par ailleurs leurs dépenses de santé détaillées par poste de soin sont disponibles : les dépenses de santé exactes des individus de 15 ans ou plus d'EHS 2019 sont connues *via* leur appariement avec le SNDS, et les individus de moins de 15 ans de SRCV 2017 se sont vu imputer des dépenses de santé issues du SNDS 2014. Le niveau de finesse de ces deux ensembles de données, imputées ou réelles, permet une microsimulation détaillée par poste de soin des remboursements que ces individus perçoivent de la part de leur complémentaire santé.

Modes d'expression des garanties

Pour chaque poste de soins, l'enquête OC propose aux complémentaires santé de détailler les garanties proposées sous plusieurs formats possibles. Le remboursement peut ainsi s'exprimer en pourcentage de la base de remboursement de la sécurité sociale (BRSS), proportion pouvant excéder 100 % si le contrat rembourse des dépassements d'honoraires. Il peut aussi s'exprimer en pourcentage des frais réels, c'est-à-dire en pourcentage de l'ensemble de la dépense de santé. Il est également fréquent que le remboursement puisse s'exprimer sous forme d'un forfait monétaire par acte, ou d'un forfait annuel. Certains contrats proposent un remboursement forfaitaire pour la dépense qui excède le pourcentage remboursé sur la BRSS ou sur les frais réels. Par exemple, un contrat peut rembourser l'intégralité du ticket modérateur d'une visite chez le spécialiste ainsi que les dépassements d'honoraires à hauteur de 20 euros.

Tableau 11 Modes d'expression des garanties possibles dans l'enquête OC 2019 selon le poste de soins

Poste de soins	RAC opposable si responsable	% de la BRSS	% des frais réels	Forfait monétaire acte par acte	Forfait annuel	Plafond éventuel
Hôpital frais de séjour	x					
Honoraires de praticiens hospitaliers		x	x	x		
Généralistes et auxiliaires	x		x	x		
Médecins spécialistes		x	x	x		
Radiologie		x	x			
Médicaments		x				
Analyses biologiques		x	x			
Transports et matériel	x					
Monture et verres		x	x		x	
Lentilles					x	
Soins dentaires conservateurs		x	x	x		
Prothèses céramo-métalliques visibles, Dentiers, Bridges, Inlay-Onlay, Inlay-Core		x	x	x		x
Orthodontie		x	x		x	x
Implantologie / parodontologie				x		
Prothèses auditives		x	x		x	

Lecture > Les dépenses de santé chez le médecin spécialiste peuvent être remboursées sous la forme soit d'un pourcentage de la BRSS, soit d'un pourcentage de l'ensemble de la dépense, soit d'un forfait monétaire pour chaque acte. Le mode de remboursement opéré dans le modèle dépend de ce qui a été déclaré par l'OCAM dans l'enquête OC pour le contrat imputé.

Source > DREES, enquête OC 2019.

Concernant l'optique, les OCAM ont la possibilité de rembourser l'ensemble monture plus verres avec un unique forfait, lequel diffère selon la nature simple ou complexe des verres achetés. Dans les très rares cas où l'individu aurait consommé un verre simple et un verre complexe sur une même monture, l'ensemble est considéré comme étant de l'optique complexe. Enfin, pour l'orthodontie, les organismes complémentaires peuvent préciser un plafond

semestriel de remboursement ne pouvant pas être dépassé, traité comme un forfait annuel. Le tableau 11 récapitule les différents modes d'expression des garanties possibles dans l'enquête OC pour chaque poste de soins.

Niveau de détail des simulations

Le plus fin niveau de détail des postes de soins disponible dans l'enquête OC est mobilisé lors des simulations de remboursements. En particulier, pour les postes où les garanties peuvent être exprimées sous forme de forfait monétaire acte par acte, les dépenses de santé ont été conservées à un niveau de détail acte par acte lors de l'agrégation des données du SNDS. De même, la simulation respecte le niveau de détail proposé par l'enquête OC sur les garanties appliquées aux consultations de spécialiste, lesquelles sont différenciées selon que la visite respecte ou non le parcours de soins et selon que le spécialiste adhère ou non à l'Option de pratique tarifaire maîtrisée (Optam).

Pour certains postes de soins, l'enquête OC ne donne aucune précision sur les garanties appliquées : médecins généralistes, sages-femmes, auxiliaires, transports, matériel médical, soins dentaires conservateurs. Pour ces postes, il est considéré que les contrats responsables remboursent l'intégralité du RAC opposable mais ne remboursent pas les franchises médicales. Les dépassements d'honoraires éventuels sont traités comme s'il s'agissait de consultations de spécialistes. Concernant les contrats non responsables, lesquels couvrent moins de 2 % de la population assurée par une complémentaire privée, l'enquête OC n'indique pas le niveau de remboursement offert pour ces postes. Il est alors supposé que ces contrats ne remboursent pas le RAC opposable.

Les organismes complémentaires ont la possibilité de déclarer des niveaux de garantie différents selon que le soin a été consommé dans ou hors du réseau de professionnels qui leur est propre. Il n'est bien entendu pas possible de savoir si une certaine dépense de santé observée dans les données a été effectuée dans ou hors du réseau de soins lié à la complémentaire santé du consommant. L'hypothèse faite est que l'ensemble des dépenses de santé ont lieu au sein du réseau lorsque celui-ci existe. Ceci est en accord avec l'objectif d'étudier le remboursement maximal auquel les individus peuvent prétendre s'ils font les démarches idoines.

Remboursements de la CMU-C et de l'ACS

L'appariement d'EHIS 2019 avec le SNDS comporte déjà des informations sur les remboursements supplémentaires effectués par la CMU-C et par les OCAM au titre de contrats ACS. Il n'est donc pas toujours nécessaire de simuler les remboursements supplémentaires de la CMU-C et des OCAM au titre de contrats ACS pour leurs bénéficiaires dans EHIS 2019, si ceux-ci sont repérés comme bénéficiaires dans le SNDS.

En revanche il est nécessaire de calculer les remboursements supplémentaires de la CMU-C pour les individus auxquels a été imputé le bénéfice de la CMU-C mais qui ne sont pas repérés comme tels dans le SNDS. Pour ceux-là, on ne dispose pas de remboursements supplémentaires de la CMU-C. Il est également nécessaire de compléter ces remboursements pour les individus bénéficiaires de la CMU-C dans le SNDS, mais sans montant positif de remboursement de la part de la CMU-C.

Rappelons qu'à l'issue de la phase de redressement des déclarations de CMU-C dans l'échantillon d'EHIS 2019 apparié avec le SNDS, 783 individus de 15 ans ou plus d'EHIS 2019 sont identifiés comme bénéficiaires de la CMU-C, dont 546 consomment des soins en 2019 et sont reconnus comme tels dans le SNDS. 64 individus ne consomment pas de soins en 2019 : il est donc nécessaire de calculer intégralement les remboursements supplémentaires de la CMU-C pour 173 individus d'EHIS 2019. De même, 245 individus de 15 ans ou plus d'EHIS 2019 sont identifiés comme bénéficiaires de l'ACS, dont 209 sont reconnus comme tels dans le SNDS ; il est donc nécessaire de calculer intégralement les remboursements supplémentaires des OCAM au titre des contrats ACS pour 36 individus d'EHIS 2019.

En revanche, l'appariement d'EHIS 2014 avec le SNDS ne comprend pas d'informations sur les remboursements supplémentaires effectués par les OCAM au titre des contrats ACS. Ces variables ne sont donc pas imputées aux individus de moins de 15 ans de SRCV 2017 lors du pseudo-appariement avec EHIS 2014. Il est donc nécessaire de calculer tous les remboursements supplémentaires des OCAM au titre des contrats ACS pour les moins de 15 ans d'Omar 2019.

Remboursement des équipements dentaires et optiques par la CMU-C

La simulation des remboursements des équipements auditifs, dentaires et optiques par la CMU-C obéit à une logique différente de celle des remboursements de consultations ou de séjours hospitaliers. Cette partie traite uniquement des remboursements des équipements optiques et dentaires par la CMU-C, dans la mesure où le nombre d'équipements auditifs consommés par des bénéficiaires de la CMU-C, qui sont majoritairement jeunes, est extrêmement faible.

Principe de remboursement

L'État a institué à partir de 2014 des paniers d'équipements soumis à un prix limite de vente, réservés aux bénéficiaires de la CMU-C, et sur lesquels ceux-ci n'ont pas de reste à charge. À partir de 2019 est progressivement entrée en vigueur la réforme du « 100 % santé », qui permet l'accès à un panier sans reste à charge pour l'ensemble des bénéficiaires d'une complémentaire santé (publique ou privée) en optique, en prothèses auditives et dentaires (voir par exemple Lapinte *et al.*, 2024). De ce fait, plusieurs paniers ont été en vigueur tour à tour sur l'année 2019 pour les bénéficiaires de la CMU-C :

- En optique : un premier panier existant depuis 2014 a été en vigueur jusqu'en octobre 2019. Un nouveau panier d'équipements optiques a été instauré en novembre 2019, courant jusqu'à la fin de l'année et l'entrée en vigueur du 100% santé pour les équipements optiques au 1er janvier 2020.
- En dentaire : un panier de prothèses dentaires a été instauré en janvier 2019, en remplacement d'un panier instauré en 2017. Il a été complété par un deuxième panier en avril 2019, puis par un troisième panier en novembre 2019.

Chaque panier spécifie une base de remboursement de la sécurité sociale (BRSS) et un prix limite de vente (PLV) pour chaque bien qu'il contient. La Sécurité sociale prend à son compte 70 % de la BRSS (60 % en optique) ; la CMU-C prend à son compte 30 % de la BRSS (40 % en optique) plus l'écart entre la BRSS et le PLV.

Imputation d'un prix limite de vente

Or le PLV d'un équipement ne figure pas dans le SNDS. Pour imputer le PLV d'un équipement consommé par un bénéficiaire de la CMU-C, on recherche sa BRSS, fournie par le SNDS, dans le panier sans reste à charge de la CMU-C. Dans certains cas, on ne retrouve pas la BRSS de l'équipement consommé parmi les BRSS des biens du panier. Cela peut signifier que l'individu a consommé en dehors du panier CMU-C, s'exposant ainsi à un reste à charge positif. Cela peut aussi signifier que l'individu qui, dans le modèle, est compté comme bénéficiaire de la CMU-C, ne l'est pas en réalité et n'a pas pu consommer un bien du panier.

Dans ces cas-là, il est néanmoins décidé d'imputer un PLV pour les biens dont la BRSS ne correspond à celle d'aucun bien des paniers : c'est le cas pour 12 % des équipements optiques consommés par des bénéficiaires de la CMU-C, et pour 22 % des prothèses dentaires consommées par ceux-ci.

Complétion du remboursement de la CMU-C

Le SNDS fournit une variable qui comprend les remboursements de la CMU-C, pour les équipements comme pour les autres dépenses de santé. Cette variable est renseignée et valable pour les individus qui sont effectivement bénéficiaires de la CMU-C, et pour les dépenses de santé effectuées pendant une période de souscription de la CMU-C. Nous préférons ne pas altérer les informations provenant du SNDS.

Ainsi dans Ines-Omar 2019 la variable « REMB_CMU », qui contient les remboursements supplémentaires de la CMU-C, n'est redressée que si $REMB_CMU=0$ et le total des remboursements observés dans le SNDS est strictement inférieur au PLV. Cela représente 20 % des équipements optiques consommés par des bénéficiaires de la CMU-C, et 16 % des prothèses dentaires consommées par des bénéficiaires de la CMU-C. Dans ces cas-là, REMB_CMU se voit imputer $0,7*BRSS + (PLV - BRSS)$ pour les prothèses dentaires, et $0,6*BRSS + (PLV - BRSS)$ pour les équipements optiques.

Dans quelques cas, l'imputation porte à la fois sur le PLV de l'équipement, qui ne correspond à celui d'aucun bien des paniers d'équipements, et sur REMB_CMU, qui est nulle. C'est le cas de 3 % des consommations d'équipements d'optique par des bénéficiaires de la CMU-C et de 5 % des prothèses dentaires consommées par des bénéficiaires de la CMU-C. Ces cas de figure correspondent à des individus considérés comme bénéficiaires de la CMU-C dans le modèle, mais qui n'en sont pas dans le SNDS au moment de leur consommation d'équipement, ou qui ont consommé en dehors du panier. Ainsi, le modèle simule ce qu'aurait été le remboursement théorique de la CMU-C pour une telle consommation de soins.

Ce procédé est cohérent avec la simulation des remboursements par les complémentaires privées. En effet, on estime le remboursement des OCAM sans tenir compte du fait que certaines dépenses n'ont pas été portées au remboursement des OCAM. Le modèle conduit à estimer un remboursement maximal théorique de la part des OCAM comme de la CMU-C, quitte à surestimer un peu les remboursements complémentaires.

Les tableaux 12 et 13 présentent le bilan des cas de figures possibles et la proportion des équipements concernés en optique :

Tableau 12 Bilan des traitements réalisés sur les équipements optiques consommés par les bénéficiaires de la CMU-C

Remboursement initial AMO + CMU-C	Remboursement de la CMU-C = 0	Remboursement de la CMU-C > 0
< PLV	Le remboursement de la CMU-C est manquant => imputation. 20 % des cas.	Incohérence entre le PLV imputé et les remboursements dans le SNDS. On ne corrige pas le remboursement de la CMU-C, car il est issu du SNDS. 20 % des cas.
= PLV	-	PLV et remboursements cohérents. 45 % des cas.
> PLV	-	Incohérence entre le PLV imputé et les remboursements dans le SNDS. On ne corrige pas le remboursement de la CMU-C, car il est issu du SNDS. 15 % des cas.

Tableau 13 Bilan des traitements réalisés sur les prothèses dentaires consommées par les bénéficiaires de la CMU-C

Remboursement initial AMO + CMU-C	Remboursement de la CMU-C = 0	Remboursement de la CMU-C > 0
< PLV	Le remboursement de la CMU-C est manquant => imputation. 16 % des cas.	Incohérence entre le PLV imputé et les remboursements dans le SNDS. On ne corrige pas le remboursement de la CMU-C, car il est issu du SNDS. 82 % des cas.
= PLV	Il semblerait que le remboursement de la CMU-C ait été compté avec celui de l'AMO. Pas de correction. 2 % des cas.	-
> PLV	-	-

Il est à noter qu'à la suite de ces corrections, le reste à charge après remboursement de l'AMO et de la CMU-C reste positif sur une partie des équipements auditifs et optiques consommés par des bénéficiaires de la CMU-C. C'est le cas de 38 % des équipements optiques, et de 8 % des équipements dentaires ; alors le plus souvent le PLV est inférieur au montant effectif de la dépense. Cet écart entre le PLV imputé et le montant de la dépense peut être dû au fait que l'équipement a été consommé en dehors du panier CMU-C, ou au fait qu'un individu a bénéficié de la CMU-C pendant une partie de l'année seulement et a consommé des équipements durant le reste de l'année. Cet écart, pour les individus de moins de 15 ans, peut aussi s'expliquer par le fait que le montant des dépenses de santé qui leur ont été imputées à partir d'EHIS 2014 – SNDS ne correspond pas aux prix limites de ventes figurant dans les paniers d'équipements de 2019.

En définitive, le fait que des bénéficiaires de la CMU-C supportent un reste à charge sur des équipements optiques ou dentaires n'invalide pas le modèle, puisque cela peut s'expliquer entre autres par des consommations en dehors du panier CMU-C. Le modèle vise avant tout à estimer la prise en charge maximale et théorique de la CMU-C sur la consommation de ses bénéficiaires. De plus, en l'absence de source externe au modèle sur l'ampleur des restes à charge des bénéficiaires de la CMU-C sur leurs achats d'équipements, il est préférable de ne pas procéder à des corrections hasardeuses de ces restes à charge positifs.

■ LE FINANCEMENT DE L'ASSURANCE MALADIE OBLIGATOIRE

À ce stade de la construction du modèle il est possible de simuler les dépenses de santé, les remboursements par l'AMO, les remboursements versés par les OCAM - et donc ce qui reste à la charge des ménages après intervention des complémentaires - et les cotisations de complémentaire santé. Le modèle Omar est donc achevé. Néanmoins, il ne permet pas à lui seul d'avoir une vision d'ensemble de la redistribution opérée par le système de santé. Pour ce faire, un pseudo-appariement au modèle Ines 2019, géré conjointement par l'Insee, la DREES et la CNAF, est nécessaire afin de connaître la contribution de chaque individu au financement de l'AMO (voir encadré 1).

Encadré 1 Le modèle Ines

Le modèle Ines, développé par la DREES, l'Insee et la Caisse nationale des allocations familiales (CNAF), applique la législation sociofiscale à un échantillon représentatif de la population vivant en France métropolitaine en logement ordinaire (ne couvrant donc pas les personnes vivant en institution, comme par exemple les étudiants vivant en résidence étudiante ou les personnes âgées ou dépendantes vivant en établissement).

Il permet de simuler, pour chaque ménage, les contributions au financement de l'AMO. Les principales contributions prises en compte sont :

- la part des cotisations sociales à la charge des employeurs affectée au risque maladie (hors part correspondant au financement des indemnités journalières) ;
- les cotisations des indépendants et la contribution sociale généralisée (CSG) affectées au risque maladie ;
- la contribution pour le remboursement de la dette sociale (CRDS) ;
- la part des prélèvements indirects affectée au financement de l'AMO (TVA, taxes sur le tabac, TSA, etc.).

Le modèle est adossé à l'enquête Revenus fiscaux et sociaux (ERFS) de l'Insee qui réunit les informations sociodémographiques de l'enquête Emploi, les informations administratives de la CNAF et le détail des revenus déclarés à l'administration fiscale pour le calcul de l'impôt sur le revenu.

Pseudo-appariement au modèle Ines 2019

Le sens de l'imputation : les ménages Ines sont le socle du modèle Ines-Omar

Deux stratégies d'imputation sont possibles : soit transmettre aux individus du modèle Omar des données de contribution au financement de l'AMO issues d'un donneur côté Ines ; ou, inversement, transmettre aux individus du modèle Ines les dépenses de santé et les éléments de complémentaire santé issus du modèle Omar. La première méthode a l'avantage de transférer un plus faible nombre de variables. Néanmoins, le modèle Ines comptant près de 120 000 individus, soit environ sept fois plus d'individus que le modèle Omar (17 000 individus), la seconde méthode apporte plus de puissance statistique en permettant de mobiliser l'ensemble des données d'Ines.

En outre, les méthodes d'imputation reposant nécessairement sur une part de hasard, le fait que chaque individu d'Omar soit apparié en moyenne à sept individus d'Ines permet de réduire l'influence de l'aléa par rapport à une solution inverse où seul un individu d'Ines sur sept aurait été apparié à un individu d'Omar. C'est donc la seconde option qui est retenue. Le modèle Omar peut ainsi être considéré comme un module complémentaire du modèle Ines, d'où l'appellation « Ines-Omar ».

Étant donné que certaines étapes de construction du modèle Omar comportent également de l'aléa (attribution d'un contrat de complémentaire santé et pseudo-appariement EHS 2014-SRCV 2017 notamment), cette logique de « *bootstrap* » est poussée un cran plus loin. Au point de départ de la construction, la base EHS 2019 et la population des individus de moins de 15 ans de SRCV 2017 sont dupliquées en un total de cinq exemplaires pour obtenir une base d'une taille plus proche de celle du modèle Ines. Le nombre d'itérations du *bootstrap* a été fixé à cinq, et non à sept, par continuité avec Ines-Omar 2017. Les cinq individus de moins de 15 ans identiques de SRCV 2017 vont alors connaître des parcours divergents : pseudo-appariement avec des individus d'EHS 2014 différents, attribution de contrats de complémentaire santé différents, recours ou non à l'ACS lorsqu'il y a éligibilité, etc. Le parcours des cinq individus identiques d'EHS 2019 va simplement diverger lors de l'attribution d'un contrat de complémentaire santé²⁴. Au moment du pseudo-appariement avec Ines, ces cinq individus représentent donc

²⁴ Le modèle Ines-Omar 2019 comporte moins d'aléa que le modèle Ines-Omar 2017, notamment du fait que l'appariement EHS 2019 - SNDS fournisse les vraies dépenses de santé des 15 ans et plus, et que les bénéficiaires de l'ACS dans l'échantillon d'EHS 2019 soient connus et non tirés au sort.

autant de réalisations divergentes de quelques-uns des multiples parcours possibles, réduisant par conséquent quelque peu l'influence du hasard dans la construction du modèle d'ensemble final.

Le pseudo-appariement par « hot-deck »

L'imputation de données d'Omar aux individus d'Ines étant actée, il reste à choisir la méthode d'imputation. Le nombre de variables à transférer est élevé et ces variables interagissent fortement entre elles : les dépenses de santé ont une structure très différente en cas d'ALD ou de recours à la CMU-C par exemple ; la cotisation de complémentaire santé et les remboursements perçus sont directement corrélés ; la base de remboursement doit correspondre à la somme du ticket modérateur et du remboursement AMO ; le reste à charge AMC doit correspondre au reste à charge opposable moins les remboursements AMC, etc. Des méthodes de régression ou d'apprentissage statistique qui chercheraient à imputer dans Ines les variables une par une par simulation perdraient partiellement ces corrélations, lesquelles sont essentielles aux analyses qui seront menées à l'aide du modèle. C'est pourquoi il est choisi de pseudo-apparier des individus et de transmettre telles quelles l'ensemble des données d'un donneur d'Omar vers un receveur d'Ines.

Les données à transmettre sont de plusieurs ordres : état de santé, consommations de soins et couverture complémentaire. Tout en étant partiellement corrélés, ces éléments ont une part d'indépendance. Or les méthodes d'appariement supervisées par un algorithme d'apprentissage statistique, comme celles présentées en annexe 5, nécessitent de sélectionner une variable expliquée à l'exclusion des autres. Certes, il serait possible d'expliquer plusieurs variables simultanément mais cela rendrait l'algorithme particulièrement gourmand en données. Or l'échantillon disponible est de dimension plutôt restreinte (17 000 individus dans Omar) et les variables communes aux deux bases à pseudo-apparier sont certes nombreuses mais très redondantes : la plupart concernent l'estimation très fine des revenus ou du statut d'occupation. C'est pourquoi les méthodes d'appariement supervisées par un algorithme d'apprentissage ne semblent pas appropriées pour ce pseudo-appariement.

À l'inverse, la méthode du « hot-deck » - celle utilisée pour l'appariement d'Ines et d'Omar - n'est pas supervisée, mais laisse le statisticien choisir les variables qui lui semblent le mieux expliquer les variables à imputer, ainsi que leur ordonnancement. Cette démarche est plus adaptée à un contexte où de nombreuses variables sont à imputer et que peu de variables disponibles à la fois dans les bases des « donneurs » et des « receveurs » peuvent les expliquer. Le hot-deck consiste à stratifier la base des donneurs et celle des receveurs de la même manière, au moyen des variables explicatives communes choisies. Chaque receveur d'une strate donnée est ainsi associé aux individus appartenant à la strate équivalente dans la base des donneurs, avec lesquels il partage un certain nombre de caractéristiques. L'appariement consiste à tirer aléatoirement, pour chaque receveur, un donneur parmi la strate équivalente.

Il est donc nécessaire de sélectionner les variables de stratification. Cinq variables communes aux deux modèles sont sélectionnées, les plus importantes pour les analyses envisagées : type de couverture complémentaire santé, niveau de vie par vingtièmes, classe d'âge en tranches décennales, sexe, statut d'occupation (salarié du secteur privé, retraité, étudiant, etc...). Cette dernière variable ne permet que de distinguer les personnes à la CMU-C et à l'ACS, la couverture privée étant inconnue dans le modèle Ines. Le recours à la CMU-C et à l'ACS a été attribué aléatoirement dans Ines 2019 aux personnes éligibles, en respectant les proportions de bénéficiaires à âge et sexe donnés²⁵.

Afin de réduire l'influence de l'aléa, les strates parmi lesquelles le tirage est effectué doivent avoir une taille minimale. Ce seuil minimal est laissé à l'appréciation du statisticien. Les strates obtenues par le croisement des cinq variables de stratification (type de couverture x niveau de vie x classe d'âge x statut d'occupation x sexe) qui seraient de taille insuffisante devront être fusionnées, soit entre strates de taille insuffisante jusqu'à atteindre la taille minimale, soit avec des strates de taille suffisante. Il existe plusieurs manières de fusionner les strates de taille insuffisante, et parmi lesquels le statisticien peut choisir. L'annexe 6 présente la méthode usuelle de traitement des strates de taille insuffisante, et une nouvelle méthode, employée dans le modèle Ines-Omar 2019.

Comme le montre le tableau 22 en annexe 6, avec la méthode de traitement des strates employée dans Ines-Omar 2019, le hot-deck apparie près des deux tiers des individus d'Ines en conservant l'ensemble des cinq variables de stratification, et 97 % d'entre eux avec au moins les variables de type de couverture complémentaire, de niveau de vie et d'âge.

Il est important de préciser que le pseudo-appariement est effectué entre individus et non entre ménages, dans Ines-Omar 2017 comme 2019. Il n'y a donc dans Ines-Omar aucune garantie que la structure de couverture complémentaire soit cohérente au sein d'un ménage : un ménage peut par exemple se composer d'une personne considérée ayant droit d'un contrat d'entreprise et d'une autre couverte par un contrat individuel. Cela altère la corrélation entre la couverture complémentaire des individus d'un même ménage, et donc également la corrélation entre leurs données de santé. Ceci n'est pas problématique tant que l'on mène des études au niveau individuel, ou

²⁵ Ces proportions par âge et par sexe sont issues de SRCV 2017 pour les bénéficiaires de la CMU-C de moins de 15 ans, d'EHIS 2019 pour les bénéficiaires de la CMU-C de 15 ans ou plus, et du SNDS 2018 pour l'ACS. On veille par ailleurs à ce que la répartition des bénéficiaires de la CMU-C et de l'ACS entre vingtièmes de niveau de vie soit conforme à celle observée dans Omar 2019.

si l'on étudie les moyennes de variables imputées depuis Omar. Il n'est en revanche pas possible d'étudier la distribution entre ménages des variables de santé imputées, comme les RAC AMC. Par ailleurs, il convient de produire des statistiques dans Ines-Omar, au niveau individuel comme au niveau du ménage, sur des sous-populations délimitées au moyen des variables d'appariement entre Ines et Omar (couverture complémentaire, âge, niveau de vie, sexe, statut professionnel). *A contrario*, il n'est pas recommandé avec Ines-Omar, par exemple, de produire des statistiques au niveau individuel par ZEAT ou par diplôme, ou bien au niveau du ménage par configuration familiale. Notons que le modèle Omar 2019 ne permet pas du tout de mener des études au niveau du ménage, puisque les individus qui y sont observés ne sont pas organisés en ménages.

Calages macro-économiques

Le modèle Ines-Omar nécessite un calage des montants de dépense de santé pour trois raisons. En premier lieu, les données de santé des individus de moins de 15 ans du modèle datent de 2014, tandis que le modèle décrit la situation de 2019. Étant donné la croissance rapide des dépenses de santé, il est impératif de mettre à jour les montants au niveau de 2019. En deuxième lieu, les données de santé disponibles, dont celles des individus de 15 ans ou plus, ne concernent qu'un peu plus de 17 000 individus. Cela est insuffisant pour garantir que les niveaux de dépense agrégés estimés soient corrects en raison de la variance associée à un échantillon de taille restreinte. Enfin, en troisième lieu, les personnes hospitalisées au moment de l'enquête, donc ayant vraisemblablement de très fortes dépenses de santé, ne sont pas présentes dans les données.

Pour chaque poste de soins, le montant agrégé de dépense est calé sur une version retraitée des comptes de la santé de 2019 dans leur édition de 2022 : seules sont conservées les dépenses remboursables, individualisables et incluses dans le champ du modèle (Arnaud, Lefebvre, 2023). Le calage est mené séparément pour les données de santé des moins de 15 ans, qui remontent à 2014, et celles des 15 ans ou plus, qui datent de 2019. Les dépenses de santé agrégées des moins de 15 ans et des 15 ans ou plus en 2019 sont estimées en appliquant aux dépenses de santé provenant des comptes de la santé de 2019 des clés de répartition, entre moins et plus de 15 ans, calculées dans la base RAC 2018.

Les actes et consultations externes (ACE) réalisés à l'hôpital public, ainsi que les honoraires de l'activité libérale des professionnels de santé pratiquée à l'hôpital privé, sont placés au sein des soins de ville dans le modèle Ines-Omar. Cependant les comptes de la santé comptent les ACE avec les soins réalisés à l'hôpital public, tandis que les honoraires des professionnels de santé à l'hôpital privé y forment une catégorie à part. Les dépenses de santé figurant dans le modèle Ines-Omar ne peuvent donc pas être calées telles quelles sur les comptes de la santé. Il est néanmoins possible d'estimer, au moyen de la base RAC 2018, la part de l'activité hospitalière correspondant aux ACE, détaillée par professions médicales. En revanche, la base RAC ne permet pas de repérer assez précisément les honoraires de professionnels de santé perçus à l'hôpital privé ; les honoraires perçus à l'hôpital privé repérés par le modèle sont donc conservés à part des soins en ville pour l'étape de calage, et calés sur le montant des honoraires perçus à l'hôpital privé issu des comptes de la santé (CNS). Ainsi, cela permet de disposer de cibles de calage de la dépense de la santé cohérentes avec la ventilation des soins dans le modèle.

Le tableau 14 compare pour chaque poste de soins les montants de dépense agrégés avant et après calage, pour les individus de moins de 15 ans et pour ceux de 15 ans ou plus. Le tableau 15 présente la CSBM calculée par les comptes de la santé sur différents périmètres : la CSBM individualisable et remboursable, qui correspond au champ du modèle Ines-Omar, la CSBM individualisable, et la CSBM totale.

Les dépenses à l'hôpital public comptabilisées dans le modèle Ines-Omar avant le calage des dépenses de santé sont de 35,9 milliards d'euros, inférieures de moitié aux dépenses de santé à l'hôpital public prises pour cible du calage (70,4 milliards d'euros). Cet écart important se manifestait déjà en comparant la distribution des dépenses à l'hôpital (provenant pour l'essentiel de l'hôpital public) dans la base RAC 2018 et dans les échantillons répondants des enquêtes EHIS 2014 et EHIS 2019 appariés au SNDS (*graphique 1* et *graphique 2*, II).

Tableau 14 Dépense agrégée de santé avant et après calage du modèle Ines-Omar 2019, pour les individus de moins de 15 ans et ceux de 15 ans ou plus

En milliards d'euros

Poste de soins	Dépenses des moins de 15 ans		Dépenses des 15 ans ou plus en 2019	
	Avant calage, en 2014	Après calage, en 2019	Avant calage	Après calage
Hôpital public (ACE exclus)	1,5	4,9	34,4	65,5
Hôpital privé séjour	0,3	0,1	11,4	14,4
Hôpital privé honoraires	0,1	0,1	6,2	6,5
Médecins et sages-femmes	1,7	3,1	18,5	20,9

Auxiliaires médicaux	0,8	0,8	12,3	13,1
Soins dentaires	1,2	0,6	8,3	9,7
Laboratoires d'analyses	0,1	0,2	4,0	4,8
Transports sanitaires	0,1	0,3	3,4	4,5
Médicaments	0,9	1,4	24,6	25,8
Autres biens médicaux	0,6	1,1	15,2	15,9
Total	7,3	12,7	138,8	181,1

ACE : Actes et Consultations Externes.

Note > Dans le modèle Ines-Omar 2019, à la différence de ce tableau, les honoraires pratiqués en hôpital privé sont répartis entre les médecins, les auxiliaires médicaux et la biologie médicale selon le professionnel de santé concerné. Les soins dentaires recouvrent ici les soins dentaires conservateurs et les prothèses dentaires, et les autres biens médicaux désignent les biens médicaux hors optique, l'optique et les cures thermales.

Lecture > Avant calage macro-économique, le modèle estime le total de dépenses de santé en soins dentaires des individus de 15 ans ou plus à 8,3 milliards d'euros sur l'année 2019. Après calage sur le champ des dépenses individualisables et remboursables des individus de 15 ans ou plus, estimées à partir des données des comptes de la santé et de la base RAC 2018, le total de dépenses des adultes en soins dentaires est de 9,7 milliards d'euros en 2019.

Champ > Ensemble des dépenses de santé individualisables et remboursables pour les colonnes « Avant calage » et « Après calage ».

Source > DREES, Ines-Omar 2019 ; DREES, comptes de la santé 2022 pour l'année 2019 ; DREES, base RAC 2018. Traitements DREES.

Tableau 15 Champ des dépenses de santé du modèle Ines-Omar comparé au champ des dépenses individualisables et au champ de la dépense totale, dans les comptes de la santé 2019

En milliards d'euros

Poste de soins	Ines-Omar, calé sur la CSBM individualisable et remboursable	CSBM individualisable	CSBM totale
Hôpital public (ACE inclus)	74,5	74,5	74,5
Hôpital privé séjour	14,6	15,1	15,7
Hôpital privé honoraires	6,6	6,6	7,0
Médecins et sages-femmes	20,6	20,8	23,7
Auxiliaires médicaux	14,0	14,9	15,9
Soins dentaires	10,3	11,8	12,0
Laboratoires d'analyses	4,5	4,5	4,5
Transports sanitaires	4,8	4,9	5,1
Médicaments	27,2	30,4	30,7
Autres biens médicaux	17,0	18,0	18,0
Total	193,9	201,6	207,1

ACE : Actes et Consultations Externes

Note > Les soins dentaires recouvrent ici les soins dentaires conservateurs et les prothèses dentaires, et les autres biens médicaux désignent les biens médicaux hors optique, l'optique et les cures thermales. Les écarts de dépenses de santé par poste dans le modèle Ines-Omar entre le tableau 14 et le tableau 15 proviennent du transfert, dans le tableau 15, des ACE auprès de médecins et sages-femmes (3,4 milliards d'euros d'ACE), d'auxiliaires médicaux (0,1 milliards d'euros) et de laboratoires d'analyses (0,5 milliards d'euros) vers les dépenses à l'hôpital public.

Lecture > Le total de dépenses individualisables et remboursables en soins dentaires est de 10,3 milliards d'euros en 2019 (Actes et consultations externes à l'hôpital public exclus). En ajoutant la dépense non remboursable, les comptes de la santé comptabilisent 11,8 milliards d'euros de soins dentaires en 2019, et 12,0 milliards d'euros en incluant la dépense non individualisable.

Source > DREES, comptes de la santé de 2019 dans l'édition 2022. Traitements DREES.

Le calage ne concerne que le niveau des dépenses de santé. La répartition de la dépense entre individus, ainsi que la décomposition de la dépense entre remboursement AMO, remboursement AMC, liberté tarifaire, RAC opposable, etc. sont celles estimées dans le modèle, sans calage. Le tableau 16 présente les contributions de chacun

des trois financeurs dans le modèle Ines-Omar 2019 après calage, et dans les comptes de la santé 2019 non retraités. Les écarts observés sont en bonne partie dus à des différences de champ. En effet, le modèle ne prend en compte que les soins remboursables par l'AMO et exclut donc les dépenses de santé non remboursables par l'AMO, telles que les chambres particulières à l'hôpital, l'homéopathie, la prévention, etc. Or les dépenses de santé qui ne figurent pas dans le modèle sont généralement moins bien remboursées par l'AMO, donc reposent plus fortement sur les ménages, ce qui aboutit à une répartition du financement différente.

Ainsi, l'estimation du RAC AMC issue des comptes de la santé diffère de celle produite par le modèle Ines-Omar. Selon les CNS, en 2019 le RAC AMC total sur l'ensemble de la population française était de 15 674 millions d'euros, soit 233 euros par habitant (Arnaud, Lefebvre, 2023 ; fiche 21), tandis que le RAC AMC moyen par habitant est estimé par le modèle Ines-Omar 2019 à 170 euros (Lapinte *et al.*, 2024).

Le modèle ne prend en compte que les prestations des OCAM qui concernent les soins remboursables par l'AMO et exclue donc une partie de leurs prestations. Il est nécessaire de mettre en cohérence le champ des cotisations des OCAM, qui correspondent implicitement à l'ensemble du champ des prestations des OCAM. Pour cela, le total de cotisations est calé proportionnellement au total de prestations prises en compte dans le modèle. L'écart entre prestations et cotisations correspond aux taux de frais de gestion et de marge des OCAM tels qu'estimés par le rapport sur la situation financière des organismes complémentaires assurant une couverture santé (rapport OC) portant sur l'année 2019 (Adjérad, 2020). Ces taux de frais et de marge sont différenciés pour les contrats individuels (28 %) et d'entreprise (13 %). Le total de cotisations hors taxes passe ainsi de 37,2 milliards d'euros avant calage à 32,1 milliards d'euros après calage.

Le total des cotisations hors taxes simulées dans Ines-Omar 2019, avant calage, apparaît légèrement inférieur à l'estimation du rapport OC (voir *tableau 16*). En revanche, le total des cotisations hors taxes d'après Ines-Omar est plus proche du montant total des cotisations hors taxes d'après le fonds CSS, qui est de 37,5 milliards d'euros en 2019. Le fonds CSS étudie un champ un peu plus restreint que le rapport OC, qui ne recouvre que l'activité santé à l'exclusion de la prévoyance et des contrats premier euro.

Tableau 16 Dépense agrégée de santé, contributions des trois financeurs et total de primes AMC dans le modèle Ines-Omar 2019, dans les comptes de la santé de 2019 et dans le rapport OC 2020 (cotisations AMC)

	Ines-Omar 2019 (après calage sur la CSBM individualisable et remboursable)	Comptes 2019 (CSBM totale)
Dépenses totales	193,9 Mds €	207,1 Mds €
Part de remb. AMO	81,5 %	79,0 %
Part de remb. AMC	12,7 %	13,4 %
RAC AMC	5,7 %	7,6 %
Cotisations AMC hors taxes	32,1 Mds €	38,3 Mds €

Note > La part de remboursement par l'AMO indiquée ici inclut les remboursements effectués par la sécurité sociale et ceux effectués par l'État, notamment les remboursements supplémentaires au titre de la CMU-C.

Lecture > Le modèle Ines-Omar 2019 estime que 81,5 % des dépenses de santé individualisables et remboursables sont remboursées par l'assurance maladie obligatoire. Les comptes de la santé estiment que 79,0 % de toutes les dépenses de santé sont remboursées par l'AMO en 2019.

Champ > Ensemble des dépenses de santé individualisables et remboursables pour Ines-Omar 2019 ; ensemble des dépenses de santé dans les comptes.

Source > DREES, Ines-Omar 2019 ; DREES, comptes de la santé de 2019 dans l'édition de 2022, rapport OC 2020.

Distribuer le financement de l'AMO

Le modèle Ines apporte toutes les informations nécessaires au calcul de la contribution des ménages au financement de l'AMO. Il contient en effet, pour chaque ménage, les montants de prélèvement suivants : cotisations sociales maladies (des employeurs²⁶, des indépendants et des inactifs), cotisation sociale généralisée (CSG) et contribution pour le remboursement de la dette sociale (CRDS). En outre, le module complémentaire sur la taxation indirecte, fondé sur les structures de consommation déclarées dans l'enquête Budget de Famille 2017, fournit pour chaque ménage une estimation des taxes payées au titre de la taxe sur la valeur ajoutée (TVA), des accises sur le tabac et des taxes sur l'alcool (André *et al.*, 2016). L'ajout de ce module complémentaire était une nouveauté du modèle Ines-Omar 2017, d'autant plus importante que le profil de ces taxes indirectes est nettement moins progressif que le reste des sources de financement. Ces données issues d'Ines s'ajoutent aux sources de financement qui peuvent être estimées à partir du seul modèle Omar : la taxe de solidarité additionnelle (TSA), qui s'applique

²⁶ Depuis 2019, les cotisations maladie salariales ont disparu, seules subsistent les cotisations maladie patronales.

proportionnellement aux primes de complémentaire santé, et la taxe sur les médicaments, qui pèse *in fine* sur les RAC AMO des ménages en pharmacie.

Grâce à ces données, il devient possible de distribuer l'ensemble du financement de l'AMO sur les ménages. L'objectif est de savoir comment sont financés les remboursements AMO pris en compte dans le champ du modèle, soit un total de 157 milliards d'euros²⁷, afin que celui-ci soit équilibré entre ses prestations et ses financements. Par conséquent le financement des dépenses de santé non prises en compte dans le modèle, notamment les dépenses non individualisables tels que les frais généraux de fonctionnement des hôpitaux ou les coûts de gestion de la CNAM, doit être écarté. Il n'existe pas d'affectation spécifique de certaines sources de financement en direction de certains types précis de prestations. C'est pourquoi l'ensemble des sources de financement de l'AMO doivent être prises en compte, y compris la génération de dettes, et réparties proportionnellement à la part qu'elles représentent dans le budget de la Sécurité sociale en 2019. Les comptes de recettes nettes 2019 de la Sécurité sociale²⁸ sont mobilisés pour estimer la part du financement issue de chaque source. Ces différentes sources, ainsi que leurs équivalents mobilisés dans le modèle, sont présentées dans le tableau 17. La part calculée dans le tableau exclut les quelques recettes que le modèle ne permet pas de distribuer, lesquelles ne comptent que pour 4 % du financement total de l'Assurance maladie.

Quelques précisions sont nécessaires pour bien comprendre le tableau :

- Dans les comptes de la Sécurité sociale, l'écart entre les produits et les charges correspond à de l'endettement, lequel doit alors être pris en compte en tant que source de financement. Le remboursement de la dette s'effectuant essentiellement par la CRDS et la CSG, celles-ci sont donc des sources de recettes à hauteur de 1,5 milliards d'euros en 2019 ;
- En 2019, la branche Accidents du travail et maladies professionnelles (AT-MP) a contribué au financement de la branche maladie à hauteur d'un milliard d'euros. Cette contribution est financée conformément au financement du risque AT-MP, c'est-à-dire presque exclusivement par des cotisations salariales²⁹ ;
- La Caisse nationale de solidarité pour l'autonomie (CNSA) finance certaines dépenses de santé relevant de la branche maladie tout en recevant des fonds de celle-ci. Au global, il s'agit d'une contribution nette de la CNSA vers la branche maladie, c'est-à-dire d'une source de financement. Hormis les versements par la branche maladie, le financement de la CNSA repose sur les contributions de solidarité pour l'autonomie (CSA à 40 % et CASA à 15 %) et sur la CSG (à 45 %)³⁰. Les contributions de solidarité pour l'autonomie ne sont pas connues dans le modèle Ines, mais la CSA est proche des cotisations sociales maladie patronales tandis que la contribution additionnelle CASA s'applique uniformément aux pensions de retraite et d'invalidité. C'est donc la distribution de ces cotisations patronales et de ces pensions qui sont prises comme équivalents à la CSA et à la CASA dans le modèle.

Tableau 17 Sources de financement de l'assurance maladie obligatoire prises en compte et leur équivalent mobilisé dans le modèle

En milliards d'euros

Source de financement	Équivalent dans Ines-Omar	Montant comptes de la Sécurité sociale	Montant Ines-Omar	Part
Cotis. Sociales des salariés	Cotis. Sociales maladie des salariés	66,8	50,0	32,0 %
Cotis. Sociales des non-salariés	Cotis. Sociales maladie des non-salariés	4,8	3,6	2,3 %
Cotis. Sociales des inactifs	Cotis. Sociales maladie des inactifs	0,8	0,6	0,4 %
Cotisation sociale généralisée affectée à la maladie	CSG	71,8	53,7	34,5 %
Taxe sur les salaires	Cotis. Sociales maladie patronales	1,3	1,0	0,6 %
Taxe de solidarité additionnelle	Primes OCAM	2,3	1,8	1,1 %
Taxe sur les médicaments	RAC AMO en pharmacie	0,9	0,7	0,5 %
TVA + produits alimentaires	Module de taxation indirecte	41,0	30,7	19,7 %
Accises sur le tabac		13,0	9,7	6,2 %
Taxe sur l'alcool		1,6	1,2	0,8 %
Transfert de la branche AT-MP	Cotis. Sociales maladie des salariés	1,0	0,7	0,5 %

²⁷ Ce montant n'inclut pas les remboursements supplémentaires de la CMU-C.

²⁸ Plus précisément, le tableau 2 en partie 4.2 du rapport de septembre 2020 de la Commission des comptes de la Sécurité sociale.

²⁹ Voir le tableau 1 de la partie 4.3 du rapport de septembre 2020.

³⁰ Voir le budget exécuté 2019 de la CNSA.

Contribution nette à la CNSA	Cotis. Sociales maladie patronales (40 %), re- traites et invalidité (15 %), CSG (45 %)	1,4	1,0	0,7 %
Endettement	CRDS (47 %) et CSG (53 %) ³¹	1,5	1,1	0,7 %
Total		208,4	157,0	100 %

Lecture > La taxe sur les salaires est approximée dans le modèle comme étant distribuée au sein de la population de la même façon que les cotisations sociales maladie patronales. Elle rapporte un total de 1,3 milliards d'euros en 2019, soit 0,6 % du total des contributions prises en compte dans le modèle, de sorte qu'elle finance les prestations AMO prises en compte dans le champ du modèle à hauteur de 1,0 milliard d'euros.

Champ > Ensemble des contributions au financement de l'assurance maladie obligatoire.

Source > DREES, Ines-Omar 2019 ; DSS, comptes de la Sécurité sociale 2020. Traitements DREES.

³¹ D'après le rapport annuel 2019 de la Caisse d'amortissement de la dette sociale (CADES), ses ressources ont été de 7,6 milliards d'euros de CRDS, 8,6 milliards d'euros de CSG et un versement de 2,1 milliards d'euros par le Fonds de réserve pour les retraites (FRR). Étant donné que les ressources du FRR proviennent d'investissements économiques fructueux, ils ne peuvent pas être distribués entre les ménages. Par ailleurs, en 2019, cette contribution compense presque parfaitement les intérêts versés par la CADES à ses investisseurs. Il est donc considéré que l'endettement est entièrement financé par la CRDS et la CSG.

■ POUR EN SAVOIR PLUS

- Adjerad, R.** (2020, décembre). [La situation financière des organismes complémentaires assurant une couverture santé. Rapport 2020.](#) DREES.
- Adjerad, R., Courtejoie, N.** (2021, mai). [Hospitalisation : des restes à charge après AMO plus élevés en soins de suite et de réadaptation et en psychiatrie qu'en médecine, chirurgie, obstétrique et odontologie.](#) DREES, *Études et Résultats*, 1192.
- André, M., Biotteau, A.-L., Duval, J.** (2016, novembre). [Module de taxation indirecte du modèle Ines – Hypothèses, principes et aspects pratiques.](#) DREES, *Documents de travail*, 60.
- Arnaud, F., Lefebvre, G.** (dir.) (2023). [Les dépenses de santé en 2022 – Résultats des comptes de la santé – Édition 2023.](#) DREES, coll. Panoramas de la DREES-Santé.
- Barlet, M., Gaini, M., Gonzalez, L., et al.** (dir.) (2019). [La complémentaire santé – Acteurs, bénéficiaires, garanties – Édition 2019.](#) DREES, coll. Panoramas de la DREES-Santé.
- Bithorel, P., Reduron, V.** (2023, mai). [Dépenses de santé et restes à charge pour les patients : comment expliquer les disparités entre départements ?](#) DREES, *Études et Résultats*, 1265.
- Bricard, D.** (2024, avril). [Les soins et services de santé sont-ils financièrement accessibles ? De nouvelles bases factuelles sur la protection financière en France.](#) Organisation Mondiale de la Santé, Bureau régional pour l'Europe.
- Fouquet, M.** (2020, octobre). [Une hausse modérée de la couverture de la population après la généralisation de la complémentaire santé d'entreprise.](#) DREES, *Études et Résultats*, 1166.
- Fouquet, M.** (2021, décembre). [Le modèle de micro-simulation Ines-Omar 2017 – Note méthodologique détaillée.](#) DREES, *Les Dossiers de la DREES*, 87.
- Fredon, S., Sicsic, M.** (2020, juin). [Ines, le modèle qui simule l'impact des politiques sociales et fiscales.](#) Insee, *Courrier des statistiques*, 4.
- Guilloton, V., Lapinte, A.** (2022, mars). [Enquête auprès des organismes offrant une couverture complémentaire 2019 - Analyse des changements méthodologiques, révisions des années antérieures et premiers résultats.](#) DREES, *Les Dossiers de la DREES*, 93.
- Lapinte, A., Pollak, C., Solotareff, R.** (dir.) (2024). [La complémentaire santé – Acteurs, bénéficiaires, garanties – Édition 2024.](#) DREES, coll. Panoramas de la DREES-Santé.
- Leduc, A., Deroyon, T., Rochereau, T., et al.** (2021, avril). [Premiers résultats de l'enquête santé européenne \(EHIS\) 2019 - Métropole, Guadeloupe, Martinique, Guyane, La Réunion, Mayotte.](#) DREES, *Les Dossiers de la DREES*, 78.
- Leduc, A.** (2022, novembre). [Enquête santé européenne \(EHIS\) 2019 : Bilan méthodologique - France métropolitaine et DROM.](#) DREES, *DREES Méthodes*, 7.
- Loiseau, R.** (2020, février). [Aide au paiement d'une complémentaire santé : un niveau de couverture équivalent à celui des contrats du marché individuel en 2018.](#) DREES, *Études et Résultats*, 1142.
- Perronnin, M.** (2019, novembre). [L'enquête Protection sociale complémentaire d'entreprise 2017.](#) Irdes, Les rapports de l'Irdes, 572.
- Pierre, A.** (2022, octobre). [Marché de la complémentaire santé, inégalités et préférences en matière de couverture : les effets de la généralisation de la complémentaire santé d'entreprise.](#) IRDES, *Document de travail*, 89.

Annexe 1. Questionnaire couverture complémentaire d'EHIS 2019

Seules sont retranscrites ici les questions importantes pour la construction du modèle, et notamment pour la connaissance du statut de couverture dans l'enquête EHIS 2019 de la DRES. Des questions additionnelles portant sur les arrêts de travail, les raisons de la non-couverture ou encore le niveau de garanties sont posées dans l'enquête mais non retranscrites ici.

COUV1

Bénéficiez-vous [PRENOM bénéficie-t-il(elle)] actuellement de la couverture maladie universelle complémentaire, appelée plus couramment CMU ?

Enquêteur : La Couverture maladie universelle complémentaire, appelée couramment CMU, propose gratuitement une couverture en plus de la Sécurité sociale pour les personnes dont les ressources sont faibles.

- Oui..... 1 → SOUSC (p. 27)
- Non..... 2
- NSP..... 99
- Refus..... 97

Si COUV1 ≠ 1

COMPSANT

Bénéficiez-vous [PRENOM bénéficie-t-il(elle)] d'une complémentaire santé, appelée couramment « mutuelle », qui complète les remboursements de la Sécurité sociale ?

Enquêteur : Si l'enquêté a plusieurs contrats, les questions suivantes concerneront le contrat le plus souvent utilisé. Il s'agit ici de la mutuelle santé qui couvre les frais de santé en complément de la Sécurité sociale, par exemple pour les dépenses de santé chez le médecin, pour les lunettes ou les prothèses dentaires...

- Oui..... 1 → COUVENT (p. 27)
- Non..... 2
- NSP..... 99 → MODULE 4 (p. 30)
- Refus..... 97 → MODULE 4 (p. 30)

Si COMPSANT = 1

COUVENT

S'agit-il :

CITER – UNE SEULE REPONSE POSSIBLE.

Enquêteur : Les mutuelles d'entreprise sont souscrites par un employeur pour ses salariés.

- D'une mutuelle de la fonction publique ? 1
- D'une mutuelle d'entreprise souscrite par l'intermédiaire d'un employeur du privé ? 2
- D'une mutuelle obtenue individuellement ? 3
- NSP..... 99
- Refus..... 97

Si [COUV = 1] ou [COMPSANT = 1]

SOUSC

Qui a souscrit cette mutuelle (la CMU si COUV1 = 1) ?

CITER – UNE SEULE REPONSE POSSIBLE.

Enquêteur : Si plusieurs contrats, les questions suivantes concerneront le contrat le plus souvent utilisé.

- [Afficher les prénoms des membres du ménage]..... [] NOI
- Une personne hors foyer.....
- NSP..... 99
- Refus..... 97

Si [COUV = 1] ou [COMPSANT = 1]

QUICOUV

Quelles sont les personnes couvertes par cette mutuelle (la CMU si COUV1 = 1) ?

CITER – PLUSIEURS REPONSES POSSIBLES.

- [Afficher les prénoms des membres du ménage et cocher ceux couverts par le contrat] [] NOI
- Une ou des personnes hors foyer.....
- NSP..... 99
- Refus..... 97

Si COMPSANT = 1

COMB

Combien votre ménage paye-t-il par mois pour l'ensemble des personnes couvertes par cette mutuelle ?

QUESTION OUVERTE NUMERIQUE

Enquêteur : Indiquer le coût pour l'ensemble des personnes rattachées à cette mutuelle.

- Si l'enquêté a plusieurs contrats, les questions suivantes concernent celui qui est le plus souvent utilisé.
- Pour les contrats d'entreprise, il s'agit du montant qui est à votre charge, déduction faite de la participation de l'employeur. Il est de 0 € lorsque l'entreprise finance intégralement la couverture.

- [0 à 500 €] [] **→ SATISDEP (p. 28)**
- NSP..... 99
- Refus..... 97

Si [COMPSANT = 1] et [COMB = 97 ou 99]

COMBTR

A combien environ l'estimez-vous par mois ?

CITER – UNE SEULE REPONSE POSSIBLE.

- Moins de 40 €..... 1
- Entre 40 € et 59 €..... 2
- Entre 60 € et 89 €..... 3
- Entre 90 € et 119 €..... 4
- 120 € ou plus..... 5
- NSP..... 99
- Refus..... 97

Annexe 2. Modifications par rapport à Ines-Omar 2017

Tableau 18 Modifications de la méthodologie d'Ines-Omar entre les millésimes 2017 et 2019

	Ines-Omar 2017	Ines-Omar 2019
Préparation des bases		
CMU-C	SRCV 2017 : sous-déclaration de la CMU-C chez les jeunes, sur-déclaration chez les +60 ans.	EHIS 2019 et SRCV 2017 : redressement des déclarations de CMU-C au moyen d'un modèle <i>logit</i> entraîné sur EHIS 2014. EHIS 2014 et EHIS 2019 : croisement avec le bénéfice de la CMU-C dans le SNDS
ACS	SRCV 2017 : bénéficiaires de l'ACS tirés au sort, ainsi que le niveau de l'ACS (A, B ou C).	EHIS 2019 : bénéficiaires de l'ACS identifiés dans le SNDS. Mais niveau d'ACS encore imputé selon la prime déclarée.
Apurement du type de complémentaire	SRCV 2017 : le statut professionnel prime sur le type de complémentaire.	SRCV 2017 : idem. EHIS 2019 : le statut professionnel prime sur le type de complémentaire si l'ouvrant droit a lui-même répondu et donc son statut ne fait pas de doute.
Traitement des données de santé		
SNDS (PMSI et DCIRS)	Bénéfice de la CMU-C disponible.	Bénéfice de la CMU-C, de l'ACS et d'une ALD disponible.
ALD	Issue de SRCV 2017 (disponible pour les 16 ans et plus).	Pour les 15 ans et plus d'EHIS 2019 : issue du SNDS. Indisponible pour les moins de 15 ans.
Agrégation des soins	Un seul poste pour toutes les prothèses dentaires.	Prothèses dentaires détaillées entre prothèses céramo-métalliques visibles, dentiers, bridges, inlay-Onlay, inlay-Core.
Imputation des données de santé à SRCV 2017	Appariement avec EHIS-ESPS 2014 séparément pour les bénéficiaires de la CMU-C, les non couverts, et les autres.	Appariement avec EHIS-ESPS 2014 pour les moins de 15 ans uniquement, séparément pour les bénéficiaires de la CMU-C et les autres.
Tirage des contrats d'OCAM et simulation des primes		
Représentativité de l'enquête OC	L'enquête OC recueille les caractéristiques d'un échantillon représentatif des trois contrats les plus fréquemment souscrits de chaque type (individuel / collectif) des organismes enquêtés.	L'enquête OC recueille les caractéristiques d'un échantillon représentatif des dix contrats les plus fréquemment souscrits de chaque type des organismes enquêtés.
Retraitement de l'enquête OC 2019	Les contrats réservés à une branche sont considérés comme des contrats individuels.	Les contrats réservés à une branche sont désormais considérés comme des contrats collectifs réservés aux salariés du privé.
Stratification de l'enquête OC 2019	20 strates ; mêmes seuils de score pour les contrats individuels et collectifs.	15 strates ; seuils de score différents entre contrats individuels et collectifs.
Probabilités de tirage	La probabilité de tirage est différenciée selon que l'ouvrant droit a des bénéficiaires dans certaines strates seulement.	La probabilité de tirage est différenciée selon que l'ouvrant droit a des bénéficiaires dans toutes les strates de l'enquête OC 2019.
Imputation de la qualité du contrat	Imputation des valeurs manquantes de prime ajustée du nombre d'assurés par tirage dans une distribution gaussienne.	Imputation des valeurs manquantes de prime ajustée du nombre d'assurés par un modèle de régression linéaire.
Modulation de la prime selon la région	La prime moyenne par ZEAT est supposée suivre le RAC AMO moyen des 40-50 ans. Après leur simulation, les primes sont calées de sorte que leur niveau relatif par région suive le niveau relatif du RAC AMO des 40-50 ans.	La prime moyenne par ZEAT est supposée suivre le RAC AMO opposable moyen des 40-50 ans. Après leur simulation, les primes sont calées de sorte à retrouver le niveau relatif des primes par ZEAT selon EHIS 2019.
Niveau des primes avant calage	Primes individuelles simulées très proches des primes déclarées en moyenne ; primes simulées en collectif trop basses.	Primes individuelles simulées plus fortes que les primes déclarées en moyenne ; primes simulées en collectif trop basses.
Remboursement de l'AMC		

Remboursement des équipements par la CMU-C	Remboursement supplémentaire de la CMU-C calculé à partir des garanties de la CMU-C avant mise en place de paniers sans reste à charge.	Remboursement supplémentaire de la CMU-C pour les biens des paniers sans reste à charge (réservés aux bénéficiaires de la CMU-C en 2019) calculé à partir du prix limite de vente.
Remboursements de l'ACS	Toujours calculés à partir des garanties de l'ACS.	Le plus souvent fournis par le SNDS, autrement calculés à partir des garanties de l'ACS.

Constitution d'Ines-Omar

Appariement par hot-deck	Fusion des strates de taille insuffisante entre elles : produit quelques strates très hétérogènes.	Fusion des strates de taille insuffisante avec des strates similaires : produit des strates plus volumineuses mais plus homogènes. Il en résulte moins d'appariements d'individus très différents entre Ines et Omar.
Tirage des bénéficiaires de la CMU-C et de l'ACS dans Ines	Tirage parmi les éligibles de sorte à retrouver la répartition des bénéficiaires par tranche d'âge observée dans Omar.	Tirage parmi les éligibles de sorte à retrouver la répartition des bénéficiaires par tranche d'âge et par niveau de vie observée dans Omar. Aucun tirage de bénéficiaires de l'ACS au-delà du 9 ^e vingtième de niveau de vie.

Finalisation du modèle

Calage macro-économique des dépenses	Calage simultané des dépenses de tous les individus.	Calage séparé pour les dépenses des individus de moins de 15 ans, provenant d'EHIS 2014-SNDS, et celles des 15 ans ou plus, provenant d'EHIS 2019-SNDS.
Calage des prestations de la CMU-C et de l'ACS	Calage des remboursements de la CMU-C sur le coût moyen de la CMU-C, correction des prestations de l'ACS avec le coefficient appliqué aux remboursements de la CMU-C.	Calage des remboursements de la CMU-C sur le coût moyen de la CMU-C, et des prestations de l'ACS sur le coût moyen de l'ACS.

Annexe 3. Extraits du questionnaire de l'enquête OC 2019 sur les primes

Seules sont présentées ici les questions de l'enquête OC 2019 portant sur les primes payées par les assurés.

Cette partie du questionnaire porte sur le montant mensuel de la cotisation **TTC** dont devrait s'acquitter un nouvel adhérent qui souscrirait aux contrats décrits dans l'enquête.

Par défaut, on considère que le **nouvel adhérent "de référence"** :

- est un homme **vivant seul et sans enfant**
- est âgé de 40 ans
- est rattaché au régime général de la sécurité sociale ou un régime offrant les mêmes remboursements (RSI, fonctionnaires, etc.)
- est déjà couvert par un organisme concurrent au moment de son adhésion
- **travaille à temps plein** et perçoit un salaire correspondant à la plus basse tranche de revenus (si les cotisations dépendent des revenus)
- vit dans la zone géographique où les cotisations sont les plus élevées (si les cotisations dépendent du lieu de résidence).

	(A)	(B)	(C)
Quel serait le montant mensuel de la cotisation de l'adhérent de référence (décrit dans l'encadré ci-dessus)?	(1) €	€	€
Ce contrat est-il lié à un régime spécial incluant déjà une complémentaire obligatoire (type Alsace-Moselle, Industries électriques et gazières...)?	(2) <input type="radio"/> oui <input type="radio"/> non	<input type="radio"/> oui <input type="radio"/> non	<input type="radio"/> oui <input type="radio"/> non

Contrats individuels et surcomplémentaires:

Les cotisations dépendent-elles de l'âge ?

(3)	<input type="radio"/> oui <input type="radio"/> non	<input type="radio"/> oui <input type="radio"/> non	<input type="radio"/> oui <input type="radio"/> non
-----	---	---	---

Contrats individuels et surcomplémentaires:

Si oui, quel serait le montant mensuel de la cotisation de l'adhérent de référence, s'il était (*) :

Âgé de 20 ans ?	(4) €	€	€
Âgé de 60 ans ?	(5) €	€	€
Âgé de 65 ans ?	(6) €	€	€
Âgé de 75 ans ?	(7) €	€	€
Âgé de 85 ans ?	(8) €	€	€

Comment sont pris en compte les enfants ayants droit dans la cotisation ?

0. Pas de prise en charge des enfants

1. Tous les enfants cotisent (y compris si le tarif est dégressif)

2. Gratuité à partir du 2^e, 3^e ou 4^e enfant

3. Forfait famille quel que soit le nombre d'enfants

(9)	<input type="radio"/> 0	<input type="radio"/> 0	<input type="radio"/> 0
	<input type="radio"/> 1	<input type="radio"/> 1	<input type="radio"/> 1
	<input type="radio"/> 2	<input type="radio"/> 2	<input type="radio"/> 2
	<input type="radio"/> 3	<input type="radio"/> 3	<input type="radio"/> 3

Quel serait le montant mensuel de la cotisation de la famille de l'adhérent de référence s'il était marié à une femme de 40 ans (*) :

- Sans enfant ?
- Avec un enfant ?
- Avec deux enfants ?
- Avec trois enfants ?

	(A)	(B)	(C)
(10)	€	€	€
(11)	€	€	€
(12)	€	€	€
(13)	€	€	€

Contrats individuels et surcomplémentaires:

Les cotisations dépendent-elles du lieu de résidence ?

(14)	<input type="radio"/> oui <input type="radio"/> non	<input type="radio"/> oui <input type="radio"/> non	<input type="radio"/> oui <input type="radio"/> non
------	---	---	---

Si oui, quel serait le montant mensuel de la cotisation de l'adhérent de référence s'il vivait dans la zone géographique où les cotisations sont les moins élevées ?

(15)	€	€	€
------	---	---	---

Les cotisations dépendent-elles des revenus ?

(16)	<input type="radio"/> oui <input type="radio"/> non <input type="radio"/> NSP	<input type="radio"/> oui <input type="radio"/> non <input type="radio"/> NSP	<input type="radio"/> oui <input type="radio"/> non <input type="radio"/> NSP
------	---	---	---

Si oui, quel serait le montant mensuel de la cotisation de l'adhérent de référence s'il touchait un salaire brut correspondant à deux fois le plafond mensuel de la sécurité sociale (2 x 3 377 en 2019, soit 6 754 euros) (ou, à défaut, le salaire correspondant à la plus haute tranche de revenus de votre grille salariale) ?

(17)	€	€	€
------	---	---	---

Contrats individuels et surcomplémentaires:

Les cotisations dépendent-elles de l'ancienneté dans le contrat ?

(18)	<input type="radio"/> oui <input type="radio"/> non	<input type="radio"/> oui <input type="radio"/> non	<input type="radio"/> oui <input type="radio"/> non
------	---	---	---

Si oui, quel serait le montant mensuel de la cotisation de l'adhérent de référence s'il était couvert par ce contrat depuis 10 ans ? (*)

(19)	€	€	€
------	---	---	---

Annexe 4. Comparaison avec l'enquête Budget de Famille 2017

Les seules sources à ce jour qui décrivent les restes à charge après remboursement de l'assurance maladie complémentaire (RAC AMC) de l'ensemble de la population française sont le modèle de microsimulation Ines-Omar de la DREES, et l'enquête Budget de Famille 2017 (BdF) produite par l'Insee, qui questionne directement les ménages sur leurs dépenses de santé. Ces deux sources reposent sur des méthodes différentes pouvant conduire à des estimations de RAC AMC divergentes. De plus, elles portent sur un champ différent, puisque l'enquête BdF inclut également les dépenses de santé non remboursables par l'assurance maladie obligatoire (AMO) ou non présentées au remboursement.

Cette annexe présente la manière dont l'enquête BdF 2017 mesure les dépenses de santé des ménages, et compare les principaux résultats de l'enquête BdF 2017 avec ceux du millésime 2017 du modèle Omar. Dans l'ensemble, il apparaît que l'enquête BdF 2017, telle que retraitée par l'Insee, donne un niveau de RAC AMC moyen par habitant proche de celui d'Omar 2017 pour les statistiques portant sur l'ensemble de la population ou sur des sous-groupes de taille importante.

En revanche, les dépenses de santé de BdF ne sont pas fiables à un niveau individuel et ne doivent pas être interprétées en distribution. Cela est dû à la manière dont sont recueillies les dépenses de santé : les individus enquêtés sont interrogés sur leurs dépenses de santé au cours des deux derniers mois, et non pas sur la dernière année écoulée. Cette période restreinte vise à éviter les biais de réponse dus à la difficulté de se rappeler des dépenses anciennes. Cependant, la méthode retenue par l'Insee pour annualiser les dépenses déclarées consiste à multiplier par six les dépenses des deux derniers mois pour les ménages ayant eu une dépense sur tel ou tel poste de soins : cela a pour effet d'accentuer les écarts de dépenses de santé entre ménages et de fausser la distribution des RAC AMC.

Par ailleurs, les individus enquêtés peuvent indiquer, pour chaque poste de soins, soit le prix total des actes ou consultations ou biens, soit le reste à charge après remboursement de l'AMO et avant intervention de la complémentaire santé, soit le reste à charge final après remboursements de l'AMO et de la complémentaire santé éventuelle. Or, les individus enquêtés semblent confondre la deuxième et la troisième modalité de réponse, comme détaillé ci-dessous.

Les questions sur les dépenses de santé dans l'enquête BdF 2017 semblent difficiles à renseigner pour les ménages interrogés

En 2017, les questions consacrées aux dépenses de santé dans l'enquête BdF se composent de trois sous-modules relatifs au type de couverture par une complémentaire santé, aux frais de santé, et à la location ou l'achat d'appareils médicaux. Les trois sous-modules sont administrés par un enquêteur à une seule personne répondant pour l'ensemble du ménage.

Le premier sous-module recueille la couverture par une complémentaire santé de chaque membre du ménage, le type de contrat de complémentaire, la satisfaction à l'égard de la complémentaire et son coût. Dans l'enquête BdF, le taux de souscription de la CMU-C est trop faible et le taux de non-couverture par une complémentaire santé trop élevé, comparé avec les données de l'enquête SRCV 2017 et du fonds CMU-C. Les taux de couverture par un contrat individuel et par un contrat collectif sont proches dans l'enquête BdF et dans SRCV au niveau agrégé, ainsi que par tranche d'âge.

Tableau 19 Taux de couverture par types de contrat en population générale selon la source

	BdF 2017	SRCV 2017 – DREES, ER n°1166	Fonds CMU-C sur 2017
CMU-C	4,7 %	7 %	7,4 %
Contrat collectif	36,7 %	38 %	-
Contrat individuel	50,4 %	51 %	-
Non couverture	8,1 %	4 %	-

Champ > BdF 2017 et SRCV 2017 : population vivant en ménage ordinaire en France métropolitaine ; Fonds CMU-C : ensemble de la population vivant en France métropolitaine.

Source > Insee, enquête Budget de Famille 2017, calculs DREES ; Insee, SRCV 2017, calculs DREES ; Fonds CMU-C.

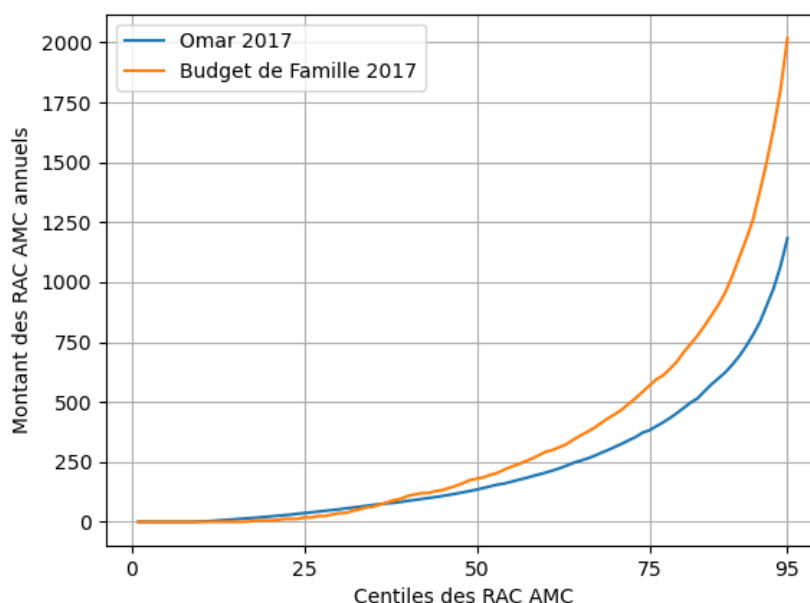
Le second sous-module interroge l'individu enquêté sur les dépenses de santé de l'ensemble de son ménage dans cinq catégories de soins : consultations de généralistes, de spécialistes, examens en laboratoire médical, visites en hôpital et hospitalisations, achats en pharmacie. L'enquête BdF laisse la possibilité aux enquêtés de déclarer, pour chacune de ces catégories de soins, la fréquence de recours, le type de dépense qu'ils ont réglé, et le montant

correspondant. Ainsi, les répondants choisissent le montant qui leur semble le plus simple à donner entre : le prix total avant tout remboursement, ou le reste à charge après remboursement par l'assurance maladie obligatoire (le RAC AMO) mais avant intervention de la complémentaire santé éventuelle, ou le reste à charge final après remboursements de l'assurance maladie obligatoire et de l'assurance maladie complémentaire éventuelle (RAC AMC), qui peut être nul. Lorsque l'enquête indique une dépense totale ou un RAC AMO, l'Insee doit procéder à un traitement de la dépense déclarée pour parvenir à un RAC AMC.

Il n'est pas évident que les ménages comprennent bien la différence entre les différents modes de réponse, aient l'information, et qu'ils déclarent des montants cohérents. D'après des calculs de la DREES sur l'enquête BdF, la distribution des RAC AMO est très proche de celle des RAC AMC pour toutes les catégories de soins (sauf les examens en laboratoire), au lieu d'être centrée autour de valeurs supérieures. Cela laisse penser que les répondants confondent les notions de RAC AMO et de RAC AMC.

Par ailleurs, les dépenses pour ces catégories de soins sont demandées pour les deux derniers mois (à l'exception des hospitalisations qui sont sur l'année écoulée), menant à des retraitements qui altèrent la distribution. En effet, les RAC AMC annuels sont obtenus en multipliant les déclarations de chaque ménage par 6. Comme l'enquête BdF 2017 est menée tout au long de l'année, les dépenses de santé ne souffrent pas d'effets saisonniers au niveau agrégé. En revanche la variabilité des dépenses de santé entre ménages est accentuée : l'estimation de la dépense annuelle pour un individu ayant consommé des soins au cours du dernier bimestre sera très élevée ; au contraire pour un individu qui n'a pas consommé de soins depuis deux mois, cela conduit à considérer qu'il n'en a pas consommé du tout sur l'année écoulée. BdF n'est pas construite pour être exploitable à un niveau individuel ou par ménage, elle ne permet donc pas de calculer la distribution des RAC AMC, mais uniquement leur moyenne (sur l'ensemble de la population ou sur des catégories de personnes), à l'exception des dépenses pour l'hôpital qui y sont recueillies sur 12 mois. Ainsi, les RAC AMC par ménage apparaissent bien plus dispersés dans l'enquête BdF 2017 que dans le modèle Omar 2017 (*graphique 7*).

Graphique 7 Distribution par centiles du montant total par ménage de RAC AMC. Comparaison entre l'enquête BdF 2017 et le modèle Omar 2017



Note > Les centiles de dépenses sont affichés jusqu'au 95^e inclus pour ne pas écraser la courbe.

Lecture > Le troisième quartile (75^e centile sur les figures) des RAC AMC les plus élevées s'élève à 568 euros en 2017 d'après l'enquête BdF 2017, et à 382 euros en 2017 d'après le modèle Omar 2017.

Champ > Ménages ordinaires en France métropolitaine, dépenses remboursables et présentées au remboursement (Omar 2017) vs. ensemble des dépenses y compris non-remboursables par l'assurance maladie obligatoire (Budget de Famille 2017).

Source > DREES, modèle Omar 2017 ; Insee, enquête BdF 2017. Traitements DREES.

Enfin le troisième sous-module recueille les dépenses afférentes à la location ou à l'achat de matériels médicaux (lunettes et lentilles, appareils dentaires, prothèses auditives, attelles et chaussures orthopédiques, fauteuil roulant, lit médicalisé, appareil respiratoire, ou autre) sur les douze derniers mois. De même, l'enquête BdF laisse la possibilité aux enquêtés de déclarer leurs dépenses en matériels médicaux selon trois formes différentes : soit la dépense totale, soit le RAC AMO, soit le RAC AMC (qui peut être nul).

Le RAC AMC moyen est assez proche dans l'enquête BdF 2017 et Omar 2017, à l'exclusion des dépenses de pharmacie

En dépit des limites des déclarations des ménages dans l'enquête BdF 2017 liées à leur bonne compréhension des différentes notions de reste à charge et à la période d'observation, le travail d'imputation et de redressement des RAC AMC de l'Insee, diffusé auprès des utilisateurs des données anonymisées, peut améliorer la qualité des informations. Nous comparons ici les RAC AMC moyens par ménage imputés et diffusés par l'Insee pour l'enquête BdF à ceux simulés par le modèle de microsimulation Omar 2017, à des niveaux d'agrégation des postes de soins similaires (*tableau 20*).

Tableau 20 Restes à charge (RAC AMC) annuels moyens des ménages par postes de soin selon la source, en euros

		BdF 2017	Omar 2017
	Total <i>(hors autres produits pharmaceutiques et autres dépenses de santé dans BdF)</i>	476	349
1.	Produits pharmaceutiques à ingurgiter et traitements, compléments alimentaires, vitamines et minéraux	132	
	Pharmacie		35
2.	Services et soins hospitaliers	26	9
3.	Total hors pharmacie et hôpital	318	305
3.1.	Appareils et matériel thérapeutique (lunettes, prothèses)	97	
	Audioprothèses + optique + prestations sanitaires diverses ³² + soins dentaires prothétiques ³³ (dont orthodontie)		213
3.2	Généralistes & spécialistes	103	60
3.3	Soins dentaires conservateurs et orthodontie	62	
	Soins dentaires conservateurs		7
3.4	Laboratoire et cabinet de radiologie	25	13
3.5	Auxiliaires médicaux	30	11
3.6	Services extrahospitaliers (ambulance, location matériel)	1	
	Transport		2
	Total module santé BdF	510	
	y.c. autres produits pharmaceutiques	25	
	y.c. autres dépenses de santé : personnes vivant hors du domicile au moins un jour par semaine, cadeau offert	9	

Champ > Ménages ordinaire en France métropolitaine, dépenses remboursables et présentées au remboursement (Omar 2017) vs. ensemble des dépenses y compris non-remboursables par l'assurance maladie obligatoire (Budget de famille 2017).

Source > Insee, enquête Budget de familles 2017, calculs DREES ; DREES, modèle de microsimulation Omar 2017 basé sur l'enquête SRCV 2017.

Il s'avère que malgré la difficulté pour les enquêtés de BdF de comprendre les notions de dépenses avant/après remboursement par l'assurance maladie obligatoire et complémentaire, les RAC AMC moyens par ménage sont relativement proches dans l'enquête BdF et dans le modèle Omar 2017, notamment sur le champ de l'ensemble des soins hors pharmacie et hors hôpital (autour de 310 euros par ménage sur ce champ). Par poste de soins certaines différences émergent, en pharmacie surtout. En effet, l'ajout des dépenses non remboursables par

³² Le poste d'Omar 2017 « prestations sanitaires diverses » comprend le matériel médical et les cures thermales.

³³ Le poste d'Omar 2017 « soins dentaires prothétiques » comprend l'orthodontie, l'implantologie et les prothèses dentaires. Comme on ne peut dissocier les consultations d'orthodontie des autres soins dentaires prothétiques, on est contraint de les laisser dans le poste Omar lors de la comparaison avec les appareils et matériel thérapeutique de BdF.

l'assurance maladie obligatoire et notamment des médicaments non remboursés, qui ne sont pas dans le champ d'Omar se basant sur le SNDS, augmente fortement le RAC AMC moyen dans l'enquête BdF.

Au niveau individuel, le RAC AMC moyen par individu dans l'enquête BdF 2017 est estimé à 237 €, dont 14 € de RAC AMC pour les soins hospitaliers et 64 € de médicaments (Bricard, 2024³⁴). Hors soins hospitaliers et hors médicaments, cela correspond à 159 euros par personne, soit un montant très proche du RAC AMC moyen par individu estimé dans Ines-Omar 2019 (170 € par habitant tous soins remboursables confondus, et 158 € hors médicaments et hors soins hospitaliers).

La part de ménages exposés à des restes à charge élevés est probablement surestimée en utilisant l'enquête BdF par rapport à une estimation via le modèle Ines-Omar

En revanche, le RAC AMC selon le niveau de vie diffère entre BdF 2017, tel qu'estimée par Bricard (2024), et Ines-Omar 2019. Selon Bricard (2024), le reste à charge moyen par personne en 2017 oscillait entre 92 euros pour les personnes faisant partie du cinquième le moins aisé et 457 euros pour ceux faisant partie du cinquième le plus aisé. Or Ines-Omar 2019 estime le RAC AMC moyen par individu à 103 euros dans le premier quintile de niveau de vie, et à 249 euros dans le dernier cinquième de niveau de vie. L'écart de RAC AMC moyen des personnes les plus aisées entre Ines-Omar 2019 et l'enquête BdF 2017 pourrait s'expliquer par le fait qu'une plus grande part de leur RAC AMC provient de médicaments et de tests de diagnostic que pour les personnes les plus modestes, d'après Bricard (2024). Or l'écart de RAC AMC moyen entre Omar 2017 et BdF 2017 provient avant tout des produits pharmaceutiques (voir *tableau 20*).

Concernant la part de ménages exposés à des RAC dits « catastrophiques » (c'est-à-dire des RAC AMC ayant un poids important dans le revenu disponible du ménage), son calcul à partir de BdF serait biaisé et surestimé en raison de la méthode de collecte. Il s'avère par ailleurs que l'estimation par l'enquête BdF 2017 de la part de ménages exposés à des RAC « catastrophiques » est plus importante que celle par le modèle Omar : selon les calculs de la DREES, 6,2 % des ménages ont un RAC AMC représentant 5 % ou plus de leur revenu disponible dans l'enquête BdF 2017, contre 3,6 % dans Omar 2017.

Le modèle Ines-Omar ne permet pas de calculer la part de ménages subissant des RAC catastrophiques tels que définis par Bricard (2024) : ce sont des restes à charge pour des soins et services de santé représentant plus de 40 % de la capacité du ménage à payer pour des soins et services de santé. Cette capacité à payer des soins est définie comme la consommation totale du ménage à laquelle est soustrait un montant standard permettant de couvrir les besoins essentiels. Bricard, 2024, estime que 2,1 % des ménages supportaient en 2017 des RAC « catastrophiques » par rapport à leurs ressources. Le modèle Ines-Omar ne contient pas l'affectation du revenu disponible à différents types de dépenses, et ne permet donc pas de calculer la capacité résiduelle du ménage à payer pour des soins et services de santé.

³⁴ Bricard, D. (2024, avril). Les soins et services de santé sont-ils financièrement accessibles ? De nouvelles bases factuelles sur la protection financière en France. Organisation mondiale de la santé. Bureau régional de l'Europe.

Annexe 5. Pseudo-appariement par forêt aléatoire

Cette annexe présente la méthode de pseudo-appariement par forêt aléatoire employée pour appairer les individus de moins de 15 ans d'EHIS-ESPS 2014 à ceux de SRCV 2017. Les variables d'appariement et les paramètres d'appariement utilisés sont présentés dans le corps du texte.

Le principe général d'un certain nombre de techniques de pseudo-appariement est de partitionner en strates la base des « donneurs », ici EHIS-ESPS 2014, et la base des « receveurs », ici SRCV 2017. Puis chaque « receveur » se voit attribuer les données d'un individu tiré au sort dans la strate correspondante de la base des « donneurs ». Les différentes méthodes de pseudo-appariement se distinguent par la manière dont sont définies les strates, et par les modalités du tirage (avec ou sans remise, avec pondération, ...).

Comme dans Ines-Omar 2017, le pseudo-appariement d'EHIS-ESPS 2014 et de SRCV 2017 a été réalisé au moyen de la méthode des forêts aléatoires (voir Fouquet, 2021). Celle-ci est un raffinement de la méthode de pseudo-appariement par arbre de régression. À la différence du *hot-deck* (utilisé pour pseudo-appairer Ines et Omar 2019, présenté en partie VIII et en annexe 6), ces deux méthodes sont supervisées, c'est-à-dire que la définition des strates est dictée par un algorithme.

Le pseudo-appariement par arbre de régression

Le pseudo-appariement par arbre de régression, qui n'est pas appliqué dans Ines-Omar, est présenté ici à titre d'introduction au pseudo-appariement par forêt aléatoire.

Un arbre de régression est avant tout une méthode d'apprentissage statistique, qui permet de prédire une variable quantitative. Le principe de l'arbre de régression est de découper l'échantillon d'apprentissage en strates, au sein desquelles la dispersion de la variable à prédire est minimale, et entre lesquelles l'écart de la variable à prédire est maximale. À partir de cette méthode de prédiction, il est possible de construire une méthode de pseudo-appariement.

Dans notre exemple, la variable quantitative prédite est la dépense de santé totale d'un individu, sur l'échantillon des individus de moins de 15 ans d'EHIS 2014, et les variables transmises par le pseudo-appariement sont les dépenses de santé détaillées par poste des individus de moins de 15 ans d'EHIS 2014.

L'échantillon de départ est découpé entre un échantillon d'apprentissage sur lequel l'algorithme va s'entraîner et un échantillon de vérification qui permet d'évaluer la prédiction obtenue par l'arbre de régression. L'arbre est généré automatiquement par l'algorithme suivant :

- Pour chaque variable explicative disponible, trouver le seuil de coupure permettant de scinder au mieux l'échantillon en deux, c'est-à-dire en maximisant l'homogénéité de la variable expliquée, ici le montant total de dépenses de santé de l'individu, au sein de chaque sous-échantillon. Par exemple estimer que c'est l'âge de 5 ans qui sépare le mieux les enfants ayant des dépenses de santé élevées des enfants ayant des dépenses de santé faibles ;
- Déterminer quelle variable est la plus explicative de la dépense de santé, puis séparer l'échantillon d'apprentissage en deux sous-groupes de part et d'autre du seuil optimal précédemment trouvé. Par exemple, si la séparation des enfants en-deçà et au-dessus du 2^e dixième de niveau de vie explique mieux la dépense de santé que la séparation entre les deux sexes, ou que la séparation entre enfants de part et d'autre de 5 ans, alors on scinde l'échantillon selon le niveau de vie ;
- On applique de nouveau les deux étapes précédentes à chacun des sous-groupes obtenus ;
- L'algorithme s'arrête lorsqu'un sous-groupe (par exemple les enfants du 1^{er} et du 2^e dixième de niveau de vie, de plus de 3 ans, et d'état de santé perçu dégradé) atteint une taille trop faible pour être à nouveau subdivisé, ou que de nouvelles subdivisions n'améliorent plus la prédiction. L'ensemble des individus du groupe font alors partie d'une même strate. La taille minimale d'une strate est fixée par le statisticien ;
- La subdivision en strates obtenue est appliquée aux données de l'échantillon de vérification afin que le statisticien puisse vérifier les bonnes performances prédictives du modèle et modifier le paramétrage de l'arbre si nécessaire.

Les paramètres d'un arbre de régression sont donc : la liste de variables de stratification choisies, la taille minimale d'une strate, et l'amélioration minimale de la prédiction exigée pour qu'une strate puisse être scindée. Ces paramètres sont choisis par essais successifs, de sorte à produire l'algorithme le plus performant. Rappelons que la performance de l'algorithme est mesurée sur les données de vérification, qui sont parfaitement distinctes des données d'apprentissage.

Comment passer d'un arbre de régression, qui est une méthode de prédiction, à un algorithme de pseudo-appariement ? L'algorithme de pseudo-appariement utilise la partition en strates de l'échantillon d'apprentissage (ici l'échantillon des « donneurs », les individus de moins de 15 ans d'EHIS 2014) réalisée par l'algorithme de prédiction. L'échantillon des « receveurs », SRCV 2017, est stratifié de la même façon, et les strates identiques dans les

deux bases sont utilisées pour réaliser l'appariement. La méthode de tirage (avec ou sans remise, pondérée ou non) est ensuite à la discrétion du statisticien.

Cette méthode ressemble donc fortement au *hot-deck*, comme elle consiste aussi à répartir les individus dans des strates homogènes. Cependant les données de l'échantillon des « donneurs » (EHIS 2014-SNDS) sont mobilisées pour déterminer quelles variables expliquent la dépense de santé, comment les subdiviser et comment les hiérarchiser entre elles ; tandis que le *hot-deck* est intégralement paramétré par le statisticien. C'est pour cela que la méthode de pseudo-appariement par arbre de régression est dite supervisée.

Cette méthode de pseudo-appariement est particulièrement indiquée quand les variables transmises d'un « donneur » à un « receveur » sont corrélées à la variable qui est prédite par l'arbre de régression. C'est le cas des données transmises ici, à savoir les dépenses de santé détaillées par poste, corrélées par construction avec la dépense de santé totale.

Le pseudo-appariement par forêt aléatoire

De même que le pseudo-appariement par arbre de régression, la méthode de pseudo-appariement par forêt aléatoire repose sur un algorithme de prédiction, la forêt aléatoire. C'est la méthode qui est utilisée dans le modèle Ines-Omar pour appairier EHIS-ESPS 2014 et SRCV 2017.

La forêt aléatoire vise à faire de meilleures prédictions qu'un simple arbre de régression avec la même quantité de données. Au lieu de générer un arbre unique et complexe, on fait pousser une grande quantité d'arbres plus simples. Ces arbres sont générés en tirant au sort avec remise les individus de l'échantillon d'apprentissage et en ne conservant que certaines variables explicatives également tirées au sort parmi une liste proposée. À l'issue, la dépense de santé prédite est la moyenne de la dépense prédite par chaque arbre. En résumé, la forêt aléatoire consiste à multiplier les points de vue possibles afin de fournir une interprétation plus robuste des données.

Les paramètres de la méthode de prédiction par forêt aléatoire sont : le nombre d'arbres de régression, la liste de variables de stratification proposées, le nombre de variables utilisées par chaque arbre et tirées parmi la liste, la taille minimale d'une strate, et l'amélioration minimale de la prédiction pour scinder une strate.

Cet algorithme permet de prédire pour chaque individu la valeur de la variable expliquée, en l'occurrence le montant de dépenses de santé total pour chaque individu de moins de 15 ans d'EHIS 2014. Or notre objectif est d'imputer aux individus de SRCV 2017 non seulement un montant global, mais des données de santé fines décomposées par poste de soins entre remboursements, restes à charge, liberté tarifaire, etc... C'est pourquoi il ne s'agit pas de prédire la valeur d'une variable mais d'associer à chaque individu de SRCV un donneur issu de EHIS 2014, duquel seront tirées l'ensemble des dépenses de santé fines. Ceci nécessite de modifier le fonctionnement de la forêt aléatoire pour en faire un outil de pseudo-appariement plutôt qu'une méthode de prédiction.

Pour ce faire, pour un individu de SRCV donné, chacun des arbres de régression calculés pour générer la forêt sélectionne un ensemble d'individus de EHIS (une strate) qui semblent lui correspondre. Il est ensuite calculé pour chaque individu de EHIS 2014 combien de fois il a été associé à un même individu de SRCV. Enfin, un donneur est tiré au sort parmi les dix individus de EHIS 2014 les plus souvent pseudo-appariés au receveur des données. Comme dans la méthode de pseudo-appariement par arbre de régression, les modalités du tirage (pondération, avec ou sans remise, ...) sont à la discrétion du statisticien.

Ici il est procédé à un tirage avec remise, en pondérant le tirage d'un donneur par le nombre d'apparitions communes avec le receveur dans une strate. Ce sont les données de santé fines de ce donneur qui sont imputées à l'individu de SRCV considéré. Les paramètres de la méthode de pseudo-appariement par forêt aléatoire sont les mêmes que pour la méthode de prédiction par forêt aléatoire, auxquels s'ajoute le nombre de « donneurs » le plus souvent associés au sein d'une même strate à un « receveur » (ici fixé arbitrairement à dix).

Annexe 6. Pseudo-appariement par hot-deck

Définition des strates du « hot-deck »

Afin de réduire l'influence de l'aléa, les strates parmi lesquelles le tirage est effectué doivent avoir une taille minimale. La méthode usuelle pour pallier l'existence de strates de taille insuffisante est la suivante : si, en mobilisant l'ensemble des variables de stratification, certaines strates sont de taille insuffisante, les individus faisant partie de ces strates seront pseudo-appariés lors d'un deuxième tour, en utilisant une variable de stratification de moins. On procède ensuite à un troisième tour, puis à un quatrième, etc. en supprimant chaque fois une nouvelle variable de stratification, et ce jusqu'à ce que tous les individus soient pseudo-appariés. Les variables d'imputation sont donc ordonnées selon leur importance pour les analyses que l'on désire mener avec le modèle : type de couverture en premier, puis vingtième de niveau de vie, classe d'âge, statut d'occupation et enfin sexe.

Tableau 21 Simulation de l'imputation par hot-deck des données d'Omar 2019 vers Ines 2019 par la méthode usuelle

		1 ^{er} tour	2 ^e tour	3 ^e tour	4 ^e tour	5 ^e tour
Nombre et part d'individus d'Ines appariés		90 605	4 023	18 771	4 277	732
		76,5 %	3,4 %	15,9 %	3,6 %	0,6 %
Variables de stratification	Type de couverture	x	x	x	x	x
	Niveau de vie	x	x	x	x	
	Classe d'âge	x	x	x		
	Statut d'occupation	x	x			
	Sexe	x				

Lecture > En utilisant la méthode usuelle de hot deck, utilisée par Ines-Omar 2017, 4 023 individus du modèle Ines 2019 auraient été pseudo-appariés lors du deuxième tour, soit 3,4 % des individus. Toutes les variables de stratification sauf le sexe auraient été mobilisées dans ce deuxième tour.

Champ > Ensemble des individus du modèle Ines 2019.

Source > Ines 2019 ; Omar 2019. Traitements DREES.

La méthode usuelle revient donc à ne fusionner ensemble que des strates de donneurs insuffisamment peuplées et qui partagent certaines variables de stratification. Cependant, cette méthode présente l'inconvénient de parfois fusionner des strates éloignées : par exemple la strate des bénéficiaires de la CMU-C du 5^e vingtième de niveau de vie entre 30 et 39 ans, assez peu nombreux, pourrait être fusionnée avec celles des bénéficiaires de la CMU-C du 10^e au 20^e vingtième de niveau de vie, très peu nombreux. Dans ce cas, des individus de niveaux de vie très différents seraient placés dans la même strate.

Une autre méthode, que nous adoptons dans Ines-Omar 2019, consiste à fusionner les strates de donneurs insuffisamment peuplées avec les strates de donneurs qui leur sont les plus proches (que ces dernières soient suffisamment peuplées ou non) au fur et à mesure des tours d'appariement. L'esprit de cette nouvelle méthode est de constituer des strates homogènes du point de vue des variables d'appariement essentielles (couverture, niveau de vie, âge) quitte à obtenir des strates volumineuses, plutôt que les plus petites strates possibles au risque d'introduire de l'hétérogénéité entre individus de certaines strates.

La mise en place des fusions entre strates proches dans l'échantillon des donneurs n'obéit pas à une règle systématique et reste à l'appréciation du statisticien. Voici la procédure de fusions des strates d'Omar 2019 adoptée dans Ines-Omar 2019 :

- Au premier niveau de stratification, les strates sont définies par le croisement de cinq variables (couverture, niveau de vie, âge, type d'occupation, sexe).
- On fusionne les strates de taille insuffisante avec leur strate voisine, définie par le sexe opposé. De ces fusions résultent des strates de niveau couverture x niveau de vie x âge x occupation. Si celles-ci atteignent la taille minimale, elles sont conservées ; autrement ces strates seront fusionnées ultérieurement. On ne fusionne pas entre elles les strates de niveau couverture x niveau de vie x âge x occupation de taille insuffisante et de même racine couverture x niveau de vie x âge.
- Le troisième niveau de stratification, portant sur les individus hors d'une strate de taille suffisante, croise les variables couverture x niveau de vie x âge. Tout d'abord, on tente de fusionner chaque strate insuffisamment peuplée avec une strate suffisamment peuplée précédemment créée et partageant les mêmes trois premiers niveaux : il en résulte une strate de niveau couverture x niveau de vie x âge.

- Puis on tente de fusionner les strates insuffisamment peuplées restantes de niveau couverture x niveau de vie x âge avec une ou plusieurs strates de même couverture x niveau de vie et de classe d'âge adjacente, elles-mêmes de taille insuffisante. De cette fusion résulte une strate au niveau couverture x niveau de vie. Si la fusion présente une taille satisfaisante, celle-ci est conservée.
- À défaut, on fusionne les strates de taille insuffisante restante avec la plus proche strate de taille suffisante, qui partage les niveaux couverture x niveau de vie, voire uniquement le niveau couverture.
- À l'issue de cette étape, il n'existe plus de strates de donneurs insuffisamment peuplées. Subsistent quelques strates de niveau couverture x niveau de vie x âge de taille insuffisante dans la base des receveurs, et qui n'ont pas de strate équivalente dans la base des donneurs. Celles-ci sont fusionnées avec la plus proche strate de taille suffisante.

La taille minimale d'une strate est fixée à cinquante donneurs potentiels. Comme le montrent les tableaux 21 et 22, cette nouvelle méthode de *hot-deck* apparie 3 % des individus d'Ines avec une ou deux variables de stratification seulement, contre 4,2 % avec la méthode usuelle. Ainsi, la nouvelle méthode d'appariement réduit le risque que l'appariement de certains individus se fasse sur un nombre très restreint de variables. En revanche la nouvelle méthode apparie moins précisément les individus en fonction de leur sexe et de leur type d'occupation.

Tableau 22 Déroulement de l'imputation par *hot-deck* des données d'Omar 2019 vers Ines 2019, avec la méthode d'Ines-Omar 2019

		Strates à 5 niveaux	Strates à 4 niveaux	Strates à 3 niveaux	Strates à 2 niveaux	Strates à 1 niveau
Nombre et part d'individus d'Ines appariés		75 192	10 446	29 206	2 864	700
		63,5 %	8,8 %	24,7 %	2,4 %	0,6 %
Variables de stratification	Type de couverture	x	x	x	x	x
	Niveau de vie	x	x	x	x	
	Classe d'âge	x	x	x		
	Statut d'occupation	x	x			
	Sexe	x				

Lecture > 75 192 individus du modèle Ines 2019 ont été pseudo-appariés dans des strates avec cinq niveaux de stratification, c'est-à-dire mobilisant toutes les variables de stratification, soit 63,5 % des individus.

Champ > Ensemble des individus du modèle Ines 2019.

Source > DREES-Insee, Ines 2019 ; DREES, Omar 2019. Traitements DREES.

Comparer simplement le niveau des strates du *hot-deck* entre la méthode usuelle et la nouvelle méthode d'imputation ne rend pas compte du fait que les strates constituées par la nouvelle méthode sont plus homogènes, à niveau de stratification identique. Ainsi le nombre d'individus d'Ines qui sont appariés avec un individu d'Omar dont l'âge ou le niveau de vie sont très éloignés diminue fortement : avec la méthode usuelle d'appariement, 2 298 individus d'Ines auraient été appariés avec un individu d'Omar avec au moins 20 ans d'écart, contre 873 avec la nouvelle méthode ; avec la méthode usuelle 172 individus d'Ines auraient eu un écart de niveau de vie avec leur donneur d'au moins 10 vingtiles de niveau de vie, contre 10 avec la nouvelle méthode.

DREES MÉTHODES
N° 18 • septembre 2024

Le modèle de microsimulation Ines-Omar 2019

Directeur de la publication
Fabrice Lengart

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 www.data.drees.sante.fr
