

L'espérance de vie par catégorie sociale et par diplôme jusqu'en 2020-2022 - Méthode

Documents de travail

N° 2024-17 - Juillet 2024





Institut national de la statistique et des études économiques

N°2024-17

L'espérance de vie par catégorie sociale et par diplôme jusqu'en 2020-2022 - Méthode

Nathalie BLANPAIN*

Juillet 2024

Direction des Statistiques Démographiques et Sociales - Timbre F001
88, avenue Verdier - CS 70058 - 92541 MONTRouGE CEDEX - France
Tél. : 33 (1) 87 69 62 82 - E-mail : dg75-f001@insee.fr - Site Web Insee : <http://www.insee.fr>

*Ces documents de travail ne reflètent pas la position de l'Insee et n'engagent que leurs auteurs.
Working papers do not reflect the position of Insee but only their author's views.*

* Insee, division Enquêtes et études démographiques

Résumé

L'objectif de ce document est de présenter la méthode permettant de calculer les espérances de vie par catégorie sociale et diplôme. La source est l'échantillon démographique permanent (EDP), qui permet de suivre la mortalité d'un échantillon d'individus, au fil du temps, en fonction de leurs caractéristiques sociodémographiques. Tout d'abord, un travail d'imputation et de correction des caractéristiques sociodémographiques a été réalisé afin d'assurer l'homogénéité de ces caractéristiques dans le temps. Puis, pour les périodes anciennes, la sous-estimation des quotients de mortalité aux grands âges a été corrigée par la méthode de Brass. Pour les périodes récentes, les quotients de mortalité ne sont plus sous-estimés, seul un lissage était nécessaire : il a été effectué par la méthode des splines. Enfin, un calage sur les espérances de vie du bilan démographique a été effectué.

Une synthèse des résultats est présentée dans l'Insee Première n° 2005 « Les écarts d'espérance de vie entre cadres et ouvriers : 5 ans chez les hommes, 3 ans chez les femmes ». Les tables de mortalité par catégorie sociale et diplôme sont disponibles dans l'Insee Résultats « Tables de mortalité par catégorie sociale et par diplôme jusqu'en 2020-2022 ».

Mots-clés : table de mortalité, espérance de vie par catégorie sociale, espérance de vie par diplôme, quotient de mortalité, France, 2022

Summary

The aim of this document is to present the method used to calculate life expectancy by social category and qualification. The source is the permanent demographic sample (EDP), which tracks the mortality of a sample of individuals over time, according to their sociodemographic characteristics. Firstly, sociodemographic characteristics were imputed and corrected to ensure that they were consistent over time. Then, for older periods, the underestimation of mortality rates at older ages was corrected using the Brass method. For recent periods, the mortality rates are no longer underestimated, and only smoothing was necessary: this was carried out using the spline method. Finally, a calibration on the benchmark life expectancies was carried out.

A summary of the results is available in Insee Première No. 2005, "Les écarts d'espérance de vie entre cadres et ouvriers: 5 ans chez les hommes, 3 ans chez les femmes". Mortality tables by social category and qualification are available in Insee Résultats "Tables de mortalité par catégorie sociale et par diplôme jusqu'en 2020-2022".

Key-words: life table, life expectancy by social category, life expectancy by qualification, mortality rate, France, 2022

Table des matières

INTRODUCTION.....	6
PARTIE 1 : LES DONNÉES.....	8
1) L'ÉCHANTILLON DEMOGRAPHIQUE PERMANENT (EDP).....	8
2) LA CATÉGORIE SOCIALE.....	9
A. CONCEPT.....	9
B. REDRESSEMENT.....	12
C. ÉVOLUTION DE LA STRUCTURE PAR CATÉGORIE SOCIALE DE LA POPULATION.....	15
3) LE DIPLÔME.....	16
A. CONCEPT.....	16
B. REDRESSEMENT.....	16
C. ÉVOLUTION DE LA STRUCTURE PAR DIPLÔME DE LA POPULATION.....	17
PARTIE 2 : LA MÉTHODE.....	18
1) QUOTIENTS DE MORTALITÉ.....	19
2) L'ESPÉRANCE DE VIE.....	21
3) LISSAGE ET ESTIMATION DES QUOTIENTS DE MORTALITÉ PAR LA MÉTHODE DE BRASS POUR LES PÉRIODES ANCIENNES (DE 1976-1984 À 2009-2013).....	22
A. PRINCIPE GÉNÉRAL.....	22
B. AMÉNAGEMENT DE LA MÉTHODE DE BRASS.....	23
C. EXEMPLE DE LISSAGE PAR LA MÉTHODE DE BRASS.....	25
D. EFFET DE LA MÉTHODE DE BRASS SUR L'ESTIMATION DES ESPÉRANCES DE VIE.....	26
4) LISSAGE ET ESTIMATION DES QUOTIENTS DE MORTALITÉ PAR LA MÉTHODE DES SPLINES POUR LES PÉRIODES RÉCENTES (2017-2019 ET 2020-2022).....	27
A. MÉTHODE.....	27
B. EFFET DU LISSAGE PAR LA MÉTHODE DES SPLINES SUR L'ESTIMATION DES ESPÉRANCES DE VIE...30	
5) CALAGE SUR LES TABLES DE RÉFÉRENCE DU BILAN DÉMOGRAPHIQUE.....	31
A. EFFET DE LA DATE DE NAISSANCE DES INDIVIDUS EDP.....	31
B. EFFET DU LIEU DE NAISSANCE.....	32
C. CALAGE SUR LE BILAN DÉMOGRAPHIQUE.....	32
6) INTERVALLE DE CONFIANCE A 90 % DES ESPÉRANCES DE VIE.....	32
PARTIE 3. LES ÉCARTS D'ESPÉRANCE DE VIE ENTRE CADRES ET.....	34
OUVRIERS : 5 ANS CHEZ LES HOMMES, 3 ANS CHEZ LES FEMMES.....	34

INTRODUCTION

Comment ont évolué les écarts d'espérance de vie par catégorie sociale ou par diplôme au fil du temps ? L'espérance de vie des hommes et des femmes est publiée chaque année par l'Insee dans le bilan démographique, à partir des estimations de population et des statistiques d'état civil¹. L'évolution de l'espérance de vie par catégorie sociale et par diplôme est analysée à partir de l'Échantillon Démographique Permanent (EDP), produit par l'Insee. Un document de travail a été publié par l'Insee en 2005² sur l'espérance de vie par catégorie sociale pour la période de la fin des années 1970 au milieu des années 1990. En 2016, un document de travail³ a actualisé ces travaux jusqu'aux années 2009-2013 et a étendu l'analyse aux écarts d'espérance de vie par diplôme. L'objectif de ce document est de mettre à jour les espérances de vie par catégorie sociale et diplôme pour deux périodes récentes (2017-2019 et 2020-2022) et de présenter la méthode de construction des tables de mortalité.

La source utilisée est l'EDP millésime 2021⁴. Il s'agit d'un échantillon de personnes pour lesquelles sont compilées les informations des recensements (notamment la catégorie socioprofessionnelle et le diplôme) ainsi que celles de l'état civil (notamment les décès).

Les tables de mortalité, compilant pour chaque âge le quotient de mortalité, la probabilité de survie et l'espérance de vie, sont estimées pour chaque catégorie sociale ou diplôme, sur une période d'observation de l'état vital de quelques années. Ainsi, les quotients de mortalité sont calculés selon :

- la CS 1975 sur la période 1976-1984,
- la CS 1982 sur la période 1983-1991,
- la CS et le diplôme 1990 sur la période 1991-1999,
- la CS et le diplôme 1999 sur la période 2000-2008,
- la CS et le diplôme de 2004 à 2012 sur la période 2009-2013,
- la CS et le diplôme de 2010 à 2018 sur la période 2017-2019,
- la CS et le diplôme de 2013 à 2020 sur la période 2020-2022.

La première partie de ce document décrit la source utilisée, détaille les redressements effectués pour assurer l'homogénéité dans le temps de la catégorie sociale et du diplôme, puis décrit la répartition de la population par CS et diplôme.

¹ Papon S., « Bilan démographique 2023 - En 2023, la fécondité chute, l'espérance de vie se redresse », Insee Première n°1978, 2024.

² Monteil C., Robert-Bobée I., « Quelles évolutions des différentiels sociaux de mortalité pour les femmes et les hommes », Insee, Documents de travail n° F0506, 2005.

³ Blanpain N., « L'espérance de vie par catégorie sociale et par diplôme - méthode et principaux résultats », Insee, Document de travail n° F1602, 2016.

⁴ Ce millésime inclut les EAR jusqu'en 2020 et les décès jusqu'en 2023.

La deuxième partie est centrée sur la méthode de construction des tables de mortalité par catégorie sociale ou diplôme. Elle présente :

- la méthode de Brass qui permet à la fois de redresser les quotients de mortalité des jeunes adultes et des personnes âgées et de les lisser pour les cinq périodes les plus anciennes,
- la méthode des splines qui permet de lisser les quotients de mortalité pour les deux périodes les plus récentes,
- le calage de l'espérance de vie calculée à partir de l'EDP sur l'espérance de vie de référence issue du bilan démographique,
- le calcul des intervalles de confiance de l'espérance de vie par catégorie sociale et diplôme.

Les principaux résultats sont présentés dans l'Insee Première n°2005 « Les écarts d'espérance de vie entre cadres et ouvriers : 5 ans chez les hommes, 3 ans chez les femmes ? ».

Les tables de mortalité par catégorie sociale et diplôme sont disponibles dans l'Insee Résultats « Tables de mortalité par catégorie sociale et par diplôme jusqu'en 2020-2022 ».

PARTIE 1 : LES DONNÉES

1) L'ÉCHANTILLON DEMOGRAPHIQUE PERMANENT (EDP)

L'Échantillon Démographique Permanent (EDP) est un panel d'individus qui compile des informations puisées dans plusieurs sources pour des individus nés certains jours de l'année. C'est un outil qui permet notamment d'étudier les comportements sociodémographiques.

L'EDP mobilise cinq sources : les bulletins d'état civil depuis 1967 ; les cinq recensements de population de 1968 à 1999 et les enquêtes annuelles de recensement (EAR) depuis 2004 ; le fichier électoral depuis 1990 ; les panels d'actifs depuis 1967 ; les données socio-fiscales issues de Fidéli (Fichier démographique d'origine fiscale sur les logements et les individus) et de FiLoSoFi (Fichier localisé social et fiscal) depuis 2011.

Le critère d'échantillonnage est basé sur le jour de naissance : tous les individus nés du 2 au 5 janvier, du 1^{er} au 4 avril, du 1^{er} au 4 juillet et du 1^{er} au 4 octobre font partie du champ de l'EDP⁵. Pour cette étude, l'échantillon concerne les personnes recensées.

Le renouvellement de l'échantillon est assuré par les naissances ou par la venue en France de nouvelles personnes. Si elles sont nées un jour EDP, ces dernières rejoignent l'EDP à l'occasion d'un recensement ou d'un événement enregistré dans un acte de l'état civil. À l'inverse, le suivi d'un individu cesse, de fait, en cas de décès ou pendant une période de vie à l'étranger.

L'EDP n'a intégré les personnes domiciliées dans les départements d'outre-mer (hors Mayotte) que depuis 2004 pour l'état civil et les EAR. Pour cette étude, le champ est donc la France métropolitaine de 1976-1984 à 2000-2008, la France métropolitaine et la France hors Mayotte en 2009-2013, et la France hors Mayotte en 2017-2019 et en 2020-2022.

L'EDP permet de disposer de la CS et du diplôme des individus grâce au recensement, ainsi que de la date de naissance⁶ et de la date de décès éventuelle grâce aux données d'état civil. On peut donc estimer, à l'aide d'une même source, les quotients de mortalité par CS et par diplôme pour des différentes périodes, et analyser ainsi l'évolution des inégalités sociales de mortalité. Pour la catégorie sociale, sept périodes de mortalité sont étudiées (1976-1984, 1983-1991, 1991-1999, 2000-2008, 2009-2013, 2017-2019, 2020-2022).

Les milieux des périodes d'observation des décès sont éloignés de 8 ans en moyenne (1980, 1987, 1995, 2004, 2011, 2018). Un espacement relativement important a été choisi, pour permettre d'analyser les évolutions d'espérance de vie par CS ou diplôme d'une période à l'autre, qui, étant donné les erreurs dues à l'échantillonnage, aux redressements et aux imputations, sont à prendre

⁵ Historiquement, l'échantillon ne comprenait que les personnes nées les quatre premiers jours d'octobre. Les individus EDP nés en janvier, avril et juillet sont observés à partir de dates variables selon la source.

⁶ On retient la date de naissance de l'état civil, car elle est plus fiable que celle du recensement.

avec précaution. Pour les deux dernières périodes, il a néanmoins été choisi deux périodes assez proches afin d'avoir des données sur une période antérieure à la pandémie de Covid-19 (2017-2019) et sur une période où la pandémie a entraîné une surmortalité (2020-2022).

Les deux dernières périodes de mortalité incluent 3 années (2017-2019, 2020-2022), contre 5 années pour 2009-2013. Cette réduction de la période est possible grâce à l'élargissement de l'échantillon à partir de l'EAR 2008 (les individus nés 16 jours particuliers dans l'année, contre 4 jours auparavant). On compte 4,3 millions d'années-personnes pour la période 2009-2013, 5,4 millions pour 2017-2019 et 5,1 millions pour 2020-2022.

Pour le diplôme, seules les cinq périodes les plus récentes sont analysées (à partir de 1991-1999). En effet, les diplômes ne sont connus que pour le quart de la population pour les recensements de 1975 et 1982. Les effectifs ne sont donc pas suffisants pour assurer la robustesse des quotients de mortalité par âge et par diplôme pour les périodes plus anciennes.

La source de toutes les figures de ce document est l'échantillon démographique permanent.

2) LA CATÉGORIE SOCIALE

A. CONCEPT

La catégorie sociale (CS) est issue du recensement. Les chômeurs ayant déjà travaillé et les retraités sont classés dans la CS de leur dernier emploi. Au niveau agrégé (CS à un chiffre), la population est répartie en sept catégories :

- agriculteurs,
- artisans commerçants et chefs d'entreprise,
- cadres et professions intellectuelles supérieures,
- professions intermédiaires,
- employés,
- ouvriers,
- inactifs non retraités.

Les « inactifs non retraités » (poste 8 de la PCS) regroupent :

- Les chômeurs n'ayant jamais travaillé (poste 81 de la PCS)
- les inactifs non retraités (poste 82 de la PCS). Il peut s'agir de personnes n'ayant jamais travaillé ou bien de personnes ayant déjà travaillé, mais inactives non retraités au moment du recensement. Il n'y a pas de condition d'âge⁷.

⁷ Le document de travail de l'Insee, n° F1108, paru en septembre 2011, présente une comparaison de l'espérance de vie des femmes et des hommes selon la CS, selon que les inactives ayant déjà retravaillé sont reclassées ou non dans la CS du dernier emploi occupé.

La nomenclature a connu plusieurs rénovations : en 1982, en 2003 et en 2020, cette dernière version n'étant cependant pas encore déployée dans l'EPD millésime 2021.

En 1975, la nomenclature de la CS (à un chiffre) est légèrement différente de celle utilisée à partir de 1982. En 1975, les professions intermédiaires regroupent les cadres moyens (instituteurs, professions intellectuelles diverses, services médicaux et sociaux, techniciens, cadres administratifs moyens). Les autres catégories (artistes, clergé, armée et police) ont été exclues. La nomenclature de 2003 est identique à celle de 1982 pour la CS à un chiffre.

L'espérance de vie a aussi été calculée au niveau CS détaillé (à deux chiffres) en 2017-2019 et 2020-2022. La population est répartie en 12 catégories pour les femmes (F) et 15 pour les hommes (H) :

- Artisans (poste 21 de la nomenclature pcs 2020) - H
- Commerçants et assimilés (22) - F/H
- Cadres de la fonction publique, professions intellectuelles et artistiques (33, 34 et 35) - F/H
- Cadres administratifs et commerciaux (37) - F/H
- Cadres techniques d'entreprise (38) - H
- Professions de l'enseignement primaire et professionnel (42) - F/H
- Intermédiaires de la santé et du travail social (43) - F
- Intermédiaires des entreprises (46) - F/H
- Techniciens (47) - H
- Agents de maîtrise de production (48) - H
- Employés de la fonction publique (52) - F/H
- Policiers, militaires et agents de sécurité privée (53) - H
- Employés administratifs d'entreprise (54) - F/H
- Employés de commerce (55) - F
- Personnels des services aux particuliers (56) - F
- Ouvriers qualifiés de type industriel ou artisanal (62 et 63) - F/H
- Conducteurs du transport (64) F/H
- Ouvriers peu qualifiés de type industriel ou artisanal (67 et 68) - F/H

Ces modalités et regroupements ont été retenus afin de disposer d'effectifs suffisants pour calculer les espérances de vie à 35 ans des femmes et des hommes.

Lorsque les effectifs sont petits, comme par exemple pour les agricultrices (2,78 % des femmes âgées de 35 ans ou plus recensées entre 2013 et 2020), l'espérance de vie doit être analysée avec précaution. L'intervalle de confiance à 95 % est de +/-0,8 ans (cf. partie 6). De plus, l'écart entre les données brutes et lissées est de 0,9 ans : des quotients de mortalité sont nuls à certains âges, en raison d'effectif réduit (cf. partie 4-B).

Pour la CS détaillée, on retient uniquement les catégories représentant au moins 2,57 % des femmes ou des hommes de 35 ans ou plus recensés entre 2013 et 2020 (catégories surlignées en vert dans les tableaux ci-après). En dessous de ce seuil, la courbe des quotients de mortalité par âge est chaotique et le calcul de l'espérance de vie est soumis à une grande incertitude, du fait des erreurs dues à l'échantillonnage, des erreurs de déclarations des enquêtés, des erreurs de correction et d'imputations (la CS est manquante et donc imputée pour 19 % des personnes de 80 ans ou plus). Par exemple, l'intervalle de confiance (qui mesure seulement l'erreur due à l'échantillonnage de

l'espérance de vie des conducteurs d'engin (2,53 % des hommes) est de +/-0,8 ans. De plus, du fait de l'incertitude, le niveau de l'espérance de vie des conducteurs d'engin est assez différent en 2017-2019 et 2020-2022. Le calcul des espérances de vie en dessous du seuil de 2,57 % n'a donc pas été retenu. La part de la population non couverte par l'analyse de la CS détaillée est de 9,6 % chez les femmes et 11,7 % chez les hommes.

Le calcul des espérances de vie par CS (à 4 chiffres) n'est pas possible en raison d'effectif insuffisant pour la quasi-totalité des professions. Chez les hommes, aucune profession ne représente au moins 2,57 % de la population. Chez les femmes, seules deux professions atteignent ce seuil : les secrétaires et les employées des services comptables et financiers.

Le calcul des espérances de vie par CS (à 4 chiffres) n'est pas possible en raison d'effectif insuffisant pour la quasi-totalité des professions. Chez les hommes, aucune profession ne représente au moins 2,57 % de la population. Chez les femmes, seules deux professions atteignent ce seuil : les secrétaires et les employées des services comptables et financiers.

Répartition des femmes par catégorie sociale détaillée (en %)

libellé	code	répartition	répartition
Agriculteurs	10	2,78	2,78
Artisans	21	1,55	1,55
Commerçants et assimilés	22	2,63	2,63
Chefs d'entreprise de plus de 10 personnes	23	0,20	0,20
Professions libérales	31	1,03	1,03
Cadres de la fonction publique	33	1,21	3,72
Professeurs et professions scientifiques	34	1,92	
Professions de l'information, de l'art et des spectacles	35	0,59	
Cadres administratifs et commerciaux	37	2,80	2,80
Cadres techniques d'entreprise	38	1,02	1,02
Professions de l'enseignement primaire et professionnel	42	5,25	5,25
Intermédiaires de la santé et du travail social	43	6,73	6,73
Religieux	44	0,03	0,03
Intermédiaires de la fonction publique	45	1,95	1,95
Intermédiaires des entreprises	46	7,29	7,29
Techniciens	47	1,26	1,26
Agents de maîtrise de production	48	0,63	0,63
Employés de la fonction publique	52	13,20	13,20
Policiers, militaires et agents de sécurité privée	53	0,45	0,45
Employés administratifs d'entreprise	54	11,90	11,90
Employés de commerce	55	5,06	5,06
Personnels des services aux particuliers	56	10,01	10,01
Ouvriers qualifiés de type industriel	62	1,37	3,00
Ouvriers qualifiés de type artisanal	63	1,63	
Conducteurs du transport	64	0,38	0,38
Conducteurs d'engins et magasiniers	65	0,32	0,32
Ouvriers peu qualifiés de type industriel	67	4,65	5,74
Ouvriers peu qualifiés de type artisanal	68	1,09	
Ouvriers agricoles	69	0,83	0,83
Chômeurs n'ayant jamais travaillé	81	0,16	0,16
Personnes sans activité professionnelle de moins de 60 ans	85	6,37	6,37

Personnes sans activité professionnelle de 60 ans et plus	86	3,69	3,69
Ensemble		100,00	100,00

Champ : cumul des femmes âgées de 35 ans ou plus nées en France, recensées entre 2013 et 2020, décédées entre 2020 et 2022 ou vivantes au 1^{er} janvier entre 2020 et 2022

Répartition des hommes par catégorie sociale détaillée (en %)

libellé	code	répartition	répartition
Agriculteurs	10	4,34	4,34
Artisans	21	5,43	5,43
Commerçants et assimilés	22	3,60	3,60
Chefs d'entreprise de plus de 10 personnes	23	0,89	0,89
Professions libérales	31	1,67	1,67
Cadres de la fonction publique	33	1,63	4,16
Professeurs et professions scientifiques	34	1,73	
Professions de l'information, de l'art et des spectacles	35	0,80	
Cadres administratifs et commerciaux	37	4,59	4,59
Cadres techniques d'entreprise	38	5,60	5,60
Professions de l'enseignement primaire et professionnel	42	2,65	2,65
Intermédiaires de la santé et du travail social	43	1,82	1,82
Religieux	44	0,08	0,08
Intermédiaires de la fonction publique	45	1,31	1,31
Intermédiaires des entreprises	46	5,70	5,70
Techniciens	47	6,37	6,37
Agents de maîtrise de production	48	3,78	3,78
Employés de la fonction publique	52	3,85	3,85
Policiers, militaires et agents de sécurité privée	53	3,32	3,32
Employés administratifs d'entreprise	54	2,57	2,57
Employés de commerce	55	1,28	1,28
Personnels des services aux particuliers	56	0,88	0,88
Ouvriers qualifiés de type industriel	62	6,86	15,1
Ouvriers qualifiés de type artisanal	63	8,24	
Conducteurs du transport	64	4,65	4,65
Conducteurs d'engins et magasiniers	65	2,53	2,53
Ouvriers peu qualifiés de type industriel	67	6,30	8,46
Ouvriers peu qualifiés de type artisanal	68	2,16	
Ouvriers agricoles	69	1,27	1,27
Chômeurs n'ayant jamais travaillé	81	0,09	0,09
Personnes sans activité professionnelle de moins de 60 ans	85	3,13	3,13
Personnes sans activité professionnelle de 60 ans et plus	86	0,89	0,89
Ensemble		100,00	100,00

Champ : cumul des hommes âgés de 35 ans ou plus nés en France, recensés entre 2013 et 2020, décédés entre 2020 et 2022 ou vivants au 1^{er} janvier entre 2020 et 2022.

B. REDRESSEMENT

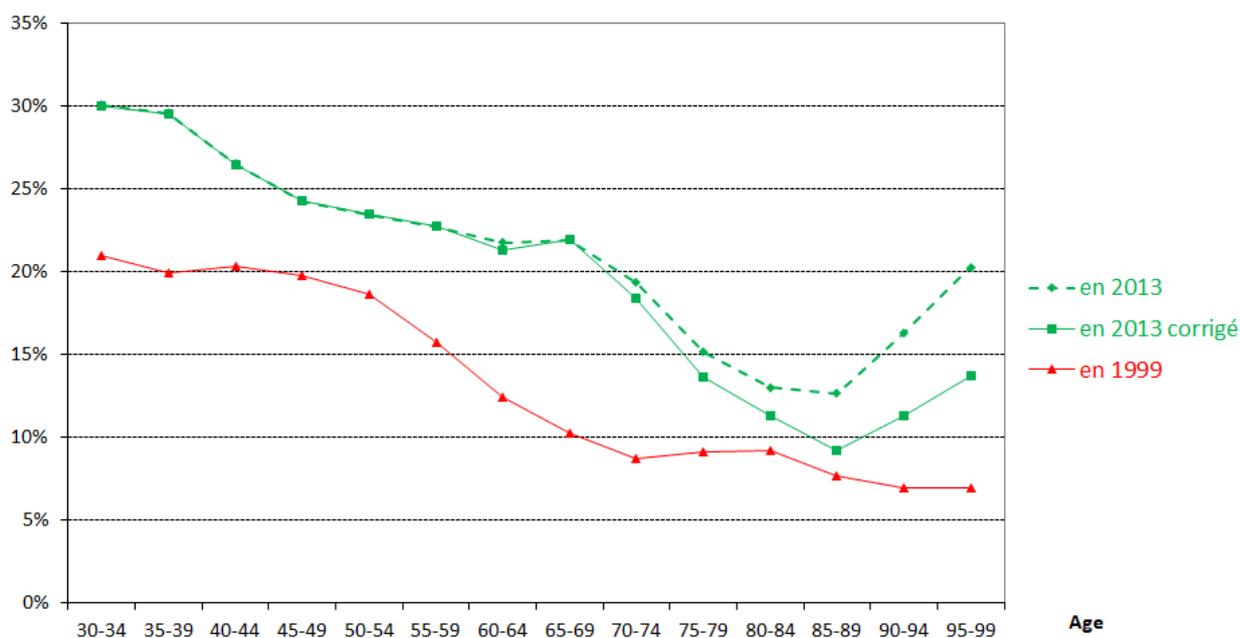
Des redressements spécifiques sont effectués sur la CS, soit pour améliorer les redressements effectués par les traitements du recensement, soit pour corriger la non-réponse partielle.

Dans les RP de 1975 et 1982, la dernière CS d'une petite fraction des retraités n'est pas connue (1,0 % des hommes et 1,5 % des femmes en 1975 ; 2,5 % des hommes et 4,7 % des femmes en 1982). Lorsque la CS déclarée au recensement précédent était disponible, elle a été retenue.

Dans les EAR de 2004 à 2017⁸, la catégorie sociale (à 1 chiffre) a été corrigée de façon à homogénéiser les données mobilisées sur longue période.

Une première correction a concerné les personnes qui déclarent être retraitées et qui déclarent par ailleurs n'avoir jamais travaillé. Sur ce point, les apurements et les imputations ont été réalisés de manière différente dans les recensements jusqu'à 1999, puis dans les enquêtes annuelles de recensement. Pour les EAR de 2004 à 2017, la proportion d'inactives non retraitées est stable au-delà de 65 ans alors qu'en 1999 elle augmentait avec l'âge. Cela provient de corrections différentes entre le recensement de 1999 et les enquêtes annuelles de recensement qui ont suivi. Une personne qui, au RP 1999, déclare n'avoir jamais travaillé et qui déclare par ailleurs être retraitée est classée en « inactif non retraité ». De 2004 à 2017, une CS lui est imputée. L'imputation s'effectue par hotdeck parmi les retraités de même lot de saisie⁹, de même sexe et de même tranche d'âge (60-64 ans, 65 ans ou plus). Afin d'avoir une approche homogène dans le temps, on a effectué le redressement suivant : les personnes qui déclarent n'avoir jamais travaillé ont été reclassées en « inactives non retraitées ». La proportion des inactives non retraitées est ainsi corrigée à la hausse aux âges élevés. Le profil de la courbe est ainsi plus proche de celui observé en 1999 et plus conforme à ce qui est attendu (hausse des femmes au foyer pour les générations les plus anciennes).

Proportion de femmes professions intermédiaires selon l'âge



Champ : femmes âgées de 30 ans ou plus nées en France.

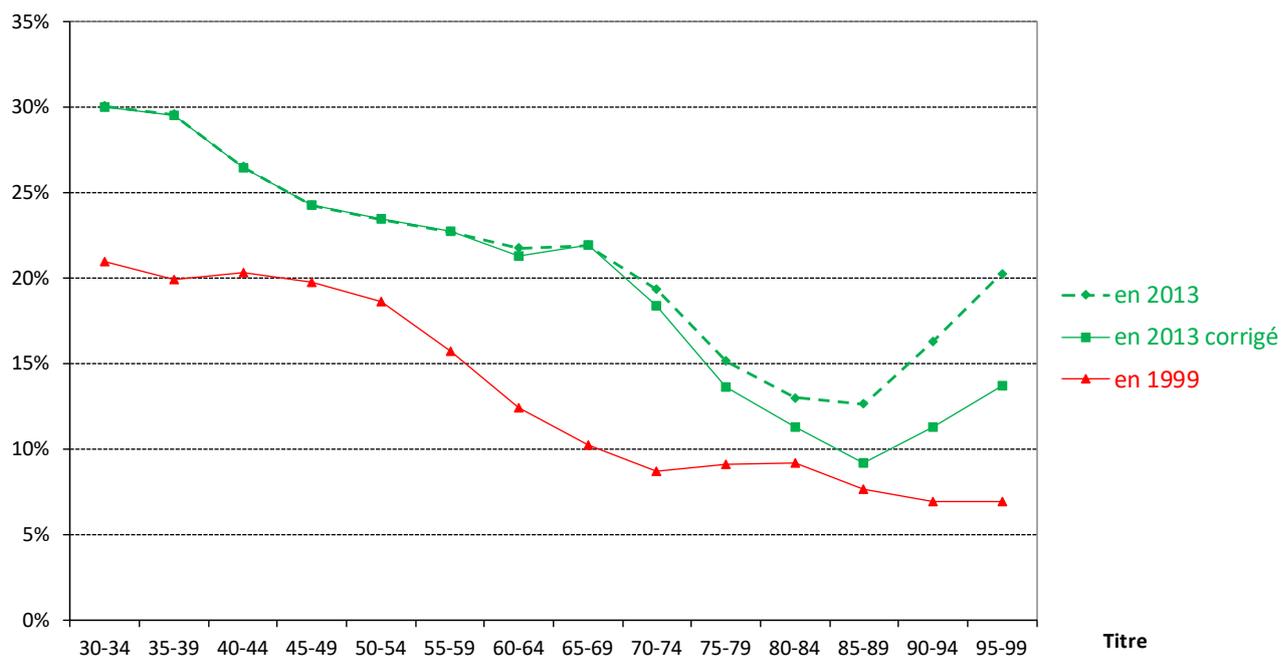
⁸ À partir de 2018, les apurements et les imputations de la CS dans les EAR deviennent conformes à ceux effectués dans le passé, i.e en 1999.

⁹ Les questionnaires sont saisis par paquets, appelés lots de saisie.

Note : La proportion d'inactifs non retraités diminue nettement aux alentours de 60 ans. Les retraités sont en effet classés dans leur ancienne CS.

La seconde correction concerne l'imputation des retraités déclarant avoir travaillé dont la CS n'est pas renseignée. La proportion de femmes professions intermédiaires est plus élevée entre le recensement de 1999 et les EAR de 2004 à 2017 pour des générations identiques. Cela provient de la hausse de cette CS entre les deux périodes, mais aussi de traitements différents dans le codage de la catégorie sociale. Dans le cas où l'information est manquante dans le recensement, la catégorie sociale est imputée. De 2004 à 2017, l'imputation de la CS pour les retraités s'effectue dans l'EAR par hotdeck parmi les retraités de même lot de saisie, de même sexe et de même tranche d'âge (60-64 ans, 65 ans ou plus). La tranche d'âge retenue pour l'imputation est trop large et conduit à imputer trop souvent une CS élevée aux personnes très âgées. En effet, la répartition par CS a changé au fil des générations, et les femmes très âgées avaient des CS moins élevées que celles proches de 65 ans. La méthode d'imputation des CS des retraités a donc été modifiée. On récupère l'information si elle est disponible dans les autres EAR (de 2004 à 2020) ou bien dans le recensement de 1999, sinon elle est imputée en ajoutant, par rapport à la méthode des EAR 2004-2017, des classes d'âges et des variables supplémentaires très liées à la CS (sexe, diplôme). On a ainsi imputé la catégorie sociale des retraités par hotdeck parmi les retraités de même sexe, de même âge quinquennal (les 95 ans ou plus étant regroupés) et de même diplôme (en 4 modalités : sans diplôme ; brevet, CAP, BEP ; baccalauréat ; supérieur au baccalauréat). Cette correction permet d'avoir des classements plus homogènes entre 1999 et 2004-2017. Elle a pour effet de corriger l'espérance de vie des femmes professions intermédiaires, qui était sous-évaluée avant correction (l'espérance de vie des femmes professions intermédiaires évoluait peu entre les périodes 2000-2008 et 2009-2013).

Proportion de femmes professions intermédiaires selon l'âge



Champ : femmes âgées de 30 ans ou plus nées en France.

La catégorie sociale détaillée (à deux chiffres, cf. partie 1 2.a) a également été corrigée de 2004 à 2018¹⁰. Elle a été corrigée et imputée de la même manière que la CS à un chiffre, sauf dans le cas où la CS à un chiffre est correctement renseignée et la CS détaillée manquante. Dans ce cas, l'imputation de la CS détaillée s'effectue parmi les personnes de même sexe, de même classe d'âge (0-35 ans, 36-64 ans, 65-75 ans, 76 ans ou plus) et de même CS à un chiffre.

C. ÉVOLUTION DE LA STRUCTURE PAR CATÉGORIE SOCIALE DE LA POPULATION

La structure de la population par catégorie sociale a évolué au fil du temps. Les cadres et les professions intermédiaires sont plus nombreux en 2022 qu'en 1975, pour les femmes comme pour les hommes, mais la hausse est plus importante chez les femmes. En 1975, 1,7 % des femmes âgées de 35 ans ou plus étaient (ou avaient été) cadres et 6,5 % exerçaient (ou avaient exercé) une profession intermédiaire. En 2020, 10,3 % étaient cadres et 23,8 % exerçaient une profession intermédiaire. Chez les hommes, seuls 9,3 % étaient cadres en 1975, contre 18,6 % en 2020 et ces pourcentages sont respectivement de 14,2 % et 21,6 % pour les professions intermédiaires.

À l'inverse, les hommes ouvriers ainsi que les agriculteurs ou agricultrices sont nettement moins nombreux en 2022 qu'en 1975. Chez les femmes, la proportion d'inactives non retraitées a beaucoup baissé : 44,5 % des femmes de 35 ans ou plus étaient inactives non retraitées en 1975 et 9,8 % seulement en 2020.

Répartition de la population de 35 ans ou plus par catégorie sociale (en %)

Catégorie sociale	Femme						Homme					
	1975	1982	1990	1999	2013	2020	1975	1982	1990	1999	2013	2020
Agriculteur	9,2	9,7	7,3	4,9	3,2	2,3	14,4	12,1	9,1	6,6	4,4	3,6
Artisan, commerçant	5,8	6,3	5,8	4,9	4,7	4,3	10,7	10,8	10,4	10,0	9,9	9,9
Cadre	1,7	2,6	4,3	5,4	8,5	10,3	9,3	11,6	14,1	14,5	17,3	18,6
Prof. intermédiaire	6,5	8,9	11,4	15,0	21,1	23,8	14,2	16,8	17,6	19,4	21,0	21,6
Employé	20,9	24,7	31,0	36,2	40,0	39,8	11,5	12,1	11,9	12,5	11,8	12,2
Ouvrier	11,4	11,1	12,3	12,4	10,7	9,7	36,9	33,0	34,0	33,3	31,2	29,5
Inactif non retraité	44,5	36,7	27,9	21,2	11,8	9,8	3,0	3,6	2,9	3,8	4,4	4,6
Ensemble	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Champ : personnes âgées de 35 ans ou plus nées en France.

¹⁰ La catégorie sociale à deux chiffres est renseignée et imputée comme par le passé à partir de 2019.

3) LE DIPLÔME

A. CONCEPT

Le diplôme est celui déclaré lors du recensement. Seul le plus élevé des diplômes est retenu. Une seule question est formulée dans l'EAR à partir de 2015 : « Quel(s) diplôme(s) avez-vous ? ».

Les modalités sont les suivantes :

- 01 Pas de scolarité ou arrêt avant la fin du primaire
- 02 Aucun diplôme et scolarité interrompue à la fin du primaire ou avant la fin du collège
- 03 Aucun diplôme et scolarité jusqu'à la fin du collège ou au-delà
- 11 CEP (certificat d'études primaires)
- 12 BEPC, brevet élémentaire, brevet des collèges
- 13 CAP, BEP ou diplôme de niveau équivalent
- 14 Baccalauréat général ou technologique, brevet supérieur capacité en droit, DAEU, ESEU
- 15 Baccalauréat professionnel, brevet professionnel, de technicien ou d'enseignement, diplôme équivalent
- 16 BTS, DUT, DEUG, DEUST, diplôme de la santé ou du social de niveau bac+2 ou équivalent
- 17 Licence, licence pro, maîtrise, diplôme équivalent de niveau bac+3 ou bac+4
- 18 Master, DEA, DESS, diplôme grande école niveau bac+5, doctorat de santé
- 19 Doctorat de recherche (hors santé)

La population est répartie en cinq catégories :

- Sans diplôme (modalités 01, 02, 03)
- Brevet, CEP (modalités 11, 12)
- CAP, BEP (modalité 13)
- Baccalauréat (modalités 14, 15)
- Supérieur au baccalauréat (modalités 16, 17, 18, 19)

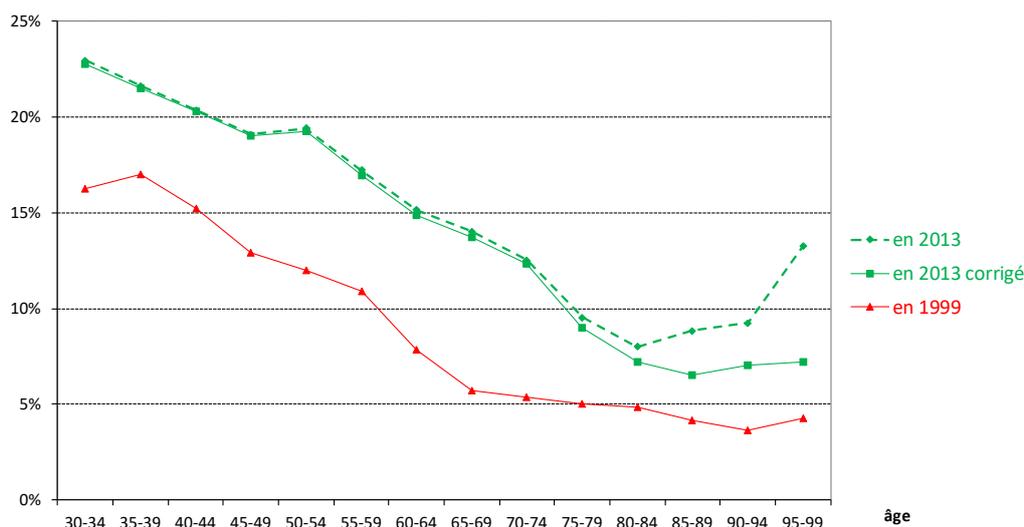
B. REDRESSEMENT

Le diplôme a été corrigé de 2004 à 2014¹¹, de façon à homogénéiser les données sur longue période. Au-delà de 80 ans, la proportion de femmes dont le diplôme le plus élevé obtenu est le baccalauréat augmente dans les EAR, ce qui n'est pas le cas dans le RP 1999. Ceci s'explique par des traitements différents dans le codage. Dans le cas où le diplôme est manquant, celui-ci est imputé. Dans les EAR de 2004 à 2014, l'imputation s'effectue par hotdeck parmi les donneurs du même lot de saisie, de même nationalité (français, étranger) et de même tranche d'âge (25 à 44 ans, 45-64 ans, 65 ans ou plus). Le sexe n'est pas pris en compte et la tranche d'âge supérieure est trop large. On impute donc trop souvent le diplôme d'un homme de 65 ans à une femme de 80 ans par exemple. Or, les femmes de 80 ans sont moins diplômées que les hommes de 65 ans.

¹¹ À partir de 2015, les imputations du diplôme dans les EAR redeviennent conformes à ceux effectués dans le passé.

La méthode d'imputation a donc été modifiée. On a récupéré l'information si elle est disponible dans les autres EAR. Sinon, on a imputé le diplôme par hotdeck parmi les donneurs de même sexe et de même âge (le même âge jusqu'à 96 ans, puis 97 ans ou plus). Cette correction permet d'avoir des classements plus homogènes entre 1999 et 2004-2014. Cela a pour effet de corriger l'espérance de vie des femmes diplômées qui était sous-évaluée avant correction (par exemple, au-delà de 80 ans, les quotients de mortalité des femmes ayant un diplôme supérieur au baccalauréat étaient plus élevés que ceux de l'ensemble des femmes du même âge).

Proportion de femmes diplômées du baccalauréat selon l'âge



Champ : femmes âgées de 30 ans ou plus nées en France.

C. ÉVOLUTION DE LA STRUCTURE PAR DIPLÔME DE LA POPULATION

Sur la période 1990-2020, les femmes comme les hommes sont de plus en plus diplômés : les personnes sans diplôme ou titulaires du CEP ou du brevet sont de moins en moins nombreuses au profit des titulaires du CAP ou BEP et du baccalauréat ou d'un diplôme du supérieur. La hausse est particulièrement marquée pour les femmes diplômées du supérieur qui sont 4 fois plus nombreuses en 2020 qu'en 1990 (31,2 % en 2020 contre 7,5 % en 1990). Les hommes diplômés du supérieur sont près de trois fois plus nombreux (30,1 % contre 11,4 %).

Répartition de la population de 35 ans ou plus par diplôme (%)

Diplôme	Femme				Homme			
	1990	1999	2013	2020	1990	1999	2013	2020
Sans diplôme	30,5	20,1	16,4	11,8	25,2	17,8	14,8	11,0
Brevet, CEP	40,2	36,6	23,3	17,5	31,2	26,3	14,9	11,3
CAP, BEP	13,0	19,9	20,6	22,9	21,8	30,5	31,4	32,4
Baccalauréat	8,8	10,3	15,6	16,6	10,4	10,2	14,7	15,2
Supérieur au baccalauréat	7,5	13,1	24,1	31,2	11,4	15,2	24,3	30,1
Ensemble	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Champ : personnes âgées de 35 ans ou plus nées en France.

PARTIE 2 : LA MÉTHODE

Les tables de mortalité sont estimées pour chaque catégorie sociale, sur une période d'observation de l'état vital de quelques années. Les quotients de mortalité sont calculés selon :

- la CS 1975 sur la période 1976-1984,
- la CS 1982 sur la période 1983-1991,
- la CS et le diplôme 1990 sur la période 1991-1999,
- la CS et le diplôme 1999 sur la période 2000-2008,
- la CS et le diplôme de 2004 à 2012 sur la période 2009-2013,
- la CS et le diplôme de 2010 à 2018 sur la période 2017-2019,
- la CS et le diplôme de 2013 à 2020 sur la période 2020-2022.

L'analyse n'intègre pas les décès survenus l'année d'observation de la catégorie sociale pour deux raisons : pour éviter des effets à très court terme dus au changement de catégorie sociale suite à un problème de santé (passage d'ouvrier à inactif par exemple si la santé de l'ouvrier ne lui permet plus de travailler par exemple) ; et pour tenir compte du fait que pendant l'année du recensement, les décès du début de l'année (avant le recensement) ne sont, par construction, pas comptabilisés, ce qui induit une sous-estimation mécanique de la mortalité sur cette année incomplète.

Pour l'analyse de la mortalité par diplôme, les deux premières périodes sont exclues. En effet, les diplômes ne sont connus que pour le quart de la population lors des recensements de 1975 et 1982. Les effectifs ne sont pas suffisants pour assurer la robustesse des quotients de mortalité par âge et par diplôme.

Les quotients de mortalité sont calculés sur le champ des personnes résidant et nées en France. Entre les recensements et la période d'observation des décès, une partie des personnes nées à l'étranger migrent. Or, une partie des décès ayant eu lieu à l'étranger ne figurent pas dans l'EDP¹². Les quotients de mortalité des personnes nées à l'étranger sont donc sous-estimés dans l'EDP.

Afin d'être représentatif de l'espérance de vie de l'ensemble des personnes résidant en France, un calage sur les tables de mortalité de référence, issues du bilan démographique, est effectué (voir la partie « calage sur les tables de mortalité de référence »). Les tables de mortalité de référence sont calculées à partir de tous les décès ayant eu lieu en France rapportés à la population résidant en France.

¹² L'EDP ne porte que sur les décès qui ont été intégrés dans la base des répertoires des personnes physiques (BRPP). Selon les cas, les décès concernant des ressortissants français à l'étranger peuvent être déclarés et enregistrés directement à l'ambassade. Sinon, les événements d'état civil sont enregistrés selon les procédures du pays concerné. Dans ce cas, les actes d'état civil étrangers peuvent ensuite être transcrits dans les registres d'état civil français. Certains actes donnent également lieu à une mention en marge dans la mairie du lieu de naissance si la personne concernée est née en France, mention en marge qui peut alors donner lieu à l'établissement d'un bulletin de mention en marge transmis à l'Insee.

1) QUOTIENTS DE MORTALITÉ

Les quotients de mortalité sont estimés pour chaque catégorie sociale et diplôme, par sexe et âge (en différence de millésimes).

Pour les quatre périodes les plus anciennes (1976-1984, 1983-1991, 1991-1999, 2000-2008), on tient compte, pour des raisons d'effectifs, de l'ensemble des décès observés sur une période de 9 années suivant le recensement. Concrètement, on retient les individus ayant répondu au recensement, puis on observe leur statut vital sur les 9 années suivantes.

Par exemple, pour la période 2000-2008, le quotient de mortalité à 39 ans est le rapport entre la somme du nombre de personnes recensées en 1999 et décédées à l'âge de 39 ans entre 2000 et 2008, et la somme du nombre de personnes recensées en 1999 et vivantes à l'âge de 39 ans au 1^{er} janvier de chaque année :

$$Q_{2000-2008}^{39} = \frac{Décès_{2000}^{39} + Décès_{2001}^{39} + \dots + Décès_{2008}^{39}}{Pop_{01/01/2000}^{39} + Pop_{01/01/2001}^{39} + \dots + Pop_{01/01/2008}^{39}}$$

Une personne recensée en 1999 et vivante en 2008 intervient ainsi dans le calcul de 9 quotients. Si elle a 30 ans en 1999 (c'est-à-dire qu'elle est née en 1969), elle intervient dans le quotient de mortalité à 31 ans, mais aussi dans les quotients de mortalité de 32 ans à 39 ans.

		année									
RP 1999		2000	2001	2002	2003	2004	2005	2006	2007	2008	
âge	30	31	32	33	34	35	36	37	38	39	
	31	32	33	34	35	36	37	38	39	40	
	32	33	34	35	36	37	38	39	40	41	
	33	34	35	36	37	38	39	40	41	42	
	34	35	36	37	38	39	40	41	42	43	
	35	36	37	38	39	40	41	42	43	44	
	36	37	38	39	40	41	42	43	44	45	
	37	38	39	40	41	42	43	44	45	46	
	38	39	40	41	42	43	44	45	46	47	

Pour les trois périodes les plus récentes (2009-2013, 2017-2019 et 2020-2022), les quotients de mortalité sont calculés en cumulant 8 ou 9 enquêtes annuelles de recensement. La méthode de recensement a changé, passant d'un recensement exhaustif jusqu'en 1999 à des enquêtes annuelles de recensement (EAR) à partir de 2004.

Par exemple, pour le calcul du quotient de mortalité à 39 ans pour la période 2020-2022, on cumule les décès survenus à l'âge de 39 ans pour les personnes recensées :

- en 2013 ayant eu 39 ans au cours de l'une des 3 années de 2020 à 2022 ;
- en 2014 ayant eu 39 ans au cours de l'une des 3 années de 2020 à 2022 ;
- en 2015 ayant eu 39 ans au cours de l'une des 3 années de 2020 à 2022 ;

- en 2016 ayant eu 39 ans au cours de l'une des 3 années de 2020 à 2022 ;
- en 2017 ayant eu 39 ans au cours de l'une des 3 années de 2020 à 2022 ;
- en 2018 ayant eu 39 ans au cours de l'une des 3 années de 2020 à 2022 ;
- en 2019 ayant eu 39 ans au cours de l'une des 3 années de 2020 à 2022 ;
- en 2020 ayant eu 39 ans au cours de l'une des 2 années de 2021 à 2022.

On divise cet effectif de décès par les effectifs correspondant de personnes vivantes au 1^{er} janvier de chaque année. Les personnes recensées à différentes EAR interviennent plusieurs fois. Chaque enquête annuelle de recensement est représentative de la population totale.

		année									
		ear 2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
âge	30								37	38	39
	31								38	39	40
	32								39	40	41
	33								40	41	42
		ear 2014	2015	2016	2017	2018	2019	2020	2021	2022	
âge	31								37	38	39
	32								38	39	40
	33								39	40	41
	34								40	41	42
		ear 2015	2016	2017	2018	2019	2020	2021	2022		
âge	32								37	38	39
	33								38	39	40
	34								39	40	41
	35								40	41	42
		ear 2016	2017	2018	2019	2020	2021	2022			
âge	33								37	38	39
	34								38	39	40
	35								39	40	41
	36								40	41	42
		ear 2017	2018	2019	2020	2021	2022				
âge	34								37	38	39
	35								38	39	40
	36								39	40	41
	37								40	41	42
		ear 2018	2019	2020	2021	2022					
âge	35								37	38	39
	36								38	39	40
	37								39	40	41
	38								40	41	42
		ear 2019	2020	2021	2022						
âge	36								37	38	39
	37								38	39	40
	38								39	40	41
	39								40	41	42
		ear 2020	2021	2022							
âge	37								37	38	39
	38								38	39	40
	39								39	40	41
	40								40	41	42

2) L'ESPÉRANCE DE VIE

L'espérance de vie à un âge x est le nombre moyen d'années restant à vivre qu'aurait une génération fictive soumise aux conditions de mortalité d'une année donnée, au-delà de cet âge x . Si l'espérance de vie à la naissance des femmes est de 86 ans en 2024, cela ne signifie pas que les filles nées en 2024 vivront en moyenne jusqu'à cet âge, puisqu'elles seront très probablement soumises au cours de leur vie à des conditions de mortalité différentes de celles de 2024. Ainsi, l'espérance de vie n'est pas une prévision, mais un indicateur synthétique du niveau de la mortalité d'une année donnée. Elle permet donc de comparer les conditions de mortalité à différentes dates ou pour différents groupes (pays, niveau d'études).

De la même façon, l'espérance de vie par CS ou diplôme n'est pas une prévision de la durée de vie moyenne qu'aurait une personne selon sa CS ou son diplôme à 35 ans. Par exemple, l'espérance de vie à 35 ans des ouvriers correspond au nombre moyen d'années restant à vivre qu'aurait une génération soumise, au-delà de 35 ans, aux conditions de mortalité des ouvriers d'une année donnée. Cette génération est fictive puisque les conditions de mortalité varient au fil des années. De plus, toutes les personnes conserveraient la même CS, or des changements de CS peuvent intervenir en cours de vie professionnelle : un ouvrier de 35 ans ne le restera pas nécessairement jusqu'à l'âge de la retraite. Néanmoins, l'espérance de vie par CS est un indicateur synthétique du niveau de la mortalité actuelle des cadres ou des ouvriers.

Les espérances de vie et les quotients de mortalité sont disponibles à partir de l'âge de 30 ans dans l'Insee Résultats « Tables de mortalité par catégorie sociale et par diplôme jusqu'en 2020-2022 ». Dans l'Insee Première n° 2005 « Les écarts d'espérance de vie entre cadres et ouvriers : 5 ans chez les hommes, 3 ans chez les femmes ? » et dans ce document de travail, on analyse l'espérance de vie à 35 ans par soucis d'homogénéité avec les publications précédentes.

L'espérance de vie se calcule de la manière suivante :

Age	Quotient de mortalité	Survie	Espérance de vie
30	q_{30}	$S_{30} = 100\ 000$	$\frac{\sum_{31}^{105} S_x}{S_{30}}$
31	q_{31}	$S_{31} = 100\ 000 * (1 - q_{30})$	$\frac{\sum_{32}^{105} S_x}{S_{30}}$
...			
X	q_x	$S_x = S_{x-1} * (1 - q_{x-1})$	$\frac{\sum_{x+1}^{105} S_x}{S_{30}}$
...			
105 ans			

L'espérance de vie à 30 ans s'écrit :

$$E_{30} = \frac{\sum_{31}^{105} S_x}{S_{30}} = \frac{0 \text{ an} * (S_{30} - S_{31}) + 1 \text{ an} * (S_{31} - S_{32}) + 2 \text{ ans} * (S_{32} - S_{33}) + \dots}{S_{30}}$$

3) LISSAGE ET ESTIMATION DES QUOTIENTS DE MORTALITÉ PAR LA MÉTHODE DE BRASS POUR LES PÉRIODES ANCIENNES (DE 1976-1984 À 2009-2013)

De 1976-1984 à 1991-1999, les quotients de mortalité par CS et diplôme sont assez fiables sur la tranche d'âge 40-80 ans, pour laquelle nous disposons en général d'effectifs suffisants de population et de décès. En revanche, après l'âge de 80 ans, certains décès semblent manquants : les quotients de mortalité diminuent avec l'âge au lieu d'augmenter. De plus, avant 40 ans, les décès ne sont pas en nombre suffisant (faible mortalité) et les quotients de mortalité ne sont donc pas estimés précisément. Toutefois, les quotients de mortalité aux jeunes âges ont assez peu d'incidence sur l'espérance de vie, car ces âges contribuent moins à son niveau que les âges avancés.

La méthode de Brass permet à la fois de lisser les quotients aux âges où les effectifs sont suffisants (la plupart du temps entre 40-80 ans) et de prolonger les quotients aux âges où les estimations sont plus incertaines (notamment aux âges élevés où la mortalité est sous-estimée).

En 2000-2008 et 2009-2013, l'estimation des quotients est de meilleure qualité : les effectifs sont plus nombreux et les décès semblent mieux mesurés (les quotients ne diminuent pas avec l'âge). La méthode de Brass est également appliquée, mais les quotients sont estimés sur une tranche d'âge plus grande par exemple 45-95 ans, afin de s'approcher au plus près des quotients de mortalité bruts aux âges élevés.

A. PRINCIPE GÉNÉRAL

La méthode de Brass¹³ consiste à déduire des tables de mortalité les unes des autres par simple transformation. Brass suppose (par observation empirique) l'existence d'une relation affine entre les logits des quotients cumulés de deux tables valides pour une plage d'âge donnée (ou, ce qui revient au même, une relation entre les logits des fonctions de survie). Le logit du quotient cumulé à l'âge x est défini par¹⁴ :

$$Y(x) = \log_e = 1/2 * \log \left(\frac{Q(x)}{1 - Q(x)} \right) = 1/2 * \log \left(\frac{S_{31} - S_x}{S_x} \right)$$

Où $Q(x) = 1 - \frac{S_x}{S_{31}}$ est la probabilité de décéder entre 31 ans et l'âge x

et S_x est le nombre de survivants à l'âge x

¹³ Leridon H., Toulemon L. (1997) « Démographie, approche statistique et dynamique des populations », Economica, page 222.

¹⁴ L'inclusion d'un facteur 1/2 est historique et ne change rien aux estimations.

Brass admet l'existence d'une relation affine entre les Logits des quotients de deux populations en fonction de l'âge :

$$Y_2(x) = a + b Y_1(x)$$

a : indicateur de niveau

b : indicateur de pente

L'interprétation de cette formule est la suivante :

Si $a > 0$: le niveau de mortalité dans la population 2 est supérieur à celui dans la population 1 (en d'autres termes, l'âge médian au décès est plus bas dans la population 2 que dans la population 1)

Si $b > 1$: la mortalité augmente avec l'âge plus vite dans la population 2 que dans la population 1

Cette méthode est couramment utilisée pour lisser des tables de quotients estimés. Dans le cas présent, la table de mortalité de la population 2 est la table de mortalité estimée pour une catégorie sociale et un sexe sur une période donnée. La table de mortalité de la population 1 prise comme référence est la table de mortalité moyenne sur la période, pour l'ensemble des hommes ou l'ensemble des femmes (toutes CS confondues). Elle est issue du bilan démographique établi chaque année par l'Insee à partir de l'état civil et des estimations de population. Les coefficients a et b sont estimés par le biais d'une régression en prenant en compte uniquement les âges pour lesquels les quotients par CS ou par diplôme sont correctement estimés (de 40 à 80 ans par exemple). La régression est pondérée par le nombre de présents (survivants) à chaque âge pour limiter l'influence sur l'estimation de points atypiques qui concerneraient peu d'individus. Les quotients lissés se déduisent ensuite des Logits déduits de la régression, pour tous les âges pour lesquels la relation linéaire est supposée être valide, que les quotients par CS ou diplôme aient pu être correctement estimés ou non.

$$\hat{Y}_2(x) = \hat{a} + \hat{b} Y_1(x)$$

B. AMÉNAGEMENT DE LA MÉTHODE DE BRASS

Les représentations graphiques des Logits des quotients cumulés estimés (avant lissage), pour un groupe donné, en fonction des Logits des quotients cumulés de la table de référence, permettent de valider l'utilisation de la méthode de Brass pour le lissage de certains groupes sociaux pour lesquels on peut admettre une relation linéaire entre les Logits par âge. Mais, pour d'autres groupes, supposer une telle relation linéaire n'est pas valide, la relation n'étant pas la même à chaque âge. Par exemple, pour les hommes cadres en 1999, la pente de la courbe est légèrement plus élevée avant 85 ans (1,073) qu'après cet âge (0,896).

Pour résoudre ce problème, on part du principe de la méthode de Brass, mais en autorisant *a priori* des changements de pente autour d'un certain âge s (pentes différentes avant et après cet âge). Si les pentes estimées avant et après cet âge sont proches, cela revient à utiliser la méthode de Brass exposée précédemment. Si au contraire les pentes diffèrent, la méthode retenue ici permet de prendre en compte des relations différentes entre les Logits des quotients de la table de référence et de la table qu'on cherche à lisser selon deux groupes d'âge et d'améliorer ainsi le lissage.

La nouvelle relation à estimer est alors la suivante :

$$Y_2(x) = a + bY_1(x) \text{ si } x < s$$

$$Y_2(x) = a_2 + b_2Y_1(x) \text{ si } x \geq s$$

et $a + bY_1(s) = a_2 + b_2Y_1(s)$

ce qui conduit à estimer la relation linéaire suivante :

$$Y_2(x) = a + b \left[Y_1(x) + (Y_1(s) - Y_1(x)) I_{x >= s} \right] + b_2 \left[(Y_1(x) - Y_1(s)) I_{x >= s} \right]$$

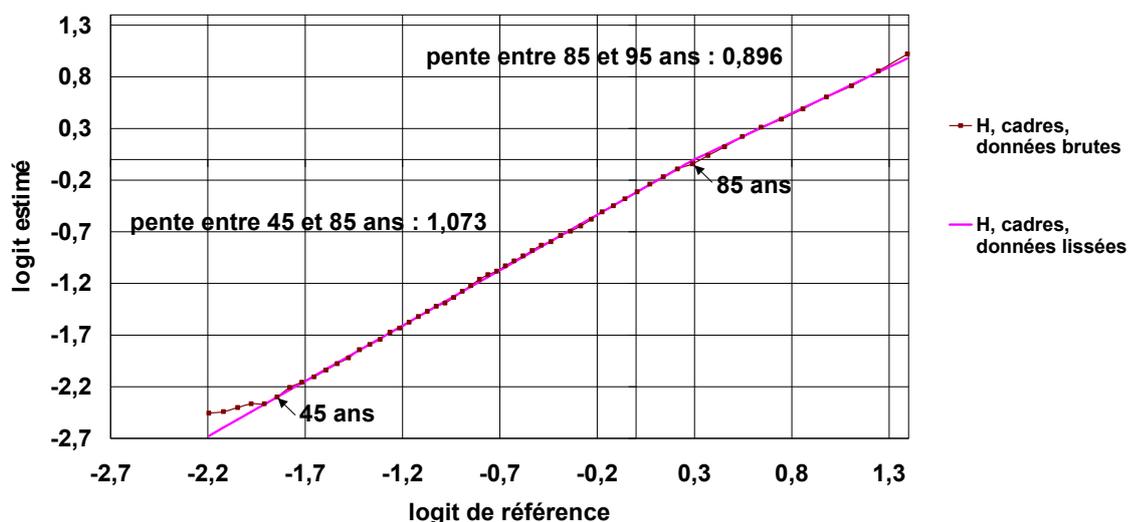
Les âges de rupture de pente sont choisis de façon *ad hoc*, afin que les quotients lissés soient les plus proches possibles des quotients bruts. En général, on a retenu des estimations entre 40 et 80 ans avec une rupture de pentes à 55 ans pour les CS pour les trois périodes les plus anciennes, et pour les diplômés en 1991-1999. Pour les deux périodes les plus récentes de cette partie (2000-2008 et 2009-2013), on a retenu en général des estimations entre 45 et 95 ans avec des âges de rupture différents selon les CS et les diplômés (65, 70, 80 ou 85 ans). Comme pour le principe général, on estime les coefficients a , b et b_2 à l'aide de régressions pondérées par le nombre de présents.

Pour les trois périodes les plus anciennes, on suppose que la relation estimée est également valide entre 30 et 40 ans et pour des âges supérieurs à 80 ans. On déduit ensuite les quotients estimés par âge atteint dans l'année, de 30 ans à 100 ans. La rupture de pentes entraîne une petite non-linéarité des variations des quotients autour de l'âge de rupture retenu. On lisse donc les quotients estimés sur plusieurs âges (entre 50 et 60 ans par exemple pour une rupture à 55 ans) par interpolation linéaire sur les logarithmes de ces quotients. Le prolongement aux « âges jeunes » (avant 40 ans en général) conduit parfois à des quotients de mortalité décroissants avec l'âge (quotient à 32 ans légèrement plus petit qu'à 31 ans). Dans ce cas, on suppose que le rapport des quotients entre l'âge ($a+1$) et l'âge (a) est le même avant et après 40 ans.

Ci-après deux exemples d'application de la méthode de Brass avec et sans rupture de pente :

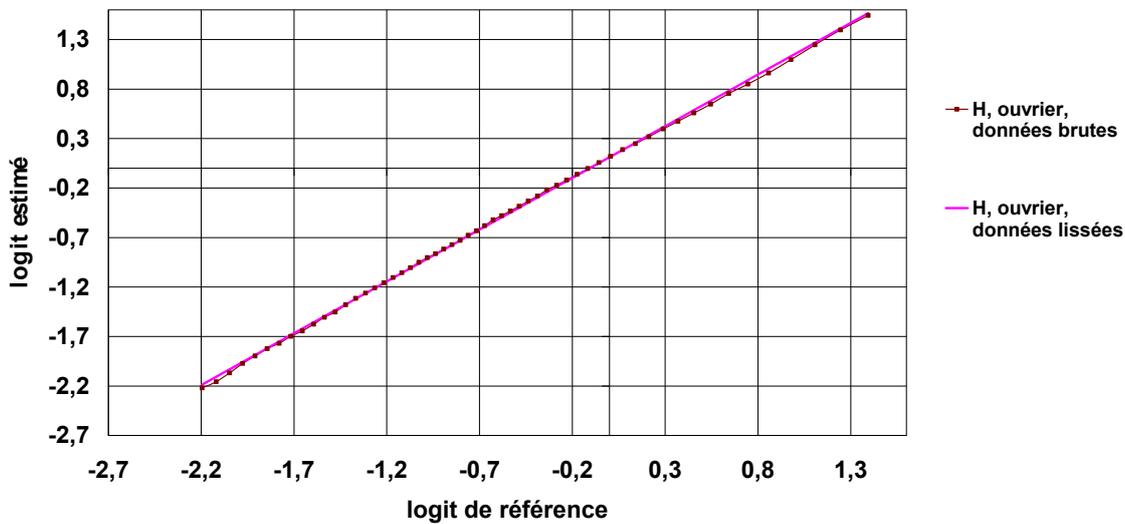
Méthode de Brass avec rupture de pente

Logits estimés en fonction des logits de références entre 40 et 95 ans pour les hommes cadres en 1999



Méthode de Brass sans rupture de pente

Logits estimés en fonction des logits de références
entre 40 et 95 ans pour les hommes ouvriers en 1999

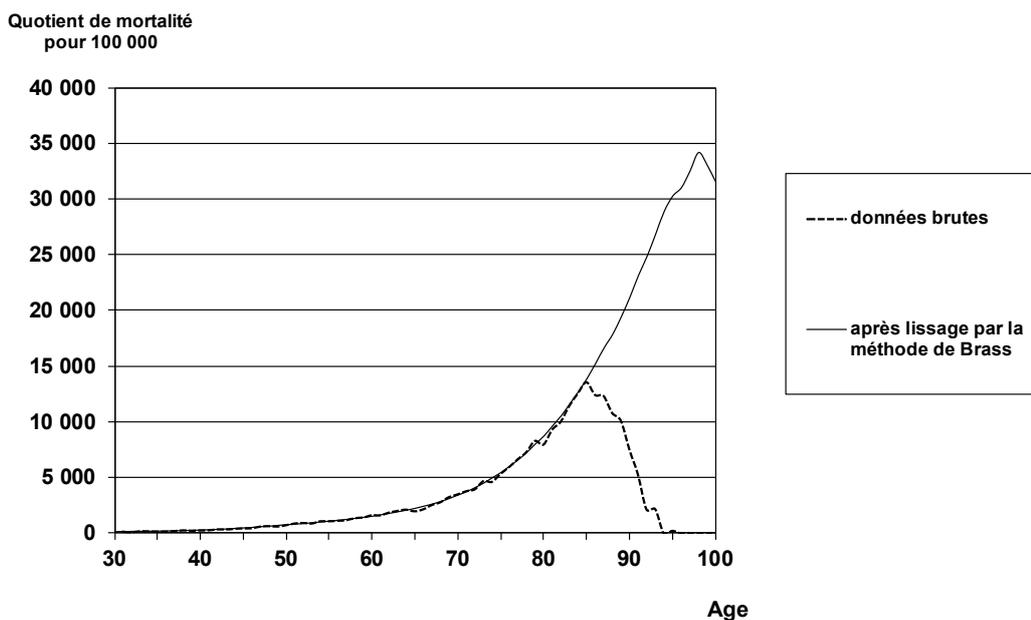


Note : les quotients de référence sont ceux des femmes ou des hommes dans le bilan démographique

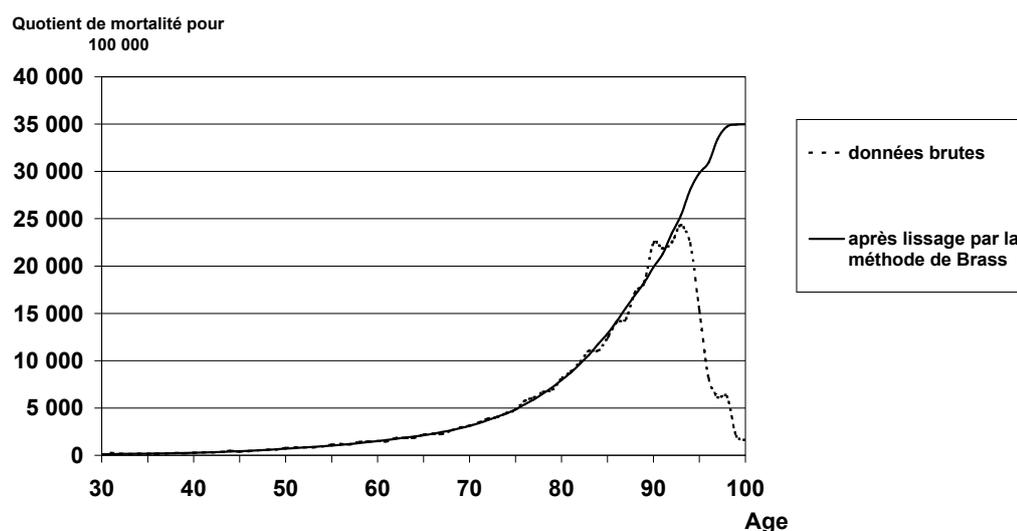
C. EXEMPLE DE LISSAGE PAR LA MÉTHODE DE BRASS

Les figures suivantes illustrent l'application de la méthode de Brass à l'ensemble des hommes. Pour les périodes anciennes, la méthode permet de corriger la sous-estimation des décès aux âges élevés et de lisser les données. Pour la période 2009-2013, la sous-estimation des décès ne concerne que les personnes âgées de 95 ans ou plus et la méthode permet surtout de lisser les quotients.

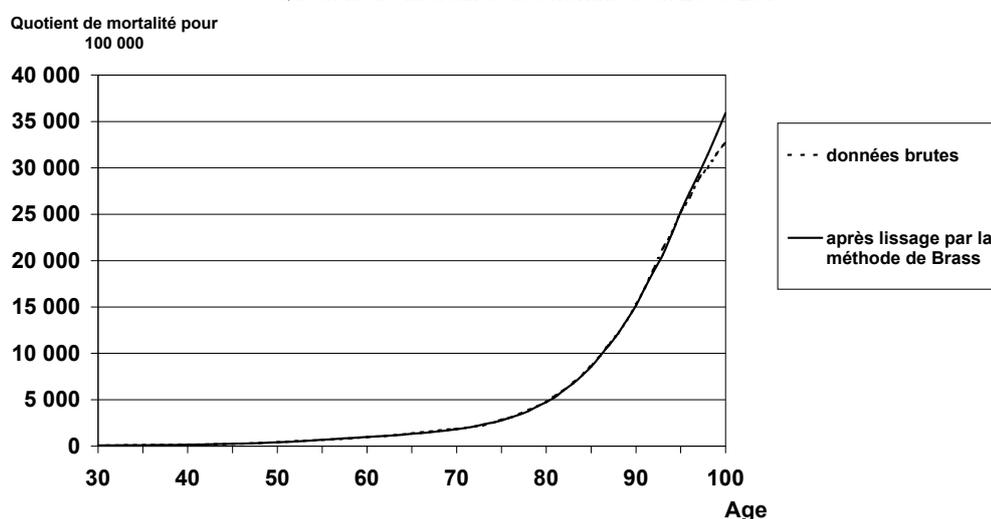
Quotient de mortalité des hommes en 1976-1984



Quotient de mortalité des hommes en 1983-1991



Quotient de mortalité des hommes en 2009-2013



D. EFFET DE LA MÉTHODE DE BRASS SUR L'ESTIMATION DES ESPÉRANCES DE VIE

Espérance de vie à 35 ans selon le sexe avant et après application de la méthode de Brass

	1976-1984	1983-1991	1991-1999	2000-2008	2009-2013
Femmes					
Estimée, brute	48,0	47,2	48,6	49,8	50,8
Estimée, après méthode de Brass	46,2	47,0	48,5	49,8	50,7
Écart	-1,8	-0,2	-0,1	-0,0	-0,1
Hommes					
Estimée, brute	39,9	39,7	41,3	43,3	44,7
Estimée, après méthode de Brass	39,0	39,6	41,1	43,2	44,7
Écart	-0,9	-0,1	-0,2	-0,1	0,0

En 1976-1984, les espérances de vie à 35 ans après lissage sont éloignées de celles avant lissage (-1,8 an pour les femmes et -0,9 an pour les hommes). Ceci est dû à la baisse des quotients de mortalité au-delà de 85 ans en raison de décès manquants. En revanche, dès 1983-1991, les espérances de vie avant et après lissage sont relativement proches : les quotients de mortalité sont

sous-estimés à partir de 94 ans, mais la correction a peu de conséquence sur le niveau de l'espérance de vie étant donné le peu de survivants à cet âge.

4) LISSAGE ET ESTIMATION DES QUOTIENTS DE MORTALITÉ PAR LA MÉTHODE DES SPLINES POUR LES PÉRIODES RÉCENTES (2017-2019 ET 2020-2022)

A. MÉTHODE

En 2017-2019 et 2020-2022, l'espérance de vie calculée à partir des données brutes de l'EDP est très proche de celle du bilan démographique : il n'y a plus la sous-estimation des décès aux grands âges, qui motivait l'emploi de la méthode de Brass pour les périodes plus anciennes. Toutefois, les quotients de mortalité bruts par CS ou diplôme sont chahutés ; ils varient assez nettement d'un âge à l'autre. Les quotients de mortalité ont donc été lissés afin d'obtenir des estimations de quotients suffisamment robustes pour être diffusés. Pour le lissage, on utilise la méthode des splines. Une courbe spline est une fonction continue par morceau de polynômes.

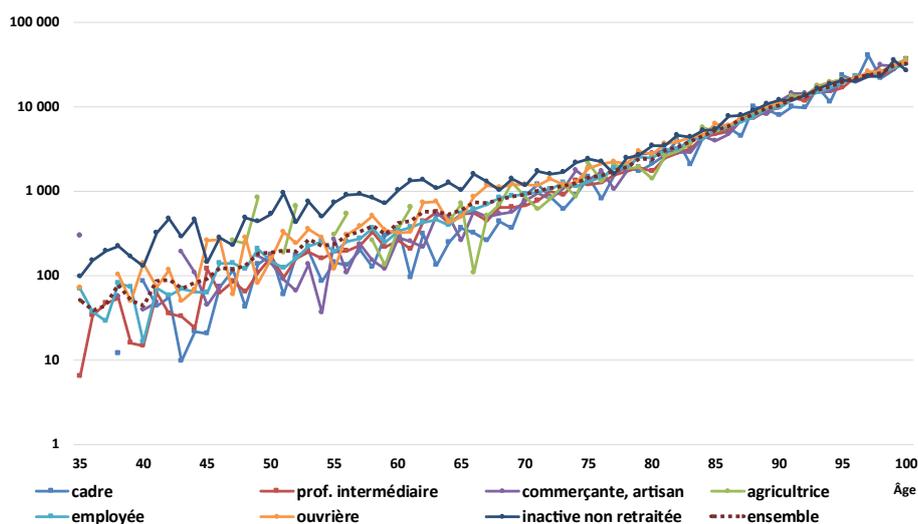
Le package mis en oeuvre s'intitule MortalitySmooth [Camarda, C. G. (2012). MortalitySmooth: An R Package for Smoothing Poisson Counts with P-Splines. Journal of Statistical Software. 50, 1-24. [http://www.jstatsoft.org/v50/i01/.](http://www.jstatsoft.org/v50/i01/)].

La méthode 3 a été retenue. Elle permet de choisir manuellement les 4 paramètres :

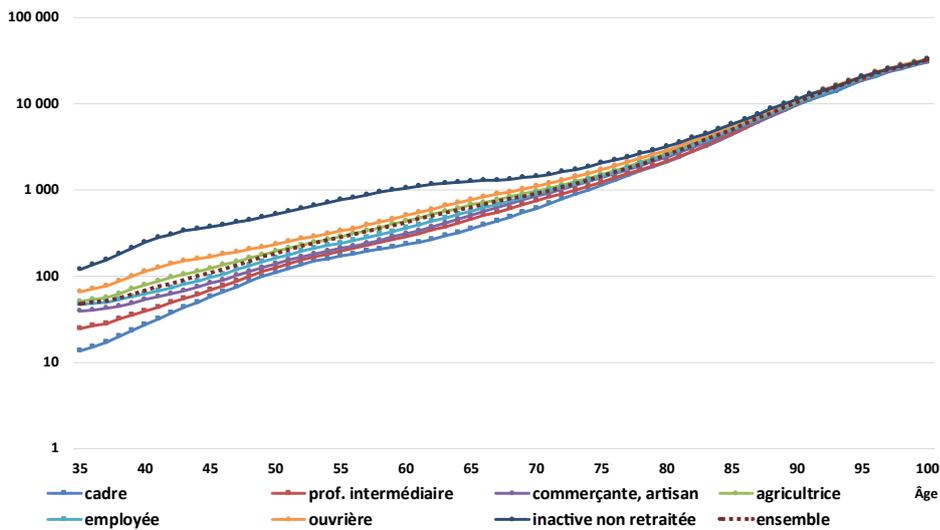
- 2 nœuds pour chaque axe (âge et CS ou diplôme)
- 2 paramètres de lissage.

Les paramètres ont été choisis de manière à ce que l'espérance de vie lissée reste la plus proche possible de l'espérance de vie brute.

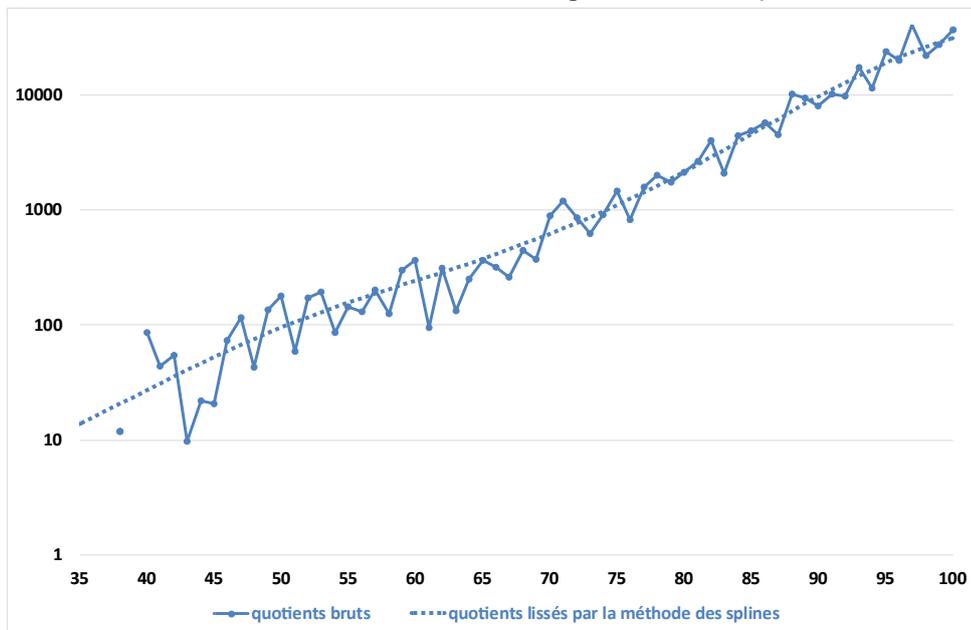
Quotient de mortalité des femmes par catégorie sociale selon l'âge en 2020-2022 (données brutes)



Quotient de mortalité des femmes par catégorie sociale selon l'âge en 2020-2022 (données lissées)

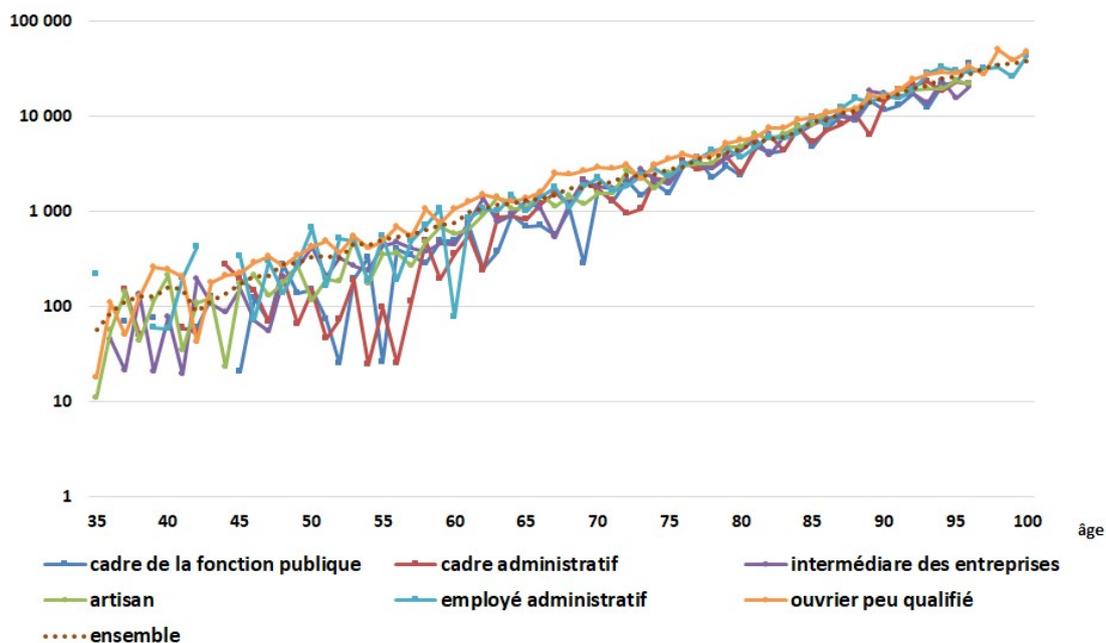


Quotient de mortalité des femmes cadres selon l'âge en 2020-2022 (données brutes et lissées)

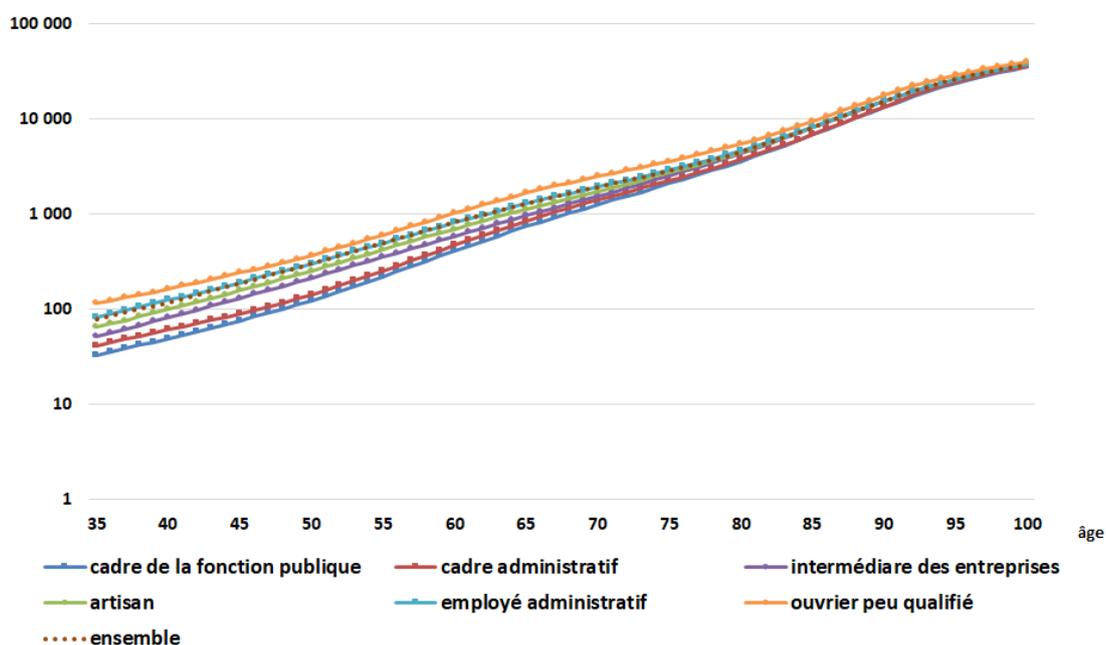


Le lissage des quotients de mortalité bruts par catégorie sociale détaillée a été effectué en ajoutant les quotients de mortalité lissés par catégorie sociale (à un chiffre) dans les données d'entrée du package MortalitySmooth. Cela permet d'éviter en partie les croisements des courbes de quotients de mortalité : les quotients de mortalité lissés d'une catégorie sociale restent la plupart du temps supérieurs ou inférieurs à ceux d'une autre catégorie pour tous les âges.

Quotient de mortalité des hommes pour certaines CS détaillées selon l'âge en 2020-2022 (données brutes)



Quotient de mortalité des hommes pour certaines CS détaillées selon l'âge en 2020-2022 (données lissées)



B. EFFET DU LISSAGE PAR LA MÉTHODE DES SPLINES SUR L'ESTIMATION DES ESPÉRANCES DE VIE

Les écarts d'espérances de vie avant et après lissage sont très proches par catégorie sociale (+/- 0,1 point), sauf pour les agriculteurs (-0,9 point pour les femmes et -0,4 point pour les hommes). L'espérance de vie brute des agriculteurs est sous-estimée en raison de la faiblesse des effectifs : il n'y a aucun décès à certains âges dans l'échantillon. Le lissage par la méthode des splines permet de corriger cette sous-estimation.

Espérance de vie à 35 ans par sexe et catégorie sociale avant et après lissage par la méthode des splines en 2020-2022

	Cadre	Profession intermédiaire	Artisan, commerçant	Employé	Agriculteur	Ouvrier	Inactif non retraité	Ensemble
Femme								
Estimée, brute	52,9	52,3	51,9	51,3	51,5	49,8	46,0	50,8
Estimée, après lissage par spline	53,0	52,3	51,8	51,3	50,6	49,7	45,9	50,9
écart	0,1	0,0	-0,1	0,0	-0,9	-0,1	-0,1	0,1
Homme								
Estimée, brute	49,0	47,4	46,7	45,4	47,7	43,9	34,8	45,5
Estimée, après lissage par spline	49,0	47,5	46,6	45,3	47,3	43,9	34,9	45,5
écart	0,0	0,1	-0,1	-0,1	-0,4	0,0	0,1	0,0

Les écarts d'espérances de vie avant et après lissage sont très proches par diplôme (de -0,1 à 0,2 point).

Espérance de vie à 35 ans par sexe et diplôme avant et après lissage par la méthode des splines en 2020-2022

	Diplôme supérieur au baccalauréat	Baccalauréat	CAP, BEP	Brevet, CEP	sans diplôme	Ensemble
Femme						
Estimée, brute	52,6	51,7	50,9	50,5	47,2	50,8
Estimée, après lissage par spline	52,6	51,7	51,1	50,4	47,3	50,8
écart	0,0	0,0	0,2	-0,1	0,1	0,0
Homme						
Estimée, brute	48,5	46,7	45,2	43,9	41,0	45,5
Estimée, après lissage par spline	48,6	46,6	45,2	43,9	41,0	45,5
écart	0,1	-0,1	0,0	0,0	0,0	0,0

5) CALAGE SUR LES TABLES DE RÉFÉRENCE DU BILAN DÉMOGRAPHIQUE

Les espérances de vie déduites des quotients estimés après lissage sont supérieures ou égales à celles calculées à partir des quotients de référence (issus du bilan démographique) sur les mêmes périodes. L'écart est important en début de période, mais diminue au fil du temps. À partir de 2009-2013, l'estimation de l'espérance de vie à partir de l'EDP est très proche de celle du bilan démographique.

Espérance de vie à 35 ans des hommes et des femmes

	1976-1984	1983-1991	1991-1999	2000-2008	2009-2013	2017-2019	2020-2022
Femmes							
Estimation brute	48,0	47,2	48,6	49,8	50,8	51,1	50,8
Estimation après lissage, avant calage sur le bilan démographique (a)	46,2	47,0	48,5	49,8	50,7	51,1	50,9
Bilan démographique (b)	45,0	46,4	48,0	49,4	50,5	51,1	50,8
Écart (b-a)	-1,2	-0,6	-0,5	-0,4	-0,2	0,0	-0,1
Hommes							
Estimation brute	39,9	39,7	41,3	43,3	44,7	45,8	45,5
Estimation après lissage, avant calage sur le bilan démographique (a)	39,0	39,6	41,1	43,2	44,7	45,8	45,5
Bilan démographique (b)	37,8	39,2	40,8	42,8	44,5	45,7	45,3
Écart (b-a)	-1,2	-0,4	-0,3	-0,4	-0,2	-0,1	-0,2

Champ : Jusqu'en 2009-2013, personnes nées en France et résidant en France métropolitaine pour l'EDP, France métropolitaine pour le bilan démographique. En 2017-2019 et 2020-2022, personnes nées en France et résidant en France hors Mayotte pour l'EDP, France hors Mayotte pour le bilan démographique.

A. EFFET DE LA DATE DE NAISSANCE DES INDIVIDUS EDP

Pour les quatre premières périodes (de 1976-1984 à 2000-2008), cet écart s'explique en partie (-0,25 an) par le fait que les personnes de l'EDP sont nées au début du mois d'octobre. Les quotients de mortalité étant calculés en âge atteint dans l'année, elles sont donc plus jeunes d'environ 3 mois que la moyenne des personnes de leur génération (qui sont, quant à elles, nées tout au long de l'année, de janvier à décembre). Elles ont donc, à âge donné, une probabilité de mourir dans l'année un peu plus faible que l'ensemble des personnes de leur génération.

En 2009-2013, l'écart dû au mois de naissance est nul. En effet, pour cette période, on utilise les EAR de 2004 à 2012. Entre 2004 et 2007, les personnes sont plus jeunes de trois mois en moyenne. À partir de 2008, les personnes de l'EDP sont nées les quatre premiers jours de chaque trimestre et sont plus âgées d'un mois et demi que la moyenne. Comme la mortalité de la période 2009-2013 est calculée à partir des CS de 2004 à 2012, les deux effets du mois de naissance se compensent : la moyenne sur l'ensemble de l'échantillon correspond au 23 juin, soit quasiment au milieu de l'année.

Pour la mortalité des périodes 2017-2019 (CS de 2010 à 2018) et **2020-2022** (CS de 2013 à 2020), les individus sont très légèrement plus âgés que la moyenne (un mois et demi) parce qu'ils sont nés en début de trimestre.

Pour toutes les périodes, l'écart s'explique également par le fait que les indicateurs sont estimés sur des champs différents selon qu'ils sont estimés à partir de l'EDP ou issus du bilan démographique : personnes résidant en France hors Mayotte et nées en France pour les estimations à partir de l'EDP ; ensemble des personnes résidant en France hors Mayotte pour le bilan démographique. Les personnes nées en France possèdent des caractéristiques (lieu de résidence, niveau de vie...) différentes de celles de l'ensemble des personnes résidant en France, ce qui peut expliquer un écart d'espérance de vie.

C. CALAGE SUR LE BILAN DÉMOGRAPHIQUE

Pour avoir des résultats relatifs à l'ensemble des personnes résidant en France hors Mayotte et éliminer le biais généré par l'usage de la date de naissance dans l'EDP, les quotients estimés à partir de l'EDP sont calés sur les espérances de vie du bilan démographique (toutes CS ou diplômes confondus). Un coefficient correctif par sexe et âge a été calculé : il s'agit du rapport entre le quotient de mortalité du bilan démographique et le quotient estimé à partir de l'EDP. Ce rapport est lissé par moyenne mobile sur 5 années (d'âge). On applique ce coefficient correctif aux quotients par CS et par diplôme. On fait donc l'hypothèse que l'écart (en %) est le même quels que soient la CS ou le diplôme.

On applique ce coefficient correctif, même lorsque l'écart avec le bilan démographique est faible, par exemple pour les femmes en 2017-2019. En effet, même lorsque l'écart est faible sur l'ensemble, les quotients de mortalité de l'EDP ont tendance à être inférieurs à ceux du bilan démographique avant 71 ans et légèrement supérieurs après cet âge.

6) INTERVALLE DE CONFIANCE A 90 % DES ESPÉRANCES DE VIE

Les intervalles de confiance ont été estimés par Bootstrap sur la période 2020-2022. Nous avons tiré, avec remise, 1 000 échantillons d'individus de la même taille que l'échantillon initial. Au final, on obtient 1 000 estimations d'espérance de vie, calculées à partir des quotients de mortalité bruts. L'analyse de la distribution (le 5^e centile et le 95^e centile) fournit l'intervalle de confiance à 90 %.

Cette estimation de l'incertitude ne prend en compte que celle générée par l'échantillonnage dans l'estimation des quotients de mortalité bruts. Le lissage et le calage sur le bilan démographique ne sont pas pris en compte dans le bootstrap.

À partir de la période 2009-2013, où les EAR sont utilisées, cette méthode conduit à sous-estimer légèrement la variance, car seul l'aléa de sondage lié à l'EDP, dont l'échantillon est défini au niveau individuel, est pris en compte, alors que celui lié à l'EAR, dont l'échantillon est défini à un niveau géographique plus agrégé, ne l'est pas.

Intervalle de confiance à 90 % de l'espérance de vie à 35 ans par CS en 2020-2022 (données brutes)

Espérance de vie à 35 ans	Médiane femme	IC 90 + ou -	Médiane homme	IC 90 + ou -
Agriculteur	51,6	0,8	47,7	0,6
Artisan commerçant, dont	52,0	0,4	46,7	0,4
Artisan	ns	ns	46,5	0,5
Commerçant et assimilé	51,5	0,7	46,6	0,5
Cadre	52,9	0,5	49,0	0,3
Cadre de la fonction publique, profession intellectuelle et artistique	53,2	0,7	49,3	0,6
Cadre technique d'entreprise	ns	ns	48,8	0,5
Cadre administratif et commercial	52,1	0,8	49,1	0,5
Profession intermédiaire	52,4	0,3	47,4	0,3
Profession de l'enseignement primaire et professionnel	52,8	0,5	48,7	0,7
Technicien	ns	ns	46,8	0,6
Intermédiaire des entreprises	52,2	0,5	47,0	0,5
Intermédiaire de la santé et du travail social	52,2	0,5	ns	ns
Agent de maîtrise de production	ns	ns	46,8	0,6
Employé	51,3	0,2	45,4	0,4
Employé de commerce	51,6	0,5	Ns	Ns
Policier, militaire et agent de sécurité privée	ns	ns	45,8	0,7
Employé de la fonction publique	51,5	0,3	45,6	0,6
Employé administratif d'entreprise	51,5	0,3	45,8	0,8
Personnel des services aux particuliers	50,4	0,3	ns	ns
Ouvrier	49,8	0,4	43,9	0,2
Conducteur du transport	ns	Ns	44,3	0,6
Ouvrier qualifié de type industriel ou artisanal	50,9	0,6	44,0	0,3
Ouvrier peu qualifié de type industriel ou artisanal	49,1	0,5	43,5	0,4
Inactif non retraité	46,0	0,4	34,9	0,6
écart cadre – ouvrier	3,1	0,6	5,1	0,4

Intervalle de confiance à 90 % de l'espérance de vie à 35 ans par diplôme en 2020-2022 (données brutes)

Espérance de vie à 35 ans	Médiane femme	IC 90 + ou -	Médiane homme	IC 90 + ou -
Sans diplôme	47,1	0,4	41,0	0,4
Brevet, CEP	50,5	0,4	43,9	0,5
CAP, BEP	50,9	0,3	45,2	0,2
Baccalauréat	51,7	0,3	46,7	0,3
Supérieur au baccalauréat	52,6	0,3	48,5	0,3
écart « sup. au baccalauréat » - sans diplôme	5,5	0,5	7,5	0,5

Cette estimation de l'incertitude avait déjà été effectuée pour 2009-2013. Les intervalles de confiance sont très proches en 2009-2013 et en 2020-2022. Ils se sont légèrement réduits pour les catégories dont la part s'accroît (cadres, titulaires du baccalauréat, ...) et se sont légèrement accrus pour celles dont la part baisse (agriculteurs, ouvrières, sans diplôme, ...).

Il est à noter que l'intervalle de confiance ne prend pas en compte les erreurs dues :

- aux non réponses ou réponses erronées pour la catégorie sociale et le diplôme,
- aux choix de méthode de correction et d'imputation de la catégorie sociale et du diplôme
- aux choix de méthode de lissage, de calage sur le bilan démographique.

PARTIE 3. LES ÉCARTS D'ESPÉRANCE DE VIE ENTRE CADRES ET OUVRIERS : 5 ANS CHEZ LES HOMMES, 3 ANS CHEZ LES FEMMES

Insertion de la maquette de l'ip