

**LES DOSSIERS DE LA DREES**

**N° 95 • mars 2022**

# **Caractéristiques socio-économiques des individus aux formes sévères de Covid-19 au fil des vagues épidémiques**

**Exploitation d'un appariement de données d'hospitalisation (SI-VIC) et de données socio-économiques (Fidéli)**

Lino Galiana, Olivier Meslin (Insee), Noémie Courtejoie, Simon Delage (DREES)



# **Caractéristiques socio-économiques des individus aux formes sévères de Covid-19 au fil des vagues épidémiques**

**Exploitation d'un appariement de données d'hospitalisation (SI-VIC) et de données socio-économiques (Fidéli)**

Lino Galiana, Olivier Meslin (Insee), Noémie Courtejoie, Simon Delage (DREES)

Remerciements : Sébastien Roux, Mathilde Gaini, Benoît Ourliac, Lionel Wilner, Nicolas Carnot, Jehanne Richet, Florence Jusot, Michael Sicsic, Mathias André, Odran Bonnet, Albert Vuagnat, Diane Naouri, Élisabeth Fery-Lemonier, Pierre Lamarche

---

Retrouvez toutes nos publications sur : [drees.solidarites-sante.gouv.fr](https://drees.solidarites-sante.gouv.fr)

Retrouvez toutes nos données sur : [data.drees.sante.gouv.fr](https://data.drees.sante.gouv.fr)



**LES DOSSIERS DE LA DREES**

**N° 95 • mars 2022**

# **Synthèse Caractéristiques socio- économiques des individus aux formes sévères de Covid-19 au fil des vagues épidémiques**

**Exploitation d'un appariement de données  
d'hospitalisation (SI-VIC) et de données  
socio-économiques (Fidéli)**

Lino Galiana, Olivier Meslin (Insee), Noémie Courtejoie, Simon Delage (DREES)

Retrouvez toutes nos publications sur : [drees.solidarites-sante.gouv.fr](https://drees.solidarites-sante.gouv.fr)

Retrouvez toutes nos données sur : [data.drees.sante.gouv.fr](https://data.drees.sante.gouv.fr)

# SYNTHÈSE

---

La crise sanitaire met en lumière certaines inégalités sociales et territoriales de santé. Les risques d'infection et de développement de formes sévères sont inégalement répartis au sein de la population, du fait d'hétérogénéités dans l'exposition au virus, dans la distribution des facteurs de risque de complications et dans l'accès aux soins. Pourtant, le risque de développer des formes sévères de Covid-19 a principalement été étudié sous l'angle médical et les études existantes, en France comme à l'étranger, analysent essentiellement le rôle des facteurs cliniques, faute de données socio-économiques au niveau individuel. La présente étude entend compléter cette littérature en mettant en relation certaines caractéristiques socio-économiques (âge, sexe, conditions de logement, niveau de vie, lieu de naissance, lieu de résidence) avec le risque de développer une forme sévère de Covid-19, définie comme conduisant à une hospitalisation, éventuellement suivie d'un décès à l'hôpital ou d'une admission en soins critiques (comprenant la réanimation, les soins intensifs et les soins continus). Elle s'appuie sur les données relatives à plus de 67 millions d'individus, dont près de 382 000 individus hospitalisés atteints de Covid-19, issues de l'appariement inédit, au niveau individuel, entre les données SI-VIC (relatives aux patients hospitalisés et atteints du Covid-19) et les données Fidéli (fichiers démographiques de l'Insee sur les logements et les individus). La mise en regard de ces données éclaire la manière dont les caractéristiques de la population développant une forme sévère du Covid-19 diffèrent de celles du reste de la population. La comparaison des différentes vagues de l'épidémie permet également d'étudier l'évolution de ces caractéristiques, compte tenu de l'évolution du contexte sanitaire et du déploiement de la vaccination. Les données mobilisées ne contenant pas d'information sur le statut vaccinal des personnes, le déploiement de la vaccination est appréhendé par le taux de vaccination au niveau des intercommunalités.

Si l'âge et le sexe (dans une moindre mesure), en tant qu'indicateurs de l'état de santé général, restent les principaux facteurs de risque de développer une forme sévère de Covid-19, certaines conditions de vie et de travail sont associées à un risque accru d'hospitalisation et de décès à l'hôpital. Un premier résultat a trait au niveau de vie : à mesure qu'il décroît, le risque d'hospitalisation augmente, ce qui peut s'expliquer par des contrastes d'exposition au virus et d'état de santé. L'effet du surpeuplement dans le logement, qui favorise l'exposition au virus, est également un résultat notable : le risque d'hospitalisation augmente à mesure que diminue la surface disponible du logement rapportée au nombre d'occupants, ce encore davantage pour les occupants du parc social. Enfin, le risque d'hospitalisation est plus important pour certaines populations nées à l'étranger, notamment en Afrique et en Asie. Parmi les personnes hospitalisées, les conditions de logement n'ont pas d'effet sur le risque de complications (décès ou admission en soins critiques). Le risque de décéder à l'hôpital est légèrement moins élevé chez les personnes aux plus hauts revenus, ce qui reflète notamment une distribution inégale des comorbidités associées à des formes sévères de Covid-19. L'analyse par vague souligne, enfin, l'originalité de la quatrième vague par rapport aux précédentes : le gradient de risque d'hospitalisation en fonction de l'âge devient moins prononcé, tandis que le gradient en fonction du niveau de vie se renforce, ce qui pourrait s'expliquer par une inégale répartition de la vaccination, donc de ses effets protecteurs, au sein de la population (recours à la vaccination croissant avec l'âge et le niveau de vie).

# SOMMAIRE

---

■ INTRODUCTION .....	2
■ DONNÉES .....	4
■ DÉFINITION DES VAGUES .....	5
■ CARACTÉRISTIQUES DES PATIENTS HOSPITALISÉS AVEC COVID-19 .....	8
<b>Approche descriptive</b> .....	<b>8</b>
L'âge : le premier facteur de risque .....	8
Un risque d'hospitalisation décroissant avec le niveau de vie .....	9
Une population hospitalisée plus souvent née à l'étranger.....	10
La population hospitalisée réside dans des logements plus petits et appartenant plus fréquemment au parc social.....	11
<b>Approche analytique</b> .....	<b>13</b>
Modélisation retenue .....	13
Résultats sur l'ensemble des vagues .....	14
Résultats vague par vague .....	16
Effet de la vaccination au niveau local .....	17
■ CARACTÉRISTIQUES DES PATIENTS HOSPITALISÉS AVEC COVID-19 DÉCÉDÉS OU ADMIS EN SOINS CRITIQUES .....	22
<b>Approche descriptive</b> .....	<b>22</b>
Passage en soins critiques et décès des patients Covid-19 .....	22
L'âge d'admission en soins critiques, le reflet de pratiques médicales.....	23
Les personnes âgées et les hommes sont surreprésentés au sein des décès à l'hôpital .....	23
<b>Approche analytique</b> .....	<b>23</b>
Modélisation retenue .....	23
Analyse des décès à l'hôpital .....	24
Analyse des passages en soins critiques.....	26
■ DISCUSSION GÉNÉRALE .....	29
■ CONCLUSION .....	31
■ ANNEXES .....	32
<b>Annexe 1. Spécificités des départements et régions d'outre-mer</b> .....	<b>32</b>
<b>Annexe 2. Description détaillée de l'algorithme d'appariement</b> .....	<b>35</b>
<b>Annexe 3. Résultats additionnels des modèles</b> .....	<b>35</b>
Effets départementaux.....	35
Résultats additionnels concernant les modèles décès et soins critiques .....	38
Robustesse des effets mesurés .....	43
<b>Annexe 4. Retraitement des données SI-VIC</b> .....	<b>45</b>
Abréviations .....	45
Contexte.....	45
Bilan du nettoyage .....	45
■ BIBLIOGRAPHIE .....	47

## ■ INTRODUCTION

La crise sanitaire met en lumière certaines inégalités sociales et territoriales de santé. Les risques d'infection et de développement de formes sévères sont inégalement répartis au sein de la population, du fait d'hétérogénéités dans l'exposition au virus, dans la distribution des facteurs de risque de complications et dans l'accès aux soins ([Dubost, et al., 2020](#)). L'exposition au virus est favorisée par la multiplication des contacts à risque et la promiscuité, qui dépendent des conditions de vie et de travail. Le risque d'exposition est notamment accru parmi les professions au contact avec le public, contraints de travailler hors de chez eux y compris en périodes de confinement, ou encore parmi les personnes vivant dans des zones densément peuplées et dans des conditions de logement défavorables (surpeuplement). L'accès aux équipements de protection individuelle (gel hydroalcoolique, masques) dépend, quant à lui, des conditions matérielles. Ces facteurs d'exposition au virus sont souvent cumulatifs, étant donné que les travailleurs dont l'activité a été maintenue en présentiel pendant la pandémie, tels que les aides-soignants ou les employés des secteurs essentiels, font davantage partie des groupes à revenus moyens ou faibles, utilisant davantage les transports publics, disposant généralement de moins d'espace dans leur habitat et vivant dans des zones plus denses ([Wachtler, et al., 2020](#)). En plus des différences dans le risque d'être infecté par le SARS-CoV-2, certaines comorbidités associées au développement de formes sévères de la maladie sont inégalement réparties entre groupes socio-économiques. Ces comorbidités sont notamment plus répandues chez les ouvriers et les personnes aux revenus les plus faibles que dans le reste de la population ([Dubost, et al., 2020](#)). En raison de la répartition socialement inégale de ces facteurs de risque, les populations défavorisées sur le plan socio-économique semblent présenter d'une part, un risque plus élevé d'infection au SARS-CoV-2 et, d'autre part, un risque plus grand d'évolution défavorable de la maladie en cas d'infection.

Les facteurs socio-économiques associés aux formes sévères de Covid-19 ont été étudiés à l'étranger. Les résultats de ces études ne sont, par définition, pas complètement transposables à la France ([Wachtler, et al., 2020](#)). Ces études restent par ailleurs pour la plupart « écologiques », c'est-à-dire cherchant à établir des corrélations spatiales entre la variabilité de l'incidence des hospitalisations ou des décès et la distribution spatiale de potentiels facteurs de risque calculés à l'échelle de groupes d'individus : peu d'entre elles reposent sur le croisement d'informations d'ordre cliniques et socio-économiques au niveau individuel. Les risques d'hospitalisation et de décès semblent négativement associés au niveau de revenu et d'éducation et positivement associés au niveau de pauvreté, de désavantage social, de chômage et de surpeuplement, ainsi qu'à certaines professions particulièrement exposées (personnels de santé et professions peu qualifiées).

Les facteurs de risque associés à l'infection au SARS-CoV-2 en France ont été étudiés par le biais d'enquêtes de séroprévalence en population générale. Parmi elles, celle ayant le plus d'ampleur, réalisée sur un échantillon représentatif de la population, est l'enquête EpiCoV. En mai 2020 ([Warszawski, et al., 2020](#)), vivre dans une commune à forte densité urbaine, exercer une profession dans le domaine du soin ou vivre avec un nombre élevé de personnes dans le même logement, sont associés à un risque plus élevé d'avoir été infecté, ce indépendamment des autres caractéristiques sociodémographiques considérées. Le risque d'infection est également plus élevé pour les personnes dont le niveau de vie est le plus élevé ou le plus faible. La plupart de ces facteurs de risque restent associés à un plus fort taux d'infection en novembre 2020 ([Warszawski, et al., 2021](#)), même si les écarts de taux d'infection par niveau de vie se sont atténués.

Ces enquêtes n'ont cependant pas permis d'étudier les facteurs de risque de développement de formes sévères (conduisant à des hospitalisations éventuellement suivies de décès), qui restent des événements rares, d'autant que les personnes encore hospitalisées au moment de l'enquête, ou décédées avant, n'ont par construction pas été interrogées. Certaines études se sont intéressées aux facteurs de risque de formes sévères de Covid-19 en France, sans mobiliser de données socio-économiques au niveau individuel. Il s'agit, dans la majeure partie, d'études « écologiques », comme défini supra ([Goutte, et al., 2020](#) ; [Pilkington, et al., 2021](#) ; [Gaudart, et al., 2021](#)). Elles ont permis d'identifier certains indicateurs socio-économiques (niveau de revenu et d'éducation, taux de chômage) comme facteurs de risque, sans pour autant parvenir à un consensus, ces effets ne restant pas significatifs dans toutes les zones géographiques ([Gaudart, et al., 2021](#)), ni au cours de toutes les vagues épidémiques ([Pilkington, et al., 2021](#)). La caractérisation de l'impact des divers facteurs de risque de formes sévères peut être améliorée par l'analyse de données au niveau individuel. Cependant, les données existantes sont soit des données d'enquêtes portant sur de petits échantillons et sur de courtes périodes temporelles, ne permettant pas d'étudier séparément les différentes vagues épidémiques ([Sesé, et al., 2020](#)), soit des données exhaustives issues du Système national des données de santé (SNDS) ([Semenzato, et al., 2021](#)), qui ne contiennent que très peu d'informations sur le profil socio-économique des individus. Le principal indicateur permettant d'aborder l'aspect socio-économique dans le SNDS est l'indice de désavantage social<sup>1</sup>, qui n'est pas individualisé, mais calculé au niveau du code postal de la commune de résidence. Ce dernier a été identifié comme positivement associé au risque d'hospitalisation et de décès à l'hôpital ([Semenzato, et al., 2021](#)).

---

<sup>1</sup> L'indice de désavantage social est la synthèse de quatre indicateurs issus des données de l'Insee : le revenu médian par unité de consommation dans le ménage, le pourcentage de bacheliers dans la population de plus de 15 ans, le pourcentage d'ouvriers dans la population active et le taux de chômage.



Compte tenu de la littérature existante à ce jour, la présente étude s'avère être une exploitation sans précédent de données individuelles : (i) exhaustives sur l'ensemble des hospitalisations et décès à l'hôpital liés au Covid-19, sur une longue période temporelle couvrant les quatre premières vagues de l'épidémie ; et (ii) enrichies d'informations socio-économiques détaillées.

## ■ DONNÉES

Les données exploitées dans cette étude proviennent de trois sources :

- SI-VIC (système d'information pour le suivi des victimes d'attentats et de situations sanitaires exceptionnelles) est un système de collecte de données originellement créé en 2016 à la suite des attentats à Paris, afin de renseigner l'identité des victimes d'attentat et leur prise en charge hospitalière. SI-VIC a ensuite été utilisé pendant la crise sanitaire pour renseigner des informations liées à l'hospitalisation des patients atteints de Covid-19. Cette base de données présente un identifiant unique anonymisé (numéro SINUS) et les différents statuts d'hospitalisation du patient (hospitalisation conventionnelle, hospitalisation en soins critiques, hospitalisation en soins de suite et de réadaptation, retour à domicile, décès, etc.), ainsi que les dates associées à ces statuts. Au cours de la période du 1<sup>er</sup> mars 2020 au 10 novembre 2021, environ 530 000 personnes sont enregistrées dans SI-VIC, c'est-à-dire qu'elles ont été hospitalisées avec un diagnostic d'infection au SARS-CoV-2 établi sur la base d'un test RT-PCR ou d'un scanner (TDM) thoracique. Cependant, SI-VIC ne permet pas de différencier de façon fiable, sur l'ensemble de la période, les personnes hospitalisées à cause du Covid-19 de celles qui ont été prises en charge pour une autre pathologie mais qui ont été, à cette occasion, testées positives (Covid incident).
- La source Fidéli (fichiers démographiques sur les logements et les individus), constituée par l'Insee à partir des données fiscales sur les individus et les logements en 2019, décrit les logements et leurs occupants (état civil, composition des ménages, revenus fiscaux et sociaux).
- Les données de vaccination disponibles par EPCI (établissement public de coopération intercommunale), par tranche d'âge et par semaine depuis début 2021, publiées chaque semaine par la Caisse nationale de l'assurance maladie (CNAM).

Ces sources de données ont été traitées en cinq étapes :

- Les données SI-VIC ont été appariées avec la source Fidéli de l'Insee. Cet appariement consiste à rechercher, parmi les individus de Fidéli, un individu dont l'état civil est identique ou très similaire à celui de ceux figurant dans les données SI-VIC. Les variables utilisées pour l'appariement sont les suivantes : le nom de naissance, le nom de conjoint le cas échéant, les prénoms, la date de naissance, l'adresse et la commune de résidence (des éléments supplémentaires sont disponibles en [Annexe 2](#)). Environ 452 000 personnes enregistrées dans SI-VIC ont ainsi été appariées à un individu de Fidéli (soit un taux d'appariement de 85 %). Il est à noter que le taux d'appariement est nettement plus faible pour le début de la première vague (mars et début avril 2020) car les informations identifiantes (nom, prénom) n'ont été que partiellement remplies au début de l'utilisation de SI-VIC pour le suivi de l'épidémie.
- L'identifiant hospitalier (numéro SINUS) est, en principe, unique. L'appariement à la source Fidéli a permis de détecter, dans les données SI-VIC, de potentiels doublons. En effet, le remplissage manuel des informations d'identité des patients dans SI-VIC par les établissements et les consignes données au début de la crise sanitaire de ne pas détailler ces informations, ont engendré une multiplicité de pseudonymes de certains patients dans SI-VIC ([Annexe 4](#)).
- Les données appariées ont fait l'objet de deux restrictions. D'une part, elles ont été restreintes aux patients pris en charge en hospitalisation conventionnelle ou en soins critiques<sup>2</sup> (comprenant la réanimation, les soins intensifs et les soins continus). D'autre part, les individus dont le parcours hospitalier était incohérent ont été retirés des données (des éléments supplémentaires sont disponibles en [Annexe 4](#)). La table finale comprend 408 000 personnes provenant de SI-VIC, correspondant à 394 000 individus distincts dans Fidéli.
- Les personnes avec un SINUS correspondant à des patients des départements et régions d'outre-mer (DROM) ont été exclus du champ de l'analyse pour deux raisons. D'une part, les informations socio-économiques disponibles sur les DROM sont moins riches que celles portant sur la métropole (par exemple, le niveau de vie des ménages ultramarins n'est connu que pour les ménages de Guadeloupe et de La Réunion). D'autre part, les DROM sont surreprésentés dans la quatrième vague de l'épidémie, ce qui induit un effet de composition, lié au fait que la population hospitalisée dans les DROM présente des caractéristiques différentes de la population hospitalisée en métropole. Des éléments supplémentaires spécifiques à ces territoires sont disponibles en [Annexe 1](#).
- Enfin, les données individuelles ont été complétées avec les taux de vaccination mesurés pour l'EPCI de résidence et la tranche d'âge de l'individu. Les deux informations utilisées sont la part d'individus ayant reçu une première dose de vaccin au 1<sup>er</sup> avril 2021 (date intermédiaire de la troisième vague) et la part d'individus ayant un primo-schéma vaccinal complet au 1<sup>er</sup> août 2021 (date intermédiaire de la quatrième vague). Ces informations sont disponibles pour les classes d'âge suivantes : de 0 à 19 ans, de 20 à 39 ans, de 40 à 54 ans, de 55 à 64 ans, de 65 à 74 ans, 75 ans ou plus.

À l'issue de ces traitements, les données comprennent 67,32 millions d'individus dont 381 968 hospitalisés, sur le champ de la France métropolitaine.

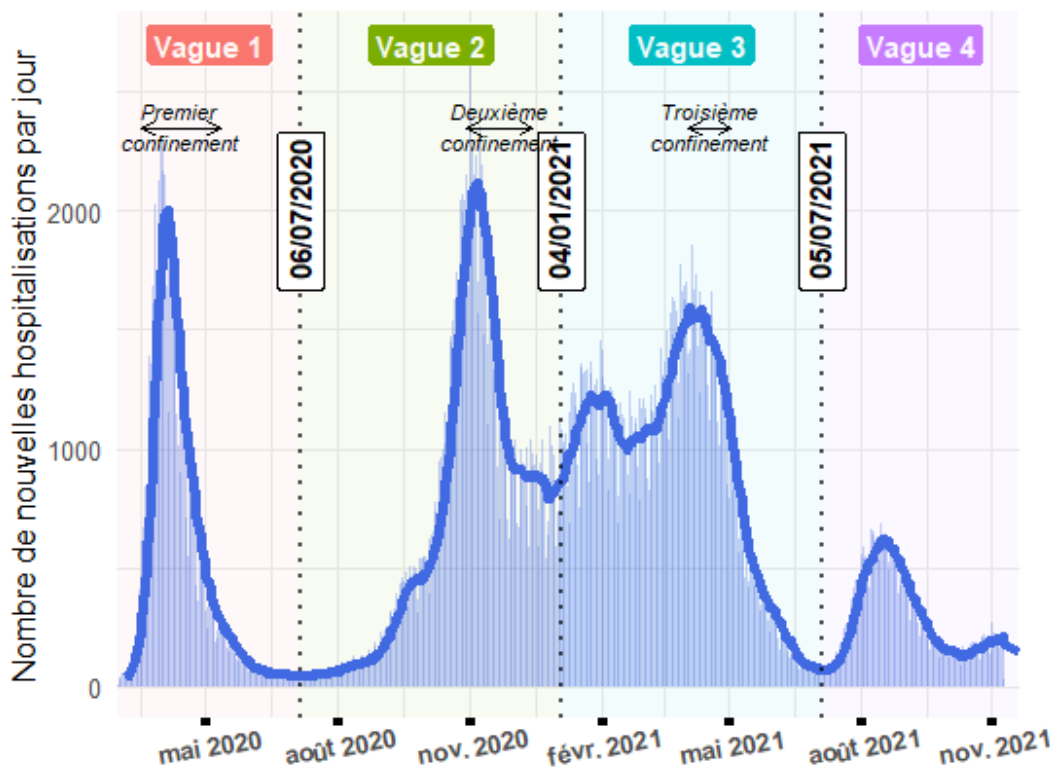
---

<sup>2</sup> On regroupe l'hospitalisation conventionnelle et les soins critiques sous appellation : Médecine, Chirurgie, Obstétrique (MCO)

## ■ DÉFINITION DES VAGUES

La période d'étude a été décomposée en quatre vagues correspondant aux différentes phases de l'épidémie (graphique 1). Les vagues sont définies comme se succédant sans période intermédiaire. Au cours de la première vague, qui s'étend du 1<sup>er</sup> mars 2020 au 5 juillet 2020, 60 942 individus ont été hospitalisés. Cette vague a été principalement localisée dans les régions Grand Est et Île-de-France (graphique 2). Au cours de la deuxième vague, qui s'étend du 6 juillet 2020 au 3 janvier 2021, 121 972 individus ont été hospitalisés. La répartition de cette vague sur le territoire est plus uniforme, même si l'ouest de la France a été globalement moins touché que le reste de la France métropolitaine. La troisième vague, qui s'étend entre le 4 janvier et le 4 juillet 2021 est la plus importante, avec 158 353 individus hospitalisés. Celle-ci a également concerné l'ensemble du territoire avec une incidence plus marquée dans certaines régions de l'est de la France. La quatrième vague, qui s'étend entre le 5 juillet et le 10 novembre 2021, est celle dont le pic est le moins marqué (32 585 hospitalisations). Sa spécificité est que les DROM y sont beaucoup plus représentés qu'au cours des autres vagues.

**Graphique 1** • Dynamique des nouvelles hospitalisations au cours des quatre vagues

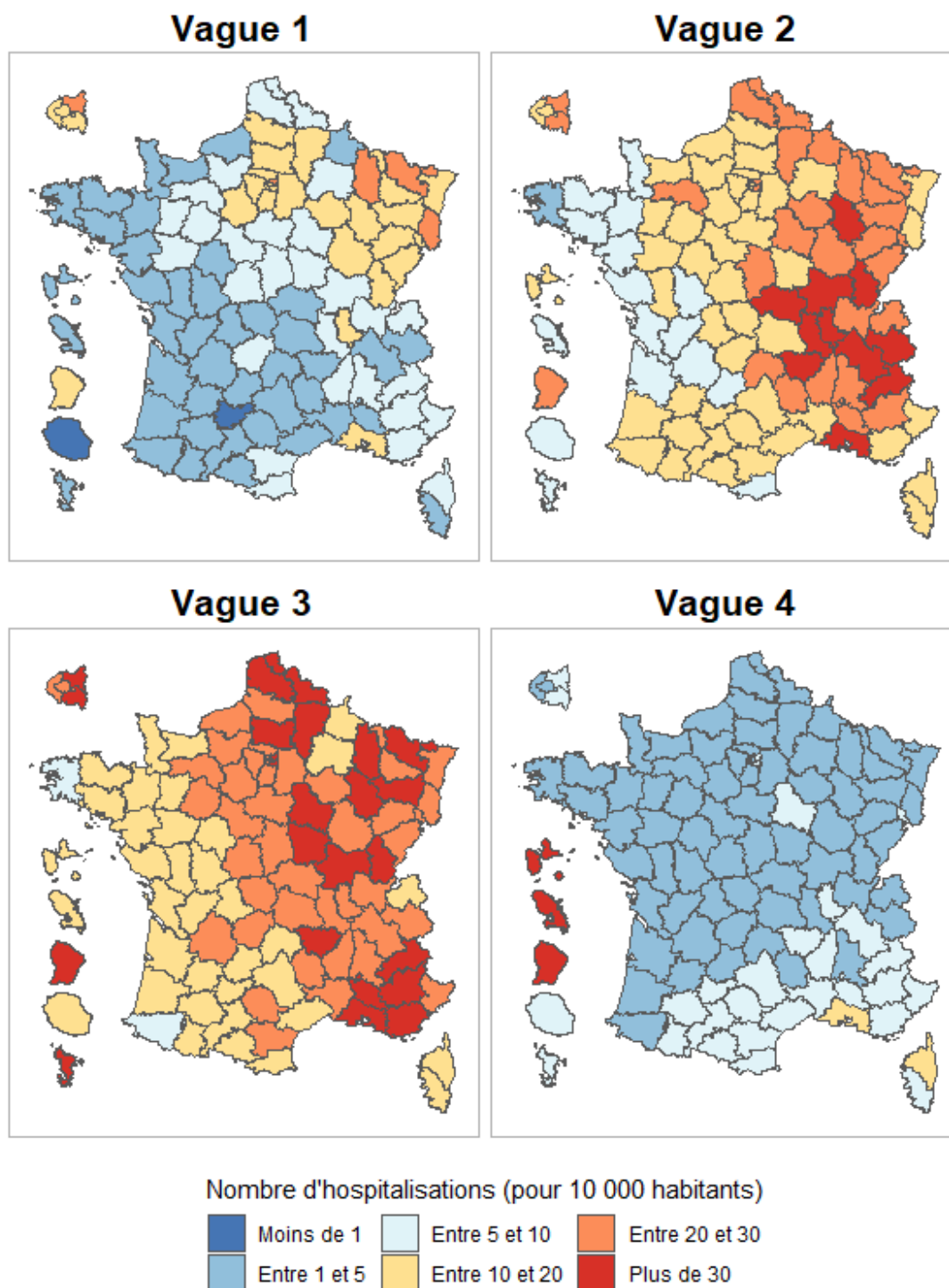


**Lecture** > Le 1<sup>er</sup> novembre 2020, date comprise dans la deuxième vague, environ 1 700 nouvelles hospitalisations ont été observées. Sur une semaine, la moyenne quotidienne était d'environ 2 000 nouvelles hospitalisations.

**Champ** > Individus hospitalisés ayant été identifiés dans la source Fidéli.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 2** • Nombre d'hospitalisations pour 10 000 habitants, par département et par vague



**Lecture** > Lors de la première vague, 23 habitants de la Seine-Saint-Denis sur 10 000 ont été hospitalisés.

**Champ** > Individus hospitalisés ayant été identifiés dans la source Fidéli.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Tableau 1 • Nombre de personnes hospitalisées par vague**

Vague	Tous individus	France Métropolitaine	Champ des analyses économétriques
Vague 1	63 466	62 708	52 974
Vague 2	127 075	124 966	108 654
Vague 3	166 563	162 792	144 357
Vague 4	36 869	31 502	25 492
<b>Toutes vagues</b>	<b>393 973</b>	<b>381 968</b>	<b>331 477</b>

1. Champ de l'étude : Individus hospitalisés pour lesquels l'appariement avec la source Fidéli a pu être effectué, résidant en France métropolitaine dans un logement ordinaire, âgés de 35 ans ou plus et dont le niveau de vie est connu et positif.

**Lecture** > 63 466 individus ont été hospitalisés et retrouvés dans Fidéli pour la première vague dont 62 708 personnes en France métropolitaine. Parmi celles-ci, nous restreignons l'approche analytique à 52 974 personnes ayant plus de 35 ans, appartenant à un ménage ordinaire et dont le niveau de vie est connu et positif.

**Champ** > Individus hospitalisés ayant été identifiés dans la source Fidéli.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

Le tableau 1 présente les informations disponibles dans les données. En complément des informations individuelles comme l'âge et le sexe, la source Fidéli contient des informations sur les caractéristiques du logement occupé par l'individu lorsqu'il s'agit d'un logement ordinaire<sup>3</sup> (surface, nombre de pièces, appartenance au parc privé ou à celui du logement social) ainsi que les caractéristiques du ménage auquel appartient l'individu (composition par sexe et âge, niveau de vie, composition du revenu). L'âge et le sexe utilisés dans cette étude sont issus de la source Fidéli et non des données SI-VIC, afin de traiter une unique source pour tous les individus, hospitalisés ou non. En revanche, la source Fidéli ne contient pas d'informations sur la profession. Il n'est donc pas possible d'utiliser cette source pour déterminer si certaines professions ont été particulièrement exposées aux risques de développer des formes sévères du Covid-19.

**Tableau 2 • Informations disponibles**

Source	Types d'information	Caractéristiques disponibles
SI-VIC	Parcours hospitalier	Type d'hospitalisation, durée de séjour, mode de sortie (décès, retour à domicile)
Fidéli	Caractéristiques démographiques	Âge, sexe, pays de naissance
Fidéli	Caractéristiques économiques individuelles	Niveau de vie (dixième d'appartenance)
Fidéli	Conditions de logements	Type de logement (maison, appartement hors parc social, logement social), taille du logement (surface par occupant)
Fidéli	Facteurs environnementaux additionnels	Commune et IRIS <sup>1</sup> de résidence, variables sociodémographiques au niveau de l'IRIS
CNAM	Informations sur la vaccination	Taux de vaccination par EPCI, par tranche d'âge et par semaine (1 254 EPCI, 6 tranches d'âges)

1. Les IRIS (îlots regroupés pour l'information statistique) sont les plus petites unités spatiales statistiques, inférieures à l'échelon communal. Les communes d'au moins 10 000 habitants et la plupart des communes de 5 000 habitants sont découpées en IRIS.

<sup>3</sup> Le logement ordinaire est un logement défini par opposition à un logement en résidence offrant des services spécifiques (résidences pour personnes âgées, pour étudiants, de tourisme, à vocation sociale, pour personnes handicapées...).

## ■ CARACTÉRISTIQUES DES PATIENTS HOSPITALISÉS AVEC COVID-19

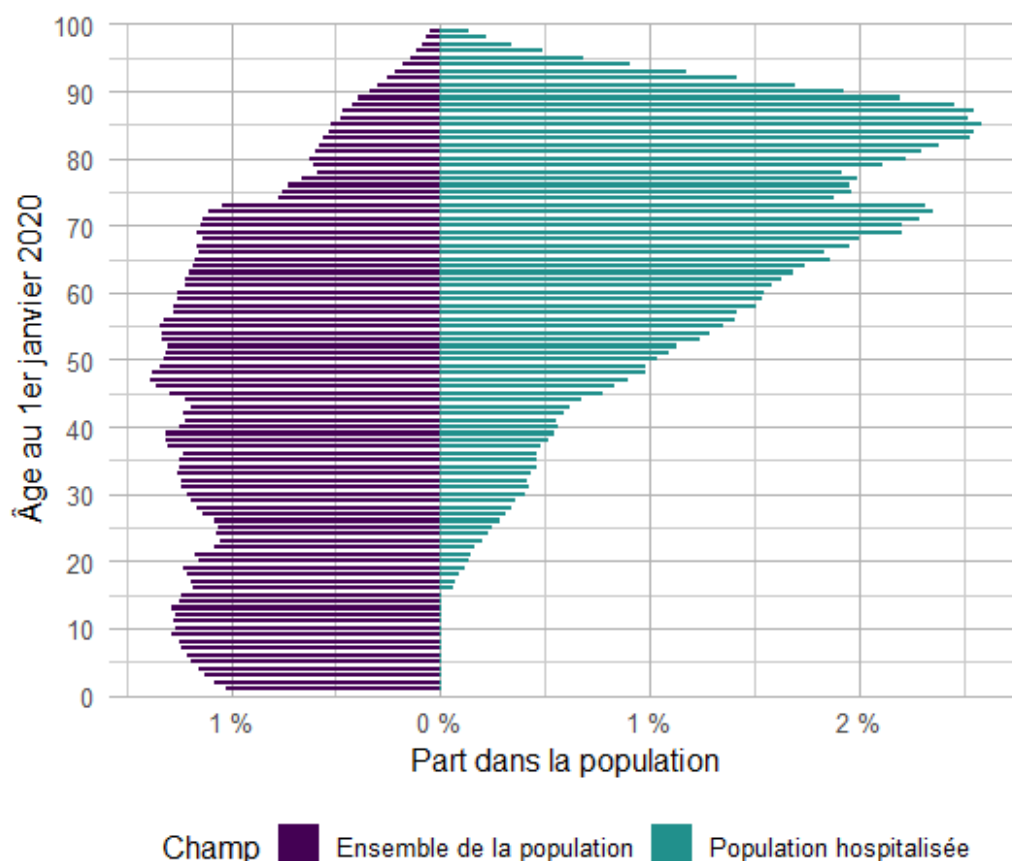
La population hospitalisée au cours de l'épidémie de Covid-19 se distingue du reste de la population par des caractéristiques documentées dans cette section. Tout d'abord, une approche descriptive précise quelles sont les particularités de la population hospitalisée, ensuite une approche analytique modélise le risque d'hospitalisation afin d'isoler le rôle propre à chaque facteur socio-économique toutes choses égales par ailleurs. Tous les résultats de cette section portent sur la France métropolitaine. Le champ est précisé sous chaque tableau ou figure.

### Approche descriptive

#### L'âge : le premier facteur de risque

L'âge est le facteur prépondérant de risque de développer une forme sévère de Covid-19 ([Williamson, et al., 2020](#) ; [Gao, et al., 2021](#)) ; ce fait se retrouve dans la présente étude. La pyramide des âges de la population hospitalisée du fait de l'épidémie est ainsi substantiellement différente de celle de la population résidente en France métropolitaine : 72 % des individus hospitalisés ont 60 ans ou plus, contre 27 % de l'ensemble de la population. Le fait d'être de sexe masculin constitue également un facteur de risque, bien que d'une façon plus limitée : 52 % des individus hospitalisés de plus de 60 ans sont des hommes, contre 45 % des individus de plus de 60 ans.

**Graphique 3** • Distribution de la population résidente et de la population hospitalisée par âge



**Note** > La source Fidéli contient tous les individus résidant en France, y compris les mineurs. En revanche, cette source contient l'état civil des seuls adultes (15 ans ou plus) connus de l'administration fiscale. Ainsi, un individu de moins de 15 ans hospitalisé pour Covid-19 figurera dans SI-VIC et dans la source Fidéli, mais les deux enregistrements ne pourront pas être appariés car Fidéli ne contient pas l'état civil de cet individu.

**Lecture** > Environ 1,3 % des individus résidant en France métropolitaine sont âgés de 50 ans ; environ 1 % des individus hospitalisés sont âgés de 50 ans.

**Champ** > Individus résidant en France métropolitaine.

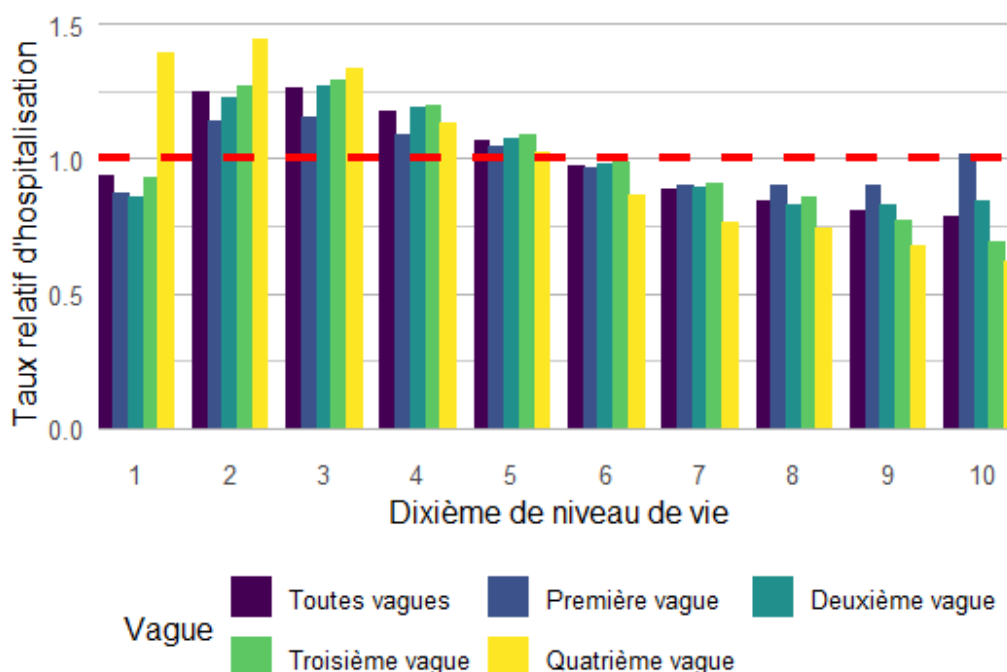
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

Les caractéristiques démographiques ne sont néanmoins pas les seuls facteurs pouvant influencer les risques de développer une forme sévère du Covid-19. La nouveauté de cette étude est qu'elle mobilise des données exhaustives pour mesurer la corrélation entre certaines caractéristiques socio-économiques et le risque de développer une forme sévère du Covid-19.

### Un risque d'hospitalisation décroissant avec le niveau de vie

La population hospitalisée a un niveau de vie plus faible que l'ensemble de la population. Ainsi, le niveau de vie moyen des individus hospitalisés est inférieur de 6 % au niveau de vie moyen de l'ensemble de la population et 57 % des individus hospitalisés ont un niveau de vie inférieur à la médiane. Cette surreprésentation des populations les plus modestes au sein de la population hospitalisée s'est accrue au cours de l'épidémie : la part des individus ayant un niveau de vie inférieur à la médiane parmi les individus hospitalisés s'élève respectivement à 57 % et 63 % lors des troisième et quatrième vagues, contre 53 % lors de la première vague (graphique 4). Le premier dixième de niveau de vie présente un taux relatif d'hospitalisation plus faible que les dixièmes suivants car ce dixième comprend un grand nombre de personnes de moins de 40 ans (notamment des étudiants) pour lesquels le risque d'hospitalisation est plus faible.

**Graphique 4 • Taux relatif d'hospitalisation par niveau de vie et par vague**



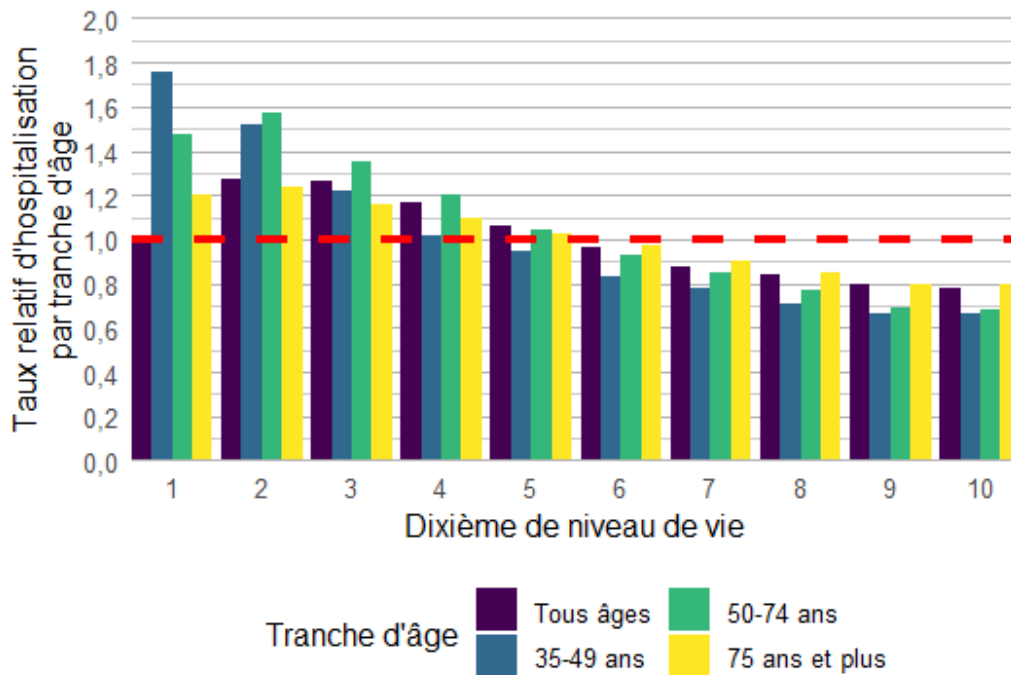
**Lecture** > Lors de la deuxième vague de l'épidémie, le taux d'hospitalisation des individus appartenant au deuxième dixième de niveau de vie était 1,2 fois plus élevé que le taux moyen d'hospitalisation dans l'ensemble de la population

**Champ** > Individus appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

Par ailleurs, cette surreprésentation des populations les plus modestes se retrouve au sein de chaque tranche d'âge, et est d'autant plus marquée que les individus hospitalisés sont jeunes (graphique 5). Au sein de la population âgée de 35 à 49 ans, le taux d'hospitalisation des individus appartenant aux 20 % les plus modestes est 1,6 fois plus élevé que le taux d'hospitalisation moyen de cette tranche d'âge ; ce risque relatif s'élève à 1,5 pour les 50-74 ans, et 1,2 pour les 75 ans ou plus. Cette corrélation plus marquée du risque d'hospitalisation avec le niveau de vie pour les populations plus jeunes peut s'expliquer de multiples façons : un niveau de vie plus faible peut par exemple être corrélé à un état de santé moins bon, à des comorbidités plus fréquentes, à des conditions de logement moins confortables (suroccupation du logement) ainsi qu'à l'exercice de métiers plus exposés au risque de contamination.

**Graphique 5 • Taux relatif d'hospitalisation par niveau de vie et par tranche d'âge**



**Lecture** > Le taux d'hospitalisation des individus âgés de 35 à 49 ans et appartenant au premier dixième de niveau de vie est 1,76 fois plus élevé que le taux d'hospitalisation moyen de cette tranche d'âge.

**Champ** > Individus appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

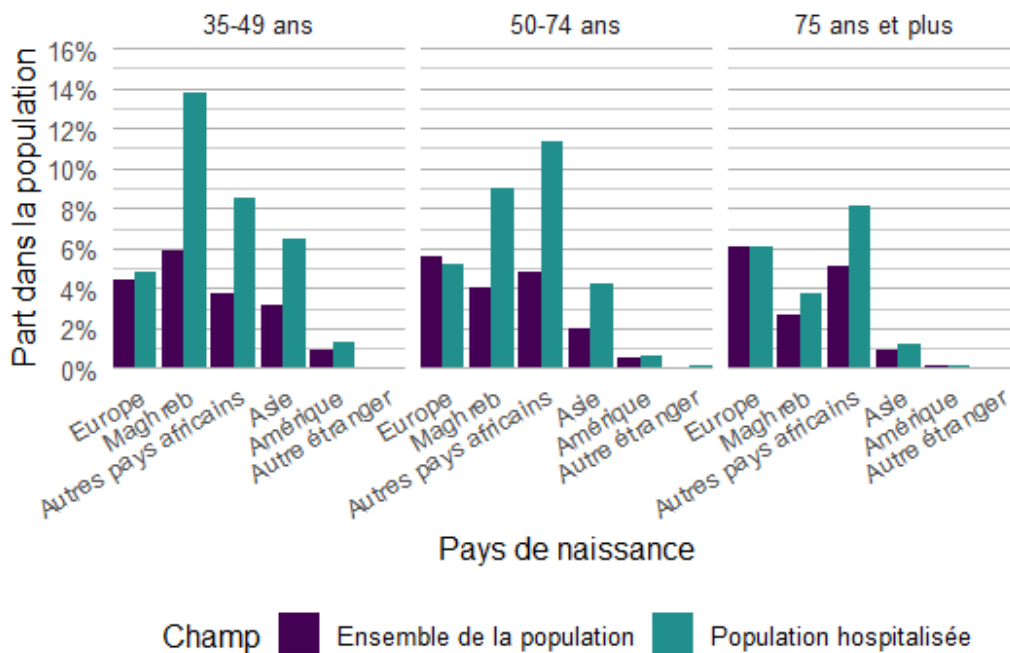
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

### Une population hospitalisée plus souvent née à l'étranger

Les individus hospitalisés sont plus fréquemment nés à l'étranger : 26 % de la population hospitalisée de 35 ans ou plus est née à l'étranger, contre 17 % de l'ensemble de la population de 35 ans ou plus. Cette surreprésentation des personnes nées à l'étranger se retrouve au sein de toutes les tranches d'âge. Toutefois, cette surreprésentation ne concerne que les individus nés hors d'Europe (graphique 6), en particulier les personnes nées en Afrique (Maghreb et Afrique subsaharienne) : 17 % des individus hospitalisés de 35 ans ou plus sont nés en Afrique (Maghreb et Afrique subsaharienne), contre 9 % des personnes résidentes en France métropolitaine de 35 ans ou plus. Cette corrélation entre pays de naissance et risque d'hospitalisation peut s'expliquer comme la corrélation entre niveau de vie et risque d'hospitalisation : état de santé moins bon, comorbidités plus fréquentes, conditions de logement dégradées et exercice de métiers plus exposés. Elle peut également être liée au fait que certains territoires comprenant une part élevée d'individus nés à l'étranger ont été fortement touchés par l'épidémie (notamment la Seine-Saint-Denis).



**Graphique 6 • Distribution de la population hospitalisée par pays de naissance (hors France)**



**Note** > La part d'individus nés en France n'est pas représentée par souci de lisibilité.

**Lecture** > 11 % des individus hospitalisés et âgés de 50 à 74 ans sont nés au Maghreb, contre 4 % de l'ensemble des individus de 50 à 74 ans.

**Champ** > Individus appartenant à un ménage ordinaire résidant en France métropolitaine.

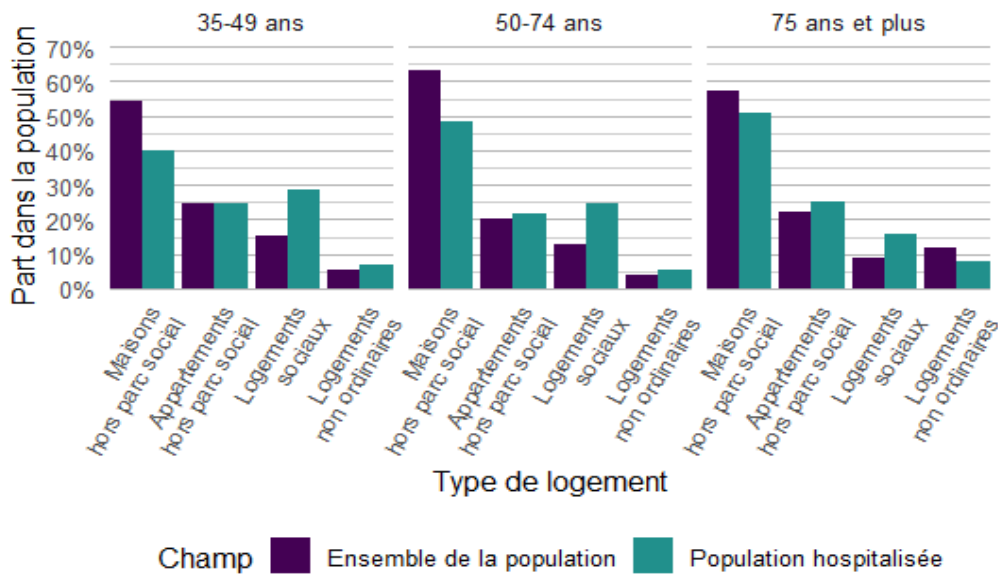
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

### La population hospitalisée réside dans des logements plus petits et appartenant plus fréquemment au parc social

Les conditions de logement de la population hospitalisée sont différentes de celles de la moyenne de la population. Tout d'abord, les individus hospitalisés vivent moins souvent dans des maisons individuelles. Ensuite, ils résident plus fréquemment dans des logements sociaux. Ainsi, 21 % des individus hospitalisés résident dans un logement social, contre 15 % de l'ensemble de la population. Cet écart est encore plus prononcé lorsqu'on décompose la population par tranche d'âge (graphique 7) : il dépasse 11 points de pourcentage pour les individus âgés de 35 à 49 ans et de 50 à 74 ans, et s'établit à 7 points de pourcentage pour les 75 ans ou plus. Enfin, les individus hospitalisés résident dans des logements plus densément occupés que la moyenne (graphique 8) : 33 % des individus hospitalisés et âgés de 50 à 74 ans résident dans un logement dont la surface par personne est inférieure à 30 mètres carrés, contre 24 % de l'ensemble des individus de cette classe d'âge.

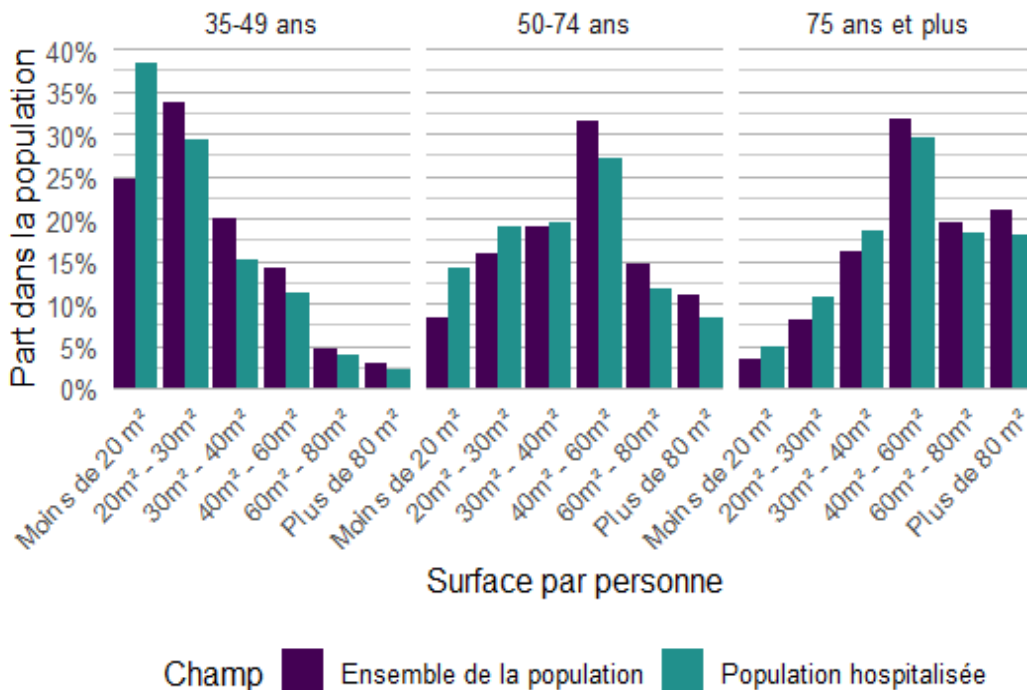
La sous-représentation des personnes de 75 ans ou plus vivant en logement non ordinaire parmi les personnes hospitalisées (graphique 7) peut être due au fait que les cas les plus graves parmi les résidents d'Ehpad ne donnent pas systématiquement lieu à une hospitalisation du fait de la prise en charge de ces patients dans leur lieu de vie. À cet égard, 74 % des décès liés au Covid-19 de résidents d'Ehpad en première vague et 73 % en seconde vague se sont produits au sein de l'établissement et non dans le cadre d'une hospitalisation ([Costemalle, et al., 2020](#)).

**Graphique 7 • Distribution de la population hospitalisée par type de logement et par tranche d'âge**



**Lecture** > 25 % des individus hospitalisés et âgés de 50 à 74 ans résident dans un logement social, contre 13 % de l'ensemble des individus âgés de 50 à 74 ans.  
**Champ** > Individus appartenant à un ménage ordinaire résidant en France métropolitaine.  
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 8 • Distribution de la population hospitalisée par densité d'occupation des logements et par tranche d'âge**



**Lecture** > 14 % des individus hospitalisés et âgés de 50 à 74 ans résident dans un logement comprenant moins de 20 m<sup>2</sup> par occupant, contre 8 % de l'ensemble des individus âgés de 50 à 74 ans.  
**Champ** > Individus appartenant à un ménage ordinaire résidant en France métropolitaine.  
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

## Approche analytique

L'approche analytique vise à déterminer si les observations issues de la comparaison précédente entre la structure de la population française et celle des individus hospitalisés résistent à un raisonnement « toutes choses égales par ailleurs ». Cette approche permet ainsi de prendre en compte les effets de composition entre facteurs de risque. Par exemple, la place d'un individu dans la distribution du revenu dépend en partie de son âge, mais ces deux facteurs pourraient influencer de manière différente le risque d'être hospitalisé. Seule une modélisation statistique permet de séparer ces effets pour déterminer, toutes choses égales par ailleurs, l'effet de l'âge à revenu donné et l'effet du revenu à âge donné.

L'approche analytique permet également de hiérarchiser les différents facteurs pouvant influencer le risque d'occurrence d'un événement (hospitalisation, passage en soins critiques, décès). Grâce à la connaissance du lieu de vie des individus présents dans la source Fidéli, il est enfin possible de quantifier à la fois le rôle de caractéristiques individuelles et de facteurs géographiques liés à ce lieu de vie.

Les modélisations qui suivent portent sur les individus hospitalisés pour lesquels l'appariement avec la source Fidéli a pu être effectué, résidant en France métropolitaine dans un logement ordinaire, âgés de 35 ans ou plus et dont le niveau de vie est connu et positif (dernière colonne du tableau 1). Cette restriction de champ permet de se concentrer sur les individus pour lesquels toutes les informations socio-économiques sont disponibles

## Modélisation retenue

Une régression logistique permet de modéliser la probabilité d'occurrence d'un événement (hospitalisation, passage en soins critiques ou décès) en fonction de divers facteurs explicatifs : des facteurs individuels ( $X_i$ ) et des facteurs locaux, communs à tous les individus habitant dans la même zone ( $Z_c$ ). On considère ainsi l'équation suivante, qui, en logarithme, spécifie le risque d'être hospitalisé rapporté à celui de ne pas l'être, en fonction des facteurs précédents :

$$\log \left( \frac{\mathbb{P}(\text{hospit}|X_i, Z_c)}{1 - \mathbb{P}(\text{non hospit}|X_i, Z_c)} \right) = \beta X_i + \gamma Z_c \quad (1)$$

L'avantage de la régression logistique est qu'il est possible de transformer les coefficients, lorsqu'ils sont relatifs à des variables explicatives sous forme d'indicatrice, en *odds ratio* afin d'interpréter les effets. Le ratio  $p/(1-p)$ , parfois appelé *odds*, correspond au risque d'occurrence de l'événement pour un groupe donné en contrôlant par la taille de la population en question. L'*odds ratio* (rapport de chances) consiste à comparer cette grandeur pour deux groupes dont les caractéristiques diffèrent. Il s'agit donc d'un ratio de la forme  $p(1-q)/q(1-p)$ . En supposant que la seule différence entre les deux groupes est une caractéristique particulière (appartenance à une classe d'âge, à un groupe de revenu...), la transformation exponentielle des coefficients permet de simplifier l'expression de l'*odds ratio* en raisonnant toutes choses égales par ailleurs.

Le modèle est estimé par une méthode statistique usuelle, le maximum de vraisemblance. Cette méthode permet d'obtenir les coefficients, les *odds ratios* et les intervalles de confiance qui correspondent, à condition de disposer de suffisamment d'observations, comme c'est le cas ici.

Lorsque le nombre de cas positifs est faible, comme dans cette situation où la part des individus hospitalisés représente 0,5 % de la population, les concepts d'*odds ratio* ( $p(1-q)/q(1-p)$ ) et de risque relatif ( $p/q$ ) se confondent approximativement<sup>4</sup>. Cette dernière notion est plus simple d'interprétation que l'*odds ratio*. Dans ce cadre, la régression logistique permet ainsi de décomposer, toutes choses égales par ailleurs, l'effet d'une variable sur le risque d'être hospitalisé.

Tous les modèles présentés dans cette section comprennent les variables explicatives suivantes : âge (par tranche d'âge de cinq ans), sexe, niveau de vie (par dixième), pays de naissance, conditions de logement (indicatrice de logement social et surface disponible par personne dans le logement regroupé en tranches). Les modèles sont estimés avec des effets départementaux qui permettent de contrôler de facteurs inobservés spécifiques à ces départements, notamment l'intensité locale de l'épidémie<sup>5</sup>.

<sup>4</sup> L'*odds ratio* est le rapport entre la chance qu'un événement, par exemple l'hospitalisation, arrive dans un groupe d'individus avec celle que ce même événement se révèle dans un autre groupe d'individu. Cette notion permet ainsi de tenir compte de la différence de composition de la population entre deux groupes qui peuvent être, potentiellement, de taille très différente. En notant,  $p$  la probabilité qu'un événement arrive dans le premier groupe et  $q$  que le même événement arrive dans le deuxième groupe, l'*odds ratio* est égal à  $\frac{p(1-p)}{q(1-q)}$ . Lorsque  $p$  (resp.  $q$ ) est faible (l'événement est rare),  $p/(1-p) \approx p$  (resp.  $q/(1-q) \approx q$ ). Autrement dit, l'*odds ratio* est approximativement égal à  $\frac{p}{q}$ . Cette dernière notion est le risque relatif.

<sup>5</sup> La méthode proposée par [Stammann, Heiss, et McFadden \(2016\)](#) pour estimer le plus précisément possible ces effets dans un modèle non linéaire a été mise en œuvre.

## Résultats sur l'ensemble des vagues

Les résultats du modèle estimé sur l'ensemble des vagues de l'épidémie (graphique 9) restent qualitativement inchangés en omettant les effets départementaux ([Annexe 3](#)). Ces effets, inclus pour tenir compte de facteurs locaux inobservés, captent essentiellement des différences d'incidence, et donc d'exposition au virus, affectant par là même le risque d'hospitalisation.

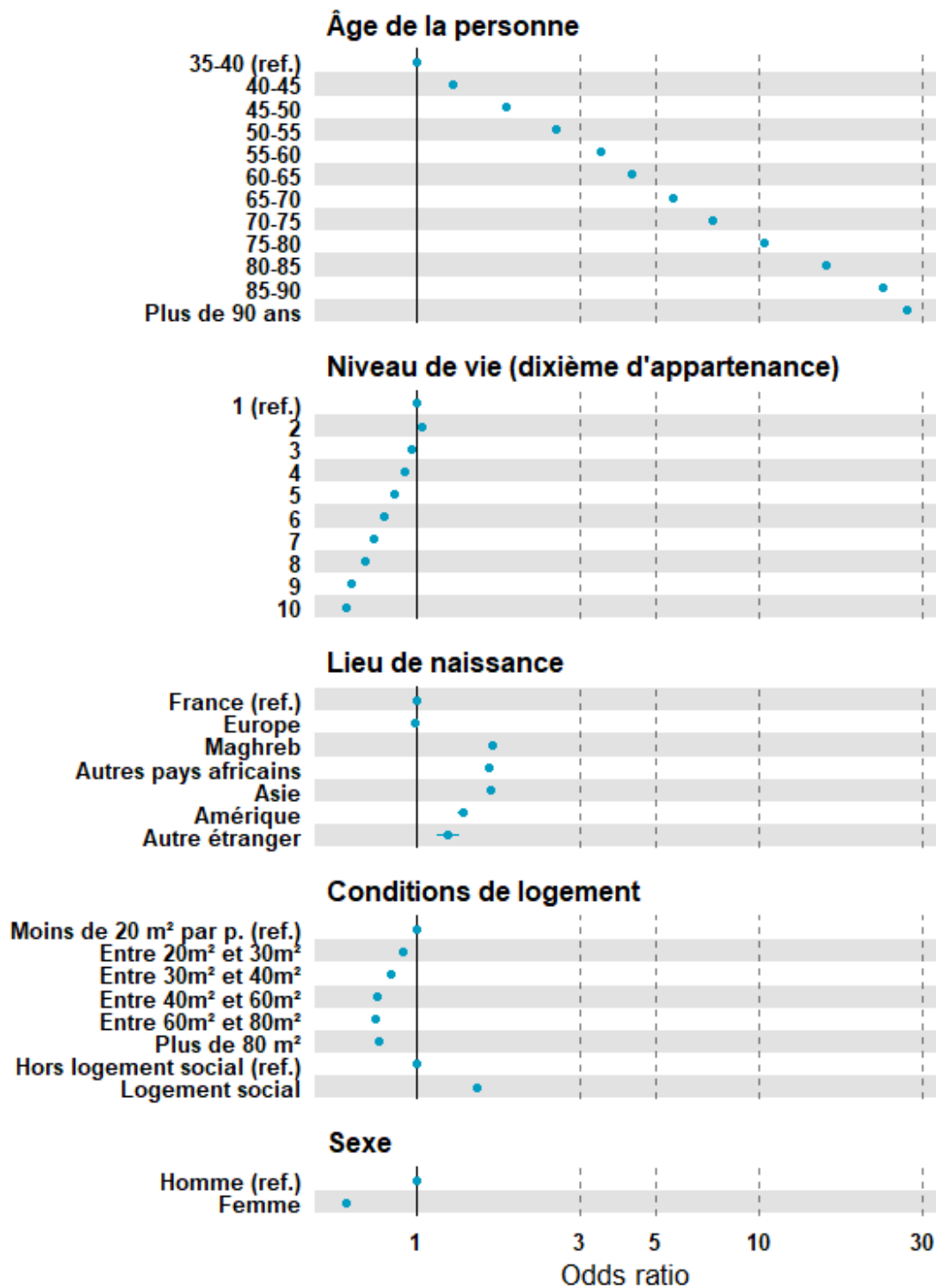
L'approche analytique confirme les analyses descriptives présentées précédemment. Le risque d'hospitalisation croît fortement avec l'âge : toutes choses égales par ailleurs, les personnes ayant entre 80 et 85 ans ont 15,7 fois plus de risques d'être hospitalisées que les personnes ayant entre 35 et 40 ans. Le risque d'hospitalisation est plus élevé chez les hommes : toutes choses égales par ailleurs, ces derniers présentent un risque d'hospitalisation 1,6 plus élevé, que les femmes.

Par ailleurs, les individus nés hors de l'Europe ont davantage de risque d'être hospitalisés : être né au Maghreb multiplie ainsi le risque d'être hospitalisé par 1,7 par rapport au fait d'être né en France. Ce risque relatif s'élève à 1,6 pour les individus nés en Afrique hors Maghreb. En revanche, le fait d'être né en Europe mais hors de France n'augmente pas significativement le risque d'être hospitalisé.

Les conditions de logement affectent également le risque d'être hospitalisé. D'une part, le risque d'hospitalisation décroît au fur et à mesure que la surface disponible par personne augmente, jusqu'à un seuil de 60 m<sup>2</sup> par personne au-delà duquel l'accroissement de la surface par personne a un effet marginal nul. Les individus résidant dans un logement dont la surface disponible par personne est comprise entre 40 et 60 m<sup>2</sup> présentent un risque d'hospitalisation réduit de 0,8 par rapport aux individus ayant les mêmes caractéristiques et résidant dans un logement exigü (moins de 20 m<sup>2</sup> par personne). D'autre part, le fait de vivre dans un logement social multiplie le risque d'hospitalisation par 1,5. La construction d'un modèle alternatif, contenant un terme d'interaction entre la surface par habitant et l'appartenance au parc social, confirme un excès de risque pour les habitants du parc social quasi constant en fonction de la surface disponible par habitant.

Les résultats présentés dans le graphique 9 mettent enfin en évidence une corrélation négative entre risque d'hospitalisation et niveau de vie. Toutes choses égales par ailleurs, le risque d'hospitalisation des individus appartenant au dernier dixième de niveau de vie (les 10 % les plus aisés) est 1,6 fois plus faible que le risque d'hospitalisation des individus du premier dixième (les 10 % les plus modestes). Plus d'éléments concernant la robustesse des effets mesurés sont disponibles en [Annexe 3](#).

**Graphique 9** • Facteurs de risque d'hospitalisation (ensemble de la période d'étude)



**Lecture** > Toutes choses égales par ailleurs, les personnes nées au Maghreb ont un risque d'être hospitalisées 1,7 fois plus important que celles nées en France. Les intervalles de confiance des coefficients sont représentés sous forme de segments, lorsqu'ils sont conséquents. Le nombre d'observations étant important, les bornes des intervalles de confiance sont généralement proches du coefficient estimé.

**Champ** > Individus ayant 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

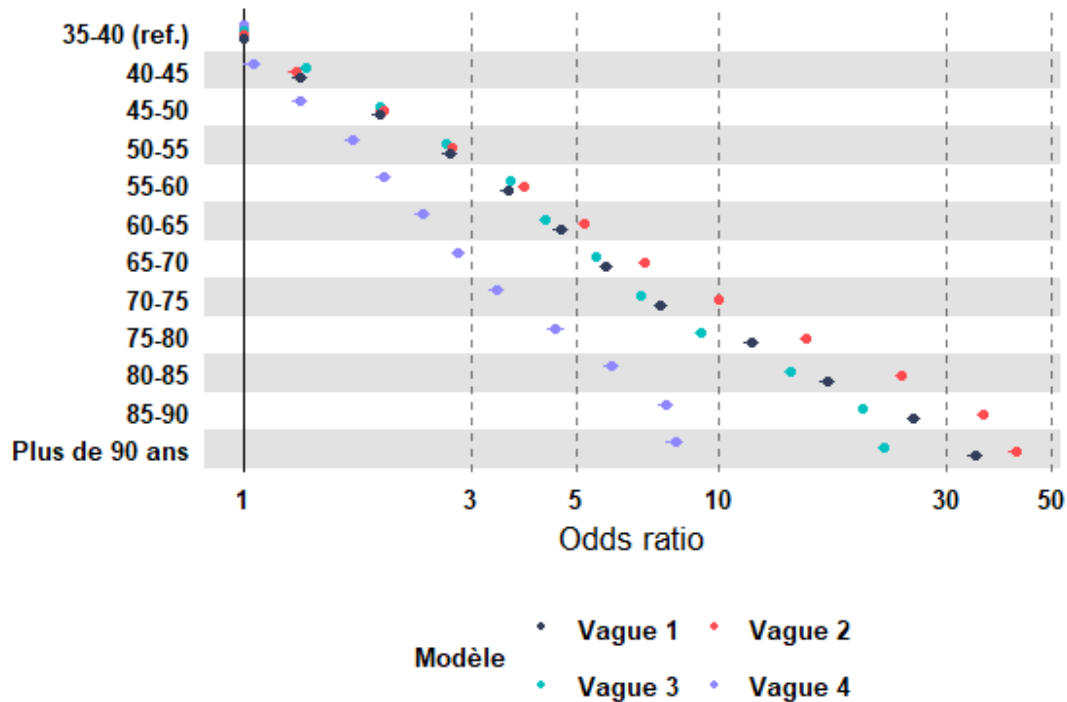
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

## Résultats vague par vague

La composition de la population hospitalisée pour une forme sévère de Covid-19 ainsi que l'importance relative des facteurs de risque ont pu évoluer au cours de l'épidémie. Le modèle présenté dans l'équation (1) a donc été estimé séparément pour chaque vague. Le modèle estimé pour chaque vague inclut les mêmes variables que le modèle précédent (notamment des effets départementaux).

Seuls les facteurs les plus affectés, à savoir âge et niveau de vie, sont présentés graphique 10 et graphique 11. Les résultats concernant les autres facteurs de risque sont reportés en [Annexe 3](#). Les effets départementaux, reportés en [Annexe 3](#) (graphique 25), capturent principalement les différences locales d'incidence<sup>6</sup>.

**Graphique 10** • Résultats des modélisations du risque d'hospitalisation vague par vague, selon l'âge



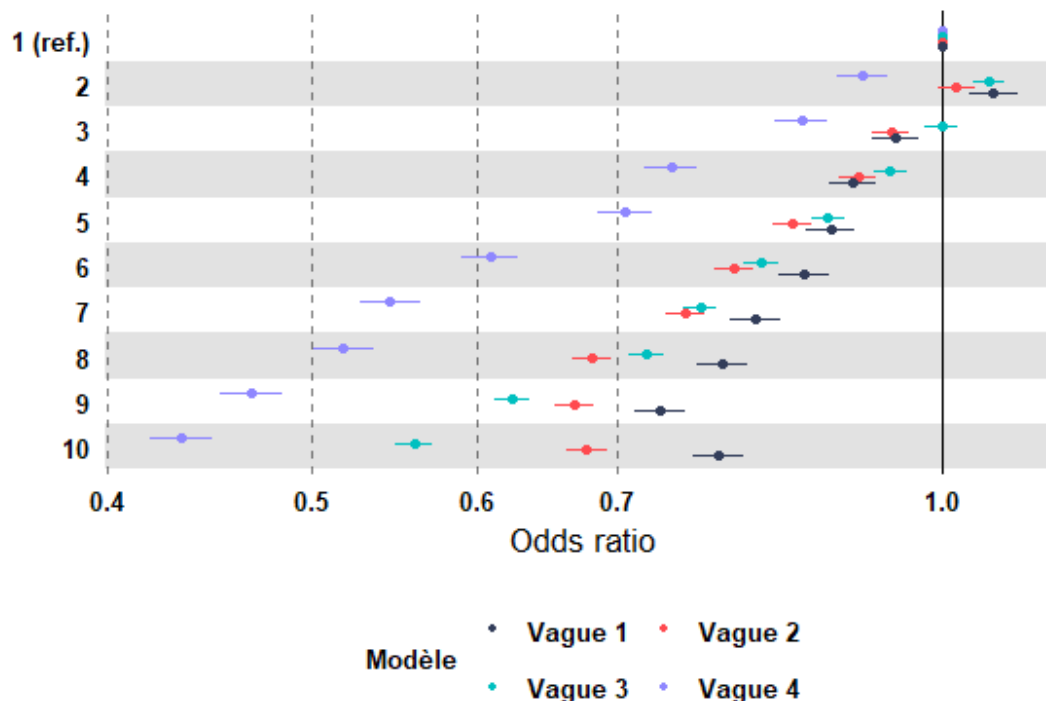
**Lecture** > Toutes choses égales par ailleurs, avoir entre 80 et 85 ans multiplie le risque d'être hospitalisé par 17,8 (resp. 5,7) par rapport au fait d'avoir entre 35 et 40 ans, lors de la première vague (resp. quatrième vague).

**Champ** > Individus ayant 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

<sup>6</sup> La corrélation entre les effets fixes présentés dans le graphique 25 et les taux d'incidence du graphique 2 est supérieure à 0,9 pour toutes les vagues.

## Graphique 11 • Résultats des modélisations du risque d'hospitalisation vague par vague selon le niveau de vie



**Lecture** > Toutes choses égales par ailleurs, appartenir au dernier dixième de niveau de vie plutôt qu'au premier dixième divisait par 1,3 (1/0,78) le risque d'être hospitalisé lors de la première vague. Ce facteur était de l'ordre de 2,3 (1/0,44) lors de la quatrième vague.

**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

Les deux principaux résultats mis en évidence pour l'ensemble des vagues se retrouvent lorsque les vagues sont étudiées séparément : le risque d'hospitalisation croît fortement avec l'âge, et décroît avec le niveau de vie. Toutefois, la quatrième vague se distingue des autres vagues sur deux points. Premièrement, l'âge influence de manière moins marquée le risque d'être hospitalisé : le risque d'hospitalisation des individus ayant entre 80 et 85 ans est (toutes choses égales par ailleurs) 5,7 fois plus élevé que celui des personnes ayant entre 35 et 40 ans lors de la quatrième vague, alors que ce même risque était 17,8 fois plus élevé lors de la première vague. Deuxièmement, la corrélation du niveau de vie avec le risque d'hospitalisation est plus marquée lors de la quatrième vague : toutes choses égales par ailleurs, le risque d'hospitalisation des individus appartenant aux 10 % les plus aisés est 2,3 fois plus important que celui des individus appartenant aux 10 % les plus modestes, alors que ce différentiel s'établissait à 1,3 lors de la première vague. Les deux premières vagues se caractérisaient, quant à elles, par un gradient de revenu en cloche : le risque d'hospitalisation du dernier dixième de revenu n'était pas, toutes choses égales par ailleurs, inférieur à celui du neuvième dixième. La différence est plus marquée pour la première vague que pour la deuxième.

Ces deux particularités de la quatrième vague (atténuation de l'effet de l'âge et changement de la structure de l'effet du niveau de vie) étaient déjà perceptibles lors de la troisième vague, mais dans une mesure nettement moindre et uniquement aux âges les plus élevés (plus de 80 ans) et pour les niveaux de vie les plus élevés (les 20 % les plus aisés).

### Effet de la vaccination au niveau local

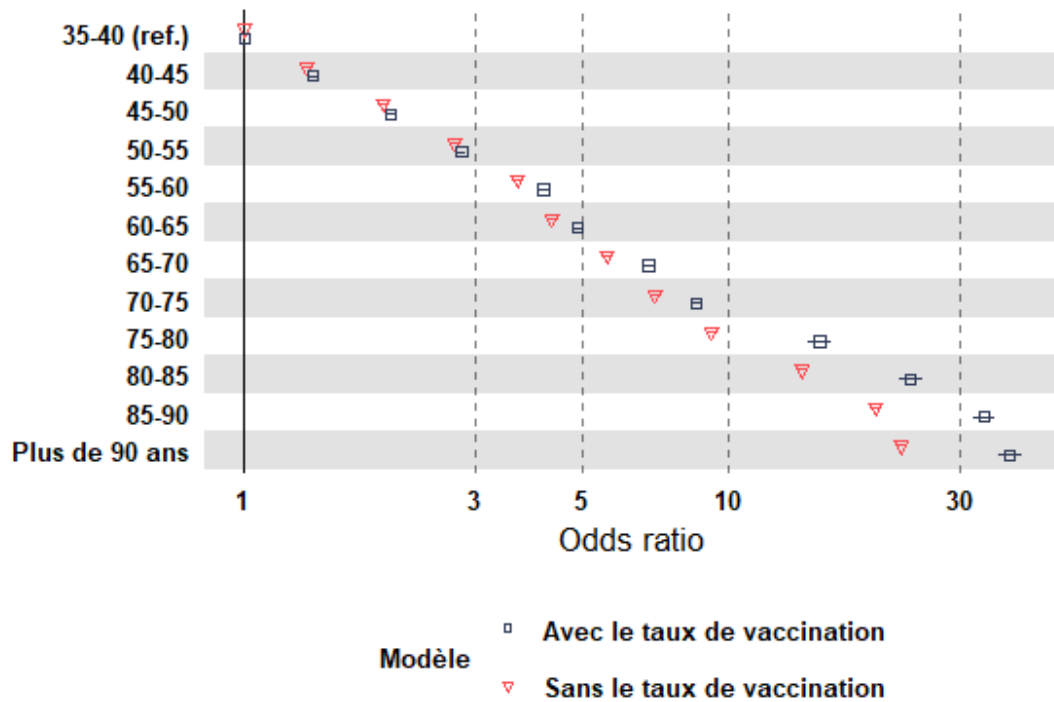
Les troisième et quatrième vagues de l'épidémie ayant été concomitantes du déploiement à grande échelle de la vaccination, il est naturel de se demander dans quelle mesure les particularités de la quatrième vague, et dans une moindre mesure de la troisième vague, mises en évidence dans le paragraphe précédent (et notamment l'aplanissement du gradient en fonction de l'âge) peuvent être interprétées comme un effet de la vaccination. Comme expliqué précédemment, les données exploitées dans cette étude ne contiennent aucune information sur le statut vaccinal des individus et il est donc impossible de mettre en évidence un effet causal au niveau individuel. Néanmoins, il est possible d'exploiter les taux de vaccination disponibles par tranche d'âge au niveau de chaque établissement public de coopération intercommunale (EPCI). Pour enrichir les modèles par vague précédemment présentés, nous avons exploité deux informations relatives à la vaccination : la part d'individus ayant reçu une première dose de vaccin au 1<sup>er</sup> avril 2021 et la part d'individus ayant un primo-schéma vaccinal complet au 1<sup>er</sup> août 2021. Ces deux variables peuvent s'interpréter comme des mesures approchées du statut vaccinal individuel.

Les graphiques 13 et 14 présentent la corrélation estimée entre âge et risque d'hospitalisation lors des troisième et quatrième vagues, selon que le modèle prend en compte les informations sur la vaccination ou non. Le graphique 14 propose la même

comparaison, en y ajoutant les gradients estimés pour les deux premières vagues. La corrélation entre âge et risque d'hospitalisation lors de la quatrième vague est nettement plus forte lorsqu'on tient compte du taux de vaccination par EPCI et par tranche d'âge ; elle devient alors similaire à celle constatée lors des premières vagues. Pour la troisième vague, le gradient d'âge qui s'écartait de celui des vagues précédentes chez les plus âgés (à partir de 80 ans), redevient également similaire à celui des vagues précédentes après l'introduction de la vaccination.

En revanche, introduire les informations relatives à la vaccination dans le modèle n'affecte que marginalement les autres effets estimés pour les troisième et quatrième vagues (en particulier l'effet lié au dixième de niveau de vie).

**Graphique 12 • Effet de la vaccination sur le risque estimé du facteur d'âge lors de la troisième vague**



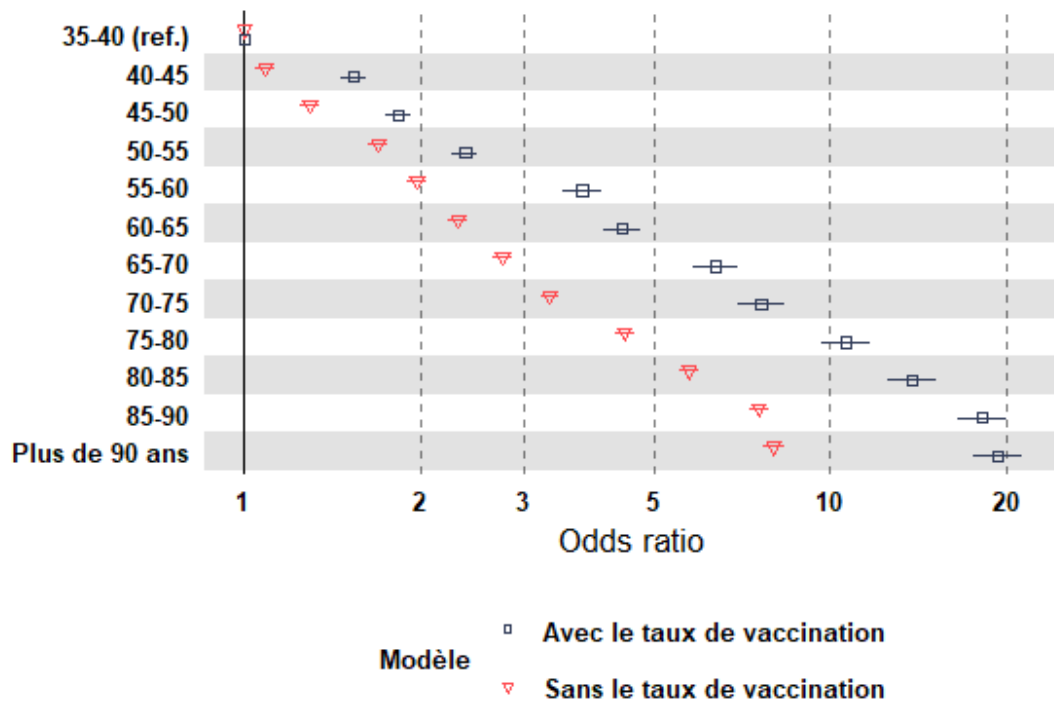
**Lecture** > Lorsqu'on ne prend pas en compte le taux local de vaccination, toutes choses égales par ailleurs, avoir entre 80 et 85 ans multiplie le risque d'être hospitalisé par 14,1 par rapport au fait d'avoir entre 35 et 40 ans. Lorsqu'on tient compte du taux local de vaccination, le facteur multiplicatif est de 23,7.

**Champ** > Individus ayant plus de 35 ans, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.



**Graphique 13** • Effet de la vaccination sur le risque estimé du facteur d'âge lors de la quatrième vague

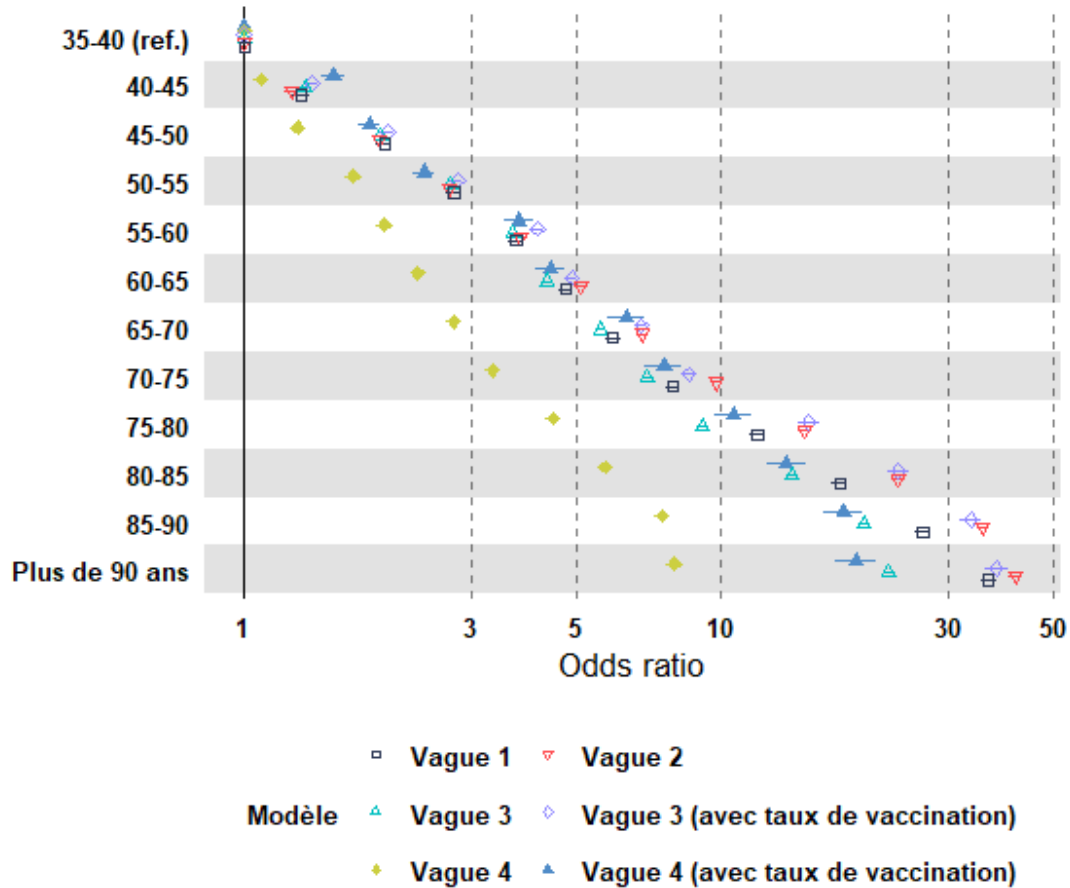


**Lecture** > Lorsqu'on ne prend pas en compte le taux local de vaccination, toutes choses égales par ailleurs, avoir entre 80 et 85 ans multiplie le risque d'être hospitalisé par 5,7 par rapport au fait d'avoir entre 35 et 40 ans. Lorsqu'on tient compte du taux local de vaccination, le facteur multiplicatif est de 13,8.

**Champ** > Individus ayant 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 14 • Effet de la vaccination sur le risque estimé du facteur d'âge**



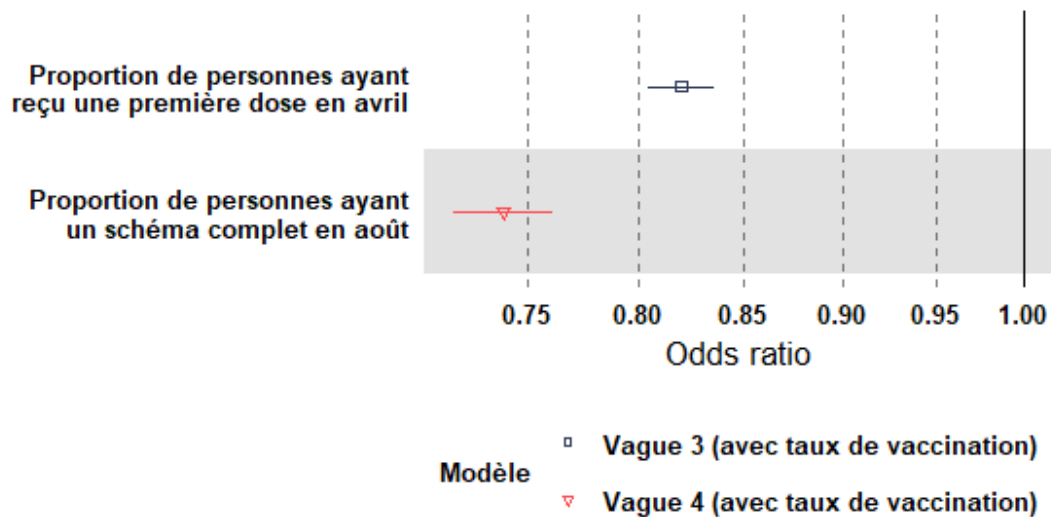
**Lecture** > Lorsqu'on ne prend pas en compte le taux local de vaccination, toutes choses égales par ailleurs, avoir entre 80 et 85 ans multiplie le risque d'être hospitalisé par 5,7 par rapport au fait d'avoir entre 35 et 40 ans. Lorsqu'on tient compte du taux local de vaccination, le facteur multiplicatif est de 13,8. Le même facteur de risque était, par exemple, estimé à 17,8 lors de la première vague.

**Champ** > Individus ayant 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

Par ailleurs, le graphique 15 illustre la corrélation estimée entre risque d'hospitalisation et taux de vaccination par EPCI et par tranche d'âge. On constate que, toutes choses égales par ailleurs, une hausse d'un écart-type du taux local de vaccination dans la classe d'âge d'appartenance, soit une hausse de 20 points de pourcentage du taux de vaccination en troisième vague réduit le risque individuel d'hospitalisation de 22 % ( $1/0,82 - 1$ ). En ce qui concerne la quatrième vague, le risque individuel d'hospitalisation était réduit de 35 % ( $1/0,74 - 1$ ) pour chaque variation d'un écart-type du taux local de vaccination (soit une hausse de 14 points de pourcentage du taux de vaccination).

**Graphique 15** • Effet de la vaccination sur le risque estimé d'être hospitalisé : comparaison des vagues



**Lecture** > Toutes choses égales par ailleurs, une hausse d'un écart-type du taux local de vaccination dans la classe d'âge d'appartenance, soit une hausse de 20 points de pourcentage du taux de vaccination en troisième vague réduit le risque individuel d'hospitalisation de 22 % ( $1/0,82 - 1$ ) lors de la troisième vague. En ce qui concerne la quatrième vague, le risque individuel d'hospitalisation était réduit de 35 % ( $1/0,74 - 1$ ) pour chaque variation d'un écart-type du taux local de vaccination (soit une hausse de 14 points de pourcentage du taux de vaccination).

**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

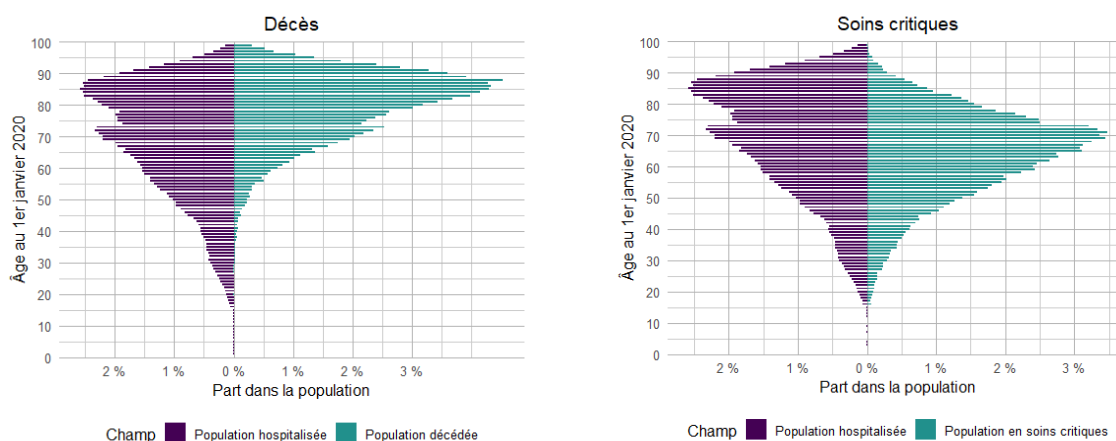
# CARACTÉRISTIQUES DES PATIENTS HOSPITALISÉS AVEC COVID-19 DÉCÉDÉS OU ADMIS EN SOINS CRITIQUES

La première partie de cette étude a permis d'identifier certains facteurs associés au risque d'hospitalisation. Ces facteurs reflètent à la fois une exposition différentielle au virus et une probabilité inégale, en cas d'infection, de développer une forme sévère donnant lieu à hospitalisation. La seconde étape de cette étude est consacrée au parcours hospitalier des patients, une fois qu'ils ont été admis à l'hôpital pour une forme sévère de la maladie. Les différences d'exposition au virus sont ici secondaires car la population d'étude est restreinte aux individus infectés. Il convient de se demander si certains facteurs socio-économiques restent associés au risque de complications, conditionnellement au fait d'être hospitalisé.

## Approche descriptive

### Passage en soins critiques et décès des patients Covid-19

Graphique 16 • Distribution de la population par âge (décès et soins critiques)



**Lecture** > Les individus âgés de 78 ans représentent 2,6 % des individus hospitalisés et 4,3 % des personnes décédées à l'hôpital.

**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

Le parcours hospitalier des individus hospitalisés avec Covid-19 est caractérisé par les soins reçus et l'issue du séjour hospitalier. On distingue en particulier : (i) les personnes prises en charge au moins une fois en soins critiques de celles uniquement prises en charge en hospitalisation conventionnelle ; (ii) les personnes décédées à l'hôpital, de celles retournées à leur domicile.

Pour l'étude des caractéristiques des patients décédés à l'hôpital ou admis en soins critiques, on se restreint aux admissions à l'hôpital survenues dans la période allant du 1<sup>er</sup> mars 2020 au 15 octobre 2021, pour pouvoir suivre les patients pendant au moins 45 jours suivant leur hospitalisation<sup>7</sup>. Au sein du champ de référence, en se restreignant aux 35 ans ou plus, parmi les 342 614 individus hospitalisés en MCO avec Covid-19 au cours des quatre premières vagues de l'épidémie, 73 261 (21 %) ont été pris en charge au moins une fois en soins critiques. 64 432 individus (19 %) sont décédés, dont 45 262 (70 %) sans passer par les soins critiques. La majorité des décès ont lieu en dehors des soins critiques, et cette proportion atteint plus de 93 % chez les personnes âgées de 85 ans ou plus.

<sup>7</sup> L'extraction de la base SI-VIC utilisée pour l'appariement avec Fidéli a été actualisée pour la dernière fois le 1<sup>er</sup> décembre 2021.

## L'âge d'admission en soins critiques, le reflet de pratiques médicales

L'âge médian des individus hospitalisés en MCO est de 71 ans. 23 % de la population hospitalisée a 85 ans ou plus. Cependant, les personnes âgées sont moins fréquemment admises en soins critiques. En effet, l'âge médian en soins critiques est de 65 ans, les 75 ans ou plus ne représentant que 21 % des patients dans ces services, alors qu'ils représentent 42 % des patients en hospitalisation conventionnelle. La proportion de personnes admises en soins critiques augmente avec l'âge jusqu'à 72 ans où elle atteint son maximum, du fait de l'augmentation physiologique du risque de forme sévère avec l'âge. Au-delà de 72 ans, la proportion de personnes admises en soins critiques diminue (graphique 16). Cette baisse est le reflet de pratiques médicales différenciées selon l'âge et non d'une moindre gravité de la maladie chez les patients âgés : la ventilation et l'intubation sont en effet des pratiques invasives à risque, à la suite desquelles la question du sevrage des machines est délicate, ce particulièrement chez les plus âgés ([Guidet, et al., 2017](#)). L'âge est un facteur de risque de formes sévères en soi, mais également un indicateur du degré d'autonomie et de la fréquence des comorbidités ([Barhoumi, et al., 2020](#)). La décision d'admettre ou non les patients en soins critiques relève ainsi d'un arbitrage entre la sévérité de la pathologie et l'état de santé du patient préalablement à la maladie : n'y sont admis que les patients aux formes sévères de Covid-19, dont l'état de santé leur permet tout de même de bénéficier des soins prodigués.

Les hommes, déjà surreprésentés parmi les personnes hospitalisées (53 %), le sont encore davantage parmi les personnes prises en charge en soins critiques (66 %). Des tendances similaires s'observent sur les trois premières vagues, avec une proportion de personnes prises en charge en soins critiques de 20 %, 18 % et 23 % respectivement, pour les première, deuxième et troisième vagues. La quatrième vague diffère des autres avec une proportion de 26 %. Cela s'explique en partie par une proportion plus faible de patients très âgés (qui ont moins tendance à être pris en charge en soins critiques) parmi ceux hospitalisés au cours de la quatrième vague (12 % de 85 ans ou plus en quatrième vague contre 22 % pour les trois premières vagues).

## Les personnes âgées et les hommes sont surreprésentés au sein des décès à l'hôpital

La proportion de personnes décédées durant leur séjour à l'hôpital s'élève à environ 20 % sur l'ensemble des vagues. Ce taux de mortalité à l'hôpital est stable pour les trois premières vagues avec une proportion de 20 %, 20 % et 18 % respectivement, pour les première, deuxième et troisième vagues, puis chute à 14 % au cours de la quatrième vague.

Par ailleurs, le taux de mortalité à l'hôpital présente de fortes disparités en fonction de l'âge et du sexe des patients. Le principal facteur de variation est l'âge : si l'âge médian des personnes décédées est de 82 ans, 60 % de l'ensemble des décès en MCO ont lieu chez les plus de 80 ans, et le taux de mortalité passe de 3 % chez les 35-50 ans à 13 % chez les 60-70 ans et à 32 % chez les 80 ans ou plus. La mortalité hospitalière est également plus élevée chez les hommes (12 %) que chez les femmes (8 %). Elle atteint 42 % chez les hommes de 80 ans ou plus.

La variabilité de la mortalité et de l'admission en soins critiques dans le temps, déjà documentée par ailleurs ([Courtejoie et Dubost 2020](#) ; [Lefrancq, et al., 2021](#)) peut refléter plusieurs facteurs : l'évolution du profil des individus hospitalisés, mais aussi, à caractéristiques comparables, une évolution de la connaissance de la maladie et des modes de prise en charge, la pathogénicité différente de certains variants, voire un effet de la vaccination. Il est donc essentiel de développer une approche analytique pour étudier plus finement l'effet des multiples facteurs de risque de mortalité, et de passage en soins critiques, parmi les patients hospitalisés avec Covid-19.

## Approche analytique

---

### Modélisation retenue

L'étude des facteurs de risque des formes les plus sévères de la maladie parmi les patients hospitalisés implique la définition préalable des parcours hospitaliers défavorables. L'analyse des hospitalisations s'achevant par un décès est incontournable. La question du passage en soins critiques est plus délicate, étant donné que l'admission en soins critiques ne mesure pas uniquement la sévérité de la maladie. Le passage en soins critiques et le décès à l'hôpital reflètent des réalités différentes, qui doivent donc être étudiées séparément<sup>8</sup>.

Des modèles de régression logistique distincts ont été estimés, afin d'étudier respectivement le risque d'admission en soins critiques et le risque de décès à l'hôpital, sur l'ensemble de la période et pour chacune des vagues épidémiques. L'étude des facteurs de risque de décès à l'hôpital est effectuée à partir de l'ensemble de la population des patients hospitalisés. L'étude des facteurs de risque d'admission en soins critiques est réalisée sur la population restreinte aux individus pris en charge en

---

<sup>8</sup> Ces événements ne sont pas complètement disjoints, car des patients décèdent en soins critiques.

soins critiques ou rentrés à domicile, afin de comparer autant que possible les individus pris en charge en soins critiques avec ceux qui présentent des formes moins sévères de la maladie. Sont ainsi exclus les individus décédés à l'hôpital sans passer par les soins critiques. En complément, la population d'étude a été scindée en deux et des modèles distincts ont été estimés sur chacune des sous-populations : les moins de 75 ans et les 75 ans ou plus. L'âge pivot de 75 ans permet en effet d'identifier deux sous-populations qui diffèrent en termes d'état de santé d'une part, de prise en charge à l'hôpital d'autre part. C'est notamment l'âge qui a été défini pour identifier les personnes fragiles prioritaires lors du déploiement de la vaccination. Ce découpage permet d'autoriser l'effet des variables explicatives à différer selon l'âge.

Les modèles portant sur le passage en soins critiques des patients hospitalisés et le risque de décès à l'hôpital intègrent les mêmes variables explicatives que les modèles portant sur le risque d'hospitalisation, à une exception près : le taux de vaccination à l'EPCI par tranche d'âge n'a pas été retenu parmi les variables d'intérêt. En effet, l'approximation de l'effet de la vaccination par cette variable, qui a été faite en population générale, ne semble pas applicable à la sous-population des personnes hospitalisées, qui est très particulière : étant donné que la vaccination protège contre les formes sévères de la maladie, le taux de vaccination parmi les personnes hospitalisées n'est pas le même qu'en population générale, particulièrement pendant la phase de montée en charge de la vaccination (ciblée en priorité sur les individus les plus fragiles, et en particulier sur les plus âgés). Ainsi, au début de la quatrième vague (fin juillet 2021), alors que les personnes non vaccinées constituaient 45 % de la population générale, elles représentaient 87 % des personnes admises en soins critiques et 82 % des personnes décédées à l'hôpital ([DREES, 2021](#)). De plus, la seconde partie de cette étude porte sur le parcours fin et individuel des patients à l'hôpital. Or l'effet de la vaccination contre le risque d'admission en soins critiques et de décès varie en fonction du nombre de doses reçues et du temps écoulé depuis l'injection de ces doses ([Suarez-Castillo, et al., 2021](#)). L'analyse de l'effet de la vaccination sur la survenue d'éventuelles complications à l'hôpital ne peut donc pas se faire à partir des données à l'EPCI, mais nécessiterait d'avoir connaissance du statut de vaccination de chaque individu hospitalisé.

Enfin, ces modèles incluent des effets départementaux. Toutefois, ces effets sont cette fois-ci nettement plus homogènes et ne modifient que marginalement les résultats, contrairement aux modèles portant sur le risque d'hospitalisation.

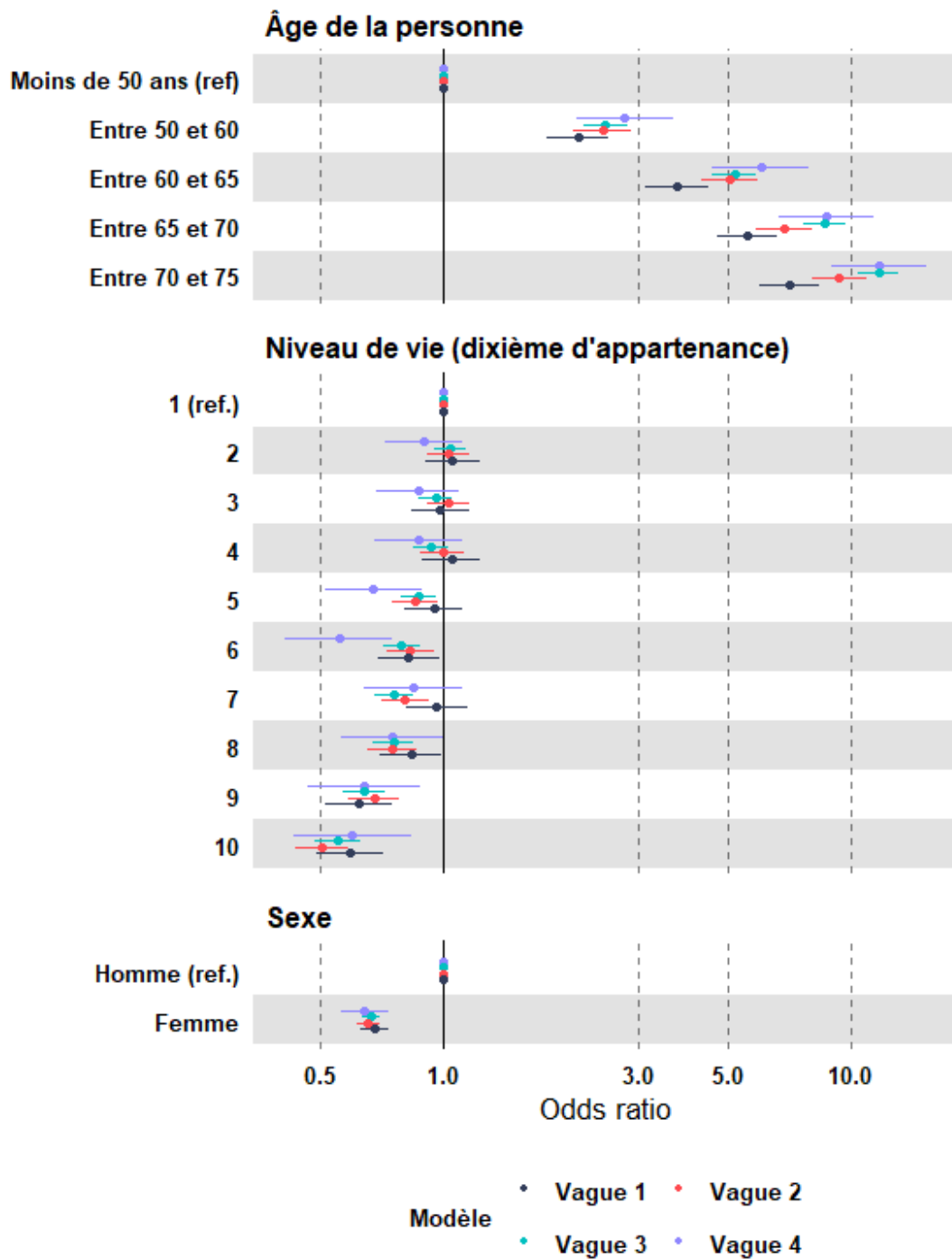
### Analyse des décès à l'hôpital

La quantification des facteurs de risque de décès à l'hôpital (graphiques 17 et 18) souligne que l'effet de l'âge est, une fois encore, prépondérant. Le risque de décéder croît nettement avec l'âge et ce gradient semble s'accroître d'une vague à l'autre. Il s'agit ici d'une évolution relative : l'excès de risque de décéder à l'hôpital lié au fait d'être âgé, par comparaison aux individus les plus jeunes, augmente légèrement d'une vague à l'autre. Cependant, la probabilité absolue de décéder à l'hôpital a plutôt tendance à baisser dans le temps ([El Rais, et al., 2021](#)). On retrouve aussi que le sexe masculin est un facteur de risque, particulièrement chez les personnes âgées.

Chez les moins de 75 ans, le risque de décéder est corrélé avec le niveau de revenu : toutes choses égales par ailleurs, plus le niveau de revenu augmente, plus le risque de décéder décroît. En revanche, cet effet ne se retrouve pas chez les 75 ans ou plus. Il en est de même pour les conditions de logement (le fait d'habiter dans un logement social, ou dans un logement surpeuplé), qui ne sont pas corrélées avec le risque de décéder du Covid-19 conditionnellement au fait d'être hospitalisé, quel que soit l'âge ([Annexe 3](#)).

Le fait d'être né à l'étranger, qui affectait le risque d'hospitalisation, ne modifie quasiment pas le risque de décès, conditionnellement au fait d'être hospitalisé ; les intervalles de confiance obtenus sont larges et les résultats volatils d'une vague à l'autre ([Annexe 3](#)).

**Graphique 17** • Résultats des modélisations du risque de décès (personnes âgées de moins de 75 ans)



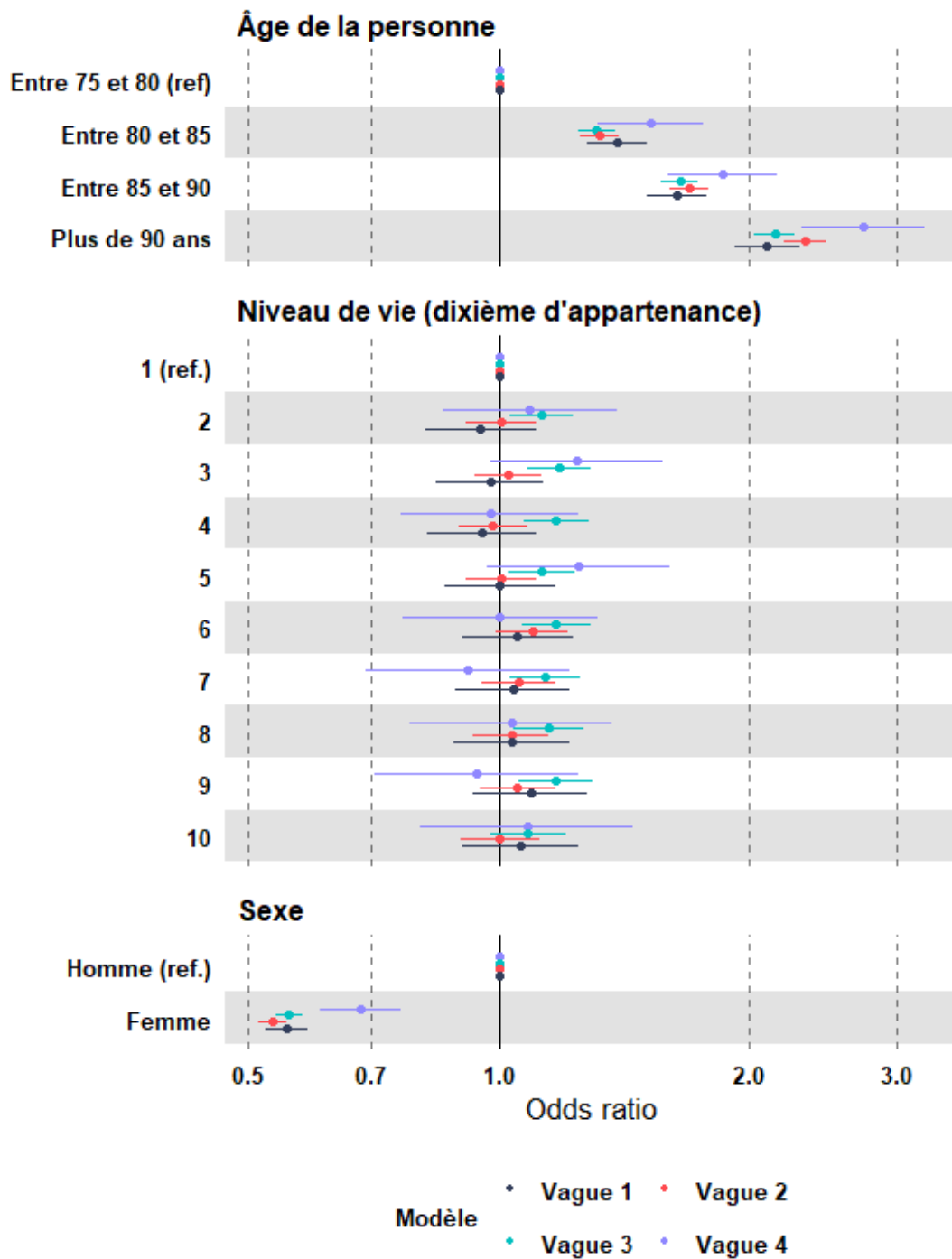
Autres contrôles: lieu de naissance et conditions de logement (voir annexe)

**Lecture** > Lors de la première vague, toutes choses égales par ailleurs, la mortalité est plus fréquente parmi les personnes hospitalisées ayant entre 70 et 75 ans relativement aux personnes hospitalisées de moins de 50 ans (*odds ratio* de 7).

**Champ** > Individus hospitalisés de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 18** • Résultats des modélisations du risque de décès (personnes âgées de 75 ans ou plus)



Autres contrôles: lieu de naissance et conditions de logement (voir annexe)

**Lecture** > Lors de la première vague, toutes choses égales par ailleurs, la mortalité est plus fréquente parmi les personnes hospitalisées ayant 90 ans ou plus relativement aux personnes hospitalisées ayant entre 75 et 80 ans (*odds ratio* de 2,1).

**Champ** > Individus hospitalisés de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

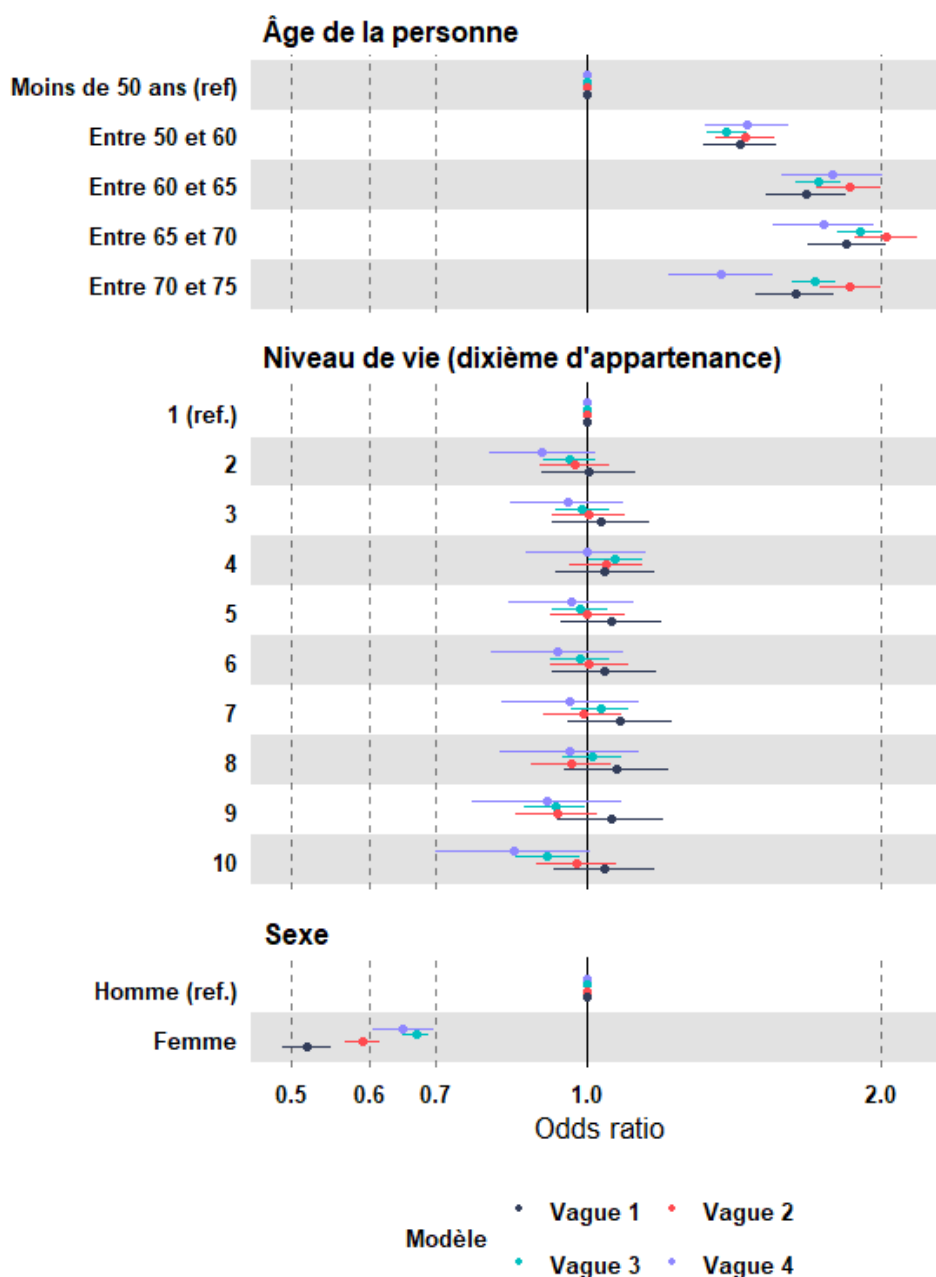
### Analyse des passages en soins critiques

Le modèle statistique confirme les principaux résultats de l'analyse descriptive, à savoir une augmentation du risque de passage en soins critiques qui atteint son maximum chez les 65-70 ans, puis décroît par la suite (graphiques 19 et 20). L'effet âge n'est pas significativement différent entre vagues épidémiques, bien que les *odds ratios* semblent plus élevés lors de la première vague chez les individus âgés de 75 ans ou plus. On retrouve l'effet attendu du sexe, les hommes étant une fois encore



plus à risque, toutes choses égales par ailleurs. Le niveau de vie n'a pas d'effet significatif sur le risque de passage en soins critiques chez les moins de 75 ans. Chez les plus âgés, ce risque semble corrélé au niveau de vie au cours des première et deuxième vagues et plus modérément au cours de la troisième vague avec une légère augmentation du risque d'admission parmi les individus des derniers dixièmes de revenu. Cet effet disparaît au cours de la quatrième vague. On constate un effet du pays de naissance, plus modéré que pour le risque d'hospitalisation, avec un risque légèrement accru dans certaines populations, à la fois chez les plus jeunes et chez les plus âgés, particulièrement chez les personnes nées en Asie ([Annexe 3](#)). Les conditions de logement (le fait de vivre dans un logement social, ou dans un logement surpeuplé) ne sont pas corrélées avec le risque de passer en soins critiques, conditionnellement au fait d'être hospitalisé ([Annexe 3](#)).

**Graphique 19** • Résultats des modélisations du risque d'admission en soins critiques (personnes âgées de moins de 75 ans)



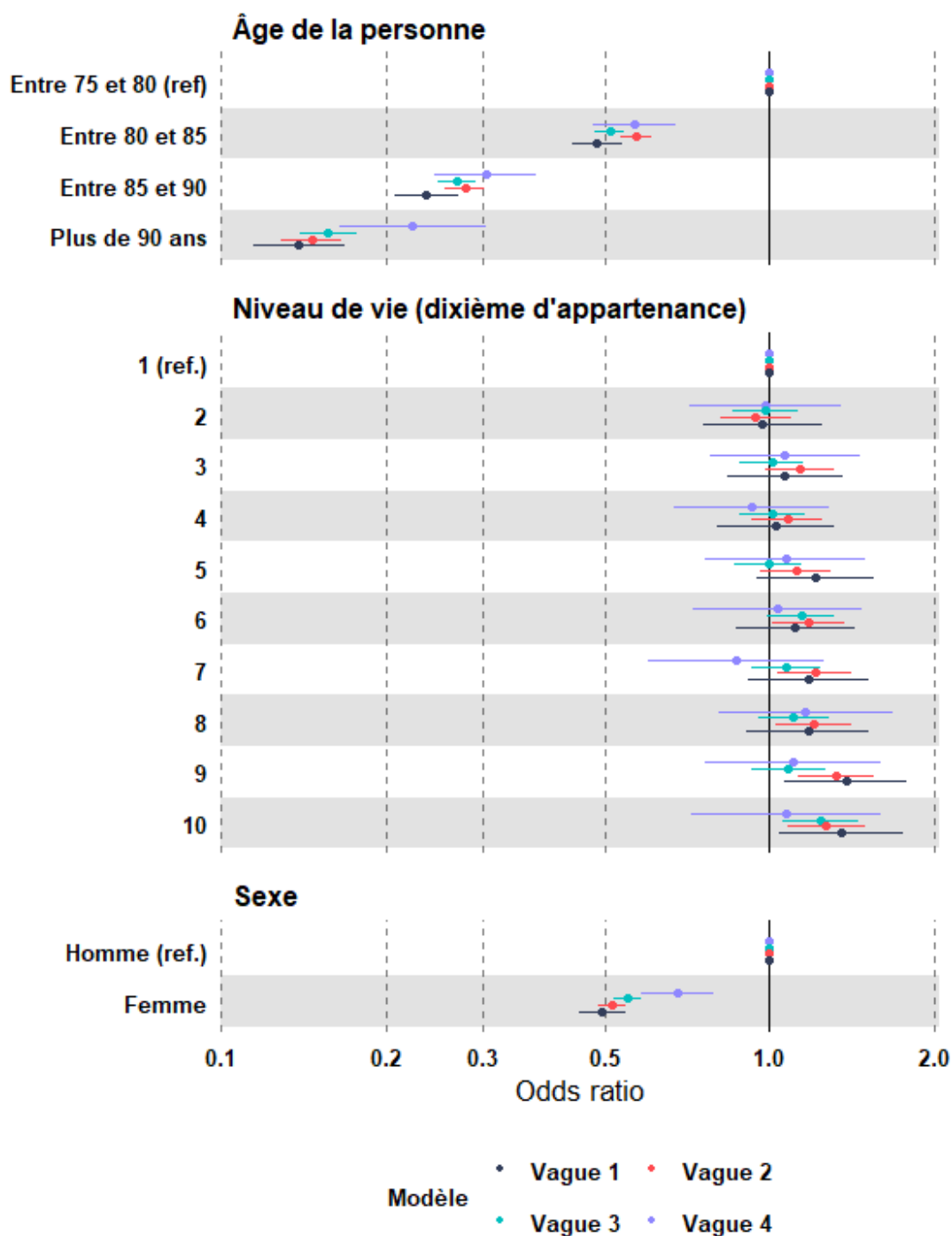
Autres contrôles: lieu de naissance et conditions de logement (voir annexe)

**Lecture** > Lors de la première vague, toutes choses égales par ailleurs, l'admission en soins critiques est plus fréquente parmi les personnes hospitalisées ayant entre 70 et 75 ans relativement aux personnes hospitalisées ayant moins de 50 ans (*odds ratio* de 1,6)

**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 20** • Résultats des modélisations du risque d'admission en soins critiques (personnes âgées de 75 ans ou plus)



Autres contrôles: lieu de naissance et conditions de logement (voir annexe)

**Lecture** > Lors de la première vague, toutes choses égales par ailleurs, l'admission en soins critiques est moins fréquente parmi les personnes hospitalisées ayant 90 ans ou plus relativement aux personnes hospitalisées ayant entre 75 et 80 ans (*odds ratio* de 0,14)

**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

## ■ DISCUSSION GÉNÉRALE

L'analyse en deux temps a permis de rechercher les facteurs socio-économiques associés au risque d'hospitalisation pour Covid-19 d'une part, et au risque de complications à l'hôpital (admission en soins critiques ou décès) d'autre part.

La grande majorité des études portant sur le risque de formes sévères de Covid-19 se concentre sur les facteurs cliniques, à savoir l'âge, le sexe et la présence de comorbidités. La présente étude aborde le sujet à travers le prisme des caractéristiques socio-économiques des personnes hospitalisées, au moyen d'un appariement inédit des données d'hospitalisation des patients atteints de Covid-19, et de données fiscales au niveau individuel. Les données d'hospitalisation ont pour principale limite l'absence d'informations sur la présence de comorbidités, dont la répartition inégale dans la population est indirectement captée par d'autres variables du modèle, notamment l'âge et le sexe, mais aussi le niveau de revenu ou encore le pays de naissance. De même pour la profession, qui est un facteur d'exposition différentielle au SARS-CoV-2, et qui ne figure pas dans les données fiscales.

L'âge reste le facteur de risque prépondérant de formes sévères de Covid-19 identifié dans cette analyse, indépendamment des autres variables étudiées. Un très net gradient s'observe pour les risques d'hospitalisation et de complication à l'hôpital (admission en soins critiques ou décès). La seule exception concerne le risque d'admission en soins critiques qui diminue chez les plus de 70-75 ans, en lien avec les pratiques médicales. Par ailleurs, l'effet de l'âge sur l'admission en soins critiques parmi les 75 ans ou plus varie légèrement au cours du temps, du fait notamment de l'évolution des modes de prise en charge. À cet égard, l'augmentation de l'âge d'admission en soins critiques entre la première et la deuxième vague, constatée par ailleurs, pourrait être attribuable à la diminution du recours à certaines pratiques de ventilation particulièrement invasives ([Naouri et Vuagnat, 2022](#)).

D'une part, la proportion d'individus atteints de comorbidités les rendant vulnérables au Covid-19 (obésité, hypertension artérielle, diabète, maladies coronariennes, pathologies pulmonaires chroniques) augmente avec l'âge ([Dubost, et al., 2020](#)). D'autre part, l'âge est en soi un facteur aggravant de la pathologie, du fait notamment de défenses immunitaires réduites ([Gao, et al., 2021](#)). Une étude de cohorte menée parmi les patients hospitalisés en France a permis de montrer que l'effet de l'âge reste majeur, y compris après la prise en compte des comorbidités ([Semenzato, et al., 2021](#)). Dans cette étude, l'âge est donc principalement une mesure approximative de l'état de santé général des individus, qui conditionne leur vulnérabilité à l'infection. L'effet de l'âge sur le risque d'hospitalisation est cependant plus ou moins prononcé d'une vague à l'autre. Cela pourrait refléter une exposition différentielle au virus du fait de comportements individuels variant avec l'âge et au cours du temps (en termes de protection individuelle ou de recours à la vaccination notamment).

Le sexe masculin a lui aussi déjà été identifié comme un facteur de risque de formes sévères ([Williamson, et al., 2020](#) ; [Gao, et al., 2021](#)). Le risque d'hospitalisation plus élevé chez les hommes peut refléter le fait que les hommes ont pu être davantage exposés au virus, notamment par leur profession, étant donné qu'ils ont continué, plus que les femmes, à travailler à l'extérieur y compris en période de confinement ([Boring et Moroni, 2021](#)). Cependant, les hommes restent plus durement touchés que les femmes conditionnellement au fait d'être hospitalisé, ce qui reflète essentiellement la prévalence différente des facteurs de risque, mais également le fait que le sexe pourrait être en lui-même un facteur de risque de forme sévère ([Nguyen, et al., 2021](#)), notamment du fait de processus hormonaux ou immunitaires spécifiques ([Gao, et al., 2021](#)).

L'effet des conditions de logement est un résultat majeur de cette étude. Le risque d'hospitalisation est plus élevé pour les personnes résidant dans des logements surpeuplés et pour les occupants du parc social. Plus précisément, le risque d'hospitalisation augmente à mesure que diminue la surface disponible par habitant et il est plus élevé pour les occupants du parc social, indépendamment de la surface disponible. Les conditions de logement n'affectent cependant pas le risque de complication à l'hôpital, ce qui confirme qu'elles captent probablement un risque accru d'exposition au virus dans des espaces confinés où les contacts sont plus fréquents et les gestes barrières plus difficiles à respecter. L'effet propre de l'appartenance au parc social pourrait refléter la densité des logements autour, ou encore la qualité intrinsèque des logements (aération). Si l'existence de corrélations entre de mauvaises conditions de logement et un risque accru d'incidence de Covid-19 a déjà été démontrée, cela n'a été fait qu'au niveau local ; pas à partir de données individuelles exhaustives, ni sur une aussi longue période temporelle ([Ahmad, et al., 2020](#) ; [Deguen et Kihal-Talantikite, 2021](#)).

Le risque d'hospitalisation augmente régulièrement à mesure que décroît le niveau de vie. La surexposition des bas revenus est robuste à l'introduction de variables corrélées au revenu mais jouant un rôle plus direct dans l'exposition au virus, comme le fait de vivre dans un logement surpeuplé. Le revenu permet ici de capter un éventuel effet résiduel de conditions de vie et de travail associées à un risque accru d'exposition au virus, et non capté par les autres variables considérées. Il pourrait également capter une distribution inégale des comorbidités associées à des formes sévères de Covid-19, plus répandues parmi les bas revenus ([Dubost, et al., 2020](#) ; [Fouquet et Pollak, 2022](#)). Une étude de cohorte menée en France a d'ailleurs mis en évidence un risque accru d'hospitalisation et de décès associés au Covid-19 parmi les habitants des communes les plus défavorisées, qui persiste en contrôlant de la présence de comorbidités ([Semenzato, et al., 2021](#)). L'effet protecteur des hauts revenus contre le risque de décéder à l'hôpital se retrouve chez les individus de moins de 75 ans uniquement. Il est possible que la distribution inégale des comorbidités captée par l'effet revenu soit particulièrement manifeste chez les plus

jeunes, dont la santé peut être dégradée du fait de conditions de vie ou de travail défavorables. Chez les plus âgés, l'âge devient le principal indicateur de l'état de santé général. Parmi les admissions en soins critiques, on observe au contraire un risque accru d'admission pour les individus aux plus hauts revenus chez les 75 ans ou plus uniquement. Ce résultat, qui peut sembler contre-intuitif, souligne le fait que la population des individus pris en charge en soins critiques est particulière. Les plus âgés y sont moins fréquemment admis, mais les admis sont ceux qui bénéficieront le plus d'un passage des soins spécifiques à ces services, pas ceux dont l'état de santé est le plus dégradé.

L'effet du pays de naissance fait écho aux nombreuses études internationales ayant identifié l'origine ethnique comme un facteur de risque majeur d'infection et de développement de formes sévères de Covid-19 ([Pan, et al., 2020](#)), ce qui pourrait être en grande partie le reflet de contextes socio-économiques spécifiques, de pratiques et de comportements propres à certaines populations, ainsi que de la concentration de certaines comorbidités. À cet égard, dans le premier volet de l'enquête EpiCoV (mai 2020) [[Warszawski, et al., 2020](#)], le taux d'infection qui apparaissait à première vue plus élevé chez les personnes immigrées, ne l'est plus après la prise en compte des conditions de vie socio-économiques et d'habitat. Il le reste cependant dans le second volet (juillet 2020) [[Warszawski, et al., 2021](#)], malgré la prise en compte de ces caractéristiques. Dans cette étude, l'effet du pays de naissance est marqué avec un risque accru d'hospitalisation pour les personnes nées hors d'Europe, et particulièrement celles nées en Afrique, qui persiste après prise en compte des autres variables étudiées, et qui pourrait refléter une inégale répartition des facteurs (i) d'exposition (profession, respect des gestes barrière) ou (ii) de vulnérabilité (présence de comorbidités) non captées par les variables disponibles. Ces premiers facteurs sont sans doute prépondérants, étant donné que l'effet du pays de naissance est beaucoup moins prononcé conditionnellement au fait d'être hospitalisé. À partir des troisième et quatrième vagues, l'effet du pays de naissance peut également refléter une inégale propension à se faire vacciner, comme constaté dans le troisième volet de l'étude EpiCoV (juillet 2021) [[Bajos, et al., 2022](#)].

L'analyse par vague épidémique souligne l'originalité de la quatrième vague et, dans une moindre mesure, de la troisième vague, par rapport aux précédentes. La quatrième vague se distingue notamment par une augmentation plus modérée du risque d'hospitalisation avec l'âge, qui se retrouve également au cours de la troisième vague, mais uniquement chez les 80 ans ou plus. Ces observations pourraient s'expliquer par l'effet protecteur de la vaccination, d'abord réservée aux plus âgés. Au milieu de la troisième vague (4 avril 2021), seulement 14 % de la population avait reçu au moins une dose de vaccin, mais cette proportion atteignait 61 % chez les personnes de 75 ans ou plus, priorisées dans la campagne de vaccination. Au milieu de la quatrième vague (mi-septembre), 64 % de la population, et 88 % des personnes de 75 ans ou plus, avaient reçu un schéma de primo-vaccination complet. Les gradients selon l'âge des risques d'hospitalisation au cours des troisième et quatrième vagues redeviennent similaires à ceux des deux vagues précédentes après l'intégration des taux de vaccination par âge à l'EPCI dans le modèle. La vaccination, ainsi introduite, apparaît comme un facteur protecteur contre le risque d'hospitalisation au niveau individuel, dans la mesure où le taux de vaccination par EPCI et par classe d'âge constitue un indicateur du statut vaccinal individuel. Cela est cohérent avec des résultats déjà montrés par ailleurs ([Suarez-Castillo, et al., 2021](#)). Du fait de la vaccination progressive par classe d'âge décroissante, l'effet protecteur de la vaccination n'est d'abord visible que parmi les plus âgés au cours de la troisième vague. Il s'observe ensuite plus largement au cours de la quatrième vague, bien que restant plus marqué chez les plus âgés. L'interprétation de ces résultats est cependant délicate, étant donné que l'information sur la vaccination disponible n'est pas individuelle. Le proxy utilisé pour la vaccination pourrait également capter un effet local, dans la mesure où la transmission du virus pourrait être réduite dans les zones aux populations fortement vaccinées, réduisant ainsi les risques d'hospitalisation, y compris pour les personnes non vaccinées. Le vaccin protège à la fois vis-à-vis du risque d'infection, du développement de formes sévères menant à l'hospitalisation et du risque de complications à l'hôpital ([Suarez-Castillo, et al., 2021](#)). Cependant, dans cette étude, l'impact de la vaccination sur le risque de parcours défavorable à l'hôpital n'a pas pu être évalué, faute de données disponibles au niveau individuel. Cela pourrait être exploré plus finement à l'aide de l'appariement des données SI-VIC-Fidéli à celles du référentiel VAC-SI, rassemblant les informations exhaustives sur les personnes vaccinées en France.

La quatrième vague se distingue aussi par un gradient de niveau de vie plus marqué dans le risque d'hospitalisation, avec un risque encore davantage accru parmi les bas revenus. Ce gradient n'est pas modifié par l'inclusion de la vaccination à l'EPCI, mais pourrait capter des déterminants de recours individuels à la vaccination (non captés par la variable agrégée). Le troisième volet de l'étude EpiCoV (juillet 2021) [[Bajos, et al., 2022](#)] a en effet mis en évidence que le recours à la vaccination croît avec le niveau de vie, les plus aisés étant aussi les plus vaccinés. Par ailleurs, le variant Delta, apparu à l'été 2021 et devenu majoritaire au cours de la quatrième vague, est plus contagieux et plus virulent que les variants ayant circulé auparavant. La particularité de la quatrième vague, qui touche plus spécialement les plus modestes, pourrait donc refléter la distribution inégale de la couverture vaccinale dans la population, ou la surexposition associée aux conditions de vie et de travail, exacerbée par la plus grande contagiosité du variant Delta. Les données disponibles ne permettent pas de dissocier ces deux effets.

## ■ CONCLUSION

L'exploitation inédite de l'appariement au niveau individuel entre les données d'hospitalisation pour Covid-19 et les données démographiques et fiscales éclaire la manière dont les caractéristiques de la population développant une forme sévère de Covid-19 diffèrent de celles du reste de la population.

Comme attendu, les principaux facteurs de risque d'hospitalisation et de complications à l'hôpital sont d'ordre clinique : âge et sexe. L'âge, dont l'effet est prépondérant, est une mesure approchée de l'état de santé et reflète très probablement des comorbidités plus nombreuses chez les plus âgés.

Les conditions de logement défavorables, le faible niveau de vie et l'appartenance à certaines populations nées à l'étranger (en Afrique et en Asie notamment) sont associés à un risque accru d'hospitalisation, très probablement du fait de l'inégale répartition du risque d'exposition au virus dans la population.

L'effet du surpeuplement est un résultat majeur de cette étude : le risque d'hospitalisation décroît en effet à mesure que la surface disponible augmente et il est plus élevé pour les occupants du parc social.

Le risque de décéder à l'hôpital reste légèrement moins élevé chez les personnes aux plus hauts revenus et parmi les populations d'origine européenne, ce qui reflète probablement une distribution inégale des comorbidités associées à des formes sévères de Covid-19.

L'originalité de la quatrième vague épidémique et, dans une moindre mesure, de la troisième par rapport aux précédentes, pourrait s'expliquer par des effets protecteurs différentiels des campagnes de vaccination. La vaccination semble aplanir le risque d'hospitalisation dans la population en limitant l'accroissement de ce risque avec l'âge, ce particulièrement pour les plus âgés, qui étaient aussi les plus vaccinés à l'époque. À l'inverse, le fait que le recours à la vaccination croisse avec le niveau de vie semble augmenter les inégalités en fonction du niveau de vie.

Cette étude permet de mieux comprendre la construction des inégalités sociales de santé dans le contexte de la crise sanitaire et d'objectiver le rôle de certaines conditions de vie. Celles-ci reflètent notamment l'inégale répartition de l'exposition au virus, mais aussi des facteurs de vulnérabilité aux formes sévères de Covid-19, particulièrement répandues dans certaines populations à l'état de santé dégradé. Enfin, ces résultats suggèrent que le moindre recours à la vaccination parmi les populations les plus pauvres, comparées aux plus aisées, accroît le différentiel de risque de développement de formes sévères.

## ■ ANNEXES

### Annexe 1. Spécificités des départements et régions d'outre-mer

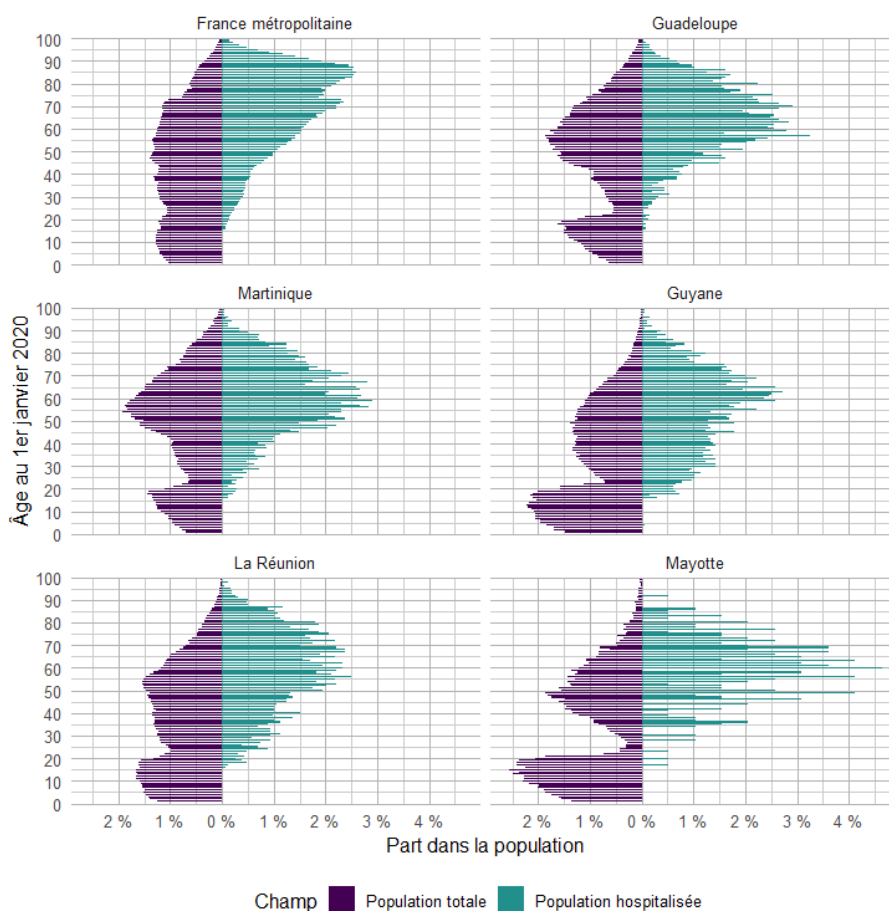
---

Les départements et régions d'outre-mer (DROM) ont été exclus du champ de l'analyse en raison des spécificités suivantes :

- La structure par âge de la population (résidente comme hospitalisée) diffère nettement entre la métropole et les DROM, que l'on regarde la population résidente ou la population hospitalisée (graphique 21). La population des DROM est sensiblement plus jeune que celle de la métropole : l'âge médian de la population (respectivement de la population hospitalisée) s'établit à 37 ans (resp. 59) dans les DROM, contre 42 ans (resp. 71) en métropole ;
- la population hospitalisée et résidant dans les DROM est nettement plus concentrée dans le bas de la distribution des niveaux de vie que la population hospitalisée et résidant en métropole (graphique 22) ;
- les individus résidant dans les DROM sont largement surreprésentés lors de la quatrième vague de l'épidémie (graphique 23) ;
- les données de niveau de vie ne sont que partiellement renseignées dans Fidéli, pour les individus résidant dans les DROM ;
- les données SI-VIC collectées dans les DROM sont de moins bonne qualité.

La conjonction de ces spécificités explique que conserver les DROM dans le champ de l'analyse aurait compliqué l'analyse toutes choses égales par ailleurs en nécessitant de mieux neutraliser les effets de composition. Il a donc été jugé préférable de restreindre l'analyse à la population métropolitaine.

**Graphique 21 • Structure de la population par âge, France métropolitaine et DROM**

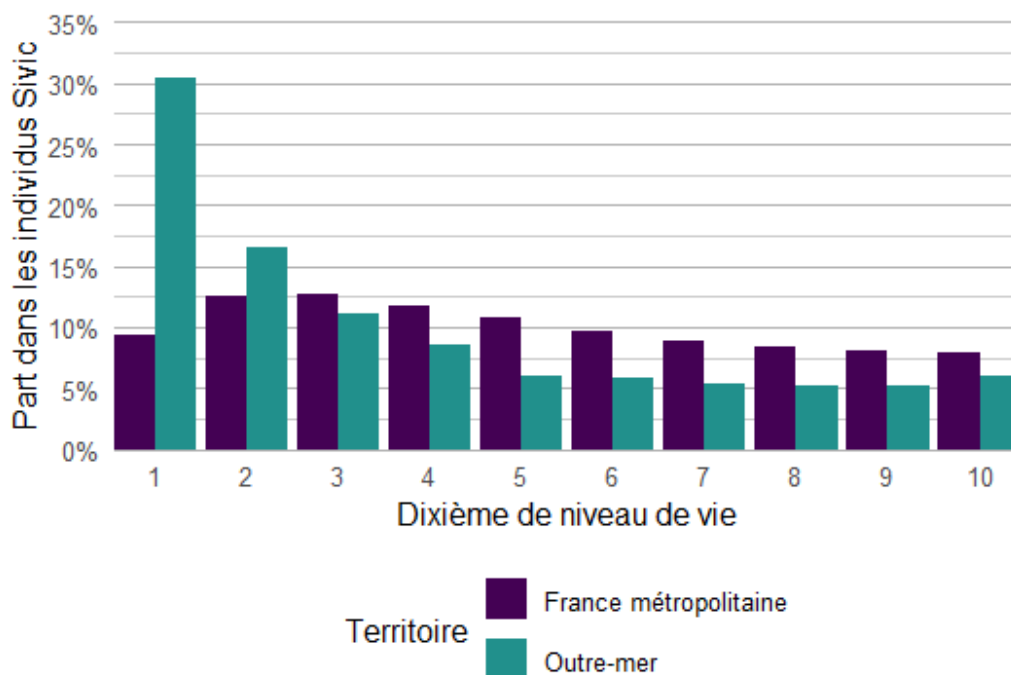


**Lecture** > Environ 1,3 % des individus résidant en France métropolitaine sont âgés de 50 ans ; environ 1 % des individus hospitalisés et résidant en France métropolitaine sont âgés de 50 ans. Environ 1,5 % des individus résidant à La Réunion sont âgés de 50 ans ; environ 1,9 % des individus hospitalisés à La Réunion sont âgés de 50 ans.

**Champ** > Individus résidant en France.

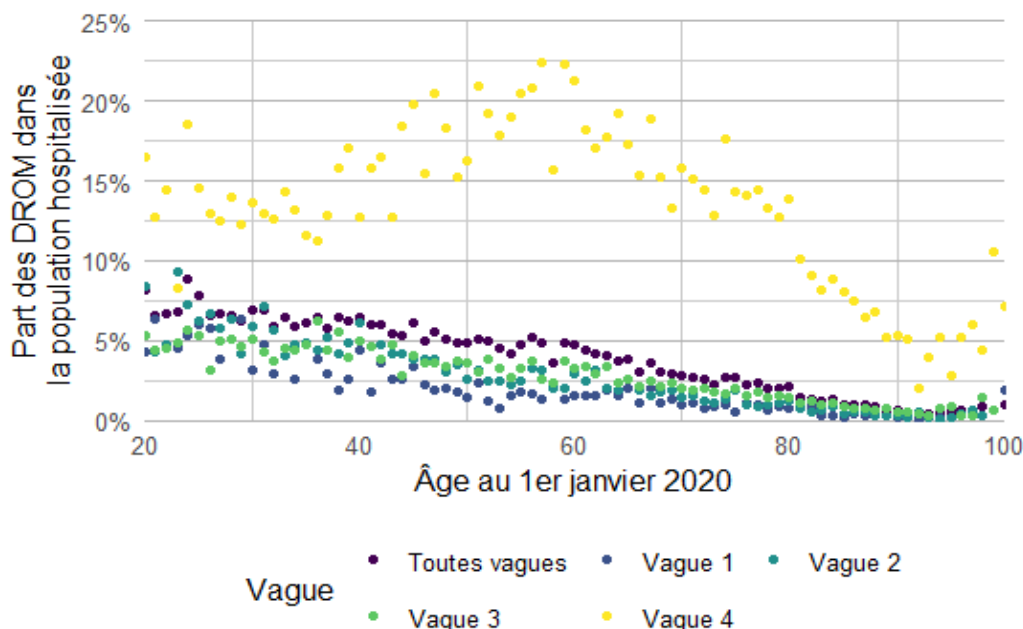
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 22** • Structure de la population hospitalisée par niveau de vie, France métropolitaine et outre-mer (La Réunion et Martinique)



**Lecture** > Environ 9,4 % des individus hospitalisés et résidant en France métropolitaine appartiennent au premier dixième de niveau de vie ; environ 30,4 % des individus hospitalisés et résidant dans les départements et régions d'outre-mer appartiennent au premier dixième de niveau de vie.  
**Champ** > Individus résidant en France.  
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 23** • Part des DROM dans les individus hospitalisés, par âge et par vague



**Lecture** > Environ 4,8 % des individus hospitalisés sur l'ensemble des vagues résident dans les départements et régions d'outre-mer (DROM). Environ 16,3 % des individus hospitalisés au cours de la quatrième vague résident dans les DROM.  
**Champ** > Individus résidant en France.  
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.



## Annexe 2. Description détaillée de l'algorithme d'appariement

---

L'appariement entre les données SI-VIC et la source Fidéli consiste à rechercher dans la table des individus de la source Fidéli un individu dont l'état civil est identique ou très similaire à celui des individus figurant dans les données SI-VIC. Cette recherche est réalisée de façon séquentielle : chaque étape cherche à appairer les enregistrements encore non appariés à l'issue de l'étape précédente. L'appariement en tant que tel comprend sept étapes ; une huitième étape traite les enregistrements non appariés.

Les variables utilisées pour l'appariement sont les suivantes : le nom de naissance, le nom de conjoint le cas échéant, les prénoms, la date de naissance, l'adresse et la commune de résidence. Il est à noter que le lieu de naissance n'est pas utilisé car il est trop peu renseigné dans les données SI-VIC.

Chaque étape d'appariement comprend des **variables d'appariement** et des **variables de validation**, et se déroule en trois temps :

- Les variables d'appariement (également appelées variables de blocage) sont utilisées pour réaliser la jointure entre les données SI-VIC et la table des individus de la source Fidéli. Exemple : pour un patient SI-VIC, on recherche les individus de la source Fidéli ayant exactement le même nom de naissance, le même premier prénom, et la même année de naissance.
- Les variables de validation sont utilisées pour vérifier que l'individu Fidéli retrouvé est bien le bon, et pour départager les individus Fidéli lorsque plusieurs d'entre eux répondent aux critères d'appariement (échos multiples). Pour ce faire, on détermine si l'état civil de l'individu Fidéli vérifie des conditions logiques restrictives (qui sont définies à partir des variables de validation), puis on calcule un score dont la valeur est d'autant plus élevée que l'individu Fidéli a vérifié des conditions restrictives. Le score est nul lorsque l'état civil de l'individu Fidéli ne vérifie aucun des filtres de validation. Les variables utilisées pour définir les critères de validation sont par définition différentes des variables d'appariement.
- Exemple : le score vaut la somme de :
  - 1 000 si le nom d'épouse de l'individu Fidéli est identique au nom d'usage du patient SI-VIC ;
  - 100 si le jour et le mois de naissance ainsi que la commune de résidence de l'individu Fidéli sont identiques à celui du patient SI-VIC ;
  - 10 si seuls le jour et le mois de naissance sont identiques ;
  - 1 si seule la commune de résidence est identique.
- On obtient ainsi un score compris entre 0 (aucun critère de validation n'est vérifié) et 1 111 (tous les critères de validation sont vérifiés).
- Un appariement est réalisé entre un patient SI-VIC et un individu de la source Fidéli uniquement si le score le plus élevé n'a été obtenu que par un seul individu Fidéli, et si le score est strictement positif (donc si au moins un critère de validation est vérifié). Aucun appariement n'est réalisé lorsque deux ou plusieurs individus Fidéli distincts ont obtenu le score le plus élevé, ou si aucun individu Fidéli n'a un score strictement positif.
- Exemple 1 : aucun des individus Fidéli n'obtient un score de 1 000 ni de 100 ; un seul obtient un score de 10, et deux autres individus ont un score de 1. On apparie alors le patient SI-VIC avec l'individu Fidéli ayant obtenu le score de 10.
- Exemple 2 : aucun des individus Fidéli n'obtient un score de 1 000 ni de 100 ; deux individus Fidéli obtiennent un score de 10, et un autre individu obtient un score de 1. Aucun appariement n'est effectué, car l'algorithme n'est pas parvenu à départager les deux individus Fidéli ayant obtenu le score de 10.

## Annexe 3. Résultats additionnels des modèles

---

### Effets départementaux

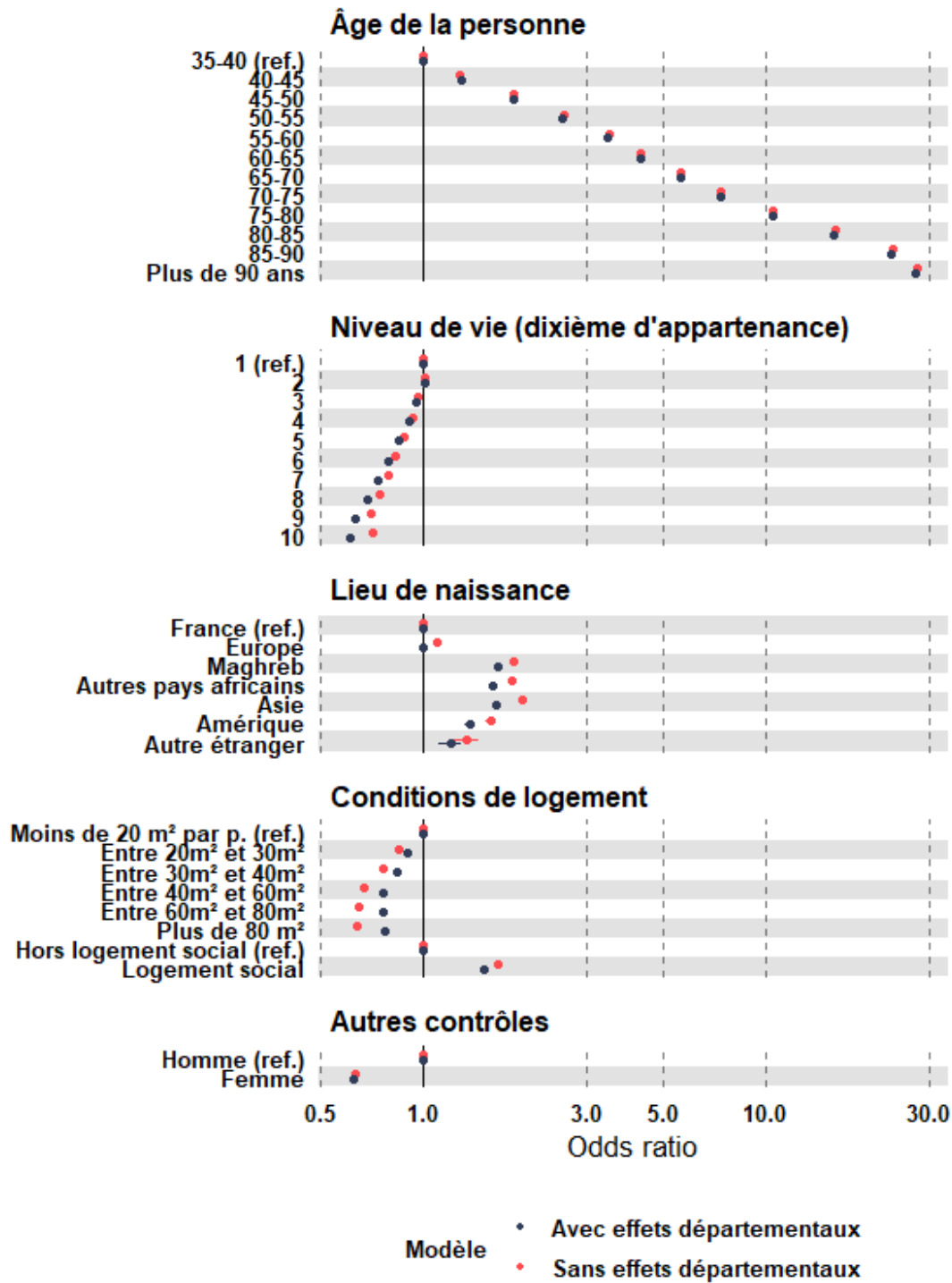
Les modèles de référence utilisent des effets départementaux. Ils sont fixes dans le temps. Lorsque le modèle est estimé sur l'ensemble de la période, un seul facteur par département est introduit. En revanche, dans les modèles par vague, ces effets départementaux sont spécifiques à chaque vague. Les régressions par vague permettent ainsi de contrôler de l'évolution de la circulation épidémique entre les vagues.

Les effets départementaux permettent de tenir compte de facteurs locaux non observés qui pourraient biaiser la mesure des facteurs de risque introduits dans le modèle statistique. En particulier, le taux d'incidence, qui varie entre les départements, est un facteur qui affecte le risque d'être hospitalisé. Sans l'introduction d'effets départementaux, la variabilité des taux d'hospitalisation pourrait être imputée à tort à une différence de structure de la population entre départements plutôt qu'à une circulation épidémique plus ou moins importante selon les départements.

Le graphique 24 illustre l'effet de l'introduction d'effets départementaux sur les résultats de la modélisation du risque d'hospitalisation sur l'ensemble des vagues. Les conclusions concernant les principaux facteurs de risque individuel, à savoir l'âge et le genre, sont robustes à l'introduction d'effets départementaux. De même, le gradient de niveau de vie n'est que marginalement modifié (il devient alors légèrement plus prononcé) par l'introduction d'effets départementaux. En revanche, l'introduction de ces effets départementaux réduit l'importance du pays de naissance et des conditions de logement. Dans les deux cas, cela illustre le fait que ces effets capturent des différences de structure de population ou de taille du parc de logements entre les départements.

Le graphique 25 illustre, quant à lui, la manière dont ces effets fixes permettent, dans un modèle estimé vague par vague, de contrôler de l'hétérogénéité des taux d'incidence, au niveau départemental, qui affectent le risque d'être hospitalisé. La similitude entre les effets départementaux par vague et le graphique 2 est très forte : la hiérarchie des effets départementaux correspond bien à celle de l'intensité épidémique.

**Graphique 24** • Effet de l'introduction d'effets départementaux sur les coefficients estimés



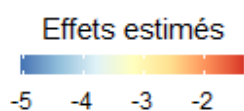
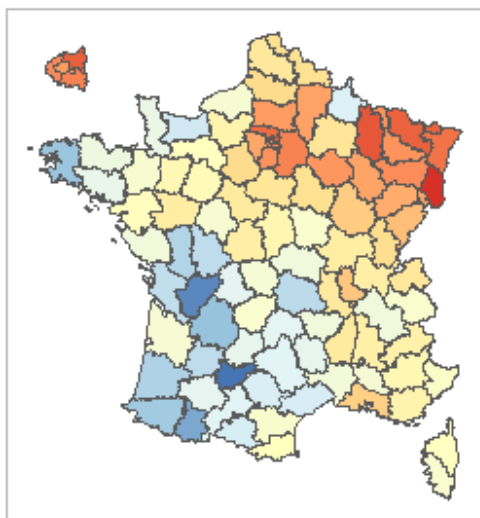
**Lecture** > En l'absence d'effets fixes départementaux, être né au Maghreb multiplie le risque d'hospitalisation par 1,83 par rapport au fait d'être né en France. En présence d'effets fixes départementaux, être né au Maghreb multiplie le risque d'hospitalisation par 1,66 par rapport au fait d'être né en France.

**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

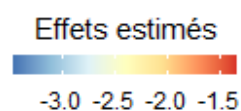
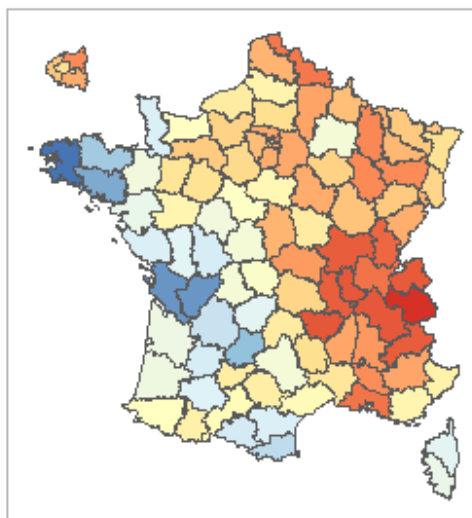
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 25** • Effets départementaux vague par vague

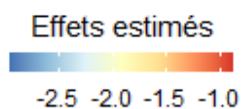
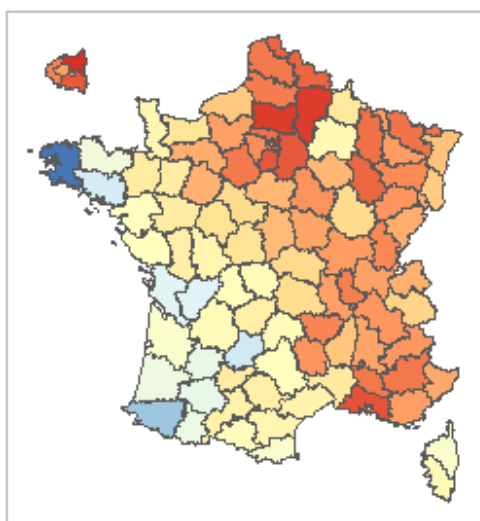
Vague 1



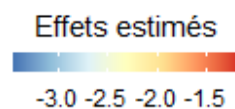
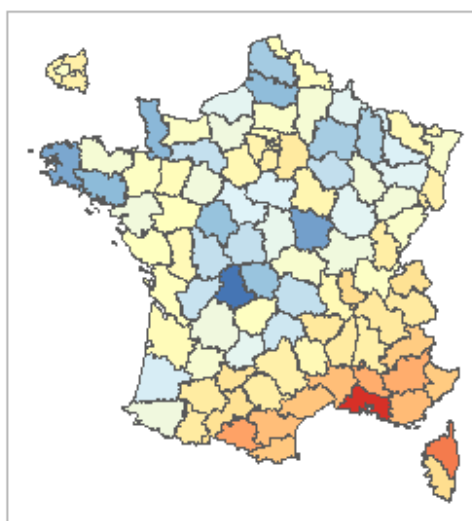
Vague 2



Vague 3



Vague 4

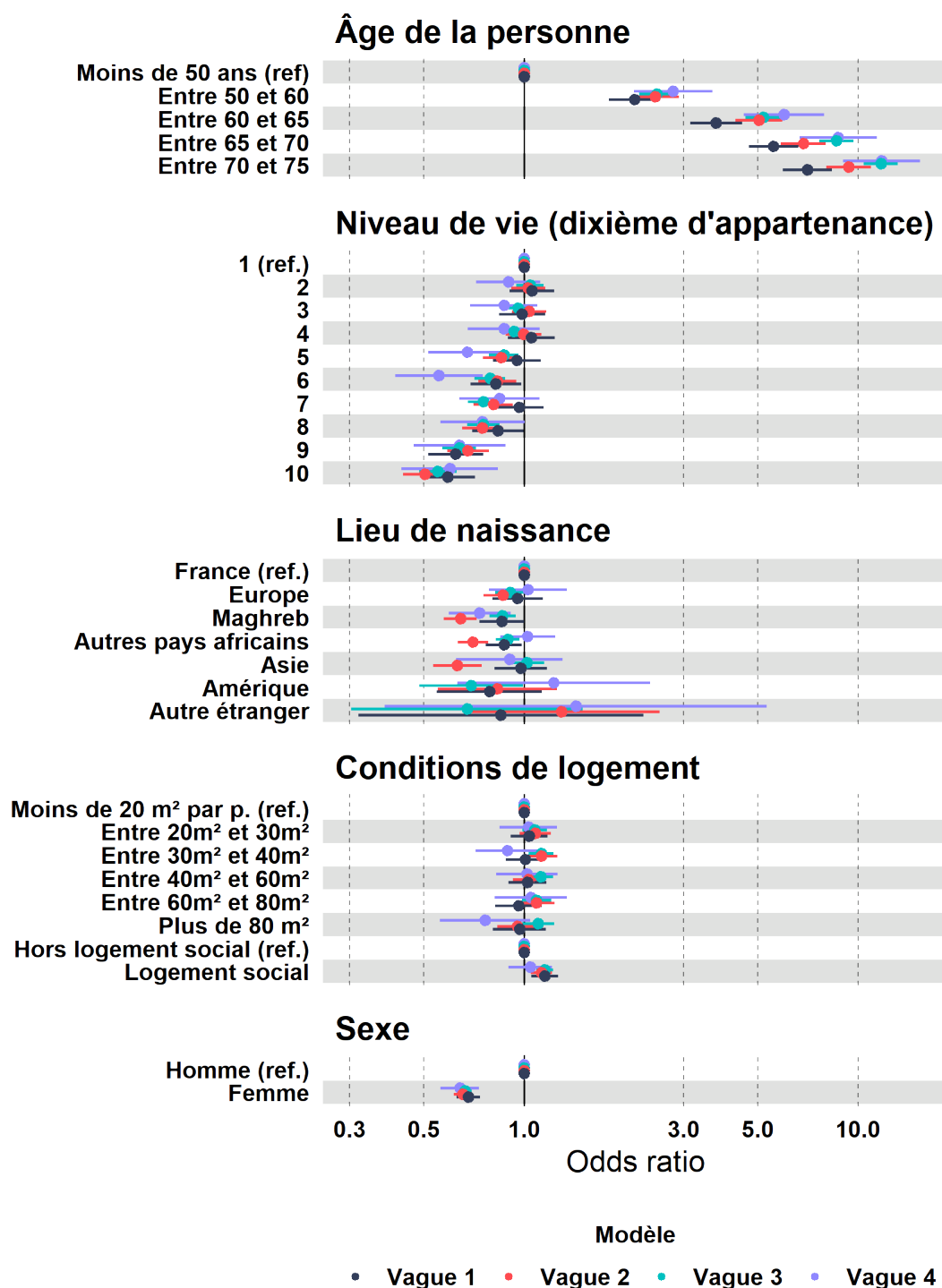


**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.  
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

### Résultats additionnels concernant les modèles décès et soins critiques

Les graphiques 17 à 20 présentent les principaux résultats concernant les événements postérieurs à l'hospitalisation. Les figures ci-dessous (26 à 29) proposent ces mêmes résultats en représentant également, cette fois, les résultats sur les variables de logement et de pays de naissance.

**Graphique 26** • Résultats des modélisations du risque de décès (personnes âgées de moins de 75 ans)

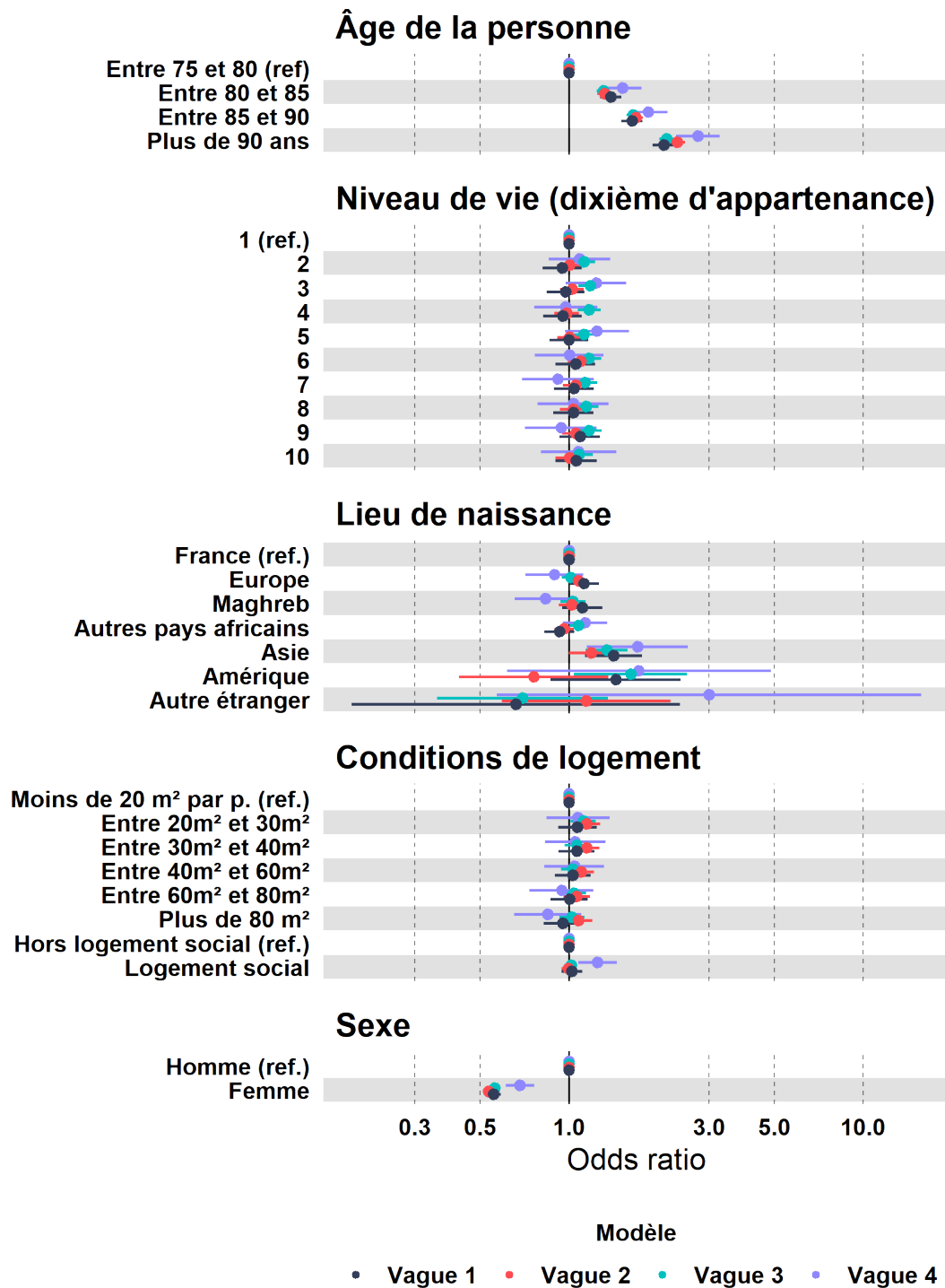


**Lecture** > Lors de la première vague, toutes choses égales par ailleurs, la mortalité est plus fréquente parmi les personnes hospitalisées ayant entre 70 et 75 ans relativement aux personnes hospitalisées de moins de 50 ans (*odds ratio* de 7).

**Champ** > Individus hospitalisés de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 27** • Résultats des modélisations du risque de décès (personnes âgées de 75 ans ou plus)

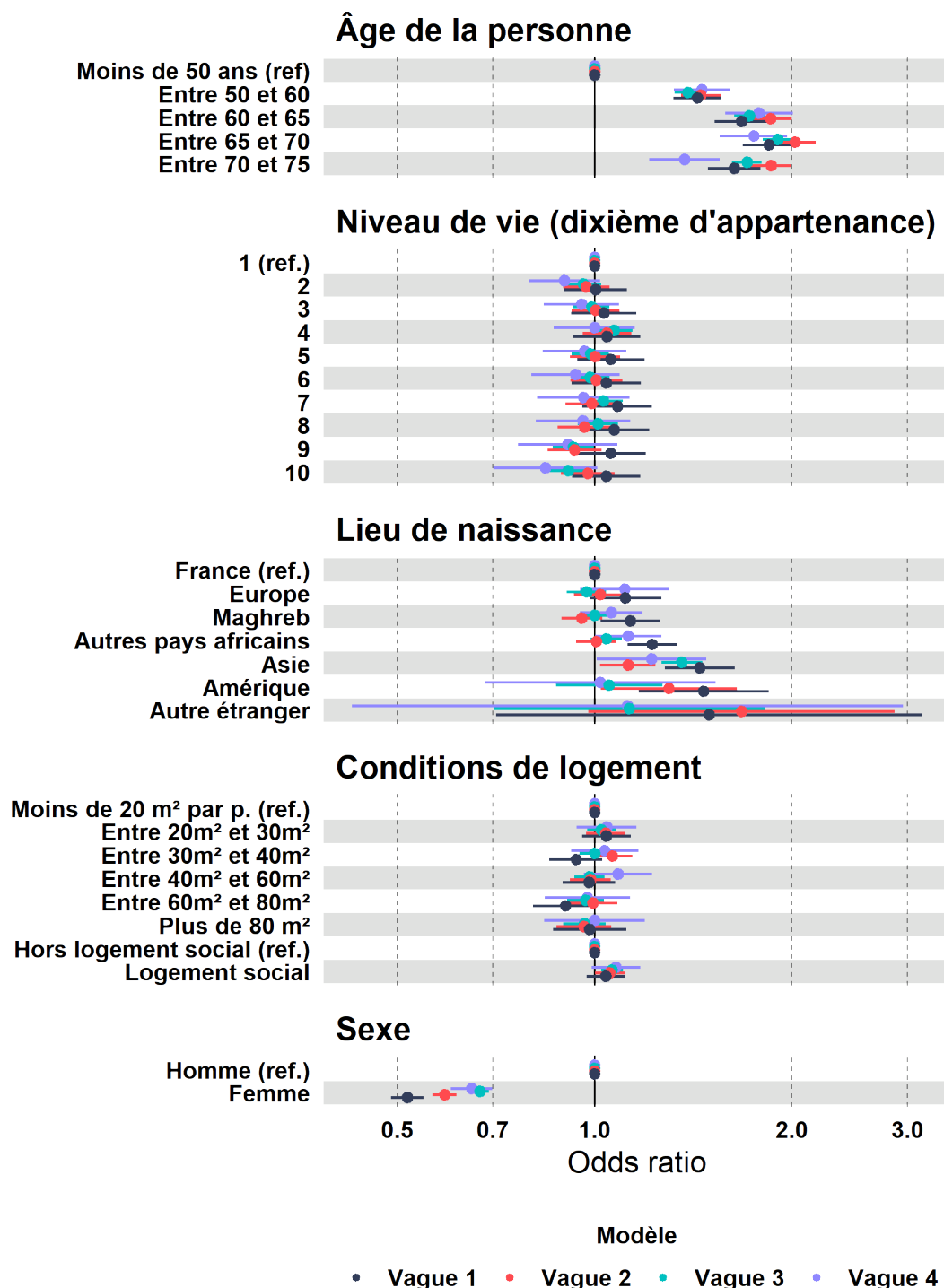


**Lecture** > Lors de la première vague, toutes choses égales par ailleurs, la mortalité est plus fréquente parmi les personnes hospitalisées ayant 90 ans ou plus relativement aux personnes hospitalisées ayant entre 75 et 80 ans (*odds ratio* de 2,1).

**Champ** > Individus hospitalisés de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 28** • Résultats des modélisations du risque d'admission en soins critiques (personnes âgées de moins de 75 ans)

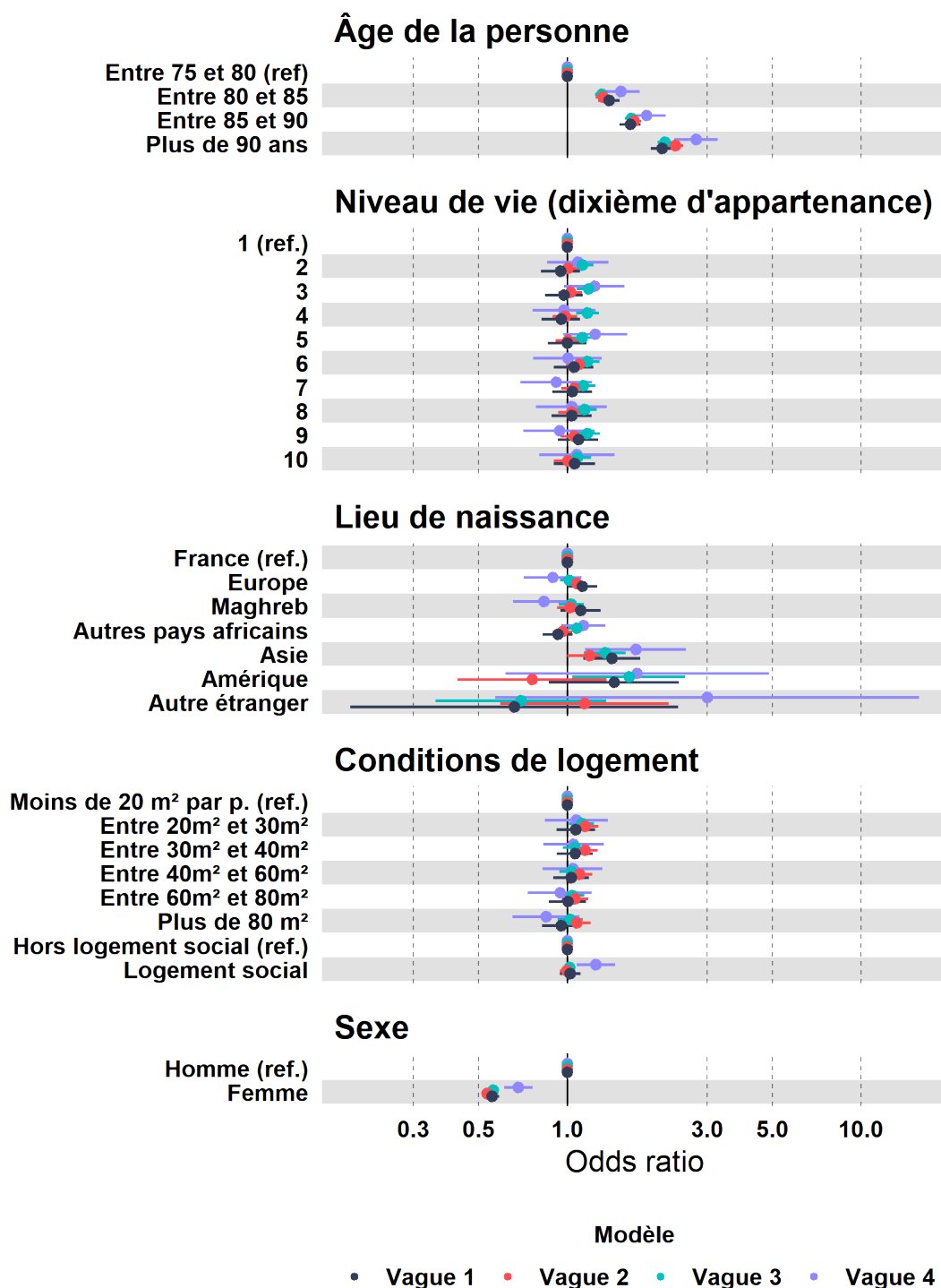


**Lecture** > Lors de la première vague, toutes choses égales par ailleurs, l'admission en soins critiques est plus fréquente parmi les personnes hospitalisées ayant entre 70 et 75 ans relativement aux personnes hospitalisées ayant moins de 50 ans (*odds ratio* de 1,6).

**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

**Graphique 29** • Résultats des modélisations du risque d'admission en soins critiques (personnes âgées de 75 ans ou plus)



**Lecture** > Lors de la première vague, toutes choses égales par ailleurs, l'admission en soins critiques est moins fréquente parmi les personnes hospitalisées ayant 90 ans ou plus relativement aux personnes hospitalisées ayant entre 75 et 80 ans (*odds ratio* de 0,14).

**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

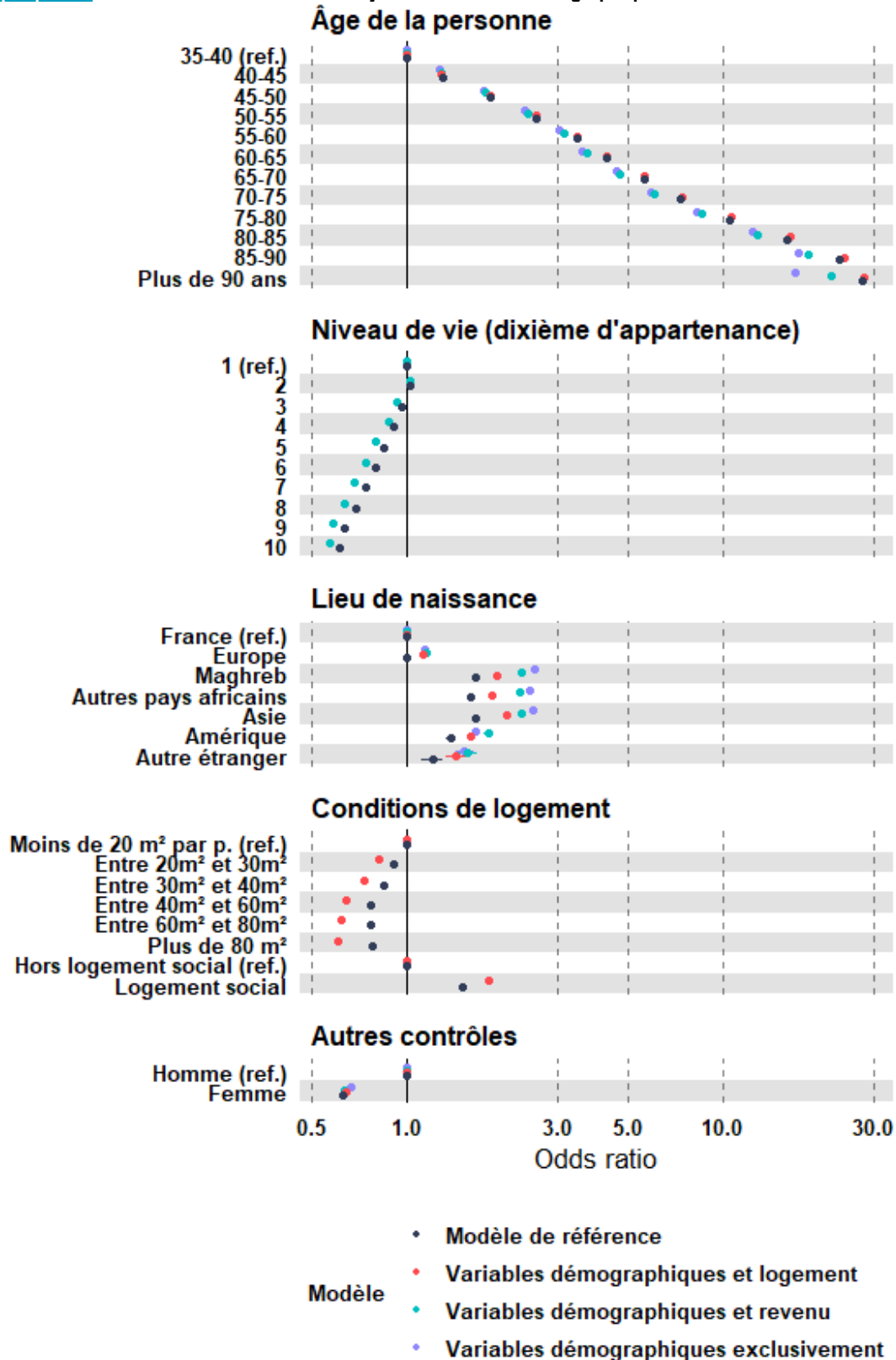
**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.



### Robustesse des effets mesurés

Le graphique 30 mesure l'effet de se restreindre à un sous-échantillon de facteurs. Le modèle de référence est comparé à des modèles comprenant seulement un sous-ensemble des facteurs de risques. Ceci permet d'évaluer la robustesse des estimations présentées dans le corps de l'étude. Dans l'ensemble, le fait d'introduire des facteurs socio-économiques dans le modèle de régression permet de corriger un biais du modèle sur les âges élevés. Sans variables socio-économiques, le gradient d'âge est moins prononcé, notamment aux âges élevés. La place dans la distribution de revenu et le lieu de naissance étant corrélés, ne prendre en compte qu'un de ces facteurs amène à surévaluer le rôle de ce facteur.

**Graphique 30** • Robustesse des modèles à l'ajout de variables démographiques



**Lecture** > Dans le modèle de référence, être né au Maghreb multiplie le risque d'hospitalisation par 1,65 par rapport au fait d'être né en France. Ce facteur multiplicatif s'élève à 1,91 dans un modèle prenant en compte les caractéristiques démographiques et les conditions de logement, à 2,3 dans un modèle prenant en compte les caractéristiques démographiques et le niveau de vie, et à 2,52 dans un modèle prenant en compte uniquement les caractéristiques démographiques.

**Champ** > Individus de 35 ans ou plus, appartenant à un ménage ordinaire résidant en France métropolitaine et dont le niveau de vie est connu et positif.

**Sources** > SI-VIC, Fidéli 2019 ; calculs Insee et DREES.

## Annexe 4. Retraitement des données SI-VIC

### Abréviations

Afin de pouvoir présenter quelques exemples synthétiques, nous proposons quelques abréviations portant sur les données hospitalières :

- HC : hospitalisation conventionnelle
- SC : soins critiques
- MCO : médecine, chirurgie, obstétrique et odontologie (rassemblant HC et SC)
- HAD : hospitalisation à domicile
- SU : soins d'urgences
- RAD : retour à domicile
- DC : décès

### Contexte

L'identifiant des enregistrements figurant dans les données SI-VIC, appelé « numéro SINUS », devrait en principe permettre d'identifier les individus de manière unique. En pratique, certains individus ont été enregistrés à plusieurs reprises dans les données SI-VIC sous différents identifiants. Cela s'explique notamment par le remplissage manuel des informations d'état civil des patients dans SI-VIC par les établissements, ou encore par les consignes de codage et leur évolution. Par exemple, les consignes données au début de la crise sanitaire étaient de ne pas détailler certaines informations individuelles, compliquant la recherche des individus concernés dans la base en cas de transfert ou de réhospitalisation. Par ailleurs, il n'avait pas été clairement précisé, au début de l'épidémie, qu'un patient devait se voir réattribuer le même numéro SINUS en cas de réadmission à l'hôpital, et les pratiques de codage dans ce cas de figure étaient variables d'un établissement à l'autre. L'appariement sur traits d'identité avec la source Fidéli permet de retrouver les multiples observations dans SI-VIC qui correspondent en réalité à un même individu, et donc de fiabiliser les informations sur les parcours hospitaliers. Dans l'ensemble, 3,8 % des individus présents dans Fidéli ont un ou plusieurs identifiants SI-VIC associés.

Pour ne garder que les parcours cohérents, un nettoyage des statuts d'hospitalisation de SI-VIC est ensuite nécessaire. D'une part, les statuts redondants ou inutiles dans un parcours ont été supprimés. D'autre part, les parcours associés aux individus hospitalisés hors du champ MCO (médecine, chirurgie, obstétrique et odontologie) ont été supprimés, afin de limiter la prise en compte des Covid incidents. En effet, les personnes hospitalisées pour une forme sévère de Covid-19 sont théoriquement prises en charge en MCO. Certaines personnes ont contracté le Covid-19 alors qu'elles étaient déjà hospitalisées en psychiatrie ou en soins de suite et de réadaptation (SSR) et ont été traitées sur place, sans être transférées vers des établissements de MCO. Étant donné que les consignes données pour le renseignement de SI-VIC par les structures de psychiatrie et de SSR ont été hétérogènes dans le temps et selon les régions, nous ne sommes pas en mesure de savoir s'il s'agissait de formes sévères de la maladie.

### Bilan du nettoyage

Plusieurs séjours peuvent être concernés par plusieurs règles de nettoyage. Au total, environ 40 000 numéros SINUS ont été supprimés. Le détail est donné dans le tableau ci-dessous.

Étapes	Explications	Nombres arrondis de statuts/identifiants concernés
<b>Problèmes d'appariement</b>	Les SINUS pour lesquels aucun identifiant Fidéli n'a été retrouvé ne sont pas pris en compte.	35 000 SINUS concernés
<b>Restriction du champ d'étude</b>	Les séjours commençant avant le 1 <sup>er</sup> mars 2020 ne sont pas pris en compte <sup>1</sup> .	1 200 SINUS concernés
<b>Statut manquant</b>	Si le statut manquant à la même date de début que le statut précédent ou suivant, le statut manquant est supprimé. Sinon, le séjour entier est supprimé, étant donné que ce statut pourrait être important dans le	700 statuts supprimés et 3 100 identifiants Fidéli concernés

	parcours (comme un passage en soins critiques par exemple).	
<b>Décès au milieu</b>	Si un décès se trouve en milieu de parcours, le parcours est entièrement supprimé.	100 identifiants
<b>Retour à domicile suivi d'un décès</b>	Le parcours est entièrement supprimé <sup>2</sup> .	500 identifiants Fidéli concernés
<b>Parcours sans MCO</b>	Les parcours se déroulant entièrement hors du champ MCO sont supprimés.	34 100 identifiants Fidéli concernés
<b>Les hospitalisations à domicile (HAD)</b>	La remontée des informations sur les patients Covid en HAD n'était pas prévue avant début 2021. Les informations n'étant pas complètes, les séjours sont tronqués de la HAD <sup>3</sup> .	100 statuts concernés
<b>Les soins d'urgences(SU)</b>	Certaines informations sur les patients Covid en soins d'urgence sont enregistrées dans SI-VIC, mais la remontée de ces informations n'était pas prévue initialement et n'est donc pas exhaustive sur la période <sup>4</sup> .	46 600 statuts concernés
<b>Retours à domicile consécutifs</b>	Seul le premier retour à domicile est conservé <sup>5</sup> .	3 800 statuts concernés
<b>Statuts d'hospitalisation consécutifs</b>	Seul le premier statut est conservé.	30 700 statuts concernés
<b>Retours à domicile en milieu de parcours</b>	On estime que si un statut d'hospitalisation a lieu moins de deux jours après un retour à domicile, ce retour n'en est pas vraiment un, et ce statut est supprimé. Autrement, ce retour à domicile est utilisé pour identifier deux séjours hospitaliers distincts. <sup>6</sup>	5 600 statuts concernés
<b>Retour à domicile au début</b>	Si un retour à domicile se trouve en première position du parcours, ce statut est supprimé.	2 200 statuts concernés
<b>Séjours non finis</b>	Les séjours non terminés (ne finissant ni par un décès, ni par un retour à domicile), 45 jours après la fin de la période d'étude sont supprimés.	2 900 identifiants Fidéli concernés

1. SI-VIC a été déployé pour le suivi des séjours Covid à partir du 16 mars 2020, avec un rattrapage des admissions antérieures, mais la qualité des premières observations reste moins bonne, du fait de l'immaturation des consignes de codage et de la situation sur le terrain, dans un contexte de saturation des hôpitaux où les équipes étaient par ailleurs très mobilisées par les soins.

2. Cette anomalie peut être due à un statut manquant entre les deux, ou à un mauvais codage d'un retour à domicile ou d'un décès.

3. Ex : HC-RAD-HAD-RAD devient HC-RAD. Si le statut avant HAD n'est pas un RAD, on remplace le statut HAD par un RAD.

4. En règle générale, le statut SU est supprimé. Néanmoins, si l'écart avec le statut suivant est de plus d'un jour, le SU est transformé en hospitalisation conventionnelle : un soin d'urgence ne devrait pas durer plus d'un jour et on estime que les soins prodigués correspondent à de l'hospitalisation conventionnelle.

5. Cette anomalie peut être due à des HAD codés en retours à domicile, ou au fait que, lors de la reconstitution des parcours avec identifiants communs, des parcours peuvent être codés en double et donc faire des doublons de statut.

6. Ce cas peut être présent pour deux raisons : (i) après un transfert vers un autre service ou établissement, un mauvais retour à domicile peut être codé ; (ii) si le patient effectue vraiment un autre séjour.

## ■ BIBLIOGRAPHIE

- Ahmad, K., Erqou, S., Shah, N. et al.** (2020, mai). Association of Poor Housing Conditions with COVID-19 Incidence and Mortality Across US Counties. *medRxiv*. <https://doi.org/10.1101/2020.05.28.20116087>.
- Bajos, N., Costemalle, V., Leblanc, S., et al.** (2022, février). Recours à la vaccination contre le Covid-19 : de fortes disparités sociales. DREES, *Études et Résultats*, 1222.
- Barhoumi, M., Jonchery, A., Lombardo, P., et al.** (2020, décembre). *Les inégalités sociales à l'épreuve de la crise sanitaire : un bilan du premier confinement*. Dans *France portrait social*. Paris : France, Insee, coll. Références.
- Boring, A., Gloria M.** (2021, mai). *Turning back the clock: Beliefs about gender roles during lockdown*. *The Economist*.
- Costemalle, V., Courtejoie, N., Miron de l'Espinay, A., et al.** (2020, décembre). À la veille du deuxième confinement, le système de dépistage détectait plus de la moitié des personnes infectées par la Covid-19. DREES, *Les Dossiers de la DREES*, 72.
- Courtejoie, N., Dubost, C.-L.** (2020, octobre). Parcours hospitalier des patients atteints de la Covid-19 lors de la première vague de l'épidémie-Ministère des Solidarités et de la Santé. DREES, *Les Dossiers de la DREES*, 67.
- Deguen, S., et Kihal-Talantikite, W.** (2021, décembre). Geographical pattern of COVID-19-Related outcomes over the pandemic period in France: a nationwide Socio-Environmental study. *International Journal of Environmental Research and Public Health* 18 (4), p. 1824.
- Dubost, C.-L., Pollack, C., Rey, S. (coord.)** (2020, juillet). Les inégalités sociales face à l'épidémie de Covid-19. État des lieux et perspectives. DREES, *Les Dossiers de la DREES*, 62.
- El Rais, H., Aflak-Kattar, M., Bleistein, L.** (2021, mai). Parcours hospitaliers des patients atteints de la Covid-19 de mars 2020 à janvier 2021. DREES, *Les Dossiers de la DREES*, 79.
- Fouquet, M., Pollack, C.** (2022, février). L'assurance maladie publique contribue fortement à la réduction des inégalités de revenu. DREES, *Études et Résultats*, 1220.
- Gao, Y., Mei D., Xiang D., et al.** (2021, février). Risk factors for severe and critically ill COVID-19 patients: a review. *Allergy* 76 (2), pp. 428-55.
- Gaudart, J., Landier, J., Huiart, L., et al.** (2021, avril). Factors associated with the spatial heterogeneity of the first wave of COVID-19 in France: a nationwide geo-epidemiological study. *The Lancet Public Health* 6 (4): e222-31.
- Goutte, S., Péran, S., Porcher, T.** (2020, juin). The role of economic structural factors in determining pandemic mortality rates: Evidence from the COVID-19 outbreak in France. *Research in International Business and Finance* 54: 101281. <https://doi.org/https://doi.org/10.1016/j.ribaf.2020.101281>.
- Guidet, B., Leblanc, G., Tabassome S., et al.** (2017, octobre). Effect of Systematic Intensive Care Unit Triage on Long-term Mortality Among Critically Ill Elderly Patients in France: A Randomized Clinical Trial. *JAMA* 318 (15): 1450-59. <https://doi.org/10.1001/jama.2017.13889>.
- Lefrancq, N., Paireau, J., Hozé, N., et al.** (2021, mars). Evolution of outcomes for patients hospitalised during the first 9 months of the SARS-CoV-2 pandemic in France: a retrospective national surveillance data analysis. *The Lancet Regional Health-Europe* 5 : 100087.
- Naouri, D., Vuagnat, A., et al.** (2022, mars). Covid-19 : prise en charge des patients en soins critiques au cours des trois premières vagues de l'épidémie. DREES, *Études et Résultats*, 1226.
- Nguyen, N.T., Chinn, J., De Ferrante, M. et al.** (2021, juillet). Male gender is a predictor of higher mortality in hospitalized adults with COVID-19. *PLoS One* 16 (7) : e0254066.
- Pan, D., Sze, S., Minhas, J.S., et al.** (2020, juin). The impact of ethnicity on clinical outcomes in COVID-19 : A systematic review. *EClinicalMedicine*, 23 : 100404. <https://doi.org/https://doi.org/10.1016/j.eclinm.2020.100404>.
- Pilkington, H., Feuillet, T., Rican, S., et al.** (2021, novembre). Spatial determinants of excess all-cause mortality during the first wave of the COVID-19 epidemic in France. *BMC public health* 21 (1), pp. 1-10.
- Semenzato, L., Botton, J., Drouin, J., et al.,** (2021, septembre). Chronic diseases, health conditions and risk of COVID-19-related hospitalization and in-hospital mortality during the first wave of the epidemic in France: a cohort study of 66 millions people. *The Lancet Regional Health - Europe* 8: 100158. <https://doi.org/https://doi.org/10.1016/j.lanepe.2021.100158>.
- Sesé, L., Nguyen, Y., Giroux, E., et al.** (2020, décembre). Impact of socioeconomic status in patients hospitalised for COVID-19 in the Greater Paris area. *European Respiratory Journal* 56 (6).

**Stammann, A., Heiss, F., McFadden, D.** (2016). *Estimating fixed effects logit models with large panel data*. Conference Paper.

**Suarez-Castillo, M., Khaoua, H., Courtejoie, N.** (2021, décembre). Estimation de l'efficacité vaccinale contre les formes graves de Covid-19 : exploitation des appariements SI-VIC-SI-DEP-VAC-SI de février à novembre 2021. DREES, *Les Dossiers de la DREES*, 90.

**Wachtler, B., Michalski, N., Nowossadeck, E., et al.** (2020, octobre). Socioeconomic inequalities and COVID-19—A review of the current international literature. *Journal of Health Monitoring* 5(S7): e1-e12.

**Warszawski, J., Bajos, N., Vianney Costemalle, Leblanc, S., et al.** (2021, juillet). « 4 % de la population a développé des anticorps contre le SARS-CoV-2 entre mai et novembre 2020 ». DREES, *Études et Résultats*, 1202.

**Warszawski, J., Bajos, N., Meyer, L., et al.** (2020, octobre). En mai 2020, 4,5 % de la population en France métropolitaine a développé des anticorps contre le SARS-CoV-2. Premiers résultats de l'enquête nationale EpiCov. DREES, *Études et Résultats*, 1167.

**Williamson, E.J., Walker, A.J., Bhaskaran, K., et al.** (2020, juillet). Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 584 (7821): pp. 430-36.



**Les dossiers de la DREES**

N° 95 • mars 2022

---

Caractéristiques socio-économiques des individus aux  
formes sévères  
de Covid-19 au fil des vagues épidémiques

---

**Directeur de la publication**

Fabrice Lenglard

**Responsable d'édition**

Valérie Bauer-Eubriet

**ISSN**

2495-120X

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

14 avenue Duquesne - 75 350 paris 07 SP  
Retrouvez toutes nos publications sur [drees.solidarites-sante.gouv.fr](https://drees.solidarites-sante.gouv.fr) et nos données sur [www.data.drees.sante.fr](https://www.data.drees.sante.fr)

---