

Les facteurs de mortalité du Covid

1	MORTALITE DU COVID : DEUX FACTEURS IMPORTANTS	2
2	REPARTITION DES DECES EN FONCTION DE L'AGE	2
3	PRISE EN COMPTE DES COMORBIDITES	4
3.1	HYPOTHESE 1.....	4
3.2	HYPOTHESE 2.....	5
3.3	RESULTATS	6
4	REMARQUES SUR LES DONNEES ET LE RESULTAT	7
4.1	LES CLASSES D'AGE	7
4.2	LA COMORBIDITE	7
5	CONCLUSIONS	7
5.1	LES DONNEES	7
5.2	UNE MEILLEURE IDENTIFICATION DES PROFILS A RISQUE	8
6	ANNEXE : EXEMPLE D'UNE FONCTION A 2 VARIABLES	8

Résumé

La mortalité due au COVID est fonction de deux facteurs : l'âge et la comorbidité, ce dernier facteur étant peu mentionné, voire ignoré, dans les médias.

L'objectif de ce document est de mettre en évidence une loi (une fonction, un modèle) qui décrive la mortalité due au Covid en fonction des deux facteurs, puis d'utiliser ce modèle pour représenter et ordonner les différentes classes d'individus concernés.

En classant les groupes d'individus par proportion de décès croissante (Figure 4 et Tableau 7), les résultats obtenus donnent une image de la mortalité liée au COVID plus proche de la réalité que celle qui est généralement présentée et qui s'appuie essentiellement, voire uniquement, sur l'âge.

Au-delà de ces résultats, et en considérant les données disponibles dans les documents de référence ci-dessous, le papier s'interroge sur la façon dont les services compétents abordent le recueil, la gestion et le traitement de l'information.

Référence 1 : Le journal des femmes Santé

<https://sante.journaldesfemmes.fr/fiches-maladies/2622115-victimes-covid-19-coronavirus-age-moyen-deces-profil-femmes-hommes-jeune-personnalites-aujour-d-hui/>

Source : Chiffres et statistiques mis à jour le vendredi 27 novembre à 11h33 avec le point épidémiologique hebdomadaire du 26 novembre 2020 du Covid-19 de Santé publique France. Ce bilan est basé sur les données épidémiologiques de surveillance du COVID-19 (SARS-CoV-2) rapportées à Santé publique France jusqu'au 24 novembre.

Référence 2 : Santé publique France

COVID-19 Point épidémiologique hebdomadaire du 26 novembre 2020"

1 Mortalité du Covid : deux facteurs importants

La mortalité due au Covid est liée à **l'âge**. C'est le principal constat que diffusent en permanence les médias. Les articles de presse sur le thème "Peut-on confiner les seules personnes âgées?" n'abordent généralement que les aspects éthiques et humains, sans préciser à partir de quel âge et sans s'interroger sur l'intérêt réel d'une telle mesure.

Une autre cause de la mortalité, parfois mentionnée dans les médias, est la **comorbidité** liée à des pathologies diverses telles que les problèmes cardiaques, l'hypertension artérielle, le diabète, l'obésité, etc.

Or lorsqu'un phénomène a plusieurs causes, son étude rigoureuse (scientifique) doit prendre en compte ces différentes causes, même si la facilité de n'en retenir qu'une est parfois tentante. Or pour étudier la mortalité en tenant compte de ces 2 facteurs (âge et comorbidité), il est indispensable de disposer de données numériques recueillies par les organismes de santé. Ces données doivent être suffisamment **globales** et **détaillées** pour afficher une image fidèle du problème, préalable indispensable à son traitement.

En observant les statistiques disponibles dans les médias "autorisés" français, on constate que ni l'âge ni la comorbidité ne sont renseignés de façon globale et détaillée. De plus, si chacun de ces facteurs est renseigné (plus ou moins bien), le lien entre les deux n'est pas suffisamment établi et exploité.

L'objectif de ce papier est d'explicitier une loi (une fonction, un modèle) qui décrive la mortalité due au Covid en fonction des deux facteurs, puis d'utiliser ce modèle pour représenter les différentes classes d'individus concernés.

Pour ceux qui ne détournent pas le regard dès qu'on parle "mathématiques", l'exercice peut se résumer comme suit :

$M = g(\text{âge})$ La mortalité est une fonction (notée **g**) de l'âge

$M = h(\text{comorbidité})$ La mortalité est une fonction (notée **h**) de la comorbidité

$M = f(\text{âge, comorbidité})$ La mortalité est une fonction (notée **f**) de l'âge et de la comorbidité

La fonction "**f**" à 2 variables contient beaucoup plus d'information que les fonctions "**g**" et "**h**" réunies, car elle contient en plus le lien entre ces deux fonctions. Elle permet par exemple de préciser le taux de mortalité dans une classe d'âge en différenciant, dans cette classe, ceux qui présentent ou non une ou plusieurs comorbidité(s).

Remarque. Ceux qui n'ont pas bien compris et qui souhaiteraient comprendre peuvent faire un détour par l'annexe du §6.

2 Répartition des décès en fonction de l'âge

Les données ci-dessous proviennent du document de référence 1

Entre le 1^{er} mars et le 24 novembre 2020, **50 237 décès** de patients COVID-19 ont été rapportés à Santé publique France : 34 399 décès étaient survenus au cours d'une hospitalisation et 15 838 décès (données au 22 novembre) parmi des résidents en établissements d'hébergement pour personnes âgées (EHPA) et autres établissements médico-sociaux (EMS).

- 93% des cas étaient âgés de 65 ans ou plus.
- L'âge **médian** des décès est de 84 ans (50% ont moins de 84 ans et 50% ont plus de 84 ans).
- Les hommes représentent 55% des décès et les femmes 45%
- 65% des décès sont associés à une comorbidité (une hypertension artérielle dans 23% des cas, une pathologie cardiaque pour 35%).

La répartition des personnes décédées **en fonction de l'âge** "certifiés par voie électronique du 1er mars au 23 novembre en France (avec ou sans comorbidités)" est définie par les 5 classes d'âge suivantes :

Age	0-14 ans	15-44 ans	45-64 ans	65-74 ans	75 ans et plus
	0%	1%	7%	14%	78%

Cette répartition ne tient pas compte de l'information contenue dans l'âge médian qui est égal à 84 ans. Pour améliorer la qualité des données et donc de l'analyse, il est nécessaire de partager la dernière classe (75 ans et plus) en 2 classes (75 à 84) et (84 et plus), cette dernière contenant 50% des décès.

La classe	75 ans et plus	est remplacée par les classes	75 à 84	et	84 et plus
	78%		28%		50%

Les 6 classes d'âge utilisées dans l'exercice sont affichées dans le tableau suivant.

Classe d'âge	0-14 ans	15-44 ans	45-64 ans	65-74 ans	75-84 ans	84 ans et +	Total
Code	C1	C2	C3	C4	C5	C6	
% décès	0	1%	7%	14%	28%	50%	
nbre décès	0	502	3 517	7 033	14 066	25 119	50 237

Tableau 1

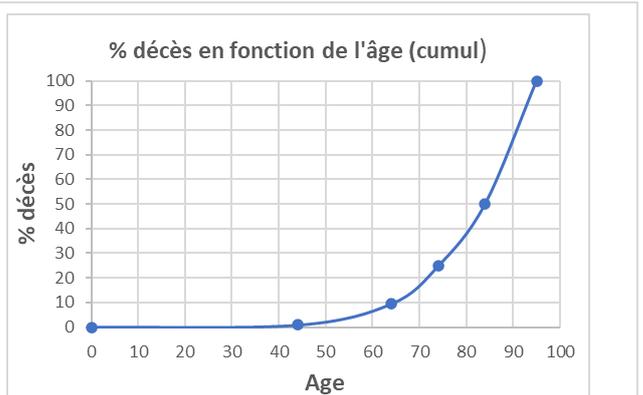
La 3^{ème} ligne du tableau (sans compter la ligne "Classe d'âge") affiche le nombre de décès dans chaque classe, calculé à partir du total (50 237) auquel on applique le % de la classe. Cette ligne représente la fonction "g" mentionnée au §1.

Remarque 1 : le document d'origine (Réf. 2) contient plusieurs répartitions des personnes décédées en fonction de l'âge. Ne sachant laquelle choisir, c'est la répartition affichée dans le document de Ref. 1 qui a été retenue.

Figure 1 : % décès en fonction de l'âge - Cumul

La courbe ci-contre est obtenue en représentant chaque classe par la borne supérieure de l'âge. La courbe représente donc le pourcentage cumulé de décès pour chaque âge. La valeur 95 ans a été retenue pour la dernière classe.

Par exemple, 22% des décès concernent des personnes dont l'âge est inférieur à 75 ans.



Nature du problème et des données

Les pourcentages de décès du tableau 1 peuvent être interprétés comme des probabilités dans le sens suivant :

La probabilité qu'une personne "k" appartienne à la classe d'âge C3 sachant qu'elle est décédée des suites du COVID est égale à 0,07. C'est une probabilité conditionnelle, la condition étant introduite par "sachant que", qui s'écrit de façon abrégée :

$$P_1(3) = \text{Proba}(k \in C3 / k \text{ décédée Covid}) = 0,07$$

Le signe \in signifie "appartient à" et la barre / signifie "sachant que"

Pour le **calcul de P₁**, la population retenue est constituée des 50 237 personnes décédées du tableau 1.

Cette probabilité est différente de la suivante:

Probabilité qu'une personne "k" décède des suites du COVID sachant qu'elle appartient à classe d'âge C3. C'est aussi une probabilité conditionnelle qui s'écrit :

$$P_2(3) = \text{Proba}(k \text{ décède du Covid} / k \in C3) = ?$$

Pour le **calcul de P₂**, la population à considérer est l'ensemble de la population française. Il est donc nécessaire d'inclure d'autres données dans l'analyse, en particulier celles du tableau ci-dessous qui donne le bilan démographique 2019 de la population française.

Population totale par sexe et âge au 1er janvier 2020, France, Bilan démographique 2019

<https://www.insee.fr/fr/statistiques/1892086?sommaire=1912926>

Population	0-14 ans	15-44 ans	45-64 ans	65-74 ans	75-84 ans	84 ans et +	67 063 703
Effectif	11 943 747	23 972 387	17 396 991	7 377 042	3 917 426	2 456 110	
%	17,8%	35,7%	25,9%	11,0%	5,8%	3,7%	1

Tableau 2

Outre les données numériques, le calcul de P₂ doit tenir compte d'éléments concernant le mode vie selon l'âge (les personnes âgées ont tendance à prendre moins de risques que les plus jeunes), le lieu de résidence, etc.

Le présent exercice ne concerne pas la probabilité P₂

3 Prise en compte des comorbidités

Certaines pathologies augmentent le risque de présenter une forme grave d'infection au Covid-19. Il s'agit par exemple des maladies cardiovasculaires, du diabète, des maladies chroniques respiratoires, de l'insuffisance rénale chronique, des cancers (personnes sous traitement), de l'obésité morbide, etc. Ces maladies sont regroupées sous le terme "comorbidité". Les personnes décédées suite au Covid peuvent présenter 0 ou 1 ou plusieurs comorbidités.

La répartition des personnes décédées avec ou sans comorbidités est globalement la suivante :

Avec comorbidités : 65% Sans comorbidités : 35%

Pour reprendre l'écriture abrégée de P_1 et P_2 ci-dessus, on peut écrire que :

Proba(une personne k présente au moins 1 comorbidité / k est décédée du Covid) = 0,65

Proba(une personne k ne présente pas de comorbidité / k est décédée du Covid) = 0,35

La distribution de la comorbidité par classes d'âge pour les données précédentes n'étant pas disponible dans les documents de référence, deux méthodes ont été utilisées pour déterminer cette distribution.

3.1 Hypothèse 1

La fonction h (comorbidité) du §1 est ici définie par les 2 valeurs (avec comorbidités = 65%) et (sans como. = 35%).

Hypothèse : Les deux pourcentages de comorbidité (65% et 35%) sont répartis sur chaque ligne proportionnellement au pourcentage de décès. On obtient alors le tableau suivant:

	Classe d'âge	0-14 ans	15-44 ans	45-64 ans	65-74 ans	75-84 ans	84 ans et +	Total
1	Code	C1	C2	C3	C4	C5	C6	
2	% décès	0	1%	7%	14%	28%	50%	
3	nbre décès	0	502	3 517	7 033	14 066	25 119	50 237
4	Avec como 0,65	0,0%	0,7%	4,6%	9,1%	18,2%	32,5%	65%
5	Sans como 0,35	0,0%	0,4%	2,5%	4,9%	9,8%	17,5%	35%
6	nbre dc avec como	0	327	2 286	4 572	9 143	16 327	32 654
7	nbre dc sans como	0	176	1 231	2 462	4 923	8 791	17 583

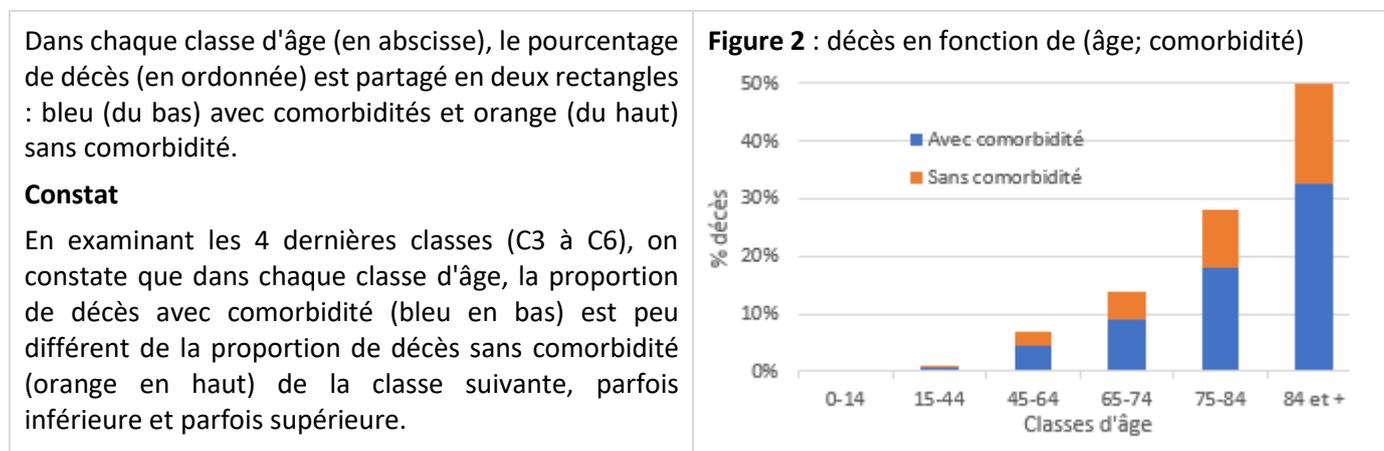
Tableau 3

A la ligne 4 (en bleu) et colonne 3 (C3), le pourcentage 4,6% est égal à : $7\% * 0,65$. Le nombre de décès dans cette classe (3 517) est alors réparti entre 2286 avec comorbidités et 1231 sans comorbidité. La somme de la ligne 4 est bien égale à 0,65.

La fonction f (âge, comorbidité) du §1 est définie ici par les lignes 4 et 5 en pourcentage ou les lignes 6 et 7 en quantité. C'est bien une fonction à 2 variables :

$f(45-64, \text{avec como}) = 4,6\%$ ou 2 286 décès et $f(45-64, \text{sans como}) = 2,5\%$ ou 1 231 décès.

La fonction "Proportion de décès" = f (âge, comorbidité) est représentée sur la figure suivante.



Ces résultats ne semblent pas corroborer cette phrase lue récemment dans un hebdomadaire "Tout le monde le sait désormais : la mortalité par Covid est liée à l'âge". **La réalité est peut-être moins simpliste que ce qu'en disent les médias.**

3.2 Hypothèse 2

Le tableau 7 du document de référence 2 (p 33) est intitulé "description des décès certifiés par voie électronique, avec une mention de COVID-19 dans les causes médicales de décès, en France, du 1er mars au 23 novembre 2020 (données au 24 novembre) (Source : Santé publique France, Inserm-CépiDC)".

Les données de ce tableau portent sur 20 766 décès en séparant pour chaque classe d'âge les décès avec et sans comorbidité. Ces données sont reproduites dans le tableau ci-dessous.

	Age	0-14	15-44	45-64	65-74	75 et +	Total
1	Classe	C1	C2	C3	C4	C5	
2	Total	3	143	1 495	2 995	16 130	20 766
3	Avec como	3	95	1 066	2 061	10 198	13 423
4	Sans como	0	48	429	934	5 932	7 343

Tableau 4

Les lignes 3 (Avec como en bleu) et 4 (Sans como en orange) de ce tableau représente la fonction "f" mentionnée au §1. Cependant elle est ici limitée à 5 classes d'âge et ne porte que sur un sous ensemble de la population considérée (20 237 décès au lieu de 50 237).

Comme pour le tableau 1, une classe est rajoutée pour prendre en compte le fait que la médiane du nombre de décès est égale à 84 ans. 50% des décès (soit 10 383) sont déplacés de la classe (75 et +) vers la nouvelle classe (84 ans et +). La même opération est effectuée pour les décès avec et sans comorbidité.

Age	0-14	15-44	45-64	65-74	75-84	84 et +	Total
Classe	C1	C2	C3	C4	C5	C6	
Total	3	143	1 495	2 995	5 747	10 383	20 766
Nbre avec	3	95	1 066	2 061	3 487	6 712	13 423
Nbre sans	0	48	429	934	2 261	3 672	7 343
% avec	0,01	0,5	5,1	9,9	16,8	32,3	65
% sans	0,0	0,2	2,1	4,5	10,9	17,7	35

Tableau 5

Pour finir le traitement de ces données, les pourcentages calculés dans le tableau 5 sont appliqués au nombre total de décès constatés au cours de la période considérée, c'est-à-dire 50 237 décès. On obtient alors le tableau suivant.

	Classe d'âge	0-14 ans	15-44 ans	45-64 ans	65-74 ans	75-84 ans	84 ans et +	Total
1	Classe	C1	C2	C3	C4	C5	C6	
2	% décès	0	1%	7%	14%	28%	50%	
3	nbre décès	7	346	3 617	7 245	13 903	25 119	50 237
4	Avec como 0,65%	0,01%	0,5%	5,1%	9,9%	16,8%	32,3%	0,65
5	Sans como 0,35%	0,0%	0,2%	2,1%	4,5%	10,9%	17,7%	0,35
6	nbre dc avec como	7	230	2 579	4 986	8 435	16 236	32 654
7	nbre dc sans como	0	116	1 038	2 260	5 469	8 882	17 583

Tableau 6

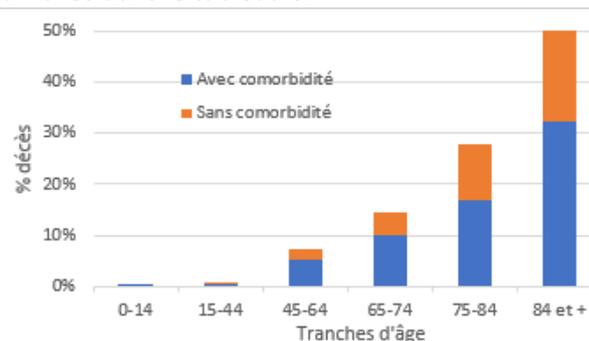
Les lignes 4 (Avec como en bleu) et 5 (Sans como en orange) de ce tableau représente la fonction "f" mentionnée au §1. Elle contient 6 classes d'âge et porte sur la totalité de la population considérée (50 237 décès). La figure ci-dessous donne une représentation graphiques des pourcentages affichés dans le tableau 6.

Figure 3 : décès en fonction de (âge; comorbidité)

En examinant les 4 dernières classes (C3 à C6), on fait le même constat qu'avec l'hypothèse 1 :

Dans chaque classe d'âge, le pourcentage de décès avec comorbidité est peu différent du pourcentage de décès sans comorbidité de la classe suivante, parfois inférieur et parfois supérieur.

Dans le cas de cette hypothèse 2, les différences sont légèrement plus importante que dans l'hypothèse 1.



3.3 Résultats

La figure 1 confirme que le nombre de décès augmente avec l'âge. Cependant, en croisant les 6 classes d'âge et les 2 valeurs du facteur "comorbidité" (avec et sans) on obtient un tableau contenant 12 cases, chacune d'entre elles contenant un couple (âge, comorbidité). Les résultats obtenus précédemment montrent que si l'âge est un facteur important, c'est aussi le cas de la comorbidité.

En effet, on constate que dans chacune des classes, C3 à C6 la proportion de décès avec comorbidité (bleu) est peu différente de la proportion sans comorbidité (orange) de la classe suivante, parfois inférieur et parfois supérieur.

Pour **visualiser ce résultat**, les classes d'âge C3 à C6 ont été partagées en 2 classes de décès, l'une sans comorbidité et l'autre avec. Les 6 classes d'âge sont donc remplacées par **10 classes de décès** qui sont représentées dans la figure 4 suivante.

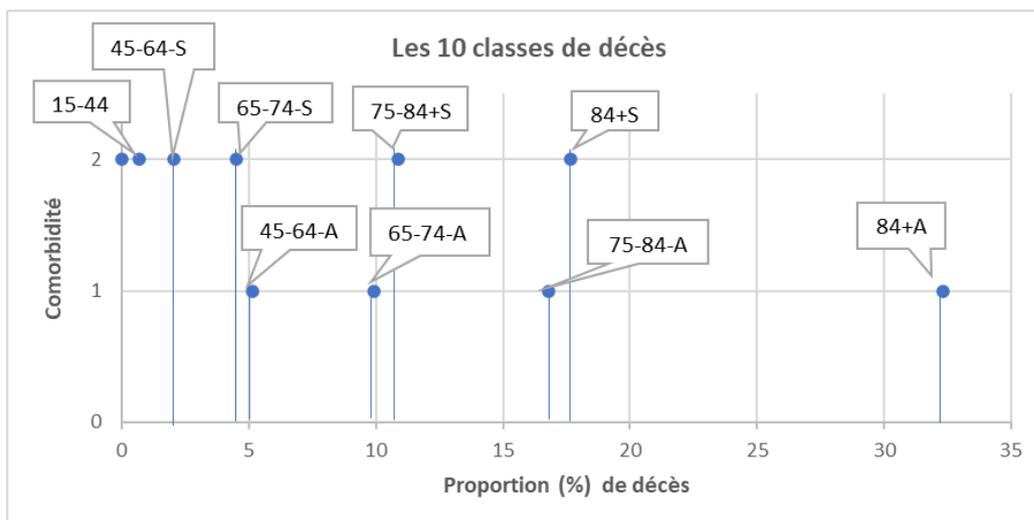


Figure 4

Dans ce graphique, l'abscisse représente la proportion de décès d'une classe et l'ordonnée permet de différencier les classes avec et sans comorbidité (1 = Avec et 2 = Sans). Le libellé de chaque classe de décès résume les 2 facteurs.

Exemple : **65-74-S** = classe d'âge 65-74 Sans comorbidité / **45-64-A** = classe d'âge 45-64 Avec comorbidités.

Au vu du graphique de la figure 4, les 10 classes de décès peuvent être ordonnées par proportion de décès croissante, en regroupant les classes qui sont voisines. On obtient alors le classement suivant.

Rang	Classes de décès	Proportion de décès (%)	Tranches d'âge concernées
1	0-14	0,01	0 à 14
2	15-44	0,7	15 à 44
3	45-64 sans	2,1	45 à 64
4	65-74 sans et 45-64 avec	4,5 à 5,1	45 à 74
5	65-74 avec et 75-84 sans	9,9 à 10,9	65 à 84
6	75-84 avec et 84 et + sans	16,8 à 17,7	75 à 84 et +
7	84 et + avec	32,3	84 et +

Tableau 7

Cette présentation de la mortalité liée au COVID est plus proche de la réalité que celle qui est généralement présentée et qui s'appuie essentiellement, voire uniquement, sur l'âge. On constate qu'à partir de la classe d'âge C3, les décès avec comorbidité(s) sont associés à la classe d'âge supérieure sans comorbidité.

L'âge n'est pas ici le critère qui permet de classer les groupes d'individus par proportion croissante de décès. Le classement obtenu ne correspond pas au seul critère de l'âge dont les tranches se chevauchent très largement.

Remarque : On obtient un graphique analogue à partir des résultats obtenus avec l'hypothèse 1 (Cf. §3.1).

4 Remarques sur les données et le résultat

4.1 Les classes d'âge

Les données initiales affichées dans le tableau 4 reposent sur 5 classes d'âge, la dernière étant [75 ans et +]. **Ce découpage :**

- d'une part **ignore** l'information très importante concernant **l'âge médian des décès** qui est égal à 84 ans;
- d'autre part crée une **dernière classe beaucoup trop large** qui ignore les différences à partir de 75 ans, **ce qui dégrade l'analyse de ces données** et réduit la pertinence des résultats.

Pour palier ces inconvénients, l'âge médian a été pris en compte en passant de 5 classes d'âge à 6.

A quel âge est-on un aîné, un sénior, un vieux, etc.

Si l'on retient le découpage en tranche d'âge retenu dans les documents de référence, les individus de 75 ans et plus sont assimilés et regroupés dans la même classe. On est donc vieux à partir de 75 ans, et pour faire des statistiques sur le COVID, **il ne semble pas utile de différencier les personnes de 75 ans et celles de 85 ou 90 ans**, sauf pour le calcul de l'âge médian.

Les médias, et en particulier les chaînes de télévision de grande audience, répètent sans relâche que la mortalité touche principalement les personnes âgées et qu'il faut protéger nos aînés, puis illustrent ces propos avec des reportages dans les EHPAD. Les EHPAD seraient-ils représentatifs des personnes de plus de 75 ans. Comme aurait dit Fernand Raynaud dans l'un de ses sketches, "y'aurait comme un défaut que ça m'étonnerais pas".

4.2 La comorbidité

Sur les 2 jeux de données utilisés, l'un porte sur l'âge de la "totalité" des personnes décédées (50 237) et fait l'objet de l'hypothèse 1. L'autre porte sur l'âge et la comorbidité d'un sous ensemble des décès (20 766) et fait l'objet de l'hypothèse 2.

Les résultats obtenus avec ces 2 hypothèses et présentés aux §3.1 et 3.2 sont presque identiques. Cela laisse penser que :

- L'hypothèse 1 selon laquelle "les deux pourcentages de comorbidité (65% et 35%) sont répartis dans les tranches d'âge proportionnellement au pourcentage de décès" est raisonnable;
- Le sous ensemble de décès utilisé avec l'hypothèse 2 est assez bien représentatif de la totalité des décès pour l'analyse réalisée.

Or selon l'hypothèse 1, la proportion de décès sans (resp. avec) comorbidité(s) n'est pas indépendante de l'âge. Il y a donc un lien entre l'âge et la comorbidité. Un découpage plus fin et des données plus riches le mettraient facilement en évidence.

5 Conclusions

5.1 Les données

Pour pouvoir analyser les données de la mortalité liée au COVID, il est nécessaire de disposer pour chaque décès des informations concernant l'âge et les pathologies à risques telles que les maladies cardiovasculaires, le diabète, les maladies chroniques respiratoires, l'obésité, etc.

Ces informations, probablement recueillies systématiquement dans les centres hospitaliers ou par les médecins qui constatent les décès qui surviennent hors de ces centres, doivent être regroupées et centralisées pour alimenter une **base de données unique**. Il est probable que ce processus de recueil global fonctionne actuellement, mais **cela n'apparaît pas** dans le document de référence 2 qui ne mentionne pas une telle base de données.

Tels que définis dans les documents de référence :

- le découpage en "tranches d'âge" ne semble pas assez fin pour donner une bonne représentation de ce critère;
- le facteur "comorbidité" renseigné au niveau de chaque patient ne doit pas se limiter aux 2 modalités, "avec" et "sans" (Cf. §3). Il faut bien sûr l'enrichir en détaillant les pathologies pour chaque décès.

L'exercice décrit dans les paragraphes précédents a été réalisé avec des données disponibles dans les documents de référence, en y ajoutant des hypothèses qui devraient être levées avec des données complètes. Les résultats

obtenus et présentés ici ne sauraient être une représentation exacte de la réalité. Ils servent essentiellement à montrer les tendances de l'impact dû au facteur comorbidité. Le poids de ce facteur (65% avec et 35% sans) est trop important pour être ignoré dans les études, les discours et les actions.

5.2 Une meilleure identification des profils à risque

Une meilleure identification de la distribution à 2 facteurs de la mortalité permet d'améliorer l'efficacité de la prévention en ciblant mieux les consignes de sécurité. Le gain est à la fois sanitaire (c'est évident) et économique, car si les consignes de sécurité sont mieux ciblées, elles ne pénalisent pas l'ensemble des acteurs économiques du pays.

De plus, laisser croire à la population que l'âge est le facteur essentiel peut déresponsabiliser les personnes qui ne se considèrent pas comme "âgées" et les amener à avoir un comportement à risque.

6 Annexe : Exemple d'une fonction à 2 variables

A l'issue d'une élection qui a eu lieu en 1973, un sondage a été réalisé sur un échantillon de 16 200 électeurs pour identifier leurs profils. Chaque électeur devait répondre à 2 questions :

- La catégorie socio-professionnelle (CSP)
- Le parti politique (PP) pour lequel il avait voté.

Les résultats du sondage sont présentés dans le tableau suivant.

CSP \ PP	PC	EG	PS	MR	MJ	DD	Total (g)
AGR	1,04	0,08	1,36	1,04	4,49	0,00	8,0
COM	0,80	0,24	1,85	1,28	3,53	0,32	8,0
LIB	0,41	0,48	0,48	1,83	3,33	0,27	6,8
EMP	3,04	0,72	4,12	2,33	6,80	0,90	17,9
OUV	11,65	1,26	8,50	2,52	6,61	0,94	31,5
INA	5,56	0,28	5,56	3,33	12,22	0,83	27,8
Total (h)	22,5	3,0	21,9	12,3	37,0	3,3	100,0

Parti communiste / Extrême gauche / Parti socialiste / Mouvement réformateur / Majorité / Divers droites
Agriculteur / Commerçant / Profession libérale / Employé / Ouvrier / Inactif

Ce tableau comporte **2 variables** (CSP et PP) et **3 fonctions**.

La **ligne du bas** définit la proportion des voix obtenue par chaque parti politique. C'est une fonction "h" à une variable (PP). Cette variable contient ici 6 modalités (PC, EG, PS, etc.).

$$h(PC) = 22,5\% \quad h(PS) = 21,9\% \quad \text{etc.}$$

La **colonne de droite** définit la proportion d'individus dans chaque catégorie socio-professionnelle. C'est une fonction "g" à une variable (CSP). Cette variable contient ici 6 modalités (AGR, COM, LIB, etc.).

$$g(AGR) = 8\% \quad g(EMP) = 17,9\% \quad \text{etc.}$$

Ces 2 fonctions donnent une information sur les partis politiques et sur les catégories socio-professionnelles. C'est utile, mais ce n'est pas suffisant.

Si l'on ignore la dernière ligne (fonction h) et la dernière colonne (fonction g), la **partie du tableau en bleu** définit la fonction "proportion de votants ou de voix" qui est une fonction "f" à 2 variables (CSP et PP).

$$f(AGR, PC) = 1,04 \quad f(EMP, MJ) = 6,80 \quad \text{etc.}$$

On constate aisément que cette fonction à 2 variables apporte beaucoup plus d'information que les deux fonctions à 1 variable réunies, car elle contient en plus le lien entre ces 2 variables, ce qui permet une analyse plus approfondie.

Les ouvriers ont d'abord votés pour le PC (11,65%), puis pour le PS (8,5%), etc. Les autres catégories socio-professionnelles ont d'abord voté pour la majorité (MJ). Les deux fonctions marginales (h : dernière ligne et g : dernière colonne) ne contiennent pas cette information.

De même, une fonction des 2 variables "âge" et "comorbidité" apporte beaucoup plus d'information que les deux fonctions traitant ces variables séparément. CQFD