

Les facteurs de mortalité du Covid

L'étude initiale a été réalisée avec des données arrêtées le 24 novembre 2020. Cette **mise à jour** est réalisée avec des données arrêtées le **19 janvier 2021**. Le présent document remplace celui du 01/01/21.

1	RESUME	2
2	MORTALITE DU COVID : DEUX FACTEURS IMPORTANTS	3
3	REPARTITION DES DECES EN FONCTION DE L'AGE	4
4	PRISE EN COMPTE DES COMORBIDITES	5
4.1	CHOIX DES DONNEES	6
4.2	ANALYSE SUR 5 CLASSES.....	6
4.3	ANALYSE SUR 6 CLASSES.....	7
4.4	RESULTATS	8
5	REMARQUES SUR LES DONNEES ET LE RESULTAT	9
5.1	LE SOUS ENSEMBLE ETUDIE	9
5.2	LES CLASSES D'AGE.....	9
6	CONCLUSIONS	9
6.1	LES DONNEES.....	9
6.2	UNE MEILLEURE IDENTIFICATION DES PROFILS A RISQUE	10

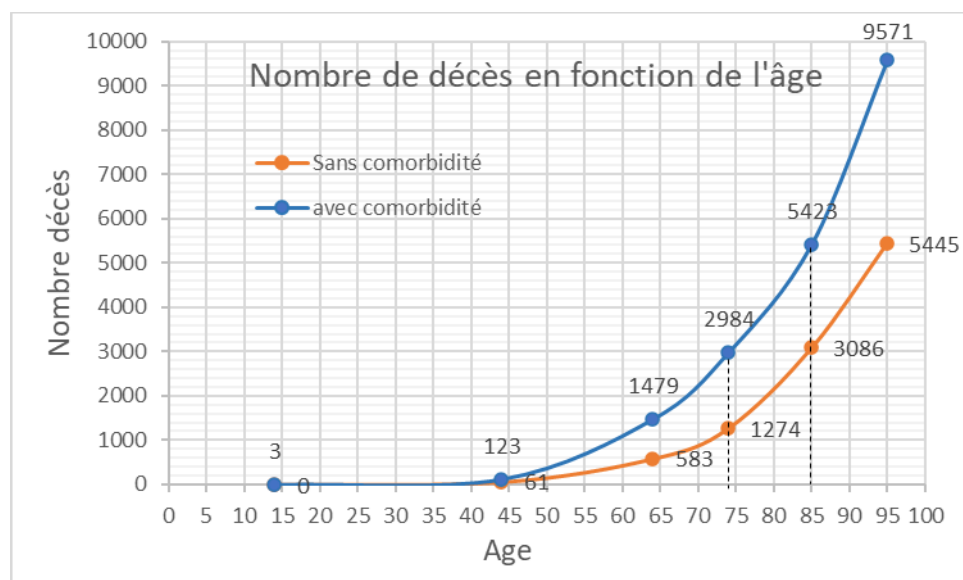
1 Résumé

La mortalité due au COVID est fonction de deux facteurs : l'âge et la **comorbidité** liée à des pathologies diverses telles que les problèmes cardiaques, l'hypertension artérielle, le diabète, l'obésité, etc. , ce dernier facteur étant peu mentionné, voire ignoré, dans les médias. A noter que **65% des décès** concernent des personnes présentant au moins une comorbidité.

L'objectif de cette étude est de mettre en évidence un **modèle** (une fonction) qui décrit la mortalité due au Covid en fonction des deux facteurs, puis d'utiliser ce modèle pour représenter et **comparer les différentes classes d'individus** concernés. Cette étude a été réalisée à partir de données extraites du "point épidémiologique hebdomadaire" du 21 janvier 2021 de "Santé publique France"

les **résultats de l'étude** 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. Ces **résultats** sont résumés dans le graphique ci-dessous dans lequel chaque classe d'âge (*) est représentée en abscisse par sa borne supérieure. Par exemple la classe d'âge "74-85" est représentée par l'âge 85 ans. Le graphique affiche 2 courbes.

Les **deux courbes** du graphique représentent en ordonnée le nombre de décès dans chaque classe d'âge. La courbe du dessus (en bleu) concerne les personnes présentant au moins une comorbidité et celle du bas (en orange), celles qui n'en présentent pas. Par exemple pour la classe d'âge '65-74', le nombre de décès avec comorbidité (resp. sans) est égal à 2984 (resp. 1274).



En examinant les deux courbes, on constate que dans chaque classe d'âge, le nombre de décès avec comorbidité (bleu) est peu différent du nombre de décès sans comorbidité (orange) de la classe suivante (\approx même ordonnée). **Présenter une comorbidité est donc équivalent à monter d'une classe d'âge.** L'âge n'est pas un paramètre suffisant pour décrire la mortalité.

La mortalité est bien une fonction à 2 variables, l'âge et la comorbidité qui est représentée ici par "avec" ou "sans". L'impact de cette deuxième variable (65% des décès) est trop important pour ne pas la mentionner et la traiter avec autant d'attention que l'âge.

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

(*) Classes d'âge : [0-14 ans], [15-44], [45-64], [65-74], [75-85[,]85 et +]

Introduction

La mise à jour de l'étude avec les données arrêtées le 19 janvier 2021 **confirme et précise les résultats obtenus** dans l'étude initiale réalisée avec les données du 24 novembre 2020.

Dans l'étude initiale, 2 jeux de données ont été utilisés. L'un porte sur l'âge de la "totalité" des personnes décédées (50 237) et a 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 étant presque identiques, l'hypothèse 1 a été supprimée de la mise à jour. Cela modifie la numérotation des paragraphes et de certains tableaux et figures.

2 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. 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 cette étude est d'expliciter 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
$M = f(\text{comorbidité}, \text{âge})$	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.

3 Répartition des décès en fonction de l'âge

Les données ci-dessous proviennent du document de référence 1.

Entre le 1er mars 2020 et le 19 janvier 2021, **71 342** décès de patients COVID-19 ont été rapportés à Santé publique France : 49 696 décès étaient survenus au cours d'une hospitalisation et 21 646 décès en établissements d'hébergement pour personnes âgées (EHPA) et autres établissements médico-sociaux (EMS) (données au 17 janvier 2021).

- 93% des cas étaient âgés de 65 ans ou plus.
- L'âge **médian** des décès est de 85 ans (50% ont moins de 85 ans et 50% ont plus de 85 ans).
- Les hommes représentent 54% des décès et les femmes 46%
- 65% des décès sont associés à une comorbidité. Le tableau suivant affiche des statistiques sur les maladies à risque.

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 85 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 85) et (85 et plus), cette dernière contenant 50% des décès.

La classe 75 ans et plus est remplacée par les classes 75 à 85 et 85 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	75-85 ans	85 ans et +	Total
Code	C1	C2	C3	C4	C5	C6	
% de décès	0	1	7	14	28	50	
Nbre décès	0	713	4 994	9 988	19 976	35 671	71 342

Tableau 1

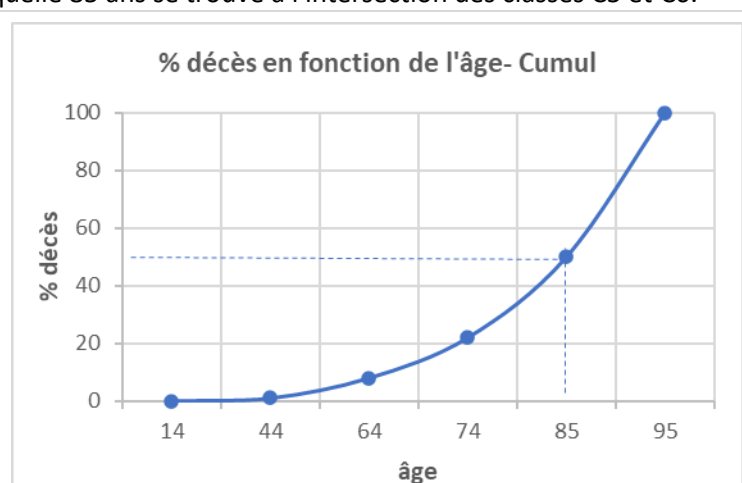
La 3^{ème} ligne du tableau (sans compter la ligne "Classe d'âge") affiche le nombre de décès dans chaque classe, calculé à partir du total (71 342) auquel on applique le % de la classe. Cette ligne représente la fonction "g" mentionnée au §1.

Remarque : L'âge médian des décès est de 85 ans, c'est-à-dire que 50% ont moins de 85 ans et 50% ont plus de 85 ans. C'est la raison pour laquelle 85 ans se trouve à l'intersection des classes C5 et C6.

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 le pourcentage cumulé de décès pour chaque âge. La valeur 95 ans a été retenue pour la dernière classe.

Par exemple, 50% des décès concernent des personnes dont l'âge est inférieur à 85 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 P_1** , la population retenue est constituée des 71 342 personnes décédées (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 P_2** , 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 P_2 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 P_2

4 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 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.

Le tableau suivant est extrait du document de référence 1.

Comorbidités	%	Comorbidités	%
Obésité	6	Hypertension artérielle	22
Diabète	16	Pathologie neurologique	9
Pathologie respiratoire	13	Pathologie rénale	13
Pathologie cardiaque	35	Immunodéficience	2

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 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

4.1 Choix des données

La distribution de la comorbidité par classes d'âge n'est pas présentée dans le document de référence 1 dont les données portent sur 71 342 personnes. Il est donc nécessaire de faire une hypothèse supplémentaire pour analyser ces données.

La 1^{ère} étude réalisée en 2020 contient 2 analyses qui utilisent ces 2 jeux de données :

- l'une porte sur l'âge de la "totalité" des personnes décédées (50 237) et fait l'objet de l'hypothèse suivante: "les deux pourcentages de comorbidité (65% et 35%) sont répartis dans les tranches d'âge proportionnellement au pourcentage de décès";
- l'autre porte sur l'âge et la comorbidité d'un sous ensemble des décès (20 766).

Les résultats obtenus avec ces 2 jeux de données étant presque identiques, cette mise à jour de l'étude se limite à une analyse avec un jeu de données qui porte sur l'âge et la comorbidité d'un sous ensemble des décès.

4.2 Analyse sur 5 classes

Le tableau 10 du document de référence 2 (p 39) 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 2020 au 18 janvier 2021 (données au 19 janvier 2021) (Source : Santé publique France, Inserm-CépiDC)".

Les données de ce tableau portent sur **30 032 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 +	Somme
1	Classe d'âge	C1	C2	C3	C4	C5	
2	Total échantillon	3	184	2 062	4 258	23 525	30 032
3	nombre avec comorbidité	3	123	1 479	2 984	14 994	19 583
4	nombre sans comorbidité	0	61	583	1 274	8 531	10 449
5	% avec	0,01	0,41	4,92	9,94	49,93	65,2
6	% sans	0,00	0,20	1,94	4,24	28,41	34,8
7	nombre avec / Total en %	100,0	66,8	71,7	70,1	63,7	65,2

Tableau 3

Les **lignes 3** (avec comorbidité en bleu) et **4** (sans comorbidité en orange) de ce tableau représente la fonction à 2 variables "f" mentionnée au §1. Si "f" est la fonction "Nombre de décès", on peut écrire :

nombre de décès(avec como, classe d'âge C4) = 2 984 = f(avec, C4)

nombre de décès(sans como, classe d'âge C3) = 583 = f(sans, C3)

Les **lignes 5 et 6** représentent la proportion de mortalité dans chaque case pour l'échantillon de 30 032 personnes. Si on définit une fonction "P" qui représente la proportion de décès, on peut alors écrire :

Proportion de décès dans la classe C3 de personnes présentant au moins une comorbidité = P(avec, C3) = 4,92 %.

La somme des proportions pour la ligne 5 est égale à 65,2 % qui correspond au % global de comorbidité.

La **ligne 7** représente la proportion de décès dans chaque classe d'âge des personnes présentant au moins une comorbidité par rapport au nombre de décès dans la classe d'âge (à ne pas confondre avec la ligne 3). Par exemple, la proportion de décès dans la classe C3 de personnes présentant au moins une comorbidité est égale à 71,7% du nombre de décès dans cette classe d'âge (= 1479/2062).

Cette valeur correspond à la **part de "avec comorbidité" dans chaque classe d'âge**.

La dernière colonne de cette ligne (sur fond vert) affiche la valeur moyenne calculée des proportions. Cette moyenne correspond bien au % global de comorbidité.

On peut constater que cette proportion n'est pas constante. A l'exception de la classe C1 qui concerne 3 décès avec comorbidité, **cette proportion est maximale pour la classe C3** et est minimale pour la classe C5.

C'est dans la classe d'âge "75 et +" que la proportion de décès avec comorbidité dans la classe (14 994/23 525) est la plus faible. Dans la même colonne, la proportion de la ligne 5 est la plus élevée car elle porte sur l'ensemble de l'échantillon.

Ce résultat surprenant peut s'expliquer par le fait que les personnes présentant au moins une comorbidité dans la classe C5 prennent plus de précautions que celles de la classe C3.

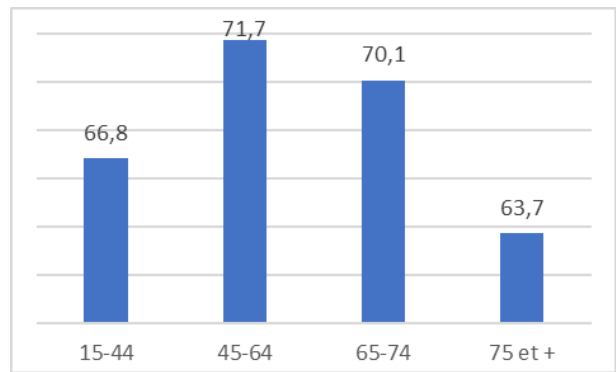


Figure 2 : proportion de comorbidité selon l'âge

4.3 Analyse sur 6 classes

Comme au §2, Les données contenues dans le tableau 3 ne tiennent pas compte de l'information très importante contenue dans l'**âge médian** qui est égal à 85 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 85[et]85 et plus), cette dernière contenant 50% des décès, c'est-à-dire 15 016, cette quantité étant retirée de la classe (75 et +).

On obtient alors le tableau suivant :

	Age	0-14	15-44	45-64	65-74	75-85	85 et +	Somme
1	Classe d'âge	C1	C2	C3	C4	C5	C6	
2	Total échantillon	3	184	2 062	4 258	8 509	15 016	30 032
3	nombre avec comorbidité	3	123	1 479	2 984	5 423	9 571	19 583
4	nombre sans comorbidité	0	61	583	1 274	3 086	5 445	10 449
5	% avec	0,01	0,41	4,92	9,94	18,06	31,87	65,2
6	% sans	0,00	0,20	1,94	4,24	10,27	18,13	34,8
7	nbre avec / Total en %	100,0	66,8	71,7	70,1	63,7	63,7	65,21

Tableau 4

Les significations des lignes sont identiques à celles du tableau 3.

Le passage de 5 à 6 classes nécessite une hypothèse concernant la ligne 7.

Hypothèse : La proportion de comorbidité de la classe "75 et +" du tableau 3, égale à 63,7%, est conservée pour les classes "75-85" et "85 et +" du tableau 4. Cette hypothèse est retenue malgré la décroissance de cet indicateur qu'illustre la figure 2, car toute autre hypothèse aurait été hasardeuse.

La valeur de la **ligne 3** "nbre avec" pour la **classe C5** (= 5 423) est obtenue en multipliant l'effectif de la classe (8 509) par 63,7 % qui représente la proportion de comorbidité de cette classe. La valeur de la case du dessous (= 3 086) est obtenue par différence (8 509 – 5 423). La même opération est effectuée pour la classe C6.

Les lignes 4 (Avec comorbidité en bleu) et 5 (Sans comorbidité en orange) de ce tableau sont les valeurs de la fonction "f" mentionnée au §1, et qui mesure le nombre de décès avec 2 variables. Par exemple :

nombre de décès(avec comorbidité, classe d'âge C3) = 1 479 = f(avec, C3)

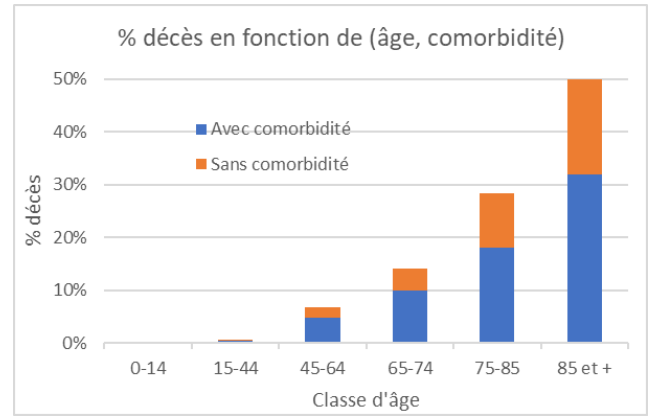
nombre de décès(sans comorbidité, classe d'âge C4) = 1 274 = f(sans, C4)

La figure ci-dessous donne une représentation graphiques de la fonction f (en fait f/30 032), c'est-à-dire des pourcentages affichés dans les lignes 3 et 4 du tableau 6.

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

En examinant les 4 dernières classes (C3 à C6), on constate que dans chaque classe d'âge, le pourcentage de décès avec comorbidité (bleu) est peu différent du pourcentage de décès sans comorbidité (orange) de la classe suivante, parfois inférieur et parfois supérieur.

Dans l'exemple ci-dessus, $f(\text{avec}, C3) > f(\text{sans}, C4)$.



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.**

4.4 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) ou (comorbidité, âge) ce qui est identique. Les résultats obtenus précédemment montre que **si l'âge est un facteur important, c'est aussi le cas de la comorbidité.**

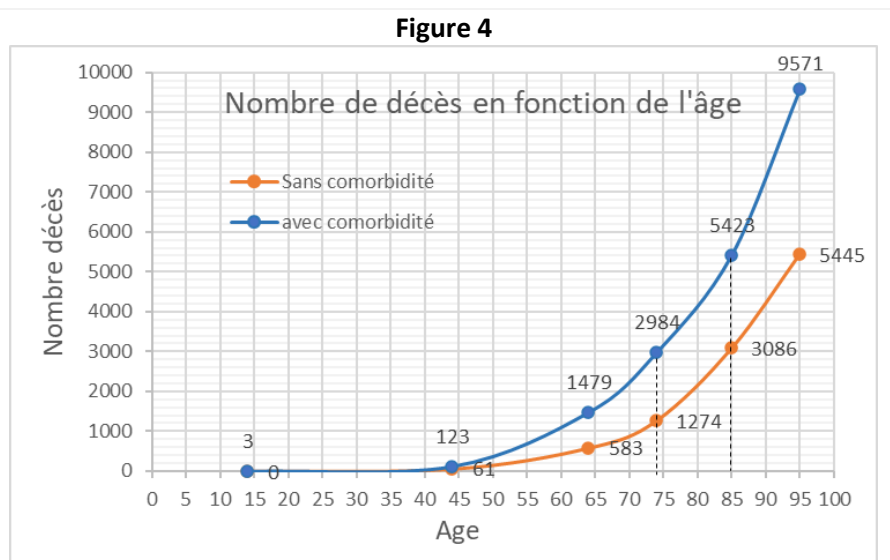
En effet, on constate que dans chacune des classe, C3 à C6 la proportion de décès avec comorbidité (bleu) est peu différente de la proportion sans comorbidité (orange) de la classe suivante, parfois inférieure et parfois supérieure.

Les **résultats de l'étude** sont résumés dans le graphique ci-dessous dans lequel chaque classe d'âge est représentée en abscisse par sa borne supérieure. Par exemple la classe d'âge "74-85" est représentée par l'âge 85 ans. Le graphique affiche 2 courbes :

Les **deux courbes** du graphique représentent en ordonnée le nombre de décès dans chaque classe d'âge. La courbe du dessus (en bleu) concerne les personnes présentant au moins une comorbidité et celle du bas (en orange), celles qui n'en présentent pas. Par exemple pour la classe d'âge '65-74', le nombre de décès avec comorbidité (resp. sans) est égal à 2984 (resp. 1274).

En examinant les deux courbes, on constate que dans chaque classe d'âge, le nombre de décès avec comorbidité (bleu) est peu différent du nombre de décès sans comorbidité (orange) de la classe suivante (\approx même ordonnée).

Présenter une comorbidité est donc équivalent à monter d'une classe d'âge.



L'âge n'est pas un paramètre suffisant pour décrire la mortalité

5 Remarques sur les données et le résultat

5.1 Le sous ensemble étudié

L'analyse des données portant sur les variables "âge" et "comorbidité" a été réalisée sur un sous ensemble de 30 032 personnes, ce qui représente moins de la totalité des décès (71 342) sur la période considérée. Il est évident que réduire ainsi l'échantillon soumis à l'étude ne peut que dégrader la qualité des résultats.

Dans la 1ère étude réalisée en 2020 (Cf. §3.1) les résultats obtenus avec les 2 jeux de données étant presque identiques, il avait été conclu que :

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

L'étude de la part de comorbidité dans chaque classe d'âge réalisée dans la présente mise à jour ne confirme pas l'hypothèse 1. La prudence incite donc à penser que la conclusion 2) ci-dessus reste à confirmer.

5.2 Les classes d'âge

Les données initiales affichées dans le tableau 3 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 85 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".

6 Conclusions

6.1 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 l'existence d'une telle base.

Tels que définis dans les documents de référence :

- le découpage en "tranches d'âge" n'est 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.

Aux variables "âge" et "comorbidité", il conviendrait d'ajouter le sexe et une variable sur le lieu de résidence (par exemple grand ville, ville moyenne, petite ville, village).

L'exercice décrit dans les paragraphes précédents a été réalisé avec des données disponibles dans les documents de référence, en y ajoutant une hypothèse qui devrait être levée 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.

6.2 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.

7 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 bleue du tableau** 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