

**Mémoire présenté devant l'Institut de Science Financière et  
d'Assurances  
pour l'obtention du diplôme du Master Actuariat  
et l'admission à l'Institut des Actuaires**

20/05/2025

Par : Ribal SLIM

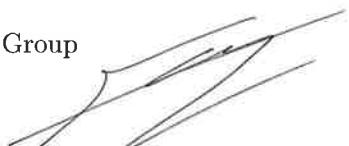
Titre : Estimation des *Value-at-Risk* de prestations décès en assurance emprunteur à l'aide de modèles agents dans le contexte d'une pandémie.

Confidentialité :  Non     Oui    (Durée :  1 an     2 ans)

*Les signataires s'engagent à respecter la confidentialité ci-dessus*

*Membres présents du jury de l'Institut  
des Actuaires :*  
Q.T. Nguyen

*Membres présents du Jury du Master  
Actuariat de l'ISFA :*  
P. Ribereau

*Entreprise :*  
Nom : Périclès Group  
Signature : 

*Directeur de Mémoire en entreprise :*  
Nom : Mouna AIT-OMAR ROSENBAUM  
Signature : 

*Autorisation de publication et de mise en ligne sur un site de diffusion de documents  
actuariels (après expiration de l'éventuel délai de confidentialité)*

*Secrétariat :*

*Signature du responsable entreprise*



*Bibliothèque :*

*Signature du candidat*



« Dans les derniers jours de novembre 1918, Edmond Rostand quitta sa propriété de Cambo et se rendit à Paris pour participer à l'allégresse générale de l'armistice. Un soir il alla au théâtre Sarah Bernhardt assister à une répétition de l'Aiglon avec la grande actrice. Il prit froid dans les coulisses et rentra chez lui frissonnant et souffrant de douleurs thoraciques. Le lendemain, 30 novembre, la fièvre atteignait 41 °C. Deux jours plus tard il était mort. »

GUÉNON L. - *Edmond Rostand et la grippe espagnole*, Thèse médecine, Aix-Marseille,  
1994

## Résumé

---

Tous les assureurs ne sont pas directement concernés par le risque pandémie, mais ceux qui le sont doivent le prendre en compte lors de l'évaluation de leurs solvabilités. Comprendre comment ils sont impactés et le modéliser peut aider à identifier et quantifier le poids d'un tel risque.

Cependant, les modèles déterministes utilisés échouent à englober l'aspect social d'une pandémie. De plus, ils ne possèdent pas de volatilité naturelle.

L'objectif de ce mémoire est d'estimer la distribution de prestations décès en assurance emprunteur à l'aide d'une autre famille de modèles, dits modèles agents, dans le contexte d'une pandémie.

Ces modèles ont été explicités à l'aide du simulateur *CovDyn* qui fut augmenté de nouvelles composantes afin de prendre en entrée un portefeuille emprunteur construit à partir de données transmises par une entreprise d'assurance. Plusieurs scénarios (central et stressés) ont été pensés.

La situation de l'assureur peut effectivement se voir détériorée en fonction de la situation dans laquelle il se retrouve. Premièrement, un taux élevé de décès engendre beaucoup plus de retombées en termes de prestations. Deuxièmement, une probabilité d'hospitalisation élevée, à elle seule, ne peut pas empirer la situation de l'assureur. Cependant, couplée d'un taux de décès élevé, la situation devient plus grave que dans le premier cas et peut mettre en péril la solvabilité de l'assureur.

Les résultats suggèrent que les acteurs du secteur de l'assurance doivent être conscients de ce risque et de l'étendue de son impact suivant les différents cas possibles. Ils doivent pouvoir modéliser ce risque et pousser les développements afin d'intégrer tous les aspects d'une pandémie dans leurs modèles.

---

*Mots-clés : Pandémie, Risque décès, Assurance emprunteur, Covid-19, Modèle agents, Value-at-Risk.*

## Abstract

---

Not all insurers are directly affected by the risk of a pandemic, but those who are must take it into account when assessing their solvency. Understanding how they are impacted and modeling it can help identify and quantify the weight of such a risk.

However, the deterministic models used fail to encompass the social aspect of a pandemic and lack natural volatility in their simulations.

The objective of this thesis is to estimate the distribution of death benefits in the case of a credit loan insurance. It will utilize another family of models, known as agent-based models, in the context of a pandemic.

These models were explained using the simulator *CovDyn*, which was enhanced with new components to take as input a credit loan portfolio constructed from data provided by an insurance company. Several scenarios (central and stressed) were considered.

The insurer's situation can indeed deteriorate depending on the situation they find themselves in. Firstly, a high death rate results in much higher payouts in terms of benefits. Secondly, a high probability of hospitalization alone cannot worsen the insurer's situation. However, coupled with a high death rate, the situation becomes more severe than in the first case and can jeopardize the insurer's solvency.

The results suggest that insurance industry players must be aware of this risk and the extent of its impact under different possible scenarios. They must be able to model this risk and push developments to integrate all aspects of a pandemic into their models.

---

*Keywords :* *Pandemic, Mortality risk, Credit loan insurance, Covid-19, Agent based models, Value-at-Risk.*

# Note de Synthèse

Certes la crise sanitaire de la Covid-19 n'a pas eu beaucoup d'impacts directs chez les assureurs emprunteurs en France. En 2020, la sinistralité au sein de cette branche n'a pas augmentée. Les assurés étant généralement jeunes, le risque de décès n'a pas dérivé[15]. Les franchises pour arrêt de travail, souvent de 90 jours, ont limité les prestations versées[8]. Mais cela ne veut pas dire qu'une autre pandémie, plus puissante à ampleur inconnue, ne pourrait pas générer plus de retombées.

Et l'estimation de ces retombées requiert un changement d'hypothèses en fonction de la branche d'activité de l'assureur. Par exemple, certains rajoutent un excès de mortalité à leurs paramètres. Or cette correction pourrait ne pas comprendre tous les aspects de la maladie sous-jacente et plus précisément le côté social. Les comportements lors d'une pandémie changent, d'une part grâce aux interventions gouvernementales et d'une autre part grâce à l'état d'alerte de la population (et qui peuvent aller dans un sens ou dans l'autre). Les assureurs ont donc besoin de modèles qui englobent cet aspect. L'épidémiologie en fournit certains.

Sous un cadre épidémiologique, l'objectif de ce mémoire consiste à réviser un modèle épidémiologique et de l'ajuster afin de le rendre pertinent pour les besoins d'un assureur emprunteur. Le modèle sera ainsi adapter pour simuler l'impact d'un choc pandémie sur la distribution des prestations et l'évaluation du risque catastrophe. Dans un premier temps, les deux principaux types de modèles épidémiologiques, les modèles compartimentaux et les modèles agents, seront présentés brièvement. Les modèles agents seront retenus dans le cadre de cette étude. Étant composés de plusieurs composantes, ces modèles seront, dans une deuxième partie, plus explicitement détaillés à l'aide du simulateur *CovDyn*. Ce dernier est un modèle agents qui sera modifié et inclura dans une dernière partie une application actuarielle. Elle se basera sur les données d'un assureur détenteur d'un portefeuille emprunteur et se focalisera sur le risque de décès uniquement. Le simulateur va se voir augmenté de nouvelles composantes afin d'inclure les éléments nécessaires et finalement les résultats seront ensuite analysés et comparés.

## Modèles épidémiologiques : compartiments *vs* agents

Les modèles épidémiologiques sont conçus pour expliquer et prédire le comportement des agents infectieux. Ces modèles se séparent principalement en deux grandes familles :

1. les **modèles compartimentaux** qui se basent principalement sur des équations différentielles pour gérer l'évolution de la pandémie. Ces modèles restent majoritairement déterministes. La base de tous les modèles compartimentaux est le modèle **SIR** qui simule une population divisée en trois états : les *Susceptibles*, ceux qui sont aptes à être contaminés, les *Infected*, ceux qui

sont infectés et qui peuvent transmettre la maladie, et finalement les ***Recovered*** ou ***Removed***, ceux qui deviennent rétablis et qui peuvent être parfois supposés comme immunisés à de futures transmissions ou infections.

- les **modèles agents** qui créent une population artificielle formée d'agents représentant chaque individu de la population considérée. Ces agents vont interagir suivant des règles prédéfinies pour se comporter comme une vraie population (et ainsi avoir des contacts) et d'autres lois simuleront les possibilités d'infection ainsi que le parcours de chaque agent dans les différents états de la maladie (incubation, hospitalisation, etc.)

Les modèles compartimentaux sont déterministes et de légères modifications des conditions initiales résulteraient en des conclusions très différentes en fonction de ces altérations. De plus, ces modèles ne possèdent pas de volatilité naturelle d'autant qu'ils n'englobent pas l'aspect sociétal de la pandémie. Cette étude priviliege alors les modèles agents qui amènent certaines solutions en échange d'un temps de calcul plus important.

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I(t) \times S(t) \\ \frac{dI(t)}{dt} = \beta I(t) \times S(t) - \gamma \times I(t) \\ \frac{dR(t)}{dt} = \gamma \times I(t) \end{cases}$$

FIGURE 1 – Système d'équations différentielles régissant le modèle SIR

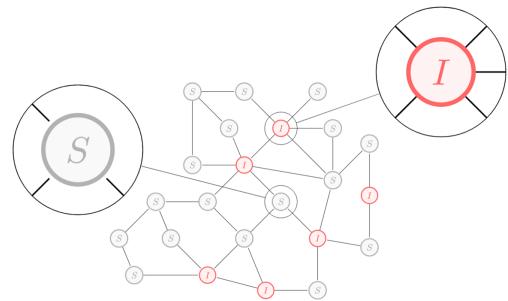


FIGURE 2 – Population artificielle d'agents à contacts et à différentes étapes de la maladie

## Modèles agents : simulateur *CovDyn*

Le simulateur utilisé dans ce mémoire est appelé *Covid Dynamics* abrégé par *CovDyn*. Il a été mis en place par les Hospices Civils de Lyon dans le cadre de la pandémie de la Covid-19. Ce **modèle agents** servira de base pour expliquer en détail la construction et le fonctionnement de cette famille de modèle et, après sa modification, en tant qu'application dans le cadre du risque de décès, risque sous-jacent d'une assurance emprunteur. Chaque modèle agents est constitué de différentes composantes et le modèle *CovDyn* n'en est pas une exception. Les composantes principales amènent du réalisme à la représentation et chacune d'elles s'occupe d'un aspect spécifique de la pandémie. Dans le cas du simulateur *CovDyn*, elles sont du nombre de quatre.

La première composante est la **composante société**. Elle crée la population artificielle et la caractérise. Cela prend en charge le nombre d'agents total qui sera simulé ainsi que leur répartition dans différentes classes d'âge considérées. A été retenue pour le modèle *CovDyn* une représentation de la population française en 2022 suivant la pyramide des âges de la même année. Cependant 64 millions d'agents virtuels demanderaient un temps de calcul bien trop long et donc la population a été réduite à l'aide d'un facteur de réduction. Initialement fixé à 1%, ce facteur a été revu afin de réduire le temps de calcul tout en gardant une bonne reproduction de l'évolution de la pandémie. Finalement, le facteur retenu vaut 0,5% et revient à simuler environ 320 000 agents. Les résultats finaux seront multipliés par 200.

Chacun de ces agents se verra assigné des attributs différents comme la classe d'âge à laquelle

il appartient. Ici, il existe 17 classes d'âge de 5 ans chacune, allant de 0 à 79 avec une classe d'âge supplémentaire pour les individus âgés de plus de 80 ans.

Ensuite vient la **composante mouvement** qui va s'occuper des interactions des agents de la première composante. Effectivement, elle mettra les différents agents en « contact » entre eux comme si c'était de vraies personnes qui interagissaient dans la vraie vie. Cela ouvre alors à la possibilité de contaminations. Le nombre de contacts par jour dépendra de sa classe d'âge. Au cœur de cette composante se trouve la **matrice de contact** qui résume toute l'information

et dont les éléments, notés  $C_{i,j}$ , représentent le nombre de contacts moyen journalier d'un agent de la classe  $i$  avec un autre agent qui lui appartient à la classe  $j$ .

	80P	0.05	0.05	0.07	0.03	0.05	0.02	0.05	0.05	0.05	0.08	0.09	0.09	0.13	0.14	0.33	0.64	0.64
75-79	0.05	0.05	0.07	0.03	0.05	0.02	0.05	0.05	0.05	0.08	0.09	0.09	0.13	0.14	0.33	0.64	0.64	0.64
70-74	0.09	0.06	0.07	0.03	0.05	0.03	0.05	0.10	0.09	0.08	0.09	0.12	0.26	0.28	0.97	0.37	0.37	
65-69	0.13	0.11	0.09	0.05	0.05	0.05	0.09	0.15	0.11	0.09	0.12	0.22	0.43	1.00	0.62	0.37	0.37	
60-64	0.19	0.18	0.10	0.07	0.08	0.11	0.18	0.26	0.18	0.15	0.27	0.51	1.41	0.52	0.66	0.28	0.28	
55-59	0.27	0.19	0.15	0.16	0.33	0.36	0.41	0.36	0.27	0.34	0.65	1.46	0.68	0.44	0.34	0.34	0.34	
50-54	0.31	0.21	0.27	0.48	0.81	0.96	0.81	0.92	1.11	1.09	1.92	0.92	0.44	0.32	0.34	0.54	0.54	
45-49	0.31	0.33	0.57	1.05	1.27	1.00	1.14	1.34	1.68	2.20	1.74	0.68	0.42	0.30	0.43	0.44	0.44	
40-44	0.50	0.85	1.08	1.02	0.97	1.26	1.46	2.10	2.95	1.57	1.59	0.79	0.52	0.43	0.51	0.45	0.45	
35-39	1.04	1.14	0.83	0.76	1.13	1.47	1.89	3.21	1.88	1.47	1.21	0.75	0.69	0.45	0.48	0.38	0.38	
30-34	1.21	0.92	0.40	0.53	1.17	1.78	2.92	1.76	1.59	1.26	1.18	0.92	0.64	0.45	0.25	0.31	0.31	
25-29	0.75	0.42	0.24	0.65	1.71	3.46	1.72	1.41	1.25	0.99	1.29	0.93	0.51	0.31	0.24	0.19	0.19	
20-24	0.34	0.16	0.29	1.36	3.93	1.94	1.01	0.76	0.92	0.95	1.02	0.61	0.34	0.21	0.16	0.15	0.15	
15-19	0.26	0.27	0.85	7.83	2.17	0.73	0.42	0.68	1.11	1.54	1.20	0.75	0.36	0.16	0.36	0.36	0.36	
10-14	0.41	1.01	6.99	2.30	0.26	0.14	0.53	0.84	0.98	0.61	0.83	0.57	0.26	0.26	0.31	0.45	0.45	
5-9	1.11	5.03	1.54	0.31	0.17	0.25	0.85	1.07	0.66	0.47	0.59	0.67	0.36	0.31	0.29	0.32	0.32	
0-4	3.80	1.06	0.24	0.12	0.20	0.53	0.72	0.70	0.31	0.41	0.25	0.54	0.40	0.20	0.11	0.24	0.24	

FIGURE 3 – Matrice de contact

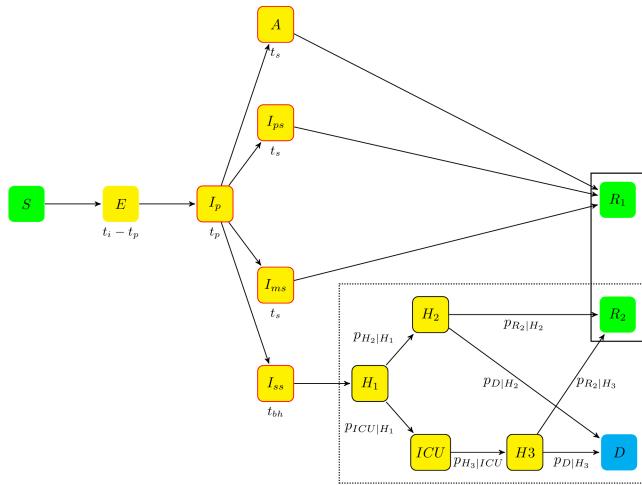


FIGURE 4 – Graphe des états

aléatoirement dans la sous-population de la classe cible et seront assignés à l'agent considéré pour la journée. Cela permet un brassage des agents des différentes classes et ainsi de suivre le comportement de chacune.

Au début de chaque jour simulé et pour chaque agent, le simulateur réalise un tirage aléatoire du nombre de contacts journaliers qu'aura cet agent avec les autres agents des différentes classes d'âges. Ce chiffre est renouvelé à chaque nouveau jour. Ce tirage se fait en construisant une loi de Poisson qui a pour paramètre l'élément  $C_{i,j}$  correspondant dans la matrice de contact. Le tirage est alors réalisé grâce à cette loi. Cette dernière a été préservée vu la pertinence de ses propriétés dans le cadre de la surveillance de contacts dans un intervalle de temps déterminé.

Les contacts seront alors tirés

Puis, la composante suivante s'appelle la **composante maladie** et peut être scindée en deux : la **sous-composante infection** qui est appelée lors d'un contact entre deux agents dont l'un est infecté et l'autre non, et réalise un tirage pour voir si une infection se transmet entre les deux ou pas. Cette probabilité dépend de la classe d'âge de l'agent infecté, notée  $i$ , du type d'intervention non-pharmaceutique mis-en-place, noté  $INP$  et qui a trois niveaux d'intensité (du moins au plus contraignant : *Relaxed*, *Intensive* et *Extended*) et l'état dans lequel est la personne infectée, noté  $Z$ . La durée moyenne d'interaction est représentée par le coefficient  $\beta_1$  qui sera corrigé par l'efficacité des interventions qui influencent différemment chaque classe d'âge, notée  $\beta_2$ . Cette probabilité s'écrit donc

$$P_{infection}(i, INP, Z) = \beta_1 \times \beta_{2,i,INP}(t) \times i_Z$$

Et lorsqu'un individu devient infecté, il suivra le parcours de la maladie de la Covid. La **sous-composante états** s'occupe du passage d'un état à l'autre. Le graphe ci-dessus résume l'ensemble des états dans lesquels un agent peut passer. À noter qu'un agent est infectieux et peut transmettre la maladie s'il se trouve dans les compartiments  $I_p$ ,  $A$ ,  $I_{ps}$ ,  $I_{ms}$  et  $I_{ss}$ . Les probabilités de passage et temps de séjour moyen se notent  $p_Z$  et  $t_Z$  pour un compartiment  $Z$ .

1.  $S$  (Susceptible) : ce sont les individus qui n'ont pas été infectés par le virus ;
2. La phase d'incubation se décompose en deux compartiments :
  - $E$  (Exposé) contient les individus qui n'ont pas encore développé de symptômes et qui ne sont pas contagieux ;
  - quand un individu appartenant à  $E$  commence à devenir contagieux, il est transféré au compartiment  $I_p$  (Phase prodromique) qui contient les individus qui sont toujours dans la phase d'incubation étant déjà contagieux mais sans symptômes apparents.
3.  $A$  (Asymptomatique) : individus qui ont complété la phase d'incubation, sont devenus contagieux mais ne possèdent pas de symptômes.
4. Les individus développant des symptômes sont stratifiés en trois compartiments pour distinguer leurs sévérités :
  - $I_{ps}$  (l'état paucisymptomatique) : individus chez qui ne peuvent être perçus que peu de symptômes concernant leur maladie ;
  - $I_{ms}$  (symptômes moyens) : individus possédant des symptômes (de la fièvre par exemple) qui ne nécessitent pas d'hospitalisation.
  - $I_{ss}$  (symptômes sévères) : individus dont les symptômes les mènent à l'hospitalisation.
 Après l'écoulement du temps d'attente dans les états  $A$ ,  $I_{ps}$  ou  $I_{ms}$ , les individus quittent ces derniers et se retrouvent dans le compartiment  $R$  (*Removed* ou rétablis) ;
5. La période d'hospitalisation se décompose en 4 compartiments.
  - $H_1$  : compartiment où se décide la répartition des agents afin de voir s'ils ont besoin de soins intensifs ou pas.
  - $H_2$  : individus qui n'ont pas besoin de soins intensifs ;
  - $ICU$  : individus hospitalisés qui requièrent des soins intensifs ;
  - $H_3$  : individus hospitalisés après les soins intensifs.

(!) Être hospitalisé est ici équivalent à un isolement de tous contacts. Les agents se trouvant dans un de ces quatre compartiments ne peuvent plus être considérés durant le tirage des contacts.
6. À l'issue de l'hospitalisation, les individus transitionnent vers l'un des deux compartiments :
  - $D$  (Décès) : individus décédés durant l'hospitalisation.
  - $R$  : agents rétablis de la maladie.

(!)  $R$  et  $D$  sont des états absorbants. Cela équivaut à une immunité à vie.

Finalement, la dernière composante est la **composante environnement**. En effet, les institutions ainsi que la population réagissent au fait qu'une pandémie a lieu et ainsi essayent de réduire sa

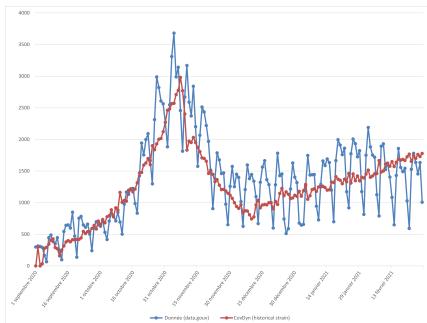


FIGURE 5 – Ancienne calibration :  
2ème vague

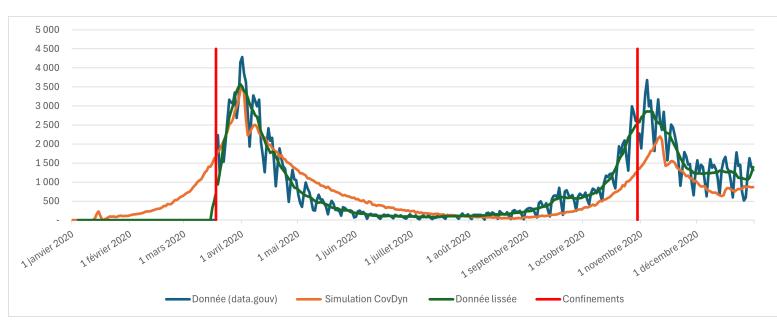


FIGURE 6 – Nouvelle calibration : une année de Covid-19 (1ère et 2ème vague)

propagation. La composante rajoute alors des contraintes à la maladie : soit par des *INP*, interventions non-pharmaceutiques, pour réduire la probabilité de transmission (comme vu dans la sous-composante infection) et en réduisant le nombre de contacts des agents en fonction de la contrainte imposée en termes de confinement ; soit par l'introduction de vaccins. Ces derniers ne vont pas être détaillés ni utilisés dans le mémoire mais à savoir qu'ils existent dans le simulateur. Il y a trois types d'interventions : les *INP-Extended* qui représentent les confinements stricts, la fermeture des écoles, des universités, des différents commerces ; les *INP-Intensive* qui comportent des mesures sanitaires strictes comme les couvre-feux, la distanciation sociale entre les gens, l'obligation du port du masque, le télétravail recommandé, la fermeture des restaurants, bars ; et les *INP-Relaxed* qui présentent un relâchement de certaines mesures (réouverture des lieux de restauration, des boutiques, plus de couvre-feu) mais maintiennent le port du masque et de la distanciation sociale. Ces coefficients seront appliqués suivant le calendrier qui a été déroulé durant l'année 2020 afin de respecter les durées de confinements.

Cependant, une modification du modèle est nécessaire. Ce dernier a été initialement calibré pour reproduire la deuxième vague de la Covid en France. Le recalibrage du nombre initial d'infectés ainsi que la durée de chaque niveau d'intervention a donc été nécessaire afin de débuter la pandémie en début d'année 2020 et observer l'évolution de celle-ci durant la période d'un an. Ce recalibrage a alors été réalisé et sa visualisation apparaît dans la figure 2.

De plus, a été ajouté une interface graphique afin d'avoir une interaction entre le simulateur et l'utilisateur. Ce dernier peut désormais choisir les différents scénarios proposés plus loin ainsi que d'avoir la main sur certains des paramètres du simulateur.

## Application : assurance emprunteur et risque décès

Pour être utilisé par des acteurs du secteur de l'assurance, des composantes additionnelles ont été ajoutées au simulateur dont la **composante assurance**. Elle adapte ainsi le simulateur afin de prendre un portefeuille emprunteur en entrée, de déterminer les caractéristiques des assurés et d'injecter ces dernières dans une sous-population d'agents qui joueront le rôle d'assurés. Ces agents vont se doter d'un nouvel attribut *Capital* qui représente le volume des capitaux sous risque.

Cette sous-population sera donc suivie dans le temps et les décès seront enregistrés. Ainsi, lors de l'occurrence de ces décès, la valeur du capital à payer sera remontée et inscrite dans un fichier. Dans le cas de ce mémoire, l'historique entier d'un portefeuille emprunteur transmis par un assureur a été utilisé comme base de construction d'un nouveau portefeuille à effectif fixé. Cette création suit les étapes suivantes :

1. Une distribution de valeur de prêts a été ajustée au sein de chaque classe d'âge. Ces distributions seront ensuite utilisées afin de réaliser des tirages aléatoires lors de la création des agents dits « assurés ». Des lois Gamma ont finalement été utilisées. À noter que les capitaux restants dûs ne figuraient pas dans la base de données transmise. Les montants initiaux de prêts ont été utilisés à la place.
2. L'effectif fixé dans cette étude est de 200 000 assurés ce qui revient à 1 000 assurés dans la population réduite. La démographie du portefeuille, stratifiée en classe d'âge comme la matrice de contact, est préservée. Les assurés sont répartis de telle sorte à retrouver les différentes proportions des différentes classes d'âge. L'utilisateur aura également la possibilité d'entrer un nombre de son choix.

Ainsi, pour le nouveau portefeuille construit, les capitaux sous risque s'élèvent à **320** millions d'euros en moyenne pour 1 000 assurés.

À l'issue d'une multitude de simulations, à savoir 600, une distribution des flux sortants du risque sous-jacent pourra être dressée et un quantile à 99,5% sera ainsi calculé. Cette VaR formera donc un élément de calcul pour trouver la valeur du SCR, car elle ne considère que le facteur coût dans son approche.

① Il est important de noter qu'une hypothèse implicite est faite ici : la pandémie se déroule sur un horizon relativement court par rapport à la durée d'un prêt. Avec cela, a été considéré que les capitaux sous risques n'évoluent pas durant cette période-là, i.e. 1 an. L'approche reste cependant prudente, car la valeur des capitaux restants dûs est décroissante car les prêts sont remboursés avec le temps.

Dans le cadre du calcul de ce quantile à 99,5%, le scénario de la Covid-19 original reprend les paramètres estimés par les épidémiologistes. Il sera utilisé comme base de référence aux autres scénarios stressés qui sont au nombre de trois. En effet, l'assureur devrait étudier l'évolution de la valeur de la VaR en fonction des situations extrêmes, surtout celles qui pourraient mettre en péril sa solvabilité. La pandémie de la Covid-19 vécue en France n'est qu'une des réalisations du déroulé de la crise sanitaire. Les conséquences et les réponses auraient pu être très différentes suivant la puissance du virus.

Donc le comportement des versements est intéressant à regarder. De plus, l'analyse des raisons qui font qu'ils se comportent d'une manière ou d'une autre devient plus simple à expliciter grâce aux modèles agents.

Scénario	VaR (en euro)
Central	825 060
Décès uniformes	2 667 912
Hospitalisations uniformes	449 511
Décès et hospitalisations uniformes	2 915 262

TABLE 1 – Valeur de la *Value-at-Risk* en fonction du scénario mis en place

Le scénario original reprend le déroulement de la pandémie comme tel et n'amène pas de changement aux paramètres du simulateur qui furent calibrés sur des données d'hospitalisation et de décès liés au virus de la Covid-19. Pour le portefeuille simulé, cela équivaut à une *Value-at-Risk* d'environ 800 mille d'euros. Cette valeur semble faible comparée au total de capitaux sous risque et cela est dû au fait que les personnes âgées étaient les plus vulnérables et elles étaient celles qui décédaient le plus souvent. Combiné avec le fait que les tranches d'âges concernées ne représentent que peu de capitaux sous risque,

la valeur de la *Value-at-Risk* semble cohérente. Le simulateur reflète donc bien cet aspect générationnel et cela ressort bien, que ça soit sur un plan épidémiologique mais également assurantiel.

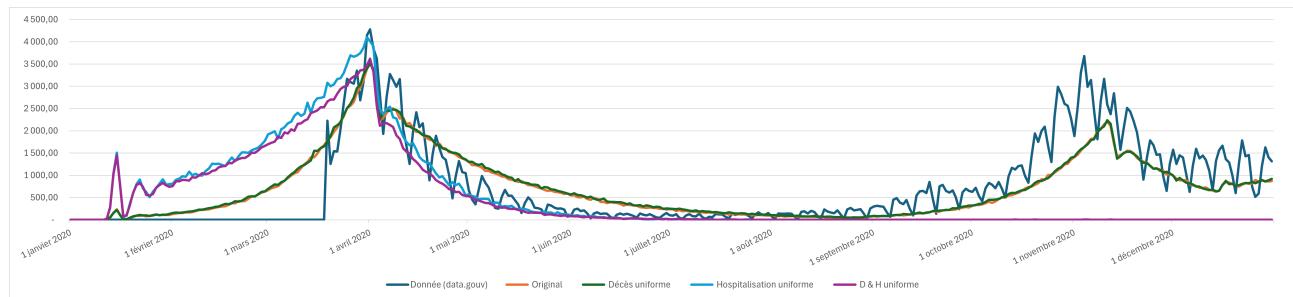


FIGURE 7 – Hospitalisations moyennes journalières pour tous les scénarios

### Taux de décès uniformes

Afin de simuler des pandémies issues de virus plus dangereux que la Covid-19, les probabilités de décès, précédemment uniques pour chaque classe d'âge, sont uniformisées en assignant à tout le monde la même probabilité de décès, à savoir les plus importantes en fonction des compartiments ( $H2$  ou  $H3$ ). Ainsi, après avoir été hospitalisé, n'importe quel agent aura la même probabilité de décéder. La courbe des hospitalisations reste, en moyenne, inchangée alors que la courbe des décès cumulés majore celle du scénario original. Cela résulte en un quantile plus élevé, 4 fois plus grand que celui du scénario précédent : plus de décès se traduit par une plus grande probabilité que l'un de ces décès provienne de l'un des agents assurés et donc plus de simulations seront non nulles en terme de prestations à verser. De plus, les jeunes adultes, qui possèdent les plus grands capitaux à risques, décèdent à la même vitesse que les personnes âgées qui en possèdent, mais à des valeurs beaucoup plus faibles.

### Taux d'hospitalisation uniforme

A été aussi testée une uniformisation de la probabilité d'hospitalisation. Dans ce scénario, les probabilités de décès sont remises à leurs valeurs initiales et le même principe appliqué dans le paragraphe précédent est appliqué à la probabilité d'hospitalisation. Ainsi, n'importe quel agent aura la même probabilité d'être hospitalisé. La courbe des hospitalisations journalières est plus haute que celle du scénario original, mais décroît beaucoup plus rapidement après avoir atteint son pic, si rapidement que la pandémie est éteinte et ne peut pas se relancer avec une deuxième vague. Cela est dû à l'hypothèse d'isolement lorsqu'un agent rentre dans les compartiments liés à l'hospitalisation. Un plateau dans le graphe des décès cumulés révèle cela aussi avec un nombre de décès qui stagne. Ce nombre réduit de décès implique un quartile à 450 mille euros. Cette valeur est moindre considérant qu'une deuxième n'a pas eu lieu. Avec cela, les limites du modèle sont atteintes : l'hypothèse d'isolement dans le bloc hospitalier devient absurde, car le nombre de lits disponibles dans les hôpitaux serait saturé. De plus, l'envergure de la crise serait accompagnée de l'effondrement du système de santé.

### Taux de décès et d'hospitalisation uniformes

Le dernier scénario est un mélange des deux derniers, avec des probabilités uniformes en termes d'hospitalisation et en termes de décès. De même, cela exhibe une croissance forte et une décroissance rapide qui limite la pandémie à une seule vague. Cependant, le nombre de décès dans ce cas est supérieur au précédent et présente le quantile le plus important avec plus de 2,9 millions d'euros. Même avec un nombre inférieur de décès cumulés que les deux premiers scénarios (Original et Décès uniforme), ce sont les « bonnes » personnes qui décèdent, c'est-à-dire les personnes qui viennent de commencer à

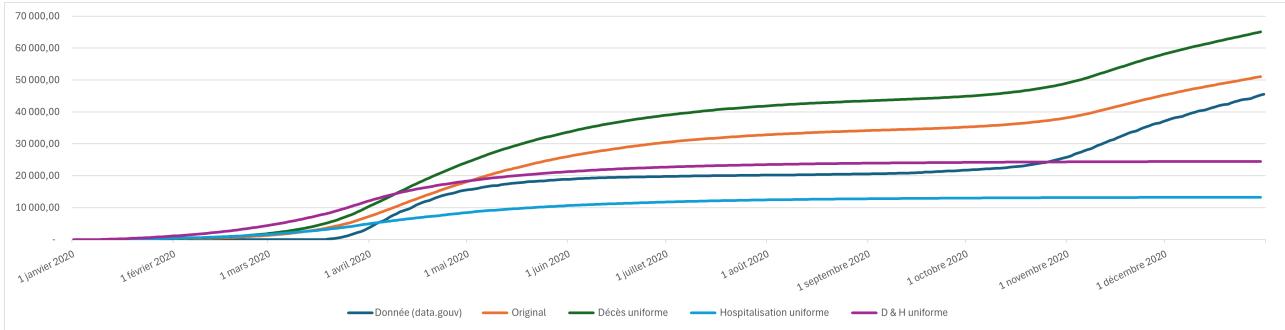


FIGURE 8 – Décès moyens journaliers pour tous les scénarios

rembourser leurs prêts, i.e. les jeunes assurés. Les capitaux sous risque sont généralement plus élevés chez les jeunes assurés et donc, en donnant la même probabilité à tout le monde de rentrer à l'hôpital et de décéder, une partie des décès représentera les classes d'âge où les capitaux sous risque ne sont pas négligeables. Dans le deuxième scénario stress (hospitalisations uniformes), tous les agents avaient effectivement la même probabilité d'être hospitalisés, mais ce ne sont que les agents des classes d'âge avancées qui sont « privilégiés à décéder » i.e. les agents dont les capitaux sous risque sont moindres.

### Covid aléatoire

Tous les paramètres du modèle CovDyn sont prédéfinis et statiques, représentant la pandémie de Covid-19 en France. Pour ajouter de la volatilité, des intervalles de valeurs sont créés autour des paramètres  $p_A$ ,  $p_H$ ,  $p_{ICU|H1}$ ,  $p_{D|H2}$  et  $p_{D|H3}$ , en ajoutant et retranchant 20% à leur valeur originale. Cette valeur est arbitraire. Deux méthodes de tirage sont envisagées : gaussienne et uniforme. La méthode uniforme est retenue pour sa simplicité et prudence, donnant le même poids à toutes les valeurs de l'intervalle. Elle inclut alors les extrêmes dans le tirage.

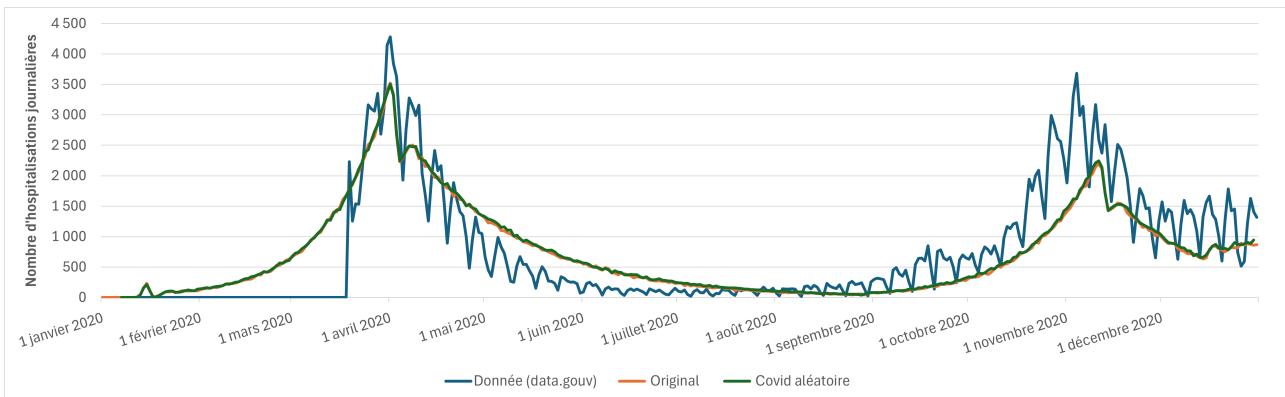


FIGURE 9 – Comportement en moyenne égal à celui du Covid-19 original en termes d'hospitalisations journalières

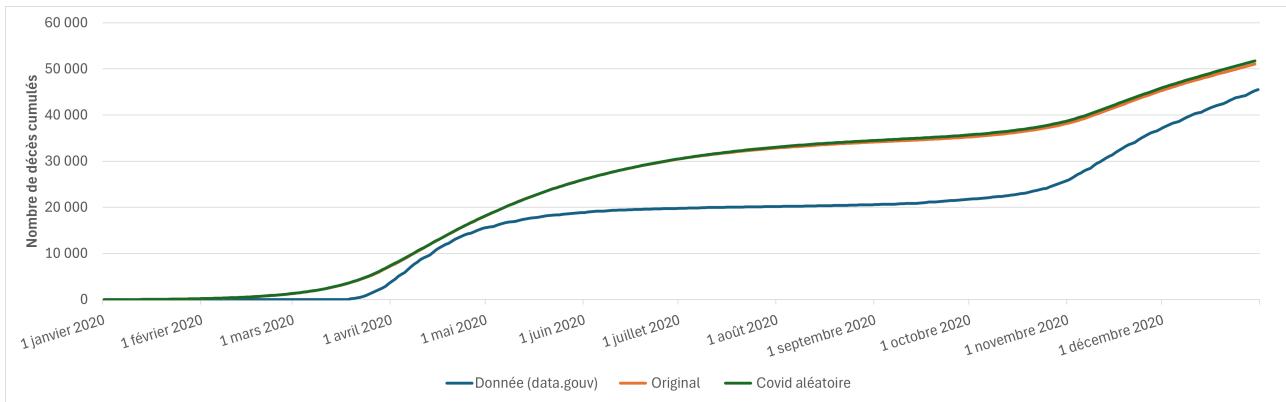


FIGURE 10 – Évolution identique, en moyenne, des décès cumulés

Le quantile respectif est légèrement plus élevé et atteint 941 614,1 euros. La valeur reste proche de celle du scénario central. La différence pourrait être amenée par cette volatilité supplémentaire que produisent les tirages des paramètres. Pour en être certain, il faudrait un nombre plus important de simulations.

### Décalage du début du confinement

Décalage	VaR (en euro)	Evolution (en %)
0 jour (Central)	825 059,8	
5 jours	1 109 803,16	+ 34
10 jours	1 139 935,23	+ 38
15 jours	2 111 002,41	+ 155

TABLE 2 – Valeur de la *Value-at-Risk* en fonction du décalage de la date du premier confinement

Relevant d'un aspect politique, la valeur de la VaR a été estimée dans le cadre de 3 décalages dans le temps : un décalage de 5 jours, un décalage de 10 jours ainsi qu'un décalage de 15 jours. C'est-à-dire, au lieu de commencer le 18 mars, le premier confinement sera déclenché le 23 mars, 28 mars ou le 2 avril suivant la simulation.

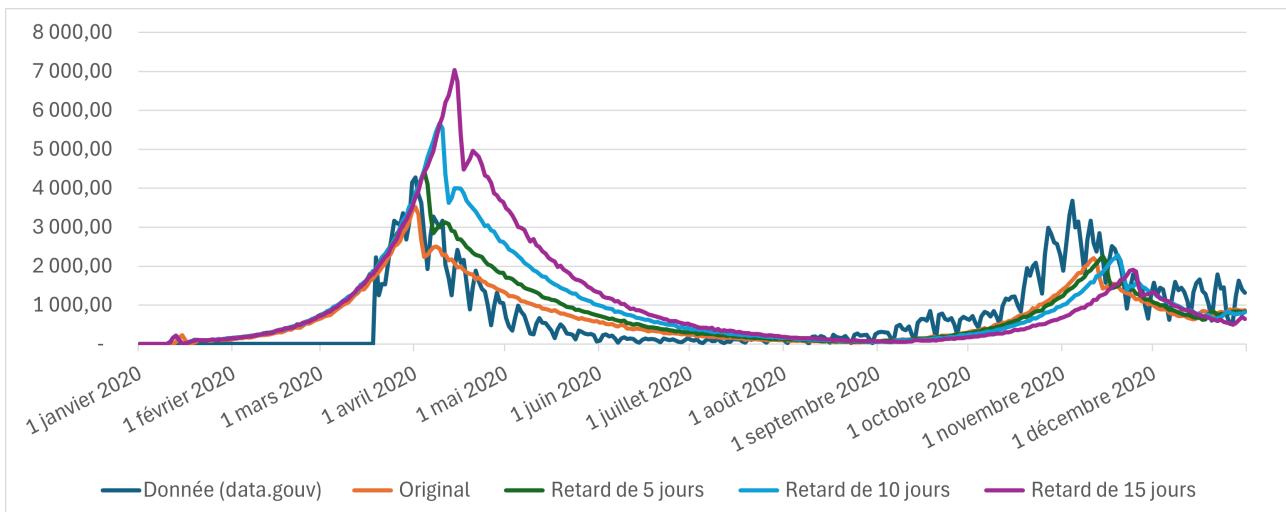


FIGURE 11 – Évolutions aux mêmes motifs mais décalé vers la droite et avec des pics plus importants

La courbe représentant le nombre d'hospitalisations journalières se déplace vers la droite, atteignant un pic plus prononcé durant la première vague, à mesure que la date du premier confinement est différée. Cela garantit le maintien de la vitesse de transmission du virus, ce qui permet l'admission d'un nombre considérable d'agents au sein du bloc hospitalier. Les impacts sur la VaR sont exprimés ci-dessus. Elle est d'autant plus élevé que la date du confinement est décalée dans le temps.

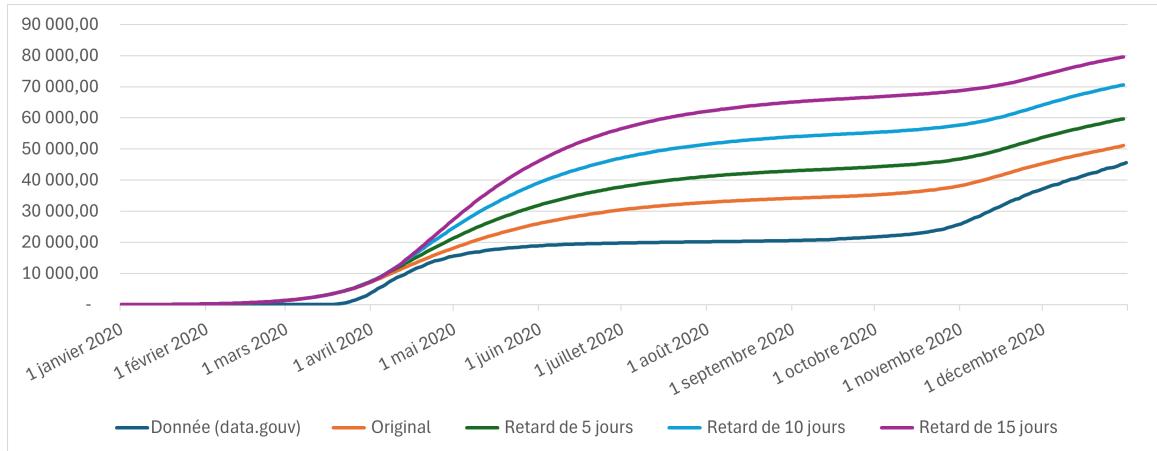


FIGURE 12 – Evolution des décès cumulés en fonction du décalage

La classe d'âge la plus touchées ici est celle des 80P qui a la plus haute probabilité d'hospitalisation et donc va être la plus vulnérable (comme dans le scénario central). Malgré une augmentation significative du nombre de décès, la valeur de la VaR ne dépasse pas celle des scénarios antérieurement discutés, car la plupart des personnes concernées appartiendraient à la catégorie des personnes âgées, préservant ainsi les segments d'âge où se trouve le principal volume de capitaux à risque.

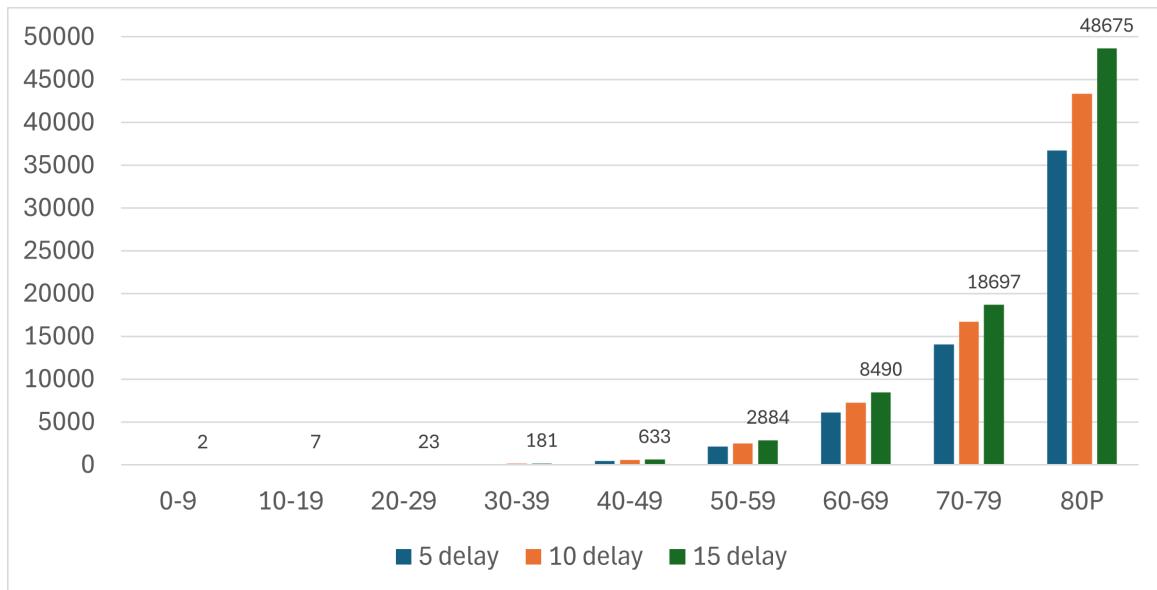


FIGURE 13 – Distribution des décès suivant le degré de décalage - Même forme de distribution que le scénario original

## Limites et conclusions

1. les décès ne peuvent survenir qu'après une hospitalisation. Les agents en dehors du bloc hospitalier ne peuvent pas se retrouver dans le compartiment décès. Et donc il pourrait y avoir une sous-estimation du nombre de décès total ;
2. les calibrations des paramètres concernant les interventions non pharmaceutiques sont celles de la deuxième vague et donc ne représentent pas tout à fait la première vague. Un meilleur ajustement peut être réalisé à l'issue d'un recalibrage de ces coefficients ;
3. certains des scénarios peuvent s'éloigner de la réalité, mais proposent quand même un schéma de pensée pour rajouter des composantes ou en affiner des existantes. Par exemple, une addition serait d'intégrer une limite au nombre d'agents pouvant être présents dans le bloc hospitalier. Un autre exemple serait d'assouplir l'hypothèse d'isolation totale dans ce bloc-là et de se questionner sur la validité de cette hypothèse ;
4. les montants des prêts initiaux ne représentent pas le risque actuel que porte l'assureur. Afin d'avoir une analyse plus fine et plus poussée, un portefeuille emprunteur contenant les capitaux restants dûs serait privilégié.

Finalement, les résultats précédents montrent que l'assureur doit avoir une vision prudentielle et doit tester sa solvabilité à court terme, surtout dans le cas d'une pandémie. Il doit voir comment des cas plus extrêmes pourraient mettre en péril cette solvabilité.

Il doit surtout réaliser des stress tests concernant des scénarios où le taux de décès est particulièrement élevé. Le scénario augmente le risque pour l'assureur surtout s'il se produit simultanément avec une situation de probabilités d'hospitalisations élevées.

Les modèles agents permettent d'une part d'intégrer une perspective axée sur l'assurance, de fournir des éléments de calcul du quantile à 99,5% ainsi que des pistes d'analyses afin de comprendre le comportement d'un certain scénario et donc les retombées en termes de prestations générées par celui-ci.

Il est donc important et nécessaire de prévoir ce genre d'exercice et d'appliquer des chocs sur les paramètres mêmes afin d'étudier l'évolution de la valeur de la VaR.



# Synthesis note

Certainly, the Covid-19 pandemic did not have many direct impacts on credit loan insurers in France. In 2020, the claims rate in this sector did not increase [15]. The insured individuals were generally young, so the risk of death did not deviate. The deductible often being 90 days often limited the number of benefits paid [8]. However, this does not mean that another, more powerful pandemic of unknown magnitude could not generate more repercussions.

Estimating these repercussions requires a change in assumptions depending on the insurer's line of business. For example, some add excess mortality to their parameters. However, this correction might not encompass all aspects of the underlying disease, particularly the social side of it. Behaviors during a pandemic change, partly due to government interventions and partly due to the population's state of alertness (which can go either way). Insurers therefore need models that include this social aspect. Epidemiology provides some of these models.

Within an epidemiological framework, the objective of this thesis is to revise an epidemiological model and adjust it to make it relevant to the needs of a credit loan insurer. The model will thus be adapted to simulate the impact of a pandemic shock on the distribution of benefits and the assessment of the catastrophe risk. First, the two main types of epidemiological models, compartmental models and agent-based models, will be briefly presented. Agent-based models will be retained for this study. Being composed of several components, these models will be more explicitly detailed in the second part using the *CovDyn* simulator as an example. It is an agent-based model that will be modified and will include an actuarial application. This application will be based on the data of an insurer holding a credit loan portfolio and will focus solely on the risk of death. The simulator will be augmented with new components to include the necessary elements, and the results will then be analyzed and compared.

## Epidemiological models: Compartmental *vs* Agent-Based

Epidemiological models are designed to explain and predict the behavior of infectious agents. These models are mainly divided into two major families:

1. **Compartmental models**, which are primarily based on differential equations that manage the evolution of the pandemic. These models are mostly deterministic. The basis of all compartmental models is the **SIR** model, which simulates a population divided into three states: *Susceptibles*, those who are susceptible to infection; *Infected*, those who are infected and can transmit the disease; and finally *Recovered* or *Removed*, those who have recovered and are sometimes assumed to be immune to future transmissions or infections.
2. **Agent-based models**, which create an artificial population composed of agents representing

each individual in the considered population. These agents interact according to predefined rules to behave like a real population (and thus have contacts), and other laws simulate the possibilities of infection as well as the progression of each agent through different states of the disease (incubation, hospitalization, etc.).

Compartmental models are deterministic, and slight modifications of the initial conditions would result in very different conclusions depending on these alterations. Moreover, these models do not possess natural volatility as they do not encompass the societal aspect of the pandemic. This study therefore favors agent-based models, which provide certain solutions in exchange for a longer computation time.

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I(t) \times S(t) \\ \frac{dI(t)}{dt} = \beta I(t) \times S(t) - \gamma \times I(t) \\ \frac{dR(t)}{dt} = \gamma \times I(t) \end{cases}$$

Figure 14 – System of differential equations governing the SIR model

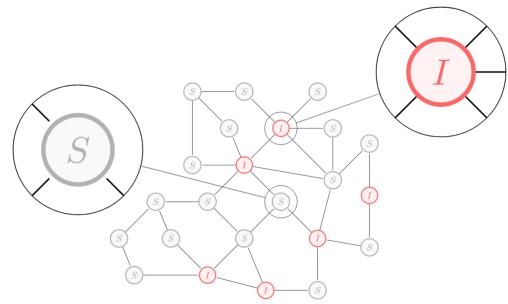


Figure 15 – Artificial population of agents with contacts at different stages of the disease

## Agent-Based Models: *CovDyn* Simulator

The simulator used in this thesis is called *Covid Dynamics*, abbreviated as *CovDyn*. It was developed by the *Hospices Civils de Lyon* during the Covid-19 pandemic. This **agent-based model** will serve as the basis to explain in detail the construction and functioning of this family of models and will be used, after some modifications, as an application in the context of the risk of death, an underlying risk of credit loan insurance. Each agent-based model is made up of different components, and the *CovDyn* model is no exception. The main components add realism to the representation, and each of them addresses a specific aspect of the pandemic. In the case of the *CovDyn* simulator, there are four main components.

The first component is the **society component**. It creates the artificial population and characterizes it. This involves the total number of agents to be simulated as well as their distribution across different age groups. For the *CovDyn* model, a representation of the French population in 2022 based on the population pyramid of that year was chosen. However, simulating 64 million virtual agents would require an excessively long computation time, so the population was reduced using a reduction factor. Originally set to 1%, the reduction factor was revisited so to reduce computation time even further while preserving a good reproduction of the evolution of the pandemic. Finally, the value retained is 0.5%, resulting in approximately 320,000 agents being simulated. The final results will be multiplied by 200.

Each of these agents will be assigned different attributes, such as the age group to which they belong. There are 17 age groups of 5 years each, ranging from 0 to 79, with an additional age group for individuals over 80 years old.

Next comes the **movement component**, which will handle the interactions of the agents from the first component. Indeed, it will put the different agents in "contact" with each other as if they

were real people interacting in real life. This opens up the possibility for infections. The number of contacts per day will depend on the agent's age group. At the heart of this component is the **contact matrix**, which summarizes all the information. The elements, denoted as  $C_{i,j}$ , represent the average daily number of contacts an agent from age group  $i$  has with another agent from age group  $j$ .

At the beginning of each simulated day and for each agent, the simulator performs a random draw of the number of daily contacts that this agent will have with other agents from different age groups. This number is renewed each day. The drawing is done by constructing a Poisson distribution with its parameter being the corresponding element  $C_{i,j}$  in the contact matrix. The drawing is then performed using this distribution. This latter has been preserved due to the relevance of its properties in the context of the occurrence of contacts within a determined time interval.

The contacts are then randomly drawn from the sub-population of the target age group and assigned to the considered agent for the day. This allows for a mixing of agents from different age groups and thus tracking the behavior of each group.

The next component is called the **disease component** and can be divided into two parts: the **infection sub-component**, which is triggered during contact between two agents, one of whom is infected and the other is not, and performs a draw to see if an infection is transmitted between them. This probability depends on the age group of the infected agent, denoted as  $i$ , the type of non-pharmaceutical intervention in place, denoted as  $NPI$ , which has three levels of intensity (from least to most restrictive: *Relaxed*, *Intensive*, and *Extended*), and the state of the infected person, denoted as  $Z$ . The average duration of interaction is represented by the coefficient  $\beta_1$ , which will be adjusted by the effectiveness of interventions that differently influence each age group, denoted as  $\beta_2$ . This probability is written as

$$P_{infection}(i, NPI, Z) = \beta_1 \times \beta_{2,i,NPI}(t) \times i_Z$$

When an individual becomes infected, they will follow the disease progression of Covid-19 and the individual's outcome. The **state sub-component** handles the transition from one state to another. The diagram summarizes all the states an agent can go through. Note that an agent is infectious and can transmit the disease if they are in compartments  $I_p$ ,  $A$ ,  $I_{ps}$ ,  $I_{ms}$ , and  $I_{ss}$ . The transition probabilities and average residence times are denoted as  $p_Z$  and  $t_Z$  for a specific compartment  $Z$ .

The following lists all the compartments considered along with a quick description of each one.

1.  $S$  (Susceptible): individuals who have not been infected by the virus;
2. The incubation phase is divided into two compartments:

80P	0.05	0.05	0.07	0.03	0.05	0.02	0.05	0.05	0.05	0.08	0.09	0.09	0.13	0.14	0.33	0.64	0.64
75-79	0.05	0.05	0.07	0.03	0.05	0.02	0.05	0.05	0.05	0.08	0.09	0.09	0.13	0.14	0.33	0.64	0.64
70-74	0.09	0.06	0.07	0.03	0.05	0.03	0.05	0.10	0.09	0.08	0.09	0.12	0.26	0.28	0.97	0.37	0.37
65-69	0.13	0.11	0.09	0.05	0.05	0.05	0.09	0.15	0.11	0.09	0.12	0.22	0.43	1.00	0.62	0.37	0.37
60-64	0.19	0.18	0.10	0.07	0.08	0.11	0.18	0.26	0.18	0.15	0.27	0.51	1.41	0.52	0.66	0.28	0.28
55-59	0.27	0.19	0.15	0.16	0.33	0.36	0.41	0.36	0.27	0.34	0.65	1.46	0.68	0.44	0.34	0.34	0.34
50-54	0.31	0.21	0.27	0.48	0.81	0.96	0.81	0.92	1.11	1.09	1.92	0.92	0.44	0.32	0.34	0.54	0.54
45-49	0.31	0.33	0.57	1.05	1.27	1.00	1.14	1.34	1.68	2.20	1.74	0.68	0.42	0.30	0.43	0.44	0.44
40-44	0.50	0.85	1.08	1.02	0.97	1.26	1.46	2.10	2.95	1.57	1.59	0.79	0.52	0.43	0.51	0.45	0.45
35-39	1.04	1.14	0.83	0.76	1.13	1.47	1.89	3.21	1.88	1.47	1.21	0.75	0.69	0.45	0.48	0.38	0.38
30-34	1.21	0.92	0.40	0.53	1.17	1.78	2.92	1.76	1.59	1.26	1.18	0.92	0.64	0.45	0.25	0.31	0.31
25-29	0.75	0.42	0.24	0.65	1.71	3.46	1.72	1.41	1.25	0.99	1.29	0.93	0.51	0.31	0.24	0.19	0.19
20-24	0.34	0.16	0.29	1.36	3.93	1.94	1.01	0.76	0.92	0.95	1.02	0.61	0.34	0.21	0.16	0.15	0.15
15-19	0.26	0.27	0.85	7.83	2.17	0.73	0.42	0.68	1.11	1.54	1.20	0.75	0.36	0.16	0.36	0.36	0.36
10-14	0.41	1.01	6.99	2.30	0.26	0.14	0.53	0.84	0.98	0.61	0.83	0.57	0.26	0.26	0.31	0.45	0.45
5-9	1.11	5.03	1.54	0.31	0.17	0.25	0.85	1.07	0.66	0.47	0.59	0.67	0.36	0.31	0.29	0.32	0.32
0-4	3.80	1.06	0.24	0.12	0.20	0.53	0.72	0.70	0.31	0.41	0.25	0.54	0.40	0.20	0.11	0.24	0.24
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80P

Figure 16 – Contact matrix

- $E$  (Exposed): contains individuals who have not yet developed symptoms and are not contagious;
  - When an individual in  $E$  starts to become contagious, they are transferred to the  $I_p$  (Prodromal phase) compartment, which contains individuals who are still in the incubation phase, already contagious but without apparent symptoms.
3.  $A$  (Asymptomatic): individuals who have completed the incubation phase, become contagious but do not have symptoms.
  4. Individuals developing symptoms are stratified into three compartments to distinguish their severity:
    - $I_{ps}$  (Pausisymptomatic state): individuals with few symptoms related to their illness;
    - $I_{ms}$  (Moderate symptoms): individuals with symptoms (such as fever) that do not require hospitalization.
    - $I_{ss}$  (Severe symptoms): individuals whose symptoms lead to hospitalization.
 After the waiting time in states  $A$ ,  $I_{ps}$ , or  $I_{ms}$ , individuals leave these states and move to the  $R$  (Removed or Recovered) compartment;
  5. The hospitalization period is divided into four compartments:
    - $H_1$ : compartment where the distribution of agents is decided to see if they need intensive care or not.
    - $H_2$ : individuals who do not need intensive care;
    - $ICU$ : individuals hospitalized who require intensive care;
    - $H_3$ : individuals hospitalized after intensive care.
 ⓘ Being hospitalized here is equivalent to isolation from all contacts. Agents in one of these four compartments can no longer be considered during the drawing of contacts.
  6. At the end of hospitalization, individuals transition to one of two compartments:
    - $D$  (Deceased): individuals who died during hospitalization.
    - $R$ : agents recovered from the disease.
 ⓘ  $R$  and  $D$  are absorbing states. This is equivalent to lifelong immunity.

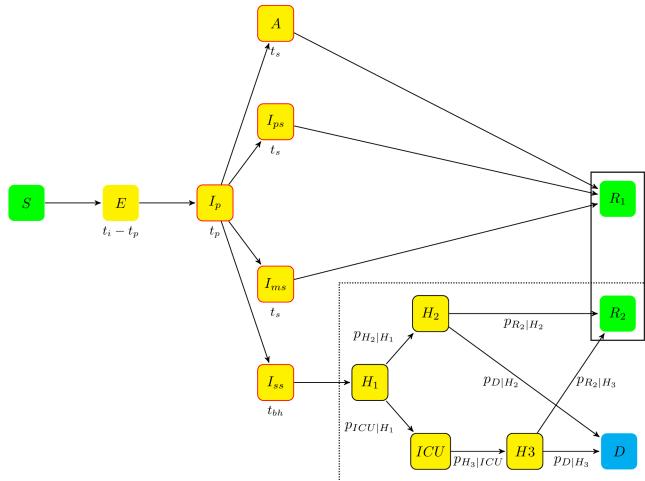


Figure 17 – State diagram

in the simulator. There are three types of interventions: *Extended* NPIs, which represent strict lockdowns, the closing of schools, universities, and various businesses; *Intensive* NPIs, which include strict health measures such as curfews, social distancing, mandatory mask-wearing, recommended

Finally, the last component is the **environment component**. Indeed, institutions and the population react to the fact that a pandemic is occurring and thus try to reduce its spread. This component adds constraints to the disease: either through non-pharmaceutical interventions (NPIs) to reduce the probability of transmission (as seen in the infection sub-component) and by reducing the number of contacts of agents depending on the imposed confinement constraint; or through the introduction of vaccines. The latter will not be detailed or used in the thesis, but it is important to know that they exist

remote working, and the closing of restaurants and bars; *Relaxed* NPIs, which involve the relaxation of certain measures (reopening of dining establishments, shops, no more curfews) but maintain mask-wearing and social distancing.

These coefficients will be applied according to the timeline that was implemented during the year 2020 to respect the durations of the lockdowns.

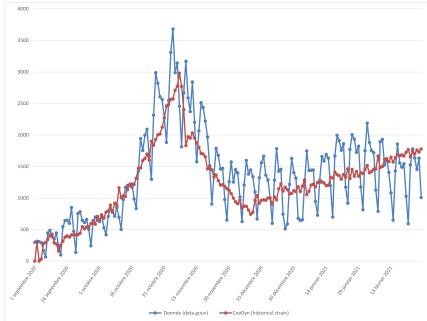


Figure 18 – Previous calibration :  
Second wave only

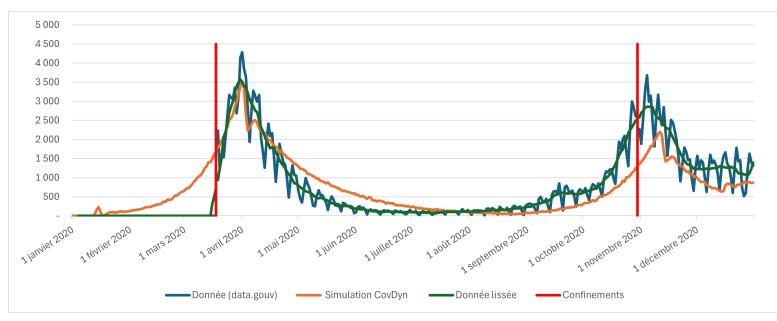


Figure 19 – New calibration : a full year of Covid-19 (1st and 2nd wave)

However, a change to the model needed to be operated. The latter was initially calibrated to reproduce the second wave of Covid in France. Therefore, recalibrating the initial number of infected individuals and the duration of each level of intervention was necessary to start the pandemic at the beginning of 2020 and observe its evolution over a one-year period. This recalibration was carried out, and its visualization appears in Figure 2.

Additionally, a graphical interface has been added to enable interaction between the simulator and the user. The user can now choose from the various scenarios (see below) provided and also have control over some of the simulator's parameters.

## Application: Risk of Death and credit Loan Insurance

To be used by actors of the insurance sector, additional components have been added to the simulator, including the **insurance component**. This component adapts the simulator to take a credit loan portfolio as an input, determine the characteristics of the insured individuals, and inject these characteristics into a sub-population of agents who will act as policyholders. These agents will be assigned a new attribute, *Capital*, which represents the volume of capital at risk.

This sub-population will be tracked over time, and deaths will be recorded. Thus, when these deaths occur, the value of the capital to be paid will be reported and recorded in a file. In the case of this thesis, the entire history of a credit loan portfolio provided by an insurer was used as the basis for constructing a new portfolio with a fixed number of individuals. This creation follows the following steps:

1. A distribution of loan values was adjusted within each age group. These distributions will then be used to perform random drawings when creating the so-called "policyholder" agents. Gamma distributions were ultimately used. It should be noted that the remaining loan capitals were not included in the transmitted database. The initial loan amounts were used instead.
2. The fixed number of insured individuals in this study is 200 000, which corresponds to 1 000 insured individuals in the reduced population. The demographics of the portfolio, stratified by

age group like the contact matrix, are preserved. The policyholders are then distributed in such a way as to reflect the different proportions of the various age groups. The user will also have the option to enter a number of their choice.

Thus, with the newly constructed portfolio, the capital at risk amounts up to an average of **320** million euros for 1 000 policyholders.

After multiple simulations, namely 600, a distribution of the liabilities of the underlying risk can be established, and a 99.5% quantile will be calculated. This *Value-at-Risk* will form an element of calculation to find the value of the SCR, as it only considers the cost factor in its approach (and not its frequency).

① It is important to note that an implicit assumption is made here: the pandemic occurs over a relatively short horizon compared to the duration of a loan. With this, it was considered that the capital at risk does not evolve during this period, i.e., 1 year. However, the approach remains cautious, as the value of the remaining loan capitals decreases over time as the loans are being paid back.

As part of the calculation of this 99.5% quantile, the original Covid-19 scenario uses the parameters estimated by epidemiologists. It will be used as a reference base for the other three stressed scenarios. Indeed, the insurer should study the evolution of the VaR value based on extreme situations, especially those that could jeopardize its solvency. The Covid-19 pandemic experienced in France is just one realization of the unfolding health crisis. The consequences and responses could have been very different depending on the virus's strength.

Therefore, the behavior of the payments is interesting to observe. Moreover, analyzing the reasons why they behave in one way or another becomes easier to explain thanks to agent-based models.

Scenario	VaR (in euros)
Original	825 059,8
Uniform deaths	2 667 912,28
Uniform hospitalizations	449 510,52
Uniform deaths and hospitalizations	2 915 262,35

Evolution of the *Value-at-Risk* depending on the scenario

The original scenario follows the course of the pandemic as it occurred and does not bring any changes to the simulator parameters, which were calibrated based on hospitalization and death data related to the Covid-19 virus. For the simulated portfolio, this equates to a *Value-at-Risk* of approximately 800 thousand euros. This value seems low compared to the total capital at risk, and this is due to the fact that the elderly were the most vulnerable and were the ones who died most frequently. Combined with the fact that the affected age groups represent only a small portion of the capital at risk, the *Value-at-Risk* value seems consistent. The simulator thus accurately reflects this generational aspect, both epidemiologically and in terms of insurance.

## Uniform Death Rates

To simulate pandemics caused by viruses more dangerous than Covid-19, the death probabilities, previously unique for each age group, are standardized by assigning everyone the same death probability, namely the highest ones depending on the compartments ( $H2$  or  $H3$ ). Thus, after being

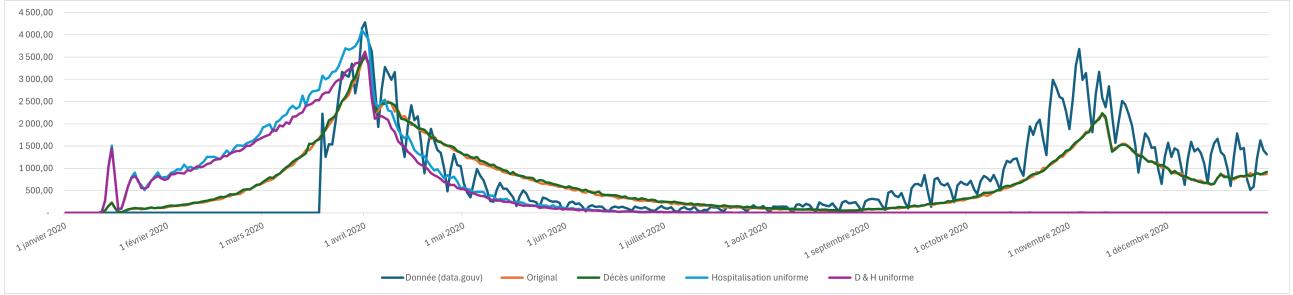


Figure 20 – Average daily hospitalizations for all scenarios

hospitalized, any agent will have the same probability of dying. The hospitalization curve remains, on average, unchanged, while the cumulative death curve exceeds that of the original scenario. This results in a higher quantile, four times greater than that of the previous scenario: more deaths translate to a higher probability that one of these deaths comes from one of the insured agents, and thus more simulations will be non-zero in terms of benefits to be paid. Additionally, young adults, who have the highest capital at risk, die at the same rate as the elderly, who have much lower values.

### Uniform Hospitalization Rate

An equalization of the hospitalization probability was also tested. In this scenario, the death probabilities are reset to their initial values, and the same principle applied in the previous paragraph is applied to the hospitalization probability. Thus, any agent will have the same probability of being hospitalized. The daily hospitalization curve is higher than that of the original scenario but decreases much more rapidly after reaching its peak, so quickly that the pandemic is extinguished and cannot restart with a second wave. This is due to the isolation assumption when an agent enters the hospitalization-related compartments. A plateau in the cumulative death graph also reveals this, with the number of deaths stagnating. This reduced number of deaths implies a quantile of 450 thousand euros. This value is lower considering that a second wave did not occur. With this, the model's limits are reached: the isolation assumption in the hospital block becomes absurd, as the number of available hospital beds would be saturated. Additionally, the scale of the crisis would be accompanied by the collapse of the healthcare system.

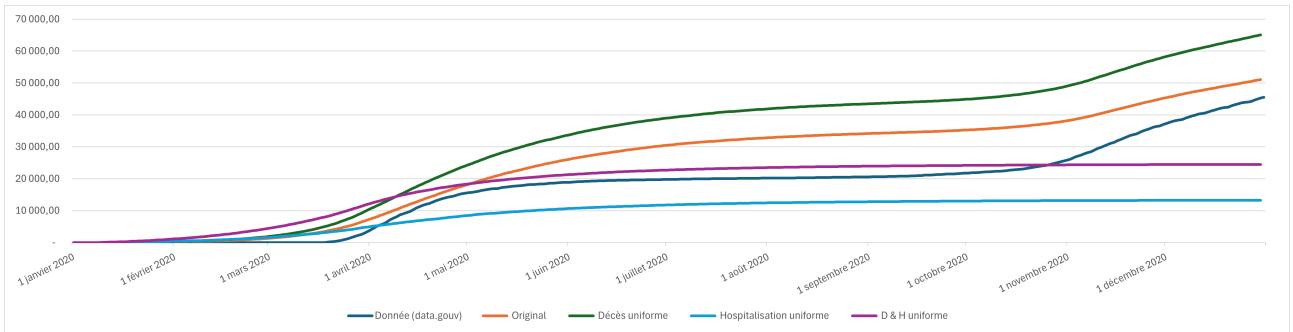


Figure 21 – Average daily deaths for all scenarios

### Uniform Death and Hospitalization Rates

The last scenario is a mix of the previous two, with uniform probabilities in terms of hospitalization and death. Similarly, it exhibits strong growth and rapid decline, limiting the pandemic to a single wave. However, the number of deaths in this case is higher than the previous one and presents the

highest quantile with more than 2.9 million euros. Even with a lower number of cumulative deaths than the first two scenarios (Original and Uniform Deaths), it is the "right" people who die, i.e., those who have just started repaying their loans, i.e., young adults. The capital at risk is generally higher among young insured individuals, and thus, by giving everyone the same probability of being hospitalized and dying, a portion of the deaths will represent the age groups where the capital at risk is not negligible. In the second stress scenario (Uniform Hospitalizations), all agents indeed had the same probability of being hospitalized, but only the agents from advanced age groups were "privileged to die," i.e., the agents whose capital at risk is lower.

## Random Covid

All parameters of the CovDyn model are predefined and static, representing the Covid-19 pandemic in France. To add volatility, value intervals are created around the parameters  $p_A$ ,  $p_H$ ,  $p_{ICU|H1}$ ,  $p_{D|H2}$ , and  $p_{D|H3}$  by adding and subtracting 20% from their original value. This value is arbitrary. Two sampling methods are considered: Gaussian and uniform. The uniform method is chosen for its simplicity and caution, giving equal weight to all values within the interval. It thus includes the extremes in the sampling.

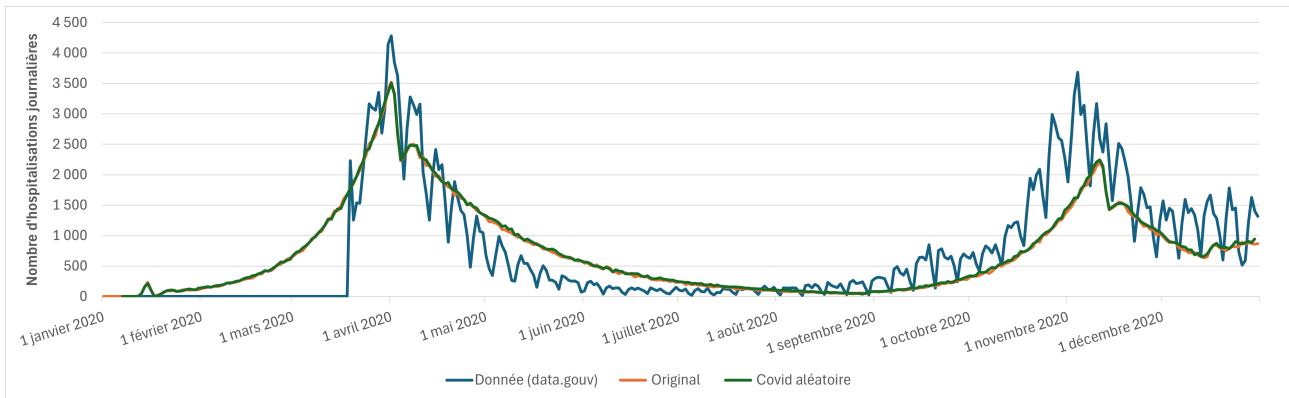


Figure 22 – Average behavior equal to that of the original Covid-19 in terms of daily hospitalizations.

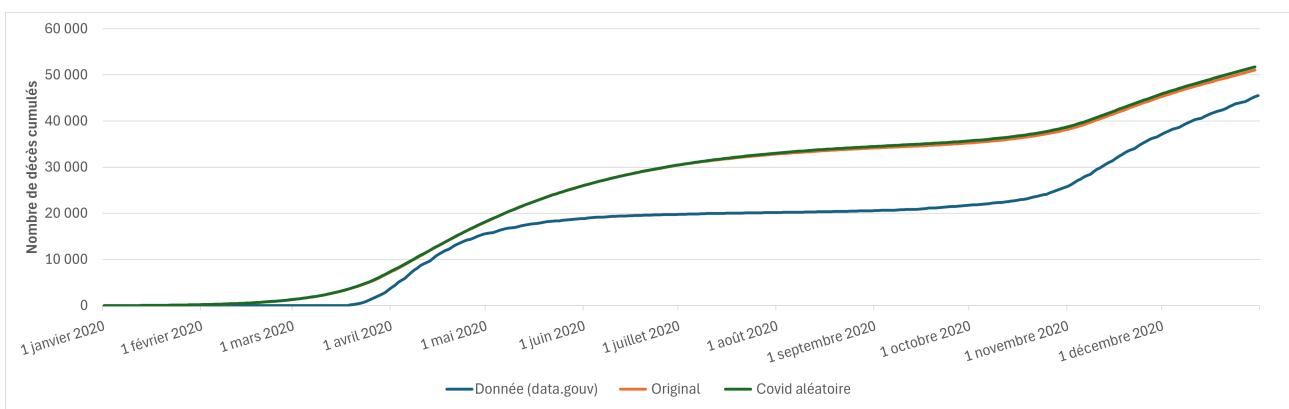


Figure 23 – Average evolution of cumulative deaths identical to that of the original Covid-19.

The respective quantile is slightly higher and reaches 941,614.1. The value remains close to that of the central scenario. The difference could be caused by the additional volatility produced by the parameter sampling. To be certain, a larger number of simulations would be needed.

## Postponement of the start of the lockdown

Delay	VaR (in euro)	Evolution (in %)
0 days (Central)	825 059,8	
5 days	1 109 803,16	+ 34
10 days	1 139 935,23	+ 38
15 days	2 111 002,41	+ 155

Table 3 – Value of the *Value-at-Risk* based on the postponement of the date of the first lockdown

From a political perspective, the value of the Value-at-Risk (VaR) was estimated based on three time delays: a delay of 5 days, a delay of 10 days, and a delay of 15 days. That is, instead of starting on March 18, the first lockdown will be triggered on March 23, March 28, or April 2 according to the simulation.

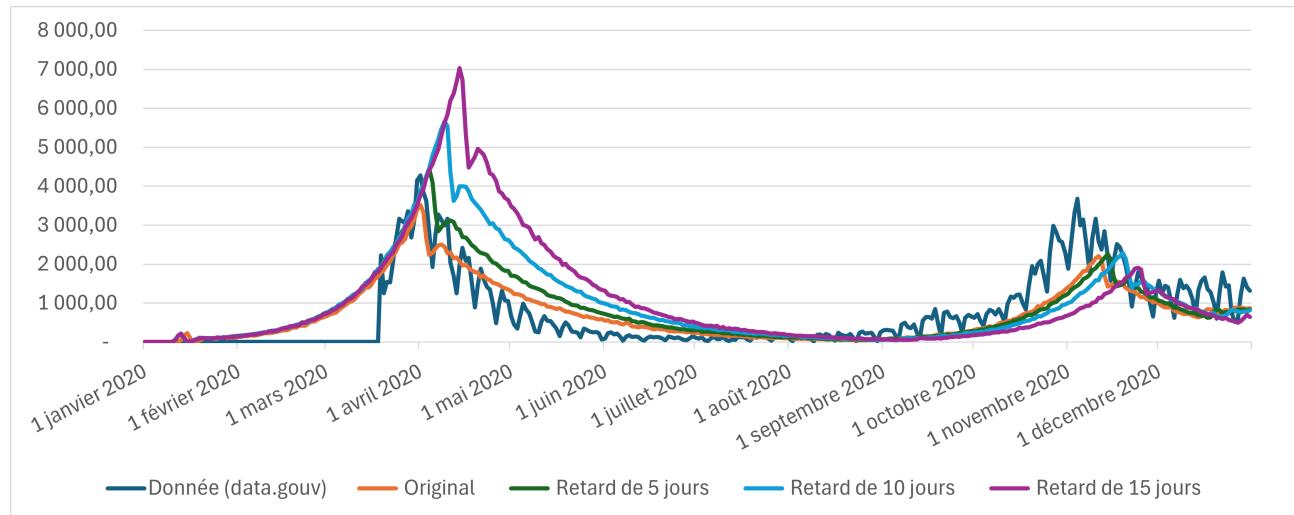


Figure 24 – Evolutions with the same patterns but shifted to the right and with higher peaks

The curve representing the number of daily hospitalizations shifts to the right, reaching a more pronounced peak during the first wave as the date of the first lockdown is postponed. This ensures the maintenance of the virus transmission rate, allowing a considerable number of patients to be admitted to the hospital. The impacts on the Value-at-Risk (VaR) are expressed above. It is higher the more the lockdown date is postponed.

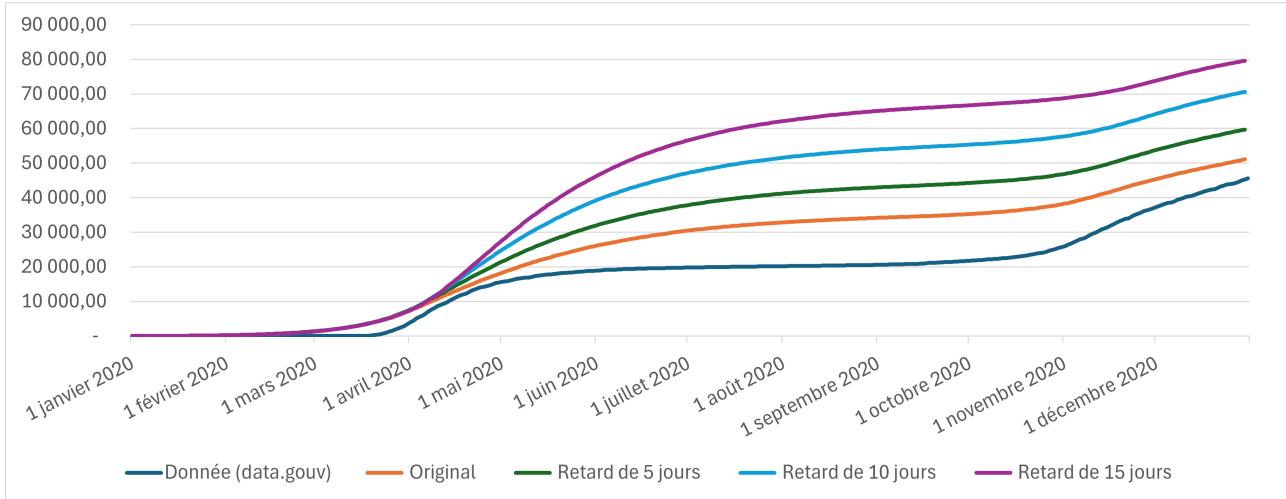


Figure 25 – Evolution of cumulative deaths based on the postponement

The age group most affected here is the 80P group, which has the highest probability of hospitalization and is therefore the most vulnerable (as in the central scenario). Despite a significant increase in the number of deaths, the *Value-at-Risk* does not exceed that of the previously discussed scenarios, as most of the affected individuals would belong to the elderly category, thus preserving the age segments where the main volume of capital at risk is located, i.e. the younger policyholders.

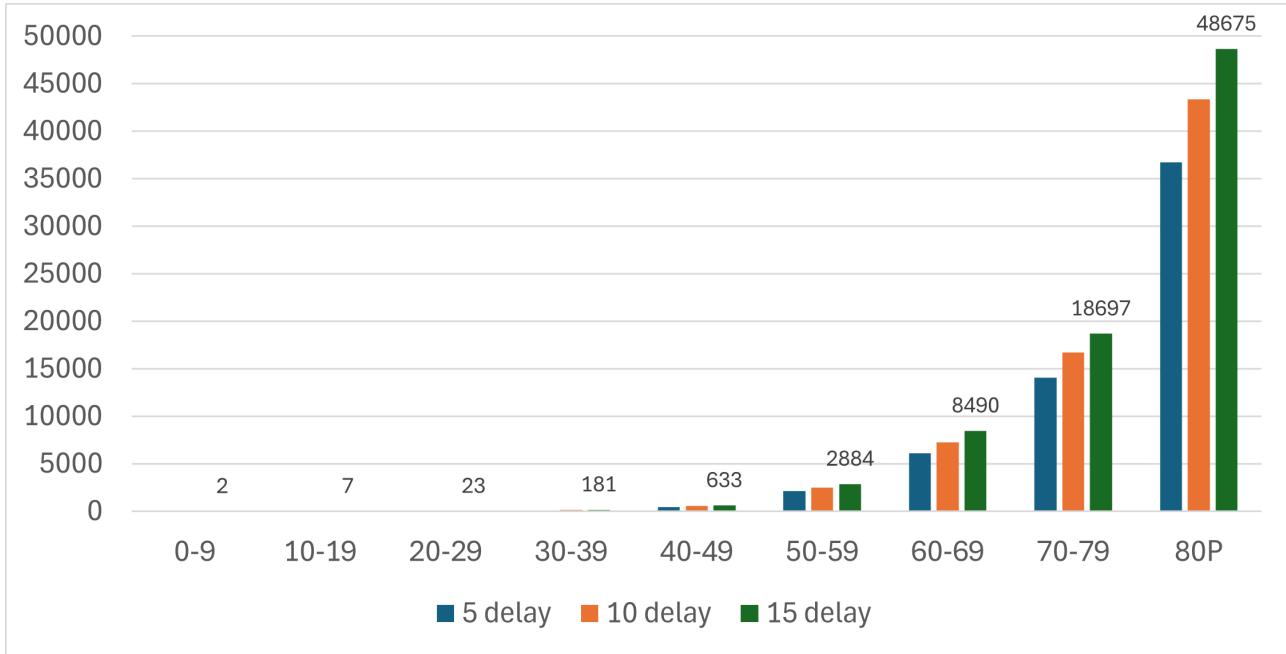


Figure 26 – Distribution of deaths according to the degree of postponement - Same distribution shape as the original scenario

## Limits and Conclusions

1. Deaths can only occur after hospitalization. Agents outside the hospital block cannot end up

- in the death compartment. Therefore, there could be an underestimation of the total number of deaths;
2. The calibration of parameters concerning non-pharmaceutical interventions is based on the second wave and therefore does not fully represent the first wave. Better adjustment can be achieved through recalibration of these coefficients;
  3. Some scenarios may deviate from reality but still propose a thought process to add components or refine existing ones. For example, an addition would be to integrate a limit on the number of agents that can be present in the hospital block. Another example would be to relax the total isolation assumption in this block and question the validity of this assumption;
  4. The initial loan amounts do not represent the current risk borne by the insurer. To have a more detailed and thorough analysis, a borrower portfolio containing the remaining loan capitals would be preferred.

Finally, the previous results show that the insurer must have a prudent vision and test its short-term solvency, especially in the case of a pandemic. It must see how more extreme cases could jeopardize this solvency.

It must especially conduct stress tests concerning scenarios where the death rate is particularly high. The scenario increases the risk for the insurer, especially if it occurs simultaneously with a situation of high hospitalization probabilities.

Agent-based models allow, on the one hand, to integrate an insurance-focused perspective, provide elements for calculating the 99.5% quantile, and offer analysis insights to understand the behavior of a certain scenario and thus the repercussions in terms of benefits generated by it.

It is therefore important and necessary to plan this type of exercise and apply shocks to the parameters themselves to study the evolution of the VaR value.



# Remerciements

Je tiens à remercier tout d'abord Romain GAUCHON qui m'a proposé le sujet de ce mémoire, qui l'a suivi du début jusqu'à sa toute fin, sillonnant entre toutes les complications, que ce soit de loin ou de près. Je tiens également à remercier tous ceux qui ont suivi ce mémoire : Anne EYRAUD pour ses conseils, Antoine BURG pour s'être embarqué dans l'aventure avec nous, Elise CALLAC qui a repris le sujet (même si l'épidémiologie ce n'était pas son truc), Aurélie TREILHOU qui a cru en moi et qui m'a soutenu pour présenter mes résultats au JoCo à Bruxelles, à Mouna AIT-OMAR ROSENBAUM pour ses retours et son accompagnement (même en vacances !) et à Elise LAURENT pour des remarques extrêmement pertinentes et les parties de pêche aux fautes d'orthographe !

Je remercie également Imran BOUKHENFER et Léo ALLERS, ainsi que tout le bureau de Périclès Actuarial et Périclès Group Development pour le milieu de travail engageant.

Je tiens aussi à remercier ma famille qui m'a soutenu depuis le début de mon parcours scolaire jusqu'à la fin de mes études. À ma mère, à mon père, à mon frère, à mes trois tantes, je vous remercie pour tout ce que vous avez pu m'offrir.

À ma cousine Léa, ma cousine Dana, mon cousin Jad, ma tante et mon oncle qui ont été ma famille de sang ici en France, loin de la mienne. À tous mes amis que j'ai pu faire depuis mon arrivée : Alexandre (de Reims), Lola, Chloé, Laureline, Aïssatou, Ahmed, Saphir, Elsa, Alexandre (de Lyon), Anaïs, Maëlle et Francky. C'est incroyable de vous avoir dans ma vie; je n'aurais pas pu souhaiter mieux.

Et un dernier merci, le plus grand des mercis à tous mes professeures et professeurs. Je termine ainsi mes études et c'est en grande partie grâce à vous que je suis là aujourd'hui. Je n'étais potentiellement qu'un élève parmi des centaines voire des milliers, mais pour moi, vous serez toujours dans mes pensées. À mes enseignants, à tous ceux qui m'ont donné envie d'apprendre, à :

ABDALLAH Mohamad, ABOU RICH Sawsan, BADER Caroline, BEJJANI Pascal, BLANC-BRUDE Gilles, BOURRACHOT Alexandre, CHALL Sari, CHAMPAGNE Caroline, CHAMSEDDINE Ghada, CLOT Denis, CLUZEL Valérie, DAGHER Roger, DEVOUGE Claire, DIB NEHME Zeina, DOROBANTU Diana, DUMAS Georges, EID Jessy, EL KHECHEN Rola, FANOUS KAROUT Amira, GHOSON Lama, GHOSON Adnan, HAIDAR Imane, HAIDAR Chams, HAIDAR Samia, HANADY Salhab, HARAJLI Zeinab, HINDAWI Mona, HOTEIT Wafaa, HOUMANI Sanaa, JACQUELIN KAB-BANI Delphine, JIAO Ying, KABBANI Mosbah, KAMBRIS Siham, KAMBRIS Samar, KATTAN Yolaine, KREIDIEH Husn, LASTAKOWSKI Henri, LE NESTOUR Baptiste, MAKKI HALLAL Ramla, MASIELLO Esterina, MEHDI Lina, MRAD Zeina, NORIS Marion, NOVAK Alice, PENOT Alexis, PEU DUVALLON Mayliss, PIETTE Pierrick, PINET Patrick, PLANCHET Frédéric, PLU Vin-

cent, PRÉVOST Thierry, RAAD Yasmina, RAWAS Leila, RIBEREAU Pierre, ROBERT Christian, SAKHER May, SHOUKRI MRAD Rayan, SIDANI Hatem Gilles, SLEIMAN Nada, SUKKARIEH Fatima, TANNOUS Mhanna, WEHBE Maha, WERQUIN Marie; à Sanaa, Hana, Rana, Mayssa, Maya, Marleine, Katia, Mariam, Najwa, Madonna et Hassana.

كَادَ الشَّعْلَمُ أَنْ يَكُونَ رَسُولاً

قُمْ لِلْمُعَلَّمِ وَفِي التَّبْجِيلِ

يَنِي وَيُنِشِّئُ أَنْفُسًا وَعُقُولًا

أَعْلَمْتُ أَشْرَفَ أَوْ أَجَلَّ مِنَ الَّذِي

أَحْمَدُ شَوْقِي

# Contents

<b>Note de Synthèse</b>	<b>5</b>
<b>Synthesis note</b>	<b>17</b>
<b>Remerciements</b>	<b>29</b>
<b>Table des matières</b>	<b>31</b>
<b>Introduction</b>	<b>33</b>
<b>1 Pandémies et assurances</b>	<b>35</b>
1.1 Épidémies, pandémies: une histoire sans fin ? . . . . .	35
1.2 Risque pandémie et Solvabilité II . . . . .	41
1.3 Calcul du SCR par formule standard . . . . .	42
1.4 Calcul du SCR par modélisation . . . . .	44
<b>2 Épidémiologie</b>	<b>45</b>
2.1 Histoire . . . . .	45
2.2 $R_0$ , $R_e$ et définitions . . . . .	47
2.3 SARS-CoV-2 . . . . .	50
2.4 Règle des trois unités: de lieu, de temps et d'action . . . . .	53
2.5 Les modèles compartimentaux . . . . .	57
2.6 Modèles agents . . . . .	60
2.7 Modèle <i>CovDyn</i> . . . . .	63
<b>3 Risque décès et assurance emprunteur</b>	<b>73</b>
3.1 Portefeuille emprunteur . . . . .	74

3.2	Modifications du simulateur . . . . .	79
3.3	Covid-19 et co. . . . .	83
3.4	Covid aléatoire . . . . .	92
3.5	Retardement des confinements . . . . .	94
<b>Conclusion</b>		<b>99</b>
<b>Bibliographie</b>		<b>101</b>
<b>Annexes</b>		<b>103</b>
<b>A</b>	<b>Visualisation du fonctionnement du SIR</b>	<b>105</b>

# Introduction

Dans le théâtre grec, la **catastrophe** est la dernière des cinq parties d'une tragédie où le héros recevait sa punition. Emprunté au latin *catastropha* « coup de théâtre » attesté au sens de « dénouement (d'une tragédie ou d'une comédie) » depuis le IV<sup>e</sup> siècle et lui-même emprunté au grec *καταστροφή* « bouleversement » et « fin, dénouement », le terme désigne un événement dont les conséquences sont particulièrement graves, voire irrémédiabes, et qui peut être d'origine naturelle ou humaine. On discerne notamment les situations qui inscrivent des pertes humaines et des dévastations à grande échelle. Hors-scène, dans les coulisses, les risques catastrophes prennent une multitude de formes à ampleurs dévastatrices survenant à des moments difficiles à déterminer à l'avance. Peuvent être citées les catastrophes naturelles, déclinées en catastrophes climatiques, inondations, raz-de-marée, catastrophes sismiques, glissements de terrain, éruptions volcaniques, et les catastrophes d'origine humaine comme les catastrophes industrielles, nucléaires, incendies, effondrements, guerres.

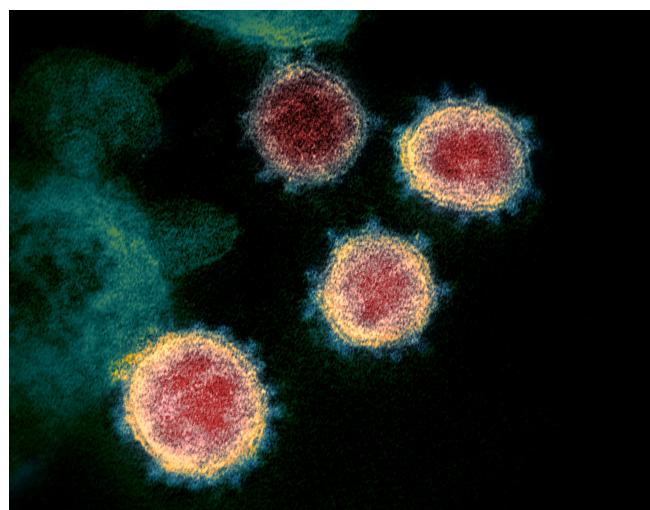


Figure 27 – Image colorisée au microscope électronique du SARS-CoV-2 isolé d'un patient aux États-Unis. Cellules cultivées en laboratoire. © NIAID-RML Creative Commons.

Une catastrophe vient bouleverser la continuité routinière de la vie, événement aux conséquences particulièrement graves qui en résulte ruine ou désastre. En d'autres mots, elle est annoncée par une interruption soudaine et rompt de façon non négligeable la vie quotidienne des victimes de cet événement. C'est ce qu'a fait la pandémie causée par le virus SARS-CoV-2 par exemple: confinements obligatoires à travers le monde, changement d'approches concernant le commerce, l'enseignement, le travail, les interactions sociales, l'assurance, etc. Elle a secoué les infrastructures sociales, les fondements des interactions à toute échelle. Malgré les survenances ultérieures sous différentes légions de maladies, dont la grippe espagnole, il reste difficile de prédire l'apparition d'une pandémie ainsi

que sa portée sur l'activité humaine. L'épidémiologie est une science qui essaye de répondre à de nombreuses questions sur la fréquence de ces évènements, leurs modes de diffusion, leurs sévérités. Il est important de pouvoir évaluer le rayon de diffusion, la sphère de ces intermissions.

Cette irrégularité exceptionnelle vient troubler le quotidien des (ré)assureurs en changeant l'environnement dans lequel ils travaillent et dans lequel leurs assurés confrontent les aléas et les imprévisibilités de la vie. Durant une pandémie, les besoins de soins augmentent avec l'afflux de maladies nouvelles. Leurs caractéristiques seront encore inconnues au début d'une pandémie, par exemple la sévérité qui souvent peut induire une augmentation de la mortalité. Cette hausse devrait être prise en compte par les assureurs pour mieux apprécier leurs risques.

Il est donc nécessaire de prendre en compte le risque catastrophe, en considérant ici le risque pandémie, pour faire face à sa survenance possible, quantifier les expositions que pourraient avoir un assureur et réduire la probabilité de ruine potentiellement induite et de maintenir la solvabilité de l'entité. Pour cela, il faudrait avoir un moyen d'évaluer les prestations ou pertes induites et les intégrer dans les différents calculs de provisions, de primes, etc. La réglementation propose certaines mesures et formules mais il serait plus efficace de prendre en compte la spécificité de chaque exposition, de leurs caractéristiques. En effet, la composition des portefeuilles peut avoir différents impacts vu la structure de létalité de la pandémie suivant l'âge par exemple, pièce fondamentale dans la modélisation du risque pandémie.

Par la perspective de l'épidémiologie, ce mémoire a pour objectif de développer un outil pour simuler l'impact d'un choc pandémique sur la distribution des prestations décès et l'évaluation du risque catastrophe. Il sera appliqué ici à la pandémie de la Covid-19 et concerne un assureur détenteur d'un portefeuille d'assurance emprunteur. Le risque étudié est le risque décès. Dans une première partie, le contexte sera présenté et la réglementation en vigueur va être rappelée. Ensuite l'épidémiologie sera présentée ainsi que le contexte épidémiologique de la crise sanitaire de la Covid-19. Seront alors détaillés les outils et modèles communément utilisés en épidémiologie avec en particulier une concentration sur le modèle *CovDyn*, un modèle agent qui sera finalement utilisé pour simuler l'impact d'une pandémie sur un assureur détenteur d'un portefeuille emprunteur.

# Chapter 1

## Pandémies et assurances

### 1.1 Épidémies, pandémies: une histoire sans fin ?

Les épidémies et pandémies ne datent pas de l'époque contemporaine même si la mémoire peut ne proposer que des exemples plus récents qui ont frappé la population mondiale. L'humanité, sur l'échelle biologique, a été façonnée par la rencontre avec les microbes qui a ravagé des populations à différentes périodes de l'existence et qui a modelé le système immunitaire de l'humain ainsi que son patrimoine génétique. Aucune survenance n'a encore été répertoriée avant le début du néolithique. Ce n'est qu'avec l'avènement de l'âge de la pierre polie et de sa révolution néolithique caractérisée par l'adoption de groupes de l'élevage et de l'agriculture que les premières endémies apparaissent. Il faudra attendre le Moyen-Âge, l'époque moderne et la renaissance pour voir apparaître des premiers exemples d'épidémies et de pandémies.

Et comme l'explique Antoine FLAHAULT, médecin épidémiologiste et directeur de l'Institut de santé globale à l'université de Genève: « tout germe peut potentiellement venir à tout endroit de la planète et déclencher des épidémies. » Près de 75% des virus responsables d'épidémies sont d'origine animale. Notamment les primates en ce qui concerne le SIDA, les rongeurs pour la variole, les poux et les puces pour le typhus, les chauves-souris pour les coronavirus, les oiseaux pour la grippe aviaire, etc. Avoir une attention toute particulière devient capital pour les appréhender. Et à l'ère du réchauffement climatique, la fréquence d'apparition pourrait augmenter.

Cette récurrence est importante à souligner car personne ne peut savoir si une maladie va atteindre ou non le stade de pandémie. Cette incertitude concerne autant les épidémiologistes, les chefs d'État mais aussi les assureurs. Ce risque reste encore mal défini dans le domaine de l'assurance. Cela pousse alors à avoir un regard vers le passé, à l'encontre des épidémies et pandémies précédentes et à collecter les enseignements acquis sur celles-ci pour préluder un tel évènement ainsi que ses répercussions.

#### 1.1.1 La grippe

La grippe est une infection respiratoire contagieuse due au virus influenza, dont la particularité est la grande variabilité génétique. Le virus se transmet par les aérosols, au moyen de microgouttelettes excrétées par un patient qui tousse. Il pénètre dans l'organisme par voie respiratoire. Les symptômes apparaissent entre 1 et 4 jours après la contamination, l'infection durant généralement une semaine au bout duquel les sujets guérissent avec un traitement symptomatique.

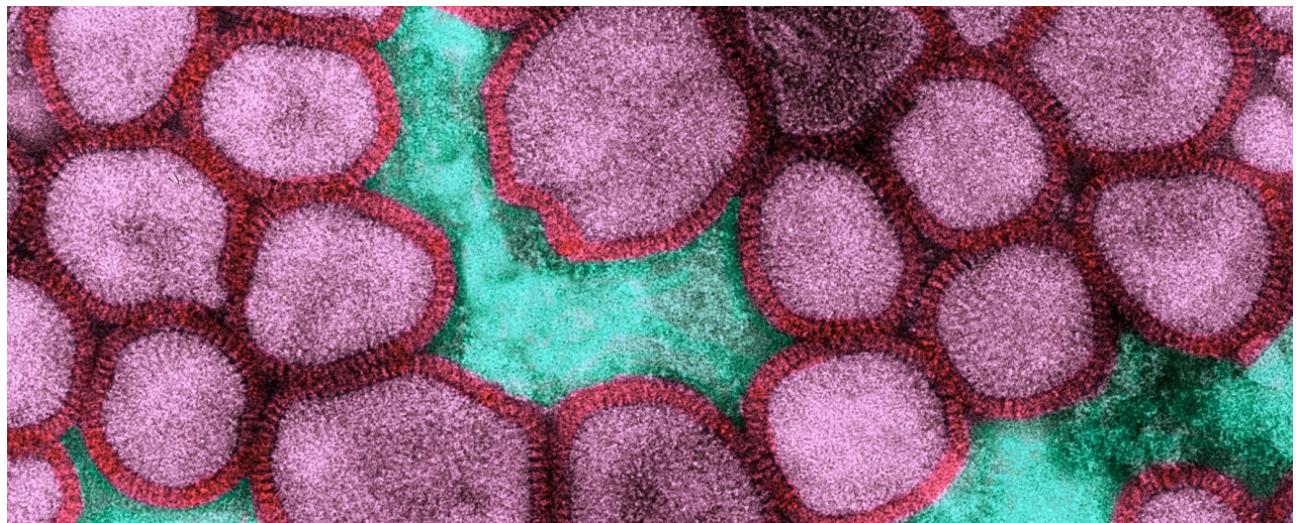


Figure 1.1 – Virus de la grippe pandémique A (H1N1) 2009 produits dans une culture de cellule musculaire humaine - © Institut Pasteur

Et cette maladie ne peut être mentionnée sans évoquer la pandémie la plus sévère: la **grippe espagnole**, tenant son nom du fait que l'Espagne, neutre durant la guerre, était la seule à publier ses chiffres et à mentionner l'épidémie survenue. Cette grippe se répandit très rapidement, touchant plus du tiers de la population mondiale entre 1918 et 1919 et devient alors plus meurtrière que la Première Guerre mondiale avec 20 à 50 millions de morts dans le monde. Sa deuxième vague à l'automne de 1918 a introduit une mutation de qui le rendit extrêmement virulent et entraîna des symptômes sévères à toutes les personnes touchées.

Les épidémies de grippe évoluent selon une certaine saisonnalité, et apparaissent durant les périodes de froid où le virus est plus puissant. En France, la grippe saisonnière touche 2 à 6 millions de personnes et est responsable d'environ 10 000 décès chaque année.

### 1.1.2 Le virus SRAS

Le SRAS (syndrome respiratoire aigu sévère) est une maladie infectieuse grave et transmissible qui a émergé en ce XXIe siècle, causée par un virus appartenant à la famille des coronavirus, le SARS-CoV-1. Au départ nommé pneumopathie atypique, le SRAS se caractérise par une fièvre supérieure à 38°C, de la toux sèche, de l'essoufflement et des difficultés respiratoires. Avec un cercle d'étude restreint, l'OMS considère le taux de létalité à 15% et qu'il peut dépasser 50% chez les personnes âgées.

La pandémie débute en 2002, en Chine, et éclate au niveau mondial en 2003 faisant plus de 8 000 cas et près de 800 morts. Grâce à une mobilisation internationale sans précédent suivant l'alerte mondiale sonnée le 12 mars 2003 par l'OMS, l'épidémie a pu être endiguée par des mesures d'isolement et de quarantaine. Le réservoir animal du coronavirus du SRAS a été identifié comme étant une chauve-souris insectivore. L'hôte intermédiaire qui a permis le passage du virus à l'homme est la civette palmiste masquée, animal sauvage vendu sur les marchés et consommé au sud de la Chine.

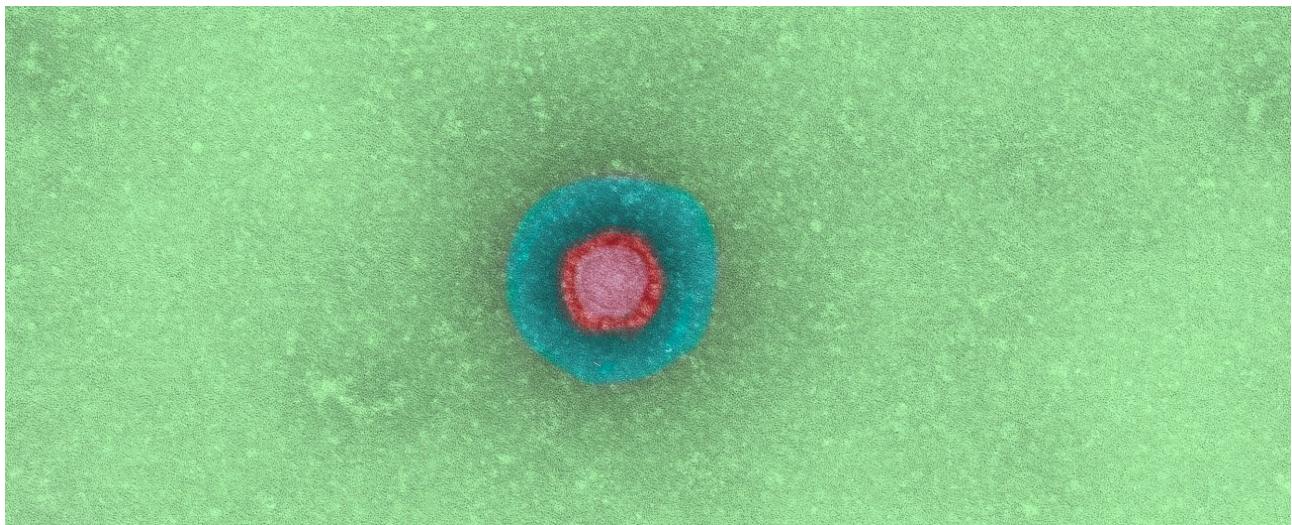


Figure 1.2 – Le coronavirus agent causal de la pneumopathie atypique ou SRAS - © Institut Pasteur

### 1.1.3 Le MERS-CoV

Le MERS-CoV (Coronavirus du Syndrome Respiratoire du Moyen-Orient) est un virus identifié pour la première fois en Arabie Saoudite en avril 2012. Comme le SARS, il a été détecté chez des chauves-souris. Celles-ci pourraient constituer le réservoir naturel mais les dromadaires agiraient comme hôtes intermédiaires qui seraient la source de l'introduction chez l'homme. Ainsi, des transmissions occasionnelles à l'homme se sont produites au contact de dromadaires ou lors de la consommation de lait cru de dromadaire. Depuis 2012, 1 219 cas de MERS-CoV ont été détectés dont 449 morts.

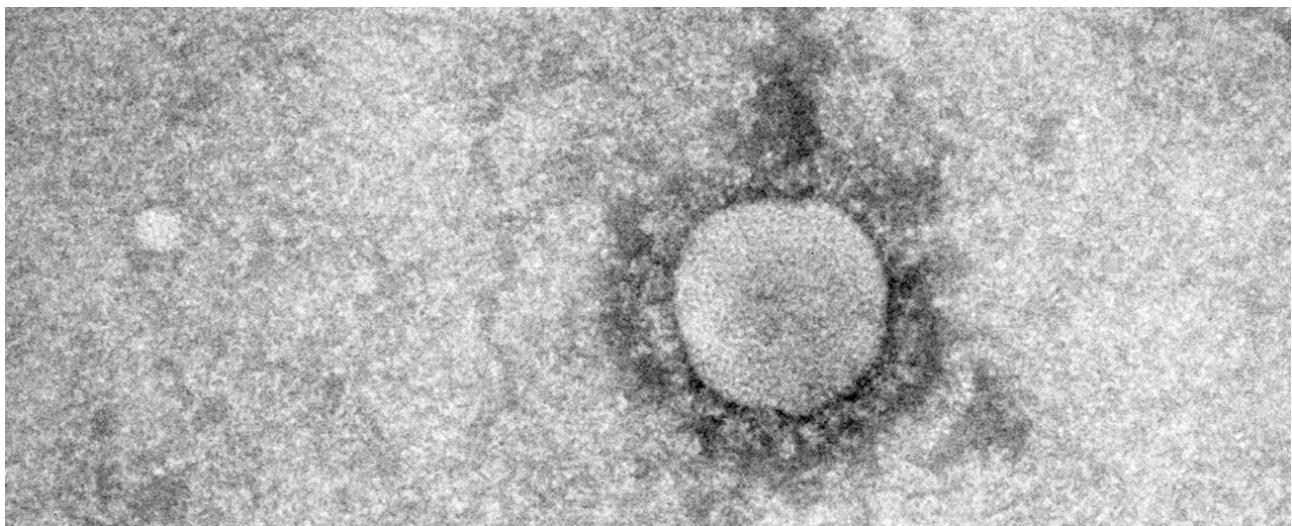


Figure 1.3 – Particule virale de Coronavirus MERS-CoV en microscopie électronique - © Institut Pasteur/Meriadeg Le Gouil

Le virus peut être très dangereux quand il est porté par les personnes déjà affaiblies par d'autres pathologies : diabète, insuffisance rénale, infection pulmonaire chronique, immunodépression. Le taux de létalité estimé est de 30% environ. Ses symptômes ressemblent à ceux du SARS.

En Corée du Sud, le premier cas avéré de MERS-CoV a été constaté le 20 mai 2015 chez un

patient de retour du Moyen-Orient. Ce cas a contaminé directement ou indirectement plus d'une centaine de personnes dans les hôpitaux qu'il a fréquentés. La transmission en lien avec ces différents établissements a été favorisée par la grande fréquentation de ce type de lieu. À partir du 16 juin 2015, aucune transmission n'a été observée en dehors des hôpitaux.

#### 1.1.4 Le VIH/SIDA

Avec environ 6 000 nouvelles contaminations recensées chaque année et 3,7 millions de personnes infectées, le VIH reste un fléau surtout en l'absence de traitement malgré le progrès de la prise en charge des patients séropositifs. Le VIH cible et détruit les lymphocytes T, globules blancs essentiels jouant un rôle important dans le mécanisme du système immunitaire. Si aucun traitement n'est mis en place, le syndrome de l'immunodéficience acquise (SIDA) apparaît en moyenne sept ans après l'infection, dernière étape de la maladie. À ce stade, la majorité des lymphocytes ont été détruits par le virus et cela laisse la place aux pathogènes normalement inoffensifs et frivoles de se développer et causer de vrais dégâts que le système immunitaire ne peut combattre.

Les symptômes de la maladie se voient évoluer selon le schéma suivant. La première phase peut se caractériser par des signes cliniques similaires à ceux de la grippe : cette phase est appelée primo-infection. Après cela, débute la phase asymptomatique qui dure plusieurs années. Le patient reste contagieux. Le virus se transmet par voie sexuelle et sanguine et peut même être transmise de la mère à l'enfant.

Depuis son apparition en 1981, les différentes organisations essayent de trouver un moyen d'éradiquer cette épidémie mondiale. Et c'est là qu'ont commencé à être mis en pratique les modèles SIR.

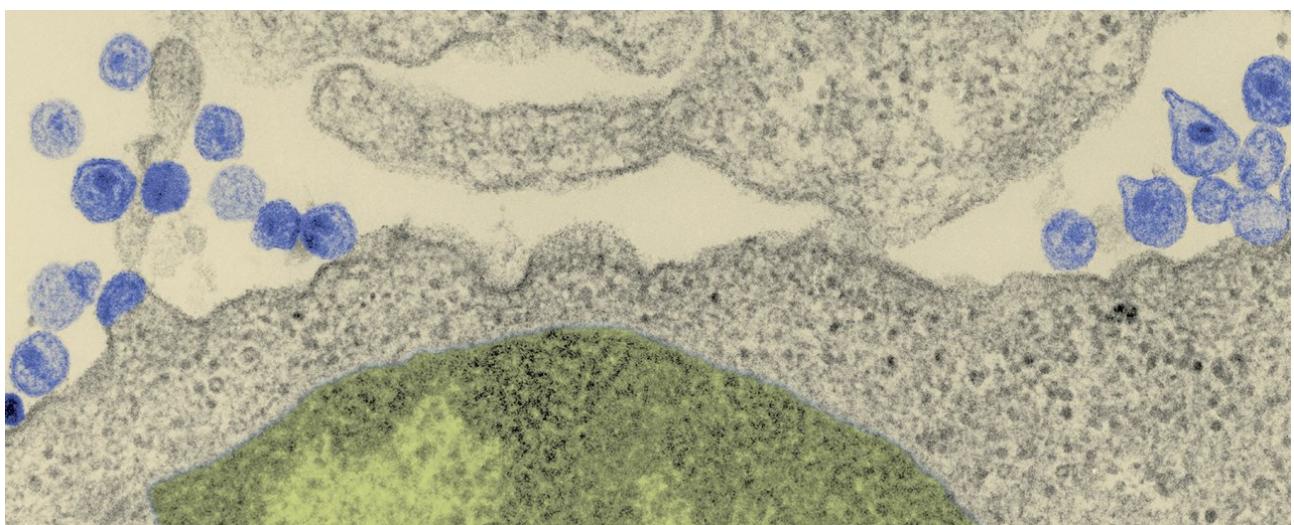


Figure 1.4 – Détail d'une section de lymphocyte T infecté par le VIH, isolé chez un malade en 1983 -  
© Institut Pasteur

Or, la pandémie récente de la Covid a montré que ces modèles SIR ne seraient pas suffisants et que l'intégration de comportements de la population est capitale.

### 1.1.5 Covid-19 : débuts, évolutions et calendrier

Démarrant comme une mystérieuse épidémie de pneumonie dans la ville de Wuhan, province de Hubei, en Chine, c'est la topologie planétaire et instantanée qui donna gravité aux conséquences de la maladie de la Covid-19. Et à cause d'un départ relativement tôt en Italie, l'Europe a été l'un des continents le plus sévèrement touché avec plus de 2,10 millions de décès confirmés [7]. Elle s'est traduite en France par quatre vagues épidémiques, au printemps et à l'automne 2020, au début de 2021 et pendant l'été 2021. Entre mars 2020 et septembre 2021, 460 000 personnes ont été hospitalisées et 116 000 sont décédées à l'hôpital ou en Ehpad lors d'une infection à la Covid-19 [12].

En ce qui concerne les mesures sanitaires pour endiguer l'épidémie, le premier confinement débute officiellement le 17 mars à midi même si certaines mesures ont déjà été mises en place quelques jours avant cette date comme la fermeture des restaurants. Durant ce confinement, les individus ne pouvaient pas quitter leur domicile et les déplacements étaient restreints au strict nécessaire comme les courses alimentaires, les nécessités de premiers soins, ainsi que pour des activités physiques individuelles, d'hygiène canine, etc. Le confinement se voit levé le 11 mai 2020. Le deuxième confinement, moins strict, s'étale entre le 30 octobre et le 15 décembre. Les établissements scolaires sont ouverts, ainsi que les classes préparatoires et les BTS. De plus, certains secteurs sont autorisés à poursuivre leur activité.

Ces mesures sanitaires ont été mises en place afin d'empêcher le virus de prendre en ampleur : la propagation du virus doit être entravée pour limiter ses conséquences. Par exemple, outre les décès d'origine Covid, les hôpitaux possèdent une capacité limitée en ce qui concerne le nombre de lits proposés: une entrée accrue en hospitalisation et de demande de soins intensifs peut surcharger l'occupation des lits.

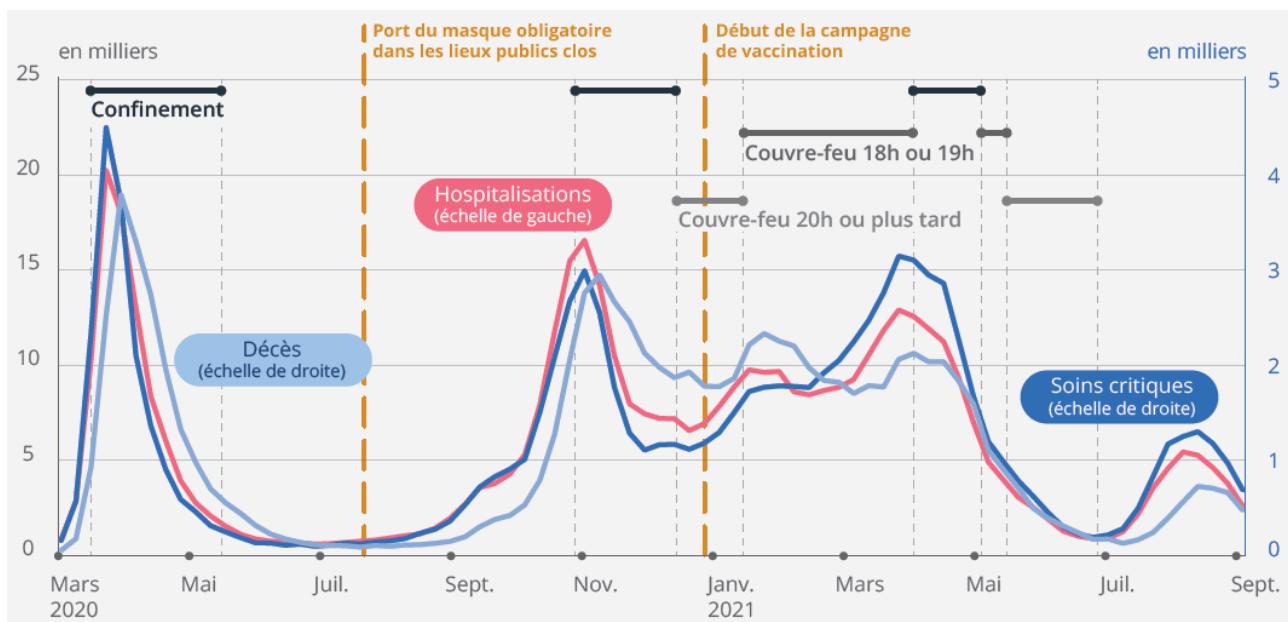


Figure 1.5 – 4 vagues et 4 pics de plus en plus bas mais qui durent plus longtemps. Concerne: le nombre hebdomadaire de nouvelles hospitalisations, d'entrées en soins critiques et de décès à l'hôpital liés à la Covid-19. [12]

Comme le montre la figure, les amplitudes des différentes vagues épidémiques se voient décroître alors que les durées des trois premières ne font que s'allonger. Les périodes durant lesquelles le nombre d'hospitalisations hebdomadaire dépasse 5 000 passent de 6 semaines lors de la première vague à 12 semaines lors de la deuxième et à 19 semaines lors de la troisième. La quatrième vague vient après les

premières campagnes de vaccination début juillet 2021 et ne dure que 2 semaines avec plus de 5 000 hospitalisations.

De plus, a été étudié le risque d'hospitalisation et de décès lors des deux premières vagues, donc précédant la disponibilité de vaccins, et l'âge a été retenu comme facteur de risque. D'après [9], une augmentation exponentielle des risques d'hospitalisation et de décès a été observés en fonction de l'âge. Par exemple, le risque d'hospitalisation chez les 85 ans et plus était 8 fois plus élevé par rapport aux 40-44 ans et le risque de décès à l'hôpital était 200 fois plus élevé.

Cette létalité qui se rajoute a aussi touché tous les coins du globe, les spectateurs devenant acteurs dans la propagation et la vie de la pandémie qu'est la Covid-19.

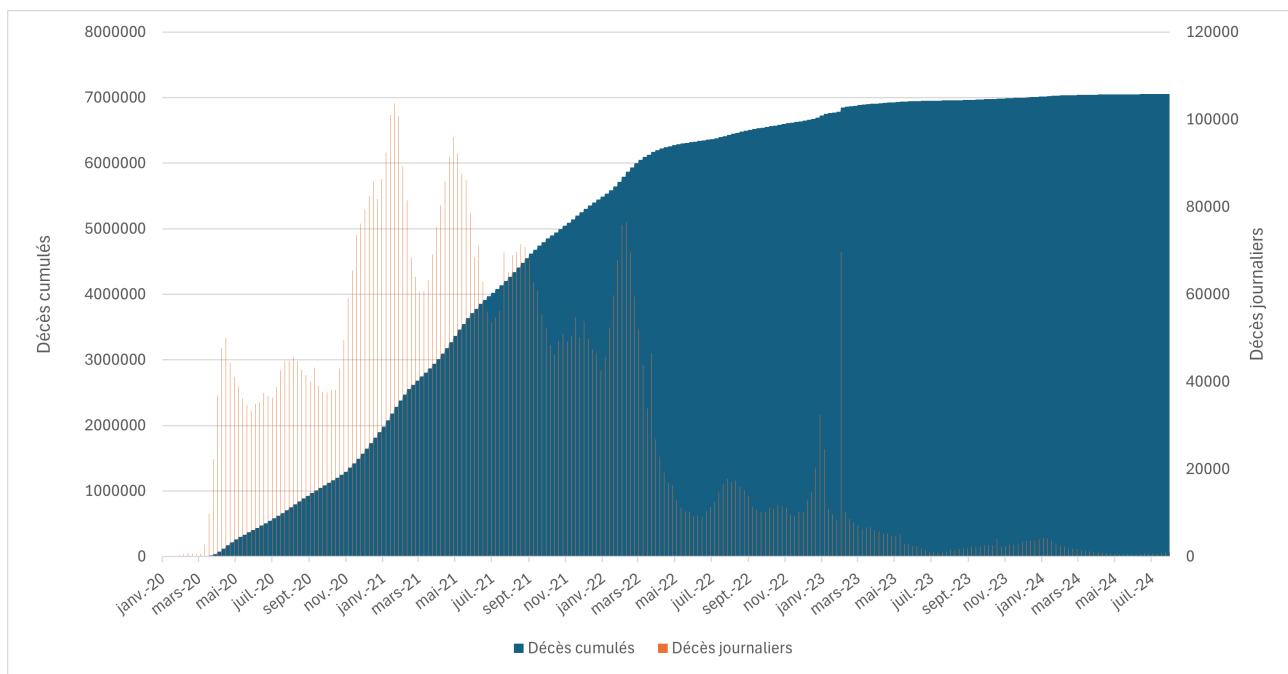


Figure 1.6 – Plus de 7 millions de décès causés par la Covid-19 à travers le monde

Cette évolution peut se voir ralentir à certains instants grâce aux confinements et au développement de vaccins et s'accélérer lors des relâchements des mesures sanitaires.

La population mondiale n'est donc pas à l'abri de nouveaux pathogènes capables de déclencher des pandémies. De plus, chaque pandémie touche des populations différentes selon l'âge ou la comorbidité. Et ces différences dépendent des caractéristiques épidémiologiques de la maladie. Il est donc primordial de pouvoir suivre les virus dans leurs trajectoires au sein des différentes populations et étudier la probabilité que ces virus déclenchent une pandémie. L'assureur est donc tenu de faire l'inventaire de ses expositions vu l'aspect global d'un tel événement et de pouvoir jauger les impacts sur sa solvabilité. Des règles ont été ainsi mises à disposition des assureurs pour faire face à ce type de risque émergent et atténuer les conséquences sur la solvabilité. La survenance récente de la pandémie de la Covid-19 appuie ce choix d'intégrer un tel risque dans le système d'assurance.

## 1.2 Risque pandémie et Solvabilité II

La norme solvabilité II a principalement deux objectifs : l'harmonisation du secteur de l'assurance et de l'ensemble des règles édictées par ce secteur au niveau européen, et permettre à tous les acteurs de se développer dans un environnement compétitif dans la zone européenne. Elle vient accompagner la norme Solvabilité I pour amener une vision plus fidèle des profils de risque des organismes assureurs. Acceptée par la Commission Européenne le 22 avril 2009, la norme Solvabilité II ne se voit appliquée qu'en 2016 et vient mieux évaluer et intégrer le risque dans les contraintes imposées aux assureurs.

Étant toujours en vigueur, les principes généraux de la directive Solvabilité II vont être rappelés et une brève présentation de ses composantes sera proposée. Ensuite, sera abordée l'évaluation du risque pandémie dans le cadre qu'elle instaure.

### 1.2.1 Solvabilité II

Pour qu'un risque soit assuré, il doit vérifier trois caractéristiques : il doit s'inscrire dans le futur et ne peut pas donc être déjà réalisé, il doit faire ressortir de l'incertitude, refléter un événement aléatoire, et que sa survenance ne peut dépendre exclusivement de la volonté de l'assuré. Un réel hasard existe sur le déroulement des contrats d'assurance et ce hasard joue de façon considérable sur la solvabilité des entreprises d'assurance.

La connaissance et la maîtrise de la volatilité des risques de l'assureur sont capitales pour pouvoir être suffisamment solvable, c'est-à-dire être assez solide financièrement pour pouvoir respecter tous les engagements envers les assurés. La construction de provisions est basée sur la mutualisation des risques apportés par les assurés, réduisant la volatilité globale. Certaines méthodes mathématiques permettent de prévoir approximativement le montant des sinistres futurs et donc d'estimer au mieux les prestations potentielles à venir.

Afin de garantir cette solvabilité, la directive impose aux sociétés d'assurances de disposer, en plus des provisions techniques qui couvrent la totalité des engagements, des fonds propres en quantité et en qualité suffisante pour faire face à des événements imprévus pouvant perturber le respect des engagements. Elle propose un mode de calcul qui s'ajoute à celui déjà existant et détermine ainsi le niveau minimal de ces capitaux nécessaires. De plus, elle pousse une meilleure gestion du risque en basant les calculs de solvabilité sur des critères plus vastes que simplement forfaitaires.

Le modèle de Solvabilité II est basé sur trois piliers : les exigences quantitatives, les activités de gouvernance et le devoir d'information aux régulateurs ainsi qu'à destination du public.

Le premier pilier veut fournir des éléments de calculs des composantes du bilan économique pour assurer la solvabilité des acteurs en précisant des montants minimums de fonds propres :

- le MCR ou *Minimum Capital Requirement*. Il correspond au niveau minimum de fonds propres en dessous duquel le régulateur ou l'autorité de contrôle intervient impérativement pour rétablir la santé de l'entité (ou pour sa liquidation);
- le SCR ou *Solvency Capital Requirement*. Il présente le capital nécessaire pour absorber des pertes extrêmes et imprévisibles à horizon de 1 an. Il s'agit de la *VaR* à 99,5% ou la perte bicentenaire.

Les quantités présentées peuvent être calculées grâce à un modèle interne développé au sein de l'entreprise, par une formule standard proposée par la Commission Européenne, convenable sans différentiation entre assureurs, par un mélange des deux ainsi appelé modèle interne partiel dont la

calibration des paramètres revient à être réalisée par l'assureur.

### 1.3 Calcul du SCR par formule standard

Les assureurs sont libres de choisir la méthode de calcul du SCR. La plupart optent pour se baser entièrement sur la formule standard proposée par la directive. Cette dernière s'appuie sur l'agrégation des risques auxquels les entreprises d'assurances sont exposées. Cet agrégat se base sur une matrice de corrélation entre les différentes branches d'activités de l'entité considérée et a pour but de lier les différents risques entre-eux.

Cette approche par modèle standard consiste à cartographier les risques en modules et sous-modules et de remonter les ramifications pour calculer le SCR global en passant par les SCR sous modulaires et les SCR modulaires. L'arborescence essaye de refléter les différents secteurs considérés. Les modules principaux sont: le risque de marché, le risque de souscription Santé, le risque de souscription Vie, le risque de souscription Non-Vie, le risque de contrepartie ou de défaut, le risque Intangible et le risque opérationnel. Pour chacun de ces risques, un SCR est calculé suivant une formule spécifiée puis agrégée.

Deux sous-modules s'intéressent au risque pandémie et vont être explicités dans la partie suivante.

#### Le risque pandémie

Le risque pandémie est bien pris en compte dans la formule standard et se fait au niveau des sous-modules catastrophes et concerne principalement l'assurance vie et l'assurance santé. Il n'est pas à négliger que d'autres modules pourraient être affectés comme le risque opérationnel, le risque de défaut, le risque de marché, étant tous vulnérable à la pandémie.

##### 1.3.1 Le risque vie CAT

Le sous-module risque vie CAT (pour catastrophe) essaye de capter le risque de surmortalité qu'induirait une catastrophe, par exemple une pandémie. Il couvre en premier lieu les diverses garanties décès comme les capitaux décès, les frais d'obsèques, les rentes éducation et rentes conjoint. La survenance d'une pandémie augmente effectivement le nombre de décès et peut générer des prestations supplémentaires non prises en compte.

Le SCR associé à ce sous-module permet à l'assureur de répondre aux prestations supplémentaires imprévisibles dont le niveau de virulence se reproduirait une fois tous les 200 ans. Le règlement délégué [10] compare, dans l'article 96, l'exigence de capital pour risque de catastrophe en vie à une hausse soudaine de 0,15 point de pourcentage des taux de mortalité. Ce coefficient s'applique de façon uniforme sur la table de mortalité choisie et la transformation sera ainsi utilisée pour le calcul des provisions. Le règlement propose également (article 96) une formule simplifiée pour le calcul de l'exigence de capital pour risque de catastrophe en vie:

$$SCR_{vieCAT} = \sum_i 0,15\% \times CAR_i$$

où  $CAR_i$  représente le capital sous risque du contrat i.

Cette hypothèse est discutable du fait qu'il est supposé qu'une pandémie interviendrait de la même façon sur toutes les classes d'âge et que l'impact sur la mortalité serait le même pour toute la population.

### 1.3.2 Le risque santé CAT

L'article 160 du règlement délégué décompose le risque santé CAT en trois risques comme explicité dans la formule proposée :

$$SCR_{healthCAT} = \sqrt{SCR_{ma}^2 + SCR_{ac}^2 + SCR_p^2} \quad (1.1)$$

où

1.  $SCR_{ma}^2$  représente l'exigence de capital du sous-module « risque d'accident de masse »;
2.  $SCR_{ac}^2$  représente l'exigence de capital du sous-module « risque de concentration d'accidents »;
3.  $SCR_p^2$  représente l'exigence de capital du sous-module « risque de pandémie ».

Le risque santé pandémie couvre les garanties incapacité, invalidité et frais médicaux. L'article 163 du règlement [10] définit cette exigence de fonds comme « égale à la perte des fonds propres de base de l'entreprise d'assurance ou de réassurance résultant de la perte soudaine d'un montant qui, sans déduction des montants recouvrables au titre des contrats de réassurance et des véhicules de titrisation, est calculé comme suit: »

$$SCR_p = 0,000\,075 \times \sum_i E_i + 0,4 \times \sum_c N_c \times M_c \quad (1.2)$$

Le premier facteur de la somme concerne le maintien de salaire en cas d'incapacité et le deuxième terme concerne les frais médicaux engendrés par la survenance d'une pandémie. Plus précisément:

- $\sum_i E_i$  représente l'exposition des entreprises d'assurance et de réassurance au risque de protection du revenu en cas de pandémie.  $E_i$  représente le montant des prestations à payer pour l'assuré i en cas d'incapacité de travail permanente provoquée par une maladie infectieuse. La valeur des prestations correspond à la somme assurée ou, si le contrat prévoit le versement de prestations récurrentes, à la meilleure estimation des versements de prestations, en supposant que l'assuré souffre d'un handicap permanent et définitif;
- la somme considère tous les pays  $c$  concernés par l'activité d'assurance ou de réassurance;
- $N_c$  représente le nombre de personnes assurées par les entreprises d'assurance et de réassurance qui sont des habitants du pays  $c$  et sont couvertes par des engagements d'assurance et de réassurance en frais médicaux, autre que des engagements d'assurance ou de réassurance d'indemnisation des travailleurs. Ces frais ne concernent que ceux liés à une maladie infectieuse;
- $M_c$  est le montant moyen de prestations ou de réassurance que paieraient les assureurs en cas de pandémie en fonction d'un pays donné  $c$ . Il se calcule de la façon suivante :

$$M_c = \sum_h H_h \times CH_{(h,c)}$$

où  $CH_{(h,c)}$  est la meilleure estimation des montants à payer par les entreprises d'assurance et de réassurance pour un assuré dans un pays  $c$ . Cela ne concerne que les frais médicaux

lors du recours à des soins de santé  $h$  en cas de pandémie.  $H_h$  représente le ratio d'assurés présentant des symptômes cliniques qui ont recours aux soins de santé  $h$  figurant à l'annexe XVI du règlement délégué [10].

De même, cette formule possède un caractère d'application uniforme, quel que soit l'âge.

## 1.4 Calcul du SCR par modélisation

Le capital de solvabilité requis (SCR) sous Solvabilité II est défini, pour une société d'assurance donnée, comme la *Value-at-Risk* à 99,5% de la variation sur une année de ses fonds propres économiques. Le SCR pourra être déterminé par un mode de calcul par défaut, dit « formule standard » ou un modèle interne développé par l'entreprise.

Dans le cadre de ce mémoire, il ne sera pas question de formule standard. L'étude se rapproche plutôt du modèle interne. Dans ce dernier cas, une distribution des charges de sinistres, généralement obtenue par simulation Monte Carlo, est nécessaire. Différents scénarios peuvent être lancés afin de retrouver une distribution des prestations futures additionnelles. La construction de tels modèles peut se baser sur des méthodes actuarielles mais des domaines différents de l'actuariat se sont déjà penchés sur la problématique de simulation de pandémies comme l'épidémiologie. Elle peut apporter des pistes et outils qui permettraient de produire des projections plus fines.

Donc la simulation d'une pandémie par modélisation devient intéressante lors du calcul des distributions des prestations concernées par les pandémies.

## Chapter 2

# Introduction à l'épidémiologie

“Faut-il que, lorsqu'un malade sera atteint d'une affection microbienne contagieuse, il ne soit plus considéré comme un homme, mais qu'on ne voie en lui que le porteur de microbes ?”

---

R. BIOT, *Politique de la santé publ.*, 1933, p. 31.

L'épidémiologie, par définition, est la discipline qui a pour objet d'étudier les maladies, qu'elles soient épidémiques ou pas, leurs modes de contagion sous l'influence de divers facteurs comme l'environnement, les modes de vie, la santé de la population, etc. Elle essaye d'expliquer le comportement d'une épidémie, d'en extraire une fréquence d'apparition, une distribution de propagation, une sévérité. Ce sont ces caractéristiques qui intéressent fortement les assureurs vue la nature aléatoire des risques qu'ils entreprennent.

Cette science tente de collecter des données concernant la maladie pour permettre une meilleure compréhension de son comportement, de se forger un avis sur la gravité de la situation et des interventions nécessaires pour contenir la propagation et limiter le nombre de victimes. Seront donc intéressants la vitesse à laquelle la maladie se transmet entre les individus, la mesure des interactions sociales entre eux pouvant accélérer ou freiner la diffusion, la létalité associée après avoir contracté la maladie et des complications considérables, les moyens de diffusions, les souches ainsi que de potentielles mutations en des variants plus puissants. Ils formeront alors des paramètres qui serviront à pousser une prédition de l'évolution de l'épidémie, de l'impact d'interventions de nature variée pour dévier la trajectoire originellement entamée.

Une reconstruction du passé de cette science va être rapidement retracée, proposant des éléments majeurs, piliers de l'épidémiologie contemporaine.

### 2.1 Histoire

Malgré des exemples des pestes du Moyen-Âge, il faudra se tourner vers l'Antiquité pour voir le premier répertoriage des maladies. Les premières descriptions apparaissent en Grèce antique dans le traité hippocratique *Épidémies*. Or Jacques Jouanna met en garde les lecteurs : la définition proposée

par Hippocrate diverge de celle utilisée maintenant. Il voit cela comme tout ce qui « atteint plusieurs personnes au même moment » et qu'une maladie sévit de manière épidémique par l'intermédiaire de « ce que nous respirons », et donc liée à une souillure malsaine de l'air, ou un miasme. Il crée l'unicité de la cause, généralement aérienne. Galien étend cette définition à l'eau, la nourriture et développe le premier concept de contagion où des personnes infectées ou des cadavres pouvaient transmettre la maladie aux individus à proximité en salissant l'air commun.

De plus, l'épidémie était avant tout pour eux causée par des causes générales, liée alors à des facteurs exogènes contrairement à la perception actuelle qui recentre le débat sur les implications individuelles sur le plan tant biologique que sociétal.

Cette idée de miasme ou d'air souillé persiste comme l'idée dominante jusqu'à ce qu'un médecin anglais propose une alternative. Après avoir lancé le domaine de l'anesthésie, John Snow † s'intéresse à la santé publique et présente une approche singulière : il propose durant la deuxième épidémie de choléra à Londres (1848-1849) que la maladie est causée par un agent, une bactérie (ou) microbe transmis par contact direct des selles, eaux contaminées et habits souillés. Cependant, sa théorie était en contradiction avec celle qui prédominait à ce moment-là : le choléra se propage dans un air chargé de miasmes provenant de la matière organique en décomposition.

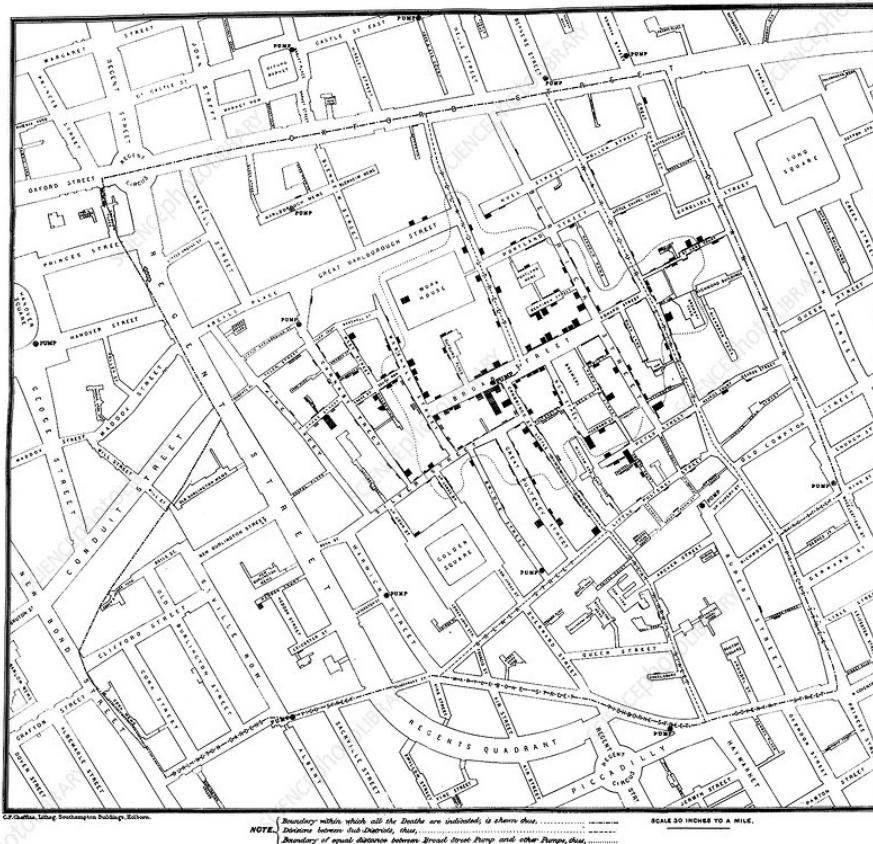


Figure 2.1 – Cartographie de l'épidémie de choléra à Soho, Londres.

Et c'est avec la troisième épidémie de choléra (1853-1855) qu'il montre une vraie approche épidémiologique découlant d'une étude qui va lui assigner une haute distinction : dans le quartier à Soho, Londres, une pompe sur Broad Street utilisée par la plupart dans l'entourage était à l'origine de

†. Non. Pas celui de la série de fantaisie *Game of Thrones*.

nombreuses morts. Il croyait que l'eau pouvait tuer. Il suspectait que le déversement des égouts dans l'eau potable était à l'origine des morts et non pas de miasmes. Habilement, il raisonne avec des cartes (voir 2.1) pour faire ressortir la source de ces décès, démontrant ainsi l'impact de l'eau contaminée. Il essaye de converger vers le foyer de l'épidémie et se sert de la répartition des morts (rectangles noir foncé sur la carte) pour délimiter une zone de recherche assez restreinte (zone marquée par un contour en pointillé). De plus, pour confirmer son hypothèse, le nombre de morts autour de la brasserie était pratiquement nul car la brasserie avait sa propre source d'eau et les ouvriers consommaient des boissons maltées et donc n'était pas contaminés. Finalement, il demande alors de retirer le bras de la pompe mais les craintes subsistent.

En 1927, William Ogilvy Kermack, biochimiste écossais, et Anderson Gray McKendrick, médecin militaire et épidémiologiste écossais, publient un article qui deviendra la base des méthodes mathématiques en épidémiologie. Suivi de deux autres en 1932 et 1933, les articles proposent une hypothèse qui permet de prédire le nombre et la distribution d'individus infectés par une maladie qui se propage à travers une population avec le temps. Malgré un texte alambiqué, la caractéristique intéressante est qu'il s'agissait d'un modèle où l'âge de l'infection modulait les taux de transmissions et de passage en état rétablis. Ce sujet sera détaillé plus tard mais en prenant des paramètres constants pour tous les âges, la dynamique conduit à celle d'un modèle compartimental standard appelé SIR. Plusieurs conclusions deviendront les éléments clés des futures recherches: un palier de densité de population existe et que si la densité est inférieure à ce palier, aucune pandémie ne peut survenir et que des petites augmentations du taux d'infection peut amener à des épidémies plus grandes. Deux aspects sont alors à retenir : la structure de la population est un aspect non négligeable et l'attribut biologique de la maladie est tout aussi important.

Le domaine de l'épidémiologie est donc fortement interdisciplinaire. Outre ses liens étroits avec les statistiques, notamment la biostatistique, elle s'appuie fortement sur d'autres concepts comme la biologie, la pathologie et la physiologie des sciences de la santé et biomédicales ainsi que sur les disciplines de l'anthropologie, de la psychologie, et sociologie dans les sciences comportementales et sociales. Ces dernières ne peuvent pas être exclues car elles jouent un rôle important dans la dynamique des maladies et spécialement des épidémies.

En effet, les interactions sociales forment une architecture de transmission qui permet de fluidifier ou contraindre la propagation de la maladie dans la population. Et donc quantifier le facteur multiplicatif d'agents malades est tout aussi important que de pouvoir estimer à l'aube d'une épidémie (voir une pandémie) ainsi que pendant son déroulement.

## 2.2 $R_0$ , $R_e$ et définitions

Pour introduire la notion du  $R_e$  et du  $R_0$ , des exemples seront donnés pour proposer une intuition et une visualisation.

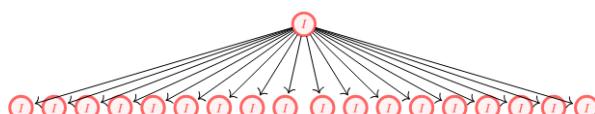


Figure 2.2 – La rougeole,  $R_0$  de 12 à 18

La rougeole est une des maladies les plus contagieuses découvertes en médecine. Une seule personne atteinte de la rougeole peut potentiellement infecter entre douze et dix-huit personnes dans son entourage. Ce nombre est appelé le **nombre de reproduction de base**, noté  $R_0$ . À noter qu'il ne s'applique qu'à une population qui est totalement susceptible d'être infectée et cela avant vaccination et immunité. En d'autres mots, c'est le nombre moyen attendu de cas directement générés par un seul cas (avant sa guérison ou son décès) dans une population où tous les individus sont sensibles à l'infection à  $t = 0$ . Le coefficient  $R_0$  reflète la force du virus même. Cependant il n'est pas unique et peut varier d'un endroit à un autre (généralement plus grand dans les villes densément peuplées que dans les campagnes).

Il est donc utilisé pour décrire la sévérité d'une certaine maladie. Pour illustrer, le nombre de reproduction de base du virus Zika, ayant touché des millions de personnes à travers le monde, vaut entre 3 et 6,6. Très contagieux comparé à celui de la grippe saisonnière qui vaut un peu plus que 1. Et pour le Covid-19, la maladie qui a causé la pandémie en 2020, le  $R_0$  du virus originel vaut à peu près 2.

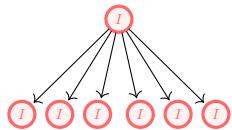


Figure 2.3 – Zika:  $R_0$  de 3 à 6,6



Figure 2.4 – Grippe,  $R_0 = 1,3$

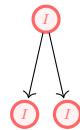


Figure 2.5 – Covid-19:  $R_0$  de 2 à 2,5

À première vue, la différence entre la grippe et la Covid-19 ne semble pas si flagrante, les rapprochements entre ces maladies au début de la pandémie étant, à priori, compréhensibles vu le niveau d'information détenu sur le nouveau virus. De plus, leurs symptômes sont similaires : de la fièvre, des frissons, de la toux, des douleurs articulaires et musculaires et de la fatigue [6]. Les deux peuvent même mener jusqu'à la pneumonie, complication qui peut être fatale. Cependant ces comparaisons sont trompeuses et il est capital de revenir sur ce  $R_0$ .

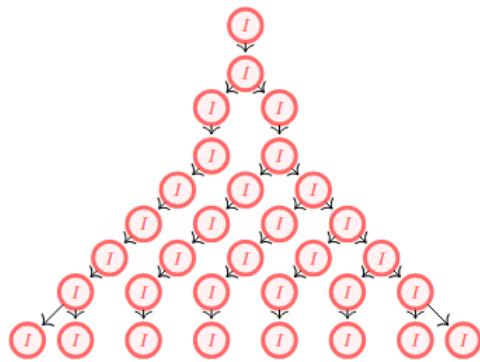
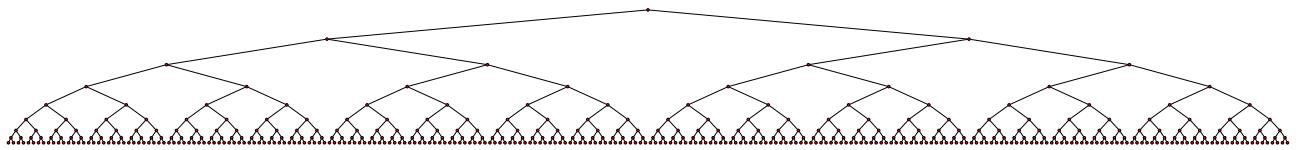
Supposons l'introduction du « patient zéro », la première personne d'une épidémie à avoir été contaminée par l'agent pathogène, dans une population entièrement contaminable. Les durées d'incubation (durée séparant la contamination et l'arrivée des symptômes) sont présumées constantes. En moyenne est supposée l'hypothèse suivante: la première vague de transmission génère  $R_0$  individus tirés aléatoirement dans la population. Ceux-ci pouvant chacun, à leur tour, contaminer  $R_0$  personnes (en supposant que tous leurs contacts sont susceptibles et qu'aucun d'eux ne soit dans l'état infecté). La deuxième vague de transmission en générera  $R_0^2$  de plus. Généralisée, pour  $n$  vagues, le nombre total de personnes infectées (toute vague confondue) se calcule de la façon suivante:

$$N_{inf} = \sum_{k=1}^n R_0^k$$

Par exemple,  $R_{0,grippe} = 1,3$  donc chaque personne, en moyenne, passe la maladie à une ou deux personnes. Après 8 vagues d'infections, le nombre d'individus infecté arriverait en moyenne à 31.

Et en faisant la même chose pour le Covid-19, après 8 générations, plus de 510 personnes seront infectées. C'est une différence non négligeable et il est important de bien pouvoir déterminer ce coefficient pour essayer de prédire la magnitude de la portée qu'un nouveau virus pourrait engendrer. En effet, une petite différence de ce chiffre amène à une dérive significative dans un sens ou un autre.

De plus, il existe un autre paramètre qui, lui, permet le suivi de l'épidémie en fonction de l'évolution

Figure 2.6 – Grippe,  $R_0 = 1,3$ Figure 2.7 – Covid-19:  $R_0 = 2$ 

de l'immunité et du nombre de personnes vaccinées: c'est le **nombre de reproduction**, noté  $R_e(t)$ . Contrairement au  $R_0$  qui s'applique à une population entièrement susceptible d'être infectée où personne n'a encore eu d'immunité que ce soit par rencontre d'un agent infectieux ou par vaccination, le  $R_e(t)$  vient suivre l'évolution en prenant en compte le nombre de personnes ne pouvant plus devenir infectés.

Soit  $1 - f(t)$  la proportion de la population considérée qui est déjà immunisée à l'instant  $t$ . Un contact avec une personne de cette proportion n'induira aucune contamination et seule la partie restante peut être contaminée. Donc, nous avons plus exactement:

$$R_e(t) = f(t) \times R_0$$

Ainsi, ont été définis trois types de situations liées à la propagation de maladie dans la société et elles se distinguent par leurs valeurs de  $R_e(t)$  :

1. **L'épidémie.** Elle est définie avec un  $R_e > 1$  c'est-à-dire qu'une personne infectée passe le virus à au moins une autre personne. Cela permet au virus de se répandre dans la population. Plus il est grand, plus la vitesse de propagation est rapide. C'est dans ce cas-là qu'il faut essayer de limiter voire d'arrêter les diverses transmissions.
2. **L'endémie.** Le virus se trouve de façon permanente dans la population sans jamais s'estomper. La maladie vit durablement dans une société et peut ne pas être un problème de santé publique (comme la lèpre ou le tétanos en Europe par exemple). Cela se passe quand  $R_e(t)$  vaut sensiblement 1.
3. Et finalement, si le  $R_e(t)$  passe sous 1 (non pas à un instant donné mais sur une plage assez longue pour y voir un effet), l'épidémie est dite **contenue** c'est-à-dire qu'il y a de moins en moins de gens qui se font infecter et donc moins de personnes qui peuvent transmettre la maladie. Lorsque suffisamment de gens sont immunisés, cela est appelé l'immunité collective. Pour la grippe, l'immunité collective est obtenue par la vaccination. Toutes les mesures contre

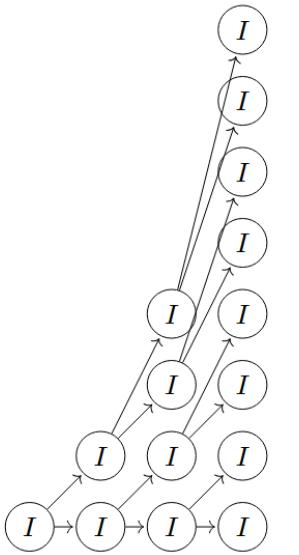


Figure 2.8 –  $R_e > 1$  :  
Epidémie

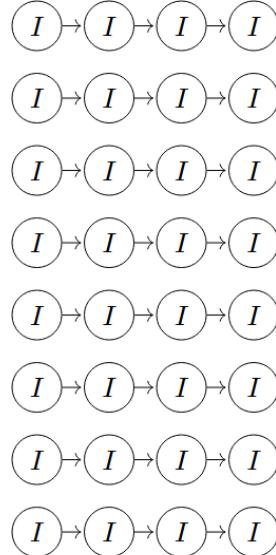


Figure 2.9 –  $R_e = 1$ :  
Endémie

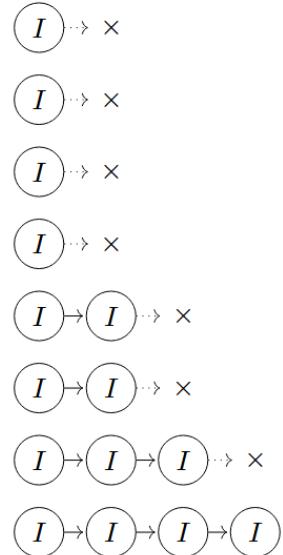


Figure 2.10 –  $R_e < 1$ :  
Contenué!

le COVID-19 communiquées au public - lavage des mains, distanciation sociale, confinements, auto-isolation, traçage des contacts, mise en quarantaine, port du masque et même « immunité collective » - ont toutes le même but : obtenir un  $R_e < 1$ .

Donc ce début de caractérisation amène à un jugement et une catégorisation des maladies à une échelle épidémiologique : certaines maladies n'auront pas le pouvoir de progresser et de passer au stade d'une épidémie alors que d'autres auront les attributs nécessaires ainsi qu'une virulence assez forte pour pouvoir être qualifiée comme tel. Donc ces coefficients sont déterminants, donnent la gravité potentiellement attendue lors de la venue d'une épidémie et permettent un début de modélisation.

## 2.3 SARS-CoV-2 - Coronavirus 2 du syndrome respiratoire aigu sévère

Initialement, l'OMS avait estimé que le  $R_0$  du SARS-CoV-2, le virus de la maladie de la Covid-19, se situait entre 1,4 et 2,2 [16] alors qu'une étude sortie en avril 2020 [4] avait estimé une contagion plus importante, une cinétique plus considérable de diffusion : compte tenu de l'état actuel des connaissances à la date de publication, les chercheurs proposent un  $R_0$  valant 5,7 avec une intervalle de confiance [3,8 – 8,9] dans la ville de Wuhan. Le nouvel an chinois aurait potentiellement eu un rôle à jouer dans l'inflation de cette valeur, la population n'ayant pas été sensibilisée à ce nouvel agent pathogène. Malgré cette quantité qu'on qualifierait maintenant d'erronée, le SARS-CoV-2 est un candidat parfait pour devenir l'aurore d'une pandémie.

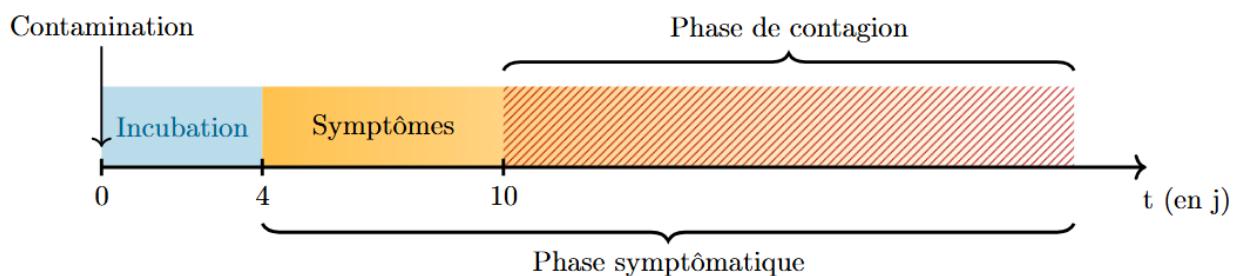
Le comportement de la maladie étant inconnu au stade de départ de la pandémie, de la recherche extensive s'est réalisée, que ça soit sur le plan biologique (et donc concernant les mécanismes de l'organisme vivant) ou le plan sociétal (et donc concernant sa propagation).

### 2.3.1 Échelle individuelle et collective

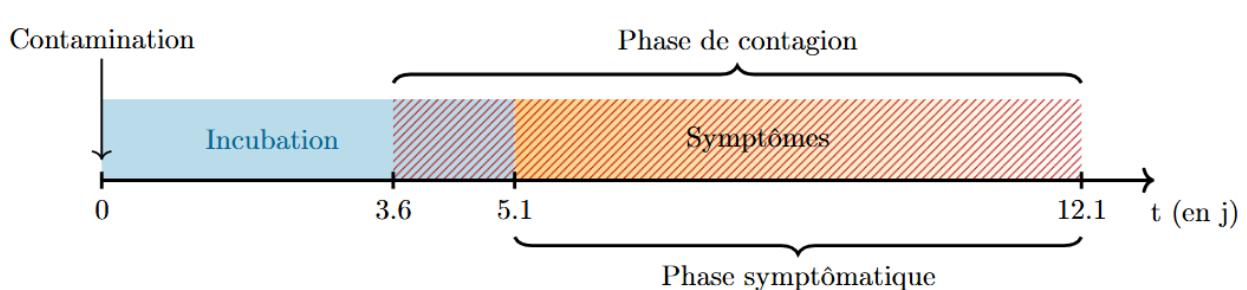
Le fonctionnement de la maladie au niveau biologique du corps ainsi que sa méthode de transmission peuvent plus ou moins propager le virus dans une population.

Sur le plan biologique, chaque virus a une période d'incubation qui sépare le moment de contamination et l'apparition de symptômes ainsi qu'une période de contagion où la maladie peut se transmettre d'une personne à une autre. Il est très important de distinguer ces deux périodes et de pouvoir les placer sur une chronologie. Cette nécessité sera exhibée par la comparaison qui suit en commençant par rappeler ce qu'est le SRAS.

Le SRAS, ou syndrome respiratoire aigu sévère, est une maladie infectieuse causée par un virus appartenant à la famille des coronaviruses, le SARS-CoV-1. Extrêmement virulent, il peut se transmettre chez les humains. Cependant, ce virus a la particularité que sa phase contagieuse débute quelques jours après l'apparition des symptômes ce qui signifie l'absence de contagion avant la manifestation des symptômes. Cela amène à une détection de la maladie et une possibilité d'intervention pour freiner sa propagation en isolant et traitant les patients atteints du virus du SARS.



En revanche, le virus SARS-CoV-2, souche virale génétiquement distincte des autres coronavirus mais appartenant à la même espèce biologique que le SARS-CoV-1, présente une superposition des états comme l'expose le schéma suivant :



Une fois la personne infectée, la durée d'incubation totale en moyenne est 5,1 jours. Durant celle-ci, au bout de 3,6 jours en moyenne, une phase prodromique apparaît. Le malade devient ainsi contagieux sans pour autant présenter des symptômes qui signaleraient l'existence du virus chez le patient. Les symptômes apparaissent vers la fin de la période d'incubation et persistent jusqu'au douzième jours après la contamination (donc durent en moyenne 7 jours). Donc une infectiosité présymptomatique existe alors et rajoute de la dangerosité au virus, cette fenêtre de contagion n'étant pas annoncée par

des symptômes.

Sur un plan collectif, la transmission se réalise essentiellement par l'intermédiaire de gouttelettes émises lors de la prise de parole, des efforts de toux, d'éternuements et cette transmission se complète après un contact rapproché et prolongé avec un agent contagieux. Une charge virale devrait se construire et quand celle-ci atteint un certain niveau, la maladie commence à se manifester. Ces gouttelettes étant de petite taille et pouvant conserver son infectiosité, le virus peut respectivement rester en suspension dans l'air, justifiant ainsi les ventilations nécessaires des pièces fermées et le port du masque, et peut créer des zones d'infection sur des surfaces inertes, poussant une bonne hygiène des mains chez les individus de la population concernée.

Le virus SARS-CoV-2 n'a pas provoqué à lui seul la pandémie de la Covid-19 mais a fait appel à des versions mutées, s'adaptant pour persévérer et assurer sa présence en devenant plus habile à éviter le système immunitaire.

### 2.3.2 Variants

Comme chaque virus, le SARS-CoV-2 n'est pas épargné par les évolutions de son patrimoine génétique et s'est finalement retrouvé avec une multitude de variants.

Un virus est un corps contenant du matériel génétique: soit de l'ADN soit de l'un de ses deux brins complémentaires: l'ARN. Ces derniers sont formés de molécules pouvant être représentées par des lettres. Chaque partie de cette séquence porte les instructions nécessaires pour la production de protéines spécifiques qui permettent au virus de fonctionner et de former les protéines sur sa surface, appelée protéines *Spike*, pour pouvoir se lier aux protéines récéptrices sur la surface de la cellule hôte.

AGCTGATGCGTATGGATCGCAATCGCATGATGTGAAATGATGTAGA

Figure 2.13 – Séquence de nucléotides

En effet, pour se multiplier, le virus doit utiliser les ressources et organites des cellules du corps et en réalisant ces reproductions, le séquençage peut devenir erroné : ces sont les mutations. Des lettres peuvent être retirées (mutation par délétion), ajoutées (mutation par insertion) ou modifiées (mutation par substitution). Ces modifications peuvent être sans conséquences, voir même affaiblir le virus et d'autres peuvent le renforcer. La protéine devient plus adaptée à se lier aux cellules du corps et donc ouvre les portes à une plus grande démultiplication.

Les variants Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1) et Delta (B.1.617.2) présentent eux quatre des mutations concernant la protéine *Spike* et sont considérés par l'OMS comme des variants dont les autorités devraient se soucier.

### 2.3.3 Vaccins et traitement

Jusqu'alors, plusieurs méthodes étaient utilisées pour le développement de vaccins: les vaccins à vecteurs viraux employaient des formes affaiblies (pour la rougeole) ou inactives (pour la grippe saisonnière) de pathogènes pour susciter une réponse immunitaire du corps. Certains usent de toxines

inertes comme les vaccins contre le tétanos alors que d'autres contiennent directement des morceaux de l'agent pathogène (par exemple la protéine *Spike*) qui déclenchera une réaction sans le risque d'introduire des unités de pathogènes. Les vaccinations relatives à l'hépatite B ont recours à cette dernière méthode. Cependant, le temps de développer des tels vaccins, réaliser des tests en laboratoires, des tests cliniques, créer une licence, les produire et les distribuer prend du temps, ce qui manquait lors de la survenance du virus de la Covid-19.

Le vaccin contre le SARS-CoV-2 utilise une alternative à ces solutions et fait appel à l'émissaire biologique de l'être vivant : l'ARN messager, version maturée de l'ARN qui est capable de quitter les cellules. En effet, toutes les fonctions du corps sont amenées par les protéines et donc les cellules sont en constante production de ces chaînes d'acides nucléiques. Sont donc injectées les instructions pour construire les protéines *Spike* du virus dans le corps et provoquer la réaction du système immunitaire. Cette méthode est utilisée par Pfizer-BioNTech et Moderna. Cependant, l'ARN messager se dégrade rapidement ce qui peut compliquer son acheminement aux centres de vaccination. Viennent alors les vaccins à adénovirus qui eux portent l'ADN, plus stable, utilisés par Oxford-AstraZeneca et Janssen.

Le temps de développement du vaccin de la Covid-19 a été relativement court mais malheureusement, cela n'a pas toujours été le cas. Le volume de ressources humaines et matérielles mobilisé pour contrer la pandémie de la Covid-19 dépasse celles appelées lors de pandémies antérieures. Les capacités de production de masques et de vaccins ont été poussées à leurs limites, les modélisations sont devenues plus précises grâce aux enseignements collectés. Or, cette avancée n'empêche pas l'apparition de virus ayant la faculté de tisser une pandémie. Il faut donc être prêt.

## 2.4 Règle des trois unités: de lieu, de temps et d'action

Les épidémies et pandémies présentées ultérieurement donnent donc de la gravité à l'estimation du nombre de reproduction qui permettrait de pouvoir se situer dans la temporalité, déterminer l'étape actuelle de l'évolution, l'évolution future ainsi que les mesures à prendre pour contenir la maladie, voire la rendre endémique.

Au milieu des années 1990, l'anthropologue Mervill Singer introduit le concept de **syndémie**, combinaisons de maladies interagissant en synergie avec des facteurs biologiques et environnementaux. En effet, l'apparition d'une pandémie n'empêche pas la circulation d'autres maladies déjà préexistantes dans la population. Et ce brassage aggrave la mortalité de la pandémie. Les approches pour contrer cet effet ont toujours été réductionnistes: identification du virus, suivi du développement de traitement curatif (anti-infectieux) et aller vers le préventif (vaccins). L'approche syndémique conduit à des stratégies de prévention prenant en compte les mentalités, les facteurs socio-économiques, comportementaux, environnementaux, des maladies existantes, du changement climatique, etc.

Il serait donc intéressant d'avoir une certaine méthodologie, un sens de travail et de projection, d'idées pour pouvoir estimer le nombre de reproduction, que ce soit à son stade initial avec le  $R_0$  ou le  $R$ , ainsi que d'autres paramètres concernant la survenance, la sévérité ou tout ce qui pourrait être d'une utilité pour cerner le comportement général de la propagation de la maladie. Trois concepts viennent alors en tête, trois axes de recherches qui orientent l'étude : le lieu, le temps et l'action.

### 2.4.1 De lieu

Tout d'abord, le **lieu** où naît, se propage la maladie et où se réalise l'étude est une notion capitale à garder en tête. En effet, la maladie se propagera de façon différente suivant le lieu, la densité de population et la topologie des interactions sociales qui peuvent freiner ou précipiter la propagation. De plus, certaines régions seront amenées à développer des nouvelles maladies plus fréquemment que d'autres suivant leurs géographie, biodiversité, économie et style de vie.

La géolocalisation affecte de manière conséquente la fréquence d'apparition de maladies. L'Asie du Sud-Est a été identifié comme un foyer de maladies émergentes, majoritairement des zoonoses. Ces dernières se transmettent par des bactéries, virus ou parasites entre l'animal et l'humain. Cette incidence accrue découle de facteurs climatiques et de l'évolution d'usage des sols. En effet, la croissance locale d'élevage de bétail, la déforestation pour faire place à des plantations commerciales, les pertes de biodiversité, les fractionnements des paysages et l'instabilité climatique dans la zone exhibent des corrélations avec l'augmentation des épidémies de maladies infectieuses comme la dengue, le typhus des broussailles et de la leptospirose. De plus, la promiscuité est étroite entre l'humain et les animaux. Dans les campagnes par exemple, les animaux vivent à proximité des maisons et ont donc un contact permanent avec les habitants pouvant mener à des transmissions plus fréquentes. Une approche intégrant donc les valeurs culturelles associées aux animaux sauvages et domestiques, à la culture économique du pays et à son exposition aux changements climatiques et environnementaux est donc nécessaire pour donner un avis sur le risque de survenance d'une pandémie.

Cela mène à considérer aussi le mode de vie quotidien: en effet, la maladie ne se propagera pas de la même manière dans les villes que dans les campagnes. La structure d'interactions en ville est plus dense que celle retrouvée en campagne, de plus que la multiplicité des rencontres est d'autant plus élevée en zone urbaine que dans les zones rurales. Suite au choix d'une circonscription géographique pour effectuer l'étude, des hypothèses démographiques viennent se confronter à la délimitation : la commodité d'avoir une population stationnaire en nombre amène une simplification mais peut rapidement devenir irréaliste lors de fluctuations non négligeables de la population concernée. La question qui se pose serait de déterminer ce qui serait considéré comme entrées et comme sorties par rapport à cette population : envisager les naissances, les décès, les déplacements vers l'intérieur ou l'extérieur du périmètre étudié, les individus en déplacements temporaires comme les touristes, etc.

Cette délimitation peut ne pas s'arrêter au sein d'une même région ou d'une même ville mais peut s'étendre au-delà des frontières et passer au rang d'une pandémie en touchant le globe en entier. La complexité redouble avec les interactions qui deviennent embrasées. Le périmètre géographique d'un portefeuille est donc indicateur de l'exposition de l'assureur ainsi que le type d'interaction de ses assurés.

Cette unité de lieu est fondamentale dans la compréhension des survenances et des modes de diffusion des maladies. Cependant, seule, l'angle d'approche serait figé à un instant donné et produira des résultats à des fins documentaires et ne pourront servir comme prédicteurs.

### 2.4.2 De temps

Couplée avec l'unité de lieu, la temporalité porte ses propres hypothèses: la maladie n'a pas vocation à être statique et peut évoluer en fonction de l'environnement dans lequel elle existe et se développe. Le **temps** procure un espace d'existence indépendamment de l'espace matériel. L'intensité peut varier et la valeur des prestations versées par les assureurs peut augmenter avec la nature des soins, des

événements en même temps que leurs nombres.

La notation du nombre de reproduction,  $R_e(t)$ , suggère et même impose que la valeur du coefficient est affectée par le choix de l'instant d'observation. En effet, les interactions permettent de répandre la maladie mais cette dernière peut avoir sa propre façon de manifester. Le temps de la phase de contagion est important: plus elle est longue, plus l'agent infecté a le temps de passer la maladie à autrui et ces derniers de faire de même en suivant le chemin d'arbre de transmission. Cette période de pouvoir de transfert vient effectivement faire évoluer ce  $R_e(t)$  car si aucune intervention n'est mise en place pour limiter les infections, il y aura plus de chance de devenir infectieux et donc augmenter la valeur du  $R_e(t)$ .

Pourrait être aussi mentionné le temps de séjour moyen en hôpital ainsi que celui de l'utilisation/occupation d'un lit en soins intensifs ou en réanimation à la suite d'une hospitalisation. La saturation des hôpitaux est un sujet qui est revenu souvent durant la pandémie de la Covid-19, la tension se posant sur le nombre limité de lits disponibles. Si le flux de personnes étant reçues à l'hôpital est plus grand que le flux de personnes en sortant, il n'y aura plus de lits à disposition pour accueillir les futurs patients. Le système hospitalier ne peut négliger le fait que des personnes atteintes d'autres pathologies doivent et ont le droit aussi d'être admises en soins intensifs ou réanimation vu leurs complications et la nécessité d'interventions et de surveillance.

Cependant, certaines périodes peuvent alléger, quoique de très peu, la valeur du nombre de reproduction: il prend la forme d'une intervention exogène dont l'intensité ne peut être contrôlée. Pour le Covid-19, une augmentation de 1° Celsius fait baisser le  $R_e$  de 1,2% [3]. La raison n'a pas été creusée et n'était pas un des objectifs de l'étude mentionnée mais certains spéculent que les gouttelettes portant le virus s'évaporeraient plus rapidement dans un environnement plus chaud et donc viendrait diminuer la charge virale transmise. Effectivement, l'été ne va pas diminuer le  $R_e$  afin de devenir inférieur à 1 mais il assiste à sa réduction et pourrait permettre une réallocation des ressources concernant les interventions qui pourraient être légèrement réduites durant l'été pour les intensifier en période hivernale. De même, la grippe « saisonnière » voit son efficacité s'effondrer en été, prospérant en revanche dans des conditions humides et froides.

Finalement, la durée de la pandémie peut rapidement avoir une conséquence sur la maladie même (qui peut évoluer) et les santés physiques et mentales de la population. Cela se matérialise par les confinements qui séparent les familles et la maladie qui creuse les tombes de milliers, séparation finale. Les intervenants de santé et hospitalier sont surchargés. Et à ne pas oublier ceux qui perdent leurs revenus et la capacité de travail, que ce soit causé de façon directe ou indirecte par la pandémie.

La dimensionnalité temporelle est d'autant plus cruciale voyant les volets sur laquelle elle agit. Donc agir rapidement est tout aussi important que de déployer ces actions à grande portée.

#### 2.4.3 D'action

L'appréhension du comportement d'une épidémie est capitale pour commencer à prendre des décisions. L'épidémirosurveillance de la maladie, la prise en main de la situation et les **actions** mises en place sont là pour tenir et contenir la propagation.

Avec les prérequis de lieu et de temps, des mesures sanitaires de deux types seront éventuellement mis en place: en premier lieu les mesures sanitaires non pharmaceutiques et ensuite les mesures

sanitaires pharmaceutiques i.e. les gestes barrières et confinements puis les vaccins. Ces premières mesures sont relativement rapides à mettre en place quand les ressources sont prêtes et mobilisées assez rapidement pour assurer le bon respect des mesures comme les confinements en vigueur, les horaires des couvre-feux. Comparées au développement d'un vaccin, elles ne requièrent pas de tests préalables car il va de soi qu'elles seront efficaces avec la coopération de la population et de gouvernement. Les mesures sanitaires pharmaceutiques, elles, sollicitent davantage de moyens, nécessitant des épreuves élaborées pour évaluer les effets secondaires plausibles d'apparaître à l'issue du développement de vaccins ainsi que l'efficacité apportée par la ou les doses obtenues. Malgré les avancées dans le domaine, cette phase peut s'étendre par précaution de qualité et de quantité même si cette dernière peut être remédiée par des fins industrielles de production.

Or simplement cette organisation n'est pas suffisante. Que ce soient les coûts générés et les masses concernées, il faudrait pouvoir apprécier l'efficience des actions prises et approuver des choix adoptés. L'épidémiologie n'est qu'un bourgeonnement alors que les modèles épidémiologiques en sont les fleurs, le travail de communautés scientifiques interdisciplinaires, et leurs prédictions en sont les fruits. Ces modèles projettent avec toutes les hypothèses mises en place des scénarios contraignants ou non et permettront de déterminer la meilleure approche suivant la situation qui se présente, les différentes simulations générées et les interprétations. Par exemples, trouver le meilleur schéma de confinement en choisissant entre un long confinement ou des circuits avec des confinements intermittents; le programme de vaccination le plus propice en choisissant quels ensembles de la population devraient l'avoir en priorité.

Cet aléa politique concerne également l'exposition de l'assureur, étant impacté par des décisions prises (variant en intensité d'un pays à l'autre), des dispositions mises en place, la vitesse de développement des vaccins, de la réaction de la population et de leur respect des confinements, etc.

Grâce à ces trois unités, un plan d'attaque tient à diriger l'approche en tenant compte des particularités de la localisation, des évolutions à intégrer afin de transformer l'épidémiologie et construire des outils de projection et d'évaluation interne des risques de solvabilité, spécifiquement du risque catastrophe.

# Les modèles épidémiologiques

“[...] all models are approximations. Essentially, all models are wrong, but some are useful. However, the approximate nature of the model must always be borne in mind [...]”

---

George Box

L'épidémiologie propose un cadre de raisonnement auquel les approches empiriques sont rapidement limitées ou incapables de donner des éléments de réponse vu les contraintes d'espace et de temps.

L'approche par modélisation, quoique jamais parfaite, permet de produire des résultats interprétables. Elle n'est qu'une simplification par choix afin d'étudier des phénomènes liés à la propagation de maladies. Le but est de pouvoir décrire l'épidémie, de donner ses caractéristiques comme la sévérité (l'évolution du nombre de décès, de personnes infectées), la temporalité (la durée sur laquelle elle pourrait s'étaler), sa localisation (si elle peut devenir une pandémie par exemple), etc. Les modèles sont là pour aider à comprendre les systèmes sur lesquels se basent les épidémies, expliquer les phénomènes observables. Ils seront alors utilisés pour de la prospection afin de prendre des décisions. Ils peuvent constituer une base de scénario de chocs applicable à la projection d'un bilan d'un assureur.

L'abstraction oblige cette réduction et mettre en œuvre ces modèles fait appel à différentes méthodes suivant le contexte et la précision. Va être ainsi appelé **simulateur** tout modèle qui permet de recréer l'environnement et le comportement d'une épidémie. En entrée, se trouveront les différents paramètres gouvernant les dynamiques d'infection, de rétablissement.

Deux types de modèles ont principalement été utilisés en France durant la crise de la Covid-19: les **modèles compartimentaux** et les **modèles agents**.

## 2.5 Les modèles compartimentaux

Les modèles compartimentaux en épidémiologie sont des outils purement théoriques qui ont connu un essor d'utilisation et de mise en pratique à partir des années 1980 avec l'épidémie du SIDA.

Ces modèles se basent sur deux concepts : les *compartiments* dans lesquels la population simulée va évoluer et les *lois* qui gouvernent ces passages d'un compartiment à un autre. Les compartiments subdivisent la population en divers états par rapport à la maladie : susceptible, infecté, asymptomatique, hospitalisé, rétablis, décédé; tout ce qui peut être modélisé dans l'architecture théorique. Le choix revient au degré de précision voulu et de la traçabilité dans les différents états. Les lois sont là

pour spécifier la proportion des individus passant d'un compartiment à un autre.

Un des modèles les plus important et relativement riche en information est le modèle *SIR*.

## Le modèle *SIR*

Une des quantités à chiffrer et considérer serait le nombre d'infectés dans la population. Suivre l'évolution de ce nombre ainsi que celui des susceptibles à avoir la maladie devrait nous donner une idée de départ. Commençons avec ce modèle de base. Les individus **Infectieux** peuvent transformer les individus **Susceptibles** de contracter la maladie en des individus **Infectieux**. Les **Infectieux** sont ici supposés ne pas guérir. Une fois infectés, on le reste (pour l'instant).

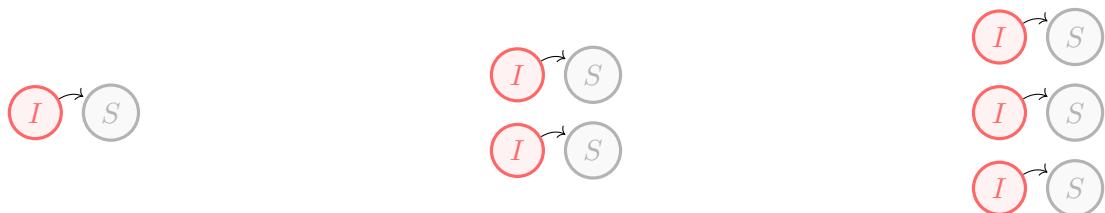


Figure 2.14 – Augmentation du nombre d'agents infectés : de 1 à 2, de 2 à 4 et de 3 à 6.

Par exemple, en choisissant  $R_0 = 2$ , on doublera le nombre d'infectés tous les quatre jours, en moyenne, estimation au début de l'épidémie de la Covid-19. En commençant qu'avec un seul infecté, au bout de 10 doublements, il y aurait déjà 1024 infecté ! Ceci est la courbe exponentielle, commençant très petit et explosant rapidement. Pour autant, cette simulation est fausse. Cette croissance ne peut pas durer infiniment car une partie de la population a déjà le virus. Donc les individus infectés créent une sorte de barrage et empêche la diffusion du virus.

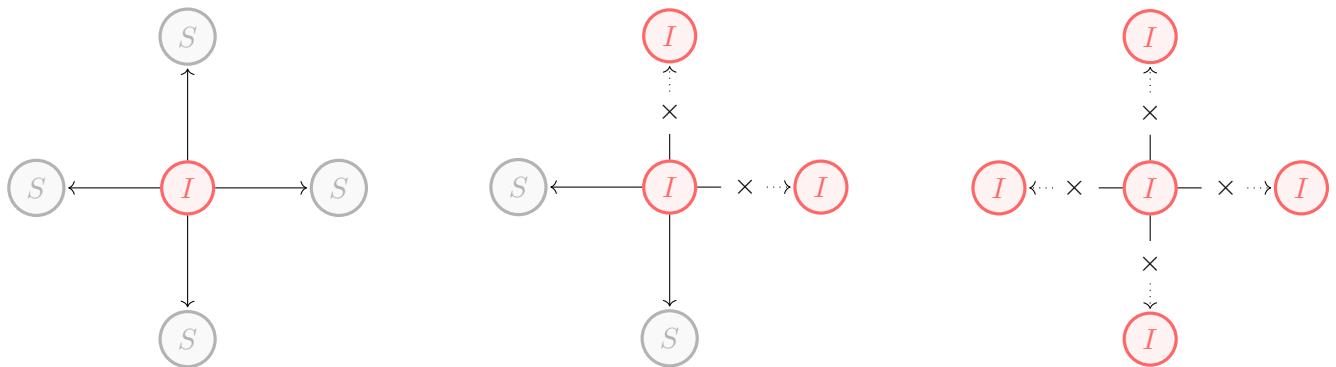


Figure 2.15 – La vitesse de transmissions devient plus lente avec l'augmentation du nombre d'infectés.

Quand les contacts de l'individu infecté sont tous susceptibles, il pourra alors infecter 100% des contacts alors que quand 50% d'entre eux sont déjà infectés, il ne restera que les autres 50%. Si tous ses contacts sont déjà infectés, il ne pourra plus infecter personne (donc 0% de ses contacts). Vient alors la naissance de l'état *Recovered* ou Rétablis. On supposera que tous les individus Infectieux deviennent guéris (sain-et-sauf, avec des lésions ou morts) et seront considérés comme immunisés.

Cette dynamique peut être résumée avec le système d'équations différentielles ordinaires suivant dont les paramètres seront explicités :

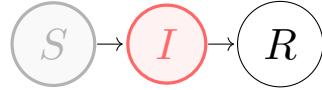


Figure 2.16 – Graphe de transition du modèle SIR : Susceptible puis Infecté puis Rétablis

$$\begin{cases} \frac{dS(t)}{dt} = -\beta \frac{I(t)}{N} \times S(t) \\ \frac{dI(t)}{dt} = \beta \frac{I(t)}{N} \times S(t) - \gamma \times I(t) \\ \frac{dR(t)}{dt} = \gamma \times I(t) \end{cases} \quad (2.1)$$

Ce système d'équation essaye de décrire la vitesse à laquelle les effectifs de ces trois compartiments pour une certaine population vont évoluer:

1.  $S(t)$ , le nombre d'individu susceptibles d'être infectés,
2.  $I(t)$ , le nombre d'individu infectés et pouvant transmettre la maladie,
3. et  $R(t)$ , le nombre de personnes rétablies qui ne peuvent redevenir infectés ni retransmettre la maladie.

Les équations sont là pour décrire le mouvement d'individus entre chaque compartiment à un instant  $t$ .  $\beta$  et  $\gamma$  sont respectivement le taux d'infection et le taux de rétablissement. Et comme vu précédemment, plus le nombre de personne infectées est élevé, plus la probabilité qu'une personne susceptible rencontre une personne infectée sera élevée. Donc, la probabilité de transition devrait être proportionnelle au nombre  $I(t)$ . De plus, le nombre total de ce genre de rencontre intervenant sur une durée  $dt$  est proportionnel au nombre de susceptibles encore présents dans la population. Donc le nombre de contaminations, sur un intervalle  $dt$ , s'élève à  $\beta \times \frac{I(t)}{N} \times S(t)$  que l'on retranche à  $S(t)$  et que l'on ajoute à  $I(t)$ .

$\frac{1}{\gamma}$  représente la durée moyenne durant laquelle une personne est contagieuse et peut transmettre la maladie et  $\frac{1}{N\beta}$ , où  $N$  est l'effectif total de la population considérée, est la durée moyenne entre deux contaminations générées par un agent du compartiment infectieux à  $t = 0$ . Donc, le rapport  $\frac{N\beta}{\gamma}$  est le nombre moyen de contamination pendant une période de contagiosité. Cela rappelle le nombre de reproduction qui peut donc se calculer de la façon suivante:

$$R_0 = \frac{N\beta}{\gamma} \approx \frac{S(0)\beta}{\gamma}$$

où  $S(0) \gg I(0)$  donc  $\frac{S(0)}{N} \approx 1$ .

Plusieurs hypothèses sont mises en place :

- la contrainte d'une population fermée, i.e. d'une population dont l'effectif total est constant. Mathématiquement, cela s'écrit:  $N = S(t) + I(t) + R(t)$ ; ne sont donc pas considérées les naissances. Les décès dus à la maladie sont considérés alors que les décès pour autre cause ne le sont pas;
- Il n'y a pas de « marche arrière » ; devenir rétablis signifie immunité permanente,
- les paramètres  $\beta$  et  $\gamma$  sont constants et ne sont pas affectés par l'âge des individus, l'avancement de l'épidémie, des mutations du virus, etc.

De plus, le modèle n'envisage pas une période de latence pour simuler l'incubation de la maladie : un individu devient infectieux dès son infection. Des modèles l'intégreront éventuellement, notamment

le modèle *SEIR* qui rajoute le compartiment *E*, *Exposed* ou Exposé, avec une phase d'incubation pendant laquelle l'individu n'a ni symptômes et n'est pas contagieux.



Figure 2.17 – Modèle SEIR: ajout du compartiment Exposé, phase latente sans contagion ni symptômes.

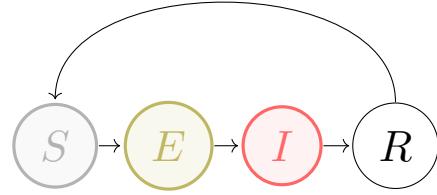


Figure 2.18 – Modèle SEIRS : retour dans le compartiment Susceptible avec une perte d'immunité.

Les modèles compartimentaux sont donc des modèles à équations, relativement simples à implémenter, et ont montré une vraie aptitude à produire des résultats qui peuvent ressembler à ceux que provoquerait une vraie épidémie.

Cependant, ce sont des modèles déterministes. D'une part, une toute petite modification des conditions initiales pourrait engendrer des résultats complètement différents. Cette sensibilité pourrait, après un temps court, faire diverger deux copies du même modèle à un point où ils ne se ressembleraient plus. D'une autre part, avec une même initialisation, le modèle reproduira toujours le même résultat. Ce résultat est vrai car il présente la trajectoire moyenne de la pandémie mais ne possède donc pas de volatilité naturelle. De plus, le mouvement de population en proportion ne permet pas de modéliser certains phénomènes comme les *superspreaders*, personnes particulièrement contagieuses qui contaminent davantage de gens que le prévoit le  $R_0$ .

Si les modèles compartimentaux sont les outils les plus communément utilisés pour anticiper les effets des interventions, des approches alternatives existent: les modèles agents.

## 2.6 Modèles agents

Une étude du psychologue social Stanley Milgram [13] montre que la population américaine semble être plus interconnectée que ce qui était imaginé. En effet, à l'issue d'une expérience qu'il réalisa à l'échelle de l'État fédéral, en choisissant deux personnes au hasard, il démontre qu'il existe en moyenne six intermédiaires entre ces deux personnes inconnues. Cela déclenche une nouvelle perspective sur les structures sociales : il existerait des clusters locaux qui possèdent des liens entre eux ce qui permet cette « séparation à six degrés ». Cette motivation lance la construction de modèles basés sur les interactions sociales et qui gagneront de l'importance dans différents secteurs comme la sociologie, la psychologie sociale, l'immunologie et l'épidémiologie : ce sont les modèles agents.

Des simulateurs basés sur des modèles agents sont des alternatives aux modèles compartimentaux et peuvent produire des scénarios plus riches que ces derniers. Les modèles agents permettent d'ôter l'hypothèse d'une population homogène et donne lieu à des interactions entre les agents et l'environnement.

Il existe plusieurs raisons d'intégrer les structures des interactions dans un modèle de diffusion

de virus et l'une d'elles est la flexibilité qui permettrait d'implémenter des interventions ciblées, des interventions qui affectent certains groupes différemment d'autres. En effet, un modèle agent peut être vu comme une population artificielle qui représente des agents autonomes qui interagissent suivant des règles prédéfinies sous des contraintes d'environnement prenant en compte les conditions spatiales et temporelles. De plus, ces interactions sont aléatoires et donc génèrent de la volatilité de façon naturelle dans les résultats.

Chaque modèle agent est constitué de composantes qui s'occupent des différents éléments qui prennent part durant une pandémie. Il y en a principalement 4 dans chaque modèle mais d'autres peuvent être rajoutées au besoin. Les 4 principales composantes de chaque modèle agent sont: la composante **société**, la composante **mouvement**, la composante **maladie** et la composante **environnement**.

La **composante société** est là pour créer et caractériser la population simulée par le modèle agent. La taille de la population est déterminée à ce stade là et les agents de cette population peuvent se voir assigner une multitude d'attributs qui peuvent les faire distinguer les uns des autres. Ces attributs peuvent par exemple être reliés au sexe de l'agent, à son âge, sa classe socio-professionnelle, bref tout ce qui peut être considéré comme pertinent par l'utilisateur et qui peut amener de la diversité dans la représentation.

Et ces agents doivent bien se rencontrer comme dans la vraie vie ! Les interactions inter-agents est déterminée par la **composante mouvement** qui va édicter la fréquence à laquelle les agents vont interagir ensemble. Cette fréquence est nuancée par rapport aux attributs de chaque agent pour essayer de refléter les comportements de chacun. Effectivement, l'entourage d'un enfant à l'école maternelle est différent de celui d'un jeune salarié et de celui d'un retraité qui profite de sa retraite. Cette composante peut proposer différents types de rencontre. Certains modèles agents mettent en place des schémas de villes en créant des sites de confluence où peuvent se retrouver les agents comme par exemple les écoles, les supermarchés, les habitations, etc.

Et lors de la rencontre de deux agents, un sain et un autre infecté, la **composante maladie** joue deux rôles. D'une part, elle est là pour gérer la transmission, ou pas, de la maladie d'un agent à un autre. La probabilité de cette transmission de la maladie va dépendre de toutes les caractéristiques de l'agent infectieux ainsi que du milieu dans lequel il se trouve. D'autre part, si la maladie est bien transmise, la composante va accompagner l'agent durant son évolution dans les différents états de la maladie. Elle décidera par exemple du temps de séjour dans chaque état, les probabilités de passage d'un état à un autre, le taux d'infectiosité maintenant que l'agent possède la maladie, son état d'hospitalisation, s'il se rétablira de la maladie ou s'il y succombera.

Finalement, la **composante environnement** vient altérer les trois composantes précédentes pour les nuancer. Elle rajoute, par exemple, l'effet de l'ensemble de politiques de restrictions des contacts humains et des déplacements en réponse à la pandémie de Covid-19 en France. En conséquence, elle change les dynamiques de population et donc touche à la composante mouvement. De plus, elle peut introduire des évolutions de la maladie sous la forme de variants plus ou moins puissants et donc peut amener des éléments supplémentaires à la composante maladie comme de nouvelles probabilités d'infections et/ou de passage d'un état à un autre. Aussi, la composante société peut se voir complétée de nouveaux attributs en lien avec la mise en place des campagnes de vaccination.

## Visualisation

À chaque agent lui est attribué un vecteur d'attributs qui lui sont spécifiques : l'âge, le sexe, le taux d'immunité, les déterminants de la contagiosité individuelle, les temps de séjour dans les différents

états de la maladie. De plus, ces caractéristiques peuvent se voir rester invariables ou évoluer dans le temps et dépendamment dans l'état dans lequel ils se trouvent. Les règles de ces évolutions et celles qui gouvernent les interactions peuvent être les mêmes pour toute la population ou configurées pour s'adapter à une hétérogénéité dues aux différents comportements sous-jacents des classes pouvant être partitionnées. Grâce à cette adaptabilité, le comportement général d'une pandémie est de porter l'attention sur la manière dont ces réactions observables affectent les résultats des simulations. Ils visent alors à sonder le lien entre les propriétés individuelles et le motif des interactions inter-agents. Donc contrairement aux modèles compartimentaux, les modèles agents se focalisent non seulement sur l'aspect global de la dynamique mais incluent aussi le côté social et individuel.

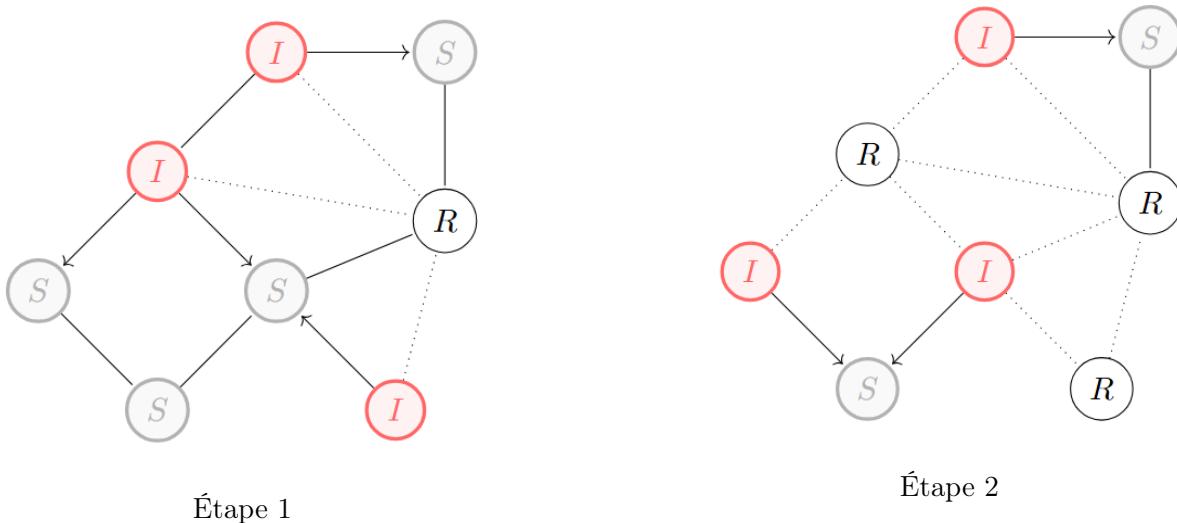


Figure 2.19 – Changement des états des individus suivant les interactions simulées.

Dans cet exemple, les agents transitionnent entre les mêmes états que ceux retrouvés dans le modèle *SIR*. Cependant, ces passages sont gouvernés par la nature des contacts qu'a chaque agent et par les lois de probabilité d'infection. Un graphe reproduit ici les liens et différents contacts possibles que peuvent avoir chaque agent. Les individus infectés peuvent contaminer, action représentée par des flèches orientées, les contacts qui sont susceptibles d'attraper la maladie et non les personnes rétablies, d'où les pointillés. L'étape suivante révèle trois choses :

1. deux individus susceptibles ont été contaminés. Ils peuvent à leur tour diffuser la maladie à leurs contacts,
2. deux des trois individus infectés à la première étape deviennent rétablis, la phase de maladie étant finie,
3. un des agents susceptibles n'est finalement pas infecté (en haut à gauche). Il reste susceptible mais risque toujours de devenir infecté à la prochaine étape.

C'est ce type de schéma que suit un modèle agent avec une population évidemment plus grande lorsque cela concerne les épidémies. De plus, des états peuvent être rajoutés comme dans les modèles compartimentaux afin d'affiner la maille et le détail du modèle, permettant alors de suivre les évolutions des différents états si nécessaire. Les modèles agents font alors partie des modèles essentiels en épidémiologie, pouvant produire des résultats quand des expériences à taille réelle ne sont pas viables.

Durant la pandémie de la Covid-19, un des rôles majeurs des modèles épidémiologiques est de comprendre le comportement d'une pandémie. Et pour le comprendre pleinement, il fallait regarder la pandémie non seulement comme un phénomène biologique ou qui n'a de lien qu'avec la santé

mais aussi comme un phénomène sociétal. Le cours d'une pandémie est tout autant impacté par la puissance du virus que par la réaction politique et les interventions non-pharmaceutiques prises. Dans cette perspective, les modèles agents complexes et réalistes ont dû être nécessaires. Chaque modèle compartimental peut éventuellement être répliable par un modèle agent. Une des raisons qui poussent à se tourner vers ces derniers est qu'ils amènent beaucoup plus de détails aux analyses et rajoutent des dimensions qui manquent aux modèles compartimentaux. Or cela se fait en échange d'un temps de calcul qui peut devenir significativement plus long à force de vouloir rentrer dans les détails, des détails qui nécessitent une puissance de calcul afin de suivre chaque agent. De plus, le code nécessaire n'est pas évident à mettre en place comme pourrait l'être le cas chez les modèles compartimentaux. Un autre désavantage est que le volume de paramètres nécessaire peut être plus élevé pour les modèles agents et cela peut poser un problème à l'instar de pandémies\*. En effet, le volume d'informations sur les caractéristiques de la maladie sont peu nombreuses voir inexistantes. Cependant, les modèles agents sont plus faciles à interpréter et à analyser que les modèles compartimentaux. Le traçage précis des agents le permet.

## 2.7 Modèle *CovDyn*

Chaque modèle possède donc ses caractéristiques suivant le degré de précision et le degré de liberté accordé à l'utilisateur. Certains modèles sont dits généraux car ils peuvent s'adapter à plusieurs types de maladies, de populations et de modes de vie alors que d'autres sont dits spécifiques lorsqu'ils ciblent une maladie, une population et des dynamiques précises.

Le modèle utilisé dans le cadre de ce mémoire, appelé modèle *CovDyn* ou *Covid Dynamics model*, se base principalement sur un modèle agent. Le modèle est une extension [2] du modèle de l'Institut national de la santé et de la recherche médicale, ou modèle Inserm, qui lui-même est une extension de celui présent dans le modèle de Di Domenico et al [1]. Le fonctionnement de celui-ci va alors être détaillé ainsi que les premières modifications opérées afin de l'adapter au contexte assurantiel.

Le modèle *CovDyn* [5], héritier des enseignements antérieurs, peut être vu comme une imbrication de plusieurs composantes : une composante de société qui construit la population virtuelle, une composante de mouvement qui régit les interactions de ses agents, une composante de maladie qui s'occupe de tous les aspects de cette dernière et enfin une composante environnement qui englobe le tout pour émuler des changements qui affectent les trois autres composantes. Il a donc 4 principales composantes.

### 2.7.1 Composante société

La première composante, appelée **composante société**, va créer une population artificielle d'agents et la caractériser. Par cela, elle va prendre en compte le nombre d'agents total qui seront simulés et va les répartir dans différentes catégories. L'effectif total et la distribution par âge de la population se basent sur une étude de l'INSEE qui considère environ 64 millions d'individus répartis en 17 classes d'âge de 5 ans chacune allant de l'âge 0 à 79 avec une catégorie supplémentaire pour les individus âgés de 80 ans et plus. La répartition est présentée dans le tableau suivant.

Or, le simulateur ne va pas gérer 64 millions d'agents ainsi que les interactions de chacun. Ne serait-ce que de les créer pour une seule simulation demanderait 64 millions d'étapes. Il faudra s'imaginer que

---

\*. Les modèles agents pourraient avoir moins de paramètres mais ils se rapprocheraient des modèles compartimentaux

Classe d'âge	Population	Classe d'âge	Population
0-4	3 740 814	45-49	4 329 120
5-9	3 970 379	50-54	4 129 128
10-14	3 925 573	55-59	3 991 112
15-19	3 668 641	60-64	3 973 525
20-24	3 797 702	65-69	3 736 593
25-29	3 978 588	70-74	2 588 969
30-34	4 000 149	75-79	2 146 168
35-39	4 273 138	80P	3 893 098
40-44	4 357 172		
		<b>Total</b>	<b>64 499 869</b>

Table 2.1 – Distribution de la population française en 2020; INSEE

chacun aura des contacts à gérer aussi, d'autres éléments à calculer. Cela fait grandir exponentiellement le nombre d'étapes et aura besoin de beaucoup de temps de calcul.

Un facteur de réduction de la taille de la population totale a donc été envisagé pour ces raisons de temps de calculs bien trop longs. Le choix retenu par l'équipe de recherche fut d'utiliser un facteur de réduction qui vaut 1% ce qui créera environ 640 000 agents.

Il a été question de revoir ce facteur de réduction afin de réduire encore plus ce temps de calcul. En effet, la problématique actuarielle posée au chapitre 3 requiert un nombre très important de simulations. Actuellement, avec une quarantaine de simulations, la trajectoire moyenne de la pandémie de la Covid-19 peut être correctement prédicté voir reproduite. Or c'est le comportement sous-jacent et sa variabilité qui servira pour ce calcul actuariel. Réduire le facteur de réduction pourrait potentiellement dégrader l'information. Il faut alors trouver une certaine valeur qui réduit le temps de calcul et qui conserve l'information de cette trajectoire moyenne. Il ne faut pas perdre de l'information.

Plusieurs valeurs inférieures à 1% ont été testées afin d'étudier la sensibilité du simulateur à ce facteur de réduction. A donc été remarqué qu'au-delà de 0,5%, la trajectoire moyenne se comportait de façon différente et sous-estimait les vagues épidémiques. En effet, les modèles agents simulent des systèmes complexes où les interactions entre agents produisent des comportements émergents. Et chaque agent joue bien un rôle avec des comportements et objectifs uniques à lui (ici le nombre et le type d'interaction qu'il a avec ses contacts par exemple). Et donc réduire le nombre d'agents ne limite pas seulement la capacité du modèle à capturer les dynamiques mais le rend moins robuste et beaucoup plus sensible aux variations. A donc été retenu un facteur de réduction valant 0,5%.

Finalement, la population qui est créée aura 0,5% de la taille de la population nationale française avec au total d'environ 320 000 agents. Des simulations seront alors réalisées, moyennées et corrigées par un facteur de 200 pour obtenir un résultat interprétable.

## 2.7.2 Composante mouvement

Ensuite vient la **composante mouvement** qui s'occupe d'initier les interactions entre les agents de la première composante. En d'autres termes, elle va mettre en « contact » les différents agents de la simulation pour reproduire des interactions de la vie réelle et ouvre donc à la possibilité de contamination et de propagation du virus de la Covid-19.

Ce nombre de mis-en-contact va varier d'un agent à un autre et dépendra principalement des

habitudes que possède chaque agent. Ces habitudes dépendent de la classe d'âge à laquelle il appartient et se décline pour chaque classe d'âge existante. Cette information se retrouve dans la **matrice de contact**, matrice dont les éléments, notés  $C_{i,j}$ , représentent le nombre de contacts moyen journalier d'un agent de la classe  $i$  avec un autre agent qui lui appartient à la classe  $j$ . Ce nombre moyen prend en compte les contacts dans 4 cadres différents : l'école, le travail, les habitations et autres lieux de brassage de populations [14].

La matrice de contact n'est pas remise en cause dans le cadre de ce mémoire. En effet, la *heatmap* ou carte de chaleur semble dégager des tendances qui sont raisonnables. Les éléments sur la diagonale principale possèdent les plus grandes valeurs (en général et par classe d'âge). Cela met en valeur le fait qu'un agent rencontrera dans les simulations plus d'agents appartenant à sa classe d'âge qu'à celle d'une autre. Elle reflète donc la fréquence des contacts entre individus du même groupe d'âge. Ces valeurs sont importantes car elles montrent à quel point les individus interagissent avec d'autres personnes de leur propre classe d'âge. Cela est pertinent pour des groupes comme les enfants dans les écoles et les personnes âgées dans les maisons de retraite par exemple.

Au début de chaque simulation, la matrice de contact (la matrice dans le simulateur est la version transposée de celle présentée et formalisée dans ce mémoire) va être utilisée pour générer les contacts quotidiens des agents suivant la classe d'âge à laquelle ils appartiennent. Un nombre total de contacts va être tiré par jour en utilisant une loi de Poisson dont le paramètre sera l'élément  $C_{i,j}$  correspondant de la matrice. Le paramètre peut être donc retrouvé en sommant ici les colonnes (dans le simulateur ça sera par ligne).

La loi de Poisson a été préservée car elle semble convenir à la situation du tirage de contact. En effet, la loi de Poisson est une loi de probabilité discrète, de support est  $\mathbb{N}$  qui associe une probabilité à un nombre d'événements, ici le nombre de contacts total, se produisant dans un intervalle de temps fixé, ici un jour. Est supposé aussi que les rencontres inter-agents se produisent indépendamment du temps écoulé depuis l'événement précédent.

Ensuite, des probabilités vont être calculées pour pouvoir redistribuer ce nombre de contacts total aux différentes classes. Elles seront calculées de façon à prendre en compte la pondération des contacts de chaque classe d'âge.

En réalité, pour une classe d'âge  $i$ ,  $i \in [1, 17]$ , à un moment  $t$ , le taux d'incidence  $\alpha_t^i$  est calculé comme tel:

$$\alpha_t^i = \beta_1 \beta_{2,i,INP}(t) \sum_{k=1}^{17} \left[ C_{i,j} \times \sum_{Z \in \{E, I_p, A, I_{ps}, I_{ms}, I_{ss}\}} i_Z \frac{Z_t^j}{N_{tot}^j} \right],$$

où :

- $C$  est la matrice de contact. L'élément  $C_{i,j}$  est le nombre moyen d'individus de la classe d'âge  $j$  rencontré, par jour, par un seul individu appartenant à la classe d'âge  $i$ ,
- $Z_t^j$  est le nombre d'individus dans le compartiment  $Z$ ,  $Z \in \{E, I_p, A, I_{ps}, I_{ms}, I_{ss}\}$ , de la classe d'âge  $j$  à l'instant  $t$ ,
- $N_{tot}^j$  est le nombre total d'individus de la classe d'âge  $j$  (infectés ou pas)
- $i_Z$  est l'infectiosité relative aux individus du compartiment  $Z$ . Quand l'agent n'est pas contagieux,  $i_Z = 0$  par exemple,
- $\beta_1$  est le taux d'incidence dans le cas de contact avec un agent infectieux pour une durée moyenne de contact,
- le taux de hasard relatif  $\beta_{2,i,INP}(t)$ ,  $INP \in \{Relaxed, Intensive, Extended\}$ , permet d'intégrer

80P	0.05	0.05	0.07	0.03	0.05	0.02	0.05	0.05	0.05	0.08	0.09	0.09	0.13	0.14	0.33	0.64	0.64	
75-79	0.05	0.05	0.07	0.03	0.05	0.02	0.05	0.05	0.05	0.08	0.09	0.09	0.13	0.14	0.33	0.64	0.64	
70-74	0.09	0.06	0.07	0.03	0.05	0.03	0.05	0.10	0.09	0.08	0.09	0.09	0.12	0.26	0.28	0.97	0.37	0.37
65-69	0.13	0.11	0.09	0.05	0.05	0.05	0.09	0.15	0.11	0.09	0.12	0.22	0.43	1.00	0.62	0.37	0.37	
60-64	0.19	0.18	0.10	0.07	0.08	0.11	0.18	0.26	0.18	0.15	0.27	0.51	1.41	0.52	0.66	0.28	0.28	
55-59	0.27	0.19	0.15	0.16	0.33	0.36	0.41	0.36	0.27	0.34	0.65	1.46	0.68	0.44	0.34	0.34	0.34	
50-54	0.31	0.21	0.27	0.48	0.81	0.96	0.81	0.92	1.11	1.09	1.92	0.92	0.44	0.32	0.34	0.54	0.54	
45-49	0.31	0.33	0.57	1.05	1.27	1.00	1.14	1.34	1.68	2.20	1.74	0.68	0.42	0.30	0.43	0.44	0.44	
40-44	0.50	0.85	1.08	1.02	0.97	1.26	1.46	2.10	2.95	1.57	1.59	0.79	0.52	0.43	0.51	0.45	0.45	
35-39	1.04	1.14	0.83	0.76	1.13	1.47	1.89	3.21	1.88	1.47	1.21	0.75	0.69	0.45	0.48	0.38	0.38	
30-34	1.21	0.92	0.40	0.53	1.17	1.78	2.92	1.76	1.59	1.26	1.18	0.92	0.64	0.45	0.25	0.31	0.31	
25-29	0.75	0.42	0.24	0.65	1.71	3.46	1.72	1.41	1.25	0.99	1.29	0.93	0.51	0.31	0.24	0.19	0.19	
20-24	0.34	0.16	0.29	1.36	3.93	1.94	1.01	0.76	0.92	0.95	1.02	0.61	0.34	0.21	0.16	0.15	0.15	
15-19	0.26	0.27	0.85	7.83	2.17	0.73	0.42	0.68	1.11	1.54	1.20	0.75	0.36	0.16	0.36	0.36	0.36	
10-14	0.41	1.01	6.99	2.30	0.26	0.14	0.53	0.84	0.98	0.61	0.83	0.57	0.26	0.26	0.31	0.45	0.45	
5-9	1.11	5.03	1.54	0.31	0.17	0.25	0.85	1.07	0.66	0.47	0.59	0.67	0.36	0.31	0.29	0.32	0.32	
0-4	3.80	1.06	0.24	0.12	0.20	0.53	0.72	0.70	0.31	0.41	0.25	0.54	0.40	0.20	0.11	0.24	0.24	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80P	

Figure 2.20 – Matrice de contact française transposée. Une personne de la classe d'âge 25-29 rencontrera, en moyenne, 1 personne de la classe d'âge 45-49 et 3,46 personnes de sa même classe d'âge.

les effets d'interventions non pharmaceutiques ou mesures et gestes barrières (port du masque, distanciation sociale, auto-isolation, confinements, couvre-feu, etc.) Il peut être vu comme un coefficient qui vient réduire le nombre moyen de contacts par jour et/ou leur infectiosité. Est supposé que  $\beta_{2,i,INP}(t)$  est décroissante avec des mesures sanitaires de plus en plus contraignantes. C'est-à-dire qu'il est improbable que les mesures mises en place exacerbent les contacts et devraient sûrement les diminuer. Ces taux de hasard relatifs seront exposés ultérieurement.

Dans le cas d'un seul groupe, i.e. une seule classe d'âge, cette équation du taux d'incidence, noté  $\alpha_t$ , devient:

$$\alpha_t = \beta_1 \beta_{2,INP}(t) \times C \sum_{Z \in \{E, I_p, A, I_{ps}, I_{ms}, I_{ss}\}} i_Z \frac{Z_t}{N_{tot}},$$

où  $C$  est le nombre moyen de contact journaliers.

Par jour simulé, les contacts de chaque agent seront tirés aléatoirement et de façon uniforme dans la sous-population de la classe cible, seront corrigés par des coefficients et seront finalement assignés à l'agent considéré pour la journée. Cette méthode favorise alors le brassage des agents des différentes classes.

Les éléments de cette composante mouvement permettent au simulateur *CovDyn* d'être général

dans le sens où la matrice de contact résume le comportement de la population française au niveau national. Et donc, se concentrer sur la population d'une ville française produirait, en moyenne, les mêmes résultats.

### 2.7.3 Composante maladie

Après avoir réalisé le tirage des contacts journaliers, la propagation de la maladie doit être mise en compte. Il faut alors simuler la réalisation, ou non, d'une transmission de la maladie d'un agent à un autre.

#### Sous-composante infection

À chaque pas de temps, ici unitaire, la **sous-composante infection** calcule la probabilité d'infection entre les deux agents concernés. Cette probabilité va être affectée par  $\beta_1$ ,  $\beta_{2,i,INP}(t)$  et  $i_Z$ . La probabilité quotidienne  $P_{infection}$  d'une personne susceptible d'être infectée appartenant à la classe d'âge  $i$  est donc décomposée de la façon suivante :

$$P_{infection}(i, INP, Z) = \beta_1 \times \beta_{2,i,INP}(t) \times i_Z$$

D'une part, la probabilité d'infection durant la durée d'exposition à un agent infecté est uniforme sur toutes les classes d'âge avec  $\beta_1 = 0.043$  [2]. De même pour l'infectiosité relative mais celle-ci varie en fonction du compartiment dans lequel se trouve cet agent infecté[2].

$i_A$	$i_{Ip}$	$i_{Ips}$	$i_{Ims}$	$i_{Iss}$
0,55	0,55	0,55	1	1

Table 2.2 – Variation de l'infectiosité relative en fonction du compartiment.

D'autre part, la probabilité d'infection dépend de l'environnement sanitaire dans lequel la population se retrouve. Cet environnement affectera les différentes classes d'âge de façon différente et cela par rapport à l'intensité des mesures entamées et les évolutions de ces dernières en fonction du moment où est considérée cette population.

#### Sous-composante états

Lorsqu'ils contractent la maladie, la **sous-composante états** fait transitionner les différents individus de la population simulée entre différents états de la maladie. Ils évoluent alors dans le graphe de ces états en suivant des probabilités de passage dépendant de leurs âges ainsi que des temps de séjours moyens. Ces deux dépendent des états dans lesquels les agents se trouvent. En termes de modélisation, ces états sont naturellement représentés par des compartiments (comme dans les modèles compartimentaux).

Le modèle *CovDyn* se voit alors doté des compartiments suivants:

1.  $S$  (Susceptible): ce sont les individus qui ne sont pas infectés par le virus;
2. La phase d'incubation comporte deux compartiments:

- quand un individu devient infecté, celui-ci passe du compartiment  $S$  au compartiment  $E$  (Exposé) qui contient les individus qui n'ont pas encore développé de symptômes et qui ne sont pas contagieux. Le temps moyen de séjour dans le compartiment  $E$  est de  $(t_i - t_p)$  où  $t_i$  est la durée d'incubation et  $t_p$  est la durée de la phase prodromique (voir l'élément suivant);
  - quand un individu appartenant à  $E$  commence à devenir contagieux, il est transféré au compartiment  $I_p$  (Phase prodromique) qui contient les individus qui sont toujours dans la phase d'incubation étant déjà contagieux mais sans symptômes apparents. Cette phase est généralement courte. Après un temps de séjour moyen de  $t_p$ , les individus rentrent dans un des quatre compartiments définis ci-dessous.
3.  $A$  (Asymptomatique): contient les individus qui ont complété la phase d'incubation, qui sont devenus contagieux mais ne possèdent aucun symptôme pouvant les distinguer. Passage avec probabilité  $p_A$ . La durée moyenne dans le compartiment est de  $t_s$ ;
  4. Les individus développant des symptômes (avec probabilité  $1 - p_A$ ) sont stratifiés en trois compartiments pour distinguer leurs sévérités :
    - $I_{ps}$  (l'état paucisymptomatique) : individus chez qui ne peuvent être perçus que peu de symptômes concernant leur maladie;
    - $I_{ms}$  (symptômes moyens): individus possédant des symptômes (de la fièvre par exemple) qui ne nécessitent pas d'hospitalisation. Le temps de séjour moyen dans les états  $I_{ps}$  et  $I_{ms}$  est le même que celui de l'état  $A$ ;
    - $I_{ss}$  (symptômes sévères): individus dont les symptômes les mènent à l'hospitalisation. Ils restent dans ce compartiment pour une durée moyenne de  $t_{bh}$  pour ensuite passer à  $H_1$ . Sachant que la maladie sera symptomatique, les probabilités de rentrer dans  $I_{ps}$ ,  $I_{ms}$  ou  $I_{ss}$  sont respectivement  $(1 - p_A)p_{I_{ps}}$ ,  $(1 - p_A)p_{I_{ms}}$  et  $(1 - p_A)p_H$  (dont la somme vaut un). Après l'écoulement du temps d'attente dans les états  $A$ ,  $I_{ps}$  ou  $I_{ms}$ , les individus quittent ces derniers et se retrouvent dans le compartiment  $R$  (*Removed* ou rétablis);
  5. La période d'hospitalisation comporte 4 compartiments.
    - $H_1$ : après avoir séjourné dans le compartiment  $I_{ss}$ , les individus passent dans le compartiment  $H_1$  avant de les répartir pour voir s'ils ont besoin de soins intensifs ou pas.
    - $H_2$ : individus qui n'ont pas besoin de soins intensifs;
    - $ICU$ : individus hospitalisés qui requièrent des soins intensifs;
    - $H_3$ : individus hospitalisés après les soins intensifs.

① Être hospitalisé est ici équivalent à un isolement de tout contact. Les agents se trouvant dans un de ces quatre compartiments ne peuvent plus être considérés durant le tirage des contacts.
  6. À l'issu de l'hospitalisation, les individus transitionnent vers l'un des deux compartiments suivants:
    - $D$  (Décès): individus décédés durant l'hospitalisation.
    - $R$  : agents rétablis de la maladie.

①  $R$  et  $D$  sont des états absorbants. Cela équivaut à une immunité à vie.

Durant toute cette phase d'hospitalisation, chaque individu suit une dynamique de chaîne de Markov, avec les probabilités de transitions qui suivent:  $p_{H_2|H_1}$ ,  $p_{ICU|H_1}$ ,  $p_{H_3|ICU}$ ,  $p_{D|H_2}$ ,  $p_{R|H_2}$ ,  $p_{D|H_3}$  et  $p_{R|H_3}$ .

① Une grande hypothèse est prise : les décès sont supposés tous découlent d'une hospitalisation. Les décès hors-hôpitaux ne sont pas pris en compte dans le modèle.

A été décidé de garder tous les compartiments vu leurs pertinences. Par exemple, le compartiment concernant la phase prodromique est essentiel. Lorsqu'il est question de maladie qui, une fois portée par un agent, peut se transmettre sans symptômes apparents, il est important de pouvoir simuler cette fenêtre de temps car elle peut avoir un impact sur la propagation de la maladie. La Covid étant une

de ces maladies, il faut garder ce compartiment.

De même, le compartiment  $A$  doit être préservé car une partie de la population pourra continuer à transmettre la maladie sans pour autant révéler des symptômes même après leurs phases prodromiques.

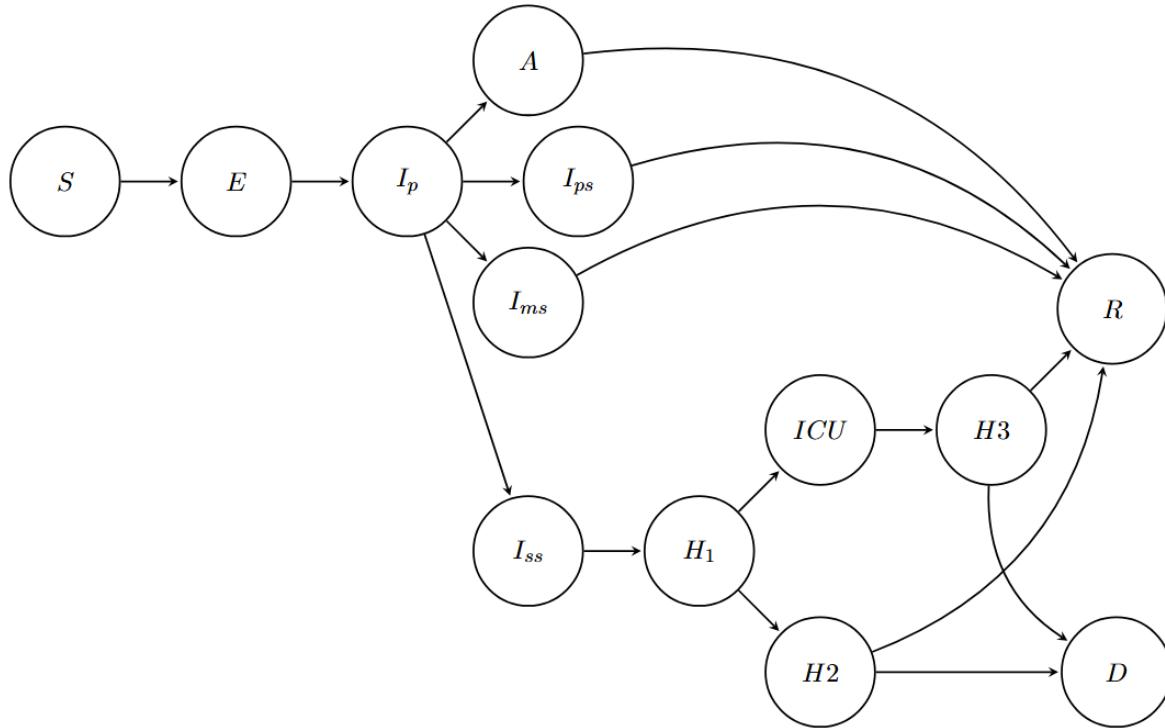


Figure 2.21 – Graphes des états

Cette dynamique s'inspire du modèle compartimental qui peut être décrit par le système d'équation qui suit:

$$\begin{aligned}
 \frac{dS(t)}{dt} &= -\alpha_t S(t) \\
 \frac{dE(t)}{dt} &= \alpha_t S(t) - \frac{1}{t_i - t_p} E(t) \\
 \frac{dI_p(t)}{dt} &= \frac{1}{t_i - t_p} E(t) - \frac{1}{t_p} I_p(t) \\
 \frac{dA(t)}{dt} &= \frac{p_A}{t_p} I_p(t) - \frac{1}{t_s} A(t) \\
 \frac{dI_{ps}(t)}{dt} &= \frac{(1 - p_A)p_{I_{ps}}}{t_p} I_p(t) - \frac{1}{t_s} I_{ps}(t) \\
 \frac{dI_{ms}(t)}{dt} &= \frac{(1 - p_A)p_{I_{ms}}}{t_p} I_p(t) - \frac{1}{t_s} I_{ms}(t) \\
 \frac{dI_{ss}(t)}{dt} &= \frac{(1 - p_A)p_H}{t_p} I_p(t) - \frac{1}{t_{bh}} I_{ss}(t) \\
 \frac{dH_1(t)}{dt} &= \frac{1}{t_{bh}} I_{ss}(t) - \frac{1}{t_{H_1}} H_1(t)
 \end{aligned}$$

$$\begin{aligned}
\frac{dH_2(t)}{dt} &= \frac{p_{H2|H_1}}{t_{H_1}} H_1(t) - \frac{1}{t_{H_2}} H_2(t) \\
\frac{dICU(t)}{dt} &= \frac{p_{ICU|H_1}}{t_{H_1}} H_1(t) - \frac{1}{t_{ICU}} ICU(t) \\
\frac{dH_3(t)}{dt} &= \frac{p_{H3|ICU}}{t_{ICU}} ICU(t) - \frac{1}{t_{H_3}} H_3(t) \\
\frac{dD(t)}{dt} &= \frac{p_{D|H_2}}{t_{H_2}} H_2(t) + \frac{p_{D|H_3}}{t_{H_3}} H_3(t) \\
\frac{dR(t)}{dt} &= \frac{1}{t_s} (A(t) + I_{ms}(t) + I_{ms}(t)) + \frac{p_{R|H_2}}{t_{H_2}} H_2(t) + \frac{p_{R|H_3}}{t_{H_3}} H_3(t)
\end{aligned}$$

Chaque compartiment est subdivisé en 17 pour accueillir les 17 classes d'âges formées arbitrairement dans la population.

Cependant, l'évolution en soit n'est pas indépendante de l'environnement de l'agent concerné. Il évolue dans le schéma de la maladie mais également dans un brassage modulé de la population. En effet, les institutions gouvernementales de santé réagissent à la présence d'une pandémie et mettent en place des dispositifs afin de réduire ses impacts et de la contenir pour qu'elle ne cause plus de dommages.

#### 2.7.4 Composante environnement

Finalement, la composante environnement vient rajouter la dernière couche de réalisme en mettant en place trois types de mesures virtuels. Mises en place par le gouvernement pour limiter la crise, les trois interventions qui ont été considérées sont: les INP-*Relaxed*, les INP-*Intensive* et les INP-*Extended*. Chacune d'elles a pris en compte un ensemble de mesures de protection pour atteindre un nombre de reproduction spécifique  $R_e(t)$ . Les trois types ont été implémentés dans le modèle et sont basés sur les restrictions et mesures mises en place en France. Du plus au moins contraignant, il y a :

1. les INP-*Extended*. Elles consistent en des mesures sanitaires très renforcées, représentant les confinements stricts, la fermeture des écoles, des universités, des différents commerces, etc.
2. les INP-*Intensive*. Elles comportent des mesures sanitaires strictes comme les couvre-feux, la distanciation sociale entre les gens, l'obligation du port du masque, le télétravail recommandé, la fermeture des restaurants, bars, etc.
3. et les INP-*Relaxed*. Ici, relâchement de certaines mesures (réouverture des lieux de restauration, des boutiques, plus de couvre-feu) mais maintien du port du masque et de la distanciation sociale.

Comme défini dans la partie 2.2, le nombre de reproduction  $R_e(t)$  dépend de la proportion d'individus encore non immunisés, susceptibles d'être infectés. En prenant en compte le type de mesure, cette proportion, ultérieurement notée  $R_e(t)$ , décroît avec le temps, suivant ainsi la décroissance du nombre d'agents dans le compartiment Susceptibles. Il s'écrit maintenant:

$$R_e(t) = R_0 \frac{S(t)}{N} \times \beta_{2,i,INP}(t)$$

Un enchaînement de ces trois types de mesures est possible et sera appelé un calendrier.

Classe d'âge	$\beta_{2,i,INP}(t)$		
	Extended	Intensive	Relaxed
0-4	0,143	0,209	0,267
5-9	0,143	0,209	0,267
10-14	0,143	0,209	0,267
15-19	0,143	0,209	0,267
20-24	0,217	0,318	0,406
25-29	0,217	0,318	0,406
30-34	0,095	0,139	0,178
35-39	0,095	0,139	0,178
40-44	0,043	0,064	0,081
45-49	0,131	0,191	0,244
50-54	0,122	0,179	0,229
55-59	0,183	0,268	0,343
60-64	0,195	0,285	0,364
65-69	0,195	0,285	0,364
70-74	1,991	2,917	3,725
75-79	1,991	2,917	3,725
80P	2,039	2,988	3,816

Table 2.3 – Valeurs du taux de hasard relatif  $\beta_{2,i,INP}(t)$  croissantes avec des niveaux de mesures de moins en moins contraignantes [2].

### 2.7.5 Composantes additionnelles : vaccins et variants

Deux autres composantes majeures existent également dans le simulateur : la première relate de **vaccin**. Elle met en place des campagnes de vaccination et propose, en quantités limitées, différents vaccins avec des efficacités variantes. Lorsqu'un agent est vacciné, plusieurs paramètres sont alors modifiés comme sa probabilité de passage à l'hôpital, sa probabilité de retomber malade, etc. La deuxième introduit un nouvel variant, le B.1.1.7 ou la variant Alpha, avec ses propres paramètres. Deux virus de la même souche peuvent alors circuler dans la population du modèle agent.

Ces deux composantes ont été désactivées manuellement dans le cadre de ce mémoire. En effet, comme le chapitre suivant va l'exposer, l'étude se réalise durant la première année d'une pandémie. Des hypothèses sont alors faites :

1. aucun variant n'apparaît durant la première année de l'apparition d'une pandémie. A été supposé que cet événement ne survient pas sur un horizon bien plus loin. Par exemple, le variant britannique de la Covid a été détecté en octobre 2020 et s'est propagé rapidement vers mi-décembre, donc relativement tard depuis le début de la pandémie.
2. aucun vaccin n'est prêt à être distribué et utilisé durant la première année. Le développement de vaccins prend du temps. Le cas de ceux de la Covid 19 furent une exception vu les fonds et ressources fournis mais ne sont quand même apparus qu'à partir de janvier 2021.



## Chapter 3

# Risque décès et assurance emprunteur

“... un incendie me faisait penser à l’assurance. Voilà une déesse qui n’est pas aimée, à beaucoup près, comme la fortune. On la redoute ; on lui porte de maigres offrandes, sans aucun enthousiasme. Et cela est aisément à comprendre ; les bienfaits de l’assurance ne se montrent qu’en même temps que le malheur.”

---

Alain, *Propos*, 1921, p. 302.

Jusqu’à présent, le simulateur gère une population d’agents dont chacun a des caractéristiques particulières et les fait vivre dans un contexte de pandémie. Donc un suivi de l’évolution peut être réalisé à l’échelle des agents de la simulation. Un assureur pourrait en tirer profit, pouvant utiliser le simulateur pour surveiller une certaine sous-population d’agents qui auraient les mêmes caractéristiques que ses assurés et étudier leurs comportements et les sinistres qu’ils pourraient engendrer. Par exemple, un agent qui développerait des symptômes pourrait potentiellement faire appel à des frais médicaux (comme les frais d’un séjour à l’hôpital), ou à des capitaux décès s’il devait décéder.

Puisque les agents d’un modèle agents peuvent être vus comme une population artificielle qui vont interagir entre-eux, le parallèle avec la vraie vie peut alors être réalisé: une population d’habitants peut être divisée entre ceux qui ont souscrit un contrat d’assurance et qui bénéficient de sa couverture, et ceux qui n’en ont pas. Pour reproduire cela, une nouvelle composante a été intégrée au simulateur afin d’ajouter cette dimension assurantielle.

Dans cette dernière partie, l’incorporation et l’usage de cette nouvelle composante sont détaillés. Cela comportera :

1. la description d’un portefeuille emprunteur transmis par un assureur,
2. l’adéquation de lois sur les distributions des capitaux par classe d’âge de ce portefeuille,
3. l’élaboration d’un nouvel portefeuille à effectif fixé à l’aide de ces lois,
4. les changements appliqués au simulateur afin de simuler une année entière de pandémie et intégrer cette dimension assurantielle,
5. le calcul de la distribution des prestations de décès dans le cadre d’un choc pandémie (scénario central),
6. et la réalisation de stress-tests.

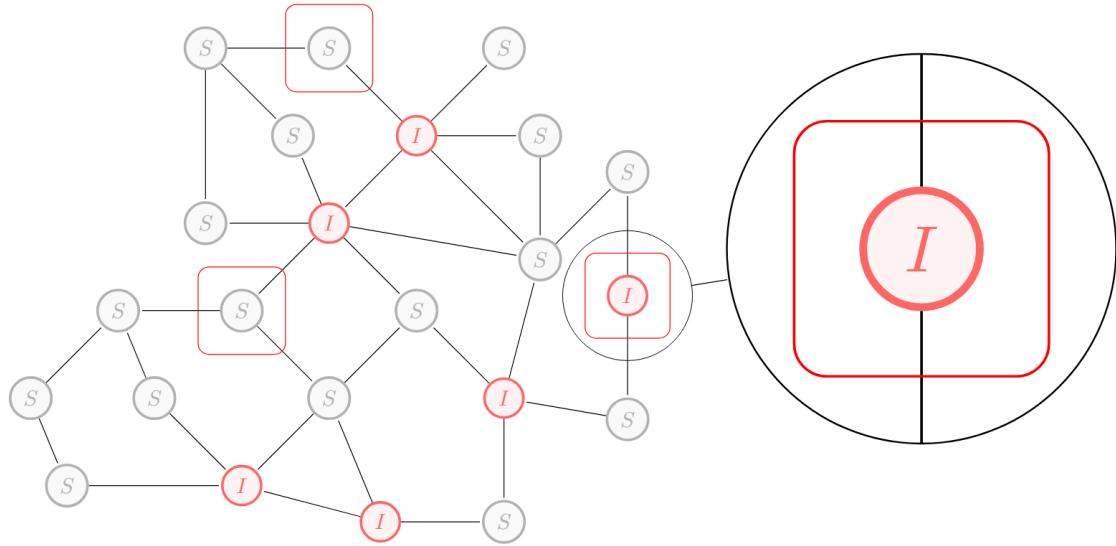


Figure 3.1 – Sélection des agents qui seront les assurés et suivi de leurs évolutions

### 3.1 Portefeuille emprunteur

Les agents du simulateur vont se comporter comme tels mais certains, ceux qui appartiendront à la sous-population assurée, posséderont des caractéristiques en plus qui intéressent l'assureur et qui lui permettront de mieux évaluer certains risques. Cette sous-population d'agents qui deviendront des agents assurés va être alors construite à l'image de celle composant le portefeuille d'un vrai assureur. Ainsi, l'assureur injecte les caractéristiques de ses vrais assurés dans la simulation et ce sont ces caractéristiques qui vont lui fournir la forme du risque qu'il entretient.

Ce mémoire va se concentrer sur de l'assurance emprunteur et se focalise sur le risque de décès sous-jacent. La sous-partie ci-dessous présente un portefeuille assureur qui servira de base afin de reconstruire un portefeuille avec un plus grand effectif d'assurés.

#### 3.1.1 Description du portefeuille assureur transmis

Le portefeuille transmis qui va être exposé repose sur de l'**assurance emprunteur**. Lorsqu'un ménage octroie un prêt, quasi-systématiquement, il va devoir l'accompagner d'une assurance. Elle n'est pas légalement obligatoire en France, mais elle est systématiquement exigée par les banques comme condition d'octroi du crédit. Elle vient couvrir le remboursement du montant du prêt restant en cas de situations qui affectent la capacité financière de l'assuré. Elle inclut le décès de l'assuré, la perte de revenu générée par une incapacité temporaire de travail, une perte totale et irréversible d'autonomie, etc. Dans ce mémoire, seul le **risque de décès** sera pris en compte.

Le portefeuille emprunteur transmis par l'entreprise d'assurance recense 15 610 prêts en total. Il a été arrêté au 1<sup>er</sup> juillet 2023. Les données transmises par l'assureur comportent :

- l'identifiant du contrat anonymisé,
- l'identifiant de l'assuré anonymisé,
- l'objet du prêt,
- la date du début du prêt,
- la date de fin du prêt,
- la durée différée de début du remboursement,
- le montant du prêt,
- le montant initial du prêt,
- et la quotité de l'assureur.

Un nettoyage des données a été effectué et des choix ont été faits. Toutes les lignes identiques ont été supprimées. Il n'y aura donc pas de redondance : un prêt ne pourra pas apparaître deux fois dans la base. De plus, les prêts ayant des montants initiaux anormaux ont été retirés, par exemple les montants négatifs, nuls ou très faibles (quelques dizaines d'euros). Les types des emprunts restants sont variés : des emprunts pour la consommation pour acheter des appareils électroménagers, payer des factures, acheter une voiture, des emprunts pour acheter une résidence principale ou secondaire, ou des emprunts professionnels utilisés par des compagnies à des fins d'investissements locatifs, acquisition de matériel, etc. Une entreprise n'est effectivement pas soumise au risque décès mais au risque de faillite durant une pandémie et en l'état, le simulateur ne permet pas de simuler ce genre de cas.

① Les montants des capitaux restants dus, ou CRD, n'étaient pas disponibles dans la base de données communiquées par l'assureur. Dans le cas où ils avaient été transmis, ils auraient été utilisés par la suite car ils représentent, de façon la plus juste, le risque actuel de l'assureur et permettraient d'obtenir de meilleurs résultats en termes des distributions de prestations. Compte tenu des données disponibles, il a été décidé d'utiliser les **montants initiaux** au lieu des montants restants dus. Donc, en prenant les montants initiaux, cela introduit une déformation par rapport à la vraie vision qu'aurait eue un assureur. Or, cette distorsion reste prudente.

Il reste en total 15 476 prêts. 10 414 des prêts ont été remboursés, soit 67% de l'effectif total, et il reste 5 062 en cours (dénoté par une date de fin de prêt *9999-12-31T00:00:00Z*).

Ensuite, la base a été croisée avec une base de données recensant tous les assurés ainsi que toutes les garanties souscrites, et pas seulement ceux possédant une assurance emprunteur. Elle a été transmise ultérieurement par l'assureur après une demande. En effet, l'âge ou la date de naissance des assurés ne figuraient pas dans la base des prêts. La démographie du portefeuille peut dorénavant être explicitée. Elle servira dans la suite d'une part afin de regrouper les assurés par classes d'âge et d'une autre part préserver la proportion des classes d'âge lors de la construction du nouvel portefeuille.

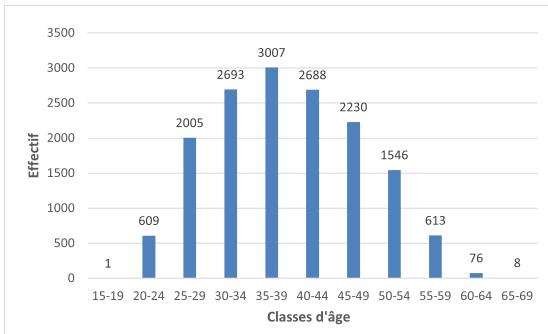


Figure 3.2 – Effectifs des assurés par rapport aux différentes classes d'âge

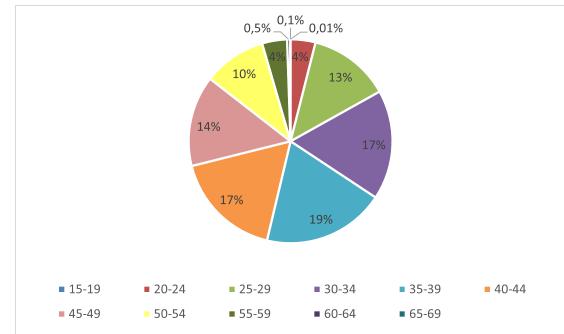


Figure 3.3 – Proportion des classes d'âge du portefeuille transmis

Le portefeuille comprend des assurés dont les âges varient de 19 à 68 ans. Plus de la moitié d'eux

appartiennent au groupe des 30-44 ans.

Et comme mentionné précédemment, le portefeuille ne sera donc pas utilisé comme tel mais servira de base afin de caractériser un nouveau portefeuille. Ce dernier servira de fichier en entrée du simulateur afin d'assigner le rôle d'assuré à un sous-ensemble d'agents.

Afin de le construire, il a fallu déterminer les lois de distributions des montants initiaux au sein de chaque classe d'âge. Elles seront ensuite utilisées pour les tirages de montants lors de la création des agents assurés dans le simulateur.

### 3.1.2 Adéquation des lois des distributions de capitaux par classe d'âge

Les assurés étant regroupés par classes d'âge, la calibration des lois concernant les montants sur chacune de ces classes d'âge peut être réalisée.

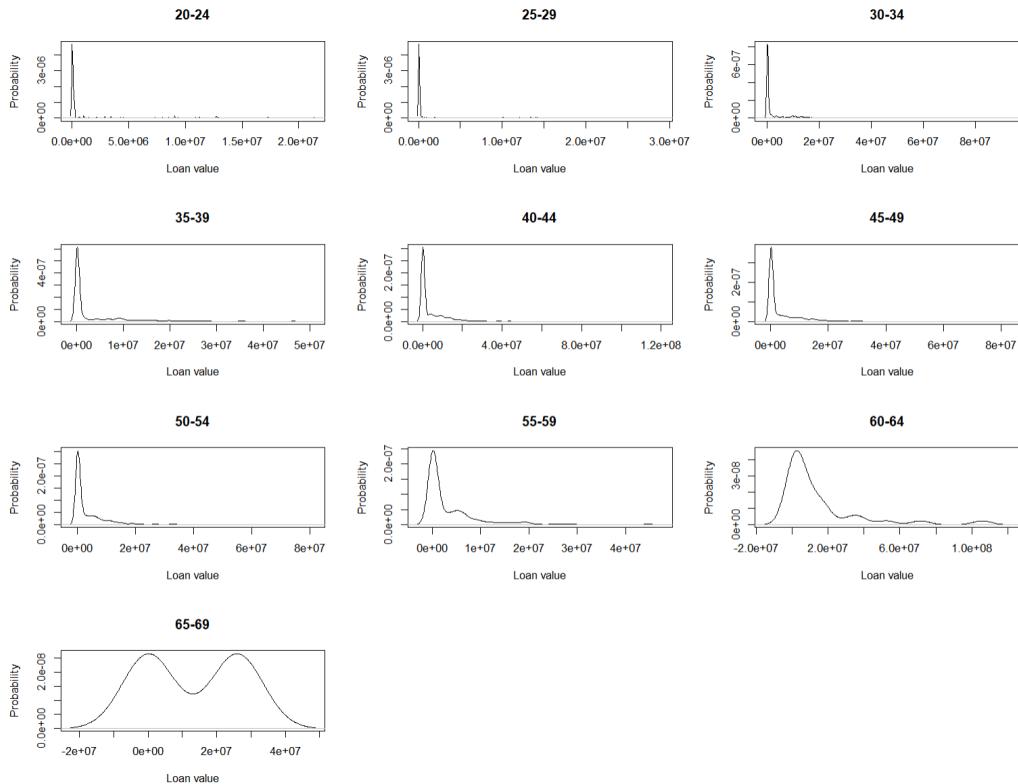


Figure 3.4 – Distribution des montants initiaux assurés par classe d'âge

A été délaissée la classe d'âge 15-19 car elle ne contenait pas assez de points, à savoir un seul. Cela ne permet pas d'avoir une distribution des montants des prêts initiaux de cette classe d'âge. Elle sera donc retirée du portefeuille.

Les formes de distributions semblent se ressembler. Une seule loi pourrait alors être utilisée avec différents paramètres pour chaque classe d'âge. Les tests d'ajustements utiliseront deux lois avec un support positif : la loi exponentielle et la loi Gamma. Il serait intéressant de reproduire ces décroissances rapides apparaissant sur les courbes de distributions. Ces tendances sont reproductibles par ces deux lois.

La loi exponentielle ne permettait pas un bon ajustement : elle pourrait être remise en question, car elle attribue un poids non-négligeable à des valeurs de montants très élevés. La loi Gamma offrait alors plus de flexibilité et a finalement donné un meilleur ajustement pour chaque classe d'âge. Les paramètres  $\alpha$  et  $\beta$  de chacune des lois Gamma ont été estimés en utilisant la moyenne et la variance des montants de chaque classe.

$$\text{Soit } X \hookrightarrow \Gamma(\alpha, \beta), \text{ alors } \mathbb{E}[X] = \frac{\alpha}{\beta} \text{ et } \mathbb{V}[X] = \frac{\alpha^2}{\beta^2} \Leftrightarrow \alpha = \frac{\mathbb{E}[X]^2}{\mathbb{V}[X]} \text{ et } \beta = \frac{\mathbb{E}[X]}{\mathbb{V}[X]}$$

L'étude statistique n'a pas été approfondie, car elle ne constitue pas l'objectif principal de ce mémoire.

Les lois ainsi trouvées, elles permettront de réaliser les tirages aléatoires des montants de prêts initiaux. Et afin de savoir combien de fois vont être réalisés ces tirages, le nombre d'assurés du nouveau portefeuille doit être déterminé et réparti dans les différentes classes d'âges.

### 3.1.3 Choix du nombre d'assuré

Une fois l'effectif total de la population assurée choisi, il sera distribué dans différentes classes d'âge, les mêmes que celles stratifiées par la matrice de contact de la composante mouvement.

① Un assureur pourrait directement construire des agents reflétant son portefeuille et ne pas passer par la construction d'un nouveau portefeuille.

Est supposé ici que le périmètre de l'assureur s'étend sur une ville de la taille de Nantes (intra-muros), d'une population assez similaire à la population réduite du simulateur avec, pour rappel, environ 322 499 agents (pour environ 325 070 d'habitants en 2022). Le nombre d'assurés finalement fixé s'élève à 200 000. Après l'application du facteur de réduction, 1 000 assurés seront finalement gérés par le simulateur. Ce nombre est alors distribué de la manière suivante :

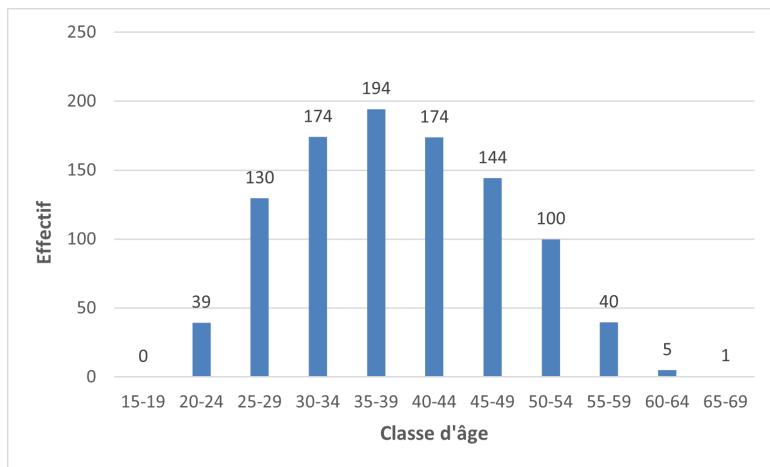


Figure 3.5 – Distribution des assurés par classe d'âge - De 20 à 69 ans

En moyenne, un portefeuille simulé par le modèle possède environ **320 millions d'euros** en tant que capitaux sous risques pour 1000 agents (ou 0.5% de 200 000).

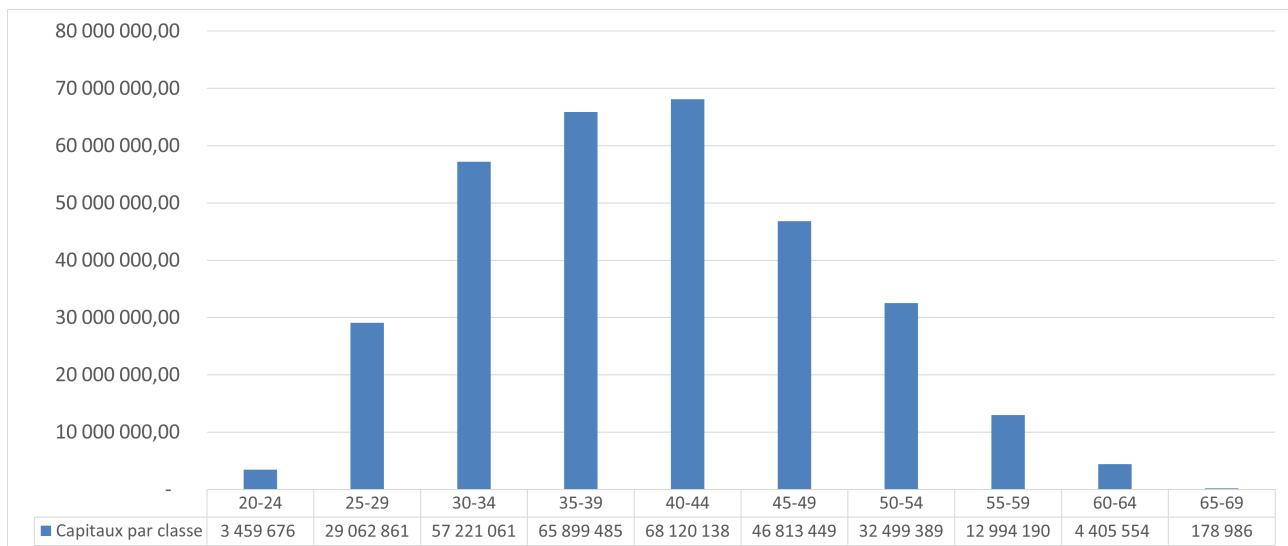


Figure 3.6 – Distribution des capitaux assurés par classe d'âge - Concentration chez les adultes

Il est important de noter qu'une hypothèse supplémentaire et implicite est faite ici : la pandémie se déroule sur un horizon relativement court par rapport à la durée d'un prêt. Avec cela, a été considéré que les montants sous risques n'évoluent pas durant cette période-là, i.e. 1 an. L'approche reste cependant prudente, car le montant des prêts sont décroissants car étant remboursés avec le temps.

Finalement, deux hypothèses sont à garder en tête :

1. ce sont les montants des prêts initiaux qui ont été utilisés,
2. et ces montants sont constants sur une durée d'un an.

Néanmoins, cela n'affecte pas les analyses des scénarios qui viendront car les hypothèses sont les mêmes pour tous.

Lorsque la composante population va créer les agents de la simulation, la **nouvelle composante assurance** va générer des montants suivants les lois de distributions estimées et va les assigner au bon sous-effectif d'agents dits assurés. À l'issue de chaque simulation, deux fichiers sont produits par le simulateur :

1. le premier, nommé *capitaux\_sous\_risque.csv*, liste les montants des prêts tirés par classe d'âge,
2. le deuxième, nommé *capitals\_evo.csv*, présente l'évolution des prestations de décès cumulées par jour et par classe d'âge. Les prestations de décès cumulées augmentent à chaque fois qu'un assuré décède et est incrémentée de la valeur assurée.

Afin d'exploiter les résultats plus facilement, la composante assurance permet aussi de réaliser l'agrégation des différents fichiers *capitals\_evo.csv*. Cela permet d'obtenir un dernier fichier qui affiche, par ligne, le montant total versé à la fin de chaque simulation. Ce fichier sera utile afin de retrouver les distributions de prestations dans différentes situations qui seront exposées prochainement.

La création de la sous-population assurée étant expliquée, il reste à lancer des simulations en utilisant le modèle *CovDyn* afin de se pencher sur le comportement de la pandémie et de déterminer la distribution des prestations à verser concernant le décès d'assurés.

## 3.2 Modifications du simulateur

Des changements ont été amenés au simulateur *CovDyn*. En effet, la version du simulateur utilisée par les Hospices Civils de Lyon, développée sur Python, a été calibrée de sorte à reproduire uniquement la deuxième vague. De plus, il n'est pas évident à modifier et manque d'interface graphique avec l'utilisateur qui manque de visibilité sur ce qui se passe lors du lancement des simulations ainsi que l'exploitation des résultats en sortie.

### 3.2.1 Construit et calibré pour la deuxième vague - Extension à la première

Le modèle *CovDyn* a été initialement calibré [2] pour reproduire la deuxième vague de la pandémie de la Covid-19. Les données d'hospitalisations et de décès disponibles ont été utilisées pour réaliser les différents calibrages concernant le taux d'hospitalisation, le taux de décès, de rétablissement ainsi que la distribution de la population dans les différents compartiments et tous les paramètres du simulateur. Cela permet alors de replacer le modèle à un stade avancé de la pandémie, juste avant cette deuxième vague. Ce scénario est dénommé *historical strain* dans le simulateur.

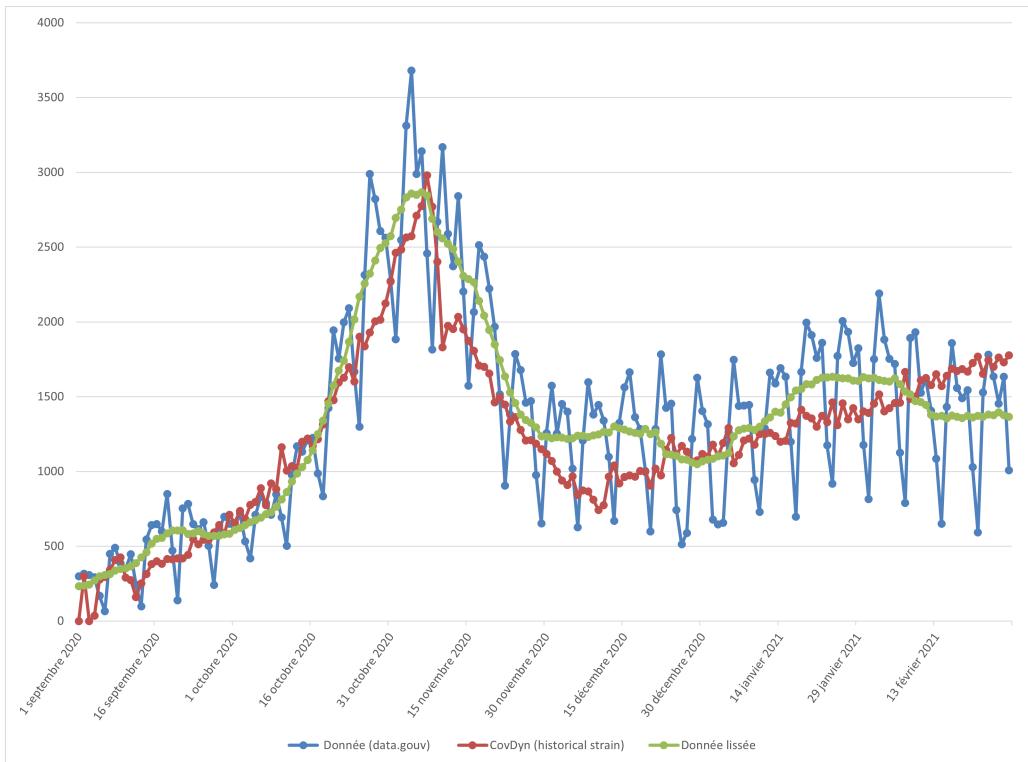


Figure 3.7 – Évolution du nombre journalier d'hospitalisations issu de simulations du scénario *historical strain*, deuxième vague de la pandémie en France

Le modèle a été lancé sous ce paramétrage et la moyenne de 40 simulations produit les résultats ci-dessus. La courbe verte est issue d'une moyenne mobile sur 7 jours appliquée sur les données extraites du site *data.gouv*. Dans ce cas, le *Mean Absolute Error*, ou *MAE*, entre les données publiques et les résultats du simulateur *CovDyn* vaut 201. C'est-à-dire, la valeur prédictive s'éloigne du vrai nombre d'hospitalisations journalières de 201 hospitalisations en moyenne, dans un sens ou dans l'autre (surestime ou sous-estime). À noter que le nombre moyen d'entrées en hôpitaux sur la période de

simulation (qui dure 179 jours ou environ 6 mois) s'élève à 1 341.

La procédure du simulateur n'est pas altérée par l'ajout du volet assurantiel et reproduit bien la deuxième vague comme prévu par le modèle.

Maintenant le but est de proposer un nouveau scénario et donc plusieurs enchaînements de mesures sanitaires non pharmaceutiques. Cela est fait dans le but d'étendre les simulations pour intégrer la première vague et pousser les simulations de cinq ou six mois à une année entière, englobant ainsi le début de la crise sanitaire en France.

Le schéma d'interventions non-pharmaceutiques n'est pas connu en avance. A été alors supposé un même déroulement que celui qui a été mis en place durant la pandémie de la Covid-19. Ainsi, le modèle commence à  $t = 0$  le 1<sup>er</sup> janvier 2020 avec des mesures de type *Relaxed* jusqu'au 18 mars 2020, le premier jour du confinement. Pendant celui-ci, des mesures *Extended* ont été placées suivies de mesures *Intensive* à partir du 12 mai pour simuler un relâchement des mesures mais pas au point de ne pas en avoir. De même pour la deuxième vague : cette dernière respecte le même schéma en terme d'intensités des différentes mesures.

① Imposer un confinement sans que la plupart des gens ne soient contaminés n'a pas de sens. Pour aller plus loin, il faudrait définir un critère objectif de passage d'une *INP* à une autre. Des tests seront alors effectués dans le cadre de cette étude pour examiner l'influence de ce second confinement et de la pertinence de le mettre en place.

Dans un second temps, il a fallu adapter le nombre d'exposés pour se replacer à l'instar d'une pandémie et non durant la deuxième vague. En modélisation, une pandémie est généralement déclenchée par 10 individus qui seront assignés au compartiment Exposé. Or ici, la pandémie a débuté beaucoup plus tôt donc ce chiffre devrait être plus important. De plus qu'avec un tel effectif d'exposés en début de simulation, une majorité des pandémies échoue, n'arrivant pas à démarrer, le confinement venant trop hâtivement/précocement interrompre l'évolution exponentielle et éteint la première vague. Par ailleurs, la distribution de ce nombre d'exposés dans les différentes classes d'âge doit être aussi pensée. Une distribution proportionnelle a donc été mise en œuvre, allouant ainsi  $E(0)$ , i.e. le nombre d'individus dans le compartiment Exposé à  $t = 0$ , à toutes les classes d'âge de ce compartiment en respectant la distribution de la pyramide des âges. Plusieurs calibrations ont été testées et la calibration retenue utilise l'initialisation  $E(0) = 2\,000$ . Elle produit en moyenne la meilleure projection par rapport au nombre d'hospitalisations fourni par les données officielles des entités médicales et gouvernementales. Bien que cela dépasse clairement le consensus épidémiologique, il est important de noter qu'à ce stade de l'étude, aucune autre méthode n'était disponible pour ajuster les paramètres épidémiologiques.

À noter que les paramètres utilisés (probabilités de transition, temps de séjour, etc.) restent inchangés car l'apparition de variants n'intervient qu'en décembre 2020.

Le simulateur étant calibré suivant le nombre d'hospitalisations et le nombre cumulé de décès, l'étude des résultats du modèle *CovDyn* comparés à la donnée officielle est primordiale à réaliser. La période de simulation est alors doublée comparée à celle du scénario *historical strain*. Elle commence le 1<sup>er</sup> janvier 2020 et se termine le 31 décembre 2020.

La courbe bleue représente la donnée officielle. Ses deux croissances \* représentent le début des deux vagues qui ont eu lieu durant la première année. Elles sont assez similaires dans leurs comportements. Elles suivent la même tendance qu'un début de vague que ce soit pour la première ou la deuxième.

---

\*. La première ne commence qu'à partir du 18 mars ; la donnée n'a commencé à être fournie qu'à partir de cette date-là. Cela explique l'absence d'une courbe croissante dès le début de l'année et le saut ce jour-là.

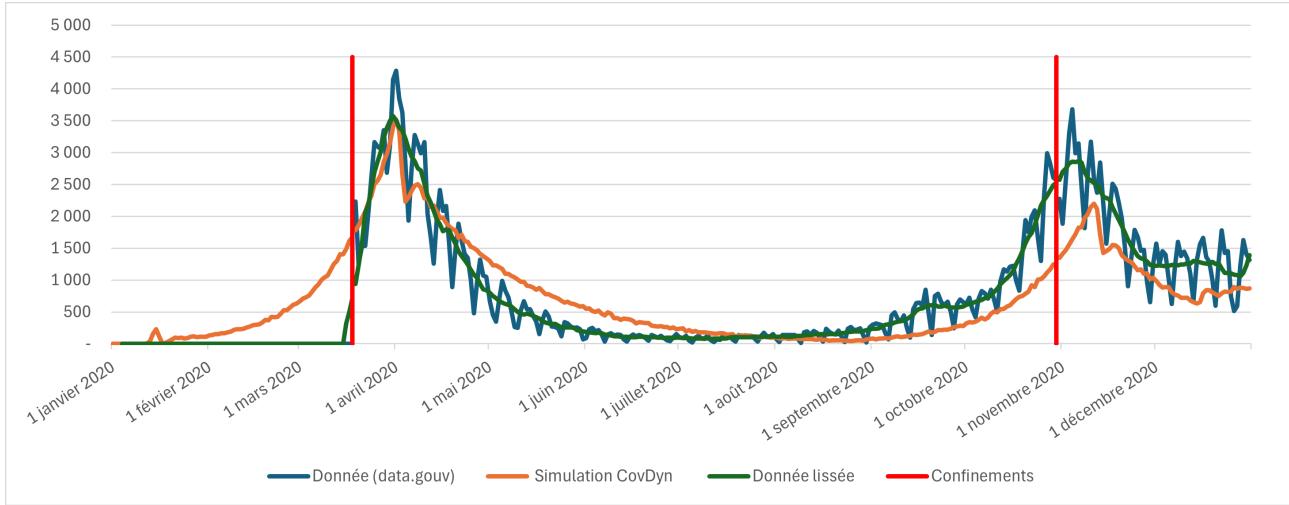


Figure 3.8 – Nombre d'hospitalisations journalières, moyenne de 400 simulations - 1 an de pandémie avec  $E(0) = 2\,000$ .

Cependant, ce sont les décroissances qui suivent qui ne sont pas identiques. En effet, la première décroissance est plus sévère que celle de la deuxième vague. Cela pourrait sous-entendre que l'intensité des interventions qui ont été mises en place afin d'endiguer la deuxième vague était moins importante que celle de la première. Et en effet, une étude de l'Inserm et du CHU de Bordeaux affirme l'efficacité des interventions et que de la deuxième vague a effectivement ralenti la survenance de nouvelles hospitalisations et de nouveaux décès (réduisant la transmission de 74% (intervalle de confiance (IC) à 95% : 69-77), mais pas autant que les interventions de la première vague (réduction de 84% (83-85))[11]. Cela est explicité car les paramètres  $\beta_{2,i,INP}(t)$  ont été **ajustés en partie à l'aide de cette deuxième vague**, découlant du fait que le simulateur a été construit dans le but de reproduire la deuxième vague. Donc les coefficients sont adaptés pour reproduire cette deuxième décroissance et non pas la première.

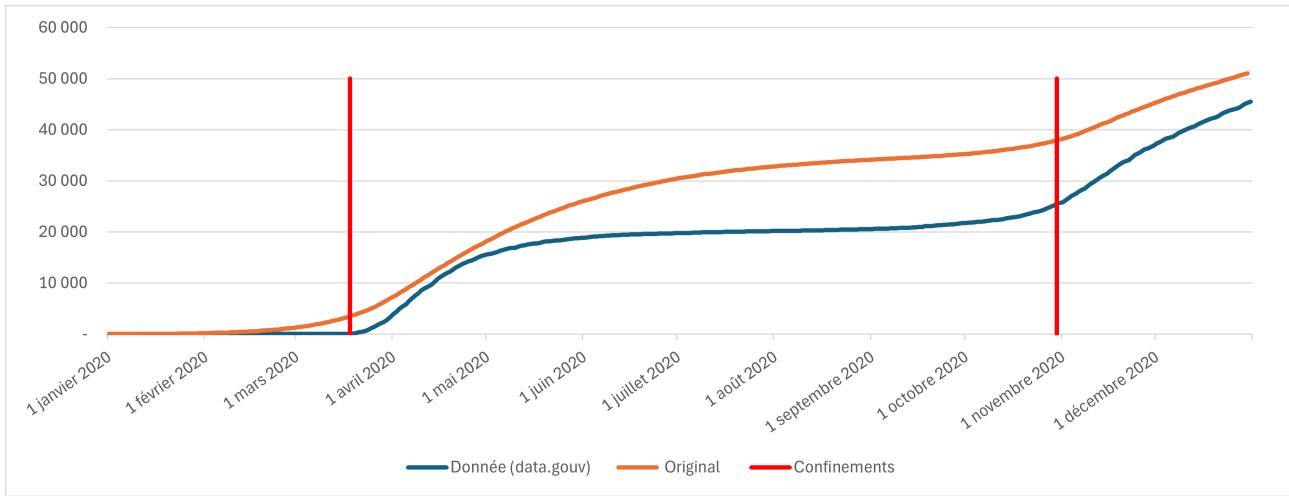


Figure 3.9 – Nombre de décès (en hôpitaux) cumulés, moyenne de 400 simulations - 1 an de pandémie avec  $E(0) = 2\,000$ .

C'est à cause de cela que la première décroissance de la courbe orange, celle qui représente en moyenne la trajectoire prédictive du simulateur, n'épouse pas exactement celle représentée par la

donnée publique. Cependant, elle reprend quand même sa forme générale. La *MAE* ici vaut 347. Le nombre d'hospitalisations est alors un peu surestimé à partir de fin mai jusqu'en fin octobre ce qui n'est pas forcément dérangeant puisqu'il garde une certaine prudence (et cela en surestimant le nombre d'hospitalisations journalières). La deuxième vague par simulation *CovDyn* est une bonne représentation de la deuxième vraie vague en termes d'hospitalisations journalières. Cela est d'autant plus rassurant, ajoutant ainsi une couche de certitude sur le fonctionnement du simulateur.

De même, le nombre de décès (issus des hôpitaux) cumulés reflète cela. Une croissance assez similaire jusqu'en mai diverge de la tendance réelle pour ensuite retrouver une même tendance croissante, néanmoins avec un écart. Fin octobre, il y a un écart d'à peu près 10 000 morts. Cela est dû aussi au fait que les paramètres  $\beta_{2,i,INP}(t)$  ont été calibrés suivant la décroissance des hospitalisations durant la deuxième vague. Et pour le décès soit comptabilisé, il faut passer par une hospitalisation. Donc, l'effet de surestimer le nombre d'hospitalisations journalières peut mécaniquement induire une hausse du nombre cumulé de décès. Finalement, au bout de 1 an, une différence d'à peu près 5000 décès se dresse entre la moyenne des simulations et le nombre total de décès (en hôpitaux) publié par les institutions publiques. Cela représente une surestimation d'environ 11% par rapport à ce qui s'est vraiment produit.

Il faut cependant garder en tête que cette donnée ne débute que le 18 mars 2020 alors que les simulations cumulent à partir du 1<sup>er</sup> janvier de la même année. Il manque alors tous les décès journaliers (comme pour les hospitalisations) entre le début de l'année et le début du confinement, i.e. le 18 mars. Les moyennes des simulations présentent donc un chiffre du même ordre de grandeur assez proche de ce qui a pu être observé.

### 3.2.2 Interaction avec l'utilisateur

Le simulateur a été modifié de telle sorte à améliorer l'interface utilisateur du simulateur. Maintenant, lorsque l'utilisateur démarre le simulateur, plusieurs questions lui sont proposées. Il devra saisir des réponses afin de préciser le type de simulation voulu. Ces questions relatent du :

1. **profilage du code.** Cela permet de lister le temps d'exécution ainsi que le nombre d'appels des différentes fonctions appartenant au code.
2. **scénario simulé.** Cela inclut le *historical strain*, la première vague et les deux vagues précédemment présentées. L'utilisateur n'aura plus besoin de changer à la main les parties du code.
3. **stress test**<sup>\*</sup>. 4 situations sont proposées : le scénario central qui englobe la Covid-19, et trois situations stressées.
4. **Covid aléatoire.** Cela crée de l'aléa chez les paramètres en entrée.
5. le **nombre d'assurés.** Cela permet de moduler la taille du portefeuille construit.

De plus, les fichiers en sortie seront directement traités pour obtenir la moyenne des simulations. Les fichiers d'intérêts ici sont les fichiers *evo\_pop.csv* que ressort le simulateur. Ils présentent l'évolution du nombre d'agents dans chaque compartiment à chaque pas de temps. Leur moyenne est calculée et est également proposée en sortie.

Une bonne reproduction du nombre d'hospitalisations journalières et le nombre de décès cumulés justifie alors une certaine viabilité du modèle. Ce dernier peut de manière fiable reproduire une année de pandémie qui se comporte comme la Covid-19. Cependant, les résultats des hospitalisations et de décès

---

\*. Ils seront détaillés dans la partie suivante.

ne sont pas connus en avance, de plus que les décisions sur le plan politique sanitaire sont difficilement envisagées qu'après leur mise en place. En supposant une réaction similaire à une arrivée future d'une pandémie, l'assureur devrait pouvoir se prémunir de la majorité des scénarios qui pourraient mettre en péril sa solvabilité.

### 3.3 Covid-19 et co.

Il sera intéressant de voir les distributions des prestations décès à la fin d'une année de pandémie ainsi que la valeur de la VaR à 99,5% concernées par le portefeuille construit. Cela se fera dans le cadre de la crise sanitaire de la Covid-19, mais également dans des scénarios stressés. Dans ces derniers, certains des paramètres vont être modifiés afin de regarder leurs impacts sur la valeur de la VaR et comparer la sévérité de ces scénarios.

Un calcul de la VaR à 99,5% a donc été implémenté et celle-ci peut former un des éléments pour le calcul du SCR concerné. En effet, ici, la VaR n'inclut que l'aspect du coût alors que le SCR devrait comprendre le coût **et** la fréquence.

Par ailleurs, il faudra au moins 200 simulations pour commencer à estimer cette quantité. La VaR est là pour prémunir l'assuré de 99,5% des événements ou d'un événement qui se produirait une fois tous les 200 ans. Donc, il faudra au moins 200 simulations pour la calculer. Finalement, 600 simulations ont été lancées. Ce nombre a été choisi notamment par des considérations de temps de calcul. Or, 600 simulations reste peu pour un SCR, surtout comme le nombre de zéros (simulations qui ne produisent pas de prestations) est très grand. Il faudra un nombre nettement supérieur afin d'avoir de consolider le résultat.

Les résultats sont analysés dans ce qui suit.

#### 3.3.1 Scénario central : crise sanitaire de la Covid-19

566 de ces simulations ne présentent aucun sinistre. Calculer le quantile à 99,5% donne une VaR qui vaut **825 059,8 euros**, soit 0,26% du montant total sous risque. Cela est concordant avec les résultats sortis par les assureurs qui révèlent des impacts mineurs qu'ils ont supportés durant la période de la pandémie de la Covid-19. Cela peut s'expliquer chez d'autres types d'assurances, mais il faut se demander pourquoi un portefeuille emprunteur, dont le risque primaire est le décès, n'a pas été autant impacté malgré une hausse de la mortalité de la population.

Une réponse est amenée par l'étude de la distribution du nombre de décès dans les différentes classes d'âge. Le journal l'Express publie l'infographie [17] suivante à partir de données proposées par Santé publique France. Elle résume la répartition du nombre total de décès à l'hôpital en France selon l'âge des patients depuis le 18 mars 2020 jusqu'au 14 avril 2021. D'après le graphique, les classes les plus impactées concernent celles liées aux personnes âgées, coïncidant ainsi avec le fait que ces personnes-là appartiennent aux groupes les plus vulnérables. En joignant cela avec le fait que dans le portefeuille ces classes d'âge n'existent pas ou que l'effectif associé est négligeable, cela donne une très faible chance que les personnes décédées appartiennent au portefeuille (cette probabilité n'est pas nulle, mais l'événement devient très rare). En d'autres mots, la pandémie a frappé de façon conséquente la « mauvaise » tranche d'âge ce qui induit un nombre relativement bas de sinistres et donc un risque moindre dans le cas d'une pandémie qui aurait les mêmes caractéristiques que la Covid-19.

① Les chiffres sont un peu plus élevés du côté de l'article de l'Express, car celui-ci s'étend un peu

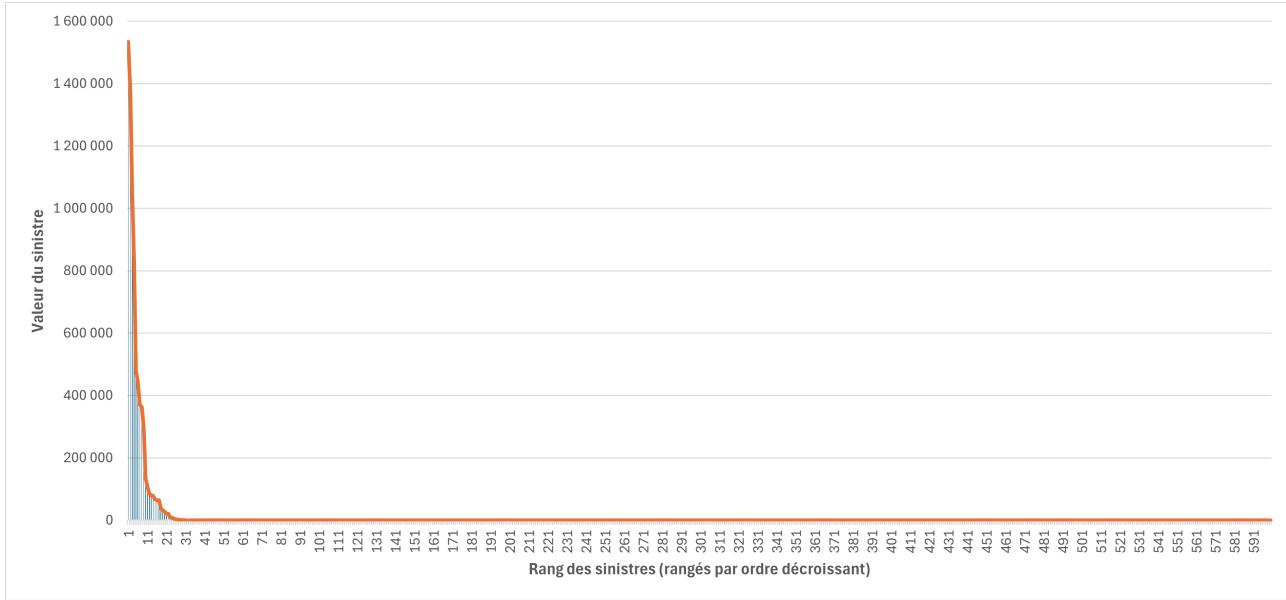


Figure 3.10 – Valeur des sinistres issus de 600 simulations, rangés en ordre décroissant.

plus longtemps dans le temps, couvrant une partie de la troisième vague qui a été la plus meurtrière (avec 40 000 morts sur une durée de 19 semaines) et qui a introduit le premier variant alpha dit alors « britannique ». Plus contagieux, il amène avec lui des couvre-feux et des contrôles aux frontières renforcés.

#### Décès en France par catégorie d'âge

Données au 14 avril ■ Nombre de décès à l'hôpital, selon l'âge des patients

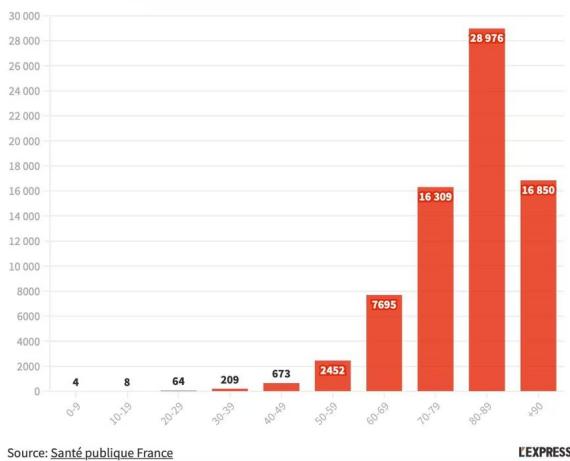


Figure 3.11 – Répartition des décès du 18 mars 2020 jusqu'au 4 avril 2021 dans les différentes classes d'âges.

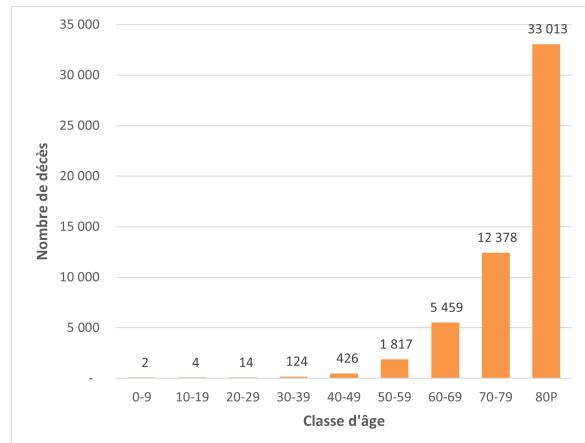


Figure 3.12 – Répartition des décès du 1 janvier 2020 au 31 décembre 2021 par classe d'âge.

En utilisant la formule proposée par la directive Solvabilité II, le SCR risque vie CAT vaut:

$$SCR_{vieCAT} = \sum_i 0,15\% \times CAR_i = 0,15\% \times 320\,654\,792,97 = 480\,982,18$$

D'ailleurs, cette valeur n'est pas directement comparable avec la VaR calculée plus haut. En effet, l'étude de ce mémoire ne se focalise que sur le coût du risque tandis que le SCR de la formule standard lui est censé couvrir le coût **et** probabilité de survenance de la pandémie.

Ce scénario n'est qu'**une seule réalisation** de la pandémie de la Covid-19. Le simulateur prend en entrée un jeu de données et de paramètres statiques, présumant la certitude de leurs valeurs. De plus, la Covid-19 aurait pu être beaucoup plus infectieuse, moins mortelle chez les personnes âgées, mais plus chez les jeunes par exemple. Les choses auraient pu se réaliser d'une autre manière : la probabilité d'infections, le taux d'asymptomatiques, les effectifs d'hospitalisations et le nombre de décès auraient pu être très différents.

Les simulations réalisées dans le scénario central peuvent servir dans le cadre d'une vision rétrospective sur la pandémie de la Covid-19 et donnent des pistes de recherches pour le calcul de SCR liés au risque pandémie. Or ce n'est qu'une reproduction du déroulement de cette dernière et il faudrait avoir une vision plus globale et pour englober des pandémies plus violentes ainsi que celles qui pourraient être moins importantes. Cela implique un changement des caractéristiques de la maladie en jeu et de voir comment se comporte le nombre de décédés par rapport à ces modifications et à l'impact de ce nouveau comportement sur la valeur du quantile utilisé pour calculer le SCR.

Changer les caractéristiques de la maladie en jeu revient alors à modifier la sévérité et/ou longévité du pathogène, que ce soit par la période symptomatique, les degrés de sévérité et leurs probabilités, le taux d'asymptomatiques, la probabilité d'hospitalisation et évidemment la probabilité de décès (ici, pour rappel, à l'hôpital). Dans le simulateur, cela se traduit par des changements des valeurs des différents paramètres en jeu qui gouvernent les interactions entre les agents et le caractère du virus.

Vont être distinguées trois situations fictives (qui pourraient ou non se réaliser un jour) qui se basent sur le scénario central (le vrai) de la pandémie de la Covid-19 en altérant certains paramètres pour tester leurs impacts respectifs sur la valeur du quartile à 99,5%. Vu leur pertinence, les probabilités d'hospitalisation ainsi que de décès après hospitalisation vont être modifiées selon certains aspects vu que ce sont eux dont se soucie le plus l'assureur dans le cadre du portefeuille emprunteur. Ces changements sont justifiés par le fait que l'événement qui déclenche la garantie est le décès. Or, une hypothèse du simulateur est que tous les décès doivent passer par une hospitalisation. Donc, le chemin menant au décès est tout aussi impacté par la probabilité d'hospitalisation que la probabilité de décéder sachant avoir été hospitalisé. Ces deux paramètres jouent alors un rôle dans la détermination du nombre d'agents morts à la fin de l'année après le début de la pandémie.

De plus, le simulateur a été calibré sur les données d'hospitalisations donc il est pertinent de jouer sur cet aspect de la modélisation. Cependant, a été supposé que les institutions gouvernementales ont mis en place les mêmes interventions non-pharmaceutiques : leurs effets ne changent pas d'ampleur ni d'une vague à l'autre ni par rapport à la puissance de la maladie. C'est-à-dire les coefficients  $\beta$  ne sont pas modifiés. De plus, la matrice de contact n'est pas affectée, car elle ne représente que les habitudes d'interactions des agents en temps normal et les coefficients viennent les moduler.

La première utilisation du simulateur *CovDyn* implique de prendre en compte des conditions qui influencent la probabilité de décès.

### 3.3.2 Situation 1 : taux de décès uniformes

Les probabilités de décès à l'issue d'une hospitalisation liés au virus de la Covid-19 ont des valeurs plus importantes chez les personnes des classes d'âge élevé. Comme vu précédemment, la maladie épargne le reste de la population. Or cela pourrait ne pas être vrai pour d'autres maladies. Pour tester ce cas de figure, le taux de décès va être uniformisé sur toutes les classes d'âge, c'est-à-dire une seule probabilité (en valeur) va être appliquée à toutes les classes d'âges. Sachant être rentré à l'hôpital, cela donnera la même « chance » de mourir à l'issue de cette dernière. Une fois que la maladie a créé des complications chez les patients, le virus est autant mortel pour les personnes âgées que pour les enfants, adolescents et jeunes adultes.

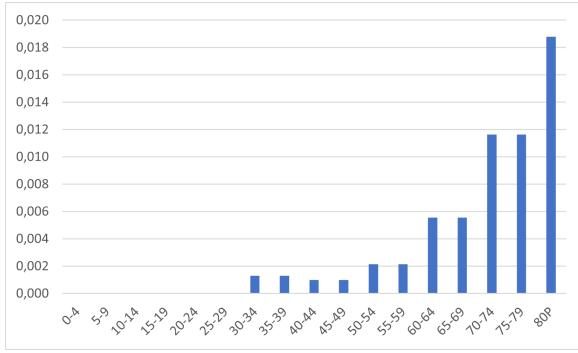


Figure 3.13 – Probabilité de décès sachant être dans le compartiment H2 pour chaque classe d'âge

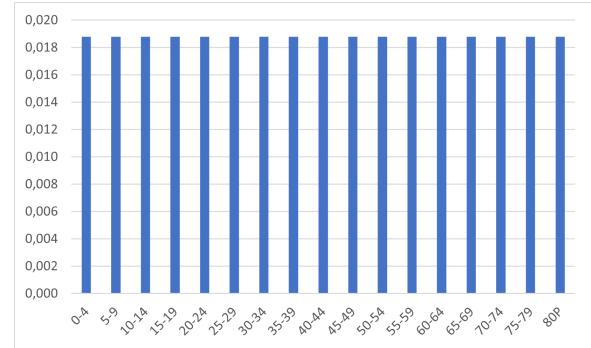


Figure 3.14 – Probabilité uniforme pour tous. Elle vaut 0,019

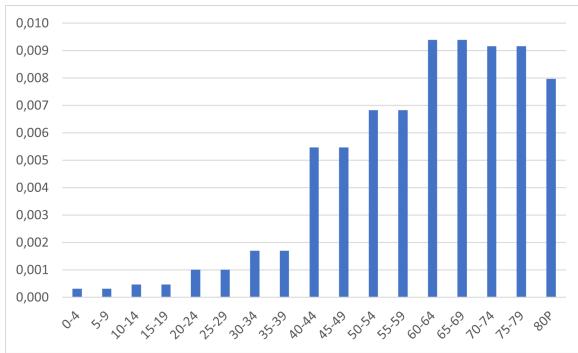


Figure 3.15 – Probabilité de décès sachant être dans le compartiment H3 pour chaque classe d'âge

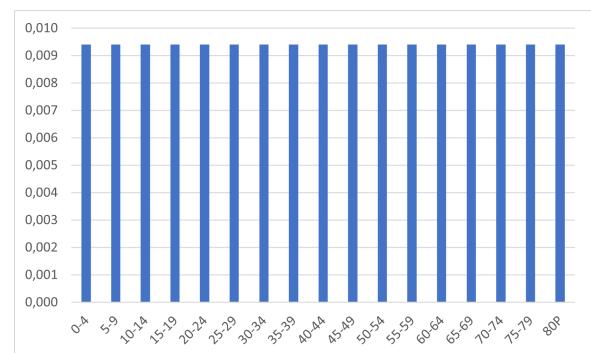


Figure 3.16 – Probabilité uniforme pour tous. Elle vaut 0,0094

Cette modification concerne deux probabilités:  $p_{D|H2}$  et  $p_{D|H3}$ , la première étant la probabilité de décéder après une hospitalisation normale tandis que la deuxième est celle de ceux qui sont passés par des soins intensifs. Les valeurs les plus élevées ont été choisies pour chacune de ces probabilités, respectivement 0,019 et 0,0094, et elles ont été assignées à toutes les classes d'âge sans distinction. Est donc retenu une probabilité de décès supérieure si l'agent ne va pas en réanimation que s'il y va. Cela découle du fait qu'une réanimation implique une intervention médicale immédiate et intensive. Cette réponse rapide et continue peut augmenter les chances de stabiliser l'état du patient. Surtout que la Covid-19 était un virus inconnu et les médecins ne savaient pas comment le traiter directement.

Comme prévu, la courbe d'évolution du nombre journalier d'hospitalisations reste inchangée. Les

probabilités modifiées n'interviennent qu'après le passage à l'hôpital et n'affectent donc pas le nombre d'entrées dans le compartiment  $H1$ .

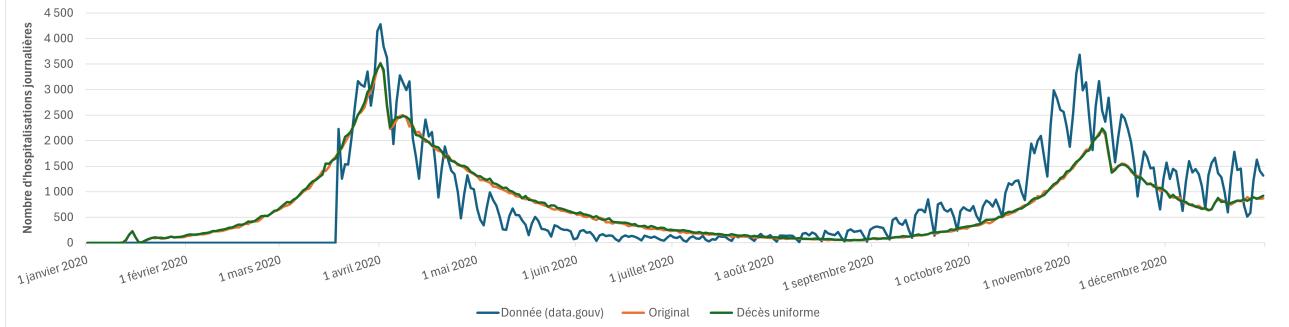


Figure 3.17 – Évolution du nombre d'hospitalisations pratiquement identique à celle du scénario central

Et contrairement à la courbe des hospitalisations, celle des décès cumulés se transforme : elle domine complètement le scénario central, que ce soit la donnée publiée ou celle issue des simulations. Cela est attendu car le nombre de décès ne devrait qu'augmenter avec des probabilités de plus en plus grandes de décès. En moyenne, une telle pandémie génère 15 000 décès supplémentaire au scénario central, faisant alors atteindre le nombre de décès cumulé à 65 000.

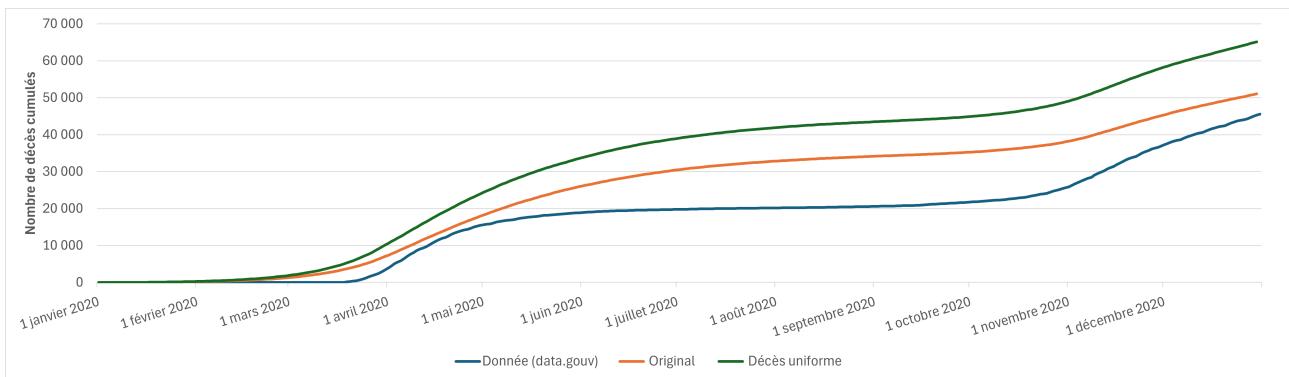


Figure 3.18 – Évolution majorant le scénario central - 65 000 au total

Et qui dit plus de décès dit une plus grande probabilité que l'un des assurés en fasse partie ! Cela engendre plus de simulations avec au moins un sinistre, donnant de plus en plus de poids à la queue de distribution des sinistres. De plus, tous les âges sont considérés de la même manière et donc un jeune adulte a la même probabilité de décéder qu'une personne âgée. Cela induit des versements de prestations beaucoup plus importants étant donné que les valeurs des montants sont largement supérieures à des âges plus jeunes.

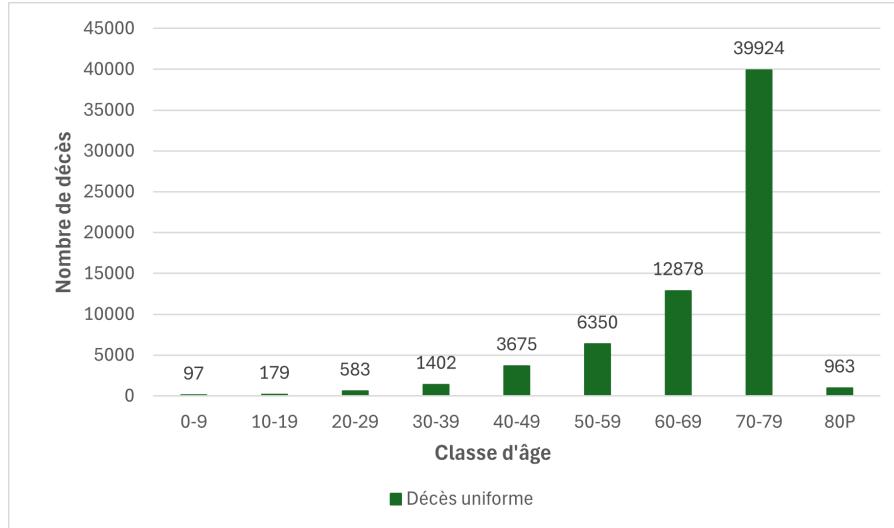


Figure 3.19 – Répartition des décès - Décès uniformes

Dans ce cas, la VaR à 99,5% vaut **2 667 912,28 euros**, plus que triplant la VaR du scénario central et correspond à une hausse de 223,4% par rapport au scénario central. L'augmentation considérable du nombre de décès\* dans chaque classe d'âge explique clairement cette inflation de la valeur de la VaR.

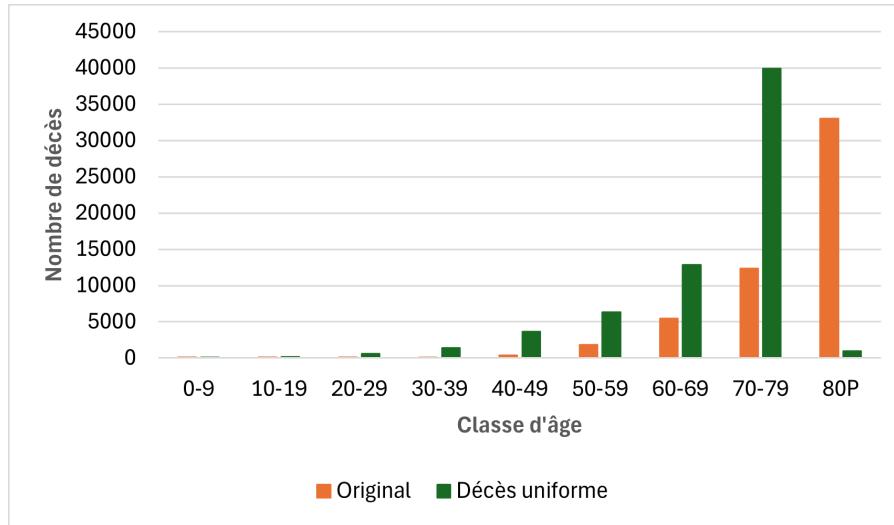


Figure 3.20 – Comparaison des répartitions décès - Original vs Décès uniformes

### 3.3.3 Situation 2 : taux d'hospitalisations uniformes

Les taux d'hospitalisation ne sont pas les mêmes pour les différentes classes d'âge. En effet, les personnes âgées sont les plus vulnérables et présentent donc la probabilité de passage dans le compartiment  $H1$  bien supérieure aux groupes restants.

\*. Il y a une chute importante dans le nombre de morts chez les 80P. Cela s'explique du fait que les 80P n'interagissent que très peu avec les autres classes d'âge excepté les 75 – 79. Ces derniers décèdent en très grand nombre et donc il en existe moins pour transmettre la maladie au groupe des 80P.

Dans cette situation numéro 2 va être considéré que tous les agents, une fois présentant des symptômes, vont rentrer dans l'état  $H1$  avec une même probabilité. La probabilité associée sera celle du groupe des 80P, et vaut 0,8767. Cela donne une version de la Covid-19 qui serait plus dangereuse au niveau symptomatique et qui requiert un suivi poussé par des agents de santé. Cette distribution crée l'aspect létal et devient prononcé chez toute la population contrairement au scénario central. À noter que les probabilités de décès sont remises à leurs valeurs initiales.

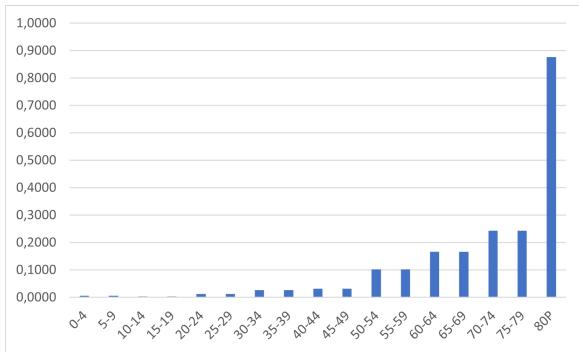


Figure 3.21 – Probabilité d'hospitalisation pour chaque classe d'âge

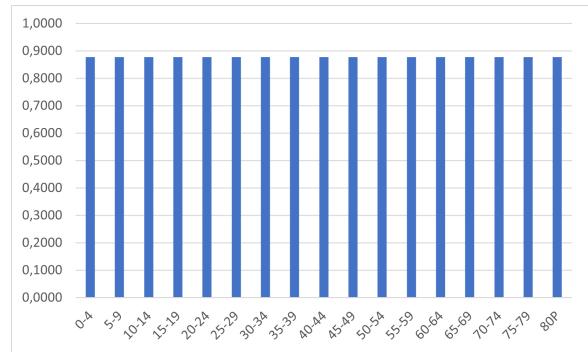


Figure 3.22 – Probabilité uniforme pour tous. Elle vaut 0,8767

Il y a un embrasement du nombre journalier d'hospitalisations durant la première vague, atteignant un pic à 4 108 hospitalisations journalières à son apogée, à savoir le 31 mars 2020. Cette forte croissance est suivie d'une décroissance exponentielle, plus rapide que celle du scénario central, et qui ne sera pas, elle, suivie de la deuxième vague. Cela s'explique par l'hypothèse d'isolement lors du passage dans les compartiments relatifs à l'hospitalisation : augmenter la probabilité  $p_H$  fait diminuer à son tour les probabilités  $p_{I_{ps}}$  et  $p_{D_{ps}}$ . Donc, une plus grande partie des personnes symptomatiques se retrouvent dans les compartiments représentant une hospitalisation et donc sont isolées des autres agents. En d'autres termes, les agents hospitalisés ne peuvent pas être des contacts des autres agents et ne peuvent pas leur transmettre la maladie. Cela réduit énormément le nombre de contacts et donc la probabilité de transmettre le virus et de le propager. Relancer une deuxième vague devient très peu probable et c'est ce qui est présenté par le graphique.

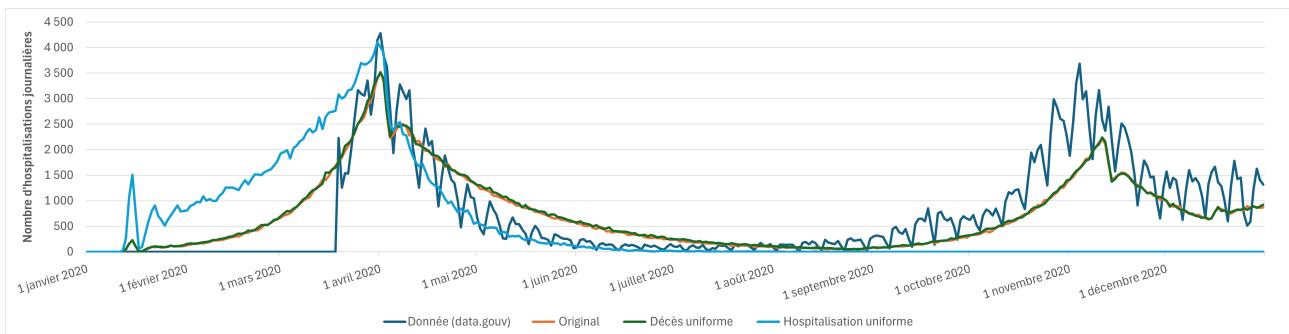


Figure 3.23 – Embrasement de la première vague et extinction de la deuxième

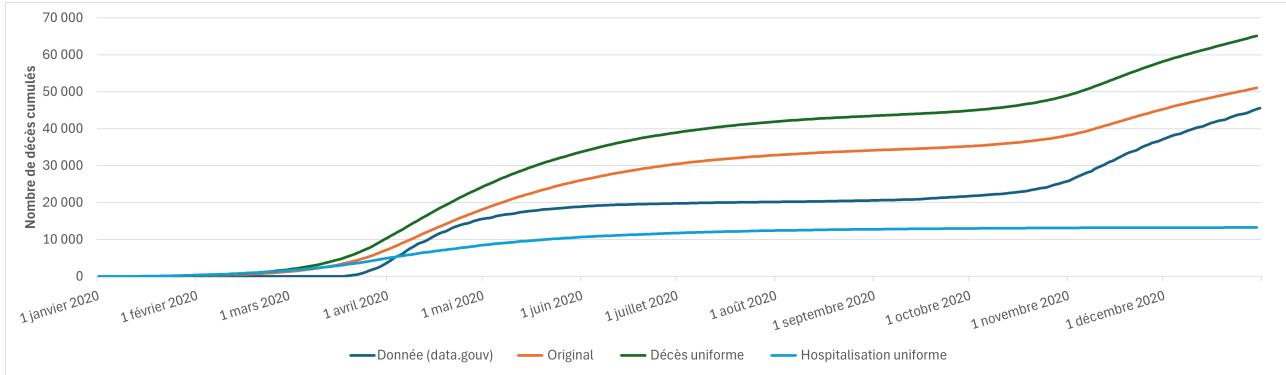


Figure 3.24 – Croissance faible puis plateau du nombre des décès cumulés

Cela peut aussi être perçu grâce au graphique des décès cumulés : la croissance ralentie de plus en plus jusqu'à faire apparaître un plateau, révélant que le nombre de décès n'augmente plus significativement.

Ce scénario résulte donc en un nombre de décès moindre que les deux autres. En effet, moins de personnes sont hospitalisées au total et les probabilités de décès restent biaisées et influencées par l'âge. Le quartile à 99,5% vaut **449 510,52 euros**, ce qui correspond à une baisse de 45,5% par rapport au scénario central.

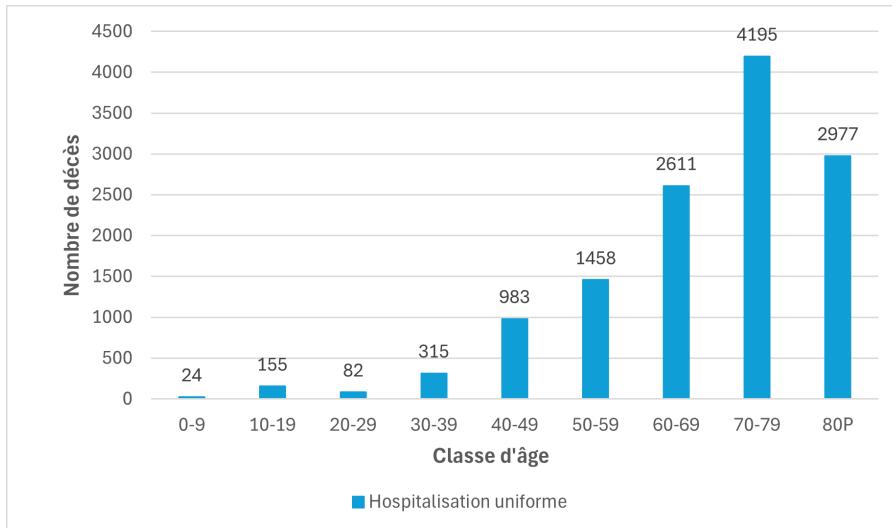


Figure 3.25 – Répartition des décès - Hospitalisations uniformes

Cependant, ce scénario dévoile une limite du modèle utilisé qui arrive en deux temps: premièrement, l'hypothèse d'isolement devient absurde car un tel événement saturerait non seulement les unités de soins intensifs mais aussi les hospitalisations normales. Cela conduirait donc à un débordement du corps médical. Tout le monde ne pourra pas bénéficier de ces soins et donc ne se retrouveraient pas à l'hôpital où ils sont supposés isolés. Deuxièmement, arriver à une telle envergure de la pandémie serait accompagné de l'effondrement du système de santé : ce dernier n'est pas prêt à accueillir un tel nombre de patients. Cette situation atteint alors une des limites du modèle.

Il est tentant de penser que le deuxième confinement en cours paralysent la relance d'une deuxième vague, laissant un doute sur leur efficacité s'il s'avère qu'il n'y a pas effectivement de seconde vague.

Ils ont été retirés afin d'évaluer s'ils influencent ou non le résultat final. Finalement, a été remarqué que les confinements n'ont pas d'effet sur le lancement de la deuxième vague et que les nombreuses hospitalisations sont l'origine de l'écrasement de la courbe et donc des transmissions.

### 3.3.4 Situation 3 : taux d'hospitalisations uniformes et taux de décès uniformes

Cette situation numéro 3 est une combinaison des deux derniers. Les paramètres ont été donc uniformisés pour le décès et l'hospitalisation.

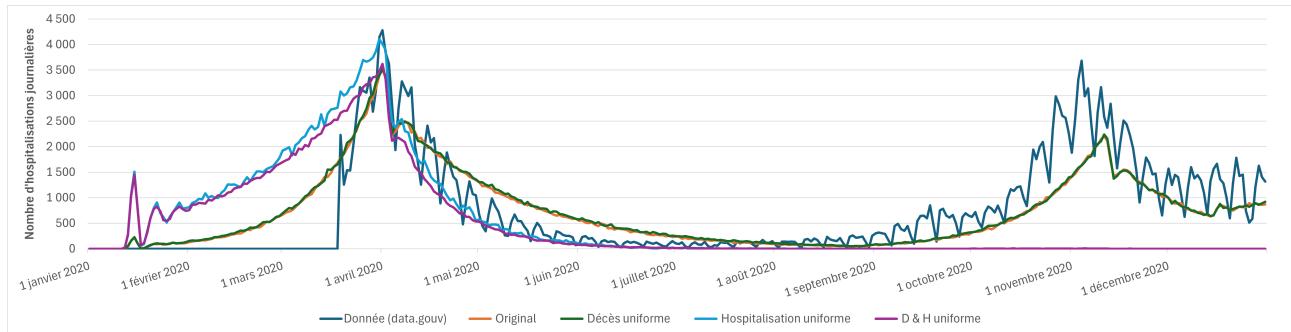


Figure 3.26 – Croissance plus rapide mais atteint quand même un plateau

Un comportement commun est détecté entre ce scénario et le précédent : les hospitalisations journalières croient assez rapidement et ne se voient pas recroître pour la deuxième. Au contraire, l'évolution du nombre de décès cumulés change un peu mais garde la même forme que celle du scénario avec les taux d'hospitalisation uniformes. Le graphique 3.21 expose une croissance plus rapide au début pour atteindre un plateau plus élevé. Cela est dû au fait que à l'entrée en hospitalisation, les agents ont une plus grande probabilité de décéder. À rappeler que le plateau se forme à cause de l'hypothèse d'isolement qui devient fragile.

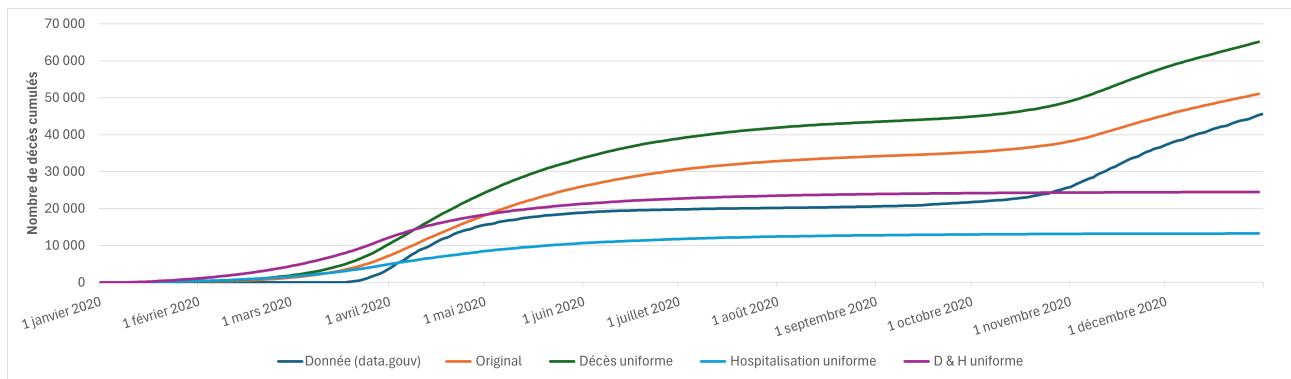


Figure 3.27 – Croissance faible puis plateau du nombre des décès cumulés

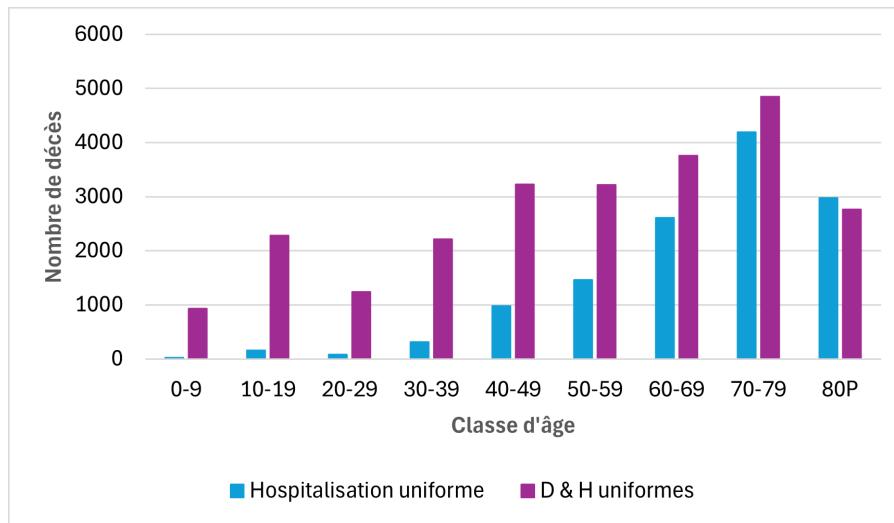
Et étrangement, c'est le scénario qui résulte avec la plus grande VaR, atteignant **2 915 262,35 euros**. Cela correspond à une hausse de 253,3% par rapport au scénario central. Comment expliquer une telle augmentation ?

En effet, régler le taux d'hospitalisation à un même niveau que celui des personnes âgées pour tous force les « bonnes » personnes aux hôpitaux, i.e. les agents appartenant aux classes d'âges entre 20

Scénario	VaR (en euro)
Central	825 059,8
Décès uniforme	2 667 912,28
Hospitalisation uniforme	449 510,52
Décès et hospitalisations uniformes	2 915 262,35

Table 3.1 – Valeur de la *Value-at-Risk* en fonction du scénario mis en place

et 65 ans. Couplant cela avec une probabilité uniforme et élevée de décès, ces agents auront une plus grande probabilité de décéder. Même restreinte à une seule vague, la pandémie touche beaucoup plus les classes d'âges qui sont en sur-représentation dans le portefeuille et donc induit un nombre supérieur de sinistres que dans les autres scénarios. Donc ce sont les groupes ayant le plus de capitaux sous risque sont nettement plus touchés que le reste de la population.

Figure 3.28 – Comparaison des répartitions décès - Hospitalisations uniformes *vs* D & H uniformes

Ces différents cas commencent à donner une idée des conséquences induites par le changement de tels paramètres mais un assureur ne peut pas prédire l'évolution de ces paramètres. Un aléa doit être instauré, même en ébauche, pour commencer à construire une opinion sur ce genre de simulateur afin de calculer les SCR (ou au moins ici la VaR).

### 3.4 Covid aléatoire

Tous les paramètres utilisés dans le modèle *CovDyn* sont prédéfinis dans un fichier *input* avant de lancer une simulation. En particulier, dans le cadre de ce mémoire, les paramètres concernant la Covid-19 sont statiques durant toute la durée de la simulation et représentent ce qui a pu être observé durant la pandémie en France. Or, cela prétend donner beaucoup de confiance en leur validité et ne permet pas d'avoir de la volatilité concernant leurs valeurs. En effet, il n'y a pas d'intervalles de confiance sur les paramètres testés un peu plus haut. Cette partie essaye donc d'ajouter un certain aléa autour de la valeur de quelques paramètres afin de capter un plus grand nombre de réalisations de la pandémie de la Covid-19.

Les paramètres concernés ici pour pouvoir générer des Covid dits « aléatoires » sont :  $p_A$ ,  $p_H$ ,  $p_{ICU|H1}$ ,  $p_{D|H2}$  et  $p_{D|H3}$ . Les probabilités  $p_{H2|H1}$ ,  $p_{R|H2}$ ,  $p_{R|H3}$ ,  $p_{I_{ps}}$  et  $p_{I_{ms}}$  sont déduites (étant les probabilités des événements complémentaires des 5 premiers).

Les chemins qui partent du compartiment des susceptibles et qui mènent au compartiment décès sont en partie déterministes et en partie aléatoire. La partie aléatoire commence à partir du compartiment  $I_{ss}$  et où les probabilités  $p_H$  et  $p_A$  entrent en jeu. Suite à l'hospitalisation, il va falloir déterminer la probabilité de se retrouver dans le compartiment décès et donc inclut les taux de passages  $p_{ICU|H1}$ ,  $p_{D|H2}$  et  $p_{D|H3}$ . C'est la raison pour laquelle ces paramètres ont été choisis dans le cadre de ce scénario.

Les intervalles des valeurs pouvant être prises par ces paramètres vont être construits de la façon suivante : une fenêtre est ouverte autour de la valeur originale du paramètre. En ajoutant et retranchant 20% à cette valeur originale, les bornes de l'intervalle sont déterminées. La valeur d'ouverture, le 20%, est arbitraire. Cet intervalle aura alors la forme  $[p_{Comp} \times 0,8; p_{Comp} \times 1,2]$ , où  $p_{Comp}$  est la probabilité du compartiment concerné.

Ensuite, pour chaque simulation, un tirage de la nouvelle valeur du paramètre doit être réalisé. Deux méthodes ont alors été pensées afin de réaliser ce tirage : un tirage gaussien ou un tirage uniforme.

1. L'idée du tirage gaussien de la nouvelle valeur est de toujours donner le plus de poids à la valeur initiale tout en permettant une certaine flexibilité autour. Une loi gaussienne tronquée sur l'intervalle va alors être utilisée.
2. Le tirage uniforme, lui, amène de la prudence. En effet, la loi uniforme donne le même poids à toutes les valeurs contenues l'intervalle. Les extrêmes ne sont pas négligés et ont autant d'importance que la valeur initiale.

La méthode retenue dans ce mémoire est celle du tirage uniforme. Un des avantages qu'elle possède est qu'elle est plus simple à mettre en place. De plus, ne sachant pas comment se comportent les distributions des valeurs que peuvent effectivement prendre les paramètres, il est plus prudent de prendre la loi uniforme afin de capter les événements extrêmes.

Les résultats d'évolution d'hospitalisations et de décès cumulés sont, en moyenne, les mêmes que ceux du scénario central.

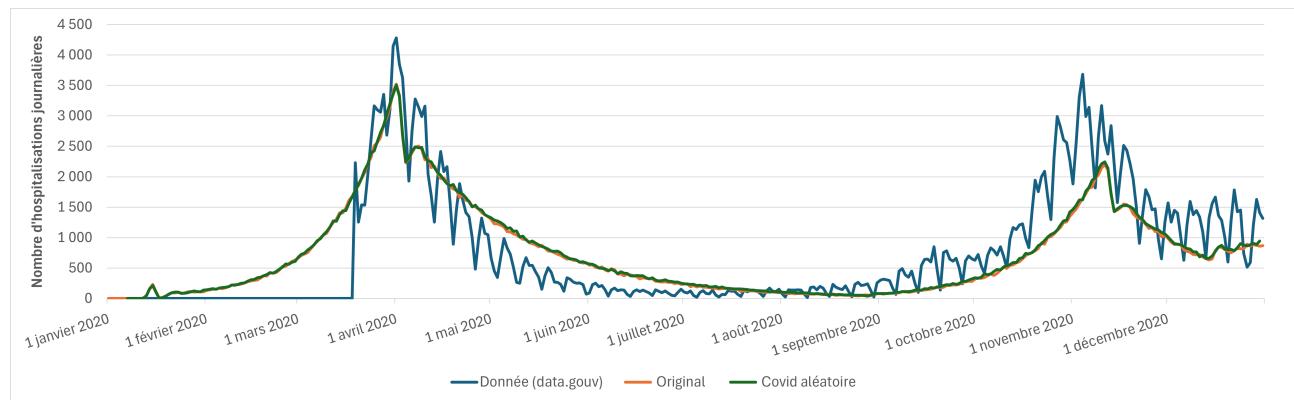


Figure 3.29 – Comportement en moyenne égal à celui du Covid-19 original en termes d'hospitalisations journalières

Cependant, le quantile devient légèrement plus élevé et atteint **941 614,1 euros**. C'est à une hausse de 14,1% par rapport au scénario central. La valeur reste proche de celle du scénario central et

cela est attendu : les paramètres tirés aléatoirement devraient, en moyenne, retomber sur leur valeur d'origine. À noter que 600 simulations reste assez faible pour être très catégorique sur ce constat.

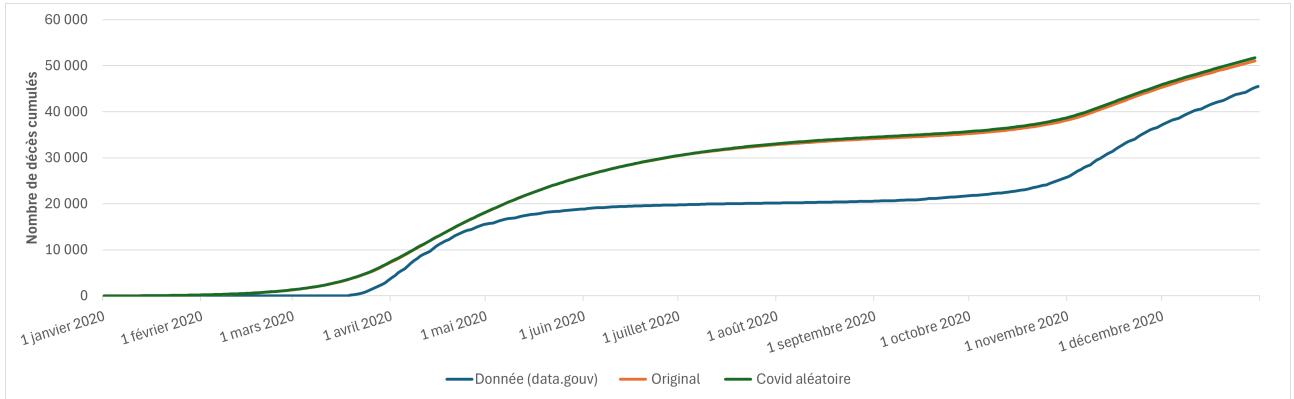


Figure 3.30 – Évolution identique, en moyenne, des décès cumulés

Cependant, cette hausse de la valeur de la VaR peut être expliquée par le fait que les tirages aléatoires apportent de la volatilité, ce qui manquait en quelque sorte aux scénarios précédemment exposés. En effet, cette volatilité permet au modèle d'atteindre une multitude de scénarios futurs et d'obtenir une meilleure vue d'ensemble sur la distribution des prestations décès. Elle permet d'exercer une évaluation du comportement du quantile et comment les conditions peuvent affecter le portefeuille.

Le simulateur est maintenant capable d'amener une couche de volatilité supplémentaire.

### 3.5 Retardement des confinements

Jusqu'ici, les principales sensibilités réalisées ont été opérées sur la base de paramètres épidémiologiques tout en gardant le même « calendrier » qui a été tenu concernant les confinements. Pour rappel, les mêmes dates de confinement et de déconfinements que celles qui ont été implémentées en France sont utilisées dans le cadre des scénarios précédents.

Relevant d'un aspect politique relatif à l'ensemble des restrictions des contacts humains et des déplacements, il est difficile de considérer quel décalage de ces dates est plus pertinent à tester que d'autres. Le choix qui a motivé le confinement du 18 mars est basé sur la consultation d'avis d'experts en santé publique, des responsables gouvernementaux, par des décisions prises par les autorités compétentes ainsi que l'analyse des alternatives en vue de l'urgence sanitaire mondiale. Cependant, il faudrait regarder l'impact d'un tel décalage sur la valeur du quantile considéré. Le confinement a voulu freiner le nombre de cas et donc le nombre de décès qui en découlent. Reporter cette date sous-entends augmenter le nombre d'hospitalisations journalières mais aussi augmenter le nombre de décès par simulation, générant alors un nombre de sinistre supérieur que dans le scénario central.

Dans cette partie, les paramètres liés à l'hospitalisation et au décès des agents seront réinitialisés comme dans le scénario central et ne traiteront pas d'aléa autour de leurs valeurs (comme dans la partie précédente). 3 décalages dans le temps seront considérés : un décalage de 5 jours, un décalage de 10 jours ainsi qu'un décalage de 15 jours. C'est-à-dire, au lieu de commencer le 18 mars, le premier confinement sera déclenché le 23 mars, 28 mars ou le 2 avril suivant la simulation.

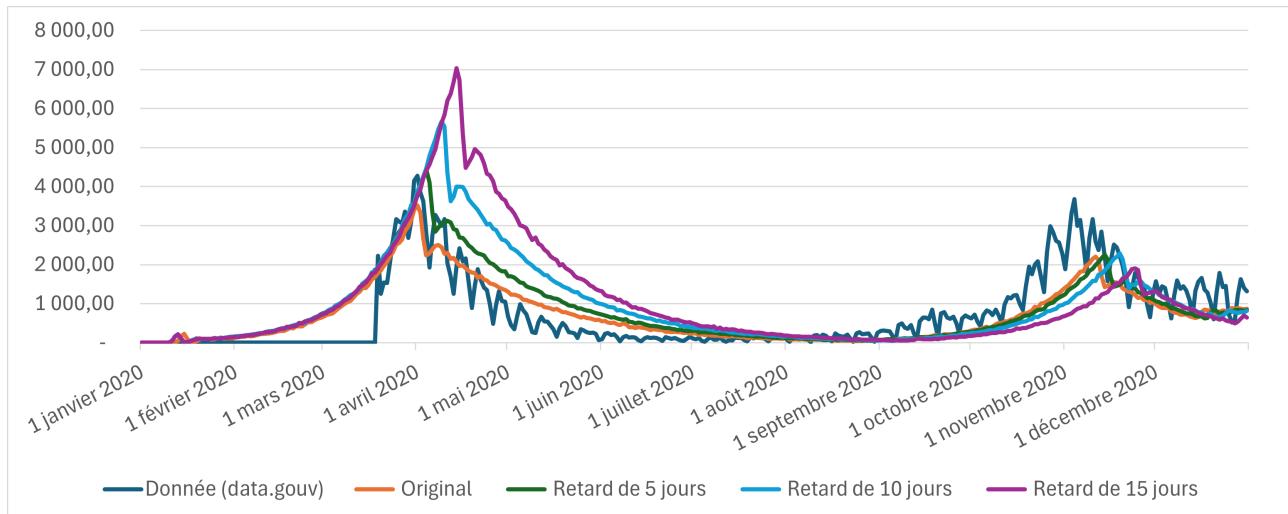


Figure 3.31 – Évolutions aux mêmes motifs mais décalé vers la droite et avec des pics plus importants

Comme anticipé, la représentation graphique du nombre d'hospitalisations journalières se déplace vers la droite, atteignant un sommet encore plus marqué durant la première vague, à mesure que la date de la première période de confinement est retardée. Cela permet de garder la vitesse de transmission du virus et donc de continuer à faire rentrer beaucoup plus d'agents dans les compartiments hospitaliers. Les impacts sur la VaR sont exprimés ci-dessous. Elle est d'autant plus élevé que la date du confinement est décalée dans le temps.

Décalage	VaR (en euro)	Evolution (en %)
0 jour (Central)	825 059,8	
5 jours	1 109 803,16	+ 34
10 jours	1 139 935,23	+ 38
15 jours	2 111 002,41	+ 155

Table 3.2 – Valeur de la *Value-at-Risk* en fonction du décalage de la date du premier confinement

Le comportement de la pandémie semble, a priori, être le même que dans les scénarios aux parties 3.3.3 et 3.3.4. Quelle est l'explication derrière la résurgence de la pandémie entraînant une deuxième vague ? Et qu'est-ce qui fait que les VaR correspondantes sont inférieures à celles présentées dans les parties mentionnées ?

Ce qui diffère ici, c'est le fait que la majorité des agents infectés sont encore libres de transmettre la maladie et ne sont pas contraints de rentrer dans le bloc hospitalier (ce qui peut étouffer les transmissions du virus comme vu dans les parties 3.3.3 et 3.3.4). Autrement dit, un plus grand nombre d'agents sont effectivement infectés et donc la proportion qui sera hospitalisée va être proportionnelle à ce nombre vu la valeur du paramètre par classe d'âge (alors que dans les parties 3.3.3 et 3.3.4, la majorité des infectés se retrouveront dans le bloc hospitalier).

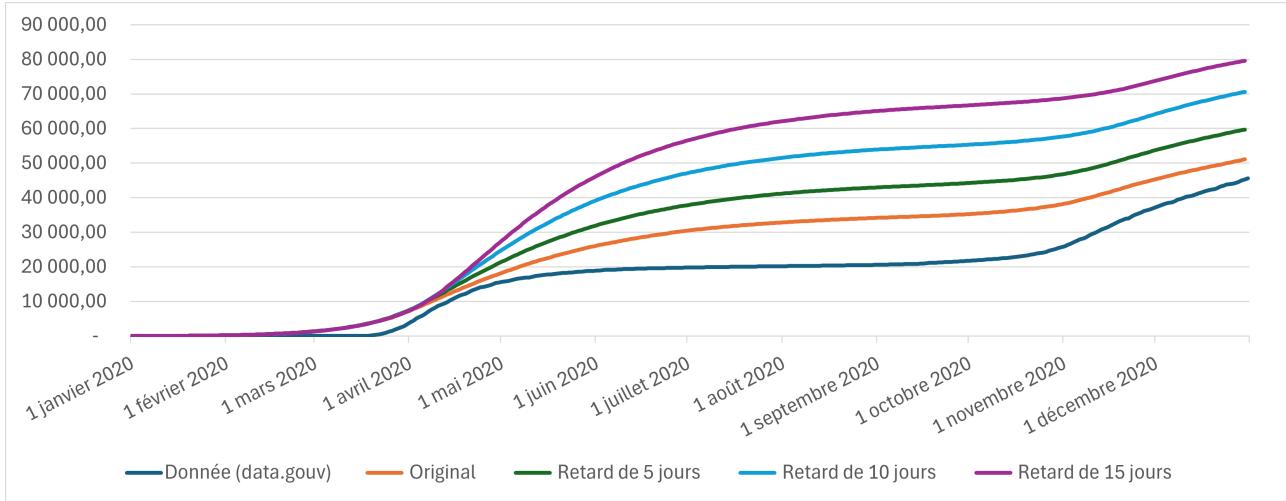


Figure 3.32 – Evolution des décès cumulés en fonction du décalage

La classe d'âge la plus touchées ici est celle des 80P qui a la plus haute probabilité d'hospitalisation et donc va être la plus vulnérable (comme dans le scénario central). Bien que le nombre de décès soit considérablement plus élevé, la valeur de la VaR n'excède pas celle des scénarios préalablement évoqués, car la majorité des personnes concernées seraient des individus âgés, ce qui permet de préserver les tranches d'âge où se concentre le plus grand volume de capitaux à risque.

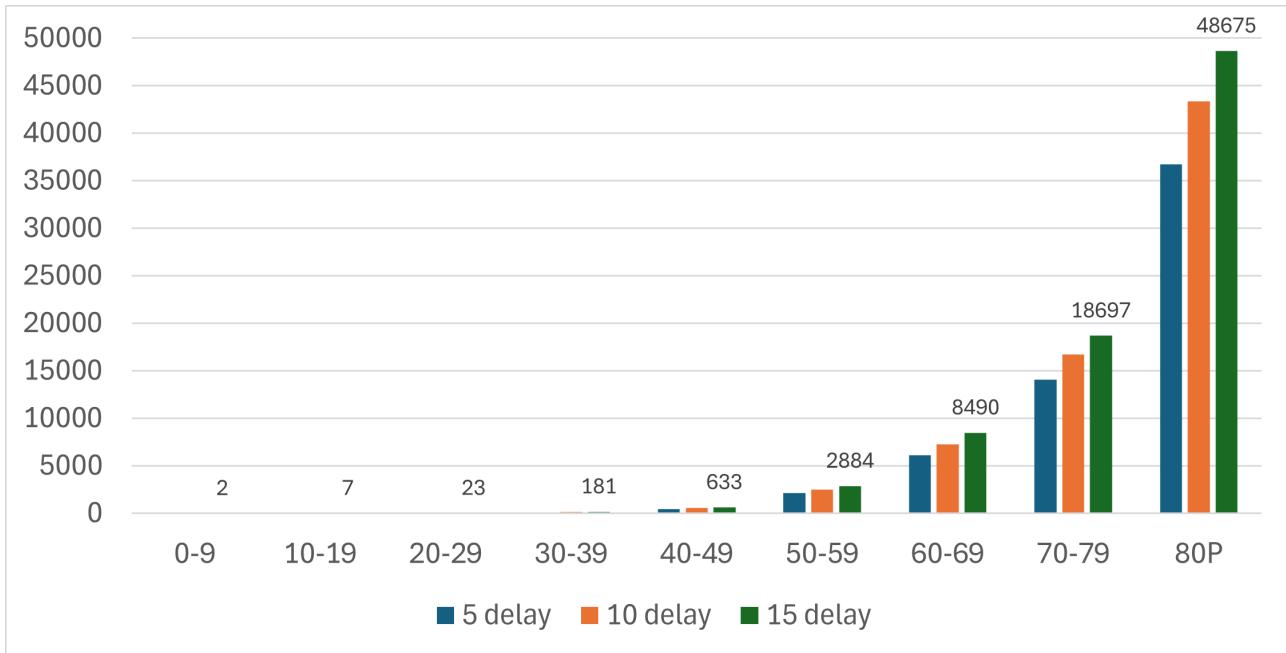


Figure 3.33 – Distribution des décès suivant le degré de décalage - Même forme de distribution que le scénario original

Le scénario central est supposé représenter le cas où l'assureur est solvable, en termes de coût uniquement ici. Et dans une vision prudentielle, l'assureur teste sa solvabilité à court terme. Dans un contexte de pandémie, il aurait pu vouloir voir comment des cas plus extrêmes pourraient mettre en péril cette solvabilité.

Effectivement, deux événements ont détérioré la situation à laquelle l'assureur est confronté : une pandémie dont le taux de décès est élevé, accompagné ou non d'un taux d'hospitalisation élevé aussi. La solvabilité peut être d'autant plus détériorée quand les deux cas extrêmes surviennent en même temps. Ce sont donc les scénarios où la probabilité de décès est importante qui devraient amener beaucoup plus de risque à l'assureur. Une probabilité d'hospitalisation de valeur considérable ne semble pas pouvoir aggraver la situation à elle-même, mais quand elle est couplée à cette probabilité de décès, elle peut rajouter du péril.

Il est donc important et nécessaire de prévoir ce genre d'exercice et d'appliquer des chocs sur les paramètres mêmes afin d'étudier l'évolution de la valeur de la VaR.



# Conclusion

La spécificité du risque pandémie dans le cadre d'un portefeuille emprunteur est importante à explorer, notamment en ce qui concerne la modélisation de la distribution des prestations à verser dans ce contexte. La nature de ces contrats, reposant sur le risque décès dans notre étude pour enclencher l'engagement de l'assureur, devrait être questionnée lors de l'avènement d'une pandémie. Ces contrats exposent la compagnie d'assurance à une surmortalité potentielle et à des versements de prestations précoces. La nécessité de modéliser ce risque est importante car elle joue un rôle essentiel dans la capacité des assureurs à respecter leurs engagements futurs tout en assurant leur solvabilité.

Ce mémoire a montré la pertinence d'utiliser les modèles agents : ils sont plus flexibles et beaucoup plus réalistes grâce aux différentes composantes qui les forment. En contrepartie d'un temps de calcul plus élevé, l'évolution de la pandémie peut être retracée plus précisément sur plusieurs niveaux que celle fournie par des modèles déterministes. En se basant sur des interactions entre agents, ce type de modèle peut révéler des tendances et permet, comme l'a été présenté, d'intégrer une sous-population d'assurés et de les suivre durant une année de pandémie.

La pandémie de la Covid-19 a été reprise dans le cadre du simulateur *CovDyn* afin d'avoir un référentiel de base. Cela a permis, à l'issue de nombreuses simulations, de déterminer une distribution de prestations décès et ainsi de calculer une *Value-at-Risk* à 99,5%.

La *Value-at-Risk* de la pandémie de la Covid-19 est calculée et est confrontée à celles de scénarios de « Covid modifiés » où les taux d'hospitalisation et de décès sont altérés. Certains exemples mettent en lumière des contextes qui agravaient le risque de l'assureur et pourraient mettre en péril sa solvabilité. Les modèles agents permettent d'analyser et de comprendre le raisonnement qui amène à de telles situations. Les résultats sont prometteurs dans le sens où il devient intuitif d'utiliser les modèles agent et d'injecter les caractéristiques propres de chaque portefeuille dans la population, et de suivre le risque ainsi que le choquer.

De plus, les tirages uniformes sur les paramètres permettent non seulement de rajouter de la volatilité aux simulations, mais également de garder une certaine prudence sur l'aléa qu'on intègre lors des tirages. Cela est d'autant plus intéressant et affine le calcul des éléments afin de retrouver un SCR.

Cependant, plusieurs hypothèses ont été émises :

1. les décès ne peuvent survenir qu'après une hospitalisation. Les agents en dehors du bloc hospitalier ne peuvent pas se retrouver dans le compartiment décès,
2. les calibrations des paramètres concernant les interventions non pharmaceutiques sont celles de la deuxième vague et donc ne représentent pas tout à fait la première vague. Un meilleur ajustement peut être réalisé à l'issue d'un recalibrage de ces coefficients,
3. certains des scénarios peuvent s'éloigner de la réalité mais proposent quand même un schéma de pensée pour rajouter des composantes ou en affiner des existantes. Par exemple, une addition

serait d'intégrer un nombre fini d'agents qui pourraient être dans le bloc hospitalier en même temps. Un autre exemple serait d'assouplir l'hypothèse d'isolation totale dans ce bloc-là et de se questionner sur la validité de cette hypothèse.

Le mémoire montre la viabilité de ces modèles agents et peut inspirer à sortir du cas spécifique de la pandémie de la Covid-19. En effet, il serait pertinent d'aller encore plus loin que de simplement créer un « Covid aléatoire » mais d'arriver à générer une maladie « aléatoire », qui n'existerait pas, avec ses propres paramètres. Ces derniers seraient tirés par rapport à des distributions de maladies déjà recensées et dont sont connus leurs caractéristiques et fonctionnements. De plus, ces modèles agents pourraient être utilisés dans une approche prévoyance dans le cadre des arrêts de travail et des dépenses de santé. Suivre l'agent durant son hospitalisation et le temps de séjour pourrait être intéressant à voir dans un contexte de pandémie afin d'estimer le coût pour le risque santé.

# Bibliography

- [1] Di Domenico et al. “Impact of lockdown on COVID-19 epidemic in Ile-de-France and possible exit strategies.” In: *BMC medicine* 18.240 (2020). DOI: [10.1186/s12916-020-01698-4](https://doi.org/10.1186/s12916-020-01698-4).
- [2] Gauchon et al. “Lessons learnt from the use of compartmental models over the COVID-19 induced lockdown in France”. In: *medRxiv* (2021). DOI: <https://doi.org/10.1101/2021.01.11.21249565>.
- [3] Jingyuan et al. “Impact of Temperature and Relative Humidity on the Transmission of COVID-19: A Modeling Study in China and the United States”. In: (2020). URL: [https://papers.ssrn.com/sol3/Papers.cfm?abstract\\_id=3551767](https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=3551767).
- [4] Sanche et al. “High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2”. In: (2020). URL: [https://wwwnc.cdc.gov/eid/article/26/7/20-0282\\_article](https://wwwnc.cdc.gov/eid/article/26/7/20-0282_article).
- [5] Simon Pageaud et al. “Expected Evolution of COVID-19 Epidemic in France for Several Combinations of Vaccination Strategies and Barrier Measures”. In: *MDPI* 9.12 (2021). DOI: <https://doi.org/10.3390/vaccines9121462>.
- [6] Améli. “Reconnaitre la grippe”. In: (2023). URL: <https://www.ameli.fr/rhone/assure/sante/themes/grippe/reconnaitre-grippe>.
- [7] Our World in Data. “Coronavirus (COVID-19) Deaths”. In: (2024). URL: <https://ourworldindata.org/covid-deaths>.
- [8] Guerric BRAS Elise CALLAC Aurélie TREILHOU. “Covid-19, un an après : synthèse des principaux impacts dans le secteur de l’assurance”. In: (2021). URL: <https://www.pericles-actuarial-blog.com/post/covid-19-un-an-apr%C3%A8s-synth%C3%A8se-des-principaux-impacts-dans-le-secteur-de-l-assurance>.
- [9] EPI-PHARE. “Covid-19 : facteurs de risques d’hospitalisation et de décès à l’hôpital lors de la 2ème vague”. In: (2021). URL: <https://www.epi-phare.fr/rapports-detudes-et-publications/covid-19-facteurs-risques-2/>.
- [10] La Commission Européenne. “Règlement délégué (UE) 2015/35 de la commission du 10 octobre 2014 complétant la directive 2009/138/CE du Parlement européen et du Conseil sur l’accès aux activités de l’assurance et de la réassurance et leur exercice (Solva bilité II).” In: (2015). URL: <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32015R0035>.
- [11] Iris Ganser et al. “Estimating the population effectiveness of interventions against COVID-19 in France: A modelling study”. In: *Epidemics* 46 (2024), p. 100744. ISSN: 1755-4365. DOI: <https://doi.org/10.1016/j.epidem.2024.100744>. URL: <https://www.sciencedirect.com/science/article/pii/S1755436524000057>.
- [12] INSEE. “En quatre vagues, l’épidémie de Covid-19 a causé 116 000 décès et lourdement affecté le système de soins”. In: (2021). URL: <https://www.insee.fr/fr/statistiques/5432509?sommaire=5435421>.

- [13] Stanley MILGRAM Jeffrey TRAVERS. "An Experimental Study of the Small World Problem". In: 32 (1969), pp. 425–443. DOI: <https://doi.org/10.1016/B978-0-12-442450-0.50018-3>.
- [14] M. Jit K. Prem A.R. Cook. "Projecting social contact matrices in 152 countries using contact surveys and demographic data". In: *PLoS Comput Biol* 13.9 (2017). DOI: <https://doi.org/10.1371/journal.pcbi.1005697>.
- [15] Raphaële KARAYAN. "Coronavirus : l'assurance emprunteur relativement épargnée par la crise". In: (2020). URL: <https://www.argusdelassurance.com/assurance-de-personnes/assurance-emprunteur/coronavirus-l-assurance-emprunteur-relativement-epargnee-par-la-crise.163511>.
- [16] OMS. "Déclaration sur la réunion du Comité d'urgence du Règlement sanitaire international (RSI) concernant la flambée de nouveau coronavirus (2019-nCoV)." In: (2020). URL: [https://www.who.int/fr/news/item/23-01-2020-statement-on-the-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/fr/news/item/23-01-2020-statement-on-the-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)).
- [17] Maxime RECOUILLE. "INFOGRAPHIES. Age, sexe... Qui sont les 100 000 morts du Covid-19 en France ?" In: (2021). URL: [https://www.lexpress.fr/sciences-sante/sante/infographies-age-sexe-qui-sont-les-100-000-morts-du-covid-19-en-france\\_2148780.html](https://www.lexpress.fr/sciences-sante/sante/infographies-age-sexe-qui-sont-les-100-000-morts-du-covid-19-en-france_2148780.html).

# Annexes



## Appendix A

# Visualisation du fonctionnement du SIR

Une simulation va faire tourner un modèle dit SIR, cela veut-dire que la population considérée est répartie en trois catégories, aussi appelées compartiments ou états:

1. ceux qui sont susceptibles d'avoir la maladie. Cela est noté par la lettre S;
2. ceux qui sont infectieux, i.e. ceux qui peuvent transmettre la maladie et l'aider à se propager dans la population susceptible. Les infectieux appartiennent au compartiment I;
3. et finalement, ceux qui sont rétablis de leurs infections. Ils sont représentés par l'état R.

Au début est considérée une population de  $N$  agents qui seront en majorité dans le compartiment S avec certains qui se retrouveront dans le compartiment I. Aucun agent ne se retrouve dans le compartiment R durant la phase d'initialisation. À chaque étape de la simulation va être calculé le nombre d'agents passant d'un compartiment à une autre, adjacent à ce dernier.

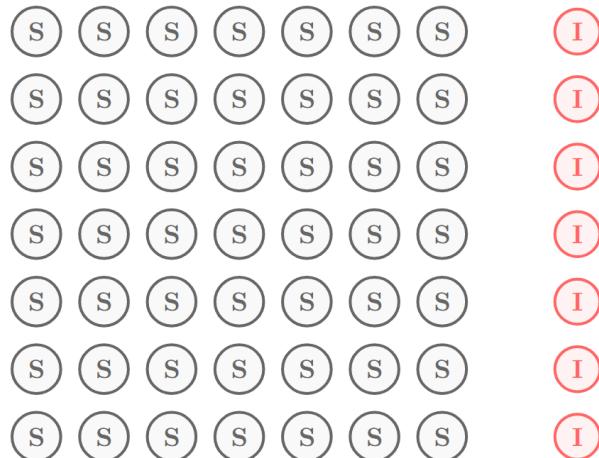


Figure A.1 – Étape 1 : Initialisation de la population

Après l'initialisation, il faut commencer par calculer la proportion de Susceptibles qui vont devenir Infectés. Cela dépend de deux facteurs :

- la probabilité de devenir infecté. C'est un paramètre constant qui est estimé par des épidémiologues, et sera noté  $\beta$ ;
- et la proportion de gens infectés dans la population totale

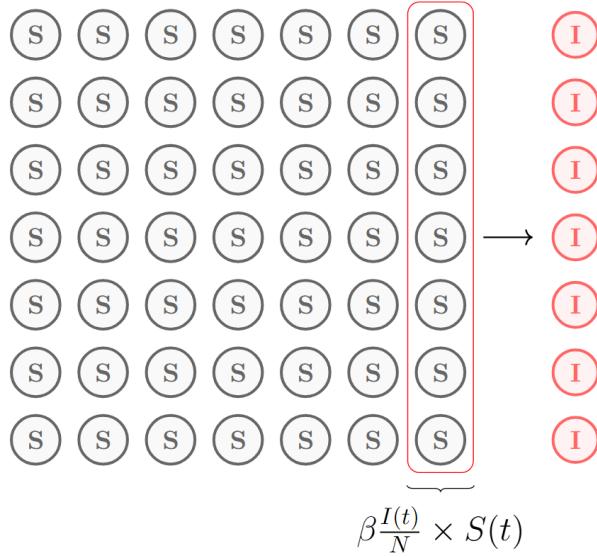


Figure A.2 – Étape 2 : Certaine proportion de Susceptibles passe en Infectés

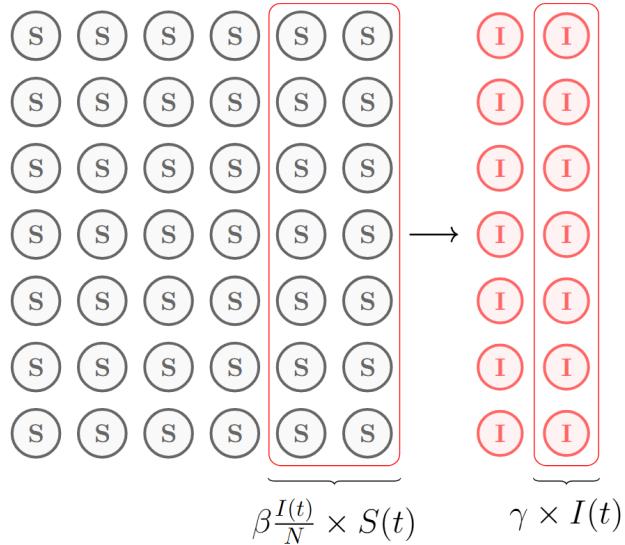


Figure A.3 – Étape 3 : Certains des Infectés sont Rétablis

Après un certain moment, quelques agents dans l'état Infectieux vont se rétablir et ne pourront plus propager la maladie. La proportion d'agents infectieux qui passeront dans l'état Rétablis va dépendre du facteur  $\gamma$ , lui aussi estimé par des spécialistes.

Après cela, le processus est réitéré jusqu'à ce qu'il n'y ait plus d'agents dans l'état Susceptibles.

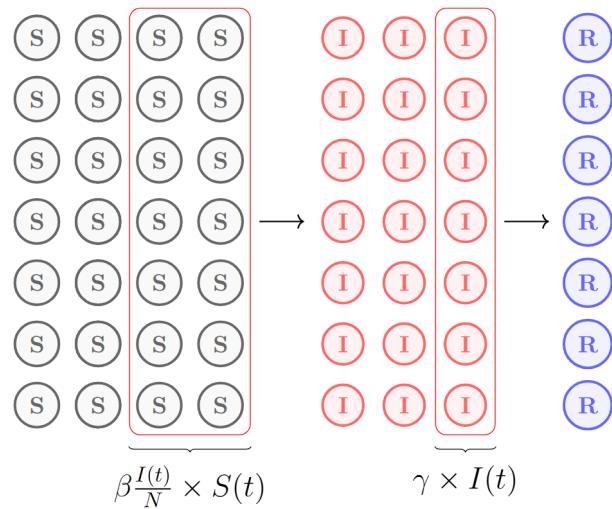


Figure A.4 – Étape 4 : On recommence jusqu'à ce qu'il n'y ait plus d'infectés