

Mémoire présenté le :

**pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA
et l'admission à l'Institut des Actuaires**

Par : Clément LAGARDE

Titre : Proposition d'un produit « garantie maladie redoutée »

Confidentialité : ☒ NON ☐ OUI (Durée : ☐ 1 an ☐ 2 ans)

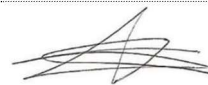
Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

*Membre présents du jury de l'Institut
des Actuaires*

signature

Entreprise :

Nom : SEABIRD CONSEIL



Signature :

Directeur de mémoire en entreprise :

Nom :

Signature :

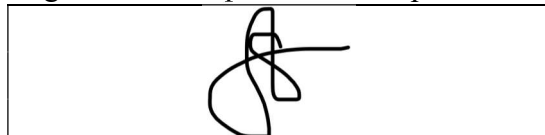
Invité :

Nom :

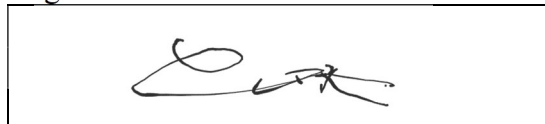
Signature :

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

Signature du responsable entreprise



Signature du candidat



Résumé

Les affections de longue durée touchent de plus en plus de Français chaque année. Selon les données de l'assurance Maladie, plus de 12 millions de personnes sont concernées, soit environ 18% de la population. Parmi ces affections, le cancer occupe une place prépondérante en tant que principale cause de décès prématurés chez les hommes et la deuxième chez les femmes.

Bien que cette maladie soit prise en charge à 100% par l'Assurance Maladie, le quotidien des personnes atteintes reste difficile et des frais non couverts peuvent affecter la qualité de leur traitement. Pour alléger leur quotidien, il existe la garantie "maladie redoutée", qui agit selon trois axes : la prévention, le soutien financier et l'accompagnement.

Au cours des 30 dernières années, le nombre de cas de cancer en France n'a cessé d'augmenter. Il est intéressant de se demander si l'utilisation des données historiques permettrait de développer des modèles prédictifs pour estimer le taux d'incidence du cancer dans le futur en France, et par conséquent, d'évaluer les répercussions sur le calcul de la prime pure de la garantie "maladie redoutée".

La garantie "maladie redoutée" est conçue pour offrir une couverture à long terme, ce qui signifie que toute augmentation du taux d'incidence du cancer influencera directement la rentabilité. Prédire le taux d'incidence permettra d'établir des primes plus précises, assurant ainsi la pérennité financière de la garantie tout en maintenant un soutien adéquat pour les personnes atteintes de cancer.

Mots-clés : Cancer, Prime pure, Incidence du cancer, Garantie "maladie redoutée".

Abstract

Long-term illnesses affect more and more French people every year. According to data from Health Insurance, over 12 million people are affected, which represents about 18% of the population. Among these conditions, cancer plays a predominant role as the leading cause of premature death in men and the second leading cause in women.

Although this illness is covered 100% by Health Insurance, the daily lives of those affected remain challenging, and uncovered costs can impact the quality of their treatment. To ease their daily lives, there is the "feared illness" insurance, which operates on three fronts: prevention, financial support, and guidance.

Over the past 30 years, the number of cancer cases in France has steadily increased. It is worth asking whether the use of historical data could help develop predictive models to estimate the future cancer incidence rate in France, and consequently, assess the impact on the calculation of the pure premium for the "feared illness" insurance.

The "feared illness" insurance is designed to provide long-term coverage, meaning any increase in the cancer incidence rate will directly influence profitability. Predicting the incidence rate will allow for more accurate premiums, thus ensuring the financial sustainability of the insurance while maintaining adequate support for individuals affected by cancer.

Keywords : Cancer, Pure premium, Cancer incidence, Critical illness insurance

Note de synthèse

Le cancer désigne une maladie où certaines cellules de notre corps subissent des mutations, entraînant leur prolifération incontrôlée et potentiellement la formation de ce que l'on appelle une tumeur. Selon une étude réalisée par l'ISOP en 2023, le cancer est la maladie qui préoccupe le plus les Français en matière de santé.

Première cause de décès prématuré en France, le nombre de cas de cancer n'a cessé d'augmenter chaque année au cours des 30 dernières années ¹ Malgré cette progression, le taux de mortalité est en constante diminution depuis 25 ans, grâce aux progrès médicaux. En France, cette maladie est prise en charge à 100 % par l'Assurance Maladie. Cependant, ces dernières années, des associations telles que la Ligue contre le cancer ont signalé que, malgré cette couverture, le parcours de soins reste un poids financier pour de nombreux patients. Entre les dépassements d'honoraires, les participations forfaitaires, les frais de transport, le besoin d'une aide à domicile dû à la fatigue causée par les traitements, ou encore les soins complémentaires non remboursés, le reste à charge peut rapidement devenir important.

Il existe toutefois la "garantie maladie redoutée", qui permet d'allouer un capital défini dans le contrat lors du diagnostic de cette maladie. En France, son développement reste assez faible et peu connu. Les assureurs ne mettent pas nécessairement cette garantie en avant, bien que, dans le contexte actuel, elle pourrait s'avérer intéressante. Aux États-Unis, où le marché est plus développé en raison d'un système de santé plus complexe, des études prévoient une croissance importante de ce type de produit dans le monde et en particulier en Europe.

Le but de notre étude est de proposer un produit d'assurance pour un contrat de type garantie maladie redoutée spécifiquement dédié au cancer. Dans cette étude, nous chercherons à déterminer la meilleure méthode pour prédire le taux d'incidence du cancer en France. Ensuite, nous estimerons le capital optimal pour un assuré en fonction du niveau de confort qu'il souhaite avoir en cas de diagnostic d'un cancer.

La première étape pour développer un tel produit consiste à analyser le taux d'incidence qui correspond au nombre de nouveaux cas dans une population donnée.

1. Vie publique. Cancers : le nombre de nouveaux cas a presque doublé en France [19]

Ce taux est généralement exprimé en « nombre de personnes pour 100 000 personnes par an », permettant ainsi de suivre son évolution dans le temps et de le comparer entre différents groupes d'âge.

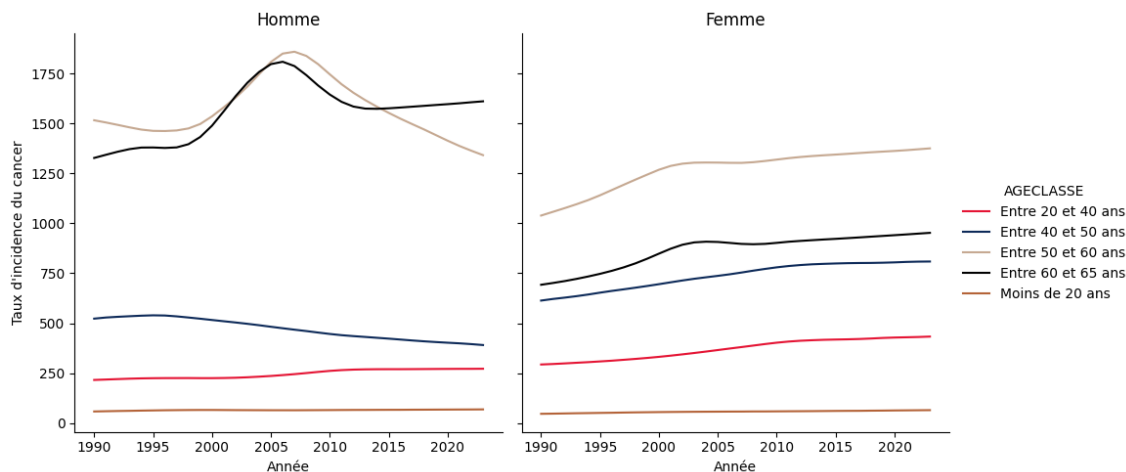


FIGURE 1 – Comparaison de l'incidence selon l'âge et le sexe

Pour nos prévisions, nous avons testé trois méthodes : ARIMA, une méthode classique pour la prédiction de séries temporelles ; la méthode de Lee-Carter, traditionnellement utilisée pour les taux de mortalité ; et une approche plus récente, utilisant les réseaux de neurones, avec la méthode RVFL.

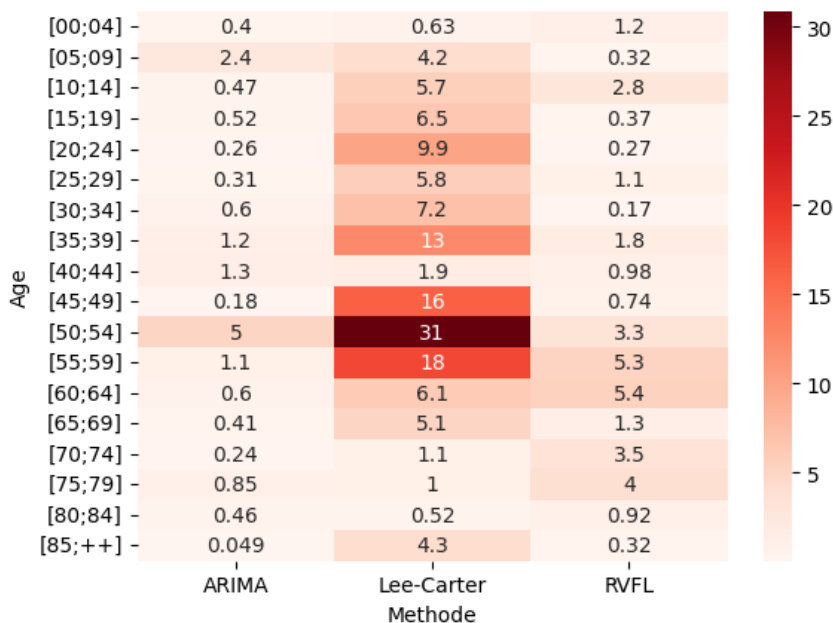


FIGURE 2 – MAE pour chaque âge en fonction des trois méthodes étudiées

Après avoir appliqué nos modèles, le modèle Lee-Carter s'est révélé être le moins performant. Nous avons estimé que cela était dû à la grande campagne de dépistage menée en France en 2005, qui a entraîné un pic dans la courbe du taux d'incidence au fil du temps. Lee-Carter, étant un modèle basé sur la tendance passée, a des difficultés à intégrer des variations dues à des événements ponctuels.

Bien que la méthode RVFL ait donné de meilleurs résultats que Lee-Carter, les prédictions n'étaient pas convaincantes. Cette méthode, qui a montré son efficacité dans le domaine des prévisions temporelles, n'a pas bien fonctionné dans notre cas en raison du manque de données : nous ne disposons que de 30 points par catégorie d'âge pour entraîner nos modèles.

Finalement, le modèle ARIMA a été le plus performant avec les données disponibles. Malgré certains points faibles et le pic d'incidence de 2005, ARIMA nous a permis d'obtenir des erreurs nettement plus faibles que celles des deux autres modèles.

Une fois l'incidence prédite, la deuxième étape a été de déterminer le capital nécessaire pour aider au mieux l'assuré. Nous avons recensé les principales dépenses annexes des personnes atteintes de cancer, puis estimé le coût total. Trois niveaux de capital ont été proposés, correspondant à différents niveaux de qualité de couverture sur une période donnée.

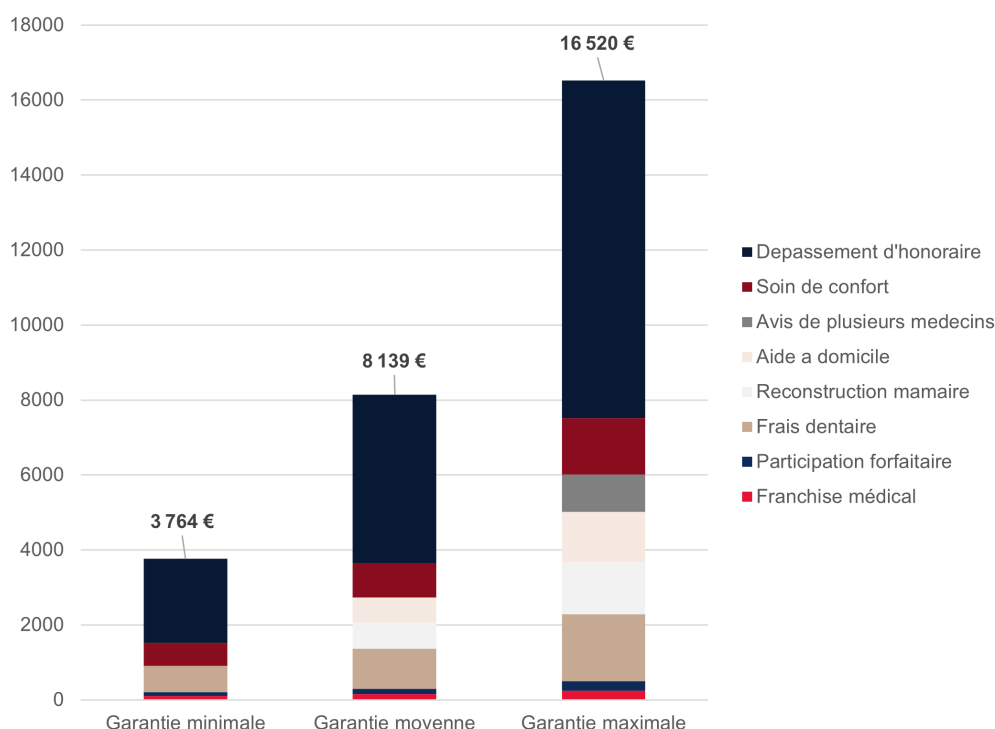


FIGURE 3 – Montant des capitaux assurés selon le type de garantie

Le premier capital vise à offrir une couverture minimale sur 2 ans, permettant d'accéder aux soins complémentaires les plus basiques.

Le deuxième capital offre une meilleure qualité de soins complémentaires, avec une couverture prévue pour une durée de 3 ans.

Le troisième niveau permet d'accéder à tous les soins complémentaires pendant une durée de 5 ans.

Une fois le taux d'incidence prédit sur un horizon de 30 ans et le capital estimé, nous avons calculé la prime pure.

Notre produit final couvre le risque de cancer pour les assurés à partir de leur souscription jusqu'à 75 ans. La couverture prend fin le jour des 75 ans de l'assuré. Le contrat devient caduc si l'assuré le résilie, s'il meurt ou s'il est diagnostiqué avec un cancer (fin du contrat après le versement de la prime assurée).

Après avoir formulé ces différentes hypothèses, nous avons obtenu les tarifs suivants :

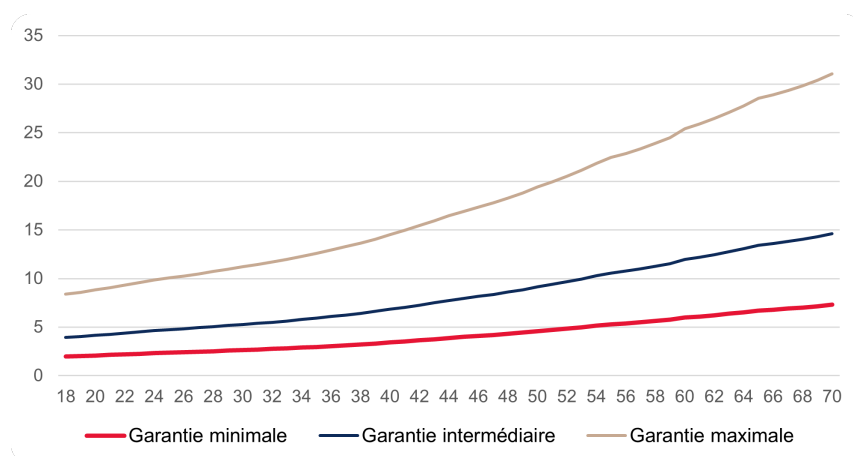


FIGURE 4 – Prime pure en fonction de l'âge et du capital assuré

Notre modèle vise à estimer les primes pures dans le contexte du cancer en France, mais il comporte des limites dues à la disponibilité et à la qualité des données :

- **Facteurs externes** : il est difficile de trouver et d'intégrer les différentes variables exogènes dans notre modèle, surtout sur un horizon de 30 ans.
- **Limites des données** : Nous disposons de seulement 32 points de données d'incidence jusqu'en 2023, avec des agrégats par classe d'âge, ce qui réduit la précision des prévisions.
- **Modélisation par type de cancer** : L'incidence globale du cancer est modélisée, mais les données par type de cancer sont partielles, ce qui empêche une analyse plus ciblée des risques.
- **Estimation du capital** : Le capital est estimé de manière prudente, mais des données plus précises pourraient affiner l'estimation et la rendre plus représentative des besoins réels.

Executive Summary

Cancer is a disease in which certain cells in the body undergo mutations, leading to uncontrolled proliferation and potentially forming what is known as a tumor. According to a study conducted by ISOP in 2023, cancer is the disease that concerns the French the most in terms of health.

As the leading cause of premature death in France, the number of cancer cases has steadily increased every year over the past 30 years¹. Despite this increase, the mortality rate has been continuously decreasing for 25 years, thanks to medical advances. In France, cancer treatment is fully covered by the National Health Insurance. However, in recent years, associations such as the Ligue contre le cancer have highlighted that, despite this coverage, the treatment journey remains a financial burden for many patients. Expenses such as excess fees, copayments, transportation costs, the need for home assistance due to treatment-induced fatigue, or complementary treatments not covered by insurance can quickly add up.

There is, however, a "critical illness insurance" policy that provides a lump sum payment upon diagnosis of such a disease. In France, its development remains relatively limited and poorly known. Insurers do not necessarily highlight this coverage, even though, in the current context, it could be relevant. In the United States, where the market is more developed due to a more complex healthcare system, studies predict significant growth for this type of product worldwide, particularly in Europe.

The objective of our study is to propose an insurance product dedicated to critical illness coverage specifically for cancer. In this study, we aim to determine the best method for predicting the cancer incidence rate in France. Then, we will estimate the optimal lump sum for a policyholder based on the level of financial support they wish to have in case of a cancer diagnosis.

The first step in developing such a product is to analyze the incidence rate, which represents the number of new cases in a given population. This rate is generally expressed as "the number of cases per 100,000 people per year," allowing us to track its evolution over time and compare it across different age groups.

1. Vie publique. Cancers: the number of new cases has almost doubled in France [19]

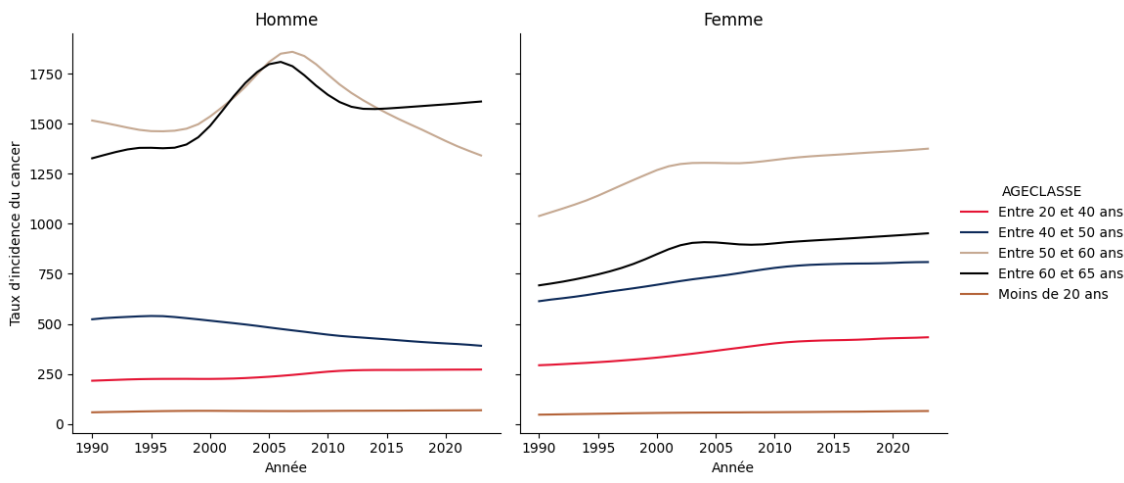


Figure 5 – Comparison of incidence by age and gender

For our forecasts, we tested three methods: ARIMA, a classical time series forecasting method; the Lee-Carter model, traditionally used for mortality rates; and a more recent approach using neural networks, the RVFL method.

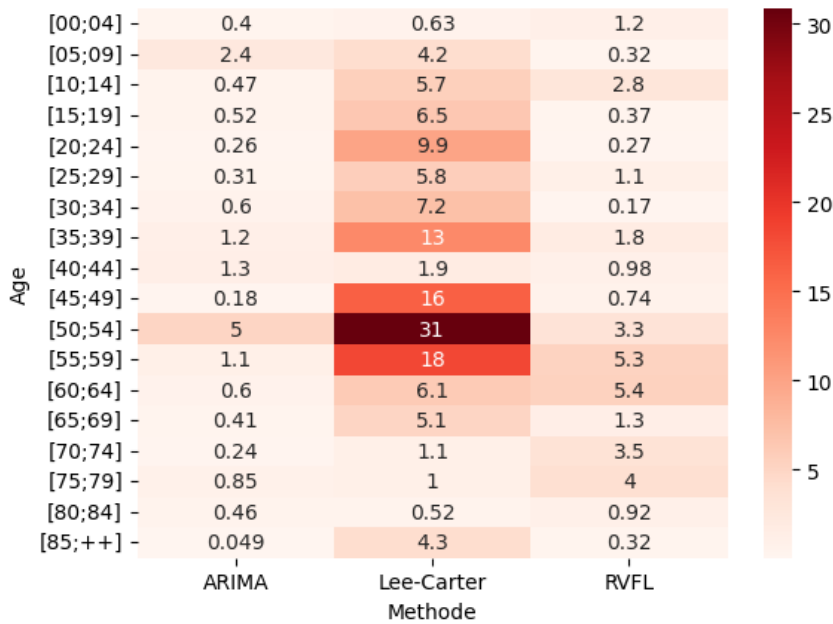


Figure 6 – MAE for each age group based on the three studied methods

After applying our models, the Lee-Carter model proved to be the least effective. We attributed this to the major screening campaign conducted in France in 2005, which caused a peak in the incidence rate over time. Since Lee-Carter is a trend-based model, it struggles to integrate variations due to one-time events.

Although the RVFL method performed better than Lee-Carter, the predictions were not convincing. This method, which has demonstrated efficiency in time series forecasting, did not perform well in our case due to data scarcity: we had only 30

data points per age category to train our models.

Ultimately, the ARIMA model was the most effective with the available data. Despite some weaknesses and the 2005 incidence peak, ARIMA allowed us to achieve significantly lower errors compared to the other two models.

Once the incidence was predicted, the second step was to determine the necessary lump sum to best support the insured. We identified the main additional expenses for cancer patients and estimated the total cost. Three levels of coverage were proposed, corresponding to different levels of financial support over a specified period.

The first coverage level provides minimal support over two years, covering the most basic complementary treatments.

The second coverage level offers better complementary care quality, with coverage for a duration of three years.

The third level grants access to all complementary treatments for a period of five years.

The calculated lump sum amounts for each coverage level are as follows:

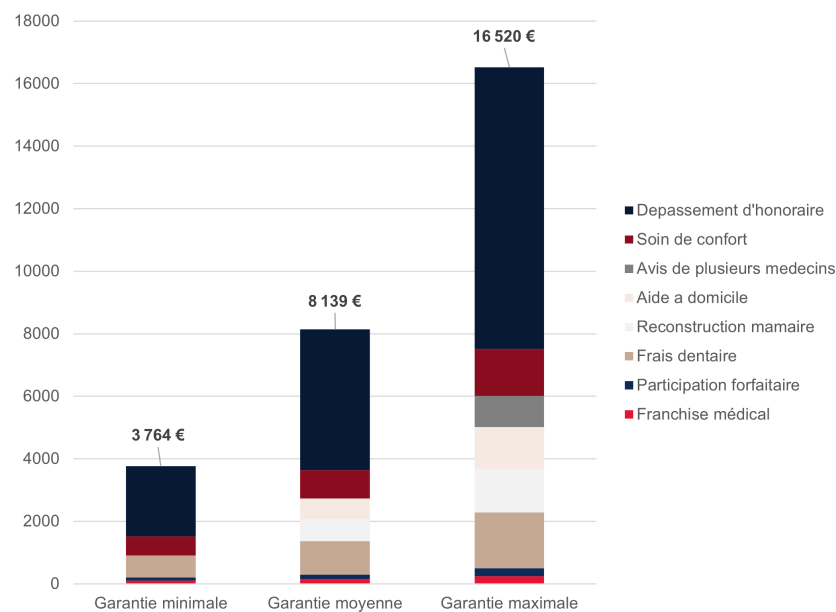


Figure 7 – Insured lump sum amounts based on coverage type

Once the cancer incidence rate was predicted over a 30-year horizon and the necessary lump sum estimated, we calculated the pure premium.

Our final product covers the cancer risk for policyholders from subscription until the age of 75. Coverage ends on the insured's 75th birthday. The contract terminates if the policyholder cancels it, passes away, or is diagnosed with cancer (leading to contract termination after the insured lump sum is paid).

After formulating these assumptions, we obtained the following pricing:

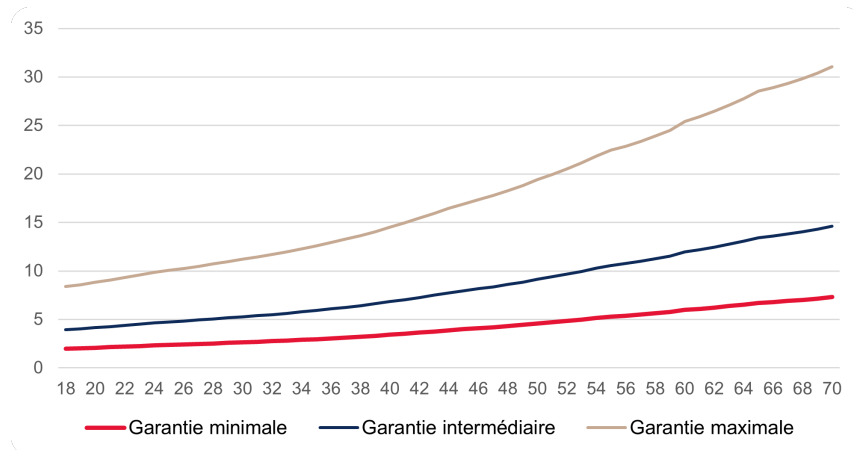


Figure 8 – Pure premium based on age and insured capital

Our model aims to estimate pure premiums in the context of cancer in France, but it has limitations due to data availability and quality:

- **External factors:** Integrating various exogenous variables into our model is challenging, especially over a 30-year horizon.
- **Data limitations:** We have only 32 incidence data points up to 2023, aggregated by age class, reducing the accuracy of forecasts.
- **Cancer type modeling:** We modeled overall cancer incidence, but data by cancer type is partial, preventing a more targeted risk analysis.
- **Lump sum estimation:** The estimated lump sum is conservative, but more precise data could refine the estimation and make it more representative of actual needs.

Remerciements

Je tiens tout d'abord à exprimer ma gratitude à Marina Choux, à l'origine de l'idée de ce mémoire, dont les conseils m'ont permis de m'engager sur ce sujet.

Un grand merci à Seabird, où j'ai eu la chance d'évoluer dans un environnement à la fois stimulant et bienveillant.

Je remercie tout particulièrement Sofia Sali pour son encadrement tout au long de cette expérience. Son suivi et ses conseils ont été d'une aide précieuse.

Je tiens également à exprimer ma reconnaissance à toute l'équipe de Seabird pour son accueil et son soutien, avec une mention spéciale aux alternants avec qui j'ai partagé cette année. Leur bonne humeur et leur esprit d'entraide ont largement contribué à rendre cette expérience aussi enrichissante que plaisante.

Enfin, merci à toutes celles et ceux qui, de près ou de loin, m'ont soutenu dans la réalisation de ce mémoire.

Contents

Résumé	3
Abstract	4
Note de synthèse	6
Executive Summary	11
Remerciements	16
Table des matières	18
1 Contexte général et assurance	21
1.1 Contexte	21
1.2 Qu'est ce que le cancer ?	22
1.2.1 Cellule normale et cellule cancéreuse	22
1.2.2 La cancérogenèse, mécanismes de la cancérisation	23
1.2.3 Les différents types de cancer	24
1.3 Le cancer en France	26
1.3.1 Panorama et chiffres-clés du cancer en France	26
1.3.2 Facteurs et causes	27
1.3.3 Traitements	28
1.4 Les coûts économiques	31
1.4.1 Les coûts globaux	31
1.4.2 Le reste charge	32

1.5	Les principaux acteurs sur le marché assurantiel	34
1.5.1	L'assurance maladie	34
1.5.2	Les organismes complémentaires de santé en France	37
1.5.3	Garantie maladie redoutée	38
1.5.4	Antisélection et aléa moral	40
2	Prédiction du taux d'incidence	42
2.1	Les données	42
2.1.1	L'INCa (Institut National du Cancer)	42
2.1.2	Les registres des cancers en France	43
2.1.3	Présentation des données	44
2.2	Modèle ARIMA	49
2.2.1	Choix de la méthode	49
2.2.2	Présentation de la méthode ARIMA	49
2.3	Méthode Lee-Carter	52
2.3.1	Choix de la méthode	52
2.3.2	Présentation de la méthode Lee-Carter	52
2.4	Modèle RVFL	55
2.4.1	Choix de la méthode	55
2.4.2	Présentation de la méthode RVFL	55
2.5	Résultats	57
2.5.1	Méthodologie	57
2.5.2	Indicateurs	59
2.5.3	Choix du modèle	61
3	Application sur le calcul de la prime pure	65
3.1	Marché de la garantie maladie redoutée	65
3.2	Capital pour notre produit	67
3.3	Méthode de calcul de la prime pure	77
3.4	Formule de calcul de la prime pure	79

3.5	Temps de couverture	80
3.6	Proposition de produit	82
3.6.1	Calcul de la prime	83
3.6.2	Test de sensibilité	85
3.6.3	Limites et améliorations possibles de notre modèle	88
Conclusion		90
Bibliographie		94

Chapter 1

Contexte général et assurance contre le cancer en France

1.1 Contexte

Ce mémoire s'inscrit dans la continuité des travaux menés récemment sur la modélisation du risque cancer en assurance, notamment dans le cadre de produits forfaitaires de type "garantie maladies redoutées". Les approches développées jusqu'ici ont principalement cherché à anticiper l'évolution de l'incidence du cancer à très long terme, sur la base de données japonaises, en mobilisant des modèles à facteurs dynamiques tels que Lee-Carter ou des modèles à facteurs variables dans le temps, avec pour objectif d'évaluer la rentabilité de ces produits dans différents scénarios économiques et épidémiologiques.

Le présent travail propose une perspective légèrement différente sur plusieurs plans. D'abord, il se distingue par un recentrage géographique, en portant l'analyse sur les données françaises, alors que la littérature récente s'est majoritairement concentrée sur les marchés asiatiques. Ce choix permet de mieux intégrer les spécificités du système de santé français, ainsi que les comportements de souscription observés localement.

Sur le plan méthodologique, ce mémoire propose une comparaison élargie des approches de prévision, en mobilisant à la fois des modèles classiques (ARIMA, Lee-Carter) et un modèle de type RVFL (Random Vector Functional Link), qui s'inspire des réseaux de neurones tout en conservant une structure relativement simple.

Par ailleurs, une autre différence notable concerne l'orientation du raisonnement actuariel. Alors que les travaux précédents s'attachent surtout à mesurer la rentabilité attendue d'un produit donné, ce mémoire propose une réflexion autour de l'estimation du capital nécessaire pour couvrir les frais liés à un diagnostic de cancer. L'objectif est ici d'approcher un capital dit "optimal", c'est-à-dire suffisant pour faire face aux charges supportées par l'assuré, tout en étant minimisé afin de garantir une prime d'assurance compétitive. Il ne s'agit pas de modéliser l'ensemble du produit dans ses moindres détails, mais de fournir un cadre de réflexion pour évaluer l'adéquation entre niveau de couverture, niveau de risque, et accessibilité tarifaire. Cette approche permet de replacer la conception du produit dans une logique d'équilibre entre protection efficace et soutenabilité économique.

En somme, ce mémoire apporte une contribution modeste mais utile à la réflexion actuarielle sur l'assurance cancer. Il s'attache à adapter les outils existants au contexte français, propose une exploration comparative des méthodes de projection, et introduit une analyse orientée vers la gestion du capital et des risques. Ces éléments, pris ensemble, visent à enrichir les travaux existants tout en restant ancrés dans les préoccupations concrètes des acteurs du secteur.

1.2 Qu'est ce que le cancer ?

1.2.1 Cellule normale et cellule cancéreuse

Notre corps est composé de trillions de cellules, chacune d'elles possède un noyau refermant des gènes qui leur indiquent comment se développer, se comporter, se multiplier et mourir. De sa création à sa mort, en passant par son développement, une cellule suit ce que l'on appelle le cycle cellulaire. À chaque cycle, la cellule mère croît, réplique son ADN et se divise pour produire deux cellules filles identiques. Lorsqu'une cellule devient trop vieille ou endommagée et ne peut plus remplir ses fonctions, elle meurt par un processus d'auto-destruction. Le cycle cellulaire comporte des points de contrôle qui empêchent le développement des cellules 'en mauvaise santé. Dans de rares cas, ces points de contrôle peuvent être défectueux, permettant à une cellule mutée ou endommagée de se développer de manière désordonnée, créant ainsi des cellules cancéreuses.

Il existe 3 différences essentielles entre une cellule saine et une cellule cancéreuse :

- **L'immortalité**

Lorsque l'on étudie le processus de division cellulaire, on remarque que même dans les conditions nécessaires à leur croissance, au-delà d'un certain nombre de cycles cellulaires, leur capacité à se diviser disparaît et leur mort survient. Cet événement, qui n'intervient pas sur les cellules cancéreuses, est programmé et c'est ce que l'on nomme l'apoptose.

- **L'autonomisation de la division cellulaire**

Une cellule saine ou normale ne se divise que quand elle en reçoit le signal, sa prolifération est contrôlée par des substances inhibitrices ou stimulantes. La cancérisation provoque un dérèglement de la division cellulaire conduisant ainsi à l'autonomisation de la cellule cancéreuse. De ce fait, elle peut se reproduire par elle-même sans un signal du corps humain.

- **Infiltration et migration**

Elles peuvent facilement se propager et elles envahissent et endommagent les tissus et organes. Lors de leur prolifération, elles vont finir par se tasser pour devenir une masse compacte que l'on appelle « tumeur maligne ».

1.2.2 La cancérogenèse, mécanismes de la cancérisation

On distingue schématiquement 3 étapes lors du développement d'un cancer : l'initiation, la promotion et la progression :

- **L'initiation**

L'initiation est la première étape qui intervient lors du développement d'un cancer, durant laquelle s'opère un changement du matériel génétique de la cellule normale. Cette modification de l'ADN de la cellule peut être d'origine endogène (lorsqu'il s'agit d'une erreur de la duplication de l'ADN) ou résulter de l'exposition externe à des facteurs environnementaux cancérogènes. Par cette altération de l'ADN, la cellule acquiert la propriété d'échapper aux régulations cellulaires, elle est alors dite « maligne ».

- **La promotion**

La seconde phase est celle de la promotion. Cette étape, relativement longue, correspond à la prolifération de la cellule initiée par le concours d'agents promoteurs. Ces agents promoteurs, qui sont des substances rencontrées dans le milieu (comme les hormones) ou des facteurs extérieurs (alimentaires ou chimiques) et qui ne sont pas cause du cancer en eux-mêmes, vont, par leur action répétitive, permettre à la cellule initiée de proliférer et provoquer un dérèglement de la multiplication cellulaire.

- **La progression ou propagation**

La troisième et dernière phase qu'on appelle la progression ou propagation, est une phase relativement longue et complexe où l'on observe une vascularisation de la tumeur. Les cellules cancéreuses, dotées des propriétés que nous avons notées plus haut, vont peu à peu former de nouveaux vaisseaux sanguins qui vont alimenter et irriguer la tumeur en oxygène et nutriments. Sans ce processus, connu sous le nom d'angiogenèse, une tumeur ne pourrait croître au-delà d'un dixième de millimètre. Et c'est par ces nouveaux vaisseaux sanguins que le cancer va pouvoir se propager et migrer vers les tissus ou les organes sains voisins, formant ainsi des métastases.

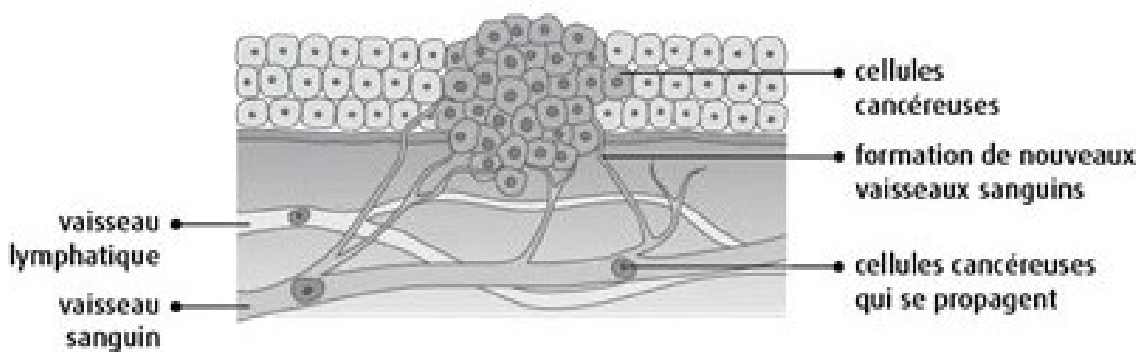


Figure 1.1 – Schéma représentant la propagation du cancer ¹

1.2.3 Les différents types de cancer

Le terme de cancer regroupe plus d'une centaine de maladies très différentes les unes des autres. Il est d'ailleurs à ce titre plus juste de parler des cancers au pluriel que du cancer. Les cancers sont classés selon la nature du tissu dans lequel ils se développent en premier lieu. On distingue ainsi les tumeurs du sang ou des tissus hématopoïétiques des tumeurs dites « solides » (les sarcomes et les carcinomes) qui sont caractérisées par une masse individualisée.

L'Institut National du Cancer les définit comme suit :

– Les carcinomes

Les cellules cancéreuses se forment initialement dans un épithélium. Un épithélium est un type de tissu composé de cellules étroitement jointes qui recouvrent les surfaces internes et externes du corps. Les surfaces internes incluent les tissus de revêtement des organes, comme les muqueuses de l'estomac et des intestins. Les surfaces externes incluent des parties du corps comme la peau, aussi appelée épiderme.

Parmi les cancers qui apparaissent dans l'épithélium, on trouve les adénocarcinomes. Les adénocarcinomes sont un type de cancer qui se développe partir de l'épithélium glandulaire. L'épithélium glandulaire est le tissu qui compose les glandes, des structures spécialisées dans la sécrétion de substances comme des enzymes ou des hormones. Par exemple, le tissu glandulaire du sein produit du lait, et celui de la prostate produit un liquide qui fait partie du sperme.

Ces cancers sont assez fréquents et se développent principalement dans les glandes mammaires et prostatiques.

– Les sarcomes

Les cellules cancéreuses peuvent également se former dans un type de tissu appelé « tissu de support ». Le tissu de support comprend les structures du corps qui soutiennent, relient ou entourent d'autres tissus et organes. Parmi ces tissus de

1. Source : Canadian Cancer Society

support, on trouve les os, la graisse et les muscles.

Dans la famille des sarcomes, on retrouve les ostéosarcomes, quand le cancer se développe dans les os, les liposarcomes, quand il se développe dans les tissus graisseux, et enfin les rhabdomyosarcomes, quand le cancer se développe dans les muscles striés, c'est-à-dire les muscles responsables des mouvements volontaires.

– Les cancers hématopoïétiques ou hématologiques

Les cellules cancéreuses apparaissent dans la moelle osseuse, qui est le tissu situé à l'intérieur des os et responsable de la production des cellules sanguines. La moelle osseuse fabrique les globules rouges (qui transportent l'oxygène dans le corps), les globules blancs (qui combattent les infections) et les plaquettes (qui aident à la coagulation du sang), ainsi que leurs cellules précurseurs.

Les cellules cancéreuses peuvent également apparaître dans d'autres organes du système lymphoïde, qui est une partie essentielle du système immunitaire. Les organes lymphoïdes comprennent le thymus (une glande située derrière le sternum), les ganglions lymphatiques (petites structures en forme de haricot réparties dans tout le corps), la rate (un organe situé dans l'abdomen qui filtre le sang) et les amygdales (des masses de tissu lymphoïde situées à l'arrière de la gorge).

Il existe trois principales familles de cancers hématologiques, qui sont des cancers affectant le sang, la moelle osseuse et le système lymphatique :

- **Les leucémies** : Ce sont des cancers des cellules sanguines et de la moelle osseuse qui entraînent une production excessive de globules blancs anormaux.
- **Les myélomes** : Ce sont des cancers des cellules plasmiques, un type de globule blanc présent dans la moelle osseuse, responsable de la production d'anticorps.
- **Les lymphomes** : Ce sont des cancers des lymphocytes, un autre type de globule blanc, qui se développent principalement dans les ganglions lymphatiques et d'autres tissus lymphoïdes.

1.3 Le cancer en France

1.3.1 Panorama et chiffres-clés du cancer en France

Le cancer est la première cause de mortalité prématurée en France chez l'homme et la seconde chez la femme (derrière les maladies de l'appareil circulatoire). En 2023, la projection estimée du nombre de cas de cancer en France est estimée à 433 136. Entre 1990 et 2023, le nombre de cas a doublé avec une augmentation de 98% des cancers chez l'homme et de 104% chez la femme. Cette augmentation s'explique selon 3 facteurs : l'évolution démographique et le vieillissement de la population qui expliquent tout deux 78% de l'augmentation du nombre de cas chez l'homme et 57% de l'augmentation du nombre de cas chez la femme. Le troisième facteur de risque responsable du reste de ces augmentations est lié directement à nos modes de vie et comportements adoptés. Aujourd'hui, on estime à 3.8 millions le nombre de personnes atteintes d'un cancer en France. Bien qu'il existe de nombreux cancers, celui de la prostate reste le plus fréquent chez l'homme (24% des nouveaux cas en 2023) et celui du sein chez la femme (33% des nouveaux cas en 2023). Ci-dessous, un schéma représentant les cancers les plus fréquents en 2023 chez l'homme ainsi que l'évolution de leur incidence entre 2010 et 2023 :

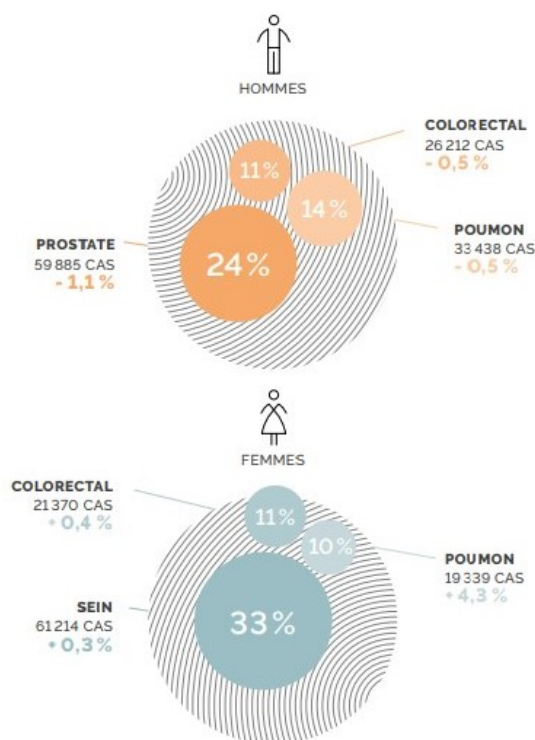


Figure 1.2 – Nombre de cas en 2023 et évolution du taux d'incidence entre 2010 et 2023¹

1. Source : Panorama des cancers, 2023

1.3.2 Facteurs et causes

La transformation d'une cellule maligne en une cellule cancéreuse dépend de nombreux facteurs liés à l'environnement, au mode de vie, à l'âge ou encore au patrimoine génétique. Chaque cancer n'est jamais la conséquence d'une cause unique mais d'un ensemble complexe de facteurs qui interagissent entre eux. On parle du cancer comme d'une maladie multifactorielle. C'est l'épidémiologie, en étudiant la fréquence des différents cancers dans des populations ayant des modes de vie différentes, qui permet d'identifier les facteurs responsables des cancers..

On distingue ainsi les facteurs externes, souvent évitables des facteurs internes.

– Les facteurs de risque externes

Parmi les facteurs externes, de nombreux dépendent de nos comportements et de nos modes de vie.

Le tabac est aujourd'hui le plus important facteur de risque de cancer connu en France et dans le monde. Cause d'environ un quart des décès par cancer, le tabac est responsable de près de 20% des cancers, ce qui représente l'apparition d'environ 70 000 nouveaux cas de cancer chaque année en France (principalement les cancers des poumons, des voies aérodigestives ¹ et les cancers de la vessie). Vient ensuite l'alcool dont la consommation excessive est responsable de 8% des nouveaux cas de cancers en France. L'alimentation, l'exposition aux UV, la sédentarisation, le surpoids et l'obésité sont également d'importants facteurs de risques de développer des cancers.

D'autres facteurs sont liés à l'environnement comme les polluants d'origine physique ou chimique que l'on trouve surtout en milieu professionnel. On retrouve parmi ces agents cancérogènes les rayonnements et radiations ionisantes (comme les rayons X lors des radiographies) et des substances comme l'amiante, la poussière de bois, le glyphosate, le benzène mais aussi des pesticides. Il faut aussi compter parmi les facteurs environnementaux la pollution atmosphérique ainsi que les agents infectieux (bactéries et virus).

Ces agents infectieux seraient responsables de 4% des cancers en France. Les principaux connus sont les papillomavirus humains (HPV), les virus des hépatites B et C, la bactérie *Helicobacter pylori*, le virus de l'immunodéficience humaine (VIH) ou encore, le virus d'Epstein-Barr.

Le Centre International de Recherche sur le Cancer, basé à Lyon (CIRC ou IARC en anglais) qui dépend de l'Organisation Mondiale de la Santé, a ainsi depuis 1971 classé 534 agents cancérogènes ou potentiellement cancérogènes³. Ces agents désignent aussi bien des : substances chimiques, mélanges complexes, expositions professionnelles, organismes biologiques, des agents physiques, des composants de l'alimentation, des médicaments. Parmi ces 534 agents, 121 sont reconnus can-

1. La bouche, le larynx, le pharynx et l'œsophage.

3. JNCI: Journal of the National Cancer Institute, Volume 103, Issue 24, 21 December 2011, Pages 1827–1839, <https://doi.org/10.1093/jnci/djr483>

cérogènes, 93 le sont probablement, et 320 le sont peut-être. Par un changement de nos modes de vie ou nos comportements, on estime que près de la moitié des cancers en France, soit environ 170 000 par ans, pourraient être évités.

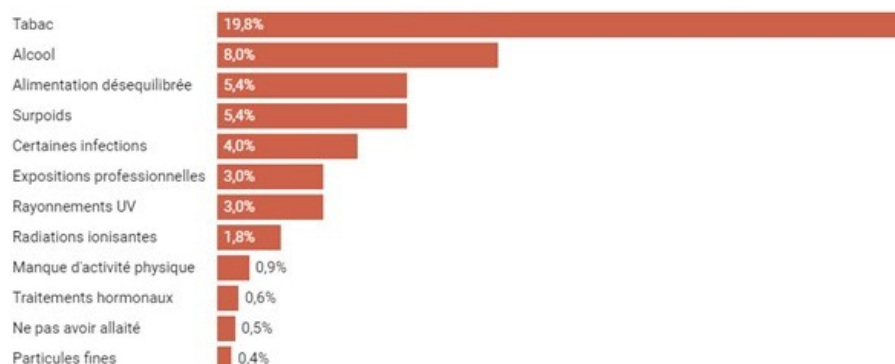


Figure 1.3 – Proportion des cancers liés aux principaux facteurs de risque⁴

– Les facteurs de risque internes

Parmi les facteurs internes, celui de l'âge est le plus important. En effet, même si un cancer peut se développer tout au long de la vie, il est bien plus fréquent à partir de 60 ans. De plus, avec le temps, les mécanismes de réparation de l'ADN sont de moins en moins efficaces. A cela s'ajoute le cumul dans le temps des agressions externes subies par les cellules. Outre l'âge, il existe aussi un facteur de risque lié à l'hérédité et au patrimoine génétique. On estime en effet que 5% à 10% des cancers sont liés à la transmission héréditaire d'une mutation génétique connue.

1.3.3 Traitements

Chaque patient est unique et il existe, comme nous l'avons expliqué, de nombreux types de cancers différents, ainsi le parcours de soins n'est jamais identique. Cependant, on retrouve souvent les mêmes traitements.

Les principaux sont :

– La chirurgie

Comme évoqué précédemment, une tumeur est une masse de cellules qui se développent anormalement. La chirurgie a pour but de la retirer par une opération directement sur l'organe ou la partie du corps touchée. Les chances de succès sont plus grandes lorsque la tumeur est encore petite et dépendent de la zone à opérer. La première chirurgie pour le cancer remonte à 1906, il y a plus de 100 ans. Aujourd'hui en France, il y a eu 419 050 opérations réalisées en 2021 selon l'INCa. Bien que cette pratique soit en légère baisse suite à l'apparition de nouvelles méthodes, elle reste prépondérante.

4. Graphique : Vie-publique.fr / DILA. Source : Institut National du Cancer.

– La chimiothérapie

La chimiothérapie est un traitement reposant sur la prise de médicaments. Ceux-ci ont pour but soit de stopper la prolifération des cellules cancéreuses, soit de les éliminer. Les médicaments sélectionnés dépendent du type de cancer. L'administration peut se faire par voie orale ou par injection intraveineuse au cours de perfusions. Née au début des années 1940, elle est aujourd'hui le deuxième traitement le plus utilisé en France, avec 363 160 personnes traitées en 2021.

– La radiothérapie

La radiothérapie vise à détruire les cellules cancéreuses par des rayonnements ciblés sur la tumeur. L'intervention peut être préopératoire, afin de réduire au maximum la tumeur, ou post-opératoire, pour éliminer les cellules cancéreuses qui n'auraient pas été complètement éliminées par l'opération. La première utilisation remonte à 1896, par un docteur dans un hôpital de Lyon. Aujourd'hui, elle représente le 3^e traitement le plus utilisé, avec 129 250 personnes traitées dans le secteur public et 113 042 traitées dans le secteur libéral.

– Les thérapies ciblées

Les thérapies ciblées sont arrivées dans les années 2000. Elles ne concernaient d'abord que certains types de cancers (du sein et du sang), et peuvent s'appliquer maintenant grâce aux progrès de la recherche à plus de 20 types de cancer différents. La thérapie ciblée repose sur l'analyse moléculaire des cellules cancéreuses afin de déterminer le médicament et la dose optimale pour empêcher la prolifération et détruire les cellules anormales.

– L'immunothérapie

Présentant des résultats très prometteurs, cette méthode est au cœur des recherches depuis le début des années 2000. Il s'agit d'utiliser les défenses immunitaires du corps humain pour détruire les cellules cancéreuses. Plusieurs méthodes existent, mais les recherches se concentrent sur la modification des gènes des cellules immunitaires pour les armer contre les cellules cancéreuses. En 2021, plus de 60 000 personnes ont eu recours à ce traitement en France.

– L'hormonothérapie

Les hormones peuvent, sur certains cancers, comme celui du sein ou de la prostate, augmenter la capacité des cellules cancéreuses à se multiplier. L'hormonothérapie consiste à bloquer la production des hormones qui contribuent à la prolifération des cellules anormales. Par exemple, pour le cancer du sein, on bloquera les œstrogènes et les androgènes pour les cancers de la prostate. Contrairement aux traitements vus précédemment cette intervention a pour seul but de stopper la multiplication des cellules cancéreuses. Son utilisation est moins fréquente,

Le principe de ces différentes interventions est souvent de limiter la propagation des cellules cancéreuses ou de les éliminer. Régulièrement, un malade aura recours à plusieurs de ces traitements dans le parcours de soin d'un cancer. Les traitements

locaux agissent sur une zone précise du corps comme la chirurgie ou la radiothérapie. Cependant, il arrive que les traitements locaux n'arrivent pas à éliminer toutes les cellules, ainsi on utilisera les traitements systémiques qui agissent sur tout le corps tels que l'hormonothérapie, la chimiothérapie ou bien l'immunothérapie.

Il n'est pas possible de trouver le pourcentage d'efficacité de chaque traitement, mais pour chaque cancer, on peut déterminer la survie nette à 5 ans. Cela correspond au pourcentage de personnes qui ont survécu au moins 5 ans après le traitement de leur cancer.

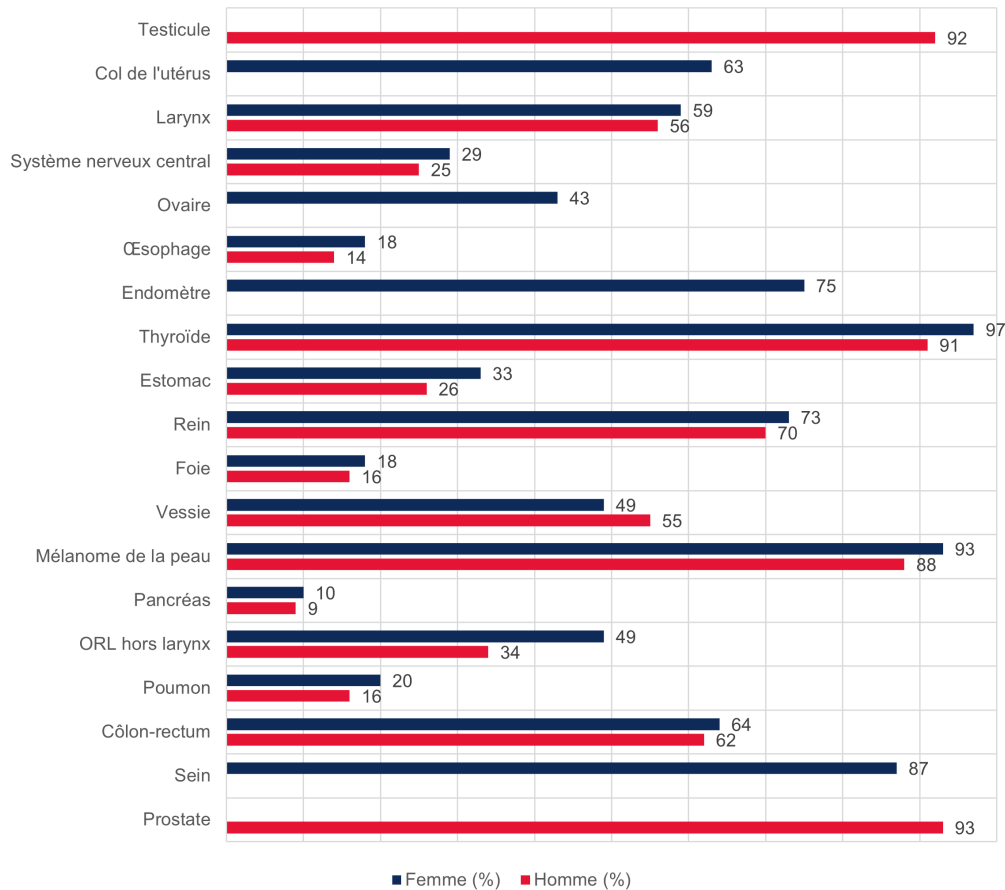


Figure 1.4 – Survie nette à 5 ans par type de cancer

1.4 Les coûts économiques

1.4.1 Les coûts globaux

Selon le *Rapport charges et produits pour 2017*, le coût annuel, en France, de la prise en charge des cancers pour le système de santé était de 16,1 milliards, dont 3,2 milliards pour les médicaments anticancéreux ¹.

En 2021, les dépenses hospitalières liées au diagnostic, au traitement ou au suivi des personnes atteintes de cancer représentaient 6,3 milliards d'euros ², et ces chiffres ne font qu'augmenter d'année en année.

Parce que le cancer est une maladie qui nécessite un suivi médical long et des traitements coûteux, il est reconnu en France comme une affection de longue durée. En effet, l'affection de longue durée (ALD) concerne une maladie dont la gravité et/ou le caractère chronique nécessitent un traitement prolongé. À ce titre, les soins sont pris en charge à 100% par l'Assurance Maladie et n'ont pas besoin d'être avancés.



Figure 1.5 – Synthèse des caractéristiques concernant l'ALD

Cette image peut se retrouver sur le site de l'assurance maladie, elle a pour but de synthétiser toutes les informations importantes sur le cas de l'ALD.

1. Dossier de presse, Observatoire Cancer Institut Curie – Viavoice, 2017

2. Panorama des cancers 2023.

Le coût annuel de la prise en charge des cancers en France représente 12,63 % des dépenses de l'assurance maladie, ce qui en fait l'une des maladies les plus coûteuses pour notre système de santé.

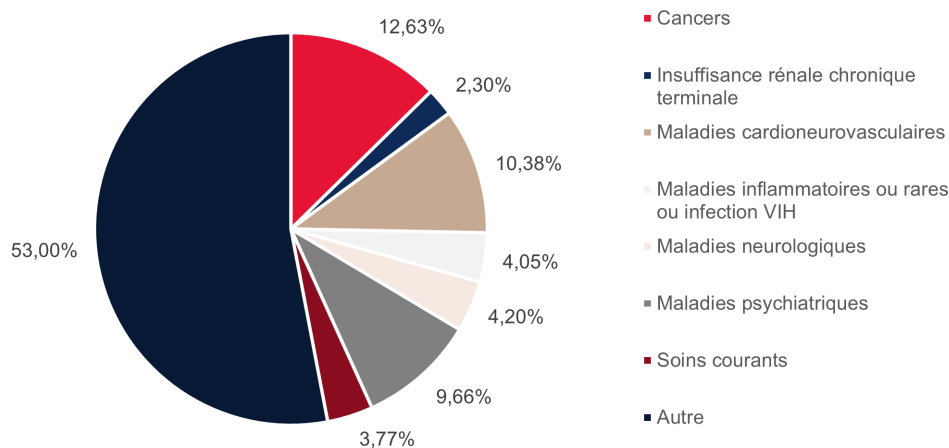


Figure 1.6 – Répartition des dépenses de l'assurance maladie

Mais, bien que le cancer soit reconnu comme une affection de longue durée et intégralement pris en charge par l'Assurance Maladie, il existe de nombreux restes à charge pour les personnes atteintes, parfois très coûteux.

1.4.2 Le reste charge

Les personnes atteintes de cancer se retrouvent confrontées, tout au long de leur parcours de soins et même parfois bien après la fin de leurs traitements, à de multiples dépenses qui peuvent s'avérer lourdes.

De nombreux organismes, comme l'Observatoire Sociétal des Cancers, ont révélé les nombreux frais restant à la charge des individus atteints de cancer, notamment en ce qui concerne les soins dits de « confort ». Ces soins, tels que les médicaments pour lutter contre les infections, les produits d'hygiène spécialisés, les vitamines et compléments alimentaires, ainsi que les traitements contre les effets secondaires (problèmes de peau, chute des ongles, repousse des cheveux, etc.), sont payants pour les malades. Bien que ces soins soient parfois indispensables pour favoriser la guérison, ils sont souvent onéreux. À cela s'ajoutent les dépassements d'honoraires, les frais de transport, les suivis psychologiques et les consultations de spécialistes comme les diététiciens et ergothérapeutes. Face à ces dépenses supplémentaires, les individus sont parfois contraints de renoncer à certains médicaments complémentaires au traitement principal, de sacrifier d'autres postes de dépenses comme les loisirs ou l'alimentation, de recourir à la solidarité familiale, voire de s'endetter ¹.

1. Article de la Ligue contre le cancer, Les restes à charge, entre incompréhension et acceptation, 2013.

Par ailleurs, pour les personnes atteintes de cancer, le retour à l'emploi est difficile : une personne sur cinq a perdu son emploi cinq ans après le diagnostic ¹. Comme l'a indiqué la Ligue contre le cancer dans son Observatoire Sociétal, les malades sont parfois contraints, faute d'accès aux soins près de chez eux, de déménager à proximité de leur lieu de traitement ², engendrant des dépenses supplémentaires et imprévues.

Pour protéger ses employés de telles dépenses, une entreprise peut ajouter à son contrat santé la garantie « maladies redoutées ».

1. Panorama des cancers 2023

2. Observatoire Sociétal des Cancers

1.5 Les principaux acteurs sur le marché assurantiel

1.5.1 L'assurance maladie

Avant les ordonnances de 1945, la France ne disposait pas de système de sécurité sociale généralisé, seules certaines professions bénéficiaient de régimes de protection spécifiques.

Initialement conçue comme un système universel, la Sécurité sociale a néanmoins rencontré des réticences de la part de certaines professions, craignant une moindre couverture. Cela explique la coexistence actuelle du régime général, englobant plus de 90 % de la population, du régime agricole et d'autres régimes sociaux.

La Sécurité sociale « est la garantie donnée à chacun qu'en toutes circonstances il disposera des moyens nécessaires pour assurer sa subsistance et celle de sa famille dans des conditions décentes »¹. Elle va ainsi préserver les citoyens des dépenses imprévues et garantir un accès équitable aux soins de santé entre les français. Elle repose sur un principe de solidarité, chacun participant à son financement selon ses moyens. Le régime de la sécurité sociale se compose de six branches.

Une branche est une entité qui a à sa charge la gestion d'un ou plusieurs « risques ». Ces risques sont définis comme des événements qui peuvent, au cours d'une vie, porter atteinte à la sécurité économique d'une personne. Ils font donc l'objet d'une prise en compte, d'une réparation ou d'une rétribution.

L'assurance maladie, comme représentée sur la figure 1.7 est composée de la branche maladie et de la branche des risques professionnels.



Figure 1.7 – Les différentes branches de la Sécurité sociale

1. Source : ordonnance du 4 octobre 1945, texte fondateur

La branche des risques professionnels gère les risques professionnels auxquels sont confrontés les travailleurs : accidents du travail, accidents de trajet et maladies professionnelles. À ce titre, elle indemnise les victimes et fixe la contribution respective des entreprises au financement du système. Elle met en œuvre une politique de prévention des risques professionnels.

Pour ce qui est de la branche maladie, elle assure la couverture des dépenses de santé des assurés et garantit l'accès aux soins, tout en favorisant la prévention et contribuant à la régulation du système de santé français. Elle englobe les risques liés à la maladie, à la maternité, à l'invalidité et au décès. Son rôle social est également essentiel, favorisant l'accès à la santé des populations les plus défavorisées et contribuant au bon fonctionnement des établissements médico-sociaux.

La gestion des risques occupe une place centrale dans sa mission. Elle vise à maîtriser l'évolution des dépenses tout en améliorant l'état de santé global de la population. Pour assurer ses fonctions, la branche Maladie est administrée par la Caisse nationale de l'Assurance maladie (Cnam) et son réseau étendu, incluant les caisses primaires d'assurance maladie (Cpam), les caisses générales de sécurité sociale (CGSS) dans les départements d'outre-mer, les directions régionales du service médical (DRSM), les caisses d'assurance retraite et de la santé au travail (Carsat), ainsi que les unions de gestion des établissements de caisse d'assurance maladie (Ugecam).

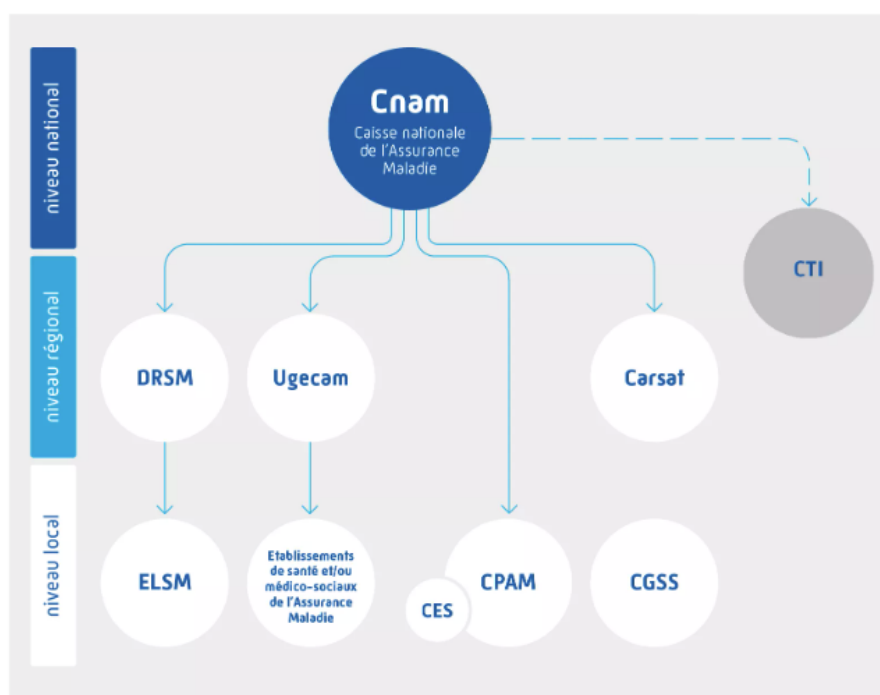


Figure 1.8 – Réseau de la Cnam

La réforme de 2016 a simplifié l'accès à la couverture maladie en instaurant la "protection universelle maladie" (Puma). Ainsi, toute personne travaillant ou résidant de manière stable et régulière en France a droit à une prise en charge continue de ses frais de santé tout au long de sa vie.

Tout comme les branches retraite et famille de la Sécurité sociale, l'Assurance Maladie se finance principalement grâce aux cotisations et contributions des travailleurs, représentant plus de 90 % des ressources de la Sécurité sociale. Ces fonds sont prélevés sur les salaires bruts des employés, qui contribuent à hauteur d'environ 15,9 % chaque mois, ainsi que sur les salaires versés par les entreprises, qui reversent environ 29 % des salaires de leurs salariés. Pour le régime général de l'Assurance Maladie, ces cotisations et contributions sont collectées par l'Urssaf. En outre, une partie des impôts collectés par l'État, provenant notamment des taxes sur le tabac et l'alcool, contribue également au financement de la Sécurité sociale, représentant 9 % de ses ressources.

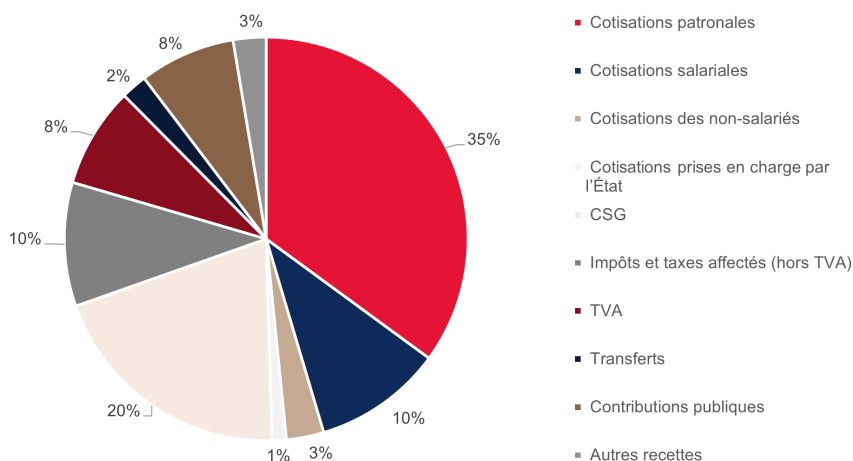


Figure 1.9 – Répartition des recettes de la Sécurité Sociale

L'assurance maladie rembourse les dépenses de ses assurés selon une base de remboursement prédéfinie. Cependant, cette couverture de base ne couvre pas toujours l'intégralité des frais réels engagés. C'est pour combler cette lacune que les organismes complémentaires de santé existent, offrant ainsi une protection supplémentaire et permettant de réduire le reste à charge pour les patients.

1.5.2 Les organismes complémentaires de santé en France

Les contrats collectifs santé

Achat de lunettes de vue, pose de prothèses dentaires ou auditives, consommation de certains médicaments..., de nombreuses dépenses de santé ne sont pas prises en charge, ou pas totalement, par l'assurance maladie obligatoire. C'est pourquoi la plupart des assurés optent pour une complémentaire santé.

Une complémentaire santé est un contrat qui a pour but de compléter les remboursements de la Sécurité sociale dans les champs de la maladie, des accidents et de la maternité. Ces contrats permettent une prise en charge de tout ou une partie de ces dépenses. Par exemple, si vous consultez votre médecin traitant qui est un médecin généraliste conventionné exerçant en secteur 1, sur le montant de 26,50 € de la consultation, l'Assurance maladie vous remboursera 70 %, soit 17,55 € (moins 1€ de la participation forfaitaire). Les 30 % restants, soit 7,95 €, seront à votre charge. Si vous possédez une complémentaire santé, c'est cette dernière qui remboursera tout ou partie de cette somme.

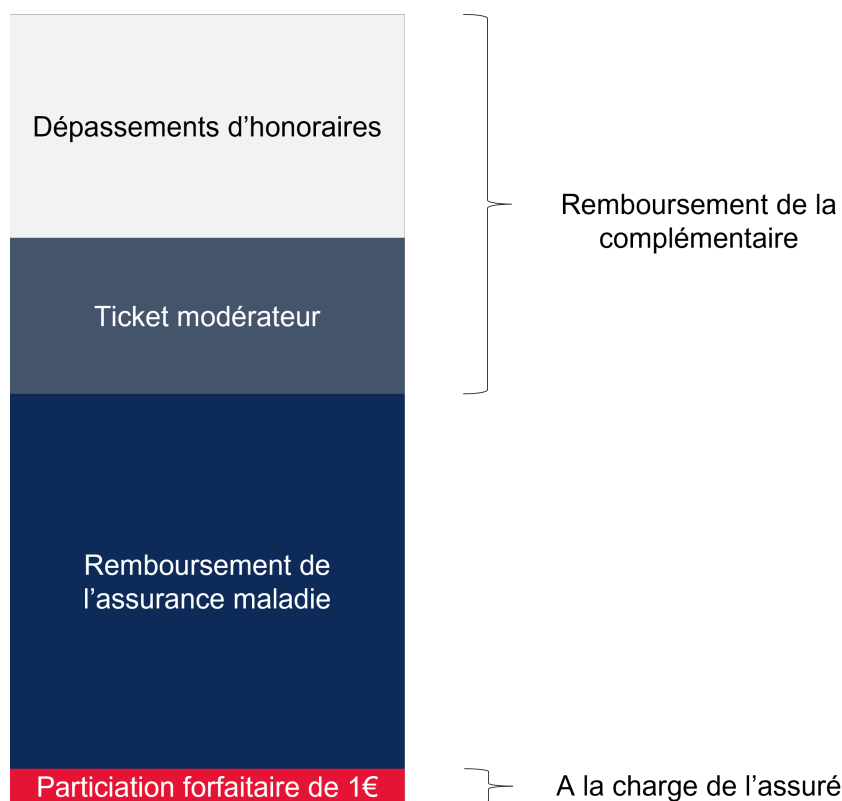


Figure 1.10 – Fonctionnement du remboursement des frais de santé

Il existe deux types de mutuelles : la mutuelle santé individuelle, qui n'est pas obligatoire mais recommandée, et la mutuelle santé collective, qui est obligatoire. On parle également de mutuelle santé d'entreprise. Depuis 2016, les entreprises ont l'obligation de proposer une couverture complémentaire de santé collective à leurs salariés qui n'en disposent pas. Elles participent au minimum à 50% au montant des

cotisations. L'employeur et/ou les partenaires sociaux peuvent ouvrir la couverture aux ayants droit (conjoint, enfants) du salarié, ce n'est pas systématique. Tous les contrats n'offrent pas les mêmes niveaux de garanties et de remboursement. Une complémentaire santé de base propose le remboursement du "ticket modérateur", soit la différence entre le coût global et la part prise en charge par la Sécurité sociale, pour les actes les plus basiques, sans dépassement d'honoraires.

1.5.3 Garantie maladie redoutée

Histoire

Le 6 août 1983, la première assurance contre les maladies graves a été introduite. Contrairement aux assurances traditionnelles, cette initiative a été lancée non par une compagnie d'assurance, mais par un chirurgien cardiaque, qui n'avait initialement aucune expérience en matière d'assurance.

Ce chirurgien a observé, dans son environnement de travail, que de nombreux patients faisaient face à des difficultés financières après le diagnostic et le traitement de maladies graves. Avec l'amélioration des traitements médicaux, les coûts associés ont également augmenté, ce qui a engendré un coût financier important pour les patients. La fatigue liée à la maladie compliquait davantage la capacité des patients à travailler et à générer des revenus pour couvrir les frais médicaux nécessaires. Ainsi les patients étaient souvent contraints de faire des sacrifices tant sur un point de vue personnel que médical. Ces défis financiers avaient un impact significatif sur leur santé et la qualité de leur traitement.

Face à cette situation, il est devenu évident que le soutien financier était essentiel pour un traitement efficace. L'analyse des produits d'assurance disponibles a révélé que les options existantes se limitaient à l'assurance-vie, qui fournit des fonds après le décès, et à l'assurance-invalidité, qui soutient financièrement les personnes incapables de travailler. Cependant, les patients atteints de crises cardiaques, d'AVC ou de cancer survivaient des années après leur maladie mais étaient appauvris. Il est donc apparu nécessaire de développer une nouvelle forme d'assurance. Ce produit devait fournir une assistance financière au moment du diagnostic de la maladie, lorsque les besoins des patients étaient les plus urgents.

Après trois années de recherche pour obtenir le soutien des assureurs, une collaboration a été établie avec Crusader Life, un assureur sud-africain, permettant le lancement de la première police d'assurance contre les maladies graves en 1983. Cette police couvrait initialement quatre conditions : crise cardiaque, cancer, AVC et chirurgie des artères coronaires. Bien que le lancement ait suscité un certain enthousiasme parmi les courtiers, les ventes initiales ont été lentes. Cependant, l'introduction de cette nouvelle couverture a entraîné la création de polices similaires par d'autres compagnies d'assurance, augmentant ainsi la sensibilisation à ce produit.

Le produit s'est d'abord développé en Afrique du Sud, puis s'est étendu au Royaume-Uni, en Irlande, en Australie, en Nouvelle-Zélande, en Extrême-Orient, et en Amérique du Nord, qui est aujourd'hui le marché principal, ainsi que dans plusieurs autres pays dont la France.

Les données recueillies montrent que la majorité des réclamations concernaient les crises cardiaques, les AVC et le cancer avec un âge moyen d'environ 41 ans. Ces caractéristiques étaient constantes dans tous les pays fournissant des statistiques. Le cancer était la maladie qui inquiétait le plus, en effet, elle représentait 50 % des demandes dans la plupart des pays, et jusqu'à 80 % dans certains.

Fonctionnement de la garantie

Une affection de longue durée (ALD) est une maladie grave dont les soins sont lourds et coûteux, et s'étalent sur une durée importante. Il existe plusieurs types de maladies dans cette catégorie dont la prise en charge et les coûts varient.

Les maladies les plus contractées en France et prises en charge par la sécurité sociale sont l'AVC invalidant, les tumeurs malignes, le diabète de type 1 ou 2, la mucoviscidose ou encore la sclérose en plaques.

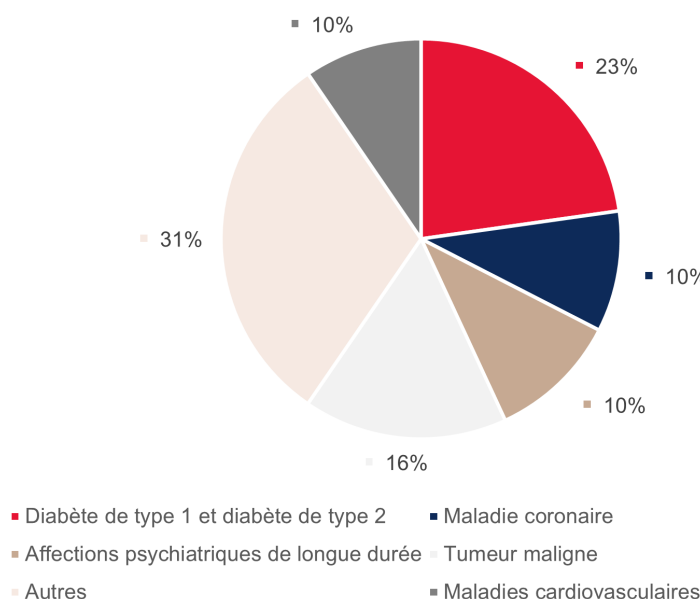


Figure 1.11 – Répartition des maladies comprises dans le cadre des ALD

Il existe 30 maladies différentes qui sont prises en charge en ALD. Même si les traitements sont remboursés à 100%, des frais annexes viennent s'y ajouter. En effet, les patients atteints d'ALD vont faire face à de nouvelles dépenses sur les frais de transport pour les consultations, des frais pour l'emploi d'une aide à domicile, des frais d'hébergement pour les soins éloignés du domicile, des séances de psychologue, diététicien, ergothérapeute, produits d'hygiène et les surcoûts liés à l'alimentation.

Afin de protéger ses employés de telles dépenses, une entreprise peut ajouter au contrat santé qu'elle propose la garantie « maladies redoutées ».

Dans la plupart des cas, elle agit sous trois angles différents. Premièrement, par la prévention en informant les employés comment leur mode de vie peut affecter les chances de contracter une maladie grave, puis financièrement, en allouant un capital lors du diagnostic de la maladie pour pallier les frais non remboursés par la sécurité sociale et enfin, en accompagnant les malades avec des consultations chez le psychologue, un accompagnateur pour expliquer les démarches et les traitements à suivre, ainsi que des consultations supplémentaires de médecins.

Cette garantie est très utile et rassurante pour les employés. En effet, ils savent que s'ils sont touchés par une maladie grave, ils auront le capital nécessaire pour accéder à de meilleurs soins et seront accompagnés dans leurs démarches. Certains assureurs proposent la garantie selon différentes formules, permettant de prendre plus ou moins de maladies en charge. De même, il est possible d'avoir plusieurs niveaux de garantie pour le montant du capital assuré. Les tarifs varient énormément en fonction du capital assuré et de l'âge. En plus de donner l'image forte d'une entreprise soucieuse du bien-être de ses employés, cette garantie permet, par la prévention, de réduire le nombre de chances d'avoir une maladie grave et elle aide à réduire l'absentéisme en aidant les salariés malades à réintégrer leur équipe.

Il existe cependant deux inconvénients. Cette garantie ne fonctionne que pour les maladies stipulées dans le contrat. Ainsi, un employé atteint de sclérose en plaques ayant la garantie "maladies redoutées" seulement pour le cancer ne se verra accorder aucun accompagnement ni capital par l'assureur.

De plus, on sait que l'assurance santé, par ses coûts souvent importants, peut coûter cher à certaines entreprises et à certains employés. Par conséquent, malgré ses qualités sociales, ce n'est pas forcément la garantie qui bénéficie de la plus grande publicité de la part des assureurs lorsqu'ils tarifient un contrat santé.

1.5.4 Antisélection et aléa moral

Antisélection

L'antisélection est un phénomène qui se produit en assurance lorsque les individus qui présentent un risque plus élevé sont plus enclins à souscrire une police d'assurance. En d'autres termes, ceux qui sont plus susceptibles de faire une réclamation sont surreprésentés parmi les assurés. Cela peut se produire lorsque les individus ont une meilleure connaissance de leur propre santé ou de leurs antécédents médicaux que l'assureur. En effet, une personne qui sait que sa famille a des antécédents de cancers va plus se renseigner sur comment se couvrir et va plus facilement souscrire à une garantie maladie redoutée.

L'antisélection peut entraîner des problèmes pour les assureurs, car elle peut déséquilibrer le portefeuille. Afin de gérer l'antisélection, les assureurs vont par exemple spécifier le type de maladies couvertes par la garantie maladie redoutée, donner un questionnaire médical à remplir pour évaluer le risque ainsi que mettre une franchise avant que la garantie soit effective.

Aléa moral

L'aléa moral est un concept dans les contrats d'assurance pour décrire le changement de comportement d'une personne une fois qu'elle est assurée. Plus précisément, une fois qu'une personne est couverte par une assurance, elle peut être tentée de prendre des risques ou d'adopter des comportements plus risqués qu'elle ne le ferait autrement, sachant qu'elle sera protégée financièrement en cas de problème. Il est possible qu'un assuré couvert par la garantie maladie redoutée soit moins effrayé par le fait de contracter une maladie grave et qu'ainsi il fasse moins attention à son comportement, qui, comme vu précédemment, a une forte influence sur la probabilité de contracter un cancer, par exemple.

L'aléa moral peut poser des défis aux assureurs en augmentant les coûts des réclamations et en affectant la rentabilité des polices d'assurance. Pour gérer l'aléa moral, les assureurs peuvent mettre en œuvre des mécanismes tels que des franchises, des limitations de couverture ou des incitations à adopter des comportements plus sûrs. Dans le cas de l'assurance maladie redoutée, l'assureur va généralement faire de la prévention auprès de ses assurés afin de les former sur les comportements à risque face aux maladies graves.

Chapter 2

Prédiction du taux d'incidence du cancer en France sur un horizon de 30 ans

L'objectif de cette partie est d'analyser en détail le taux d'incidence du cancer en France afin de calculer de manière adéquate la prime pure de la garantie maladie redoutée. Le taux d'incidence du cancer correspond au nombre de nouvelles personnes atteintes de cette maladie sur une période donnée. Dans notre étude, il est exprimé pour 100 000 personnes.

En France, depuis 1990, le taux d'incidence a presque doublé chez l'homme (+98%) et plus que doublé chez la femme (+104%). La question qui se pose est de savoir si cette progression va se poursuivre et, si oui, dans quelle mesure.

L'évolution de l'incidence du cancer dans le temps a un impact direct sur la rentabilité et donc sur le prix des produits d'assurance cancer. En effet, la garantie maladie redoutée est une assurance généralement souscrite sur le long terme, dont la tarification dépend fortement du taux d'incidence.

Ce chapitre se concentrera sur les différentes techniques permettant de prédire le taux d'incidence du cancer. Nous testerons trois méthodes et analyserons leur pertinence dans le cadre de notre étude.

2.1 Les données

2.1.1 L'INCa (Institut National du Cancer)

Créé le 9 août 2004 par la loi de santé publique, l'Institut National du Cancer (INCa) est un groupement d'intérêt public (GIP), c'est-à-dire qu'il est composé de plusieurs organismes qui peuvent aussi bien être publics que privés afin de réaliser leur mission : lutter contre le cancer en France.

L'INCa, afin de réaliser sa mission, va pouvoir agir sur différents piliers :

Coordonner les actions sur la lutte du cancer : partenariat avec les acteurs associatifs et institutionnels, collaboration avec les Agences régionales de santé (ARS) et au niveau international, pilotage de la Stratégie décennale de lutte contre les cancers

Initier et soutenir l'innovation : organisation d'appels à projets compétitifs annuels, lancement de projets conjoints avec les grandes associations, gestion des programmes hospitaliers et de recherche clinique en cancérologie

Structurer les organisations : organisation des dépistages, soins et recherche

Produire des expertises : recommandations nationales, rapports et avis, information destinée aux décideurs, professionnels de santé et population générale

Produire, analyser et évaluer des données : observation des cancers par études et analyses, diffusion de données en cancérologie

Favoriser l'appropriation des connaissances et des bonnes pratiques : outils de formation, campagnes d'information, présence en ligne via site internet et réseaux sociaux.

La loi du 8 mars 2019 a confié à l'INCa la mission d'établir une stratégie décennale de lutte contre le cancer en France. Cette initiative montre la volonté de renforcer la recherche, l'innovation, d'améliorer la prévention et les soins, et de garantir une coordination efficace entre tous les acteurs concernés. Grâce à cet organisme, il est possible de trouver la plupart des données sur le cancer en France sur un même site, notamment le taux d'incidence du cancer en France de 1990 à 2023. Ces données sont accessibles grâce aux registres des cancers en France et aux Hospices de Lyon qui ont publié leurs résultats sur l'INCa.

2.1.2 Les registres des cancers en France

Crée à partir des années 1975 sur des initiatives individuelles dans un double objectif de surveillance et de recherche, les registres des cancers jouent aujourd'hui un rôle essentiel dans la surveillance épidémiologique du cancer en France. Ils recueillent de manière continue et exhaustive les données provenant des hôpitaux sur les cas de cancer. Actuellement, il existe 32 registres évalués par le comité d'évaluation des registres (CER), chacun étant soumis à des contrôles de pertinence et de qualité des données. Il existe deux types de registre. Les registres généraux qui recensent les données sur tous les cancers en général et les registres spécialisés qui vont recenser les données sur des populations spéciales ou pour des types de cancers spéciaux.

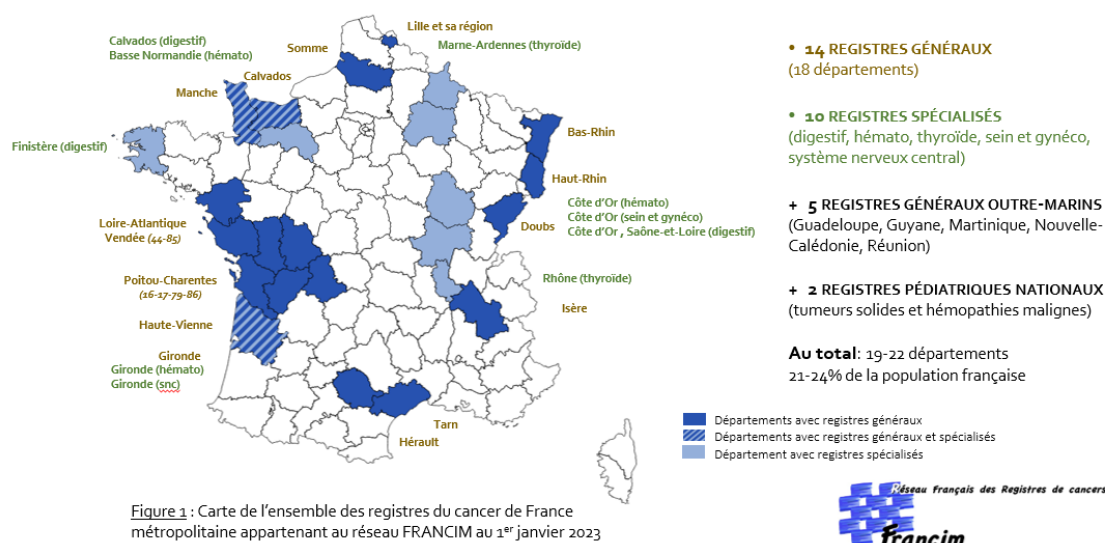


Figure 1 : Carte de l'ensemble des registres du cancer de France métropolitaine appartenant au réseau FRANCIM au 1^{er} janvier 2023

Figure 2.1 – Les différents registres en France

Ces registres sont regroupés sous le réseau FRANCIM. Son objectif principal est d'harmoniser les méthodes d'enregistrement et de codification des cancers, ainsi que de coordonner et publier les travaux réalisés sur les données de ces registres. Ce dispositif de surveillance repose sur un partenariat scientifique et financier impliquant notamment l'Institut national du cancer (INCa), le Service de biostatistique-bioinformatique des Hospices civils de Lyon (HCL), et Santé publique France.

Grâce à cette coordination, les registres des cancers, tant généraux que spécialisés, permettent de retrouver le taux d'incidence du cancer en France selon l'âge, le sexe, l'année et le type de cancer. Pour accéder à des données plus spécifiques, il est nécessaire de démontrer la motivation de son projet d'analyse et de son intérêt pour le bien public, assurant ainsi une utilisation responsable et éthique de ces informations.

2.1.3 Présentation des données

Les données que nous utiliserons proviennent de l'étude intitulée "Incidence des principaux cancers en France métropolitaine en 2023 et tendances depuis 1990", publiée dans le Bulletin Épidémiologique Hebdomadaire (BEH) en 2023. Basées sur les données réelles des registres couvrant environ 20% de la population, ces données nous donnent le taux d'incidence du cancer en France selon 19 types de cancers différents ainsi que pour l'incidence globale. Elles sont disponibles par sexe et par classe d'âge quinquennale.

Par souci de visualisation, les prochains graphiques pourront être présentés sous forme d'autres catégories.

Les 3 cancers les plus fréquents

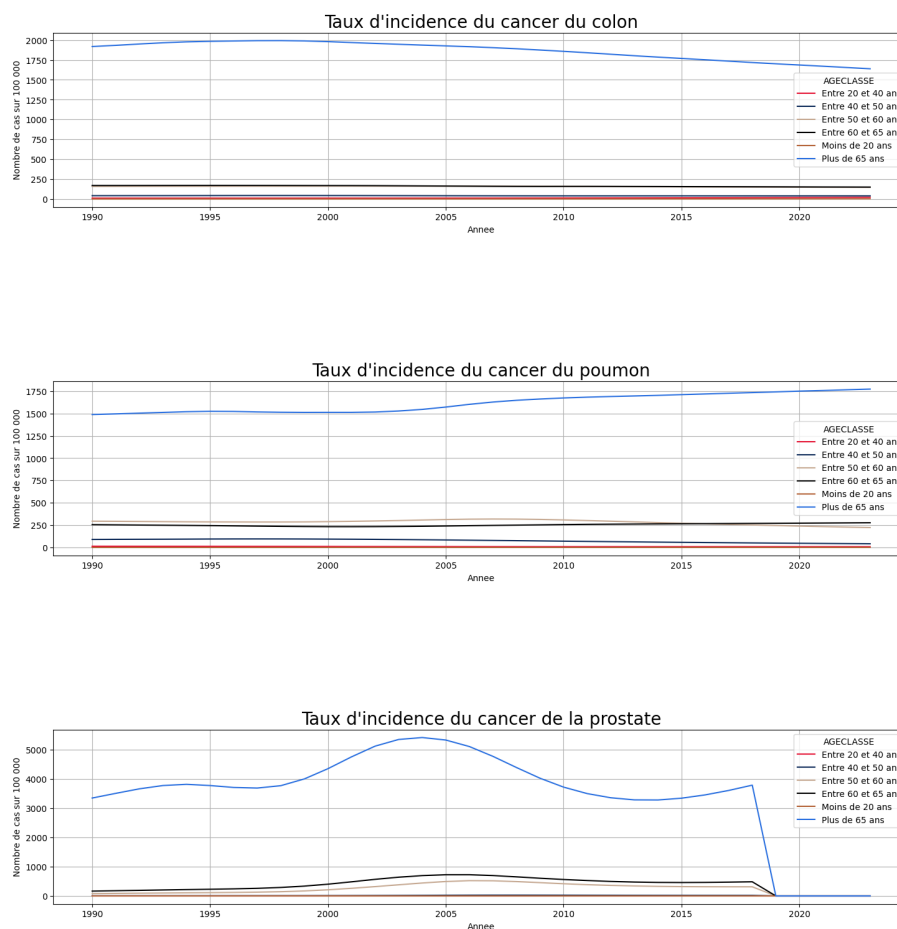


Figure 2.2 – Taux d'incidence pour les 3 cancers les plus fréquents chez l'homme

Pour le cancer du côlon chez l'homme, on observe une baisse de l'incidence, ce qui est attribuable à l'augmentation des dépistages et à l'élimination des polypes avant qu'ils ne deviennent cancéreux. Les innovations en matière de dépistage montrent ainsi leur efficacité pour ce type de cancer. Cependant, cela ne s'applique pas à tous les cancers, comme nous pouvons le constater pour ceux du poumon et de la prostate. Le cancer du poumon a vu son incidence augmenter de manière quasi linéaire depuis 1990.

Pour le cancer de la prostate, l'analyse de l'incidence est plus complexe. Le dépistage de ce cancer repose sur le taux de PSA, une technique introduite dans les années 1990, ce qui explique le pic d'incidence jusqu'en 2005. Cependant, vers les années 2000, il a été démontré que ce dépistage n'était pas toujours fiable et pouvait conduire à des erreurs. Cela a entraîné une diminution des dépistages entre 2005 et 2015, les Français se faisant moins souvent dépister. L'augmentation observée à partir de 2005 pourrait être due à des changements de comportements des Français. La courbe d'incidence s'arrête en 2018 en raison d'une évolution importante de ce cancer. Par souci d'incertitude sur leur évolution à court terme, les projections d'incidence pour ces cancers ne sont pas disponibles pour les années 2019 à 2023.

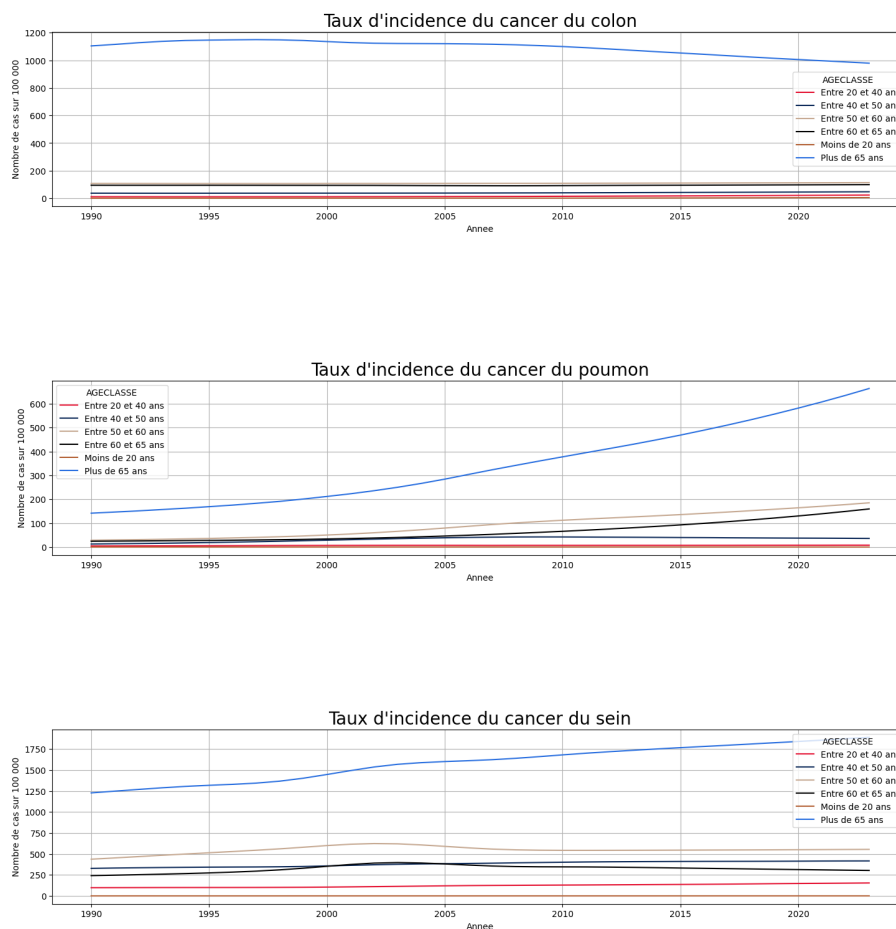


Figure 2.3 – Taux d'incidence pour les 3 cancers les plus fréquent chez la femme

Chez la femme, on observe une tendance similaire pour le cancer du côlon, avec une diminution de l'incidence grâce à l'augmentation des dépistages, comme chez l'homme. Cependant, pour les cancers du poumon et du sein, l'augmentation est significative depuis 1990, en particulier pour le cancer du poumon dont l'incidence a plus que triplé. Cette hausse est due aux comportements des femmes, qui fument et boivent plus qu'elles ne le faisaient dans les années 1990.

Analyse de l'incidence global

Avant de lancer nos modèles, il est intéressant d'analyser le comportement de l'incidence de 1990 à 2023 chez les hommes et les femmes. Cette première analyse nous permettra de mieux comprendre nos données.

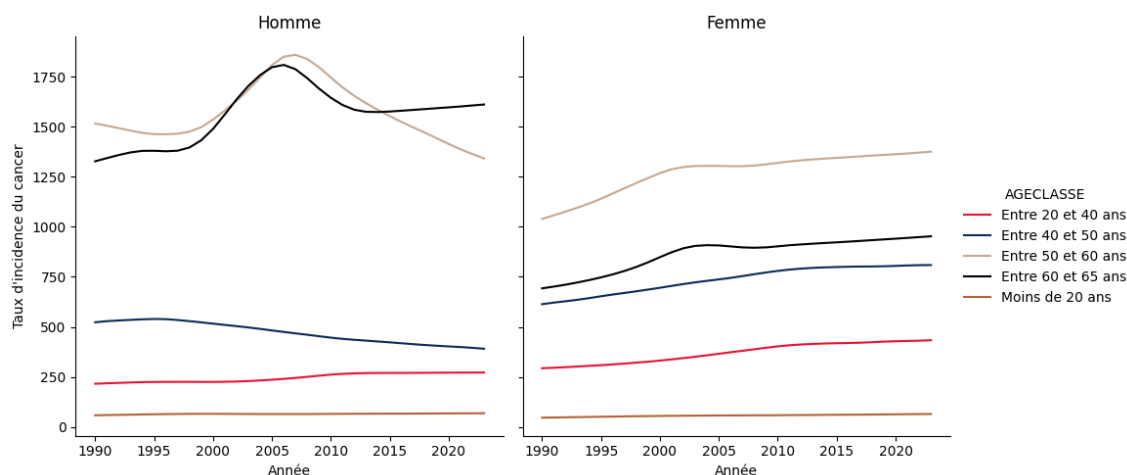


Figure 2.4 – Comparaison de l'incidence selon l'âge et le sexe

On observe chez les hommes un pic notable dans l'incidence des cancers entre les années 2005 et 2010, particulièrement pour les personnes âgées de 50 à 65 ans. Cette augmentation significative est en grande partie attribuée aux campagnes de dépistage qui ont été intensifiées au cours de cette période. Dans les années 1990, des tests pilotes ont été menés pour évaluer l'efficacité du dépistage systématique. Ces tests ont montré des résultats prometteurs, conduisant à une adoption plus large et à une mise en œuvre à l'échelle nationale au début des années 2000. En conséquence, le nombre de dépistages a considérablement augmenté.

Cette augmentation du nombre de tests a conduit à la détection de davantage de cas de cancer, expliquant ainsi l'augmentation brutale des diagnostics observée. L'objectif principal des campagnes de dépistage est de détecter les cancers le plus tôt possible. Une détection précoce permet d'améliorer les options de traitement et les chances de guérison, ce qui est crucial pour la gestion de cette maladie.

Il est important de noter que ce pic dans les statistiques du cancer ne reflète pas nécessairement une augmentation du risque réel de contracter la maladie, mais plutôt une augmentation des cas détectés grâce à un dépistage plus étendu. Après l'intensification des campagnes de dépistage, le taux d'incidence a tendance à se stabiliser et à suivre une évolution plus linéaire. Cette stabilisation peut être attribuée à un changement de comportement parmi les Français, qui est l'une des variables principales influençant le taux d'incidence du cancer. Les efforts de sensibilisation et les changements dans les habitudes de vie contribuent à cette tendance.

Chez les femmes, un phénomène similaire est observé avec un pic d'incidence dans la même tranche d'âge, bien que ce pic soit moins prononcé que chez les hommes. Cette différence pourrait s'expliquer par des variations dans les comportements et les habitudes de vie des femmes par rapport aux hommes depuis les années 1990. Par exemple, il a été observé que les femmes ont augmenté leur consommation d'alcool et de tabac, ce qui est un facteur contribuant à une augmentation de l'incidence du cancer. Ainsi, chez les femmes, l'augmentation des cas détectés peut être due à la fois à une incidence réelle accrue de la maladie et à une augmentation des dépistages.

En conclusion, les pics d'incidence du cancer observés chez les hommes et les femmes autour des années 2005 peuvent être largement expliqués par l'augmentation des dépistages. Toutefois, les comportements individuels, tels que la consommation d'alcool et de tabac, jouent également un rôle significatif dans les tendances observées, particulièrement chez les femmes. La stabilisation des taux d'incidence après l'intensification des campagnes de dépistage souligne l'importance de la détection précoce et des changements de comportement dans la lutte contre le cancer.

Format de nos données pour nos modèles

Pour les différentes méthodes que nous utiliserons dans la prochaine partie, nous avons décidé de conserver les données de base dans leur intégralité. En effet, l'âge a une influence significative sur le taux d'incidence du cancer, et il serait incorrect de réduire le nombre de classes d'âge. Une telle réduction pourrait biaiser nos analyses et modèles en ne reflétant pas adéquatement les variations de risque associées à chaque tranche d'âge.

Pour l'application de nos modèles, nous utiliserons les données sur le taux d'incidence global et non sur chaque cancer. Cela est plus intéressant, car pour certains cancers, les données que nous avons ne sont pas complètes. De plus, la garantie "maladie redoutée" ne distingue pas les types spécifiques de cancer. Elle couvre tous les types de cancers. Ainsi, dès qu'un assuré est diagnostiqué avec n'importe quel type de cancer, il reçoit le capital prévu par son assurance.

L'utilisation des données globales sans transformation de la base sur le taux d'incidence nous permettra de modéliser de manière plus fidèle et rigoureuse les risques associés à la garantie maladie redoutée. Réduire le nombre de classes d'âge pourrait nuire à la précision de nos résultats. S'il avait été possible d'avoir ces taux par âge, cela aurait été encore mieux.

2.2 Modèle ARIMA

2.2.1 Choix de la méthode

La méthode ARIMA (AutoRegressive Integrated Moving Average) a été formalisée et popularisée dans les années 1970 par les statisticiens George Box et Gwilym Jenkins, dans leur ouvrage de référence *Time Series Analysis: Forecasting and Control* (1970). Ce livre a établi la base théorique et pratique de l'ARIMA, une méthode utilisée pour modéliser et prédire les séries chronologiques.

Avant cet ouvrage, des concepts tels que l'autocorrélation et les modèles de moyenne mobile existaient déjà, mais Box et Jenkins ont standardisé leur application dans une approche unifiée. L'objectif initial de la méthode était de fournir aux chercheurs un cadre analytique pour traiter les données temporelles présentant des tendances ou des variations non stationnaires.

Cette méthode s'est très vite répandue dans de nombreux domaines par son efficacité : dans la finance pour prédire les taux de change, dans certaines chaînes de magasins pour gérer les stocks en fonction de la période de l'année, ou bien même dans la santé, comme en 2011 dans l'étude de Chaves et al., qui a utilisé un modèle ARIMA pour prévoir les épidémies de dengue au Brésil.

C'est ainsi que notre choix s'est porté sur cette méthode pour prédire le taux d'incidence en France.

2.2.2 Présentation de la méthode ARIMA

Le premier modèle que nous testerons afin de prédire le taux d'incidence en France est la méthode ARIMA (Autoregressive Integrated Moving Average). C'est un modèle statistique utilisé pour la prévision de séries temporelles. Il se base sur la décomposition d'une série en trois composantes :

- Une composante autorégressive (AR)
- Une composante de différenciation (I)
- Une composante de moyenne mobile (MA)

Ces trois composantes influencent la structure du modèle et le nombre de paramètres à estimer. Lorsqu'on parle d'un modèle ARIMA, il est souvent noté sous la forme $ARIMA(p, d, q)$, où p représente l'ordre de l'autorégression, d l'ordre de différenciation, et q l'ordre de la moyenne mobile. L'objectif est de déterminer les paramètres de manière à ce que le modèle s'ajuste au mieux aux données. Cependant, avant d'entrer dans les détails, nous allons expliquer à quoi correspond chacune de ces composantes.

La différentiation (d)

Le but de la différentiation va être de rendre la série stationnaire, que la moyenne et la variance soient constantes dans le temps.

$$\mathbb{E}[y_t] = \mu \quad \text{et} \quad \text{Var}(y_t) = \sigma^2$$

On définit l'ordre d'intégration comme le nombre de différences nécessaires pour transformer une série non stationnaire en une série stationnaire. À titre d'exemple, un modèle ARIMA(0,1,0), de premier ordre, se représente de la manière suivante :

$$y_t - y_{t-1} = \mu + \epsilon_t$$

où μ est la constante du modèle et ϵ_t le terme d'erreur.

Les modèles d'ordre supérieur représentent des différences de différences. Par exemple, pour un modèle ARIMA(0,2,0), cela correspond à $(y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$. Ainsi, le modèle s'écrira :

$$y_t = \mu + 2y_{t-1} - y_{t-2} + \epsilon_t$$

De manière générale, on utilisera l'opérateur de retard pour définir la différentiation d'ordre d . L'opérateur de retard est défini de la manière suivante :

$$By_t = y_{t-1}$$

Ainsi, dans le cas général, la différentiation d'ordre d s'écrira :

$$\Delta^d X_t = (1 - B)^d X_t$$

Une fois que l'on a réussi à rendre la série stationnaire, on peut passer à l'étape suivante et analyser la composante autorégressive.

L'auto-régression (p)

Une série est dite auto-régressive si son terme à la date t est linéairement dépendant des termes précédents. Prenons une série stationnaire auto-régressive d'ordre 2 alors :

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t$$

avec ϕ_1 et ϕ_2 qui correspondent aux coefficients d'auto-régression.

Mathématiquement, on écrira un processus auto-régressif d'ordre p ainsi :

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$$

La dernière étape dans le modèle ARIMA est la moyenne mobile.

La moyenne mobile (q)

Lorsque l'on parle d'un système à moyenne mobile, on sous-entend que la série fluctue autour d'une valeur moyenne. Mathématiquement, on va dire que la série est égale à cette moyenne plus les fluctuations qui seront représentées par la pondération des termes erreurs des précédents termes. Un modèle d'ordre 2 sera représenté de la manière suivante :

$$y_t = \mu - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \epsilon_t$$

Dans le cas général, on obtient :

$$y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$$

Le modèle ARIMA est la combinaison des trois caractéristiques que nous venons de voir. Ainsi, lorsque l'on écrit le modèle dans le cas général, on obtient l'équation suivante pour un ARIMA(p,d,q) :

$$(1 - B)^d Y_t = \mu + \sum_{k=1}^p \phi_{t-k} (1 - B)^d Y_{t-k} - \sum_{k=1}^q \theta_{t-k} \epsilon_{t-k}$$

où :

- Y_t est la variable à expliquer à l'instant t .
- d est l'ordre de différenciation pour rendre la série temporelle stationnaire.
- B est l'opérateur de retard.
- ϕ_{t-k} et θ_{t-k} sont les coefficients représentant les valeurs passées pondérées de Y_t et du terme d'erreur respectivement.
- ϵ_{t-q} sont les termes d'erreur.

Toute l'analyse repose sur le fait de trouver les bons paramètres (p, d, q) de notre modèle. Plusieurs techniques existent. Cependant, il est possible d'utiliser sous Python la fonction `auto_arima` du module `pmdarima`, qui permet de nous trouver le modèle optimal à partir de notre jeu de données, basé sur la méthode d'erreur de notre choix.

2.3 Méthode Lee-Carter

2.3.1 Choix de la méthode

La méthode Lee-Carter est une approche statistique développée en 1992 par Ronald Lee et Lawrence Carter pour modéliser et prévoir les tendances de mortalité à long terme. Leurs travaux, publiés dans l'article *Modeling and Forecasting U.S. Mortality* (Journal of the American Statistical Association), visaient à répondre à la difficulté croissante de prévoir les taux de mortalité, particulièrement dans le contexte du vieillissement de la population. Son utilisation s'est ensuite développée dans le monde de l'assurance, où elle est devenue un outil très pratique pour évaluer la mortalité.

L'analogie entre la mortalité et l'incidence du cancer peut se faire rapidement d'un point de vue mathématique. Ainsi, nous cherchons à vérifier si son applicabilité dans le domaine de la mortalité peut être transférée à la prédiction de l'incidence du cancer.

2.3.2 Présentation de la méthode Lee-Carter

Le modèle de Lee-Carter est une méthode d'extrapolation des tendances passées, devenue rapidement un standard pour la modélisation du taux instantané de mortalité. Il est formulé comme suit :

$$\log \mu_{xt} = \alpha_x + \beta_x \kappa_t + \epsilon_{xt}$$

où μ_{xt} représente le taux instantané de mortalité pour l'âge x et l'année t , α_x est l'effet moyen par âge, β_x est la sensibilité relative au changement à chaque âge, κ_t est un effet temporel commun à tous les âges, et ϵ_{xt} est le terme d'erreur aléatoire.

Le paramètre α_x capture le profil de mortalité moyen par âge, β_x mesure la sensibilité de la mortalité à l'âge x face aux variations temporelles générales représentées par κ_t , et κ_t capte l'évolution globale de la mortalité au fil du temps. On vérifie que $\frac{\partial \log(\mu_{xt})}{\partial \kappa_t} = \beta_x$, indiquant que β_x mesure la sensibilité de la mortalité instantanée à l'âge x par rapport à l'évolution générale κ_t .

Pour rendre le modèle identifiable, des contraintes sont ajoutées aux paramètres, notamment $\sum_x \beta_x = 1$ pour fixer l'échelle de κ_t , et $\sum_t \kappa_t = 0$ pour centrer la tendance temporelle et éviter des solutions arbitraires. Les paramètres sont ensuite estimés en minimisant un critère de moindres carrés non linéaire :

$$\arg \min_{\alpha_x, \beta_x, \kappa_t} \sum_{x,t} (\log(\mu_{xt}) - (\alpha_x + \beta_x \kappa_t))^2$$

Bien que ce modèle soit couramment utilisé pour prédire le taux instantané de mortalité, il peut également être appliqué pour prédire le taux d'incidence du cancer.

En effet, ces deux taux dépendent de l'âge et du temps. Dans ce cas, une matrice qui, pour chaque catégorie d'âge et chaque année, donne le taux d'incidence, est utilisée. Le taux de mortalité instantané μ_{xt} est alors remplacé par le taux d'incidence $m_{x,t}$ pour la catégorie d'âge x et l'année t . Les caractéristiques des paramètres du modèle seront analysées de manière similaire pour le taux d'incidence.

Avant d'expliquer les étapes que nous utiliserons pour implémenter la méthode Lee-Carter, nous allons définir ce qu'est la décomposition en valeurs singulières (SVD). C'est une technique mathématique permettant de décomposer une matrice $\mathbf{A}$ de dimensions $m \times n$ en trois matrices distinctes. La décomposition s'exprime comme suit :

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^*$$

où :

- $\mathbf{U}$ est une matrice unitaire de taille $m \times m$, dont les colonnes sont les vecteurs propres de $\mathbf{A}\mathbf{A}^*$.
- $\mathbf{S}$ est une matrice diagonale de taille $m \times n$ contenant les valeurs singulières de $\mathbf{A}$. Les valeurs sur la diagonale sont toujours non négatives et indiquent l'importance relative des différentes composantes de $\mathbf{A}$, capturant la variance des données.
- $\mathbf{V}^*$ est la transposée conjuguée de $\mathbf{V}$, une matrice unitaire de taille $n \times n$, dont les colonnes sont les vecteurs propres de $\mathbf{A}^*\mathbf{A}$.

Dans notre modèle de Lee-Carter, nous utilisons principalement la première valeur singulière de $\mathbf{S}$, qui représente la composante principale, capturant ainsi les principales variations du taux d'incidence au fil du temps.

Maintenant que nous savons ce qu'est la SVD, voici les étapes que nous suivrons pour obtenir les paramètres de notre modèle :

Étape 1: Calculer $\mathbf{a}_x$

Calculer $\mathbf{a}_x$, qui est la moyenne dans le temps de $\ln(\mathbf{m}_{x,t})$ pour chaque âge :

$$\mathbf{a}_x = \frac{\sum_{t=1}^T \ln(\mathbf{m}_{x,t})}{T}$$

Étape 2: Calculer $\mathbf{A}_{x,t}$

Calculer $\mathbf{A}_{x,t}$, qui sera utilisée dans la décomposition en valeurs singulières (SVD) :

$$\mathbf{A}_{x,t} = \ln(\mathbf{m}_{x,t}) - \mathbf{a}_x$$

Étape 3: Effectuer la SVD de $\mathbf{A}_{x,t}$

Effectuer la SVD de $\mathbf{A}_{x,t}$:

$$\mathbf{U}\mathbf{S}\mathbf{V}^* = \text{svd}(\mathbf{A}_{x,t})$$

Étape 4: Trouver $\mathbf{k}_t$, s_1 et $\mathbf{b}_x$

Trouver $\mathbf{k}_t$, s_1 et $\mathbf{b}_x$ depuis $\mathbf{U}$, $\mathbf{S}$ et $\mathbf{V}^*$:

$$\mathbf{k}_t = (u_{1,1}, u_{2,1}, \dots, u_{t,1})$$

$$\mathbf{b}_x = (v_{1,1}, v_{1,2}, \dots, v_{1,x})$$

Étape 5: Prédire $\mathbf{k}_t$ pour n prochaines années

Prédire $\mathbf{k}_t$ pour n prochaines années. Dans notre modèle nous utiliserons le random-walk-with-drift qui suppose une tendance constante.

Étape 6: Utiliser $\mathbf{k}_{t+n}$, $\mathbf{b}_x$ et $\mathbf{a}_x$

Utiliser $\mathbf{k}_{t+n}$ ainsi que $\mathbf{b}_x$ et $\mathbf{a}_x$ pour obtenir nos prédictions :

$$\mathbf{m}_{x,t+n} = \exp(\mathbf{a}_x + s_1 \mathbf{k}_{t+n} \mathbf{b}_x)$$

2.4 Modèle RVFL

2.4.1 Choix de la méthode

La méthode RVFL (Random Vector Functional Link) est une variante des réseaux de neurones développée pour améliorer l'efficacité des réseaux de neurones traditionnels. Elle a été initialement introduite dans les années 1990 par Pao Y.-H. dans son ouvrage *Adaptive Pattern Recognition and Neural Networks* (1992). Contrairement aux réseaux de neurones traditionnels, la méthode RVFL se distingue par une architecture simplifiée qui accélère l'apprentissage tout en maintenant de bonnes performances.

La méthode RVFL a été appliquée dans la finance pour prédire des séries temporelles comme les prix des actions. En 2019, une étude de Wang et al. a montré que l'utilisation du RVFL pour prédire les tendances des prix boursiers surpassait les modèles ARIMA traditionnels en termes de précision et de rapidité, grâce à sa capacité à capter des relations complexes non linéaires.

Dans notre étude, cherchant à tester des méthodes intégrant les réseaux de neurones, la méthode RVFL s'est révélée intéressante pour la prédiction de l'incidence du cancer en France. Sachant qu'elle a parfois surpassé ARIMA il sera intéressant dans notre cas de les comparer.

2.4.2 Présentation de la méthode RVFL

Dans cette section, nous abordons la formulation du modèle RVFL standard. Soit $X = [x_1, x_2, \dots, x_N]^T$, où $x_i \in \mathbb{R}^d$ représente l'ensemble de données d'entraînement, et $Y = [y_1, y_2, \dots, y_N]^T$, où $y_i \in \mathbb{R}^c$ est la matrice des cibles. Dans notre analyse, N correspond au nombre de classes d'âge dont nous disposons, d représente le nombre d'années dans notre échantillon d'entraînement, et c représente le nombre d'années dans nos données de test.

Le modèle RVFL est une version aléatoire d'un réseau de neurones à une seule couche cachée, avec trois couches appelées couche d'entrée, couche cachée et couche de sortie. Toutes les trois couches sont constituées de neurones connectés par des poids. Pour éviter l'implémentation de l'algorithme de rétropropagation, les poids de la couche d'entrée à la couche cachée sont générés aléatoirement et restent fixes pendant le processus d'entraînement. Seuls les poids de sortie sont calculés analytiquement par la méthode des moindres carrés.

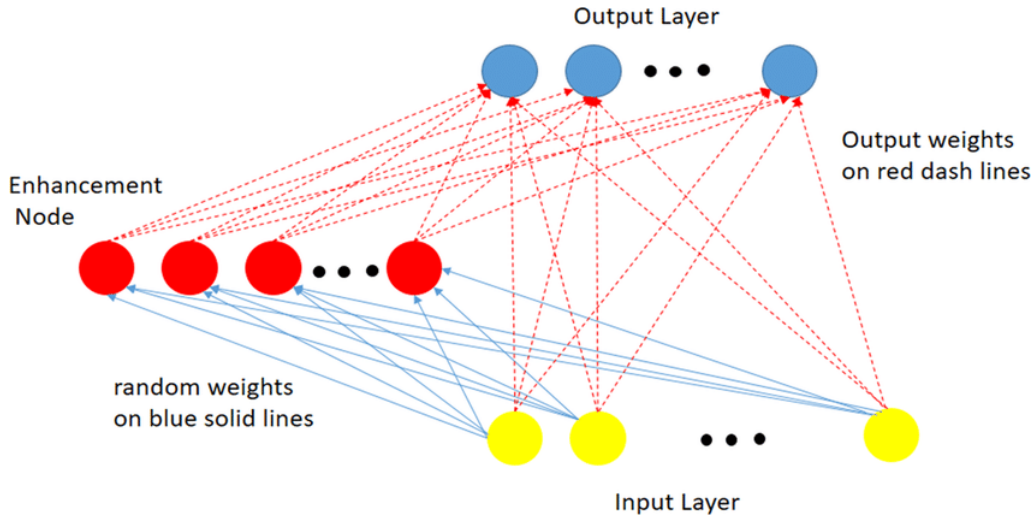


Figure 2.5 – Schéma représentant le fonctionnement d'un réseaux RVFL

Dans ce modèle, les caractéristiques originales sont également utilisées pour relier les couches d'entrée et de sortie. Les liens directs améliorent la performance de généralisation du RVFL et évitent l'overfitting.

Mathématiquement, le modèle RVFL peut être écrit comme suit avec $f : \mathbb{R}^d \rightarrow \mathbb{R}^c$:

$$f(x_i) = \sum_{k=1}^d \beta_k x_{ik} + \sum_{k=d+1}^{L+d} \beta_k \theta(\langle \mu_k, x_i \rangle + \sigma_k), \quad i = 1, 2, \dots, N$$

où :

- β_k sont les poids de sortie
- x_{ik} sont les caractéristiques d'entrée
- θ est la fonction d'activation
- $\langle \mu_k, x_i \rangle$ représente le produit scalaire entre les vecteurs μ_k et x_i
- σ_k sont les biais des neurones cachés
- L est le nombre total de neurones dans la couche cachée

Dans le cadre de notre modélisation, nous avons d'abord choisi de normaliser nos données pour assurer une meilleure convergence du modèle.

Concernant la fonction d'activation, nous avons décidé d'utiliser la fonction ReLU (Rectified Linear Unit), qui est largement utilisée dans les réseaux de neurones en raison de sa simplicité et de son efficacité. Elle est définie par la formule suivante :

$$\text{ReLU}(x) = \max(0, x)$$

2.5 Résultats

2.5.1 Méthodologie

Afin d'évaluer la qualité de nos modèles, nous allons utiliser une méthode de validation croisée. L'idée est de séparer nos données en un ensemble d'apprentissage (train) et un ensemble de validation (test). Le modèle est ensuite entraîné sur l'ensemble d'apprentissage, puis nous évaluons la qualité en comparant ses prédictions à l'ensemble de validation. Cela permet de simuler un véritable scénario de prévision où l'on utilise les données passées pour prédire les données futures. Cette méthode est utile pour évaluer la stabilité de notre modèle et, surtout, pour le comparer avec d'autres modèles.

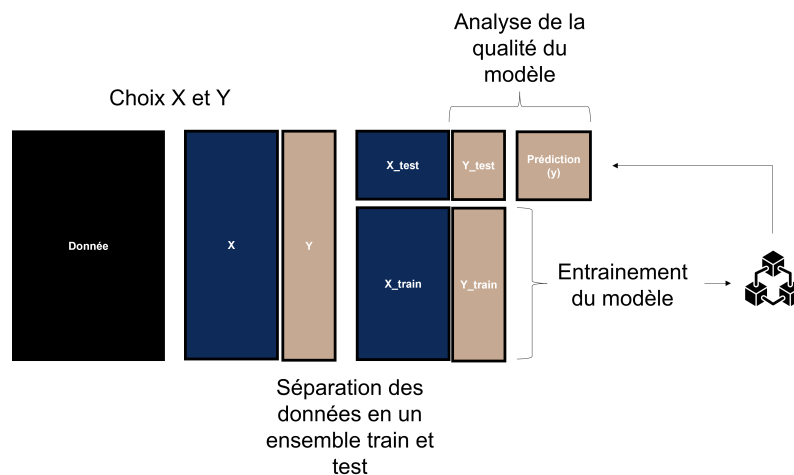
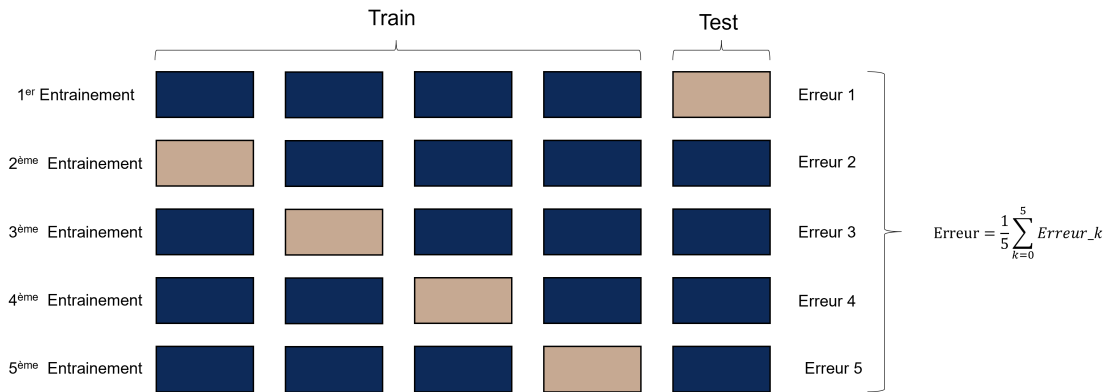


Figure 2.6 – Schéma représentant la méthodologie utilisée

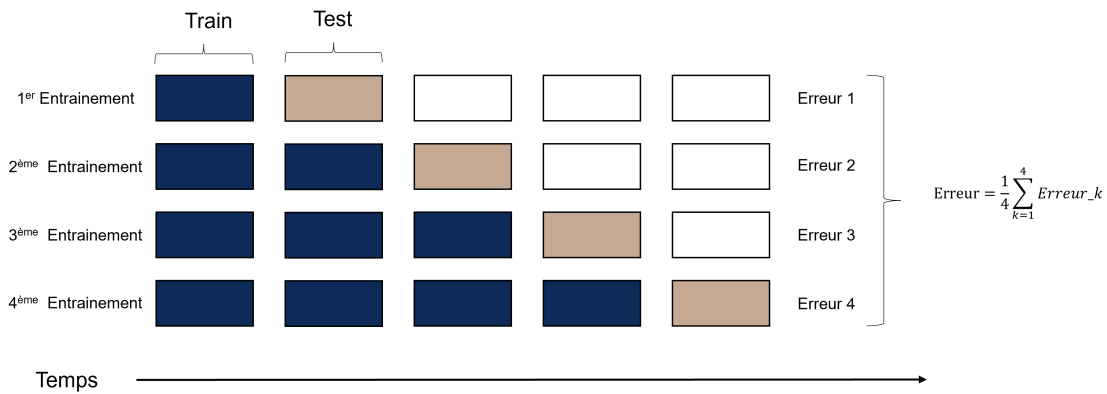
Différentes méthodes de validation croisée existent. Les plus répandues sont la méthode des k -folds et le train-test split. L'idée du k -fold est de séparer les données en k plis (folds), puis d'entraîner le modèle sur $k-1$ plis et de tester l'erreur de prédiction sur le dernier pli. Cela permet d'obtenir k modèles et de calculer une moyenne des erreurs de prédiction.

Cependant, la séparation en plis se fait aléatoirement, ce qui ne convient pas pour des séries temporelles. En effet, contrairement à des données classiques où les observations sont indépendantes et identiquement distribuées, les observations dans une série temporelle sont corrélées. Par exemple, l'observation x_t au temps t est très probablement corrélée avec l'observation x_{t+1} au temps suivant. Ainsi, la cross-validation k -folds est souvent écartée pour les séries temporelles. Pour la remplacer, on utilise la méthode déroulante.

Dans cette méthode, on débute avec un petit sous-ensemble de données servant d'ensemble d'entraînement. On utilise cet ensemble pour faire des prédictions sur les données ultérieures et on vérifie l'exactitude de ces prédictions. Les points de données correctement prévus sont ensuite ajoutés à l'ensemble d'entraînement, puis



on répète le processus pour prévoir les points de données suivants.



Cette méthode est très efficace; cependant, nous ne disposons pas d'un grand nombre de points de données, et la méthode déroulante pourrait mener à des ensembles de test trop petits, rendant l'estimation de l'erreur de notre modèle moins pertinente. Par conséquent, une division simple en utilisant les premières valeurs pour l'apprentissage et les dernières pour la validation reste la meilleure approche dans notre contexte. Nous opterons pour 80% des données pour l'apprentissage et les 20% restants pour la validation. Cette séparation, bien que plus basique, permet une évaluation plus appropriée à nos données.

2.5.2 Indicateurs

Supposons que notre ensemble de données soit noté $y_1, \dots, y_T$, et que nous le divisons en deux sections : les données d'apprentissage ($y_1, \dots, y_N$) et les données de test ($y_{N+1}, \dots, y_T$).

Pour vérifier la précision de notre méthode de prévision, nous estimerons les paramètres en utilisant les données d'apprentissage et préviendrons les données de test comme vu précédemment. Ces prévisions pourront ensuite être comparées aux données de test.

Nous utiliserons la notation suivante : la prévision h -pas en avant peut être écrite comme $\hat{y}_{N+h|N}$. La notation avec un "chapeau" indique qu'il s'agit d'une estimation plutôt que d'une valeur observée, et l'indice indique que nous estimons y_{N+h} en utilisant toutes les données observées jusqu'au temps N inclus.

Les erreurs de prévision sont la différence entre les valeurs réelles dans le jeu de test et les prévisions produites en utilisant uniquement les données du jeu d'apprentissage :

$$e_t = y_t - \hat{y}_{t|N} \quad \text{pour } t = N + 1, \dots, T$$

La mesure de l'efficacité de notre modèle ne peut se baser uniquement sur la valeur de ces e_t . Nous allons donc utiliser des métriques pour comparer la qualité des modèles à l'aide des erreurs. Les métriques les plus répandues lors d'une régression sont : l'erreur absolue moyenne (MAE : Mean Absolute Error), l'erreur quadratique moyenne (RMSE : Root Mean Squared Error) et l'erreur absolue moyenne en pourcentage (MAPE : Mean Absolute Percentage Error). Lors de la comparaison des méthodes de prévision sur un seul ensemble de données, la MAE est populaire car elle est facile à comprendre et à calculer. Cependant, dans notre modèle, nous choisirons d'utiliser le MAPE et le RMSE. Les taux d'incidence variant considérablement selon l'âge, le MAPE nous permet de mieux comparer nos modèles par catégorie d'âge, car il ne dépend pas de l'échelle des valeurs prévues. Le RMSE sera intéressant lorsque l'on compare des modèles au sein d'une même catégorie d'âge.

Erreur quadratique moyenne (RMSE)

Le carré moyen des erreurs (MSE pour Mean Square Error) est la moyenne des carrés des écarts entre les prévisions du modèle et les valeurs observées.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (e_i)^2$$

L'erreur quadratique moyenne (RMSE) est simplement la racine carrée du MSE, offrant une mesure sur la même échelle que les données initiales, facilitant ainsi l'interprétation des écarts.

$$\text{RMSE} = \sqrt{\text{MSE}}$$

Erreur absolue moyenne en pourcentage (MAPE)

L'erreur absolue moyenne en pourcentage (MAPE pour Mean Absolute Percentage Error) est une mesure de précision qui calcule la moyenne des erreurs absolues en pourcentage par rapport aux valeurs observées. Cette méthode est particulièrement utile pour comparer les performances de prévision entre différents ensembles de données, car elle est indépendante de l'échelle. Le MAPE est avantageux car il normalise les erreurs, permettant ainsi des comparaisons directes entre séries de données de tailles et d'unités différentes. La formule du MAPE est la suivante :

$$\text{MAPE} = 100 * \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{y_t} \right|$$

2.5.3 Choix du modèle

En raison de l'impossibilité de tarifier par sexe en France, nous avons opté pour l'entraînement de nos modèles en utilisant le taux d'incidence pondéré en fonction de la répartition des hommes et des femmes dans la population. Bien qu'une approche consistant à créer deux modèles distincts, un pour chaque sexe, aurait été envisageable, nous avons jugé que la moyenne des taux d'incidence avant l'entraînement permettait de réduire le pic observé en 2005, qui était plus élevé chez les hommes que chez les femmes.

Après avoir entraîné nos modèles à partir des données de 1990 à 2018, nous avons évalué leur précision en les comparant aux données de 2018 à 2023.

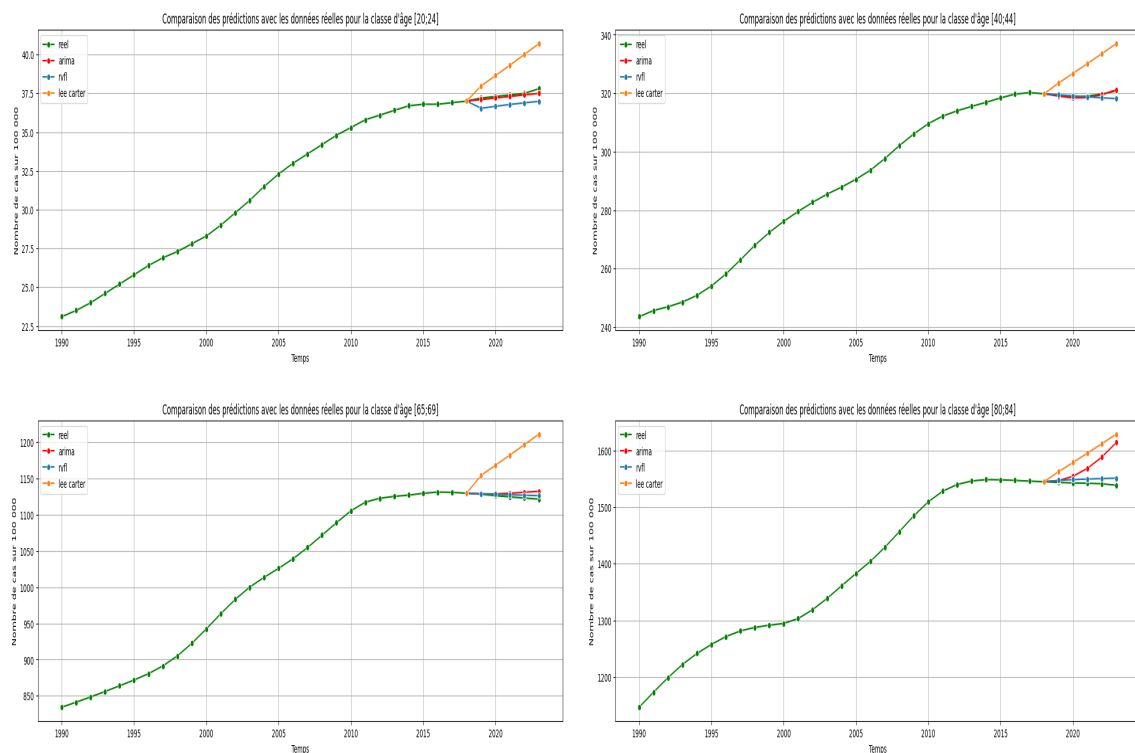


Figure 2.9 – Comparaison des méthodes utilisées par rapport aux données réelles

On remarque que le modèle Lee-Carter est souvent le moins précis parmi les méthodes que nous avons testées. Comme nous l'avons vu dans la partie dédiée à la présentation des données, il y a eu un pic d'incidence dans les années 2005, principalement dû à une augmentation du nombre de dépistages. Le modèle Lee-Carter, dont le but est de capturer la tendance des valeurs passées et de les projeter dans le futur, est influencé par ce pic, ce qui peut diminuer son efficacité.

En ce qui concerne les modèles RVFL et ARIMA, il est difficile de déterminer a priori lequel est le meilleur sans une analyse plus approfondie. Afin d'affiner notre choix de modèle, nous allons utiliser différents indicateurs de performance que nous avons présentés précédemment.

Après avoir lancé nos modèles, nous obtenons les valeurs de Mean Absolute Percentage Error (MAPE) pour chaque catégorie d'âge et en fonction de la méthode utilisée. Ces indicateurs nous permettront de comparer quantitativement les performances des modèles et de choisir celui qui offre les prévisions les plus précises.

Comme le premier graphique l'indiquait, le modèle Lee-Carter est de loin le moins efficace, surtout pour les tranches d'âge de 45 à 65 ans. C'est précisément dans cette tranche d'âge que l'on observe le pic le plus important.

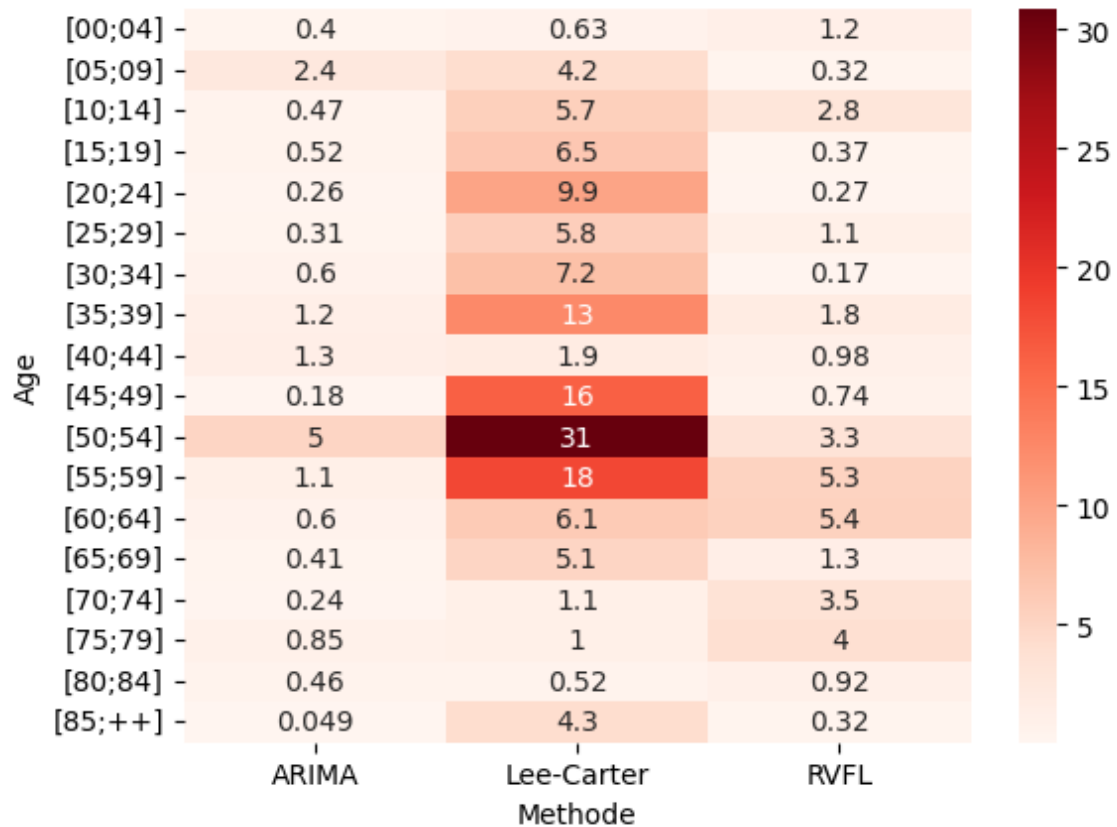


Figure 2.10 – MAPE pour chaque âge en fonction des trois méthodes étudiées

Comme le premier graphique l'indiquait, le modèle Lee-Carter est de loin le moins efficace, surtout pour les tranches d'âge de 45 à 65 ans. C'est précisément dans cette tranche d'âge que l'on observe le pic le plus important.

En ce qui concerne les deux autres méthodes, ARIMA semble mieux performer en moyenne, même si la différence n'est pas énorme. Nous allons maintenant analyser l'erreur quadratique moyenne (MSE) pour obtenir une évaluation plus précise des performances des modèles.

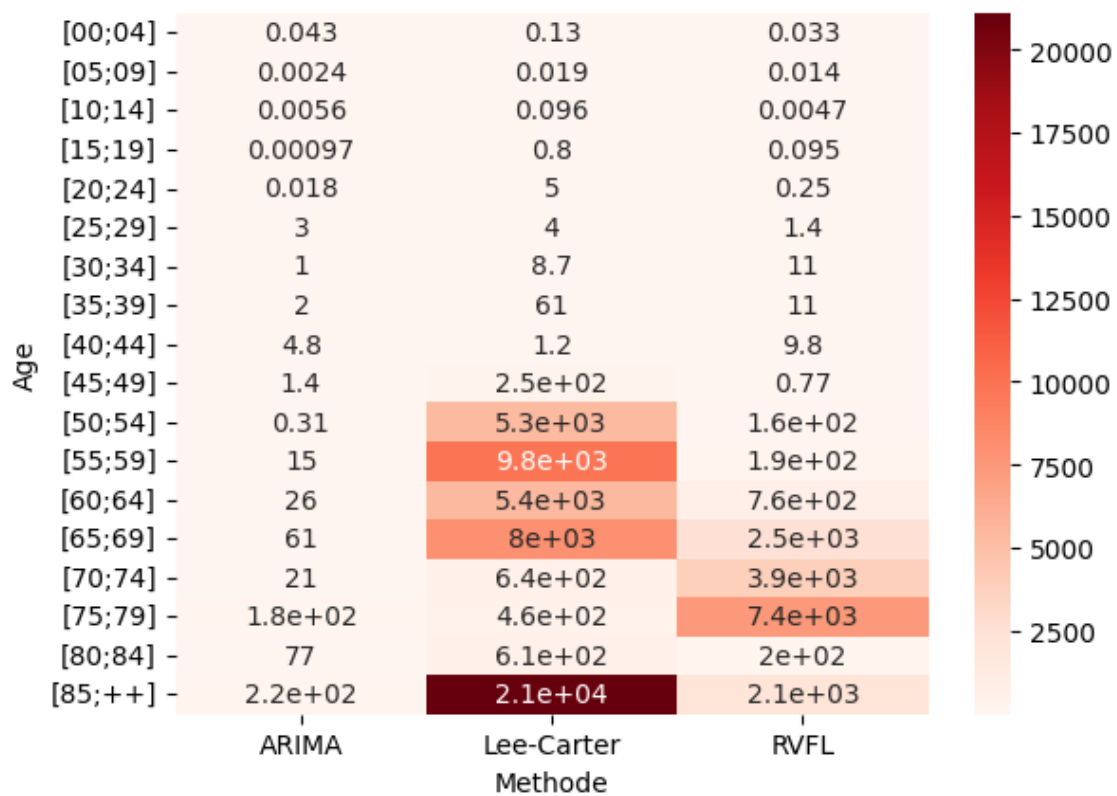


Figure 2.11 – Comparaison des methodes selon le MSE

Pour ce qui est du modèle Lee-Carter, notre analyse du MSE ne nous apporte pas de nouvelles informations. Ce modèle reste systématiquement moins performant que les deux autres méthodes. La comparaison entre ARIMA et RVFL révèle une différence plus marquée. En effet, pour certaines catégories d'âge, ARIMA se montre nettement plus performante que RVFL. Il est intéressant de noter que pour les tranches d'âge de 70 à 85 ans, RVFL est parfois même moins performant que Lee-Carter.

ARIMA semble être le modèle qui réussit le mieux à intégrer et à ajuster le pic observé dans les années 2005. Cette observation nous conduit à privilégier ARIMA comme modèle de prédiction.

Cependant, avant de prendre une décision finale, il est essentiel d'examiner comment évoluent les prédictions sur un horizon de 30 ans

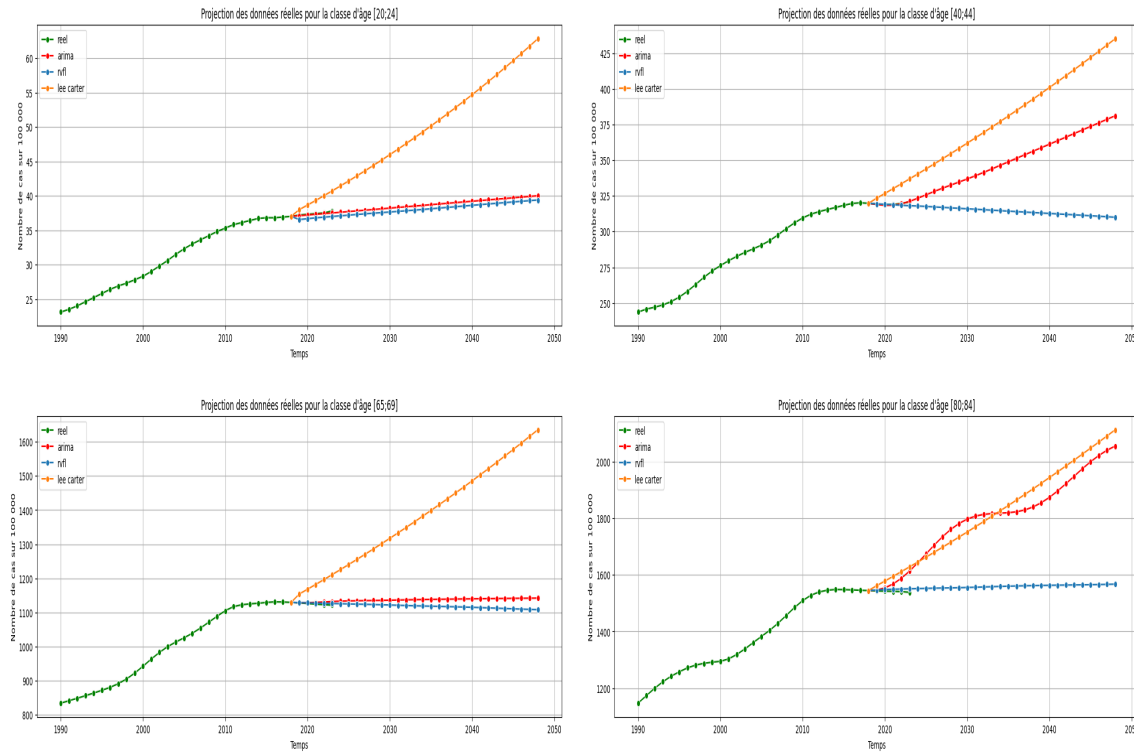


Figure 2.12 – Projection de l'incidence sur 30 ans selon les différentes méthodes

Par analyse, on remarque qu'ARIMA est un bon compromis entre le modèle de Lee-Carter, qui évalue une forte augmentation de l'incidence, et le modèle RVFL, qui se situe souvent légèrement en dessous de l'ARIMA ou prédit une baisse de l'incidence.

Par souci de prudence, nous choisirons de prendre comme modèle ARIMA, qui, dans le cadre de nos données, est le plus efficace.

Maintenant que nous avons choisi notre modèle, nous allons calculer la prime pure de la garantie maladie redoutée.

Chapter 3

Application sur le calcul de la prime pure

3.1 Marché de la garantie maladie redoutée

Contrairement aux pays anglo-saxons, la garantie maladie redoutée n'est pas très développée en France. Sur le marché, on trouve un peu moins d'une dizaine d'assureurs qui la proposent en tant que telle, bien qu'il soit possible de la trouver intégrée dans des contrats de prévoyance plus larges. Pour avoir une idée plus précise, on peut consulter le site américain Precedence Research, qui fournit des informations sur la part de chaque continent dans le marché de cette garantie.

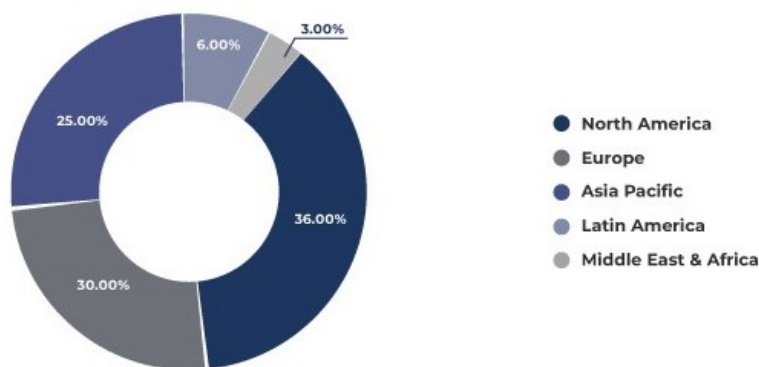


Figure 3.1 – Part de marché de la garantie maladie redoutée par région du monde

On peut noter que l'Amérique du Nord constitue le marché le plus important dans le domaine de la garantie maladie. En 2022, elle représentait 36 % de la capitalisation boursière, qui était alors de 358,6 milliards de dollars. Selon les prévisions de la même étude, ce marché devrait presque doubler d'ici 2032, atteignant 735,63 milliards de dollars. D'autres études similaires ont été réalisées et prévoient également une telle évolution pour cette garantie.

Bien que le marché américain ne soit pas identique à celui de la France, il

témoigne néanmoins d'un intérêt croissant pour cette assurance. Par ailleurs, une autre étude publiée par WantStat fournit des informations détaillées sur l'évolution du marché en Europe, ainsi que par pays, jusqu'en 2025. Ces données permettent de mieux comprendre les dynamiques régionales et les perspectives futures pour la garantie maladie redoutée en Europe.

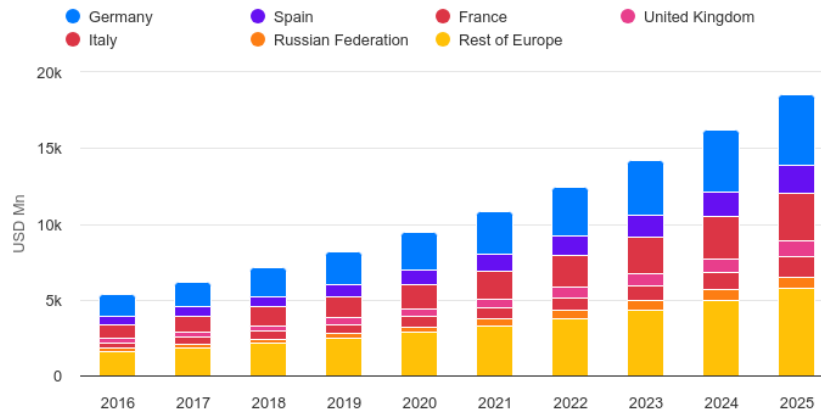


Figure 3.2 – Prédiction des parts de marché de la garantie maladie redoutée en Europe

Les études publiées ne permettent pas d'obtenir des chiffres plus précis que ceux présentés dans ce graphique. Cependant, on observe que, bien que la part de la France reste faible par rapport à des pays comme l'Italie ou l'Allemagne, les prévisions indiquent une évolution positive du chiffre d'affaires dû à cette garantie. Depuis 2016, le marché de la garantie maladie en France n'a cessé de croître jusqu'en 2023, montrant ainsi une tendance à la hausse continue.

Une des raisons de ce manque de développement pourrait être due au fait que la population française est moins inquiète par rapport aux conséquences financières liées au cancer car cette maladie fait partie des ALD et qu'elle est remboursée à 100 %. Cependant, ces dernières années, les inquiétudes en France concernant le cancer ont augmenté. Les études récentes révèlent des statistiques préoccupantes : le nombre de cancers a considérablement augmenté et, aujourd'hui, à la naissance, une personne a une chance sur deux de développer un cancer au cours de sa vie.

Bien que la maladie soit prise en charge à 100 %, les frais annexes deviennent de plus en plus importants, impactant ainsi la vie des malades. Ces frais peuvent inclure des coûts liés aux soins complémentaires, aux déplacements pour les traitements, ou encore à la perte de revenus en raison de l'incapacité de travailler. Par exemple, les traitements innovants ou les thérapies alternatives ne sont pas toujours remboursés, et les patients doivent souvent payer de leur poche pour ces options.

Le 30 mai 2024, l'Assemblée nationale a débattu d'une loi visant à mieux rembourser le cancer du sein, en proposant de dispenser les patientes de plusieurs frais annexes, tels que le forfait journalier hospitalier, le forfait urgences pour les consultations sans hospitalisation, et la franchise médicale sur les médicaments et autres prestations. Cette proposition montre une prise de conscience croissante des

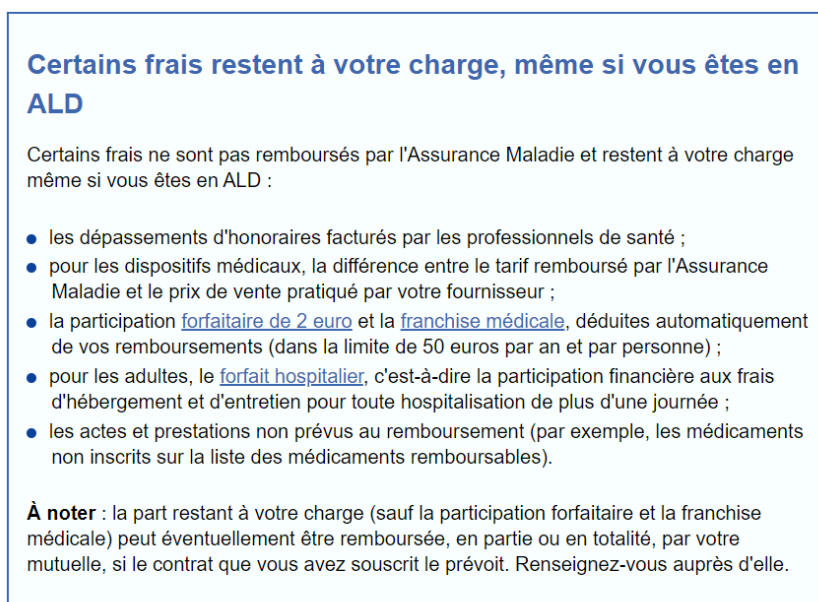
besoins des patients au-delà du simple remboursement des soins médicaux de base.

Ces facteurs peuvent conduire à une augmentation de la peur des Français par rapport au cancer. Dans ce contexte, développer la garantie maladie redoutée pourrait être une excellente opportunité pour les assureurs. Elle permettrait de couvrir non seulement les soins médicaux, mais aussi les frais annexes, offrant ainsi une protection financière plus complète et rassurante pour les patients.

En ce qui concerne les conditions des garanties, on retrouve généralement un âge compris entre 18 et 70 ans. Toutes les offres ne sont pas concernées, mais en règle générale, les assureurs ne prennent plus en charge le risque au-delà de 75 ans. En effet, à cet âge, le taux d'incidence est plus élevé et l'erreur d'estimation est plus grande. Choisir une couverture à vie peut augmenter les tarifs, tandis qu'une garantie "maladie redoutée" n'est pas forcément utile à partir d'un certain âge, lorsque les chances de survie sont faibles. Par exemple, selon Santé publique France, la survie à 5 ans pour les leucémies aiguës myéloïdes est de 69 % chez les plus jeunes (30 ans), alors qu'elle n'est que de 6 % chez les plus âgés (80 ans).

3.2 Capital pour notre produit

Lorsque l'on est atteint d'un cancer, de nombreux frais viennent s'ajouter au budget des malades. Sur le site de l'assurance maladie, Ameli, on peut d'ailleurs trouver sur la page présentant le cas des ALD que même si cette maladie est prise en charge à 100%, certains frais ne sont pas couverts :



Certains frais restent à votre charge, même si vous êtes en ALD

Certains frais ne sont pas remboursés par l'Assurance Maladie et restent à votre charge même si vous êtes en ALD :

- les dépassements d'honoraires facturés par les professionnels de santé ;
- pour les dispositifs médicaux, la différence entre le tarif remboursé par l'Assurance Maladie et le prix de vente pratiqué par votre fournisseur ;
- la participation [forfaitaire de 2 euro](#) et la [franchise médicale](#), déduites automatiquement de vos remboursements (dans la limite de 50 euros par an et par personne) ;
- pour les adultes, le [forfait hospitalier](#), c'est-à-dire la participation financière aux frais d'hébergement et d'entretien pour toute hospitalisation de plus d'une journée ;
- les actes et prestations non prévus au remboursement (par exemple, les médicaments non inscrits sur la liste des médicaments remboursables).

À noter : la part restant à votre charge (sauf la participation forfaitaire et la franchise médicale) peut éventuellement être remboursée, en partie ou en totalité, par votre mutuelle, si le contrat que vous avez souscrit le prévoit. Renseignez-vous auprès d'elle.

Figure 3.3 – Capture d'écran du site AMELI

Dans cette partie, nous allons essayer de présenter au mieux les différents frais afin d'estimer le capital nécessaire pour aider les assurés en fonction du confort qu'ils

souhaiteraient avoir en cas de diagnostic de la maladie.

Nous verrons ensuite que les contrats dits « responsables » ont un impact sur le reste à charge. On estime que près de 90 % des contrats proposés par les assureurs sont dits « responsables ». Un tel contrat doit respecter certaines règles. Pour la plupart des frais, nous ferons la distinction entre ces deux types de contrats.

À titre d'exemple, pour le cancer du sein, la Ligue nationale contre le cancer fait état d'un reste à charge compris entre 1 300 et 2 500 euros. Le professeur Pierre Marès estime qu'avec un capital de 1 000 euros par an pendant deux ans, la vie des malades et la qualité des soins seraient fortement améliorées.

Forfait hospitalier :

Le forfait hospitalier est une charge de 20 euros par jour destinée à couvrir les frais d'hébergement et d'entretien. Il n'est pas pris en charge par l'assurance maladie, mais les mutuelles peuvent choisir de le rembourser ou non. Comme mentionné précédemment, la plupart des contrats sont dits responsables, et l'une des conditions est le remboursement intégral de ce frais.

Frais d'urgence

Appelée « forfait patient urgence », il correspond à un frais que chaque personne qui se présente aux urgences, sans être hospitalisée, doit payer.

Prévue au I de l'article L. 160-13, ce forfait patient urgence est exigible depuis le 1er janvier 2022. Le montant, établi à 19,61 euros, est ramené à 8,49 euros pour les patients en ALD. Ce montant sera remboursé dans le cas d'un contrat responsable.

Frais de transport :

Ils sont remboursés dans le cadre d'une ALD (Affection de Longue Durée) uniquement si le médecin le prescrit, et ce, sous certaines conditions. Cela est possible si le patient présente des déficiences ou des incapacités à se déplacer. Cela inclut les situations où la personne doit rester allongée, nécessite un transport avec une hygiène rigoureuse, ou possède un handicap physique ou mental l'empêchant de se déplacer par ses propres moyens. Un accord supplémentaire est nécessaire si la distance dépasse 150 km ou si le trajet est récurrent. Une franchise de 4 euros par trajet, avec une limite de 8 euros par jour, est appliquée. Les montants maximaux de ces frais seront expliqués plus précisément dans le cadre de la franchise médicale.

Franchise médicale :

Elle concerne les médicaments, les actes paramédicaux ainsi que les transports sanitaires : 1,00 € par boîte de médicaments (ou toute autre unité de conditionnement, par exemple un flacon), 1 € par acte paramédical, dans la limite de 4 € par jour, et 4 € par transport sanitaire, dans la limite de 8 € par jour. Sur l'année, la franchise ne peut dépasser 50 euros. Dans le cas du traitement du cancer, on peut estimer que cette limite sera facilement atteinte. La question à se poser lors de l'estimation du capital est de savoir sur combien d'années l'assuré veut se couvrir de ces frais.

Participation forfaitaire :

Depuis le 15 mai 2024, la participation forfaitaire est de 2 euros par acte, avec une limite de 8 euros par jour par professionnel de santé et de 50 euros par an. Un contrat responsable ne rembourse pas cette participation. De la même manière que la franchise médicale, on peut, sans prendre de risque, estimer que cette limite sera atteinte.

Frais dentaires

Bien que n'étant pas directement liés au traitement du cancer, les frais dentaires s'ajoutent aux dépenses non prises en charge dans le cadre des ALD. En effet, pendant le traitement contre le cancer, des problèmes dentaires peuvent survenir. Les différents traitements auxquels les patients sont confrontés peuvent provoquer des effets secondaires touchant la région bucco-dentale, nécessitant souvent un bilan préalable pour éviter les infections. En cas d'infection, des détartrages ou des extractions dentaires peuvent être nécessaires, entraînant des coûts supplémentaires. Les frais dentaires après le traitement peuvent donc influencer le reste à charge des patients. Bien que les prothèses dentaires, telles que les dentiers, soient remboursées dans le cadre du 100 % santé, les implants dentaires ne le sont que dans des cas spécifiques, limités à un nombre précis par région de la bouche. Selon une étude de la Ligue contre le cancer, le reste à charge pour les patients atteints de cancer nécessitant des soins dentaires peut atteindre en moyenne 356 euros par an.

Reconstruction mammaire

La reconstruction mammaire est importante dans le traitement du cancer du sein chez les femmes. Elle est bien prise en charge à 100 % par l'Assurance Maladie pour les personnes en ALD. Cependant, il peut être difficile de la réaliser dans les hôpitaux publics en raison d'une offre insuffisante. De plus, les délais de prise en charge peuvent aller de plusieurs mois à plusieurs années. Pour cette raison, de nombreuses femmes se tournent vers le secteur privé. Le reste à charge moyen

déclaré par les femmes ayant recours à une chirurgie reconstructrice après un cancer du sein s'élevait à 1 391 euros en 2014, principalement en raison des dépassements d'honoraires pour l'intervention chirurgicale. Ce reste à charge peut constituer un frein, car 15 % des patientes qui envisageaient une reconstruction ont abandonné l'idée en raison du coût trop élevé.

Aide à domicile

Selon la Société canadienne du cancer, la fatigue est le symptôme le plus fréquent chez les personnes atteintes de cancer. Même avec du sommeil et du repos, la fatigue peut persister et même s'aggraver lors des traitements de chimiothérapie et de radiothérapie. L'aide à domicile permet ainsi d'améliorer les conditions de vie des patients en les soulageant de certaines tâches physiques. La chimiothérapie est le traitement le plus long, d'une durée d'environ six mois, voire plus. Si nous souhaitons offrir une aide à domicile pendant une durée de six mois, à raison de deux heures par semaine, avec un tarif de 28 € de l'heure sans crédit d'impôt, nous aurions besoin d'un budget total de 1 344 euros.

Avis de plusieurs professionnels de santé

Le diagnostic du cancer peut être très inquiétant, et de nombreux patients se sentent rassurés en obtenant plusieurs avis sur leur maladie et les traitements appropriés. Si un patient souhaite consulter un autre médecin, surtout s'il se trouve dans une autre région, cela peut entraîner des frais supplémentaires qui ne seront pas remboursés.

Pour aider à couvrir ces coûts, nous proposons un forfait de 1 000 euros. Ce forfait permettrait de financer 10 déplacements, avec un budget de 100 euros par trajet. Cette aide financière vise à réduire le stress lié aux consultations multiples, permettant ainsi aux patients de se concentrer pleinement sur leur traitement et leur rétablissement.

Soins de confort

Les effets secondaires du cancer et de ses traitements, notamment la chimiothérapie, sont nombreux et nécessitent des soins de support. Les soins de support, traduits de l'anglais "supportive care", englobent l'ensemble des soins et soutiens nécessaires aux personnes malades, parallèlement aux traitements spécifiques, tout au long des maladies graves. Ils permettent de limiter les effets secondaires des traitements ainsi que ceux de la maladie. L'importance des soins de support en oncologie a été reconnue par une circulaire du ministère de la Santé dès 2005.

En 2016, l'Institut national du cancer a publié un rapport définissant un panier de soins oncologiques de support. Ce panier comprend quatre soins de base :

la prise en charge de la douleur, les soins diététiques et nutritionnels, le soutien psychologique, et le soutien social, familial et professionnel. Il inclut également cinq soins de support complémentaires : l'activité physique, le soutien psychologique des proches et aidants, la préservation de la fertilité, la prise en charge des troubles de la sexualité, et les conseils d'hygiène de vie, ainsi que deux techniques spécifiques d'analgésie (l'hypnoanalgésie et l'analgésie intrathécale).

Pendant les traitements, ces soins de support ne sont pas pris en charge par l'assurance maladie dans le cadre du parcours de soins coordonné des patients en ALD, même lorsqu'ils sont prescrits médicalement. Les consultations de psychologue, les soins diététiques, l'activité physique adaptée, ou encore la socio-esthétique ne sont accessibles aux patients que s'ils peuvent les financer eux-mêmes ou s'ils bénéficient de séances gratuites auprès d'associations proches de leur domicile ou de leur lieu de soins.

Selon une enquête de la Ligue contre le cancer, les dépenses annuelles moyennes des patients peuvent atteindre 301 euros durant les deux à cinq années suivant la fin des traitements. Ces dépenses incluent divers besoins tels que la lingerie et les vêtements adaptés, les protections urinaires, les crèmes dermatologiques, ainsi que les consultations de professionnels de santé comme les ostéopathes, les acupuncteurs et les diététiciens, ainsi que les médicaments non remboursés.

Dépassement d'honoraire :

L'une des charges les plus importantes pour une personne en ALD suite à un cancer est le dépassement d'honoraires. Selon des études de 2013 sur le reste à charge des Français, cette charge représente environ la moitié des dépenses non couvertes. Ce dépassement et son remboursement dépendent fortement du secteur et de la convention à laquelle le médecin adhère. Il faut ainsi se renseigner avant de consulter un spécialiste.

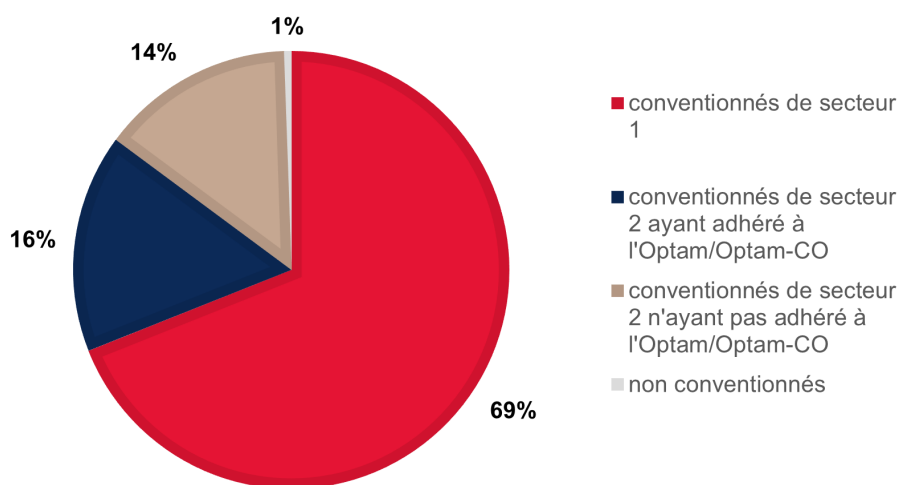


Figure 3.4 – Répartition du type de médecins en France

Sur ce graphique représentant la proportion de médecins en fonction de leur secteur et des conventions auxquelles ils adhèrent, on remarque qu'en France, une grande majorité des médecins spécialistes sont de secteur 1. Cela signifie que les actes qu'ils proposent sont au tarif conventionnel et ne pratiquent pas de dépassements d'honoraires. La Sécurité sociale remboursera une partie, et le ticket modérateur sera remboursé par la mutuelle dans le cas d'un contrat responsable.

En suivant le parcours de soins qu'avec des médecins de secteur 1 il n'y aura pas de dépassements d'honoraires cependant il est possible de vouloir voir des médecins spécifiques, de ne pas avoir beaucoup de médecins de secteur 1 dans sa région à cause des déserts médicaux. De plus on sait que selon les régions la proportion de médecins en secteur 1 est différente. Comme on peut le voir dans le graphique suivant, la proportion de médecins de secteur 1, n'est pas uniforme, en Île-de-France, il y a plus de médecins dans des conventions autres que le secteur 1.

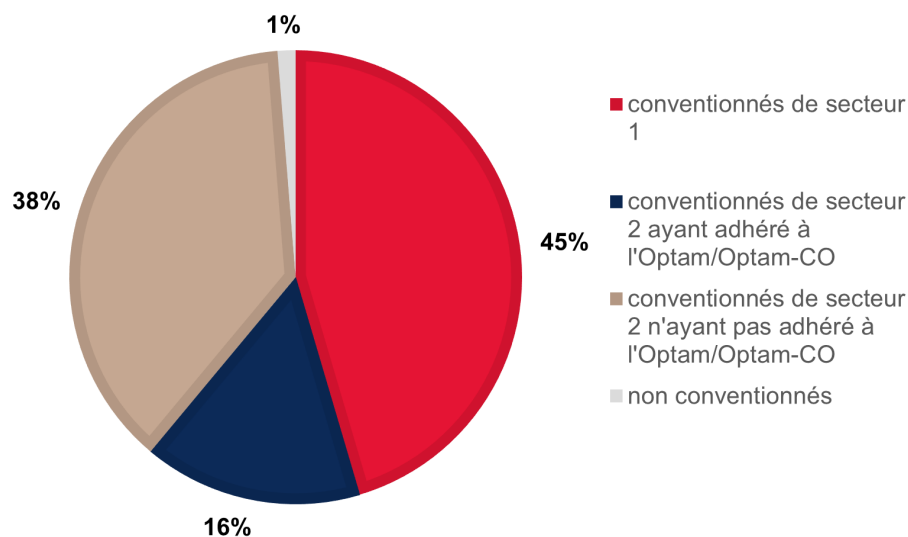


Figure 3.5 – Répartition du type de médecins en Îles de France

Dans ces régions où les médecins de secteur 1 ne sont pas majoritaires, on peut imaginer que le parcours de soins peut être ralenti par la forte demande envers les professionnels de santé de secteur 1.

Recourir à un médecin de secteur 2 peut permettre d'accélérer les soins et d'améliorer leur qualité, mais cela peut entraîner des coûts plus élevés. En effet, dans ce secteur, les dépassements d'honoraires sont autorisés. Le remboursement de ces frais dépendra du contrat de santé et de l'adhésion ou non de l'intervenant à l'OPTAM (Option Pratique Tarifaire Maîtrisée).

Le montant des prestations est libre, mais les tarifs doivent être clairement communiqués aux patients et, conformément au code de déontologie, fixés avec « tact et mesure ». Les dépassements d'honoraires ne sont pas pris en charge par la Sécurité sociale, le remboursement dépendra de la mutuelle.

Pour être qualifié de contrat responsable, un contrat d'assurance santé doit lim-

iter la prise en charge des dépassements d'honoraires à 200 % du tarif conventionnel pour les médecins non signataires de l'OPTAM et maintenir au moins 20 % d'écart entre le remboursement des dépassements d'honoraires des médecins adhérents à l'OPTAM et ceux des non-adhérents. Ainsi, s'il rembourse à 200 % les tarifs des adhérents, il doit limiter le remboursement à 180 % pour les non-adhérents pour respecter l'écart de 20 %.

La dernière catégorie comprend les médecins non conventionnés, de secteur 3, environ 500 en France, qui pratiquent comme ils l'entendent. Généralement, leurs tarifs sont très élevés et le remboursement se base sur le tarif d'autorité qui est faible. Fixé par l'article L. 162-5-10 du Code de la Sécurité sociale, il est égal à 16 % des tarifs conventionnels.

Pour les patients, avec une faible mutuelle, qui ne peuvent pas se permettre les médecins de secteur 2, leur traitement va être plus compliqué et long s'il se retrouve confrontés à un manque de médecins en secteur 1 ou si il y a une forte demande qui ralentie le processus. Cela va créer un écart de qualité de parcours de soins entre ceux qui peuvent se payer le secteur 2 et ceux qui ne le peuvent pas. Il est difficile d'estimer avec précision leurs coûts étant donné qu'ils dépendent de l'endroit où la personne vit et de la qualité des soins souhaitée.

Nous aurons besoin, pour déterminer quel montant affecter à cette section, de faire quelques hypothèses. Nous prendrons comme base une personne qui souhaite consulter des médecins de secteur 2 dans des cliniques privées. Pour ce faire, nous nous baserons sur les données disponibles sur Scan Santé, une plateforme qui répertorie les quantiles des dépassements d'honoraires des actes réalisés par des médecins selon les codes CCAM. Cette base de données nous permettra d'examiner le quantile à 75 % des dépassements d'honoraires pour les interventions principales.

Il est important de souligner que chaque parcours de soins est unique et qu'il est ainsi impossible de déterminer le montant exact du dépassement d'honoraires. En effet, un seul traitement peut parfois suffire pour certains patients, mais il est également possible que d'autres aient besoin de plusieurs interventions pour être complètement guéris. Considérer tous les soins possibles et calculer combien ils coûteraient en privé serait une grosse erreur d'estimation. En effet, le secteur public est souvent sollicité pour le traitement du cancer, et seulement certains actes sont réalisés dans le privé.

Pour estimer le capital nécessaire pour cette partie, nous ferons l'hypothèse qu'une personne souhaite effectuer toute sa radiothérapie dans une clinique privée. Selon les chiffres de l'assurance maladie, de nombreuses radiothérapies sont effectuées dans le privé. De plus, avec la chimiothérapie, c'est l'un des soins les plus coûteux et qui peut entraîner le plus de dépassements d'honoraires. Les données disponibles sur Scan Santé indiquent que 75 % des dépassements d'honoraires pour la radiothérapie sont inférieurs à 300 €.

En ce qui concerne le nombre de séances nécessaires, le site internet de l'Institut de Soin et de Recherche en Radiothérapie (ISRN) précise qu'un traitement de radiothérapie nécessite en moyenne entre 15 et 40 séances. Pour être prudents

et éviter de sous-estimer les coûts potentiels, nous prendrons comme base une estimation de 30 séances. Nous arrivons ainsi à un total de 9 000 € pour l'ensemble du traitement.

Ce calcul nous permet d'établir une estimation raisonnable du montant à prévoir pour couvrir les frais de radiothérapie dans une clinique privée. Il est crucial de noter que cette estimation est basée sur des données médianes et pourrait varier en fonction des spécificités de chaque cas individuel. De plus nous avons considéré que le cas de la radiothérapie mais cela nous permet d'avoir une bonne base qui semble raisonnable.

Récapitulatif de tous les coûts

Nous avons constaté que les restes à charge varient considérablement en fonction de la qualité des soins souhaités par une personne. Afin que la prime d'assurance ne soit pas trop élevée et corresponde bien aux besoins des assurés, nous proposons trois niveaux de couverture.

- **Garantie minimale** : Cette couverture prend en compte uniquement le minimum nécessaire pour réduire au maximum l'impact financier sur le quotidien des patients, sans toutefois améliorer considérablement leur confort.
- **Garantie moyenne** : Cette couverture permet de couvrir la quasi-totalité des frais pendant environ trois ans, mais sans offrir le confort maximal possible.
- **Garantie maximale** : Cette couverture permet de réduire les problèmes financiers au minimum et offre la possibilité de consulter presque tous les établissements de santé, à l'exception des plus élitistes.

Lorsque nous avons présenté les différents frais que l'on peut rencontrer dans son parcours de soins, nous avons conclu par une estimation annuelle ou un capital global permettant de réduire ou d'éliminer le risque de frais. Afin de définir un capital pour chaque garantie, nous avons traité chaque frais de la manière suivante. Premièrement, nous avons déterminé si le frais entre dans la garantie. Ensuite, s'il est inclus, nous avons distingué entre les frais globaux et les frais annuels. Si le frais est annuel, nous avons choisi d'assurer le risque pendant 2 ans pour la garantie minimale, 3 ans pour la moyenne, et 5 ans pour la garantie maximale. Si le frais est un capital global, nous avons choisi d'attribuer un quart pour la garantie minimale, la moitié pour la garantie moyenne et 100% pour la maximale.

Voici un tableau représentant les frais choisis pour chaque garantie, indiquant s'il s'agit de montants annuels ou globaux, avec les montants correspondants :

Frais	Type	Montant	Garantie
Dépassements d'honoraires	Forfait	9 000 €	Toutes les garanties
Dentaire	Annuel	356 €	
Soins de confort	Annuel	301 €	
Franchise médicale	Annuel	50 €	
Participation forfaitaire	Annuel	50 €	
Reconstruction mammaire	Forfait	1 391 €	Garantie moyenne et maximale
Aide à domicile	Forfait	1 344 €	
Accès à un deuxième avis médical	Forfait	1 000 €	

Table 3.1 – Tableau des frais avec détails

En utilisant les informations synthétisées dans le tableau précédent et la technique d'allocation de capital à chaque garantie, on obtient les montants suivants :

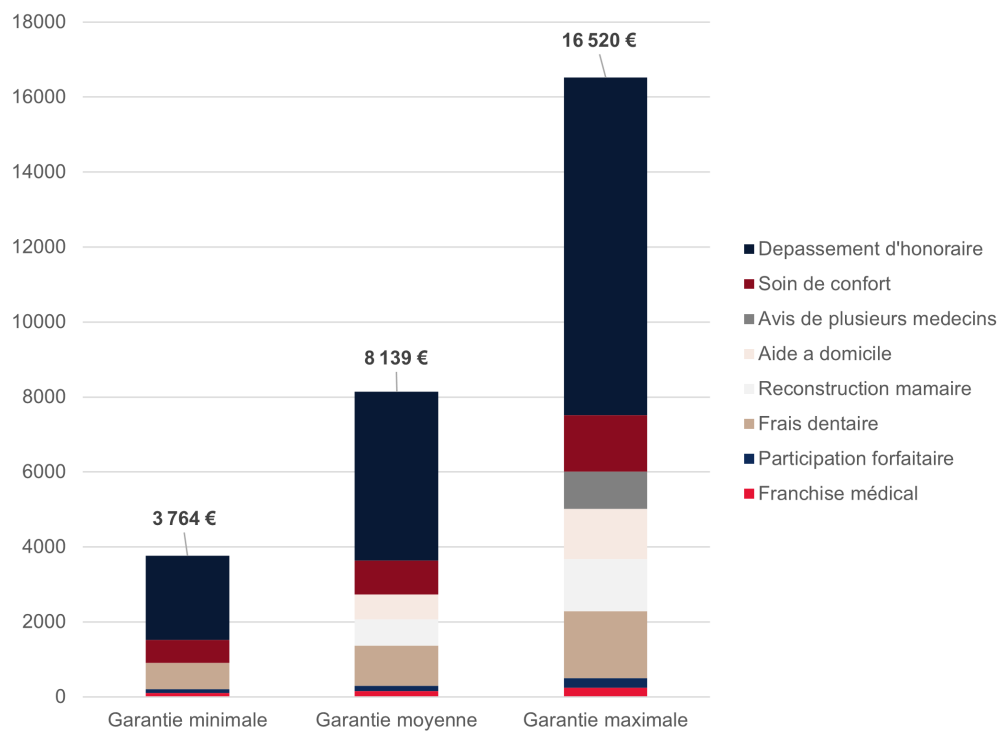


Figure 3.6 – Montant des capitaux assurés selon le type de garantie

Afin de garantir une présentation plus cohérente et pratique des assurances, nous arrondirons les montants des garanties au millier supérieur. Les trois garanties proposées seront les suivantes :

Type de garantie	Montant
Garantie minimale	4 000 €
Garantie moyenne	9 000 €
Garantie maximale	17 000 €

Maintenant que nous avons déterminé les montants des différentes garanties, nous allons établir la formule que nous utiliserons pour tarifer notre garantie.

3.3 Méthode de calcul de la prime pure

La prime pure est la partie de la prime d'assurance destinée à couvrir les risques assurés sans tenir compte des frais d'administration, des commissions ou des marges de profit. Pour une garantie maladie redoutée, le calcul de la prime pure repose sur les probabilités d'occurrence des sinistres et les coûts associés.

La première étape d'une tarification consiste à rechercher les facteurs influençant de manière significative le risque ainsi que les règles juridiques relatives au choix des paramètres. Il faut aussi prendre en compte qu'en France, il est interdit de proposer une tarification différente pour deux personnes dont la seule différence est le sexe. Cependant, il convient de noter que le sexe a un impact sur la probabilité de contracter un cancer, comme nous allons le voir.

Pour le risque lié à la garantie maladie redoutée, les deux principaux facteurs que nous allons analyser sont : le sexe et l'âge.

Le sexe

L'espérance de vie des femmes est supérieure à celle des hommes. Même si l'espérance de vie des hommes tend à rattraper celle des femmes, l'écart reste actuellement proche de 6 ans. De plus, on sait que la probabilité de développer un cancer chez la femme est bien moins importante que chez l'homme, environ deux fois moins :

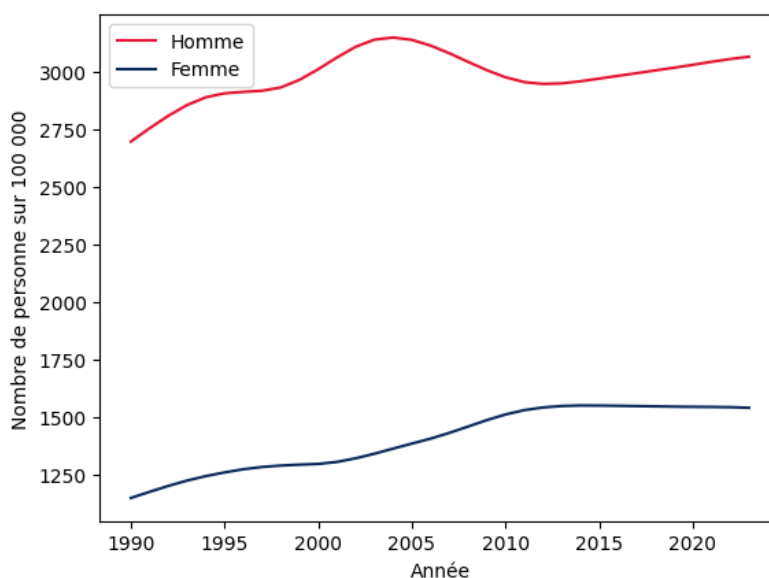


Figure 3.7 – Taux d'incidence du cancer selon le sexe pour des personnes de plus de 85 ans

Par conséquent, la répartition des individus du groupe entre hommes et femmes a donc une importance non négligeable sur la tarification de cette garantie. Le sexe étant une variable binaire, sa prise en compte dans les calculs est relativement simple.

L'âge

De façon logique, la durée de vie résiduelle, et donc la probabilité de survie, est une fonction décroissante de l'âge. De plus on sait que les chances de contracter un cancer ne sont pas les mêmes selon l'âge :

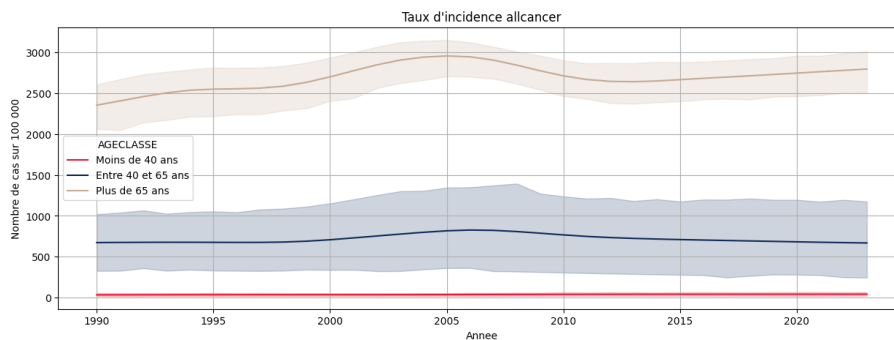


Figure 3.8 – Taux d'incidence du cancer chez l'homme selon l'âge

Autres raisons

Il existe d'autres facteurs qui pourraient influencer sur la tarification de ce produit, tels que la zone géographique dans laquelle vit l'assuré, son comportement ainsi que sa génétique. Cependant, il est difficile de disposer de données permettant d'évaluer leur impact.

Dans notre cas, comme nous ne disposons pas d'un portefeuille spécifique, nous avons, comme expliqué dans un chapitre précédent, tarifié en nous basant sur la répartition des hommes et des femmes en France selon l'INSEE.

Nous allons ensuite utiliser le taux d'incidence que nous avons projeté dans la partie précédente pour calculer notre prime pure.

3.4 Formule de calcul de la prime pure

Notation	Description
v	Facteur d'actualisation, $v = \frac{1}{1+i}$, avec i le taux d'actualisation
μ_{x+t}	Probabilité qu'un individu âgé de $x + t$ soit diagnostiqué avec un cancer au cours de l'année
${}_t c_x$	Probabilité qu'un individu âgé de x ans n'ait pas contracté le cancer jusqu'à l'âge de $x + t$
${}_t p_x$	Probabilité qu'un individu âgé de x ans survive jusqu'à l'âge de $x + t$
C	Capital assuré versé en cas de diagnostic de cancer
A_x	Valeur actuarielle actualisée des prestations d'assurance versées en cas de diagnostic de cancer
l_x	Nombre de personnes survivant jusqu'à l'âge x

Table 3.2 – Notations

L'un des principaux défis pour le calcul de la prime pure est de réussir, dès la souscription d'un contrat, à déterminer la valeur de l'engagement à long terme dont l'issue est incertaine. Pour y parvenir, on utilise le concept de valeur actuelle probable. Ce concept associe la probabilité d'occurrence (calcul de probabilité) à la valeur actuelle (mathématiques financières).

La valeur actuelle probable d'un engagement se calcule en multipliant la valeur actuelle de cet engagement par la probabilité de sa réalisation. Chaque année, il existe une probabilité non nulle qu'un assuré contracte un cancer ; ainsi, il est possible que le capital garanti doive être versé cette année-là.

Dans le cadre de la garantie maladie redoutée, nous projetons la probabilité de contracter un cancer à un certain âge, tant que l'individu est toujours en vie, en utilisant le taux d'incidence obtenu dans la partie précédente. Ce taux est ensuite multiplié par le capital assuré actualisé. Cela nous donne la formule suivante :

$$VAP(\text{Assureur}) = \sum_{t=0}^{\infty} {}_t p_x \cdot {}_t c_x \cdot \mu_{x+t} \cdot v^t \cdot C$$

La prime pure unique d'un contrat correspond à la prime nécessaire pour couvrir précisément le risque assuré, sans inclure les divers frais additionnels de l'assureur, ce qui correspond exactement à la valeur actuelle probable (VAP).

Pour déterminer la prime annuelle, on utilise la relation suivante :

$$PA = PU \cdot {}_t a_x$$

où ${}_t a_x$ vaut :

$${}_t a_x = \sum_{t=0}^{\infty} {}_t p_x \cdot {}_t c_x \cdot v^t$$

Cette formule nous permettra d'obtenir la prime annuelle qu'un assuré devra payer chaque année. Analysons maintenant l'impact de la durée de couverture afin de déterminer la couverture optimale en fonction du risque souhaité.

3.5 Temps de couverture

L'un des facteurs déterminants pour le prix de l'assurance est la durée de la couverture. En effet, plus la période de couverture est longue, plus le coût de l'assurance est élevé. De plus, au-delà d'un certain âge, il devient difficile de se fier uniquement aux taux d'incidence disponibles, en raison de l'incertitude accrue liée à cette tranche d'âge.

On se doit d'examiner les conditions spécifiques de l'assurance pour comprendre les risques associés. À partir de 80 ans, l'incidence globale des cancers est assez élevée, et l'incertitude concernant les taux d'incidence augmente considérablement. Assurer des personnes au-dessus de 80 ans pour plusieurs années représente donc un risque important pour les assureurs.

Pour atténuer ce risque, une approche possible serait de cesser de couvrir les assurés après un certain âge. Notre objectif est de déterminer, selon nos données, l'âge le plus cohérent.

Dans un premier temps, nous avons établi des graphiques représentant le coût mensuel d'une garantie maladie redoutée, en utilisant les formules présentées précédemment, en fonction de l'âge de souscription. Notre but est de comparer les primes pour chaque âge, selon que l'on couvre sur la vie entière ou jusqu'à un âge précis. Nous avons choisi de présenter les résultats pour une garantie assurant un capital de 17 000 €.

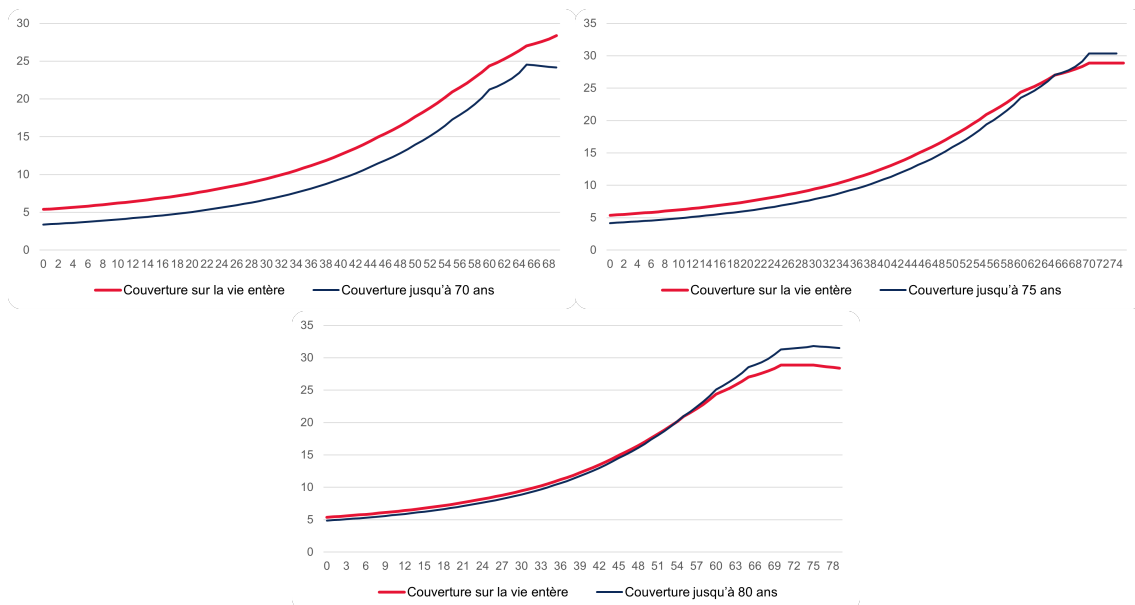


Figure 3.9 – Prime mensuelle en euros selon l'âge et le temps de couverture

Nous constatons qu'en arrêtant de couvrir le risque à 70 ans, la prime de la garantie devient nettement plus faible. Une différence de plus de 2 euros peut avoir un impact significatif sur la décision d'un individu de souscrire ou non à cette garantie. Lorsque le risque est étendu jusqu'à 75 ans, la différence persiste mais devient moins marquée. Aux alentours de 65 ans, on observe même une inversion de tendance : la prime pour une couverture à vie devient plus avantageuse. À cet âge, il pourrait être plus pertinent de proposer une couverture à vie, offrant à l'assureur un avantage concurrentiel, à la fois sur le plan tarifaire (par rapport aux concurrents) et en termes de solidité de la protection pour l'assuré.

Dans le cas où la couverture s'arrête à 80 ans, l'avantage concurrentiel en termes de coût devient presque insignifiant. À partir de 50 ans, il semble même plus judicieux pour l'assureur de proposer une couverture à vie, car cela permet de répondre aux attentes des assurés tout en restant compétitif. D'un point de vue tarifaire, limiter la couverture à 70 ans permettrait à l'assureur de proposer des primes beaucoup plus compétitives, renforçant ainsi sa position sur le marché. Cependant, cette limitation pourrait susciter des inquiétudes chez les assurés potentiels concernant la durée de leur protection.

Arrêter la couverture à 75 ans présente l'avantage d'offrir une protection plus longue tout en maintenant un avantage de coût pour l'assureur, et cela permet de réduire l'exposition au risque lié aux âges très avancés. Il est toutefois important de souligner que notre modèle, disposant de peu de données, ne permet pas d'obtenir des résultats avec une grande certitude. En effet, maintenir la couverture au-delà de 80 ans expose l'assureur à des risques accrus de paiements, et bien que les courbes de primes soient similaires pour une couverture jusqu'à 80 ans et une couverture à vie, l'exposition au risque pour l'assureur est bien plus élevée dans le cas d'une couverture à vie. Ainsi, en tenant compte du risque et des coûts, l'âge de 75 ans

semble être le plus pertinent dans notre étude.

Cette approche vise à trouver un équilibre entre la gestion des risques et la maximisation des profits pour l'assureur, tout en garantissant une protection adéquate pour les assurés, en particulier dans une tranche d'âge où les taux d'incidence sont mieux maîtrisés et plus faibles.

Après avoir analysé les caractéristiques de notre produit d'assurance et ajusté la gestion des risques en fonction de l'âge et de la durée de couverture, nous allons maintenant procéder à une synthèse des principaux éléments. Cette synthèse nous permettra de finaliser le produit en prenant en compte les objectifs financiers de l'assureur ainsi que les besoins des assurés.

3.6 Proposition de produit

Dans cette section, nous présenterons un exemple de produit "Garantie Maladie redoutée" qui fonctionnerait pour le cancer seulement. Nos calculs futurs seront pour un contrat individuel.

- **Âge minimal de souscription** : 18 ans.
- **Durée de couverture** : La garantie reste en vigueur jusqu'à ce que l'assuré atteigne l'âge de 75 ans, sous réserve que les primes requises soient payées.
- **Primes** : Les primes sont payées annuellement. L'assuré a le droit de résilier son contrat chaque année.
- **Conditions de sortie du portefeuille** : L'assuré quitte le portefeuille pour trois raisons : survenance d'une maladie grave (tel que le cancer), résiliation ou décès.
- **Capital assuré** : En cas de survenance d'une maladie grave comme le cancer, l'assuré reçoit un montant forfaitaire dépendant de la garantie qu'il estime lui convenir le mieux. Ce montant est fixe et n'est pas influencé par la nature ou la gravité de la maladie.

Avec toutes les conditions bien établies nous allons pouvoir calculer la prime pure.

3.6.1 Calcul de la prime

Une fois le taux d'incidence du cancer projeté selon la méthode retenue, un outil Excel a été développé afin de calculer, pour un âge donné, la prime pure associée à la garantie. Ce calcul repose sur les formules détaillées précédemment, en intégrant les caractéristiques techniques du contrat définies dans le cadre de l'étude (âge de souscription, durée de couverture, montant du capital assuré, etc.). L'outil permet ainsi de simuler rapidement la prime théorique en fonction des paramètres choisis, et de l'ajuster selon différents scénarios.

Afin de tarifier notre produit, comme présenté lors du calcul de la prime, nous avons choisi d'utiliser pour l'actualisation le taux sans risque publié par la Banque de France. Ce taux peut être facilement mis à jour dans notre outil, ce qui permet d'adapter la tarification aux conditions de marché. Dans le cadre de cette étude, nous utilisons la courbe des taux zéro coupon du 30 avril 2025, que nous représentons ci-dessous :

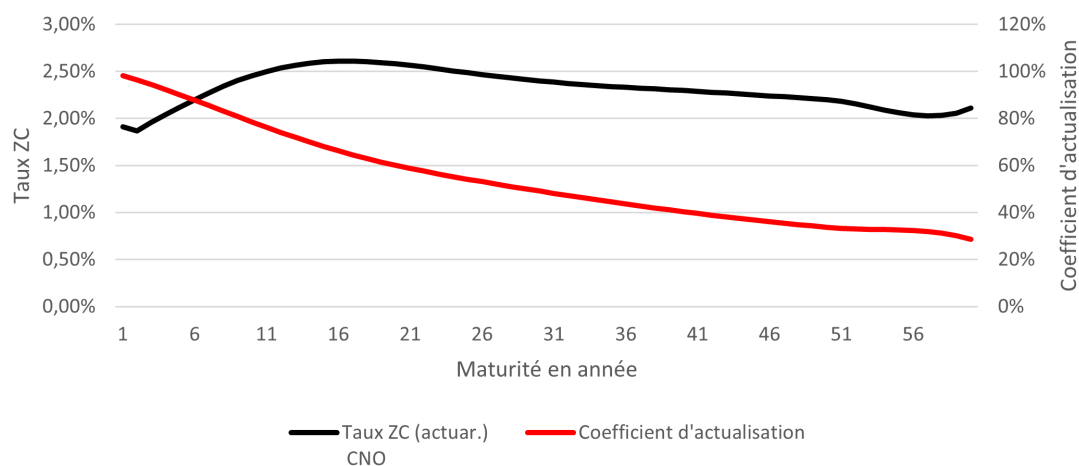


Figure 3.10 – Courbe taux zéro coupon du 30 avril 2025

En l'absence de données propres sur une population assurée, ce mémoire s'appuie sur les tables réglementaires TH 00-02 et TF 00-02, construites à partir des données de mortalité de la population française entre 2000 et 2002, publiées par l'INSEE. Ces tables distinguent la mortalité selon le sexe : la table TH pour les hommes, la table TF pour les femmes. Elles présentent, pour chaque âge entre 0 et 112 ans, la survie d'une cohorte théorique de 100 000 individus. Ce format permet de calculer les probabilités de décès et de survie nécessaires au calcul de notre prime.

Age	lx	Age	lx	Age	lx	Age	lx
0	100 000	16	99 308	32	98 260	48	95 422
1	99 566	17	99 271	33	98 175	49	95 081
2	99 530	18	99 224	34	98 083	50	94 717
3	99 506	19	99 168	35	97 984	51	94 332
4	99 487	20	99 105	36	97 876	52	93 924
5	99 471	21	99 038	37	97 760	53	93 489
6	99 457	22	98 970	38	97 633	54	93 026
7	99 445	23	98 903	39	97 495	55	92 533
8	99 433	24	98 836	40	97 343	56	92 008
9	99 422	25	98 768	41	97 175	57	91 453
10	99 411	26	98 701	42	96 989	58	90 866
11	99 399	27	98 632	43	96 784	59	90 246
12	99 387	28	98 562	44	96 558	60	89 589
13	99 373	29	98 491	45	96 310	61	88 892
14	99 357	30	98 417	46	96 037	62	88 148
15	99 336	31	98 340	47	95 741	63	87 351

Figure 3.11 – Table de mortalité

Afin d'adapter ces données à une population mixte, nous avons utilisé la table unisexe présentée ci-dessus, qui a été construite en prenant une moyenne pondérée des tables TH et TF selon la répartition homme/femme observée dans les données démographiques de l'INSEE. Cette approche permet de rester cohérent avec les caractéristiques de la population générale, tout en simplifiant le calcul de la prime dans le cas d'un tarif unique non différencié selon le sexe. Bien entendu, dans le cadre d'une mise en œuvre opérationnelle du produit, une table d'expérience serait envisageable.

Maintenant que nous avons détaillé les données que nous avons utilisées, voici le montant des primes pures mensuelles obtenues que nous avons calculé par garantie et par âge :

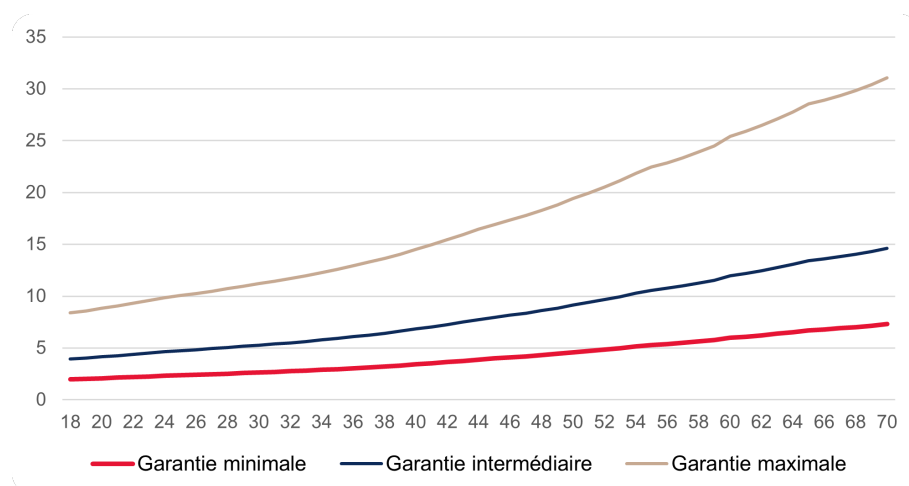


Figure 3.12 – Prime pure en fonction de l'âge et du capital assuré

3.6.2 Test de sensibilité

L'évolution de l'incidence du cancer peut être influencée par de nombreux facteurs externes, comme les conditions environnementales, les politiques de santé publique ou encore les pratiques médicales. Cette incertitude rend pertinent un test de sensibilité, dont l'objectif est d'évaluer dans quelle mesure la prime pure réagirait à une hausse de l'incidence, en cas de scénario moins favorable que celui projeté.

Pour simplifier l'analyse, nous fixons le capital assuré à 9 000 € et considérons une garantie unique. Trois scénarios ont été définis. Ils visent à représenter des contextes distincts : une erreur de prévision modérée, un choc immédiat suivi d'un retour à la normale, et un choc différé. Le choix des paramètres s'inspire notamment de l'augmentation observée en 2005 à la suite des campagnes de dépistage, où l'incidence de certains cancers avait temporairement progressé de l'ordre de 15 %.

- **Scénario 1 : augmentation uniforme de 5 %**

Ce scénario vise à tester l'impact d'une prévision légèrement sous-estimée, pouvant résulter d'un vieillissement de la population ou d'un changement progressif des facteurs de risque.

- **Scénario 2 : hausse de 15 % sur les cinq premières années**

Il s'agit ici de simuler un choc temporaire dès le lancement du produit, par exemple en cas d'intensification soudaine du dépistage, comme cela a été observé par le passé.

- **Scénario 3 : hausse de 15 % à partir de la cinquième année, pendant cinq ans**

Ce scénario représente un choc différé, pouvant correspondre à l'émergence future d'un facteur de risque ou à une modification des comportements sanitaires.

Les résultats obtenus sont présentés pour différents âges à la souscription.

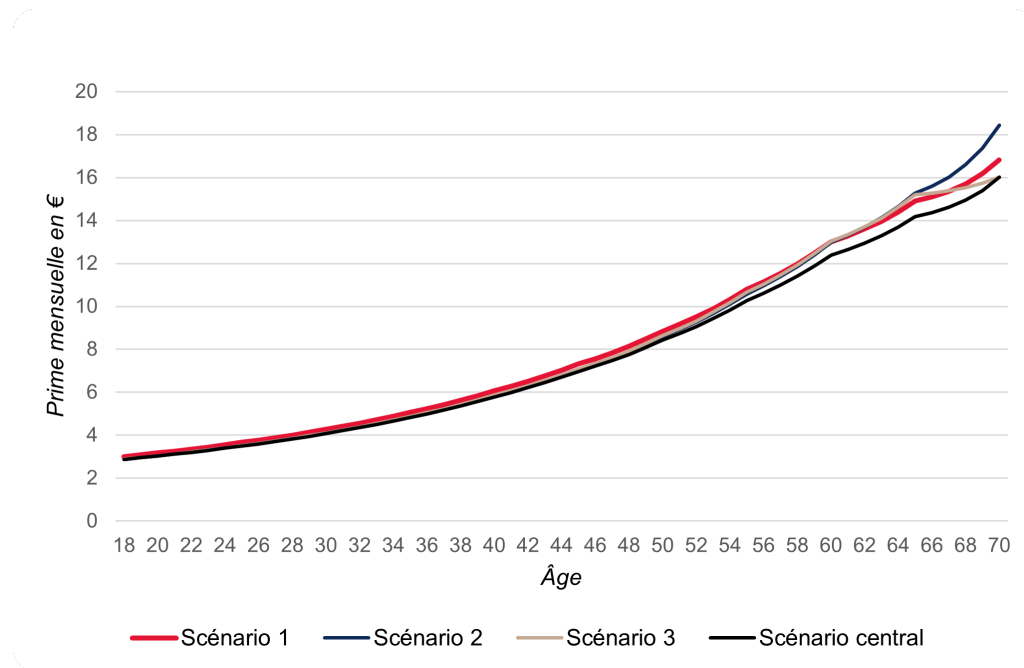


Figure 3.13 – Évolution de la prime pure selon les scénarios de choc

Afin d'enrichir cette première visualisation, un second graphique permet de comparer directement les niveaux de prime obtenus selon les différents scénarios étudiés. Cette représentation complémentaire offre une lecture synthétique de l'impact des chocs d'incidence sur le coût du produit, en fonction de l'âge de l'assuré. Dans cette représentation, nous avons choisi trois types d'assurés possibles pour voir l'impact du scénario selon le profil de l'assuré.

	20 ans	30 ans	40 ans	60 ans
Prime pure scénario central	3,03 €	4,08 €	5,78 €	12,39 €
Prime pure scénario 1	3,17 €	4,28 €	6,06 €	13,01 €
Augmentation en %	4,66%	4,73%	4,82%	5,03%
Prime pure scénario 2	3,04 €	4,11 €	5,84 €	12,99 €
Augmentation en %	0,21%	0,55%	1,01%	4,84%
Prime pure scénario 3	3,04 €	4,11 €	5,86 €	13,04 €
Augmentation en %	0,37%	0,74%	1,36%	5,24%

Figure 3.14 – Prime pure selon le scénario

De manière générale, le scénario 1, fondé sur une hausse uniforme de l'incidence dans le temps et sur l'ensemble des âges, met en évidence une relation quasi proportionnelle entre l'évolution du risque et le montant de la prime pure. Cette linéarité est attendue au regard de la structure du modèle, et elle permet d'anticiper de manière assez directe l'impact d'une variation globale du risque sur le coût du produit.

En revanche, les scénarios 2 et 3, qui introduisent des chocs d'incidence différenciés selon l'âge, révèlent une sensibilité nettement plus marquée pour les profils les plus âgés. Dans ces configurations, un choc de 15 % appliqué à l'incidence n'entraîne qu'une faible variation de la prime pure pour les assurés jeunes (20 ou 30 ans), avec des hausses souvent inférieures à 1 %. À l'opposé, pour les assurés âgés de 60 ans, la même intensité de choc peut conduire à une augmentation de la prime pouvant atteindre 5 %. Ce contraste s'explique par le fait que, chez les jeunes, la probabilité d'occurrence de la pathologie reste très faible, même après choc, ce qui limite l'impact financier. À l'inverse, chez les seniors, où l'incidence est déjà plus élevée, une variation relative identique se traduit par un surcoût plus significatif.

Ces résultats soulignent l'importance d'intégrer une approche différenciée dans la tarification, en particulier pour les tranches d'âge les plus sensibles. Afin de refléter cette hétérogénéité de réaction au risque, une marge de prudence de 5 % a été ajoutée à la prime pure des assurés âgés de 50 ans ou plus. Cette hypothèse, bien que volontairement simple, permet d'illustrer comment la structure tarifaire pourrait être adaptée pour tenir compte de l'exposition accrue à l'incertitude dans certaines classes d'âge, sans complexifier excessivement le produit.

3.6.3 Limites et améliorations possibles de notre modèle

Limites de notre modèle face au contexte et facteurs externes

Dans un contexte où le nombre de cancers augmente et inquiète les Français, le gouvernement a lancé une campagne de lutte contre le cancer en France, visant à réduire l'impact de cette maladie sur le quotidien des patients. Cette campagne s'articule autour de trois axes principaux : le dépistage, la prévention et l'amélioration du parcours de soins. Selon la feuille de route de la stratégie décennale de lutte contre le cancer, près de 40 % des cancers pourraient être évités par un changement de mode de vie. La campagne de prévention cible particulièrement le tabagisme et la consommation d'alcool, identifiés comme les deux principaux facteurs de risque. L'objectif final est de réduire de 60 000 par an le nombre de nouveaux cas de cancer. De plus, la campagne cherche à augmenter le nombre de dépistages de plus d'un million par an, afin de détecter les cancers le plus tôt possible, ce qui accroît les chances de guérison des patients. Ce contexte complexifie l'analyse de l'incidence du cancer. En effet, l'augmentation du nombre de dépistages pourrait, comme en 2005, entraîner une hausse apparente de l'incidence du cancer en France en révélant davantage de cas. Toutefois, si la prévention devient efficace, elle pourrait également contribuer à réduire le nombre de cancers et, par conséquent, faire baisser le taux d'incidence. Intégrer ces données dans notre modèle est donc délicat, sachant que cette campagne aura probablement un impact sur le taux d'incidence.

On a pu remarquer lors de notre modélisation que le taux d'incidence était très sensible à des facteurs externes indépendants du temps, comme la campagne de 2005. Le contexte actuel en France pourrait également avoir une influence importante. Toutefois, il reste difficile de l'intégrer dans notre modèle, en particulier sur un horizon de 30 ans. Pour compléter ce modèle, il serait pertinent de l'alimenter avec d'autres données exogènes, telles que le nombre de fumeurs, la consommation d'alcool, ou encore d'autres facteurs évoqués précédemment.

Limites Associées aux Données

Nous travaillons sur les incidences disponibles jusqu'à l'année 2023, ce qui représente seulement 32 points de données. De plus, la méthode de calcul utilisée pour déterminer l'incidence du cancer tous types par classe d'âge n'est pas décrite en détail dans la source. Les auteurs expliquent brièvement leur approche sans détailler les erreurs de prédiction, ce qui limite la transparence et la reproductibilité de l'analyse.

Les données sont disponibles par classe d'âge, alors que des données par âge auraient pu fournir des résultats plus précis. En effet, l'incidence du cancer peut varier considérablement entre les âges au sein d'une même classe. L'absence de granularité dans les données entrave la précision des modèles prédictifs et apporte de l'erreur à la prédiction de notre modèle.

Validation Croisée

En raison du faible nombre d'années d'historique disponibles, une validation croisée n'a pas été réalisée. Même pour une simple division en ensemble d'entraînement et de test, nous n'avions que peu de données pour l'entraînement et encore moins pour valider nos modèles.

À ce stade, réduire davantage le nombre de données d'entraînement aurait été compliqué et aurait pu compromettre la qualité des modèles. Il serait bénéfique de disposer de plus de données historiques pour permettre une validation croisée plus robuste, améliorant ainsi la fiabilité des prévisions.

Modélisation par Type de Cancer

La modélisation globale du cancer ne permet pas de connaître l'évolution individuelle de chaque type de cancer. Une modélisation par type de cancer offrirait à l'assureur une meilleure compréhension de son risque. Malheureusement, nous ne disposons pas de toutes les données par type de cancer. Par exemple, pour le cancer de la prostate, nous n'avons des données que jusqu'à 2018. Les hospices de Lyon ont décidé de ne pas donner les taux pour ces années, du fait des évolutions récentes de l'incidence de ce cancer et l'incertitude sur leur évolution à court terme.

Dans cette étude, la couverture forfaitaire appliquée dans le cadre de l'assurance cancer est indépendante du type de cancer. C'est pourquoi la modélisation du cancer tous types en France répond au mieux à la problématique. Cependant, une approche plus détaillée par type de cancer pourrait permettre de mieux cibler les couvertures d'assurance et d'offrir des produits plus adaptés aux besoins spécifiques des patients.

Modélisation du Capital

Le capital que nous avons estimé afin de répondre au mieux aux attentes des patients repose sur des hypothèses assez fortes. Nous avons toujours essayé de surestimer le capital par souci de prudence. Toutefois, une analyse plus poussée avec des données plus précises permettrait d'obtenir un capital plus adapté. Il serait pertinent d'examiner de nouvelles sources de données et d'affiner les hypothèses pour améliorer la précision des estimations.

Conclusion

Le cancer, classé comme Affection de Longue Durée (ALD) en France, est une maladie qui préoccupe de plus en plus les Français. Le nombre de cas augmente principalement en raison de trois facteurs : le vieillissement de la population, l'augmentation démographique et les comportements à risque. Malgré l'amélioration des méthodes de dépistage et des traitements, le nombre de cancers a doublé entre 1990 et aujourd'hui. Bien que cette maladie soit prise en charge à 100% dans le cadre des ALD en France, elle peut entraîner des coûts importants pour les patients. L'assurance « maladie redoutée » peut être une solution adaptée pour ces personnes.

Cette assurance permet d'octroyer un capital, choisi en fonction du confort de soins souhaité, à une personne diagnostiquée avec un cancer. Le capital a pour but d'aider à couvrir tous les frais annexes qui ne sont pas pris en charge, tels que les dépassements d'honoraires, une aide à domicile, des soins de confort, etc. Le but de cette étude était de restructurer ce produit d'assurance afin de mieux l'adapter au contexte actuel.

Pour ce faire, nous avons utilisé la base de données de l'Institut National du Cancer (INCa), qui nous fournit le taux d'incidence du cancer de 1990 à 2023 en France selon des catégories d'âge quinquennales. Les taux étaient donnés par type de cancer ou pour l'ensemble des cancers. Il aurait été possible de chercher à modéliser le taux d'incidence en modélisant les prévisions par type de cancer, mais nous avons choisi de concentrer notre analyse sur le taux général. Nous avons déterminé que la méthode ARIMA était la plus adaptée pour nos données. En effet, le nombre de points était assez limité et nous avons un pic dû aux dépistages, ce qui rendait la méthode Lee-Carter moins efficace. Les réseaux RVFL nous donnaient également des résultats intéressants, mais en raison de l'erreur possible dans nos modélisations, nous avons choisi la prédiction la plus prudente.

Pour le choix des différents forfaits proposés, nous avons répertorié les principaux frais auxquels un malade peut être confronté au cours de son traitement et créé trois catégories différentes. Un premier capital permettant de couvrir les frais principaux, un deuxième permettant d'avoir un confort important sans pour autant avoir accès aux meilleurs soins, et un dernier capital prenant en compte toutes les dépenses listées.

Notre analyse, bien que reposant sur de grandes hypothèses, permet d'établir une bonne base pour la tarification d'un tel produit. Un assureur souhaitant se lancer

dans cette garantie pourrait se servir de ces travaux comme base. Afin d'améliorer cette étude, il serait intéressant d'ajouter les données internes d'un assureur afin de mieux cibler la population visée. Trouver d'autres variables explicatives pourrait également être bénéfique, bien que l'obtention de telles données reste compliquée en France. De plus, adapter les restes à charges en fonction de chaque individu pourrait être une bonne manière de trouver le bon capital.

Cette assurance permet de réduire le poids financier, de sensibiliser et d'accompagner les personnes atteintes de cette maladie qui inquiète de plus en plus de Français. Le développement d'un tel produit par un assureur pourrait améliorer son image en montrant une volonté de lutter contre le cancer en France.

Cependant, il convient de souligner que cette étude, bien qu'offrant une base avancée pour la création d'une telle assurance, pourrait être améliorée par l'intégration de données issues des portefeuilles d'assurés. L'accès à des données internes faciliterait non seulement l'affinement des prévisions de taux d'incidence, mais aussi une estimation plus précise du capital optimal à allouer.

List of Figures

1	Comparaison de l'incidence selon l'âge et le sexe	7
2	MAE pour chaque âge en fonction des trois méthodes étudiées	7
3	Montant des capitaux assurés selon le type de garantie	8
4	Prime pure en fonction de l'âge et du capital assuré	9
5	Comparison of incidence by age and gender	12
6	MAE for each age group based on the three studied methods	12
7	Insured lump sum amounts based on coverage type	13
8	Pure premium based on age and insured capital	14
1.1	Propagation du cancer	24
1.2	Nombre de cas en 2023 et évolution du taux d'incidence entre 2010 et 2023 ¹	26
1.3	Proportion des cancers liés aux principaux facteurs de risque ⁴	28
1.4	Survie nette à 5 ans par type de cancer	30
1.5	Synthèse des caractéristiques concernant l'ALD	31
1.6	Répartition des dépenses de l'assurance maladie	32
1.7	Les différentes branches de la Sécurité sociale	34
1.8	Réseau de la Cnam	35
1.9	Répartition des recettes de la Sécurité Sociale	36
1.10	Fonctionnement du remboursement des frais de santé	37
1.11	Répartition des maladies comprises dans le cadre des ALD	39
2.1	Les différents registres en France	44
2.2	Taux d'incidence pour les 3 cancers les plus fréquent chez l'homme . .	45

2.3	Taux d'incidence pour les 3 cancers les plus fréquent chez la femme .	46
2.4	Comparaison de l'incidence selon l'âge et le sexe	47
2.5	Schéma représentant le fonctionnement d'un réseaux RVFL	56
2.6	Schéma représentant la méthodologie utilisée	57
2.7	Schéma représentant le fonctionnement de la méthode k-fold	58
2.8	Schéma représentant le fonctionnement de la méthode déroulante . .	58
2.9	Comparaison des méthodes utilisées par rapport aux données réelles .	61
2.10	MAPE pour chaque âge en fonction des trois méthodes étudiées . . .	62
2.11	Comparaison des methodes selon le MSE	63
2.12	Projection de l'incidence sur 30 ans selon les différentes méthodes . .	64
3.1	Part de marché de la garantie maladie redoutée par région du monde	65
3.2	Prédiction des parts de marché de la garantie maladie redoutée en Europe	66
3.3	Capture d'écran du site AMELI	67
3.4	Répartition du type de médecins en France	71
3.5	Répartition du type de médecins en Îles de France	72
3.6	Montant des capitaux assurés selon le type de garantie	75
3.7	Taux d'incidence du cancer selon le sexe pour des personnes de plus de 85 ans	77
3.8	Taux d'incidence du cancer chez l'homme selon l'âge	78
3.9	Prime mensuelle en euros selon l'âge et le temps de couverture	81
3.10	Courbe taux zéro coupon du 30 avril 2025	83
3.11	Table de mortalité	84
3.12	Prime pure en fonction de l'âge et du capital assuré	84
3.13	Évolution de la prime pure selon les scénarios de choc	86
3.14	Prime pure selon le scénario	86

Bibliography

- [1] Ameli. À quoi sert la prise en charge en affection longue durée (ald) exonérante ?, 2024.
- [2] Annick Berger. L'inquiétante explosion des coûts du cancer en france. *Capital*, Février 2020.
- [3] Antoine BONNANS. Cancer insurance: Forecasting incidence and impact on profitability, 2021.
- [4] Quitterie Brezillon. Augmentation des cas de cancers chez les femmes jeunes : la qualité de vie, parent pauvre des politiques publiques centrées sur la prévention et la recherche. *Rose Up*, 2023.
- [5] Gautier Defosser. Estimations nationales de l'incidence et de la mortalité par cancer en france métropolitaine entre 1990 et 2018. *INCa*, 2019.
- [6] Didier Delignières. Séries temporelles, modèles arima. 2020, 2020.
- [7] Département Prévention Cancer Environnement. Cancer de la prostate, 2024.
- [8] Juliette Favre. Dépassement d'honoraires : quel remboursement de la sécurité sociale et la mutuelle ? *Selectra*, 2024.
- [9] France Assos Santé. Dépassements d'honoraires en clinique : les chiffres de l'augmentation, 2015.
- [10] Yasmine HAJJEM. Modélisation du risque cancer pour des contrats d'assurance long terme, 2021.
- [11] Antoine Henry. 1000€ par an : le témoignage de patients sur leurs restes à charge « invisibles ». *France Assos Santé*, 2019.
- [12] Rob Hyndman. Mesuring forecast accuracy, 2014.
- [13] La ligue contre le cancer. Les retes à charges, 2014.
- [14] La Ligue contre le cancer. Impact social du cancer, 2016.
- [15] Pierre Lerest. Remboursement d'un dépassement d'honoraires : est-il pris en charge par l'assurance maladie et la mutuelle ? *Aide Sociale*, 2023.

- [16] V. Monbet. Modélisation de séries temporelles. *Université de Rennes*, 2011.
- [17] Frédéric PLANCHET. Modèles de durée, 2023.
- [18] Previsima. Mon contrat santé est-il un contrat responsable ?, 2024.
- [19] Vie publique. Cancers : le nombre de nouveaux cas a presque doublé en france, 2023. [lien](#).
- [20] Fabien Roussel. Rapport numéro 2643. *Assemblée Nationale*, 2024.
- [21] Mathilde Roy. L’assurance maladies graves en 10 questions. *ProtégezVous*, 2023.
- [22] Santé publique France. Des maladies fréquentes dont la survie augmente, 2024.
- [23] Nagaratnam Suganthan. Random vector functional link network: recent developments, applications, and future directions, 2023.
- [24] Frédéric Planchet Thierry Moudiki and Areski Cousin. Multiple time series forecasting using quasi-randomized functional link neural networks. *ISFA*, 2018.
- [25] Nicolas Vandeput. Forecast kpis: Rmse, mae, mape et bias. *Towards Data Science*, 2019.
- [26] Vie publique. Cancers : le nombre de nouveaux cas a presque doublé en france, 2023.