

**Mémoire présenté le : 7 Novembre 2022
pour l'obtention du diplôme
de Statisticien Mention Actuariat
et l'admission à l'Institut des Actuares**

Par : Madame Amélie SIRIEYS

Titre du mémoire : Réduction du droit à l'oubli : construction d'une loi d'extra-mortalité et étude de l'impact sur les engagements d'un assureur

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

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

Membres présents du jury de la
filiale :

Signature : Entreprise : Optimind

Nom : Gildas ROBERT

Signature :



Directeur de mémoire en entreprise :

Nom : Pierre MORLON

Signature : Signature :



Membres présents du jury de
l'Institut des Actuares :

Invité :

Nom :

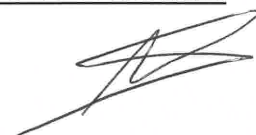
Signature :

**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

Le droit à l'oubli se définit comme étant un droit permettant à des individus touchés par une pathologie cancéreuse de ne pas la déclarer lors de la souscription à une assurance emprunteur, une fois passé un certain délai en rémission sans rechute. Historiquement, ce droit se fondait sur un délai de 10 ans pour les assurés ayant été diagnostiqués d'un cancer après 21 ans, et 5 ans pour ceux dont le diagnostic avait eu lieu avant leurs 21 ans. Depuis le 17 février 2022, avec l'adoption de la loi Lemoine, il a été uniformisé à 5 ans peu importe l'âge au diagnostic du patient. En ce sens, il implique davantage de risques d'antisélection pour les compagnies d'assurance qu'il convient de mesurer.

L'objectif de ce mémoire est alors de quantifier l'impact de la réduction du délai d'application du droit à l'oubli pour les individus diagnostiqués d'un cancer après leurs 21 ans. La base de données SEER recense un suivi du statut vital des patients atteints d'un cancer aux Etats-Unis. A partir de celles-ci, les probabilités de décès des suites du cancer diagnostiqué et traité, par genre, en fonction de l'âge de fin de traitement thérapeutique et l'ancienneté en rémission sont calculées. Finalement, à partir de ces tables, le niveau des engagements d'un assureur pour sa garantie décès peut être réévalué à la hausse, permettant une mesure économique de l'antisélection ajoutée.

Mots-clés : assurance emprunteur, loi Lemoine, droit à l'oubli, base SEER, table d'extra-mortalité des suites du cancer, niveau des engagements d'un assureur

Abstract

The right to be forgotten is defined as a right allowing individuals affected by a cancerous pathology not to declare it when taking out loan insurance, once a certain period of time in remission without relapse has passed. Historically, this right was based on a period of 10 years for policyholders who were diagnosed with cancer after the age of 21, and 5 years for those who were diagnosed before the age of 21. Since February 17, 2022, with the adoption of the Lemoine law, it has been standardized to 5 years regardless of the patient's age at diagnosis. In this sense, it implies more risks of anti-selection for the insurance companies that should be measured.

The objective of this dissertation is to quantify the impact of reducing the delay in the application of the right to be forgotten for individuals diagnosed with cancer after their 21st birthday. The SEER database tracks the vital status of cancer patients in the United States. From these, the probabilities of death from diagnosed and treated cancer, by gender, as a function of age at the end of therapeutic treatment and length of time in remission are calculated. Finally, based on these tables, the level of an insurer's commitments for its death benefit can be re-evaluated upwards, allowing an economic measure of the added anti-selection.

Keywords: loan insurance, Lemoine law, right to be forgotten, SEER base, table of extra-mortality of the consequences of cancer, level of an insurer's commitments

Note de Synthèse

Mise en contexte

Aujourd'hui, l'octroi d'un prêt est conditionné à la souscription d'une assurance emprunteur. Toutefois, pour les profils à risque élevé de santé, les compagnies d'assurance tendent à les refuser ou à appliquer une surprime. Ces profils ont alors des difficultés pour réaliser leurs projets de vie. Puisque le législateur souhaite qu'un grand nombre de Français puisse accéder à l'achat immobilier, différentes mesures ont vu le jour depuis 1991. Parmi celles-ci, la convention AERAS (S'Assurer et Emprunter avec un Risque Aggravé de Santé) permet aux personnes ayant ou ayant eu des problèmes médicaux de s'assurer pour l'attribution de leur emprunt. Révisée en 2020, elle contient notamment :

- une grille de référence, qui est une liste de pathologies pour lesquelles un accès à l'assurance emprunteur est accordé dans des conditions presque identiques à un profil d'assuré sans risque de santé après un temps en rémission défini pour chaque pathologie ;
- un droit à l'oubli, qui permet à des individus touchés par une pathologie cancéreuse de ne pas la déclarer lors de la souscription à une assurance emprunteur, une fois passé un certain délai en rémission sans rechute. Avant février 2022, ce droit se fondait sur un délai de 10 ans pour les assurés ayant été diagnostiqués d'un cancer après 21 ans, et 5 ans pour ceux dont le diagnostic avait eu lieu avant leurs 21 ans.

Le 17 février 2022, le projet de loi Lemoine, porté par la députée Patricia Lemoine, est examiné et adopté par l'Assemblée Nationale et le Sénat. Il comporte quatre avancées majeures :

- la résiliation à tout moment pour les assurés ;
- la suppression du questionnaire médical ;
- un devoir d'information de par l'assureur auprès de ses assurés ;
- la réduction du délai d'application du droit à l'oubli. Plus particulièrement, le délai d'application du droit à l'oubli est uniformisé à 5 ans pour tous les assurés, peu importe leur âge au diagnostic.

Problématique actuarielle

La réduction du délai d'application du droit à l'oubli accentue le risque d'antisélection pour les organismes assureurs. Désormais, les individus diagnostiqués d'un cancer après leurs 21 ans et ayant une ancienneté en rémission entre 5 et 10 ans entrent dans le nouveau périmètre d'application du droit à l'oubli. En ce sens, il devient primordial d'anticiper cette hausse du phénomène d'antisélection en quantifiant l'impact de la réduction du délai du droit à l'oubli de 10 à 5 ans, pour les personnes atteintes d'un cancer après leurs 21 ans. Il s'agit là de l'objectif de ce mémoire.

Il est évident qu'un assuré appliquant son droit à l'oubli au moment de la souscription d'une assurance emprunteur a un risque plus élevé de mortalité, dû au cancer, qu'un assuré n'ayant pas été touché par un cancer, toutes choses égales par ailleurs. Ainsi, une problématique claire se dégage pour évaluer les risques encourus suite à l'application du droit à l'oubli : quelle est la probabilité de décès des

suites du cancer diagnostiqué et traité d'un assuré tout au long de sa période de rémission ? A partir de ces probabilités, une compagnie d'assurance est alors en mesure de réévaluer ses risques couverts.

Afin de répondre à cette problématique, il est nécessaire de sélectionner des données qui permettent de suivre individuellement des personnes touchées par un cancer, en phase de rémission, et dont la cause de décès est précisée. Puis, à partir de ces données, les probabilités de décès des suites du cancer diagnostiqué et traité au cours de la phase de rémission peuvent être estimées avec les outils d'analyse de survie. Finalement, ces tables d'extra-mortalité sont appliquées dans le calcul des engagements d'un assureur pour la garantie décès d'un de ses produits emprunteurs.

Les probabilités de décès des suites du cancer diagnostiqué et traité

Les données SEER

En France, les données médicales individuelles de suivi des pathologies cancéreuses ne sont pas accessibles au grand public. En revanche, il existe aux Etats-Unis la base de données SEER, en open-data, qui recense un suivi du statut vital des patients atteints d'un cancer aux Etats-Unis. C'est à partir de ces données que sont évaluées les probabilités de décès des suites du cancer diagnostiqué et traité pendant la phase de rémission.

Dans la dernière version de novembre 2021, le programme SEER comporte un maximum de 267 variables et enregistre les cancers parmi 48% de l'ensemble de la population des Etats-Unis. Pour cette étude, l'extrait sélectionné couvre approximativement 26.5% de la population des Etats-Unis (d'après la population enregistrée en 2010). Ainsi, ces travaux reposent, avant retraitements, sur 7 788 941 tumeurs malignes recensées sur la période 2000-2019.

L'objectif de ces travaux est de comprendre le risque encouru par les assureurs suite à la réduction de moitié du délai d'application du droit à l'oubli. Ce délai commence à partir de l'âge de fin de traitement thérapeutique suivi. Il convient donc d'étudier, pour chaque année passée depuis la fin du traitement, la probabilité de décéder à cause du cancer diagnostiqué et traité. Ainsi, il est primordial de disposer de l'âge à la fin du protocole thérapeutique qui correspond à la date de début de la table d'extra-mortalité à construire. Toutefois, la base SEER ne contient pas cette information. A l'aide de différents retraitements sur la base et des temps moyens par thérapie, l'âge de fin de traitement de chaque individu est estimé pour la suite de l'étude.

Les hypothèses considérées

L'application du droit à l'oubli nécessite les conditions suivantes :

- l'assuré doit avoir achevé son traitement thérapeutique, entrant ainsi en phase de rémission, pour commencer le compte du délai d'application du droit à l'oubli ;
- il ne doit pas rechuter au cours de ce temps écoulé ;
- la dernière échéance de son prêt doit avoir lieu avant son 71^{ème} anniversaire.

Pour être en adéquation avec ces conditions et répondre à la problématique de ce mémoire, les hypothèses suivantes ont été considérées :

- les personnes atteintes d'un cancer métastasé rechutent des suites de cette pathologie et n'entrent jamais en phase de rémission. Elles sont donc exclues de cette étude ;
- les personnes touchées d'un cancer non métastasé ont une probabilité nulle de rechute pendant la phase de rémission ;

- si une personne est touchée par plusieurs cancers non métastasés, alors la fin de traitement thérapeutique est comptée à partir de la date du dernier diagnostic ;
- le temps de traitement thérapeutique est estimé selon les durées moyennes de la chirurgie, de la chimiothérapie et de la radiothérapie ;
- la situation épidémiologique des cancers aux Etats-Unis est identique à celle de la France.

Construction des lois d'extra-mortalité selon l'âge de fin de traitement thérapeutique et l'ancienneté en rémission

A partir du test d'indépendance du χ^2 , qui indique que la réalisation des décès des suites des cancers diagnostiqués et traités dépend du genre de l'individu, une loi d'extra-mortalité est conçue pour les femmes, et une seconde pour les hommes. Puis, l'application du test du log-rank permet de détecter une différence significative des taux d'extra-mortalité entre la période 2000-2008 et 2009-2019. Par conséquent, l'estimation des taux d'extra-mortalité repose sur la période la plus récente des diagnostics (2009-2019).

Pour l'estimation des taux bruts d'extra-mortalité, l'estimateur des moments de Hoem et l'estimateur de Kaplan-Meier sont considérés. Puisque l'estimateur des moments de Hoem repose sur l'exposition au risque de chaque individu et qu'il n'existe pas de grande différence de tendance entre ces deux estimateurs, les taux bruts calculés par l'estimateur des moments de Hoem sont retenus (figure 1).

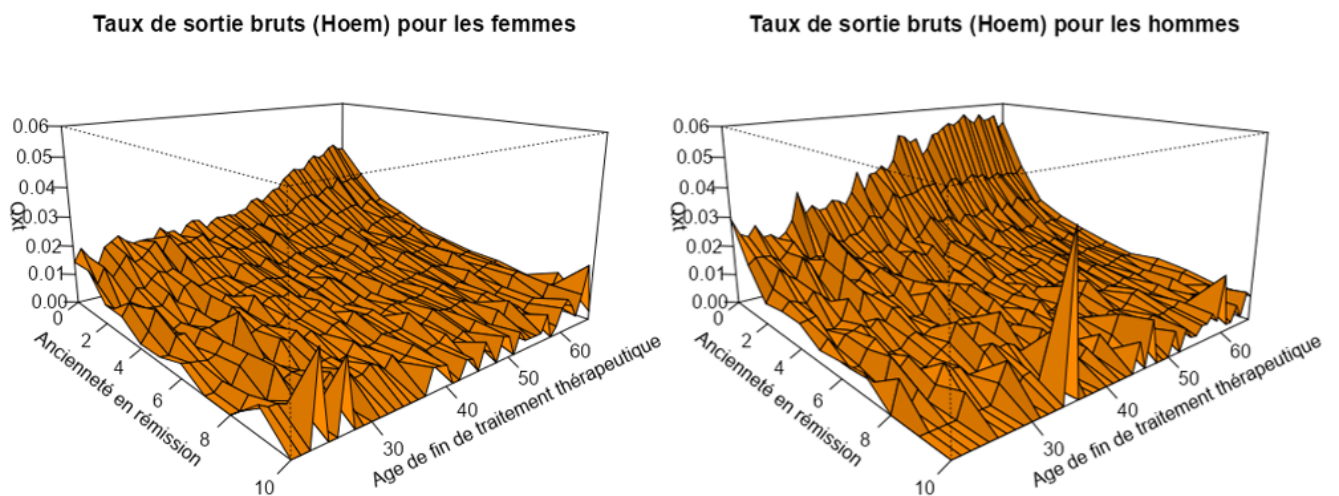


FIGURE 1 – Taux bruts retenus d'extra-mortalité des suites du cancer diagnostiqué et traité (Hoem)

Toutefois, ces taux bruts comportent des irrégularités causées par les fluctuations d'échantillonnage. Afin de les corriger, les méthodes des P-splines et de Whittaker-Henderson sont appliquées. Pour déterminer quelle méthode d'ajustement doit être sélectionnée, les tests du χ^2 d'adéquation et des changements de signes sont réalisés. Ils indiquent que la méthode de Whittaker-Henderson ajuste les taux bruts de la méthode des moments de Hoem de manière fidèle à l'extra-mortalité empirique, et vérifie le critère de régularité du lissage puisqu'il y a autant de taux sur-lissés que de taux sous-lissés. En ce sens, les probabilités finales de décès des suites du cancer diagnostiqué et traité reposent sur le calcul à l'aide de l'estimateur des moments de Hoem et la méthode de lissage de Whittaker-Henderson (figure 2).

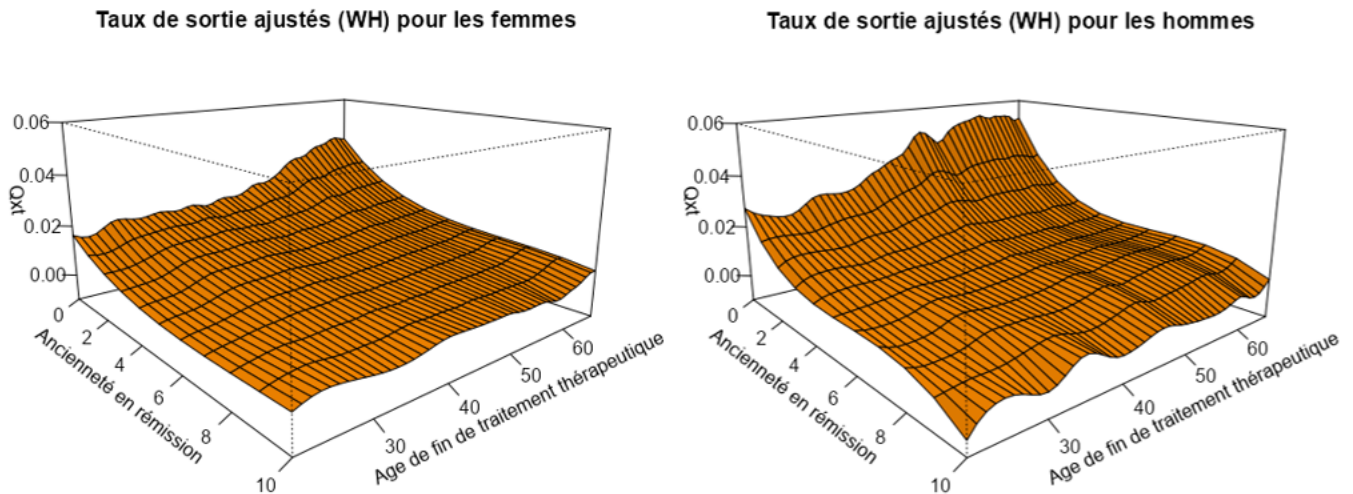


FIGURE 2 – Taux d’extra-mortalité lissés retenus avec $(\alpha, \beta, z_v, z_h) = (1, 0.001, 3, 4)$

Application des tables d’extra-mortalité dans le calcul des engagements d’un assureur

L’application actuarielle de ce mémoire repose sur des données modifiées d’un des portefeuilles emprunteur du Groupe Magnolia. En ce sens, les résultats présentés ne reflètent pas les engagements réels du groupe.

Le portefeuille étudié a les principales caractéristiques suivantes :

- Produit : assurance des crédits immobiliers pour les particuliers ;
- Age maximale de couverture du décès pour le produit souscrit : 89 ans ;
- Date de début de souscription : 01/01/2013 ;
- Date exacte d’extraction : 06/03/2022 ;
- Nombre de polices souscrites en cours au moment de l’extraction : 49 533 ;
- Répartition des sexes : 49,0% de femmes et 51,0% d’hommes.

Puisque Magnolia.fr n’a pas accès au passif de la pathologie cancéreuse de l’assuré entrant, la mesure de l’impact se fonde sur différents scénarios de parts d’assurés appliquant leur droit à l’oubli. Pour chacun de ces scénarios, le niveau des engagements total est mesuré.

De manière générale, le tableau 1 montre qu’il existe un impact sur le niveau des engagements de Magnolia.fr suite à l’application du nouveau droit à l’oubli pour la garantie décès. Cet impact varie linéairement entre 0.00% pour le scénario le plus optimiste (aucun entrant), et 0.432% pour le scénario le plus pessimiste. Bien que ces taux soient inférieurs à 1%, il s’agit, pour le scénario le plus pessimiste, d’un coût de près de 1 000 000 € pour un an et deux mois d’application de cette nouvelle législation.

TABLE 1 – Calcul des niveaux des engagements de Magnolia.fr selon les scénarios

Scénarios	% d'entrants appliquant le droit à l'oubli	Ecart moyen par assuré pondéré par les capitaux assurés (en €)	Engagement total de l'assureur (en €)	Evolution par rapport au scénario initial
0	0%		205 221 885 38	0.000%
1	1%	2 408.35	205 269 740.81	0.023%
2	2%	2 277.53	205 315 640.09	0.046%
3	3%	3 090.94	205 394 216.83	0.084%
4	4%	3 539.76	205 488 290.65	0.130%
5	5%	3 563.93	205 571 034.57	0.170%
6	6%	3 550.59	205 624 539.19	0.196%
7	7%	3 355.21	205 682 248.35	0.224%
8	8%	3 188.72	205 722 314.97	0.244%
9	9%	3 070.29	205 754 640.27	0.260%
10	10%	3 184.53	205 830 720.84	0.297%
11	11%	3 336.15	205 910 643.76	0.336%
12	12%	3 331.37	205 968 152.07	0.364%
13	13%	3 231.10	206 009 844.49	0.384%
14	14%	3 262.22	206 075 765.54	0.416%
15	15%	3 171.43	206 108 773.76	0.432%

Conclusion

Les travaux effectués montrent qu'il existe une probabilité, plafonnée à 2%, pour les personnes atteintes d'un cancer non métastasé et traité, de décéder des suites de ce même cancer entre la cinquième et la dixième année passée en rémission quelque soit leur âge à la fin du traitement thérapeutique et leur ancienneté en rémission. De ces probabilités et d'un portefeuille modifié d'un assureur, une évolution à la hausse de ses engagements a été quantifiée pour une de ses garanties décès emprunteur. Bien que cette hausse soit relativement faible, puisqu'elle est inférieure à 1% y compris dans le scénario le plus pessimiste, il est important de souligner qu'il ne s'agit que d'un des pans de la loi Lemoine et que cet impact dépend clairement de la prévalence des cancers en France.

Les probabilités ont été estimées à partir de différentes hypothèses susmentionnées. Parmi celles-ci, celle sur l'estimation de l'âge de fin de traitement thérapeutique et celle sur l'équivalence des situations épidémiologiques des cancers aux Etats-Unis et en France ont été challengées. Les résultats obtenus sont ainsi relativement sensibles à ces deux hypothèses principales. Il serait alors intéressant de définir précisément l'âge de fin de traitement thérapeutique à l'aide de connaissances plus profondes en oncologie et de quantifier exactement la différence des situations épidémiologiques entre les deux pays.

Finalement, toujours pour généraliser ces travaux, il pourrait être judicieux d'intégrer une modélisation de rechute des patients en intégrant les cancers métastasés, permettant ainsi de supprimer les trois premières hypothèses considérées.

Synthesis note

Contextualization

Today, the granting of a loan is conditional on the subscription of a borrower's insurance. However, for profiles with high health risks, insurance companies tend to refuse them or apply an additional premium. These profiles then have difficulties to realize their life projects. Since the legislator wants a large number of French people to be able to buy real estate, various measures have been taken since 1991. Among these, the AERAS convention (S'Assurer et Emprunter avec un Risque Aggravé de Santé) allows people who have or have had medical problems to insure themselves for the granting of their loan. Revised in 2020, it contains in particular:

- a reference grid, which is a list of pathologies for which access to loan insurance is granted under almost identical conditions to an insured profile without health risk after a time in remission defined for each pathology ;
- a right to forget, which allows individuals affected by a cancerous pathology not to declare it when taking out loan insurance, once a certain period of time in remission without relapse has passed. Prior to February 2022, this right was based on a 10-year period for policyholders diagnosed with cancer after age 21, and 5 years for those diagnosed before age 21.

On February 17, 2022, the Lemoine bill, introduced by Deputy Patricia Lemoine, was examined and adopted by the National Assembly and the Senate. It includes four major advances:

- the cancellation at any time for the insured persons ;
- the suppression of the medical questionnaire;
- a duty of information by the insurer to its policyholders;
- a reduction in the period of application of the right to be forgotten. More specifically, the time limit for the application of the right to be forgotten is standardized at 5 years for all policyholders, regardless of their age at diagnosis.

Actuarial issues

The reduction in the time limit for the right to forget increases the risk of anti-selection for insurers. From now on, individuals diagnosed with cancer after their 21st birthday and having been in remission for between 5 and 10 years will be included in the new scope of the right to be forgotten. In this sense, it becomes essential to anticipate this increase in the phenomenon of anti-selection by quantifying the impact of reducing the time limit for the right to be forgotten from 10 to 5 years for people diagnosed with cancer after the age of 21. This is the objective of this brief.

It is clear that a policyholder who applies the right to be forgotten at the time of taking out loan insurance has a higher risk of mortality due to cancer than a policyholder who has not been affected by cancer, all other things being equal. Thus, a clear problem emerges to evaluate the risks incurred following the application of the right to be forgotten: what is the probability of death from cancer

diagnosed and treated by an insured throughout the period of remission? Based on these probabilities, an insurance company is then in a position to reassess its covered risks.

In order to answer this question, it is necessary to select data that allows for the individual follow-up of people affected by cancer, in remission phase, and whose cause of death is specified. Then, from these data, the probabilities of death from cancer diagnosed and treated during the remission phase can be estimated with survival analysis tools. Finally, these extra-mortality tables are applied in the calculation of an insurer's commitments for the death cover of one of its loan products.

The probability of death from diagnosed and treated cancer

The SEER data

In France, individual medical data on the follow-up of cancer pathologies are not accessible to the general public. However, in the United States, there is the SEER database, in open-data, which tracks the vital status of cancer patients in the United States. It is from this data that the probability of death from cancer diagnosed and treated during the remission phase is evaluated.

In the latest version of November 2021, the SEER program has a maximum of 267 variables and records cancers among 48% of the total US population. For this study, the selected extract covers approximately 26.5% of the US population (based on the 2010 registered population). Thus, this work is based on 7,788,941 malignancies identified during the period 2000-2019 before adjustments.

The objective of this work is to understand the risk incurred by insurers following the reduction by half of the period of application of the right to be forgotten. This period starts from the age of the end of the therapeutic treatment. It is therefore necessary to study, for each year since the end of the treatment, the probability of dying because of the diagnosed and treated cancer. Thus, it is essential to have the age at the end of the therapeutic protocol which corresponds to the starting date of the extra-mortality table to be constructed. However, the SEER database does not contain this information. Using various reprocessing operations on the database and the average times per therapy, the age at the end of treatment for each individual is estimated for the rest of the study.

The assumptions considered

The application of the right to be forgotten requires the following conditions:

- the insured must have completed his therapeutic treatment, thus entering into a phase of remission, in order to start counting the period of application of the right to be forgotten;
- he must not relapse during this elapsed time;
- the last payment on his loan must be before his 71th birthday.

To be consistent with these conditions and to answer the problematic of this thesis, the following hypotheses have been considered:

- people with metastatic cancer relapse as a result of this pathology and never enter remission. They are therefore excluded from this study;
- people with non-metastatic cancer have a zero probability of relapse during the remission phase;
- if a person has more than one non-metastatic cancer, then the end of treatment is counted from the date of the last diagnosis;
- if a person is affected by several non-metastatic cancers, then the end of therapeutic treatment is counted from the date of the last diagnosis; the time of therapeutic treatment is estimated

- according to the average durations of surgery, chemotherapy and radiotherapy;
- the epidemiological situation of cancer in the United States is identical to that in France.

Construction of extra-mortality laws according to age at the end of therapeutic treatment and length of time in remission

From the χ^2 test of independence, which indicates that the realization of deaths from diagnosed and treated cancers depends on the gender of the individual, an extra-mortality law is designed for women, and a second one for men. Then, the application of the log-rank test allows to detect a significant difference in the extra-mortality rates between the period 2000-2008 and 2009-2019. Therefore, the estimation of extra-mortality rates is based on the most recent period of diagnoses (2009-2019).

For the estimation of crude extra-mortality rates, the Hoem moment estimator and the Kaplan-Meier estimator are considered. Because the Hoem moments estimator relies on each individual's risk exposure and there is not a large difference in trend between these two estimators, the crude rates calculated by the Hoem moments estimator are retained (figure 3).



FIGURE 3 – Selected crude rates of extra-mortality from diagnosed and treated cancer (Hoem)

However, these crude rates contain irregularities caused by sampling fluctuations. To correct for these, the P-splines and Whittaker-Henderson methods are applied. To determine which adjustment method should be selected, the χ^2 tests of adequacy and sign changes are performed. They indicate that the Whittaker-Henderson method adjusts the raw Hoem moments rates faithfully to the empirical extra-mortality, and verifies the smoothness regularity criterion since there are as many over-smoothed rates as under-smoothed rates. In this sense, the final probabilities of death from diagnosed and treated cancer are based on the calculation using the Hoem moments estimator and the Whittaker-Henderson smoothing method (figure 4).

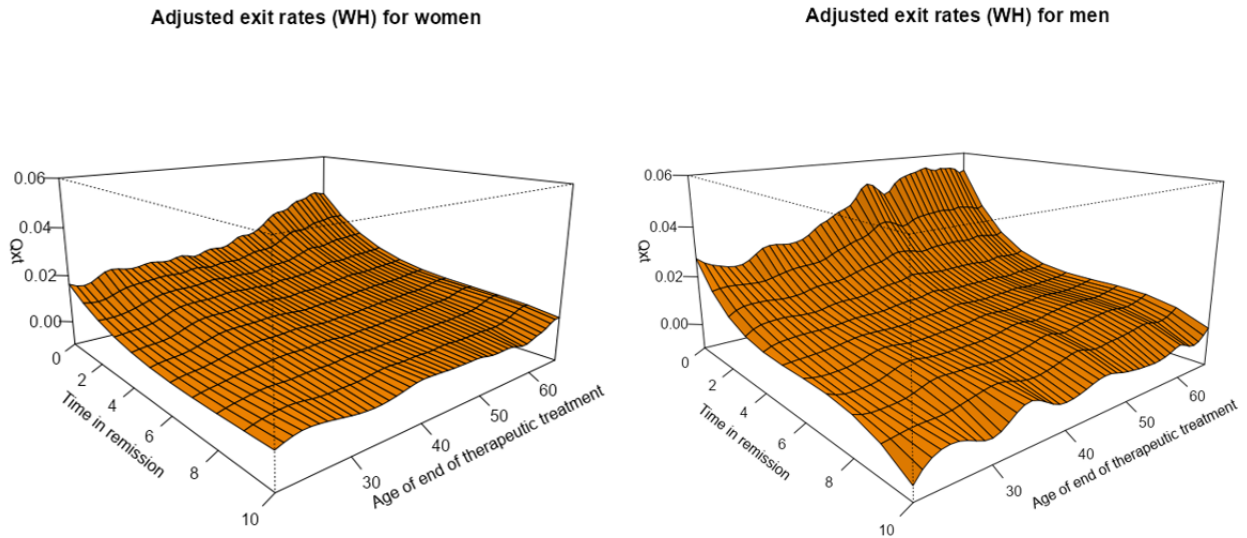


FIGURE 4 – Smoothed extra-mortality rates retained with $(\alpha, \beta, z_v, z_h) = (1, 0.001, 3, 4)$

Application of extra-mortality tables in the calculation of an insurer's liabilities

The actuarial application in this brief is based on modified data from one of the Magnolia Group's borrower portfolios. As such, the results presented do not reflect the actual liabilities of the group.

The portfolio studied has the following main characteristics:

- Product: mortgage insurance for individuals;
- Maximum age of death cover for the product underwritten: 89 years;
- Date of beginning of subscription: 01/01/2013 ;
- Exact date of extraction: 06/03/2022 ;
- Number of policies in force at the time of extraction: 49,533;
- Gender distribution: 49.0% female and 51.0% male.

Since Magnolia.fr does not have access to the liabilities of the cancerous pathology of the incoming policyholder, the measurement of the impact is based on different scenarios of shares of policyholders applying their right to be forgotten. For each of these scenarios, the level of total commitments is measured.

Generally speaking, the table 2 shows that there is an impact on the level of Magnolia.fr's commitments following the application of the new right to be forgotten for the death benefit. This impact varies linearly between 0.00% for the most optimistic scenario (no inflow), and 0.432% for the most pessimistic scenario. Although these rates are less than 1%, for the most pessimistic scenario, the cost is nearly €1,000,000 for one year and two months of application of this new legislation.

TABLE 2 – Calculation of the levels of Magnolia.fr commitments according to the scenarios

Scenarios	% of entrants applying the right to be forgotten	Average gap per insured weighted by insured capital (in €)	Total commitment of the insurer (in €)	Evolution compared to the initial scenario
0	0%		205,221,885.38	0.000%
1	1%	2,408.35	205,269,740.81	0.023%
2	2%	2,277.53	205,315,640.09	0.046%
3	3%	3,090.94	205,394,216.83	0.084%
4	4%	3,539.76	205,488,290.65	0.130%
5	5%	3,563.93	205,571,034.57	0.170%
6	6%	3,550.59	205,624,539.19	0.196%
7	7%	3,355.21	205,682,248.35	0.224%
8	8%	3,188.72	205,722,314.97	0.244%
9	9%	3,070.29	205,754,640.27	0.260%
10	10%	3,184.53	205,830,720.84	0.297%
11	11%	3,336.15	205,910,643.76	0.336%
12	12%	3,331.37	205,968,152.07	0.364%
13	13%	3,231.10	206,009,844.49	0.384%
14	14%	3,262.22	206,075,765.54	0.416%
15	15%	3,171.43	206,108,773.76	0.432%

Conclusion

The work carried out shows that there is a probability, capped at 2%, for people with non-metastasized and treated cancer to die from the consequences of this same cancer between the fifth and tenth year spent in remission, whatever their age at the end of the therapeutic treatment and their length of time in remission. From these probabilities and a modified portfolio of an insurer, an upward trend in its commitments was quantified for one of its death benefits. Although this increase is relatively small, since it is less than 1% even in the most pessimistic scenario, it is important to emphasize that this is only one aspect of the Lemoine law and that this impact clearly depends on the prevalence of cancer in France.

The probabilities were estimated on the basis of the different assumptions mentioned above. Among these, the one on the estimation of the age of the end of therapeutic treatment and the one on the equivalence of the epidemiological situations of cancers in the United States and in France have been challenged. The results obtained are thus relatively sensitive to these two main hypotheses. It would therefore be interesting to define the age of completion of therapeutic treatment precisely with the help of more in-depth knowledge of oncology and to quantify exactly the difference in epidemiological situations between the two countries.

Finally, in order to generalize this work, it might be judicious to integrate a model of patient relapse by integrating metastasized cancers, thus eliminating the first three hypotheses considered.

Remerciements

Dans un premier temps, je tiens à remercier Christophe EBERLÉ, président fondateur d'Optimind, pour m'avoir permis de réaliser mon alternance au sein de son entreprise dans les meilleures conditions. Je remercie également les *partners*, Valérie DEPPE et Gildas ROBERT, pour m'avoir accueilli au sein de leur *practice Actuarial & Financial Services*.

J'adresse également mes remerciements au Groupe Magnolia, et plus particulièrement à Marina VIVARGENT, directrice technique et data du Groupe Magnolia, pour le partage de leurs données qui a permis de réaliser l'application actuarielle de ce mémoire.

Je remercie infiniment mes tuteurs en entreprise, Rachel BELLENGER et Pierre MORLON, pour leur encadrement, leur expertise, leur disponibilité, et leurs pertinents conseils tout au long de la réalisation de ce mémoire. Je tiens à remercier également Emmanuel BERTHELÉ et André GRONDIN pour leurs conseils avisés.

Au même titre, je souhaite remercier l'ensemble des collaborateurs d'Optimind avec qui j'ai pu échanger pour leur bienveillance et leur expertise actuarielle, et notamment mes collègues alternants et stagiaires Aurélie AMARD, Amélie DE LA HAYE, Ana FERREIRA PIRES DE SOUSA et Pablo GASSIOT. J'adresse mes sincères remerciements à Elsa LAVAL pour ses précieuses relectures et sa bonne humeur quotidienne.

Merci à Guillaume BIESSY, mon tuteur académique, pour ses conseils avisés, ainsi qu'à l'ensemble du corps professoral de l'ISUP, sous la direction d'Olivier LOPEZ, pour la qualité de l'enseignement dispensé sur les deux années du Master Actuariat.

Enfin, je remercie ma famille et mes amis pour leur soutien sans faille. Plus particulièrement, j'adresse toute ma gratitude à mes parents et à ma soeur qui ne cessent de me guider et me soutenir.

Table des matières

Résumé	3
Abstract	4
Note de synthèse	5
Synthesis note	10
Remerciements	15
Table des matières	16
Introduction	22
1 Contexte de l'étude	23
1.1 Généralités sur l'assurance emprunteur	23
1.2 La réglementation du marché	28
1.3 Situation globale d'épidémiologie des cancers en France	33
2 Les données SEER	37
2.1 Introduction aux données SEER	37
2.2 Pré-sélection des données du programme SEER	38
2.3 Complétion des données de la base SEER	40
2.4 Retraitements de la base SEER	41
2.5 Retraitements des observations pré-sélectionnées	44
2.6 Construction de l'âge à la fin du traitement thérapeutique suivi	45
2.7 Analyse statistique	48
2.8 Les segmentations des taux d'extra-mortalité	57

<i>TABLE DES MATIÈRES</i>	17
3 Construction des lois	62
3.1 Notions préliminaires pour l'analyse de survie	62
3.2 Estimation des taux bruts	65
3.3 Ajustement et lissage	67
3.4 Sélection et validation de la table finale	70
3.5 Application numérique sur les données SEER retraitées	72
4 Mesure de l'impact	83
4.1 Calcul du niveau des engagements d'un assureur	83
4.2 Mesure de l'impact du droit à l'oubli pour Magnolia.fr	87
4.3 Etude de sensibilité des résultats aux hypothèses principales	92
4.4 Limites, réflexions et ouvertures des travaux	95
Conclusion	98
Bibliographie	99
Annexes	101

Table des figures

1	Taux bruts retenus d'extra-mortalité des suites du cancer diagnostiqué et traité (Hoem)	7
2	Taux d'extra-mortalité lissés retenus avec $(\alpha, \beta, z_v, z_h) = (1, 0.001, 3, 4)$	8
3	Selected crude rates of extra-mortality from diagnosed and treated cancer (Hoem)	12
4	Smoothed extra-mortality rates retained with $(\alpha, \beta, z_v, z_h) = (1, 0.001, 3, 4)$	13
1.1	Montant des cotisations en € par type d'acteur (FRANCE ASSUREURS, 2021b)	31
1.2	Répartition des demandes d'assurance de prêts par année (FRANCE ASSUREURS, 2021a)	32
1.3	Taux d'incidence des cancers en France selon l'année et par âge à l'échelle logarithmique, tous les cancers confondus (LE GUYADER-PEYROU et al., 2019)	35
1.4	Taux de mortalité des cancers en France selon l'année et par âge à l'échelle logarithmique, tous les cancers confondus (LE GUYADER-PEYROU et al., 2019)	36
2.1	Étape préliminaire pour extraire les données SEER	38
2.2	Étape préliminaire bis pour extraire les données SEER	40
2.3	Répartition des cancers diagnostiqués en 2018 selon le sexe	50
2.4	Répartition des cancers selon l'ancienneté en rémission à la sortie	51
2.5	Distribution des âges révolus à la fin du traitement thérapeutique suivi	52
2.6	Distribution des anciennetés en rémission à la sortie (décès ou censure)	53
2.7	Durée moyenne en année en rémission à la sortie (censure et décès) selon l'âge révolu à la fin du traitement thérapeutique	54
2.8	Répartition des sorties selon l'âge révolu à la fin du traitement thérapeutique	55
2.9	Diagramme en violon de l'ancienneté en rémission selon le type de sortie	56
2.10	Répartition des décès cancers selon l'ancienneté en rémission au moment du décès	57
2.11	Distribution des décès cancer selon le sexe et l'ancienneté, puis le sexe et l'âge à la fin du traitement thérapeutique	58
3.1	Exposition et nombre de décès des suites du cancer pour les femmes	73

3.2	Exposition et nombre de décès des suites du cancer pour les hommes	73
3.3	Taux bruts d'extra-mortalité des suites du cancer diagnostiqué et traité (Hoem)	74
3.4	Taux bruts d'extra-mortalité des suites du cancer diagnostiqué et traité (Kaplan-Meier)	75
3.5	Minimisation du BIC selon les variations de λ_x et λ_t	76
3.6	Taux bruts d'extra-mortalité lissés par P-splines pour les femmes avec $(\lambda_x, \lambda_t) = (100, 1000)$ et pour les hommes avec $(\lambda_x, \lambda_t) = (1000, 100)$	77
3.7	Taux d'extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 1, 2, 2)$	78
3.8	Taux d'extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 1, 3, 3)$	78
3.9	Taux d'extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 1, 3, 4)$	79
3.10	Taux d'extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 0.01, 3, 4)$	79
3.11	Taux d'extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 0.001, 3, 4)$	80
4.1	Distribution des souscriptions par année	88
4.2	Distribution des âges à la souscription	88
4.3	Distribution des capitaux assurés initiaux par tête inférieurs au quantile à 99%	89
4.4	Taux d'extra-mortalité ajustés selon l'âge à la fin du traitement thérapeutique vieilli de 6 mois	93

Liste des tableaux

1	Calcul des niveaux des engagements de Magnolia.fr selon les scénarios	9
2	Calculation of the levels of Magnolia.fr commitments according to the scenarios	14
2.1	Durées minimales et maximales fixées pour chaque traitement	46
2.2	Classement par ordre croissant des degrés de gravité des cancers établi	46
2.3	Temps de traitement de la chimiothérapie selon le degré de gravité du cancer traité	47
2.4	Répartition des observations par genre et année de diagnostic	48
2.5	Statistiques sur l'âge au diagnostic selon le sexe	49
2.6	Répartition des 10 cancers les plus représentés sur 2000-2019	49
2.7	Répartition des 10 cancers les plus létaux sur 2000-2019	50
2.8	Statistiques sur la durée de traitement estimée en mois entiers	52
2.9	Statistiques sur le nombre de mois entiers écoulés entre le diagnostic et le début du traitement	52
2.10	Statistiques sur l'âge à la fin de traitement thérapeutique	52
2.11	Statistiques sur l'ancienneté en rémission à la sortie (décès ou censure) en années	53
2.12	Effectifs par sortie	55
2.13	Résultats du test d'indépendance du khi-deux	59
2.14	Résultats des tests du log-rank sur la base des femmes et celle des hommes	60
3.1	Résultats des applications du test du χ^2	81
3.2	Résultats des applications du test des changements de signes	81
4.1	Tableau d'amortissement pour un assuré i dont le prêt est à amortissements constants	84
4.2	Tableau d'amortissement pour un assuré i dont le prêt est à mensualités constantes	85
4.3	Tableau d'amortissement pour un assuré i dont le prêt est in-fine	85
4.4	Statistiques sur les capitaux initiaux assuré par tête (CA)	89

4.5	Pourcentage d'entrants respectant la condition d'âge pour le droit à l'oubli	90
4.6	Paramètres des différents scénarios	91
4.7	Calcul des niveaux des engagements de Magnolia.fr selon les scénarios	91
4.8	Calcul des niveaux des engagements de Magnolia.fr après vieillissement de l'âge de fin de traitement thérapeutique (H1)	94
4.9	Taux standardisés monde d'incidence et de mortalité du cancer par sexe en France et aux Etats-Unis en 2020 (INTERNATIONAL AGENCY FOR RESEARCH ON CANCER, 2022)	94
4.10	Comparaison sur les taux du tableau 4.9	94
4.11	Calcul des niveaux des engagements de Magnolia.fr après ajustement au cas français (H2)	95
4.12	Dictionnaire des 48 variables sélectionnées	103
4.13	Création de la variable Cancer	104
4.14	Création de la variable Grade	106
4.15	Table d'extra-mortalité du cancer non métastaté diagnostiqué et traité pour les femmes selon l'âge à la fin du traitement thérapeutique et l'ancienneté en rémission	107
4.16	Table d'extra-mortalité du cancer non métastaté diagnostiqué et traité pour les hommes selon l'âge à la fin du traitement thérapeutique et l'ancienneté en rémission	108
4.17	Table d'extra-mortalité du cancer non métastaté diagnostiqué et traité pour les femmes selon l'âge à la fin du traitement thérapeutique et l'ancienneté en rémission - Age de fin de traitement thérapeutique vieilli de 6 mois	109
4.18	Table d'extra-mortalité du cancer non métastaté diagnostiqué et traité pour les hommes selon l'âge à la fin du traitement thérapeutique et l'ancienneté en rémission - Age de fin de traitement thérapeutique vieilli de 6 mois	110

Introduction

Le 17 février 2022, le projet de loi pour la libéralisation du marché de l'assurance emprunteur des crédits immobiliers, porté par la députée Patricia Lemoine, est examiné et adopté par l'Assemblée Nationale et le Sénat. Il est également enrichi de dispositions pour une assurance plus inclusive. Ainsi, depuis le 1^{er} juin 2022, la sélection médicale est supprimée pour certains prêts immobiliers et le délai du droit à l'oubli est réduit à 5 ans. Par conséquent, ces dernières avancées législatives accentuent le risque d'antisélection pour les organismes assureurs qui doivent alors adapter leurs offres. C'est dans ce contexte que cette étude prend son sens. Plus particulièrement, ces travaux permettent de quantifier l'impact de la réduction du délai du droit à l'oubli de 10 à 5 ans, pour les personnes atteintes d'un cancer après leurs 21 ans.

Un assuré appliquant son droit à l'oubli au moment de la souscription d'une assurance emprunteur a un risque plus élevé de mortalité, dû au cancer, qu'un assuré n'ayant pas été touché par un cancer, toutes choses égales par ailleurs. En ce sens, une problématique claire se dégage pour évaluer les risques encourus suite à l'application du droit à l'oubli : quelle est la probabilité de décès des suites du cancer diagnostiqué et traité d'un assuré tout au long de sa période de rémission ? Avec la connaissance de ces probabilités, un assureur peut alors réévaluer économiquement ses risques couverts et quantifier leur évolution suite à la nouvelle législation.

Ce mémoire s'organise en quatre chapitres. Le premier chapitre permet de mettre en contexte ces travaux en présentant le marché de l'assurance emprunteur, ses évolutions législatives et la situation épidémiologique des cancers en France. Puis, le deuxième chapitre présente les données de la base SEER, qui recense un suivi du statut vital des patients atteints d'un cancer aux Etats-Unis, ainsi que les retraitements effectués et les hypothèses considérées pour les adapter au cadre strict d'application du droit à l'oubli en France. Une fois les données remaniées, l'extra-mortalité des suites du cancer diagnostiqué et traité en phase de rémission est calculée selon l'âge à la fin du traitement thérapeutique et l'ancienneté en rémission, au cours du troisième chapitre. Finalement, pour analyser l'impact des nouveaux risques entrants avec la nouvelle législation, les tables d'extra-mortalité construites sont intégrées au calcul des engagements d'un assureur.

Chapitre 1

L'assurance emprunteur et le cancer en France

Ce chapitre initial définit, dans un premier temps, les notions fondamentales autour de l'assurance emprunteur qui sont nécessaires à la correcte appréciation du sujet. Puis, nous détaillons les évolutions réglementaires qui ont mené au projet de loi Lemoine. Enfin, un point d'attention est donné quant à la situation en France des personnes atteintes d'un cancer afin de comprendre les enjeux de la réduction du délai de droit à l'oubli.

1.1 Généralités sur l'assurance emprunteur

L'assurance emprunteur est un produit de prévoyance temporaire qui permet à un emprunteur de se protéger, sur la durée de son prêt, de son risque d'insolvabilité suite à son décès, son invalidité, son incapacité et/ou à la perte de son emploi. Cette couverture de risque n'est pas une obligation légale dans le processus de demande de crédit mais elle est exigée par l'ensemble des organismes prêteurs. En effet, elle garantit au prêteur le remboursement de la totalité, ou d'une partie, de ses échéances dans le cas où l'emprunteur n'est pas en mesure de payer ses dettes suite à la réalisation d'un ou plusieurs événements inscrits au contrat souscrit.

Cette première partie est ainsi l'occasion de présenter brièvement le fonctionnement de cette assurance, ses particularités quant à la sélection des risques à l'adhésion, et le cadre législatif en vigueur.

1.1.1 Une brève description du produit

Chaque produit d'assurance dispose de ses particularités quant à sa distribution, les risques couverts et sa tarification. Il est ainsi primordial de définir celles de l'assurance emprunteur.

1.1.1.1 Les modes de distribution

Dans le domaine de la distribution des assurances de prêt, il est d'usage de séparer les acteurs du marché en deux classes :

- les bancassureurs, qui regroupent les organismes dont l'activité principale est la vente de produits bancaires couplée à des services d'assurance ;
- les délégataires d'assurance, qui concentrent tous les acteurs, autres que les bancassureurs, dont le coeur de métier est la vente d'assurances.

Historiquement, les bancassureurs distribuaient uniquement des contrats groupes d'assurance de prêt. Aujourd'hui, pour faire face aux offres du marché de la délégation d'assurance, ils proposent également des contrats individuels. Par opposition, les délégataires d'assurance vendent exclusivement des contrats individuels.

D'une part, le contrat groupe correspond à une assurance dite de groupe à adhésion facultative. Ainsi, comme l'indique l'article L141-1 du CODE DES ASSURANCES (2007), ce type de contrat est souscrit par l'établissement prêteur en vue de l'adhésion d'un ensemble de ses emprunteurs répondant aux critères de son choix. Ce type de contrat permet une mutualisation des risques entre les emprunteurs impliquant ainsi une faible segmentation tarifaire et un processus de souscription rapide.

D'autre part, le contrat individuel correspond à un contrat d'assurance individuelle classique conclut directement entre un assuré et un organisme assureur. A la différence du contrat groupe, les tarifs proposés sont individualisés à l'aide d'un processus de souscription détaillé.

Chacun de ces contrats est profitable à un profil précis d'assuré. Les contrats groupe intéressent davantage les emprunteurs dont la définition précise de leur situation personnelle tend à leur attribuer une surprime en assurance individuelle. D'une manière contraire, les assurances individuelles avantagent les assurés avec des profils de risque faible tels que les emprunteurs jeunes, qui déclarent ne pas fumer et n'ayant aucun antécédent médical.

Enfin, le COMITÉ CONSULTATIF DU SECTEUR FINANCIER (2020) fait état, dans son bilan de l'assurance emprunteur, de la répartition des ventes d'assurance emprunteur selon ces types de contrats. En 2019, les contrats groupe et les contrats individuels (proposés par le réseau bancaire et les assureurs) représentaient respectivement 74.5% et 25.5% des produits vendus. Ainsi, malgré une concurrence accrue, les contrats groupe dominent le marché de l'assurance emprunteur.

1.1.1.2 Le niveau de couverture

En application stricte de la législation, un établissement de crédit ne peut pas obliger son emprunteur à souscrire à une assurance de prêt. Aujourd'hui, il en fait une condition d'octroi pour l'acceptation d'un emprunt puisqu'elle lui permet d'être couvert contre l'insolvabilité de son emprunteur. Il est, toutefois, en mesure d'exiger un niveau de couverture minimum. Ce dernier dépend principalement de deux paramètres : la quotité assurée et les garanties assurantielles sélectionnées.

La quotité assurée

La quotité assurée correspond à la proportion du capital emprunté qui est assuré et est appliquée généralement pour chacune des garanties sélectionnées. Pour un emprunteur unique, l'établissement prêteur demande à ce qu'elle soit au minimum égale à 100% afin que l'intégralité du capital emprunté soit couvert.

Toutefois, un emprunt peut être souscrit par plusieurs personnes physiques qui deviennent alors co-emprunteurs. Chacune d'elle est signataire du crédit et elles sont ainsi toutes solidaires de la dette. De façon concrète, chaque co-emprunteur participe au remboursement des mensualités selon la répartition définie en amont. De ce fait, tous les co-emprunteurs peuvent prétendre à une assurance de prêt. Dans

cette situation, la quotité assurée par co-emprunteur peut être inférieure à 100%. En effet, l'organisme prêteur demande seulement que la somme des quotités de chacun des co-emprunteurs doit être au minimum égale à 100% et peut atteindre jusqu'à 200%. Par exemple, dans le cas d'un emprunt en couple, si chacun décide d'être assuré à 50%, alors dans le cas du décès de l'un d'entre eux, l'organisme assureur remboursera les 50% du capital restant dû liés au co-emprunteur décédé. Les 50% restant seront à la charge du second co-emprunteur.

Les garanties assurantielles

Outre la quotité minimum assurée, l'organisme prêteur peut imposer certaines garanties, en cas de souscription à une assurance de prêt, afin d'être en adéquation avec sa politique des risques. Les garanties contre les risques de décès, de perte totale et irréversible d'autonomie, d'invalidité et d'incapacité sont alors exigibles. Celles en lien avec le risque de perte d'emploi restent optionnelles. Ainsi, un contrat d'assurance de prêt se compose d'une ou plusieurs garanties parmi les suivantes :

- la garantie décès : elle permet, en cas de décès de l'assuré, le remboursement intégral du capital restant dû et des intérêts courus, à hauteur de la quotité assurée, à l'organisme prêteur par l'assureur ;
- la garantie PTIA (perte totale et irréversible d'autonomie) : elle défère le remboursement à la banque du capital restant dû et des intérêts courus de l'assuré, modulo sa quotité, à l'assureur si l'assuré est en perte totale et irréversible d'autonomie. Cet état de santé se définit par le besoin systématique d'une tierce personne pour réaliser les actes de la vie quotidienne (se lever, se déplacer, se nourrir, se laver) ;
- la garantie IPT (invalidité permanente totale) : elle permet le remboursement en partie, ou en totalité, des échéances de l'assuré selon sa quotité par l'assureur dans le cas où son inaptitude totale à exercer son activité professionnelle, ou toute autre activité lui assurant un revenu, est reconnue par l'assureur ;
- la garantie IPP (invalidité permanente partielle) : si l'assuré est déclaré inapte partiellement à exercer son activité professionnelle tel qu'il le faisait auparavant impliquant ainsi le besoin d'un aménagement de son poste pour la poursuivre, alors la garantie IPP concède le remboursement d'une partie ou de la totalité des échéances de l'assuré à l'assureur à hauteur de sa quotité ;
- la garantie ITT (incapacité temporaire totale) : si l'assuré est dans l'impossibilité totale d'exercer son activité professionnelle, ou autre activité qui lui assure un revenu, de façon temporaire, alors l'assureur relaie l'assuré pour le paiement de ses échéances selon sa quotité sur la durée fixée ;
- la garantie ITP (incapacité temporaire partielle) : cette garantie se définit sur le même principe que l'ITT mais elle ne s'applique uniquement qu'à un nombre restreint d'activités ;
- la garantie perte d'emploi : en cas de licenciement indemnisé par Pôle Emploi, si l'assuré dispose de cette garantie, alors son assureur couvre les échéances de son prêt, à hauteur de sa quotité, pendant la période de prise en charge.

De manière générale, les organismes prêteurs exigent au minimum la garantie contre le risque décès couplée à celle contre le risque de perte totale et irréversible d'autonomie.

Pour chacune des garanties définies précédemment, les exclusions de garanties, les périodes de carence et les délais de franchise sont précisés dans les conditions générales et particulières des contrats. En effet, les exclusions de garanties permettent à l'organisme assureur de réduire son exposition au risque en prévoyant la non-couverture de certains événements. Il en est de même avec la période de carence qui correspond à un temps, à la souscription du contrat, pendant lequel les garanties ne peuvent encore être appliquées. Enfin, le délai de franchise est un nombre de jours qui, une fois écoulé, donne droit aux versements des prestations. A la différence de la période de carence, les garanties sont acquises pendant ce délai.

1.1.1.3 Les différents types de primes

De manière générale, les primes périodiques de l'assurance de prêt sont demandées au même instant que l'échéance de remboursement du prêt. L'organisme assureur est en droit d'exclure l'assuré en cas de non paiement des primes, comme l'inscrit le CODE DES ASSURANCES (2012). Elles peuvent être exprimées en fonction du capital initial ou bien en fonction du capital restant dû, et ainsi être respectivement constantes ou décroissantes au cours de l'échéancier de remboursement.

De plus, l'évaluation du coût du risque pour la tarification d'un produit emprunteur dépend du mode de distribution de celui-ci : contrat groupe ou contrat individuel. Pour le contrat groupe, le nombre de variables discriminantes est généralement inférieur à celui utilisé pour la modélisation des tarifs dans le contrat individuel. En effet, pour ce type de contrat, le principe de mutualisation des risques est davantage considéré et permet de distribuer un produit à un tarif unique nécessitant ainsi moins d'informations sur l'assuré suite à la faible segmentation tarifaire. Les modèles de tarification dépendent donc de la sélection des risques à l'entrée de chaque contrat pour chaque assureur.

1.1.2 La sélection des risques à l'adhésion

Au moment de la souscription d'une assurance de prêt, l'assuré doit renseigner des informations personnelles afin que l'organisme assureur puisse évaluer son risque encouru, lui communiquer le tarif adéquat ou l'exclure. Elles lui permettent de réduire son phénomène d'antisélection, soit l'asymétrie d'information entre son assuré et lui-même qui lui empêche une évaluation précise du risque. Par conséquent, un processus général de sélection des risques en assurance emprunteur, pour l'ensemble des profils d'assuré, a été mis en place. Pour les profils détectés comme étant à risque aggravé de santé, le législateur a également instauré deux solutions par la convention AERAS : la grille de référence et le droit à l'oubli.

1.1.2.1 Le processus de sélection des risques

Outre les informations classiques sur l'emprunteur telles que son âge, son sexe, son statut fumeur ou non et sa profession qui permettent une première segmentation tarifaire, l'assureur est en droit de demander certaines informations médicales complémentaires pour les risques santé et prévoyance. Pour l'ensemble des formalités médicales, l'assuré se doit d'être honnête dans ses réponses. La couverture des risques sera nulle si une fausse déclaration de l'assuré a été détectée par l'assureur. Ces formalités se distinguent en trois niveaux et sont réalisées selon le profil de l'emprunteur.

Le premier niveau est le questionnaire médical. Dans ce questionnaire, l'assureur demande approximativement l'ensemble des informations suivantes à l'assuré :

- son âge ;
- sa taille ;
- son poids ;
- son statut fumeur ou non ;
- son niveau de consommation d'alcool ;
- ses traitements médicaux suivis ;
- s'il souffre d'un handicap ou d'une maladie ;
- s'il a été hospitalisé ou s'il le sera prochainement ;
- s'il a eu des arrêts de travail.

Ce questionnaire médical simplifié est directement adressé aux assurés qui semblent être les moins à

risque, soit ceux dont leur âge est inférieur à 50 ans avec un capital emprunté inférieur à 400 000 euros.

Deuxièmement, l'assureur peut réclamer une étude médicale du dossier. Cette étude consiste à demander des pièces complémentaires au questionnaire médical pour détailler davantage l'état de santé de l'assuré. Ces pièces complémentaires sont, par exemple, des documents médicaux comme des comptes-rendus hospitaliers ou un rapport rédigé par le médecin traitant de l'assuré. Ce second niveau est généralement demandé pour les emprunteurs dont l'âge est supérieur à 45 ans et dont le montant emprunté est supérieur à 250 000 euros.

Enfin, le dernier niveau correspond à l'examen médical. L'assuré doit alors réaliser des examens médicaux définis par l'assureur tels qu'un bilan sanguin, un hémogramme, une analyse d'urine ou bien un électrocardiogramme. Ils sont en général demandés si l'assuré a déclaré des antécédents médicaux dans son questionnaire médical.

Pour conclure, l'assureur accepte l'assuré avec un tarif classique, une surprime justifiée, ou met en application la convention AERAS (S'Assurer et Emprunter avec un Risque Aggravé de Santé).

1.1.2.2 La convention AERAS

La convention AERAS permet à toute personne souscrivant une assurance emprunteur, et dont le profil ne permet pas de s'assurer aux conditions standards du contrat, d'être couvert. Elle peut être mise en oeuvre si la part assurée sur l'encours cumulé de prêts est inférieure à 320 000 euros et si l'échéance de l'assurance intervient avant le 71^{ème} anniversaire de l'assuré. Deux dispositifs sont mis en place dans cette convention : la grille de référence AERAS et le droit à l'oubli.

La grille de référence AERAS

La grille de référence AERAS est une liste de pathologies pour lesquelles un accès à l'assurance emprunteur est accordée dans des conditions presque identiques à un profil d'assuré sans risque de santé après un temps défini pour chaque pathologie. Cette grille est constituée de deux parties :

- Partie I : elle correspond à une liste de pathologies à déclarer à l'assureur pendant le processus de sélection médical sous peine de sanction et qui permet à l'assuré d'obtenir une proposition d'assurance sans surprime ni exclusion de garanties si la pathologie respecte un type d'histologie, un stade et un temps sans rechute depuis la fin du traitement thérapeutique ;
- Partie II : il s'agit également d'une liste de pathologies, à déclarer à l'assureur, avec conditions sur le type d'histologie, le grade et un temps donné sans rechute à respecter. Toutefois, pour ces maladies, l'organisme assureur peut émettre une proposition avec une surprime plafonnée se rapprochant ainsi des conditions standards.

Enfin, cette grille évolue au rythme des innovations thérapeutiques et des études épidémiologiques afin de garantir un accès juste à l'emprunt pour la population française.

Le droit à l'oubli

Le droit à l'oubli est un droit qui complète la grille de référence AERAS, sans la condition sur le montant de l'emprunt. Par opposition à cette-dernière où l'assuré doit déclarer sa pathologie, il est ici question de ne pas déclarer certains antécédents médicaux à son assureur passé un certain délai sans rechute et sous une certaine condition d'âge à la fin du prêt. Plus précisément, ce droit se définit comme suit dans la CONVENTION AERAS (2020) :

« Aucune information médicale relative à une pathologie cancéreuse ne pourra être sollicitée par l'assureur dès lors que le protocole thérapeutique relatif à cette pathologie est achevé depuis plus de 10 ans. Pour les cancers diagnostiqués avant l'âge de 21 ans, ce délai est ramené à 5 ans. [...] Les autres pathologies et les facteurs de risque, les situations actuelles d'incapacité, d'invalidité ou d'inaptitude au travail, en lien ou non avec l'affection relevant du « droit à l'oubli », sont à déclarer à l'assureur en réponse au questionnaire de santé et pourront faire l'objet d'une décision adaptée ou d'une tarification en tant que telle. »

Les délais et types de pathologies susmentionnés correspondent à l'application du droit dans sa version antérieure à celle adoptée par la loi Lemoine.

Dans cette définition, la notion de protocole thérapeutique achevé est importante à maîtriser pour une application irréprochable de ce droit par les assurés concernés. D'une part, la CONVENTION AERAS (2020) précise que la fin du protocole thérapeutique est *« la date de la fin du traitement actif du cancer, en l'absence de rechute, par chirurgie, radiothérapie chimiothérapie effectuées en structure autorisée, à laquelle plus aucun traitement n'est nécessaire hormis la possibilité d'une thérapeutique persistante de type hormonothérapie ou immunothérapie »*. D'autre part, par absence de rechute, elle entend qu'aucune *« manifestation médicalement constatée du cancer, qu'elle le soit par le biais d'un examen clinique, biologique ou d'imagerie »* (CONVENTION AERAS, 2020) n'a été observée.

Par conséquent, ce droit d'omission accentue l'asymétrie d'information entre assuré et assureur. Des questions autour de l'état de santé à long terme d'un assuré, anciennement touché par une pathologie cancéreuse, en émanent. Malgré celles-ci, le délai du droit à l'oubli sera réduit avec la loi Lemoine pour davantage d'inclusion.

1.1.3 Un marché dominé par la couverture des crédits immobiliers

Il existe trois catégories de prêts : les prêts immobiliers, les prêts à la consommation et les prêts professionnels. Pour chacun de ces prêts, l'organisme créateur souhaite une couverture en cas de non-remboursement du capital. L'assurance emprunteur est une des protections les plus couramment utilisées.

Le marché français de l'assurance emprunteur est dominé, en terme de montants de cotisations, par l'assurance des crédits immobiliers avec près de 8 milliards d'euros en 2021 comme l'indique FRANCE ASSUREURS (2022) dans son rapport annuel sur les données d'assurance françaises 2021. Ce montant de cotisations représente près de 67% de l'ensemble des cotisations du marché de l'assurance emprunteur. Afin de libéraliser ce marché, le législateur se préoccupe principalement de son cadre de lois avec la mise en place en février 2022 de la loi Lemoine. Par conséquent, ce mémoire se concentre sur ce produit d'assurance emprunteur.

1.2 La réglementation du marché

L'assurance des crédits immobiliers a un cadre législatif stricte.

1.2.1 Historique des réglementations jusqu'à 2020

Depuis le début des années 2000, l'assurance emprunteur connaît de grandes avancées législatives afin de protéger le pouvoir d'achat de tous les Français et de rendre le marché plus juste envers les profils d'emprunteurs présentant des risques de santé.

1.2.1.1 Des mesures pour protéger le pouvoir d'achat des Français

Le gouvernement, souhaitant protéger le pouvoir d'achat des consommateurs, a pris des mesures importantes pour encadrer les pratiques financières impactant ainsi le domaine de l'assurance emprunteur.

Publiée au JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE (2001) le 12 décembre 2001, la loi MURCEF, acronyme de Mesures Urgentes de Réforme à Caractère Économique et Financier, établit le principe de la délégation d'assurance de prêt. Cette loi requiert des banques davantage de transparence vis-à-vis des consommateurs. Pour l'assurance des crédits immobiliers, elle interdit aux organismes bancaires de lier un produit d'emprunt immobilier à une assurance pour ce dernier. Ainsi, cette législation a permis d'ouvrir à la concurrence le marché de l'assurance emprunteur.

L'Etat a ensuite voté en 2008 la loi Châtel, inscrite dans le CODE DE LA CONSOMMATION (2008), dont l'un des objectifs a été de faciliter la résiliation de l'assurance de prêt immobilier avec une obligation d'information quant à la tacite reconduction du contrat. En effet, elle exige des organismes assureurs de communiquer un à trois mois en amont de la date de résiliation sur la possibilité de reconduire ou non le contrat d'assurance avec un avis d'échéance. De plus, s'ils ne respectent pas cette obligation dans les délais, les assurés peuvent alors résilier leur contrat sans attendre son échéance et sans frais.

En complément de la loi Châtel, les pouvoirs publics continuent de délier le prêt immobilier de son assurance, développant ainsi le marché de la délégation d'assurance, avec la création en 2010 de la loi Lagarde dans le CODE DE LA CONSOMMATION (2010). Elle comporte trois versants bousculant les pratiques des banques :

- un emprunteur peut désormais refuser l'assurance proposée par la banque auprès de laquelle il réalise son crédit immobilier, s'il choisit d'être couvert par un autre acteur ;
- un organisme prêteur n'est plus en droit de refuser la proposition de substitution d'assurance si les garanties présentées sont au moins équivalentes à celles qu'il a proposées initialement : il s'agit du principe d'équivalence des garanties ;
- une banque n'est pas autorisée à modifier les conditions de taux du prêt, ni ajouter des frais de délégation dans le cas où l'emprunteur propose un contrat d'assurance de substitution.

En 2014, le législateur décide de renforcer les droits des emprunteurs avec la loi Hamon (CODE DES ASSURANCES, 2014). Cette loi permet aux emprunteurs de résilier leur contrat d'assurance à tout moment dans les douze mois suivants la signature du prêt et sans frais. Cette résiliation est conditionnée à la présentation d'un nouveau contrat respectant le principe d'équivalence des garanties défini dans la loi Lagarde.

Enfin, l'amendement Bourquin (CODE DES ASSURANCES, 2017) de la loi Sapin 2 vient davantage simplifier la résiliation d'assurance avec son application en 2018. En effet, il donne droit aux assurés de résilier annuellement à chaque date d'anniversaire du contrat, sans besoin de réception d'avis d'échéance, et avec un préavis de deux mois.

1.2.1.2 Une assurance des emprunteurs inclusive

En parallèle de l'ouverture du marché à la concurrence pour diminuer les tarifs, le législateur entend que chacun puisse avoir accès à l'assurance emprunteur afin d'obtenir le prêt immobilier auquel il prétend. Comme expliqué précédemment, les banques conditionnent l'octroi d'un prêt à la souscription d'une assurance emprunteur. Toutefois, les organismes assureurs peuvent refuser de couvrir des personnes présentant déjà, au moment de la souscription, des risques de santé. Ainsi, des mesures ont été établies afin de permettre au plus grand nombre d'accéder à l'achat immobilier.

En 1991, la convention assurances-sida est signée par l'Etat et les professionnels de l'assurance. Elle permet aux personnes séropositives de protéger leurs données médicales transmises et d'accéder à l'assurance-décès pour obtenir un crédit immobilier. Toutefois, cette convention a alors un périmètre restreint d'application. Ainsi, une deuxième convention est mise en place en 2001, nommée convention Belorgey (CONVENTION BELORGEY, 2001), conçue par les pouvoirs publics, les professionnels, les associations de consommateurs et les associations représentant les personnes malades et/ou handicapées. Elle étend les principes établis dans la convention assurances-sida pour la garantie décès à d'autres pathologies, ajoute le principe d'analyse à trois niveaux, et instaure un code de bonne conduite.

Après cinq années d'application de la convention Belorgey, les parties signataires ont souhaité la faire évoluer. En ce sens, ils créent la convention AERAS en vigueur depuis 2007. Désormais, les personnes à risques de santé peuvent se couvrir contre l'invalidité avec l'ensemble des tarifs majorés et des démarches facilitées. La convention met également en place deux commissions de suivi : la commission de médiation et la commission des études et de recherches. La commission de médiation a pour objectif de faciliter le dialogue entre un assuré ayant des risques de santé aggravés, son médecin et le médecin conseil de l'assureur. Quant à la commission des études et de recherches, elle est en charge du suivi des analyses sur les dernières données de santé qui déterminent les surprimes ou les refus de garanties des assureurs.

La convention est ensuite renouvelée en 2011 avec la création de la garantie spécifique d'assurance invalidité (GIS) qui peut être substituée par la garantie perte totale et irréversible d'autonomie (PTIA) si l'organisme assureur n'est pas en mesure de l'offrir. De plus, le devoir d'information envers les emprunteurs est enrichi et les démarches sont à nouveau simplifiées.

Enfin, en 2015, la convention AERAS est révisée afin d'y ajouter la grille de référence AERAS et le principe du droit à l'oubli (CONVENTION AERAS RÉVISÉE, 2015). Ces deux notions sont renforcées par les avenants de 2019 (CONVENTION AERAS, 2019) et de 2020 (CONVENTION AERAS, 2020) qui agrandissent notamment la liste des pathologies reconnues.

1.2.2 Un bilan mitigé des réglementations jusqu'à aujourd'hui

Malgré les évolutions réglementaires en faveur de la libéralisation du marché de l'assurance emprunteur, la part des contrats détenus par les acteurs de la délégation d'assurance n'a pas eu une croissance forte. Des avancées en faveur des assurés ont tout de même été observées pour leur pouvoir d'achat et pour un accès plus favorable à tous à l'emprunt.

1.2.2.1 Les Français, bénéficiaires de la libéralisation du marché de l'assurance emprunteur

D'une part, avec la mise en place de la loi Hamon en 2014 et celle de l'amendement Bourquin en 2018, une évolution à la hausse des souscriptions chez les délégataires d'assurance était attendue. Toutefois, la figure 1.1 montre une part quasi constante de la délégation d'assurance en terme de montant de cotisations entre 2016 et 2020, restant autour de 12% (FRANCE ASSUREURS, 2021b).

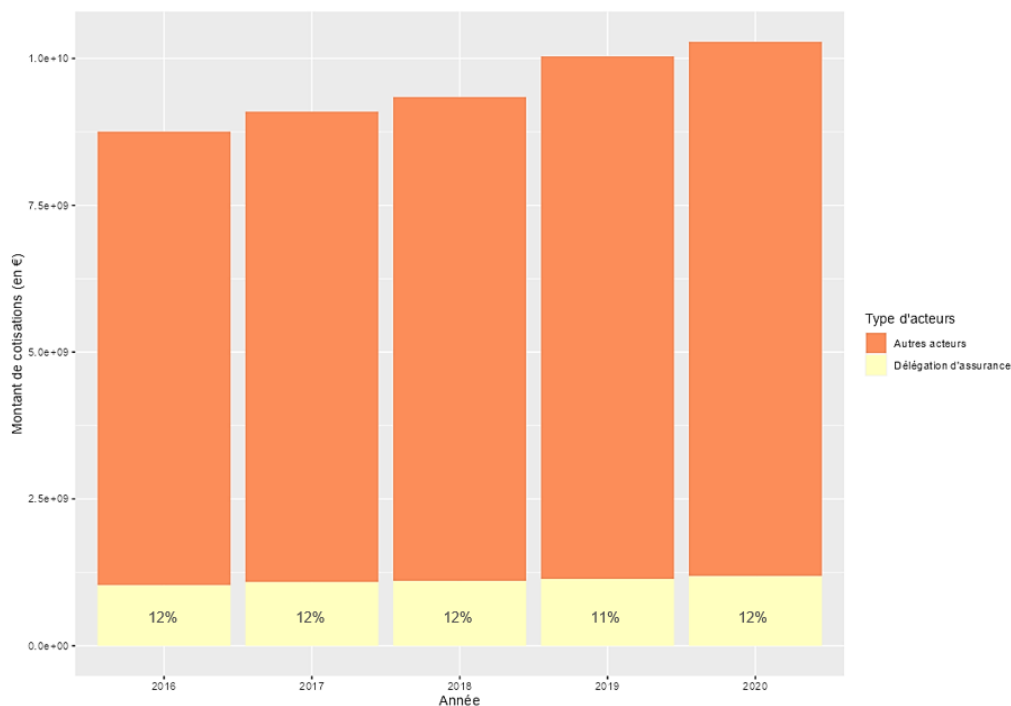


FIGURE 1.1 – Montant des cotisations en € par type d'acteur (FRANCE ASSUREURS, 2021b)

D'autre part, en novembre 2020, le COMITÉ CONSULTATIF DU SECTEUR FINANCIER (2020) (CCSF) publie, à la demande du ministre de l'Economie et des Finances, un bilan du marché de l'assurance emprunteur. Il fait état d'une baisse généralisée des tarifs depuis 2010 et d'une augmentation de la couverture des assurés. Plus précisément, il estime la baisse des tarifs pour les contrats individuels du marché de la délégation entre 20% et 41% pour l'ensemble des profils d'assuré. Quant aux contrats groupes bancaire, il constate une diminution des tarifs plus modérée entre 13% et 26% pour les profils représentant un nombre important de contrats et pour les classes d'âge les plus jeunes.

Par conséquent, les mesures prises en faveur de l'ouverture du marché de l'assurance emprunteur ont été davantage profitables au pouvoir d'achat des Français qu'à la part de marché détenue par les acteurs de la délégation d'assurance. Simultanément, les Français concernés par un risque aggravé de santé profitent de l'application de la convention AERAS.

1.2.2.2 Une assurance emprunteur pour tous : application de la convention AERAS et du droit à l'oubli

Depuis son entrée en vigueur en 2007, la convention AERAS a évolué en faveur des consommateurs pour un accès plus juste à la demande de prêt. En effet, FRANCE ASSUREURS (2021a) précise, dans son bilan d'application de la convention AERAS, que chaque année depuis 2016 près de 96% des demandes d'assurances pour des risques aggravés de santé ont reçu une proposition de couverture par les organismes assureurs. De plus, pour la garantie décès et la garantie PTIA les assureurs ont su en 2020, pour respectivement 68% et 88% des demandes de couverture, proposer un tarif sans surprime et sans exclusion de garanties. Cette convention permet donc aujourd'hui de répondre aux besoins des consommateurs au profil de risque élevé en conservant leur pouvoir d'achat.

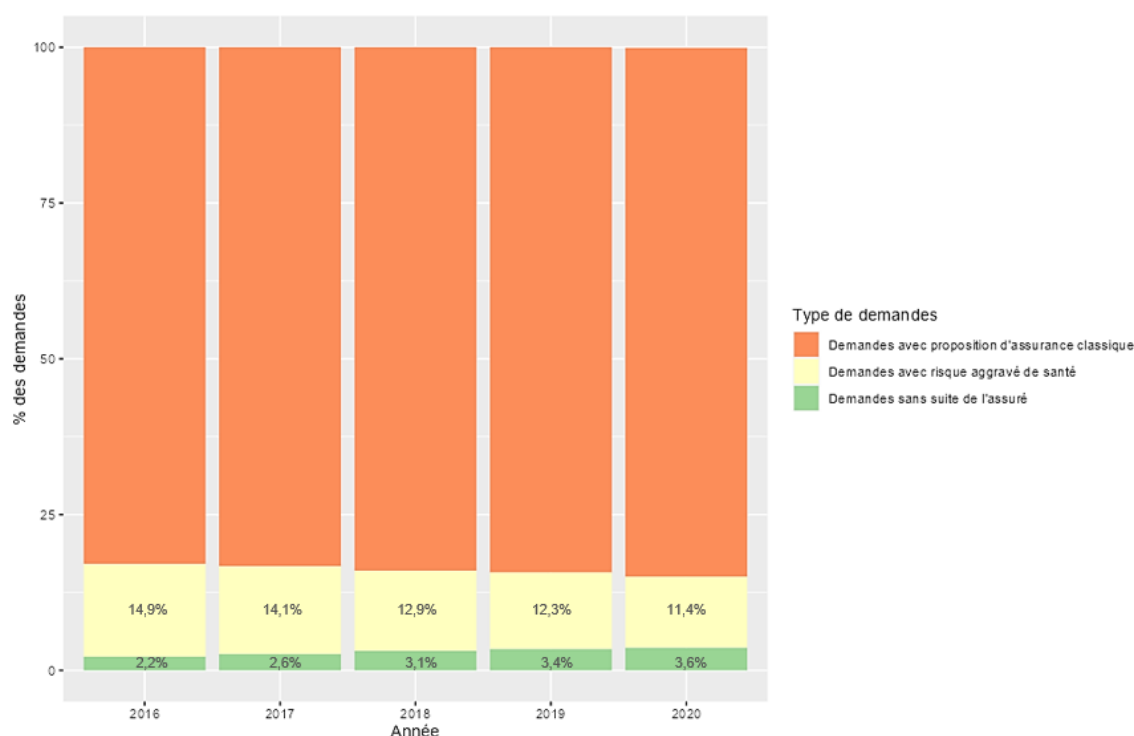


FIGURE 1.2 – Répartition des demandes d'assurance de prêts par année (FRANCE ASSUREURS, 2021a)

En outre, dans la figure 1.2, il ressort une diminution continue de 2016 à 2020 de la part des demandes de prêts présentant un risque aggravé de santé. Cette diminution peut être signe de l'application du droit à l'oubli depuis sa création en 2015 avec des profils en risque aggravé qui ne sont ainsi plus détectés. Cependant, il ne s'agit que de supposition puisqu'il est difficile de mesurer l'application de ce cadre qui consiste en des pathologies non déclarées à l'organisme assureur.

En dépit de ces données favorables à une assurance emprunteur plus juste et inclusive, la convention AERAS reste restreinte à une liste non exhaustive de pathologies. Les associations pour les personnes malades continuent ainsi de batailler pour étendre cette convention. Parallèlement, les assureurs ont continué de militer pour une plus forte libéralisation du marché avec notamment l'instauration de la résiliation infra-annuelle.

1.2.3 La loi Lemoine

Depuis 2020, le marché de l'assurance des prêts immobiliers connaît de grands rebonds politiques quant à la résiliation infra-annuelle et la sélection médicale. Ainsi, le 28 février 2022, la loi Lemoine 2022-270 a été adoptée « *pour un accès plus juste, plus simple et plus transparent au marché de l'assurance emprunteur* » (JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE, 2022). En application depuis le 1^{er} juin 2022 pour les nouveaux contrats et à compter du 1^{er} septembre 2022 pour les contrats en stock, cette loi comprend quatre avancées majeures.

La résiliation à tout moment

La loi Lemoine met en place la résiliation à tout moment sans préavis et sans frais. Ce droit est valable pour tous les assurés détenant des contrats d'assurance de prêt immobilier souscrits pour un bien à usage d'habitation ou mixte (habitation et professionnelle). Cette résiliation est effective si le nouveau contrat d'assurance respecte le principe d'équivalence des garanties. L'assureur possède un délai de 10 jours pour effectuer un avenant.

La suppression du questionnaire médical

Toute personne physique souhaitant couvrir un emprunt pour un bien à usage d'habitation ou mixte, dont la part assurée sur encours cumulé des crédits est inférieure ou égale à 200 000 euros, et dont la dernière échéance de remboursement survient avant son 60^{ème} anniversaire n'est désormais plus soumise au questionnaire médical.

Le devoir d'information

Le devoir d'information des organismes assureurs envers les assurés est renforcé avec cette nouvelle législation. Les assureurs sont ainsi dans l'obligation d'informer leurs assurés quant à la mise en oeuvre du droit de résiliation. Ils doivent également communiquer sur la date d'échéance du contrat et sur le coût de l'assurance sur 8 ans. Le législateur a prévu une amende de 15 000 euros pour tout organisme d'assurance qui ne respecterait pas ces obligations.

La réduction du délai de droit à l'oubli

Le droit à l'oubli est modifié en faveur des personnes à risques aggravés de santé. En amont de cette loi, un assuré qui avait été atteint d'un cancer avant ses 21 ans n'avait pas à le déclarer si au moins 5 ans s'étaient écoulés entre la fin du traitement thérapeutique et la demande de prêt sans rechute. Ce temps était augmenté à 10 ans pour les emprunteurs dont le cancer avait été diagnostiqué après leur 21 ans. Avec la loi Lemoine, il est aujourd'hui généralisé à 5 ans peu importe l'âge au diagnostic du cancer et également étendu aux personnes victimes d'hépatite virale C.

Avec la résiliation à tout moment et le devoir d'information, le législateur continue de libéraliser le marché. En parallèle, avec la suppression du questionnaire médical et la réduction du délai du droit à l'oubli, il permet à des profils d'assurés ayant un risque aggravé de santé de s'assurer pour emprunter : des nouveaux risques entrent alors dans les portefeuilles des assureurs.

1.3 Situation globale d'épidémiologie des cancers en France

Maintenant que l'assurance emprunteur, ses évolutions législatives et notamment le droit à l'oubli ont été présentés, la situation globale d'épidémiologie des cancers en France est décrite dans cette partie.

En France, les cancers figurent comme la première cause de décès chez les hommes et la deuxième chez

les femmes d'après SANTÉ PUBLIQUE FRANCE (2021). Les pouvoirs publics tiennent à suivre avec attention la situation et mettent en place des plans de lutte contre le cancer, le droit à l'oubli en est un exemple. En ce sens, cette revue a pour but de comprendre les enjeux amenés par cette évolution législative en faveur des personnes touchées par un cancer en revenant sur la définition de cette pathologie, puis en communiquant les statistiques françaises produites par les organes scientifiques publics.

1.3.1 Définition

Le terme « cancer » regroupe en réalité un grand nombre de maladies distinctes les unes des autres. Toutefois, elles réfèrent toutes à une multiplication anarchique de cellules qui forment une tumeur maligne. Chacune de ces maladies se distingue par l'origine des cellules cancéreuses (cancer du foie, cancer du poumon, etc.).

Cette multiplication anarchique et excessive commence localement dans l'organisme et s'étend dans les tissus proches. Si cette prolifération n'est pas arrêtée à temps, elle peut également se poursuivre dans le reste de l'organisme, à distance, et former ainsi des métastases.

Il existe deux grandes classes de cancers :

- les tumeurs solides qui regroupent les tumeurs développées dans n'importe quel tissu et représentent 90% des cas diagnostiqués (FONDATION CONTRE LE CANCER, s. d.) ;
- les tumeurs sanguines qui se développent dans le sang, la moelle osseuse ou le système lymphatique.

Aujourd'hui, ces pathologies sont connues de tous en France. Elles sont détectées soit par diagnostic précoce, soit par dépistage organisé par les pouvoirs publics. Plus le diagnostic est réalisé tôt, plus la mortalité est réduite.

A partir de cette définition et de ce contexte, les principaux indicateurs de la surveillance des cancers en France peuvent être étudiés.

1.3.2 Le suivi des cancers en France

Depuis 2003, des plans de lutte cancer sont mis en place pour des durées de 4, 5 ou 10 ans afin d'établir une stratégie précise autour de la prévention, du dépistage, de l'organisation des soins, de la recherche, de l'accompagnement du patient et de son entourage et de l'après-cancer. Ces plans donnent trois indicateurs permettant le suivi de la situation épidémiologique : le taux d'incidence, le taux de prévalence et le taux de survie.

1.3.2.1 L'incidence des cancers en France

L'incidence d'une pathologie correspond au nombre de nouveaux cas diagnostiqués de cette pathologie sur une population pour un temps fixé. Il est possible d'en déduire le taux d'incidence. Précisément, ce dernier correspond au nombre de nouveaux cas d'une pathologie observés pendant une période déterminée rapporté à la population dont ils sont issus pour cette même période. Afin de pouvoir comparer les taux d'incidence dans le temps et entre les populations, le taux d'incidence standardisé monde a été conçu. Il est calculé en supposant que la structure d'âge de la population observée est identique à celle de la population mondiale. Il est exprimé en « nombre de personnes pour 100 000 personnes par année ».

Dans leurs estimations nationales pour les tumeurs solides, LE GUYADER-PEYROU et al. (2019) indiquent qu'entre 1990 et 2018 le nombre de nouveaux cas de cancers diagnostiqués pour les hommes a augmenté de 65%, passant de 124 000 à 204 600 cas. Quant aux femmes, cette hausse s'élève à 93%, avec 91 800 cas pour 1990 et 177 400 pour 2018. Cette hausse s'explique en majeure partie par la croissance de la population et son vieillissement. En effet, l'INSEE (2022) indique que la population féminine passe approximativement de 29 755 000 en 1990 à 34 594 000 en 2018, et la population masculine de 28 241 000 à 32 398 000. Pour les parts allouables à l'augmentation du risque cancer, elle est estimée à 6% pour les hommes et 45% pour les femmes. De part ces constats, les taux d'incidence standardisés monde sont établis pour la France métropolitaine. En 2018, ils sont de 330.2 pour les hommes et de 274.0 pour les femmes.

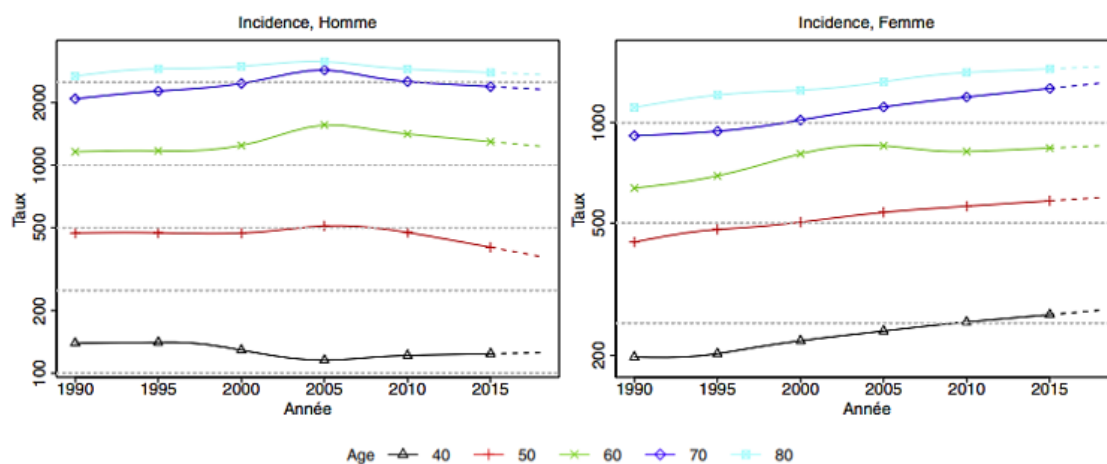


FIGURE 1.3 – Taux d'incidence des cancers en France selon l'année et par âge à l'échelle logarithmique, tous les cancers confondus (LE GUYADER-PEYROU et al., 2019)

Avec davantage d'historique, la figure 1.3 de gauche, représentant le taux d'incidence pour les hommes, montre une tendance à la baisse de ce taux au cours des années pour tous les âges confondus. Cette tendance est contraire à celle des femmes, présentée à droite de cette même figure, qui indique une tendance à la hausse.

1.3.2.2 La prévalence des cancers en France

De manière générale, la prévalence dénombre les personnes en vie qui ont été touchées par un cancer au cours de leur vie dans une population sur une période de temps fixée. Elle se distingue de l'incidence puisqu'elle recense l'ensemble des cas et non uniquement les nouvellement diagnostiqués. Le taux de prévalence en est déduit en rapportant la prévalence à la population totale de la période de temps fixée.

INSTITUT NATIONAL DU CANCER (2022) indique qu'en 2017 3.8 millions de personnes ont été touchées par un cancer au cours de leur vie en France métropolitaine. Il précise également que cette prévalence est en hausse constante grâce à l'amélioration de la survie et la hausse du nombre de cas diagnostiqués.

1.3.2.3 La mortalité des cancers en France

Quant à la mortalité, il s'agit du nombre de décès survenus dans une population pour une période de temps fixé. Ainsi, rapportée à la population totale dans cette même période de temps observée, le taux de mortalité est calculé. Comme pour l'incidence, il existe le taux de mortalité standardisé monde basé sur le même principe de construction et exprimé en « nombre de personnes pour 100 000 personnes par année ».

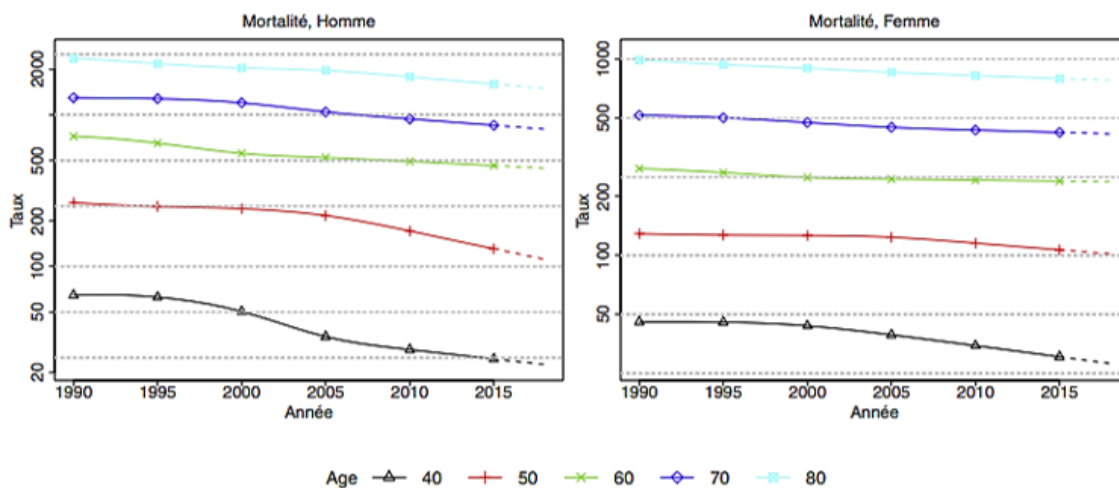


FIGURE 1.4 – Taux de mortalité des cancers en France selon l'année et par âge à l'échelle logarithmique, tous les cancers confondus (LE GUYADER-PEYROU et al., 2019)

Comme le montrent LE GUYADER-PEYROU et al. (2019) à l'aide de la figure 1.4 les taux de mortalité, tous sexes et âges confondus, tendent à diminuer au cours du temps pour la période 1990-2015. C'est également ce qui a été constaté par l'INSTITUT NATIONAL DU CANCER (2022) pour la période 2010-2018, avec une baisse générale calculée à -2% pour les hommes et -0.7% pour les femmes.

Au regard des taux d'incidence qui ont une légère tendance à la hausse et des taux de mortalité qui diminuent, la prévalence augmente comme il a été souligné en amont. En ce sens, il est supposé que le nombre de personnes diagnostiquées d'un cancer et pouvant ainsi appliquer leur droit à l'oubli augmente. Par conséquent, il devient primordial d'étudier le risque de décès des suites du cancer diagnostiqué et traité afin d'anticiper l'impact de la réduction du délai d'application du droit à l'oubli.

Chapitre 2

Présentation, analyse et valorisation des données SEER

En 1971, le Congrès américain a acté la collecte, l'analyse et le partage des données individuelles relatives aux cancers (THE NATIONAL CANCER ACT OF 1971, 1971) afin d'en réduire la mortalité en facilitant la recherche. Le programme SEER (Surveillance, Epidemiology, and End Results) est ainsi mis en place par le National Cancer Institute (NCI) en 1973. Cette base de données est mise à jour annuellement et recense, pour chaque personne diagnostiquée avec une tumeur, ses informations personnelles, médicales et biométriques, un suivi de son statut vital, ainsi que la source de cette observation. Ces données sont accessibles à toute personne qui en fait la demande au NCI. Dans sa dernière version de novembre 2021, le programme SEER comporte un maximum de 267 variables et couvre les cancers enregistrés parmi 48% de l'ensemble de la population des États-Unis. Il s'agit ainsi de la base de données de suivi des pathologies cancéreuses la plus complète qui est mise à disposition de tous.

En France, ce type de données médicales individuelles n'est pas accessible au grand public. Bien que la problématique étudiée dans ce mémoire concerne uniquement la population française ayant été diagnostiquée d'un cancer et étant en rémission, la probabilité de décès des suites du cancer diagnostiqué et traité en rémission est établie sur les données SEER extraites de la population des États-Unis. En effet, en considérant que ces populations sont similaires de part le niveau de développement élevé de ces pays, il est supposé que la situation épidémiologique des cancers aux États-Unis est identique à celle de la France. Ainsi, les taux d'extra-mortalité calculés sur la population des États-Unis à l'aide de la base SEER sont appliqués sans ajustement pour la mesure de la réduction du droit à l'oubli en France. Cette hypothèse est challengée à la fin des travaux.

Les données SEER ne sont pas directement exploitables pour étudier la mortalité due au cancer après traitement thérapeutique. C'est pourquoi il convient, dans cette seconde partie, de les introduire, de les présenter, de donner les traitements effectués et d'analyser la base finale. Enfin, la structure des tables d'extra-mortalité construites est justifiée.

2.1 Introduction aux données SEER

Les données SEER ne sont pas directement accessibles sur le site du programme <https://seer.cancer.gov/>. En effet, ce site internet permet uniquement de réaliser la demande d'accès au logiciel SEER,

où les données sont hébergées, et présente également des statistiques globales sur les cancers aux Etats-Unis.

Le programme SEER est organisé en deux sous-programmes : SEER Research et SEER Research Plus. SEER Research exclut certaines variables par souci de confidentialité. Ces dernières sont accessibles dans la version Research Plus qui nécessite une demande spécifique par l'utilisateur au NCI. Ainsi, dans le cadre de ces travaux, une demande spécifique pour recherche universitaire a été réalisée afin d'étudier les données SEER Research Plus.

Par la suite, le NCI envoie le lien de téléchargement du logiciel ainsi que les informations relatives au compte utilisateur attribué.

Une fois le logiciel SEER installé et le compte utilisateur ouvert, la fonctionnalité « Case Listing », affichée en figure 2.1, permet de concevoir une base de données de suivi individuel de chaque pathologie cancéreuse recensée.

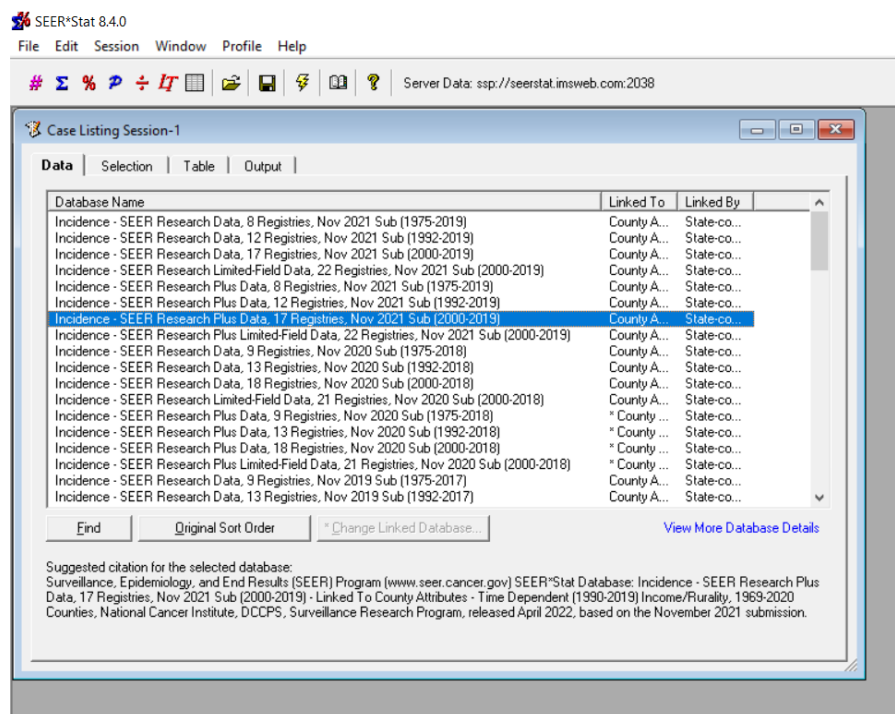


FIGURE 2.1 – Étape préliminaire pour extraire les données SEER

Une première sélection des données est à paramétrer dans chacun des onglets afin de réaliser l'extraction des lignes. Cette étape initiale permet de réduire la taille de l'extraction de la base de données à étudier en pré-sélectionnant seulement les lignes et variables intéressantes pour ces travaux.

2.2 Pré-sélection des données du programme SEER

Une étape essentielle pour accéder aux données SEER est la définition des paramètres de l'extraction qui sont à renseigner dans la fonctionnalité « Case Listing ». Ces paramètres sont : le regroupement des registres, les variables et les conditions sur les lignes à sélectionner.

2.2.1 Choix du regroupement des registres

En figure 2.1, l'onglet « Data » présente une liste des bases de données à disposition pour l'utilisateur. Les données sont initialement collectées à l'échelle des états qui sont sélectionnés selon leur capacité à maintenir une haute qualité du reporting des données. Ces différentes sources de reporting se nomment « registres ». En 1973, seules les pathologies cancéreuses du Connecticut, Iowa, Nouveau-Mexique, Utah, Hawaï et des régions métropolitaines de Detroit et San Francisco-Oakland étaient recueillies. La dernière mise à jour de la base SEER, en novembre 2021, comporte vingt-deux registres.

Afin de maintenir une cohérence dans le regroupement des données, suite à l'ajout de nouvelles sources de reporting, elles sont divisées en quatre groupes de registres :

- SEER 8 : regroupement de huit registres des cas diagnostiqués depuis 1975 jusqu'à aujourd'hui ;
- SEER 12 : regroupement de douze registres des cas diagnostiqués depuis 1992 jusqu'à aujourd'hui ;
- SEER 17 : regroupement de dix-sept registres des cas diagnostiqués depuis 2000 jusqu'à aujourd'hui ;
- SEER 22 : regroupement de vingt-deux registres des cas diagnostiqués depuis 2000 jusqu'à aujourd'hui mais limité à un nombre restreint de variables disponibles.

Ces regroupements sont identifiables dans la figure 2.1.

Dans la liste des bases de données à disposition en figure 2.1, le sous-programme (Research ou Research Plus) est également précisé. En effet, celui-ci implique un accès différent aux variables. Comme susmentionné, un accès à l'ensemble des variables via le sous-programme Research Plus a été accordé ici.

Par conséquent, cette étude repose sur SEER 17 Research Plus en vue de maximiser le nombre de variables disponibles, le nombre de cas diagnostiqués et d'éviter un recul trop important dans le temps. En ce sens, la base de données sélectionnée couvre approximativement 26.5% de la population des Etats-Unis (d'après la population recensée en 2010). Plus précisément, les registres couverts sont :

- San Francisco - Oakland SMSA ;
- Connecticut ;
- Hawaï ;
- Iowa ;
- Nouveau Mexique ;
- Seattle (détroit de Puget) ;
- Utah ;
- région métropolitaine d'Atlanta ;
- San Jose - Monterey ;
- Los Angeles ;
- Alaska ;
- Géorgie rurale ;
- grande Géorgie ;
- Californie, sauf San Francisco, San Jose - Monterey, et Los Angeles ;
- Kentucky ;
- Louisiane ;
- New Jersey.

Finalement, le nombre de tumeurs malignes recensées sur la période 2000-2019 est de 7 788 941.

2.2.2 Choix des lignes et des variables

L'onglet « Selection » permet de sélectionner les observations utiles à ces travaux. L'image 2.2 présente les conditions appliquées pour la construction de la base. Ainsi, seules les observations où l'âge est connu (« Known Age ») et qui sont liées à une tumeur maligne¹ (« Malignant Behavior ») ont été extraites pour entrer dans le cadre du droit à l'oubli. En effet, le droit à l'oubli concernant uniquement les pathologies cancéreuses, les tumeurs bénignes sont exclues. Enfin, aucune condition supplémentaire n'a été ajoutée dans cette pré-sélection.

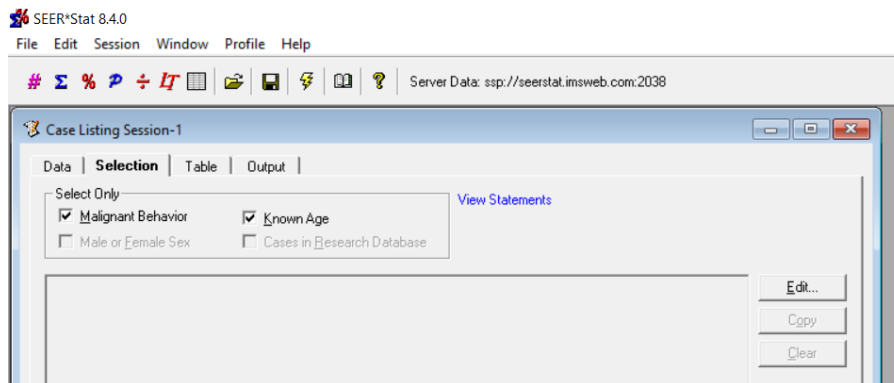


FIGURE 2.2 – Étape préliminaire bis pour extraire les données SEER

Puis, au sein du regroupement SEER 17 Research Plus, quarante huit variables ont été sélectionnées dans l'onglet « Table ». Elles traitent, de manière générale, des informations personnelles sur le patient ainsi que sa survie, sur le type de la tumeur, et sur le protocole thérapeutique mis en place. Le tableau complet des ces variables sélectionnées est disponible en annexe (4.12). Les variables exclues entraient davantage dans le détail scientifique des pathologies avec les spécificités du site touché ou encore l'extension de la maladie à l'aide de son histologie.

Par conséquent, la base extraite comporte 7 788 415 lignes et quarante-huit variables. Malgré cette étape de pré-sélection, les données ne sont pas exactement adaptées pour étudier l'extra-mortalité due au cancer après la fin du protocole thérapeutique. Ainsi, les variables de cette base doivent être complétées pour obtenir les données nécessaires à l'étude de la réduction du délai d'application du droit à l'oubli.

2.3 Complétion des données de la base SEER

L'objectif de ces travaux est de comprendre le risque encouru par les assureurs, pour une garantie décès, suite à la réduction de moitié du délai d'application du droit à l'oubli. Ce délai commence à partir de l'âge de fin de traitement thérapeutique suivi. Il convient donc d'étudier, pour chaque année passée depuis la fin du traitement, la probabilité de décéder à cause du cancer diagnostiqué et traité. Ainsi, il est primordial de disposer de l'âge à la fin du protocole thérapeutique qui correspond à la date de début de la table d'extra-mortalité à construire.

Toutefois, parmi l'ensemble des variables mises à disposition dans la base SEER Research Plus, aucune n'indique l'âge de fin de traitement, ni la durée du protocole prescrit. L'enjeu premier des retraitements

1. « Amas de cellules cancéreuses »(CANCER, s. d.[a])

de la base extraite est alors de construire l'âge à la fin du traitement thérapeutique.

Concrètement, la construction de l'âge à la fin de traitement thérapeutique repose sur l'estimation du temps de traitement selon la gravité du cancer. En ce sens, un cancer peu grave a une durée moyenne de traitement inférieure à celle d'un cancer grave, sans considérer le décès. Cette notion de gravité se définit à l'aide du stade et du grade du cancer.

Par conséquent, les retraitements présentés par la suite permettent de définir le caractère grave ou non de chaque pathologie cancéreuse et d'être au clair sur le protocole thérapeutique suivi pour traiter chacun des cancers afin d'attribuer un temps de traitement au patient.

2.4 Retraitements des variables descriptives du cancer et du protocole thérapeutique suivi

Afin de déterminer le temps de thérapie suivi pour chaque patient, il est nécessaire d'identifier le degré de gravité de chaque cancer ainsi que le protocole thérapeutique mis en place. C'est pourquoi, des retraitements sur ces variables sont réalisés.

Cette partie présente les retraitements effectués dans leur grande ligne. Pour davantage de précisions, le lecteur peut consulter la méthodologie appliquée en annexe 4.4.2.

2.4.1 Les variables descriptives du cancer

Au sein de la base SEER, chaque pathologie cancéreuse est décrite selon les codes scientifiques mis en place par les institutions médicales internationales. Ces notions correspondent à des termes médicaux précis qu'il convient de simplifier. Ainsi, il est nécessaire de créer des variables descriptives des cancers, à partir des variables existantes et sans perte d'information, qui sont aisément interprétables pour définir la gravité de chaque pathologie cancéreuse. Un cancer se définissant par sa localisation, son grade et sa classe, ce sont ces trois variables qui sont créées à partir des informations disponibles.

Création de Cancer : la localisation du cancer

De manière générale, la localisation du cancer correspond au nom attribué à la pathologie. Par exemple, un cancer se développant dans le foie porte le nom de « cancer du foie ». Ces noms sont référencés dans la classification internationale des maladies pour l'oncologie (CIM-0) élaborée par l'Organisation Mondiale de la Santé (OMS). Ce répertoire est détaillé afin de classer un grand nombre de pathologies cancéreuses. Dans ces travaux, il n'est pas nécessaire d'utiliser la liste complète et détaillée de l'OMS étant donné la faible représentation de certaines. Par conséquent, un regroupement de ces modalités est indispensable selon leur fréquence dans l'extraction réalisée.

Le tableau en annexe 4.13 présente la variable initiale, `Site.recode.ICD.0.3.WHO.2008` détaillée, et les correspondances actuellement appliquées pour la création de la variable `Cancer`. Par la suite, chacune des pathologies cancéreuses est détaillée par son grade et son stade.

Création de Grade : le grade d'un cancer

Le grade correspond à l'apparence des cellules cancéreuses pour les tumeurs solides et est compris entre 1 et 4. Il donne un degré de « normalité » de ces cellules. Si les cellules cancéreuses ressemblent aux cellules normales alors le cancer est considéré comme peu agressif et est de bas grade (proche de 1). Au contraire, si les cellules cancéreuses diffèrent des cellules classiques, alors le cancer est agressif

et est de haut grade (grade 4). Enfin, la particularité du grade des cancers du sang dans la base SEER est qu'il s'agit du nom des cellules touchées et non un chiffre. Par exemple, le grade peut être égal à « *Lymphocytes B* » ou « *Lymphocytes T* ».

Le détail des modalités de **Grade** est présenté dans le tableau en annexe 4.14 qui a été réalisé à l'aide du manuel SEER (ADAMO M, GROVES C, DICKIE L, RUHL J, 2021). Cette notion est à compléter avec le stade de la pathologie cancéreuse pour comprendre l'état d'avancement du cancer.

Création de Stade : le stade d'un cancer

Le stade du cancer donne l'avancement de celui-ci à partir de sa taille et de son extension à d'autres parties du corps. Il correspond à un chiffre compris entre 1 et 4, où 1 indique que le cancer ne s'est pas étendu et qu'il reste de taille minimale, alors que 4 désigne le cas où il se propage à d'autres parties du corps et qu'il est de taille conséquente. Ces modalités numériques sont également résumées, dans la base SEER, par les termes anglais suivants dans leur ordre de degré de dissémination : « *in situ* », « *localized* », « *regional* » et « *distant* ».

Finalement, la variable **Stade** se fonde sur les modalités « *in situ* », « *localized* », « *regional* » et « *distant* » uniquement. Chaque degré de gravité de cancer est ainsi aisément compréhensible à l'aide de la localisation, du grade et du stade. Le traitement thérapeutique vient alors compléter l'étude des lignes pour la définition des temps de thérapie appliquée.

2.4.2 Les variables descriptives du protocole thérapeutique suivi

Pour chaque observation de pathologie cancéreuse, le programme SEER reporte une information globale quant au protocole thérapeutique prescrit au moment du diagnostic. Cette dernière est enregistrée dans les variables suivantes :

- `Reason.no.cancer.directed.surgery` : variable qui communique la raison pour laquelle le patient n'a pas reçu de chirurgie ;
- `Radiation.recode` : variable qui permet de décrire la radiothérapie² prescrite ;
- `Chemotherapy.recode.yes.no.unk.` : variable qui indique si le patient a reçu une chimiothérapie³ ou non ;
- `RX.Summ..Surg.Rad.Seq` : variable complémentaire à `Radiation.recode` et `Reason.no.cancer.directed.surgery` pour indiquer si le patient a eu une combinaison des traitements radiothérapie et chirurgie ;
- `RX.Summ..Systemic.Sur.Seq` : variable complémentaire à `Chemotherapy.recode.yes.no.unk.` et `Reason.no.cancer.directed.surgery` pour indiquer si le patient a eu une combinaison des traitements thérapie systémique (sans préciser laquelle parmi la chimiothérapie, l'endocrinothérapie, les thérapies ciblées et les médicaments de soutien) et chirurgie.

Le manuel SEER (ADAMO M, GROVES C, DICKIE L, RUHL J, 2021) indique que `RX.Summ..Surg.Rad.Seq` et `RX.Summ..Systemic.Sur.Seq` sont des variables de faible qualité au regard de `Reason.no.cancer.directed.surgery`, `Radiation.recode` et `Chemotherapy.recode.yes.no.unk.` En ce sens, parmi toutes les thérapies systémiques seule la chimiothérapie est étudiée par la suite, en complément de la chirurgie et de la radiothérapie. Cette hypothèse reste représentative des traitements en place

2. « *Traitement du cancer par des rayons qui détruisent les cellules cancéreuses ou stoppent leur développement. Contrairement à la chimiothérapie qui agit sur les cellules cancéreuses dans l'ensemble du corps, la radiothérapie est un traitement local, comme la chirurgie.* » (CANCER, s. d.[b])

3. « *Traitement d'un cancer par des substances chimiques qui tuent ou affaiblissent les cellules cancéreuses. Une chimiothérapie est un traitement général qui vise à détruire les cellules cancéreuses ou à les empêcher de se multiplier dans l'ensemble du corps.* » (CANCER, s. d.[c])

puisque la chimiothérapie reflète la majorité des thérapies systémiques prescrites.

Création de la variable Radiothérapie

La variable `Radiothérapie` est conçue à l'aide des variables `Radiation.recode` et `RX.Summ..Surg.Rad.Seq`. Ainsi, `Radiothérapie` permet de préciser si le patient a suivi une radiothérapie et son type lorsque l'information est disponible.

Création de la variable Chimiothérapie

La variable `Chimiothérapie` est simplement basée sur la variable `Chemotherapy.recode..yes..no.unk..`. Elle reprend les modalités de cette dernière et informe par « *Oui* » si le patient a subi une chimiothérapie et « *Non* » dans le cas contraire.

Cette nouvelle variable permet par la suite, en complément de `Radiothérapie`, de créer la variable signalant si le patient a reçu une chirurgie ou non.

Création de la variable Chirurgie

Dans l'extraction de la base SEER, il existe trois variables indiquant la prescription ou non d'une chirurgie pour traiter la pathologie cancéreuse. Il s'agit de `Reason.no.cancer.directed.surgery`, `RX.Summ..Surg.Rad.Seq` et `RX.Summ..Systemic.Sur.Seq`. Étant donné la faible qualité des deux dernières, la variable `Chirurgie` se base initialement sur `Reason.no.cancer.directed.surgery`, puis est complétée par les informations disponibles dans `RX.Summ..Surg.Rad.Seq` et `RX.Summ..Systemic.Sur.Seq`. Ainsi, cette nouvelle variable `Chirurgie` prend uniquement les modalités « *Oui* » et « *Non* », respectivement pour un cancer traité à l'aide d'une chirurgie et un cancer traité sans.

Finalement, les variables `Chirurgie`, `Radiothérapie` et `Chimiothérapie` permettent de connaître le traitement thérapeutique appliqué pour chaque cancer recensé. Enfin, la cause de sortie de chacune des observations doit également être renseignée pour la suite de ces travaux.

2.4.3 Création de la variable du statut vital

Afin de construire la table d'extra-mortalité due au cancer, il est nécessaire de disposer du statut vital de chaque patient. Il peut être en vie, décédé du cancer recensé ou décédé d'une autre cause.

Au sein de la base SEER, le suivi du statut vital de chaque patient est réalisé à l'aide des variables suivantes :

- `Vital.status.recode..study.cutoff.used.` : variable indiquant si le patient est en vie ou décédé ;
- `SEER.other.cause.of.death.classification` : variable permettant de savoir si le patient est en vie ou s'il est décédé d'une cause autre que le cancer rattaché à ce patient ;
- `SEER.cause.specific.death.classification` : variable précisant si le patient est en vie ou s'il est décédé suite au cancer observé.

A partir de ces variables, `Sortie` est égale à « *Cancer* » lorsque le patient est décédé du cancer diagnostiqué, « *Censure* » quand le décès est dû à une autre cause, et « *Vivant* » si le patient est toujours en vie au 31/12/2019.

L'extraction de la base SEER comporte désormais sept variables supplémentaires, aisément interprétables, qui permettent de comprendre le degré de gravité de chaque observation cancéreuse, le protocole thérapeutique qui a été suivi pour la traiter ainsi que le suivi du statut vital du patient. A partir

de ces variables, il convient également de sélectionner uniquement les patients qui seront en mesure d'appliquer le droit à l'oubli.

2.5 Retraitements des observations pré-sélectionnées

A l'aide des variables créées pour l'estimation de l'âge de fin de traitement et des variables sélectionnées à l'étape de pré-sélection, il est nécessaire de sélectionner les lignes qui correspondent au cadre de ces travaux.

2.5.1 Les patients touchés par un cancer et concernés par le droit à l'oubli

L'application du droit à l'oubli est conditionnée à l'entrée du patient dans sa phase de rémission. Toutefois, certains cancers sont considérés comme invasifs, empêchant le patient d'entrer dans cette phase (LEE, s. d.). Il s'agit des cancers étendus présentant des métastases. Ces pathologies doivent être exclues de notre étude car aucune fin de traitement thérapeutique, au sens de la définition du droit à l'oubli, ne pourra être identifiée précisément. Par conséquent, puisque `Stade` = « *distant* » implique que le patient a des métastases, seules les tumeurs où la variable `Stade` est égale à « *in situ* », « *localized* » ou « *regional* » sont conservées.

La base SEER enregistre les cancers diagnostiqués à l'autopsie des patients quand ils sont la cause de leur mort. Dans la même logique que le point d'exclusion précédent, puisque la maladie est signalée à la mort, le patient n'a jamais été soigné pour cette dernière et il n'est donc pas entré en phase de rémission. Ainsi, les lignes où le cancer est déclaré à l'autopsie, soit `Reason.no.cancer.directed.surgery` = « Unknown; death certificate; or autopsy only (2003+) » ou `Reason.no.cancer.direct.ed.surgery` = « Not recommended, contraindicated due to other cond; autopsy only (1973-2002) », sont retirées de la base de données d'étude.

Enfin, il existe des patients où aucun protocole thérapeutique n'a été mis en place. Là encore, cela implique que le malade n'a pas pu entrer en phase de rémission puisque le cancer n'a pas été traité. Ces lignes, où les variables de traitement sont vides, sont donc exclues de la base d'étude.

Par conséquent, la base d'étude comporte uniquement des patients où l'application du droit à l'oubli est envisageable. Il est nécessaire de traiter également les lignes où des valeurs sont manquantes.

2.5.2 Retraitements des valeurs manquantes

Pour la construction de l'âge à la fin du traitement thérapeutique, les variables suivantes seront utilisées :

- `Stade` et `Grade` : degré de gravité de la pathologie cancéreuse recensée ;
- `Chirurgie`, `Radiotherapie` et `Chimiotherapie` : détail du protocole thérapeutique suivi ;
- `Age.recode.with.single.ages.and.100.` : âge révolu du patient au moment du diagnostic ;
- `Months.from.diagnosis.to.treatment` : nombre de mois entre le diagnostic et le début du traitement thérapeutique ;
- `Survival.months` : nombre de mois de survie à partir du diagnostic.

Les lignes où ces variables ne sont pas renseignées sont retirées de la base de données.

Ainsi, chaque observation conservée correspond à un patient dont l'application du droit à l'oubli est réalisable et où son âge à la fin du traitement thérapeutique peut être estimé. Toutefois, plusieurs

cancers sont rattachés à un unique `Patient.ID`, il convient alors d'agrèger ces lignes.

2.5.3 Agrégation des pathologies rattachées à un unique patient

Dans la définition des paramètres de l'extraction de la base SEER, il a été précisé de renseigner le suivi du patient à l'aide de la variable `Patient.ID`. En effet, un patient peut disposer de plusieurs lignes dans la base de données s'il est atteint de plusieurs cancers. En ce sens, il est nécessaire de regrouper ces lignes puisque cela implique qu'après son premier traitement thérapeutique visant à traiter sa tumeur maligne primaire, il rechute, et il sort ainsi de sa phase de rémission.

L'hypothèse considérée dans ces travaux est que le délai du droit à l'oubli court à compter de la fin du traitement thérapeutique de la dernière tumeur recensée uniquement. Concrètement, si un patient est atteint d'un premier cancer diagnostiqué en 2015, puis d'un second en 2019, alors sa phase de rémission entre 2015 et 2019 est ignorée, et la phase de rémission considérée ici est celle à compter de 2019. Par conséquent, seule la ligne du cancer le plus récent est conservée pour chaque patient.

Pour conclure, la base de données comporte désormais 4 119 072 lignes. Il convient par la suite d'estimer pour chaque patient son âge de fin de traitement thérapeutique et d'effectuer les tests de cohérence qui en émanent.

2.6 Construction de l'âge à la fin du traitement thérapeutique suivi

Les données SEER donnent les informations quant à l'âge au moment du diagnostic, le temps entre ce dernier et le début du traitement, ainsi que celles sur la thérapie appliquée. En ajoutant le temps écoulé entre l'âge au diagnostic et le début du traitement, l'âge au début du traitement est identifiable. Quant à l'âge en fin de traitement, il est reconstituable en prenant des hypothèses sur les durées de traitement à l'aide des variables sur les thérapies et la classification du cancer. Ainsi, l'objectif de cette partie est de présenter la méthodologie appliquée pour la définition de la durée du traitement ainsi que celle utilisée pour définir l'âge à la fin de traitement.

2.6.1 Estimation des durées de traitement

Dans un premier temps, la Convention AERAS définit le protocole thérapeutique comme les soins appliqués à la suite du diagnostic de la pathologie cancéreuse. Les traitements persistants comme l'immunothérapie⁴ ou l'hormonothérapie⁵ et la chirurgie reconstructrice en sont exclus et ne doivent ainsi pas être pris en considération pour le compte du délai du droit à l'oubli. Ainsi, l'hypothèse suivante est considérée : seule la durée de la thérapie primaire renseignée dans la base SEER au moment du diagnostic (**Radiothérapie**, **Chimiothérapie** et **Chirurgie**) est prise en compte pour déterminer la date de fin de traitement à partir de laquelle le délai du droit à l'oubli s'écoule.

De plus, il existe un grand nombre de traitements pour soigner un cancer. Le traitement nécessaire est adapté selon le type de cancer et son degré de gravité. De manière générale, quelques patients suivent une unique médication mais il est plus courant de prescrire une combinaison de traitements comme

4. « *Traitement qui vise à stimuler les défenses immunitaires de l'organisme contre les cellules cancéreuses.* » (CANCER, s. d.[d])

5. « *Traitement du cancer qui vise à réduire ou à empêcher l'activité ou la production d'une hormone susceptible de stimuler la croissance d'une tumeur cancéreuse.* » (CANCER, s. d.[e])

la chirurgie couplée à la chimiothérapie ou la radiothérapie. Par conséquent, une autre hypothèse est prise ici : lorsque le patient reçoit une combinaison de traitements alors la durée de sa thérapie complète est la somme de chacune des thérapies suivies.

En complément de ces deux hypothèses, il convient désormais de définir pour chaque traitement recensé dans les variables **Radiotherapie**, **Chimiotherapie** et **Chirurgie** une durée minimale moyenne et une durée maximale moyenne de temps passé dans celui-ci. Le tableau 2.1 présente ces durées.

TABLE 2.1 – Durées minimales et maximales fixées pour chaque traitement

VARIABLE	MODALITÉ	DURÉE MINIMALE	DURÉE MAXIMALE
Chirurgie	Oui	1 jour soit arrondi à 1 mois	1 jour soit arrondi à 1 mois
Chimiotherapie	Oui	3 mois	6 mois
Radiotherapie	Oui	3 semaines soit arrondi à 1 mois	9 semaines soit arrondi à 3 mois
Radiotherapie	Beam radiation = radiothérapie externe	3 semaines soit arrondi à 1 mois	9 semaines soit arrondi à 3 mois
Radiotherapie	Combinaison radiothérapie interne et externe	3 semaines soit arrondi à 1 mois	9 semaines soit arrondi à 3 mois
Radiotherapie	Implant radioactif	1 mois	3 mois
Radiotherapie	Radioisotope	1 jour soit arrondi à 1 mois	1 jour soit arrondi à 1 mois

Après avoir défini les durées minimales et maximales de chaque traitement, une interpolation linéaire entre le minimum et le maximum est réalisée afin d'estimer pour chaque degré de gravité de cancer un temps de traitement. Le tableau 2.2 recense l'ordre établi dans ces travaux pour classer les cancers du moins agressif au plus agressif selon le stade et le grade.

TABLE 2.2 – Classement par ordre croissant des degrés de gravité des cancers établi

STADE	GRADE
In situ	Grade I
In situ	Grade II
In situ	Grade III
In situ	Grade III
Localized	Grade I
Localized	Grade II
Localized	Grade III
Localized	Grade IV
Regional	Grade I
Regional	Grade II
Regional	Grade III
Regional	Grade IV
In situ - Localized - Regional	B-cell ; B-precursor cell - NK cell (natural killer cell) - Null cell ; Non-T-non-B - T-cell ; T-precursor cell - Unknown

Finalement, un temps de traitement est estimé pour chaque combinaison (Traitement, **Stade**, **Grade**). Les matrices des durées de traitement sont disponibles en annexe (??). Le tableau 2.3 présente une de ces matrices pour le cas de la chimiothérapie.

Par conséquent, à partir de ces matrices de temps de traitement et des caractéristiques de chaque observation, trois variables sont ajoutées à la base de données :

- **Temps___chirurgie** : durée de la chirurgie si le patient en a eu une, sinon 0 ;

TABLE 2.3 – Temps de traitement de la chimiothérapie selon le degré de gravité du cancer traité

STADE	GRADE	TEMPS ARRONDI AU SUPÉRIEUR
In situ	Grade I	4 mois
In situ	Grade II	4 mois
In situ	Grade III	4 mois
In situ	Grade III	4 mois
Localized	Grade I	5 mois
Localized	Grade II	5 mois
Localized	Grade III	5 mois
Localized	Grade IV	5 mois
Regional	Grade I	6 mois
Regional	Grade II	6 mois
Regional	Grade III	6 mois
Regional	Grade IV	7 mois
In situ Localized - Regional	B-cell ; B-precursor cell - NK cell (natural killer cell) - Unknown Null cell ; Non-T-non-B - T-cell ; T-precursor cell	7 mois

— Temps__radiotherapie : durée de la radiothérapie si le patient en a reçue une, sinon 0 ;

— Temps__chimiotherapie : durée de la chimiothérapie si le patient en a reçue une, sinon 0.

A partir de ces variables ajoutées, l'âge à la fin du traitement thérapeutique peut être estimé.

2.6.2 Estimation de l'âge à la fin du traitement thérapeutique

A partir des variables Temps__chirurgie, Temps__radiotherapie et Temps__chimiotherapie, chaque patient de la base dispose désormais d'une durée en mois pour le protocole thérapeutique suivi.

D'après le manuel SEER (ADAMO M, GROVES C, DICKIE L, RUHL J, 2021), l'âge au diagnostic correspond à l'âge révolu du patient au temps du diagnostic, c'est-à-dire le nombre d'années entières qui se sont écoulées depuis le jour de la naissance du patient. Afin de calculer l'âge exact à la fin du traitement thérapeutique, une hypothèse majeure est considérée ici : le jour et le mois de la date de diagnostic sont égaux au jour et au mois de la date de naissance du patient. Ainsi, l'âge exact à la fin du traitement thérapeutique est calculé de la manière suivante :

$$\text{Age_fin_traitement} = \frac{\text{Months.from.diagnosis.to.treatment} + \text{Duree.traitement}}{12} + \text{Age.recode.with.single.ages.and.100.}$$

avec

$$\text{Duree.traitement} = \text{Temps_chimiotherapie} + \text{Temps_chirurgie} + \text{Temps_radiotherapie}$$

Suite à la conception de cette variable d'âge à la fin de traitement thérapeutique, il est nécessaire de réaliser un test de cohérence sur la base de données SEER entre l'âge estimé et les informations disponibles.

2.6.3 Test de cohérence de l'âge de fin de traitement estimé

L'âge de fin de traitement thérapeutique a été estimé sans considérer le nombre de mois pendant lequel le patient a survécu. Ainsi, il existe dans la base SEER des lignes où le temps de traitement est supérieur au temps survécu. Ces lignes sont donc exclues de l'étude car cela implique que le patient soit décédé avant d'entrer en phase de rémission.

Pour conclure, la base SEER comporte désormais 3 891 413 lignes ainsi que l'ensemble des variables nécessaires à la construction de la table d'extra-mortalité due au cancer après traitement thérapeutique, soient :

- `Age__fin__traitement` : âge de début de la table d'extra-mortalité ;
- `Survival.months` : variable permettant de déterminer l'âge de sortie du patient ;
- `Sortie` : variable déterminant la cause de sortie exacte du patient de l'échantillon.

Elle peut ainsi être analysée afin de comprendre la population de pathologies cancéreuses étudiées ici.

2.7 Analyse statistique

Avant de construire la table d'extra-mortalité due au cancer attendue dans ces travaux, il est nécessaire d'analyser la base de données SEER retraitée. Cette analyse permet d'appréhender des faits ou tendances qui pourraient influencer les futurs résultats. Ainsi, une analyse globale de la population américaine observée est présentée, puis une analyse multidimensionnelle est réalisée et permet d'anticiper le futur calcul d'extra-mortalité.

2.7.1 Généralités sur l'échantillon retraité

La base SEER finalement construite comporte 3 891 413 lignes. Il convient de s'intéresser aux caractéristiques de notre population de personnes atteintes d'un cancer ainsi qu'aux pathologies recensées.

2.7.1.1 Caractéristiques de la population

La population observée après retraitements se compose de 53% femmes et 47% d'hommes sur la période 2000-2019. Le tableau 2.4 montre une répartition homogène de la population à travers les années de diagnostic et le genre.

TABLE 2.4 – Répartition des observations par genre et année de diagnostic

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Femme	2.18%	2.25%	2.26%	2.23%	2.37%	2.42%	2.46%	2.55%	2.64%	2.73%
Homme	1.95%	2.05%	2.07%	2.03%	2.21%	2.22%	2.33%	2.46%	2.47%	2.49%
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Femme	2.73%	2.80%	2.87%	2.91%	3.01%	3.10%	3.17%	3.20%	3.30%	2.05%
Homme	2.48%	2.52%	2.41%	2.41%	2.46%	2.54%	2.59%	2.65%	2.66%	1.77%

Au sein de la base SEER, l'âge au diagnostic du patient est renseigné. Le tableau 2.5 indique que l'âge médian au diagnostic pour les cancers diagnostiqués est, pour les femmes et pour les hommes, respectivement de 62 ans et 65 ans. Cela implique que la population étudiée est relativement âgée.

TABLE 2.5 – Statistiques sur l'âge au diagnostic selon le sexe

	Minimum	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Maximum
Femme	0	51.00	62.00	61.12	73.00	99.00
Homme	0	56.00	65.00	63.85	73.00	99.00

Ainsi, l'échantillon étudié de la base SEER est relativement âgé au moment du diagnostic et homogène en terme de répartition des sexes sur les années de diagnostic. Il faut par la suite étudier les cancers reportés.

2.7.1.2 Caractéristiques des pathologies cancéreuses

A travers cette partie, une analyse des pathologies cancéreuses constituant la base SEER retraitée est réalisée.

Au sein de la base SEER retraitée, les cancers du sein, de la prostate et du côlon-rectum représentent approximativement 49.4% des cancers diagnostiqués entre 2000 et 2019 (tableau 2.6). Plus particulièrement, d'après la figure 2.3, le cancer du sein reste le cancer le plus fréquemment diagnostiqué pour les femmes en 2018 dans la base SEER retraitée. Cette analyse est similaire pour le cancer de la prostate chez les hommes en 2018.

TABLE 2.6 – Répartition des 10 cancers les plus représentés sur 2000-2019

	Pourcentage
Sein	22.60%
Prostate	15.60%
Colon et rectum	11.20%
Peau hors basale et squameuse	7.50%
Vessie	6.10%
Poumon et bronche	6.00%
Systeme endocrinien	4.50%
Corpus et uterus	4.40%
Rein et pelvis renal	4.20%
Cavite buccale et pharynx	2.60%

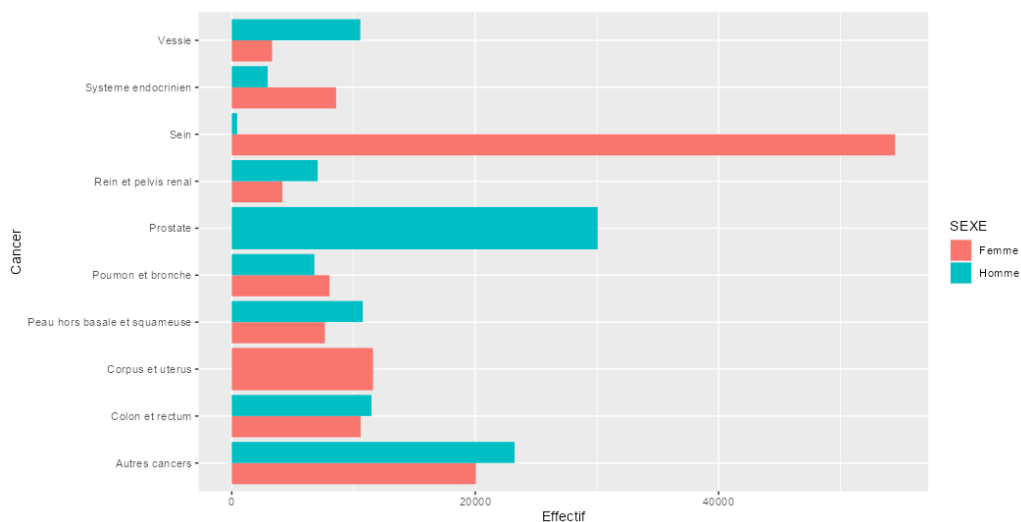


FIGURE 2.3 – Répartition des cancers diagnostiqués en 2018 selon le sexe

Puis, le tableau 2.7 souligne que le cancer le plus létal sur la période 2000 à 2019 est celui du poumon et des bronches pour tous les sexes confondus.

TABLE 2.7 – Répartition des 10 cancers les plus létaux sur 2000-2019

	Pourcentage
Poumon et bronche	16.30
Sein	14.30
Colon et rectum	13.80
Vessie	7.00
Cerveau et autre système nerveux	5.40
Prostate	4.80
Pancreas	3.80
Peau hors basale et squameuse	3.70
Cavité buccale et pharynx	3.60
Rein et pelvis renal	3.30

Enfin, d'un autre point de vue, il est intéressant d'observer la répartition des cancers selon l'ancienneté en rémission du patient à sa sortie (décès ou censure). Ainsi, la figure 2.4 reflète la présence majoritaire des cancers du sein et de la prostate tout au long du temps passé en rémission. De manière contraire, la proportion du cancer de la vessie diminue nettement après les cinq premières années en rémission. Il en est de même pour les cancers du col de l'utérus et du colon et du rectum. Puisque la sortie concerne ici le décès ou la censure, il convient de s'intéresser désormais à la sinistralité de notre base de données afin de préciser cette analyse.

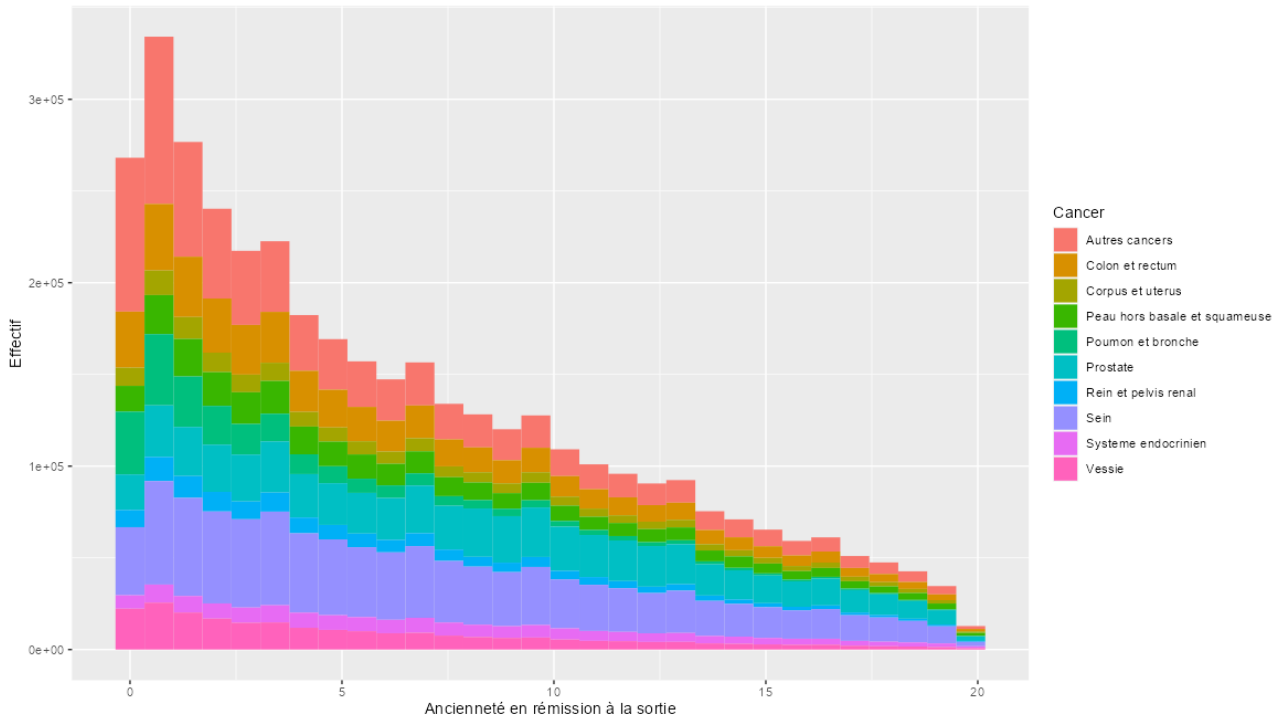


FIGURE 2.4 – Répartition des cancers selon l'ancienneté en rémission à la sortie

Par conséquent, cette brève analyse donne une description de la population observée dans la base SEER retraitée. Par la suite, il est primordial de réaliser une analyse multidimensionnelle de la base de données SEER retraitée. Cette analyse permet de comprendre, dans sa globalité, le parcours en rémission des patients puis de mettre en évidence des phénomènes qui pourraient influencer cette étude.

2.7.2 Analyse multidimensionnelle de la base finale SEER

L'objectif de ce mémoire étant de mesurer l'extra-mortalité causée par un cancer diagnostiqué et traité en phase de rémission, ce risque est étudié selon l'âge à la fin du traitement thérapeutique et la durée passée en rémission. Ainsi, cette partie permet, dans un premier temps, d'analyser la population observée selon ces variables. Puis, il convient de s'interroger également sur les sorties de la base selon ces variables de durées. Enfin, une analyse de la répartition des cancers est produite selon ces mêmes variables.

2.7.2.1 Répartition des âges à la fin de traitement thérapeutique et des durées en rémission

Dans ces travaux, l'âge à la fin de traitement thérapeutique correspond à l'âge d'entrée en rémission. Cette variable de durée est utilisée pour la construction de la table d'extra-mortalité.

Selon la méthode établie précédemment quant à la définition de la durée de traitement, le tableau 2.8 montre qu'en moyenne les traitements sont estimés à un peu plus de trois mois. A cela s'ajoute le temps entre le diagnostic et le début de ce traitement qui est donné en moyenne à un mois (tableau 2.9).

TABLE 2.8 – Statistiques sur la durée de traitement estimée en mois entiers

	Minimum	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Maximum
Durée	1	1	1	3.041	4	12

TABLE 2.9 – Statistiques sur le nombre de mois entiers écoulés entre le diagnostic et le début du traitement

	Minimum	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Maximum
Durée	0	0	1	1.056	2	24

A partir de ces deux durées étudiées, l'âge à la fin du traitement thérapeutique suivi peut être calculé. En ce sens, la figure 2.5 montre que la fin des traitements thérapeutiques se concentre sur la tranche d'âge 55 ans - 75 ans. Cette distribution est précisée par le tableau 2.10 où la médiane de notre échantillon est indiquée à 64 ans. Par conséquent, la population de personnes touchées par une pathologie cancéreuse et qui ont fini leur protocole thérapeutique est une population relativement âgée.

TABLE 2.10 – Statistiques sur l'âge à la fin de traitement thérapeutique

Minimum	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Maximum
0.08	54.08	64.08	62.74	73.25	99.92

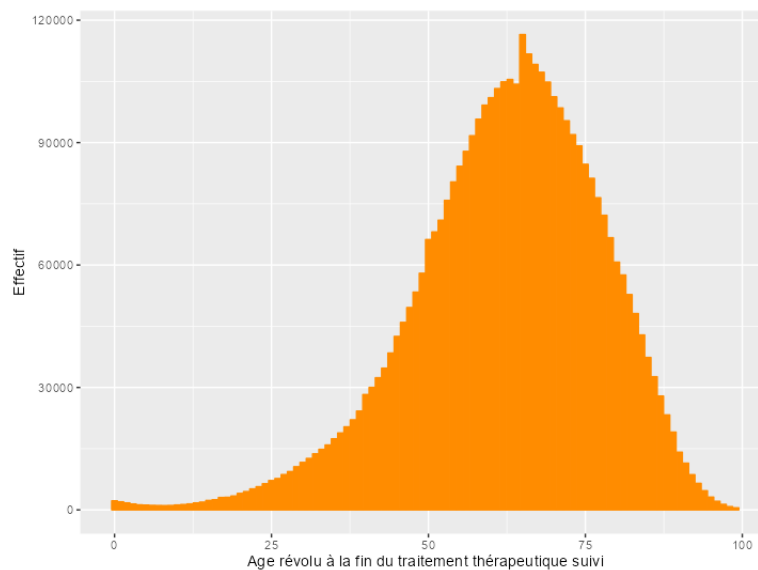


FIGURE 2.5 – Distribution des âges révolus à la fin du traitement thérapeutique suivi

Une fois le traitement thérapeutique achevé, la période de rémission commence. C'est sur chacune de ces années passées en rémission que l'extra-mortalité sera calculée pour être en adéquation avec le cadre du droit à l'oubli.

La figure 2.6 montre que le nombre de patients de la base SEER retraitée diminue nettement en fonction de l'ancienneté en rémission à la sortie (décès ou censure). Plus précisément, le tableau 2.11 indique que 50% de la population de l'échantillon a une durée en rémission de plus de cinq ans, et 25% de la population a un temps de rémission entre cinq et dix ans. Il s'agit de cette population, ayant

une ancienneté en rémission entre cinq et dix ans, qui entre dans le nouveau périmètre d'application du droit à l'oubli.

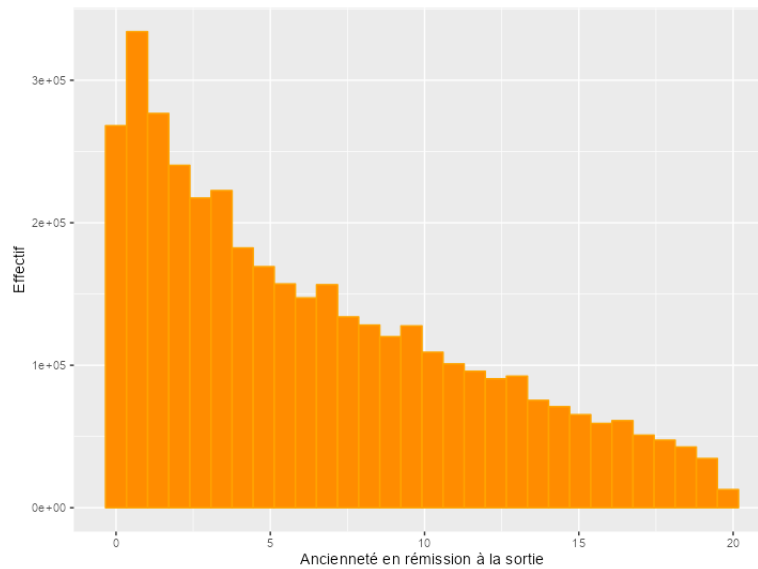


FIGURE 2.6 – Distribution des anciennetés en rémission à la sortie (décès ou censure)

TABLE 2.11 – Statistiques sur l'ancienneté en rémission à la sortie (décès ou censure) en années

Minimum	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Maximum
0	2.00	5.25	6.49	10.17	19.83

Enfin, il est intéressant d'observer la durée moyenne en rémission selon l'âge à la fin du traitement thérapeutique à l'aide du graphique 2.7. Ainsi, la durée moyenne en rémission, toutes sorties confondues (décès ou censure), décroît nettement en fonction de l'âge à la fin de traitement thérapeutique à partir de 40 ans.

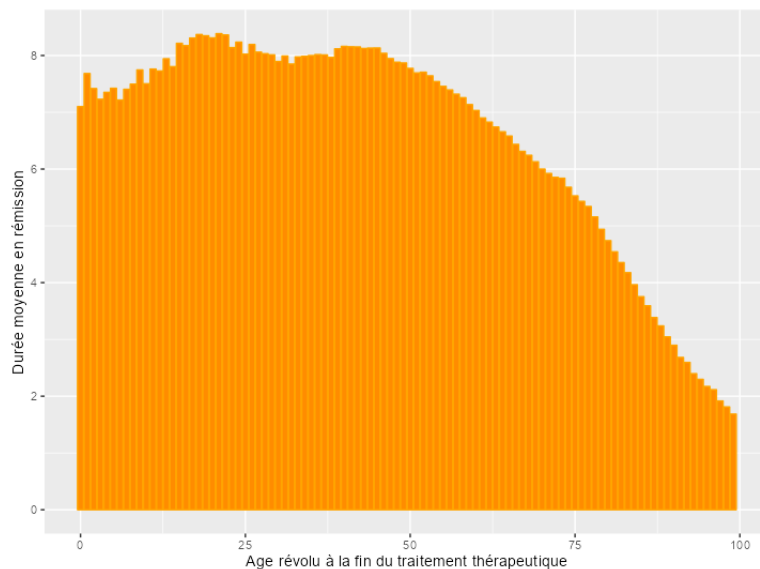


FIGURE 2.7 – Durée moyenne en année en rémission à la sortie (censure et décès) selon l'âge révolu à la fin du traitement thérapeutique

Par conséquent, cette analyse des variables de durée, qui sont utilisées pour la construction de la table d'extra-mortalité, souligne les faits suivants :

- la population étudiée est relativement âgée lors de l'entrée en rémission avec une forte concentration entre les âges 55 ans et 75 ans ;
- elle est répartie de manière décroissante selon l'ancienneté en rémission ;
- 25% de la population dispose d'une ancienneté en rémission entre cinq et dix ans ;
- la durée moyenne en rémission décroît avec l'âge à la fin du traitement thérapeutique.

De la même façon, les sorties de la population doivent être analysées avant de construire la table d'extra-mortalité.

2.7.2.2 Analyse des sorties de l'échantillon observé

Dans la théorie des modèles de durée, il convient d'analyser la cause de sortie d'une personne de l'échantillon étudié. Dans cette étude, un patient peut ne plus être observé car :

- il est décédé des suites du cancer diagnostiqué et traité ;
- il est décédé d'une autre cause que le cancer diagnostiqué et traité. Son décès des suites du cancer ne se réalisera pas et ne sera ainsi pas reporté, il s'agit d'une donnée censurée à droite ;
- il est toujours en vie. Ainsi, son potentiel décès des suites du cancer diagnostiqué et traité n'est pas encore connu. Il s'agit également d'une donnée censurée à droite.

Comme il est intéressant de l'observer dans le tableau 2.12, les décès causés par les cancers diagnostiqués et traités représentent 16% des causes de sortie de la population étudiée. Les données censurées constituent ainsi 84% de la base SEER retraitée.

TABLE 2.12 – Effectifs par sortie

	Pourcentage
Décès causé par le cancer traité	16.00%
Décès autre cause	17.70%
Vivant	66.30%

A partir de la figure 2.8, les répartitions des sorties peuvent être étudiées. Elle montre que la distribution de la variable âge d'entrée en rémission varie selon le type de sortie :

- les malades toujours vivants se concentrent sur la tranche d'âge 50-75 ans. La distribution de leur âge à la fin du traitement thérapeutique est ainsi semblable à celle de l'échantillon avec une moyenne d'âge en rémission calculée à 59 ans ;
- la distribution des âges à la fin du traitement thérapeutique des malades décédés d'une autre cause que le cancer diagnostiqué et traité se décale vers la droite impliquant ainsi un effectif plus âgé que celui de la population étudiée. La moyenne d'âge d'entrée en rémission est de 73.18 ans, le premier quartile est établi à 66.42 ans et le troisième à 81.17 ans ;
- finalement, les patients décédés du cancer diagnostiqué et traité ont des âges d'entrée en rémission inférieurs à la population qui décède d'une autre cause et est davantage dispersée. En ce sens, la moyenne d'âge de fin de traitement est de 66.68 ans, le premier quartile est à 57.75 ans et le troisième quartile à 77.17 ans.

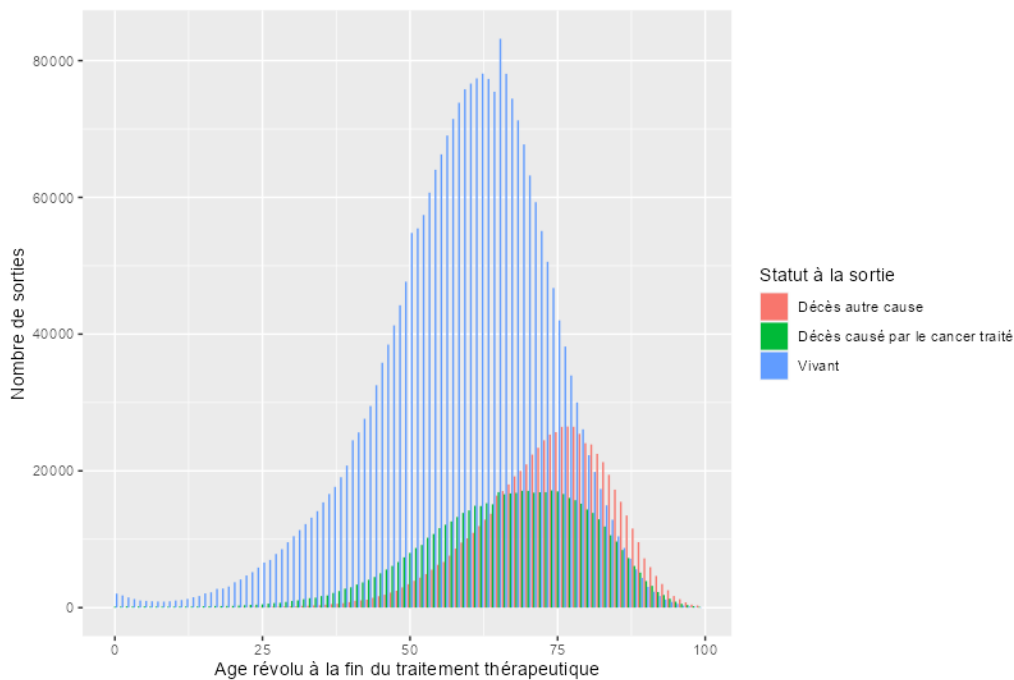


FIGURE 2.8 – Répartition des sorties selon l'âge révolu à la fin du traitement thérapeutique

Avec la réduction du délai d'application du droit à l'oubli de dix à cinq ans, ce sont les décès des suites du cancer diagnostiqué et traité entre les anciennetés cinq et dix ans qui entrent dans le périmètre des risques du droit à l'oubli déjà en place. Ainsi, ce sont ces décès qui sont au coeur de cette étude.

La figure 2.9 indique, avec la somme des accolades oranges et bleues, que la tranche 5 à 10 ans

d'ancienneté concerne un peu moins de 25% des décès causés par les cancers traités, tous les âges de fin de traitement thérapeutique confondus. La distribution de ces décès étant concentrée sur les anciennetés inférieures à 5 ans, les assureurs étaient jusqu'à présent peu exposés à ce risque avec un délai d'application du droit à 10 années passées en rémission. A la différence des anciennetés en rémission pour le décès autre cause et les patients vivants qui sont davantage réparties uniformément entre 0 et 15 ans en rémission.

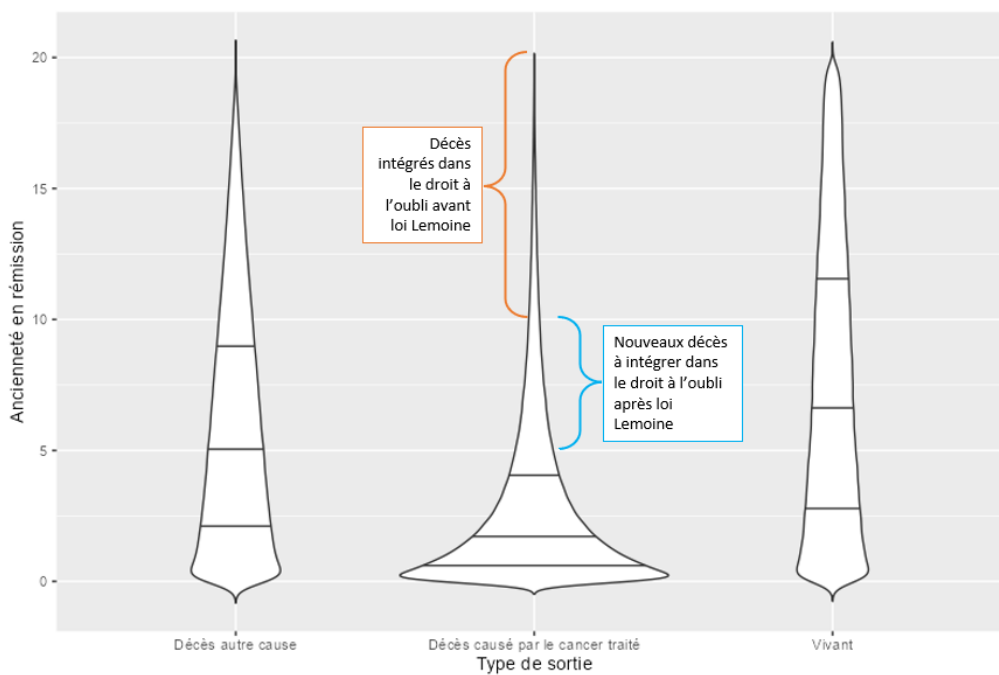


FIGURE 2.9 – Diagramme en violon de l'ancienneté en rémission selon le type de sortie

Enfin, l'étude de la répartition des décès selon l'ancienneté en rémission et selon le cancer diagnostiqué et traité permet de comprendre s'il existe des tendances de mortalité selon les cancers. En ce sens, le graphique 2.10 montre que les décès dus au cancer de la vessie ont lieu dans les cinq premières années du temps en rémission. Il en est de même pour les cancers du poumon et des bronches. De manière contraire, les patients diagnostiqués et traités des cancers du sein et de la prostate passent des temps en rémission plus longs, notamment sur la tranche 5 à 10 ans en rémission.

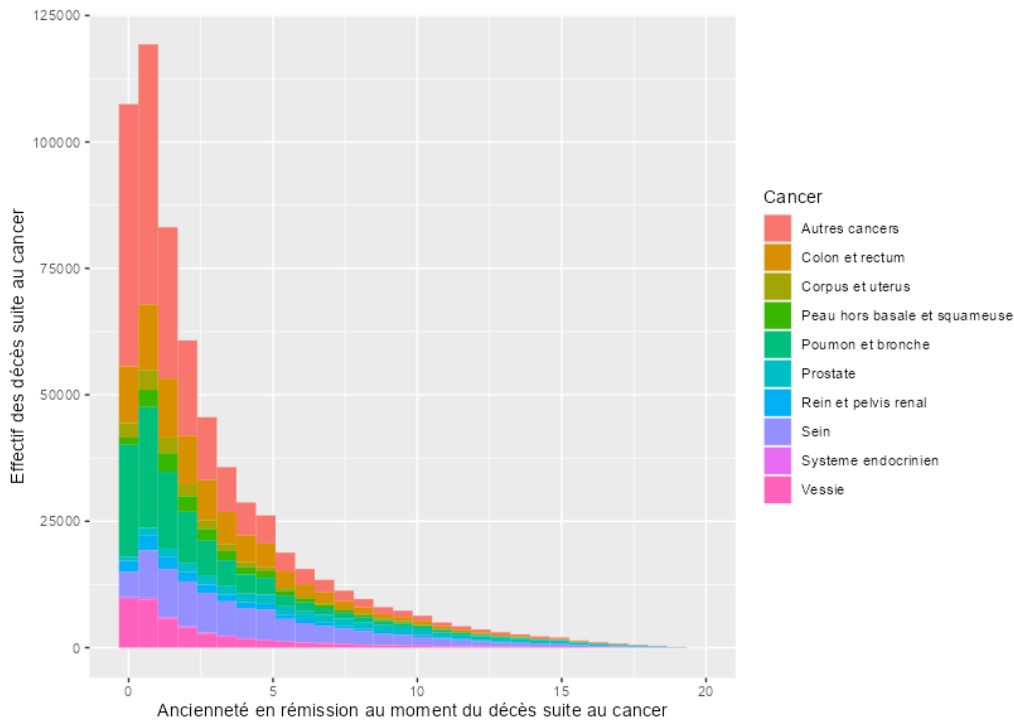


FIGURE 2.10 – Répartition des décès cancers selon l’ancienneté en rémission au moment du décès

Par conséquent, le risque de décéder d’un cancer diagnostiqué et traité existe entre la cinquième et la dixième année en rémission même si une grande partie des décès semble se réaliser avant la cinquième année de phase de guérison. Ainsi, il est essentiel de calculer les taux d’extra-mortalité du cancer pour comprendre l’impact de la réduction du délai d’application du droit à l’oubli de dix à cinq ans. Toutefois, avant d’estimer ces taux, il convient de déterminer leur structure.

2.8 Les segmentations des taux d’extra-mortalité

Les taux d’extra-mortalité des suites d’un cancer diagnostiqué et traité sont détaillés selon l’âge d’entrée en rémission et l’ancienneté en rémission. Il est également possible de les discrétiser selon le genre puisqu’il est couramment utilisé dans l’évaluation des engagements d’un assureur. De plus, les cancers diagnostiqués et traités sont recensés sur vingt ans, du 01/01/2000 au 31/12/2019. Il est alors intéressant de se demander si les tendances d’extra-mortalité sont identiques au cours de ces vingt années. Ainsi, cette partie a pour objectif de justifier ces choix de segmentation des taux et de présenter les bornes inférieures et supérieures des variables de durée.

2.8.1 Construction d’une table d’extra-mortalité pour chaque sexe

Depuis l’extension de la Gender Directive au secteur de l’assurance en 2013, il est interdit de segmenter un tarif selon le sexe de l’individu. Toutefois, les compagnies d’assurance peuvent analyser leur portefeuille selon ce dernier et segmenter ainsi leurs lois pour calculer leurs provisions. En ce sens, il est nécessaire de s’interroger sur une possible différence significative des taux d’extra-mortalité selon le genre.

Les graphiques en figure 2.11 montrent que les distributions des décès des suites du cancer diagnostiqué et traité en fonction de l'âge et en fonction de l'ancienneté diffèrent selon le genre. Pour l'âge à la fin du traitement thérapeutique, les décès pour les hommes sont davantage plus concentrés sur la tranche d'âge 60 à 75 ans au contraire de celle des femmes qui est plus étendue allant de 50 à 75 ans. De plus, au cours de la première année en rémission, le nombre de décès pour les hommes est plus important que pour les femmes (environ 12 000 décès supplémentaires), alors que cette différence semble tendre vers zéro après cinq années passées en rémission. Par conséquent, il convient de mesurer le probable lien entre le genre et le décès des suites du cancer à l'aide du test d'indépendance du khi-deux.

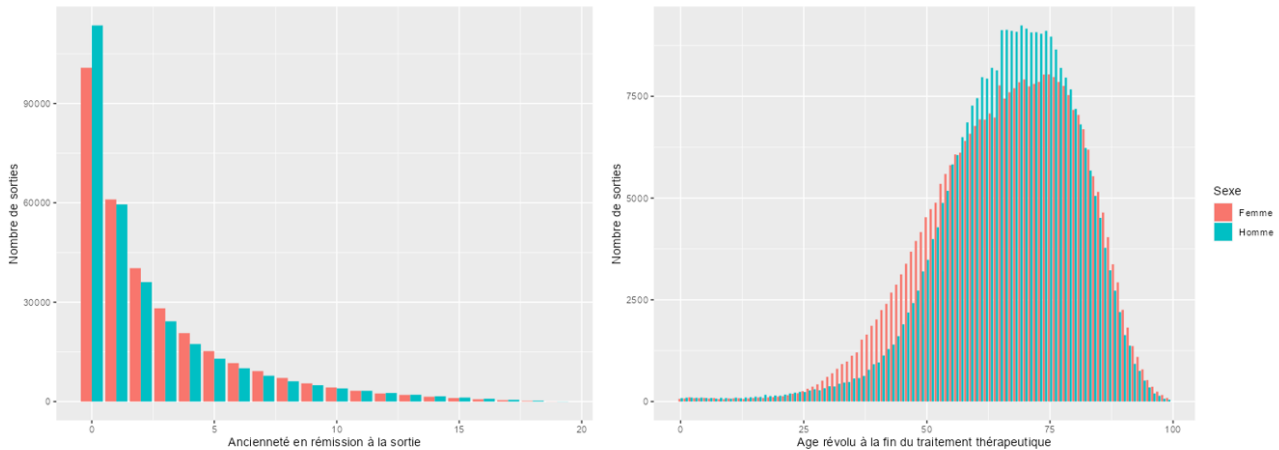


FIGURE 2.11 – Distribution des décès cancer selon le sexe et l'ancienneté, puis le sexe et l'âge à la fin du traitement thérapeutique

Le test d'indépendance du χ^2 permet de justifier l'absence de lien statistique entre deux variables aléatoires. C'est pourquoi, ce test est appliqué pour étudier la liaison entre le genre et la réalisation ou non du décès des suites du cancer diagnostiqué et traité

Principe

Soient *Sexe*, la variable aléatoire indiquant le sexe d'un individu, et *d* la réalisation de son décès des suites du cancer ou non ($d_i = 1$ si l'individu *i* est décédé des suites de son cancer diagnostiqué et traité, 0 sinon).

Les hypothèses du test d'indépendance du χ^2 sont les suivantes :

- H_0 : *Sexe* et *d* sont indépendantes ;
- H_1 : *Sexe* et *d* sont dépendantes.

Pour tester H_0 , la construction d'un indicateur d'écart K^2 est nécessaire. Il correspond à la somme, pour chaque couple $(Sexe, d)$, des écarts au carré divisé par l'effectif théorique. Comme le démontrent GREENWOOD et NIKULIN (1996), K^2 suit une loi du χ^2 à 1 degré de liberté. Ainsi,

$$K^2 = \sum_{k,l} \frac{(O_{k,l} - T_{k,l})^2}{T_{k,l}} = \sum_{k,l} \frac{O_{k,l}^2}{T_{k,l}} - N$$

où O est l'effectif observé, T l'effectif théorique, N l'effectif totale, $k \in \{1; 2\}$ et $l \in \{1; 2\}$.

A partir de cet indicateur, la probabilité suivante est calculée : $P(\chi^2 \geq K^2)$. Le risque d'erreur accepté dans cette étude est de 5%. Ainsi, si cette probabilité est inférieure à ce seuil alors l'hypothèse nulle est rejetée.

Application

Après application de ce test sur la base de données SEER retraitée, les résultats du test sont présentés dans le tableau 2.13.

TABLE 2.13 – Résultats du test d'indépendance du khi-deux

χ^2	Degrés de liberté	P-valeur
2200.6	1	< 2.2e-16

La p-valeur du test étant strictement inférieure à 0.05, le risque d'erreur toléré ici, l'hypothèse nulle est rejetée. Ainsi, il est nécessaire d'estimer la loi d'extra-mortalité des cancers des suites du cancer diagnostiqué et traité pour chaque genre. Pour la suite de l'étude, la base SEER est segmentée selon le genre.

Un phénomène identique de lien entre l'année du diagnostic et le taux d'extra-mortalité est possible étant donné le progrès médical sur les vingt dernières années. Ainsi, ce probable lien est également étudié.

2.8.2 Période d'observation

La base de données SEER étudiée rassemble les cancers diagnostiqués et traités entre le 01/01/2000 et le 31/12/2019. Il est intéressant de se questionner sur les tendances d'extra-mortalité au cours de cette période, notamment suite aux progrès scientifiques.

Afin de comprendre si la période du diagnostic a un impact sur le nombre de décès des suites des cancers diagnostiqués et traités, **un test du log-rank** est effectué. Ce test non paramétrique est appliqué, pour la base des hommes et celle des femmes, sur l'estimation de la fonction de survie par Kaplan-Meier prenant ainsi en compte la censure. Cette théorie est approfondie en troisième partie de ce mémoire.

Les bases de données SEER retraitées pour la population féminine et la population masculine sont séparées en deux périodes telles que :

- la période 1 regroupe tous les cancers diagnostiqués entre le 01/01/2000 et le 31/12/2008 ;
- la période 2 regroupe tous les cancers diagnostiqués entre le 01/01/2009 et le 31/12/2019.

Principe

Dans la suite, le principe du log-rank est présenté pour la population féminine. Il convient d'étendre cette définition à la population masculine également.

Ce test du log-rank repose sur les hypothèses suivantes :

- H_0 : les fonctions de survie des échantillons période 1 et période 2 sont égales ;
- H_1 : les fonctions de survie des échantillons période 1 et période 2 sont différentes.

Pour cette théorie, soient :

- t_i les instants des décès du cancer diagnostiqué et traité ;

- n_{1_i} le nombre de femmes exposées au risque juste avant l'instant t_i de la période 1 ;
- n_{2_i} le nombre de femmes exposées au risque juste avant l'instant t_i de la période 2 ;
- D_i le nombre de femmes décédées du cancer diagnostiqué et traité toutes périodes confondues en t_i ;
- D_{1_i} le nombre de femmes décédées du cancer diagnostiqué et traité de la période 1 en t_i ;
- D_{2_i} le nombre de femmes décédées du cancer diagnostiqué et traité de la période 2 en t_i ;
- E_{1_i} le nombre estimé en t_i de femmes décédées du cancer diagnostiqué et traité de la période 1 ;
- E_{2_i} le nombre estimé en t_i de femmes décédées du cancer diagnostiqué et traité de la période 2.

Sous l'hypothèse nulle, la probabilité estimée de décéder du cancer diagnostiqué et traité à chaque t_i est identique pour chaque période :

$$p_i = \frac{D_i}{n_{1_i} + n_{2_i}}$$

En ce sens, pour que l'hypothèse nulle soit acceptée, le nombre estimé de décès des suites du cancer diagnostiqué et traité pour chaque période à chaque instant t_i doit être égale à :

$$E_{1_i} = p_i \times n_{1_i} \text{ et } E_{2_i} = p_i \times n_{2_i}$$

Afin de tester l'hypothèse nulle, ce test repose sur la construction d'un indicateur d'écart qui est construit sur les différences asymptotiquement gaussiennes entre les taux de sortie théoriques et les taux de sortie observés pour chaque période. Ces statistiques de test sont :

$$Z_1 = \frac{[\sum_i E_{1_i} - D_{1_i}]^2}{\sum_i \text{Var}(D_{1_i})^2} \text{ et } Z_2 = \frac{[\sum_i E_{2_i} - D_{2_i}]^2}{\sum_i \text{Var}(D_{2_i})^2}$$

Comme l'indiquent PLANCHET et THÉRON (2006), la statistique de ce test peut être approchée par :

$$Z = \frac{[\sum_i E_{1_i} - D_{1_i}]^2}{E_1} + \frac{[\sum_i E_{2_i} - D_{2_i}]^2}{E_2}$$

Elle suit asymptotiquement une loi du χ^2 à 1 degré de liberté.

Enfin, il convient de se référer à la p-valeur pour conclure ce test avec un risque d'erreur toléré à 5%.

Application

Le test du log-rank est appliqué pour la base des femmes puis celles des hommes. Les résultats sont indiqués dans le tableau 2.14.

TABLE 2.14 – Résultats des tests du log-rank sur la base des femmes et celle des hommes

Sexe	χ^2	Degrés de liberté	P-valeur
Femme	5008	1	< 2.2e-16
Homme	1063	1	< 2.2e-16

La p-valeur pour chaque test est inférieure à 5%. Ces résultats impliquent que l'hypothèse nulle est rejetée, soit qu'il existe une différence significative des taux d'extra-mortalité entre la période 2000-2008 et 2009-2019. Ceci s'explique notamment par les progrès thérapeutiques réalisés sur les vingt dernières années. Par conséquent, il est primordial d'étudier les taux d'extra-mortalité sur la période

la plus récente des diagnostics (2009-2019), avec un historique suffisant sur dix ans, afin de quantifier au plus juste le risque d'extra-mortalité du cancer diagnostiqué et traité.

Par la suite, les bornes minimales et maximales des lois d'extra-mortalité par genre, calculées sur les cancers diagnostiqués entre 2009 et 2019, doivent être précisées.

2.8.3 Âges retenus

Le risque d'extra-mortalité des suites du cancer diagnostiqué et traité est mesuré selon le sexe et uniquement sur la période de diagnostic de 2009 à 2019. Il est désormais primordial de définir la structure des tables.

Pour rappel, le droit à l'oubli a longtemps été régi par l'âge au diagnostic du cancer de l'assuré qui souhaitait appliquer son droit :

- pour les emprunteurs dont l'âge à la date du diagnostic de la pathologie cancéreuse était inférieur à 21 ans, une ancienneté en rémission minimale de cinq ans était demandée ;
- pour les emprunteurs dont l'âge à la date du diagnostic de la pathologie cancéreuse était supérieur à 21 ans, une ancienneté en rémission minimale de dix ans était demandée.

Depuis le 2 mars 2022, le législateur a uniformisé le délai d'application du droit à l'oubli en le réduisant à cinq années passées en rémission, pour tous les assurés quelque soit leur âge au diagnostic du cancer. Ainsi, le nouveau risque d'extra-mortalité du cancer qui entre dans les portefeuilles des assureurs concerne uniquement les patients dont l'âge au diagnostic dépasse 21 ans. Il est supposé que les assureurs provisionnent déjà le risque d'extra-mortalité du cancer après seulement cinq ans en rémission pour les assurés dont l'âge au diagnostic était inférieur à 21 ans, étant donné le cadre de loi précédent. C'est pourquoi, seules les lignes dont l'âge au diagnostic est supérieur ou égal à 21 ans sont conservées. Ainsi, la base SEER retraitée comporte 2 266 655 lignes et l'âge à la fin du traitement thérapeutique minimum est de 21 ans.

De plus, pour appliquer son droit à l'oubli dans le cadre de l'assurance emprunteur, un assuré doit avoir moins de 71 ans à la dernière échéance de son assurance de prêt. Ainsi, il convient de fermer la table d'extra-mortalité des suites du cancer diagnostiqué et traité lorsque la somme de l'ancienneté en rémission et de l'âge à la fin du traitement thérapeutique est égale à 71 ans. Avant la loi Lemoine, la table aurait été ainsi fermée à 61 ans. Avec la réduction du délai d'application à cinq ans, il a été décidé de clôturer la loi à 66 ans.

Enfin, puisque le droit à l'oubli pour les assurés de plus de 21 ans est réduit de dix à cinq ans, l'ancienneté en rémission est étudiée sur dix années. En effet, il est supposé que les assureurs considèrent déjà ce risque pour les anciennetés supérieures à dix années, tel que c'était le cas avant le 2 mars 2022.

Ce deuxième chapitre définit clairement les données SEER utilisées et les retraitements nécessaires au cadre de cette étude. De plus, des tendances d'extra-mortalité ont été relevées selon le genre et selon l'année du diagnostic. En ce sens, il convient désormais de construire les tables par genre sur les cancers diagnostiqués entre le 01/01/2009 et le 31/12/2019.

Chapitre 3

Construction des lois d'extra-mortalité due au cancer après traitement

A partir des retraitements effectués sur les données, ces dernières sont exploitables pour étudier la probabilité de décéder d'un cancer non métastasé, diagnostiqué et traité en phase de rémission. Toutefois, la population étant restreinte à des patients atteints d'une pathologie cancéreuse en rémission et la cause de sortie exacte étant connue, il est primordial de préciser que les méthodes classiques d'analyse de mortalité peuvent être appliquées. Ainsi, le terme « extra-mortalité » est utilisé tout au long de ces travaux pour nommer le phénomène étudié mais il ne réfère en aucun cas aux notions théoriques d'analyse de surmortalité.

L'objectif de ce troisième chapitre est de présenter les notions fondamentales nécessaires à la construction d'une table d'extra-mortalité à double entrée. Puis, il convient de définir les estimateurs des taux bruts choisis, les méthodes d'ajustement et de lissage ainsi que les tests de validation à effectuer. Enfin, ces méthodes sont appliquées sur les données SEER retraitées.

3.1 Notions préliminaires pour l'analyse de survie

La table d'extra-mortalité construite dans cette étude est une table à deux dimensions. La première dimension correspond à l'âge d'entrée en rémission ($x \in [21; 66]$), également nommée âge de fin de traitement thérapeutique. Puis, la seconde dimension représente l'ancienneté en année ($t \in [0; 10]$) dans la phase de rémission. Ainsi, la théorie pour l'analyse de survie, généralement unidimensionnelle, est adaptée au cadre bidimensionnel étudié.

Cette partie est l'occasion de définir, dans un premier temps, l'ensemble des notations utilisées dans ce travail. Par la suite, les notions fondamentales nécessaires à la compréhension des outils utilisés ainsi que le problème des données incomplètes en analyse de survie sont rappelés.

3.1.1 Notations

Tout au long de ces travaux, les notations principales pour le cadre bidimensionnel de l'analyse de survie sont les suivantes :

- $q_{x,t}$ est la probabilité de décéder des suites du cancer diagnostiqué et traité pour un malade du cancer en rémission depuis l'âge x et dans l'état depuis t années ;

- $\hat{q}_{x,t}$ est le taux brut correspondant à $q_{x,t}$;
- $q_{x,t}^*$ est la valeur lissée de $\hat{q}_{x,t}$;
- $n_{x,t}$ est le nombre d'individus d'âge d'entrée en rémission x vivants au début de l'ancienneté en rémission t ;
- $D_{x,t}$ est la variable aléatoire représentant le nombre de décès observés pour l'âge d'entrée en rémission x sur l'intervalle de temps en rémission $]t, t + 1]$;
- $\hat{D}_{x,t}$ est la valeur estimée brute de $D_{x,t}$;
- $D_{x,t}^*$ est la valeur lissée de $\hat{D}_{x,t}$;
- $d_{x,t}$ est la réalisation de $D_{x,t}$;
- X_i^x est la variable aléatoire de durée de vie de l'individu i entré en rémission à l'âge x ;
- T_i^x est la variable aléatoire, à valeurs positives, de durée de survie du cancer diagnostiqué et traité de l'individu i pour un âge d'entrée en rémission x . Elle est à différencier de $T_{(j)}^x$ qui est un sous intervalle de temps (j) ;
- S^x est la fonction de survie pour l'âge de fin de traitement thérapeutique x ;
- F^x est la fonction de répartition pour l'âge de fin de traitement thérapeutique x ;
- f^x est la densité pour l'âge de fin de traitement thérapeutique x ;
- h^x est la force de mortalité, également nommée fonction de hasard, pour l'âge de fin de traitement thérapeutique x ;
- H^x est la fonction de hasard cumulé pour l'âge de fin de traitement thérapeutique x ;
- $E_{x,t}$ est l'exposition au risque pour l'âge d'entrée en rémission x et l'ancienneté en rémission t ;
- C_i^x est la variable aléatoire positive modélisant la censure pour l'individu i d'âge de fin de traitement thérapeutique x ;
- K_i^x est l'indicateur pour l'individu i d'âge de fin de traitement thérapeutique x qui permet de savoir s'il est censuré ou non ;
- $c_{x,t}$ est le nombre d'individus entrés en rémissions à l'âge x qui sont censurés sur l'intervalle de temps en rémission $]t - 1; t]$.

Ces notations sont détaillées selon l'âge à la fin du traitement thérapeutique et l'ancienneté. Elles sont équivalentes pour la population masculine et pour la population féminine, c'est pourquoi elles ne sont pas distinguées par sexe. La suite des notions théoriques présentées correspond à une population genrée et s'étend à l'identique pour la seconde modalité.

3.1.2 Notions fondamentales en analyse de survie

L'objectif de cette étude porte sur l'analyse de survie d'un patient en rémission d'un cancer diagnostiqué et traité. Ce temps de survie en rémission, noté T_i^x , est étudié pour chaque âge de fin de traitement thérapeutique x comme l'indique son exposant. La loi de cette variable se définit à l'aide des fonctions suivantes :

- la fonction de répartition : $F^x(t) = P(T^x \leq t)$;
- la fonction de survie : $S^x(t) = 1 - F^x(t) = P(T^x > t)$. Elle est continue et décroissante avec $S^x(0) = 1$ (sous l'hypothèse que $P(T^x = 0) = 0$ et $\lim_{t \rightarrow \infty} S^x(t) = 0$) ;
- la densité : $f^x(t) = \frac{d}{dt}F^x(t) = -\frac{d}{dt}S^x(t) = \lim_{h \rightarrow 0} \frac{P(t \leq T^x \leq t+h)}{h}$;
- la fonction de hasard, qui représente la mortalité instantanée : $h(t) = \frac{f^x(t)}{S^x(t)} = -\frac{S^{x'}(t)}{S^x(t)} = -\frac{d}{dt} \ln S^x(t)$;
- la fonction de hasard cumulée : $H^x(t) = \int_t^0 h^x(s) ds$ qui permet d'exprimer $S^x(t) = \exp(-H^x(t))$.

A partir de ces caractéristiques, la loi de T^x est entièrement caractérisée. La suite de la théorie des modèles de durée présentée se base sur ces définitions de la variable de survie en rémission T^x .

3.1.3 Les données incomplètes

Une particularité importante à prendre en considération pour la distribution de T^x est le fait que les données soient incomplètes. Cette caractéristique peut résulter :

- d'un problème de troncature des données, c'est-à-dire qu'un individu n'est observable que conditionnellement à la réalisation d'un autre événement ;
- d'un problème de censure des données puisque les décès des suites du cancer diagnostiqué et traité de tous les individus ne sont pas systématiquement observés.

Il convient ainsi de s'interroger sur la présence de ces phénomènes dans le cadre des données étudiées.

Troncature

Dans une étude de durée, il est d'usage d'avoir des données tronquées. Ce phénomène implique qu'un individu est observé uniquement si un événement aléatoire a lieu.

Toutefois, dans cette étude, les malades atteints d'un cancer entrent dans la base SEER dès lors qu'ils sont diagnostiqués et traités. En ce sens, les données ne sont pas tronquées.

Bien que les données étudiées ne comportent pas de troncature, elles peuvent être censurées.

Censure

Il existe deux catégories de censure :

- la censure à gauche : la réalisation de la variable d'intérêt (ici le décès des suites du cancer diagnostiqué et traité) a lieu avant le début de la période d'observation ;
- la censure à droite : la réalisation de la variable d'intérêt n'a pas eu lieu au cours de la période d'observation.

Toutefois, dans cette étude, les patients entrent dans la base de données uniquement une fois le cancer diagnostiqué et traité et ils sont suivis jusqu'à leur décès. Ainsi, seule la censure à droite est constatée. Elle peut être la conséquence d'un des deux phénomènes suivants :

- à la fin de la période d'observation, au 31/12/2019, le patient traité de son cancer est toujours vivant. Il s'agit d'une censure fixe ;
- le patient sort de la période d'observation suite à un décès d'une autre cause que le cancer diagnostiqué et traité avant le 31/12/2019. Il s'agit d'une censure aléatoire.

Il convient de préciser que la censure fixe est un cas particulier de la censure aléatoire. C'est pourquoi la suite de la définition se réfère au cas aléatoire.

Dans ce contexte, le fait connu est que la durée de survie du patient entré en rémission à l'âge x (X_i^x) est supérieure à un certain seuil aléatoire C_i^x . Pour chaque individu, au lieu d'observer directement la durée de vie individuelle X_i^x , c'est $T_i^x = \min(X_i^x, C_i^x)$ et $K_i^x = 1_{X_i^x \leq C_i^x}$ qui sont observées. Les notions fondamentales ayant été définies et le problème des données incomplètes étant détecté, l'estimation des taux bruts peut être réalisée en prenant en compte la censure des données.

3.2 Estimation des taux bruts

Les estimateurs des taux bruts de mortalité se distinguent en deux types :

- les estimateurs paramétriques qui reposent sur une hypothèse a priori de la fonction de survie ;
- les estimateurs non-paramétriques qui, au contraire, ne nécessitent pas d'hypothèse a priori quant à la forme de la loi de survie permettant ainsi une estimation à partir des données directement.

Dans ces travaux, il est décidé d'estimer les taux bruts d'extra-mortalité d'expérience à l'aide de l'estimateur de Hoem (paramétrique) et celui de Kaplan-Meier (non paramétrique) dans un cadre bidimensionnel (âge de fin de traitement thérapeutique et ancienneté en rémission). Cette partie permet de présenter la théorie nécessaire à la compréhension de la construction de chacun.

3.2.1 Estimateur des moments de Hoem

L'estimateur des moments de Hoem est un estimateur paramétrique des taux bruts. Comme l'indique PLANCHET et THÉRON (2006), il s'agit d'une généralisation de l'estimateur binomial en prenant en considération la censure et la troncature des données. Il se calcule à partir du nombre de décès et de l'exposition au risque.

Construction

Pour un âge d'entrée en rémission x fixé, cet estimateur considère qu'un patient i est exposé au risque d'extra-mortalité sur l'intervalle de temps en ancienneté $[t, t + 1[$ seulement entre deux dates α_i et β_i .

Sur cet intervalle de temps en rémission, chaque individu d'âge en rémission x et d'ancienneté dans cet état t a trois issues possibles : décéder du cancer diagnostiqué et traité, décéder d'une autre cause ou vivre. Toutefois, il convient de s'intéresser uniquement aux décès des suites du cancer diagnostiqué et traité pour cette étude d'extra-mortalité. Les sorties "décéder d'une autre cause" représentent des données censurées. Par conséquent, les seules issues considérées ici sont : décéder du cancer diagnostiqué et traité ou vivre. Ce phénomène est modalisable par la variable aléatoire X_i^x qui prend la valeur 1 si le patient entré en rémission à l'âge x décède des suites de son cancer diagnostiqué et traité et 0 sinon.

En considérant l'hypothèse que la probabilité de décéder du cancer diagnostiqué et traité est linéaire dans le temps en rémission ($\beta_i - \alpha_i q_{x,t+\alpha_i} \approx (\beta_i - \alpha_i) q_{x,t}$), X_i^x suit alors une loi de Bernoulli de paramètre $\beta_i - \alpha_i q_{x,t+\alpha_i}$.

Puis, le nombre de décès des suites du cancer diagnostiqué et traité entre les anciennetés en rémission t et $t+1$ pour l'âge d'entrée en rémission x fixé est un processus de comptage noté $D_{x,t} = \sum_{k=1}^{n_{x,t}} X_k^x$. Avec l'hypothèse d'indépendance des X_i^x , il résulte que $D_{x,t}$ suit une loi Binomiale $B(n_{x,t}, \beta_i - \alpha_i q_{x,t+\alpha_i})$.

Finalement,

$$\mathbb{E}(D_{x,t}) = \mathbb{E}\left(\sum_{k=1}^{n_{x,t}} X_k^x\right) = \sum_{k=1}^{n_{x,t}} \mathbb{E}(X_k^x) = \sum_{k=1}^{n_{x,t}} \beta_i - \alpha_i q_{x,t+\alpha_i} \approx \sum_{k=1}^{n_{x,t}} (\beta_i - \alpha_i) q_{x,t}$$

Ainsi,

$$q_{x,t} = \frac{\mathbb{E}(D_{x,t})}{\sum_{k=1}^{n_{x,t}} (\beta_i - \alpha_i)}$$

Soit,

$$\hat{q}_{x,t} = \frac{d_{x,t}}{\sum_{k=1}^{n_{x,t}} (\beta_i - \alpha_i)}$$

Propriétés

L'estimateur des moments de Hoem dispose des propriétés suivantes :

- il est sans biais, $\mathbb{E}(\hat{q}_{x,t}) = q_{x,t}$;
- il est convergent, $\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(|\hat{q}_{x,t} - q_{x,t}|) = 0$;
- si les hypothèses considérées sont proches de la loi d'extra-mortalité, l'estimateur des moments de Hoem est le plus adapté.

3.2.2 Estimateur de Kaplan-Meier

L'estimateur de Kaplan-Meier est un estimateur non-paramétrique, ne reposant ainsi sur aucune hypothèse a priori de la loi d'extra-mortalité. Il est l'estimation du maximum de vraisemblance non-paramétrique de la fonction de survie. Introduit par KAPLAN et MEIER (1958), il consiste à calculer la probabilité de survie des individus en considérant comme pas de sortie dans ces travaux le décès des suites du cancer diagnostiqué et traité. Il repose sur le comptage du nombre de survivants, du nombre de décès des suites du cancer diagnostiqué et traité et du nombre de patients censurés.

Construction

Cet estimateur se base sur le fait que pour un individu, d'âge d'entrée en rémission x , sa probabilité de survivre au-delà de l'ancienneté en rémission t correspond à survivre au-delà de l'ancienneté en rémission $t - 1$ et ne pas décéder du cancer diagnostiqué et traité sur l'intervalle de temps en rémission $]t - 1; t]$. En ce sens, la fonction de survie pour l'âge d'entrée en rémission x peut également s'écrire :

$$S^x(t) = P(T^x > t | T^x > t - 1) \times S^x(t - 1) = p_t \times S^x(t - 1)$$

L'opération peut être répétée en remplaçant systématiquement la fonction de survie dans le produit par un instant de conditionnement inférieur. Le produit limite est ainsi obtenu :

$$S^x(t) = \prod_{k=1}^t p_k$$

A partir de $p_{x,t} = 1 - q_{x,t}$ et l'estimateur naturel de mortalité $\hat{q}_{x,t} = \frac{d_{x,t}}{n_{x,t}}$, l'estimateur de Kaplan-Meier de la fonction de survie s'écrit :

$$\hat{S}^x(t) = \prod_{k \leq t} \left(1 - \frac{d_{x,k}}{n_{x,k}}\right)$$

Avec $n_{x,k} = n_{x,k-1} - d_{x,k-1} - c_{x,k-1}$, où la censure est prise en compte.

Propriétés

L'estimateur de Kaplan-Meier a les propriétés suivantes :

- il est biaisé positivement, $\mathbb{E}(\hat{S}^x(t)) \geq S^x(t)$;
- il est convergent, $\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(|\hat{S}^x(t) - S^x(t)|) = 0$;
- il ne nécessite pas d'hypothèse a priori sur la fonction de survie.

3.3 Ajustement et lissage

Les taux bruts d'extra-mortalité calculés comportent des irrégularités, en partie impliquées par les conditions de l'expérience. Ces aspérités ne représentent pas la réalité des tendances d'extra-mortalité puisque, de manière générale, la probabilité de décéder du cancer diagnostiqué et traité diminue au cours du temps passé en rémission. Il convient donc d'ajuster les taux bruts calculés pour estimer une loi régulière et fidèle au risque étudié en deux dimensions.

En analyse de survie, il existe un grand nombre de méthodes pour lisser ou ajuster des taux bruts d'extra-mortalité. Ces méthodes sont classées en quatre catégories :

- les modèles paramétriques, qui se basent sur des hypothèses a priori de la distribution de mortalité connue sur d'autres populations afin d'ajuster les taux bruts calculés ;
- les lissages paramétriques, qui ne reposent sur aucune hypothèse a priori de la distribution des taux d'extra-mortalité à la différence de la catégorie précédente. Ils consistent à utiliser une fonction dépendant d'un nombre fini de paramètres et reflétant correctement les taux bruts calculés selon les variables de durée ;
- les lissages non-paramétriques, tout comme les lissages paramétriques, n'utilisent pas des hypothèses a priori sur la forme des taux d'extra-mortalité. Ils utilisent une fonction décrite par un nombre de paramètres augmentant avec le nombre d'observations pour lisser les taux de mortalité ;
- les modèles relationnels, dont le principe est identique aux modèles paramétriques avec des hypothèses a priori sur la distribution de mortalité. Toutefois, les taux sont exprimés en fonction de taux de référence et non de la variable de durée comme pour les modèles paramétriques ;
- les méthodes non-paramétriques qui permettent de lisser la distribution obtenue à partir des observations sans considérer d'hypothèse a priori.

Dans le cadre de ces travaux, seules des méthodes de lissage paramétrique et non-paramétrique sont utilisées, respectivement la méthode des splines cubiques et la méthode de Whittaker-Henderson. Toutefois, il convient de rappeler que les taux bruts à lisser sont doublement indicés par l'âge à la fin de traitement thérapeutique et l'ancienneté en rémission. Ainsi, la théorie de chacune de ces méthodes est présentée brièvement par la suite en les adaptant au cadre bidimensionnel.

3.3.1 Méthode des P-splines

En analyse fonctionnelle, une approximation de fonction est souvent réalisée à l'aide d'un polynôme qui est un outil simple permettant de modéliser toute fonction continue. Toutefois, les taux bruts d'extra-mortalité comportent des parties plates qui sont difficilement exploitables par une unique fonction polynomiale ayant tendance à osciller. A partir de ce constat et dans l'objectif de lisser des taux bruts en deux dimensions, il a été proposé par EILERS et MARX (1996), DURBÁN et al. (2002), CURRIE et al. (2003) puis PLANCHET et WINTER (2010) d'utiliser des splines de dimension 2 pour le lissage des lois de durées bidimensionnelles.

Par définition, une spline correspond à une fonction définie par morceaux par des polynômes. Dans le

contexte de l'analyse de survie avec des ruptures dans l'évolution des taux bruts, une spline permet donc d'ajuster un polynôme par sous-intervalle de la plage de la fonction à ajuster initialement, puis de les raccorder aux points de jonction. En ce sens, avec un découpage optimal, une spline permet un ajustement de fonctions plus simples sur chaque sous-intervalle que celle qu'il aurait fallu ajuster au global.

Parmi les méthodes de lissage à l'aide des splines, EILERS et MARX (1996) proposent la méthode des P-splines qui est elle-même construite à partir des deux notions suivantes :

- le lissage B-splines ;
- la vraisemblance pénalisée qui permet de corriger les inconvénients des B-splines.

Elle est par la suite utilisée dans les travaux de DURBÁN et al. (2002) et CURRIE et al. (2003).

La construction de la méthode de lissage suivante est une adaptation des travaux réalisés par CURRIE et al. (2003) où les P-splines sont appliqués pour l'ajustement des taux de mortalité en deux dimensions.

Comme le cadre d'étude de CURRIE et al. (2003), les données sont représentées par deux matrices D et E , respectivement la matrice du nombre de décès des suites du cancer diagnostiqué et traité et la matrice des expositions au risque. Leurs lignes sont indexées par les âges à la fin du traitement thérapeutique et les colonnes par les anciennetés en rémission. En ce sens, la matrice du logarithme des taux de hasard se note $R = \log\left(\frac{D}{E}\right) = \log(D) - \log(E)$.

Par la suite, pour lisser R , DURBÁN et al. (2002) expliquent qu'un modèle linéaire généralisé pénalisé peut être utilisé pour D .

La matrice de régression B est produite par le produit de Kronecker,

$$B = B_t \otimes B_x$$

où :

- $B_x = B(x)$, la matrice de régression B-splines basée sur la variable explicative âge à la fin du traitement thérapeutique. Elle est de taille $46 \times c_x$ avec c_x correspondant au nombre de noeuds, c'est-à-dire le nombre de sous-intervalles fixés découpant la plage du B-spline ;
- $B_t = B(t)$, la matrice de régression B-splines basée sur la variable explicative ancienneté en rémission après traitement thérapeutique. Elle est de taille $11 \times c_t$ avec c_t correspondant au nombre de noeuds, c'est-à-dire le nombre de sous-intervalles fixés découpant la plage du B-spline.

Il convient de fixer le degré des polynômes utilisés dans les B-splines. Les plus couramment paramétrés en analyse de survie sont ceux de degré 2 et ceux de degré 3.

Alors, pour être en adéquation avec la matrice de régression, la matrice du nombre de décès des suites du cancer diagnostiqué et traité est vectorisée, c'est-à-dire transformée en vecteur colonne, et l'on a : $d = \text{Vec}(D)$. La vectorisation est également appliquée pour la matrice des expositions qui devient $e = \text{Vec}(E)$.

Ainsi, le modèle s'écrit sous forme d'un modèle linéaire généralisé :

$$d = \mu + \epsilon$$

où $\log(\mu) = \log(e) + Ba$, et $\log(e)$ est le décalage habituel dans un modèle log-linéaire pour les données de mortalité.

Ce modèle a tendance à être sur-paramétré. C'est pourquoi, une pénalité sur a est appliquée. Une matrice de pénalité appropriée est alors,

$$P = \lambda_x I_{c_t} \otimes {}^t M_x M_x + \lambda_t {}^t M_t M_t \otimes I_{c_x}$$

où :

- I_{c_t} est la matrice d'identité de taille c_t ;
- M_x est la matrice différence de dimensions $(c_x - p_x) \times c_x$ où p_x est l'ordre de la pénalité sur les âges à la fin de traitement thérapeutique ;
- I_{c_x} est la matrice d'identité de taille c_x ;
- M_t est la matrice différence de dimensions $(c_t - p_t) \times c_t$ où p_t est l'ordre de la pénalité sur les anciennetés en rémission ;
- λ_x et λ_t correspondent aux paramètres de lissage à déterminer de manière algorithmique selon un critère choisi tels que l'Akaike Information Criterion (AIC) ou le Bayesian Information Criterion (BIC).

Finalement, en déterminant les paramètres de lissage λ_x et λ_t selon le critère choisi (AIC ou BIC), l'ordre des matrices de différence et le degré des splines, les pénalités appliquées dans l'ajustement du modèle linéaire généralisé sont obtenues. Les taux lissés d'extra-mortalité sont alors calculés.

3.3.2 Méthode de Whittaker-Henderson

La méthode de Whittaker-Henderson est une des méthodes de lissage non paramétrique introduite au cours du XX^{ème} siècle par WHITTAKER (1922) et reprise par HENDERSON (1924). Elle consiste à ajuster les taux d'extra-mortalité en respectant la fidélité aux taux bruts et en ayant une forme finale de fonction régulière.

Cette méthode consiste à déterminer des valeurs ajustées des taux d'extra-mortalité qui minimisent la somme du critère de fidélité et du critère de régularité. Cette méthode est reprise par PLANCHET et THÉRON (2006) avec deux dimensions pour l'estimation de la loi de maintien en arrêt de travail en fonction de l'âge à l'entrée et de l'ancienneté de l'arrêt.

Dans le cadre bidimensionnel de cette étude, le critère de fidélité s'écrit :

$$F = \sum_{x=21}^{66} \sum_{t=0}^{10} w_{x,t} (q_{x,t} - \hat{q}_{x,t})^2$$

Puis, le critère de régularité est différencié selon la régularité verticale et la régularité horizontale. Toutes deux comportent l'opérateur différence $\Delta^n u(x)$ qui se définit par la formule suivante :

$$\Delta^n u(x) = \sum_{j=0}^n C_n^j (-1)^{n-j} u(x+j)$$

Alors, le critère de régularité horizontale est :

$$S_h = \sum_{t=0}^{10} \sum_{x=21}^{66-z} (\Delta_h^z q_{x,t})^2$$

Puis, le critère de régularité verticale est :

$$S_v = \sum_{x=21}^{66} \sum_{t=0}^{10-z} (\Delta_v^z q_{x,t})^2$$

Les poids $w_{x,t}$ sont ici choisis comme étant égaux au rapport de l'effectif du point considéré (x, t) sur l'effectif total.

Il convient désormais de minimiser la fonction :

$$M = (1 - \alpha - \beta)F + \alpha S_v + \beta S_h$$

Finalement, les taux lissés sont égaux à :

$$q^* = (w^* + \alpha {}^t(K_z^v)K_z^v + \beta {}^t(K_y^h)K_y^h)^{-1} w^* u$$

où u , de taille $(46, 11)$, est égal à : $u_{(46)(x-1)+t} = \hat{q}_{x,t}$ et les matrices K_z^v et K_y^h sont telles que :

$$S_v = {}^t(K_z^v)K_z^v \text{ et } S_h = {}^t(K_y^h)K_y^h$$

Lors de l'application, il est nécessaire de fixer les valeurs pour z_v, z_h, α et β .

Une fois les méthodes de lissage appliquées sur les taux bruts calculés, une table d'extra-mortalité genrée doit être validée et sélectionnée. Cette dernière étape repose également sur des notions théoriques définies dans la partie suivante.

3.4 Sélection et validation de la table finale

Afin de sélectionner les tables d'extra-mortalité qui sont les plus conformes à l'extra-mortalité empirique, il est nécessaire de mettre en place différents tests d'adéquation. Les tests utilisés dans cette étude sont le test du χ^2 et le test des changements de signes, le premier étant un indicateur de la qualité du lissage et le deuxième de sa régularité. Leurs notions théoriques sont présentées dans cette partie pour les taux d'extra-mortalité lissés $q_{x,t}^*$ par rapport aux taux bruts sélectionnés.

3.4.1 Le test du χ^2 d'adéquation

L'objectif du test du χ^2 d'adéquation est de s'assurer que les taux d'extra-mortalité lissés sont proches des taux d'extra-mortalité bruts. Ainsi, il permet de mesurer la qualité d'ajustement du lissage et repose sur les hypothèses suivantes :

- H_0 : la distribution des taux lissés $q_{x,t}^*$ est en adéquation avec les taux bruts $\hat{q}_{x,t}$;
- H_1 : la distribution des taux lissés $q_{x,t}^*$ n'est pas en adéquation avec les taux bruts $\hat{q}_{x,t}$.

Il convient de remarquer que ce test est similaire au test d'indépendance du χ^2 présenté dans le chapitre précédent. Ainsi, il se base également sur un indicateur d'écart. PLANCHET et WINTER (2010) présentent une version *ad hoc* du test du χ^2 pour une estimation de loi de maintien en arrêt de travail en deux dimensions. Des regroupements sur les âges et sur les anciennetés peuvent être réalisés

pour obtenir des classes bidimensionnelles : L est le nombre de classes créées pour l'âge à la fin du traitement thérapeutique et K est le nombre de classes pour l'ancienneté en rémission. La statistique de test s'écrit alors :

$$W_{K,L} = \sum_{l=1}^L \sum_{k=1}^K \frac{(\hat{D}_{l,k} - D_{l,k}^*)^2}{D_{l,k}^*}$$

$W_{K,L}$ est asymptotiquement distribuée comme une variable χ_{K+L-2}^2 . L'hypothèse nulle est alors rejetée si le calcul de la statistique de test donne une valeur qui est supérieure à la valeur correspondant au quantile à 95% dans la loi du χ^2 à $(K + L - 2)$ degrés de liberté.

Par conséquent, ce premier test permet de conclure quant à la qualité de l'ajustement des taux lissés aux taux bruts. Dans une démarche scientifique claire, il convient de renforcer ce résultat en testant la régularité du lissage à l'aide du test des changements de signes.

3.4.2 Le test des changements de signe

Le test des changements de signes est un test non-paramétrique qui analyse la fréquence des changements de signes des différences entre les taux d'extra-mortalité lissés et les taux d'extra-mortalité bruts. Il repose sur le constat que si les taux sont correctement lissés alors il devrait y avoir autant de taux sur-lissés que de taux sous-lissés. En ce sens, il permet de mesurer la régularité du lissage effectué.

Soit $E_{x,t}$ la différence entre le taux d'extra-mortalité lissé et le taux d'extra-mortalité brut pour l'âge de fin de traitement thérapeutique x et l'ancienneté en rémission t . Les hypothèses considérées dans ce test sont :

- $H_0 : P(E_{x,t} > 0) = \frac{1}{2}, \forall x, t;$
- $H_1 : P(E_{x,t} > 0) \neq \frac{1}{2}, \forall x, t$

Comme l'expliquent PLANCHET et THÉRON (2006), dans un cadre bidimensionnel, la statistique de test pour ces travaux est :

$$E = \sum_{x=21}^{66} \sum_{t=0}^{10} 1_{E_{x,t} > 0}$$

Puisque E suit une loi binomiale de paramètres $(46 \times 11, \frac{1}{2})$, il est possible de faire une approximation gaussienne et de travailler avec la statistique suivante :

$$S = \frac{2E - (46 \times 11 - 1)}{\sqrt{46 \times 11 - 1}}$$

L'hypothèse nulle est alors rejetée si le calcul de la statistique de test S donne une valeur en absolu qui est supérieure à la valeur correspondant au quantile à 97.5% de la loi normale centrée réduite.

Pour conclure, le test du χ^2 d'adéquation et le test des changements de signes mesurent la qualité et la régularité de la méthode de lissage appliquée sur les taux d'extra-mortalité bruts. A partir de ces mesures et des objectifs de l'étude, les tables d'extra-mortalité des suites du cancer diagnostiqué et traité genrées selon l'âge à la fin du traitement thérapeutique et l'ancienneté en rémission peuvent être sélectionnées et utilisées pour le calcul de l'impact.

La suite du chapitre se concentre désormais sur l'application des notions théoriques susmentionnées sur les données SEER retraitées.

3.5 Application numérique sur les données SEER retraitées

A l'aide des notions théoriques définies précédemment, il convient de construire la table d'extramortalité genrée avec les cancers diagnostiqués entre le 01/01/2009 et le 31/12/2019. L'ensemble des résultats présentés sont obtenus à l'aide du logiciel R.

Au cours des travaux de retraitements et d'analyse des données, les hypothèses suivantes ont été prises pour cette étude :

- les malades touchés par un cancer métastasé sont exclus de l'étude car ils ont une probabilité significative de rechute après leur traitement thérapeutique ;
- les patients atteints d'un cancer non métastasé ne rechutent pas une fois entrés en rémission ;
- si un individu est touché par plusieurs cancers, la période de rémission court à compter de la fin de traitement du dernier cancer diagnostiqué ;
- l'âge de fin de traitement thérapeutique est estimé selon des durées moyennes de traitement en fonction du stade et du grade du cancer ;
- la situation épidémiologique des cancers aux Etats-Unis est similaire à celle de la France.

Certaines de ces hypothèses sont challengées à la fin des travaux dans le chapitre suivant.

3.5.1 Construction des taux bruts

Cette première partie permet d'exposer les taux bruts estimés par la méthode d'Hoem puis celle de Kaplan-Meier.

3.5.1.1 Résultats graphiques des taux bruts avec l'estimateur de Hoem

La méthode d'estimation de Hoem repose sur le calcul de l'exposition au risque de chaque patient et sur le calcul du nombre de décès des suites du cancer diagnostiqué et traité pour pour chaque classe d'âge à la fin du traitement thérapeutique et d'ancienneté en rémission, soit chaque (x, t) . Ainsi, il est intéressant d'analyser graphiquement les expositions et les nombres de décès pour chaque sexe avant de regarder les taux bruts.

Le graphique de gauche de la figure 3.1 montre que l'exposition des femmes se concentre sur les âges à la fin du traitement thérapeutique supérieurs à 40 ans, dont les anciennetés en rémission sont inférieures à 5 ans. En parallèle, le nombre de décès des suites du cancer diagnostiqué et traité est élevé sur les anciennetés en rémission inférieures à 5 ans et dont l'âge à la fin du traitement thérapeutique appartient à l'intervalle $[45;66]$. Ainsi, il convient de s'attendre a une forte volatilité des taux bruts en dehors des intervalles cités.

Quant à la base de données des hommes, la figure 3.2 permet de constater que l'exposition se centralise sur les anciennetés en rémission inférieures à 5 ans et dont l'âge à la fin de traitement thérapeutique est au sein de l'intervalle $[50;66]$. Un grand nombre de décès des hommes des suites du cancer diagnostiqué et traité est observable pour les âges à la fin du traitement thérapeutique compris entre 50 et 66 ans et dont l'ancienneté en rémission est inférieures à 3 ans. Par conséquent, pour les hommes également, une forte volatilité des taux bruts en dehors des intervalles cités est attendue.

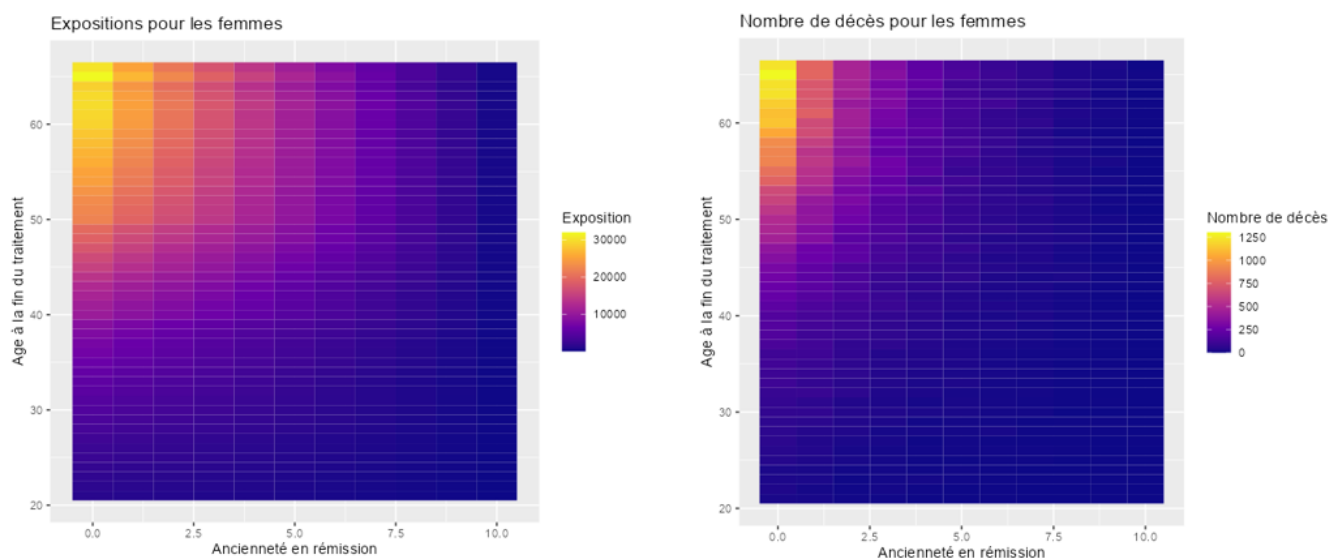


FIGURE 3.1 – Exposition et nombre de décès des suites du cancer pour les femmes

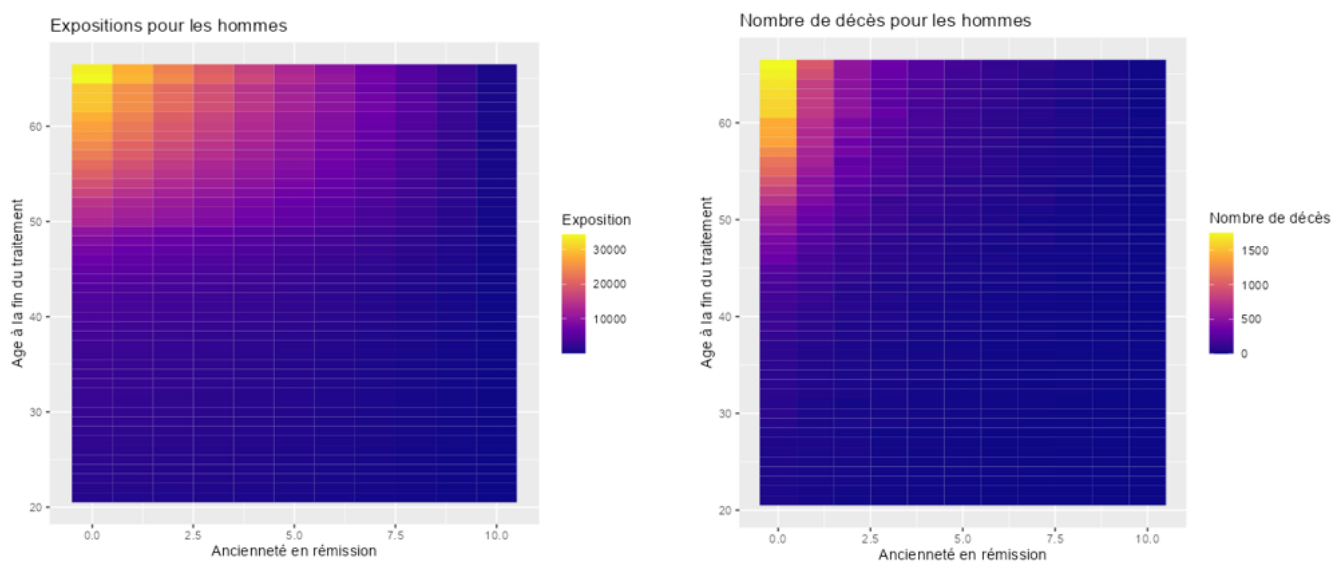


FIGURE 3.2 – Exposition et nombre de décès des suites du cancer pour les hommes

A partir des expositions au risque et du nombre de décès des suites du cancer diagnostiqué et traité calculés, les taux bruts par la méthode de Hoem sont estimés. L'allure des probabilités de décéder du cancer diagnostiqué et traité pour chaque genre sont présentées en figure 3.3.

La figure 3.3 montre que, de manière générale, les femmes ont une forte probabilité de décéder des suites du cancer diagnostiqué et traité dans les premières années du processus de rémission et notamment lorsque leur âge à la fin du traitement thérapeutique s'élève. Ce constat signifie que les taux d'extra-mortalité des suites du cancer diagnostiqué et traité augmentent selon l'âge à la fin du traitement thérapeutique puis diminuent au cours du temps passé en rémission. Cette tendance est également observée pour les taux d'extra-mortalité des hommes à droite de la figure 3.3. Toutefois, elle semble

être plus marquée avec une pente importante de la surface pour les premières années en rémission.

Quant à l'objectif de ces travaux sur l'évolution du droit à l'oubli, il est intéressant de noter que les taux d'extra-mortalité, pour tous les âges confondus, semblent être plafonnés à 2% pour les anciennetés en rémission comprises entre 5 et 10 ans pour les hommes et pour les femmes.

Enfin, comme indiqué dans les analyses graphiques des expositions et des nombres de décès, les taux sont volatiles pour les femmes et pour les hommes. Cette volatilité se remarque par des pics de valeur. Pour les femmes, cela s'observe notamment pour les âges inférieurs à 40 ans et pour les anciennetés en rémission comprises entre 8 et 10 ans. Ce phénomène reste mesuré puisque les pics sont peu fréquents et estompés. Quant aux hommes, cette volatilité est plus forte avec des pics en plus grand nombre et plus importants que pour les femmes.

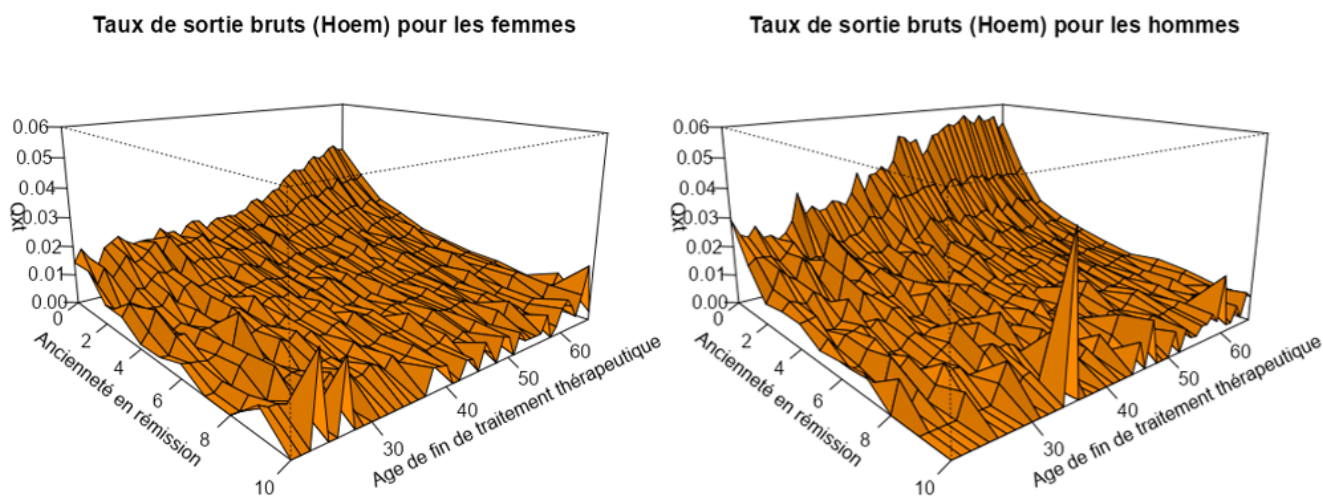


FIGURE 3.3 – Taux bruts d'extra-mortalité des suites du cancer diagnostiqué et traité (Hoem)

Par la suite, la méthode de Kaplan-Meier est appliquée sur les bases générées des données SEER retraitées afin de comparer les taux obtenus avec ceux calculés par l'estimateur des moments de Hoem.

3.5.1.2 Résultats graphiques des taux bruts avec l'estimateur de Kaplan-Meier

L'estimation des taux par la méthode de Kaplan-Meier est réalisée à l'aide du package « Survival » du logiciel R. Les fonctions de survie empiriques obtenues pour chaque sexe sont présentées en figure 3.4.

De manière générale, des remarques similaires à celles faites précédemment, pour les taux bruts estimés à l'aide de l'estimateur des moments de Hoem, peuvent être écrites. En effet, sur les deux figures ci-dessous, les taux d'extra-mortalité augmentent avec l'âge à la fin du traitement thérapeutique puis diminuent avec les années passées en rémission pour les deux sexes. Là encore, ce phénomène est plus marqué pour les hommes avec une forte pente dans les premières années en rémission.

Bien que les taux semblent globalement plus faibles avec Kaplan-Meier qu'avec l'estimateur des moments de Hoem, il semble que la probabilité de décéder du cancer diagnostiqué et traité est toujours plafonnée à 2% pour les anciennetés en rémission comprises entre 5 et 10 ans.

Une différence majeure, avec les taux bruts de la méthode des moments de Hoem, est la réduction du

phénomène de volatilité des taux avec la méthode de Kaplan-Meier pour les deux sexes. En effet, les pics des bords de table observés pour les femmes et pour les hommes dans les graphiques de Hoem sont nettement réduits dans cette nouvelle estimation.

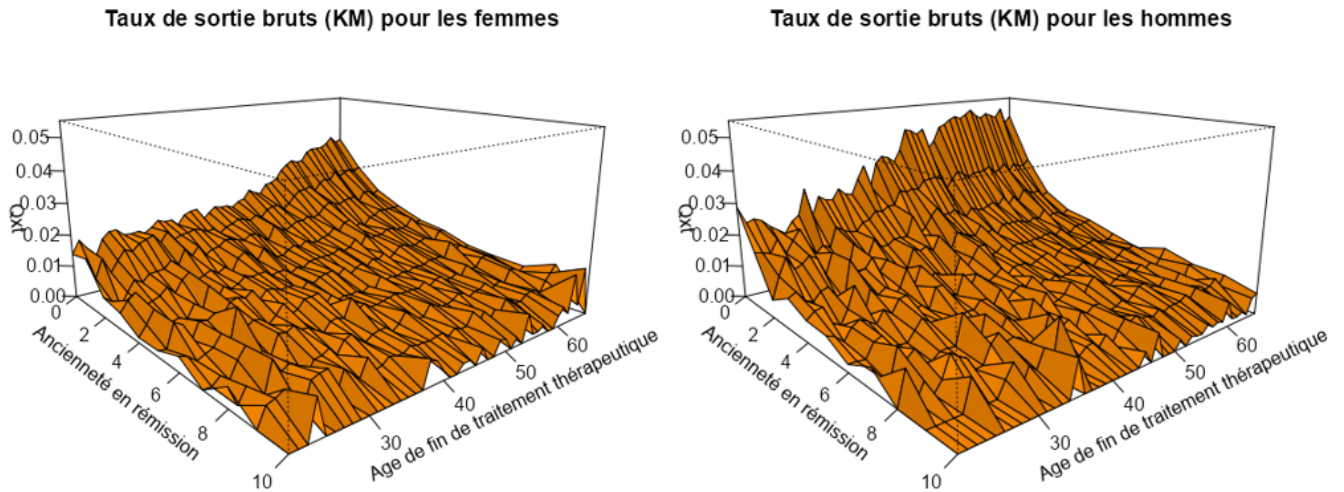


FIGURE 3.4 – Taux bruts d’extra-mortalité des suites du cancer diagnostiqué et traité (Kaplan-Meier)

Par conséquent, puisque les tendances d’extra-mortalité observées sur les taux bruts sont analogues entre l’estimateur de Hoem et celui de Kaplan-Meier, il convient de poursuivre ces travaux avec une seule méthode d’estimation des taux bruts. Toutes deux peuvent être utilisées dans le cadre de ces travaux. Cependant, les taux bruts déterminés à l’aide de la méthode des moments de Hoem ont l’avantage d’être plus prudents puisqu’ils considèrent l’exposition au risque des patients à la différence de ceux chiffrés par la méthode non-paramétrique de Kaplan-Meier. C’est pourquoi, l’étude est finalement réalisée avec les taux bruts estimés par la méthode des moments de Hoem qui sont ensuite ajustés.

3.5.2 Ajustement et lissage des taux bruts

Les taux bruts d’extra-mortalité obtenus à l’aide de la méthode des moments comporte des irrégularités causées par les fluctuations d’échantillonnage. Afin de les corriger, les méthodes des P-splines et de Whittaker-Henderson sont appliquées.

3.5.2.1 Application de la méthode des P-splines

Afin d’obtenir les taux lissés par la méthode des P-splines, il est nécessaire de calibrer les paramètres de la méthode :

- pour le degré des splines, puisqu’il est d’usage dans l’analyse de survie d’utiliser des polynômes de degré 3, ils sont également fixés à 3 ;
- le nombre total de noeuds est fixé à 100, soit $c_t = 5$ et $c_x = 20$;
- quant aux ordres des matrices de différence, ils sont fixés tels que $(p_x, p_t) = (2, 2)$. En effet, CURRIE et al. (2003) indiquant que ces ordres n’ont pas d’impact significatif sur le lissage mais uniquement sur l’extrapolation qui n’est pas utile ici, il convient de conserver le paramétrage usuel $(p_x, p_t) = (2, 2)$ extrait de la littérature ;

— enfin, λ_x et λ_t sont déterminés en optimisant le modèle selon la valeur du BIC.

La figure 3.5 montre la minimisation du BIC en fonction des combinaisons de valeurs de λ_x et λ_t . Les combinaisons suivantes sont testées :

- $\lambda_x = 10^i \forall i = 0, \dots, 10$;
- $\lambda_t = 10^i \forall i = 0, \dots, 10$

Ces combinaisons sont exprimées sur l'axe des abscisses avec la dimension effective. Pour les femmes, la dimension effective minimisant le critère du BIC est 15.8, correspondant à $(\lambda_x, \lambda_t) = (100, 1000)$. Quant aux hommes, il s'agit du point dont la dimension effective vaut 14.8, soit $(\lambda_x, \lambda_t) = (1000, 100)$.

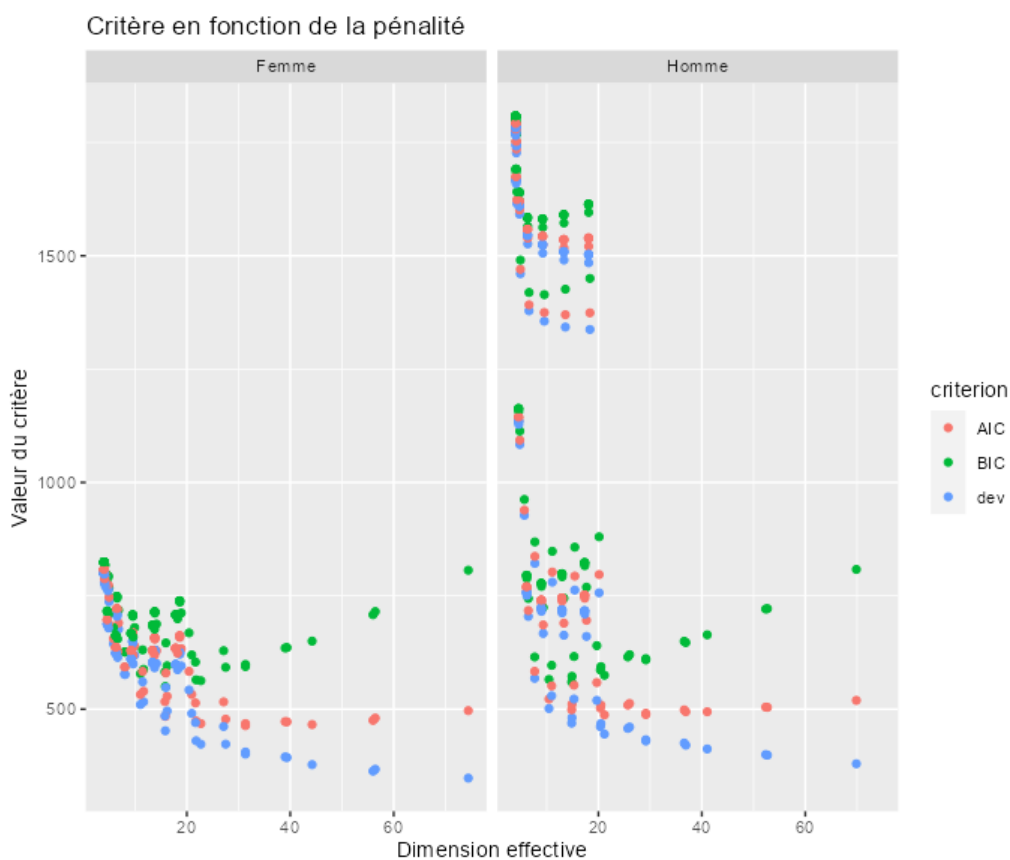


FIGURE 3.5 – Minimisation du BIC selon les variations de λ_x et λ_t

Finalement, avec $(\lambda_x, \lambda_t) = (100, 1000)$ pour les femmes les taux d'extra-mortalité lissés par la méthode P-splines sont observés à gauche en figure 3.6 et ceux pour les hommes avec $(\lambda_x, \lambda_t) = (1000, 100)$ à droite de la même figure. Ces deux surfaces ont des taux nuls puisqu'il s'agit de combinaisons (x, t) où le nombre de sorties des suites de cancer diagnostiqué et traité est nul. De manière générale, pour les deux sexes, la surface semble plus régulière que fidèle à l'extra-mortalité empirique.

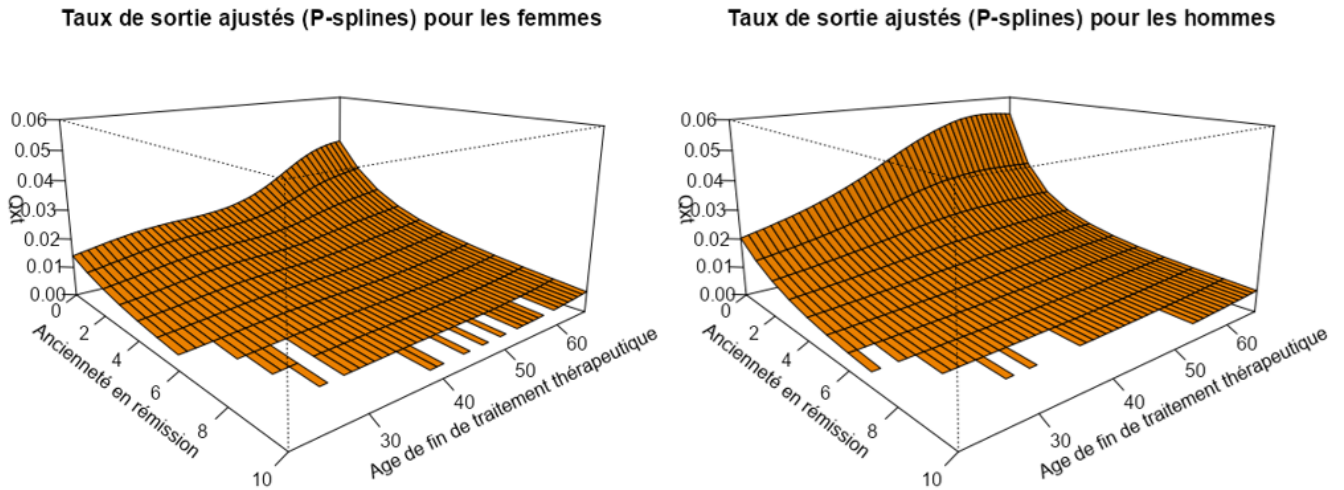


FIGURE 3.6 – Taux bruts d’extra-mortalité lissés par P-splines pour les femmes avec $(\lambda_x, \lambda_t) = (100, 1000)$ et pour les hommes avec $(\lambda_x, \lambda_t) = (1000, 100)$

Dans la suite de l’étude, il est intéressant de lisser les taux d’extra-mortalité bruts avec la méthode non-paramétrique de Whittaker-Henderson.

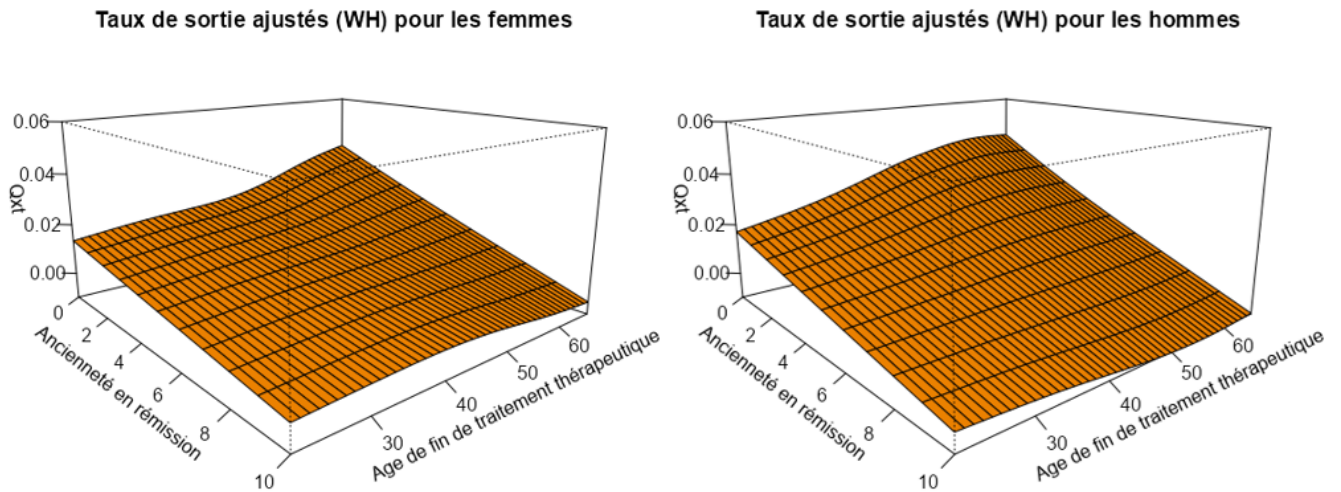
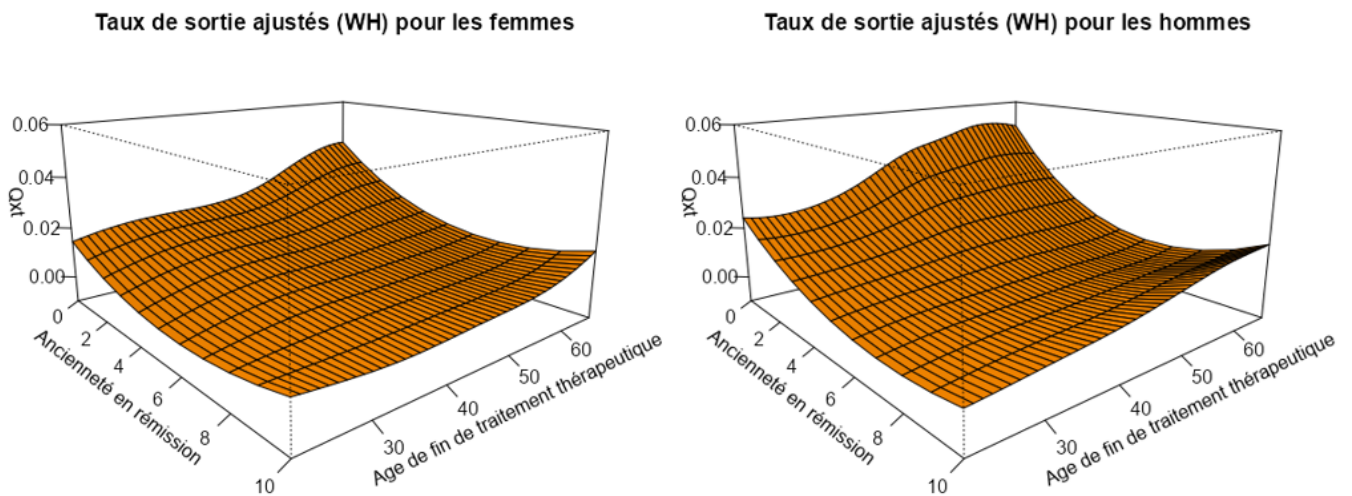
3.5.2.2 Application de la méthode de Whittaker-Henderson

Le lissage de Whittaker-Henderson nécessite de fixer les paramètres z_v , z_h , α et β pour obtenir les taux bruts d’extra-mortalité lissés. En ce sens, plusieurs combinaisons de paramètres ont été testées afin de définir la combinaison la plus adéquate selon l’analyse graphique de la surface obtenue. Ce processus de calibrage par analyse graphique est semblable à celui réalisé par WINTER (2005) dans ses travaux de lissage bidimensionnel.

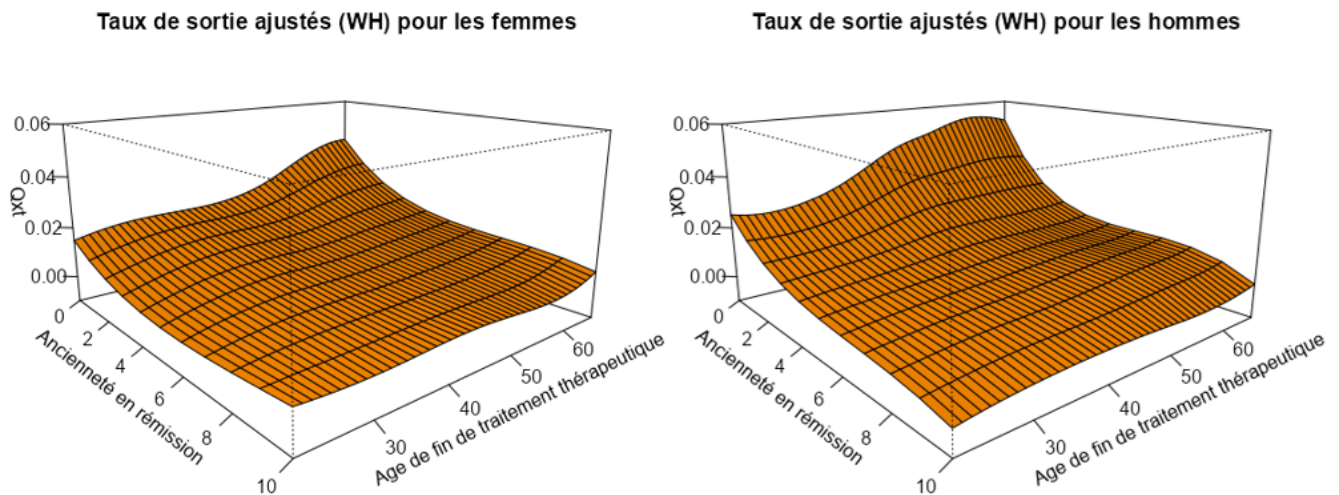
Cette définition repose sur les principes suivants :

- un compromis pour les valeurs α et β doit être réalisé puisqu’ils représentent le poids du critère de régularité, respectivement verticale et horizontale. Ainsi, si les valeurs sont élevées alors le lissage sera fortement régulier mais peu fidèle aux données, et inversement pour des valeurs faibles ;
- quant aux ordres de régularité verticale et horizontale, z_v et z_h , ils permettent de déterminer la forme de la surface influant ainsi également sur la régularité des taux lissés.

Dans un premier temps, pour α et β fixés à 1, il convient de faire varier les ordres de régularité verticale et horizontale. En comparant les figures 3.7 et 3.8, l’importance des valeurs de l’ordre vertical et de l’ordre horizontal est soulignée pour $(\alpha, \beta) = (1, 1)$. En effet, lorsque $(z_h, z_v) = (2, 2)$, la surface est nettement régulière effaçant ainsi les vagues d’extra-mortalité constatées dans les données brutes. Au contraire, les valeurs $(z_h, z_v) = (3, 3)$ permettent un ajustement plus fidèle aux taux bruts d’extra-mortalité en révélant certaines irrégularités empiriques lissées.

FIGURE 3.7 – Taux d’extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 1, 2, 2)$ FIGURE 3.8 – Taux d’extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 1, 3, 3)$

L’importance de ces travaux étant de capter les fluctuations des taux d’extra-mortalité selon l’ancienneté en rémission pour faire preuve de prudence face à la réduction du délai du droit à l’oubli, il est intéressant d’augmenter la valeur de l’ordre de régularité verticale à 4. Comme il est présenté en figure 3.9, le lissage obtenu pour $(z_h, z_v) = (3, 4)$ apporte davantage d’adéquation aux vagues repérées dans les taux bruts calculés à l’aide de la méthode des moments de Hoem. Par conséquent, il est décidé de fixer $(z_h, z_v) = (3, 4)$ pour les deux sexes pour la suite de l’étude.

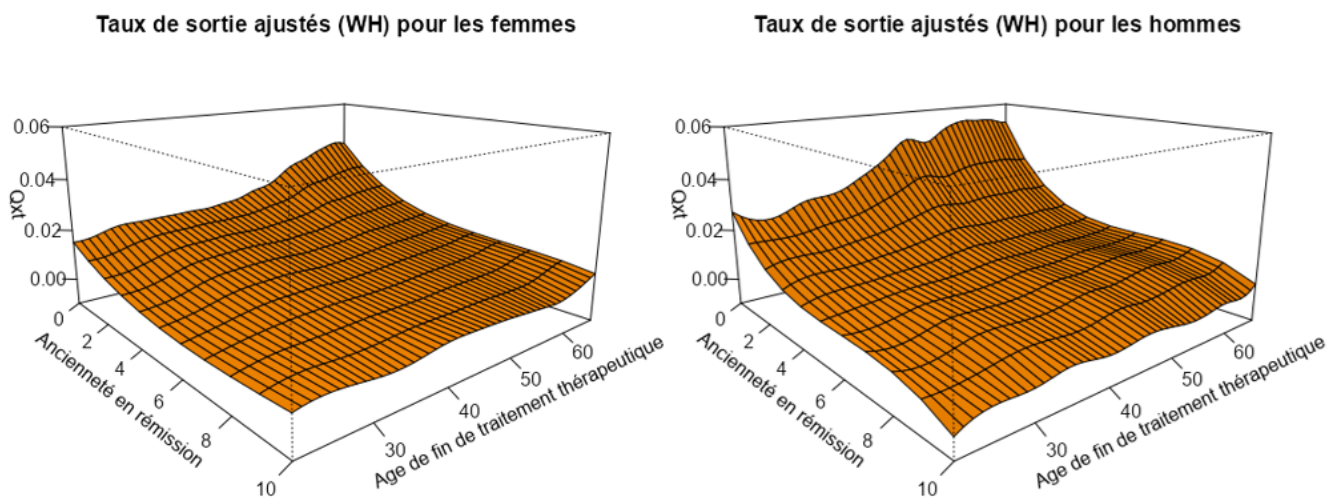
FIGURE 3.9 – Taux d’extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 1, 3, 4)$

A partir de $(z_h, z_v) = (3, 4)$, il est nécessaire de faire fluctuer α et β pour déterminer leur valeur optimale. Les comportements attendus de la surface selon la fluctuation de α et β sont les suivants :

- plus α est faible, plus les taux lissés seront fidèles aux taux empiriques selon l’axe horizontale, soit l’âge à la fin de traitement thérapeutique, et inversement ;
- plus β est faible, plus les taux lissés seront fidèles aux taux empiriques selon l’axe vertical, soit l’ancienneté en rémission, et inversement.

Comme susmentionné quant à l’intérêt de conserver les fluctuations des taux d’extra-mortalité selon l’ancienneté en rémission, il convient de privilégier une valeur faible de β . Une valeur plus élevée pour α vient alors apporter la régularité aux taux lissés.

Les figures 3.10 et 3.11 montrent que pour α , z_h et z_v fixés, la diminution de la valeur de β entraîne plus de fidélité aux données empiriques selon l’ancienneté en rémission. Par conséquent, les taux lissés présentés en figure 3.11 avec $(\alpha, \beta, z_v, z_h) = (1, 0.001, 3, 4)$ sont ceux sélectionnés pour la suite de l’étude.

FIGURE 3.10 – Taux d’extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 0.01, 3, 4)$

Comme le montre la figure 3.11 pour les hommes, la méthode de lissage produit des taux légèrement négatifs. Pour des raisons de logique, ils sont fixés à 0%.

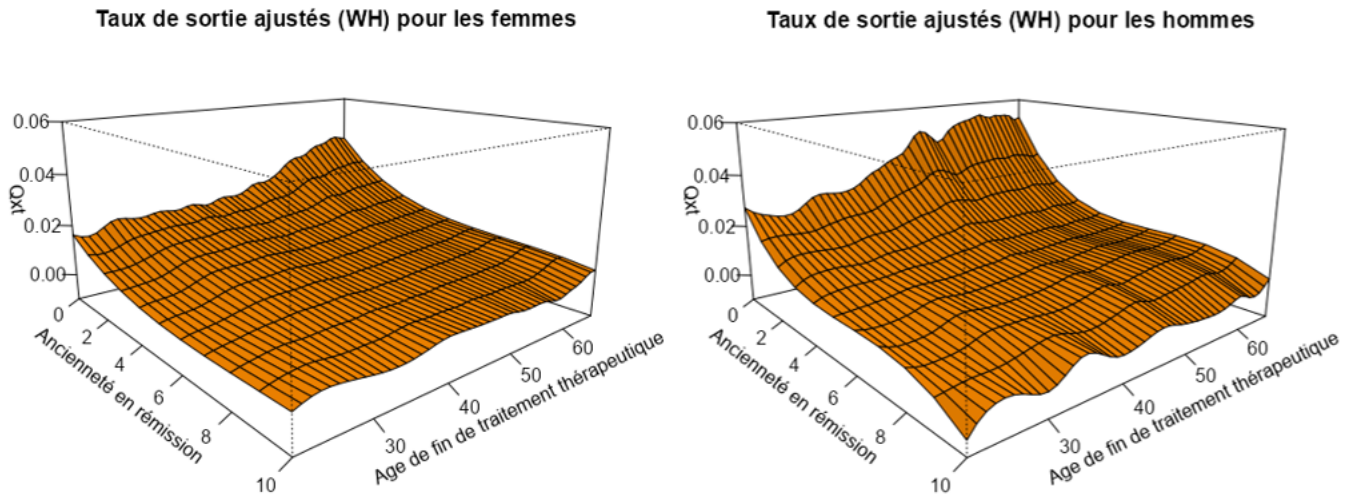


FIGURE 3.11 – Taux d’extra-mortalité lissés avec $(\alpha, \beta, z_v, z_h) = (1, 0.001, 3, 4)$

Finalement, les taux d’extra-mortalité générés estimés par la méthode des moments de Hoem puis lissés par la méthode de Whittaker-Henderson avec $(\alpha, \beta, z_v, z_h) = (1, 0.001, 3, 4)$ doivent être comparés à ceux obtenus à l’aide de la méthode des P-splines. L’application des tests d’adéquation sur chacune des tables permet de sélectionner la table d’extra-mortalité finale pour chaque sexe.

3.5.3 Validation et sélection des tables ajustées

Afin de sélectionner la méthode de lissage la plus adéquate, appliquée sur les taux bruts d’extra-mortalité estimés avec la méthode des moments de Hoem, le test du χ^2 d’adéquation et le test des changements de signes sont appliqués.

3.5.3.1 Application du test du χ^2 d’adéquation

Pour réaliser le test du χ^2 d’adéquation, il faut regrouper les âges de fin traitement thérapeutique ainsi que les anciennetés en rémission. Ainsi, le regroupement suivant est effectué, à l’aide du découpage usuel par tranche de 5 années :

- 9 classes d’âges ($L = 9$) : $[21;26[$, $[26;31[$, $[31;36[$, $[36;41[$, $[41;46[$, $[46;51[$, $[51;56[$, $[51;56[$, $[56;61[$ et $[61;66[$;
- 2 classes d’anciennetés ($K = 2$) : $[0;5[$ et $[5;10[$.

A la suite de ce regroupement, la statistique de test $W_{9,2}$ suit asymptotiquement une loi du χ^2 à 9 degrés de liberté. Si la valeur de la statistique, calculée pour chaque table lissée, est supérieure au quantile à 95% de la loi du χ^2 à 9 degrés de liberté alors l’hypothèse nulle est rejetée. Ainsi, la région de rejet de l’hypothèse nulle est $[16.92; +\infty[$.

Le tableau 3.1 présente les valeurs des statistiques pour chaque sexe et chaque méthode de lissage ainsi que l’hypothèse retenue sachant la région de rejet de l’hypothèse nulle. Il permet de remarquer que

pour l'ensemble des tables lissées avec la méthode de Whittaker-Henderson H_0 est acceptée à 95%. Cependant, pour le cas du lissage par la méthode des P-splines pour les femmes, l'hypothèse nulle est rejetée indiquant un mauvais ajustement des taux lissés par rapport aux taux bruts. Enfin, les valeurs des statistiques pour chaque sexe étant inférieures avec la méthode de Whittaker-Henderson à celles calculées pour la méthode des P-splines, il convient de conclure que la méthode de lissage de Whittaker-Henderson est la plus adaptée dans ces travaux pour lisser les taux d'extra-mortalité en restant fidèle à l'extra-mortalité empirique.

TABLE 3.1 – Résultats des applications du test du χ^2

Sexe	Méthode d'ajustement	$W_{2,9}$	Hypothèse retenue
Femme	P-splines	17.71	H_1
Femme	Whittaker-Henderson	6.8	H_0
Homme	P-splines	12.05	H_0
Homme	Whittaker-Henderson	8.06	H_0

Le test des changements de signes permet, par la suite, de s'assurer de la régularité des méthodes de lissage.

3.5.3.2 Application du test des changements de signes

Une fois l'écart calculé entre chaque taux brut et taux lissé, la statistique de test S est calculée pour chaque table genrée et ajustée. Avec l'approximation gaussienne, la région de rejet de l'hypothèse nulle est $] -\infty; -1.96] \cup [1.96; +\infty[$.

Le tableau 3.2 présente les résultats de cette application. Pour l'ensemble des tables d'extra-mortalité lissées, le test des changements de signes indique qu'il y a autant de taux sur-lissés que de taux sous-lissés. En ce sens, le critère de régularité est vérifié pour chaque méthode de lissage appliquée sur les taux bruts estimés par la méthode des moments de Hoem.

TABLE 3.2 – Résultats des applications du test des changements de signes

Sexe	Méthode d'ajustement	E	S	Hypothèse retenue
Femme	P-splines	238	- 1.29	H_0
Femme	Whittaker-Henderson	245	- 0.67	H_0
Homme	P-splines	233	- 1.74	H_0
Homme	Whittaker-Henderson	247	- 0.49	H_0

3.5.3.3 Sélection de la table d'extra-mortalité pour chaque sexe

A partir des résultats du test du χ^2 d'adéquation et du test des changements de signe, la table d'extra-mortalité lissée est sélectionnée pour chaque sexe.

Au regard du test du χ^2 d'adéquation, la méthode d'ajustement des taux bruts, calculés par la méthode des moments de Hoem, la plus fidèle aux phénomènes empiriques est la méthode de Whittaker-Henderson pour chaque sexe. De plus, le test des changements de signes assure que toutes les méthodes d'ajustement testées respectent le critère de régularité pour chaque sexe.

Par conséquent, les **taux bruts d'extra-mortalité estimés par la méthode des moments de Hoem et lissés par la méthode de Whittaker-Henderson** sont sélectionnés pour la construction

de la table d'extra-mortalité de chaque sexe. Le lecteur intéressé peut consulter la table d'extra-mortalité des suites du cancer non métastaté diagnostiqué et traité pour les femmes en annexe 4.15, et en annexe 4.16 pour les hommes.

Ce troisième chapitre permet d'estimer, pour chaque sexe, la probabilité de décéder du cancer non métastaté diagnostiqué et traité pour chaque année passée en rémission selon l'âge à la fin du traitement thérapeutique. A partir de ces probabilités, l'impact de la réduction de moitié du délai d'application du droit à l'oubli, pour les assurés âgés de 21 ans ou plus au moment du diagnostic, peut être mesuré.

Chapitre 4

Mesure de l'impact de la réduction du droit à l'oubli sur les engagements d'une garantie décès d'un assureur

L'objectif de ce dernier chapitre est de mesurer, pour la garantie décès, l'impact de la réduction à 5 ans du délai d'application du droit à l'oubli pour les emprunteurs dont le cancer a été diagnostiqué à partir de 21 ans. Cet impact est mesuré à l'aide du niveau des engagements d'un assureur sur sa garantie décès de son produit assurance emprunteur. Précisément, les données de portefeuille sont un extrait modifié d'un des portefeuilles d'assurance emprunteur du courtier Magnolia.fr.

Après avoir présenté les notions théoriques de calcul des engagements d'un assureur sur sa garantie décès emprunteur, la prise en considération des taux produits d'extra-mortalité est explicitée. Par la suite, le calcul est effectué pour divers scénarios de pourcentage d'assurés appliquant leur droit à l'oubli. Enfin, la sensibilité aux hypothèses considérées des résultats obtenus est étudiée.

4.1 Calcul du niveau des engagements d'un assureur

Afin d'établir le niveau de risque encouru dans ses portefeuilles, un assureur calcule régulièrement son niveau d'engagement. Plus précisément, le niveau d'engagement correspond, pour un assuré, au montant probable que l'assureur devra verser en cas de réalisation du risque couvert. De manière générale, ce niveau d'engagement dépend de la probabilité de réalisation du risque couvert de chaque individu et du montant de prestation garanti.

Cette partie a pour objectif de présenter la méthode de calcul du niveau d'engagement d'un assureur dans le cadre d'une garantie décès pour une assurance emprunteur. Par la suite, cette formule est adaptée en considérant les lois d'extra-mortalité des suites du cancer diagnostiqué et traité produites dans le chapitre précédent.

4.1.1 Notions préliminaires

En assurance emprunteur, la prestation versée en cas de décès correspond au capital emprunté restant dû au moment du décès de l'assuré. En ce sens, il est coutume de calculer le niveau des engagements

d'un assureur par pas mensuel de projection, comme le montant de remboursement des échéances du prêt.

Le niveau des prestations

Afin de déterminer le montant des prestations pour chaque pas de projection, le tableau d'amortissement du prêt de chaque assuré doit être conçu.

Il existe deux grandes catégories de remboursement d'un emprunt :

- prêt amortissable : à chaque échéance, l'emprunteur rembourse une partie du capital emprunté et une partie des intérêts. Il existe des prêts à amortissements constants et des prêts à mensualités constantes. Dans le premier cas, la part du capital remboursé à chaque échéance est constante et les mensualités fluctuent selon le montant des intérêts. Dans le second, les mensualités sont constantes impliquant une répartition non fixe entre intérêts dûs et part du capital remboursé. D'une manière progressive, l'assuré paye de moins en moins d'intérêts et de plus en plus de capital ;
- prêt in-fine : contrairement au cas précédent, seuls les intérêts sont remboursés à chaque échéance. Le capital est remboursé en une fois à l'échéance du prêt.

Ces différentes catégories de prêts impliquent un tableau d'amortissement différent. En ce sens, les méthodes de calcul appliquées pour la valeur du capital restant dû à chaque pas de projection sont indiquées dans les tableaux 4.1, 4.2 et 4.3. Pour l'ensemble des tableaux, les notations sont les suivantes :

- i correspond à l'indice pour l'emprunteur i ;
- $CRD_{i,t}$ est le capital restant dû de l'emprunteur i au début de la date t ;
- $Mens_{i,t}$ est la mensualité payée par l'emprunteur i à la date t ;
- $I_{i,t}$ est le montant des intérêts réglé par l'emprunteur i à la date t ;
- $Amort_{i,t}$ est le montant du capital amorti par l'emprunteur i entre la date $t - 1$ et la date t .

TABLE 4.1 – Tableau d'amortissement pour un assuré i dont le prêt est à amortissements constants

Date	CRD début de la période	Mensualité	Intérêts	Amortissement	CRD fin de la période
t	$CRD_{i,t}$	$Mens_{i,t} =$ $Amort_{i,t} + I_{i,t}$	$I_{i,t} =$ $CRD_{i,t} \times r_i$	$Amort_{i,t} =$ $\frac{CRD_{i,t}}{\text{durée restante}_{i,t}}$	$CRD_{i,t+1} =$ $CRD_{i,t} -$ $Amort_{i,t}$
t+1	$CRD_{i,t+1}$	$Mens_{i,t+1} =$ $Amort_{i,t+1} +$ $I_{i,t+1}$	$I_{i,t+1} =$ $CRD_{i,t+1} \times$ r_i	$Amort_{i,t+1} =$ $Amort_{i,t}$	$CRD_{i,t+2} =$ $CRD_{i,t+1} -$ $Amort_{i,t+1}$
t+2	$CRD_{i,t+2}$	$Mens_{i,t+2} =$ $Amort_{i,t+2} +$ $I_{i,t+2}$	$I_{i,t+2} =$ $CRD_{i,t+2} \times$ r_i	$Amort_{i,t+2} =$ $Amort_{i,t}$	$CRD_{i,t+3} =$ $CRD_{i,t+2} -$ $Amort_{i,t+2}$
...					
Dernière échéance du prêt (T)	$CRD_{i,T}$	$Mens_{i,T} =$ $Amort_{i,T} + I_{i,T}$	$I_{i,T} =$ $CRD_{i,T} \times r_i$	$Amort_{i,T} =$ $Amort_{i,T-1}$	0

TABLE 4.2 – Tableau d'amortissement pour un assuré i dont le prêt est à mensualités constantes

Date	CRD début de la période	Mensualité	Intérêts	Amortissement	CRD fin de la période
t	$CRD_{i,t}$	$Mens_{i,t}$	$I_{i,t} = CRD_{i,t} \times r_i$	$Amort_{i,t} = Mens_{i,t} - I_{i,t}$	$CRD_{i,t+1} = CRD_{i,t} - Amort_{i,t}$
t+1	$CRD_{i,t+1}$	$Mens_{i,t+1} = Mens_{i,t}$	$I_{i,t+1} = CRD_{i,t+1} \times r_i$	$Amort_{i,t+1} = Mens_{i,t+1} - I_{i,t+1}$	$CRD_{i,t+2} = CRD_{i,t+1} - Amort_{i,t+1}$
t+2	$CRD_{i,t+2}$	$Mens_{i,t+2} = Mens_{i,t+1}$	$I_{i,t+2} = CRD_{i,t+2} \times r_i$	$Amort_{i,t+2} = Mens_{i,t+2} - I_{i,t+2}$	$CRD_{i,t+3} = CRD_{i,t+2} - Amort_{i,t+2}$
...					
Dernière échéance du prêt (T)	$CRD_{i,T}$	$Mens_{i,T} = Mens_{i,T-1}$	$I_{i,T} = CRD_{i,T} \times r_i$	$Amort_{i,T} = Mens_{i,T} - I_{i,T}$	0

TABLE 4.3 – Tableau d'amortissement pour un assuré i dont le prêt est in-fine

Date	CRD début de la période	Mensualité	Intérêts	Amortissement	CRD fin de la période
t	$CRD_{i,t}$	$Mens_{i,t} = I_{i,t}$	$I_{i,t} = CRD_{i,t} \times r_i$	$Amort_{i,t} = 0$	$CRD_{i,t+1} = CRD_{i,t}$
t+1	$CRD_{i,t+1}$	$Mens_{i,t+1} = I_{i,t+1}$	$I_{i,t+1} = CRD_{i,t+1} \times r_i$	$Amort_{i,t+1} = 0$	$CRD_{i,t+2} = CRD_{i,t+1}$
t+2	$CRD_{i,t+2}$	$Mens_{i,t+2} = I_{i,t+2}$	$I_{i,t+2} = CRD_{i,t+2} \times r_i$	$Amort_{i,t+2} = 0$	$CRD_{i,t+3} = CRD_{i,t+2}$
...					
Dernière échéance du prêt (T)	$CRD_{i,T}$	$Mens_{i,T} = I_{i,T} + CRD_{i,T}$	$I_{i,T} = CRD_{i,T} \times r_i$	$Amort_{i,T} = CRD_{i,T}$	0

A partir des tableaux d'amortissement précédents, la valeur des prestations à verser en cas de décès de l'assuré est connue pour chaque pas de temps de projection. Ainsi, le niveau des engagements peut être apprécié.

Calcul du niveau des engagements

La formule du niveau des engagements d'une garantie décès emprunteur pour un assuré i est la suivante :

$$EA_i = \sum_{j=1}^{\text{Temps max}} {}_jP_x \times q_{x+j} \times CRD_{i,j} \times \text{quotité}_{DC,i} \times \nu_j$$

Avec, en complément des notations susmentionnées et déjà utilisées, les définitions des notations suivantes :

- EA_i est la valeur de l'engagement d'un assureur pour son assuré i ;
- j correspond aux pas de temps mensuels ;

- *Temps max* est le temps jusqu'au quel l'assuré est couvert pour la garantie décès. Il est égal au minimum entre la durée résiduelle en mois et l'âge limite de couverture du décès ;
- ${}_jP_x$ est la probabilité que l'assuré soit vivant à l'âge $x + j$ sachant qu'il est vivant à l'âge x . Cette valeur repose sur la table de mortalité réglementaire TH 00-02 si l'assuré est un homme et TF 00-02 si l'assuré est une femme ;
- q_{x+j} est la probabilité pour un assuré d'âge x et vivant en $x + j$ de décéder entre $x + j$ et $x + j + 1$. Là encore, cette valeur se détermine avec la table de mortalité réglementaire TH 00-02 si l'assuré est un homme et TF 00-02 si l'assuré est une femme ;
- $quotité_{DC,i}$ est la quotité de l'assuré i pour le risque décès ;
- ν_j correspond au facteur d'actualisation des flux qui repose sur le taux technique vie en vigueur à la date d'estimation des engagements.

Ainsi, le niveau total des engagements d'un assureur, noté EA , sur son portefeuille emprunteur pour la garantie décès est :

$$EA = \sum_{i=1}^{\text{Nombre total d'assurés}} EA_i$$

Ces formules sont par la suite adaptées pour mesurer le niveau des engagements d'un assureur dans le cadre d'application du droit à l'oubli.

4.1.2 Considération des lois d'extra-mortalité des suites du cancer diagnostiqué et traité

Lorsqu'un assuré souscrit une assurance emprunteur en appliquant son droit à l'oubli pour une pathologie cancéreuse, sa réelle probabilité de décéder est sous-estimée si l'assureur estime son niveau d'engagement avec la formule susmentionnée comprenant les taux de mortalité réglementaires.

Les travaux menés dans le troisième chapitre de ce mémoire montrent qu'il existe une probabilité significative de décéder des suites du cancer diagnostiqué et traité pendant la période de rémission selon l'âge auquel le traitement thérapeutique a été achevé.

Ainsi, pour un assuré appliquant son droit à l'oubli, l'hypothèse suivante est considérée dans cette étude : sa probabilité de décès est égale à la somme de sa probabilité de décès communiquée par les tables réglementaires TH 00-02 et TF 00-02 et la probabilité de décès des suites du cancer diagnostiqué et traité estimée dans cette étude. Soit,

$$q'_{x+j} = q_{x',R+j}^{Cancer} + q_{x+j}$$

Avec,

- q'_{x+j} est la réelle probabilité de décès d'un assuré appliquant son droit à l'oubli ;
- x' est l'âge de fin de traitement thérapeutique de l'assuré ;
- R correspond à l'ancienneté en rémission de l'assuré au moment de la souscription ;
- $q_{x',R+j}^{Cancer}$ est la probabilité de décès des suites du cancer diagnostiqué et traité sachant que l'âge à la fin du traitement thérapeutique est x' et sa rémission $R + j$.

A partir de cette hypothèse, la formule du niveau d'engagement pour un assuré appliquant son droit à l'oubli devient :

$$EA_i^{Cancer} = \sum_{j=1}^{Temps\ max} {}_jP'_x \times q'_{x+j} \times CRD_{i,j} \times quotité_{DC} \times \nu_j$$

Avec,

- EA_i^{Cancer} le niveau des engagements pour l'assuré i appliquant son droit à l'oubli ;
- ${}_jP'_x$ est la probabilité que l'assuré soit vivant à l'âge $x + j$ sachant qu'il est vivant à l'âge x . Cette valeur repose sur la probabilité réelle de décès de l'assuré appliquant son droit à l'oubli q'_{x+j} .

En considérant que le portefeuille d'un assureur se compose d'assurés sains et d'assurés appliquant leur droit à l'oubli, alors le niveau total de ses engagements pour la garantie décès est :

$$EA = \sum_{s=1}^{Nombre\ total\ d'assurés\ sains} EA_s + \sum_{c=1}^{Nombre\ total\ d'assurés\ cancer} EA_c^{Cancer}$$

4.2 Mesure de l'impact du droit à l'oubli pour Magnolia.fr

A partir de la formule du niveau des engagements susmentionnée intégrant le droit à l'oubli, l'impact de la réduction du délai d'application de ce droit est mesuré. Plus particulièrement, cette évaluation est réalisée sur le niveau des engagements pour la garantie décès d'une assurance emprunteur proposée par Magnolia.fr.

Après avoir analysé le portefeuille modifié de Magnolia.fr sur lequel est calculé l'impact, les résultats de ce dernier sont présentés.

4.2.1 Analyse du portefeuille modifié

Magnolia.fr contribue à ce mémoire en procurant un extrait de portefeuille d'un de leurs produits d'assurance emprunteur. Toutefois, pour des raisons de confidentialité, les données de ce portefeuille ont été significativement modifiées. En ce sens, les analyses suivantes ne reflètent pas les affaires détenues par cet organisme de courtage en assurance.

Le portefeuille étudié a les principales caractéristiques suivantes :

- Produit : assurance des crédits immobiliers pour les particuliers ;
- Age maximale de couverture du décès pour le produit souscrit : 89 ans ;
- Date de début de souscription : 01/01/2013 ;
- Date exacte d'extraction : 06/03/2022 ;
- Nombre de polices souscrites en cours au moment de l'extraction : 49 533 ;
- Répartition des sexes : 49,0% de femmes et 51,0% d'hommes.

Il est primordial de préciser que le portefeuilles étudié se fonde uniquement sur les polices souscrites toujours en cours à la date d'extraction.

Au sein du portefeuille d'assurés, la répartition des polices d'assurance souscrites, et toujours en cours, peut être observée selon l'année de souscription à l'aide de la figure 4.1. Il est normal d'observer peu de souscriptions pour les années 2013 à 2015 puisque la durée réelle d'un crédit immobilier, et donc de son assurance, varie entre 8 et 10 ans.

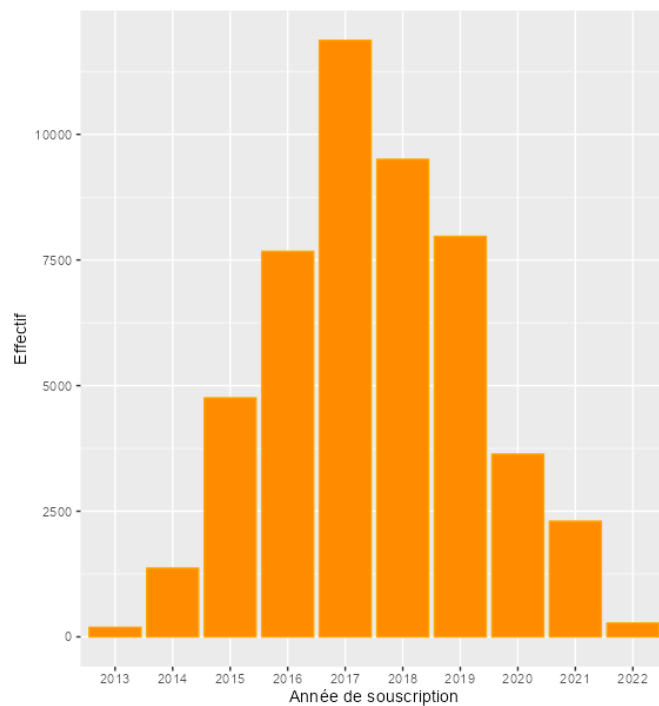


FIGURE 4.1 – Distribution des souscriptions par année

Par la suite, comme la figure 4.2 le montre, l'âge à la souscription s'étend de 18 à 74 ans dans le portefeuille modifié. De plus, 50% des assurés ont un âge à la souscription inférieur à 44 ans.

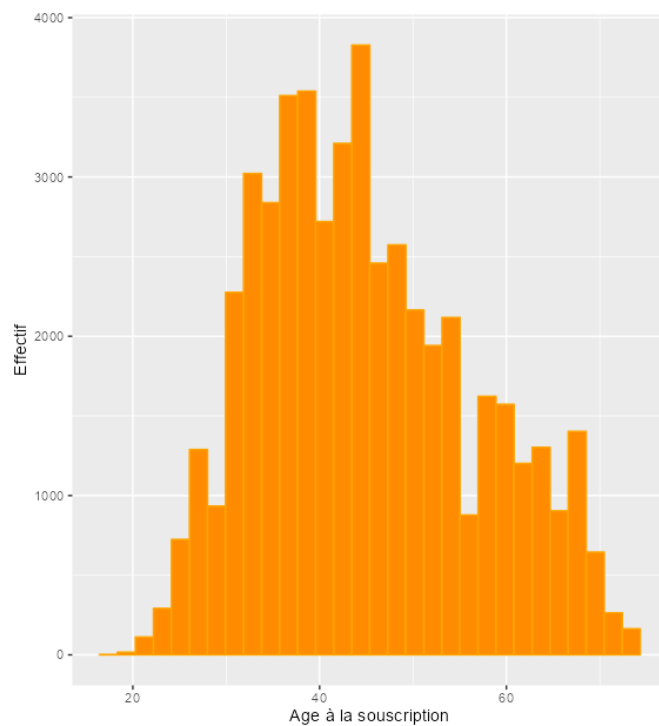


FIGURE 4.2 – Distribution des âges à la souscription

Parmi les informations communiquées, Magnolia.fr renseigne le capital assuré initial par tête. Il correspond à la quotité du décès multiplié par le montant total emprunté, soit le montant de capital sous risque qu'il doit verser en cas de décès. Il est compris entre 306 € et 2 580 000 € comme l'indique le tableau 4.4. Afin d'observer sa distribution, l'ensemble des capitaux inférieurs au quantile à 99% sont représentés dans la figure 4.3. En ce sens, il convient de souligner que la majorité des capitaux sous risque par assuré se concentre sur la tranche [50 000 € ; 200 000 €].

TABLE 4.4 – Statistiques sur les capitaux initiaux assuré par tête (CA)

	Minimum	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Quantile à 99%	Maximum
CA	306	51 442	87 000	104 719	138 229	366 774.4	2 580 000

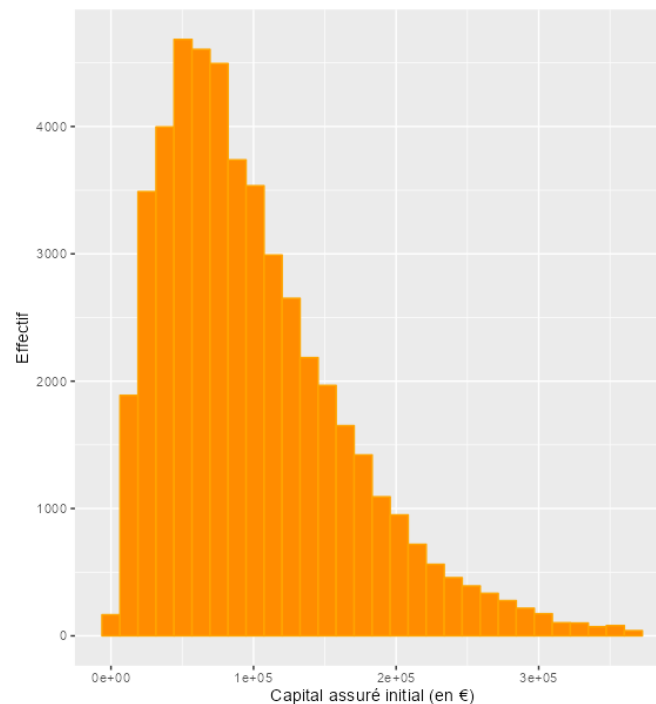


FIGURE 4.3 – Distribution des capitaux assurés initiaux par tête inférieurs au quantile à 99%

Finalement, puisque des scénarios de pourcentage d'entrants appliquant leur droit à l'oubli sont simulés par la suite, il est intéressant de connaître la part des assurés, par année de souscription, pouvant appliquer leur droit à l'oubli. Pour rappel, outre les conditions sur la pathologie et le statut médical de l'assuré, ce dernier peut appliquer son droit à l'oubli si l'échéance de son emprunt immobilier a lieu avant son 71^{ème} anniversaire. Le tableau 4.5 recense par année de souscription, parmi les contrats en cours, la part d'assuré respectant ce dernier critère. Il est donc intéressant de conclure qu'au minimum 7 assurés sur 10 peuvent appliquer leur droit à l'oubli en se référant uniquement à la condition sur l'âge à la fin du prêt.

TABLE 4.5 – Pourcentage d'entrants respectant la condition d'âge pour le droit à l'oubli

Année de souscription	% dont l'âge de fin < 71 ans
2013	84%
2014	90.2%
2015	90.9%
2016	90.1%
2017	89.1%
2018	81.9%
2019	75.1%
2020	73.5%
2021	74.4%
2022	79.9%

A partir de ce portefeuille, la mesure d'impact de la réduction de moitié du délai d'application du droit à l'oubli peut être étudiée.

4.2.2 Présentation des résultats

Puisque Magnolia.fr n'a pas accès au passif de la pathologie cancéreuse de l'assuré entrant, la mesure de l'impact se fonde sur différents scénarios de parts d'assurés appliquant leur droit à l'oubli. Pour chacun de ces scénarios, le niveau des engagements total est mesuré.

Démarche

La part d'entrants appliquant leur droit à l'oubli est considérée uniquement sur la dernière année pleine de souscription, soit du 01/01/2021 au 06/03/2022. Dans ce cadre, il est ainsi supposé que la réduction du droit à l'oubli est effective à compter du 01/01/2021, et son impact est mesuré après un an et deux mois d'application pour ce portefeuille.

Dans les statistiques 2020 fournies par FRANCE ASSUREURS (2021a), la part des demandes d'assurance de prêt présentant un risque aggravé de santé est estimée à 11.4%. Les assurés touchés par un cancer et en rémission entre 5 et 10 ans étant considérés au moment de cette étude en risque aggravé de santé, c'est à partir de ce constat qu'est défini l'échelle de pourcentages d'entrants appliquant leur droit à l'oubli. En ce sens, les scénarios sont simulés pour un pourcentage allant de 1% à 15%, 15% étant le scénario le plus pessimiste.

Pour chaque scénario, parmi les assurés ayant souscrit une garantie décès emprunteur entre le 01/01/2021 et le 06/03/2022 et ayant un âge de fin de prêt strictement inférieur à 71 ans, un pourcentage d'assurés est considéré comme en rémission d'un cancer diagnostiqué et traité. Aléatoirement, une durée en rémission au moment de la souscription est alors attribuée pour chaque assuré considéré en rémission d'un cancer. Pour correspondre exclusivement aux nouveaux risques impliqués par l'évolution législative, elle est comprise entre 5 et 10 ans. De cette durée en rémission et de l'âge à la souscription, l'âge de fin de traitement thérapeutique est estimé.

L'augmentation du pourcentage d'assuré appliquant leur droit à l'oubli se base systématiquement sur le scénario précédent. Ainsi, un assuré considéré comme en rémission d'un cancer et dont une durée en rémission lui a déjà été attribuée, a toujours ce même statut dans le scénario suivant.

Résultats

Enfin, le calcul des engagements est réalisé. Le tableau 4.6 indique le statut médical moyen généré (durée en rémission au moment de la souscription et âge de fin de traitement thérapeutique) pour

chaque scénario, et le tableau 4.7 communique le niveau des engagements de Magnolia.fr. Pour rappel, les chiffres présentés dans le tableau 4.7 ne reflètent pas le niveau réel des engagements de Magnolia.fr puisque le portefeuille sur lequel il repose a été modifié.

TABLE 4.6 – Paramètres des différents scénarios

Scénarios	% d'entrants appliquant le droit à l'oubli	Age moyen de fin traitement thérapeutique	Age moyen de fin traitement thérapeutique pondéré par le capital assuré	Durée moyenne en rémission	Durée moyenne en rémission pondérée par le capital assuré
0	0%				
1	1%	37.29	36.14	7.32	7.82
2	2%	37.51	36.12	7.25	7.73
3	3%	37.91	36.66	7.39	7.59
4	4%	37.41	36.35	7.30	7.44
5	5%	37.27	35.78	7.34	7.46
6	6%	36.59	35.68	7.46	7.51
7	7%	36.80	35.77	7.49	7.57
8	8%	35.87	35.01	7.57	7.64
9	9%	36.11	35.24	7.58	7.66
10	10%	36.26	35.45	7.58	7.63
11	11%	35.85	35.03	7.62	7.63
12	12%	35.64	34.74	7.61	7.63
13	13%	35.46	34.61	7.61	7.65
14	14%	35.63	34.81	7.60	7.64
15	15%	35.87	35.09	7.64	7.68

TABLE 4.7 – Calcul des niveaux des engagements de Magnolia.fr selon les scénarios

Scénarios	% d'entrants appliquant le droit à l'oubli	Ecart moyen par assuré pondéré par les capitaux assurés (en €)	Engagement total de l'assureur (en €)	Evolution par rapport au scénario initial
0	0%		205 221 885 38	0.000%
1	1%	2 408.35	205 269 740.81	0.023%
2	2%	2 277.53	205 315 640.09	0.046%
3	3%	3 090.94	205 394 216.83	0.084%
4	4%	3 539.76	205 488 290.65	0.130%
5	5%	3 563.93	205 571 034.57	0.170%
6	6%	3 550.59	205 624 539.19	0.196%
7	7%	3 355.21	205 682 248.35	0.224%
8	8%	3 188.72	205 722 314.97	0.244%
9	9%	3 070.29	205 754 640.27	0.260%
10	10%	3 184.53	205 830 720.84	0.297%
11	11%	3 336.15	205 910 643.76	0.336%
12	12%	3 331.37	205 968 152.07	0.364%
13	13%	3 231.10	206 009 844.49	0.384%
14	14%	3 262.22	206 075 765.54	0.416%
15	15%	3 171.43	206 108 773.76	0.432%

Pour les paramètres des différents scénarios, il est intéressant de regarder les paramètres pondérés par

les capitaux assurés initiaux puisque ce sont également les niveaux de capitaux sous risque qui font évoluer le niveau des engagements. En effet, si un assuré a une durée en rémission de 5 ans avec un capital sous risque élevée alors Magnolia.fr aura un coût du risque élevé pour cet individu, pondéré par la probabilité appliquée de décès comprenant le risque cancer. Pour l'ensemble des scénarios, la durée moyenne en rémission pondérée est autour de 7 ans et demi avec un âge de fin traitement moyen pondéré fluctuant autour de 35 ans.

De manière générale, le tableau 4.7 montre qu'il existe un impact sur le niveau des engagements de Magnolia.fr suite à l'application du nouveau droit à l'oubli pour la garantie décès. Cet impact varie linéairement entre 0.0% pour le scénario le plus optimiste (aucun entrant appliquant leur droit à l'oubli), et 0.432% pour le scénario le plus pessimiste. Bien que ces taux soient inférieurs à 1%, il s'agit, pour le scénario le plus pessimiste, d'un coût de près de 1 000 000 € pour un an et deux mois d'application de cette nouvelle législation.

Ces résultats reposent sur les lois d'extra-mortalité calculées dans le chapitre précédent où différentes hypothèses ont été prises. Ainsi, il est nécessaire d'étudier la sensibilité des résultats aux deux principales hypothèses considérées, soient :

- la situation épidémiologique des cancers aux Etats-Unis est similaire à celle de la France ;
- la durée du traitement thérapeutique est comprise entre un et douze mois.

4.3 Etude de sensibilité des résultats aux hypothèses principales

Pour réaliser cette étude, l'ensemble des hypothèses suivantes ont été prises :

- si un individu est touché par plusieurs cancers, la période de rémission court à compter de la fin de traitement du dernier cancer diagnostiqué ;
- les malades touchés par un cancer métastasé sont exclus de l'étude car ils ont une probabilité significative de rechute après leur traitement thérapeutique ;
- les patients atteints d'un cancer non métastasé ne rechutent pas une fois entrés en rémission. Ainsi, le taux de rechute pour les cancers non métastasés est nul ;
- l'âge de fin de traitement thérapeutique est estimé selon des durées moyennes de traitement en fonction du stade et du grade du cancer. Ainsi, la durée du traitement thérapeutique est comprise entre un et douze mois ;
- la situation épidémiologique des cancers aux Etats-Unis est similaire à celle de la France.

Les probabilités de décès des suites du cancer diagnostiqué et traité calculées varient notamment en fonction des deux dernières hypothèses susmentionnées. C'est pourquoi, il est nécessaire d'étudier la sensibilité des lois construites à celles-ci et leur influence sur les résultats précédents.

Chaque hypothèse est étudiée séparément dans la suite de l'étude et seul le scénario pessimiste est remis en question. En effet, il est considéré que l'impact sur les résultats est identique pour chacun des scénarios.

Pour la suite, les raccourcis suivants sont utilisés :

- Hypothèse 1 : l'âge de fin de traitement thérapeutique est estimé selon des durées moyennes de traitement en fonction du stade et du grade du cancer. Ainsi, la durée du traitement thérapeutique est comprise entre un et douze mois ;
- Hypothèse 2 : la situation épidémiologique des cancers aux Etats-Unis est similaire à celle de la France.

4.3.1 Hypothèse 1

La durée des traitements thérapeutiques est estimée à partir des durées moyennes de chaque traitement principale : chirurgie, chimiothérapie et radiothérapie. La durée finalement calculée varie d'un mois à douze mois. Toutefois, il est possible que cette durée de traitement ait été sous-estimée. En effet, il n'existe pas de durée unique de traitement nécessaire pour chaque type de cancer. Le traitement est adapté pour chaque patient selon son état de santé, sa sensibilité aux traitements et l'avancée de son cancer.

Par la suite, l'âge de fin de traitement thérapeutique est ajusté en ajoutant six mois supplémentaires à chaque individu. Les lois réalisées à l'aide de l'estimateur des moments de Hoem et du lissage de Whittaker-Henderson sont recalculées. Le lecteur intéressé peut observer le détail des lois produites aux annexes 4.17 et 4.18, respectivement pour les femmes et les hommes.

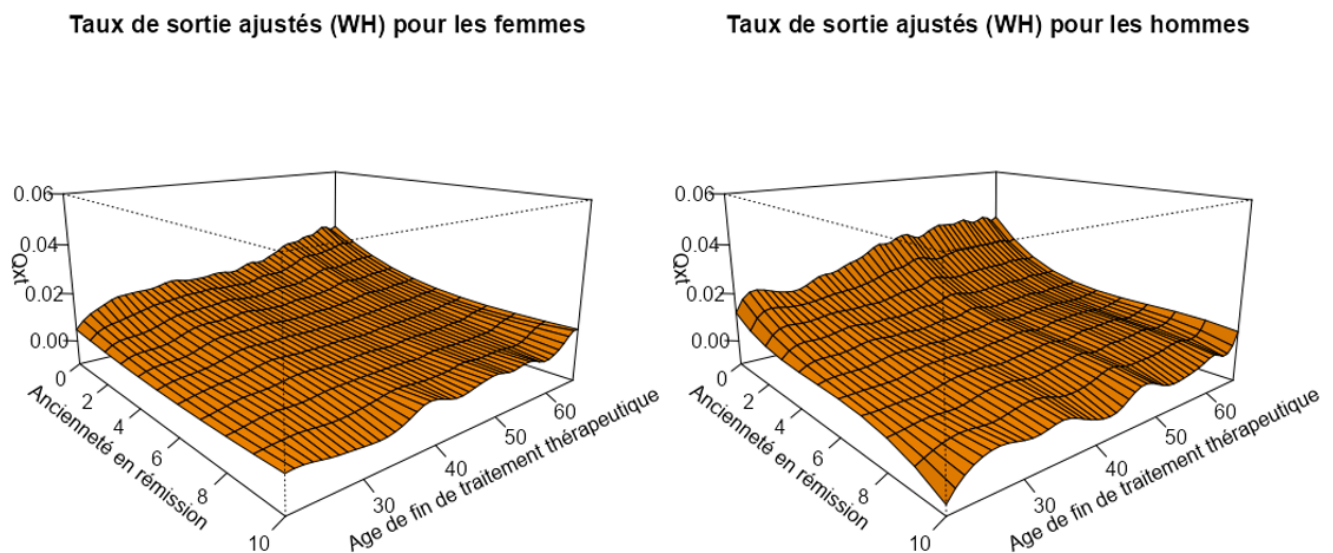


FIGURE 4.4 – Taux d'extra-mortalité ajustés selon l'âge à la fin du traitement thérapeutique vieilli de 6 mois

D'après la figure 4.4, en comparaison des taux finalement produits et sélectionnés au chapitre précédent, la volatilité des lois pour chaque sexe est réduite. De plus, cette figure souligne que la probabilité de décès des suites du cancer diagnostiqué et traité diminue nettement pour les premières années en rémission avec le vieillissement de l'âge de fin de traitement thérapeutique de six mois pour les deux sexes. Au regard des changements notés, le niveau des engagements de Magnolia.fr pour le scénario pessimiste doit être déterminé.

Le tableau 4.11 présente le calcul des engagements pour le scénario le plus pessimiste, soit 15% d'entrants appliquant le droit à l'oubli, en vieillissant l'âge de fin de traitement thérapeutique de six mois. Comme observé sur la figure 4.4, les probabilités de décès des suites du cancer diagnostiqué et traité étant revues à la baisse pour chaque sexe, le niveau des engagements de Magnolia.fr est légèrement diminué. Il passe de 206 108 773.76 € à 206 101 832.83 €, soit une baisse de 0,0034%.

TABLE 4.8 – Calcul des niveaux des engagements de Magnolia.fr après vieillissement de l'âge de fin de traitement thérapeutique (H1)

Scénarios	% d'entrants appliquant le droit à l'oubli	Ecart moyen par assuré pondéré par les capitaux assurés (en €)	Engagement total de l'assureur (en €)	Evolution par rapport au scénario initial
0	0%		205 221 885 38	0.000%
15	15%	3 171.43	206 108 773.76	0.432%
15 H1	15%	3 160.25	206 101 832.83	0.429%

4.3.2 Hypothèse 2

Pour calculer les probabilités de décès des suites d'un cancer diagnostiqué et traité pour un patient en rémission, il est nécessaire d'obtenir un suivi médical de chaque patient. En France, ce type de données n'est pas accessible au grand public à la maille individu. C'est pourquoi, il a été décidé d'utiliser la base de données des suivis de cancers par patient produite et mise en ligne par les Etats-Unis. De ce fait, ce sont des lois d'extra-mortalité par sexe produites sur une population américaine qui sont utilisées pour estimer le risque encouru sur un portefeuille d'assurés français. En ce sens, il est considéré que la situation épidémiologique des cancers aux Etats-Unis est identique à celle en France.

Il est judicieux de remettre en question cette hypothèse et son impact sur le niveau des engagements finaux de Magnolia.fr. Pour cela, les taux standardisés d'incidence et de mortalité du cancer de chaque pays doivent être étudiés.

TABLE 4.9 – Taux standardisés monde d'incidence et de mortalité du cancer par sexe en France et aux Etats-Unis en 2020 (INTERNATIONAL AGENCY FOR RESEARCH ON CANCER, 2022)

Pays	Sexe	Incidence	Mortalité
France	Femme	297.8	82.3
USA	Femme	333.2	76.9
France	Homme	397.2	139.2
USA	Homme	400.9	98.1

De manière générale, le tableau 4.9 indique que la France a un taux standardisé monde d'incidence du cancer inférieur à celui des Etats-Unis en 2020, et un taux standardisé monde de mortalité du cancer supérieur à celui des Etats-Unis en 2020 pour les deux sexes. Plus précisément, le tableau 4.10 référence ces évolutions chiffrées. Ce constat implique que la France diagnostique moins de cancers mais davantage de décès des suites d'un cancer pour chaque sexe.

TABLE 4.10 – Comparaison sur les taux du tableau 4.9

Pays	Sexe	Incidence	Mortalité
France/USA	Femme	- 10.6%	+ 7.0%
France/USA	Homme	- 0.9%	+ 41.9%

Afin de calculer la sensibilité des engagements aux lois d'extra-mortalité construites reposant sur l'hypothèse 2, il est nécessaire de faire évoluer la population sous risque de la différence d'incidence remarquée puis identiquement pour le nombre de décès des suites du cancer diagnostiqué et traité. A partir du tableau 4.10, il vient les taux d'ajustement pour chacune des tables suivantes :

- Taux d’ajustement pour les femmes = $\frac{1+(-10.6\%)}{(1+7.0\%)} = 119.7\%$;
- Taux d’ajustement pour les hommes = $\frac{1+(-0.9\%)}{(1+41.9\%)} = 143.2\%$.

Une fois les lois d’extra-mortalité pour chaque sexe majorées des taux susmentionnés, le niveau des engagements de Magnolia.fr pour sa garantie décès est calculé dans le cadre du scénario le plus pessimiste.

TABLE 4.11 – Calcul des niveaux des engagements de Magnolia.fr après ajustement au cas français (H2)

Scénarios	% d’entrants appliquant le droit à l’oubli	Ecart moyen par assuré pondéré par les capitaux assurés (en €)	Engagement total de l’assureur (en €)	Evolution par rapport au scénario initial
0	0%		205 221 885 38	0.000%
15	15%	3 171.43	206 108 773.76	0.432%
15 H2	15%	4 120.38	206 368 928.37	0.535%

A l’aide des résultats présentés dans le tableau 4.11, il est important de souligner que le niveau des engagements augmente de 0.130%, soit un impact final sur le taux d’évolution par rapport au scénario initial de 25%. Il est donc évident que les résultats présentés dans la partie précédente sont sensibles à l’hypothèse 2 en terme d’impact final.

La loi d’extra-mortalité par sexe étant sensible à la définition de l’âge de fin traitement thérapeutique et à la situation épidémiologique précise du pays, l’impact de la réduction de moitié du délai d’application du droit à l’oubli calculé dans cette étude peut être discuté.

4.4 Limites, réflexions et ouvertures des travaux

Les travaux menés jusque là montrent qu’il existe une probabilité non négligeable, pour les personnes atteintes d’un cancer non métastasé et traité, de décéder des suites de ce même cancer entre la cinquième et la dixième année passée en rémission. De ces probabilités et d’un portefeuille modifié de Magnolia.fr, une évolution à la hausse de ses engagements a été quantifiée. Toutefois, il convient d’analyser les limites et les perspectives de ces résultats.

4.4.1 Les limites de l’application

Cette étude comporte des limites quant à l’utilisation des données SEER et les hypothèses considérées dans la méthodologie utilisée.

La complexité des données de cancérologie

L’étude de l’extra-mortalité des suites du cancer diagnostiqué et traité nécessite une analyse approfondie de la base SEER. Plus particulièrement, le pragmatisme de ces travaux se base sur les variables principales quant à la définition du cancer, son stade d’avancement, le traitement suivi et le statut vital du patient.

Un grand nombre de variables liées à la description biologique de la pathologie cancéreuse peuvent ap-

porter davantage de précisions quant au traitement spécifique appliqué à chaque patient. Par exemple, une variable renseigne sur la dose prescrite de rayons pour une radiothérapie externe. En ce sens, elles pourraient être utilisées pour préciser l'âge à la fin du traitement thérapeutique pour chaque patient avec une connaissance suffisante en oncologie.

Limites méthodologiques de l'étude

La construction des tables d'extra-mortalité à partir des données SEER repose sur les hypothèses simplificatrices rappelées ci-dessous :

- les personnes atteintes d'un cancer métastasé rechutent des suites de cette pathologie et n'entrent jamais en phase de rémission. Elles sont donc exclues de cette étude ;
- les personnes touchées d'un cancer non métastasé ont une probabilité nulle de rechute pendant la phase de rémission ;
- si une personne est touchée par plusieurs cancers non métastasés, alors la fin de traitement thérapeutique est comptée à partir de la date du dernier diagnostic ;
- le temps de traitement thérapeutique est estimé selon les durées moyennes de la chirurgie, de la chimiothérapie et de la radiothérapie ;
- la situation épidémiologique des cancers aux Etats-Unis est identique à celle de la France.

Ces hypothèses permettent de s'adapter au cadre stricte d'application du droit à l'oubli qui impose :

- l'officialisation par un médecin que le patient a correctement achevé son traitement thérapeutique et qu'il entre en phase de rémission tel qu'elle est définie dans la Convention AERAS ;
- la validation, toujours par un médecin, que le patient n'a pas rechuté au cours de sa phase de rémission pendant cinq années consécutives ;
- que la dernière échéance du prêt du patient ait lieu avant ses 71 ans.

Les hypothèses concernant la construction de l'âge de fin de traitement thérapeutique et l'équivalence France et Etats-Unis ayant été remises en question, il pourrait être intéressant d'ajouter une modélisation de rechute des patients en intégrant les cancers métastasés pour généraliser cette étude, permettant ainsi de supprimer les trois premières hypothèses citées précédemment.

4.4.2 Réflexion sur les résultats

Bien que l'impact calculé soit attribuable uniquement pour un des portefeuilles de la société de courtage en assurance emprunteur Magnolia.fr, il est intéressant de réfléchir à une généralisation de ces travaux. Les tables d'extra-mortalité construites peuvent être utilisées par d'autres acteurs du marché. En effet, avec ces probabilités, ils sont en mesure de réévaluer les risques couverts et ainsi ajuster leurs primes sous contrainte de solvabilité.

De plus, une assurance emprunteur ne se restreint pas à la garantie décès. Afin de tenir compte de l'impact de la réduction du droit à l'oubli dans son ensemble, il est nécessaire d'étudier les probabilités de passage en incapacité et en invalidité des personnes touchées par un cancer diagnostiqué et traité, puisque ce risque donne également droit à des prestations.

Finalement, les probabilités d'extra-mortalité ont été établies avec les cancers diagnostiqués entre le 01/01/2009 et le 31/12/2019, et les scénarios simulés pour mesurer l'impact selon la situation épidémiologique des cancers à date. Il convient de s'interroger sur l'impact du droit à l'oubli à horizon 2040 si la prévalence des cancers en France continue son ascension. Il y aurait alors davantage de personnes assurables à l'aide du droit à l'oubli.

Ce dernier chapitre représente l'application actuarielle de ce mémoire. Il permet de prendre en considération les tables d'extra-mortalité construites au chapitre précédent dans le calcul des engagements d'un assureur. Il en résulte que même s'il existe une probabilité non nulle de décès des suites du cancer diagnostiqué et traité, son intégration dans la formule des engagements d'un assureur implique une hausse inférieure à 1% dans le scénario le plus pessimiste d'entrants appliquant leur droit à l'oubli. Aussi, ce chapitre permet de discuter des perspectives d'évolutions de ces travaux. Ces derniers peuvent désormais être conclus.

Conclusion

L'objectif de ce mémoire était de quantifier l'impact de la réduction du délai d'application du droit à l'oubli de 10 à 5 ans, pour les personnes atteintes d'un cancer après leurs 21 ans. Pour parvenir à ces fins, les travaux ont débuté par une mise en contexte des enjeux de cette avancée législative avec une présentation de l'assurance emprunteur, ses évolutions réglementaires et de la situation épidémiologique des cancers en France. Puis, les données de la base SEER ont été analysées et retraitées afin de correspondre avec le cadre d'application du nouveau droit à l'oubli. Ce deuxième chapitre a notamment permis de considérer des hypothèses pour l'estimation des taux d'extra-mortalité des suites du cancer diagnostiqué et traité réalisée dans le troisième chapitre de ce mémoire. Finalement, les tables d'extra-mortalité genrées construites ont été prises en compte dans le calcul des engagements d'un assureur pour une garantie décès emprunteur.

Bien qu'il existe une probabilité non nulle de décès des suites du cancer diagnostiqué et traité pour les individus touchés par une pathologie cancéreuse après leur 21 ans, sa prise en considération dans le niveau des engagements d'un assureur a un impact relativement faible. Dans le scénario le plus pessimiste, qui correspond à 15% d'entrants appliquant leur droit à l'oubli, la hausse du niveau des engagements est estimée à 0.432%. Il est important de souligner que cette mesure dépend fortement du type de portefeuille emprunteur étudié et ne représente qu'un pan de la loi Lemoine. L'organisme assureur doit compléter cette mesure de la hausse du risque d'antisélection avec les autres impacts anticipés de la loi Lemoine. Aussi, l'impact quantifié est lié à la prévalence des cancers en France : si elle continue son ascension, alors il y aura davantage de personnes assurables à l'aide du droit à l'oubli et l'impact sera d'autant plus fort pour les assureurs.

Une perspective de développement pour ce mémoire repose notamment sur les hypothèses considérées pour estimer les probabilités de décès des suites du cancer diagnostiqué et traité. Avec des connaissances supplémentaires en oncologie, il pourrait être intéressant de préciser l'estimation de l'âge de fin de traitement thérapeutique pour chaque individu. Enfin, il pourrait être pertinent d'ajouter une modélisation de rechute des patients, en intégrant les cancers métastasés, pour généraliser cette étude.

Bibliographie

- ADAMO M, GROVES C, DICKIE L, RUHL J (2021). . SEER Program Coding and Staging Manual 2022. Rapp. tech. National Cancer Institute, Bethesda, MD 20892.
- CANCER, I. N. D. (s. d.[a]). Dictionnaire. URL : <https://www.e-cancer.fr/Dictionnaire/T/tumeur-maligne>.
- CANCER, I. N. D. (s. d.[b]). Dictionnaire. URL : <https://www.e-cancer.fr/Dictionnaire/R/radiotherapie>.
- CANCER, I. N. D. (s. d.[c]). Dictionnaire. URL : <https://www.e-cancer.fr/Dictionnaire/C/chimiotherapie>.
- CANCER, I. N. D. (s. d.[d]). Dictionnaire. URL : <https://www.e-cancer.fr/Dictionnaire/I/immunotherapie>.
- CANCER, I. N. D. (s. d.[e]). Dictionnaire. URL : <https://www.e-cancer.fr/Dictionnaire/H/hormonotherapie>.
- CODE DE LA CONSOMMATION (2008). Loi n°2008-3 du 3 janvier 2008 - art. 33. C. consom., art. L136-1.
- CODE DE LA CONSOMMATION (2010). Loi n° 2010-737 du 1er juillet 2010 portant réforme du crédit à la consommation. C. consom., art. L136-1.
- CODE DES ASSURANCES (2007). Modifié par Loi n°2006-1770 du 30 décembre 2006 - art. 65 (V) JORF 31 décembre 2006 en vigueur le 1er octobre 2007. C. assur., art. L141-1.
- CODE DES ASSURANCES (2012). Ordonnance n°2011-839 du 15 juillet 2011 - art. 8. C. assur., art. L175-16.
- CODE DES ASSURANCES (2014). Loi n° 2014-344 du 17 mars 2014, dite loi consommation ou loi Hamon. C. assur., art. L113-15-2.
- CODE DES ASSURANCES (2017). Loi n° 2017-203 du 21 février 2017. C. assur., art. L113-12.
- COMITÉ CONSULTATIF DU SECTEUR FINANCIER (2020). Bilan de l'assurance emprunteur. Réalisé par Actélior pour le CCSF.
- CONVENTION AERAS (2019).
- CONVENTION AERAS (2020).
- CONVENTION AERAS RÉVISÉE (2015). Convention AERAS révisée à la suite de la mise en place d'un droit à l'oubli.
- CONVENTION BELORGEY (2001). Convention visant à améliorer l'accès à l'emprunt et à l'assurance des personnes présentant un risque de santé aggravé.
- CURRIE, I., DURBÁN, M. et EILERS, P. (2003). Using P-splines to extrapolate two-dimensional Poisson data. *Proceedings of 18th International Workshop on Statistical Modelling, Leuven, Belgium*, p. 97-102.
- DURBÁN, M., CURRIE, I. et EILERS, P. (2002). Using P-splines to smooth two-dimensional Poisson data. *Proceedings of 17th International Workshop on Statistical Modelling, Chania, Crete*, p. 207-214.
- EILERS, P. H. et MARX, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical science* 11.2, p. 89-121.
- FONDATION CONTRE LE CANCER (s. d.). Les types de cancers. URL : <https://www.santepubliquefrance.fr/maladies-et-traumatismes/cancers>.
- FRANCE ASSUREURS (2021a). Convention AERAS : statistiques 2020.
- FRANCE ASSUREURS (2021b). L'assurance française, les données clés 2020.

- FRANCE ASSUREURS (2022). L'assurance française, les données clés 2021.
- GREENWOOD, P. E. et NIKULIN, M. S. (1996). A guide to chi-squared testing. T. 280. John Wiley & Sons.
- HENDERSON, R. (1924). A new method of graduation. *Transactions of the Actuarial Society of America* 25, p. 29-40.
- INSEE (2022). Population par sexe. URL : <https://www.insee.fr/fr/statistiques/2381466#tableau-figure1>.
- INSTITUT NATIONAL DU CANCER (2022). Panorama des cancers en France - Edition 2022.
- INTERNATIONAL AGENCY FOR RESEARCH ON CANCER (2022). Cancer today. URL : <https://gco.iarc.fr/today/>.
- JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE (2001). Loi n° 2001-1168 du 11 décembre 2001 portant mesures urgentes de réformes à caractère économique et financier.
- JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE (2022). Loi n° 2022-270 du 28 février 2022 pour un accès plus juste, plus simple et plus transparent au marché de l'assurance emprunteur.
- KAPLAN, E. L. et MEIER, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association* 53.282, p. 457-481.
- LE GUYADER-PEYROU, S., DEFOSSEZ, G., DANTONY, E., MOUNIER, M., CORNET, E., UHRY, Z., COWPLI-BONY, A., MAYNADIÉ, M., TROUSSARD, X., DELAFOSSE, P. et al. (2019). Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018. *Etude à partir des registres des cancers du réseau Francim 2*.
- LEE, S. (s. d.). Qu'est-ce que le cancer métastatique? URL : <https://cancer.ca/fr/cancer-information/cancer-types/metastatic/what-is-metastatic-cancer>.
- PLANCHET, F. et THÉRON, P. (2006). Modèles de durée. *Economica*.
- PLANCHET, F. et WINTER, P. (2010). L'utilisation des splines bidimensionnels pour l'estimation de lois de maintien en arrêt et de travail. *arXiv preprint arXiv :1001.1907*.
- SANTÉ PUBLIQUE FRANCE (2021). Cancers. URL : <https://www.santepubliquefrance.fr/maladies-et-traumatismes/cancers>.
- THE NATIONAL CANCER ACT OF 1971 (1971).
- WHITTAKER, E. T. (1922). On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society* 41, p. 63-75.
- WINTER, P (2005). Méthodes bidimensionnelles pour l'ajustement de lois de maintien d'expérience en arrêt de travail. *ISFA, Mémoire d'actuariat*.

Annexes

Retraitements des données SEER

Création de Grade : le grade d'un cancer

Pour ces travaux, seules les variables suivantes, relatives à la définition du grade, sont utilisées pour créer une version élémentaire du grade de chaque cancer :

- `Grade..thru.2017.` : grade basé sur le répertoire CIM-0 pour les cancers reportés jusqu'à 2017 inclus ;
- `Grade.Clinical..2018..` : grade attribué au cours de la période clinique pour les cancers enregistrés à compter de 2018.

A l'aide des variables susmentionnées, la variable `Grade` a été conçue de la manière suivante :

- pour les tumeurs enregistrées avant 2018, `Grade = Grade..thru.2017.` ;
- pour celles reportées après 2018, `Grade = Grade.Clinical..2018..`

Ainsi, `Grade` repose simplement sur la concaténation des variables du grade différentes dans le temps.

Création de Stade : le stade d'un cancer

Le stade de chaque cancer est renseigné par les variables suivantes dans la base SEER extraite :

- `Summary.stage.2000..1998.2017.` : stade enregistré pour l'ensemble des tumeurs reportées entre 1998 et 2017. Il est basé sur les termes anglais cités précédemment ;
- `Combined.Summary.Stage..2004..` : stade renseigné pour les tumeurs à compter de 2004 avec les termes anglais également.

A partir de ces deux variables et du manuel SEER (ADAMO M, GROVES C, DICKIE L, RUHL J, 2021), la variable `Stade` est élaborée avec les conditions énoncées ci-après :

- `Grade = Summary.stage.2000..1998.2017.` ;
- puis si `Grade` est vide, alors `Grade = Combined.Summary.Stage..2004..`

Création de la variable Radiothérapie

La variable `Radiothérapie` est conçue à l'aide des variables `Radiation.recode` et `RX.Summ..Surg.R ad.Seq` selon les conditions suivantes :

- `Radiothérapie = Radiation.recode` ;
- pour les cas où `Radiation.recode = Recommended, unknown if administered` : si `RX.Summ..Surg.Rad.Seq` indique qu'il y a eu une radiothérapie alors `Radiothérapie = Oui`.

Création de la variable Chirurgie

La variable **Chirurgie** repose sur les variables **Reason.no.cancer.directed.surgery**, **RX.Summ..Surg.Rad.Seq** et **RX.Summ..Systemic.Sur.Seq**. Le processus de conception est le suivant :

- **Chirurgie** = **Reason.no.cancer.directed.surgery** ;
- pour les cas où **Reason.no.cancer.directed.surgery** = *Recommended, unknown if administered* : si **RX.Summ..Surg.Rad.Seq** et **RX.Summ..Systemic.Sur.Seq** signalent qu'il y a eu une chirurgie alors **Chirurgie** = *Oui* ;
- pour les lignes où **Reason.no.cancer.directed.surgery** = *Recommended, unknown if administered* et **RX.Summ..Surg.Rad.Seq** = *No radiation and/or cancer – directed surgery* : si la nouvelle variable **Radiotherapie** indique qu'il y a bien une radiothérapie en place, alors **Chirurgie** = *Non* ;
- pour les lignes où **Reason.no.cancer.directed.surgery** = *Recommended, unknown if administered* et **RX.Summ..Systemic.Sur.Seq** = *No systemic therapy and/or surgical procedure* : si la nouvelle variable **Chimiotherapie** indique qu'il y a bien une chimiothérapie en place, alors **Chirurgie** = *Non*.

Création de la variable **Sortie** : le statut vital du patient

La variable **Sortie** est alors construite telle que :

- pour les lignes où **Vital.status.recode..study.cutoff.used.** = *Alive* : **Sortie** = *Vivant* ;
- pour les lignes où **SEER.other.cause.of.death.classification** = *Dead (attributable to causes other than this cancer dx)* : **Sortie** = *Censure* pour préciser que le patient est sorti de l'échantillon de la population observée suite à son décès non attribué au cancer recensé ;
- pour les lignes où **SEER.cause.specific.death.classification** = *Dead (attributable to this cancer dx)* : **Sortie** = *Cancer* pour indiquer que le patient est décédé du cancer recensé.

Annexe des tableaux

TABLE 4.12 – Dictionnaire des 48 variables sélectionnées

Variables	
Age recode with ≥ 1 year olds	Months from diagnosis to treatment
Sex	Tumor Size Summary (2016+)
Year of diagnosis	CS tumor size (2004-2015)
Site recode ICD-O-3/WHO 2008	COD to site recode
Behavior code ICD-O-3	SEER cause-specific death classification
TNM 7/CS v0204+ Schema recode	SEER other cause of death classification
Primary Site - labeled	Survival months
Histologic Type ICD-O-3	Survival months flag
Grade (thru 2017)	COD to site rec KM
Grade Clinical (2018+)	Vital status recode (study cutoff used)
Grade Pathological (2018+)	Sequence number
Laterality	First malignant primary indicator
Diagnostic Confirmation	Primary by international rules
ICD-O-3 Hist/behav, malignant	Record number recode
Histology recode - broad groupings	Total number of in situ/malignant tumors for patient
Site recode - rare tumors	Age recode with single ages and 100+
Combined Summary Stage (2004+)	Month of diagnosis
Summary stage 2000 (1998-2017)	Month of diagnosis recode
RX Summ-Surg Prim Site (1998+)	SS seq # - mal+ins (most detail)
RX Summ-Surg/Rad Seq	Site - mal+ins (most detail)
Reason no cancer-directed surgery	Site - malignant (most detail)
Radiation recode	Site - malignant (least detail)
Chemotherapy recode (yes, no/unk)	Patient ID
RX Summ-Systemic/Sur Seq	Type of Reporting Source

TABLE 4.13 – Création de la variable Cancer

Site.recode.ICD.O.3.WHO.2008	Cancer (regroupement et traduction)
Breast	Sein
Prostate	Prostate
Lung and Bronchus	Poumon et bronche
Melanoma of the Skin	Peau hors basale et squameuse
Urinary Bladder	Vessie
Miscellaneous	Divers
Kidney and Renal Pelvis	Rein et pelvis renal
Corpus Uteri	Corpus et uterus
NHL - Nodal	Lymphome
Pancreas	Pancreas
Thyroid	Systeme endocrinien
Rectum	Colon et rectum
Sigmoid Colon	Colon et rectum
Liver	Foie
Stomach	Estomac
Cecum	Colon et rectum
Ovary	Ovaire
NHL - Extranodal	Lymphome
Myeloma	Myelome
Brain	Cerveau et autre systeme nerveux
Ascending Colon	Colon et rectum
Chronic Lymphocytic Leukemia	Leucemie
Esophagus	Oesophage
Cervix Uteri	Col de l'uterus
Acute Myeloid Leukemia	Leucemie
Tongue	Cavite buccale et pharynx
Soft Tissue including Heart	Tissus mous dont le coeur
Larynx	Systeme respiratoire a l'exclusion des poumons et des bronches
Rectosigmoid Junction	Colon et rectum
Transverse Colon	Colon et rectum
Testis	Systeme genital masculin a l'exclusion de la prostate
Hodgkin - Nodal	Lymphome
Small Intestine	Intestin grele
Large Intestine, NOS	Colon et rectum
Tonsil	Cavite buccale et pharynx
Other Non-Epithelial Skin	Peau hors basale et squameuse
Descending Colon	Colon et rectum
Other Biliary	Vesicule biliaire et autres voies biliaires
Anus, Anal Canal and Anorectum	Anus, canal anal et anorectum
Chronic Myeloid Leukemia	Leucemie
Acute Lymphocytic Leukemia	Leucemie
Gum and Other Mouth	Cavite buccale et pharynx
Hepatic Flexure	Colon et rectum

Site.recode.ICD.O.3.WHO.2008	Cancer (regroupement et traduction)
Vulva	Systeme genital feminin a l'exclusion du col de l'uterus, de l'ovaire et du corps et de l'uterus
Salivary Gland	Cavite buccale et pharynx
Gallbladder	Vesicule biliare et autres voies biliaires
Appendix	Colon et rectum
Intrahepatic Bile Duct	Voie biliare intrahepatique
Splenic Flexure	Colon et rectum
Mesothelioma	Mesotheliome
Bones and Joints	Os et articulations
Eye and Orbit	Oeil et orbite
Other Endocrine including Thymus	Systeme endocrinien
Lip	Cavite buccale et pharynx
Nose, Nasal Cavity and Middle Ear	Systeme respiratoire a l'exclusion des poumons et des bronches
Other Female Genital Organs	Systeme genital feminin a l'exclusion du col de l'uterus, de l'ovaire et du corps et de l'uterus
Nasopharynx	Cavite buccale et pharynx
Hypopharynx	Cavite buccale et pharynx
Other Digestive Organs	Autres organes digestifs
Peritoneum, Omentum and Mesentery	Peritoine, epiploon et mesentere
Floor of Mouth	Cavite buccale et pharynx
Ureter	Systeme urinaire a l'exclusion de la vessie et des reins et du bassinnet
Kaposi Sarcoma	Sarcome de Kaposi
Aleukemic, Subleukemic and NOS	Leucemie
Oropharynx	Cavite buccale et pharynx
Uterus, NOS	Corpus et uterus
Other Lymphocytic Leukemia	Leucemie
Cranial Nerves Other Nervous System	Cerveau et autre systeme nerveux
Vagina	Systeme genital feminin a l'exclusion du col de l'uterus, de l'ovaire et du corps et de l'uterus
Retroperitoneum	Retroperitoneal
Penis	Systeme genital masculin a l'exclusion de la prostate
Other Urinary Organs	Systeme urinaire a l'exclusion de la vessie et des reins et du bassinnet
Other Acute Leukemia	Leucemie
Other Oral Cavity and Pharynx	Cavite buccale et pharynx
Acute Monocytic Leukemia	Leucemie
Trachea, Mediastinum and Other Respiratory Organs	Systeme respiratoire a l'exclusion des poumons et des bronches
Other Myeloid/Monocytic Leukemia	Leucemie
Other Male Genital Organs	Systeme genital masculin a l'exclusion de la prostate
Hodgkin - Extranodal	Lymphome
Pleura	Systeme respiratoire a l'exclusion des poumons et des bronches

TABLE 4.14 – Création de la variable **Grade**

Modalités des variables initiales	Grade
1	Grade I
2	Grade II
3	Grade III
4	Grade IV
5	T-cell ; T-precursor cell
6	B-cell ; B-precursor cell
7	Null cell ; Non-T-non-B
8	NK cell (natural killer cell)
9	Unknown
Well differentiated ; Grade I	Grade I
Moderately differentiated ; Grade II	Grade II
Poorly differentiated ; Grade III	Grade III
Undifferentiated ; anaplastic ; Grade IV	Grade IV
T-cell	T-cell ; T-precursor cell
Null cell ; non T-non B	Null cell ; Non-T-non-B
NK cell ; natural killer cell (1995+)	NK cell (natural killer cell)
B-cell ; pre-B ; B-precursor	B-cell ; B-precursor cell
Unknown	Unknown
A	Grade I
B	Grade II
C	Grade III
D	Grade IV
E	Grade III
H	Grade IV
L	Grade II
M	Grade III
S	Grade IV

TABLE 4.15 – Table d’extra-mortalité du cancer non métastasé diagnostiqué et traité pour les femmes selon l’âge à la fin du traitement thérapeutique et l’ancienneté en rémission

	0	1	2	3	4	5	6	7	8	9	10
21	1.63%	1.16%	0.80%	0.54%	0.36%	0.25%	0.20%	0.19%	0.20%	0.23%	0.26%
22	1.50%	1.10%	0.80%	0.57%	0.42%	0.32%	0.27%	0.26%	0.28%	0.32%	0.37%
23	1.46%	1.09%	0.81%	0.60%	0.45%	0.36%	0.31%	0.31%	0.33%	0.38%	0.45%
24	1.53%	1.14%	0.84%	0.62%	0.47%	0.38%	0.34%	0.35%	0.38%	0.44%	0.50%
25	1.71%	1.24%	0.90%	0.65%	0.50%	0.41%	0.38%	0.39%	0.43%	0.47%	0.51%
26	1.88%	1.36%	0.98%	0.72%	0.55%	0.46%	0.43%	0.44%	0.47%	0.49%	0.49%
27	1.95%	1.45%	1.07%	0.81%	0.64%	0.54%	0.50%	0.48%	0.49%	0.48%	0.45%
28	1.91%	1.49%	1.17%	0.93%	0.75%	0.64%	0.56%	0.52%	0.48%	0.45%	0.41%
29	1.85%	1.53%	1.26%	1.04%	0.86%	0.73%	0.62%	0.54%	0.47%	0.42%	0.36%
30	1.82%	1.56%	1.32%	1.12%	0.94%	0.80%	0.67%	0.56%	0.47%	0.39%	0.32%
31	1.85%	1.58%	1.35%	1.15%	0.98%	0.84%	0.71%	0.60%	0.49%	0.39%	0.28%
32	1.91%	1.62%	1.37%	1.17%	1.00%	0.86%	0.73%	0.62%	0.51%	0.39%	0.26%
33	1.95%	1.63%	1.38%	1.17%	1.00%	0.86%	0.74%	0.63%	0.52%	0.40%	0.25%
34	1.92%	1.62%	1.37%	1.17%	1.00%	0.87%	0.74%	0.63%	0.52%	0.40%	0.26%
35	1.87%	1.61%	1.38%	1.19%	1.02%	0.87%	0.74%	0.63%	0.52%	0.41%	0.29%
36	1.88%	1.63%	1.42%	1.22%	1.05%	0.89%	0.76%	0.64%	0.53%	0.43%	0.34%
37	1.96%	1.70%	1.46%	1.26%	1.08%	0.92%	0.78%	0.66%	0.56%	0.47%	0.40%
38	2.00%	1.73%	1.48%	1.27%	1.09%	0.93%	0.79%	0.68%	0.59%	0.51%	0.46%
39	1.95%	1.69%	1.46%	1.25%	1.07%	0.91%	0.78%	0.68%	0.60%	0.54%	0.51%
40	1.86%	1.64%	1.42%	1.23%	1.05%	0.89%	0.76%	0.66%	0.58%	0.54%	0.54%
41	1.91%	1.67%	1.44%	1.23%	1.04%	0.87%	0.74%	0.63%	0.56%	0.53%	0.53%
42	2.03%	1.73%	1.46%	1.22%	1.02%	0.85%	0.72%	0.62%	0.55%	0.52%	0.52%
43	2.09%	1.74%	1.45%	1.20%	1.00%	0.84%	0.72%	0.63%	0.57%	0.53%	0.51%
44	2.08%	1.72%	1.42%	1.18%	0.99%	0.85%	0.74%	0.65%	0.59%	0.54%	0.49%
45	2.06%	1.72%	1.44%	1.21%	1.03%	0.88%	0.76%	0.66%	0.59%	0.53%	0.47%
46	2.12%	1.80%	1.52%	1.27%	1.07%	0.90%	0.76%	0.65%	0.57%	0.51%	0.47%
47	2.26%	1.87%	1.55%	1.29%	1.07%	0.90%	0.77%	0.66%	0.58%	0.52%	0.47%
48	2.43%	1.96%	1.58%	1.29%	1.06%	0.90%	0.78%	0.69%	0.62%	0.55%	0.47%
49	2.49%	1.99%	1.60%	1.30%	1.08%	0.92%	0.80%	0.72%	0.64%	0.57%	0.47%
50	2.47%	2.00%	1.62%	1.33%	1.11%	0.95%	0.83%	0.73%	0.64%	0.56%	0.45%
51	2.54%	2.04%	1.65%	1.36%	1.14%	0.97%	0.85%	0.74%	0.64%	0.53%	0.39%
52	2.68%	2.14%	1.73%	1.41%	1.18%	1.01%	0.87%	0.76%	0.65%	0.52%	0.36%
53	2.86%	2.28%	1.82%	1.48%	1.22%	1.04%	0.89%	0.77%	0.66%	0.52%	0.35%
54	3.11%	2.44%	1.92%	1.53%	1.25%	1.05%	0.90%	0.78%	0.66%	0.53%	0.36%
55	3.30%	2.58%	2.03%	1.61%	1.30%	1.07%	0.91%	0.78%	0.66%	0.53%	0.37%
56	3.39%	2.65%	2.07%	1.63%	1.32%	1.09%	0.92%	0.79%	0.67%	0.53%	0.35%
57	3.41%	2.62%	2.03%	1.60%	1.31%	1.10%	0.95%	0.82%	0.69%	0.52%	0.28%
58	3.49%	2.66%	2.05%	1.61%	1.31%	1.11%	0.97%	0.84%	0.70%	0.51%	0.22%
59	3.69%	2.78%	2.12%	1.64%	1.32%	1.11%	0.97%	0.84%	0.71%	0.52%	0.23%
60	3.83%	2.88%	2.18%	1.69%	1.35%	1.13%	0.98%	0.86%	0.74%	0.56%	0.30%
61	3.87%	2.93%	2.23%	1.72%	1.38%	1.15%	1.00%	0.90%	0.78%	0.63%	0.39%
62	3.94%	2.95%	2.23%	1.73%	1.40%	1.19%	1.07%	0.97%	0.86%	0.70%	0.45%
63	4.09%	3.03%	2.26%	1.74%	1.41%	1.22%	1.12%	1.04%	0.96%	0.81%	0.54%
64	4.16%	3.07%	2.30%	1.77%	1.44%	1.25%	1.16%	1.10%	1.04%	0.91%	0.67%
65	4.09%	3.05%	2.30%	1.79%	1.46%	1.27%	1.18%	1.12%	1.06%	0.96%	0.78%
66	4.05%	3.07%	2.35%	1.84%	1.50%	1.29%	1.17%	1.10%	1.03%	0.94%	0.78%

TABLE 4.16 – Table d’extra-mortalité du cancer non métastasé diagnostiqué et traité pour les hommes selon l’âge à la fin du traitement thérapeutique et l’ancienneté en rémission

	0	1	2	3	4	5	6	7	8	9	10
21	2.74%	1.72%	1.06%	0.70%	0.53%	0.50%	0.52%	0.50%	0.38%	0.07%	0.00%
22	2.52%	1.77%	1.24%	0.89%	0.66%	0.53%	0.45%	0.38%	0.27%	0.10%	0.00%
23	2.37%	1.80%	1.34%	0.99%	0.73%	0.55%	0.41%	0.31%	0.23%	0.14%	0.04%
24	2.28%	1.78%	1.37%	1.03%	0.77%	0.56%	0.41%	0.30%	0.23%	0.18%	0.16%
25	2.20%	1.75%	1.37%	1.05%	0.79%	0.58%	0.42%	0.31%	0.24%	0.20%	0.20%
26	2.16%	1.73%	1.37%	1.06%	0.81%	0.61%	0.45%	0.33%	0.25%	0.20%	0.18%
27	2.19%	1.72%	1.34%	1.04%	0.81%	0.63%	0.50%	0.39%	0.29%	0.20%	0.09%
28	2.32%	1.74%	1.31%	1.01%	0.81%	0.67%	0.57%	0.48%	0.37%	0.21%	0.00%
29	2.53%	1.80%	1.31%	1.00%	0.82%	0.72%	0.67%	0.60%	0.47%	0.24%	0.00%
30	2.73%	1.88%	1.33%	1.00%	0.83%	0.77%	0.74%	0.70%	0.57%	0.30%	0.00%
31	2.81%	1.95%	1.38%	1.04%	0.86%	0.79%	0.77%	0.73%	0.62%	0.37%	0.00%
32	2.81%	2.01%	1.46%	1.10%	0.90%	0.80%	0.75%	0.70%	0.61%	0.43%	0.10%
33	2.76%	2.04%	1.52%	1.16%	0.92%	0.79%	0.71%	0.66%	0.59%	0.49%	0.30%
34	2.74%	2.06%	1.55%	1.18%	0.93%	0.77%	0.68%	0.63%	0.59%	0.54%	0.45%
35	2.77%	2.08%	1.56%	1.18%	0.93%	0.76%	0.67%	0.62%	0.59%	0.55%	0.48%
36	2.85%	2.11%	1.57%	1.18%	0.93%	0.76%	0.67%	0.61%	0.57%	0.51%	0.40%
37	2.99%	2.20%	1.62%	1.21%	0.94%	0.77%	0.68%	0.61%	0.54%	0.43%	0.25%
38	3.21%	2.34%	1.71%	1.27%	0.98%	0.81%	0.70%	0.62%	0.51%	0.35%	0.08%
39	3.40%	2.46%	1.79%	1.33%	1.03%	0.84%	0.72%	0.62%	0.49%	0.29%	0.00%
40	3.51%	2.56%	1.87%	1.38%	1.07%	0.86%	0.73%	0.62%	0.49%	0.28%	0.00%
41	3.68%	2.67%	1.94%	1.43%	1.09%	0.88%	0.74%	0.63%	0.50%	0.29%	0.00%
42	3.83%	2.76%	1.99%	1.45%	1.10%	0.89%	0.76%	0.66%	0.53%	0.34%	0.02%
43	3.93%	2.83%	2.03%	1.48%	1.13%	0.91%	0.78%	0.69%	0.58%	0.40%	0.09%
44	4.06%	2.95%	2.14%	1.57%	1.20%	0.97%	0.84%	0.74%	0.64%	0.47%	0.18%
45	4.37%	3.15%	2.26%	1.65%	1.25%	1.02%	0.89%	0.80%	0.70%	0.54%	0.25%
46	4.76%	3.33%	2.32%	1.64%	1.23%	1.01%	0.91%	0.86%	0.78%	0.62%	0.29%
47	4.95%	3.40%	2.31%	1.60%	1.18%	0.98%	0.90%	0.88%	0.82%	0.65%	0.28%
48	4.87%	3.34%	2.27%	1.58%	1.17%	0.96%	0.89%	0.86%	0.79%	0.60%	0.21%
49	4.62%	3.21%	2.22%	1.56%	1.16%	0.95%	0.86%	0.80%	0.71%	0.51%	0.14%
50	4.38%	3.07%	2.15%	1.53%	1.15%	0.95%	0.84%	0.77%	0.67%	0.46%	0.09%
51	4.50%	3.14%	2.18%	1.55%	1.17%	0.97%	0.88%	0.81%	0.70%	0.48%	0.06%
52	4.83%	3.29%	2.23%	1.56%	1.17%	0.99%	0.92%	0.88%	0.77%	0.51%	0.02%
53	5.03%	3.39%	2.27%	1.56%	1.18%	1.00%	0.95%	0.92%	0.81%	0.54%	0.00%
54	5.14%	3.43%	2.28%	1.56%	1.18%	1.01%	0.97%	0.95%	0.84%	0.55%	0.00%
55	5.22%	3.47%	2.30%	1.58%	1.20%	1.04%	1.01%	0.99%	0.88%	0.58%	0.00%
56	5.27%	3.50%	2.30%	1.58%	1.20%	1.06%	1.03%	1.02%	0.92%	0.61%	0.00%
57	5.37%	3.54%	2.31%	1.56%	1.18%	1.04%	1.02%	1.03%	0.94%	0.64%	0.03%
58	5.44%	3.55%	2.28%	1.52%	1.14%	1.01%	1.01%	1.03%	0.96%	0.69%	0.10%
59	5.33%	3.51%	2.28%	1.53%	1.15%	1.00%	1.00%	1.02%	0.96%	0.72%	0.20%
60	5.27%	3.51%	2.32%	1.58%	1.18%	1.02%	0.98%	0.99%	0.93%	0.72%	0.26%
61	5.35%	3.59%	2.39%	1.64%	1.22%	1.03%	0.98%	0.97%	0.89%	0.66%	0.18%
62	5.31%	3.55%	2.36%	1.61%	1.21%	1.04%	1.00%	0.98%	0.88%	0.60%	0.06%
63	5.31%	3.52%	2.32%	1.59%	1.21%	1.06%	1.04%	1.03%	0.92%	0.62%	0.01%
64	5.25%	3.47%	2.28%	1.57%	1.21%	1.08%	1.07%	1.07%	0.97%	0.68%	0.08%
65	5.04%	3.37%	2.26%	1.58%	1.24%	1.12%	1.11%	1.12%	1.05%	0.79%	0.25%
66	5.14%	3.49%	2.37%	1.69%	1.32%	1.19%	1.18%	1.19%	1.14%	0.93%	0.47%

TABLE 4.17 – Table d’extra-mortalité du cancer non métastasé diagnostiqué et traité pour les femmes selon l’âge à la fin du traitement thérapeutique et l’ancienneté en rémission - **Age de fin de traitement thérapeutique vieilli de 6 mois**

	0	1	2	3	4	5	6	7	8	9	10
21	0.49%	0.40%	0.32%	0.26%	0.21%	0.18%	0.16%	0.16%	0.17%	0.20%	0.24%
22	0.80%	0.65%	0.51%	0.41%	0.32%	0.26%	0.23%	0.21%	0.21%	0.23%	0.26%
23	1.02%	0.80%	0.62%	0.49%	0.39%	0.33%	0.29%	0.27%	0.26%	0.26%	0.26%
24	1.17%	0.87%	0.66%	0.51%	0.42%	0.37%	0.34%	0.33%	0.31%	0.28%	0.22%
25	1.31%	0.93%	0.68%	0.52%	0.43%	0.40%	0.39%	0.38%	0.36%	0.30%	0.17%
26	1.45%	1.01%	0.71%	0.54%	0.45%	0.42%	0.43%	0.43%	0.40%	0.31%	0.12%
27	1.54%	1.08%	0.77%	0.59%	0.49%	0.46%	0.45%	0.44%	0.40%	0.29%	0.09%
28	1.56%	1.14%	0.86%	0.68%	0.57%	0.51%	0.48%	0.45%	0.39%	0.27%	0.07%
29	1.58%	1.22%	0.97%	0.79%	0.67%	0.59%	0.53%	0.47%	0.38%	0.24%	0.05%
30	1.60%	1.29%	1.06%	0.90%	0.77%	0.67%	0.59%	0.50%	0.38%	0.23%	0.02%
31	1.60%	1.34%	1.14%	0.98%	0.86%	0.75%	0.65%	0.54%	0.40%	0.23%	0.00%
32	1.62%	1.36%	1.17%	1.02%	0.91%	0.81%	0.71%	0.59%	0.44%	0.25%	0.00%
33	1.68%	1.39%	1.19%	1.04%	0.92%	0.83%	0.74%	0.63%	0.48%	0.28%	0.01%
34	1.76%	1.43%	1.20%	1.03%	0.91%	0.82%	0.74%	0.64%	0.51%	0.32%	0.07%
35	1.80%	1.47%	1.22%	1.04%	0.90%	0.80%	0.72%	0.63%	0.52%	0.37%	0.16%
36	1.80%	1.48%	1.24%	1.05%	0.91%	0.79%	0.70%	0.61%	0.52%	0.41%	0.28%
37	1.73%	1.48%	1.27%	1.09%	0.94%	0.80%	0.70%	0.60%	0.53%	0.47%	0.42%
38	1.69%	1.51%	1.32%	1.15%	0.98%	0.83%	0.70%	0.60%	0.54%	0.52%	0.54%
39	1.69%	1.52%	1.34%	1.15%	0.98%	0.83%	0.70%	0.60%	0.55%	0.54%	0.59%
40	1.70%	1.50%	1.30%	1.12%	0.95%	0.81%	0.69%	0.60%	0.55%	0.53%	0.55%
41	1.72%	1.47%	1.26%	1.08%	0.92%	0.80%	0.69%	0.61%	0.55%	0.51%	0.47%
42	1.77%	1.48%	1.24%	1.06%	0.91%	0.80%	0.71%	0.63%	0.55%	0.47%	0.38%
43	1.79%	1.48%	1.24%	1.05%	0.91%	0.80%	0.72%	0.64%	0.55%	0.45%	0.32%
44	1.77%	1.47%	1.23%	1.05%	0.91%	0.80%	0.71%	0.63%	0.55%	0.44%	0.31%
45	1.72%	1.46%	1.25%	1.07%	0.93%	0.81%	0.71%	0.62%	0.54%	0.45%	0.36%
46	1.74%	1.52%	1.31%	1.12%	0.96%	0.82%	0.69%	0.60%	0.52%	0.46%	0.43%
47	1.86%	1.61%	1.38%	1.17%	0.97%	0.81%	0.68%	0.58%	0.51%	0.48%	0.48%
48	1.97%	1.67%	1.40%	1.17%	0.97%	0.81%	0.69%	0.59%	0.54%	0.51%	0.53%
49	2.02%	1.68%	1.40%	1.17%	0.98%	0.83%	0.71%	0.64%	0.59%	0.57%	0.57%
50	1.97%	1.67%	1.41%	1.19%	1.01%	0.86%	0.75%	0.67%	0.61%	0.58%	0.57%
51	2.00%	1.70%	1.44%	1.22%	1.04%	0.89%	0.77%	0.67%	0.60%	0.55%	0.52%
52	2.13%	1.79%	1.50%	1.26%	1.07%	0.91%	0.78%	0.67%	0.59%	0.52%	0.46%
53	2.25%	1.88%	1.57%	1.31%	1.10%	0.92%	0.78%	0.67%	0.58%	0.51%	0.45%
54	2.38%	2.00%	1.67%	1.39%	1.15%	0.95%	0.79%	0.67%	0.58%	0.52%	0.50%
55	2.56%	2.15%	1.78%	1.47%	1.20%	0.98%	0.80%	0.67%	0.59%	0.54%	0.54%
56	2.64%	2.24%	1.87%	1.54%	1.25%	1.01%	0.82%	0.68%	0.58%	0.54%	0.55%
57	2.66%	2.24%	1.86%	1.54%	1.26%	1.02%	0.83%	0.69%	0.59%	0.52%	0.49%
58	2.69%	2.21%	1.82%	1.50%	1.24%	1.03%	0.86%	0.72%	0.61%	0.50%	0.39%
59	2.82%	2.27%	1.84%	1.50%	1.24%	1.03%	0.87%	0.73%	0.61%	0.47%	0.31%
60	2.89%	2.32%	1.87%	1.52%	1.25%	1.04%	0.87%	0.74%	0.61%	0.48%	0.33%
61	2.89%	2.37%	1.93%	1.58%	1.29%	1.07%	0.90%	0.76%	0.65%	0.54%	0.42%
62	3.05%	2.45%	1.97%	1.60%	1.32%	1.11%	0.96%	0.84%	0.74%	0.65%	0.54%
63	3.22%	2.54%	2.02%	1.63%	1.34%	1.15%	1.01%	0.92%	0.85%	0.79%	0.72%
64	3.20%	2.56%	2.05%	1.66%	1.37%	1.16%	1.03%	0.95%	0.92%	0.92%	0.93%
65	3.05%	2.51%	2.05%	1.68%	1.40%	1.19%	1.05%	0.98%	0.96%	1.01%	1.09%
66	3.17%	2.56%	2.07%	1.70%	1.43%	1.24%	1.13%	1.07%	1.05%	1.05%	1.07%

TABLE 4.18 – Table d’extra-mortalité du cancer non métastasé diagnostiqué et traité pour les hommes selon l’âge à la fin du traitement thérapeutique et l’ancienneté en rémission - **Age de fin de traitement thérapeutique vieilli de 6 mois**

	0	1	2	3	4	5	6	7	8	9	10
21	1.20%	0.79%	0.55%	0.42%	0.38%	0.37%	0.35%	0.28%	0.12%	0.00%	0.00%
22	1.71%	1.25%	0.92%	0.69%	0.54%	0.44%	0.36%	0.27%	0.15%	0.00%	0.00%
23	1.95%	1.48%	1.11%	0.84%	0.63%	0.48%	0.37%	0.28%	0.19%	0.09%	0.00%
24	2.00%	1.54%	1.17%	0.88%	0.66%	0.50%	0.38%	0.30%	0.24%	0.19%	0.14%
25	1.94%	1.51%	1.15%	0.88%	0.66%	0.50%	0.39%	0.32%	0.28%	0.26%	0.27%
26	1.84%	1.46%	1.13%	0.87%	0.66%	0.50%	0.39%	0.32%	0.29%	0.30%	0.34%
27	1.74%	1.43%	1.15%	0.90%	0.69%	0.52%	0.40%	0.32%	0.28%	0.30%	0.37%
28	1.69%	1.43%	1.18%	0.95%	0.74%	0.56%	0.42%	0.32%	0.27%	0.27%	0.34%
29	1.66%	1.45%	1.24%	1.02%	0.81%	0.62%	0.46%	0.34%	0.26%	0.24%	0.29%
30	1.70%	1.49%	1.28%	1.07%	0.87%	0.69%	0.53%	0.40%	0.30%	0.24%	0.22%
31	1.81%	1.53%	1.28%	1.07%	0.89%	0.73%	0.59%	0.47%	0.36%	0.27%	0.18%
32	1.96%	1.58%	1.28%	1.04%	0.87%	0.73%	0.62%	0.53%	0.43%	0.33%	0.19%
33	2.10%	1.62%	1.27%	1.01%	0.83%	0.71%	0.63%	0.57%	0.49%	0.39%	0.25%
34	2.15%	1.65%	1.28%	1.01%	0.84%	0.72%	0.64%	0.59%	0.53%	0.44%	0.32%
35	2.16%	1.68%	1.31%	1.05%	0.87%	0.74%	0.66%	0.60%	0.55%	0.47%	0.35%
36	2.24%	1.70%	1.31%	1.04%	0.86%	0.74%	0.67%	0.62%	0.55%	0.46%	0.31%
37	2.38%	1.76%	1.32%	1.02%	0.83%	0.72%	0.65%	0.60%	0.53%	0.41%	0.21%
38	2.48%	1.83%	1.36%	1.03%	0.83%	0.70%	0.62%	0.55%	0.46%	0.32%	0.09%
39	2.56%	1.90%	1.42%	1.09%	0.86%	0.72%	0.62%	0.53%	0.41%	0.24%	0.00%
40	2.67%	1.99%	1.50%	1.15%	0.92%	0.77%	0.65%	0.54%	0.40%	0.19%	0.00%
41	2.83%	2.10%	1.57%	1.20%	0.96%	0.81%	0.70%	0.58%	0.42%	0.18%	0.00%
42	3.02%	2.18%	1.60%	1.21%	0.96%	0.82%	0.72%	0.62%	0.46%	0.20%	0.00%
43	3.14%	2.23%	1.60%	1.19%	0.95%	0.81%	0.73%	0.64%	0.50%	0.24%	0.00%
44	3.10%	2.26%	1.66%	1.26%	1.01%	0.86%	0.76%	0.66%	0.52%	0.30%	0.00%
45	3.14%	2.36%	1.79%	1.39%	1.11%	0.93%	0.81%	0.70%	0.57%	0.38%	0.10%
46	3.30%	2.50%	1.90%	1.47%	1.18%	0.98%	0.85%	0.75%	0.64%	0.49%	0.27%
47	3.39%	2.55%	1.91%	1.46%	1.15%	0.96%	0.84%	0.76%	0.68%	0.57%	0.38%
48	3.41%	2.51%	1.85%	1.39%	1.09%	0.91%	0.80%	0.73%	0.67%	0.57%	0.39%
49	3.27%	2.42%	1.80%	1.35%	1.05%	0.86%	0.75%	0.67%	0.60%	0.50%	0.32%
50	3.18%	2.35%	1.75%	1.32%	1.04%	0.85%	0.74%	0.65%	0.56%	0.42%	0.20%
51	3.26%	2.38%	1.75%	1.32%	1.05%	0.89%	0.79%	0.70%	0.58%	0.39%	0.07%
52	3.43%	2.47%	1.80%	1.36%	1.10%	0.95%	0.86%	0.78%	0.65%	0.40%	0.00%
53	3.55%	2.53%	1.83%	1.38%	1.12%	0.98%	0.90%	0.82%	0.68%	0.41%	0.00%
54	3.65%	2.59%	1.87%	1.40%	1.12%	0.98%	0.89%	0.81%	0.67%	0.41%	0.00%
55	3.61%	2.61%	1.90%	1.44%	1.16%	1.00%	0.90%	0.81%	0.67%	0.43%	0.01%
56	3.62%	2.61%	1.91%	1.45%	1.17%	1.02%	0.93%	0.84%	0.70%	0.46%	0.05%
57	3.59%	2.61%	1.93%	1.47%	1.19%	1.03%	0.93%	0.85%	0.72%	0.49%	0.11%
58	3.67%	2.64%	1.92%	1.45%	1.17%	1.02%	0.94%	0.87%	0.76%	0.55%	0.17%
59	3.74%	2.66%	1.92%	1.43%	1.15%	1.00%	0.93%	0.88%	0.79%	0.60%	0.26%
60	3.63%	2.65%	1.94%	1.47%	1.17%	1.00%	0.90%	0.84%	0.76%	0.62%	0.36%
61	3.54%	2.67%	2.01%	1.53%	1.21%	0.99%	0.86%	0.78%	0.70%	0.59%	0.42%
62	3.61%	2.70%	2.01%	1.53%	1.20%	0.99%	0.85%	0.76%	0.67%	0.54%	0.33%
63	3.74%	2.73%	2.00%	1.51%	1.20%	1.01%	0.90%	0.82%	0.71%	0.53%	0.22%
64	3.68%	2.70%	2.00%	1.52%	1.23%	1.06%	0.96%	0.88%	0.77%	0.57%	0.24%
65	3.51%	2.59%	1.94%	1.50%	1.23%	1.08%	1.01%	0.96%	0.89%	0.75%	0.48%
66	3.67%	2.70%	2.02%	1.56%	1.28%	1.15%	1.11%	1.11%	1.12%	1.10%	0.99%