

Proposition d'élargissement de l'accès à l'emprunt en cas de pathologie grave à partir de l'open data¹

Version 1.5 du 28/10/2021

Édouard Debonneuil²

Marie Péju³

Frédéric Planchet⁴

ISFA - Laboratoire SAF^β

Université de Lyon - Université Claude Bernard Lyon 1

Résumé

En France, l'accès à l'emprunt nécessite l'obtention d'une assurance emprunteur et la présence d'une pathologie chez le demandeur peut être un motif de refus. Améliorer la connaissance des risques de santé et mutualiser davantage les risques sont deux pistes d'élargissement de l'accès à l'emprunt. Nous tentons d'analyser ces possibilités en nous appuyant sur l'open data et des scénarios de mutualisation des risques. Élargir l'accès à l'emprunt permettrait les projets de vie associés, ce qui serait bénéfique pour les patients et pour le développement de l'économie associée à ces projets.

1. INTRODUCTION	2
2. ASSURABILITÉ DE PATHOLOGIES GRAVES : APPORT DE L'OPEN DATA.....	3
a. Mortalité.....	3
i. Point de repères : mortalité pour diverses populations françaises.....	5
ii. Mortalité de la population générale en fonction de divers facteurs de risque	7
iii. Transposition à des emprunteurs	8
b. Incapacité	11
iv. Incidence et maintien en arrêt de travail parmi la population salariée.....	11
v. Transposition à des emprunteurs	12
c. Invalidité.....	13
vi. Incidence en invalidité parmi la population salariée.....	13
vii. Transposition à des emprunteurs	14
d. Calcul de primes	15
3. APPLICATION DEUX PATHOLOGIES, LE CANCER DU SEIN ET LE DIABÈTE	17
e. Cancer du sein	17
viii. Mortalité par facteur de risque.....	17
ix. Incapacité en fonction de l'âge	20
x. Invalidité en fonction de l'ancienneté du diagnostic.....	20
xi. Design du modèle.....	20
xii. Résultats.....	21
f. Diabète	24
xiii. Mortalité par type de diabète.....	24

¹ Ce travail a bénéficié du support du FUI ASPRET, cofinancé par BPSI et Sogecap.

² Édouard Debonneuil est gérant de ActuRx (<https://acturx.com/>).

³ Marie Péju est *data scientist* indépendante.

⁴ Frédéric Planchet est Professeur à l'ISFA et actuaire associé chez PRIM'ACT. Contact : frederic@planchet.net.

^β Univ Lyon, Université Claude Bernard Lyon 1, Institut de Science Financière et d'Assurances (ISFA), Laboratoire SAF EA2429, F-69366, LYON, France.

xiv. Incapacité en fonction de l'âge et de l'ancienneté du diagnostic.....	27
xv. Invalidité en fonction de l'ancienneté du diagnostic.....	28
xvi. Design du modèle.....	28
xvii. Résultats.....	30
4. CONCLUSION	31
5. RÉFÉRENCES.....	31

1. Introduction

Chaque année, des emprunteurs se voient refuser leur demande d'emprunt, pour l'achat d'un bien immobilier ou pour la création d'une entreprise, du fait de leur état de santé. L'impact humain est d'autant plus fort que ces patients bien-sûr n'ont pas choisi d'être malades et que l'impossibilité de compléter ses ressources financières par un prêt pour un projet de vie est aujourd'hui perçu comme une discrimination. Ces refus ont également un impact économique, au-delà du prêt lui-même, puisque ces projets de vie auraient contribué au développement de l'économie.

Une partie de ces refus vient de la connaissance insuffisante des risques de santé par les assureurs. Ce refus est la réponse de l'assureur face à un risque de perte financière jugé trop fort, du fait des risques de décès, incapacité ou invalidité. Chez les assureurs, les actuaires savent généralement tarifier les contrats sur base de statistiques des assurés ; mais pour des pathologies aujourd'hui refusées ils ne disposent pas des statistiques correspondantes. Trouver les informations externes n'est pas évident car pour des raisons de protection de la vie privée en matière de santé ils n'ont pas accès aux données permettant de construire ces statistiques à partir de leurs données d'assurance santé ni prévoyance, ni à partir des données nationales de santé. L'appréciation à dire d'expert des risques liés aux pathologies n'est pas évidente : les pathologies et facteurs de risque associés sont nombreux, les avancées scientifiques et les évolutions comportementales modifient les niveaux de risque. Face à ce manque de connaissance, les assureurs prennent des marges de prudence (et ils en sont tenus pour éviter des faillites et le non-respect de leurs garanties) et n'ont en pratique pas d'autre choix que de remplacer des tarifs trop élevés par un refus pour éviter un risque d'image important (et ils sont tenus de ne pas dépasser un taux d'usure).

Par ailleurs, mutualiser les risques élevés avec des risques plus faibles se heurte aux comportements populationnels. Dans le cadre existant, un assureur qui créerait un contrat emprunteur sans sélection médicale, donc maximale inclusif vis-à-vis des pathologies, risquerait de voir affluer les patients qui ont vu leur demande de prêt refusée ailleurs, sans que cet afflux ne soit compensé par un afflux suffisant de personnes à risques faibles : l'expérience tournerait court.

Ainsi, le cadre existant ne favorise pas la connaissance fine des risques associés à des pathologies, ni la mutualisation des risques de santé élevés avec des risques plus faibles. Nous tentons ici d'utiliser l'open data et nos connaissances en assurance pour trouver des voies de sortie de la situation actuelle. Nous prenons comme exemples d'application le cancer du sein et le diabète. D'un point de vue des risques de mortalité, incapacité, invalidité, nous commençons par présenter les approches inclusives. Nous poursuivons en apportant notre technicité sur l'utilisation de l'open data pour améliorer la connaissance des risques,

avec divers degrés possibles d'inclusion des patients. Nous espérons que ces éléments pourront contribuer à élargir l'accès à l'emprunt.

2. Assurabilité de pathologies graves : apport de l'open data

Affiner la connaissance du risque d'une pathologie ou d'un ensemble de pathologies permet de réduire l'incertitude sur les risques et d'élargir l'accès à l'emprunt pour tout ou partie des formes de la pathologie considérée. Comme indiqué en introduction, devant la difficulté à obtenir des données fines pour affiner la connaissance du risque, nous nous appuyons ici sur l'open data.

Il s'agit de segmenter le tarif en fonction des caractéristiques de risque de la pathologie, d'une manière qui couvre plus de cas qu'actuellement, mais de refuser la demande lorsque le risque est élevé et le tarif est peu acceptable. Cette approche ouvre donc moins l'accès à l'emprunt que la mutualisation des risques élevés avec les risques moins élevés ; mais peut être mise en place dans le cadre existant. Elle nécessite une connaissance fine du risque, ce qui *via* l'open data nécessite des modélisations et recherches de données bien plus sophistiquées.

Notons d'emblée que le travail d'actuariat est complexe. Internet foisonne d'informations erronées, décalées par rapport aux besoins, ou présentées avec un contexte imprécis. Le travail de filtrage et d'assemblage est donc fastidieux et nécessite une multiplicité d'angles pour conforter les résultats. Alors qu'un actuariaire aurait tendance à d'abord définir un modèle puis à chercher dans les données clients les informations nécessaires au modèle, les données disponibles sur Internet nécessitent une modélisation adaptée.

Nous commençons par préciser l'obtention de données adéquates pour la mortalité, garantie emprunteur la plus importante, puis donnons des sources et une méthodologie associée pour estimer les risques d'incapacité et d'invalidité de façon moins précise mais avec moins d'effort. Nous ne modélisons pas le fait que, que d'un point de vue comportemental, les patients avec des pathologies lourdes sont potentiellement moins enclins à demander un prêt ; de ce point de vue le cadre de modélisation présenté ici est prudent.

a. Mortalité

Pour une cinquantaine de pathologies, les fiches sur les pathologies du site Ameli (Ameli, 2018a) fournissent une information sur le niveau de mortalité dans la population générale. Elles donnent (en haut de leur deuxième page) les taux de mortalité par tranche d'âge.

En prenant comme exemple le cancer du sein, la Table 1 indique les risques de mortalité selon Ameli, en combinant les cancers du sein actifs et sous surveillance (AMELI, 2018b et AMELI, 2018c) ; selon Ameli, un cancer est dit *actif* lorsque le cancer a nécessité une prise en charge autre que pour un bilan au cours des 2 dernières années ; dans le cas contraire, il est dit *sous surveillance*). Nous observons que la présence d'un cancer du sein ne se présente pas comme un multiplicateur de risque de la mortalité en fonction de l'âge, mais plutôt un surrisque additif – en première approche.

Tab. 1. Éléments de modélisation de la surmortalité liée au cancer du sein⁵

Âge	Mortalité y. c. féminine	En cas de cancer du sein (et surmortalité)	Distinction : cancers actifs	Distinction : cancers sous-surveillance
15-34	0,04% 0,03%	2,xx% (+2,xx% +2.x%)	3,02% (+2,98% +3,0%) N=2 600	N/A (+N/A +NA) N=1 100
35-54	0,19% 0,13%	1,57% (+1,38% +1,4%)	3,20% (+3,01% +3,1%) N=48 100	0,21% (+0,02% +0,1%) N=58 000
55-64	0,67% 0,43%	1,83% (+1,16% +1,4%)	4,71% (+3,04% +4,3%) N=38 700	0,51% (-0,16% +0,1%) N=84 800
65-74	1,34% 0,89%	2,29% (+0,95% +1,4%)	5,56% (+4,22% +4,7%) N=44 700	0,98% (-0,36% +0,1%) N=111 700

Ces nombres sont très approximatifs par rapport aux besoins car ils ne distinguent pas l'ancienneté de la maladie et sont propres à la population générale, par opposition aux catégories socio-professionnelles de la population emprunteuse. Mais en première approche les cancers du sein sous surveillance ont des surrisques faibles, les cancers du sein actifs ont des risques élevés : une mutualisation des cancers du sein actifs s'accompagnerait comme surprime d'un « pourmillage » de 30 pour mille (3 %) de la somme à rembourser (« capital restant dû »). Un tel niveau de pourmillage est élevé, il est possible que cela nuierait à la réputation de l'assureur. Si le cadre évitait que les bons risques n'aillent à la concurrence et que l'assureur ne se retrouve qu'avec les mauvais risques, il serait possible de couvrir l'ensemble des cancers du sein avec un pourmillage inférieur à 15, ce qui est plus acceptable.

Notons qu'une telle mutualisation de l'ensemble des cancers du sein entrerait en conflit avec la convention AERAS de mars 2021, indiquant une surprime double (pourcentage de 100) voire nulle pour certains cancers du sein peu risqués (de type pT1NoMo, au bout de trois et sept ans après la fin du protocole et sans rechute), et avec le droit à l'oubli pour les cancers (au bout de 5 ans avant l'âge de 21 ans puis au bout de 10 ans). Il conviendrait que le cadre change pour le remplacer par celui-là. Notons aussi que lorsque le surrisque est essentiellement additif avec l'âge, établir des surprimes en pourcentage rend les contrats intéressants pour les assureurs pour les âges élevés mais pas les âges jeunes ; mieux vaut donc pour le cancer du sein appliquer des surprimes en pourmillages.

En prenant à présent l'exemple du diabète, la Table 2 indique les risques de mortalité selon Ameli (AMELI, 2018d). Contrairement au cancer du sein, nous n'avons pas ici de relation simple avec la mortalité non-spécifique, comme l'indique la troisième colonne. Une relation approximative est que le risque de mortalité des diabétiques est environ 50 % supérieur à

⁵ Deux premières colonnes : mortalité par tranches d'âge de la population générale française en 2018 (issu de la table de mortalité 2017-2019; mortalité tous sexes confondus en gras). Troisième colonne : mortalité en cas de cancer du sein (moyenne pondérée à partir des cancers du sein actifs et sous surveillance) et entre parenthèses la surmortalité additive par différence avec la deuxième colonne (en gras ou pas suivant que la référence soit tous sexes confondus). Quatrième et cinquième colonne, les données Ameli utilisées pour calculer la troisième colonne (taux de mortalité en cas de cancer du sein actif et sous surveillance et prévalence de ces cancers) et surmortalité par rapport à la population générale.

celui de la population générale, plus surrisque additif relativement faible (de l'ordre de 0,1 %). Une interprétation est la suivante (cohérente pour d'autres pathologies étudiées de façon similaire) : les cancers du sein sont une maladie assez souvent brutale, pouvant se développer assez rapidement, alors que le diabète est une maladie beaucoup plus souvent progressive, qui s'installe avec le temps et dégrade la santé progressivement, un peu comme le corps se dégrade en vieillissant. Ainsi un contrat inclusif du diabète correspondrait environ à une surprime de 50 pour cent et de un pour mille.

Tab. 2. Éléments de modélisation de la surmortalité liée au diabète⁶

Âge	Mortalité	En cas de diabète (et surmortalité)
15-34	0,04%	0,17% (+0,1% ou x 4,0 ou x 1,5 + 0,11%)
35-54	0,19%	0,52% (+0,3% ou x 2,7 ou x 1,5 + 0,23%)
55-64	0,67%	1,12% (+0,5% ou x 1,7 ou x 1,5 + 0,12%)
65-74	1,34%	2,03% (+0,7% ou x 1,5 ou x 1,5 + 0,01%)

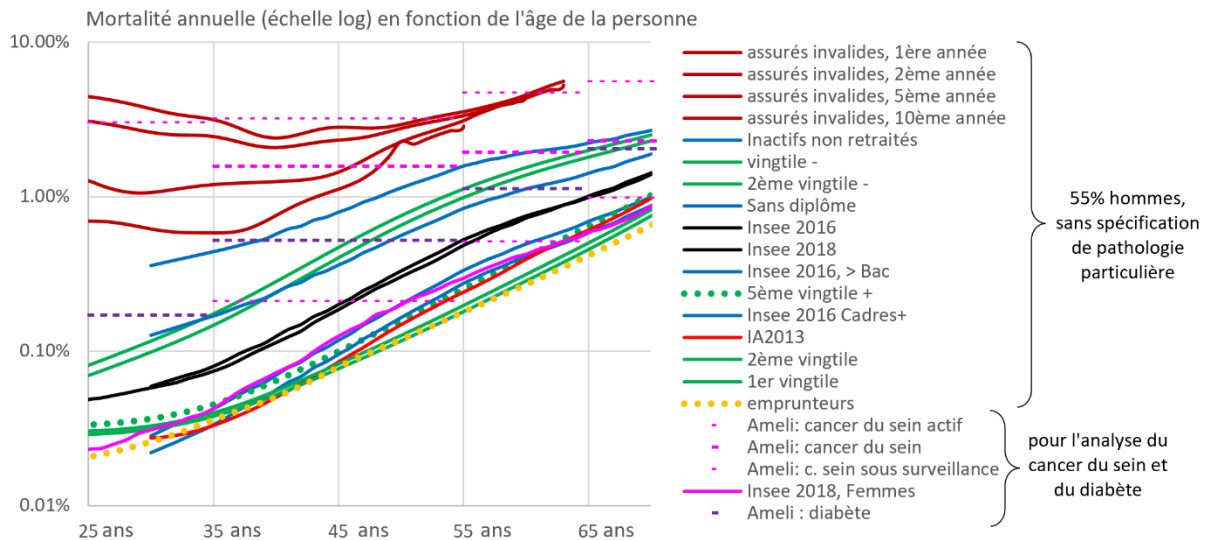
Le cas du diabète est probablement moins important que celui du cancer du sein dans la mesure où les refus liés à la pathologie y sont plus rares. Mais par la suite nous déroulons l'analyse technique en tant qu'exemple de pathologie.

i. Point de repères : mortalité pour diverses populations françaises

Dans le cadre de l'utilisation de l'open data, il convient de comparer les risques de diverses sources, pour s'assurer de la cohérence des résultats et éviter ainsi une mauvaise compréhension des sources, mais aussi apprécier au cas par cas le besoin en informations complémentaires. À la Figure 1, les risques de mortalité que nous venons d'étudier, en cas de cancer du sein et en cas de diabète, sont mis en regard des risques de mortalité de diverses populations françaises non définies par une pathologie.

⁶ Deux premières colonnes : mortalité par tranches d'âge de la population générale française en 2018 (issu de la table de mortalité 2017-2019⁶). Troisième colonne : mortalité en cas de diabète et, entre parenthèses, trois représentations de la surmortalité. La première est une surmortalité additive. La deuxième une surmortalité multiplicative. La troisième consiste à observer que le risque de mortalité en cas de diabète est environ 50 % supérieur au risque de mortalité de la population générale, plus un petit surrisque additif dépendant de l'âge.

Fig. 1: Taux annuel de mortalité en France entre 25 et 65 ans, pour différentes populations, pour 55% d'hommes et 45% de femmes⁷



Nous pouvons observer la cohérence des risques de populations non définies par une pathologie. La population générale française a bien un risque (représenté en noir) similaire en 2016 (Insee, 2016) et 2018 (Insee 2018), avec une tendance à la baisse. Lorsque segmentée en vert par vingtile de revenu (Insee 2018) et en bleu par diplômes (Insee 2016), le niveau de mortalité baisse bien lorsque le niveau de catégorie socio-professionnelle augmente. En rouge, le risque moyen d'assurés (IA 2013) est de l'ordre de celui des cadres et professions intellectuelles supérieures (courbe bleue la plus basse). En pointillés jaunes, une moyenne de tables actuarielles en assurance emprunteurs (LSAF 2021), correspondant à la souscription à des revenus élevés et une absence de pathologie sévère, montre un risque plus faible que celui en rouge d'assurés en général, en moyenne moins sélectionnés, voir plus faible que le premier vingtile en termes de revenus, lequel représente certes. Le risque de décès en cas d'invalidité (BCAC, 2013), lui, est particulièrement grand. S'il décroît les premières années (suite à un accident ou une maladie, vraisemblablement du fait du décès des personnes les plus à risque et d'une stabilisation des états de santé pour les autres), avant de monter avec l'âge, il reste plus élevé que pour les inactifs non retraités (invalides ou non). Nous avons l'habitude de l'usage de ces dernières courbes et savons que l'ordre de grandeur est correct.

La cohérence observée entre ces courbes de référence nous permet de reconstruire mentalement et approximativement ce que seraient les risques de mortalité donnés par les fiches pathologie Amelie si elles étaient disponibles par tranches d'âge plus fines (les droites par morceau seraient penchées comme les courbes de références proches, avec approximativement les mêmes centres). Nous voyons que les risques liés à des cancer du sein actifs sont de l'ordre de ceux d'assurés invalides dans les années qui suivent la cause de l'invalidité, et que les risques de la population générale avec un cancer du sein ou du diabète

⁷ Pour les risques d'assurés cette répartition de sexe est à dire d'expert, les données à notre disposition ne distinguant pas le sexe

sont fortement supérieurs à ceux de la population emprunteuse, même en cas de cancer du sein sous surveillance. Il convient donc d'étudier d'une part plus en détail les risques associés à différentes formes des pathologies et d'autre part la transposition du risque aux catégories socio-professionnelles de type emprunteurs.

Notons que la Figure 1 permet d'apprécier ce que pourraient être les risques des emprunteurs en l'absence de sélection médicale, si un dispositif évitait que les pathologies sévères ne soient sur-représentées. Les demandeurs d'emprunt étant majoritairement des cadres et professions intellectuelles supérieures, surtout en pondérant par les montants, le risque devrait être proche de la courbe bleue la plus basse. Du fait de l'absence de sélection médicale, le risque devrait être similaire à la courbe rouge (tendance à être plus élevé de ce fait, tendance à être plus bas du fait des ressources financières nécessaires pour emprunter). Associant ces catégories professionnelles à un vingtile, cela correspondrait alors au 5^{ème} vingtile de population française, représenté en pointillés vert, soit une augmentation du risque de l'ordre de 40% par rapport aux emprunteurs actuels (courbe jaune en pointillés), ou moins puisque les patients en état grave de santé ont vraisemblablement moins de velléités à chercher un emprunt. Mais vu le facteur presque 100 entre ces risques et ceux des personnes avec un cancer du sein actif par exemple, cette vision théorique se heurte au risque que le cadre soit insuffisant et que les pathologies sévères soient sur-représentées chez un ou quelques assureurs plutôt que distribuées de manière similaire entre assureurs.

ii. Mortalité de la population générale en fonction de divers facteurs de risque

Nous voyons que dans le cadre actuel, il est indispensable d'affiner la connaissance du risque en fonction de facteurs liés à la pathologie. L'information est délicate à trouver pour des données françaises récentes. Il convient alors de chercher des données plus anciennes ou étrangères, puis d'ajuster les risques au niveau de mortalité observé dans les fiches pathologie Ameli. Au sein de cette complexité, des mot-clé utiles et des manières d'analyser l'information facilitent le travail tels que « survie », « âge », « taux de mortalité », « taux de décès », « mortalité », « décès », « cancer du sein », « diabète de type 1 » ou leurs équivalents en anglais.

Dans les articles scientifiques et plus encore dans les autres éléments, les éléments sur Internet sont sujets à erreur, il convient de vérifier la cohérence des informations trouvées, à la manière de la Figure 1 pour la mortalité de populations en France, et de parcourir les articles scientifiques cités dans les articles non scientifiques pour vérifier que les chiffres énoncés s'y retrouvent et sont utilisés dans le bon contexte.

Dans le cadre du cancer du sein, un rapport de Santé Publique France (ROGEL et al, 2016) indique que la survie est de 87% sur 5 ans pour les femmes diagnostiquées entre 2005 et 2010, que l'âge médian de décès est 73 ans, et que depuis 2015 la mortalité baisse de 1.5% par an. Nous pouvons conforter ces données avec les niveaux de mortalité vus précédemment. En notant q le taux moyen annuel de mortalité, la survie sur 5 ans est $S = (1 - q)^5$ donc $q = 1 - S^{1/5}$: le taux de mortalité annuel moyen pour les femmes diagnostiquées avec un

cancer du sein vers 65 ans (73 ans moins 5 à 10 ans) est de $(1 - 87 \%^{1/5}) (1 - 1,5 \%)^{13} = 2.26\%$ en 2018, ce qui rappelle le 2,29 d'Ameli dans le tableau 1 (Ameli 2018b, 2018c).

Une autre source⁸ précise que ce taux de survie à 5 ans de 87 % est en réalité supérieur à 92 % pour des tumeurs de moins de 5 cm, de l'ordre de 60 % pour des tumeurs touchant plusieurs ganglions lymphatiques ou les tissus voisins des seins, et de 20 % pour les cancers ayant métastasé à distance. Ceci permet de suite de voir qu'une approche peut consister à accepter tous les cancers du sein sauf les cancers du sein métastatiques.

Ces analyses confortent dans la modélisation du risque de mortalité en cas de cancer du sein un article datant de 2008 (DABAKUYO et al, 2008). Cet article effectue une analyse multivariée de la mortalité sur quinze ans pour environ 3000 femmes atteintes d'un cancer du sein, en fonction de l'âge au diagnostic, du stade TNM, du grade SBR, et du fonctionnement des récepteurs œstrogènes. Nous y retrouvons une mortalité en fonction de l'âge compatible à celle des données d'Ameli, et un risque important associé aux ganglions lymphatiques. La modélisation est détaillée en Annexe.

Pour affiner davantage la modélisation du risque, nous pouvons nous appuyer sur le « SEER »⁹, une application gratuite qui fournit des statistiques très précises sur les cancers aux Etats-Unis, produit *via* le. Sur son site notamment les statistiques de mortalité et de cohortes atteintes de cancer aux Etats-Unis en fonction de nombreux paramètres tels que la localisation du cancer, son stade de sévérité, l'âge de la personne et le temps écoulé depuis le diagnostic. La taille de la population américaine permet d'avoir des statistiques solides comparées à celles d'articles français. Les risques aux Etats-Unis ne sont pas forcément les mêmes qu'en France, mais si à une grosse maille les risques de mortalité en France et aux Etats-Unis ont des ressemblances, il est possible de s'inspirer de ces statistiques américaines pour affiner les risques en France.

Afin de présenter quelques aspects méthodologiques complexes, prenons à présent le cas du diabète de type 1. Il est très difficile de trouver des taux de mortalité spécifique au diabète de type 1 en France. Les recherches à l'étranger nous amènent à une étude danoise (JØRGENSEN et al., 2013) portant sur 4 821 individus ayant un diabète de type 1 suivis entre 2002 et 2011. Cette étude modélise la mortalité en fonction de l'âge au diagnostic, le temps écoulé depuis, le sexe, et l'absence ou non de néphropathie, et montre notamment l'importance de ce dernier facteur de risque. N'ayant trouvé de source claire sur le niveau de mortalité français, nous comparons cette mortalité à la mortalité de la population générale danoise, le ratio entre les deux nous permettant d'aboutir à une modélisation de la mortalité de diabétiques de type 1 français avec un niveau socio-professionnel typique des emprunteurs, suivant les formules de la section qui suit.

iii. Transposition à des emprunteurs

Les risques de mortalité obtenus avec l'approche ci-dessus correspondent à ceux de la population générale. Or, comme vu en Figure 1, la mortalité des assurés et assurés

⁸ Survie attendue des patients atteints de cancer en France : état des lieux (2010). Table 3, p15.

⁹ National Cancer Institute. surveillance, Epidemiology, and End Results (SEER). seer.cancer.gov

emprunteurs actuels est en moyenne de l'ordre de 2 à 2,5 fois moindre que celle de la population générale.

Dès lors, les demandeurs d'emprunt avec une pathologie ont-ils le risque normal des patients avec cette pathologie ou un risque environ deux fois moindre ?

Il est logique que la mortalité en cas de pathologie soit plus faible chez les demandeurs d'emprunt que la population générale : ils sont, a priori, en moyenne plus éduqués face à la santé donc se font dépister plus tôt, ont de meilleures conditions de vie et moins de comorbidités et donc se remettent plus facilement de soucis de santé. Ainsi, plus la survie d'une pathologie dépend de ces éléments, plus la survie des assurés sera élevée par rapport à celle de la population générale.

Un point de vue mathématique permet d'avoir une idée a priori de la quantification pour certaines pathologies. Les populations de type assurés ont typiquement deux fois moins de risque de décès toutes causes confondues. Cela se traduit en ratio de référence pour les causes de mortalité : le taux de décès lié à une cause donnée (taux de décès relatif à la taille de la population, et non pas à la population diagnostiquée) est $2+x$ fois moins que celui de la population générale, où x est centré d'une certaine manière autour de zéro (tantôt positif ou négatif suivant les pathologies).

En particulier, si une pathologie est très fréquente, x a de fortes chances d'être petite en valeur absolue (car si x était grand, les autres causes de mortalité devraient avoir en moyenne un x grand dans l'autre sens, pour que la mortalité toutes causes soit ce qu'elle est). La mortalité liée à cette pathologie se décompose elle-même comme le produit de l'incidence et de la mortalité en cas de diagnostic. Là, en fonction de la pathologie considérée, la connaissance du lien qualitatif entre incidence et catégorie socio-professionnelle permet d'avoir une idée de la mortalité post diagnostic. Ainsi, dans le cadre du cancer du sein, on peut s'attendre à ce que le taux de décès (relatif à la taille de population) des assurées soit environ la moitié de celui des femmes de la population générale. En considérant que les dépistages sont beaucoup plus fréquents et précoces dans les catégories socio-professionnelles élevées, l'incidence au sens diagnostic et non pas par rapport à un stade d'avancement de la pathologie devrait avoir tendance à être plus grande chez les assurées : on peut s'attendre à ce que la mortalité en cas cancer du sein soit moins de la moitié de celle dans la population générale.

Les cancers sont particuliers dans le sens où leur détection précoce change radicalement la survie associée - d'où les approches de dépistage du cancer du sein, des mélanomes ou du cancer du côlon. Pour d'autres pathologies, nous attendons une incidence moindre chez les assurés que dans la population générale du fait de la meilleure santé en moyenne. Les populations de type assurés ont tout de même a priori tendance à voir le médecin plus tôt donc il y a tout de même une tendance à ne pas être très éloigné de l'incidence de la population générale. Ainsi le facteur deux de mortalité toute cause se retrouverait plus sur la mortalité en cas de pathologie que sur l'incidence.

Ainsi, la référence est que le risque relatif entre la mortalité en présence ou absence d'une pathologie spécifique devrait être en moyenne un peu plus grand pour les assurés que la population générale, et plus petit dans le cadre de pathologies comme les cancers pour

lesquels les assurés font des dépistages particulièrement précoces. C’est pourquoi nous définissons le “multiplicateur de risque relatif” par le ratio entre ces deux risques relatifs, qui devrait donc en moyenne être un peu supérieur à 1 mais inférieur à 2, et vraisemblablement inférieur à 1 pour le cancer du sein.

Considérons ces populations A, B, C et D et, par abus de langage, utilisons les mêmes lettres pour représenter leur risque annuel de mortalité :

Tab. 3. Multiplicateurs de risque

	Population avec une pathologie	Population “saine” (*)	Surmortalité liée à une pathologie
Population CSP+	A	B	A/B à calculer
Population générale (PG)	C	D	C/D (type de métrique générée par l’outil ActuRx)

(*) le mot “saine” est un abus de langage pour indiquer que la population n’a pas été sélectionnée pour une pathologie particulière.

Nous cherchons la valeur de ce multiplicateur :

$$\text{multiplicateurRR} = \frac{A/B}{C/D}$$

Il est très difficile de trouver A/B sur Internet. Cependant, les inégalités sociales étant un sujet important pour les politiques sociales et de santé, il est ainsi plus facile de trouver A/C et B/D , ce qui permet aussi d’obtenir le multiplicateur:

$$\text{multiplicateurRR} = \frac{A/C}{B/D}$$

Cette conversion amène à s’appuyer sur ce ratio de risques relatif plutôt par exemple qu’une différence de risques relatifs. L’obtention de A/C et B/D n’est toutefois pas évidente d’autant que ces deux sous-ratios doivent définir les catégories socio-professionnelles de la même manière (diplômes, métiers, salaires...).

Le taux de mortalité des demandeurs d’emprunt avec une pathologie utilise alors la définition du multiplicateur :

$$A = B \times \frac{C}{D} \times \text{multiplicateurRR}$$

La notion de catégorie socio-professionnelle supérieure est ici celle des demandeurs d’emprunt : A est la mortalité des demandeurs d’emprunt avec la pathologie, C est la mortalité des patients dans la population générale, B/D est le ratio entre la mortalité des assurés emprunteurs et de la population générale.

Écrivons la formule avec des termes plus explicites :

$$qx \text{ emprunteur, pathologie} = qx \text{ emprunteur} \frac{qx \text{ pop. gen., pathologie}}{qx \text{ pop. gen.}} \text{ multiplicateur RR}$$

Dans cette formule,

- $qx \text{ emprunteur}$ et $qx \text{ pop. gen.}$ sont fonction de l'âge et éventuellement du sexe,
- $qx \text{ pop. gen., pathologie}$ dépend de l'âge et, éventuellement, du sexe, mais aussi de variables liées à la pathologie comme l'ancienneté du diagnostic et un élément de sévérité de la pathologie.

La formule est rudimentaire au sens où le même multiplicateur est appliqué quel que soit l'âge ou d'autres variables selon lesquelles les divers termes sont disponibles. Nous supposons cela suffisant pour tarifier des contrats s'il s'agit d'un ajustement de mortalité. Si le multiplicateur était supérieur à 150 % ou inférieur à 75 % par exemple, il conviendrait de vérifier cette hypothèse (au prix de recherches difficiles sur Internet). Il est utile si possible d'établir ce multiplicateur pour divers pays afin de s'assurer de sa robustesse par des valeurs similaires malgré des environnements (pays) différents.

Dans le cas du cancer du sein, nous obtenons un multiplicateur légèrement inférieur à 1, de 83 % pour des données françaises et 76 % ou 80 % selon l'analyse pour des données danoises. Pour le diabète, nous obtenons un multiplicateur légèrement supérieur à 1, de 115 % pour la France, 108 % pour la Corée, 111 % pour les Etats-Unis, et 123 % pour l'Ecosse. Le détail amenant à ces résultats est en annexe. Notons que ces multiplicateurs sont conformes aux attentes les plus probables exprimées en début de section : des multiplicateurs légèrement inférieurs à 1 lorsque le diagnostic précoce chez les catégories socio-professionnelles supérieures fait une différence sur les risques, et légèrement supérieurs à 1 pour des pathologies fréquentes.

À ce stade, il est donc possible pour les assureurs d'établir des tables de mortalité pour une approche inclusive ou partiellement inclusive pour diverses pathologies, sachant que le risque de mortalité est la garantie prépondérante des contrats emprunteurs.

b. Incapacité

iv. Incidence et maintien en arrêt de travail parmi la population salariée

Comme pour la mortalité, pour chaque pathologie la première étape est de rechercher des sources de données ou des études permettant de modéliser le passage en incapacité ou tout du moins en arrêt de travail (*disability* en anglais) et le maintien en arrêt de travail. Cette recherche n'est pas évidente, les données disponibles mentionnant une pathologie spécifique en lien avec l'arrêt de travail étant souvent des indicateurs très macroscopiques.

Nous avons pu identifier la solution consistant à s'appuyer sur la base de données américaine en open data mise à disposition par le National Center for Health Statistic (NHIS, 2016). S'agissant d'une base de données américaine, les risques d'incapacité y sont moindres et la prévalence de certaines pathologies comme le diabète y est plus élevée. Il s'agit alors de comparer les indicateurs macroscopiques pour apprécier les écarts, de

modéliser le risque aux Etats-Unis, et de l'ajuster aux indicateurs macroscopiques français, dans le cadre de la compréhension obtenue à l'occasion de la comparaison des indicateurs.

Cette base de données américaine a été constituée à partir d'un questionnaire envoyé chaque année entre 2010 et 2018 à des individus différents et choisis de manière aléatoire. Elle est composée de 284 809 individus dont 55 % de femmes. 3,3 % des femmes ont été diagnostiquées avec un cancer du sein et les âges moyen et médian de ces femmes sont respectivement de 68 et 70 ans. L'âge moyen au diagnostic étant de 56 ans.

Il est possible de tester différentes combinaisons de variables et ainsi de sélectionner les variables les plus pertinentes.

Nous créons deux modèles :

1. une régression logistique donnant l'incidence d'un arrêt de travail ;
2. une régression gamma donnant la durée d'un arrêt de travail.

À partir des coefficients obtenus nous pouvons calculer la probabilité d'avoir un arrêt maladie dans l'année (incidence) par régression logistique :

$$P(\text{arrêt}|X) = \frac{e^{\beta_0 + \beta_1 \times X_1 + \dots + \beta_n \times X_n}}{1 + e^{\beta_0 + \beta_1 \times X_1 + \dots + \beta_n \times X_n}}$$

Les coefficients obtenus avec cette régression nous permettent de calculer la distribution conditionnelle du nombre de jours passés en arrêt à l'aide de la formule suivante :

$$N \sim \Gamma(k = k_{\text{valeur}}, \mu = e^{\beta_0 + \beta_1 \times X_1 + \dots + \beta_3 \times X_3})$$

L'utilisation de régressions permet d'améliorer la compréhension des relations entre incapacité et facteurs de risque. Il est important de rechercher les variables explicatives les plus pertinentes en fonction de la pathologie étudiée.

v. Transposition à des emprunteurs

Les modèles d'incidence ("ix") et de durée ("dx") d'arrêt de travail aux Etats-Unis sont calibrés avec et sans la pathologie (la présence de la pathologie est une variable du modèle). De là, pensant que des populations plus à risque de décéder sont aussi plus à risque d'incapacité et d'invalidité, nous appliquons le multiplicateur utilisé lors de la transposition aux assurés des risques de mortalité :

$$ix_{\text{emprunteur, pathologie}} = ix_{\text{emprunteur}} \frac{ix_{\text{USA, pathologie}}}{ix_{\text{USA}}} \text{ multiplicateur RR}$$

$$dx_{\text{emprunteur, pathologie}} = dx_{\text{emprunteur}} \frac{dx_{\text{USA, pathologie}}}{dx_{\text{USA}}}$$

Le multiplicateur n'est pas appliqué à la durée afin de ne pas le double-compter, le risque d'incapacité étant représenté par l'incidence et la durée. "ix emprunteur" est une moyenne de tables emprunteur de place, calculée par le laboratoire selon la même approche que pour la mortalité. "dx emprunteur" est la durée moyenne d'incapacité selon les tables du BCAC.

c. Invalidité

vi. Incidence en invalidité parmi la population salariée

Comme pour l'incapacité, l'idéal est de trouver une base de données ou une étude permettant de modéliser le passage en état d'invalidité en fonction des facteurs de risque spécifiques à la pathologie étudiée. Mais comme pour l'incapacité, cela n'est pas évident.

Nous avons pu identifier cette solution générique en s'appuyant sur une étude française (CUERQ et al. 2008) permettant de calculer la probabilité annuelle de passage en état d'invalidité en fonction de la pathologie et du temps écoulé depuis le diagnostic.

La probabilité de passage en état d'invalidité est calculée à partir de données agrégées par pathologie. Nous annualisons les taux d'invalidité à 3 et 10 ans afin d'obtenir une probabilité de passage en fonction du temps écoulé depuis le diagnostic.

Tab. 4. Taux de patients en invalidité, 3 et 10 ans après le début de l'affectation de longue durée (ALD) en France

	Taux (%)	
	3 ans après	10 ans après
Sclérose en plaques (ALD 25)	14,0	23,4
Accident vasculaire cérébral invalidant (ALD 1)	19,8	21,9
Polyarthrite rhumatoïde évolutive grave (ALD 22)	10,1	17,1
Artériopathies chroniques avec manifestations ischémiques (ALD 3)	10,6	17,0
Maladie coronaire (ALD 13)	9,9	15,1
Insuffisance cardiaque, cardiopathies graves (ALD 5)	9,6	14,3
Néphropathie chronique grave et syndrome néphrotique (ALD 19)	7,6	13,2
Formes graves des affections neurologiques, épilepsie grave (ALD 9)	8,8	13,0
Affections psychiatriques de longue durée (ALD 23)	9,9	13,0
Insuffisance respiratoire chronique grave (ALD 14)	9,0	12,6
Spondylarthrite ankylosante grave (ALD 27)	8,0	12,5
Tumeurs malignes (ALD 30)	8,4	10,8
Hypertension artérielle sévère (ALD 12)	5,1	9,8
Maladies chroniques actives du foie et cirrhoses (ALD 6)	5,9	8,6
Diabète de type 1 et diabète de type 2 (ALD 8)	3,3	7,6
Maladie de Crohn et Rectocolite hémorragique évolutive (ALD 24)	2,2	4,7
Déficit immunitaire primitif grave, sida (ALD 7)	1,9	3,6

L'annualisation se faire de la manière suivante :

$$P_{inval}(t) = 1 - (1 - P_3)^{1/3} \text{ pour } t \text{ entre } 0 \text{ et } 3 \text{ ans}$$

$$P_{inval}(t) = 1 - \left(\frac{1 - P_{10}}{1 - P_3}\right)^{1/7} \text{ pour } t \text{ au-delà de } 3 \text{ ans}$$

La probabilité de passage en état d'invalidité ne dépend alors que du temps écoulé depuis le diagnostic. Elle incorpore implicitement dans sa définition des niveaux de mortalité.

Il est possible d'explicitier l'intervention de la mortalité dans les statistiques d'invalidité P_3 et P_{10} . Considérons un niveau de mortalité moyen, par exemple pour une personne de 40 ans avec une sévérité moyenne de la pathologie. Etre en invalidité trois ans après le diagnostic (probabilité P_3) revient à n'être pas décédé les trois premières années (probabilité $(1-q_1)(1-q_2)(1-q_3)$) ni n'être resté hors invalidité (probabilité $1-(1-P)^3$) :

$$P_3 = ((1 - q_1)(1 - q_2)(1 - q_3))(1 - (1 - P_{inval}(t))^3) \text{ pour } t \text{ entre } 0 \text{ et } 3 \text{ ans}$$

De cette équation se déduit la probabilité annuelle de passage en invalidité les trois premières années suivant le diagnostic, sans y incorporer d'hypothèse de la mortalité:

$$P_{inval}(t) = 1 - [1 - P_3 / ((1 - q_1)(1 - q_2)(1 - q_3))]^{1/3} \text{ pour } t \text{ entre 0 et 3 ans}$$

Il convient pour cette équation de prendre un taux de mortalité moyen et non pas défini par l'âge ou d'autres variables, car une mortalité faible sera alors accompagnée d'une modélisation élevée d'invalidité et une mortalité élevée d'une invalidité faible, ce qui ne fait pas de sens.

De P_3 et P_{10} on peut similairement déduire la probabilité annuelle de passage invalidité, après les 3 premières années post diagnostic.

Par simplicité et parce qu'il n'est pas pertinent de chercher plus de précision face à l'utilisation de ces données macroscopiques d'invalidité, pour le calcul de primes d'incapacité et invalidité nous nous appuyons sur les formules encadrées.

vii. Transposition à des emprunteurs

Pensant que des populations plus à risque de décéder sont aussi plus à risque d'invalidité, nous appliquons le multiplicateur utilisé lors de la transposition aux assurés des risques de mortalité :

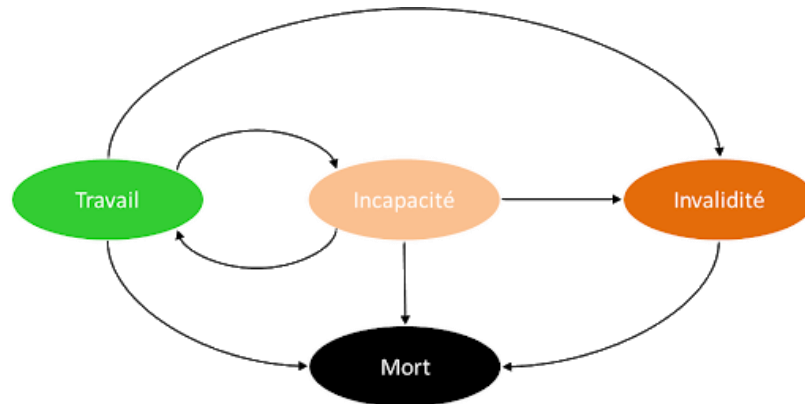
$$P_{inv \text{ emprunteur, pathologie}} = P_{inv \text{ emprunteur}} \frac{P_{inv \text{ pop. gen., pathologie}}}{P_{inv \text{ pop. gen.}}} \text{ multiplicateurRR}$$

“ $P_{inv \text{ emprunteur}}$ ” est une moyenne de tables emprunteur de place, calculée par le laboratoire selon la même approche que pour la mortalité. “ $P_{inv \text{ pop. gen., pathologie}}$ ” est le $P_{inval}(t)$ définit dans cette section (“pop. gen.” indique ici la population générale au sens travailleurs) “ $P_{inv \text{ pop. gen.}}$ ” est défini selon l'incidence en arrêt de travail en France, à laquelle le maintien en incapacité et passage en invalidité est appliqué selon les tables du BCAC 2013 comme effectué par (BAGUI, 2013).

d. Calcul de primes

Un modèle de Markov constitue une approximation simple pour assembler les risques de mortalité, incapacité et invalidité¹⁰. Ce modèle est schématisé à la Fig. 2 :

Fig. 2 : Schématisation d'un modèle de Markov



Chaque flèche correspond à une probabilité : probabilité de décès depuis l'un des trois autres états, probabilité de passer en incapacité, de passer en invalidité avant la fin des 3 ans en incapacité (au bout de 3 ans le passage est automatique), la probabilité de revenir à l'état de travail. Ces probabilités seraient typiquement des fonctions de l'âge, de caractéristiques de la pathologie, et de l'ancienneté dans l'état ou l'ancienneté du diagnostic dans un cadre semi-markovien.

En pratique, les données trouvées nécessitent des ajustements. Nous ajustons le modèle ainsi, avec un pas de temps d'un an :

- Nous n'utilisons pas des lois de mortalité spécifiques à chaque état. Ainsi la prime pour la garantie décès se calcule sans considération d'incapacité et d'invalidité, sans passer par le modèle multi-états (cela donnerait le même s'il utilise la même loi de mortalité pour chaque état en vie, ce que nous avons vérifié numériquement).
- Nous regroupons les états « Travail » et « Incapacité » en un seul état. Ainsi chaque année restant dans cet état, la sinistralité en incapacité correspondant à l'incidence et la durée moyenne en incapacité est comptabilisée.
- En utilisant les formules de passage en invalidité qui incorporent implicitement la mortalité, la probabilité de décéder depuis cet état est mise à zéro pour ne pas compter la mortalité en double. Les montants à régler en cas d'invalidité intègrent alors ceux à verser en cas de décès, multipliés par la probabilité de décès. A condition de faire cela, nous avons vérifié que l'incorporation de la mortalité ou non dans la probabilité de passage en invalidité n'a pas d'impact matériel sur les primes.

En $t=0$, la proportion des emprunteurs est 1 dans l'état « Travail ou incapacité » et 0 dans les deux autres états. Elle est donc représentée par un vecteur $(1; 0; 0)$. A chaque pas de temps suivant, la proportion des emprunteurs dans chaque état est le produit entre d'une part la

¹⁰ Les probabilités conditionnelles de transitions dépendant de la durée passée dans l'état, les modèles semi-markoviens sont mieux adaptés, mais compte tenu du contexte de la présente étude, cette simplification est acceptable.

matrice de passage constituée des probabilités annuelles de transition et d'autre part le vecteur de proportion des états chez les emprunteurs du pas de temps précédent.

Notons que la matrice est comme suit et que sa valeur numérique est différente pour chaque date ($t=0, 1, \text{etc.}$ sur la durée du contrat), en fonction de l'âge et des caractéristiques initiales de la pathologie.

Tab. 5. Primes et pathologies (matrice de passage¹¹)

1- <i>qx</i> emprunteur, pathologie - <i>Pinv</i> emprunteur, patholog	<i>Pinv</i> emprunteur, patholog	<i>qx</i> emprunteur, pathologie
0	1- <i>qx'</i> emprunteur, pathologie	<i>qx'</i> emprunteur, pathologie
0	0	1

La prime pure est la somme des montants de sinistralité que l'assureur doit verser dans l'année pour chaque état, multipliés par la proportion d'emprunteurs dans l'état. L'exemple le plus simple est celui d'un emprunt couvert par une assurance emprunteur ne portant que sur le décès, car ne dépendant pas des probabilités d'incapacité et d'invalidité. Chaque année $t \geq 0$, une proportion $S_t = \prod_{x=x_0}^{x=x_0+(t-1)} (1 - q_{x_0+t} \text{ emprunteur, pathologie})$ de patients ayant emprunté un montant C à l'âge x_0 pour une durée D avec un taux d'intérêt r , est en vie ($S_0 = 1$) et risque de décéder avec une probabilité $q_{x_0+t} \text{ emprunteur, pathologie}$. En ce cas l'assurance doit prendre en charge le versement du reste à payer, ou "Capital Restant Dû" $CRD_t = [C(1+r)^d/d](d-t)$ (en supposant dans ce prêt que le même montant, entre crochets, est à rembourser tous les ans). Alors la prime pure est :

$$\text{prime pure décès}_t = S_t q_{x_0+t} \text{ emprunteur, pathologie } CRD_t$$

Notons que dans un modèle à deux états (vivant, décédé) S_t s'obtient par la lecture de la proportion dans le premier état; dans le modèle à trois états (travail-ou-incapacité, invalidité, décès) S_t est la somme des proportions des deux premiers états.

L'inclusion de chargements (frais par exemple de distribution, de gestion de sinistres et de réassurance) donnerait une prime commerciale, plus élevée que la prime pure.

Cette prime pure pour un emprunteur en présence d'une pathologie peut être comparée à la prime pure pour un emprunteur dans les conditions habituelles en prenant dans la matrice de passage $qx \text{ emprunteur, pathologie} = qx \text{ emprunteur}$,

¹¹ La notation "qx'" avec une apostrophe indique une valeur nulle si $Pinv \text{ emprunteur, pathologie}$ est défini par la formule intégrant implicitement le décès.

i_x emprunteur, pathologie = i_x emprunteur et P_{inv} emprunteur, pathologie = P_{inv} emprunteur. Cela permet de définir une surprime en pourcentage ou en pourmillage :

Pourcentage

$$= (prime\ emprunteur,\ pathologie / prime\ emprunteur,\ pathologie - 1) * 100$$

$$Pourmillage = (prime\ emprunteur,\ pathologie - prime\ emprunteur,\ pathologie) \times 1000$$

Des exemples sont présentés ci-après, pour les cas du cancer du sein et du diabète.

3. Application deux pathologies, le cancer du sein et le diabète

Nous appliquons cette méthode au cancer du sein et au diabète de type 1.

e. Cancer du sein

viii. Mortalité par facteur de risque

Afin de modéliser la mortalité chez les femmes présentant un cancer du sein actifs infiltrants, nous nous sommes appuyés sur l'article de Dabakuyo (DABAKUYO et al, 2008). Cette étude s'appuie sur les données de 3 831 françaises atteintes d'un cancer du sein. Les données ont été récupérées entre 1982 et 1997, la durée de suivi médiane est de 9 ans. De plus, les femmes considérées dans l'étude sont âgées de 19 à 99 ans au diagnostic, avec un âge moyen et médian de 61 ans. Il est important de noter que le modèle est construit sur 2 615 femmes.

Cette étude fournit les ratios de risque (HR) en fonction des facteurs de risque suivants :

- l'âge au diagnostic ;
- le fonctionnement des récepteurs d'œstrogène ;
- le stade TNM ;
- le grade SBR.

Ces derniers ont été calculés à l'aide d'un modèle de Cox qui nous permet d'estimer la probabilité de décès de la population de référence. La population de référence de cet article est définie selon les critères suivants :

- âge entre 45 et 59 ans ;
- stade T qui vaut 1 ;
- stade N qui vaut 0 ;
- stade M qui vaut 0 ;
- dysfonctionnement des récepteurs œstrogènes ;
- grade SBR qui vaut 1.

Comme nous l'avons évoqué plus tôt, il est nécessaire de créer un multiplicateur permettant de passer de la population générale à la population assurée.

Pour la population ayant un cancer du sein, nous nous sommes appuyés sur l'étude de (GENTIL-BREVET et al, 2008). Cette étude s'appuie sur un échantillon composé de 1 150 femmes françaises ayant un cancer du sein invasif diagnostiqué entre 1995 et 1997. Elles ont été suivies jusqu'en 2006.

La répartition des catégories socioprofessionnelle est la suivante :

- 42,5 % sont dans le groupe des CSP+ défini par « executives, middle professional group and clerical employees » ;
- 40,9 % sont dans le groupe des CSP- défini par « famers, artisans, manual workers, unemployed » ;
- 16,6 % n'ont pas de CSP renseignée.

Nous avons donc $\frac{A}{C} = 66.4\%$ à 5 ans et $\frac{A}{C} = 68,1 \%$ à 7 ans.

Pour la population saine, nous nous sommes appuyés sur les tables de mortalité par sexe et quintile de niveau de vie, fournies par l'INSEE. Comme le fait l'article (Gentil-Brevet J., Colonna M., Danzon A. et al, 2008), nous avons séparé les catégories socioprofessionnelles en deux (CSP+ et CSP-). Nous avons utilisé la même pondération que l'article pour la population touchée par le cancer du sein : 25 % des femmes ont moins de 50 ans, environ 50 % ont entre 50 et 70 ans et 25 % entre 71 et 85 ans.

Nous obtenons ainsi des taux de mortalité de 686 et 860 pour 100 000 personnes les CSP+ et CSP- respectivement.

Le risque relatif associé à la population saine est de $\frac{B}{D} = \frac{686}{860} = 80 \%$.

Le multiplicateur à appliquer aux surprimes du cancer du sein générées par l'outil ActuRx se calcule de la manière suivante :

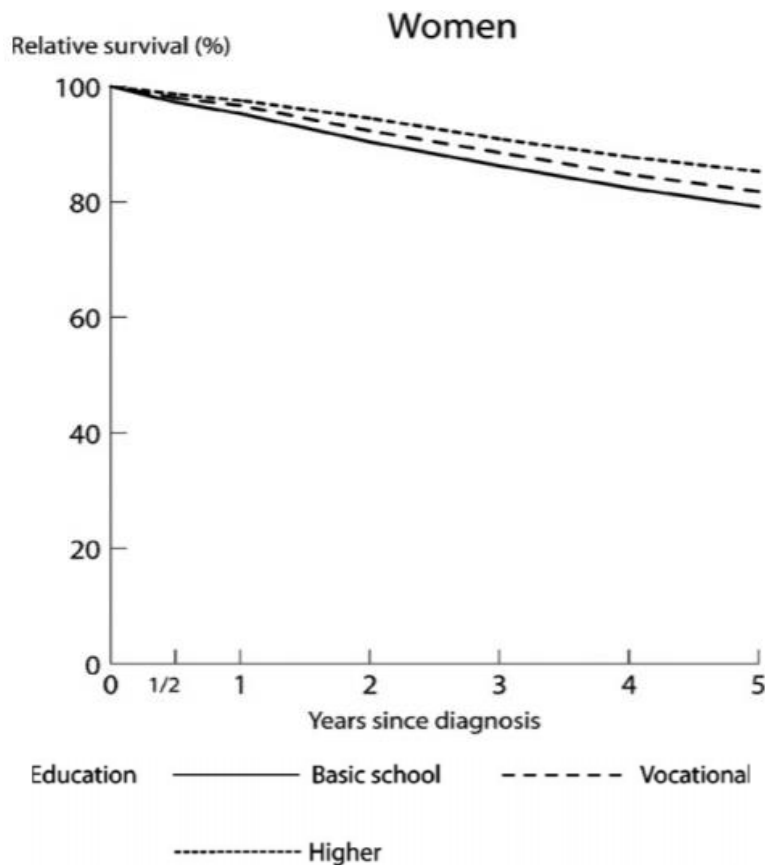
$$\frac{\frac{A}{C}}{\frac{B}{D}} = \frac{66,4 \ %}{80 \ \%} = 83\%$$

Si nous avons pu trouver des articles de divers pays sur la mortalité des femmes diagnostiquées comme ayant un cancer du sein, et de nombreux articles croisant la CSP avec les dépistages et statistiques de typologies de cancer de cancer du sein, réussir à croiser les statistiques de mortalité liées et la CSP chez les femmes s'est révélé redoutablement difficile. Mais i) de nombreuses études indiquent que les CSP+ se font diagnostiquer beaucoup plus tôt et ont de bien meilleurs pronostic ce qui corrobore l'idée d'un multiplicateur $RR < 1$, et ii) nous avons fini par trouver une étude Danoise qui convient.

Une étude (CARLSEN et al, 2008) i) s'appuie sur un grand nombre de femmes (plus d'un million et demi sans cancer du sein et plus de 25 mille avec un cancer du sein entre 2004 et 2013 - 72% des cas de cancers du sein au Danemark durant la période), ii) et croise cancer du sein ou non, CSP+ ou non et mortalité post-cancer du sein sur les mêmes données, ce qui amène une comparabilité meilleure que d'obtenir A/C par un article et B/D par un autre.

L'indicateur de la figure 5 (Figure 2 dans l'article) n'est pas la survie 5 cinq ans mais directement le ratio de survie à 5 ans entre femmes atteintes d'un cancer ou non. Et ceci, par CSP. Cette survie en cas de cancer *relative* à la survie sans cancer du sein est plus élevée pour les CSP+, la mortalité est donc relativement plus faible : le multiplicateur est inférieur à 1.

Fig. 3 : Survie relative à l'absence de cancer du sein, des femmes au Danemark (Carlsen 2008)



Pour les hauts niveaux d'études, cette survie relative à 5 ans est de 84%. La mortalité relative annuelle qui correspond est $1-(84\%)^{(1/5)} = 3,43\%$. En faisant de même pour l'ensemble des niveaux et en pondérant par la distribution des niveaux d'études indiquée dans l'article, la mortalité relative pour la population générale est 4,50%. Le ratio des deux donne un multiplicateur de 76%. La même approche peut être faite depuis l'article en fonction du revenu (par quartiles), et non obtenons un multiplicateur de 80%. Ces valeurs sont proches du 83% trouvé pour la France, 83% est plus prudent donc nous retenons 83%.

ix. Incapacité en fonction de l'âge

Les régressions 1 et 2 présentent les résultats obtenus à partir de la base de données américaine. Les résultats ont ensuite été ajustés au niveau français (KUSNICK-JOINVILLE, 2006).

Tab. 6. Sortie de la régression logistique de l'incidence d'arrêt maladie

```
Call:
glm(formula = MISS_WORK ~ AGE_P + HAS_CAN5 + SEX, family = binomial(logit),
     data = df)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8428 -0.5470 -0.4496 -0.3685  2.4803

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.1781293  0.0248435 -47.422 < 2e-16 ***
AGE_P       -0.0231627  0.0003492 -66.335 < 2e-16 ***
HAS_CAN5TRUE 0.5292755  0.0713211  7.421 1.16e-13 ***
SEX          0.1183334  0.0122204  9.683 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tab. 7. Sortie de la régression gamma de la sévérité des arrêts de travail

```
Call:
glm(formula = WKDAYR ~ AGE_P + HAS_CAN5 + SEX, family = Gamma(link = "log"),
     data = df_MW)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.8876 -1.0767 -0.8168 -0.2102  7.0034

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.2865798  0.0487372  46.916 < 2e-16 ***
AGE_P        0.0177350  0.0007704  23.022 < 2e-16 ***
HAS_CAN5TRUE 0.7315317  0.1246944  5.867 4.49e-09 ***
SEX          -0.0878692  0.0215931 -4.069 4.73e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

x. Invalidité en fonction de l'ancienneté du diagnostic

Comme nous l'avons expliqué dans la présentation de la modélisation nous avons annualisé les taux d'invalidité à 3 et 10 ans.

- $P_{inval}(t) = 1 - (1 - 0.084)^{1/3} = 2.88\%$ pour t entre 0 et 3 ans
- $P_{inval}(t) = 1 - \left(\frac{1-0.108}{1-0.084}\right)^{1/7} = 0.38\%$ pour t au-delà de 3 ans

Nous considérons qu'au-delà de 3 ans après le diagnostic, le risque de passer en invalidité pour une femme atteinte d'un cancer du sein n'est pas supérieur à celui d'un individu de la population générale.

xi. Design du modèle

Comme nous l'avons développé, nous avons ensuite modélisé la mortalité pour la population malade invalide. Pour cela, nous avons modélisé la réalité de la manière suivante :

- le taux de mortalité des femmes ayant un cancer du sein est de 1 %
- le taux de mortalité des femmes invalides ayant un cancer du sein est 2,35 %
- le taux de mortalité des femmes valides ayant un cancer du sein 0,85 %

L'ensemble est cohérent pour environ 10 % d'invalides ce qui est compatible avec les informations dont nous disposons. L'incidence d'invalidité parmi les femmes valides ayant un cancer du sein est de 1 % par an. Avec ces coefficients, nous avons simulé une cohorte entre le diagnostic et 15 ans après, mesuré le pourcentage d'invalides 7 ans post diagnostic et mesuré le nombre d'années invalides et nombre de décès pour une cohorte représentant les assurées : valides 5 ans post-diagnostic et suivis jusqu'à 15 ans post-diagnostic.

Après avoir modélisé les différents scénarios, nous avons retenu la modélisation qui considère l'incidence invalidité constante dans le temps et qui annule la mortalité des invalides. Celle-ci est prudente d'environ 4% pour la mortalité et elle est également prudente sur la prime arrêt de travail de 1 %.

xii. Résultats

L'outil ActuRx calcule les probabilités de décès en fonction de facteurs de risque (âge, classification TNM, grade SBR et le bon fonctionnement des récepteurs œstrogènes). Dans un fichier séparé, nous avons calculé des surprimes à partir des probabilités de mourir pour la population malade et pour la population saine, nous considérons la population assurée dans les deux cas. Les probabilités de décès varient beaucoup, il en va de même pour les surprimes générées.

Nous recommandons aux assureurs de fixer une surprime maximale en fonction de l'âge par exemple. Il est possible de générer des surprimes pour tous les cas à l'aide de l'outil ActuRx, mais nous ne recommandons pas pour autant de les mettre en place (certaines surprimes étant bien trop élevées, etc.).

Mortalité

Dans la Table suivante, les surprimes sont proposées à une femme ayant 42 ans au moment de la demande et ayant été diagnostiquée il y a 3 ans.

Tab. 8. Présentation de certains résultats de l'outil ActuRx

Stade T	Stade N	Stade M	Grade SBR	Fonctionnement des récepteurs œstrogènes	Surprimes (en pour mille)
1	0	0	1	Positif	2
2	0	0	1	Positif	5
3	0	0	1	Positif	7
1	0	0	2	Positif	5
1	0	0	1	Négatif	3
1	0	0	2	Négatif	6
1	1	0	1	Positif	5
2	1	0	1	Positif	12

3	1	0	1	Positif	17
2	0	0	2	Positif	12
3	0	0	2	Positif	16
1	1	0	2	Positif	11
2	1	0	2	Positif	26
3	1	0	2	Positif	35
2	0	0	1	Négatif	7
3	0	0	1	Négatif	10
1	1	0	2	Négatif	6
2	1	0	1	Négatif	16
3	1	0	1	Négatif	22
2	0	0	2	Négatif	16
3	0	0	2	Négatif	21
1	1	0	2	Négatif	15
2	1	0	2	Négatif	34
3	1	0	2	Négatif	45

Seules les surprimes, par facteurs de risque, pour les cancers du sein non métastatiques sont présentés dans la Table 13. Il est possible de proposer des surprimes mutualisées pour des groupes plus ou moins larges. Le calcul d'une surprime unique pour un groupe non homogène permet de mutualiser le risque et ainsi d'élargir l'accès à l'assurance pour une population actuellement refusée.

La table suivante présente les surprimes mutualisées à partir de différents facteurs de risque (toujours pour les cancers non métastatiques). La pondération utilisée pour le calcul des surprimes moyennes est le découpage de SEER Stats.

Nous considérons les surprimes des femmes âgées de 25 à 50 ans avec une ancienneté de 1 à 10 ans au moment de la demande.

Tab. 9. Présentation de certains résultats de l'outil ActuRx

Stade T	Stade N	Stade M	Grade SBR	Fonctionnement des récepteurs œstrogènes	Surprises (moyenne sur 8 ans) En pour mille
-	-	0	-	-	13.45
1	-	0	-	-	5.82
2	-	0	-	-	21.10
3	-	0	-	-	31.71
4	-	0	-	-	36.73
-	1	0	-	-	7.20
-	0	0	-	-	24.88
-	-	0	1	-	3.49
-	-	0	2	-	11.51
-	-	0	3	-	21.12
-	-	0	-	Négatif	21.41
-	-	0	-	Positif	11.30

Incapacité/invalidité

Les Figures 4 et 5 présentent les surprimes d'incapacité et d'invalidité en fonction de l'âge actuel. Les surprimes d'incapacité et d'invalidité en pourcentages sont présentées dans les Figures 4 et 5. Elles sont présentées en fonction de l'âge du proposant au moment de la demande. Nos surprimes diminuent avec le temps, pour présenter quelque chose de plus interprétable nous présentons les surprimes moyennes sur la durée de 8 ans.

Fig. 4 : Surprises d'incapacité en fonction de l'âge

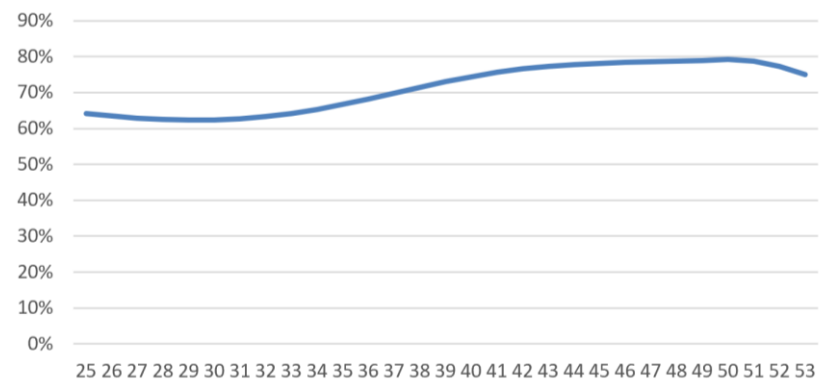
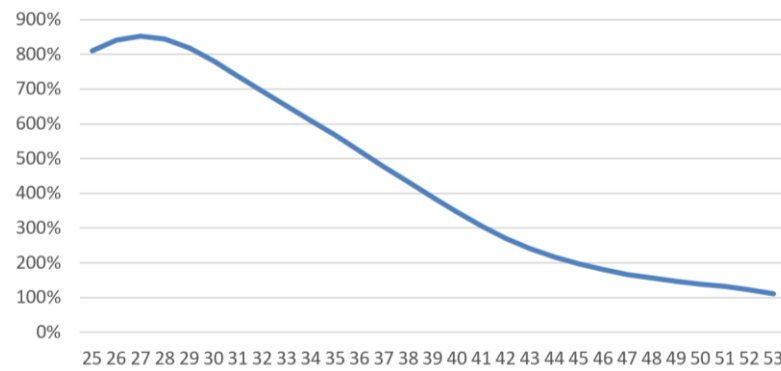


Fig. 5 : Surprises d'invalidité en fonction de l'âge



La surprime d'incapacité/invalidité est la moyenne des deux surprimes en pourcentage présentées dans les Figure 4 et 5. La surprime d'incapacité étant relativement stable en fonction de l'âge et la surprime d'invalidité diminuant avec l'âge, nous savons que la surprime d'incapacité / invalidité diminue avec l'âge.

f. Diabète

Nous présentons dans la suite la modélisation du diabète de type 1 indépendamment du diabète de type 2. Le diabète de type 2 représentant 90% des diabétiques, nous pourrions utiliser le modèle sans type de diabète pour le diabète de type 2.

xiii. Mortalité par type de diabète

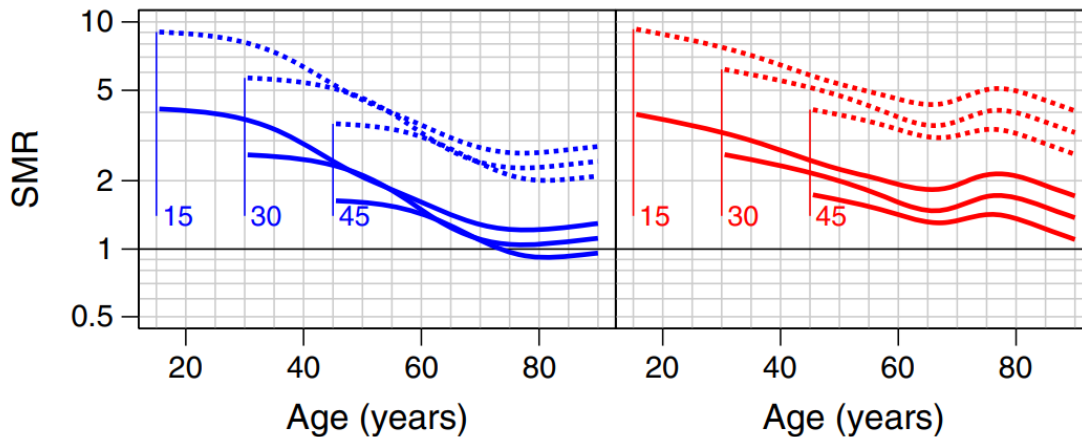
Ici, nous présentons la possibilité de considérer le diabète de type 1 indépendamment du diabète de type 2. Le diabète de type 2 représentant 90 % des diabétiques, nous pourrions utiliser le modèle sans type de diabète pour le diabète de type 2.

Après avoir cherché des articles français pour modéliser la mortalité due au diabète de type 1, sans succès, nous avons recherché des articles de pays voisins et similaires à la France.

L'étude danoise sélectionnée porte sur l'évolution des taux de mortalité relatifs et absolus d'une population diabétique de type 1 composée de 4 821 individus et suivis entre 2002 et 2011 (JØRGENSEN et al., 2013). Les patients suivis sont traités au Steno Diabetes Center, 54 % sont des hommes, l'âge médian est de 44 ans avec une durée médiane de suivi de 18 ans.

Les SMR sont calculés par rapport à la population générale. Comme nous le souhaitons, ils sont calculés par sexe, âge actuel et âge au diagnostic, mais aussi la présence de néphropathie est également une variable considérée. Les SMR sont présentés dans la Figure 6.

Fig. 6 : SMR en fonction de l'âge au diagnostic et de l'âge au moment de l'étude¹²



Notons que les SMR sont plus élevés pour les personnes diagnostiquées jeunes et les personnes avec et une néphropathie. La néphropathie diabétique est une complication qui peut évoluer en insuffisance rénale. Ce critère n'a pas été retenu car nous souhaitons utiliser des facteurs similaires aux guides d'assurance de référence. Pour modéliser les taux de mortalité des diabétiques en France, nous appliquons les SMR danois (JØRGENSEN et al., 2013) aux taux de mortalité de la population générale française.

Comme nous l'avons évoqué plus tôt dans cet article, nous devons à présent calculer un multiplicateur permettant de passer d'une population générale à une population assurée.

Pour la population diabétique, nous nous sommes appuyés sur l'article un article (PIFFARETTI et al, 2016) qui présente les risques relatifs de mortalité des diabétiques en fonction de la catégorie socioprofessionnelle. La population étudiée dans cet article est composée des 7 218 individus diabétiques de type 2 ayant répondu à l'auto-questionnaire des études Entred 2001 et Entred 2007. Le suivi est réalisé jusqu'en 2013.

Nous faisons ici l'hypothèse que la surmortalité de la population générale diabète de type 1 par rapport à la population assurée diabétique de type 1 est la même que pour le diabète de type 2. Nous pouvons donc appliquer le multiplicateur.

La Figure 7 présente les risques relatifs en fonction de la catégorie socioprofessionnelle. Les risques relatifs sont calculés avec les cadres comme référence.

¹² En bleu et rouge pour les hommes et les femmes respectivement. Les lignes pleines représentent les personnes qui n'ont pas de néphropathie (c'est la population qui nous intéresse). Les nombres associés aux courbes (15, 30 et 45) sont les âges au diagnostic.

Fig. 7 : Risques relatifs en fonction de la catégorie socioprofessionnelle avec pour référence les cadres (source : Piffaretti, 2016)

Déterminants de la mortalité des personnes diabétiques de type 2 ayant répondu à l'auto-questionnaire patient d'une des deux cohortes Entred (France) : modèle « patient »[§] (n=6 010)

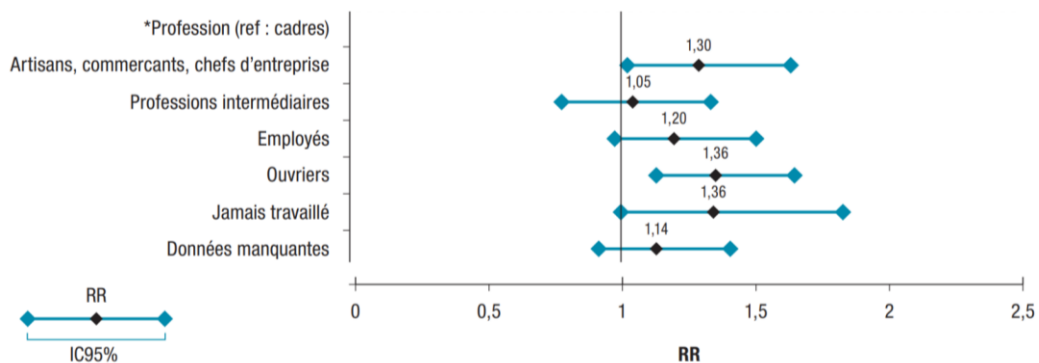


Fig. 8 : Répartition des catégories socioprofessionnelles

Caractéristiques à l'inclusion des cohortes Entred 2001 et Entred 2007 et taux de mortalité des personnes diabétiques de type 2 ayant répondu à l'auto-questionnaire patient (N=7 218), France

Profession et catégorie socioprofessionnelle	Fréquence (%)	Taux de mortalité pour 1 000 PA	p*
Artisans, commerçants et chefs d'entreprise	9,0	41,0	0,0004
Cadres	9,6	29,2	
Professions intermédiaires	11,7	24,4	
Employés	18,4	22,8	
Ouvriers	26,7	31,7	
Jamais travaillé	6,3	30,4	
Données manquantes	18,3	36,5	

En utilisant les risques relatifs présentés dans la Figure 7 et la répartition des catégories socioprofessionnelles de la Figure 8, nous avons calculé le risque relatif lié au diabète $\frac{A}{C}$ pour les cadres.

La référence pour les risques relatifs de la Figure 7 étant les cadres, nous avons calculé les risques relatifs sur la base de la population générale et ainsi obtenu le risque relatif associé aux cadres de la manière suivante :

$$RR = \frac{1}{1 * 9.6\% + 1.36 * 9.0\% + 1.05 * 11.7\% + 1.20 * 18.4\% + 1.36 * 26.7\% + 1.36 * 6.3\% + 1.14 * 18.3\%} = 82\%$$

Pour la population générale, nous avons calculé le ratio $\frac{B}{D}$ à partir des données de l'INSEE (2016). Afin de pouvoir comparer les risques relatifs moyens, nous avons utilisé la pondération par âge et sexe de l'article utilisé pour le diabète (Piffaretti et al, 2016). La cohorte Entred considère des données entre 2001 et 2007, il y a donc plus d'observations à partir de 2008, nous avons donc également pondéré la période 2009-2013 deux fois.

Le risque relatif est de 71.6% pour la population générale.

Le multiplicateur à appliquer aux surprimes générées par l'outil ActuRx se calcule en divisant le risque relatif (comparant les diabétiques aux non-diabétiques) associé aux cadres par le risque relatif associé à la population générale.

Nous obtenons un multiplicateur de $82\%/71,6\% = 115\%$.

Afin de valider la valeur du multiplicateur, nous avons souhaité comparer le multiplicateur français avec des multiplicateurs d'autres pays. Nous avons trouvé des articles intéressants pour l'Ecosse, les Etats-Unis et la Corée. Les multiplicateurs obtenus sont présentés dans la Table 10.

Ces pays étant différents de la France, nous avons uniquement cherché à valider l'ordre de grandeur du multiplicateur français obtenu.

Tab. 10. Multiplicateur par pays

Pays	Multiplicateur
Corée	108%
Etats-Unis	111%
Ecosse	123%
France	115%

Table 6 - Multiplicateurs de passage de la population générale à la population assurée

Les multiplicateurs coréen, écossais et américain sont calculés à partir des articles suivants :

- États Unis : Educational disparities in mortality among adults with diabetes in the US. (DRAY-SPIRA et al, 2010)
- Corée : Relative and combined effects of socioeconomic status and diabetes on mortality: a nationwide cohort study (KIM et al, 2016)
- Écosse : Effect of socioeconomic status on mortality among people with type 2 diabetes: a study from the Scottish Diabetes Research (WALTER et al, 2011)

Nous observons que le multiplicateur français est du même ordre de grandeur que les autres multiplicateurs.

xiv. Incapacité en fonction de l'âge et de l'ancienneté du diagnostic

Dans le cadre de la modélisation de l'incapacité du diabète de type 1, nous nous sommes appuyés sur les années 2016 et 2017 de la base de données américaine car la variable type de diabète n'était présente que pour ces deux années.

Entre 2016 et 2017, 60 000 américains ont répondu aux questionnaires, ce qui permet de capturer l'impact du diabète de type 1 sur l'incidence et le maintien en incapacité. Les variables explicatives sont l'âge du proposant, la présence de diabète de type 1 et le temps écoulé depuis le diagnostic (0 pour les individus qui n'ont pas de diabète de type 1).

Tab. 11. Formule 1 – Sortie de la régression logistique de l'incidence d'arrêt maladie

```
Call:
glm(formula = MISS_WORK ~ AGE_P + TIME_TYPE1 + HAS_DIB1, family = binomial(logit),
    data = df)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7795 -0.6112 -0.5986 -0.5880  1.9271

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.745520   0.053917 -32.374 < 2e-16 ***
AGE_P        0.003248   0.001254   2.591  0.00957 **
TIME_TYPE1   0.004041   0.010030   0.403  0.68698
HAS_DIB1TRUE 0.137766    0.234522   0.587  0.55691
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tab. 12. Formule 2 – Sortie de la régression gamma de la sévérité des arrêts de travail

```
Call:
glm(formula = WKDAYR ~ AGE_P + TIME_TYPE1 + HAS_DIB1, family = Gamma(link = "log"),
    data = df_MW)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4593 -1.0109 -0.7223 -0.1585  7.0854

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.799068   0.088289  20.377 <2e-16 ***
AGE_P        0.024276   0.002049  11.845 <2e-16 ***
TIME_TYPE1  -0.019203   0.014599  -1.315  0.188
HAS_DIB1TRUE 0.321003    0.359609   0.893  0.372
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 3.152745)
```

xv. Invalidité en fonction de l'ancienneté du diagnostic

Comme nous l'avons expliqué dans la présentation de la modélisation nous avons annualisé les taux d'invalidité à 3 et 10 ans.

- $P_{inval}(t) = 1 - (1 - 0.033)^{1/3} = 1.11\%$ pour t entre 0 et 3 ans
- $P_{inval}(t) = 1 - \left(\frac{1-0.076}{1-0.033}\right)^{1/7} = 0.65\%$ pour t au-delà de 3 ans

xvi. Design du modèle

Afin de déterminer le taux de mortalité des diabétiques invalides, nous avons appliqué la méthode décrite précédemment. Nous avons modélisé la réalité telle que :

- le taux de mortalité des diabétiques est de 1 %;
- le taux de mortalité des diabétiques invalides est 3 %;
- le taux de mortalité des diabétiques non invalides est de 0,9 %.

L'ensemble est cohérent pour environ 7% d'invalides ce qui est compatible avec ce que nous connaissons. L'incidence d'invalidité parmi les patients diabétiques non invalides est de 0.7% par an. Avec ces coefficients, nous avons simulé une cohorte entre le diagnostic et 15 ans après, mesuré le pourcentage d'invalides 7 ans post-diagnostic. Nous avons mesuré le

nombre d'années en invalidité et le nombre de décès pour une cohorte représentant les assurés : valides 5 ans post-diagnostic et suivis jusqu'à 15 ans post-diagnostic.

D'un point de vue risque de mortalité, nous voyons que les approches qui considèrent la mortalité des diabétiques sans prise en compte de l'invalidité sont à peu près prudentes de 4 %, celles qui considèrent cela uniquement chez les valides et appliquant une mortalité nulle chez les invalides sont prudentes de 1 %. D'un point de vue risque d'invalidité, les quatre approches donnent des résultats allant d'un manque de prudence de 1% à une prudence de 4 %.

Au global, les modélisations mettant à zéro la mortalité des invalides sont les plus adéquates, avec 1 % de prudence pour la mortalité et 2 % ou 4 % de prudence pour l'invalidité suivant l'interpolation d'incidence.

Cependant comme nous l'avons vu, hors considération de diabète la mortalité des invalides est entre 10 et 20 fois plus grande que celle des non invalides. De fait, pour une mortalité de diabétiques de 1 %, qui ont un risque intermédiaire entre les populations à faible risque et invalide, les hypothèses de réalité sont probablement plus justes avec un taux de décès chez les diabétiques invalides de 8% par exemple. Suivant que 7 % ou 10 % des diabétiques soient invalides (ce qui impacte la mortalité des valides de façon à retomber sur 1 %),

- Quelle que soit la méthode, la prime de décès est 40 % à 95 % trop grande. Cette prudence excessive vient de la différence entre "mortalité des diabétiques" et "mortalité des diabétiques valides", laquelle est relativement importante ; mettre à zéro la mortalité des invalides réduit cette prudence d'environ 5 %.
- Avec 10 % d'invalides (au lieu de 7 %) et/ou avec un taux annuel d'incidence invalidité réelle de 1 % (au lieu de 0,7 %), la seule méthode prudente pour l'invalidité est celle qui annule la mortalité des invalides et qui considère l'incidence invalidité constante. *Nous retenons cette méthode, qui fonctionne dans tous les cas testés.*

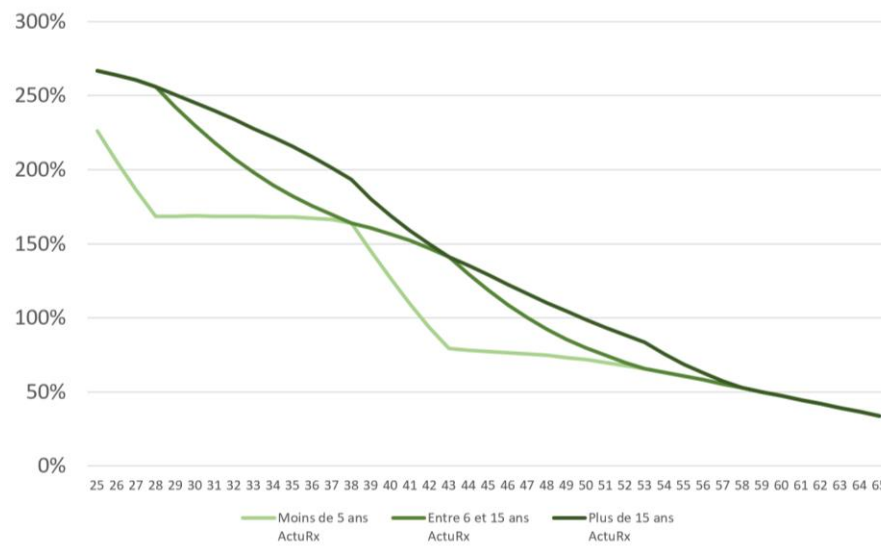
Nos primes calculées concernent des assurés (diabétiques et non diabétiques), pour lesquels l'incidence d'invalidité et donc la part d'invalides est beaucoup plus faible (deux à trois fois plus faible). La méthode retenue devrait être prudente à la fois pour la prime décès et la prime invalidité, mais de peu.

xvii. Résultats

Mortalité

Nous avons généré différents cas afin d'avoir une vue globale des surprimes de mortalité proposées par l'outil ActuRx. Les résultats obtenus sont présentés dans la Figure 10.

Fig. 9 : Surprimes de mortalité en fonction de l'âge et du temps écoulé depuis le diagnostic



Nous observons que la surprime de mortalité diminue avec l'âge et augmente avec le temps écoulé depuis le diagnostic.

L'augmentation des surprimes de mortalité avec le temps écoulé depuis le diagnostic s'explique par le fait que le diabète de type 1 est une maladie auto-immune. Le diabète de type 1 évolue avec le temps. Selon l'INSERM (cf. INSERM [2019]), il existe des risques de complication à long terme (entre 10 et 20 ans après le début des événements). Ces complications concernent principalement le cœur et les vaisseaux, qui sont les premiers lésés par une concentration excessive et permanente de glucose dans le sang.

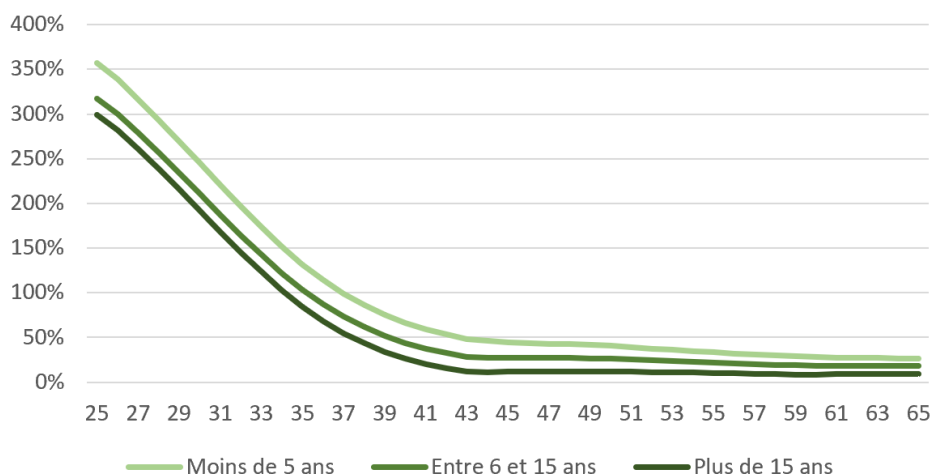
Les risques de mortalité de la population générale saine augmentant avec l'âge, la surmortalité et donc les surprimes diminuent logiquement avec l'âge. Ce que nous pouvons observer dans la Figure 10.

Incapacité/invalidité

Les surprimes obtenues par l'outil ActuRx sont présentées dans la Figure 11. L'outil ActuRx permettant de générer une probabilité de passage en état d'incapacité et une probabilité de passage en état d'invalidité, nous obtenons alors deux surprimes distinctes.

À partir d'un fichier de tarification, nous avons observé que la répartition entre les primes de mortalité, d'incapacité et d'invalidité était 46%, 25% et 29%, donc la prime d'incapacité et d'invalidité sont du même ordre de grandeur. Afin d'obtenir une unique surprime d'incapacité/invalidité, nous calculons la moyenne des deux surprimes, et nous obtenons les surprimes suivantes en fonction du temps écoulé depuis le diagnostic et de l'âge.

Fig. 10 : Surprises d'incapacité et d'invalidité en fonction de l'âge



4. Conclusion

On a montré dans ce travail qu'à l'aide de données issues de l'open data, il est possible d'affiner sensiblement la finesse de description des risques de décès et d'arrêt de travail pour des pathologies graves, conduisant actuellement à de nombreux refus d'assurance. Une appréciation plus précise de ces risques permet *de facto* d'inclure plus de personnes dans l'assurance avec des sursurrisques qui permettent de contenir la prime à un niveau raisonnable.

Les exemples de pathologies pris suggèrent qu'il est possible de couvrir l'ensemble des cancers du sein non métastatiques ou encore l'ensemble des cas de diabète sans néphropathie sans nécessiter de mutualiser des risques sévères avec des risques faibles. La méthodologie présentée permet d'étendre le travail à bien d'autres pathologies, même si chaque pathologie nécessite des efforts bibliographiques et de recoupement d'informations importants.

Pour une inclusion encore plus grande des patients, un cadre serait nécessaire pour mutualiser les risques sévères avec les risques faibles, afin de couvrir par exemple l'ensemble des cas de cancers du sein et de diabète, voire supprimer la sélection médicale. Ce dernier choix diluerait le sursurrisque à travers l'ensemble des assurés, avec une augmentation des primes de base potentielle mais limitée, mais se heurte au risque que les risques élevés soient concentrés chez le ou les premiers assureurs qui tenteraient l'aventure.

5. Références

AMELI [2018a] « Fiches sur les pathologies » disponibles sur le site Ameli.fr. Actuellement disponibles ici : <https://assurance-maladie.ameli.fr/etudes-et-donnees/entree-par-theme/pathologies/cartographie-assurance-maladie/donnees/fiches-pathologies/fiches-pathologies>

AMELI [2018b] Fiche pathologie - Femmes prises en charge pour cancer du sein actif

AMELI [2018c] Fiche pathologie - Femmes prises en charge pour cancer du sein sous surveillance

AMELI [2018d] Fiche pathologie - Diabète

BAGUI H. [2013] [Refonte des lois de maintien en incapacité temporaire de travail](#), mémoire de master, ISFA.

CARLSEN K., HOYBYE M.T., DALTON S.O. [2008] Social inequality and incidence of and survival from breast cancer in a population-based study in Denmark, 1994–2003. *European Journal of Cancer*, 2008, vol. 44, no 14, p. 1996-2002.

CUERQ A., PAÏTA M., RICORDEAU P. [2008] Points de repère Ameli n°16: les causes médicales de l'invalidité en 2006.

DABAKUYO T. S., BONNETAIN F., ROIGNOT P. [2008] Population-based study of breast cancer survival in Cote d'Or (France): prognostic factors and relative survival. *Annals of Oncology*, 2008, vol. 19, no 2, p. 276-283.

DRAY-SPIRA R., GARY-WEBB T.L., BRANCATI F.L. [2010] Educational disparities in mortality among adults with diabetes in the US. *Diabetes care*, 2010, vol. 33, no 6, p. 1200-1205

GENTIL-BREVET J., COLONNA M., DANZON A. [2008] The influence of socio-economic and surveillance characteristics on breast cancer survival : a French population-based study. *British journal of cancer*, 2008, vol. 98, no 1, p. 217-224

JØRGENSEN M. E, ALMDAL T.P., CARSTENSEN B. [2013] Time trends in mortality rates in type 1 diabetes from 2002 to 2011.

INSERM [2019], [Diabète de type 1 – Une maladie auto-immune de plus en plus fréquente](#)

KAMEGA A., PLANCHET F., WOLFRUM R. [2014] [Présentation et comparaison des nouvelles tables BCAC](#). Prim'Act, Caritat Recherche & Formation. Pages 28 (incapacité) et 38 (invalidité).

KIM, N.H., KIM, T.J., KIM, N.H. [2016] Relative and combined effects of socioeconomic status and diabetes on mortality: a nationwide cohort study. *Medicine*, 2016, vol. 95, no 30

KUSNICK-JOINVILLE O. [2006] Points de repère Ameli n°5: déterminants de l'évolution des indemnités journalières maladie

NATIONAL HEALTH SURVEY [2016]) Centers for Disease Control and Prevention

PIFFARETTI C., FAGOT-CAMPAGNA A., REY G. [2016] Déterminants de la mortalité des personnes diabétiques de type 2. Cohortes Entred, France, 2002-2013. *Bulletin Epidémiologique Hebdomadaire-BEH*, 2016, no 37-38, p. 681-690

ROGEL A. HAMERS F., QUINTIN C. [2016]. Incidence et dépistage du cancer du sein en France. Dernières données disponibles : octobre 2016 (rapport de Santé Publique France)

WALTER J.J., LIVINGSTONE S.J., COLHOUN H.M. [2011] Effect of socioeconomic status on mortality among people with type 2 diabetes: a study from the Scottish Diabetes Research Network Epidemiology Group. *Diabetes care*, 2011, vol. 34, no 5, p. 1127-1132