

MESURE DU RISQUE DE PERTE D'AUTONOMIE TOTALE EN FRANCE MÉTROPOLITAINE

Quentin
Guibert*

Frédéric
Planchet^α

Michaël
Schwarzinger^β

CEREMADE, UMR 7534, Université Paris-Dauphine, PSL Research University, 75016 Paris, France
 Univ Lyon - Université Claude Bernard Lyon 1, ISFA, Laboratoire SAF EA2429, F-69366, Lyon, France
 Translational Health Economics Network (THEN), 75010 Paris, France
 Infection, Antimicrobials, Modeling & Evolution (IAME), UMR 1137, INSERM - Université Paris Diderot, Sorbonne Paris
 Cité, Paris, France

Prim'Act, 42 avenue de la Grande Armée, 75017 Paris, France

Version 1.6 du 06/02/2019

Ce papier s'intéresse à la construction de loi d'incidence pour la perte d'autonomie totale sur la période 2010-2012 à partir des données des bases nationales d'hospitalisation (PMSI 2008-2013). Nos résultats sont décomposés selon deux types de dépendance : dépendance cognitive (ou démence) et dépendance physique. Les femmes présentent un risque légèrement plus important d'entrée en démence, alors que le risque lié à la dépendance physique est plus marqué chez les hommes. L'incidence en dépendance « toutes causes » est comparable entre hommes et femmes. Les résultats suggèrent un ralentissement de l'incidence aux delà de 90 ans et une convergence des hommes et des femmes aux grands âges. Les implications de ces résultats pour l'extrapolation aux grands âges sont discutées.

I.	INTRODUCTION	2
II.	DESCRIPTION DES DONNÉES UTILISÉES	3
III.	MÉTHODOLOGIES D'ESTIMATION ET DE LISSAGE	5
III.a.	Définition des taux d'incidence en dépendance totale.....	5
III.b.	Estimation des taux d'incidence bruts	6
III.c.	Redressement sur les expositions nationales	7
III.d.	Régularisation et extrapolation des lois d'incidence	9
IV.	TAUX D'INCIDENCE BRUTS ET LISSÉS	9
IV.a.	Démence et dépendance physique.....	9
IV.b.	Dépendance totale « toutes causes »	11
IV.c.	Validation de l'ajustement	12
V.	SENSIBILITÉ À LA MÉTHODE D'EXTRAPOLATION	13
VI.	CONCLUSION	16
VII.	RÉFÉRENCES	17

* Quentin Guibert est professeur associé à l'Université Paris Dauphine et affilié au CEREMADE, consultant au sein du cabinet Prim'Act et chercheur associé au laboratoire SAF (EA n°2429). Contact : q.guibert@hotmail.fr.

* Frédéric Planchet est chercheur au laboratoire SAF (EA n°2429) et consultant au sein du cabinet Prim'Act. Contact : frederic@planchet.net.

^α Michaël Schwarzinger est le fondateur/président de la société THEN et chercheur associé à l'Inserm. Contact : michael.schwarzinger@transhealthecon.net

^β Michaël Schwarzinger est le fondateur de la société THEN. Contact : michael.schwarzinger@transhealthecon.net
 Travaux réalisés pour le groupe d'étude **QalyDays**, qui inclut, outre les auteurs, Sylvain Baillot (THEN, Paris) ; Stéphane Luchini (CNRS GREQAM Marseille) ; Jürgen Rehm (Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON).

I. INTRODUCTION

L'objectif de ce travail est de proposer des lois d'incidence par sexe pour la population française en « bonne santé » pour le risque de perte d'autonomie totale (ou « dépendance totale »). Le lecteur pourra consulter sur ces aspects la présentation du risque dépendance et les enjeux de cette étude proposée dans le présent volume (Castaneda & Lussou, 2018; Schwarzingger, 2018).

Les techniques statistiques utilisées pour l'estimation des lois biométriques en assurance dépendance s'appuient classiquement sur des modèles multi-états (Christiansen, 2012; Deléglise *et al.*, 2009; Denuit & Robert, 2007; Haberman & Pitacco, 1998; Rickayzen & Walsh, 2002), dans un cadre markovien (Fleischmann, 2015; Fong *et al.*, 2015; Helms *et al.*, 2005; Levantesi & Menziatti, 2012; Pritchard, 2006), semi-markovien ou non markovien (Biessy, 2017; Denuit & Robert, 2007; Fuino & Wagner, 2018; Guibert & Planchet, 2018). Concernant l'incidence en dépendance, elle peut être traitée à partir de modèles à risques concurrents, vus comme un cas particulier de modèles multi-états (Dupourque *et al.*, 2019; Guibert & Planchet, 2014). Aussi, si les techniques d'estimation sont bien connues, leur application à l'estimation de lois d'incidence reste encore relativement rare compte tenu de la difficulté à disposer de données publiques ou dans le cadre d'assurances privées.

Dans cet article, nous appuyons sur la méthodologie développée dans Guibert & Planchet (2014) pour construire des lois d'incidence en dépendance totale par âge entier, à partir d'un modèle à risques concurrents à quatre états, distinguant l'incidence en dépendance cognitive (ou « démence ») ou sinon en dépendance physique (ou « état grabataire »). Les modèles sont estimés sur la période 2010-2012 dans une large population âgée de plus de 50 ans et en « bonne santé » (i.e., sans pathologie grave ou dépendance) identifiée à l'hôpital. Les résultats obtenus sont extrapolés à la population générale en « bonne santé » à partir des expositions nationales fournies par l'Insee sur la même période (Schwarzingger, 2018).

La décomposition du risque de dépendance totale apparaît comme relativement nouvelle, d'une part parce que les données issues de l'assurance privée ne permettent généralement pas de relier la dépendance à une cause médicale (exemple : utilisation de tests), et d'autre part compte tenu de la précision des définitions retenues. Les études précédentes en science actuarielle ont en effet cherché à décomposer le risque dépendance selon différents degrés de dépendance au sens des termes d'un contrat ou bien selon le type de prise en charge pour des données publiques (ex : prise en charge en institution, à domicile, ...). Le résultat principal de cette étude est donc de fournir des lois d'incidence par sexe et par type de dépendance totale.

Cet article s'articule de la manière suivante. La Section II décrit les données utilisées pour estimer le modèle, tandis que les méthodes d'estimation et de lissage des taux

d'incidence sont présentées dans la Section III. La Section IV décrit les résultats de cette étude et examine les lois estimées. La Section V fournit une analyse de sensibilité à la méthode d'extrapolation retenue. Enfin, la Section VI conclut cette étude en examinant les limites de notre approche et les perspectives à venir.

Les lois estimées sont disponibles en ligne sur la page <http://www.ressources-actuarielles.net/qalydays>.

II. DESCRIPTION DES DONNÉES UTILISÉES

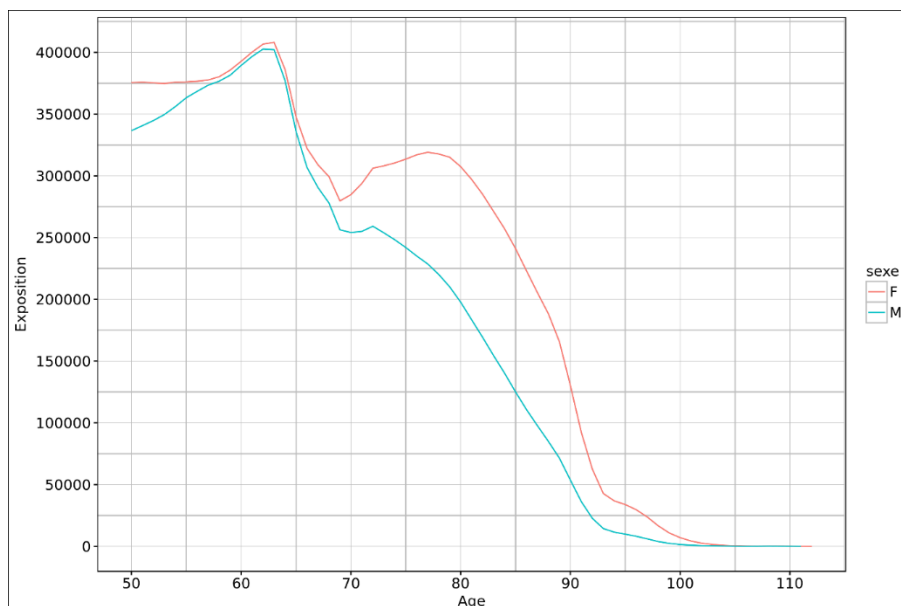
La présente étude s'appuie sur les données des bases nationales d'hospitalisation (PMSI) pour la population âgée de plus de 50 ans, résidant en France métropolitaine et en « bonne santé » (i.e., sans pathologie grave ou perte d'autonomie avérée) observée sur la période comprise entre le 01/01/2010 et le 31/12/2012. Compte tenu des volumes de données relativement importants, les estimations sont stratifiées selon le sexe.

Les définitions de la dépendance cognitive et physique retenues sont celles présentées par Schwarzinger (2018). Les décès pris en compte sont à la fois ceux observés à l'hôpital et ceux imputés à la sortie de l'hôpital selon la méthodologie décrite par la même étude.

Pour les données ainsi sélectionnées, nous observons individuellement l'âge (entier) de l'enregistrement d'une unique cause de dépendance totale (cognitive ou sinon physique), la survenance du décès ou bien l'âge de la dernière observation à l'hôpital. L'âge des patients est donc soumis à troncature à gauche (âge du premier enregistrement à l'hôpital) et à censure droite (date de la dernière observation ou fin de la période d'observation). L'imputation des décès et les investigations conduites par Schwarzinger (2018) sur l'identification des dépendants permettent de limiter le caractère informatif que pourrait prendre cette censure.

La Figure 1 présente la durée d'exposition au risque par âge pour les sous-populations féminine et masculine. Compte tenu, de la meilleure survie de la population féminine à partir de 70 ans, la population exposée pour les femmes apparaît plus étendue de plus de 21,1%, notamment pour la tranche d'âges 70-100 ans. Les volumes disponibles apparaissent significativement plus importants que ceux d'autres études équivalentes (Fong *et al.*, 2015; Fuino & Wagner, 2018; Guibert & Planchet, 2014), ce qui permet d'assurer une certaine robustesse aux résultats, notamment aux âges élevés où l'exposition est généralement plus difficile à obtenir.

Figure 1. Durée d'exposition aux risques (en année) pour les femmes (F) et les hommes (M) sur la période d'observation.



Le Tableau 1 fournit des statistiques concernant la survenance du risque de dépendance totale par sexe et selon le type de dépendance. Le risque démence survient en moyenne à 80 ans chez les hommes et plus de 3,6 ans après chez les femmes. Le nombre de cas observés est plus de deux fois plus importants pour celles-ci. La dépendance physique survient nettement moins souvent que la démence et concernent davantage les hommes jeunes.

Strate	Démence		Dépendance physique		Dépendance totale	
	Nombre survenances	Age survenance	Nombre survenances	Age survenance	Nombre survenances	Age survenance
Femmes	249 046	83,6 (7,7)	36 236	81,2 (11,3)	285 282	83,3 (8,3)
Hommes	122 953	80,0 (9,5)	26 850	75,8 (12,1)	149 803	79,2 (10,1)

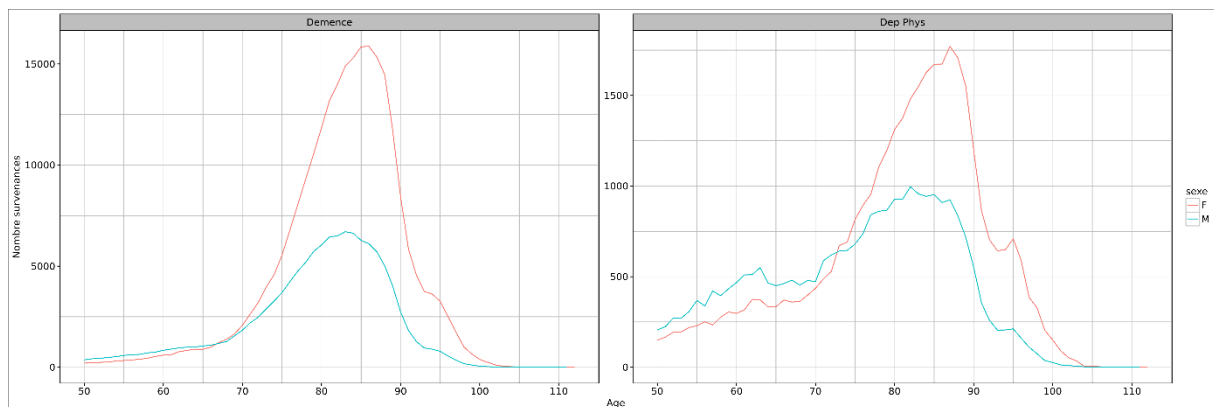
Tableau 1. Nombre d'incidences en dépendance par sexe et type de dépendance. La colonne âge de survenance précise l'âge de survenance moyen pour chaque type de dépendance avec l'écart-type correspondant entre parenthèse.

La Figure 2 illustre le nombre de survenances observées par âge entier selon le type de dépendance et selon le sexe sur l'ensemble de la période d'observation. Elle donne une idée plus précise des tranches d'âges concernées par les deux types de dépendance. S'agissant de la démence, pour une exposition comparable entre hommes et femmes, le nombre de survenances de démence reste quasiment identique jusqu'à 70 ans. Le nombre de survenances croît ensuite de manière importante jusqu'à 85 ans, mais avec une amplitude nettement plus forte pour la population féminine. Ceci ne traduit pas

nécessaire une hausse significative de l'incidence féminine, mais résulte davantage de la mortalité masculine qui est plus précoce et réduit fortement l'exposition aux risques à ces âges. Au-delà de 85 ans, la survenue de la démence est contrebalancée par l'augmentation des décès sans dépendance venant réduire fortement l'exposition aux risques.

La dépendance physique est caractérisée par d'avantage de survéances avant 75 ans chez les hommes, puis la situation s'inverse compte tenu de la plus grande exposition des femmes à ces âges.

Figure 2. Survenance de cas de démence et de dépendance physique (Dep Phys) pour les femmes (F) et les hommes (M) sur la période d'observation.



Ces premiers éléments font apparaitre deux risques de dépendance dont les propriétés sont différentes et qui affectent différemment hommes et femmes.

III. MÉTHODOLOGIES D'ESTIMATION ET DE LISSAGE

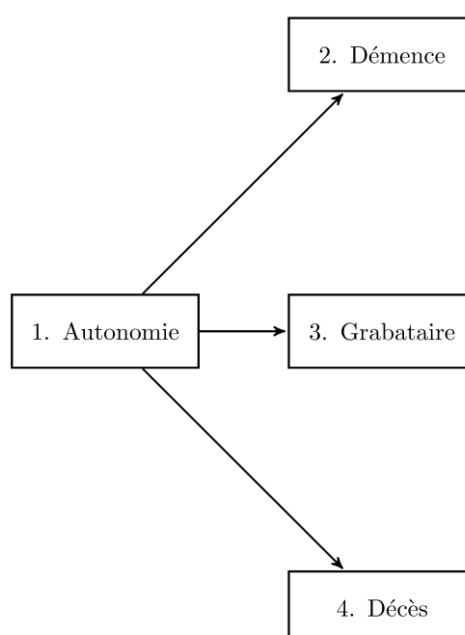
Cette section décrit le cadre retenu pour l'estimation des taux d'incidence pour la perte d'autonomie totale. Dans un premier temps, nous décrivons succinctement les cadres de modélisation et d'estimation des taux d'incidence bruts retenus en se référant pour cela à Guibert et Planchet (2014) et à Dupourque *et al.* (2019). Nous expliquons ensuite comment adaptés des taux, estimés sur la seule population observée à l'hôpital (Schwarzinger, 2018) à l'ensemble de la population nationale. Enfin, nous présentons l'approche retenue pour lisser et extrapoler les taux nationaux redressés.

III.a. Définition des taux d'incidence en dépendance totale

L'estimation des lois d'incidence en dépendance totale est examinée au moyen d'un modèle à risques concurrents (Andersen *et al.*, 1993, 2002). Nous introduisons pour cela un modèle multi-états $(X_t)_{t \geq 0}$ en temps continu prenant ses valeurs dans un espace d'états $\mathcal{S} = \{1, 2, 3, 4\}$ (Figure 3), où l'état 1 représente un état sans dépendance totale (ou autonomie), l'état 2 correspond à un état de dépendance cognitive(ou démence), l'état 3

correspond à un état de dépendance physique (ou état grabataire), et enfin l'état 4 correspond au décès sans dépendance totale. Ce dernier état est spécifiquement identifié dans le modèle puisque ne pas l'isoler comme une cause spécifique de sortie de l'état 1 conduirait à l'intégrer comme une cause de censure, potentiellement informative. Dans la suite de cette étude, nous nous focalisons plus particulièrement sur les états 2 et 3, ainsi que sur la réunion des deux, i.e. qualifiée de « dépendance totale ».

Figure 3. Modèle multi-états utilisé pour mesurer l'incidence en dépendance totale.



Nous notons également par T la durée de vie dans l'état 1 et $V = X_T$ la cause de sortie de cet état. Une première quantité d'intérêt est appelée fonction d'incidence cumulée et s'écrit pour tout $t \geq 0$ et $j \in \{2, 3, 4\}$

$$F^{(j)}(t) = \Pr(T \leq t, V = j).$$

Nous en déduisons alors les taux d'incidence annuels relative à une cause $j \in \{2, 3, 4\}$ à partir de la relation suivante

$$q^{(j)}(t) = \frac{F^{(j)}(t+1) - F^{(j)}(t)}{S(t)},$$

où $S(t) = \Pr(T > t)$ correspond à la fonction de survie associée à T .

III.b. Estimation des taux d'incidence bruts

En présence de données censurées à droite et tronquées à gauche indépendantes et non informatives, une estimation non paramétrique des taux d'incidence est conduite en

utilisant l'estimateur d'Aalen-Johansen pour un modèle à risques concurrents (Andersen *et al.*, 1993). Nous observons pour cela les grandeurs suivantes

$$\begin{cases} \tilde{T} = T \wedge C \\ \gamma = \mathbf{1}_{\{T \leq C\}} \\ L \\ \gamma V \end{cases},$$

où \tilde{T} correspond à la durée observée dans l'état 1, C correspond à la variable de censure droite, L est la date de troncature gauche et γ est l'indicatrice de non-censure. Les taux d'incidence annuels bruts sont alors estimés pour une population de taille n par

$$\hat{q}^{(j)}(t) = \sum_{\{i: t < \tilde{T}_i \leq t+1\}} \frac{\hat{S}(\tilde{T}_i -) \Delta N_{1j}(\tilde{T}_i)}{\hat{S}(t) Y(\tilde{T}_i)}, \quad j = 2, 3, t \geq 0,$$

où $Y(t) = \sum_{k=1}^n \mathbf{1}_{\{L_k < t \leq \tilde{T}_k\}}$ correspond à l'exposition au risque, $N_{1j}(t) = \sum_{k=1}^n \mathbf{1}_{\{L_k < T_k \leq C_k, \tilde{T}_k \geq t, V_k = j\}}$ est le nombre de sortie dues à la cause $j = 2, 3$, et $\hat{S}(t)$ est l'estimateur de Kaplan-Meier de la fonction de survie (Kaplan & Meier, 1958)

$$\hat{S}(t) = \prod_{\{i: \tilde{T}_i \leq t\}} \left(1 - \frac{\Delta N_1(\tilde{T}_i)}{Y(\tilde{T}_i)} \right),$$

avec $N_1(t) = \sum_{j=1}^3 N_{1j}(t)$. Les estimateurs classiques de type Greenwood de l'écart-type $\hat{\sigma}^{(j)}(t)$ associés aux estimateurs des taux bruts sont considérés pour la construction d'intervalles de confiance de niveau α dont les bornes sont approximées par (Andersen *et al.*, 1993)

$$\hat{q}^{(j)}(t) \exp \left(\pm \Phi^{-1} \left(\frac{\alpha \hat{\sigma}^{(j)}(t)}{2 \hat{q}^{(j)}(t)} \right) \right),$$

avec Φ la fonction de répartition d'une loi normale centrée réduite.

III.c. Redressement sur les expositions nationales

Les estimateurs des taux d'incidence sont estimés sur la population en « bonne santé » observée à l'hôpital. Ils doivent cependant être ajustés pour pouvoir être extrapolés à la population générale en « bonne santé ». La sélection des observations à l'hôpital conduit en effet à reconsidérer les expositions au risque d'une part, et le nombre d'individus entrés en dépendance totale d'autre part.

Concernant les durées d'exposition au risque, les individus qui ne sont pas observés à l'hôpital sont a priori en « bonne santé » (i.e., sans pathologie grave nécessitant une hospitalisation). Ainsi, le nombre total d'individus exposés au risque peut être aisément redressé en utilisant les durées d'expositions au risque par âge entier calculées par l'Insee $E^{ref}(t)$ après décompte des individus avec pathologie(s) grave(s) identifiés à l'hôpital (Schwarzinger, 2018). Nous notons $E(t) = \sum_{i=1}^n e_i(t)$, la durée d'exposition au risque à l'âge entier $t \geq 0$ de la population enregistrée à l'hôpital, avec pour exposition individuelle $e_i(t) = (\tilde{T}_k \wedge t - (t-1))^+$. Ainsi, le nombre d'individus réellement exposés au risque pour la population générale est estimé au moyen de la formule suivante

$$Y^{ref}(t) = Y(t) \frac{E^{ref}(t)}{E(t)}.$$

Concernant la survenance de la dépendance totale, il n'existe pas à notre connaissance de données françaises informatives sur la couverture du risque à l'hôpital. Nous faisons l'hypothèse que le risque dépendance est observé de manière exhaustive à l'hôpital sur la période considérée. Plusieurs arguments permettent d'étayer cette hypothèse.

Concernant la démence, il convient de noter que les taux d'incidence estimés à partir des bases nationales d'hospitalisation sont supérieurs entre 65 et 74 ans à ceux de la cohorte de référence en France (Letenneur *et al.*, 1999), puis dépendent du sexe et de l'âge tout en notant les fortes fluctuations d'échantillonnage dans la cohorte de référence après 75 ans¹. Par ailleurs, une étude de la prévalence de la démence en 2014 à partir de l'ensemble des données médico-administratives (dont le PMSI) a montré que 73,5 % des cas sont identifiées à l'hôpital, 14,4 % par la prescription d'un médicament anticholinestérasique² et 12,1% par une inscription en « affection de longue durée » (Carcaillon-Bentata *et al.*, 2016).

Concernant la dépendance physique (identifiée ici sans démence), elle complique toujours l'évolution d'une ou plusieurs pathologies graves qui nécessitent l'hospitalisation. Cependant la survie en dépendance physique est réduite avec une très forte mortalité les premiers mois (Guibert *et al.*, 2018). Ainsi, il est possible qu'une proportion (indéterminée) d'individus en fin de vie sortent de l'hôpital en dépendance partielle (non observée) et évoluent rapidement vers la dépendance totale (e.g., en EHPAD) sans que la courte durée dans l'état permette son observation (i.e., au cours d'une ré hospitalisation).

¹ La cohorte PAQUID a inclus 3 777 personnes âgées de plus de 65 ans et résidant dans les départements de Dordogne et Gironde.

² La prescription de médicaments anticholinestérasiques est autorisée pour les seules maladies d'Alzheimer avec un déficit cognitif léger à modéré et ne concerne donc pas la dépendance totale (déficit cognitif sévère). L'inscription exclusive en affection de longue durée (sans aucune hospitalisation) ne permet pas de connaître le niveau du déficit cognitif.

En conclusion, les deux hypothèses utilisées pour le redressement des estimations conduisent :

- à un réajustement du nombre d'individus exposés au risque à chaque âge entier, en supposant que le ratio nombre d'individus exposés à l'hôpital sur population générale en bonne santé est égal à celui des durées d'exposition par tranche d'âge annuelle ;
- à une possible sous-estimation du nombre d'individus entrés en dépendance totale (notamment physique) en se limitant aux cas les plus graves ayant nécessités une hospitalisation.

III.d. Régularisation et extrapolation des lois d'incidence

Les taux d'incidence bruts sont lissés par une méthode de splines cubiques (Chambers & Hastie, 1992; Delwarde & Denuit, 2005), ce qui permet également d'extrapoler les résultats. Comme tenu des données disponibles, les taux d'incidence bruts sont lissés sur la tranche d'âges 50-95 ans en pondérant le poids de chaque point à partir des durées exposition au risque, puis ils sont extrapolés entre 96 et 100 ans³.

IV. TAUX D'INCIDENCE BRUTS ET LISSÉS

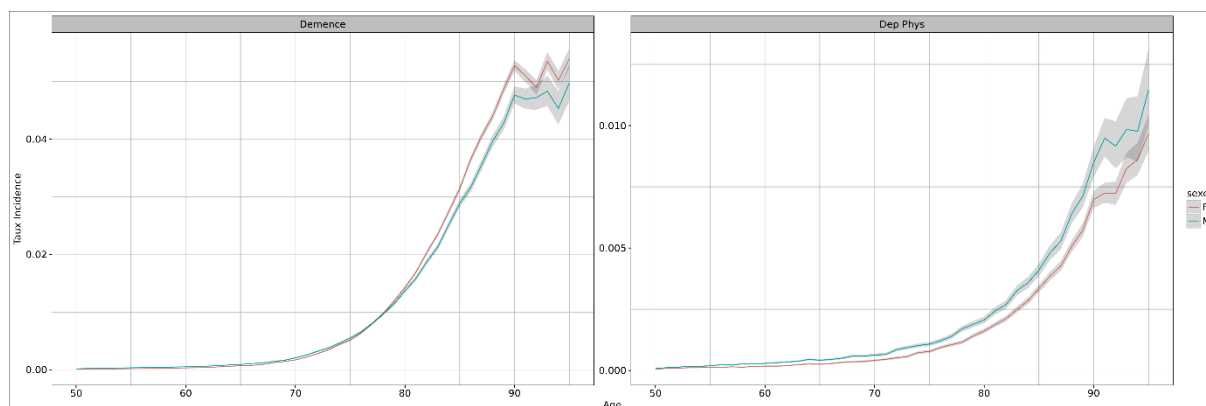
Dans cette section, nous décrivons les taux bruts estimés après redressement sur les expositions nationales, puis les taux d'incidence lissés pour les différents types de perte d'autonomie. Les estimations ont été réalisés à partir dans le langage R (R Core Team, 2018) et utilisent le package *mstate* (De Wreede *et al.*, 2011).

IV.a. Démence et dépendance physique

En suivant la méthode présentée à la Section III, la Figure 4 présente les taux bruts pour l'incidence en démence et en dépendance physique selon le sexe, ainsi que les intervalles de confiance à 95 % correspondants. Les volumes de données disponibles permettent d'obtenir des estimations relativement robustes des taux d'incidence sur la plage d'âges 50-95 ans.

³ Pour souci de simplicité, un critère de lissage unique égale à 0,5 est retenu. Il permet d'obtenir un bon compromis en régularité et fidélité aux données.

Figure 4. Taux d'incidence bruts redressés pour la démence et la dépendance physique pour les femmes (F) et les hommes (M) sur la période d'observation et intervalles de confiance à 95 %.



Ces taux bruts sont ensuite lissés et extrapolés jusqu'à 100 ans à partir de la méthode décrite dans la Section III.d. Les lois ainsi obtenues sont représentées à la Figure 5. Les fonctions d'incidence cumulée sont quant à elles représentées à la Figure 6.

Figure 5. Taux d'incidence lissés et extrapolés pour la démence et la dépendance physique pour les femmes (F) et les hommes (M) sur la période d'observation.

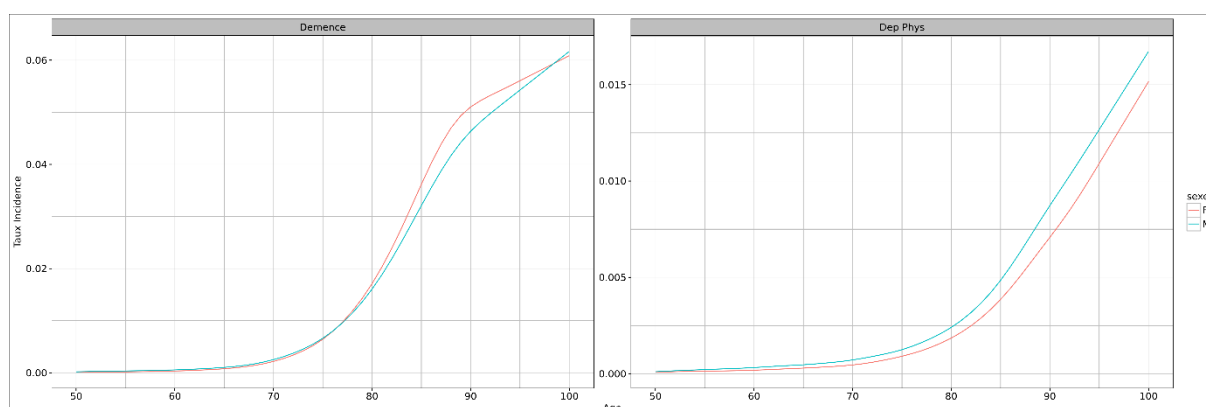
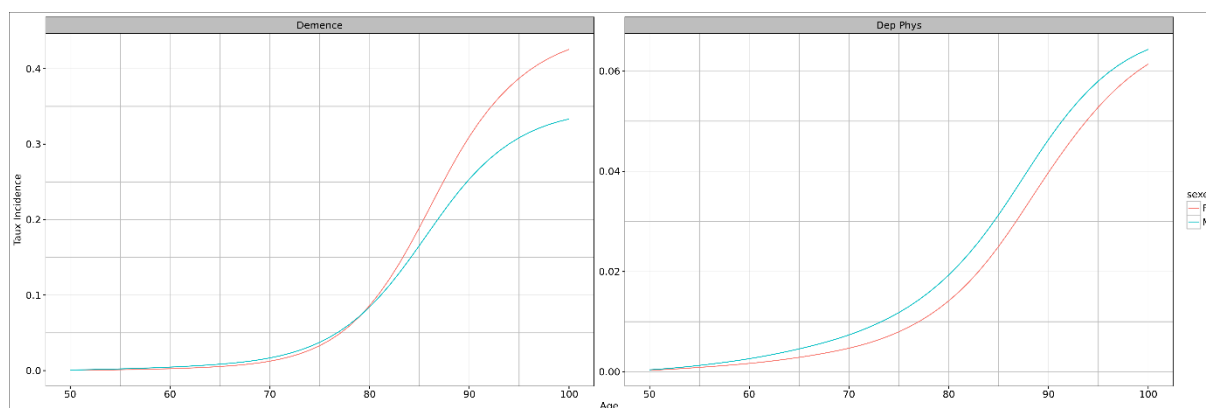


Figure 6. Incidences cumulées à 50 ans pour la démence et la dépendance physique pour les femmes (F) et les hommes (M) sur la période d'observation.



Les taux incidences annuels ont caractérisés par un net accroissement à partir de 75-85 ans, ce qui est conforme à littérature (Fong *et al.*, 2015; Fuino & Wagner, 2018; Guibert & Planchet, 2014). Par ailleurs, un ralentissement de l'incidence en démence est observé à partir de 90 ans, alors que celle en dépendance physique semble croître au même rythme aux grands âges. Ce ralentissement est observés sur les taux bruts dans certaines études, sans pour autant être attribués à la démence (Fong *et al.*, 2015; Fuino & Wagner, 2018). Il est également retrouvé dans Schwarzinger *et al.* (2018) avec une définition similaire de la démence.

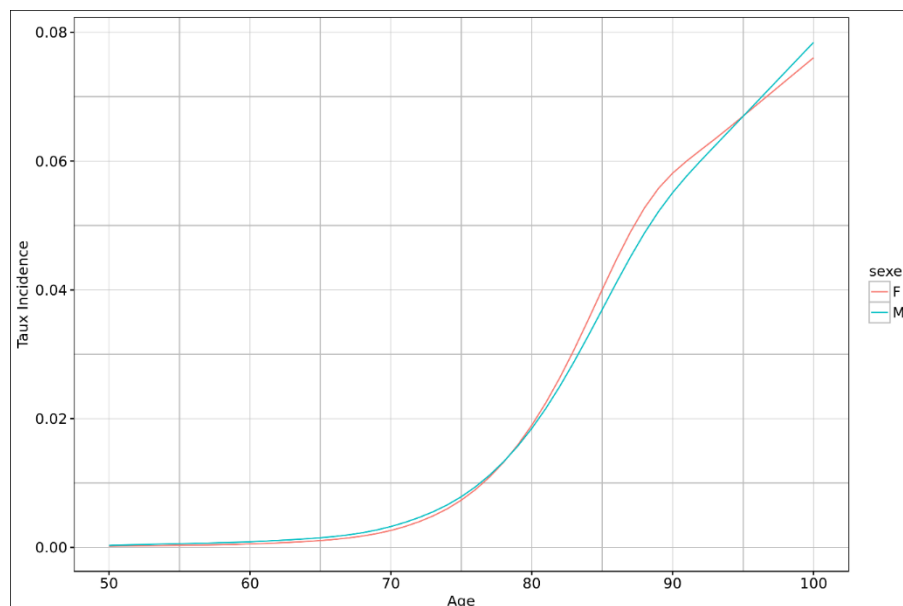
La méthode d'extrapolation retenue permet de prolonger ces tendances et conduit à faire converger le taux d'incidence des hommes en démence vers celui des femmes après 95 ans. Ce faible écart entre les hommes et les femmes est à souligner et est cohérent avec plusieurs études récentes dans le domaine médical (Yang *et al.*, 2013) et en science actuarielle (Biessy, 2016). Il permet d'expliquer la plus faible prévalence de la démence chez les hommes par une plus faible survie, notamment aux âges avancés (cf. Section II).

La décomposition de la dépendance totale entre les composantes physique et cognitive permet de faire apparaître une exposition différente des hommes et des femmes à ces deux risques. Les femmes sont en effet caractérisées par une incidence de la démence légèrement supérieure à celle des hommes, ce qui se traduit néanmoins par une probabilité d'être affecté par la démence beaucoup plus importante pour les femmes au cours de leur vie compte tenu de leur plus grande longévité (Figure 6). Ainsi, la probabilité d'être affecté par une démence est estimée à 42,6 % jusqu'à 100 ans pour une femme et 33,3 % pour un homme. Les écarts sont plus faibles pour la dépendance physique : 6,1 % jusqu'à 100 ans pour une femme et 6,4% pour un homme.

IV.b. Dépendance totale « toutes causes »

L'incidence en dépendance totale est obtenue par sommation des taux d'incidence en démence et en dépendance physique (Figure 7). Les compensations observées entre les deux types de dépendance conduisent des taux d'incidence en dépendance relativement proches entre les populations masculines et féminines, avec une prédominance de la dépendance physique aux âges jeunes, puis une augmentation de la part de la démence au-delà de 80 ans.

Figure 7. Taux d'incidence lissés et extrapolés pour la dépendance « toutes causes » pour les femmes (F) et les hommes (M) sur la période d'observation.



IV.c. Validation de l'ajustement

Nous examinons enfin la pertinence de l'ajustement effectué sur la tranche d'âges 50-95 ans en comparant les entrées en dépendance physique et en démence observés et celles prévues par le modèle. Nous utilisons pour ce faire l'indicateur *Standardized Mortality Ratio* (SMR), correspondant au ratio du nombre de cas observés par rapport au nombre de de cas prédits. Cette comparaison est effectuée par type de dépendance et par sexe. Nous examinons également la capacité du modèle à restituer le nombre de cas observés âge par âge.

Les ratios SMR observés pour la démence sont 99,9% pour les femmes et 99,6 % pour les hommes. La Figure 8 représente les entrées en démence observées et prédites sur la période d'observation et confirme le très bon ajustement du modèle aux données observés. S'agissant de la dépendance physique, ces ratios s'établissent à 99,6 % pour les femmes et 99,2 % pour les hommes et la Figure 9 illustre un constat similaire pour ce type de dépendance.

Figure 8. Comparaison des nombres d'entrées en démence observées et prédites.

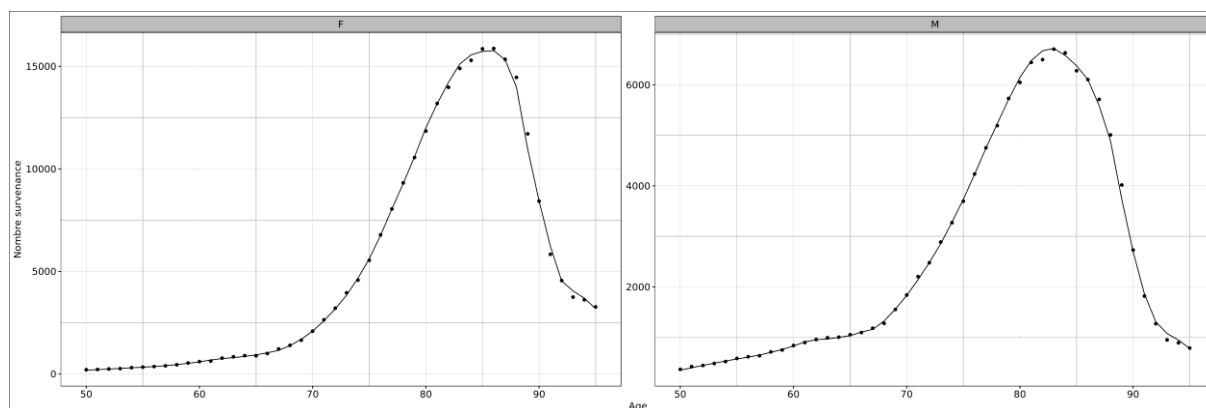
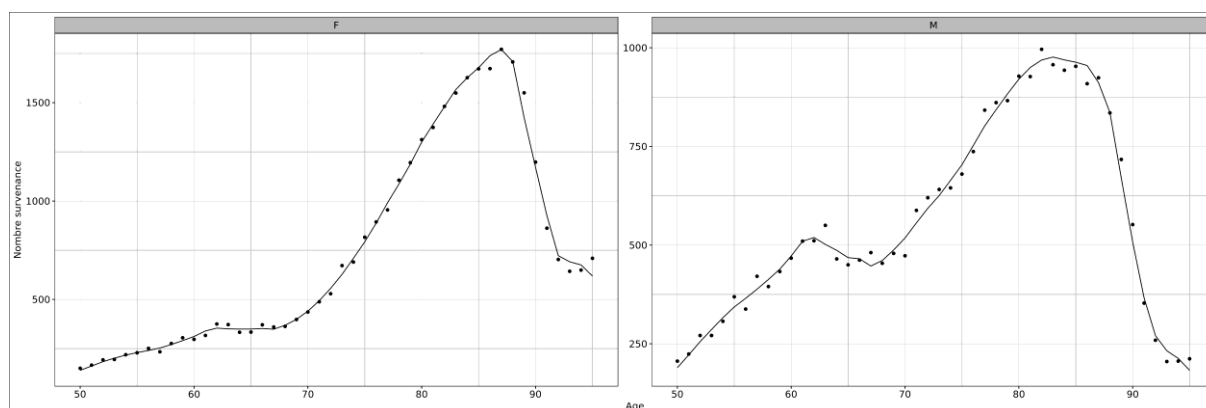


Figure 9. Comparaison des nombres d'entrées en dépendance physique observées et prédites.



V. SENSIBILITÉ À LA MÉTHODE D'EXTRAPOLATION

La question de l'extrapolation des taux d'incidence en démence et en dépendance physique est une question importante dans la perspective globale du vieillissement de la population. L'évolution de la prévalence et de l'incidence de la démence aux grands âges a été largement discutée (Slavin *et al.*, 2013), certaines études suggérant un recul du déclin cognitif aux âges élevés et des inégalités importantes entre hommes et femmes (Kliegel *et al.*, 2004). L'hypothèse la plus fréquemment retenue prévoit de prolonger la croissance exponentielle des taux d'incidence en démence ou pour la maladie d'Alzheimer (Corrada *et al.*, 2010; Grasset *et al.*, 2016; Rocca *et al.*, 2011). Les tendances obtenues présentent cependant des écarts importants sur le niveau des taux d'incidence, ce qui peut s'expliquer par les faibles volumes observés, le mode de sélection des individus ou encore l'hétérogénéité des tests cognitifs effectués pour diagnostiquer la démence (Brodaty *et al.*, 2016).

La dépendance physique peut compliquer de nombreuses pathologies graves dont la contribution évolue avec l'âge et le sexe. De ce point de vue, il apparaît plus délicat de trouver une littérature médicale ou en biostatistiques qui examinerait une définition comparable de cette forme de dépendance.

Dans la littérature actuarielle, les taux d'incidence lissés ou extrapolés aux grands âges suivent généralement une croissance exponentielle. Au vu de ces éléments et du ralentissement observés à partir de 90 ans dans la Section IV, nous examinons dans cette section l'impact d'une extrapolation des taux d'incidence à partir de 90 ans. En constatant que les données disponibles dans d'autres études ne permettent généralement pas de disposer d'informations suffisamment volumineuses au-delà de 90 ans, nous comparons les résultats décrits dans la Section IV à ceux obtenus par des approches que les praticiens pourraient légitimement retenir faute de meilleure information. Nous considérons dès lors deux approches :

- une extrapolation des taux d'incidence par splines cubiques à partir de 90 ans, i.e. via la méthodologie décrit dans la Section III.d ;
- une extrapolation des taux de d'incidence à partir de 90 ans via le modèle de Kannisto (Thatcher *et al.*, 1998), classiquement utilisé pour extrapoler les taux de décès $\mu(x)$ à l'âge x tel que

$$\mu(x) = c + \frac{ae^{bx}}{1 + ae^{bx}},$$

avec (a, b, c) les paramètres du modèle, que nous transposons aux taux d'incidence.

La Figure 10 illustre les écarts observés entre les taux d'incidence en démence obtenus en utilisant les données disponibles jusqu'à 95 ans et les deux alternatives avec des données jusqu'à 90 ans. L'utilisation de splines cubiques jusqu'à 90 ans conduit à une extrapolation quasi-linéaire des taux d'incidence aux grands âges, surestimant les taux de référence au moins jusqu'à 95 ans. L'extrapolation des taux par la méthode de Kannisto produit quant à elle une croissance exponentielle des taux d'incidence, nettement supérieure à celle observée et dont la forme est comparable à certains travaux pour la dépendance toutes causes (Fong *et al.*, 2015; Fuino & Wagner, 2018; Pitacco, 2016) ou pour la démence (Grasset *et al.*, 2016). Cette approche conduit également à sous-estimer légèrement l'incidence en démence avant 87 ans. S'agissant de la démence physique (Figure 11), des résultats comparables sont observés mais avec des écarts entre les taux d'incidence de référence et les deux alternatives considérées nettement plus réduits.

Figure 10. Taux d'incidence lissés et extrapolés après 95 ans par splines cubiques, et 90 ans par splines cubiques et par le modèle de Kannisto pour la démence pour les femmes (F) et les hommes (M).

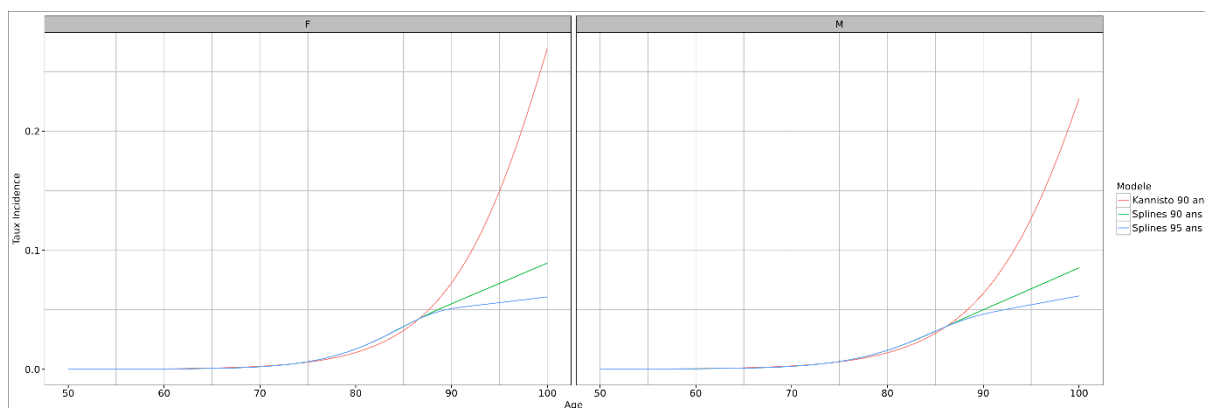
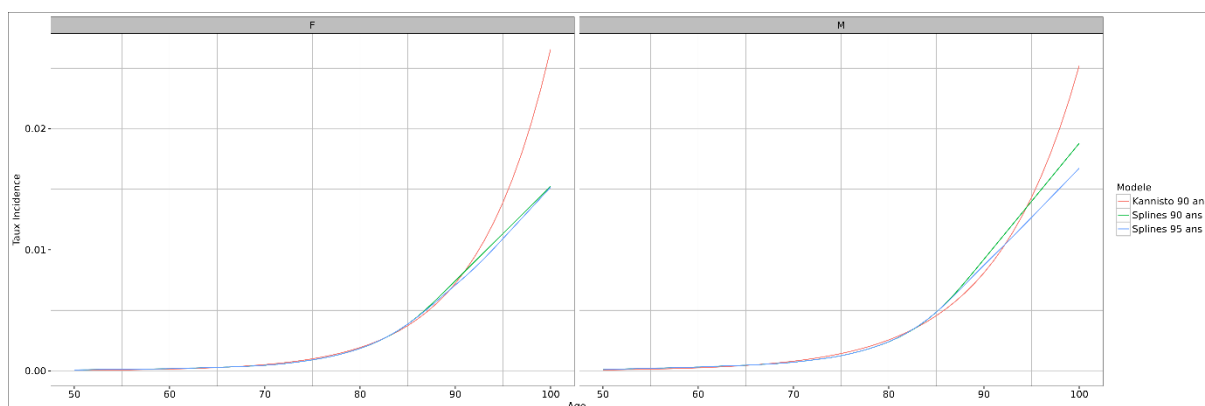
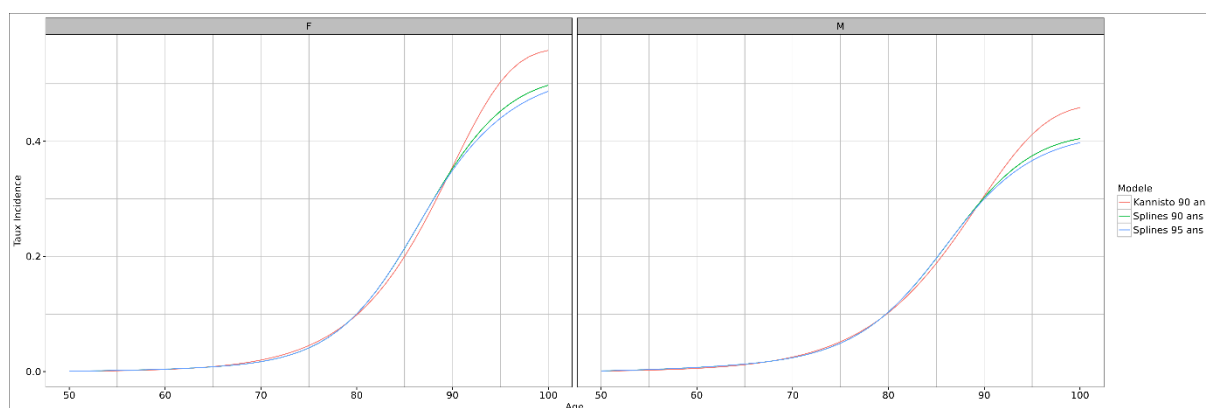


Figure 11. Taux d'incidence lissés et extrapolés après 95 ans par splines cubiques, et 90 ans par splines cubiques et par le modèle de Kannisto pour la dépendance physique pour les femmes (F) et les hommes (M).



L'effet de l'extrapolation aux grands âges dépend de la population exposée à ces âges. Pour le mesurer, nous examinons l'impact sur la fonction d'incidence cumulée pour la dépendance totale toute cause confondue, en notant que la majeure partie de l'effet provient de la démence (Figure 12). Malgré une exposition relativement faible de la population aux delà de 90 ans, la surestimation de l'incidence en démence par l'extrapolation de Kannisto conduit à une incidence cumulée de 55,8 % à 100 ans pour les femmes et de 45,9 % pour les hommes contre 48,7 % pour les femmes et 39,8 % pour les hommes avec l'approche de référence utilisant des splines cubiques jusqu'à 95 ans, soit une surestimation d'environ 15 %. Les deux approches utilisant de splines cubiques conduisent quant à elles à des résultats similaires.

Figure 12. Incidences cumulées à 50 ans pour la démence totale pour les femmes (F) et les hommes (M) obtenues par splines cubiques à partir de 90 ans et 95 ans et par le modèle de Kannisto.



Nos résultats tendent à montrer que la question de l'extrapolation aux grands âges de la du risque de démence ne doit pas être négligée dans la perspective d'applications actuarielles. La seule extrapolation de type exponentielle, fréquente dans la littérature actuarielle lorsque les modèles utilisés sont paramétriques, conduirait pour nos données à introduire une couche de prudence significative dans les tarifs ou le provisionnement. Dans le cadre de notre étude, l'extrapolation linéaire des tendances semble être un meilleur compromis pour pallier le manque d'information, en introduisant un biais limité tout en restant extrêmement simple.

VI. CONCLUSION

Ce travail s'attache à fournir des lois d'incidence pour la perte d'autonomie totale par sexe et par type de dépendance (cognitive ou sinon physique). Les lois construites peuvent ainsi être utiles pour projeter la survenance de la dépendance pour la population française jusqu'à des âges avancés. La méthodologie décrite pour mettre en œuvre cette construction s'appuie sur des volumes de données importants et utilise des estimations non paramétriques des taux d'incidence, ce qui permet d'obtenir des résultats pour lesquels les biais d'estimation sont limités. L'extrapolation de ces estimateurs, s'appuyant sur la population observée à l'hôpital, à la population nationale s'appuie sur l'hypothèse d'une identification exhaustive de la dépendance totale à l'hôpital.

Nos résultats viennent compléter plusieurs études réalisées en France sur des données d'assurances privées (Biessy, 2017; Guibert & Planchet, 2014). Les volumes de données dont nous disposons nous permettent d'examiner l'évolution des taux d'incidence aux grands âges, et font apparaître notamment un ralentissement de l'incidence en démence à partir de 90 ans. Ces travaux permettent ainsi de nuancer l'hypothèse habituelle consistant à extrapoler de manière exponentielle les taux d'incidence aux âges élevés, qui peut avoir des impacts notables dans la perspective d'applications actuarielles. Nous

montrons en outre que les taux d'incidence sont relativement comparables pour les hommes et les femmes, ce qui n'est pas le cas dans toutes les études (Fong *et al.*, 2015; Fuino & Wagner, 2018).

Une des limites principales de notre étude concerne son caractère instantané, les données étant observées sur la période 2010-2012. Les études récentes semblent mettre en évidence une tendance à la baisse de l'incidence en démence dans les pays à haut revenu (Prince *et al.*, 2016). Bien que les populations d'étude soient peu nombreuses et hétérogènes, ces tendances doivent être confirmées et font apparaître la nécessité de recherches plus approfondies sur cette thématique, notamment aux âges avancés.

REMERCIEMENTS

Les auteurs remercient Fabio Castaneda et Michel Denuit pour leurs conseils avisés et leurs recommandations. Ils remercient également l'UMR 1137 (IAME) de l'INSERM - Université Paris Diderot pour la mise à disposition de serveurs de calcul sécurisés ayant permis la réalisation des calculs.

VII. RÉFÉRENCES

ANDERSEN, P. K., ABILDSTROM, S. Z. & ROSTHØJ, S., (2002), Competing risks as a multi-state model, *Statistical Methods in Medical Research*, vol. 11, n°2, p. 203-215.

ANDERSEN, P. K., BORGAN, ØRNULF, GILL, R. D. & KEIDING, N., (1993), *Statistical Models Based on Counting Processes*, Springer-Verlag New York Inc., 767 p.

BIESSY, G., (2016, août), *A semi-Markov model with pathologies for Long-Term Care Insurance*.

BIESSY, G., (2017), Continuous-time semi-Markov inference of biometric laws associated with a long-term insurance portfolio, *ASTIN Bulletin: The Journal of the IAA*, vol. 47, n°2, p. 527-561.

BRODATY, H., WOOLF, C., ANDERSEN, S., BARZILAI, N., BRAYNE, C., CHEUNG, K. S.-L., ... SACHDEV, P. S., (2016), ICC-dementia (International Centenarian Consortium - dementia): an international consortium to determine the prevalence and incidence of dementia in centenarians across diverse ethnoracial and sociocultural groups, *BMC Neurology*, vol. 16, n°1, p. 52.

CARCAILLON-BENTATA, L., QUINTIN, C., MOUTENGOU, É., BOUSSAC-ZAREBSKA, M., MOISAN, F., HA, C. & ELBAZ, A., (2016), Peut-on estimer la prévalence de la maladie d'Alzheimer et autres démences à partir des bases de données médico-administratives ? Comparaison aux données de cohortes populationnelles, *Bulletin épidémiologique hebdomadaire*, n°28-29, p. 459-67.

CASTANEDA, F. & LUSSON, F., (2018), Un panorama de l'assurance dépendance en France, *Bulletin Français d'Actuariat*, vol. 18, n°35, .

CHAMBERS, J. M. & HASTIE, T. J., (1992), *Statistical Models in S*, Wadsworth & Brooks/Cole, 600 p.

CHRISTIANSEN, M. C., (2012), Multistate models in health insurance, *Advances in Statistical Analysis*, vol. 96, n°2, p. 155-186.

CORRADA, M. M., BROOKMEYER, R., PAGANINI-HILL, A., BERLAU, D. & KAWAS, C. H., (2010), Dementia incidence continues to increase with age in the oldest old: The 90+ study, *Annals of Neurology*, vol. 67, n°1, p. 114-121.

DE WREEDE, L. C., FIOCCO, M. & PUTTER, H., (2011), mstate: An R Package for the Analysis of Competing Risks and Multi-State Models, *Journal of Statistical Software*, vol. 38, n°7, .

DELÉGLISE, M. P., HESS, C. & NOUET, S., (2009), Tarification, provisionnement et pilotage d'un portefeuille Dépendance, *Bulletin Français d'Actuariat*, vol. 9, n°17, p. 70-108.

DELWARDE, A. & DENUIT, M., (2005), *Construction de tables de mortalité périodiques et prospectives*, Economica, Paris, 428 p.

DENUIT, M. & ROBERT, C., (2007), *Actuariat des assurances de Personnes - Modélisation, tarification et provisionnement*, Economica, Paris, 405 p.

DUPOURQUE, E., PLANCHET, F. & SATOR, N. (ÉD.), (2019), *Actuarial Aspects of Long-Term Care*, Springer Actuarial Series, Springer.

FLEISCHMANN, A., (2015), Calibrating intensities for long-term care multiple-state Markov insurance model, *European Actuarial Journal*, p. 1-28.

FONG, J. H., SHAO, A. W. & SHERRIS, M., (2015), Multistate Actuarial Models of Functional Disability, *North American Actuarial Journal*, vol. 19, n°1, p. 41-59.

FUINO, M. & WAGNER, J., (2018), Long-term care models and dependence probability tables by acuity level: New empirical evidence from Switzerland, *Insurance: Mathematics and Economics*, vol. 81, p. 51-70.

GRASSET, L., BRAYNE, C., JOLY, P., JACQMIN-GADDA, H., PERES, K., FOUBERT-SAMIER, A., ... HELMER, C., (2016), Trends in dementia incidence: Evolution over a 10-year period in France, *Alzheimer's & Dementia*, vol. 12, n°3, p. 272-280.

GUIBERT, Q. & PLANCHET, F., (2014), Construction de lois d'expérience en présence d'événements concurrents – Application à l'estimation des lois d'incidence d'un contrat dépendance, *Bulletin Français d'Actuariat*, vol. 13, n°27, p. 5-28.

GUIBERT, Q. & PLANCHET, F., (2018), Non-parametric inference of transition probabilities based on Aalen-Johansen integral estimators for acyclic multi-state models: application to LTC insurance, *Insurance: Mathematics and Economics*, vol. 82, p. 21-36.

GUIBERT, Q., PLANCHET, F. & SCHWARZINGER, M., (2018), Mesure de l'espérance de vie en dépendance totale en France, *Bulletin Français d'Actuariat*, vol. 18, n°35, .

HABERMAN, S. & PITACCO, E., (1998), *Actuarial Models for Disability Insurance*, Chapman and Hall/CRC, 280 p.

HELMS, F., CZADO, C. & GSCHLÖBL, S., (2005), Calculation of LTC Premiums Based on Direct Estimates of Transition Probabilities, *ASTIN Bulletin: The Journal of the International Actuarial Association*, vol. 35, n°02, p. 455–469.

KAPLAN, E. L. & MEIER, P., (1958), Nonparametric Estimation from Incomplete Observations, *Journal of the American Statistical Association*, vol. 53, n°282, p. 457-481.

KLIEGEL, M., MOOR, C. & ROTT, C., (2004), Cognitive status and development in the oldest old: a longitudinal analysis from the Heidelberg Centenarian Study, *Archives of Gerontology and Geriatrics*, vol. 39, n°2, p. 143-156.

LETENNEUR, L., GILLERON, V., COMMENGES, D., HELMER, C., ORGOGOZO, J. M. & DARTIGUES, J. F., (1999), Are sex and educational level independent predictors of dementia and Alzheimer's disease? Incidence data from the PAQUID project, *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 66, n°2, p. 177-183.

LEVANTESI, S. & MENZIETTI, M., (2012), Managing longevity and disability risks in life annuities with long term care, *Insurance: Mathematics and Economics*, vol. 50, n°3, p. 391-401.

PITACCO, E., (2016), Premiums for Long-Term Care Insurance Packages: Sensitivity with Respect to Biometric Assumptions, *Risks*, vol. 4, n°1, p. 3.

PRINCE, M., ALI, G.-C., GUERCHET, M., PRINA, A. M., ALBANESE, E. & WU, Y.-T., (2016), Recent global trends in the prevalence and incidence of dementia, and survival with dementia, *Alzheimer's Research & Therapy*, vol. 8, n°1, p. 23.

PRITCHARD, D. J., (2006), Modeling disability in long-term care insurance, *North American Actuarial Journal*, vol. 10, n°4, p. 48-75.

R CORE TEAM, (2018), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

RICKAYZEN, B. D. & WALSH, D. E. P., (2002), A Multi-State Model of Disability for the United Kingdom: Implications for Future Need for Long-Term Care for the Elderly, *British Actuarial Journal*, vol. 8, n°2, p. 341-393.

ROCCA, W. A., PETERSEN, R. C., KNOPMAN, D. S., HEBERT, L. E., EVANS, D. A., HALL, K. S., ... WHITE, L. R., (2011), Trends in the incidence and prevalence of Alzheimer's disease, dementia, and cognitive impairment in the United States, *Alzheimer's & dementia: the journal of the Alzheimer's Association*, vol. 7, n°1, p. 80-93.

SCHWARZINGER, M., (2018), Étude QalyDays : données source et retraitements pour l'étude du risque de perte d'autonomie, *Bulletin Français d'Actuariat*, vol. 18, n°35, .

SCHWARZINGER, M., POLLOCK, B. G., HASAN, O. S. M., DUFOUIL, C., REHM, J., BAILLOT, S., ... LUCHINI, S., (2018), Contribution of alcohol use disorders to the burden of dementia in France 2008–13: a nationwide retrospective cohort study, *The Lancet Public Health*, vol. 3, n°3, p. e124-e132.

SLAVIN, M. J., BRODATY, H. & SACHDEV, P. S., (2013), Challenges of Diagnosing Dementia in the Oldest Old Population, *The Journals of Gerontology: Series A*, vol. 68, n°9, p. 1103-1111.

THATCHER, A. R., KANNISTO, V. & VAUPEL, J. W., (1998), *The Force of Mortality at Ages 80–120*, Odense: Odense University Press.

YANG, Z., SLAVIN, M. J. & SACHDEV, P. S., (2013), Dementia in the oldest old, *Nature Reviews Neurology*, vol. 9, n°7, p. 382-393.