

PROJECTIONS D'INDICATEURS DE SANTÉ POUR MALADIES CHRONIQUES: MODÈLE ILLNESS-DEATH ET HYPOTHÈSE SEMI-MARKOVIENNE

Mathilde Wanneveich ^{1,2,*} & Hélène Jacqmin-Gadda ^{1,2} & Pierre Joly ^{1,2}

¹ *University of Bordeaux, ISPED, Centre Inserm U1219, 33000 Bordeaux*

² *Inserm, ISPED, Centre Inserm U1219, 33000 Bordeaux*

* *mathilde.wanneveich@isped.u-bordeaux2.fr*

Résumé. Les maladies chroniques font parties des problèmes de santé publique d'un intérêt grandissant. Leur fardeau économique social et démographique est alarmant pour les années à venir. Jusqu'à présent, la méthode utilisée pour faire des projections afin d'évaluer le fardeau futur de ces maladies utilisait des hypothèses Markoviennes sous un modèle 'illness death'. L'âge et le temps calendaire ont été pris en compte dans l'estimation des paramètres, cependant le temps passé avec la maladie n'a pas été considéré. Ce travail a pour but de développer cette méthode en utilisant une hypothèse semi-Markovienne pour modéliser la mortalité des malades en considérant le temps passé avec la maladie. La méthode est appliquée pour estimer plusieurs indicateurs de santé pour la démence en France en 2030.

Mots-clés. Modèle multi-état, semi-Markov, Projection, Maladie Chronique . . .

Abstract. Chronic diseases are a growing public health problem and their economic, social and demographic burden is alarming in years to come. Up to now, the method used to make projections and assess the future disease burden assumes non-homogeneous Markov assumption in an illness-death model. Both age and calendar time have been taken into account in all parameters estimations, nevertheless the time spent with the disease was not considered. This work develops the method with a semi-Markov assumption to model mortality among the diseased and considering the time spent with the disease. The method is applied to estimate several health indicators for dementia in France in 2030.

Keywords. Multi-states model, Semi-Markov, Projection, Chronic disease . . .

1 Introduction

De nos jours, les maladies chroniques représentent un réel problème de Santé Publique aussi bien d'un point de vue économique que démographique [1]. De plus, la population vieillissante laisse présager une augmentation du fardeau de ces maladies pour les années à venir. La démence comme l'Alzheimer en est un bon exemple [2]. En effet, des projections

suggèrent que la prévalence de la démence augmenterait d'un facteur de 1.76 en France entre 2010 et 2030 [3,4], et de 3.7 aux États Unis entre 1997 et 2047 [5].

Les modèles multi-états basés sur des processus de Markov sont couramment utilisés pour modéliser l'incidence et la mortalité de la démence [6], et la macro simulation permet d'obtenir des projections d'indicateurs de santé tels que la prévalence [5, 7, 8]. Les probabilités de transition utilisées dans ces modèles peuvent dépendre de l'âge et du temps calendaire. Cependant, pour la plupart des maladies chroniques, la mortalité des malades est susceptible d'augmenter en fonction du temps passé avec la maladie, comme c'est le cas pour le Parkinson [9]. Bien que de récents travaux prennent en compte la durée de la maladie dans leur modélisation [10], aucun ne la considère pour fournir des projections.

Dans ce travail nous proposons une alternative au modèle développé par Joly et al. [3] en introduisant une approche semi-Markovienne pour modéliser la mortalité des sujets malades et ainsi prendre en compte le temps passé avec la maladie pour fournir des projections de divers indicateurs de santé.

2 Méthodes

2.1 Modèle Illness-death & et hypothèse semi-Markovienne

Le modèle illustré par la figure 1, est un modèle à 3 états appelés illness-death.

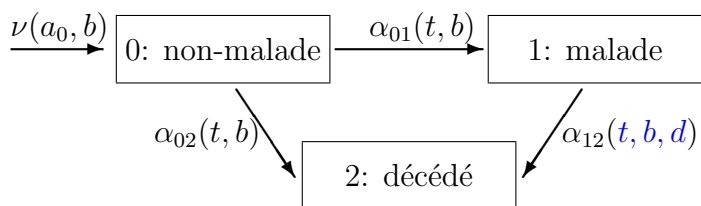


Figure 1: Modèle illness-death.

Ce modèle permet de distinguer 3 intensités de transition. Au début les sujets sont non-malades (état “0”), puis soit ils décèdent (état “2”) soit ils deviennent malades (état “1”) puis décèdent. Ainsi la transition α_{01} est interprétée comme le taux d’incidence, α_{02} et α_{12} représentent respectivement la mortalité des non-malades et la mortalité des malades. Ces intensités dépendent de t le temps calendaire et b l’année de naissance (donc $t - b$ est l’âge au temps t), mais α_{12} dépend également de d le temps passé avec la maladie. De plus, à l’entrée du modèle, nous définissons $\nu(a_0, b)$ la taille de la population à risque de développer la maladie à partir de l’âge a_0 en fonction de l’année de naissance b . Enfin, on notera $\alpha_2(t, b)$ le taux de mortalité général au temps t pour les sujets nés en b . Du

fait de la nature même des maladies chroniques, les cas de rémission ne sont pas supposés dans ce modèle.

2.2 Hypothèses sur les paramètres

La méthode repose sur plusieurs hypothèses. Tout d'abord l'incidence est supposée nulle avant un certain âge a_0 fixé (la valeur de a_0 dépend de la maladie d'intérêt).

Ensuite, pour la mortalité des non malades, souvent il est pertinent de la distinguer de la mortalité générale: $\alpha_{02}(t, b) \neq \alpha_2(t, b)$ (surtout pour les maladies avec une forte prévalence).

Enfin, pour la mortalité des malades, nous proposons un modèle additif qui permet de prendre en compte l'âge et le temps passé avec la maladie. Ainsi, nous l'écrivons comme la somme de la mortalité des non-malades et d'un sur-risque de décès α^d fonction de d :

$$\alpha_{12}(t, b, d) = \alpha_{02}(t, b) + \alpha^d(d) \quad (1)$$

2.3 Estimation des intensités de transition

Des données de cohorte sont utilisées pour estimer α_{01} et α^d en ajustant un modèle illness-death non paramétrique avec une approche par vraisemblance pénalisée et une approximation par fonctions splines [6]. Cette méthode traite le risque semi-compétitif entre devenir malade et/ou décéder ainsi que la censure par intervalle de l'âge du début de la maladie. Les projections démographiques nationales fournissent les valeurs de $\alpha_2(t, b)$ et $\nu(a_0, b)$, et nous permettront de prendre en compte l'évolution au cours du temps de chaque mortalité.

Du coup, α_{02} est estimée en résolvant une équation différentielle de type: $y'(s) = c(s) - e^{-y(s)}a(s)$, impliquant α_2 , α_{01} et α^d . La résolution (pour chaque année de naissance b) est analytique et permet d'écrire la mortalité des non-malades avec des quantités connues.

Enfin, α_{12} est estimée en utilisant l'équation (1).

2.4 Indicateurs de santé

Une fois les intensités de transition estimées on utilise les intensités cumulées (entre deux âges) A_{01} , A_{02} et A_{12} pour calculer des probabilités (par année de naissance b) telles que:

- P_{00} , la probabilité pour des sujets nés en b et vivant à l'âge a_0 d'être non-malades à l'âge $t - b$:

$$P_{00}(a_0, t - b|b) = e^{-A_{01}(a_0, t-b|b) - A_{02}(a_0, t-b|b)}$$

- P_{11} , la probabilité pour des sujets nés en b , devenant malade à l'âge a_d (avec $a_d > a_0$) d'être vivant à l'âge $t - b$:

$$P_{11}(a_d, t - b|b) = e^{-A_{12}(a_d, t-b|b, d)}$$

- P_{01} , la probabilité pour des sujets nés en b , vivant et non-malades à l'âge a_0 d'être malade et vivant à l'âge $t - b$:

$$P_{01}(a_0, t - b|b) = \int_{a_0}^{t-b} e^{-A_{01}(a_0, u|b) - A_{02}(a_0, u|b)} \alpha_{01}(u|b) e^{-A_{12}(u, t-b|b, d)} du$$

Ensuite, ces probabilités servent à calculer divers indicateurs de santé pour une année de projection t (et pas année de naissance b), afin d'avoir des informations statistiques sur le fardeau futur de la maladie. On note:

- $LE_{00}(x|b)$, *l'espérance de vie sans la maladie* à l'âge x , qui est le nombre d'année que les non-malades à l'âge x peuvent espérer vivre sans la maladie.
- $LE_{\cdot}(x|b)$, *l'espérance de vie générale* qui est la moyenne pondérée de l'espérance de vie des malades et des non-malades à l'âge x .
- $F_{01}(x|b)$ qui est le risque global de développer la maladie avant de décéder pour les sujets d'âge x vivants et non-malades.
- $Prev(a_0|t)$, *la prévalence de la maladie* qui est le nombre de cas malades entre a_0 et 99 ans au temps de projection t .

Pour exemple $LE_{00}(x|b) = \int_x^{\infty} P_{00}(x, u|b) du$

3 Application sur la démence en France et projections pour 2030.

Deux sources de données sont utilisées pour estimer les paramètres du modèle. Tout d'abord, la cohorte Française PAQUID initiée en 1988 dans le but d'étudier le vieillissement de la population permet d'estimer l'incidence de la démence (α_{01}) et le sur-risque de décès des déments (α^d). L'échantillon contient 3777 individus de plus de 65 ans et est représentatif de la population française en terme d'âge et de proportion hommes femmes [11]. Actuellement dans cette cohorte 862 cas de démence ont été diagnostiqué ainsi que 3238 décès (dont 719 cas diagnostiqués). A partir de ces données nous avons fait l'hypothèse que l'incidence est homogène au cours du temps (elle ne dépend que de l'âge et du sexe) et également nulle avant $a_0 = 65$ (car les cas de démence précoces sont rares). Ensuite, l'institut national de la statistique et des études économiques française (INSEE), fournit des projections démographiques [12], incluant la mortalité générale (α_2) par âge et sexe, et la population à risque de développer la démence (ν) à 65 ans par sexe. Enfin, la mortalité des non-déments (α_{02}) dépendant de l'âge et du temps calendaire est calculée par résolution analytique d'une équation différentielle pour chaque année b , puis

la mortalité des déments (α_{12}) est estimée en sommant cette mortalité des non-déments au sur-risque de décès des déments (α^d).

Des résultats, notamment sur la mortalité des déments en fonction de l'âge et du temps passé avec la démence seront présentés, ainsi que l'évolution des indicateurs de santé entre 2015 et 2030. La discussion portera sur le choix du modèle et les perspectives envisagées.

Bibliographie

- [1] Brayne, C. (2007), The elephant in the room - healthy brains in later life, epidemiology and public health, Nature Review Neuroscience, 8, 233-239.
- [2] World Alzheimer report. The global impact of Dementia. (2015), Alzheimer's Disease International.
- [3] Joly, P. Touraine, C. Georget, A. Dartigues, J. F. Commenges, D. and Jacqmin-Gadda, H. (2013), Prevalence projections of chronic disease and impact of public health intervention, Biometrics, 69, 109-117.
- [4] Jacqmin-Gadda, H. Alperovitch, A. Montlathuc, C. et al. (2013), 20-years prevalence projections for dementia and impact of preventive policy about risk factors, European Journal of Epidemiology, 28, 493-502.
- [5] Brookmeyer, R. Gray, S. Kawas, C. (1998), Forecasting the global burden of Alzheimer's disease, NAmerican journal of public health, 88, 1337-1342.
- [6] Joly, P. Commenges, D. Helmer, C. Letenneur, L. (2002), A penalized likelihood approach for an illness-death model with interval-censored data: application to age-specific incidence of dementia, Biostatistics, 3, 433-443.
- [7] Hebert, L. E. Scherr, P. A. Bienias, J. L. Bennett, D. A. Evans D. A. (2003), Alzheimer disease in the United States (2010-2050) estimated using the 2010 census, Neurology, 60, 1119-1122.
- [8] Wanneveich, M. Jacqmin-Gadda, H. Dartigues, F. Joly, P. (2016), Impact of intervention targeting risk factors on chronic disease burden, Statistical Methods in Medical Research, in press.
- [9] Elbaz, A. Bower, J. H. Peterson, B. J. (2003), Survival study of Parkinson disease in Olmsted country, Minnesota, Neurology, 60, 91-96.
- [10] Brinks, R. Landwehr, S. (2014), Age- and time-dependent model of the prevalence on non-communicable disease and application to dementia in Germany, Theoretical population biology, 92, 62-68.
- [11] Dartigues, J. F. Gagnon, M., Michel, P. et al. (1991), Le programme de recherche Paquid sur l'épidémiologie de la démence. Méthodologie et résultats initiaux, Revue Neurologique, 147, 225-230.
- [12] Blanpain, N. Chardon, O. (2010), Projections de populations 2007-2060 pour la France métropolitaine: méthode et principaux résultats, F1008.