

14-07

# COMPAS: UN MODÈLE DE MICROSIMULATION SANTÉ POUR LE QUÉBEC

CAHIER DE RECHERCHE  
WORKING PAPER

David Boisclair, Aurélie Côté-Sergent, Jean-Yves Duclos,  
Alexandre Lekina, Steeve Marchand et Pierre-Carl Michaud

Décembre / December 2014



Faculté des sciences sociales

ESG UQÀM



La Chaire de recherche Industrielle Alliance sur les enjeux économiques des changements démographiques est une chaire bi-institutionnelle qui s'appuie sur un partenariat avec les organisations suivantes :

- **Centre interuniversitaire de recherche en analyse des organisations (CIRANO)**
- **Industrielle Alliance, Assurance et services financiers inc.**
- **Régie des rentes du Québec**

Les opinions et analyses contenues dans les cahiers de recherche de la Chaire ne peuvent en aucun cas être attribuées aux partenaires ni à la Chaire elle-même et elles n'engagent que leurs auteurs.

Opinions and analyses contained in the Chair's working papers cannot be attributed to the Chair or its partners and are the sole responsibility of the authors.

© 2014 David Boisclair, Aurélie Côté-Sergent, Jean-Yves Duclos, Alexandre Lekina, Steeve Marchand et Pierre-Carl Michaud.  
Tous droits réservés. All rights reserved. Reproduction partielle permise avec citation du document source, incluant la notice ©.  
Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.

Dépôt légal : 4<sup>e</sup> trimestre 2014  
ISSN 2368-7207



# COMPAS: Un modèle de microsimulation santé pour le Québec

David Boisclair\*  
Aurélie Côté-Sergent†  
Jean-Yves Duclos‡  
Alexandre Lekina§  
Steeve Marchand¶  
Pierre-Carl Michaud||

Décembre 2014

## Résumé

Nous présentons dans ce document la conception et la structure de COMPAS, un modèle de microsimulation dynamique portant sur l'état et l'utilisation de soins de santé des Québécois. Ce modèle comporte deux dimensions importantes : un composante dynamique, qui permet une évolution de l'état de santé des individus simulés tout au long de leur cycle de vie, et une composante transversale, qui associe l'utilisation de ressources médicales en lien avec l'état de santé des individus simulés. Le modèle utilise la population de 2010 comme point de départ, et plusieurs sources de données pour modéliser les dimensions de l'état de santé et la mortalité ainsi que les diverses transitions entre les états. Il permet en outre d'analyser une seule cohorte à la fois, ou encore l'ensemble de la population telle qu'elle évolue au fil du temps avec ses nouveaux entrants. Nous présentons quelques résultats de simulations préliminaires réalisées avec COMPAS, notamment en matière d'état de santé et d'espérance de vie.

*Mots clés* : microsimulation, santé, espérance de vie, Québec.

*Codes JEL* : C63, I14, J11.

---

\*ESG UQAM

†ESG UQAM

‡Université Laval, CIRPÉE, CIRANO

§ESG UQAM

¶Université Laval

||ESG UQAM, CIRPÉE, CIRANO, RAND. Correspondance: Département des sciences économiques, Université du Québec à Montréal, C.P. 8888, Succ. Centre-ville, Montréal (QC) H3C 3P8, Canada.

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Structure du modèle</b>	<b>5</b>
1.1 Dynamique de COMPAS . . . . .	5
1.2 Structure du code FORTRAN . . . . .	7
1.2.1 Paramètres . . . . .	8
1.2.2 Agents . . . . .	9
1.2.3 Initialisation . . . . .	9
1.2.4 Transitions . . . . .	9
1.2.5 Dynamique populationnelle . . . . .	9
1.2.6 Nouvelles cohortes . . . . .	10
1.2.7 Immigration . . . . .	10
1.2.8 Tableau de bord . . . . .	10
1.2.9 Exécution du programme . . . . .	10
1.3 Organisation des fichiers et <i>subversion</i> . . . . .	11
<b>2 Données</b>	<b>11</b>
2.1 Survol . . . . .	11
2.2 ENSP . . . . .	11
2.2.1 Pondération . . . . .	12
2.2.2 Attrition et sélection . . . . .	13
2.2.3 Mortalité . . . . .	14
2.3 Construction des variables . . . . .	15
2.3.1 Problèmes de santé auto-déclarés . . . . .	16
2.3.2 Invalidité et institutionnalisation . . . . .	18
2.3.3 Facteurs de risque . . . . .	21
2.3.4 Utilisation des soins de santé . . . . .	22
2.3.5 Autres variables socio-économiques . . . . .	25
2.4 ESCC . . . . .	25
<b>3 Initialisation</b>	<b>29</b>
3.1 Survol . . . . .	29
3.2 Imputation . . . . .	30
3.2.1 Imputation de l'âge . . . . .	30
3.2.2 Imputation du niveau d'éducation . . . . .	30
3.2.3 Imputation de l'invalidité . . . . .	33

3.2.4	Imputation de l'Alzheimer et des autres démences . . . . .	33
3.2.5	Caractéristiques de la population initiale . . . . .	34
<b>4</b>	<b>Transitions</b>	<b>35</b>
4.1	Survol . . . . .	35
4.2	Modèles économétriques . . . . .	36
4.2.1	Modèles pour les maladies . . . . .	36
4.2.2	Modèle pour la mortalité . . . . .	39
4.2.3	Modèles pour le tabagisme . . . . .	41
4.2.4	Modèle pour l'invalidité . . . . .	41
4.2.5	Modèle pour l'obésité . . . . .	43
4.3	Conclusions et améliorations futures au modèle . . . . .	44
<b>5</b>	<b>Renouvellement</b>	<b>45</b>
5.1	Survol . . . . .	45
5.2	Modélisation . . . . .	45
5.3	Mise en application . . . . .	48
5.3.1	Tendances historiques . . . . .	48
5.3.2	Projections . . . . .	49
<b>6</b>	<b>Démographie</b>	<b>50</b>
6.1	Mortalité . . . . .	50
6.1.1	Définitions et statistiques . . . . .	50
6.1.2	Méthodes d'estimation . . . . .	52
6.1.3	Intégration dans COMPAS . . . . .	53
6.2	Immigration . . . . .	55
6.3	Conclusion . . . . .	57
<b>7</b>	<b>Utilisation des soins de santé</b>	<b>57</b>
7.1	Survol . . . . .	57
7.2	Modèles économétriques . . . . .	57
7.2.1	Régression de type binomiale négative . . . . .	57
7.2.2	Régression de type logistique . . . . .	58
7.3	Résultats . . . . .	59
7.3.1	Régression de type binomiale négative . . . . .	59
7.3.2	Régression de type logistique . . . . .	60
<b>8</b>	<b>Résultats de base en mode cohorte</b>	<b>63</b>

8.1	Scénario de référence . . . . .	63
8.2	Simulation et résultats . . . . .	64
<b>9</b>	<b>Résultats de base en mode population</b>	<b>77</b>
9.1	Scénario de référence . . . . .	77
9.1.1	Un retour dans COMPAS . . . . .	77
9.1.2	Hypothèses du scénario de référence . . . . .	78
9.2	Simulation et résultats . . . . .	80
	<b>Conclusion</b>	<b>89</b>
	<b>Bibliographie</b>	<b>93</b>

## Liste des figures

1.1	Dynamique du modèle COMPAS avec un cycle de simulation de 2 ans . . . . .	6
1.2	Structure du code FORTRAN . . . . .	7
2.1	Comparaison des taux de mortalité : ENSP vs. Human Mortality Database (HMD). Intervalle de confiance de 95% en pointillés. . . . .	15
3.1	Probabilité d’avoir un diplôme universitaire pour les individus ayant complété des études post-secondaires — EPA 2010 à grande diffusion . . . . .	32
8.1	Évolution de la taille de la cohorte née en 1979-1980 — Scénario de référence . . . . .	65
8.2	Prévalences simulées des problèmes de santé dans la cohorte née en 1979-1980 — Scénario de référence . . . . .	66
8.3	Prévalence simulée de l’obésité dans la cohorte née en 1979-1980 — Scénario de référence	67
8.4	Évolution de l’invalidité dans la cohorte née en 1979-1980 — Scénario de référence . .	68
8.5	Utilisation des soins de santé par la cohorte née en 1979-1980 — Scénario de référence	70
8.6	Proportion d’utilisateurs de soins de santé dans la cohorte née en 1979-1980 — Scénario de référence . . . . .	71
9.1	Population du Québec — Scénario de référence . . . . .	81
9.2	Parts des « plus de 65 ans » et des « plus de 75 ans » dans la population québécoise de 30 à 110 ans — Scénario de référence . . . . .	81
9.3	Taux de mortalité et espérance de vie périodique au Québec — Scénario de référence .	83
9.4	Espérance de vie résiduelle au Québec — Scénario de référence . . . . .	84
9.5	Différence entre les espérances de vie au Québec — Scénario de référence . . . . .	85
9.6	Prévalence des maladies au Québec — Scénario de référence . . . . .	86
9.7	Évolution de l’invalidité chez les 30 à 110 ans — Scénario de référence . . . . .	87
9.8	Évolution de l’obésité chez les 30 à 110 ans — Scénario de référence . . . . .	88
9.9	Utilisation des soins de santé au Québec par les individus de 30 à 110 ans hors-institution — Scénario de référence . . . . .	88

## Liste des tableaux

2.1	Nombre d'observations pour le Québec et le Canada dans l'ENSP (1994-2011) . . . . .	12
2.2	Taux de réponse et attrition par cycle dans l'ENSP . . . . .	13
2.3	Tests statistiques de sélection, ENSP (1994-2011) : Différence de proportions entre les individus répondant à tous les cycles et ceux pour qui il manque des cycles. La statistique t est représentée. Une valeur négative indique une prévalence plus élevée chez les individus qui vont quitter. . . . .	14
2.4	Taux de mortalité annuels par cycle de l'ENSP, population de 30 ans et plus . . . . .	14
2.5	Dimensions de l'état de santé prises en compte dans COMPAS . . . . .	16
2.6	Dimensions de l'utilisation des soins de santé prises en compte dans COMPAS . . . . .	16
2.7	Prévalence auto-déclarée des maladies dans l'ENSP (1994-2011) . . . . .	17
2.8	Incidence auto-déclarée des maladies dans l'ENSP (1994-2011) . . . . .	18
2.9	État d'invalidité selon l'ENSP (1994-2011) . . . . .	19
2.10	Transitions entre les états d'invalidité selon l'ENSP (1994-2011) . . . . .	20
2.11	Tabagisme dans l'ENSP (1994-2011) . . . . .	21
2.12	Taux d'initiation au tabagisme et de cessation selon l'ENSP (1994-2011) . . . . .	22
2.13	Indice de masse corporelle (IMC) selon l'ENSP (1994-2011) . . . . .	23
2.14	Transitions entre les états d'obésité selon l'ENSP (1994-2011) . . . . .	23
2.15	Utilisation des soins de santé par âge selon l'ENSP (1994-2011) . . . . .	24
2.16	Utilisation des soins de santé selon la présence de maladies, ENSP (1994-2011) : Toutes ces différences sont statistiquement significatives au seuil de 5%. . . . .	26
2.17	Caractéristiques socio-économiques tirées de l'ENSP (1994-2011) . . . . .	27
2.18	Prévalence auto-déclarée des maladies dans l'ESCC 2010 . . . . .	27
2.19	Tabagisme dans l'ESCC 2010 . . . . .	28
2.20	Indice de masse corporelle (IMC) selon l'ESCC 2010 . . . . .	28
2.21	Caractéristiques socio-économiques tirées de l'ESCC 2010 . . . . .	29
3.1	Effets marginaux moyens des variables sur la probabilité d'avoir un diplôme universitaire pour les individus ayant complété des études post-secondaires — EPA 2010 à grande diffusion . . . . .	31
3.2	Effets marginaux moyens des variables sur la probabilité d'être en institution — ENSP 2010-2011 . . . . .	33
3.3	Effets marginaux moyens des variables sur la probabilité d'être atteint d'Alzheimer ou d'une autre démence — ENSP 2010 . . . . .	34
3.4	Description des variables de la population initiale . . . . .	35
4.1	Effets permis des maladies les unes sur les autres . . . . .	36
4.2	Effets marginaux moyens des variables sur les probabilités d'incidence des maladies . . . . .	38



4.3	Effets marginaux moyens sur les probabilités de décès . . . . .	40
4.4	Effets marginaux moyens sur les probabilités de transition entre les états de tabagisme . . . . .	41
4.5	Effets marginaux moyens sur les probabilités de transition vers les états d'invalidité . . . . .	43
4.6	Effets marginaux moyens sur les probabilités de transition vers les états d'obésité . . . . .	44
5.1	Types de variables associés aux caractéristiques des individus entrants . . . . .	46
5.2	Taux de croissance annuel moyen (TCAM) des proportions de certaines caractéristiques individuelles, population canadienne âgée de 25 à 34 ans, 2000 à 2012 . . . . .	49
5.3	Codification de certaines des variables caractérisant un individu . . . . .	51
5.4	Répartition de certaines caractéristiques individuelles dans les cohortes entrantes (en %), aux cycles de simulation $t \in \{0, 1, 5, 10, 15, 20\}$ . . . . .	51
6.1	Taux annuels de réduction des taux de mortalité chez les hommes (source : Régie des rentes du Québec (2013, tab. 14)) . . . . .	56
6.2	Taux annuels de réduction des taux de mortalité chez les femmes (source : Régie des rentes du Québec (2013, tab. 14)) . . . . .	56
7.1	Régression de type binomiale négative : effets marginaux moyens des différentes variables sur l'utilisation de soins de santé. L'utilisation est calculée sur une base annuelle. . . . .	61
7.2	Régression de type logistique : effets marginaux moyens des différentes variables sur l'utilisation de soins de santé. L'utilisation est calculée sur une base annuelle. . . . .	62
8.1	Comparaison des prévalences à différents âges (%) : COMPAS vs. ENSP . . . . .	66
8.2	Espérance de vie à 30 ans en fonction du niveau de scolarité et du sexe . . . . .	69
8.3	Espérance de vie à 30 ans en fonction du niveau de scolarité et du statut de tabagisme . . . . .	69
8.4	Espérance de vie à 30 ans en fonction du niveau de scolarité et du statut d'obésité . . . . .	69
8.5	Nombre moyen viager de nuits à l'hôpital en fonction du niveau de scolarité et du sexe . . . . .	72
8.6	Nombre moyen viager de nuits à l'hôpital en fonction du niveau de scolarité et du statut de tabagisme . . . . .	72
8.7	Nombre moyen viager de nuits à l'hôpital en fonction du niveau de scolarité et du statut d'obésité . . . . .	72
8.8	Nombre moyen viager de consultations chez les généralistes en fonction du niveau de scolarité et du sexe . . . . .	73
8.9	Nombre moyen viager de consultations chez les généralistes en fonction du niveau de scolarité et du statut de tabagisme . . . . .	73
8.10	Nombre moyen viager de consultations chez les généralistes en fonction du niveau de scolarité et du statut d'obésité . . . . .	73
8.11	Nombre moyen viager de consultations chez les spécialistes en fonction du niveau de scolarité et du sexe . . . . .	74
8.12	Nombre moyen viager de consultations chez les spécialistes en fonction du niveau de scolarité et du statut de tabagisme . . . . .	74

8.13	Nombre moyen viager de consultations chez les spécialistes en fonction du niveau de scolarité et du statut d'obésité . . . . .	74
8.14	Nombre d'années d'utilisation de services à domicile en fonction du niveau de scolarité et du sexe . . . . .	75
8.15	Nombre moyen d'années d'utilisation de services à domicile en fonction du niveau de scolarité et du statut de tabagisme . . . . .	75
8.16	Nombre moyen d'années d'utilisation de services à domicile en fonction du niveau de scolarité et du statut d'obésité . . . . .	75
8.17	Nombre moyen d'années d'utilisation de médicaments en fonction du niveau de scolarité et du sexe . . . . .	76
8.18	Nombre moyen d'années d'utilisation de médicaments en fonction du niveau de scolarité et du tabagisme . . . . .	76
8.19	Nombre moyen d'années d'utilisation de médicaments en fonction du niveau de scolarité et du statut d'obésité . . . . .	76

# Introduction

## Problématique

La part du revenu national consacrée aux soins de santé ne cesse d'augmenter dans le temps. Au Canada, celle-ci se chiffrait respectivement à 7,1 % et 11,7 % du PIB en 1981 et 2011, contre 8,6 % et 12,6 % au Québec ([Institut canadien d'information sur la santé, 2012](#)). Depuis quelques décennies, le même constat s'applique — mais de façon plus marquée — à la place des dépenses de santé dans le budget de l'État. Cela est particulièrement vrai au Canada et au Québec, où les trois quarts environ de ces dépenses sont effectuées publiquement. Au Québec, la part des revenus gouvernementaux consacrée à la santé est ainsi passée de près de 32 % en 2000 à plus de 45 % en 2014.

Dans la plupart des pays industrialisés, la hausse des dépenses s'accompagne toutefois d'une amélioration de l'espérance de vie. Ainsi, un Canadien né en 1960 avait une espérance de vie à la naissance de 71 ans, contre 75,3 ans s'il naissait en 1980 et 80,8 ans s'il naissait en 2008 ([Banque Mondiale, 2014](#)). Une bonne partie des gains a été réalisée après 50 ans : chez les hommes, entre 1991 et 2008, l'espérance de vie à 50 ans a bondi de 27,6 à 31 ans, tandis que chez les femmes elle connaissait une hausse appréciable, passant de 32,8 à 34,7 ans ([Bohnert, 2013](#)). Les âges plus avancés n'étaient pas en reste, les hommes de 75 ans ayant vu leur espérance de vie passer de 9,8 à 11,4 ans et les femmes, de 12,6 à 13,6 ans ([Bohnert, 2013](#)). Au Québec, la situation est similaire : l'espérance de vie à la naissance est passée de 78 ans en 1997 à plus de 81 ans en 2012, tandis que celle à 65 ans a augmenté d'environ 2,5 années durant la même période ([Régie des rentes du Québec, 2013](#); [Statistique Canada, 2014a,b](#)).

Des tendances importantes sont à l'oeuvre. D'abord, les tendances démographiques lourdes comme le vieillissement de la population et l'allongement de l'espérance de vie font pression sur les finances publiques et sur le système de santé dans son ensemble. Ensuite, le progrès technologique en médecine et la capacité de maintenir les personnes âgées à domicile sont aussi des facteurs clés de l'évolution future des ressources consacrées à la santé.

Par ailleurs, on observe une progression des maladies chroniques associées à l'obésité au Québec, une progression qui accroît la pression sur le système de santé et sur le budget de l'État. Ainsi, la prévalence de l'hypertension a connu une croissance annuelle moyenne de 5,1% entre 2000 et 2007 ([Institut national de santé publique du Québec, 2011](#)). Au sein de la population âgée de 65 ans et plus, la prévalence de maladies cardiaques, elle, a augmenté à un taux annuel moyen de 11,0% entre 1994 et 2003 ([Institut national de santé publique du Québec, 2006](#)). Il est impératif de pouvoir modéliser et analyser adéquatement ces tendances pour arriver à des projections fiables de la situation future.

Certaines études, comme celle de [Clavet \*et al.\* \(2013\)](#), utilisent le modèle de microsimulation développé par l'équipe de SIMUL (voir ci-après) et projettent que les dépenses de santé représenteront au moins 56,5 % des revenus en 2030 et pourraient même dépasser 68% selon un scénario plausible. À l'aide d'une méthodologie faisant appel à des données agrégées, [Godbout \*et al.\* \(2014\)](#) obtiennent des résultats d'un même ordre de grandeur (à 58,6 %), tout comme plusieurs travaux antérieurs du même type remontant au début des années 2000. Or, l'évaluation de politiques visant à atténuer ces pressions, tout en continuant à améliorer l'état de santé de la population, nécessite des modèles capables de tenir compte de façon détaillée de l'ensemble des facteurs précités.

## Microsimulation de long terme

Les tendances observées étant multidimensionnelles, la modélisation agrégée au niveau populationnel ne permet pas de traiter adéquatement de ces questions en tenant compte simultanément d'un grand nombre de dimensions. Elle ne peut notamment pas tenir compte de l'hétérogénéité grandissante des trajectoires de vie des personnes âgées. C'est un des éléments qui furent soulevés au terme des travaux d'un panel visant à réunir les chercheurs et organismes s'affairant à étudier l'évolution future des dépenses de santé aux États-Unis ([National Research Council, 2012](#)). Le *Future Elderly Model* (FEM), développé par une équipe de chercheurs de la RAND Corporation, est l'un des seuls modèles de long terme aux États-Unis et dans le monde à incorporer les interactions complexes entre les diverses tendances énoncées ici-haut. Initialement construit à l'aide d'une base de données administratives portant sur 100 000 bénéficiaires du programme Medicare, donc des Américains âgés de 65 ans et plus, le modèle couvre désormais toute la population de 50 ans et plus des États-Unis et de sept pays européens. Il permet de projeter l'état de santé, l'utilisation de ressources médicales ainsi que les revenus de travail, l'offre de travail et les décisions de retraite des individus, ainsi que l'impact de changements futurs dans l'état de santé, la longévité et les technologies médicales. Le modèle a été utilisé dans plusieurs études analysant des questions différentes (récemment, voir par exemple [Goldman et al., 2013](#); [Michaud et al., 2012](#)).

Ce type de modèle part toujours du principe initial qu'il n'y aura pas à l'avenir de changement dans les politiques gouvernementales.<sup>1</sup> Ces modèles sont néanmoins très utiles, à la fois pour considérer des scénarios de moyen et de long terme et pour évaluer les bénéfices potentiels de politiques visant à améliorer la santé de la population à l'aide de scénarios « alternatifs » construits par les utilisateurs. Ces bénéfices dépassent largement la réduction des dépenses de santé puisqu'ils couvrent aussi les bénéfices économiques découlant d'une amélioration de la santé et des effets sur les revenus et dépenses des différents programmes publics. C'est dans cette optique que nous avons entrepris de développer un modèle québécois de microsimulation axé sur la santé.

## Microsimulation reliée au vieillissement et à la santé au Canada

Un certain nombre de modèles de microsimulation ont été développés et appliqués au vieillissement ou à la santé au Canada au fil des ans. Certains sont de nature générale, comme Lifepaths de Statistique Canada, tandis que d'autres étaient dès l'origine orientés vers la santé, comme POHEM (également développé chez Statistique Canada). À notre connaissance, toutefois, aucun des modèles existants n'a été développé spécifiquement pour le Québec, et aucun ne modélise des transitions simultanées entre différents états de santé définis par la présence de plusieurs maladies et facteurs de risques. Ces modèles existants sont décrits brièvement ci-après, sur la base de l'inventaire réalisé par [Décarie et al. \(2012\)](#) de même que du résumé inclus dans [Clavet et al. \(2012\)](#).

Le plus ancien modèle de microsimulation recensé au Canada est DYNACAN, mis sur pied dans les années 1970 au sein de ce qui est aujourd'hui Emploi et Développement social Canada. Ce modèle dynamique, basé sur le Recensement de 1971 et sur le modèle américain CORSIM, est axé sur les

---

1. L'idée qui sous-tend les scénarios de base (ou dits « de référence ») n'est pas que les politiques ne changeront effectivement pas, mais plutôt que les projections de long terme doivent être effectuées en fonction de ce qui est connu des politiques gouvernementales au moment de la réalisation des simulations.

programmes publics de pension (Sécurité de la vieillesse, Régime de pensions du Canada) et sur l'effet des politiques sur ceux-ci. DYNACAN a été délaissé à la fin des années 2000, mais a tout de même servi d'inspiration partielle au modèle socio-économique développé par l'équipe de SIMUL (voir ci-après), aussi basé sur un échantillon représentatif de ménages.

Le modèle actuellement en usage au gouvernement du Canada — et utilisé par un certain nombre de chercheurs — est le modèle Lifepaths, développé chez Statistique Canada depuis 25 ans. Contrairement aux autres mentionnés ici, ce modèle est basé sur le concept de personne dominante dans un ménage, et simule l'ensemble de la situation de ces personnes. Le modèle inclut des dimensions comme la retraite et l'entrée en institution, mais très peu l'état de santé et pas du tout l'utilisation des soins (seule une mesure basée sur l'état de santé général auto-déclaré est incluse ainsi qu'une mesure d'invalidité). Un grand nombre d'articles et d'analyses ont été produits à l'aide de Lifepaths, mais ce dernier intègre peu ou pas de spécificités provinciales (Clavet *et al.*, 2012).

À notre connaissance, le seul véritable modèle de microsimulation de la santé existant au Canada se nomme POpulation HEalth Model (POHEM). Depuis les années 1990, Statistique Canada a construit plusieurs modules de ce modèle afin de modéliser différentes maladies et facteurs de risque. À ce jour, toutefois, les maladies et facteurs de risque sont modélisés isolément et permettent uniquement une analyse « par maladie ». POHEM fait appel à un grand nombre de sources de données, dont plusieurs grandes enquêtes, le Recensement, les registres de maladies et d'hospitalisations, etc.

La combinaison dans POHEM d'une variété de types et de sources de données est aussi l'esprit dans lequel notre équipe a souhaité développer un modèle dynamique plus général de microsimulation relié à la santé pour le Québec. Celui-ci cherche à répondre à plusieurs préoccupations par rapport aux modèles existants, et vise, en quelque sorte, à améliorer encore davantage l'excellente infrastructure d'analyse et de recherche développée à l'aide de ces modèles.

## COMPAS

Notre équipe a déjà élaboré, avec la collaboration du ministère des Finances du Québec, un modèle de microsimulation permettant d'analyser les conséquences des changements démographiques et des politiques économiques pour les finances publiques et le niveau de vie des Québécois (Clavet *et al.*, 2012). Ce modèle ne considère toutefois pas les soins de santé, qui comptent pour près de 50 % du budget de l'État québécois.

Nous avons donc entrepris d'élaborer un modèle de microsimulation dynamique en santé en nous basant sur les sources existantes de données administratives et d'enquête. Puisque l'équipe de SIMUL a déjà développé l'infrastructure de microsimulation démographique et économique, nous avons créé un modèle complémentaire relié à la santé qui permettra d'étendre les champs d'application de SIMUL. Ce modèle se nomme COMPAS. COMPAS a été créé en 2013 et son développement se poursuit sur une base continue.

Le modèle santé comporte deux composantes importantes : a) une composante dynamique qui permet de faire progresser les individus simulés dans leur cycle de vie, permettant ainsi une évolution de leur état de santé, capté par un grand nombre de marqueurs comme la présence de maladies, la présence de limitations d'activités, etc. ; et b) une composante transversale visant à quantifier les coûts d'utilisation

des ressources associées à l'état de santé des personnes vivantes pour une année donnée. Cette deuxième composante pourra profiter de la riche infrastructure de données existante, notamment les données provenant de la RAMQ ainsi que de MED-ECHO.

Chaque année, la population en vie est caractérisée par un ensemble de variables démographiques, économiques et d'état de santé. Les variables d'état de santé permettent d'imputer une utilisation de soins de santé et les dépenses afférentes pour l'année en cours. Cette population passe ensuite par une phase de transition où l'état de santé peut changer. Par exemple, un individu âgé de 80 ans faisant de l'hypertension aura un ensemble de probabilités de subir un incident cardio-vasculaire. Ces probabilités dépendent de l'état de santé courant. Durant cette transition, chaque individu court le risque de décéder ou de développer des incapacités qui feront en sorte qu'il devra ensuite, par exemple, réduire sa participation au marché du travail ou entrer dans un établissement de soins de longue durée.

C'est cette richesse dans les états et les transitions qui permet de faire des prévisions de long terme à partir de microdonnées et d'analyser différents scénarios pouvant affecter une fraction ou l'entièreté de la population. Par exemple, que se passerait-il si on parvenait à retarder l'entrée dans un établissement de soins de longue durée? Ou à réduire le risque de mortalité pour les individus ayant souffert d'un AVC? Ou encore à réduire la prévalence du diabète? Puisque le modèle de santé sera relié au modèle socio-économique déjà développé par SIMUL, il sera possible d'analyser l'impact de ces scénarios sur les revenus et les autres dépenses du gouvernement.

À ce jour, toutefois, COMPAS produit déjà des résultats de base qui sont présentés à la fin de ce rapport. Par exemple, COMPAS prédit que la prévalence de l'hypertension en 2050 sera de l'ordre de 21% pour les individus âgés de 70 à 74 ans. Le modèle capte également le vieillissement de la population québécoise. Entre 2010 et 2050, la part d'individus âgés de plus de 75 ans (dans la population de 30 ans et plus) augmentera de 10% à un peu moins de 30%. On note aussi une augmentation de l'espérance de vie entre 2010 et 2050. Les individus âgés de 30 ans en 2050 vivront, en moyenne, 3,3 années de plus que les individus âgés de 30 ans en 2010. En ce qui concerne les individus âgés de 60 ans, cette différence sera de 2,8 années.

## **Objectifs et progrès**

Les objectifs de recherche associés au développement du modèle sont les suivants :

1. Développer une base de données longitudinales permettant d'estimer les transitions entre états de santé ainsi que la mortalité au niveau populationnel. Cette base est construite à partir de l'échantillon québécois de l'Enquête nationale sur la santé de la population (ENSP).
2. Estimer des relations dynamiques entre états de santé en utilisant la base de données créée au point 1 et une modélisation économétrique appropriée. Ces relations sont multidimensionnelles et permettent de préserver l'hétérogénéité des trajectoires de santé dans la population.
3. Développer une base de données de coûts et d'utilisation des ressources selon les états de santé identifiés au point 1. Cette base de données devra intégrer la rémunération des actes, les dépenses en médicaments et les dépenses budgétaires allouées pour chaque DRG pour les épisodes de soins en établissements hospitaliers. Elle devra permettre de créer une mesure de dépenses annuelles totales pour chaque individu faisant partie de la population cible.

4. Estimer des relations de coûts et d'utilisation des ressources à l'aide de la base de données créée au point 3 et de modèles économétriques appropriés.
5. Construire le modèle de simulation en se basant sur l'infrastructure démographique et économique du modèle SIMUL et en intégrant les relations estimées aux points 2 et 4.
6. Calibrer et valider le modèle de simulation sur une base prospective en se basant sur les projections de l'ISQ et de la RRQ quant à la mortalité et sur celles des ministères pour la prévision des dépenses de santé au Québec.
7. Évaluer la sensibilité des projections futures aux politiques économiques et sociales, aux tendances en santé, à l'amélioration de la longévité et aux conditions économiques.

Au moment de la rédaction de ce rapport, nous avons réalisé les objectifs 1 et 2 ainsi que la partie des objectifs 3 et 4 portant sur l'utilisation des soins de santé. De plus, nous avons complété la construction initiale du modèle (objectif 5 réalisé de façon préliminaire), et nous sommes à même de présenter des résultats préliminaires. Par ailleurs, nous avons complété une grande part du travail relatif à l'objectif 6. Pour l'objectif 3, la partie qui concerne les données administratives requiert un travail important afin d'obtenir l'accès aux données. À ce jour, nous avons accès aux données de la RAMQ concernant les consultations auprès de d'omnipraticiens et de médecins spécialistes ; il nous faut faire une demande auprès de la RAMQ afin d'avoir également accès aux données sur les dépenses pharmaceutiques. Ce travail sera effectué à l'automne 2014. Du côté des hospitalisations de courte durée, nous avons accès aux données de MED-ECHO, ce qui nous permettra d'estimer le coût des hospitalisations. Il reste du travail à faire concernant le coût de l'institutionnalisation et celui des services à domicile ; la recherche de bases de données permettant un tel travail a débuté à l'été 2014.

## 1 Structure du modèle

### 1.1 Dynamique de COMPAS

Le modèle de microsimulation comprend plusieurs composantes. La structure du modèle est presque entièrement calquée sur la dynamique de la population. Le modèle suit des agents, caractérisés par des attributs socio-économiques et de santé, qui transitent d'une période à l'autre. À chaque période, de nouveaux agents entrent dans la simulation parce qu'ils ont l'âge de départ, que nous fixons à 30 ans, ou parce qu'ils immigreront.<sup>2</sup> Leur vie dans le modèle se termine au moment de leur décès ou lorsqu'ils atteignent l'âge maximal permis dans le modèle, que nous fixons à 110 ans. À chaque cycle de simulation<sup>3</sup>, nous pouvons observer non seulement leur état de santé, mais aussi l'utilisation qu'ils font des soins de santé. Un groupe d'agents nés la même année — ou les deux mêmes années, lorsqu'un cycle de simulation compte deux années — est nommé cohorte. Le modèle permet de faire des simulations ou bien pour une cohorte seulement ou bien pour une population, composée de plusieurs cohortes.

Une population est constituée du nombre d'agents en vie à la fin d'une année donnée. À chaque cycle de simulation, la population évolue. D'abord, elle perd des agents en raison de la mortalité. Ensuite,

---

2. L'âge d'entrée dans le modèle est fixé à 30 ans afin d'éviter d'avoir à modéliser les parcours et les choix d'éducation des individus, et parce que les maladies et invalidités considérées dans le modèle sont peu présentes avant cet âge.

3. Dans COMPAS, un cycle de simulation correspond à une période deux ans. Ce choix est non restrictif et il peut être modifié selon les objectifs de l'analyse.

elle gagne des agents parce qu'il y a renouvellement, c'est-à-dire qu'une nouvelle cohorte d'agents âgés de 30 ans (ou 30 – 31 ans lorsque les cycles de simulation sont de 2 ans) fait son entrée dans le modèle. Finalement, elle gagne et elle perd des membres à cause de la migration. Chaque agent simulé a un poids démographique donné. Ainsi, les statistiques d'intérêt peuvent être calculées au niveau populationnel.

Le modèle débute en 2010, qui est la dernière année de données fiables pour un grand nombre des sources d'information que nous utilisons. Pour cette année-là, nous créons une population initiale sur la base d'une enquête existante. Cette phase d'initialisation donne un échantillon statistiquement représentatif du Québec de 2010. Ensuite débute le processus démographique d'évolution de la population, jusqu'à ce que l'on atteigne l'année cible : 2050. En fonction des résultats voulus, le modèle peut continuer à estimer ses variables jusqu'à l'année 2130 s'il reste des agents en vie. Ceci pourrait être utile, par exemple, pour analyser l'évolution de l'espérance de vie de chacune des cohortes entrant dans le modèle. En effet, les cohortes qui entrent dans le modèle entre 2010 et 2050 compteront une grande proportion d'agents encore en vie en 2050 (les plus vieux n'auront alors que 70 ans). La figure 1.1 résume le fonctionnement du modèle en mode population et la dynamique de COMPAS.

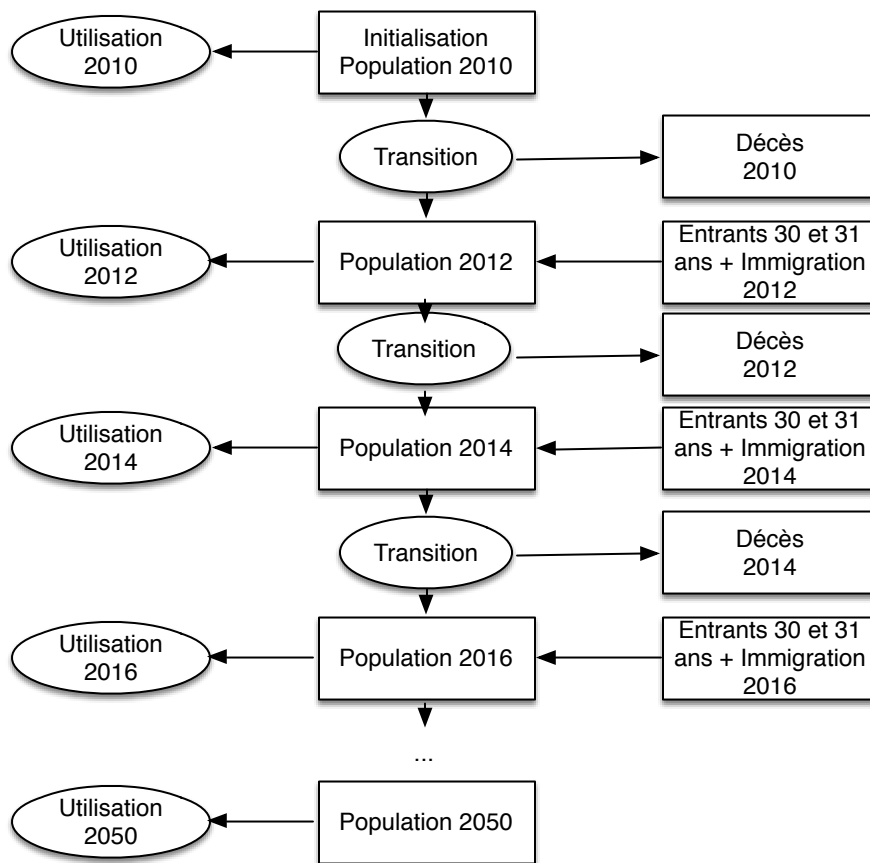


FIGURE 1.1 – Dynamique du modèle COMPAS avec un cycle de simulation de 2 ans



Dans le cas du mode cohorte, la composante de renouvellement (**newcohort.f95** ; voir ci-après) est appelée une seule fois, tandis que les composantes d’immigration (**immigration.f95**) et d’initialisation (**initialize.f95**) sont inactives. Ce mode ne génère qu’une cohorte d’individus âgés de 30 ans, ou de 30 – 31 ans avec des cycles de simulation de 2 ans (voir chapitre 3).

Puisque les simulations comportent un élément stochastique, le modèle permet de faire plusieurs répliques afin de s’assurer que le résultat obtenu n’est pas le fruit d’un pur hasard. Par défaut, 100 répliques sont utilisées dans nos travaux.

## 1.2 Structure du code FORTRAN

Cette section documente la structure du modèle et fournit un guide pour l’utilisateur avancé. Elle n’est donc pas nécessaire à la compréhension des chapitres suivants, et sa lecture peut être omise sans préjudice au lecteur ne souhaitant pas s’informer des caractéristiques informatiques du modèle.

COMPAS est principalement construit en FORTRAN, un langage simple qui permet une grande flexibilité et une grande adaptabilité aux différents besoins inhérents à la modélisation. La conception du modèle est modulaire, ce qui permet d’en modifier les paramètres et les composantes de façon séparée et sans nécessairement — ou immédiatement — affecter l’ensemble du modèle. Cette structure modulaire du code FORTRAN est reproduite dans la figure 1.2.

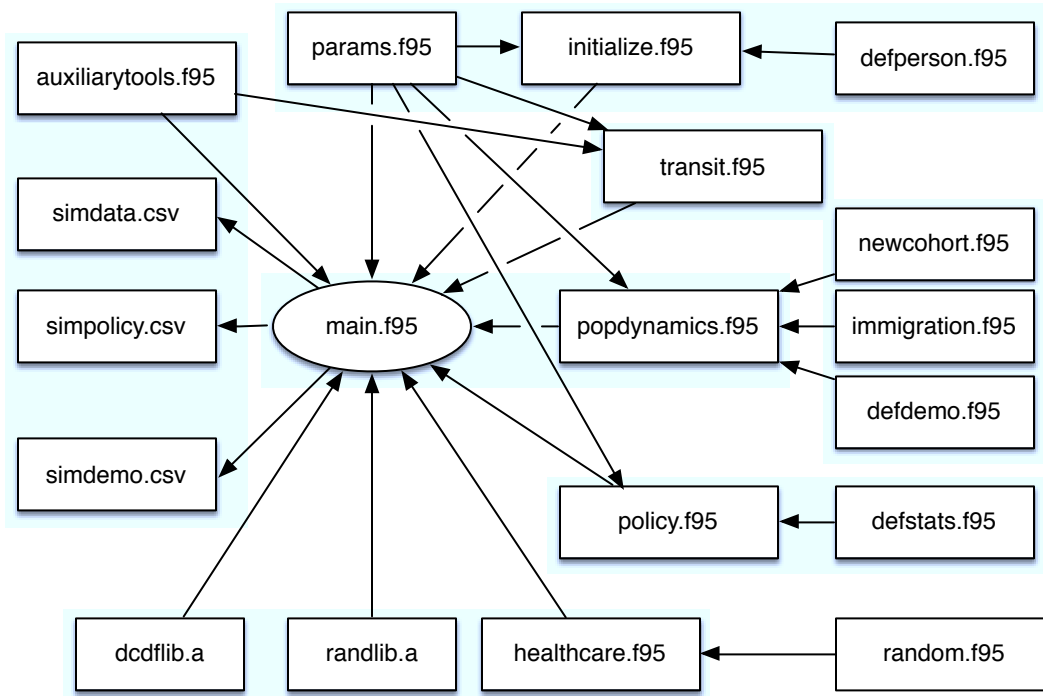


FIGURE 1.2 – Structure du code FORTRAN

Le programme principal du modèle est donc **main.f95**. Il appelle un ensemble de modules FORTRAN

se rapportant aux différentes composantes du modèle tel que présenté à la figure 1.1, plus haut. Les sous-sections qui suivent présentent les différents modules.

### 1.2.1 Paramètres

Le module *params* contient les différents paramètres du modèle. Ceux-ci sont utilisés dans presque tous les modules. L'année de départ (*startyear*) et l'année de fin (*stopyear*) de la simulation, tout comme l'année où la dernière cohorte entre (*stopyear\_enter*), sont des paramètres qui peuvent être changés par l'utilisateur. Tel que mentionné, les valeurs par défaut de ces paramètres sont 2010, 2130 et 2050, respectivement.

Le paramètre *nyears* correspond non pas au nombre d'années, mais plutôt au nombre de cycles dans la simulation. Celui-ci peut notamment varier en fonction de la dernière année de simulation : si *stopyear\_enter* prévaut, alors la simulation s'arrête à la dernière année où une cohorte entre (par défaut 2050) ; si *stopyear* prévaut, alors la simulation se poursuit jusqu'à ce que tous les agents soient décédés ou jusqu'à ce que l'année *stopyear* (par défaut 2130) soit atteinte. De façon importante, *nyears* dépend également du nombre d'années entre chaque transition démographique (*gapyears* ; voir plus bas pour les transitions), lui-même fixé par défaut à 2. C'est donc dire que pour une période allant de 2010 à 2050, le modèle effectuera par défaut 20 cycles de simulation ; ici, on a donc  $nyears = (stopyear - startyear) / gapyears$ .

Les paramètres *startage* et *stopage* déterminent l'âge minimal et maximal permis dans la simulation (par défaut 30 et 110 ans), tandis que *nages* correspond au nombre d'âges différents ou à la différence entre *stopage* et *startage*.

La taille de l'échantillon de départ (*nstartpop*) et celle des nouvelles cohortes (*nstartcohort*) sont ensuite déterminées. Ceci n'est pas lié à la taille de la population de départ. Puisque chaque agent aura un poids populationnel qui sera fonction de ses attributs, celui-ci s'ajustera en fonction de la taille de l'échantillon de départ. Plus les valeurs de *nstartpop* et de *nstartcohort* seront importantes, plus les résultats de simulation seront précis et stables. La simulation prendra cependant plus de temps.

Les quatre paramètres suivants déterminent des choix dans le modèle. Le paramètre *cohort* permet de contrôler le mode de simulation : mode cohorte ou mode population. On peut aussi choisir de sauvegarder les données individuelles simulées (*savedata*), les indicateurs de tableau de bord (*savepolicy*) et les données démographiques (*demo*).

Par ailleurs, on peut choisir de faire plusieurs réplifications des simulations (*doreps*). Le paramètre *nreps* indique le nombre de réplifications.

Les trois derniers paramètres concernent des hypothèses démographiques. Le paramètre *migration* détermine s'il y a migration ou non tandis que *immigrate* détermine le taux d'immigration nette en proportion de la population. Finalement, le paramètre *qxgain* détermine l'amélioration annuelle des taux de mortalité, en %.

### 1.2.2 Agents

Le modèle de microsimulation fonctionne à partir d'une population d'agents. Ces agents sont définis dans le fichier **defperson.f95**. Le type « person » est une structure (en langage C++) qui définit les attributs d'une personne; on y trouve toutes les caractéristiques de cette personne. Par exemple, on peut déclarer une population d'agents constituée de  $n$  personnes en créant « type (person) pop(n) ». Par la suite, on peut obtenir l'âge de la personne  $i$  à l'aide de la commande « pop(i)%age ». La combinaison de *id* et *byear* constitue un indicateur unique pour chaque personne. L'attribut « wgt » est le poids populationnel de l'agent.

### 1.2.3 Initialisation

Le module *initialize* est celui qui crée la population de l'année de départ du modèle à partir de l'ESCC 2010. Le chapitre 3 explique plus en détails la manière par laquelle l'initialisation est effectuée. Le modèle tire donc de cette enquête un échantillon avec remplacement de taille *nstartpop*. Il calibre ensuite la distribution d'âges tirée de l'enquête sur celle produite par l'ISQ et ajuste les poids de chaque agent en conséquence.

Le programme qui initialise la population se trouve dans **src/initialize.f95**.

### 1.2.4 Transitions

Le module *transit* est appelé par **main.f95** pour donner la population de départ de l'année  $t + 1$  à partir de la population de l'année  $t$ , avant que la transition démographique ne soit effectuée (c.-à-d. avant que les morts soient supprimés et que les nouveaux entrants et les immigrants soient ajoutés). Ce module contient la définition des probabilités de transition pour chacun des attributs de l'agent. C'est dans ce module qu'on fait la mise à jour des caractéristiques de l'agent. Si l'agent présente la valeur *alive=.false.*, il sera par la suite éliminé dans le module *popdynamics*. Le chapitre 4 présente les différents modèles utilisés afin de calculer les probabilités de transition.

Le programme du module est donné dans **src/transit.f95**.

### 1.2.5 Dynamique populationnelle

Le module *popdynamics* est appelé par **main.f95** pour faire la transition démographique et pour calculer les statistiques démographiques. Dans ce module, les morts sont d'abord éliminés. Par la suite le module *newcohort* est utilisé pour générer une nouvelle cohorte. Finalement, le module *immigration* est appelé pour ajuster les poids des immigrants de telle sorte qu'ils tiennent compte de l'immigration nette au Québec. Les statistiques démographiques sont calculées *avant* chaque transition démographique; ces statistiques sont définies dans le module **defdemo.f95**.

Le programme qui accomplit la transition démographique se trouve dans **src/popdynamics.f95**.

### 1.2.6 Nouvelles cohortes

Le module *newcohorts* génère les nouvelles cohortes en deux temps. D'abord, il génère un échantillon de taille *nnewcohort* à partir de la distribution conjointe de probabilités des conditions initiales dans l'échantillon des individus ayant 30 et 31 ans dans l'ESCC. Par la suite, les poids sont ajustés afin de tenir compte de la taille des cohortes entrantes projetée par l'ISQ dans ses projections démographiques. Ce module est appelé par *popdynamics*.

Le programme qui génère les nouvelles cohortes se trouve dans **src/newcohorts.f95**.

### 1.2.7 Immigration

Le module *immigration* ajuste les poids de la population actuelle pour tenir compte de l'entrée de nouveaux immigrants. Ainsi, il n'y a pas de nouvelles observations qui s'ajoutent à la population afin de tenir compte de cette entrée. C'est plutôt une repondération des observations existantes. Le module *immigration* est appelé par *popdynamics*.

### 1.2.8 Tableau de bord

Le tableau de bord est contrôlé par le module *policy*. Ce module utilise comme intrant une population d'agents. Il calcule différentes statistiques portant sur leur santé mais aussi sur leur utilisation de soins de santé.

Le programme du module est donné dans **src/policy.f95**.

Le programme qui définit quelles statistiques sont disponibles se trouve dans **src/defstats.f95**.

### 1.2.9 Exécution du programme

Le programme FORTRAN utilise deux bibliothèques de support, **dcdflib.a** et **ranlib.a**, qui regroupent des fonctions de probabilité et de tirage aléatoire. Celles-ci se retrouvent dans le répertoire **libs/** et des fichiers bash **create.sh** peuvent créer ces bibliothèques pour l'architecture désirée (PC, Mac, Linux 32 et 64 bit).

En Linux et en Mac, le fichier qui crée l'exécutable du modèle se trouve dans **build/compile\_linux.sh**. Il contient une collection de commandes bash qui feront la compilation des fichiers de code FORTRAN. Le compilateur FORTRAN utilisé est gfortran. Si le programme s'exécute correctement, il ne reste qu'à exécuter le modèle avec la commande `./runmodel`.

Le modèle crée trois fichiers d'output : **simdata.csv**, **simpolicy.csv** et **simdemo.csv**. Ces fichiers peuvent être lus avec un logiciel comme Stata pour produire des graphiques, etc.

### 1.3 Organisation des fichiers et *subversion*

Le répertoire « sante » est en *version control*, ce qui signifie que les versions du modèle sont sauvegardées dans un répertoire Assembla (<http://www.assembla.com>). Ceci permet de conserver les anciennes versions sans avoir à sauvegarder des copies du modèle. Une introduction à *subversion* est disponible à [http://dev.nozav.org/intro\\_svn.html](http://dev.nozav.org/intro_svn.html).

## 2 Données

### 2.1 Survol

Ce chapitre présente la préparation des données servant à la construction du modèle de microsimulation. Nous utilisons diverses banques de données pour construire le modèle. De loin l'enquête la plus importante est l'Enquête nationale sur la santé de la population (ENSP), qui est la seule enquête longitudinale sur la santé au Canada. Nous l'utilisons pour construire le module de transitions ainsi que le module d'utilisation des soins de santé. L'Enquête sur la santé dans les collectivités canadiennes (ESCC) de 2010 est utilisée pour construire la population initiale du modèle. Elle est représentative de la population canadienne de 2010. Finalement, nous utilisons des données de l'Institut de la statistique du Québec (ISQ) concernant la taille de la population par âge afin de reproduire la distribution de la population par âge dans la population initiale du modèle. Nous donnons plus de détails sur les sources complémentaires de données dans les chapitres suivants.

### 2.2 ENSP

L'ENSP est une enquête longitudinale effectuée auprès d'un échantillon de la population canadienne. Le questionnaire a été administré au même échantillon de répondants tous les deux ans à partir de 1994-1995. Il y a donc neuf cycles à cette enquête, le dernier datant de 2010-2011. Bien qu'il y ait trois volets à l'ENSP (le volet Ménages, le volet Établissements de soins santé et le volet Nord), seul le volet Ménages sera utilisé ici. Le volet Nord n'est plus développé par l'ENSP depuis 2000-2001, tandis que le volet Établissements de soins de santé a dû être interrompu en 2002-2003 car une trop grande proportion de l'échantillon était décédée ([Réseau de recherche en santé et en sécurité du travail du Québec, 2014](#)).

Une particularité du volet Ménages est que tous les répondants interrogés vivent forcément en ménages privés lors du premier cycle de l'enquête. L'enquête exclut donc les résidents des réserves indiennes, des terres de la Couronne et des régions éloignées du Québec et de l'Ontario en plus des membres à temps plein des Forces canadiennes et des personnes vivant en institutions ([Statistique Canada, 2012a](#)). Il est à noter que s'il s'avère qu'un répondant du volet Ménages transite vers une institution entre deux cycles de l'enquête, il est tout de même suivi dans le volet Ménages.

En 1994, il y a 17 276 répondants au Canada, dont 2 740 ont 65 ans ou plus. En comparaison, il y a 3 000 répondants au Québec dont 369 ont 65 ans ou plus. Le [tableau 2.1](#) présente le nombre d'observations dans l'ENSP par groupe d'âge pour le Québec et le reste du Canada. On peut conclure que le nombre

d’observations au Québec est faible, surtout après 65 ans. Étant donné que certains problèmes de santé sont relativement rares, il nous apparaît plus prudent de considérer l’échantillon canadien total et de permettre des déviations au besoin pour le Québec.

	Québec	Reste du Canada
20 à 24 ans	172	1047
25 à 29 ans	254	1068
30 à 34 ans	290	1361
35 à 39 ans	286	1264
40 à 44 ans	264	1049
45 à 49 ans	218	1012
50 à 54 ans	189	784
55 à 59 ans	151	731
60 à 64 ans	148	670
65 à 69 ans	138	685
70 à 74 ans	117	670
75 à 79 ans	68	492
80 à 84 ans	25	356
85 à 89 ans	22	136

Tableau 2.1 – Nombre d’observations pour le Québec et le Canada dans l’ENSP (1994-2011)

### 2.2.1 Pondération

Afin de procéder à des estimations à partir des données de l’ENSP, il faut utiliser des poids d’échantillonnage. Ces derniers permettent à l’échantillon sélectionné de représenter l’ensemble de la population canadienne. Chaque personne de l’échantillon représente un certain nombre d’individus au sein de la population totale. Un poids initial est calculé puis ajusté pour tenir compte de certaines particularités de l’enquête. Ce premier poids est l’inverse de la probabilité de sélection. Ainsi, plus un individu a de chances d’être sélectionné dans l’échantillon, plus son poids est faible. Ce poids est calculé au premier cycle de l’enquête (1994-1995) et ne varie pas dans le temps. Puisqu’il n’y a pas ré-échantillonnage au sein de l’ENSP, ce poids permet de représenter la population d’intérêt telle qu’elle était en 1994, lorsque les observations de départ étaient présentes ([Statistique Canada, 2012a](#)).

[Statistique Canada \(2012a\)](#) calcule ensuite d’autres poids qui permettent de rendre un sous-ensemble de l’échantillon représentatif de la population canadienne. Le deuxième poids s’applique seulement aux individus ayant répondu à l’enquête pour les neuf cycles (ou qui sont décédés ou en établissement lors de certains cycles). En effet, certains individus ne répondent pas à tous les cycles de l’enquête.

Une enquête longitudinale portant sur la santé nécessite un bon suivi des répondants n’ayant pu être interviewés. En particulier, il est important de pouvoir bien identifier les répondants qui sont décédés. Il est également important de vérifier la fréquence à laquelle des gens quittent l’ENSP et de voir si leurs caractéristiques sont différentes de celles des individus qui restent. Nous abordons donc ce dernier point ci-après avant de discuter de la mortalité.

## 2.2.2 Attrition et sélection

Le tableau 2.2 montre les taux de réponse cumulatifs, les taux d'attrition (la proportion de ceux ayant répondu au cycle précédent qui ne répondent pas au cycle courant) et la proportion d'entrants (des répondants ayant quitté entre deux cycles mais qui sont revenus par la suite). Pour les entrants, le logiciel d'enregistrement des réponses de Statistique Canada s'assure qu'il n'y a pas d'incohérences dans les réponses des individus entre le cycle précédent leur absence de l'enquête et le cycle suivant celle-ci (Statistique Canada, 2012a).

	Taux de réponse	Taux d'attrition	Proportion d'entrants
1994-1995	100,00%	.	.
1996-1997	92,86%	7,14%	0,00%
1998-1999	87,55%	6,75%	1,44%
2000-2001	83,61%	7,24%	3,29%
2002-2003	79,62%	7,57%	3,59%
2004-2005	74,66%	8,53%	3,57%
2006-2007	73,13%	6,50%	4,97%
2008-2009	65,62%	10,05%	2,54%
2010-2011	64,79%	6,50%	5,66%

Tableau 2.2 – Taux de réponse et attrition par cycle dans l'ENSP

Les taux d'attrition sont généralement faibles et se comparent à ce qui est observé dans d'autres enquêtes longitudinales fréquemment utilisées, comme le Health and Retirement Study (HRS) et le Panel Study of Income Dynamics (PSID) américains. L'attrition au sein du HRS est de 35,7 % entre 1992 et 2002 (Kapteyn *et al.*, 2006). Ainsi, en 2002, 64,3 % de l'échantillon de départ a répondu à l'ensemble des cycles. Quant au PSID, l'attrition observée est de 2,5 à 3% par année entre 1968 et 1989 (Fitzgerald *et al.*, 1998). En comparaison, entre 1994 et 2004, l'attrition au sein de l'ENSP n'est que de 22,4 %. Enfin, dans l'ENSP, une bonne partie des répondants ayant quitté revient dans le panel.

Nous analysons au tableau 2.3 si l'état de santé des individus demeurant dans le panel de façon continue de 1994-1995 à 2010-2011 est significativement différent de celui des individus non décédés qui ne répondent pas lors d'un cycle ou plus. Le tableau présente des statistiques de Student de différence de moyennes calculées sur les données de 1994. Par exemple, une valeur supérieure à 1,96 ou inférieure à -1,96 pour une variable signifie qu'on rejette à un seuil de confiance de 5% l'hypothèse nulle que la moyenne de cette variable pour les individus qui répondront tout au long de l'enquête est identique à celle des individus qui cesseront de répondre. Dit autrement, une valeur supérieure à 1,96 ou inférieure à -1,96 indique une différence entre les deux groupes, à un seuil de 5%. Des différences significatives pourraient causer problème, car la non-réponse entraînerait alors une perte de représentativité de l'échantillon.

Les tests se révèlent significatifs au seuil de 5% pour quelques variables. Pour certains groupes d'âge, la présence du diabète (65 ans et plus) et des maladies cardiaques (40 à 64 ans) est plus élevée en 1994 pour ceux qui cesseront plus tard de répondre à l'enquête. Cela est également vrai pour les fumeurs plus âgés, pour les anciens fumeurs de moins de 65 ans, ainsi que pour les individus souffrant de plusieurs incapacités.

	moins de 40 ans	40 à 64 ans	65 ans et plus
Diabète	-1,0222	-0,3774	-2,1553
Hypertension	-0,2972	-1,9783	-0,7707
Cancer	0,8821	-1,5148	-1,5080
Maladies cardiaques	0,3777	-2,2666	-1,6501
AVC	1,4179	-0,3828	-0,3224
Maladies pulmonaires	0,4293	0,0259	-0,9210
Démences	.	0,3665	-0,4008
Fumeurs actuels	-1,0222	-0,3774	-2,1553
Anciens fumeurs	-5,2900	-4,1568	-1,5910
Aucune incapacité	0,5909	-1,0680	4,9315
1 ADL ou trouble cognitif	0,8370	0,9817	-2,0228
2+ ADL	-2,5082	0,4724	-4,4845

Tableau 2.3 – Tests statistiques de sélection, ENSP (1994-2011) : Différence de proportions entre les individus répondant à tous les cycles et ceux pour qui il manque des cycles. La statistique t est représentée. Une valeur négative indique une prévalence plus élevée chez les individus qui vont quitter.

### 2.2.3 Mortalité

L’ENSP offre deux façons de détecter la mortalité. La première consiste pour un intervieweur à obtenir cette information lorsqu’il tente d’entrer en contact avec le répondant. La deuxième passe par une validation auprès de la banque nationale des décès. Pour que cette validation soit possible, le répondant doit avoir consenti à un jumelage avec les données administratives. Environ 90% des répondants ont consenti à ce jumelage.

Le tableau 2.4 présente les taux de mortalité annuels de l’ENSP par cycle pour la population de 30 ans et plus. Pour les premiers cycles, les taux augmentent d’année en année. Une raison qui fait que les taux sont initialement bas est que la population initiale de l’ENSP est uniquement constituée d’individus qui ne sont pas en institution. Le taux de mortalité diminue ensuite vers les derniers cycles de l’enquête, en raison notamment de problèmes avec la validation auprès de la banque nationale des décès.

	Taux de mortalité
1994-1995	0,98%
1996-1997	1,13%
1998-1999	1,13%
2000-2001	1,15%
2002-2003	1,28%
2004-2005	1,25%
2006-2007	1,05%
2008-2009	0,85%
Total	1,10%

Tableau 2.4 – Taux de mortalité annuels par cycle de l’ENSP, population de 30 ans et plus

La figure 2.1 présente les taux de mortalité par âge provenant de l’ENSP de 1994 à 2006, ainsi que l’intervalle de confiance de 95% autour de ces taux. Nous comparons ces taux avec ceux provenant de tables de mortalité périodiques de la [Human Mortality Database](#) pour une période similaire, soit



de 1995 à 2004. L'ensemble des données concernant le Canada proviennent de Statistique Canada. La correspondance est bonne jusqu'à 75 ans environ. Toutefois, il y a divergence après 75 ans, soit à partir du moment où le nombre d'individus en institution augmente. Puisque les individus qui répondent à l'ENSP ne sont pas en institution lors du premier cycle d'enquête, ils sont probablement en meilleure santé que la population canadienne dans son ensemble. Ceci peut expliquer que le taux de mortalité de l'ENSP soit plus faible que celui calculé à l'aide de tables de mortalité périodiques à partir de 75 ans. Nous recommandons donc un calibrage éventuel de la mortalité de l'ENSP en utilisant les tables de survie.

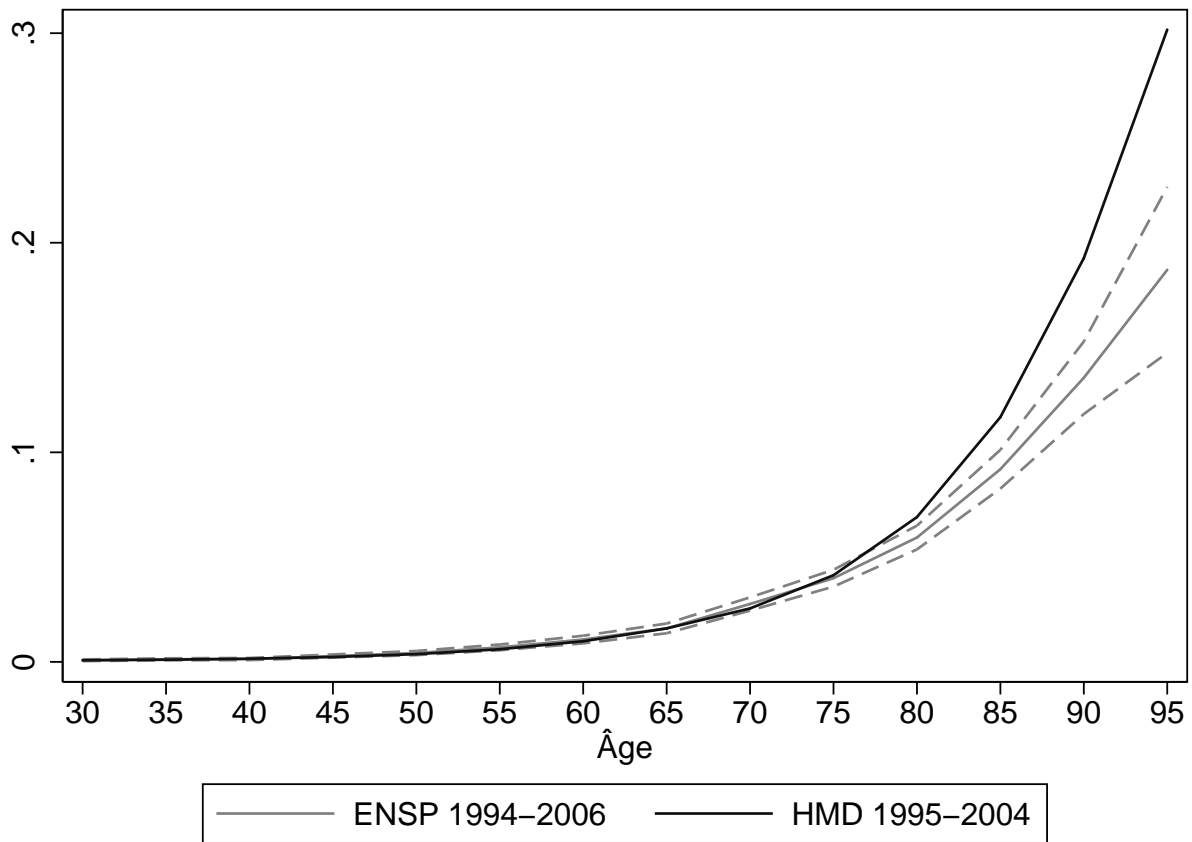


FIGURE 2.1 – Comparaison des taux de mortalité : ENSP vs. Human Mortality Database (HMD). Intervalle de confiance de 95% en pointillés.

### 2.3 Construction des variables

Le questionnaire de l'ENSP est riche et les questions sont, pour la grande majorité, identiques d'un cycle à l'autre. Nous divisons cette section en trois. En premier lieu, nous présentons la construction des variables relatives aux problèmes de santé auto-déclarés. Ces variables comportent des mesures d'incidence (commencer à être atteint d'une maladie entre deux points dans le temps) et de prévalence

(souffrir d’une maladie à un moment donné). Par la suite, nous décrivons les mesures d’invalidité. Nous présentons aussi les mesures de facteurs de risque. Finalement, nous décrivons les variables d’utilisation de soins de santé.

Les dimensions de la santé qui sont prises en compte sont présentées au tableau 2.5. L’état de santé d’une personne comporte plusieurs dimensions dans le modèle.

Diabète
Hypertension
Cancer
AVC
Maladies cardiaques
Maladies pulmonaires
Alzheimer et autres démences
Obésité
Tabagisme
Nombre d’incapacités
Institutionnalisation

Tableau 2.5 – Dimensions de l’état de santé prises en compte dans COMPAS

Pour ce qui est des variables d’utilisation de soins de santé, nous utilisons pour le moment les dimensions apparaissant dans le tableau 2.6. Dans les versions futures du modèle, nous raffinerons ces variables et considérerons d’autres dimensions comme les coûts, etc.

Nombre de visites chez un généraliste
Nombre de visites chez un spécialiste
Nombre de nuits d’hospitalisation de courte durée
Utilisation de médicaments (oui / non)
Utilisation de services à domicile (oui / non)

Tableau 2.6 – Dimensions de l’utilisation des soins de santé prises en compte dans COMPAS

Nous utilisons les poids dans les statistiques qui suivent. Dans cette version, nous n’avons pas stratifié par province ou par sexe. Cependant, afin de vérifier les différences qui peuvent exister entre le Québec et le Canada, nous présentons pour tous les résultats des tests de comparaison de Student, qui indiquent si la proportion pour le reste du Canada est significativement différente de celle pour le Québec.

### 2.3.1 Problèmes de santé auto-déclarés

Le modèle comporte sept mesures de santé auto-déclarées : présence de diabète, d’hypertension, de cancer, de maladies cardiaques, d’AVC, de maladies pulmonaires et présence de démences (Alzheimer et autres). Ces maladies sont les plus prévalentes au sein de la population. Les questions posées dans l’ENSP prennent la forme suivante : « *Rappelez-vous qu’on s’intéresse à des problèmes de santé qui ont été diagnostiqués par un professionnel de la santé. Souffrez-vous de diabète ?, ...d’hypertension ?, etc.* » (Statistique Canada, 2012b). Une fois agrégées, les réponses de l’ensemble des répondants à ces questions mènent donc directement à des mesures de prévalence. Pour développer les mesures d’incidence, nous utilisons la réponse positive au cycle courant dans le cas où la réponse était négative au cycle précédent.

Le tableau 2.7 donne les taux de prévalence par groupe d'âge de 5 ans pour chacune des maladies. Bien sûr, on remarque que la prévalence augmente avec l'âge, sauf pour certaines maladies (diabète, cancer) dont la prévalence diminue aux âges avancés. La maladie la plus fréquente chez les 40-44 ans est l'hypertension, avec un taux de prévalence de 6,8%. Chez les 65-69 ans, la maladie la plus fréquente est également l'hypertension, avec 39,7%.

On observe des différences entre le Québec et le Canada en ce qui concerne la prévalence des maladies. Comme l'indiquent les statistiques  $t$  présentées dans la deuxième partie du tableau 2.7, chez les individus de moins de 40 ans, la prévalence du diabète, de cancers et de démences est plus élevée au Québec que dans le reste du Canada, tandis que c'est l'inverse pour l'hypertension.<sup>4</sup> À partir de 65 ans, la prévalence du diabète est toujours plus importante au Québec que dans le reste du Canada. Toutefois, une plus faible proportion de Québécois souffre de cancers, de maladies cardiaques et d'AVC que les Canadiens qui habitent à l'extérieur du Québec.

Âge	Diabète	Hypertension	Cancer	Maladies cardiaques	AVC	Maladies pulmonaires	Démences
20 à 24 ans	0,59%	0,94%	0,50%	1,28%	0,21%	5,10%	0,02%
25 à 29 ans	0,86%	2,30%	0,65%	1,14%	0,24%	4,84%	0,17%
30 à 34 ans	0,93%	3,37%	0,85%	1,32%	0,24%	4,79%	0,26%
35 à 39 ans	1,23%	5,11%	1,03%	1,93%	0,26%	4,99%	0,25%
40 à 44 ans	2,18%	6,78%	1,54%	2,59%	0,35%	5,21%	0,20%
45 à 49 ans	3,23%	10,22%	1,98%	3,53%	0,60%	5,12%	0,24%
50 à 54 ans	4,61%	16,91%	3,03%	5,18%	0,93%	5,33%	0,26%
55 à 59 ans	6,97%	25,33%	4,24%	7,50%	1,51%	6,09%	0,33%
60 à 64 ans	9,40%	34,01%	6,10%	11,65%	2,65%	7,56%	0,72%
65 à 69 ans	11,77%	39,68%	7,46%	16,37%	3,36%	8,36%	1,07%
70 à 74 ans	14,48%	44,95%	8,71%	21,08%	5,59%	10,73%	1,84%
75 à 79 ans	16,45%	49,15%	9,80%	26,28%	6,93%	11,98%	3,70%
80 à 84 ans	17,63%	52,55%	11,05%	32,66%	9,25%	13,72%	6,01%
85 à 89 ans	16,27%	53,99%	15,23%	35,96%	11,92%	14,35%	11,28%
90 à 94 ans	15,06%	52,15%	14,22%	37,76%	13,33%	13,73%	14,68%
95 à 99 ans	14,79%	55,45%	13,18%	37,13%	20,47%	20,20%	18,81%
Total	5,17%	17,12%	3,41%	7,39%	1,75%	6,41%	0,90%
Test t de la différence entre les taux pour le Québec et le Canada							
moins de 40 ans	-4,3785	4,758	-4,2359	-1,8561	0,7613	1,3227	-3,4427
40 à 64 ans	2,4845	4,591	2,6713	0,347	-3,3479	1,35	-2,9588
65 ans et plus	-3,1319	-0,1904	7,401	2,5295	3,997	-4,6766	1,2087

Tableau 2.7 – Prévalence auto-déclarée des maladies dans l'ENSP (1994-2011)

L'interprétation des taux de prévalence par âge est difficile en raison de deux facteurs importants : a) les effets de cohorte (les différences intergénérationnelles pour un âge donné) dans l'incidence des maladies ; et b) la mortalité plus élevée de ceux ayant une maladie. Ainsi, on pourrait très bien observer que l'incidence augmente avec l'âge alors que la prévalence diminue.

Au tableau 2.8, nous présentons les taux d'incidence par groupe d'âge et pour chacune des maladies.

4. Dans tous les tableaux de ce chapitre, une statistique  $t$  négative indique une proportion plus élevée au Québec qu'au Canada, tandis qu'une statistique  $t$  positive indique une proportion plus faible au Québec. La significativité de ces différences dépend de la taille de la statistique ; p.ex., une statistique supérieure à 1,96 en valeur absolue indique une différence significative à un seuil de 5%. La direction de la différence dépend du signe de la statistique, tel qu'expliqué.

Chez les 40-44 ans, l'incidence la plus élevée est pour l'hypertension avec 1,6%. Cette dernière présente d'ailleurs l'incidence la plus élevée jusqu'à 80 ans. On remarque que l'incidence augmente généralement avec l'âge, bien qu'elle diminue pour certaines maladies aux âges avancés. Plusieurs raisons peuvent expliquer ceci, dont la sélection et le processus biologique menant à ces maladies.

Âge	Diabète	Hypertension	Cancer	Maladies cardiaques	AVC	Maladies pulmonaires	Démences
20 à 24 ans	0,15%	0,45%	0,11%	0,13%	.	0,63%	0,06%
25 à 29 ans	0,11%	0,71%	0,15%	0,15%	0,03%	0,52%	0,05%
30 à 34 ans	0,17%	1,00%	0,09%	0,41%	0,03%	0,63%	0,03%
35 à 39 ans	0,33%	1,19%	0,27%	0,43%	0,04%	0,59%	0,02%
40 à 44 ans	0,67%	1,61%	0,37%	0,54%	0,13%	0,52%	0,07%
45 à 49 ans	0,69%	2,80%	0,44%	0,99%	0,18%	0,54%	0,03%
50 à 54 ans	1,12%	4,24%	0,84%	1,18%	0,15%	0,66%	0,03%
55 à 59 ans	1,55%	5,68%	1,09%	1,74%	0,42%	1,07%	0,18%
60 à 64 ans	1,96%	5,52%	1,21%	2,45%	0,64%	0,96%	0,22%
65 à 69 ans	1,68%	6,44%	1,49%	2,76%	0,71%	1,37%	0,38%
70 à 74 ans	2,00%	5,46%	2,36%	4,51%	1,22%	1,43%	1,10%
75 à 79 ans	2,55%	5,74%	1,83%	4,24%	1,78%	1,75%	1,69%
80 à 84 ans	1,24%	4,68%	2,74%	4,99%	2,22%	1,83%	3,80%
85 à 89 ans	2,27%	4,31%	1,67%	5,18%	3,00%	2,13%	6,08%
90 à 94 ans	1,10%	4,53%	1,94%	6,70%	4,99%	4,31%	7,09%
95 à 99 ans	.	4,90%	.	5,41%	4,68%	2,89%	8,16%
Total	0,89%	2,82%	0,73%	1,34%	0,40%	0,83%	0,39%
Test t de la différence entre les taux pour le Québec et le Canada							
moins de 40 ans	-1,3992	0,3709	-2,3159	-0,2394	0,1002	1,0728	-1,2011
40 à 64 ans	1,1932	-0,2533	0,9679	-0,9549	-1,7444	-1,9855	0,1098
65 ans et plus	-1,0975	-1,2604	2,7038	0,5042	3,1001	-1,6169	0,8102

Tableau 2.8 – Incidence auto-déclarée des maladies dans l'ENSP (1994-2011)

### 2.3.2 Invalidité et institutionnalisation

Nous avons voulu construire une mesure d'invalidité qui comprenait les incapacités cognitives et physiques ainsi que l'institutionnalisation. Ainsi, nous avons d'abord utilisé les variables du questionnaire pour construire une mesure d'incapacité cognitive, une mesure d'incapacité physique et une mesure d'institutionnalisation. Nous présentons chacune de ces mesures.

#### *Incapacités cognitives*

Nous utilisons une variable auto-déclarée qui indique si l'individu a fréquemment des problèmes de mémoire qui nuisent à ses activités.

#### *Nombre de limitations d'activités quotidiennes (ADL)*

Nous créons une variable qui dénote le nombre de limitations d'activités quotidiennes déclarées par l'individu.

#### *Institutionnalisation*

L'ENSP indique clairement si l'individu est en institution. Les institutions, ou établissements, sont définis comme suit dans l'ENSP :

1. Établissements pour personnes âgées, y compris les établissements de soins accueillant des personnes âgées et les hôpitaux de soins prolongés ;
2. Établissements de soins cognitifs, y compris les établissements de soins accueillant des enfants présentant des troubles affectifs et des personnes présentant des troubles mentaux ou des troubles de retard du développement et les hôpitaux psychiatriques ;
3. Autres établissements de réadaptation, y compris les centres de réadaptation, les hôpitaux pédiatriques et d'autres hôpitaux spécialisés, les hôpitaux généraux comprenant des unités de soins de longue durée, ainsi que les établissements de soins accueillant des personnes physiquement handicapées.

*Variable d'incapacités construite pour le modèle*

La variable créée pour le modèle combine ces trois mesures. Nous définissons 4 états : 1) aucune incapacité ; 2) Une seule ADL ou incapacité cognitive ; 3) 2 ADL ou plus, avec ou sans incapacité cognitive ; 4) en institution.

Le tableau 2.9 présente la distribution de cette variable par groupe d'âge. Chez les 40-44 ans, peu de répondants souffrent d'incapacités, soit moins de 5%. Ce taux augmente à 9,7% pour les 65-69 ans. Entre 90 et 94 ans, environ 58% des individus souffrent d'incapacités. Finalement, au Québec, la proportion d'individus qui n'ont pas d'incapacités est plus élevée que dans le reste du Canada, et ce pour tous les groupes d'âges. Une telle conclusion s'applique également à la proportion de Québécois âgés de 40 à 64 ans qui est en institution.

Âge	Aucune incapacité	1 ADL ou trouble cognitif	2+ ADL	Institution
20 à 24 ans	96,34%	2,68%	0,97%	.
25 à 29 ans	95,85%	3,09%	1,03%	.
30 à 34 ans	96,43%	2,51%	1,06%	.
35 à 39 ans	95,92%	2,56%	1,52%	.
40 à 44 ans	95,28%	2,78%	1,92%	.
45 à 49 ans	94,33%	2,90%	2,66%	0,10%
50 à 54 ans	93,51%	3,52%	2,90%	0,07%
55 à 59 ans	92,47%	3,80%	3,64%	0,08%
60 à 64 ans	91,23%	4,76%	3,93%	0,09%
65 à 69 ans	90,30%	4,73%	4,82%	0,15%
70 à 74 ans	85,56%	7,43%	6,44%	0,58%
75 à 79 ans	78,27%	10,12%	9,96%	1,65%
80 à 84 ans	68,50%	12,18%	15,32%	4,00%
85 à 89 ans	53,69%	13,85%	23,93%	8,53%
90 à 94 ans	41,81%	12,37%	30,37%	15,45%
95 à 99 ans	21,84%	9,14%	41,01%	28,01%
Total	91,80%	4,10%	3,62%	0,48%
Test t de la différence entre les taux pour le Québec et le Canada				
moins de 40 ans	-4,5711	4,5356	1,3265	1,1331
40 à 64 ans	-6,6771	4,5927	5,1438	-2,1051
65 ans et plus	-1,6175	0,9178	0,0735	2,2091

Tableau 2.9 – État d'invalidité selon l'ENSP (1994-2011)

Dans le tableau 2.10, nous montrons deux matrices de transitions entre les états que nous avons définis. L'une est calculée pour les répondants âgés de moins de 65 ans et l'autre, pour ceux de 65 ans ou plus. Ces taux de transitions sont calculés pour les répondants ayant survécu d'un cycle à l'autre. Parmi les personnes âgées de moins de 65 ans, il y a peu de transitions d'une période à l'autre lorsque les individus n'ont pas d'invalidité. En effet, 98% des personnes qui n'ont pas d'incapacité lors de la période courante n'en ont pas lors de la période suivante. Cependant, seules 42% des personnes ayant une incapacité demeurent dans cet état lors de la période suivante. À partir de 65 ans, les transitions à travers différents états d'incapacités sont plus nombreuses. Tel que l'indique le tableau, seules 38% des personnes ayant une incapacité à la période courante en ont encore une à la période suivante. Les autres ont soit au moins deux limitations (21%) ou une aucune invalidité (37%). Ainsi, il est possible pour un individu qui a une incapacité physique à la période courante de ne plus l'avoir à la période suivante. Étant donné que les incapacités physiques proviennent de limitations dans des activités telles que la préparation des repas, il est possible que la prise de médicaments, par exemple, permette à un individu de ne plus avoir de problèmes à effectuer une certaine activité. On note également qu'il y a peu de différences entre le Québec et le reste du Canada en ce qui concerne les transitions d'une période à l'autre, et ce peu importe l'âge.

<b>Moins de 65 ans</b>				
Invalidité à la période courante	Invalidité à la période suivante			
	Aucune incapacité	1 ADL	2+ ADL	Institution
Aucune incapacité	97,55%	1,69%	0,75%	0,01%
1 ADL ou trouble cognitif	49,93%	42,19%	7,78%	0,10%
2+ ADL	27,56%	12,53%	59,56%	0,35%
Institution	.	.	.	.
Test t de la différence entre les taux pour le Québec et le Canada				
Aucune incapacité	-1,293	1,6794	0,0974	-0,2918
1 ADL ou trouble cognitif	-2,0808	2,2628	0,4195	0,8894
2+ ADL	-4,3156	1,0613	3,1151	0,0207
Institution	0,6176	0,6176	0,6176	-1,1072
<b>65 ans et plus</b>				
Invalidité à la période courante	Invalidité à la période suivante			
	Aucune incapacité	1 ADL	2+ ADL	Institution
Aucune incapacité	88,61%	5,88%	4,82%	0,70%
1 ADL ou trouble cognitif	37,38%	38,48%	21,00%	3,14%
2+ ADL	16,86%	10,88%	62,92%	9,33%
Institution	6,47%	.	.	92,25%
Test t de la différence entre les taux pour le Québec et le Canada				
Aucune incapacité	0,6783	0,5198	1,6862	0,2358
1 ADL ou trouble cognitif	0,0340	0,4836	0,5498	0,1221
2+ ADL	-1,1897	-0,2968	0,0194	1,6029
Institution	1,8798	.	2,1114	2,5498

Tableau 2.10 – Transitions entre les états d'invalidité selon l'ENSP (1994-2011)

### 2.3.3 Facteurs de risque

Nous avons considéré deux facteurs de risque pour l’instant, soit l’obésité et le tabagisme. Pour ce qui est du tabagisme, nous avons défini trois états possibles : 1) jamais fumé ; 2) fumeur actuel ; et 3) ancien fumeur. Le tableau 2.11 montre la répartition des répondants selon leur âge. Au tableau 2.12, nous montrons par groupe d’âge les taux d’initiation et de cessation, ou transitions entre états de tabagisme. On voit que le taux de tabagisme est de près de 25% entre 20 et 24 ans, puis atteint un pic à plus de 30% entre 30 et 34 ans pour redescendre ensuite avec l’âge. La proportion d’anciens fumeurs, elle, croît jusqu’à 80 ans pour diminuer par la suite, ce qui reflète probablement des effets de cohorte dans les habitudes de tabagisme. Les taux d’initiation sont plus élevés chez les jeunes, tandis que les taux de cessation sont plus grands aux deux extrémités de la distribution d’âges — une dynamique qui masque aussi sans doute certains effets de cohorte.

Âge	Fumeurs actuels	Anciens fumeurs
20 à 24 ans	24,26%	14,04%
25 à 29 ans	28,48%	19,95%
30 à 34 ans	30,63%	21,35%
35 à 39 ans	28,95%	25,95%
40 à 44 ans	28,54%	28,87%
45 à 49 ans	26,89%	32,11%
50 à 54 ans	25,04%	35,29%
55 à 59 ans	22,42%	38,66%
60 à 64 ans	19,37%	41,23%
65 à 69 ans	16,26%	44,69%
70 à 74 ans	13,73%	45,90%
75 à 79 ans	11,03%	47,30%
80 à 84 ans	7,10%	45,83%
85 à 89 ans	4,60%	40,45%
90 à 94 ans	3,02%	36,14%
95 à 99 ans	.	40,61%
Total	.	31,35%
Test t de la différence entre les taux pour le Québec et le Canada		
moins de 40 ans	-5,4252	-1,0811
40 à 64 ans	-7,7491	-4,5301
65 ans et plus	-8,8095	-6,2277

Tableau 2.11 – Tabagisme dans l’ENSP (1994-2011)

Il y a des différences notables (et statistiquement significatives) en ce qui concerne la proportion de fumeurs et d’anciens fumeurs au Québec et au Canada. Quel que soit l’âge considéré, la proportion de fumeurs et d’anciens fumeurs est plus élevée au Québec qu’au Canada.

En ce qui concerne l’obésité, nous avons construit une variable basée sur l’indice de masse corporelle (IMC) avec trois catégories (moins de 30, 30-35 et 35+). Dans le vocabulaire de l’Organisation mondiale de la Santé, un individu ayant un IMC de 30 ou plus est considéré obèse. Dans notre modèle, l’obésité de classe I regroupe les individus ayant un IMC entre 30 et 35 (classe I de l’OMS), et celle de classes II-III regroupe ceux ayant un IMC de 35 et plus (classes II et III de l’OMS) ([Organisation mondiale de la Santé, 2006](#)). Le tableau 2.13 donne la répartition des répondants par groupe d’âge. On voit que l’obésité de classe I atteint un sommet dans la soixantaine, alors que l’obésité de classes II-III est à son

Âge	Initiation	Cessation	Recommencement
20 à 24 ans	4,67%	20,37%	18,94%
25 à 29 ans	3,09%	15,63%	20,69%
30 à 34 ans	2,72%	13,19%	10,96%
35 à 39 ans	2,54%	12,15%	9,08%
40 à 44 ans	1,95%	12,40%	8,35%
45 à 49 ans	2,23%	10,90%	6,33%
50 à 54 ans	2,40%	12,92%	4,79%
55 à 59 ans	2,18%	14,10%	3,84%
60 à 64 ans	1,65%	14,58%	3,37%
65 à 69 ans	2,56%	18,13%	3,54%
70 à 74 ans	2,25%	17,94%	2,94%
75 à 79 ans	2,19%	20,85%	1,56%
80 à 84 ans	1,05%	16,66%	1,65%
85 à 89 ans	1,10%	24,81%	1,56%
Total	2,60%	14,26%	6,71%
Test t de la différence entre les taux pour le Québec et le Canada			
< 40	0,6266	1,1721	1,5307
40-64	-1,402	2,109	1,9207
>65	-0,6881	0,4028	-1,3191

Tableau 2.12 – Taux d’initiation au tabagisme et de cessation selon l’ENSP (1994-2011)

niveau le plus haut dans la cinquantaine.

Le tableau 2.14 présente les transitions individuelles entre les niveaux d’obésité. Les lignes correspondent au niveau d’obésité au cycle courant, alors que les colonnes correspondent au niveau au cycle suivant. En moyenne, un peu plus de 3% des individus non-obèses deviennent obèses à chaque cycle de 2 ans ; 17,7% des individus atteints d’obésité de classe I deviennent non-obèses à chaque cycle, tandis que 8,1% passent à un niveau d’obésité de classes II-III ; et 17,2% des individus atteints d’obésité de classes II-III passent à l’obésité de classe I tandis que 2,9% deviennent non-obèses.

### 2.3.4 Utilisation des soins de santé

L’ENSP est riche en données d’utilisation de soins de santé (Statistique Canada, 2012b). Aux fins de modélisation, cinq variables ont été créées. Les deux premières variables correspondent au nombre de visites ou de consultations par téléphone au cours des 12 derniers mois auprès, d’une part, d’un médecin de famille, d’un pédiatre ou d’un omnipraticien, et, d’autre part, auprès d’un autre médecin ou spécialiste (comme un chirurgien, un allergologue, un orthopédiste, un gynécologue ou un psychiatre). Ces deux variables ont été censurées au 99<sup>e</sup> centile, car les dernières observations présentaient un nombre de consultations élevé susceptible de fausser les résultats à la moyenne.

Les consultations auprès de médecins ne comprenant pas les séjours dans un établissement de soins, une troisième variable a été créée pour en tenir compte, soit le nombre de nuits passées à l’hôpital, dans un foyer de soins infirmiers ou dans une maison de convalescence. Tout comme pour les variables de consultations médicales et pour la même raison, cette variable a été censurée au 99<sup>e</sup> centile.

La quatrième variable est un indicateur de l’utilisation de médicaments, consistant en une variable



Âge	IMC<30	$30 \leq \text{IMC} < 35$	IMC $\geq 35$
20 à 24 ans	91,39%	5,65%	1,98%
25 à 29 ans	87,24%	8,97%	2,53%
30 à 34 ans	85,36%	10,29%	3,10%
35 à 39 ans	83,46%	11,41%	3,87%
40 à 44 ans	81,93%	12,46%	4,35%
45 à 49 ans	81,08%	13,13%	4,64%
50 à 54 ans	78,90%	14,41%	5,52%
55 à 59 ans	78,88%	14,68%	5,25%
60 à 64 ans	78,05%	15,95%	5,10%
65 à 69 ans	78,98%	15,67%	4,35%
70 à 74 ans	81,69%	14,25%	3,33%
75 à 79 ans	84,20%	12,75%	2,38%
80 à 84 ans	89,16%	9,08%	1,31%
85 à 89 ans	92,35%	6,34%	1,20%
90 à 94 ans	92,48%	6,13%	1,17%
Total	83,16%	11,93%	3,82%
Test t de la différence entre les taux pour le Québec et le Canada			
moins de 40 ans	-9,8752	7,4653	6,1378
40 à 64 ans	-12,5711	8,7684	8,3062
65 ans et plus	-0,2379	-0,2379	0,9002

Tableau 2.13 – Indice de masse corporelle (IMC) selon l’ENSP (1994-2011)

Obésité au cycle courant	Obésité au cycle suivant		
	IMC<30	$30 \leq \text{IMC} < 35$	IMC $\geq 35$
IMC<30	96,47%	3,18%	0,16%
$30 \leq \text{IMC} < 35$	17,65%	74,15%	8,09%
IMC $\geq 35$	2,90%	17,21%	79,76%
Total	85,37%	10,90%	3,56%
Test t de la différence entre les taux pour le Québec et le Canada			
IMC<30	-5,1179	5,1754	1,7965
IMC 30-35	-0,9912	-1,6353	0,804
IMC>35	-0,4775	-0,3086	1,7589

Tableau 2.14 – Transitions entre les états d’obésité selon l’ENSP (1994-2011)

binaire indiquant la consommation d’au moins un médicament au cours des 12 derniers mois. Nous utilisons cette variable à défaut d’avoir accès au nombre de médicaments consommés.

La dernière variable concerne les services à domicile. Puisque l’objectif principal est ici d’évaluer la consommation de ressources publiques, la variable que nous créons est un indicateur d’utilisation de services à domicile au cours des 12 derniers mois, mais dont le coût est au moins en partie supporté par le gouvernement. L’ENSP donne des exemples de services à domicile, ces derniers pouvant être aussi divers que des soins infirmiers, de l’aide pour prendre un bain ou la livraison de repas.

Le tableau 2.15 donne la répartition par âge. Le nombre de consultations auprès d’un médecin généraliste augmente de manière relativement constante avec l’âge, surtout à partir de 45 ans (on peut penser qu’avant cet âge, la grossesse hausse le nombre de visites). Les individus consultent, en moyenne, trois fois un médecin généraliste au cours d’une période de référence de 12 mois. Lorsqu’il s’agit de consulter un spécialiste, le nombre de visites moyen chute à moins de une au cours d’une période de 12 mois.

Le nombre de nuits d'hospitalisation augmente de manière plus importante à partir de 75 ans. Les individus âgés de 85 à 89 ans passent, en moyenne, plus de 3 nuits dans un établissement de courte durée au cours d'une période de 12 mois. Entre 80 % et 93 % de l'ensemble de la population a répondu avoir consommé au moins un médicament pendant la période de référence. Enfin, à partir de 85 ans, c'est plus de 25 % de la population qui reçoit des services à domicile.

Âge	Nb. consultations		Nb. nuits hosp.
	généraliste	spécialiste	
20-24	2,64	0,55	0,36
25-30	2,75	0,69	0,38
30-34	2,77	0,72	0,46
35-39	2,68	0,68	0,40
40-44	2,54	0,65	0,36
45-49	2,64	0,66	0,41
50-54	2,98	0,76	0,64
55-59	3,15	0,77	0,68
60-64	3,32	0,80	0,73
65-69	3,76	0,85	1,10
70-74	4,19	0,88	1,43
75-79	4,46	0,86	2,11
80-84	5,06	0,84	2,42
85-89	5,05	0,69	3,57
90-94	5,29	0,52	3,53
95-99	4,35	0,42	3,98
Total	3,07	0,72	0,71

Âge	1+ médicament	Services à domicile (oui)
20-24	80,91%	0,50%
25-30	80,58%	1,19%
30-34	79,77%	1,19%
35-39	79,96%	1,30%
40-44	81,60%	0,91%
45-49	81,70%	1,16%
50-54	82,93%	1,40%
55-59	85,10%	1,72%
60-64	87,27%	2,60%
65-69	89,20%	4,01%
70-74	90,37%	7,04%
75-79	91,32%	11,35%
80-84	93,34%	17,91%
85-89	90,88%	25,45%
90-94	90,95%	38,41%
95-99	92,17%	33,73%
Total	83,51%	2,93%

Tableau 2.15 – Utilisation des soins de santé par âge selon l'ENSP (1994-2011)

Le tableau 2.16 indique que le nombre de consultations moyen auprès d'un généraliste est environ deux fois plus élevé pour les individus qui souffrent d'un des problèmes de santé considérés. La différence est statistiquement significative dans chacun des cas. Les individus qui souffrent du diabète, d'un cancer, de maladies cardiaques, d'un AVC ou de démences consultent, en moyenne, un spécialiste plus d'une fois par année. Cette différence est statistiquement significative. Ceux qui souffrent d'une certaine forme de démence, d'une maladie cardiaque ou qui ont eu un AVC sont hospitalisés plus de cinq fois

plus de nuits que ceux qui n'en souffrent pas. Cette différence est également significative. C'est parmi ceux aux prises avec le diabète, l'hypertension, une maladie cardiaque ou un AVC que la prévalence de gens ayant consommé au moins un médicament est la plus élevée, soit plus de 96%. Il peut sembler que la consommation d'au moins un médicament augmente peu avec l'âge. Toutefois, en vertu de la définition de cette variable, un répondant qui est globalement en santé mais qui consomme un seul médicament d'ordonnance ne peut être distingué d'un répondant qui souffre de plusieurs maladies et consomme plusieurs médicaments. Ainsi, 80% des individus âgés de 20 à 24 ans ont consommé au moins un médicament durant la période de référence. L'augmentation de plus 10 points de pourcentage entre 20 et 99 ans de la proportion d'individus qui consomment au moins un médicament est, une fois la définition de la variable prise en considération, somme toute à assez importante. Environ 21% des gens qui ont eu un AVC ou qui souffrent d'une certaine forme de démence reçoivent des services à domicile, soit environ 10 fois plus que ceux qui n'en ont jamais eu — encore là une différence significative.

### 2.3.5 Autres variables socio-économiques

Nous avons construit des variables socio-économiques qui seront utiles dans l'analyse.

- Immigrant : variable indicatrice égale à 1 si l'individu est un immigrant et 0 sinon.
- Femme : variable indicatrice égale à 1 si l'individu est une femme et 0 sinon.
- Éducation : trois variables indicatrices du plus haut niveau atteint parmi un diplôme d'études secondaires, un diplôme d'études collégiales et un diplôme universitaire.

Nous avons choisi de ne modéliser que ces variables puisqu'il est raisonnable de les supposer constantes à partir de 30 ans. Si nous avons inclus d'autres variables, telles que l'état matrimonial ou le revenu, il aurait fallu modéliser les comportements individuels dans le temps. À ce stade de la conceptualisation du modèle, une telle modélisation aurait été trop complexe.

Le tableau 2.17 présente des statistiques descriptives tirées de l'ENSP en lien avec ces variables. On note que la proportion d'immigrants augmente légèrement avec l'âge, tout comme la proportion de femmes — surtout à partir de 60 ans dans ce dernier cas. En raison d'importants effets de cohorte, la proportion de diplômés universitaires diminue avec l'âge, mais à partir de 30 ans, soit un âge où les individus sont susceptibles d'avoir complété la majeure partie de leurs études.

## 2.4 ESCC

Nous utilisons également les données « publiques » de l'ESCC, c'est-à-dire les données à grande diffusion ([Statistique Canada, 2010](#)). Cette enquête est utilisée pour créer la population de départ en 2010, âgée de 30 à 110 ans. Nous utilisons seulement les observations québécoises de l'ESCC, puisqu'elles sont suffisamment nombreuses (environ 11 000) pour créer la population initiale. Nous ne pouvons pas utiliser l'ENSP parce que sa population était représentative de la population canadienne de 1994 seulement.

L'ESCC 2010 contient plusieurs variables communes à celles de l'ENSP, que nous pouvons utiliser directement. Nous avons ainsi accès à toutes les variables concernant la présence des maladies considérées dans le modèle, à l'exception de l'Alzheimer et des autres démences. Le tableau 2.18 présente la prévalence des maladies au sein de l'ESCC 2010. Similairement à ce qui était observé dans l'ENSP,

Maladies	Nb. consultations généraliste	Nb. consultations spécialiste	Nb. nuits hosp.
<b>Diabète</b>			
Non	2,78	0,63	0,57
Oui	5,40	1,06	2,48
<b>Hypertension</b>			
Non	2,59	0,60	0,47
Oui	4,66	0,98	1,74
<b>Cancer</b>			
Non	2,84	0,62	0,59
Oui	4,77	1,81	2,59
<b>Maladies cardiaques</b>			
Non	2,73	0,61	0,50
Oui	5,22	1,27	2,88
<b>AVC</b>			
Non	2,85	0,65	0,59
Oui	6,04	1,12	4,75
<b>Maladies pulmonaires</b>			
Non	2,80	0,63	0,60
Oui	4,46	1,00	1,42
<b>Démences</b>			
Non	2,88	0,65	0,63
Oui	5,61	1,13	4,90
Maladies	1+ médicament	Services à domicile (oui)	
<b>Diabète</b>			
Non	82,40%	2,04%	
Oui	97,86%	9,95%	
<b>Hypertension</b>			
Non	80,53%	1,61%	
Oui	96,62%	6,86%	
<b>Cancer</b>			
Non	82,79%	2,14%	
Oui	93,64%	10,56%	
<b>Maladies cardiaques</b>			
Non	82,10%	1,76%	
Oui	96,87%	11,36%	
<b>AVC</b>			
Non	82,92%	2,12%	
Oui	97,91%	21,01%	
<b>Maladies pulmonaires</b>			
Non	82,56%	2,11%	
Oui	92,14%	6,77%	
<b>Démences</b>			
Non	83,10%	2,30%	
Oui	93,03%	21,63%	

Tableau 2.16 – Utilisation des soins de santé selon la présence de maladies, ENSP (1994-2011) : Toutes ces différences sont statistiquement significatives au seuil de 5%.

la prévalence de la plupart des maladies augmente avec l'âge. La prévalence de maladies pulmonaires augmente jusqu'à 75 ans puis amorce une légère diminution. La maladie la plus fréquente est l'hypertension, avec un taux de prévalence de 48,6% chez les personnes âgées de 65 à 69 ans.

Âge	Immigrants	Femmes	Dip. secondaire	Dip. collégial	Dip. universitaire
20-24	9,84%	48,36%	54,23%	17,01%	8,63%
25-29	12,65%	49,33%	42,39%	22,97%	21,48%
30-34	16,93%	50,29%	43,39%	24,75%	20,96%
35-39	19,35%	50,52%	44,63%	24,21%	19,49%
40-44	20,99%	50,37%	43,44%	23,92%	19,51%
45-49	22,39%	49,33%	42,06%	23,05%	20,31%
50-54	23,76%	48,59%	40,64%	20,31%	19,97%
55-59	25,11%	49,22%	38,86%	17,59%	18,89%
60-64	24,85%	51,06%	36,57%	15,81%	15,69%
65-69	25,15%	53,92%	34,05%	14,24%	11,95%
70-74	28,16%	56,13%	31,59%	13,43%	10,11%
75-79	27,62%	59,99%	30,71%	10,93%	7,90%
80-84	28,30%	62,58%	29,51%	8,88%	8,28%
85-89	25,64%	66,32%	29,40%	8,23%	8,46%
90-94	28,09%	69,60%	31,11%	8,68%	7,52%
95-99	30,27%	65,47%	32,38%	3,43%	6,77%
Total	20,72%	51,17%	41,20%	19,83%	16,98%

Tableau 2.17 – Caractéristiques socio-économiques tirées de l’ENSP (1994-2011)

Âge	Diabète	Hypertension	Cancer	Maladies cardiaques	AVC	Maladies pulmonaires
20 à 24 ans	1,01%	3,02%	0,41%	0,78%	0,00%	0,00%
25 à 29 ans	0,64%	3,59%	1,22%	0,57%	0,12%	0,00%
30 à 34 ans	0,63%	5,24%	1,57%	1,51%	0,02%	0,00%
35 à 39 ans	1,52%	7,03%	1,89%	1,11%	0,39%	1,53%
40 à 44 ans	3,16%	12,63%	2,69%	0,93%	0,15%	1,00%
45 à 49 ans	2,94%	12,11%	4,46%	1,90%	0,27%	3,71%
50 à 54 ans	4,70%	23,39%	6,06%	3,29%	1,02%	2,79%
55 à 59 ans	8,76%	36,54%	7,53%	9,66%	0,68%	4,52%
60 à 64 ans	9,74%	38,95%	12,64%	6,69%	1,47%	5,41%
65 à 69 ans	16,05%	48,61%	15,83%	18,12%	2,75%	6,33%
70 à 74 ans	16,01%	49,51%	18,97%	20,34%	2,66%	9,79%
75 à 79 ans	17,33%	63,16%	20,59%	23,61%	5,20%	8,74%
80 ans et plus	17,26%	56,81%	22,09%	25,11%	5,98%	7,30%
Total	6,10%	23,02%	7,00%	6,34%	1,10%	3,21%

Tableau 2.18 – Prévalence auto-déclarée des maladies dans l’ESCC 2010

Les facteurs de risque (tabagisme et indice de masse corporelle) sont aussi présents dans l’ESCC. Tel que l’indique le tableau 2.19, la proportion de fumeurs est de 32,3% parmi les personnes âgées de 45 à 49 ans. C’est au sein de ce groupe d’âge que la proportion de fumeurs est la plus élevée.

Le tableau 2.20 présente la prévalence de l’obésité dans l’ESCC. La prévalence de l’obésité de classes II-III augmente, par groupe d’âge, entre 35 et 59 ans, puis diminue graduellement. C’est un schéma semblable à celui observé dans l’ENSP, quoique dans ce dernier cas la prévalence de l’obésité de classes II-III augmente, par groupe d’âge, de 20 à 54 ans. La prévalence de l’obésité de classe I est particulièrement élevée entre 70 et 74 ans, à 15,5%.

Nous pouvons également calculer le nombre de limitations d’activités quotidiennes, bien que celles-ci ne soient pas parfaitement cohérentes avec celles de l’ENSP, ce qui fait que nous ne pouvons pas avoir

Âge	Fumeurs actuels	Anciens fumeurs
20 à 24 ans	24,85%	24,24%
25 à 29 ans	30,67%	31,34%
30 à 34 ans	27,42%	32,58%
35 à 39 ans	30,77%	30,10%
40 à 44 ans	26,04%	39,32%
45 à 49 ans	32,28%	39,81%
50 à 54 ans	24,49%	50,52%
55 à 59 ans	28,52%	48,62%
60 à 64 ans	19,02%	59,06%
65 à 69 ans	15,36%	59,37%
70 à 74 ans	14,92%	50,69%
75 à 79 ans	8,47%	58,09%
80 ans et plus	5,96%	51,15%
Total	24,39%	42,76%

Tableau 2.19 – Tabagisme dans l’ESCC 2010

une mesure d’invalidité directement comparable à celle de l’ENSP. De plus, on n’observe pas le statut d’institutionnalisation dans l’ESCC. Ces statistiques ne sont pas présentées ici.

Âge	IMC < 30	30 ≤ IMC < 35	IMC ≥ 35
20 à 24 ans	89,80%	6,88%	3,31%
25 à 29 ans	86,17%	10,45%	3,38%
30 à 34 ans	84,15%	12,99%	2,86%
35 à 39 ans	84,20%	13,02%	2,78%
40 à 44 ans	83,03%	13,51%	3,46%
45 à 49 ans	85,64%	8,76%	5,61%
50 à 54 ans	80,60%	12,97%	6,43%
55 à 59 ans	78,64%	13,25%	8,11%
60 à 64 ans	81,47%	12,40%	6,12%
65 à 69 ans	82,59%	12,97%	4,44%
70 à 74 ans	81,36%	15,53%	3,11%
75 à 79 ans	82,66%	12,93%	4,40%
80 ans et plus	88,50%	10,08%	1,42%
Total	83,56%	11,87%	4,57%

Tableau 2.20 – Indice de masse corporelle (IMC) selon l’ESCC 2010

On connaît finalement le niveau d’éducation des individus mais, contrairement à l’ENSP, aucune distinction n’est faite entre les diplômes collégial et universitaire. Le tableau 2.21 présente certaines caractéristiques socio-économiques des individus au sein de l’ESCC. On remarque que la proportion d’individus ayant un diplôme collégial ou universitaire augmente, par groupe d’âge, entre 20 et 34 ans puis diminue. Ceci est semblable aux statistiques concernant l’éducation au sein de l’ENSP.

Le chapitre 3 explique plus en détails la façon dont nous résolvons les problèmes liés aux variables de l’ESCC et la façon dont nous utilisons cette enquête comme point de départ de la simulation.

Âge	Immigrants	Femmes	Dip. secondaire	Dip. collégial / universitaire
20 à 24 ans	13,38%	44,87%	25,35%	65,06%
25 à 29 ans	15,29%	52,58%	14,35%	78,12%
30 à 34 ans	22,07%	51,75%	12,45%	83,26%
35 à 39 ans	20,77%	46,80%	9,99%	82,15%
40 à 44 ans	18,71%	49,56%	13,19%	75,19%
45 à 49 ans	12,59%	49,33%	18,27%	74,59%
50 à 54 ans	10,57%	49,12%	27,33%	60,96%
55 à 59 ans	14,92%	51,36%	20,02%	57,67%
60 à 64 ans	10,62%	52,94%	19,94%	58,80%
65 à 69 ans	12,34%	47,44%	18,56%	50,18%
70 à 74 ans	17,72%	56,58%	19,70%	36,36%
75 à 79 ans	10,47%	56,96%	18,64%	30,45%
80 ans et plus	12,39%	65,94%	16,80%	25,87%
Total	14,91%	50,83%	18,18%	64,41%

Tableau 2.21 – Caractéristiques socio-économiques tirées de l’ESCC 2010

## 3 Initialisation

### 3.1 Survol

L’initialisation consiste en la création d’une base de données pour l’année initiale de la simulation (par défaut 2010). Nous ne pouvons utiliser l’ENSP comme base de données initiale car cette enquête était conçue pour être représentative de la population canadienne au moment de l’échantillonnage (1994). Elle ne dresse donc plus un portrait représentatif de l’état de santé de la population en 2010. Qui plus est, l’échantillon de répondants québécois est trop petit pour nous permettre de créer la population initiale de façon fiable.

Nous utilisons donc plutôt l’ESCC de 2010 comme base pour construire la population initiale. Cette enquête a l’avantage de dresser un portrait à jour de l’état de santé de la population canadienne. De plus, elle contient un grand nombre d’observations (11 271 pour le Québec seulement), ce qui nous permet d’utiliser seulement les observations québécoises de l’enquête. Le désavantage de l’ESCC par rapport à l’ENSP est que, contrairement à cette dernière, la première n’a pas un caractère longitudinal. Cette caractéristique n’est toutefois pas nécessaire pour la création d’une base de données initiale.

Utiliser cette deuxième enquête entraîne cependant une difficulté : la base initiale doit contenir toutes les variables nécessaires aux modèles de transition et d’utilisation des soins de santé. Pour la majorité des variables que nous utilisons, ceci n’est pas un problème puisque les variables de l’ENSP et de l’ESCC sont identiques. Toutefois, dans certains cas, il est nécessaire d’imputer des valeurs à certaines variables auxquelles nous n’avons pas directement accès dans l’ESCC. De plus, nous utilisons la version de l’ESCC à grande diffusion, ce qui fait que l’information est limitée pour certaines variables. Ce chapitre décrit les hypothèses permettant la création de la base de données — et donc de la population — initiale à partir de l’ESCC 2010.

## 3.2 Imputation

### 3.2.1 Imputation de l'âge

L'ESCC à grande diffusion catégorise l'âge des individus par intervalle de 3 à 4 ans.<sup>5</sup>

Pour pallier ce problème, nous attribuons à chaque individu un âge précis à l'intérieur de sa classe d'âge, de façon aléatoire et uniforme puis nous calibrons les poids de l'enquête de façon à reproduire la distribution d'âge de la population du Québec par sexe provenant de l'ISQ (voir [Institut de la statistique du Québec 2009](#)).

### 3.2.2 Imputation du niveau d'éducation

Les niveaux d'éducation dont on dispose dans l'ESCC à grande diffusion sont « moins qu'un diplôme d'études secondaires », « diplôme d'études secondaires » et « études post-secondaires ». Pour avoir des niveaux semblables à ceux de l'ENSP, on doit donc imputer des diplômes universitaires parmi les individus se trouvant dans la dernière catégorie. Bien qu'il eût été possible d'utiliser l'ENSP afin de faire l'imputation du niveau d'éducation, le faible nombre de répondants québécois au sein de l'enquête nous en a dissuadés.

Nous utilisons plutôt les données à grande diffusion de l'Enquête sur la population active (EPA) de 2010 de Statistique Canada, qui fournit un portrait détaillé du niveau d'éducation de la population en 2010. Nous gardons seulement les observations québécoises qui ont au moins un diplôme d'études collégiales puis nous estimons un modèle logit qui permet d'obtenir les effets des caractéristiques individuelles sur la probabilité de détenir un diplôme universitaire, conditionnellement à avoir au moins un diplôme collégial. La structure du modèle est la suivante :

$$y_i^* = \beta \mathbf{x}_i + \epsilon_i \quad (1)$$

$$y_i = \begin{cases} 1 & \text{si } y_i^* \geq 0 \\ 0 & \text{si } y_i^* < 0 \end{cases}, \quad (2)$$

où  $y_i$  est égale à 1 si l'individu  $i$  a un diplôme universitaire, 0 sinon ;  $y_i^*$  est la variable latente ;  $\mathbf{x}_i$  est le vecteur de variables explicatives de l'individu  $i$  ; et  $\epsilon_i$  est un terme aléatoire qui suit une loi logistique.

Le tableau 3.1 présente les variables utilisées et les résultats de l'estimation. Puisque le modèle utilisé n'est pas linéaire, ce ne sont pas les paramètres eux-mêmes qui sont intéressants pour analyser l'impact des variables explicatives, mais bien les effets marginaux, soit le changement de l'espérance conditionnelle de  $y_i$  quand la valeur d'une variable  $x_i$  change d'une unité. Dans le cadre de COMPAS, nous utilisons la moyenne des effets marginaux (MEM), aussi appelée effet marginal moyen. La MEM est une moyenne sur l'ensemble des individus de l'effet du changement d'une unité de la variable  $x_i$  sur

---

5. Nous utilisons les données à grande diffusion de l'ESCC pour la construction de la base initiale afin de pouvoir réaliser l'initialisation de façon répétée à l'extérieur des murs d'un Centre de données de recherche (CDR) de Statistique Canada. L'utilisation des données à accès restreint obligerait l'utilisateur à faire fonctionner le modèle dans un CDR.



l'espérance de  $y_i$ .<sup>6</sup> Nous incluons un noeud à chaque groupe d'âge pour capter la non linéarité de la variable d'âge, un effet constant de sexe féminin, ainsi que des effets d'interactions entre le sexe et ces effets d'âge. Ces effets d'interaction permettent que l'effet d'être une femme sur la probabilité d'avoir un diplôme universitaire s'ajuste selon l'âge. Cela est important vu les changements significatifs observés entre les générations concernant l'écart de niveau d'éducation entre les hommes et les femmes.

	Effet marginal
Âge (si 25 ans et moins)	0,0745 ***
Âge (si 26 à 30 ans)	0,0113 ***
Âge (si 31 à 35 ans)	-0,0090 *
Âge (si 36 à 40 ans)	0,0077
Âge (si 41 à 45 ans)	-0,0181 ***
Âge (si 46 à 50 ans)	-0,0012
Âge (si 51 à 55 ans)	0,0148 ***
Âge (si 56 à 60 ans)	-0,0012
Âge (si 61 à 65 ans)	0,0026
Âge (si 66 ans et plus)	-0,0090 *
Femme	0,3304 **
Femme*Âge (si 25 ans et moins)]	-0,0089
Femme*Âge (si 26 à 30 ans)]	-0,0085 **
Femme*Âge (si 31 à 35 ans)]	0,0145 **
Femme*Âge (si 36 à 40 ans)]	-0,0180 ***
Femme*Âge (si 41 à 45 ans)]	0,0008
Femme*Âge (si 46 à 50 ans)]	0,0203 ***
Femme*Âge (si 51 à 55 ans)]	-0,0246 ***
Femme*Âge (si 56 à 60 ans)]	0,0197 ***
Femme*Âge (si 61 à 65 ans)]	-0,0075
Femme*Âge (si 66 ans et plus)]	-0,0040
Nombre d'observations	110 626

Tableau 3.1 – Effets marginaux moyens des variables sur la probabilité d'avoir un diplôme universitaire pour les individus ayant complété des études post-secondaires — EPA 2010 à grande diffusion

La figure 3.1 présente plus clairement les probabilités estimées selon l'âge et le sexe. Les effets d'âge pour les plus jeunes sont positifs et significatifs; cela capte simplement le fait que les étudiants plus âgés sont plus avancés dans leurs études. L'effet d'âge est ensuite négatif, reflétant l'augmentation du niveau d'éducation chez les cohortes plus jeunes. Les effets d'âge sont finalement ambigus pour les âges plus avancés, car les personnes les plus éduquées ont tendance à vivre plus longtemps. L'effet d'être une femme est positif et significatif pour les plus jeunes, car les femmes plus jeunes ont tendance à aller davantage à l'université que les hommes. L'effet d'être une femme diminue cependant avec l'âge et devient négatif après 50 ans.

Les coefficients du modèle sont ensuite utilisés pour imputer les niveaux d'éducation dans l'ESCC. Pour

6. Les modèles utilisés dans ce chapitre ainsi que dans le chapitre 4 sont tous de forme non linéaire. Ainsi, nous présentons toujours la MEM plutôt que les paramètres issus de l'estimation.

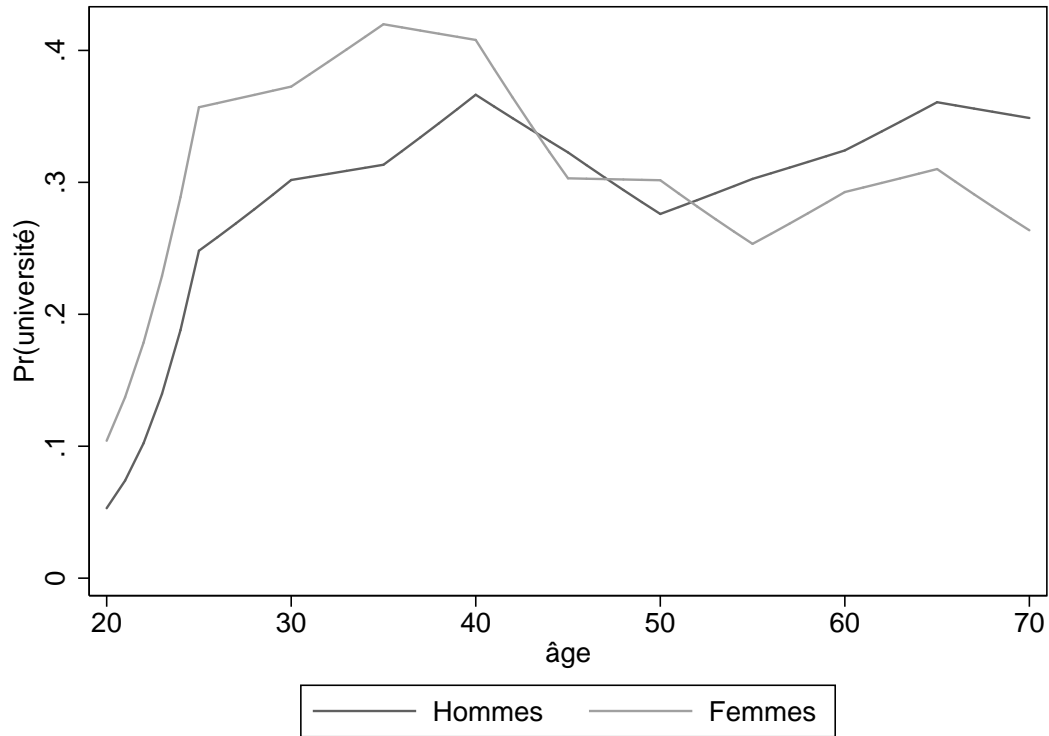


FIGURE 3.1 – Probabilité d’avoir un diplôme universitaire pour les individus ayant complété des études post-secondaires — EPA 2010 à grande diffusion

chaque observation  $i$  ayant des études post-secondaires, on estime la probabilité d’avoir un diplôme universitaire à l’aide de la formule suivante :

$$\hat{P}(univ_i = 1 | \mathbf{x}_i) = \frac{\exp(\mathbf{x}_i' \hat{\boldsymbol{\beta}})}{1 + \exp(\mathbf{x}_i' \hat{\boldsymbol{\beta}})} \quad (3)$$

On tire ensuite une variable aléatoire d’une loi uniforme (0,1) pour chaque observation et on impute un diplôme universitaire si celle-ci est inférieure à la probabilité calculée plus haut.

### 3.2.3 Imputation de l'invalidité

L'ESCC nous permet de compter le nombre de limitations d'activités quotidiennes parmi les suivantes :

- besoin d'aide pour préparer les repas ;
- besoin d'aide pour faire les courses ;
- besoin d'aide pour les soins personnels ;
- besoin d'aide pour se déplacer dans la maison.

Par rapport à celles que nous utilisons dans l'ENSP pour calculer les transitions (voir chapitre suivant), la seule limitation manquante est le besoin d'aide pour accomplir les tâches ménagères. Nous ignorons ce problème pour l'instant et calculons le nombre de limitations autres pour obtenir les mêmes variables d'invalidité que dans l'ENSP (une seule limitation ou problème cognitif ; deux limitations ou plus ; en institution).

Nous n'observons cependant pas si la personne est en institution dans l'ESCC. Nous devons donc imputer le statut d'institutionnalisation. L'imputation se fait encore une fois à l'aide d'un modèle logit, cette fois-ci estimé sur les données 2010-2011 de l'ENSP. Nous incluons un noeud à 75 ans afin de permettre à l'effet de l'âge de changer à partir de 75 ans ; un effet de sexe ; un effet de statut d'immigrant ; et deux variables binaires pour les deux premières formes d'invalidité (une seule limitation ou problème cognitif ; deux limitations ou plus). Le tableau 3.2 présente les résultats.

	Effet marginal
Âge (si moins de 75 ans)	0,0003 **
Âge (si 75 ans ou plus)	0,0008 ***
Immigrant	-0,0033 *
Femme	-0,0014
Invalidité : 1 ADL ou problème cognitif	0,0097 ***
Invalidité : 2 ADL ou +	0,0150 ***
Nombre d'observations	14 598

Tableau 3.2 – Effets marginaux moyens des variables sur la probabilité d'être en institution — ENSP 2010-2011

On voit que les effets d'âge sont significatifs. L'effet pour les femmes n'est pas significatif, mais celui pour les immigrants est légèrement négatif. Les variables indicatrices pour les autres états d'invalidité ont des coefficients élevés, car les individus qui souffrent de limitations physiques ou cognitives ont davantage tendance à être en institution.

Nous utilisons ensuite les coefficients estimés pour calculer la probabilité d'être en institution pour chaque individu observé dans l'ESCC et pour procéder à l'imputation.

### 3.2.4 Imputation de l'Alzheimer et des autres démences

La prévalence de l'Alzheimer et des autres démences est le seul problème de santé chronique que nous n'observons pas dans l'ESCC à grande diffusion. Nous imputons donc la variable à l'aide d'un logit

estimé à l'aide de l'ENSP de 2010-2011. Le modèle est semblable à celui utilisé pour l'institutionnalisation, sauf qu'il comprend en plus une variable binaire pour le statut d'institutionnalisation comme variable explicative. Les résultats de l'estimation sont présentés dans le tableau 3.3.

On trouve que la probabilité de souffrir de démences augmente fortement avec l'âge et cesse d'augmenter significativement après 75 ans. Le sexe et le statut d'immigrant n'ont pas d'effet significatif. Comme on pourrait s'y attendre, les statuts d'invalidité sont pour leur part fortement associés à la présence de démences.

	Effet marginal
Âge (si moins de 75 ans)	0,0007 ***
Âge (si 75 ans ou plus)	0,0002
Immigrant	0,0010
Femme	-0,0038 *
Invalidité : 1 ADL ou trouble cognitif	0,0122 ***
Invalidité : 2+ ADL	0,0250 ***
Invalidité : en institution	0,0225 ***
Nombre d'observations	14 598

Tableau 3.3 – Effets marginaux moyens des variables sur la probabilité d'être atteint d'Alzheimer ou d'une autre démence — ENSP 2010

Les coefficients du modèle sont ensuite utilisés afin d'imputer la présence de démences dans l'ESCC. Ainsi, pour chaque observation  $i$ , on estime la probabilité de souffrir de démences à l'aide de la formule suivante :

$$\hat{P}(\text{démences}_i = 1 | \mathbf{x}_i) = \frac{\exp(\mathbf{x}'_i \hat{\boldsymbol{\beta}})}{1 + \exp(\mathbf{x}'_i \hat{\boldsymbol{\beta}})} \quad (4)$$

où  $\mathbf{x}_i$  est le vecteur de variables explicatives de l'individu  $i$  et  $\boldsymbol{\beta}$  est le vecteur de coefficients. Afin de déterminer si une observation  $i$  se voit imputer une démence, on tire une variable aléatoire d'une loi uniforme (0,1) pour chaque observation. Si cette dernière est inférieure à la probabilité calculée, l'observation  $i$  se voit imputer une démence.

### 3.2.5 Caractéristiques de la population initiale

La population initiale du modèle, suite aux imputations effectuées sur les données de l'ESCC 2010, contient, à une exception près, toutes les variables nécessaires aux modèles de transition et d'utilisation de soins de santé. Le tableau 3.4 présente certaines statistiques concernant les variables de cette population.

	Moyenne	Écart-type	Minimum	Maximum
Âge	53,80	15,14	30	110
Immigrants	14,94%	35,65%	0	1
Femmes	51,30%	49,98%	0	1
Aucun diplôme	19,38%	39,53%	0	1
Diplôme d'études secondaires	17,76%	38,22%	0	1
Diplôme d'études collégiales	42,43%	49,42%	0	1
Diplôme universitaire	20,44%	40,32%	0	1
Présence de diabète	7,03%	25,57%	0	1
Présence d'hypertension	27,23%	44,51%	0	1
Présence de maladies cardiaques	7,54%	26,41%	0	1
Présence d'AVC	1,04%	10,15%	0	1
Présence de cancer	8,26%	27,53%	0	1
Présence de maladies pulmonaires	3,89%	19,33%	0	1
Présence de démences	1,32%	11,40%	0	1
Jamais fumé	30,23%	45,93%	0	1
Fumeurs actuels	23,59%	42,46%	0	1
Anciens fumeurs	46,18%	42,46%	0	1
Poids santé	82,76%	37,78%	0	1
Obésité de classe I	12,54%	33,12%	0	1
Obésité de classes II et III	4,70%	21,17%	0	1
Invalidité : aucune incapacité	89,76%	30,32%	0	1
Invalidité : 1 ADL ou trouble cognitif	4,63%	21,02%	0	1
Invalidité : 2+ ADL	4,87%	21,52%	0	1
Invalidité : en institution	0,74%	8,58%	0	1

Tableau 3.4 – Description des variables de la population initiale

Tel que l'indique le tableau, la moyenne d'âge au sein de la population initiale est d'environ 54 ans (rappel : les moins de 30 ans sont exclus) et la proportion d'hommes et de femmes est semblable. Suite à l'imputation, nous trouvons que 20% des individus ont un diplôme universitaire, ce qui est semblable à la proportion d'individus ayant un diplôme universitaire dans l'ENSP. La prévalence des maladies se situe en deçà de 10%, abstraction faite de l'hypertension qui affecte plus de 25% des individus. Il n'y a que 1,3% de la population initiale qui souffre de démences.

Du côté des facteurs de risque, on note que près de 24% de la population initiale fume et que 17,2% des individus souffrent d'obésité. La proportion de personnes qui ont une ou plusieurs incapacités est inférieure à 10%.

## 4 Transitions

### 4.1 Survol

Les modèles de transitions permettent de calculer la probabilité qu'un individu change d'état de santé ou de comportements en fonction de ses caractéristiques individuelles. De tels modèles requièrent d'estimer, au préalable, les paramètres de modèles économétriques de choix discrets ayant comme variables dépendantes les transitions que l'on veut modéliser. Dans COMPAS, ces paramètres sont estimés à partir de données provenant de l'ENSP car la dimension longitudinale de cette enquête permet

d’observer les transitions des individus entre les différents états. Ensuite, les paramètres estimés sont utilisés pour calculer les probabilités de transitions des individus de la simulation en fonction de leurs caractéristiques individuelles et ainsi simuler l’évolution de leurs conditions de santé dans le futur.

Dans la section 4.2, nous présentons les modèles économétriques utilisés dans COMPAS. Ceux-ci estiment les effets des caractéristiques individuelles sur les transitions se produisant entre les cycles de l’ENSP. Par défaut, ces cycles sont réalisés tous les deux ans ; ainsi, les probabilités individuelles de changer d’état sont calculées pour un intervalle de deux années.

## 4.2 Modèles économétriques

### 4.2.1 Modèles pour les maladies

Un modèle de transition est utilisé pour estimer les probabilités d’incidence de chaque maladie considérée. Les probabilités sont fonction de l’âge, de différents facteurs de risque, de caractéristiques socio-économiques et, dans certains cas, de la prévalence d’autres maladies. Comme restrictions, nous imposons que certaines maladies n’aient aucun effet sur l’incidence d’autres maladies, de façon à ne pas créer la présence de liens factices entre ces maladies. Les restrictions ont été évaluées par un panel d’experts dans le cadre du Future Elderly Model (FEM) et sont basées sur de la recherche médicale quant à la relation entre les différentes maladies (Goldman *et al.*, 2005).

Le tableau 4.1 présente les différents effets permis. Par exemple, un « × » à l’intersection de la rangée diabète et de la colonne hypertension signifie que nous permettons que la prévalence du diabète chez un individu ait un effet sur sa probabilité de commencer à faire de l’hypertension. En outre, les maladies sont considérées comme étant des états « absorbants ». Ceci implique qu’une fois qu’un individu souffre d’une maladie, il en est atteint jusqu’à la fin de sa vie (voir les explications sur la prévalence fournies au chapitre 2).

	Diabète	Hypertension	Cancer	Maladies pulmonaires	Maladies cardiaques	AVC	Démences
Diabète	s.o.	x			x	x	
Hypertension		s.o.			x	x	
Cancer			s.o.			x	
Maladies pulmonaires				s.o.			
Maladies cardiaques					s.o.	x	
AVC						s.o.	
Démences							s.o.

Tableau 4.1 – Effets permis des maladies les unes sur les autres

Soit la spécification la suivante :

$$inc_{i,j,t+1}^* = \lambda_j \mathbf{y}_{i,t} + \beta_j \mathbf{x}_{i,t} + \epsilon_{i,j,t}, \quad (5)$$

avec

$$inc_{i,j,t+1} = \begin{cases} 1 & \text{si } inc_{i,j,t+1}^* > 0 \\ 0 & \text{si } inc_{i,j,t+1}^* \leq 0 \end{cases} \quad (6)$$

où :

- $inc_{i,j,t+1}^*$  est une variable latente de l'incidence de la maladie  $j$  pour l'individu  $i$  à la période  $t + 1$  ;
- $inc_{i,j,t+1}$  est une variable binaire indiquant l'incidence de la maladie  $j$  pour l'individu  $i$  à la période  $t + 1$  ;
- $\mathbf{y}_{i,t}$  est un vecteur de variables binaires indiquant la présence ou l'absence de chaque maladie chez l'individu  $i$  à la période  $t$  ;
- $\boldsymbol{\lambda}_j$  est un vecteur qui inclut les effets des différentes maladies sur la probabilité d'incidence de la maladie  $j$ , en tenant compte des permissions décrites dans le tableau 4.1 ;
- $\mathbf{x}_{i,t}$  comprend toutes les variables explicatives prises en compte ;
- $\boldsymbol{\beta}_j$  comprend les effets de ces variables sur la maladie  $j$  ;
- $\epsilon_{i,j,t}$  est un terme aléatoire spécifique à l'individu  $i$  et à la maladie  $j$  à la période  $t$ .

On estime les paramètres uniquement sur la population qui ne souffre pas encore de la maladie  $j$  à la période  $t$ . On suppose que la fonction de répartition de  $\epsilon_{i,j,t}$  suit la répartition d'une loi de valeur extrême, ce qui permet d'utiliser le modèle log-log complémentaire. Ce modèle se distingue des modèles de choix discrets logit et probit, qui sont plus fréquemment utilisés, en relâchant l'hypothèse de symétrie des  $\epsilon_{i,j,t}$  autour de 0, ce qui en fait une meilleure spécification quand l'incidence de la variable dépendante survient rarement. Les paramètres ainsi estimés permettent ensuite de projeter les probabilités de transition individuelles dans la simulation par la formule suivante :

$$\hat{P}(inc_{i,j,t+1} = 1 | \mathbf{y}_i, \mathbf{x}_i) = 1 - \exp(-\exp(\boldsymbol{\lambda}_j \mathbf{y}_{i,t} + \boldsymbol{\beta}_j \mathbf{x}_{i,t})). \quad (7)$$

	Diabète	Hypertension	Cancer	Maladies cardiaques
Âge (si 50 ans ou moins)	0,0008 ***	0,0024 ***	0,0007 ***	0,0010 ***
Âge (si plus de 50 ans)	0,0002 ***	0,0003 ***	0,0003 ***	0,0006 ***
Fumeur	-0,0002	0,0046 **	0,0004	0,0046 ***
Ancien fumeur	0,0010	0,0037 **	0,0005	0,0019 *
Obésité de classe I	0,0101 ***	0,0158 ***	-0,0006	0,0025 *
Obésité de classes II et III	0,0169 ***	0,0208 ***	0,0041 ***	0,0060 ***
Femme	-0,0022 ***	-0,0022	0,0002	-0,0033 ***
Immigrant	0,0002	-0,0002	-0,0019 **	-0,0021 *
Diplôme d'études secondaires	0,0003	-0,0029 *	0,0009	-0,0017
Diplôme d'études collégiales	-0,0003	-0,0012	0,0020 **	-0,0008
Diplôme universitaire	-0,0037 ***	-0,0036	-0,0001	-0,0031 *
Réside au Québec	0,0010	0,0015	-0,0005	0,0006
Présence de diabète	—	0,0074 ***	—	0,0050 ***
Présence d'hypertension	—	—	—	0,0075 ***
Présence de maladies cardiaques	—	—	—	—
Présence de cancer	—	—	—	—
Nombre d'observations	105 502	98 313	106 099	103 918
Incidence moyenne	0,89%	2,82%	0,73%	1,34%

	AVC	Maladies pulmonaires	Démences
Âge (si 50 ans ou moins)	0,0004 ***	0,0001	0,0001
Âge (si plus de 50 ans)	0,0003 ***	0,0003 ***	0,0005 ***
Fumeur	0,0026 ***	0,0069 ***	0,0008
Ancien fumeur	0,0008	0,0031 ***	0,0011 *
Obésité de classe I	-0,0003	0,0003	-0,0019 **
Obésité de classes II et III	-0,0029 **	0,0037 **	-0,0014
Femme	-0,0004	0,0024 ***	-0,0002
Immigrant	0,0009 **	-0,0029 ***	0,0010 ***
Diplôme d'études secondaires	-0,0005	-0,0026 ***	-0,0008
Diplôme d'études collégiales	-0,0016 *	-0,0032 ***	-0,0014 *
Diplôme universitaire	-0,0037 ***	-0,0040 ***	-0,0027 ***
Réside au Québec	-0,0005	0,0008	0,0003
Présence de diabète	0,0030 ***	—	—
Présence d'hypertension	0,0004	-	—
Présence de maladies cardiaques	0,0017 ***	—	—
Présence de cancer	-0,0009	—	—
Nombre d'observations	107 222	104 807	107 709
Incidence moyenne	0,40%	0,83%	0,39%

Tableau 4.2 – Effets marginaux moyens des variables sur les probabilités d'incidence des maladies



Le tableau 4.2 présente les effets marginaux moyens des estimations pour les différentes maladies. Les deux premiers effets sont des effets d'âge, créés de la façon suivante :

- jusqu'à 50 ans :  $\min(\text{age}, 50)$  ;
- après 50 ans :  $\max([\text{age} - 50], 0)$ .

On impose donc un noeud à 50 ans, ce qui implique que l'effet de l'âge peut changer à partir de 50 ans. On voit dans le tableau que jusqu'à 50 ans, la probabilité d'incidence de la plupart des maladies augmente significativement avec l'âge, sauf pour les maladies pulmonaires et les démences. Après 50 ans, l'âge continue de faire augmenter la probabilité d'incidence des maladies mais moins rapidement qu'avant — sauf pour les maladies pulmonaires et les démences, deux catégories de maladies pour lesquelles l'ampleur relative des pentes est donc inversée.

Être fumeur affecte positivement et significativement la probabilité d'incidence de certaines maladies, comme l'hypertension, les maladies cardiaques, les AVC et les maladies pulmonaires. Être un ancien fumeur a également des effets positifs sur les probabilités d'incidence, mais qui sont moindres que pour les fumeurs. L'obésité a un effet généralement positif sur les probabilités d'incidence, à l'exception de l'Alzheimer et des autres démences.

Les femmes semblent moins affectées par le diabète et les maladies cardiaques mais davantage par les maladies pulmonaires. Les immigrants ont quant à eux moins de chances d'être affectés par le cancer, les maladies cardiaques et les maladies pulmonaires mais plus de chances d'être atteints d'AVC ou de démences. Un niveau d'éducation plus élevé est généralement associé à des probabilités d'incidence des maladies plus faibles, avec une exception pour le cancer. Les estimations des effets d'éducation peuvent capter deux effets qui se contrebalancent possiblement : premièrement, une santé généralement meilleure chez les gens éduqués et, deuxièmement, une tendance plus grande des gens plus éduqués à consulter un professionnel de la santé, ce qui augmente leur taux d'incidence des maladies observé dans les données d'enquête.

La variable indicatrice pour les résidents du Québec est non-significative dans tous les cas, malgré le fait que nous trouvions des différences significatives entre les taux d'incidence du Québec et du reste du Canada pour certaines maladies dans nos statistiques descriptives (voir le tableau 2.8 du chapitre 2). Cela suggère que les différences observées dans l'incidence des maladies entre le Québec et le Canada sont principalement captées par les autres variables incluses dans les modèles et qu'il est donc approprié d'utiliser les données de l'ENSP pour l'ensemble du Canada plutôt qu'uniquement celle provenant du Québec.

Les effets de la prévalence d'autres maladies sur l'incidence des maladies sont généralement positifs et significatifs quand nous permettons des effets non-nuls. Les exceptions sont que la prévalence du cancer et de l'hypertension ne semblent pas affecter significativement la probabilité d'incidence d'un AVC.

#### 4.2.2 Modèle pour la mortalité

La mortalité est modélisée de la même façon que l'incidence des maladies, sauf qu'on inclut en plus comme variables explicatives les statuts d'invalidité et qu'on n'impose aucune restriction sur les effets de la prévalence des différentes maladies.

Le tableau 4.3 présente les effets marginaux des variables sur la probabilité de mourir dans les deux années suivantes. Bien entendu, la probabilité de mourir augmente significativement avec l'âge, tant avant qu'après 50 ans. Il peut paraître étonnant que l'augmentation soit moins rapide après 50 ans, mais cela peut s'expliquer par le fait que les décès des personnes plus âgées sont mieux prédits par d'autres variables qui sont prises en compte dans le modèle et qui sont corrélées avec l'âge. En effet, la présence de maladies affecte positivement et, dans la plupart des cas, significativement la probabilité de décès (diabète, cancer, maladies cardiaques, Alzheimer et autres démences).

	Effet marginal
Âge (si 50 ans ou moins)	0,0013 ***
Âge (si plus de 50 ans)	0,0009 ***
Présence du diabète	0,0039 ***
Présence d'hypertension	0,0000
Présence du cancer	0,0078 ***
Présence de maladies cardiaques	0,0039 ***
Présence d'AVC	0,0015
Présence de maladies pulmonaires	0,0009
Présence de démences	0,0027 *
Fumeur	0,0101 ***
Ancien fumeur	0,0047 ***
Obésité de classe I	-0,0035 ***
Obésité de classes II et III	-0,0005
Femme	-0,0069 ***
Immigrant	-0,0005
Diplôme d'études secondaires	-0,0034 ***
Diplôme d'études collégiales	-0,0048 ***
Diplôme universitaire	-0,0061 ***
Réside au Québec	-0,0005
Invalité : 1 ADL ou trouble cognitif	0,0089 ***
Invalité : 2+ ADL	0,0152 ***
Invalité : en institution	0,0178 ***
Nombre d'observations	124 958
Mortalité moyenne	1,1%

Tableau 4.3 – Effets marginaux moyens sur les probabilités de décès

Être fumeur ou ancien fumeur affecte aussi significativement la probabilité de décès. Le lien entre l'obésité et les décès est ambigu : cela peut s'expliquer par la perte de poids survenant souvent dans les dernières périodes de la vie.

Être une femme diminue la probabilité de décès, ce qui est cohérent avec l'espérance de vie plus élevée de celles-ci. Un niveau d'éducation plus élevé entraîne également une diminution significative de la probabilité de décès. L'effet de la variable indicatrice pour le Québec n'est pas significatif, ce qui correspond au fait que l'espérance de vie des Québécois est comparable à celle des Canadiens si on prend en compte les autres différences.

Notons que, tel que discuté dans le chapitre 2, les taux de mortalité observés dans l'ENSP sont légèrement différents des taux de mortalité officiels, surtout pour les âges avancés. Par conséquent, le nombre de décès par année découlant de notre modèle pourrait être biaisé. Dans une version ultérieure de COMPAS, nous calibrerons les probabilités estimées de manière à reproduire les taux de mortalité provenant de la [Human Mortality Database](#), ce qui correspond aux taux de mortalité observés par

### 4.2.3 Modèles pour le tabagisme

Nous estimons trois modèles de transition pour le tabagisme :

1. devenir fumeur (estimé pour les individus qui n'ont jamais fumé),
2. cesser de fumer (estimé pour les individus qui fument),
3. recommencer à fumer (estimé pour les anciens fumeurs).

Nous utilisons encore une fois le modèle log-log complémentaire pour les trois transitions. La présence de maladies est toutefois exclue des variables explicatives pour ne pas créer un mécanisme de simultanéité dans le modèle. En effet, le tabagisme augmente la probabilité que les individus développent certaines maladies (maladies pulmonaires, cancers) et qu'ils décèdent des suites de ces dernières. Le tableau 4.4 présente les effets marginaux moyens.

	Initiation	Cessation	Recommencement
Âge (si 50 ans ou moins)	-0,0005 ***	-0,0029 ***	-0,0033 ***
Âge (si plus de 50 ans)	-0,0001	0,0044 ***	-0,0024 ***
Obésité de classe I	-0,0042	0,0146	0,0066
Obésité de classes II et III	-0,0014	0,0307 **	0,0067
Femme	-0,0198 ***	-0,0133 **	-0,0010
Immigrant	0,0013	0,0289 ***	-0,0144 **
Diplôme d'études secondaires	-0,0045 *	0,0451 ***	-0,0109 *
Diplôme d'études collégiales	-0,0129 ***	0,0756 ***	-0,0181 ***
Diplôme universitaire	-0,0213 ***	0,1217 ***	-0,0342 ***
Réside au Québec	0,0035	0,0021	-0,0053
Nombre d'observations	52 555	32 168	40 235
Incidence moyenne	2,60%	14,26%	6,71%

Tableau 4.4 – Effets marginaux moyens sur les probabilités de transition entre les états de tabagisme

Jusqu'à 50 ans, l'âge a un effet négatif sur toutes les probabilités de transitions. Après 50 ans, l'âge augmente la probabilité d'arrêter de fumer et diminue celle de recommencer. Pour ce qui est de l'obésité, on observe seulement un effet significatif pour l'effet de l'obésité de classes II-III sur la cessation. Être une femme diminue les probabilités des trois transitions, et être un immigrant augmente les chances d'arrêter et diminue celles de recommencer. Finalement, un niveau d'éducation plus élevé est associé à de meilleures chances d'arrêter de fumer et de plus faibles probabilités de commencer ou de recommencer.

### 4.2.4 Modèle pour l'invalidité

Nous modélisons les probabilités de transitions entre tous les états d'invalidité du modèle (« aucune incapacité », « une seule limitation d'activité quotidienne ou trouble cognitif », « deux limitations ou plus avec ou sans trouble cognitif », et « en institution »). L'estimation est faite à l'aide d'un modèle logit multinomial, qui permet d'estimer les probabilités qu'un individu se retrouve dans chacun des états d'invalidité possibles au cycle suivant.

On suppose que les caractéristiques individuelles affectent les probabilités de se retrouver dans chaque état d'invalidité par le biais de quatre variables latentes (une pour chaque état) :

$$inv_{i,j,t+1}^* = \beta_j \mathbf{x}_{i,t} + \epsilon_{i,j,t}, \quad \forall j = 1, 2, 3, 4 \quad (8)$$

avec

$$Pr(inv_{i,t+1} = j) = Pr(inv_{i,j,t+1}^* > inv_{i,k,t+1}^*), \quad \forall k \neq j \quad (9)$$

où :

- $inv_{i,j,t+1}^*$  est la variable latente de l'état d'invalidité  $j$  pour l'individu  $i$  à la période  $t + 1$  ;
- $inv_{i,t+1}$  est la variable catégorique d'invalidité de l'individu  $i$  à la période  $t + 1$  ;
- $\mathbf{x}_{i,t}$  est le vecteur de caractéristiques individuelles de l'individu  $i$  à la période  $t$  ;
- $\beta_j$  est le vecteur des paramètres des effets des variables  $\mathbf{x}_{i,t}$  sur la variable latente de l'état d'invalidité  $j$ ,
- $\epsilon_{i,j,t}$  est un terme aléatoire spécifique à l'individu  $i$  et à l'invalidité  $j$  à la période  $t$ .

Le vecteur  $\mathbf{x}_{i,t}$  inclut la présence de toutes les maladies et des facteurs de risques, les caractéristiques socio-économiques, ainsi qu'une variable indicatrice pour chaque état d'invalidité lors de la période précédente. On suppose que  $\epsilon_{i,j,t}$  suit une distribution de valeur extrême standard de type 1, ce qui fait que la différence entre les termes d'erreur de deux états suit une loi logistique. Comme on utilise la différence entre les valeurs des variables latentes pour calculer les probabilités, on doit normaliser à 0 le vecteur de paramètres pour l'un des états. On peut cependant obtenir les effets marginaux pour tous les états.

La [tableau 4.5](#) présente les estimations des effets marginaux moyens pour chaque état d'invalidité. L'âge augmente la probabilité de devenir invalide, surtout après 50 ans. La présence de maladies a généralement des effets positifs et significatifs sur les probabilités de souffrir des différents niveaux d'invalidité. On observe cependant que certaines maladies n'affectent pas significativement les chances d'être en institution. C'est le cas pour l'hypertension, le cancer, les maladies cardiaques et les maladies pulmonaires.

Être fumeur augmente les risques de souffrir de tous les niveaux d'invalidité, alors qu'être un ancien fumeur augmente seulement le risque de souffrir du premier niveau. Les effets de l'obésité ne sont pas significatifs.

Être une femme augmente significativement les probabilités de souffrir d'invalidité, alors que les immigrants semblent se retrouver moins souvent en institution. Un niveau d'éducation élevé est associé à des risques plus faibles de souffrir d'invalidité. Les effets des variables indicatrices d'invalidité sont évidemment forts car les individus ont tendance à demeurer dans leur état d'invalidité présent à la période suivante. Ainsi, la probabilité d'avoir une incapacité ou un problème cognitif à la période suivante augmente de 9,75% pour les individus qui ont une incapacité ou un problème cognitif à la période courante. Les individus en institution à la période courante ont 12,70% plus de chances d'avoir au moins deux incapacités lors de la prochaine période. Enfin, les Québécois ont tendance à moins souffrir du premier niveau d'invalidité que le reste des Canadiens.

	Aucune incapacité	1 ADL ou prob. cogn.	2+ ADL	En institution
Âge (si 50 ans ou moins)	-0,0003 *	-0,0001	0,0003 ***	0,0000
Âge (si plus de 50 ans)	-0,0022 ***	0,0008 ***	0,0010 ***	0,0004 ***
Diabète	-0,0155 ***	0,0050 ***	0,0096 ***	0,0009 ***
Hypertension	-0,0092 ***	0,0070 ***	0,0028 **	-0,0005
Cancer	-0,0102 ***	0,0073 **	0,0032 *	-0,0003
Maladies cardiaques	-0,0190 ***	0,0116 ***	0,0078 ***	-0,0005
AVC	-0,0283 ***	0,0120 ***	0,0149 ***	0,0013 ***
Maladies pulmonaires	-0,0213 ***	0,0116 ***	0,0104 ***	-0,0008
Démences	-0,0381 ***	0,0160 ***	0,0170 ***	0,0051 ***
Fumeur	-0,0177 ***	0,0117 ***	0,0046 ***	0,0014 ***
Ancien fumeur	-0,0047 **	0,0044 **	0,0000	0,0004
Obésité de classe I	-0,0034	0,0024	0,0011	-0,0001
Obésité de classes II et III	-0,0071 *	0,0048	0,0043	-0,0020
Femme	-0,0192 ***	0,0134 ***	0,0052 ***	0,0005 **
Immigrant	-0,0013	-0,0018	0,0044 ***	-0,0013 **
Diplôme d'études secondaires	0,0096 ***	-0,0046 ***	-0,0042 ***	-0,0008 **
Diplôme d'études collégiales	0,0118 ***	-0,0048 **	-0,0061 ***	-0,0008 **
Diplôme universitaire	0,0241 ***	-0,0098 ***	-0,0119 ***	-0,0023 ***
Invalidité : 1 ADL ou prob. cogn.	-0,1412 ***	0,0975 ***	0,0417 ***	0,0019 ***
Invalidité : 2+ ADL	-0,1698 ***	0,0663 ***	0,0987 ***	0,0047 ***
Invalidité : en institution	-0,2250 ***	0,0688 **	0,1270 ***	0,0289 ***
Réside au Québec	0,0057 **	-0,0052 ***	-0,0005	0,0000
Nombre d'observations			108 355	
Incidence moyenne	91,20%	4,26%	4,06%	0,34%

Tableau 4.5 – Effets marginaux moyens sur les probabilités de transition vers les états d'invalidité

#### 4.2.5 Modèle pour l'obésité

Nous modélisons les transitions entre les différents états possibles d'obésité ( $IMC < 30$ ,  $30 \leq IMC < 35$ , et  $IMC \geq 35$ ) de la même manière que pour l'invalidité. Cependant, le modèle n'inclut pas la présence de maladies et les états d'invalidité comme variables explicatives, car l'obésité est considérée comme un facteur de risque qui affecte ces variables et non comme une conséquence de celles-ci. Le tableau 4.6 présente les effets marginaux moyens.

On trouve qu'une année d'âge avant 50 ans augmente de 0,06% par année la probabilité de commencer à souffrir d'obésité de classe I, mais n'affecte pas la probabilité de commencer à souffrir d'obésité de classes II-III. Après 50 ans, l'âge diminue la probabilité d'être obèse, tant pour l'obésité de classe I que pour celle de classes II-III. Être fumeur augmente significativement la probabilité de devenir obèse de classes II-III, alors qu'être un ancien fumeur a un effet positif et significatif sur l'incidence des deux types d'obésité. Cette dernière relation est souvent observée en raison de la prise de poids pouvant survenir suite à l'arrêt de la consommation de tabac.

Les effets des variables indicatrices d'états d'obésité sont bien sûr très significatifs, car les individus ont tendance à demeurer dans leur état actuel à la période suivante. Les femmes ont légèrement moins tendance à devenir obèses que les hommes, tout comme les immigrants deviennent moins obèses que les non-immigrants. Un niveau d'éducation plus élevé a généralement un effet négatif sur la probabilité

de devenir obèse. Finalement, résider au Québec est associé à une probabilité légèrement plus faible de devenir obèse.

	IMC < 30	$30 \leq \text{IMC} < 35$	IMC $\geq 35$
Âge (si 50 ans ou moins)	-0,0006 ***	0,0006 ***	0,0000
Âge (si plus de 50 ans)	0,0009 ***	-0,0005 ***	-0,0005 ***
Fumeur	-0,0001	-0,0005	0,0005 ***
Ancien fumeur	-0,0078 ***	0,0047 *	0,0032 ***
Obésité de classe I	-0,2324 ***	0,2012 ***	0,0311 ***
Obésité de classes II et III	-0,2715 ***	0,1779 ***	0,0936
Femme	0,0045 **	-0,0108 ***	0,0063
Immigrant	0,0058 **	-0,0061 **	0,0003
Diplôme d'études secondaires	0,0040 *	-0,0031	-0,0009 **
Diplôme d'études collégiales	0,0068 **	-0,0060 *	-0,0008 ***
Diplôme universitaire	0,0134 ***	-0,0097 ***	-0,0037
Réside au Québec	0,0086 ***	-0,0053 **	-0,0033
Nombre d'observations		107 566	
Incidence moyenne	83,3%	11,93%	3,82%

Tableau 4.6 – Effets marginaux moyens sur les probabilités de transition vers les états d'obésité

### 4.3 Conclusions et améliorations futures au modèle

Les conclusions principales de ce chapitre concernent les facteurs de risque et les différences entre le Québec et le Canada.

On note que les facteurs de risque (tabagisme et obésité) ont un impact important sur les transitions vers différentes maladies. À titre d'exemple, les fumeurs et les anciens fumeurs ont davantage de chances de développer de l'hypertension, un cancer, une maladie cardiaque, un AVC ou une démence que les personnes qui n'ont jamais fumé. Similairement, le risque de développer du diabète, de l'hypertension, des maladies cardiaques et des maladies pulmonaires est plus élevé parmi les personnes qui souffrent d'obésité que parmi celles qui n'en souffrent pas.

En outre, le tabagisme et l'obésité sont associés à une plus grande probabilité de développer au moins une incapacité. Les personnes dont l'IMC est supérieur à 30 ont moins de chances de commencer à fumer que celles dont l'IMC est inférieur à 30, alors que le risque de décéder entre deux cycles de l'ENSP est plus élevé pour les fumeurs et les anciens fumeurs que pour ceux qui n'ont jamais fumé.

Les modèles de transition vers les maladies et vers le décès indiquent qu'il n'y a pas de différences significatives entre le Québec et le Canada en ce qui concerne les probabilités de transition. De manière analogue, la variable indicatrice du Québec n'est pas statistiquement significative quand on regarde les transitions entre différents états de tabagisme. Néanmoins, il subsiste quelques différences entre le Québec et le Canada dans les modèles de transition vers l'invalidité et l'obésité. Les individus qui vivent au Québec ont une probabilité plus faible de développer une incapacité et de devenir obèses de classe I que les individus du reste du Canada. Résider au Québec diminue de 0,52% la probabilité annuelle de développer une incapacité ou un problème cognitif. Les probabilités de transition vers l'obésité de classes II-III, vers l'institutionnalisation et vers deux incapacités ou plus ne diffèrent pas entre le Québec et le Canada.

Les résultats du modèle de transition vers les états d'obésité indiquent qu'il reste certaines améliorations à apporter à la persistance de l'obésité dans notre modèle. En effet, la probabilité de souffrir d'obésité de classe I à la période suivante n'augmente que de 20% pour un individu qui souffre d'obésité de classe I à la période courante. La probabilité de souffrir d'obésité de classes II-III à la période suivante, elle, ne dépend pas de l'obésité de classes II-III à la période courante. Il semble donc que les individus simulés changent de catégorie d'IMC relativement facilement. Or, certaines études (par exemple [Daouli \*et al.\* \(2014\)](#)) indiquent plutôt qu'il est difficile pour un individu qui souffre d'obésité de changer de catégorie d'IMC. L'estimation de la dynamique de l'obésité au sein du modèle devra donc être revue afin de représenter de manière plus adéquate la présence d'obésité au long du cycle de vie d'un individu. En raison du manque de persistance, il est possible que la prévalence de l'obésité au sein de COMPAS soit sous-estimée.

La modélisation des cohortes entrantes nécessite également davantage de travail. À ce stade-ci du développement du modèle, ce qui distingue une cohorte d'une autre est l'ensemble des caractéristiques à 30 ans. Par exemple, la prévalence de l'obésité à 30 ans diffère entre les cohortes. Cependant, les probabilités de transitions vers différents états sont les mêmes pour chaque cohorte. L'inclusion d'effets de cohorte pour l'incidence de différentes maladies et facteurs de risque nous permettrait de prendre en considération des changements entre les différentes cohortes en ce qui a trait à la probabilité de changer d'états de santé. En guise d'exemple, ceci impliquerait que la probabilité de devenir obèse de classes II-III pourrait varier entre les cohortes. Les données de L'ENSP nous permettraient de calculer l'incidence moyenne par cohorte entre 1994 et 2010. Ainsi, s'il s'avère que l'incidence d'une maladie a diminué au sein des cohortes plus jeunes, nous pourrions réduire la probabilité d'incidence des maladies au sein des cohortes entrantes dans le modèle. De tels changements pourraient provenir d'avancées médicales ou de changements comportementaux entre les cohortes.

## 5 Renouvellement

### 5.1 Survol

Dans COMPAS, l'entrée d'une nouvelle cohorte d'individus âgés de 30 et 31 ans s'effectue à chaque cycle de la simulation soit, par défaut, tous les 2 ans. Chaque individu entrant possède plusieurs caractéristiques. Toutes proportions gardées, celles-ci doivent refléter à tout moment la loi jointe des conditions initiales (à  $t = 0$ ) de la population cible des 30 et 31 ans, ainsi que d'éventuels changements de tendances que nous aurons appliqués. Dans ce chapitre, nous commençons par décrire notre méthodologie à la section 5.2 puis nous présentons, à la section 5.3, quelques résultats de notre modèle de renouvellement.

### 5.2 Modélisation

À chaque cycle de simulation, on fait entrer des individus d'un âge fixe dans COMPAS. Dénotons par  $y = (y_1, \dots, y_M)$  leurs  $M$  caractéristiques individuelles. Le tableau 5.1 donne la liste des caractéristiques d'un individu entrant. Chacune de ces caractéristiques correspond à un seul type de variable : binaire, ordonné ou entier.

Variabes	Types
Âge	entier
Année de naissance	entier
Sexe	binaire
Immigrant	binaire
Niveau de scolarité	ordonné
Diabète	binaire
Hypertension	binaire
Maladies cardiaques	binaire
AVC	binaire
Cancer	binaire
Maladies pulmonaires	binaire
Démences	binaire
Tabagisme	ordonné
Obésité	ordonné
Invalidité	ordonné
Consultations auprès d'un généraliste	entier
Consultations auprès d'un spécialiste	entier
Nuits d'hospitalisation	entier
Consommation d'au moins un médicament	binaire
Services à domicile	binaire

Tableau 5.1 – Types de variables associés aux caractéristiques des individus entrants

Chaque année, les caractéristiques moyennes de la cohorte entrante changent. Ces caractéristiques sont corrélées entre elles. Soit  $F_t(y)$  la loi jointe de ces caractéristiques à l'année  $t$  et dénotons par  $F_m(y_m)$  la loi marginale de la variable  $y_m$ . Il convient de souligner que plusieurs de ces caractéristiques ont une loi dégénérée, c.-à-d qu'elles prennent une seule valeur avec probabilité 1, par exemple 0<sup>7</sup>.

Nous souhaitons modéliser l'évolution dans le temps des cohortes entrantes. Pour ce faire, à partir de l'information obtenue des fichiers de microdonnées à grande diffusion discutés au chapitre 2, nous nous intéressons à l'évolution de  $\{F_{m,t}(y_m), m = 1, \dots, M\}$ . Nous y parvenons en modélisant l'évolution de l'espérance de  $y_m$ . Cette modélisation dépend du type de la variable représentant une caractéristique individuelle. Dans la présente section, nous discutons uniquement de certaines variables de type binaire ou ordonné. À ce titre, nous ne traitons pas des variables de consultations auprès de médecins ni de la variable qui capte le nombre de nuits d'hospitalisation. Par ailleurs, tous les individus entrants ayant 30 ou 31 ans, il n'y a aucun intérêt à modéliser l'âge et l'année de naissance. Ainsi, nous tenons uniquement compte des corrélations entre les variables binaires et ordonnées du tableau 5.1, et ce sont celles-ci uniquement qui sont préservées au sein des cohortes entrantes.

Soit  $\{\tau_{m,t}, m = 1, \dots, M\}$  une suite telle que :

$$E_{m,t}[y_m] = \tau_{m,t} E_{m,0}[y_m], \quad (10)$$

où  $E_{m,0}[y_m]$  est la moyenne de  $y_m$  à l'année d'initialisation du modèle ( $t = 0$ ) et  $E_{m,t}[y_m]$  est la valeur

7.  $X$  est une variable aléatoire dégénérée si et seulement si  $X$  est presque sûrement égale à une constante.



attendue de  $y_m$  à  $t$ . Dans COMPAS, l'année d'initialisation correspond à l'année 2010. Intéressons-nous premièrement aux variables binaires. Ces variables peuvent s'écrire comme

$$y_m = \mathbb{1}_{(\beta_m + \epsilon_m > 0)}, \quad (11)$$

où  $\mathbb{1}$  est une variable indicatrice,  $\beta_m$  est un seuil et  $\epsilon_m$  est un terme d'erreur normalement distribué. Ainsi, nous avons

$$E[y_m] = \Phi(\beta_m), \quad (12)$$

où  $\Phi(\cdot)$  est la fonction de répartition d'une loi normale.

En ce qui concerne les variables ordonnées avec  $K$  catégories, nous avons  $y_m = k$  si

$$\beta_{m,k} > \epsilon_m > \beta_{m,k-1}, \quad (13)$$

où  $\beta_0 = -\infty$  et  $\beta_K = \infty$ . Ainsi, si  $y_{m,k} = \mathbb{1}_{(y_m=k)}$ , il en découle que :

$$E[y_{m,k}] = \Phi(\beta_{m,k}) - \Phi(\beta_{m,k-1}). \quad (14)$$

Nous supposons dans ce cas que le vecteur des résidus inobservables est  $\epsilon = (\epsilon_1, \dots, \epsilon_M) \sim \mathcal{N}(0, \Omega)$ .

Pour les caractéristiques binaires, nous ajustons chaque année les paramètres de telle sorte que

$$\begin{cases} \tau_{m,t} E[y_m] &= \Phi(\beta_m^*) \\ \beta_m^* &= \Phi^{-1}(\tau_{m,t} E[y_m]) \end{cases} \quad (15)$$

où  $\beta_m^*$  est un seuil et  $\tau_{m,t}$  est un facteur de changement tel que  $\tau_{m,0} = 1$ . Les facteurs de changements  $\tau_{m,t}$  se basent sur des tendances historiques observées au sein de la population canadienne. Tel qu'expliqué au chapitre 9, nous utilisons l'Enquête sur la population active (EPA) et l'ESCC pour calculer ces tendances historiques, concernant par exemple le niveau d'éducation. Celles-ci sont ensuite utilisées pour calculer les facteurs de changements  $\tau_{m,t}$ .

De même, pour les variables ordonnées, on a pour  $k = 1$

$$\begin{cases} \tau_{m,1,t} E[y_{m,1}] &= \Phi(\beta_{m,1}^*) \\ \beta_{m,1}^* &= \Phi^{-1}(\tau_{m,1,t} E[y_{m,1}]) \end{cases} \quad (16)$$

où  $\beta_{m,1}^*$  est un seuil et  $\tau_{m,1,t}$  est un facteur de changement tel que  $\tau_{m,1,0} = 1$ . De manière récursive, on peut alors ajuster les autres paramètres en utilisant

$$\tau_{m,k,t} E[y_{m,k}] = \Phi(\beta_{m,k}^*) - \Phi(\beta_{m,k-1}^*). \quad (17)$$

Ainsi, nous avons

$$\beta_{m,k}^* = \Phi^{-1}(\tau_{m,k,t} E[y_{m,k}] + \Phi(\beta_{m,k-1}^*)). \quad (18)$$

Avec les nouveaux paramètres, on peut alors faire le tirage d'une nouvelle cohorte en utilisant un tirage aléatoire provenant de la distribution normale multivariée. En effet, si  $\tilde{\eta}$  est un vecteur de tirage indépendant de la variable aléatoire gaussienne standard, alors un tirage de  $\tilde{\epsilon}$  est donné par  $L_\Omega \tilde{\eta}$  où  $L_\Omega$  est la matrice triangulaire issue de la factorisation de Cholesky de la matrice  $\Omega$ .

Pour résumer, si l'on connaît les facteurs de changement  $\{\tau_{m,t}, m = 1, \dots, M\}$  estimés à l'aide de données officielles, alors une caractéristique de type binaire pour un individu entrant à la période  $t$  est donnée par :

$$\tilde{y}_m = \mathbf{1}_{(\beta_m^* + \sum_{j=1}^m L_{\Omega}(m,j)\eta > 0)}. \quad (19)$$

De façon similaire, si l'on connaît  $\{\tau_{m,k,t}, m = 1, \dots, M\}$ , alors la caractéristique de type ordonné avec  $k = 1, \dots, K$  catégories est donnée par :

$$\tilde{y}_m = k, \quad \text{si } \beta_{m,k}^* < \sum_{j=1}^m L_{\Omega}(m,j)\eta < \beta_{m,k+1}^*. \quad (20)$$

La taille et les caractéristiques individuelles des nouvelles cohortes peuvent alors être ajustées à l'aide de projections réalisées à l'aide de données officielles — ici, des données provenant de Statistique Canada (voir section 5.3).

### 5.3 Mise en application

Dans COMPAS, nous utilisons principalement les microdonnées à grande diffusion publiées par Statistique Canada. Si l'ENSP est principalement utilisée pour calculer les transitions des individus déjà dans le modèle, nous utilisons plutôt l'ESCC et l'Enquête sur la population active (EPA) pour construire, sur la base des tendances historiques, des facteurs de changement à appliquer aux tendances futures. Seules les quatre caractéristiques exogènes que sont l'immigration, le niveau de scolarité, le tabagisme et l'obésité font l'objet de tels ajustements. À ce propos, nous nous référons à [Agence de la santé publique du Canada et Institut canadien d'information sur la santé \(2011\)](#).

#### 5.3.1 Tendances historiques

Afin d'identifier les tendances historiques concernant l'obésité, l'éducation et le tabagisme, nous utilisons les fichiers de microdonnées à grande diffusion de l'ESCC et de l'EPA de 2000 et de 2012. L'échantillon considéré est composé de l'ensemble de la population canadienne âgée de 25 à 34 ans. Ceci revient à supposer implicitement que l'évolution de ces trois caractéristiques individuelles est la même au Québec que dans le reste du Canada, bien que leur niveau ne le soit pas forcément. Nous retenons les individus âgés de 25 à 34 ans pour avoir suffisamment d'observations tout en restant le plus proche possible de l'âge des cohortes entrantes, soit 30 et 31 ans. Seules les tendances moyennes sont considérées, c.-à-d. que nous supposons les mêmes tendances tant pour les hommes que les femmes.

Selon [Statistique Canada \(2000b, 2012d\)](#), voir tableau 5.2), entre 2000 et 2012, le taux d'obésité a augmenté pour la population âgée de 25 à 34 ans. Le taux de croissance annuel moyen (TCAM) a été de 1,09%. Cette croissance est due, essentiellement, à une importante hausse de l'obésité de classes II-III — soit un peu plus de 2% par an — puisque l'obésité de classe I n'a augmenté que de 0,12% par année.

Statistique Canada (2000a, 2012c, voir tableau 5.2) indique une importante diminution de la proportion de Canadiens âgés de 25 à 34 ans qui ne possèdent aucun diplôme. La diminution annuelle a été de 3,41%. Inversement, il n'est pas surprenant de constater que la proportion d'individus possédant un diplôme universitaire a été en hausse, soit de 2,10% par année.

Alors que la tendance est à la hausse pour l'obésité, l'inverse est observé pour le tabagisme (Statistique Canada, 2000b, 2012d, voir tableau 5.2). La proportion de fumeurs a diminué, annuellement, d'environ 2,6%. Bien que la proportion d'anciens fumeurs soit relativement stable, celle de Canadiens n'ayant jamais fumé a augmenté au cours de la période étudiée : le TCAM a été de 0,93%.

Le tableau 5.2 présente un résumé des TCAM calculés à partir des données d'enquête.

Sources de données	Caractéristiques	TCAM
ESCC	Obésité de classe I	0,12%
	Obésité de classes II-III	2,11%
	Obésité totale	1,09%
EPA	Aucun diplôme	-3,41%
	Diplôme d'études secondaires	-1,46%
	Diplôme d'études collégiales	0,38%
	Diplôme universitaire	2,10%
ESCC	Fumeur actuel	-2,57%
	Ancien fumeur	0,10%
	Jamais fumé	0,93%

Tableau 5.2 – Taux de croissance annuel moyen (TCAM) des proportions de certaines caractéristiques individuelles, population canadienne âgée de 25 à 34 ans, 2000 à 2012

### 5.3.2 Projections

Les TCAM historiques obtenus pour l'obésité, l'éducation et le tabagisme ne sont pas utilisés tels quels à chaque cycle de la simulation. Les différents taux de croissance sont ajustés afin, d'une part, de refléter une certaine incertitude quant à l'avenir et, d'autre part, de prendre en considération certaines particularités québécoises.

Entre 1990 et 2004, le TCAM de la proportion d'obèses de 25 à 34 ans a été de 8,34% au Québec (Audet, 2007). Ce taux est plus élevé que le TCAM trouvé, entre 2000 et 2012, en considérant uniquement les Québécois âgés de 25 à 34 ans (Statistique Canada, 2000b, 2012d). Il semble donc y avoir eu un ralentissement de la croissance de l'obésité au cours des dernières années. À notre avis, ceci tend à indiquer que les cohortes futures présenteront vraisemblablement des taux d'obésité à peine plus élevés comparativement à leurs prédécesseuses (Institut national de santé publique du Québec, 2012).

Un effet semblable est observé pour l'éducation. Les observations québécoises de l'EPA de 1990 à 2000 illustrent que, pour l'ensemble des catégories, les TCAM étaient plus importants, en valeur absolue,

qu'entre 2000 et 2012. Il y a donc signe d'un ralentissement dans la variation de la proportion d'individus ayant un niveau de scolarité donné.

Conséquemment, pour les variables d'obésité et d'éducation, le TCAM 2000-2012 présenté dans le tableau 5.2 est maintenu entre 2010 et 2020 seulement. Il est par la suite réduit de moitié en 2020, 2030 et à nouveau en 2040. Le TCAM de 2040 est ensuite maintenu constant entre 2040 et 2050.

Les tendances concernant le tabagisme au Québec sont relativement stables depuis 2009. La prévalence du tabagisme est stable dans la population générale, soit autour de 17 – 20% (Santé Canada, 2012). Ceci semble indiquer que la baisse de la prévalence enregistrée entre 2000 et 2012 est principalement la conséquence d'une diminution importante du nombre de fumeurs au début des années 2000. Il n'y a donc pas raison de croire que la proportion de fumeurs diminuera au sein de chaque cohorte future entre 2010 et 2050. De ce fait, chaque nouvelle cohorte entrante dans le modèle a approximativement, à 30 et 31 ans, la même proportion de fumeurs dans ses rangs que les cohortes précédentes. Autrement dit, les tendances sont stables pour le tabagisme dans le modèle : le TCAM est de zéro. C'est bien ce que montre le tableau 5.4 où sont présentées les proportions au sein des cohortes entrantes.

Contrairement à l'éducation, à l'obésité et au tabagisme, le traitement de l'immigration dans le modèle ne se base pas sur des tendances historiques qui sont ajustées puis projetées dans le futur. Ce sont plutôt les prévisions de la Régie des rentes du Québec (2013) qui sont utilisées. Celle-ci prévoit que le solde migratoire total, qui inclut autant les migrations entre provinces que l'immigration et l'émigration internationales, diminuera légèrement, passant de 38 600 en 2012 à 34 800 en 2015. Il demeure ensuite à son niveau de 2015 jusqu'en 2062. Bien que l'ajout de vagues migratoires s'arrête en 2050 dans COMPAS, ces hypothèses sont utilisées telles quelles dans la microsimulation.

Au tableau 5.3, nous avons codifié les caractéristiques incluses au tableau 5.2. Au tableau 5.4, nous présentons la répartition de certaines caractéristiques individuelles dans les cohortes entrantes. Celle-ci sont ajustées à l'aide des équations 15 et 16.

## 6 Démographie

Le présent chapitre décrit les hypothèses relatives à la mortalité et à l'immigration ainsi que les modèles démographiques servant de base à COMPAS. On y décrit la modélisation des améliorations de la mortalité.

### 6.1 Mortalité

#### 6.1.1 Définitions et statistiques

Le taux (brut) de mortalité est le rapport du nombre de décès de l'année à la population totale moyenne de l'année. Autrement dit, c'est la probabilité de mourir au cours d'une année. Celle-ci varie notamment selon la structure par âge de la population, le sexe et l'année. D'après les statistiques de l'ISQ (Payeur, 2013), le taux de mortalité est resté stable autour de 7,4 pour mille entre 2001 et 2012, s'établissant à 7,5 pour mille en 2012.

Variables	Codes	Caractéristiques
	0	Non immigrant
Immigration	1	Immigrant
	1	Aucun diplôme
	2	Diplôme d'études secondaires
Niveau de scolarité	3	Diplôme d'études collégiales
	4	Diplôme universitaire
	1	Jamais fumé
Tabagisme	2	Fumeur actuel
	3	Ancien fumeur
	1	Absence d'obésité
Obésité	2	Obésité de classe I
	3	Obésité de classes II-III

Tableau 5.3 – Codification de certaines des variables caractérisant un individu

Cycle $t$	Année	Immigration		Niveau de scolarité				Tabagisme			Obésité		
		0	1	1	2	3	4	1	2	3	1	2	3
0	2010	78,53	21,47	4,07	12,39	55,46	28,09	39,62	28,09	32,29	84,30	13,06	2,64
1	2012	78,58	21,42	4,47	12,03	53,94	29,56	41,06	27,86	31,08	84,93	12,43	2,64
5	2020	81,17	18,83	3,31	11,05	54,74	30,90	37,97	29,83	32,20	83,05	12,75	4,20
10	2030	81,48	18,52	2,55	9,57	56,35	31,53	40,38	26,79	32,83	81,13	13,51	5,37
15	2040	82,96	17,04	2,15	8,99	61,94	26,92	39,58	28,62	31,80	80,14	13,28	6,57
20	2050	80,90	19,10	2,37	9,79	58,99	28,85	40,61	27,33	32,07	80,46	14,18	5,37

Tableau 5.4 – Répartition de certaines caractéristiques individuelles dans les cohortes entrantes (en %), aux cycles de simulation  $t \in \{0, 1, 5, 10, 15, 20\}$

Bien que cette probabilité présente un intérêt certain, pour analyser avec plus de précision l'évolution de la mortalité, d'autres indicateurs statistiques, tels que l'espérance de vie à la naissance, leur sont préférés (Payeur, 2013). Pour les Nations Unies, l'espérance de vie à la naissance indique le nombre d'années qu'un nouveau-né devrait vivre si les conditions générales de mortalité au moment de sa naissance devaient rester les mêmes tout au long de sa vie (Organisation des Nations Unies, 2008) : *c'est une caractérisation de la mortalité indépendamment de la structure par âge*. En 2010, l'espérance de vie à la naissance au Québec était de 79,4 ans chez les hommes et de 83,5 ans chez les femmes d'après les statistiques de la Régie des rentes du Québec (2013). Selon les données provisoires de l'ISQ, en 2012, l'espérance de vie à la naissance s'établissait à 79,8 ans chez les hommes et à 83,8 ans chez les femmes (Payeur, 2013).

Dans les données historiques, on constate une nette amélioration de la mortalité entre 1962 et 2010

(Statistique Canada, 2013, tab. 102-0512), (Régie des rentes du Québec, 2013, tab. 14 et 15). Dans COMPAS, puisque l'entrée de jeunes cohortes de 30 et 31 ans dans le processus de modélisation se poursuit jusqu'en 2050, l'efficacité même de notre modèle requiert des projections du taux de mortalité sur une longue période. Par ailleurs, étant donné qu'un individu peut atteindre 110 ans dans COMPAS, les projections devraient idéalement inclure des groupes d'âge pour les aînés des aînés, soit les 90 ans et plus.

Il nous paraît important de souligner que cette procédure de modélisation soulève une problématique d'endogénéité. L'amélioration globale de la mortalité dépend de deux aspects principaux : (i) les tendances liées aux maladies et aux facteurs de risque, et (ii) le progrès technologique, notamment médical. La contribution des maladies et des facteurs de risque est en bonne partie modélisée dans COMPAS, et il est vraisemblable que celle-ci soit, au net, quasi nulle à cause des divers effets antagonistes. Ainsi, à long terme l'amélioration significative de la mortalité provient essentiellement du progrès technologique.

Dans la section suivante, nous présentons quelques méthodes permettant d'établir des projections du taux de mortalité.

### 6.1.2 Méthodes d'estimation

De par le monde, afin de quantifier, par exemple, l'impact de l'allongement de la durée de vie sur le coût des pensions de retraite, des tables de mortalité prospectives ont été établies (Macdonald, 1997; McDonald *et al.*, 1998, pour un descriptif). Des nombreux modèles existants, l'élaboration des projections existantes repose, généralement, sur le modèle *Lee-Carter* introduit dans Lee et Carter (1992), ainsi que sur certaines de ses variantes, notamment le modèle de Poisson log-bilinéaire (Brouhns *et al.*, 2002; Renshaw *et al.*, 1996).

Le modèle Lee-Carter consiste tout d'abord à décomposer la mortalité en deux composantes, l'âge et le temps calendaire, puis à extrapoler dans le futur les tendances constatées dans le passé. Sous sa forme log-bilinéaire, ce modèle s'écrit (Renshaw et Haberman, 2003c) :

$$\ln \mu_x(t) = \alpha_x + \beta_x \kappa_t + \epsilon_{x,t}, \quad (21)$$

où :

- $\mu_x(\cdot)$  est le taux instantané, supposé constant par morceaux, de la mortalité à la date  $t$  pour l'âge  $x$  ;
- $\alpha_x$  représente la composante spécifique à l'âge  $x$  : c'est le comportement moyen des logarithmes de  $\mu_x(\cdot)$  au cours du temps ;
- $\kappa_t$  décrit l'évolution générale du taux de mortalité  $\mu_x(\cdot)$  au cours du temps ;
- $\beta_x$  indique la sensibilité de la mortalité instantanée par rapport à l'évolution générale de la mortalité : c'est l'écart des  $\ln(\mu_x(\cdot))$  par rapport au comportement moyen  $\alpha_x$  ;
- $\epsilon_{x,t}$  est un terme d'erreur, par hypothèse nulle, qui reflète les particularités propres à l'âge  $x$  ou à la date  $t$  qui ne sont pas capturées par le modèle. Les erreurs  $\epsilon_{x,t}$  sont supposées i.i.d.  $\sim \mathcal{N}(0, \sigma^2)$ .

Dans les faits, le logarithme du taux instantané de la mortalité observé est généralement plus volatil aux grands âges étant donné le faible nombre d'observations : on a un nombre absolu de décès plus

petit chez les personnes les plus âgées dû à des effectifs plus restreints. Par conséquent, l'hypothèse sur l'homoscédasticité reste difficile à valider. Afin de remédier à ce problème et améliorer la qualité de l'estimation, [Brouhns et al. \(2002\)](#) proposent de modéliser le nombre de décès constatés à l'âge  $x$  durant l'année  $t$ , que l'on notera  $D_{x,t}$ , via une loi de Poisson de paramètre  $E_{x,t}\mu_x(t)$  :

$$D_{x,t} \sim \text{Poisson}(E_{x,t}\mu_x(t)), \quad (22)$$

où  $E_{x,t}$  est le nombre moyen de personnes d'âge  $x$  durant l'année  $t$  tandis que  $\mu_x(\cdot)$  est le taux instantané défini par :

$$\mu_x(t) = \exp(\alpha_x + \beta_x \kappa_t). \quad (23)$$

Comparativement au modèle Lee-Carter, le modèle de [Brouhns et al. \(2002\)](#) permet non seulement de lever l'hypothèse sur l'homoscédasticité mais en outre, il prend en compte le caractère entier du nombre de décès  $D_{x,t}$  et capture mieux la variance plus grande du taux instantané de la mortalité aux âges élevés due aux effectifs réduits. Des ramifications des modèles définis par les équations (21) à (23) ont été proposées respectivement dans [Renshaw et Haberman \(2003a\)](#) et [Renshaw et Haberman \(2003b,c\)](#).

Nous présentons ci-après l'intégration d'une table de mortalité prospective dans le modèle COMPAS. Nous utilisons principalement les prévisions de taux de mortalité de la [Régie des rentes du Québec \(2013, tab. 14\)](#), dont un extrait apparaît aux tableaux 6.1 et 6.2. Ce choix, qui n'est pas irrévocable, est toutefois cohérent avec la méthode choisie : la RRQ a recours à une extension du modèle Lee-Carter, modèle devenu un standard malgré des critiques (voir par exemple [Guterman et Vanderhoof, 1998](#)).

### 6.1.3 Intégration dans COMPAS

Soit  $s \in S = \{h, f\}$  avec  $S$  un ensemble représentant le genre humain où les lettres  $h$  et  $f$  font référence, respectivement, aux hommes et aux femmes. Soient  $\{ga_i, i = 1, \dots, n\}$  les groupes d'âge observés dans la population et  $\{pp_j, j = 1, \dots, m\}$  les périodes de projection. Dénotons par  $ga_{1,n} < \dots < ga_{n,n}$  et  $pp_{1,m} < \dots < pp_{m,m}$  les échantillons correspondants ordonnés en ordre croissant. Notons par  $M_{s,n,m}$  la table de mortalité prospective par sexe, groupe d'âge croissant et période croissante. Soit  $g$  le nombre d'années entre deux transitions démographiques décrites au chapitre 4. Dénotons par  $nages$  et  $nyears$  le nombre de tranches d'âge et le nombre de cycles dans la simulation, respectivement, dans le modèle COMPAS :

$$nyears = 1 + (stopyear - startyear) / g, \quad (24)$$

$$nages = stopage - startage + 1, \quad (25)$$

où  $startyear$ ,  $stopyear$ ,  $startage$  et  $stopage$  sont respectivement l'année d'initialisation du modèle, l'année de fin de la simulation, l'âge minimal que peut avoir un individu, et l'âge de décès maximal autorisé par le modèle pour un individu. Il est important de préciser que les notions de « groupe d'âge

de la population » et de « tranche d'âge » dans le modèle ne renvoient pas nécessairement aux mêmes quantités mathématiques.

Sous l'hypothèse qu'il y ait déplacement graduel de la réduction des taux de mortalité des plus jeunes âges vers les âges plus avancés (voir ci-après), on définit la diminution annuelle de la probabilité de décès, notée  $q_x(s, a, \cdot)$ , de la façon récursive suivante :

$$q_x(s, a, 1) = 1, \quad (26)$$

$$q_x(s, a, t) = q_x(s, a, t-1) (1 - M_{s,i,j})^g, \quad (27)$$

avec  $a = 1, \dots, nages$  et  $t = 2, \dots, nyears$ .

D'une part, notons que l'âge minimal, *startage*, que peut avoir un individu dans COMPAS, la tranche d'âge courante  $a$  et l'âge réel d'un individu sont liés par la relation :

$$age = startage + a - 1. \quad (28)$$

Ainsi, l'indice matriciel  $i$  dans (27) est lié avec l'âge réel d'un individu et le groupe d'âge de la population dans la table de mortalité prospective  $M_{s,n,m}$  via la relation :

$$i = \sum_{k=1}^n k \times \mathbb{1}_{(age \in ga_{k,n})} = \begin{cases} 1 & \text{si } age \in ga_{1,n} \\ 2 & \text{si } age \in ga_{2,n} \\ \vdots & \\ n & \text{si } age \in ga_{n,n} \end{cases}$$

. D'autre part, l'année d'initialisation du modèle *startyear*, le nombre d'années entre deux transitions démographiques  $g$ , le cycle courant  $t$  et l'année courante *year* dans la simulation sont liés par :

$$year = startyear + g(t-1). \quad (29)$$

Ainsi, l'indice matriciel  $j$  dans (27) est lié avec l'année courante *year* et la période de projection dans la table de mortalité prospective  $M_{s,n,m}$  par la relation :

$$j = \sum_{k=1}^m k \times \mathbb{1}_{(year \in pp_{k,m})} = \begin{cases} 1 & \text{si } year \in pp_{1,m} \\ 2 & \text{si } year \in pp_{2,m} \\ \vdots & \\ m & \text{si } year \in pp_{m,m} \end{cases}$$

On définit alors la probabilité de rester en vie d'un individu, en tenant compte de ses caractéristiques individuelles et après une transition démographique, comme

$$p := p_{s,a,t} = 1 - (1 - p_{i',j',t})q_x(s, a, t), \quad (30)$$

avec  $p_{i',j',t} = 1 - \exp(-\exp(\epsilon_{i',j',t}))$  la fonction de répartition de la loi de Gompertz où  $\epsilon_{i',j',t}$  est un terme aléatoire spécifique à l'individu  $i'$  et à la maladie  $j'$ , incluant son statut d'invalidité, au temps  $t$ ; pour des détails le lecteur se référera aux sections 4.2.1 et 4.2.2. Cette probabilité dépend donc du sexe, de la période de projection, de la tranche d'âge et de l'état de santé individuel.



Partant du constat suivant lequel la mortalité des jeunes s’est nettement améliorée, atteignant un niveau très bas pouvant difficilement continuer à diminuer au même rythme que par le passé, la RRQ applique une modification au modèle Lee-Carter afin de tenir compte du déplacement graduel de la réduction des taux de mortalité des plus jeunes âges vers les âges plus avancés. Puisque notre modèle repose principalement sur les prévisions de réduction de taux de mortalité de la RRQ, nous avons introduit, à l’équation (27), le facteur  $(1 - M_{s,i,j})^g$  afin de prendre cette observation en considération : ce facteur aura tendance à réduire le rythme de décroissance de la probabilité de décès des plus jeunes comparativement aux aînés.

L’extension du modèle utilisé par la RRQ n’incorpore pas d’avis sur une éventuelle évolution de la mortalité en raison des progrès de la médecine, de l’apparition de nouvelles maladies ou facteurs de risques ou encore de changements dans le style de vie. À ce propos, notons que la RRQ le reconnaît implicitement en ces termes : « *les retombées des nouvelles technologies, des recherches en génétique et des nouveaux médicaments sur l’espérance de vie sont difficiles à évaluer. Les habitudes de vie et l’environnement semblent, à ce jour, être des facteurs plus déterminants sur la mortalité des individus* ». Or dans COMPAS, nous modélisons les améliorations dues aux habitudes de vie, mais pas celles dues au progrès technologique. Ce dernier aspect est pris en compte en introduisant une amélioration « exogène » de la mortalité, décrite au chapitre 9.

Nous avons légèrement adapté les tableaux 6.1 et 6.2 afin de prendre en compte deux spécificités de COMPAS : (i) la simulation s’arrête en 2130 ; et (ii) un individu peut vivre jusqu’à 110 ans au maximum. Ainsi, nous avons supposé que les taux annuels de réduction des taux mortalité au-delà de 2060 seront identiques aux projections de la RRQ pour la période 2050 – 2060, quel que soit le groupe d’âge. Par ailleurs, après l’âge de 89 ans, nous considérons que les taux annuels de réduction des taux mortalité sont stables quelle que soit la période de projection.

## 6.2 Immigration

Afin de mieux refléter la réalité démographique, il convient de prendre en considération le déplacement d’individus d’un pays vers un autre — *migration internationale* — ou d’une province vers une autre — *migration interprovinciale*.

Tel qu’indiqué au chapitre 5, dans COMPAS nous nous référons simplement aux chiffres sur la migration nette de la Régie des rentes du Québec (2013). Il s’agit du solde migratoire net, c’est-à-dire, la différence entre l’immigration et l’émigration au cours d’une année. Ces chiffres qui incluent la migration des résidents non permanents montrent que le solde migratoire net a représenté environ 0,48% de la population entre 2001 et 2014. En 2015, les projections indiquent une diminution graduelle de ce solde jusqu’à 0,42%, lequel restera stable jusqu’à l’horizon 2061. En 2062, le solde migratoire net ne représentera plus que 0,36% de la population.

Notons que les projections ci-dessus supposent que (Régie des rentes du Québec, 2013) :

- les politiques québécoises et canadiennes relatives à l’immigration, et en particulier les cibles d’immigration internationale du ministère de l’Immigration et des Communautés culturelles connues en décembre 2012, restent stables ;
- le solde migratoire interprovincial est maintenu stable et correspond au niveau moyen observé au cours des 15 années précédant l’année 2012.

Groupe d'âge	Observé				Projeté				
	1980-1990	1990-2000	2000-2010		2010-2020	2020-2030	2030-2040	2040-2050	2050-2060
	%	%	%	%	%	%	%	%	%
0 à 9 ans	4,8	6,4	2,5	2,5	2,2	1,0	0,5	0,5	0,5
10 à 19 ans	2,4	4,4	5,1	5,1	1,7	0,7	0,5	0,5	0,5
20 à 29 ans	2,0	2,5	3,4	3,4	2,3	0,8	0,5	0,5	0,5
30 à 39 ans	0,1	3,4	3,1	3,1	1,6	0,6	0,5	0,5	0,5
40 à 49 ans	2,5	2,4	2,6	2,6	2,2	1,2	0,7	0,6	0,5
50 à 59 ans	3,2	2,7	2,5	2,5	2,8	2,5	1,9	1,4	1,3
60 à 69 ans	2,1	2,9	3,1	3,1	2,5	2,4	2,1	1,9	1,6
70 à 79 ans	1,2	2,1	3,4	3,4	1,9	1,8	1,6	1,3	1,1
80 à 89 ans	0,6	1,2	2,5	2,5	1,0	0,8	0,7	0,6	0,5

Tableau 6.1 – Taux annuels de réduction des taux de mortalité chez les hommes (source : Régie des rentes du Québec (2013, tab. 14))

Groupe d'âge	Observé				Projeté				
	1980-1990	1990-2000	2000-2010		2010-2020	2020-2030	2030-2040	2040-2050	2050-2060
	%	%	%	%	%	%	%	%	%
0 à 9 ans	4,0	7,6	1,0	1,0	2,5	1,6	0,8	0,6	0,5
10 à 19 ans	4,2	1,7	1,9	1,9	2,1	1,5	0,9	0,6	0,5
20 à 29 ans	3,4	1,5	1,1	1,1	0,9	0,6	0,5	0,5	0,5
30 à 39 ans	1,0	1,8	2,4	2,4	1,3	1,1	0,8	0,6	0,5
40 à 49 ans	2,4	1,1	1,8	1,8	1,3	1,2	1,0	0,8	0,7
50 à 59 ans	2,7	1,1	1,1	1,1	1,4	1,4	1,4	1,4	1,3
60 à 69 ans	2,0	1,5	1,5	1,5	1,6	1,6	1,5	1,4	1,3
70 à 79 ans	1,9	1,6	1,7	1,7	1,5	1,5	1,3	1,1	0,9
80 à 89 ans	1,2	1,3	1,6	1,6	1,1	1,0	0,8	0,6	0,5

Tableau 6.2 – Taux annuels de réduction des taux de mortalité chez les femmes (source : Régie des rentes du Québec (2013, tab. 14))

## 6.3 Conclusion

Ainsi que l'illustre la figure 9.1, en incorporant le modèle de mortalité et les projections de solde migratoire net dans COMPAS et en combinant le résultat avec les prévisions de la RRQ pour les moins de 30 ans<sup>8</sup>, on obtient une population totale d'un peu plus de 9 millions d'habitants en 2050.

## 7 Utilisation des soins de santé

### 7.1 Survol

Le module d'utilisation des soins de santé permet d'évaluer, pour chaque année de simulation, la quantité de ressources médicales utilisées par la population. Afin d'obtenir ces résultats, il faut cependant établir une relation entre l'état de santé et l'utilisation de ressources.

Les modèles utilisés pour ce faire doivent prendre en considération une des particularités des données d'utilisation des ressources médicales, soit le fait qu'elles sont généralement caractérisées par plusieurs valeurs nulles et une distribution asymétrique. Dit autrement, ce qui précède signifie que plusieurs individus, au cours d'une période de référence, ne font aucune utilisation de ressources médicales et que d'autres en consomment beaucoup plus que la moyenne (Frees *et al.*, 2011). Cette constatation s'applique aussi bien au nombre de consultations auprès d'un médecin qu'au nombre de nuits passées en hospitalisation.

Afin de pouvoir tenir compte d'une telle distribution de données, nous utilisons une régression de type binomiale négative. Deux des variables de santé considérées dans COMPAS, soit la consommation d'au moins un médicament et l'utilisation (ou non) de services à domicile, ne peuvent prendre que deux valeurs (nulle ou positive). Pour les modéliser, un modèle de régression logistique est plutôt utilisé. La théorie économétrique des modèles présentés se base sur Cameron et Trivedi (2005).

### 7.2 Modèles économétriques

#### 7.2.1 Régression de type binomiale négative

La régression de type binomiale négative est généralement utilisée afin d'analyser des données discrètes à support dénombrable, soit des données qui ne peuvent prendre que des valeurs entières et non négatives, ce qui représente bien les données d'utilisation de soins de santé. Ce type de régression, en permettant la surdispersion des données (variance supérieure à la moyenne), peut capter l'asymétrie dans la distribution des données d'utilisation de soins de santé. Utiliser la loi de Poisson, qui suppose l'égalité entre la moyenne et la variance, aurait moins bien représenté la distribution des données.

La régression binomiale négative suppose que les observations sont générées à partir d'une distribution binomiale négative, dont les deux premiers moments sont :

---

8. Cette opération est nécessaire ici puisque COMPAS ne simule que la population des 30 ans et plus.

$$\mathbb{E}(y_i|\mu_i, \alpha) = \mu_i = \exp(x_i'\beta)$$

$$\text{Var}(y_i|\mu_i, \alpha) = \mu_i(1 + \alpha\mu_i)$$

où :

- $\mathbf{x}_i$  est un vecteur de variables explicatives pour l'individu  $i$  ;
- $\beta$  est un vecteur de coefficients ;
- $\alpha$  est le paramètre de surdispersion.

Si  $\alpha > 0$ , il découle automatiquement de la définition de la variance qu'il y a surdispersion, puisque la variance est plus grande que la moyenne ( $\mu_i$  étant positif). Cette définition de la variance est celle qui est utilisée par la commande *nbreg* de Stata qui effectue la régression de type binomiale négative. L'estimation des coefficients s'effectue par maximum de vraisemblance.

Toutefois, les paramètres estimés ne donnent pas directement l'effet d'une variable  $x$  sur l'espérance conditionnelle de  $y$ . Ce ne sont donc pas les paramètres eux-mêmes qui sont intéressants pour analyser l'impact des variables explicatives, mais bien les effets marginaux, soit le changement de l'espérance conditionnelle de  $y_i$  quand la valeur d'une variable  $x_i$  change d'une unité. Tel que mentionné dans le chapitre 3, dans le cadre de COMPAS, nous utilisons la moyenne des effets marginaux (MEM). En guise de rappel, la MEM est une moyenne sur l'ensemble des individus de l'effet du changement d'une unité de la variable  $x_i$  sur l'espérance de  $y_i$ .

### 7.2.2 Régression de type logistique

La régression logistique, tel que mentionné précédemment, s'utilise avec des variables qui ne peuvent prendre que deux valeurs, par exemple la consommation ou non d'au moins un médicament. La variable dépendante  $y$  peut donc prendre une valeur de 1 avec probabilité  $p$  et une valeur de 0 avec probabilité  $1 - p$ . Cette probabilité est modélisée de la façon suivante :

$$p \equiv \mathbb{P}(y_i = 1|x_i) = \frac{\exp(x_i'\beta)}{1 + \exp(x_i'\beta)}$$

où :

- $\mathbf{x}_i$  est un vecteur de variables explicatives pour l'individu  $i$  ;
- $\beta$  est un vecteur de coefficients.

L'estimation des paramètres s'effectue encore une fois par maximum de vraisemblance. Tout comme pour les régressions binomiales négatives, les paramètres estimés n'ont pas une interprétation directe. C'est la MEM qui est utilisée pour représenter l'impact des variables explicatives sur la probabilité que  $y_i$  prenne une valeur de 1.

Les deux types de régression décrits dans cette section nous permettent de modéliser une relation entre l'utilisation de ressources médicales et la présence de sept maladies, de deux facteurs de risque et de l'invalidité.

## 7.3 Résultats

Cette section présente les résultats des différents modèles de régression discutés ci-dessus appliqués aux données de l'ENSP (voir le chapitre 2 à ce sujet). Bien que le type de régression change selon le type de soins de santé considéré, les variables explicatives demeurent les mêmes dans toutes les régressions. Les régressions sont uniquement effectuées à l'aide des répondants qui sont en ménages privés à la période courante, car il n'y a pas suffisamment d'information sur l'utilisation des soins pour ceux qui sont en institution. À titre d'exemple, si un répondant est en institution lors du dernier cycle de l'enquête, ses réponses aux cycles précédents sont tout de même utilisées (seules ses réponses au dernier cycle sont exclues). Les variables socio-démographiques utilisées sont le sexe, l'âge, le statut d'immigration, l'éducation et un indicateur pour les répondants vivant au Québec. Cette dernière variable permet de capter une différence possible entre les répondants québécois et ceux du reste du Canada.

L'effet de l'âge sur l'utilisation de ressources ne semble pas être linéaire. Plus précisément, il semble y avoir un changement à partir de 50 ans. Afin de capter cet effet, un noeud y a été créé ([Goldman et al., 2005](#)). Il y a donc deux variables d'âge, une pour les individus âgés de moins de 50 ans et une pour ceux qui ont plus de 50 ans. Cette seconde variable capte le nombre d'années écoulées depuis que l'individu a atteint 50 ans (elle prend une valeur de 5 pour quelqu'un qui a 55 ans). Le coefficient estimé pour la première variable d'âge capte ainsi l'effet de l'âge jusqu'à 50 ans sur la variable dépendante; similairement, le coefficient de la seconde variable est l'effet de l'âge après 50 ans sur l'utilisation de ressources en santé.

Les variables d'invalidité et de présence des sept maladies étudiées sont présentes dans les régressions. Les facteurs de risque (tabagisme et obésité) sont également inclus. Le tabagisme est représenté par deux variables binaires, soit une pour les anciens fumeurs et une pour les fumeurs. Deux variables binaires sont présentes dans la régression pour capter l'effet de l'obésité de classes I ( $30 \leq \text{IMC} < 35$ ) et II-III ( $\text{IMC} \geq 35$ ). Les catégories de référence sont donc les individus n'ayant jamais fumé et ceux ayant un IMC inférieur à 30.

La pondération utilisée afin de rendre l'échantillon représentatif de la population canadienne est le « poids longitudinal carré » du fichier maître, soit le premier poids calculé par [Statistique Canada \(2012a\)](#). Ce poids est calculé lors du premier cycle de l'ENSP et ne change jamais. Il permet de rendre l'échantillon représentatif de la population canadienne, même si lors de cycles subséquents certains répondants ne participent plus à l'ENSP ou ne répondent plus à l'ensemble des questions.

### 7.3.1 Régression de type binomiale négative

L'effet des différentes maladies sur l'utilisation de ressources médicales est présenté dans le tableau 7.1. L'utilisation est calculée sur une période de 12 mois. Ainsi, si un individu a consulté un médecin spécialiste 3 fois, ceci indique qu'il a consulté un médecin spécialiste 3 fois au cours d'une période de 12 mois. Tel qu'attendu, les maladies ont un effet important sur le nombre de consultations auprès d'un médecin généraliste. La présence du diabète, de l'hypertension, du cancer, de maladies cardiaques ou de maladies pulmonaires augmente le nombre de consultations auprès d'un généraliste de plus de 1,2, en moyenne. L'obésité de classes I et II-III augmente également le nombre de consultations, de 0,4 et 0,7 respectivement. Ce sont toutefois les incapacités dont souffrent les individus qui entraînent

l'augmentation la plus importante. Un individu ayant au moins 2 incapacités effectue en moyenne 2,5 consultations de plus qu'un individu qui n'a pas d'incapacité. En moyenne, les Québécois consultent moins souvent que les individus du reste du Canada. L'écart est d'environ 1 consultation.

L'effet des maladies sur le nombre de consultations auprès d'un médecin spécialiste est également positif et significatif. Ainsi, la présence du diabète, de l'hypertension, du cancer, de maladies cardiaques ou de maladies pulmonaires augmente le nombre de consultations. La présence d'une certaine forme de cancer est liée à une hausse de près de 2 consultations par année. L'effet des démences n'est pas significatif, mais ceci pourrait être dû au faible nombre d'individus qui souffrent de démences dans l'échantillon utilisé. L'obésité de classes II-III augmente de 0,1 le nombre de consultations auprès d'un spécialiste, alors que les variables de tabagisme ont peu d'effets, les anciens fumeurs effectuant 0,1 consultation de plus que les individus qui n'ont jamais fumé.

La présence de démences n'a pas d'effet significatif sur le nombre de nuits d'hospitalisation de courte durée, mais toutes les autres maladies ont des effets considérables. La présence d'un cancer ou d'un AVC augmentent en moyenne le nombre de nuits d'hospitalisation de plus de 2. Les individus souffrant de 2 incapacités ou plus ont, en moyenne, 4,2 nuits d'hospitalisation de plus que ceux n'ayant pas d'incapacité. Le coefficient de l'obésité de classe I n'est pas significatif et est près de zéro, mais les individus dont l'IMC est de 35 ou plus (classes II-III) sont hospitalisés en moyenne 0,4 nuit de plus que les individus de poids santé.

### 7.3.2 Régression de type logistique

Le tableau 7.2 présente les effets marginaux moyens de chacune des variables explicatives sur les variables binaires d'utilisation de ressources, soit la consommation d'au moins un médicament et l'utilisation (ou non) de services à domicile.

Toutes les maladies, à l'exception des démences, ont un effet positif et significatif sur la probabilité de consommer au moins un médicament. L'effet le plus important est pour l'hypertension, les individus qui en souffrent ayant 14,2% plus de chances que les individus qui n'en souffrent pas de consommer au moins un médicament. L'impact de l'obésité est également significatif. L'obésité de classes II-III augmente la probabilité de consommer au moins un médicament de 4,3%, alors que pour celle de classe I, l'augmentation est de 2,7%. Tous les types d'invalidité augmentent la probabilité de consommer au moins un médicament. Les individus ayant un diplôme d'études secondaires ou collégiales ou un diplôme universitaire ont plus de chances de consommer au moins un médicament que ceux n'ayant pas de diplôme. Résider au Québec diminue de 2,3% la probabilité de consommer au moins un médicament. Les anciens fumeurs ont 3,6% plus de chances de consommer au moins un médicament que les individus qui n'ont jamais fumé, mais être un fumeur n'a pas d'effet significatif.

La dernière variable d'utilisation de soins de santé étudiée est celle concernant les services à domicile. Les effets des états de santé sur la probabilité de bénéficier de services à domicile sont généralement assez faibles, soit de moins de 1%. Toutefois, la probabilité d'utiliser des services à domicile augmente de manière importante pour les individus qui ont au moins 2 incapacités. Pour ces derniers, la probabilité d'utiliser des services à domicile augmente de 17,8% par rapport aux individus sans incapacité. La présence d'un cancer augmente la probabilité de recevoir des services à domicile de 3,3% alors que la présence d'un AVC, elle, augmente la probabilité de recevoir des services à domicile de 1,5%. On

Variable	Nb. de consultations : généraliste	Nb. de consultations : spécialiste	Nb. de nuits hospitalisé
Femme	0,861*** (0,05)	0,272*** (0,02)	0,211* (0,10)
Âge (si moins de 50 ans)	-0,015*** (0,00)	-0,000 (0,00)	-0,008 (0,01)
Âge (si 50 ans ou plus)	0,013*** (0,00)	-0,005*** (0,00)	0,027*** (0,01)
Immigrant	0,320*** (0,07)	0,010 (0,03)	-0,216 (0,12)
Diplôme d'études secondaires	-0,256*** (0,07)	0,216*** (0,03)	-0,133 (0,11)
Diplôme d'études collégiales	-0,255** (0,08)	0,205*** (0,04)	-0,206 (0,11)
Diplôme universitaire	-0,380*** (0,08)	0,395*** (0,05)	-0,242 (0,16)
Réside au Québec	-0,931*** (0,05)	0,092** (0,03)	0,636*** (0,18)
Présence du diabète	1,216*** (0,11)	0,179*** (0,04)	0,533*** (0,15)
Présence de l'hypertension	1,228*** (0,06)	0,167*** (0,03)	0,297** (0,10)
Présence du cancer	1,353*** (0,15)	1,929*** (0,12)	2,851*** (0,44)
Présence de maladies cardiaques	1,306*** (0,10)	0,636*** (0,05)	1,639*** (0,19)
Présence d'un AVC	0,667*** (0,19)	-0,079 (0,06)	2,059** (0,75)
Présence de maladies pulmonaires	1,307*** (0,14)	0,253*** (0,06)	0,370* (0,19)
Présence de démences	0,057 (0,21)	0,118 (0,12)	0,163 (0,27)
Obésité de classe I	0,364*** (0,07)	0,056 (0,03)	-0,037 (0,11)
Obésité de classes II-III	0,698*** (0,12)	0,112* (0,05)	0,405* (0,19)
Fumeur actuel	0,176* (0,07)	0,001 (0,03)	0,294 (0,18)
Ancien fumeur	0,194*** (0,06)	0,090*** (0,02)	0,127 (0,11)
Invalidité : 1 ADL ou trouble cognitif	1,847*** (0,12)	0,713*** (0,07)	2,682*** (0,41)
Invalidité : 2+ ADL	2,541*** (0,14)	0,941*** (0,08)	4,166*** (0,38)
Nombre d'observations	82 715	82 924	76 628
Nombre moyen de consultations ou de nuits	3,11	0,75	0,86
Légende	* p<0,05 ; **p<0,01 ; *** p<0,001		

Tableau 7.1 – Régression de type binomiale négative : effets marginaux moyens des différentes variables sur l'utilisation de soins de santé. L'utilisation est calculée sur une base annuelle.

remarque également que les individus qui sont obèses de classes II-III ont 1% plus de chances de recevoir des services à domicile que ceux qui sont de poids santé.

Variable	1+ médicament	Services à domicile (oui)
Femme	0,099*** (0,00)	0,007*** (0,00)
Âge (si moins de 50 ans)	-0,000 (0,00)	-0,000 (0,00)
Âge (si 50 ans ou plus)	0,001** (0,00)	0,002*** (0,00)
Immigrant	-0,046*** (0,01)	-0,007** (0,00)
Diplôme d'études secondaires	0,029*** (0,00)	-0,002 (0,00)
Diplôme d'études collégiales	0,037*** (0,01)	-0,002 (0,00)
Diplôme universitaire	0,048*** (0,01)	-0,004 (0,00)
Réside au Québec	-0,023*** (0,00)	0,002 (0,00)
Présence du diabète	0,117*** (0,00)	0,009** (0,00)
Présence d'hypertension	0,142*** (0,00)	-0,001 (0,00)
Présence du cancer	0,062*** (0,01)	0,033*** (0,00)
Présence de maladies cardiaques	0,130*** (0,00)	0,012*** (0,00)
Présence d'un AVC	0,095*** (0,01)	0,015*** (0,00)
Présence de maladies pulmonaires	0,081*** (0,01)	0,009** (0,00)
Présence de démences	-0,044 (0,03)	0,002 (0,00)
Obésité de classe I	0,027*** (0,00)	-0,001 (0,00)
Obésité de classes II-III	0,043*** (0,01)	0,010* (0,00)
Fumeur actuel	0,007 (0,00)	0,001 (0,00)
Ancien fumeur	0,036*** (0,00)	0,001 (0,00)
Invalité : 1 ADL ou trouble cognitif	0,064*** (0,01)	0,079*** (0,00)
Invalité : 2+ ADL	0,083*** (0,01)	0,178*** (0,01)
Nombre d'observations	82 199	83 524
Proportion moyenne d'utilisateurs	84,58%	3,40 %
Légende	* p < 0,05 ; **p < 0,01 ; ***p < 0,001	

Tableau 7.2 – Régression de type logistique : effets marginaux moyens des différentes variables sur l'utilisation de soins de santé. L'utilisation est calculée sur une base annuelle.

Ce qu'on peut retenir de ces analyses est que les maladies considérées dans le modèle sont des facteurs importants lorsqu'il s'agit d'expliquer l'utilisation de ressources médicales. L'invalité a également des effets conséquents. Enfin, les répondants qui habitent au Québec se distinguent des répondants du reste



du Canada en ce qui concerne les consultations auprès d'un omnipraticien et les nuits d'hospitalisation. Les Québécois totalisent près d'une consultation supplémentaire et environ 0,6 nuit d'hospitalisation de plus que les répondants des autres provinces.

## 8 Résultats de base en mode cohorte

Ce chapitre est consacré aux résultats de COMPAS en mode cohorte. Ce mode permet de fournir des mesures sur toute une vie (« lifetime meures ») d'un, voire plusieurs, sous-ensemble(s) d'individus aux caractéristiques précises. Comme mesures, on peut citer les espérances de vie ou l'utilisation des soins de santé. Présenter des résultats pour ce mode requiert de définir préalablement un scénario dit « de référence » qui prenne en compte des statistiques officielles, incluant des projections, portant sur le profil socio-démographique et l'état de santé.

### 8.1 Scénario de référence

En mode cohorte, l'intérêt du scénario de référence est de représenter, pour le groupe d'individus concerné, la situation telle qu'elle serait en l'absence de changements dans l'effet et l'importance des déterminants des transitions vers différents états de santé. Le lecteur se référera à la section 9.1 pour la définition et les hypothèses servant de base à la construction d'un tel scénario.

Comme les problèmes de santé considérés dans le modèle sont peu courants chez les jeunes, et afin d'éviter de devoir modéliser les choix d'éducation, COMPAS utilise et simule uniquement des individus âgés de 30 ans ou plus, tel qu'indiqué au chapitre 1. Nous nous intéressons donc ici à la cohorte d'individus nés en 1979 et 1980, soit ceux ayant 30 ou 31 ans en 2010. En mode cohorte, la simulation s'effectue en deux étapes seulement :

1. Le *module d'initialisation* crée une cohorte représentative de la population âgée de 30 et 31 ans en 2010 au Québec. Pour l'essentiel, celle-ci possède les caractéristiques suivantes :
  - caractéristiques socio-démographiques : âge, année de naissance, sexe, statut d'immigration, niveau de scolarité ;
  - maladies : diabète, hypertension, maladies cardiaques, accidents vasculaires cérébraux (AVC), cancer, maladies pulmonaires, démences ;
  - facteurs de risque : tabagisme, obésité ;
  - invalidité : limitations d'activités, troubles cognitifs, institutionnalisation.
2. Le *module de transition* fait vieillir les individus d'une année, en calculant la probabilité de transition vers différents états. Ainsi, chaque individu peut développer une maladie ou une invalidité, perdre ou gagner du poids, commencer à ou cesser de fumer, entrer en institution ou mourir. Suite à cette transition, les individus décédés sont retirés de la cohorte.

L'étape 2 est ensuite répétée jusqu'à ce qu'il n'y ait plus d'individus en vie, ou jusqu'à ce que les derniers survivants atteignent l'âge de 110 ans. Après chaque année de simulation, différentes statistiques sont calculées, dont la prévalence de chaque maladie et des différents facteurs de risque ainsi que l'utilisation des différents types de soins de santé.

## 8.2 Simulation et résultats

Sur la base de notre scénario de référence (voir section 9.1), les étapes 1 et 2 décrites précédemment à la sous-section 8.1 sont répétées 100 fois. Effectuer un plus grand nombre de répliques génère des résultats très similaires. Afin de présenter des résultats sur l'évolution de l'état de santé, une série de figures et tableaux est proposée ci-après. En ce qui concerne les tableaux sur l'utilisation des soins de santé, nous formulons l'hypothèse que l'utilisation est constante sur un cycle de simulation de 2 ans. De façon plus précise, cela signifie que si un individu effectue 4 consultations chez les généralistes en 2014, alors nous présumons que cet individu consultera de nouveau 4 fois en 2015 — soit un total de 8 fois durant ces deux années.

À la figure 8.1, nous présentons l'évolution de la taille moyenne des 100 cohortes (soit une cohorte pour chaque réplique de la simulation), représentatives de la population, constituées d'individus nés en 1979 et 1980. Nous avons superposé les courbes de cette évolution selon que l'on applique le taux de mortalité projeté par la RRQ ou celui de notre modèle, tel que décrit à la sous-section 6.1.3. On observe que les individus décèdent beaucoup plus rapidement lorsque notre modélisation du taux de mortalité est appliquée. Cet écart est plus conséquent aux âges avancés : à 80 et 100 ans par exemple, les différences sont respectivement de 4647 et 5961 individus. Un certain écart est attendu dans la mesure où COMPAS modélise explicitement plusieurs maladies et facteurs de risque importants.

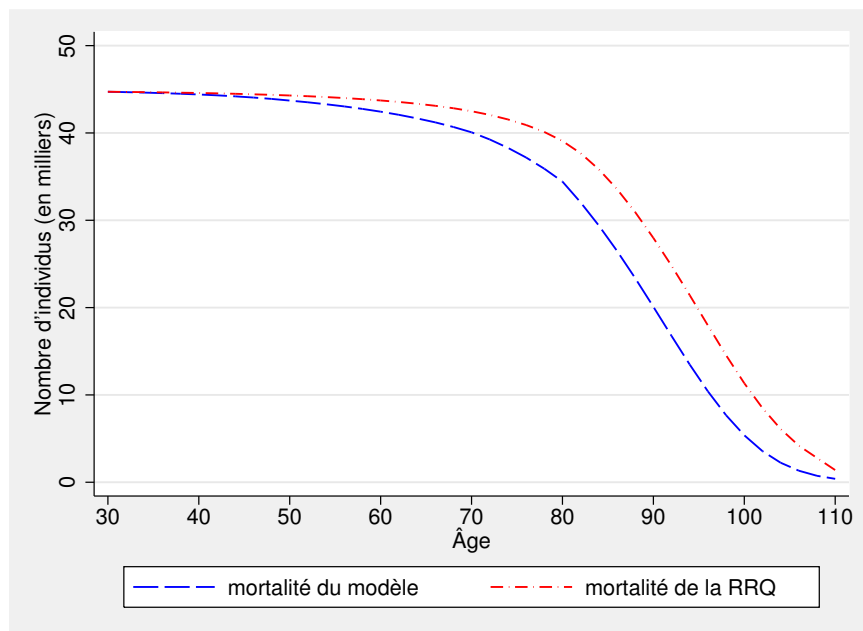


FIGURE 8.1 – Évolution de la taille de la cohorte née en 1979-1980 — Scénario de référence

Le figure 8.2 présente l'évolution de la prévalence des différentes maladies considérées. Exception faite du diabète, on observe une augmentation de la prévalence de l'ensemble des maladies entre 30 et 110 ans. Le diabète augmente jusqu'à 96 ans puis décline jusqu'à 110 ans. L'hypertension est le problème de santé qui touche le plus grand nombre : sa prévalence est de 5,4%, 22,0% et 67,6% respectivement à 30, 50 et 80 ans. Faible à 30 ans, la proportion d'individus souffrant de maladies cardiaques passe de 7,5% à 41,9% entre 50 et 80 ans. Cette tendance à la hausse se maintient jusqu'à 110 ans, moment où la proportion d'atteints est de 76,9%. La prévalence du diabète et des maladies pulmonaires augmente respectivement de 26,6 et 21,1 points de pourcentage entre 30 et 80 ans. La prévalence des maladies pulmonaires continue de croître jusqu'à 110 ans. Auparavant faible, la prévalence des démences et de l'AVC augmente rapidement après 85 ans : elle atteint respectivement 22,4% et 19,0% à 90 ans.

Le tableau 8.1 donne une comparaison des prévalences, à différents âges, des problèmes de santé dans compas avec les prévalences observées dans l'ENSP. On constate que les prévalences simulées provenant de COMPAS le sont pour la cohorte âgée de 30 ans en 2010, les prévalences de l'ENSP se rapportent à un ensemble de cohortes âgées de 30 à 95 ans sur la période 1994 – 2011. Or, ces cohortes ont des caractéristiques observables différentes. Deuxièmement, les questions de l'ENSP en 1994 demandent aux individus s'ils souffrent actuellement de la maladie en question. Or, si une personne a déjà eu une maladie dans le passé mais n'en souffre plus, et qu'elle répond en conséquence, nous sous-estimons la prévalence de celle-ci dans la population initiale. Malgré l'ajustement que nous faisons dans les cycles subséquents (une personne qui a une maladie l'a toujours par la suite, voir les chapitres 2 et 4), nous sommes tout de même susceptibles de sous-estimer la prévalence. Une troisième raison est que nous ne captions pas parfaitement l'hétérogénéité inobservée commune qui affecte l'incidence des maladies.

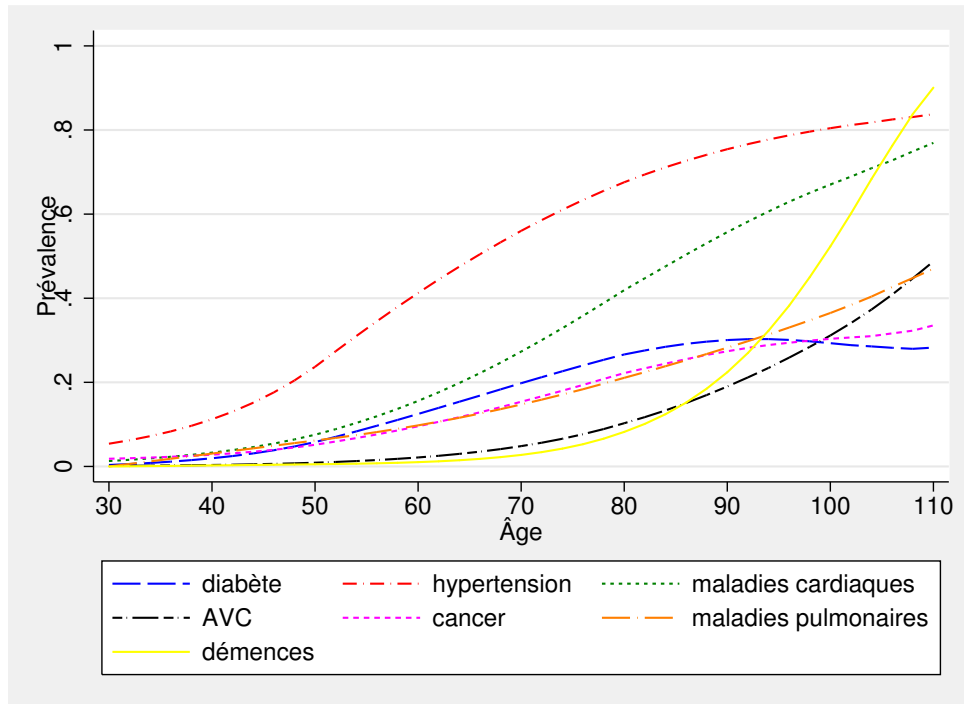


FIGURE 8.2 – Prévalences simulées des problèmes de santé dans la cohorte née en 1979-1980 — Scénario de référence

Ainsi, nous supposons que les risques de contracter les maladies sont indépendants alors qu'ils ne le sont pas. Dans un modèle markovien comme COMPAS, ceci mène à un gonflement des prévalences simulées puisque trop d'individus ont au moins une maladie et trop peu n'en ont aucune. Une des prochaines étapes du développement du modèle consistera à vérifier la plausibilité de ces trois explications.

Âges	Modèles	Années	Diabète	Hypertension	AVC	Cancer	Maladies cardiaques	Maladies pulmonaires
30 — 34 ans	COMPAS	2010 — 2014	0,59	6,28	0,17	1,96	1,60	0,61
	ENSP	1994 — 2011	0,93	3,37	0,24	0,85	1,32	4,79
50 — 54 ans	COMPAS	2030 — 2034	7,05	27,35	1,08	5,94	8,91	6,80
	ENSP	1994 — 2011	4,61	16,91	0,93	3,03	5,18	5,33
70 — 74 ans	COMPAS	2050 — 2054	21,20	58,51	5,68	16,67	30,04	15,89
	ENSP	1994 — 2011	14,48	44,95	5,59	8,71	21,08	10,73
90 — 94 ans	COMPAS	2070 — 2074	30,29	76,45	20,91	28,06	57,87	29,60
	ENSP	1994 — 2011	15,06	52,15	13,33	14,22	37,76	13,73

Tableau 8.1 – Comparaison des prévalences à différents âges (%) : COMPAS vs. ENSP

Afin d'illustrer l'évolution d'un facteur de risque en mode cohorte, la figure 8.3 présente celle de l'obésité. Dans l'ensemble, les proportions ne restent pas stables. La proportion d'individus atteints d'obésité des classes II et III augmente jusqu'à 55 ans environ, avant de diminuer. La prévalence de l'obésité de classe I diminue légèrement entre 30 et 40 ans, et atteint son sommet environ 5 ans plus tard que celle des classes II et III. Cependant, l'obésité de classe I reste toujours plus présente que celle

de classes II et III.

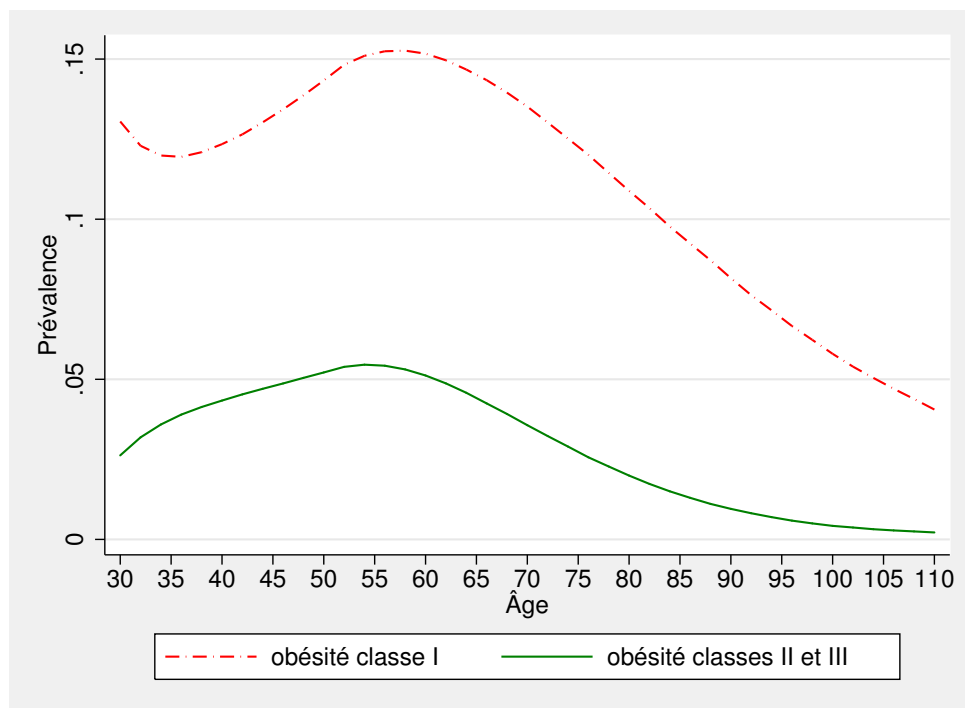


FIGURE 8.3 – Prévalence simulée de l’obésité dans la cohorte née en 1979-1980 — Scénario de référence

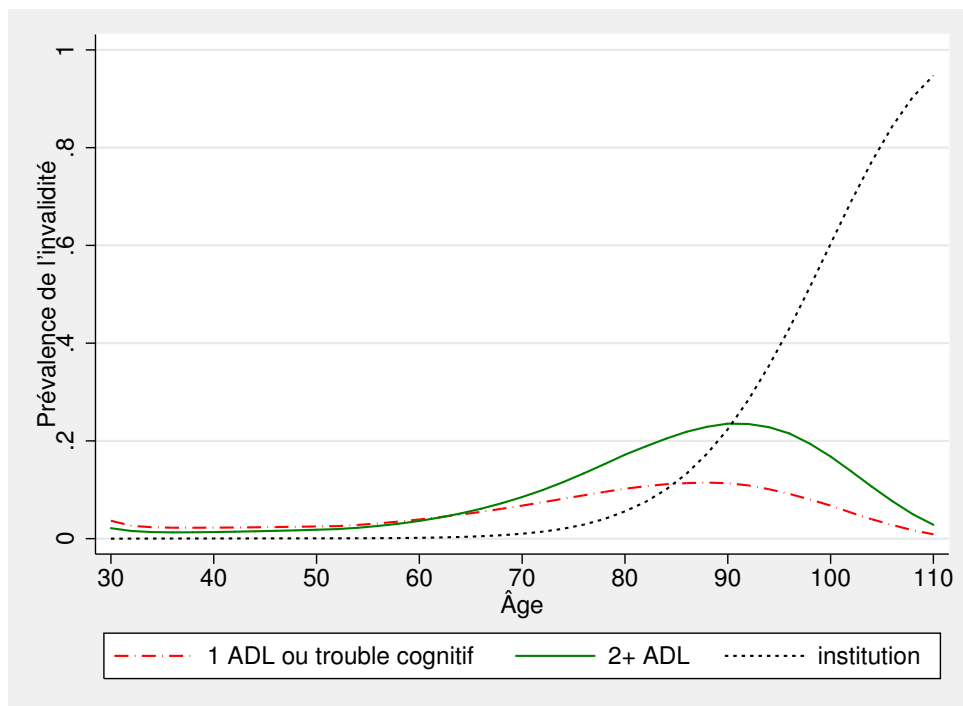


FIGURE 8.4 – Évolution de l’invalidité dans la cohorte née en 1979-1980 — Scénario de référence

La figure 8.4 montre l’évolution de l’invalidité entre 30 et 110 ans. La proportion des individus présentant un type quelconque d’invalidité augmente de façon importante entre 30 et 90 ans, passant de 6% à 57%. La proportion d’individus ayant une seule ADL ou incapacité cognitive reste relativement stable jusqu’à 56 ans puis augmente graduellement jusqu’à 92 ans ; par la suite, elle diminue jusqu’à 110 ans. La proportion des individus ayant deux ADL ou plus avec ou sans incapacité cognitive suit la même tendance. Stable jusqu’à 62 ans, la proportion des individus en institution augmente quant à elle rapidement entre 90 et 110 ans, soit une hausse de 73 points de pourcentage.

Les tableaux 8.2, 8.3 et 8.4 présentent les espérances de vie totale et en santé à 30 ans, respectivement, en fonction du niveau de scolarité et du sexe, du niveau de scolarité et du tabagisme puis du niveau de scolarité et du statut d’obésité. L’espérance de vie en santé capte le nombre d’années pendant lesquelles un individu peut espérer vivre avec moins de deux incapacités tout en demeurant en ménage privé (il n’est donc pas en institution). De façon générale, les espérances de vie augmentent avec le niveau d’éducation. D’après les résultats du tableau 8.2, les femmes vivent en moyenne plus longtemps que les hommes à niveau de scolarité équivalent. Les résultats du tableau 8.3 montrent que les non fumeurs vivent plus longtemps que les fumeurs ou les anciens fumeurs à niveau de scolarité comparable. Par exemple, les non fumeurs titulaires d’un diplôme universitaire vivent en moyenne 3,2 années de plus que leurs homologues fumeurs, et 2,9 années de plus que leurs semblables anciens fumeurs. Le tableau 8.4 montre qu’à niveau de scolarité équivalent, les espérances de vie ne suivent pas une tendance claire en fonction du statut d’obésité.

Niveau scolarité \ Sexe	Homme		Femme	
	Totale	En santé	Totale	En santé
Aucun diplôme	51,71	(45,89)	54,61	(45,68)
Diplôme d'études secondaires	54,59	(49,31)	57,71	(49,24)
Diplôme d'études collégiales	55,56	(50,55)	58,82	(50,40)
Diplôme universitaire	58,48	(54,45)	61,86	(54,80)

Tableau 8.2 – Espérance de vie à 30 ans en fonction du niveau de scolarité et du sexe

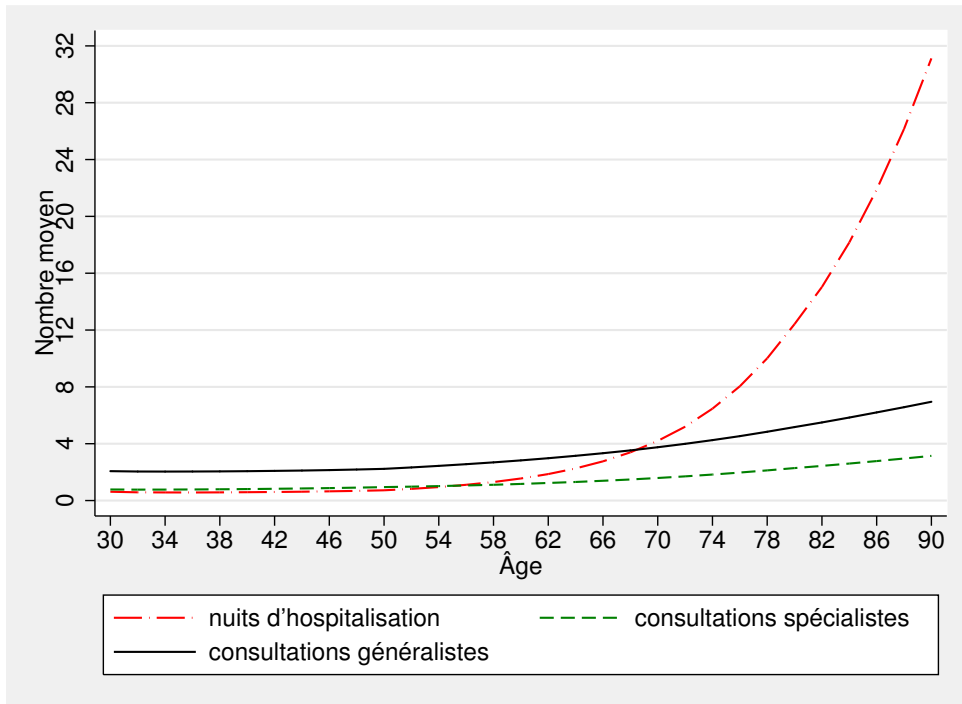
Niveau scolarité \ Tabagisme	Jamais fumé		Fumeur		Ancien fumeur	
	Totale	En santé	Totale	En santé	Totale	En santé
Aucun diplôme	55,22	(47,70)	51,64	(43,77)	52,00	(45,18)
Diplôme d'études secondaires	57,22	(50,24)	55,24	(48,46)	55,40	(48,81)
Diplôme d'études collégiales	58,72	(51,72)	55,81	(49,56)	56,25	(49,56)
Diplôme universitaire	61,92	(56,01)	58,68	(53,64)	59,05	(53,77)

Tableau 8.3 – Espérance de vie à 30 ans en fonction du niveau de scolarité et du statut de tabagisme

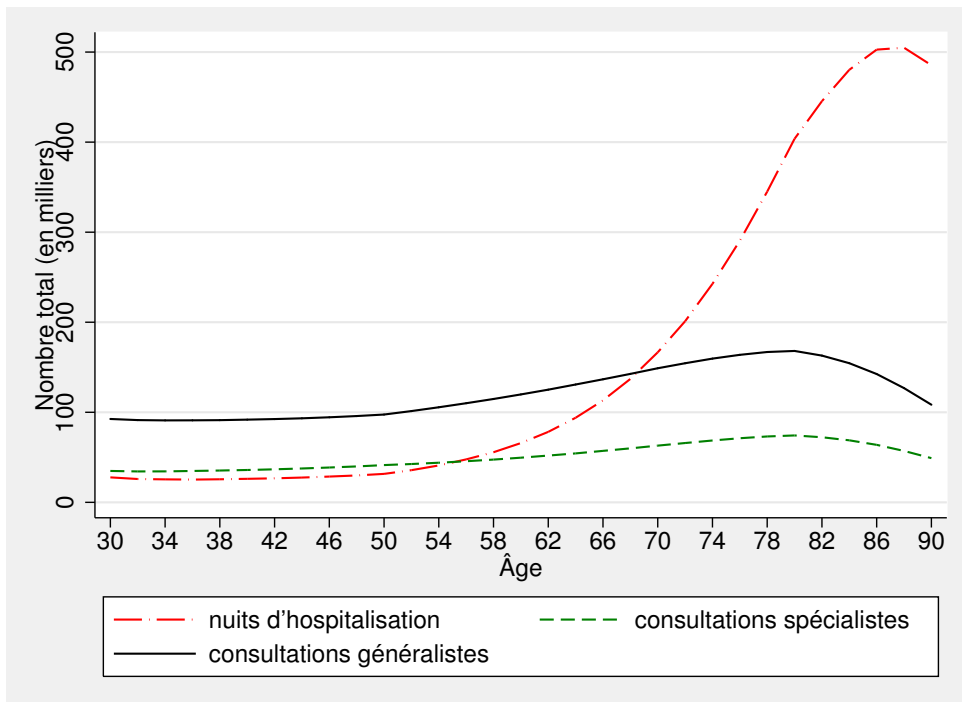
Niveau scolarité \ Statut d'obésité	Absence d'obésité		Classe I		Classes II et III	
	Totale	En santé	Totale	En santé	Totale	En santé
Aucun diplôme	53,17	(45,62)	53,61	(46,92)	51,60	(45,52)
Diplôme d'études secondaires	56,15	(49,38)	55,75	(48,63)	55,40	(49,52)
Diplôme d'études collégiales	57,13	(50,54)	56,95	(50,15)	56,89	(50,17)
Diplôme universitaire	60,00	(54,55)	60,72	(55,13)	58,85	(53,88)

Tableau 8.4 – Espérance de vie à 30 ans en fonction du niveau de scolarité et du statut d'obésité

La figure 8.5(a) présente l'évolution de l'utilisation moyenne des soins de santé entre 30 et 90 ans. À partir de 56 ans, le nombre moyen de nuits d'hospitalisation augmente rapidement de façon monotone et atteint 31 à l'âge de 90 ans. Entre 30 et 46 ans le nombre moyen de consultations chez les médecins généralistes est stable puis augmente uniformément. Stable aussi jusqu'à 46 ans, le nombre moyen de consultations chez les médecins spécialistes augmente cependant moins vite par la suite. À 0,78 à 30 ans et 0,88 à 46 ans, il atteint 3,15 à 90 ans. La figure 8.5(b) montre l'utilisation totale des soins de santé. Globalement, on observe les mêmes tendances qu'à la figure 8.5(a) jusqu'à 82 ans pour les consultations et 88 ans pour les nuits d'hospitalisation. À partir de ces âges, on observe toutefois un infléchissement, puisque la baisse du nombre d'individus toujours en vie fait plus que compenser la hausse de l'utilisation moyenne par individu.



(a) Utilisation moyenne



(b) Utilisation totale

FIGURE 8.5 – Utilisation des soins de santé par la cohorte née en 1979-1980 — Scénario de référence



La figure 8.6 donne la proportion d'individus dans la cohorte qui utilise chaque type de soins de santé entre 30 et 90 ans. De façon générale, toutes les proportions sont à la hausse entre 30 ans et 90 ans. En particulier, la proportion d'individus consommant des médicaments et celle des individus consultant un médecin généraliste croissent de 20 points de pourcentage. À partir de 40 ans, davantage d'individus consomment des médicaments comparativement à ceux consultant des généralistes. À 76 ans, la proportion d'individus passant des nuits à l'hôpital dépasse celle consultant des médecins spécialistes. Comme on pourrait s'y attendre, la proportion d'individus consultant des médecins généralistes reste toujours supérieure à celle consultant des médecins spécialistes. Stable jusqu'à 50 (resp. 60) ans, la proportion d'individus passant des nuits à l'hôpital (resp. institutionnalisés) augmente rapidement, soit une hausse de 73 (resp. 23) points de pourcentage à 90 ans par rapport à 30 ans.

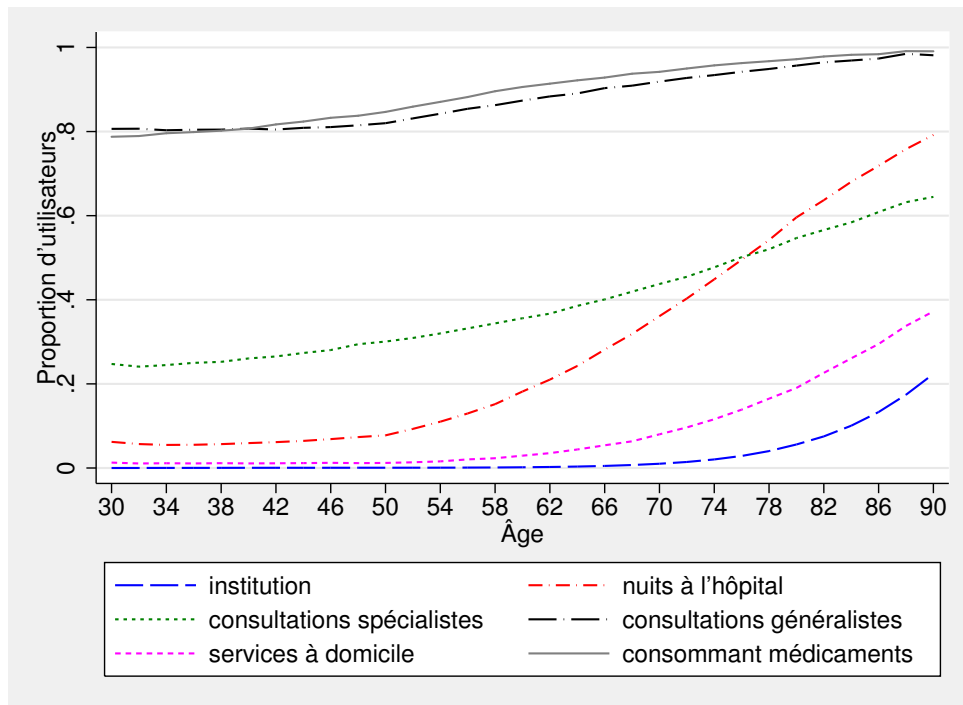


FIGURE 8.6 – Proportion d'utilisateurs de soins de santé dans la cohorte née en 1979-1980 — Scénario de référence

Les tableaux 8.5, 8.6 et 8.7 présentent le nombre total (à vie) de nuits d'hospitalisation, en fonction respectivement du niveau de scolarité et du sexe, du niveau de scolarité et du tabagisme puis du niveau de scolarité et du statut d'obésité. On observe que les individus ayant un diplôme universitaire sont hospitalisés moins de nuits. Le tableau 8.5 indique que les femmes passent plus de nuits à l'hôpital que les hommes. D'après le tableau 8.6, à niveau de scolarité équivalente, le nombre de nuits passé à l'hôpital par des individus n'ayant jamais fumé est nettement inférieur à celui des fumeurs ou des anciens fumeurs. À niveau de scolarité équivalent, le tableau 8.7 indique que le nombre total de nuits passées à l'hôpital en fonction du statut d'obésité ne suit pas une tendance uniforme.

Niveau scolarité \ Sexe	Sexe	
	Homme	Femme
Aucun diplôme	252,98	405,6
Diplôme d'études secondaires	264,8	264,8
Diplôme d'études collégiales	247,88	392,7
Diplôme universitaire	202,44	318,30

Tableau 8.5 – Nombre moyen viager de nuits à l'hôpital en fonction du niveau de scolarité et du sexe

Niveau scolarité \ Tabagisme	Tabagisme		
	Jamais fumé	Fumeur	Ancien fumeur
Aucun diplôme	309,80	340,36	346,68
Diplôme d'études secondaires	293,32	342,56	326,90
Diplôme d'études collégiales	299,22	321,18	332,88
Diplôme universitaire	248,88	262,56	262,64

Tableau 8.6 – Nombre moyen viager de nuits à l'hôpital en fonction du niveau de scolarité et du statut de tabagisme

Niveau scolarité \ Statut d'obésité	Statut d'obésité		
	Absence	Classe I	Classes II et III
Aucun diplôme	332,80	324,22	281,60
Diplôme d'études secondaires	308,50	377,26	321,14
Diplôme d'études collégiales	314,46	318,34	365,04
Diplôme universitaire	256,64	265,24	231,32

Tableau 8.7 – Nombre moyen viager de nuits à l'hôpital en fonction du niveau de scolarité et du statut d'obésité

Les tableaux 8.8, 8.9 et 8.10 récapitulent le nombre total (à vie) de consultations chez les généralistes, en fonction respectivement du niveau de scolarité et du sexe, du niveau de scolarité et du tabagisme puis du niveau de scolarité et du statut d'obésité. D'après le tableau 8.8, les femmes consultent plus de généralistes que les hommes. L'analyse du tableau 8.9 montre que les non fumeurs consultent, en moyenne, moins de généralistes que les fumeurs ou les anciens fumeurs à niveau de scolarité comparable. En l'absence d'obésité, le tableau 8.10 indique que les individus consultent, au total, moins de généralistes que ceux dont le statut d'obésité est de classe(s) I ou II-III.

Niveau scolarité \ Sexe	Sexe	
	Homme	Femme
Aucun diplôme	153,42	222,36
Diplôme d'études secondaires	149,14	210,88
Diplôme d'études collégiales	152,40	217,98
Diplôme universitaire	145,56	209,54

Tableau 8.8 – Nombre moyen viager de consultations chez les généralistes en fonction du niveau de scolarité et du sexe

Niveau scolarité \ Tabagisme	Tabagisme		
	Jamais fumé	Fumeur	Ancien fumeur
Aucun diplôme	187,36	187,24	190,44
Diplôme d'études secondaires	172,50	185,84	179,82
Diplôme d'études collégiales	178,40	186,74	186,56
Diplôme universitaire	171,52	177,30	179,70

Tableau 8.9 – Nombre moyen viager de consultations chez les généralistes en fonction du niveau de scolarité et du statut de tabagisme

Niveau scolarité \ Statut d'obésité	Statut d'obésité		
	Absence	Classe I	Classes II et III
Aucun diplôme	186,96	195,42	196,64
Diplôme d'études secondaires	176,54	189,32	191,92
Diplôme d'études collégiales	181,76	189,72	203,98
Diplôme universitaire	174,42	183,30	180,54

Tableau 8.10 – Nombre moyen viager de consultations chez les généralistes en fonction du niveau de scolarité et du statut d'obésité

Les tableaux 8.11, 8.12 et 8.13 résument le nombre de consultations total (à vie) chez les spécialistes, en fonction respectivement du niveau de scolarité et du sexe, du niveau de scolarité et du tabagisme puis du niveau de scolarité et du statut d'obésité. D'une façon générale, le nombre de consultations augmente avec le niveau de scolarité. Les femmes consultent plus de spécialistes que les hommes. D'après le tableau 8.12, les non fumeurs consultent moins de spécialistes que les fumeurs ou les anciens fumeurs à niveau de scolarité comparable. Le tableau 8.13 indique que les individus non obèses consultent généralement moins de spécialistes que ceux souffrant d'obésité de classe(s) I ou II-III.

Niveau scolarité \ Sexe	Sexe	
	Homme	Femme
Aucun diplôme	38,64	65,74
Diplôme d'études secondaires	55,44	89,16
Diplôme d'études collégiales	57,74	93,58
Diplôme universitaire	66,52	107,70

Tableau 8.11 – Nombre moyen viager de consultations chez les spécialistes en fonction du niveau de scolarité et du sexe

Niveau scolarité \ Tabagisme	Tabagisme		
	Jamais fumé	Fumeur	Ancien fumeur
Aucun diplôme	50,12	50,88	56,44
Diplôme d'études secondaires	68,18	74,58	73,00
Diplôme d'études collégiales	71,22	76,74	77,06
Diplôme universitaire	81,64	88,94	88,70

Tableau 8.12 – Nombre moyen viager de consultations chez les spécialistes en fonction du niveau de scolarité et du statut de tabagisme

Niveau scolarité \ Statut d'obésité	Statut d'obésité		
	Absence	Classe I	Classes II et III
Aucun diplôme	51,76	55,26	57,28
Diplôme d'études secondaires	70,34	78,88	72,54
Diplôme d'études collégiales	74,02	76,94	84,30
Diplôme universitaire	85,78	87,84	82,32

Tableau 8.13 – Nombre moyen viager de consultations chez les spécialistes en fonction du niveau de scolarité et du statut d'obésité

Les tableaux 8.14, 8.15 et 8.16 montrent le nombre d'années d'utilisation de services à domicile, en fonction respectivement du niveau de scolarité et du sexe, du niveau de scolarité et du tabagisme puis du niveau de scolarité et du statut d'obésité. L'analyse de ces trois tableaux suggère que le nombre d'années d'utilisation de services à domicile ne dépend pas du niveau de scolarité. Par ailleurs, les femmes reçoivent des services à domicile sur plus d'années que les hommes peu importe leur niveau de scolarité respectif. Cette observation pourrait être due au différentiel d'espérance de vie entre hommes et femmes. D'après le tableau 8.15, indépendamment du niveau de scolarité, les individus n'ayant jamais fumé recourent plus longtemps aux services à domicile que les fumeurs et les anciens fumeurs.

Niveau scolarité \ Sexe	Sexe	
	Homme	Femme
Aucun diplôme	3,34	5,14
Diplôme d'études secondaires	3,52	5,05
Diplôme d'études collégiales	3,48	5,26
Diplôme universitaire	3,32	5,15

Tableau 8.14 – Nombre d'années d'utilisation de services à domicile en fonction du niveau de scolarité et du sexe

Niveau scolarité \ Tabagisme	Tabagisme		
	Jamais fumé	Fumeur	Ancien fumeur
Aucun diplôme	4,42	3,98	4,28
Diplôme d'études secondaires	4,30	4,28	4,17
Diplôme d'études collégiales	4,44	4,26	4,23
Diplôme universitaire	4,40	4,03	4,05

Tableau 8.15 – Nombre moyen d'années d'utilisation de services à domicile en fonction du niveau de scolarité et du statut de tabagisme

Niveau scolarité \ Statut d'obésité	Statut d'obésité		
	Absence	Classe I	Classes II et III
Aucun diplôme	4,29	4,01	4,36
Diplôme d'études secondaires	4,20	4,59	4,01
Diplôme d'études collégiales	4,30	4,40	4,51
Diplôme universitaire	4,17	4,31	4,02

Tableau 8.16 – Nombre moyen d'années d'utilisation de services à domicile en fonction du niveau de scolarité et du statut d'obésité

Les tableaux 8.17, 8.18 et 8.19 récapitulent le nombre d'années de consommation de médicaments, en fonction respectivement du niveau de scolarité et du sexe, du niveau de scolarité et du tabagisme puis du niveau de scolarité et du statut d'obésité. On observe que la durée de la consommation de médicaments augmente avec le niveau de scolarité. L'analyse du tableau 8.17 montre qu'à un niveau de scolarité identique, les femmes consomment des médicaments durant plus d'années que les hommes — un effet qui pourrait ici encore être dû à leur espérance de vie plus grande. Le tableau 8.18 indique que les individus non fumeurs consomment des médicaments durant plus d'années que les fumeurs et les anciens fumeurs à niveau de scolarité équivalent. D'après le tableau 8.19, lorsque les individus sont titulaires d'un diplôme d'études secondaires ou post-secondaires, la durée de la consommation de médicaments des non obèses est inférieure à celle des individus souffrant d'obésité.

Niveau scolarité	Sexe	
	Homme	Femme
Aucun diplôme	40,09	46,36
Diplôme d'études secondaires	44,42	49,77
Diplôme d'études collégiales	45,77	51,34
Diplôme universitaire	48,88	54,96

Tableau 8.17 – Nombre moyen d'années d'utilisation de médicaments en fonction du niveau de scolarité et du sexe

Niveau scolarité	Tabagisme		
	Jamais fumé	Fumeur	Ancien fumeur
Aucun diplôme	44,21	42,27	42,98
Diplôme d'études secondaires	47,04	46,82	47,02
Diplôme d'études collégiales	48,81	47,95	48,30
Diplôme universitaire	52,26	51,08	51,72

Tableau 8.18 – Nombre moyen d'années d'utilisation de médicaments en fonction du niveau de scolarité et du tabagisme

Niveau scolarité	Statut d'obésité		
	Absence	Classe I	Classes II et III
Aucun diplôme	43,19	43,82	43,17
Diplôme d'études secondaires	46,89	47,36	47,80
Diplôme d'études collégiales	48,30	48,82	49,75
Diplôme universitaire	51,58	52,84	51,63

Tableau 8.19 – Nombre moyen d'années d'utilisation de médicaments en fonction du niveau de scolarité et du statut d'obésité

## 9 Résultats de base en mode population

Dans ce chapitre, nous présentons quelques résultats de base du modèle de microsimulation COMPAS en mode population. Ce mode incorpore la dynamique des populations ; il englobe donc en quelque sorte le mode cohorte, puisqu'il répète celui-ci à plusieurs reprises, de façon à la fois séquentielle (étape 1 du mode cohorte) et simultanée (étape 2 du mode cohorte). Sa mise en œuvre nécessite de définir, ici aussi, un scénario dit « de référence » semblable à celui du mode cohorte, qui prend en compte des statistiques officielles, incluant des projections sur la démographie (absentes du mode cohorte), le profil socio-démographique et l'état de santé de la population.

### 9.1 Scénario de référence

L'idée inhérente au scénario de référence est de représenter la situation telle qu'elle serait entre 2010 et 2050 en l'absence de changements dans l'effet et l'importance des déterminants des transitions vers différents états de santé, y compris les politiques gouvernementales. Dans la pratique, ce scénario sert également de base comparative à des scénarios alternatifs. Avant de définir notre scénario de référence, il paraît utile de faire un résumé succinct du fonctionnement de COMPAS.

#### 9.1.1 Un retour dans COMPAS

Une des particularités de COMPAS est de générer des projections sur la santé, entre 2010 et 2130, des individus âgés d'au moins 30 ans et résidant au Québec<sup>9</sup>. Étant donné que les problèmes de santé considérés dans le modèle sont peu courants chez les plus jeunes, l'exclusion de cette tranche de la population a peu d'effets sur la santé globale de la population québécoise. La microsimulation s'effectue ici en trois étapes :

1. Le *module d'initialisation* crée la population initiale du modèle. Représentative de la population du Québec âgée de 30 ans et plus en 2010, celle-ci possède plusieurs caractéristiques :
  - caractéristiques socio-démographiques : âge, année de naissance, sexe, statut d'immigration, niveau de scolarité ;
  - maladies : diabète, hypertension, maladies cardiaques, accidents vasculaires cérébraux (AVC), cancer, maladies pulmonaires, démences ;
  - facteurs de risque : tabagisme, obésité ;
  - invalidité : limitations d'activités, troubles cognitifs, institutionnalisation.
2. Le *module de transition* fait vieillir les individus d'une année, en calculant la probabilité de transition vers différents états. Ainsi, chaque individu peut développer une maladie ou une invalidité, perdre ou gagner du poids, commencer à ou cesser de fumer, entrer en institution ou mourir. Suite à cette transition, les individus décédés sont retirés de la population.
3. Le *module de renouvellement* ajoute à la population une nouvelle cohorte<sup>10</sup> d'individus âgés de

---

9. Sauf mention contraire, c'est ce groupe que désigne le terme « Québécois » dans la suite de ce chapitre.

10. Le terme « cohorte » désigne ici un groupe d'agents nés la même année et qui entrent dans le modèle au même moment.

30 et 31 ans. Les individus de la nouvelle cohorte ont, tout comme les individus de la population initiale, différentes caractéristiques.

Les étapes 2 et 3 sont répétées jusqu'en 2050, date d'entrée de la dernière cohorte de 30 – 31 ans. Après 2050, seule la deuxième étape est répétée chaque année. Toutefois, les individus simulés peuvent continuer de vieillir et de changer d'état de santé jusqu'en 2130, l'âge maximal permis dans le modèle étant de 110 ans. En d'autres termes, sachant qu'un individu ne peut vivre plus de 110 ans dans COMPAS, la microsimulation prend fin lorsque tous les individus sont décédés, soit au plus tard en 2130. Après chaque année de simulation, différentes statistiques sont calculées. Parmi celles-ci, on note la prévalence, dans la population, de chaque maladie et des différents facteurs de risque ; le nombre d'individus de chaque groupe d'âge au sein de la population ainsi que le taux de mortalité et l'espérance de vie ; et l'utilisation des différents types de soins de santé.

Nous explicitons ci-après les hypothèses servant de base à la construction du scénario de référence.

### **9.1.2 Hypothèses du scénario de référence**

Notre scénario de référence est construit sur la base de trois grandes hypothèses, portant sur (i) les caractéristiques des cohortes entrantes ; (ii) la mortalité ; et (iii) la migration.

#### *Hypothèses sur les caractéristiques des cohortes entrantes*

Afin de tenir compte dans le futur de certaines tendances observées au Québec entre 2000 et 2012, des hypothèses sont posées quant aux caractéristiques des cohortes entrantes. Il serait en effet peu approprié de supposer que, par exemple, les Québécois âgés de 30 ans en 2030 seront identiques à ceux qui avaient 30 ans en 2010. Les trois hypothèses que nous énonçons ici ont déjà été abordées jusqu'à un certain point au chapitre 5.



(a) *Hypothèses sur la prévalence de l'obésité*

Nous supposons que la prévalence de l'obésité à 30 et 31 ans augmente au sein de chaque nouvelle cohorte. Ainsi, à son arrivée dans le modèle, la cohorte de 2022 est plus touchée par l'obésité que celle de 2021, qui est elle-même plus touchée que celle de 2020. Toutefois, le risque de devenir obèse au sein de la population de plus de 30 ans, soit la probabilité de transition vers l'obésité, est le même pour chaque cohorte une fois que celle-ci est entrée dans le modèle. Nous supposons ainsi que la prévalence de l'obésité de classe I augmentera de 0,1% par année entre 2010 et 2020, alors que celle de classes II et III augmentera de 2% par année au cours de la même période. Entre 2020 et 2030, les taux de croissance sont coupés de moitié, impliquant que la prévalence de l'obésité de type I augmentera de 0,05% par année, alors que celle de classes II-III augmentera de 1% par année. Entre 2030 et 2040, les taux de croissance sont encore une fois coupés de moitié, puis ils demeurent constants à partir de 2040. Ces taux se basent sur le taux de croissance annuel moyen (TCAM) de l'obésité observé au Canada entre 2000 et 2012 chez les 25 à 34 ans ([Statistique Canada, 2000b, 2012d](#)).

(b) *Hypothèses sur le niveau d'éducation*

Étant donnée l'augmentation de la proportion d'individus ayant obtenu un diplôme d'études postsecondaires entre 2000 et 2012, le scénario de référence postule une augmentation annuelle moyenne de 2%, entre 2010 et 2020, de la proportion de Québécois qui, à 30 ans, ont un diplôme universitaire. Cette augmentation est par la suite coupée de moitié en 2020, puis de nouveau en 2030. L'augmentation de la proportion de diplômés du collégial est plus faible, soit de l'ordre de 0,3% par année entre 2010 et 2020. Similairement à ce qui a été fait pour la proportion d'individus avec des diplômes d'études universitaires, le taux de croissance est coupé de moitié en 2020, puis en 2030. Inversement, la proportion d'individus n'ayant pas de diplôme ou seulement un diplôme d'études secondaires chute entre 2010 et 2050. Ces hypothèses se basent sur des TCAM calculés à l'aide de l'Enquête sur la population active de 2000 et celle de 2012 ([Statistique Canada, 2000a, 2012c](#)).

(c) *Hypothèse sur la prévalence du tabagisme*

La proportion de fumeurs et d'anciens fumeurs est supposée constante au sein des différentes cohortes. Autrement dit, la proportion d'individus qui fument à 30 ans en 2040 sera la même qu'en 2030. Ceci s'explique par la relative stabilité observée au Québec depuis 2009 dans ces proportions ([Santé Canada, 2012](#)). La diminution de la prévalence du tabagisme enregistrée depuis 2000 semble principalement être le résultat d'une forte baisse au tout début de la décennie.

*Hypothèse sur la mortalité*

En sus de la tendance inhérente aux effets de composition (ou effets de cohorte) que comportent les projections de notre modèle, nous supposons une amélioration exogène de la mortalité au cours de la période de simulation. Celle-ci incorpore entre autres les effets des avancées technologiques sur le taux de mortalité général. L'espérance de vie a augmenté de manière régulière depuis 1971 ([Institut de la statistique du Québec, 2009](#)) et une des causes de cette amélioration, à tout le moins aux États-Unis, semble être le progrès technologique curatif en santé ([Cutler et al., 2006](#)). Il semble raisonnable de supposer qu'un tel phénomène s'est aussi produit et continue de se produire au Québec. Nous supposons que le taux de mortalité diminue au cours de la période de simulation au-delà de ce qui s'explique par la variation dans la prévalence des maladies et des facteurs de risque considérés. La réduction de la

mortalité utilisée dans le modèle est celle postulée par la RRQ pour la période allant de 2010 à 2050 ([Régie des rentes du Québec, 2013](#)). La RRQ estime que la baisse de la mortalité sera plus faible pour les femmes que pour les hommes, du moins dès l'âge de 20 ans. Chez les hommes, la diminution de la mortalité est particulièrement importante pour les 40 à 69 ans (réduction annuelle de la mortalité d'environ 2,5% entre 2010 et 2020 et de 2% entre 2010 et 2030), alors que pour les femmes on anticipe plutôt des effets entre 50 et 79 ans (réduction annuelle de la mortalité d'environ 1,5% entre 2010 et 2030).

### *Hypothèses sur la migration*

Afin de mieux refléter la réalité démographique du Québec, COMPAS incorpore les projections de solde migratoire net utilisées par la [Régie des rentes du Québec \(2013\)](#). À ce propos, tel qu'expliqué au chapitre 6, il est supposé que (i) le solde migratoire interprovincial est maintenu stable et correspond au niveau moyen observé au cours des 15 années précédant l'année 2012, et (ii) les politiques québécoises et canadiennes relatives à l'immigration — et en particulier les cibles d'immigration internationale du ministère de l'Immigration et des Communautés culturelles connues en décembre 2012 — restent stables.

## **9.2 Simulation et résultats**

Sur la base du scénario de référence construit précédemment, nous avons répété, 100 fois, les étapes 1 à 3 décrites à la sous-section 9.1.1. Tout comme dans le cas du mode cohorte, les résultats obtenus pour un plus grand nombre de répétitions sont similaires et ne sont donc pas présentés dans ce rapport. Afin de présenter ces résultats de base, une série de 9 figures est proposée.

La figure 9.1 présente les résultats du scénario de référence en termes de distribution de la population par groupe d'âge. COMPAS prédit que la population totale du Québec augmentera uniformément entre 2010 et 2038 puis se stabilisera autour de 9 millions entre 2038 et 2050 ; de façon plus précise, elle est de 8,95 millions en 2036 et de 9,03 millions en 2050. Ces statistiques sont similaires aux prévisions de l'ISQ et de la RRQ, soit une population du Québec atteignant 9 millions de personnes en 2050. Il convient de noter — à titre de rappel — que COMPAS prédit uniquement la population des 30 ans et plus. Afin d'obtenir la population totale, nous avons inclus, dans nos résultats, les projections de l'ISQ pour les moins de 30 ans. Par ailleurs, la figure 9.1 illustre également le vieillissement de la population québécoise : on note une hausse de la proportion que représentent les 70 ans et plus.

La figure 9.2 présente la part de la population des « 65 ans et plus » et des « 75 ans et plus » parmi les 30 et 110 ans. Comme précédemment, on constate une nette augmentation de la part de ces deux groupes d'âge dans la population (ici celle des 30 ans et plus).

Dans la figure 9.3(a), nous avons superposé les taux de mortalité projetés par COMPAS et ceux de la RRQ à différents âges. Pour une meilleure visualisation, nous avons opté pour l'échelle logarithmique. Comparativement aux taux de la RRQ, notre modèle, décrit au chapitre 6, projette des taux de mortalité plus élevés pour les 40 ans et plus. Cependant, pour les 30 ans, nous obtenons des résultats similaires à ceux de la RRQ. La figure 9.3(b) présente les résultats de l'espérance de vie périodique à différents âges. Globalement, on note une augmentation de l'espérance de vie. Celle-ci est strictement monotone pour les individus de 30 à 50 ans. En ce qui concerne les 60 à 80 ans, l'espérance de vie

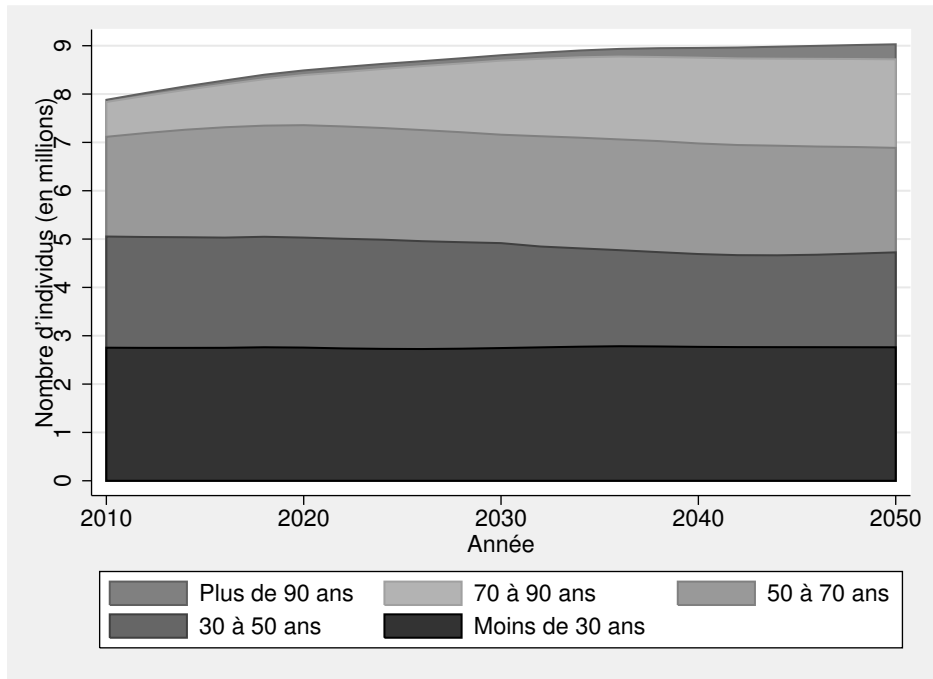


FIGURE 9.1 – Population du Québec — Scénario de référence

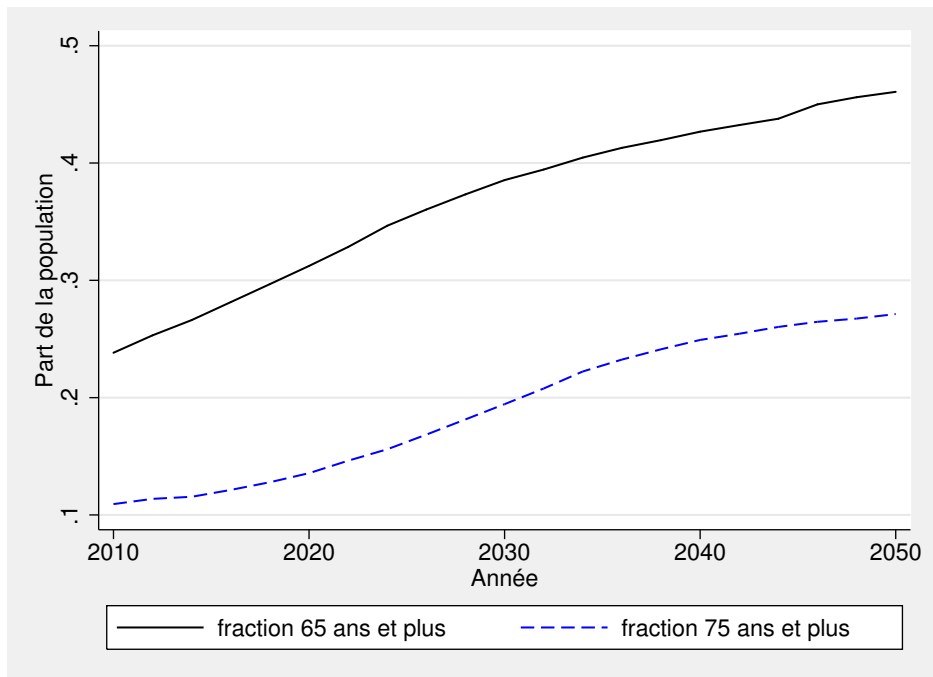


FIGURE 9.2 – Parts des « plus de 65 ans » et des « plus de 75 ans » dans la population québécoise de 30 à 110 ans — Scénario de référence

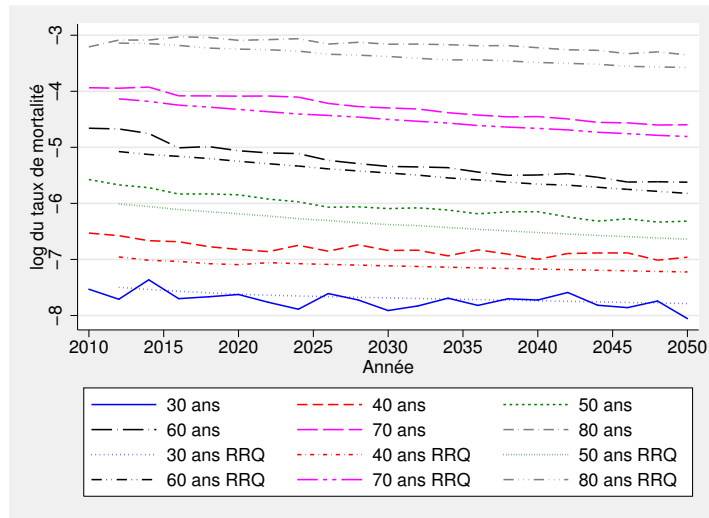
augmente pendant la majorité de la période de simulation, avec des minima en 2014 (resp. 2018) pour les 60 et 70 ans (resp. 80 ans). Si les conditions socio-médicales qui prévalent au Québec cette année-là demeurent identiques tout au long de leur vie, les individus âgés de 30 ans en 2050 vivront en moyenne 3,3 années de plus que ceux ayant eu le même âge en 2010. En ce qui concerne les 40, 50, 60, 70 et 80 ans, ces différences sont respectivement de 3,2, 3,1, 2,8, 2,2 et 1,3 années.

Enfin, la figure 9.3(c) récapitule les résultats de l'espérance de vie en santé périodique à différents âges. De prime abord, on observe que toutes les courbes de l'espérance de vie en santé périodique sont convexes. Il y a une croissance monotone de l'espérance de vie en santé périodique après 2014, 2016 et 2022 respectivement pour les individus de 30 à 50 ans, ceux de 60 ans et, enfin, pour les personnes âgées de 70 à 80 ans. En santé, les individus âgés de 30, 40, 50, 60, 70 et 80 ans en 2050 vivront en moyenne, respectivement, 3,0, 2,9, 2,7, 2,2, 1,4 et 1,2 années de plus que ceux ayant eu le même âge en 2010. Notons qu'après 2022, l'espérance de vie en santé périodique des 80 ans est presque stable : se situant à 6,2 ans en 2022, elle est de 6,8 ans en 2050, soit une augmentation de 0,6 année.

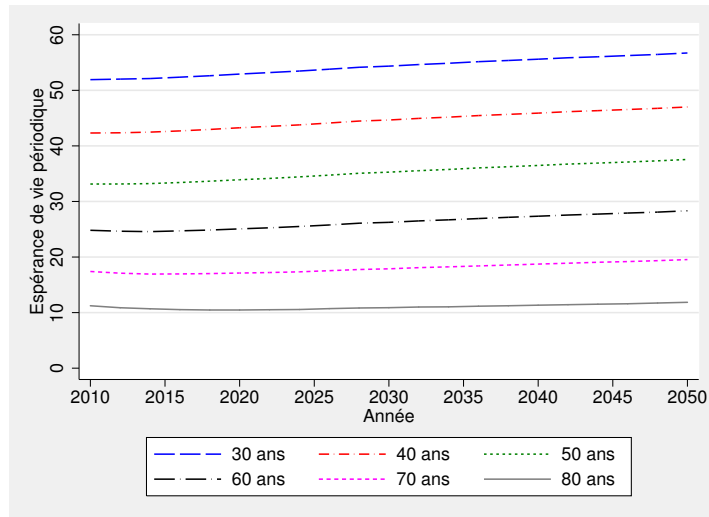
Les figures 9.4(a) et 9.4(b) présentent respectivement, à différents âges, les résultats de l'espérance de vie résiduelle périodique et ceux de l'espérance de vie résiduelle en santé périodique. La première observation montre que ces deux figures exhibent la même tendance que les figures 9.3(b) et 9.3(c). Ainsi, dans l'ensemble, il y aura une augmentation monotone des espérances de vie résiduelles (resp. espérances de vie résiduelles en santé) à partir de 2018 (resp. 2022) pour tous les âges. Cependant, la figure 9.4(a) indique une croissance monotone de l'espérance de vie résiduelle périodique des 30, 40 et 50 ans dès 2010. Les courbes de l'espérance de vie résiduelle périodique des 60, 70 et 80 ans ont une allure convexe avec des minima se situant entre 2010 et 2020. Par exemple, la courbe de l'espérance de vie résiduelle périodique des 60 et 70 (resp. 80) ans est strictement décroissante entre 2010 et 2014 (resp. 2018) et strictement croissante après 2014 (resp. 2018). Ceci indique que contrairement aux 30, 40 et 50 ans, l'augmentation de l'espérance de vie périodique des 60 et 70 (resp. 80) ans ne sera perceptible qu'après 2014 (resp. 2018). En outre, après 2036, l'espérance de vie résiduelle périodique des 60 ans dépassera celle des autres tranches d'âge. L'observation de la figure 9.4(b) montre que toutes les courbes de l'espérance de vie résiduelle en santé périodique sont convexes avec des minima entre 2010 et 2024. Dans l'ordre, la courbe des 30 ans atteint son minimum en premier, ensuite celle des 40 ans et celle des 80 ans en dernier. Autrement dit, l'espérance de vie résiduelle en santé périodique s'améliorera, en moyenne, d'abord chez les plus jeunes.

La figure 9.5(a) présente, à différents âges, la différence entre l'espérance de vie périodique et l'espérance de vie en santé périodique, ce qui correspond à « l'espérance de vie avec invalidité », tandis que la figure 9.5(b) présente leur différence résiduelle relativement à l'année 2010. De façon plus précise, si  $ex_{(age,annee)}$  et  $ehx_{(age,annee)}$  représentent, par âge et année, l'espérance de vie périodique et l'espérance de vie en santé périodique respectivement, alors leur différence résiduelle relative au temps  $t$ , notée  $rex_{(age,annee,t)}$ , est définie par

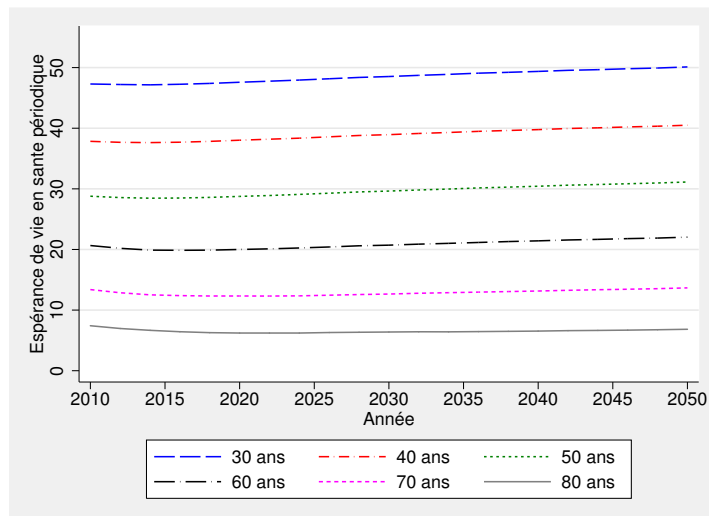
$$rex_{(age,annee,t)} = \frac{ex_{(age,annee)}}{ex_{(age,t)}} - \frac{ehx_{(age,annee)}}{ehx_{(age,t)}}.$$



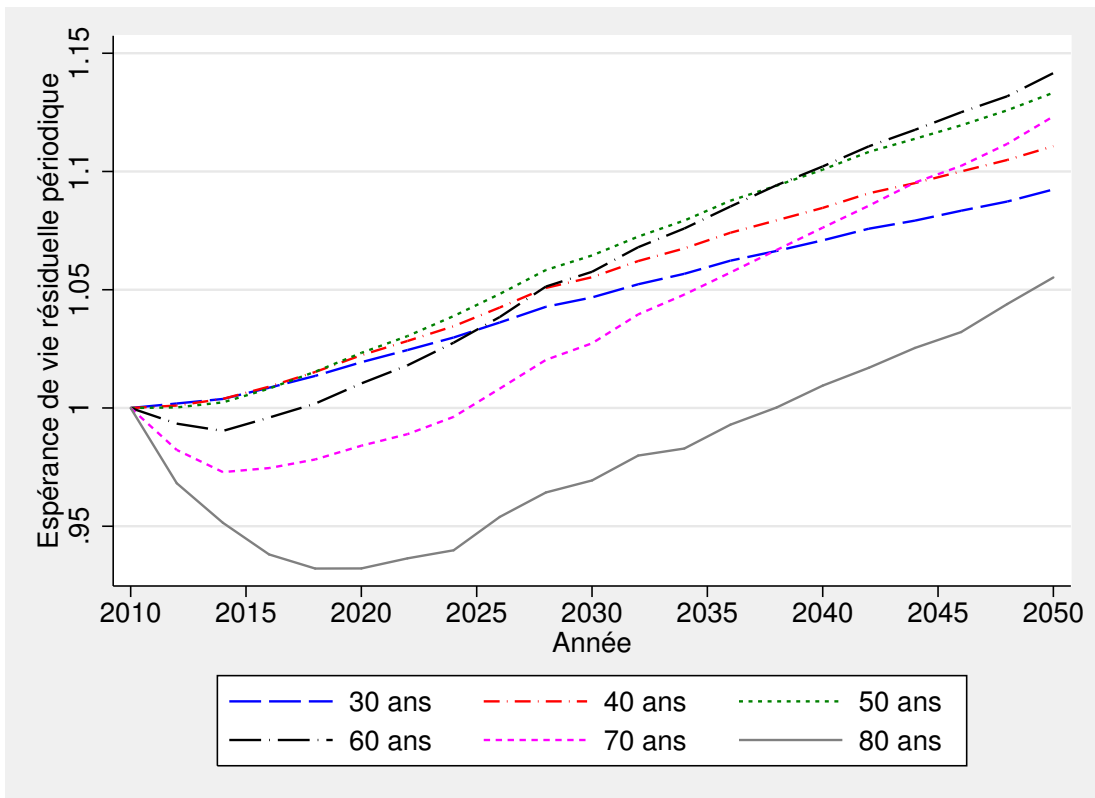
(a) Logarithme du taux de mortalité de COMPAS et de la RRQ à différents âges



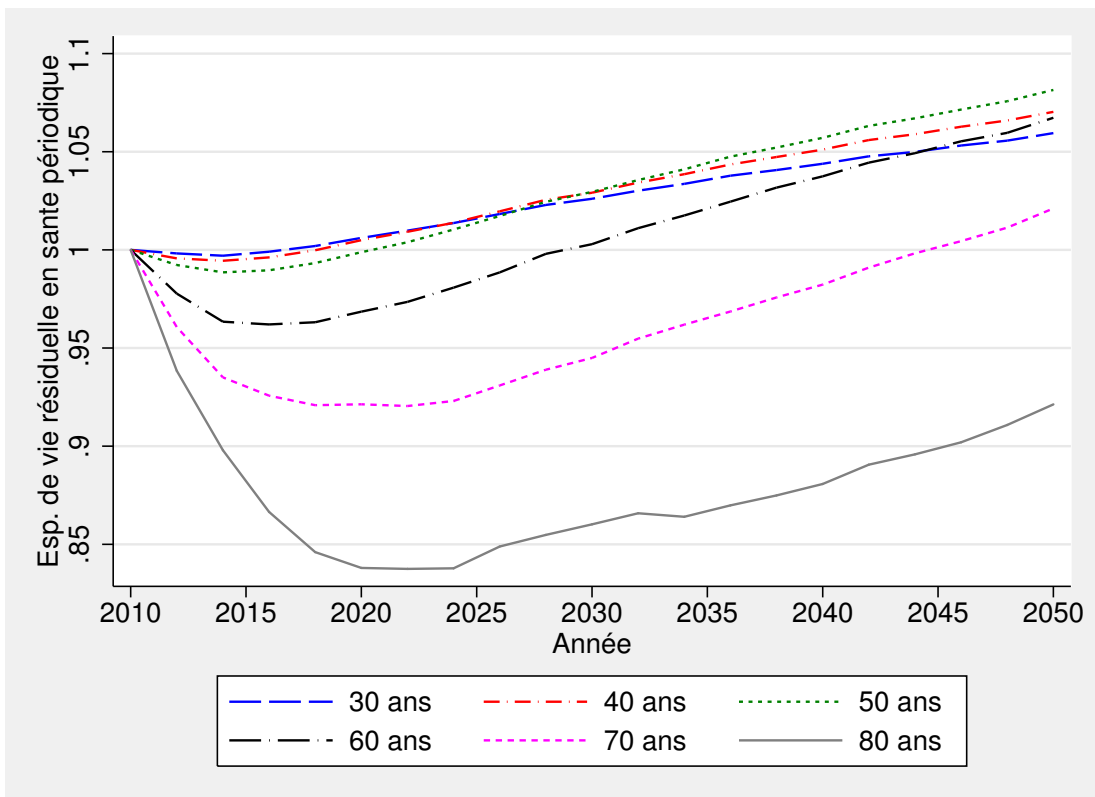
(b) Espérance de vie périodique à différents âges



(c) Espérance de vie en santé périodique à différents âges

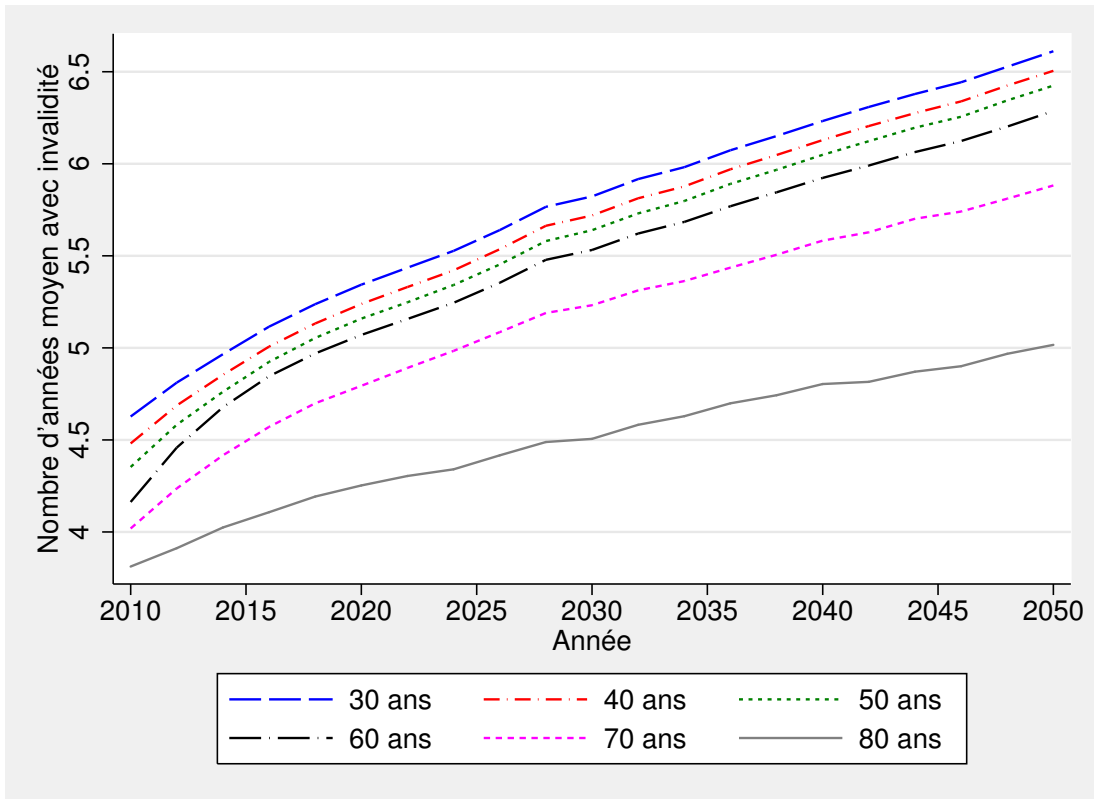


(a) Espérance de vie résiduelle périodique à différents âges

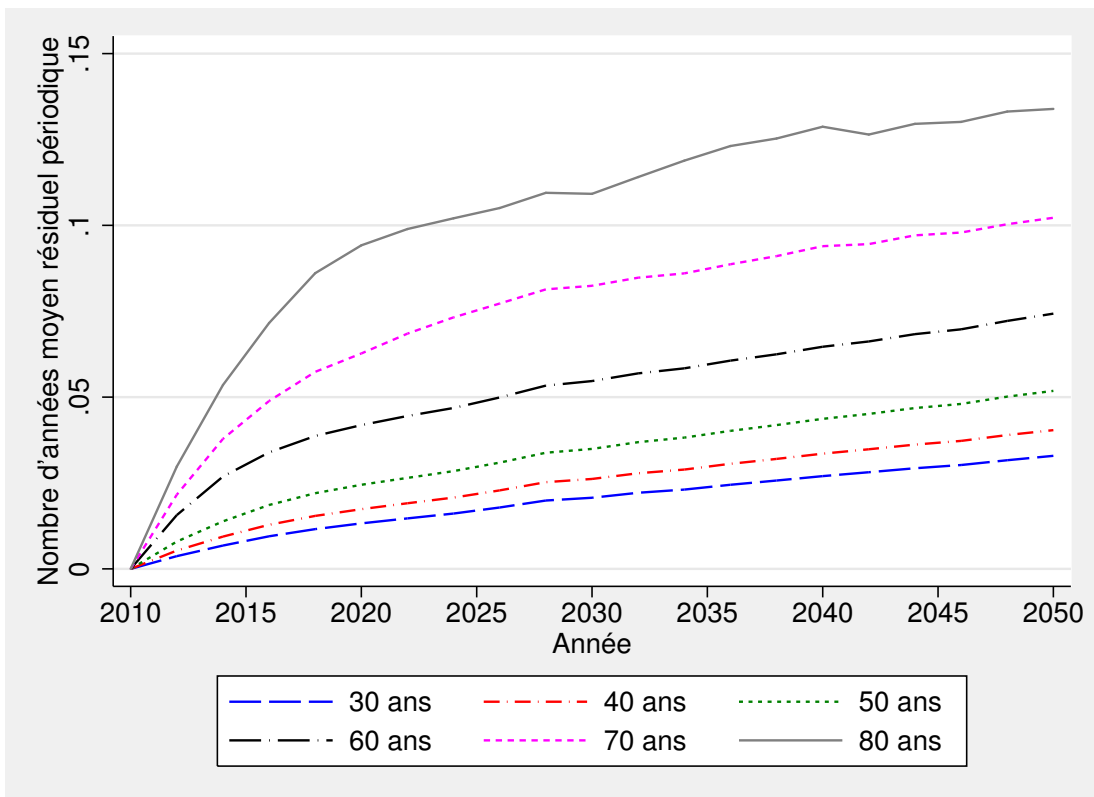


(b) Espérance de vie résiduelle en santé périodique à différents âges

FIGURE 9.4 – Espérance de vie résiduelle au Québec — Scénario de référence



(a) Différence entre l'espérance de vie périodique et l'espérance de vie en santé périodique



(b) Différence résiduelle relative à 2010<sup>85</sup> entre les espérances de vie périodique

FIGURE 9.5 – Différence entre les espérances de vie au Québec — Scénario de référence

D'allure similaire, les deux figures sont croissantes et monotones. D'après la figure 9.5(a), en 2050 les individus de 30, 40, 50, 60, 70 et 80 ans, respectivement, pourraient s'attendre à vivre en moyenne 2,0, 2,0, 2,1, 2,1, 1,9 et 1,2 années de plus avec une invalidité, comparativement à ceux ayant eu le même âge en 2010. Bien que la différence entre les espérances de vie périodiques (non résiduelles) soit décroissante des plus jeunes vers les moins jeunes, les 40 ans sont ceux qui enregistrent le plus fort gain entre 2010 et 2050.

La figure 9.6 présente les résultats en termes de prévalence des différentes maladies considérées dans le modèle. On observe une augmentation de la prévalence de l'ensemble des maladies au cours de la période étudiée. L'hypertension est le problème de santé qui touche le plus grand nombre au Québec en 2050, soit plus de 40% de la population. Bien que sa croissance soit élevée, elle demeure tout de même similaire à celle de la prévalence des maladies cardiaques au cours de la simulation. En effet, la proportion d'individus souffrant d'hypertension passe de 27,2% à 41,2% entre 2010 et 2050 ; la prévalence des maladies cardiaques augmente quant à elle de 7,5% à 21,1%. Le diabète et les maladies pulmonaires touchent eux aussi davantage de Québécois, leur prévalence augmentant respectivement de 7,1 et 7,5 points de pourcentage. La croissance de la prévalence des autres maladies est moins importante, soit autour de 4 points de pourcentage en 40 ans.

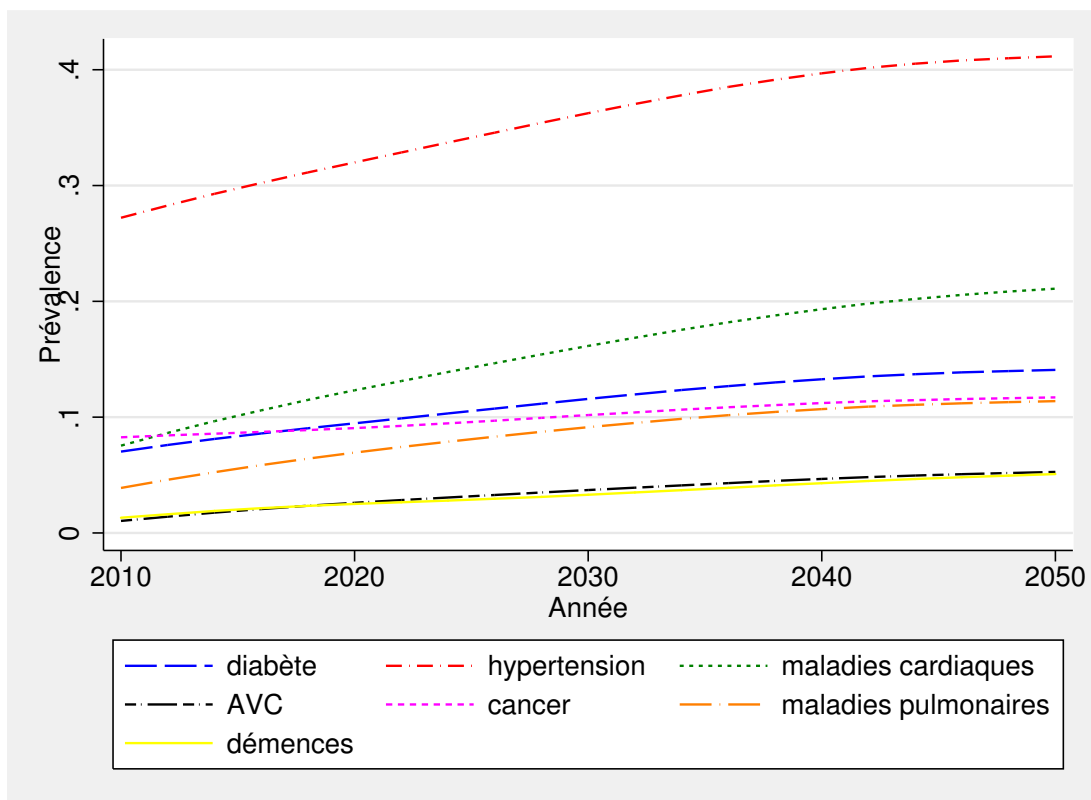


FIGURE 9.6 – Prévalence des maladies au Québec — Scénario de référence

La figure 9.7 donne l'évolution de la prévalence de l'invalidité chez les 30 ans et plus au Québec entre 2010 et 2050. On note une certaine diminution de la proportion d'individus ne présentant aucune incapacité au cours de la période étudiée : entre 2010 et 2050, celle-ci passe de 89,8% à 82,6% (non



illustrée). La proportion d'individus ayant une seule ADL ou trouble cognitif reste stable au cours de la période, avec une augmentation de 0,8 point de pourcentage. La proportion d'individus ayant deux ADL ou plus et celle d'individus en institution augmentent fortement, soit respectivement de 3,0 et 3,4 points de pourcentage entre 2010 et 2050. Dans ce dernier cas, la part de la population de 30 ans et plus en institution est multipliée par 4 durant la période.

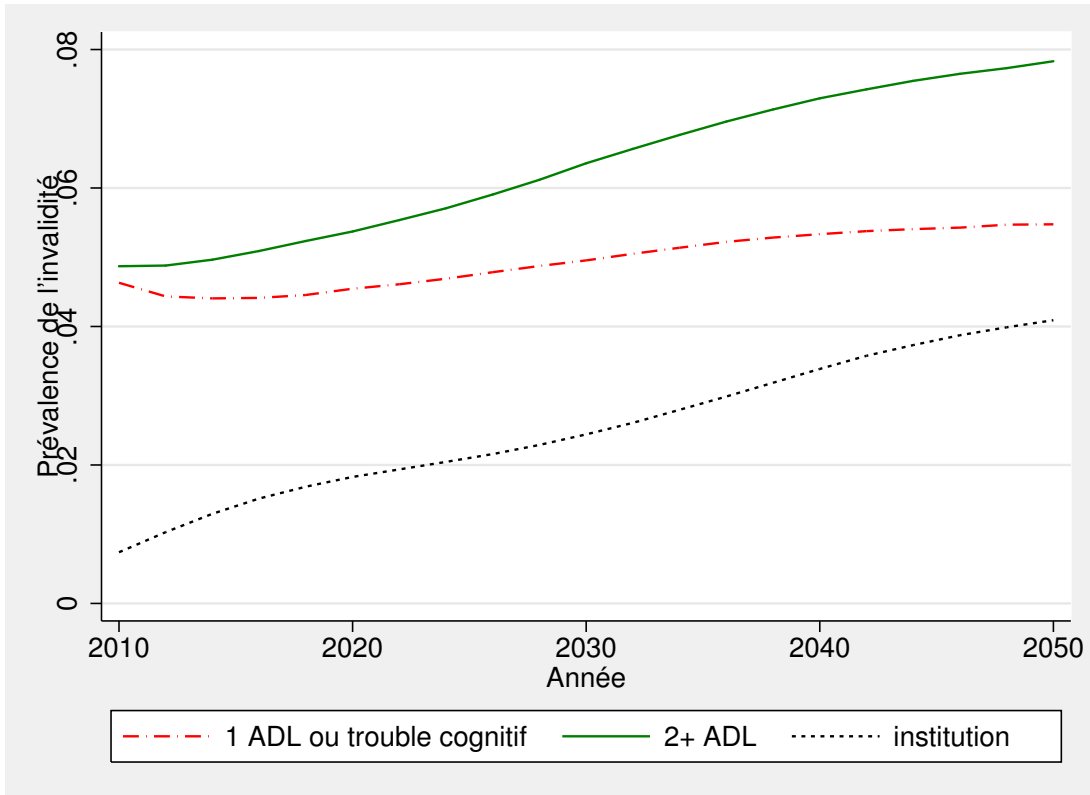


FIGURE 9.7 – Évolution de l'invalidité chez les 30 à 110 ans — Scénario de référence

La figure 9.8 présente l'évolution de l'obésité chez les 30 ans et plus au Québec. Dans l'ensemble, les proportions varient peu. Par exemple, les personnes obèses, qui représentent au total 17,2% de la population en 2010, seront 16,6% en 2050. Il faut toutefois rappeler que, tel qu'indiqué ailleurs dans ce rapport, la modélisation de l'obésité dans COMPAS devra être améliorée dans un avenir rapproché.

Enfin, la figure 9.9 montre l'utilisation des soins de santé des Québécois hors-institution entre 2010 et 2050. De façon générale, celle-ci augmente uniformément entre 2010 et 2050. Durant toute la période, les Québécois ont toujours moins recours à des consultations auprès de spécialistes qu'à des consultations auprès de généralistes ; l'écart va même toujours en s'agrandissant. Entre 2010 et 2020, le nombre de consultations chez les médecins généralistes demeure supérieur au nombre de nuits d'hospitalisation. Après 2020, le nombre de nuits d'hospitalisation dépasse le nombre de consultations chez les généralistes. Les figures 9.4(a) et 9.4(b) semblent indiquer qu'il s'agit d'une conséquence de l'augmentation de l'espérance de vie résiduelle (en santé ou non) périodique dès 2020 pour toutes les tranches d'âge.

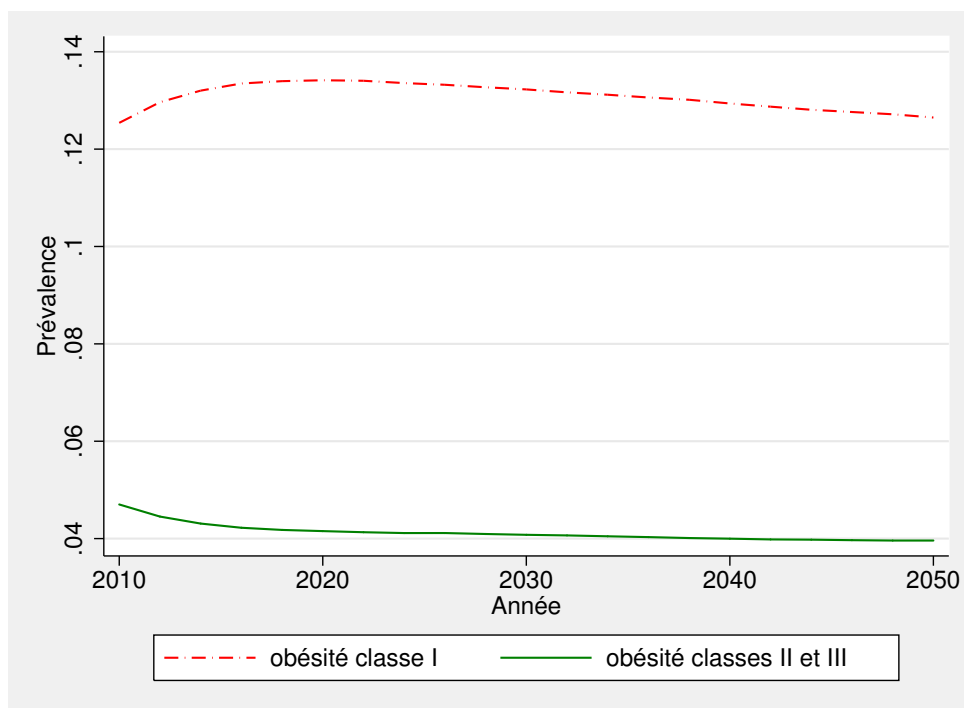


FIGURE 9.8 – Évolution de l’obésité chez les 30 à 110 ans — Scénario de référence

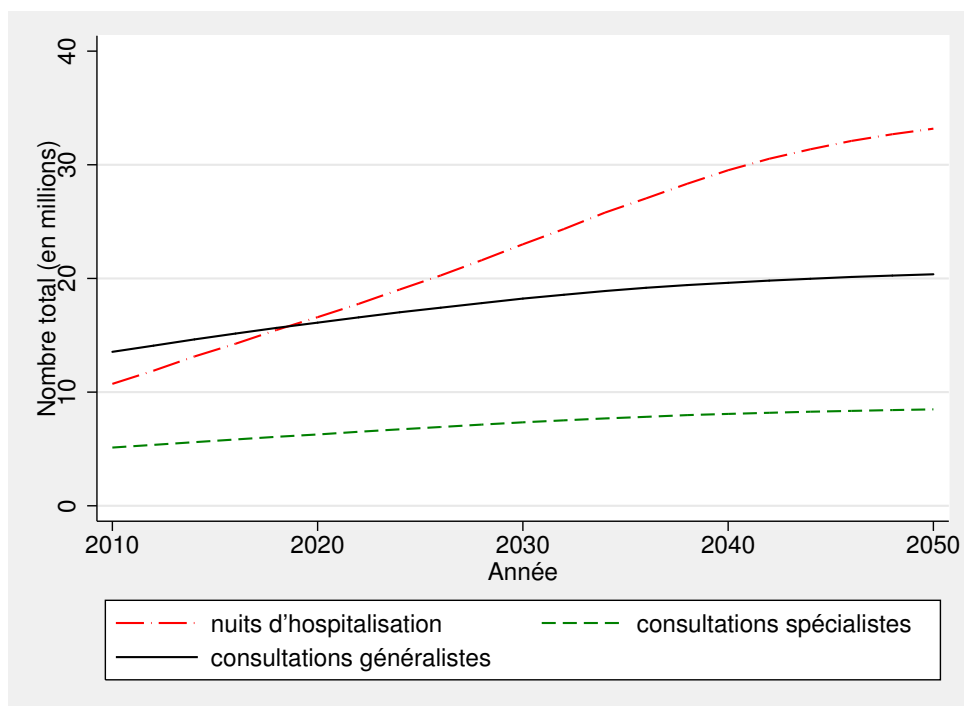


FIGURE 9.9 – Utilisation des soins de santé au Québec par les individus de 30 à 110 ans hors-institution — Scénario de référence

## Conclusion

La création du modèle de microsimulation COMPAS est récente : celle-ci a débuté en 2013, et son développement se poursuit depuis en profitant entre autres de l'expérience d'un membre de l'équipe ayant participé à la construction du *Future Elderly Model* américain. Jusqu'ici, l'équipe a jeté les bases d'une structure à la fois performante et souple, ce qui permet de raffiner progressivement les différents éléments déjà incorporés au modèle et d'en incorporer de nouveaux au fil des mois et années à venir.

COMPAS est principalement construit en FORTRAN, un langage simple qui permet une grande flexibilité et une grande adaptabilité aux différents besoins inhérents à la modélisation. La conception du modèle est modulaire, ce qui permet d'en modifier les paramètres et les composantes de façon séparée et sans nécessairement — ou immédiatement — affecter l'ensemble du modèle.

Pour le moment, COMPAS utilise principalement des données de l'ESCC de 2010 pour la construction de la population initiale du modèle, et les différents cycles de l'ENSP pour le calcul des différents modèles de transition servant à générer la dynamique d'utilisation des soins. Des données d'autres enquêtes, comme l'EPA, sont également utilisées pour construire différentes hypothèses et effectuer des imputations.

Le modèle se base en grande partie sur des hypothèses démographiques officielles, notamment celles de la RRQ, et utilise un modèle démographique standard pour générer la dynamique populationnelle. Certains ajustements de tendances sont par ailleurs effectués par l'équipe afin de tenir compte des tendances récentes dans l'évolution de certaines variables, notamment les facteurs de risque considérés (tabagisme, obésité). Ces ajustements peuvent être eux-mêmes modifiés de façon simple et à la discrétion de l'utilisateur.

Le comportement de COMPAS paraît adéquat dans l'état actuel de son développement. Bien que la mortalité semble surévaluée par rapport aux projections de la RRQ, par exemple, ces écarts ne sont pas inattendus. Ils pourront faire l'objet de recherche et d'améliorations à moyen terme mais les différents résultats du scénario de référence en matière de santé (état de santé, espérances de vie, utilisation des soins) semblent raisonnables et conformes aux attentes.

Le prochain développement majeur de COMPAS, déjà en cours, consistera à incorporer les coûts des soins de santé utilisés par les individus. Le développement suivant, lui aussi déjà en cours, consistera à arrimer COMPAS au modèle de microsimulation socio-économique mis sur pied antérieurement par l'équipe (voir [Clavet et al., 2012](#)). Ce dernier comportera également une modélisation améliorée de l'épargne et des revenus de retraite, ce qui ouvre la porte dans un avenir rapproché à des analyses fines qui pourront intégrer les dimensions état de santé, coûts de santé et comportements socio-économiques — y compris les comportements socio-démographiques, le travail, la consommation, l'épargne et la retraite — et l'impact de ces variables sur les conditions de vie des familles, la demande de services publics et les finances publiques.

Au final, COMPAS se démarque surtout des modèles de microsimulation dynamique existants par :

- son accent mis sur l'état de santé et les soins de santé, y compris les coûts associés à ces derniers ;
- son caractère d'abord québécois (le modèle pourra éventuellement être étendu à l'ensemble du Canada) ;

- son développement mené par le milieu universitaire, mais en collaboration avec différents partenaires non universitaires ;
- son arrimage éventuel avec un modèle de microsimulation socio-économique développé pour le Québec, qui inclut un simulateur fiscal détaillé tenant compte des particularités du régime fiscal et des principaux programmes de prestations en vigueur au Québec.

L'équipe espère que ce premier modèle de microsimulation québécois axé sur la santé, dans ses versions actuelle et futures, s'avérera un outil d'intérêt à la fois pour la communauté scientifique et pour l'analyse des politiques.

## Références

- AGENCE DE LA SANTÉ PUBLIQUE DU CANADA et INSTITUT CANADIEN D'INFORMATION SUR LA SANTÉ (2011). Obésité au Canada, Déterminants et facteurs contributifs. Rapport technique.
- AUDET, N. (2007). L'évolution de l'excès de poids chez les adultes québécois de 1990 à 2004 : mesures directes. Rapport technique, Institut de la statistique du Québec.
- BANQUE MONDIALE (2014). *Données : pays : Canada* [tableau sommaire]. Fichier de données.
- BOHNERT, N. (2013). Mortalité : aperçu, 2008 et 2009. Rapport technique, Statistique Canada.
- BROUHNS, N., DENUIT, M. et VERMUNT, J. K. (2002). A poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance : Mathematics and Economics*, 31(3):373 — 393.
- CAMERON, A. et TRIVEDI, P. (2005). *Microeconometrics : Methods and Application*. New York : Cambridge University Press.
- CLAVET, N.-J., DUCLOS, J.-Y., FORTIN, B. et MARCHAND, S. (2012). Le Québec, 2004-2030 : une analyse de micro-simulation. Rapport de projet CIRANO 2012RP-16.
- CLAVET, N.-J., DUCLOS, J.-Y., FORTIN, B., MARCHAND, S. et MICHAUD, P.-C. (2013). Les dépenses en santé du gouvernement du Québec, 2013-2030 : projections et déterminants. Cahier scientifique CIRANO 2013s-45.
- CUTLER, D., DEATON, A. et LLERAS-MUNEY, A. (2006). The determinants of mortality. *Journal of Economic perspectives*, 20(3):97–120.
- DAOULI, J., DAVILLAS, A., DERMOUSSIS, M. et GIANNAKOPOULOS, N. (2014). Obesity persistence and duration dependence : Evidence from a cohort of US adults. *Economics & Human Biology*, 12:pp.30–44.
- DÉCARIE, Y., BOISSONNEAULT, M. et LÉGARÉ, J. (2012). An inventory of Canadian microsimulation models. SEDAP Research Paper No. 298.
- FITZGERALD, J., GOTTSCHALK, P. et MOFFITT, R. (1998). An analysis of sample attrition in panel data : The Michigan panel study of income dynamics. *The Journal of Human Resources*, 33(2):251–299.

- FREES, E., GAO, J. et ROSENBERG, M. (2011). Predicting the frequency and amount of health care expenditures. *North American Actuarial Journal*, 15(3):377–382.
- GODBOUT, L., ST-CERNY, S., ARSENEAU, M., DAO, N. H. et FORTIN, P. (2014). La soutenabilité budgétaire des finances publiques du gouvernement du Québec,. Document de travail 2014/01 de la Chaire de recherche en fiscalité et en finances publiques de l'Université de Sherbrooke.
- GOLDMAN, D., CUTLER, D., ROWE, J., MICHAUD, P.-C., SULLIVAN, J., PENEVA, D. et OLSHANSKY, S. (2013). Substantial health and economic returns from delayed aging may warrant a new focus for medical research. *Health Affairs*, 32(10):pp.1698–1705.
- GOLDMAN, D., SHEKELLE, P., BHATTACHARYA, J., HURD, M., JOYCE, G., LAKDWALLA, D., MATSUI, D., NEWBERRY, S., PANIS, C. et SHANG, B. (2005). Health status and medical treatment of the future elderly. Rapport technique, RAND. Récupéré le 12 janvier 2014 de [http://www.rand.org/content/dam/rand/pubs/technical\\_reports/2005/RAND\\_TR169.pdf](http://www.rand.org/content/dam/rand/pubs/technical_reports/2005/RAND_TR169.pdf).
- GUTTERMAN, S. et VANDERHOOF, I. T. (1998). Forecasting changes in mortality : a search for a law of causes and effects. *North American Actuarial Journal*, 2(4):135–138.
- HUMAN MORTALITY DATABASE (2005). Complete data series. Rapport technique.
- INSTITUT CANADIEN D'INFORMATION SUR LA SANTÉ (2012). Tendances des dépenses nationales de santé, 1975-2012. Rapport technique.
- INSTITUT DE LA STATISTIQUE DU QUÉBEC (2009). Perspectives démographiques du Québec et des régions, 2006-2056. Rapport technique.
- INSTITUT NATIONAL DE SANTÉ PUBLIQUE DU QUÉBEC (2006). Les maladies du coeur et les maladies vasculaires cérébrales : prévalence, morbidité et mortalité au québec. Rapport technique, Canada.
- INSTITUT NATIONAL DE SANTÉ PUBLIQUE DU QUÉBEC (2011). Surveillance de l'hypertension au Québec : incidence, prévalence et mortalité. Rapport technique, Québec.
- INSTITUT NATIONAL DE SANTÉ PUBLIQUE DU QUÉBEC (2012). Surveillance du statut pondéral chez les adultes québécois : portrait et évolution de 1987 à 2010. Rapport technique.
- KAPTEYN, A., MICHAUD, P., SMITH, J. et SOEST, A. (2006). Effects of attrition and non-response in the health and retirement study. WR-407.
- LEE, R. D. et CARTER, L. R. (1992). Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association*, 87(419):pp. 659–671.
- MACDONALD, A. (1997). *The Second Actuarial Study of Mortality in Europe : The Companion Disk*. Groupe consultatif des associations d'actuaire des pays des Communautés Européennes.
- MCDONALD, A., CAIRNS, A., GWILT, P. et MILLER, K. (1998). An international comparison of recent trends in mortality. *British Actuarial Journal*, 4:3–141.
- MICHAUD, P.-C., LAKDAWALLA, D., GOLDMAN, D., ZHENG, Y. et GAILEY, A. (2012). The value of medical and pharmaceutical interventions for reducing obesity. *Journal of Health Economics*, 21(4):pp.630–640.

- NATIONAL RESEARCH COUNCIL (2012). Improving health care cost projections for the medicare population. Rapport technique.
- ORGANISATION DES NATIONS UNIES (2008). Perspectives relatives à la population mondiale : révision de 2008. Rapport technique, Division des Nations Unies pour la population.
- ORGANISATION MONDIALE DE LA SANTÉ (2006). BMI classification. Rapport technique, Nations Unies.
- PAYEUR, F. F. (2013). La mortalité et l'espérance de vie au Québec en 2012. *Institut de la statistique du Québec*, 26.
- RÉGIE DES RENTES DU QUÉBEC (2013). Évaluation actuarielle du Régime de rentes du Québec au 31 décembre 2012.
- RENSHAW, A. et HABERMAN, S. (2003a). Lee–Carter mortality forecasting : a parallel generalized linear modelling approach for england and wales mortality projections. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 52(1):119–137.
- RENSHAW, A., HABERMAN, S. et HATZOPOULOS, P. (1996). The modelling of recent mortality trends in United Kingdom male assured lives. *British Actuarial Journal*, 2:449–477.
- RENSHAW, A. E. et HABERMAN, S. (2003b). Lee–Carter mortality forecasting with age-specific enhancement. *Insurance : Mathematics and Economics*, 33(2):255–272.
- RENSHAW, A. E. et HABERMAN, S. (2003c). On the forecasting of mortality reduction factors. *Insurance : Mathematics and Economics*, 32(3):379–401.
- RÉSEAU DE RECHERCHE EN SANTÉ ET EN SÉCURITÉ DU TRAVAIL DU QUÉBEC (2014). Enquête nationale sur la santé de la population. Rapport technique.
- SANTÉ CANADA (2012). La prévalence du tabagisme 1999 — 2012. Rapport technique.
- STATISTIQUE CANADA (2000a). Enquête sur la population active, 2000. Fichier de données.
- STATISTIQUE CANADA (2000b). Enquête sur la santé dans les collectivités canadiennes, 2010, cycle 1.1, composante annuelle. Fichier de données.
- STATISTIQUE CANADA (2010). Enquête sur la santé dans les collectivités canadiennes, 2010, cycle 5.1, composante annuelle. Fichier de données.
- STATISTIQUE CANADA (2012a). Enquête nationale sur la santé de la population — volet ménages — longitudinal (ENSP). Rapport technique.
- STATISTIQUE CANADA (2012b). Enquête nationale sur la santé de la population, 1994-2010, cycles 1-9, volet ménages- longitudinal. Fichier de données.
- STATISTIQUE CANADA (2012c). Enquête sur la population active, 2012. Fichier de données.
- STATISTIQUE CANADA (2012d). Enquête sur la santé dans les collectivités canadiennes, 2012, composante annuelle. Fichier de données.

STATISTIQUE CANADA (2013). Tables de mortalité, Canada, provinces et territoires, 2007 à 2009. *Statistique Canada*, 84-537-X(003).

STATISTIQUE CANADA (2014a). Tableau 102-0218 -Espérance de vie, table de mortalité abrégée, selon le groupe d'âge et le sexe, moyenne de trois ans, Canada, provinces, territoires, régions sociosanitaires et groupes de régions homologues. Consulté le 8 août 2014.

STATISTIQUE CANADA (2014b). Tableau 102-0512 — Espérance de vie, à la naissance et à 65 ans, selon le sexe, Canada, provinces et territoires, annuel (années). Consulté le 8 août 2014.