



**Université de Montréal**

**La mortalité différentielle aux âges adultes et avancés selon le  
groupe linguistique au Québec : une étude de suivi sur la  
période 1991-2011**

Par :

**Cecilia Ah-kion**

Département de démographie  
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de Maîtrise ès sciences (M.Sc.)  
en Démographie

Avril 2019

Cecilia Ah-kion, 2019

**Université de Montréal**  
Faculté des études supérieures

Ce mémoire intitulé

**La mortalité différentielle aux âges adultes et avancés selon le groupe linguistique au  
Québec : une étude de suivi sur la période 1991-2011**

Présenté par :  
**Cecilia Ah-kion**

a été évalué par un jury composé des personnes suivantes :

Nadine Ouellette  
(Directrice de recherche)

Robert Bourbeau  
(Co-directeur de recherche)

Yves Carrière  
(Président - rapporteur)

Marc Termote  
(Membre du jury)

# Résumé

La situation linguistique distincte du Québec en Amérique du Nord nous amène à nous interroger sur le rôle que joue l'appartenance à un groupe linguistique sur les inégalités observées en matière de mortalité. La nouvelle base de données *Cohorte santé et environnement du recensement canadien de 1991 (CSERCan)*, issue du jumelage du recensement de 1991, des données de l'état civil et des fichiers administratifs de 1991 à 2011, offre une opportunité unique d'étudier avec plus de justesse la mortalité différentielle selon la langue au Québec. L'objectif principal de ce travail est de comparer les distributions des décès par âge, sexe, région et langue parlée le plus souvent à la maison, soit le français, l'anglais ou une langue autre. Nous dérivons ces distributions à partir des courbes de taux de mortalité correspondantes et lissées en utilisant une approche par P-splines. Nous calculons ensuite quatre indices sur la mortalité adulte et aux grands âges : l'espérance de vie à 25 ans ( $\hat{e}_{25}$ ), à 65 ans ( $\hat{e}_{65}$ ) et l'âge modal au décès ( $\widehat{M}$ ), ainsi que la mesure de dispersion associée à  $\widehat{M}$  ( $SD(\widehat{M} +)$ ). Nos résultats montrent qu'il existe bel et bien des différences significatives de mortalité aux âges adultes ( $\hat{e}_{25}$ ) selon l'appartenance linguistique au Québec, à l'exception des femmes francophones et allophones. Nous constatons cependant que les écarts d' $\hat{e}_{25}$  obtenus sont moins importants qu'annoncés dans la littérature. Les différences de mortalité observées aux âges adultes persistent chez les 65 ans ( $\hat{e}_{65}$ ) et plus. Nous n'observons toutefois pas d'écart significatif de mortalité au niveau de  $\widehat{M}$  ou de ( $SD(\widehat{M} +)$ ), mis à part chez les hommes allophones qui conservent leur avantage d'âge modal au décès. Enfin, pour les francophones et les allophones, les niveaux de mortalité ne sont pas les mêmes sur l'île de Montréal et dans le reste du Québec. Ces résultats offrent un éclairage inédit sur le sujet compte tenu de la plus grande fiabilité de la variable *langue* issue du recensement de 1991 dans la *CSERCan*, et des améliorations apportées au traitement de données et à l'analyse, qui incluent la prise en compte de l'émigration, la considération des autochtones et l'analyse par région.

**Mots-clés** : démographie linguistique, mortalité différentielle, âge modal au décès, dispersion des durées de vie, espérances de vie, lissage par P-splines, Québec, Île de Montréal, Reste du Québec, régions

# Abstract

Quebec's unique linguistic situation in North America leads us to question the effect of belonging to a linguistic group on observed inequalities in mortality. The new 1991 *Canadian Census Health and Environment Cohorts (CanCHEC)* database, which is derived from the pairing of the 1991 census, vital statistics and administrative files from 1991 to 2011, provides an unparalleled opportunity to study mortality differentials by language in Quebec with more accuracy. The main objective of this work is to compare the distributions of deaths by age and language most often spoken at home (i.e. French, English or another language). We begin by smoothing the age-specific mortality rate curves by sex, language and region using the P-splines approach. From the smoothed density functions, we derive four indicators of adult mortality and old-age mortality which are the life expectancy at age 25 ( $\hat{e}_{25}$ ), at age 65 ( $\hat{e}_{65}$ ), and the modal age (i.e. most common age) at death ( $\hat{M}$ ) as well as the dispersion measure associated to  $\hat{M}$  ( $SD(\hat{M} +)$ ). Our results show that in Quebec, there are indeed significant differences between linguistic groups in adult mortality ( $\hat{e}_{25}$ ), except for Francophone and Anglophone women. However, mortality gaps for the  $\hat{e}_{25}$  are less important than stated in the literature. Although these observed differences in adult mortality persist among those 65 years of age and older ( $\hat{e}_{65}$ ), they do not carry on among those who reach the modal age at death  $\hat{M}$ , with the exception of Allophone men who retain their mortality advantage in very old age. In addition, the dispersion of deaths beyond the modal age ( $SD(\hat{M} +)$ ) is not significantly different between linguistic groups. We also find different levels of mortality between the Island of Montreal and the rest of Quebec, for Francophones and Allophones. These results shed new light on the subject given the greater reliability of the linguistic variable from the 1991 Census in *CanCHEC*, and the improvements made in terms of data processing and analysis, which include the consideration of emigration and the indigenous population, and region-based analyses.

**Keywords:** linguistic demography, mortality differentials, modal age at death, lifespan disparity, life expectancies, P-splines smoothing, Quebec, Island of Montreal, Rest of Quebec, regions

# Table des matières

Résumé.....	iii
Abstract.....	iv
Table des matières.....	v
Liste des figures.....	viii
Liste des tableaux.....	xii
Dédicace.....	xiii
Remerciements.....	xiv
Introduction.....	1
Chapitre 1 : Problématique, recension des écrits et questions de recherche.....	3
1.1    Problématique.....	3
1.2    Recension des écrits.....	4
1.2.1    Différentiels de mortalité selon le groupe linguistique au Québec.....	4
1.2.2    Analyse comparative de la mortalité selon la source de données.....	6
1.2.3    Outils de mesure de la mortalité adulte et aux grands âges, résumant la tendance centrale et la dispersion des décès.....	7
1.2.4    Choix des indicateurs de mortalité.....	10
1.3    Compte rendu critique des indicateurs linguistiques et spatiaux.....	13
1.3.1    Choix de l'indicateur linguistique.....	13
1.3.2    Choix des catégories linguistiques.....	15
1.3.3    Choix du découpage territorial.....	15
1.4    Questions de recherche.....	16
Chapitre 2 : Sources de données et méthodologie.....	18
2.1    Sources de données.....	18
2.1.1    Cohorte Santé et Environnement du Recensement Canadien de 1991.....	18
2.1.2    Base de données sur la longévité canadienne (BDLC).....	24
2.2    Méthodologie.....	25

2.2.1	Traitement des données.....	26
2.2.2	Taux de mortalité et force de mortalité.....	34
2.2.3	Modèle de régression de Poisson.....	39
2.2.4	Lissage par P-splines.....	40
2.2.5	Les B-splines et les P-splines.....	40
2.2.6	Fonctions de survie et de densité.....	41
2.2.7	Construction des intervalles de confiance.....	42
Chapitre 3 : Présentation et interprétation des résultats.....		43
3.1	Validation de la méthodologie.....	43
3.1.1	Les taux spécifiques de mortalité.....	44
3.1.2	Validation du lissage par P-splines.....	46
3.2	Mortalité différentielle au Québec.....	48
3.2.1	Mortalité différentielle aux âges adultes.....	50
3.2.2	Mortalité différentielle aux grands âges.....	52
3.3	Mortalité différentielle selon la région.....	57
3.3.1	Différences régionales chez les femmes.....	58
3.3.2	Différences régionales chez les hommes.....	61
Discussion et conclusion.....		65
Validation et actualisation de la recherche.....		65
Mortalité différentielle aux grands âges.....		70
Profils sociodémographiques et économiques.....		71
Limites.....		73
Conclusion.....		75
ANNEXE A.....		77
Catégories linguistiques.....		77
ANNEXE B.....		79
Émigration.....		79
ANNEXE C.....		80
Validation de la CSERCan avec la cohorte BDLC.....		80

ANNEXE D .....	81
Résultats pour la province du Québec .....	81
ANNEXE E.....	87
Résultats pour l'île de Montréal et le reste du Québec .....	87
ANNEXE F.....	96
Validation de e25 et profils des groupes linguistiques.....	96
Bibliographie.....	101



# Liste des figures

Figure 1.1 Répartition des durées de vie féminines avant et après un siècle de transition en Finlande.....	8
Figure 1.2 Représentation de l'âge modal au décès (M) et de l'écart type, relatif à M, des durées de vie situées au-delà de M ( $SD(M^+)$ ), femmes, Japon, 2000-2004 .....	9
Figure 1.3 Durées modales et moyenne ( $e_0$ ) de la vie selon le sexe en Finlande, 1750 à 1995	10
Figure 2.1 Effectifs estimés par catégorie linguistique, CSERCan, Québec, 1991 à 2011 .....	28
Figure 2.2 Répartition des membres de la CSERCan selon leur dernière année de déclaration de revenus, Québec, 1991 à 2010 .....	32
Figure 2.3 Nombre annuel d'émigrants observé (1995-2005) et projeté (2006-2011), CSERCan, Québec .....	33
Figure 2.4 Estimations du nombre annuel d'émigrants, CSERCan, Québec, 1991 à 2011 .....	34
Figure 2.5 Extrait de l'univers d'analyse des données de la CSERCan, Québec, 1991 à 2011	35
Figure 2.6 Calcul des années vécues selon les quatre cas possibles.....	37
Figure 3.1 Comparaison des taux de mortalité selon l'âge et le sexe des cohortes BDLC et CSERCan avec et sans émigration, Québec, 1991 à 2011 .....	45
Figure 3.2 Comparaison entre les proportions des décès tirées de la table de mortalité (points et lignes superposés) et les fonctions de densité lissées (lignes pleines) résultant de l'approche par P-splines, selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011 .....	47
Figure 3.3 Fonction de densité lissée selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011 .....	49
Figure 3.4 Fonction de survie lissée selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011 .....	50
Figure 3.5 Estimations de l'espérance de vie à 25 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011.....	51
Figure 3.6 Estimations de l'espérance de vie à 65 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011.....	53
Figure 3.7 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011 .....	54

Figure 3.8 Estimations de l'écart-type au-delà du mode et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011 .....	56
Figure 3.9 Distribution (en %) des membres de la CSERCan par sexe, groupe linguistique et région, Île de Montréal et Reste du Québec, 1991 à 2011 .....	57
Figure 3.10 Fonction de densité lissée pour les femmes selon la région et le groupe linguistique, CSERCan, 1991 à 2011 .....	58
Figure 3.11 Estimations de l'espérance de vie à 25 ans (a) et 65 ans (b), et des intervalles de confiance bootstrap à 95% pour les femmes selon la région et le groupe linguistique, CSERCan, 1991 à 2011 .....	60
Figure 3.12 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% pour les femmes selon la région et le groupe linguistique, CSERCan, 1991 à 2011.....	61
Figure 3.13 Fonction de densité lissée pour les hommes selon la région et le groupe linguistique, CSERCan, 1991 à 2011 .....	62
Figure 3.14 Estimations de l'espérance de vie à 25 ans (a) et 65 ans (b), et des intervalles de confiance bootstrap à 95% pour les hommes selon la région et le groupe linguistique, CSERCan, 1991 à 2011 .....	63
Figure 3.15 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% pour les hommes selon la région et le groupe linguistique, CSERCan, 1991 à 2011.....	64
Figure 3.16 Fonction de densité lissée pour les hommes selon le groupe linguistique, CSERCan, Reste du Québec, 1991 à 2011 .....	64
Figure C.1 Proportion des décès selon l'âge et le sexe au Québec: Comparaison de la CSERCan et la cohorte BDLC, 1991 à 2011 .....	80
Figure D.1 Comparaison entre les proportions des décès tirées de la table de mortalité et la fonction de densité lissée par P-splines selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011.....	81
Figure D.2 Fonction de densité lissée selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011.....	82
Figure D.3 Fonction de survie lissée selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011.....	83
Figure D.4 Estimations de l'espérance de vie à 25 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011 .....	84

Figure D.5 Estimations de l'espérance de vie à 65 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011 .....	84
Figure D.6 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011 .....	85
Figure D.7 Estimations de l'écart-type au-delà du mode et des intervalles de confiance bootstrap à 95% selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011 .....	86
Figure E.1 Fonction de densité lissée selon le groupe linguistique (3) et le sexe, CSERCan, Île de Montréal, 1991 à 2011 .....	87
Figure E.2 Fonction de densité lissée selon le groupe linguistique (4) et le sexe, CSERCan, Reste du Québec, 1991 à 2011.....	88
Figure E.3 Fonction de survie lissée selon le groupe linguistique (4) et le sexe, CSERCan, Île de Montréal, 1991 à 2011 .....	89
Figure E.4 Fonction de survie lissée selon le groupe linguistique (4) et le sexe, CSERCan, Reste du Québec, 1991 à 2011.....	90
Figure E.5 Estimations de l'écart-type au-delà du mode et des intervalles de confiance bootstrap à 95% selon la région, le groupe linguistique et le sexe, CSERCan, 1991 à 2011 .....	91
Figure E.6 Estimations de l'espérance de vie à 25 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011 .....	91
Figure E.7 Estimations de l'espérance de vie à 65 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011 .....	92
Figure E.8 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011 .....	93
Figure E.9 Estimations de l'espérance de vie à 25 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011 .....	93
Figure E.10 Estimations de l'espérance de vie à 65 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011 .....	94
Figure E.11 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011 .....	95
Figure F.1 Comparaison de $e_{25}$ selon la source de données et le sexe, Québec.....	96
Figure F.2 Lieu de naissance selon le groupe linguistique et le sexe, CSERCan, Québec , 1991 à 2011 .....	97

Figure F.3 Plus haut niveau de scolarité atteint selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011.....	97
Figure F.4 État matrimonial selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011.....	98
Figure F.5 Plus haut niveau de scolarité atteint selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011 .....	98
Figure F.6 État matrimonial selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011 .....	99
Figure F.7 Plus haut niveau de scolarité atteint selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011 .....	99
Figure F.8 État matrimonial selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011 .....	100

# Liste des tableaux

Tableau I. Effectifs estimés par sexe, région et catégorie linguistique, CSERCan, Québec, 1991 à 2011 .....	29
Tableau II. Estimations de l'espérance de vie à l'âge de 25 ans par sexe et groupe linguistique selon trois sources de données, Québec .....	66
Tableau III. Écarts d'espérance de vie à 25 ans entre les sexes selon plusieurs sources de données, Québec .....	68
Tableau A.I : Langues incluses dans la catégorie linguistique allochtone, CSERCan .....	77
Tableau A.II : Langues incluses dans la catégorie linguistique autochtone, CSERCan .....	78
Tableau B.I : Estimation d'émigrants, CSERCan, Québec, 1991 à 2011 .....	79

# Dédicace

À mes parents, Jean et Denise,  
à ma sœur Caroline et mon frère Christopher,  
et à mes neveux, Evan et Charlie.

# Remerciements

Ce mémoire n'aurait pas été possible sans le soutien indescriptible et la précieuse confiance que m'ont accordés ma directrice de recherche Nadine Ouellette et mon co-directeur Robert Bourbeau. Je me considère très chanceuse d'avoir été leur auxiliaire de recherche et d'enseignement, et de faire partie de l'équipe de recherche *Mortalité et longévité* du Département de démographie, qui est à la fois dynamique et inclusive.

D'abord, j'ai eu la chance énorme d'être dirigée par Nadine Ouellette, une mentor hors pair que j'admire beaucoup et de qui j'ai énormément appris. Sa passion contagieuse pour la démographie ainsi que sa grande minutie et disponibilité ont incontestablement contribué à bonifier ce mémoire. Merci infiniment pour tes précieux conseils, ton temps, tes encouragements et ton écoute attentive.

Je souhaite aussi exprimer mon immense gratitude envers mon co-directeur Robert Bourbeau qui m'a offert un encadrement et un soutien inestimables durant mon parcours à la maîtrise. Ses encouragements et ses conseils ont été une source additionnelle de motivation tout au long de la réalisation de ce mémoire.

Un grand merci également au Centre Interuniversitaire québécois de statistiques sociales (CIQSS) de m'avoir octroyé l'accès aux données de la CSERCan et aux bourses étudiantes. De même, je tiens à remercier mes directeurs de recherche et la Faculté des études supérieures et postdoctorales de l'Université de Montréal d'avoir soutenu financièrement ce mémoire.

D'autre part, je suis très reconnaissante envers mes chers collègues de l'équipe de recherche *Mortalité et longévité* (Lorena, Anthony, Marilyn, Ahmed, Linh) et du Département de démographie, qui ont rendu cette expérience universitaire particulièrement agréable et stimulante. Un merci tout spécial à Abigail, Yacine, Lambert, Guillaume et Allyx. Je ne m'attendais pas à rencontrer des gens aussi extraordinaires qui sont devenus de véritables amis durant mon parcours à la maîtrise. À mes tendres amis, les *popos*, Van, Elsa, Sam, Xin, Gab, et tous ceux qui sauront se reconnaître, merci pour votre précieuse amitié qui m'est très chère.

Merci à Sen, mon meilleur ami avec qui j'apprends tous les jours. Enfin, un énorme merci du fond du cœur à ma famille incroyable que j'aime inconditionnellement.

# Introduction

Le Québec se distingue particulièrement en Amérique du Nord, considérant la taille et la concentration de sa population francophone au sein d'un territoire majoritairement anglophone. Selon les données du recensement de 2016, la population québécoise compte 80,61% de francophones<sup>1</sup> (N = 6 502 823), 10,75% d'anglophones<sup>2</sup> (N = 866 849) et 8,64% d'allophones<sup>3</sup> (N = 696 888) lorsqu'on utilise la langue parlée le plus souvent à la maison pour déterminer l'appartenance linguistique. À titre de comparaison, dans le reste du Canada, les francophones représentent seulement 2,32 % de la population (N = 618 633), les anglophones, 82,55% (N = 22 041 702) et les allophones, 15,13% (N = 4 040 355). La situation linguistique singulière du Québec et le fait que de nouvelles données pertinentes pour aborder le sujet soient désormais accessibles nous poussent à nous interroger sur le rôle que joue l'appartenance à un groupe linguistique sur les inégalités observées en termes de mortalité. Selon les auteurs s'étant penchés sur les tendances passées, il existe des écarts importants notamment en matière de taux de mortalité par âge et d'espérance de vie entre les principaux groupes linguistiques du Québec, qui méritent d'être actualisées et approfondies.

En effet, la baisse considérable de la mortalité depuis le début du XXe siècle observée au Québec comme dans d'autres pays et régions développés a été accompagnée par de grands changements dans la distribution des décès selon l'âge, en l'occurrence par des décès survenant de façon plus importante après l'âge de 65 ans plutôt que durant les premières années de vie. Le terme « démocratisation de la vieillesse » (Légaré, 2009) fait référence au fait qu'une part importante (plus de 90%) de la population québécoise prise à la naissance peut, suivant les conditions actuelles de mortalité, espérer vivre jusqu'aux grands âges, soit au-delà de 65 ans (Base de données sur la longévité canadienne, 2018). Alors qu'il n'y a pas si longtemps, les personnes

---

<sup>1</sup> Inclut la moitié (50%) des individus ayant répondu le français et une autre langue non officielle, et un tiers (33,3%) de ceux ayant répondu le français, l'anglais et une langue non officielle.

<sup>2</sup> Inclut la moitié (50%) des individus ayant répondu l'anglais et une autre langue non officielle, et un tiers (33,3%) de ceux ayant répondu le français, l'anglais et une langue non officielle.

<sup>3</sup> Inclut la moitié (50%) des individus ayant répondu le français et une autre langue non officielle, la moitié (50%) de ceux ayant répondu l'anglais et une autre langue non officielle, et un tiers (33,3%) de ceux ayant répondu le français, l'anglais et une langue non officielle.



âgées étaient principalement issues d'une sous-population homogène hautement sélectionnée, les sociétés vieillissantes ressemblent de plus en plus à l'ensemble de la population, soit à un groupe hétérogène tant au niveau socio-économique que culturel (Légaré, 2004). La vieillesse étant maintenant perçue comme « accessible à la grande majorité de la population », nous entendons nous questionner non seulement sur les différentiels de mortalité *aux âges adultes*, mais également sur l'existence de différences qui persisteraient dans les tendances de la mortalité *aux grands âges* selon l'appartenance linguistique.

L'objectif principal du présent travail est de comparer à l'aide de la nouvelle base de données *Cohorte santé et environnement du recensement canadien de 1991 (CSERCAN)*<sup>4</sup>, fruit du jumelage du recensement de 1991, des données de l'état civil et de fichiers administratifs de 1991 à 2011, les distributions des décès par âge pour les diverses combinaisons des variables sexe et langue (soit le français, l'anglais ou une langue autre parlée le plus souvent à la maison), grâce à une approche de lissage par P-splines.

Ce mémoire est composé de trois principaux chapitres. Le premier aborde la problématique et présente la recension des écrits ainsi que les questions de recherche. Le deuxième chapitre se consacre à une description exhaustive des sources de données et de la méthodologie, principalement liée au lissage par P-splines duquel nous dérivons quatre indices sur la mortalité adulte et aux grands âges : l'espérance de vie à 25 ans ( $e_{25}$ ), à 65 ans ( $e_{65}$ ) et l'âge modal (i.e. le plus fréquent) au décès ( $M$ ), et la mesure de dispersion associée à  $M$  ( $SD(M+)$ ). Le troisième et dernier chapitre a pour but de présenter les principaux résultats. Enfin, nous clôturons avec une discussion qui apporte une compréhension plus approfondie des résultats et avec l'identification des futures pistes à poursuivre dans la recherche sur la mortalité différentielle selon le groupe linguistique.

---

<sup>4</sup> Cette base de données a été renommée deux fois; l'appellation la plus récente était « *Étude canadienne de suivi de la mortalité selon le recensement, 1991 à 2006* », et la version originale, qui était uniquement accessible aux employés de Statistique Canada, s'intitulait *Banque de données jumelées de Statistique Canada (BDJ-SC)*.

# Chapitre 1 : Problématique, recension des écrits et questions de recherche

Ce premier chapitre présente la problématique, se poursuit avec une recension des écrits portant sur la mortalité selon le groupe linguistique au Québec et sur les outils de mesure de la mortalité aux âges adultes et aux grands âges, et se termine par un compte rendu critique des indicateurs linguistiques et spatiaux.

## 1.1 Problématique

La plupart des études précédentes sur la mortalité différentielle selon la langue au Québec s'appuyaient sur la langue d'usage à la maison telle qu'inscrite sur les bulletins de décès. Bien que les données aient été jugées acceptables (Trempe et al., 2014), il existe plusieurs sources de biais inhérentes à cette source d'information, en particulier pour la détection des décès parmi les personnes allophones.<sup>5</sup>

D'abord, il faut considérer la validité des informations linguistiques sur les bulletins de décès, remplis par un médecin dans la majorité des cas. L'enregistrement de la langue d'usage à la maison est sujette à un potentiel d'erreurs important, entre autres parce qu'il est fort possible que la langue inscrite manuellement sur les bulletins de décès soit la langue d'usage du patient à l'hôpital plutôt que la langue d'usage à la maison (Termote, 2011).

De plus, le taux de non-réponse<sup>6</sup> sur ces bulletins de décès s'élève à 11,3% (Trempe et al., 2014), ce qui est non négligeable et n'assure pas une fiabilité certaine des informations. En effet, le risque pour un individu allophone de faire partie d'une telle catégorie inconnue est vraisemblablement plus élevé que pour les individus francophones ou anglophones. La répartition des cas de langue inconnue selon les cas de langue connue risque de conduire à une sous-estimation des décès parmi les individus allophones.

---

<sup>5</sup> L'appartenance linguistique est déterminée dans ce mémoire par la langue parlée le plus souvent à la maison.

<sup>6</sup> Pour la période de 1990 à 2007.

Or, nous profitons de l'accès à une nouvelle base de données, la *Cohorte santé et environnement du recensement canadien de 1991* (CSERCan), qui nous offre une opportunité unique d'étudier avec plus de justesse la mortalité différentielle selon la langue au Québec étant donné la grande fiabilité des informations concernant la variable *langue*, qui proviennent directement du recensement de 1991 dans la CSERCan. De plus, nous abordons pour la première fois les différentiels de mortalité selon la langue *aux grands âges*, ce qui contribue à l'originalité du présent mémoire.

## 1.2 Recension des écrits

### 1.2.1 Différentiels de mortalité selon le groupe linguistique au Québec

Selon l'analyse des données relativement récentes, provenant de diverses sources d'informations telles que le registre des événements démographiques de l'Institut de la statistique du Québec, les recensements canadiens et la banque de données jumelées de Statistique Canada (BDJ-SC)<sup>7</sup>, plusieurs auteurs confirment que l'espérance de vie des allophones est nettement supérieure à celle des francophones, et que les anglophones vivent en moyenne plus longtemps que les francophones (Wilkins et al., 2008; Termote, 2011; Trempe et al., 2014).

Pour expliquer la sous-mortalité des allophones, qui sont majoritairement composés d'immigrants et de leurs descendants, du moins à Montréal, nous savons que ceux-ci sont soumis à des procédures d'immigration très sélectives liées à leur état de santé avant d'être admis sur le territoire. De plus, l'âge moyen à l'arrivée étant de 27 ans, ces individus ont également échappé à tous les risques de décéder avant cet âge (Termote, 2011). La sous-mortalité des allophones serait donc principalement expliquée par le biais de sélection lié à l'effet « d'immigrants en bonne santé » (Trovato, 1985; Bourbeau, 2002; Wilkins et al., 2008). D'ailleurs, cette hypothèse est fortement appuyée par l'étude d'Omariba et Vissanjée (2014),

---

<sup>7</sup> Comme mentionné précédemment, la BDJ-SC correspond à une version antérieure de la CSERCan, dont l'accès n'était à cette époque pratiquement restreint qu'aux personnels de Statistique Canada.

qui utilise des données d'une version antérieure mais plus récente que la BDJ-SC de CSERCan, soit *l'Étude canadienne de suivi sur la mortalité selon le recensement, 1991 à 2006*. Leur analyse, s'appuyant sur des modèles de régression de Cox, démontre que les différences de mortalité existent bel et bien entre les immigrants et les natifs canadiens, même après contrôle des conditions démographiques et socio-économiques, et que l'avantage de mortalité des immigrants tend à s'estomper au fur et à mesure que s'allonge leur durée de résidence au Canada. La perte de cet avantage de mortalité s'observe toutefois après plusieurs années de résidence seulement, et ne s'applique pas pour les immigrants de certaines régions telles que les Philippines (femmes), l'Inde (hommes) et les Caraïbes (hommes), qui conservent leur avantage. De plus, les auteurs réfutent l'hypothèse du « biais du saumon » selon laquelle la mortalité des immigrants est sous-estimée car certains décèderaient à l'étranger. Ils citent à cet effet la grande qualité du jumelage des données jusqu'à l'âge de 100 ans (Bourbeau et Lebel, 2000), l'obligation légale au Canada d'enregistrer les décès pour l'enterrement, l'importance de « régler la succession », et l'étude de Turra et Elo (2008) qui démontre que le biais du saumon a un impact minime sur l'avantage de mortalité des immigrants aux États-Unis.

Pour ce qui est de l'espérance de vie plus élevée des anglophones par rapport aux francophones, cette mortalité différentielle, observable autant dans l'ensemble du Québec que dans chaque région distincte<sup>8</sup>, mérite une attention particulière. Dans les perspectives démologiques du Québec et de la région de Montréal pour la période 2006-2056, Termote (2011) note que l'écart d'espérance de vie des francophones par rapport aux anglophones s'est significativement réduit, passant de 4,1 ans en 1976-1981 à 2,1 ans en 2011. L'amenuisement de l'écart d'espérance de vie entre ces deux groupes linguistiques pourrait provenir du fait qu'en raison du niveau d'espérance de vie plus faible des francophones, leur gain potentiel était plus important comparé aux anglophones.

Enfin, dans l'article publié par Wilkins et collègues (2008), les auteurs dressent le bilan de la mortalité par sexe en termes de taux de mortalité standardisés sur l'âge selon 17 caractéristiques de la cohorte d'une version antérieure de CSERCan, soit *l'Étude canadienne de suivi de la mortalité selon le recensement, 1991 à 2001*. Ces caractéristiques sont d'ordre socio-

---

<sup>8</sup> Ces régions sont l'île de Montréal, le reste de la RMR de Montréal et le reste du Québec (Termote, 2011)

économique (scolarité, revenu, type d'emploi), culturel (minorité visible, origine autochtone), démographique (statut matrimonial, lieu de naissance, année d'immigration), géographique (taille de la communauté, zone d'influence métropolitaine) et linguistique (langue maternelle, connaissances des langues officielles). Cependant, la langue parlée le plus souvent à la maison n'est pas prise en compte dans leur analyse, alors qu'elle constitue à notre avis la variable linguistique la plus pertinente pour mesurer la mortalité différentielle au Québec puisqu'elle reflète le mieux les habitudes de vie actuelles de l'individu (voir la section 1.3.1 pour plus de détails).

## **1.2.2 Analyse comparative de la mortalité selon la source de données**

En tenant compte des possibles sources de biais inhérentes aux informations tirées des bulletins de décès, décrites dans la section 1.1, Termote (2011) remet en question la fiabilité de l'information sur la langue, en particulier pour le groupe d'allophones. L'auteur note l'écart considérable de mortalité dans le reste du Québec où l'espérance de vie à la naissance des allophones s'élève à 87 ans pour les deux sexes réunis<sup>9</sup>, soit cinq et sept ans de plus que celle des anglophones et des francophones respectivement. En considérant le fait que le groupe d'allophones comporte une part importante d'autochtones dans cette région et en sachant que la mortalité de ces derniers est nettement plus élevée que la moyenne québécoise, l'espérance de vie des allophones est d'autant plus suspecte et mérite d'être remise en question.

Dans le but d'évaluer l'ampleur d'une possible sous-estimation de la mortalité des allophones et subséquemment d'une éventuelle surestimation de la mortalité des francophones et des anglophones dans les études sur la mortalité différentielle selon la langue au Québec, Termote (2011) a comparé les espérances de vie (sexes réunis) selon la langue d'usage à la maison, la région et la source de données. L'auteur a comparé ces indicateurs estimés à partir du registre des événements démographiques de l'Institut de la statistique du Québec (ISQ)<sup>10</sup> pour les

---

<sup>9</sup> L'auteur ne distingue pas les sexes dans son analyse afin de minimiser les fréquences des petits effectifs.

<sup>10</sup> Le registre des événements démographiques de l'ISQ fournit la langue d'usage à la maison, telle qu'enregistrée sur les certificats de décès.

périodes 2000-2002 et 2005-2007 avec ceux s'appuyant sur la compilation spéciale de l'*Étude canadienne de suivi de la mortalité, 1991 à 2011*<sup>11</sup> (Wilkins et al., 2008) pour la période 2006-2011. Selon cette comparaison, une importante sous-estimation de la mortalité des allophones est effectivement observée avec les données provenant du registre des décès de l'ISQ, en particulier dans le reste du Québec où l'espérance de vie s'élève à 86,6 ans comparativement à 82,7 ans, et de façon moindre sur l'île de Montréal (86,9 contre 86,1 ans).<sup>12</sup> Pour les francophones, leur mortalité est surestimée, surtout sur l'île de Montréal (78,7 contre 79,9 ans) et aussi dans le reste du Québec (80,0 contre 80,8 ans). Quant aux anglophones, leur mortalité est uniquement surestimée sur l'île de Montréal (81,7 contre 83,2 ans).

Enfin, quelle que soit la source de données utilisée pour estimer les espérances de vie selon la langue d'usage à la maison au Québec, les disparités régionales sont minimales comparativement aux disparités linguistiques (Termote, 2011).

### **1.2.3 Outils de mesure de la mortalité adulte et aux grands âges, résumant la tendance centrale et la dispersion des décès**

Parmi les mesures les plus souvent utilisées pour étudier les niveaux et les tendances de la mortalité, nous retrouvons, entre autres, l'espérance de vie à la naissance ou à tout autre âge, les taux de mortalité par âge et les probabilités de décès ou de survie par âge. L'espérance de vie à la naissance, soit le nombre moyen d'années vécues par l'ensemble de la population depuis la naissance, ne convient pas comme indicateur de tendance centrale pour mesurer la mortalité aux *âges adultes* ou aux *grands âges* tout spécialement puisqu'elle est fortement impactée par les événements qui surviennent en début de vie (Kannisto, 2001). Les chercheurs s'en remettent ainsi à l'espérance de vie à 25 ans et à 65 ans par exemple, qui ciblent davantage les adultes et les personnes âgées respectivement, bien que la détermination de ces limites d'âges demeure

---

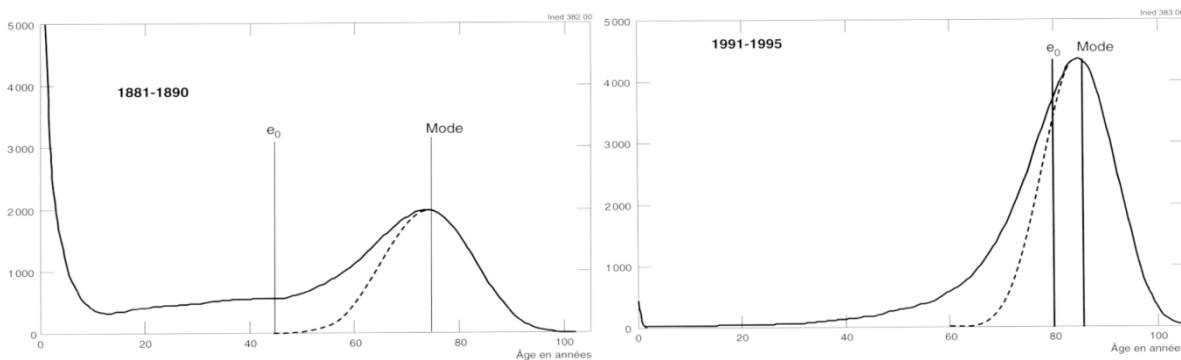
<sup>11</sup> L'*Étude canadienne de suivi de la mortalité, 1991 à 2011*, comporte l'information sur la langue parlée le plus souvent à la maison tirée du recensement détaillé de 1991.

<sup>12</sup> Pour toutes les comparaisons ici, la première valeur d'espérance de vie est celle pour la période 2005-2007 et calculée à partir des données de l'ISQ, alors que la seconde provient de la modélisation des données de l'*Étude canadienne de suivi de la mortalité, 1991 à 2011*, pour la période 2006-2011.

arbitraire. Ces indicateurs mesurent respectivement le nombre moyen d'années vécues par l'ensemble de la population à compter de l'âge de 25 ans et de 65 ans.

Aux âges les plus élevés, les travaux récents proposent, en complément à ces indicateurs usuels de la mortalité, une mesure qui ne nécessite pas d'imposer de limites d'âges arbitraires, soit l'âge modal au décès, qui est devenu un indicateur clé de la longévité (Horiuchi et al., 2013). Par sa définition, l'âge modal au décès (M) représente la durée de vie la plus commune (ou fréquente) des adultes dans une population donnée (Kannisto, 2000). Le concept des durées de vie normales provient de Lexis (1878), qui qualifiait M comme étant l'indicateur le plus central et naturel de la longévité humaine, dont les limites minimales et maximales sont déterminées par les lois de la nature. Selon Lexis, la répartition par âge des décès normaux est identique de part et d'autre de l'âge modal et suit donc une loi normale (ou gaussienne), ce qui permet de distinguer les décès normaux des décès prématurés. Ce constat peut être visualisé par les courbes (traits pointillés) en forme de cloche et ayant comme valeur centrale l'âge modal au décès dans les deux panels de la figure 1.1.

Figure 1.1 Répartition des durées de vie féminines avant et après un siècle de transition en Finlande



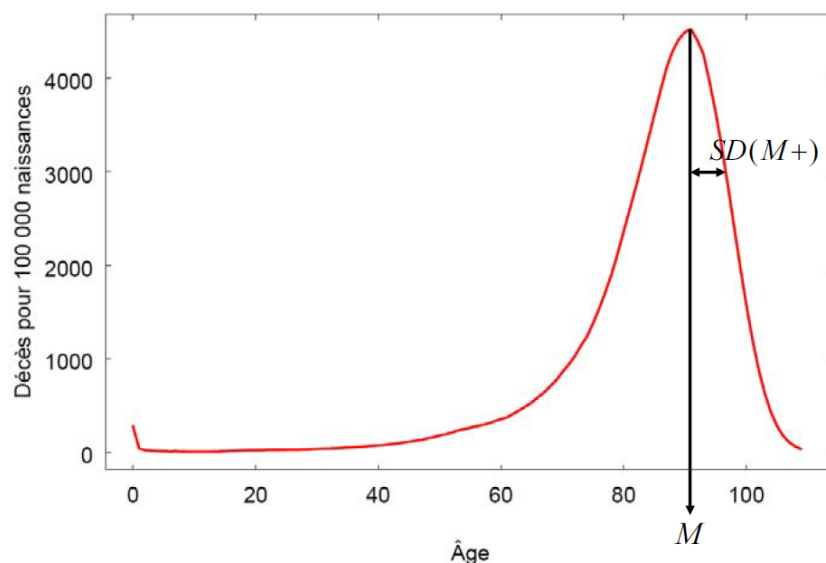
Source : Kannisto, 2001, p.185, Figure 1

Dans cette même figure, nous remarquons d'abord le décalage vers la droite (i.e. vers des âges plus élevés) des deux indicateurs reflétant l'augmentation de l'espérance de vie et de la durée de vie la plus commune. L'augmentation plus rapide de l'espérance de vie à la naissance, fortement influencée par les événements en début de vie, est provoquée par la réduction de la mortalité infantile et adulte. Nous observons également qu'avec la diminution considérable de la part des décès aux jeunes âges entre les périodes 1881-1890 et 1991-1995 en Finlande, l'ensemble des durées de vie devient plus homogène; ces dernières se concentrant autour d'un

intervalle d'âge remarquablement plus court. Ce phénomène communément appelé la *compression (globale) de la mortalité* est fortement lié au processus de rectangularisation de la courbe de survie (Wilmoth et Horiuchi, 1999).

Plus récemment, les chercheurs se sont intéressés au phénomène de la *compression de la mortalité aux âges élevés* tout particulièrement, et ce dernier phénomène a été observé dans plusieurs régions telles que le Canada, la Suisse, la France, les États-Unis et le Japon (Cheung et Robine, 2007; Cheung et al., 2009; Ouellette, 2011). Ouellette (2011) a constaté qu'au Japon, l'ère du *shifting mortality* (Kannisto, 1996) a succédé à celle de la compression de la mortalité aux grands âges depuis le début des années 1990; en d'autres mots, depuis ce temps, l'âge modal au décès continue d'augmenter, sans toutefois que les décès soient concentrés dans un intervalle d'âge de plus en plus court. Afin de mesurer la dispersion des décès aux âges élevés, les chercheurs utilisent l'écart-type des durées de vie situées au-delà de l'âge modal de décès, noté  $SD(M+)$ , soit l'écart quadratique moyen par rapport à  $M$ , comme illustré dans la figure 1.2.

Figure 1.2 Représentation de l'âge modal au décès ( $M$ ) et de l'écart type, relatif à  $M$ , des durées de vie situées au-delà de  $M$  ( $SD(M+)$ ), femmes, Japon, 2000-2004



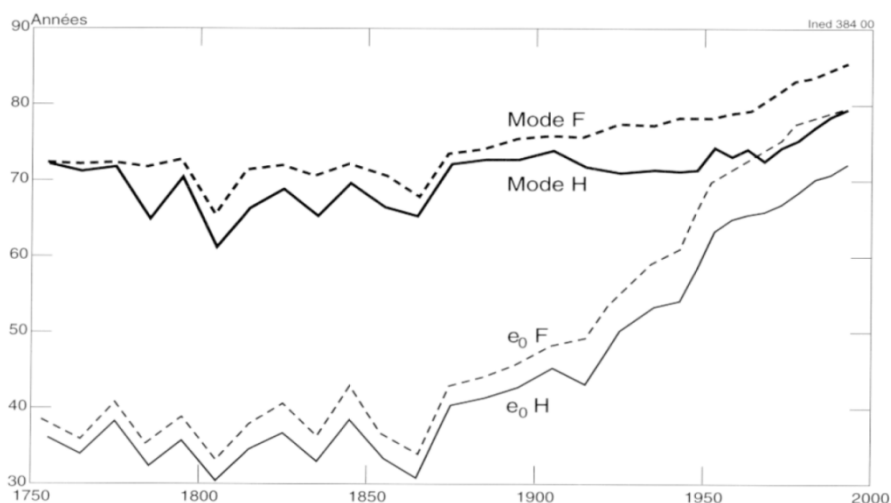
Source : Ouellette, 2011, p.16, Figure 1.2

L'évolution détaillée de l'âge modal au décès peut être observée dans la figure 1.3, en comparaison avec l'espérance de vie à la naissance et en prenant de nouveau la Finlande pour exemple. La durée de vie la plus commune y fluctue de façon marquée sur de courtes périodes, tout en restant aux alentours de 70 ans, de 1750 jusqu'à la fin du 19<sup>e</sup> siècle. Depuis les années



1970 cependant, l'âge modal augmente de façon continue à une vitesse remarquable chez les deux sexes. Effectivement, suite au déclin important de la mortalité infantile au début du vingtième siècle, auquel l'espérance de vie est très sensible contrairement à l'âge modal au décès, le prolongement de la longévité observé dans les pays à faible mortalité est principalement attribuable à l'amélioration de la survie aux grands âges. Or, l'âge modal au décès représente un indicateur qui convient précisément aux particularités d'une période de longévité accrue.

Figure 1.3 Durées modales et moyenne ( $e_0$ ) de la vie selon le sexe en Finlande, 1750 à 1995



Source : Kannisto, 2001, p.186, Figure 2

## 1.2.4 Choix des indicateurs de mortalité

Cette récente hausse de l'âge modal au décès constituait du jamais vu dans les pays à faible mortalité (Kannisto, 2001) et justifie en grande partie l'intérêt d'étudier la mortalité aux grands âges à partir de la paire d'indicateurs ( $M$ ,  $SD(M^+)$ ), comme cela se fait dans un nombre grandissant d'études. Ces deux indicateurs s'avèrent utiles pour suivre non seulement les transformations de la distribution des décès aux grands âges dans les pays à faible mortalité, mais également pour révéler les écarts de mortalité aux âges avancés entre diverses sous-populations. Notamment, la contribution de Ouellette et collègues (2013) montre que malgré la grande réduction de la mortalité depuis le début du 20<sup>e</sup> siècle à travers toutes les provinces canadiennes, les disparités importantes entre provinces au niveau de la mortalité aux grands âges sont restées constantes. Pour sa part, Lecours (2011) démontre qu'au Québec, bien que l'indice

de défavorisation ne permette pas de dévoiler des différences significatives dans les durées de vie modales des femmes, des différences notables au niveau de l'âge modal au décès et de la compression de la mortalité sont observées entre les hommes des sous-groupes les plus favorisés et les plus défavorisés. Puis, Brown et collègues (2012) confirment le lien direct entre le degré de compression de la mortalité aux grands âges et le niveau de scolarité aux États-Unis, en particulier chez les femmes. Enfin, toujours sur l'influence du niveau de scolarité sur la mortalité mais cette fois au Canada, Canon (2017) utilise la base de données CSERCan, comme nous proposons de le faire pour le présent projet de mémoire. Ses résultats confirment que l'âge modal au décès des individus ayant un faible niveau de scolarité est significativement inférieur à celui des plus éduqués, et qu'il existe un gradient de mortalité suivant le niveau de scolarité qui persiste jusqu'aux âges avancés.

Jusqu'à maintenant, la mortalité aux grands âges n'a pas encore été étudiée en fonction du groupe linguistique. Nous ignorons si les différentiels de mortalité selon la langue documentés pour les âges adultes dans plusieurs études antérieures s'observent également chez les personnes âgées, d'où notre intérêt de mieux comprendre les possibles disparités de mortalité selon le sexe et l'appartenance linguistique au Québec, avec l'âge modal au décès, un indicateur clé de la mortalité aux grands âges.

Enfin, à des fins d'analyse comparative avec les recherches antérieures, nous avons également décidé d'estimer les espérances de vie à 25 ans ( $e_{25}$ ) et à 65 ans ( $e_{65}$ ). Ces espérances de vie sont conditionnelles à la survie jusqu'aux seuils d'âge choisis (ici 25 et 65 ans respectivement) et sont largement utilisées comme mesures résumant la tendance centrale des distributions des décès aux âges adultes et avancés.

Dans un premier temps, les travaux de Trempe et collègues (2014) ont estimé  $e_{25}$  selon le sexe et le groupe linguistique au Québec, dans le but de valider la fiabilité de l'information sur la langue parlée à la maison, inscrite sur les certificats de décès du Registre des décès du Québec. Les auteurs ont comparé leurs estimations pour  $e_{25}$  avec celles calculées à partir de la compilation spéciale de l'*Étude canadienne de suivi de la mortalité selon le recensement 1991-*

2001 (Wilkins et al., 2008). L'utilisation de deux variables linguistiques différentes<sup>13</sup> constitue une limite méthodologique importante, que les auteurs justifient comme étant non-problématique étant donné leur unique intention d'évaluer si les écarts entre groupes linguistiques sont les mêmes. À partir de cette comparaison, les auteurs en concluent qu'une imputation multiple des données manquantes aide à réduire le biais provenant du taux élevé de non-réponse (11,4%), rendant l'information sur la langue d'usage à la maison issue du Registre des décès suffisamment exacte pour dresser un portrait réaliste de la mortalité par groupe linguistique. Toutefois, notre doute persiste quant au degré de fiabilité de cette source de données car comme abordé dans la section 1.1, les sources de biais liées à la validité de l'information sont multiples. Nous tenons donc à étudier la mortalité différentielle selon la langue au Québec avec plus d'exactitude en substituant l'information sur la langue issue des fichiers de décès par celle fournie dans le recensement détaillé de 1991 dans la CSERCan, une source nettement plus fiable. Nous jugeons pertinent d'estimer  $e_{25}$  afin de pouvoir mieux comparer et interpréter les différentiels de mortalité aux âges adultes avec les résultats de Trempe et collègues (2014) et de Wilkins et collègues (2008).

Dans un deuxième temps, avec les données de la CSERCan, nous estimons également  $M$  et  $e_{65}$ , qui concernent la survie aux grands âges. Horiuchi et collègues (2013) ont constaté des tendances sensiblement différentes entre ces deux indicateurs. Ils démontrent que de manière générale,  $M$  s'accroît exactement au même rythme que la courbe de mortalité lorsque celle-ci se déplace vers les âges plus avancés, alors que l'espérance de vie à 65 ans (ou à tout autre âge), augmente plus lentement. Nous souhaitons vérifier si les tendances contrastantes des deux indicateurs, notées par Horiuchi et collègues, s'appliquent également dans le contexte linguistique du Québec. Nous jugeons par ailleurs pertinent d'ajouter  $e_{65}$  à notre analyse puisqu'il s'agit d'un indicateur de mortalité grandement utilisé pour mesurer la mortalité aux grands âges, mais n'a, jusqu'à ce jour, pas été calculée en fonction du groupe linguistique au Québec.

---

<sup>13</sup> La langue maternelle est utilisée pour estimer  $e_{25}$  dans les résultats de Wilkins.

# 1.3 Compte rendu critique des indicateurs linguistiques et spatiaux

## 1.3.1 Choix de l'indicateur linguistique

Dans le recensement canadien de 1991, les trois questions relatives à la langue posées aux individus ayant reçu le recensement détaillé concernaient la langue maternelle, la connaissance des langues officielles et la langue d'usage à la maison. Ces indicateurs linguistiques se complètent et représentent chacun une dimension linguistique différente. Pour choisir judicieusement l'indicateur linguistique à utiliser pour nos analyses, il est important de bien comprendre ce que ces indicateurs représentent et les problèmes qu'ils peuvent contenir.

D'abord, la langue maternelle est définie par Statistique Canada comme étant la première langue apprise dans l'enfance et encore comprise au moment du recensement de 1991, ce qui la différencie de la définition donnée par les Nations Unies, soit la langue parlée durant la prime enfance, sans nécessairement être encore comprise (Nations Unies, 2009; Termote, 2011). La langue maternelle représente ainsi l'héritage culturel des individus. Nous pouvons mentionner plusieurs limites de cet indicateur. D'abord, dans l'éventualité où la personne juge qu'elle ne comprend plus la première langue apprise, le questionnaire du recensement canadien demande de déclarer la deuxième langue apprise (et encore comprise), ce qui a pour conséquence de dissimuler un certain nombre de transferts linguistiques. De plus, la notion de « comprendre la langue » est également subjective. Enfin, le recensement de 1991 permet des réponses multiples pour cette question, à condition que l'individu considère avoir appris ses langues maternelles simultanément durant l'enfance en plus de les comprendre au moment du recensement.

Quant à la variable « connaissance des langues officielles », elle est évaluée à partir de la capacité d'un individu à soutenir une conversation dans les langues officielles, soit l'anglais et le français. La binarité de cette variable et la compréhension différentielle de l'expression « soutenir une conversation » selon le groupe linguistique, ajoutent de l'incertitude autour de cette variable (Termote, 2011). De plus, il devient plus ardu de répartir avec exactitude les nombreux cas où un individu a répondu connaître les deux langues officielles. Ce constat annule

presqu'immédiatement la valeur de cette variable pour notre recherche. De manière similaire, la langue d'usage public, qui avait été proposée à la fin des années 1990 au Québec, est le reflet direct de la concurrence des langues officielles (Béland, 1999). Considéré comme un indicateur essentiel pour mesurer l'usage du français dans la sphère publique québécoise, cette variable n'est toutefois pas disponible dans le recensement de 1991 et fait face aux mêmes limites citées pour la connaissance des langues officielles.

Enfin, la langue d'usage à la maison se réfère pour sa part à la langue parlée le plus souvent à la maison, soit le résultat de l'interdépendance du domaine privé et du domaine public. L'expression « le plus souvent » a été reprochée dans le passé pour inciter des réponses uniques (Béland, 1999). Toutefois, dans le recensement de 1991, les réponses multiples sont permises pour tous les indicateurs linguistiques, et la part des cas de bilinguisme et de trilinguisme pour cette variable est la même que celle pour la langue maternelle. À l'opposé de la langue maternelle, la langue d'usage a l'avantage de se référer à la langue parlée à un moment précis dans le temps selon sa fréquence et son lieu d'utilisation (Bourbeau et Tremblay, 1985). Cette variable est uniquement disponible pour les répondants du questionnaire détaillé, soit ceux qui composent la CSERCan à l'étude dans la présente recherche. Dans le recensement de 1991, le taux de non-réponse pour cette question était de 1,7% au Québec et identique à celui pour la langue maternelle. Depuis 1971, Statistique Canada se base sur d'autres informations connues du répondant pour attribuer par imputation une langue maternelle ou d'usage aux personnes ayant omis de répondre à ces questions, ce qui fait qu'officiellement, il n'y a pas de non-réponse pour ces variables. Il peut être pertinent de noter qu'il n'est pas possible de mesurer la mobilité linguistique à partir d'une comparaison entre les effectifs de cette variable et celle de la langue maternelle pour la simple raison que les transferts linguistiques peuvent se produire à n'importe quel moment de la vie et pas seulement d'une période censitaire à l'autre (Termote, 2011).

Il convient de rappeler qu'en analysant la mortalité différentielle selon le groupe linguistique, nous voulons cibler ce qui différencie une sous-population d'une autre, au niveau de la langue. Bien que la langue maternelle ait pu jouer un rôle dans le développement comportemental de l'individu par le transfert intergénérationnel de la culture et de valeurs, nous jugeons qu'il est plus pertinent de choisir la langue parlée le plus souvent à la maison, qui reflète le mieux les comportements actuels de l'individu, à travers des habitudes de vie propres à l'ensemble de

valeurs liées à la culture francophone, anglophone, et « allophone » (Bourbeau et Tremblay, 1985; Termote, 2011).

### **1.3.2 Choix des catégories linguistiques**

Trois catégories de groupes linguistiques sont considérées dans la plupart des études démologiques au Québec, soit les francophones, les anglophones et les allophones. Les allophones se réfèrent à la population dont la langue d'usage est autre que le français ou l'anglais. Selon les perspectives démologiques pour la période 2006 à 2056 de Termote (2011) réalisées à partir des données des recensements, le groupe allophone, bien qu'hétérogène, ne pouvait être désagrégé en raison des petits effectifs par âge et par région. Dans le présent mémoire, nous décomposons cette catégorie linguistique en différenciant les langues allochtones des langues autochtones. Le groupe autochtone regroupe ainsi l'ensemble des langues autochtones parlées au Québec, alors que les langues allochtones se réfèrent aux autres langues non officielles issues de l'immigration (autre que l'anglais et le français). Cette différenciation se justifie par le comportement démographique nettement différent des groupes autochtones et allochtones. Effectivement, nous savons que les taux de mortalité de la population aborigène (Indiens inscrits, Indiens non-inscrits, Inuit et Métis) sont considérablement plus élevés que ceux du reste de la population canadienne (Tjepkema et al., 2009), et que c'est plutôt le contraire pour la mortalité des immigrants, qui composent une partie importante du groupe des allochtones. Nous jugeons donc primordial de séparer les personnes parlant une langue autochtone à la maison, de ceux parlant une autre langue non officielle, afin d'éviter le plus possible une bipolarité de cette catégorie linguistique. Nous supposons bien entendu que les individus dont la langue d'usage à la maison est une langue autochtone sont d'origine autochtone, le lien entre le groupe ethnique et le groupe linguistique correspondant étant relié assez fortement, dépendamment de la situation (Bourbeau et Tremblay, 1985).

### **1.3.3 Choix du découpage territorial**

Les données publiées par Statistique Canada différencient la division de recensement « île de Montréal » de l'ensemble de la région métropolitaine de recensement (RMR) de Montréal. L'île de Montréal est libellée comme étant la « Communauté urbaine de Montréal » dans la

variable de division de recensement de résidence, et n'est pas affectée par les changements des limites des municipalités dans le temps, ce qui lui donne un avantage temporel dans le cas où des études de suivi dans le temps seraient faites (Termote, 2001, 2011). Principalement, l'immigration internationale et interrégionale, en particulier l'étalement urbain, joue un rôle important dans la structure démographique de l'île de Montréal (MTL) contrairement au « reste du Québec » (RDQ), où l'incidence de ces deux facteurs est mineure, ou au reste de la RMR de Montréal, où l'immigration internationale joue un rôle secondaire mais où nous observons l'exode des immigrants de l'île vers cette région (Termote, 2011).

Dans un monde idéal, afin de mieux déceler les différentes dynamiques démographiques présentes dans chaque région, il aurait fallu séparer la province en trois régions distinctes, soit l'île de Montréal, le reste de la RMR de Montréal (i.e. la RMR moins l'île) et le reste du Québec (i.e. le Québec excluant la RMR de Montréal), comme l'a fait Termote (2011). Toutefois, en raison de petits effectifs, il nous est impossible de conduire une analyse de la mortalité comparative par âge et par sexe selon le groupe linguistique pour ces trois régions séparément.

C'est dans ce contexte que nous comptons mener nos analyses selon deux régions distinctes : l'île de Montréal et le reste du Québec. Après avoir minutieusement considéré les conditions visant à mieux représenter les dynamiques démographiques du Québec, ces deux régions de découpage territorial ont été sélectionnées en tenant compte de la disponibilité des données, de la puissance statistique liée aux effectifs, et des caractéristiques démographiques de chaque zone territoriale. Cette dernière condition vise à reconnaître la différente composition des groupes linguistiques qui existe dans l'île de Montréal par rapport au reste du Québec, révélatrice de dynamiques et de comportements démographiques diversifiés.

## **1.4 Questions de recherche**

Selon plusieurs auteurs, il existe une mortalité différentielle entre les francophones, les anglophones et les allophones du Québec qui mérite d'être mieux documentée et comprise. Les différences de mortalité entre ces groupes linguistiques sont caractérisées par une surmortalité des francophones et une sous-mortalité des allophones, par rapport aux anglophones (Tremblay et Bourbeau, 1985; Robitaille et al., 1992; Termote, 1992, 1999, 2011; Trempe et al., 2014). Comme ce projet de recherche s'intéresse d'abord à actualiser ces observations à partir d'une

source plus fiable, soit la base de données CSERCan, et ensuite à évaluer s'il existe des différences notables dans la répartition des décès aux grands âges parmi les francophones, les anglophones et les allophones au Québec, les principaux indicateurs retenus sont : l'espérance de vie à 25 ans ( $e_{25}$ ), celle à 65 ans ( $e_{65}$ ), et l'âge modal au décès (M) en plus de sa mesure de dispersion des décès associée. Bien entendu, un changement de perspective s'impose dans l'interprétation des indicateurs avec des seuils d'âge, soit  $e_{25}$  et  $e_{65}$ , et de M qui est, au contraire, exempt de toute limite arbitraire d'âge. En effet, l'âge modal au décès mesure la durée de vie la plus commune et se focalise tout particulièrement sur les personnes très âgées, en évitant l'essentiel de la mortalité prématurée. Il ne peut donc pas être interprété de la même façon que les indices  $e_{25}$  et  $e_{65}$ , qui ciblent pour leur part les adultes et les personnes âgées respectivement, et qui sont tributaires de limites d'âges arbitraires. Spécifiquement, les questions de recherche sont les suivantes :

- Existe-t-il des différences significatives de mortalité selon la langue au Québec ?
- Dans l'affirmative, est-ce que ces différences s'observent à la fois chez les 25 ans et plus, chez les 65 ans et plus et parmi l'ensemble des personnes atteignant l'âge modal au décès ?
- Comment se comparent les dispersions des décès au-delà de l'âge modal selon l'appartenance linguistique ?
- Les différences de mortalité selon l'appartenance linguistique sont-elles les mêmes sur l'île de Montréal et dans le reste du Québec ?

Nous estimons qu'il existe effectivement des différences significatives de mortalité aux âges adultes selon l'appartenance linguistique au Québec, et que ces écarts persistent aux âges avancés. Les profils sociodémographiques et économiques des divers groupes linguistiques à l'étude contribuent, à notre avis, à expliquer cette mortalité différentielle.



# Chapitre 2 : Sources de données et méthodologie

Ce deuxième chapitre présente les aspects méthodologiques du mémoire de manière approfondie. Nous débutons par une description des sources de données employées en commentant la qualité des données et leurs limites. Nous poursuivons avec la méthodologie appliquée pour le traitement et l'analyse des données.

## 2.1 Sources de données

Pour la réalisation de ce mémoire, deux sources de données ont été utilisées. La principale source s'intitule *Cohorte santé et environnement du recensement canadien de 1991* (CSERCan), successeur de *l'Étude canadienne de suivi de la mortalité et du cancer selon le recensement, 1991 à 2001 et 1991 à 2006*. Nous avons eu accès à la CSERCan, qui provient de Statistique Canada, via le Centre interuniversitaire québécois de statistiques sociales (CIQSS). La source de données secondaire à laquelle nous recourrons se nomme la *Base de données sur la longévité canadienne* (BDLC) disponible sur le site Internet suivant : [www.bdlc.umontreal.ca](http://www.bdlc.umontreal.ca).

### 2.1.1 Cohorte Santé et Environnement du Recensement Canadien de 1991

Cette recherche s'appuie principalement sur la nouvelle base de données *Cohorte Santé et Environnement du Recensement du Canada de 1991* (CSERCan) comportant 2,5 millions de Canadiens suivis du 4 juin 1991 au 31 décembre 2011.

La CSERCan est issue d'un jumelage de bases de données qui a été approuvé par le Conseil exécutif de gestion de Statistique Canada. La cohorte, suivie durant la période 1991-2011, constitue un échantillon représentatif de la population-cible, c'est-à-dire la population résidente, âgée de 25 ans et plus en date du 4 juin 1991, et ne vivant pas en institution à cette date-là (n = 2 644 370, 15% de la population). Cette cohorte est composée des individus non pensionnaires

d'un établissement ayant répondu au questionnaire détaillé et produit une déclaration de revenus en 1990 ou 1991. L'échantillon utilisé pour notre analyse du Québec compte 682 210 individus. Des analyses tirées de cet échantillon représentatif du Québec ont été publiées, entre autres dans les *Perspectives démographiques du Québec et de la région de Montréal pour la période 2006-2056* (Termote, 2011).

La version de la CSERCan utilisée dans le cadre de ce mémoire est dérivée d'un couplage de six sources de données citées dans le Guide de l'utilisateur préparé par la Division de l'analyse de la santé (DAS) de Statistique Canada (2016), qui sont:

- le Recensement de la population canadienne de 1991, questionnaire détaillé (versions 2B et 2D);
- la Base canadienne de données sur la mortalité (BCDM) de 1991 à 2011;
- la Base canadienne de données sur le cancer (BCDC) de 1969 à 1991;
- le Registre canadien du cancer (RCC) de 1992 à 2010;
- le Fichier de données fiscales sommaires historiques (FDFSH) de 1984 à 2011;
- le Fichier maître des particuliers T1 (FMP T1) de 1981 à 1983.

Nous présentons ci-dessous une brève description de chacune de ces sources de données (Statistique Canada, 2016).

### ***Le Recensement de la population canadienne de 1991, questionnaire détaillé***

L'échantillon d'individus inclus dans la CSERCan comporte plusieurs caractéristiques sociodémographiques et économiques qui sont principalement tirées du questionnaire détaillé, soit les versions 2B et 2D du Recensement de la population de 1991. Conduit à tous les cinq ans, ce questionnaire détaillé vise à dresser un portrait approfondi de la population du Canada. Parmi les informations recueillies dans les versions 2B et 2D se trouvent l'âge, le sexe, la langue parlée le plus souvent à la maison, la langue maternelle, le niveau de scolarité, l'état matrimonial, le lieu de naissance, l'activité sur le marché du travail, les caractéristiques du logement, la mobilité, etc. Il convient de mentionner que le dénombrement de tous les ménages dans les régions éloignées du Nord du Canada et dans les réserves indiennes s'est fait avec le questionnaire 2D, administré directement par l'agent recenseur. Comme discuté précédemment, entre autres dans la section 1.3.1, la langue parlée le plus souvent à la maison constitue la

variable principale de ce mémoire, car elle reflète le mieux les comportements de l'individu en date de la tenue du recensement de 1991.

### ***La Base canadienne de données sur la mortalité (BCDM) de 1991 à 2011***

La BCDM comporte des données sur la totalité des décès survenus au Canada de 1950 à ce jour. L'information est fournie annuellement par les bureaux de l'état civil des provinces et territoires canadiens, et est transmise à Statistique Canada à des fins d'analyse. Les données, compilées à partir de fichiers administratifs, sont complètes compte tenu des lois provinciales et territoriales sur la statistique de l'état civil qui portent sur l'enregistrement obligatoire de l'ensemble des naissances vivantes, des mortinaissances, des mariages et des décès survenus au sein de leur circonscription.

### ***La Base canadienne de données sur le cancer (BCDC) de 1969 à 1991***

Créée et utilisée par la Division de la statistique de la santé de Statistique Canada, la BCDC est un fichier de données historiques servant à assurer un suivi aux fins des projets de recherche sur les risques à long terme pour la santé des Canadiens. Ces données ont été extraites du Système national de déclaration des cas de cancer de 1969 à 1991, et du Registre canadien du cancer à partir de 1992. Le fichier analytique renferme les variables sur le diagnostic et sa date, la morphologie, la topographie et certaines variables démographiques. Dans le cadre de ce mémoire, les variables citées provenant de cette base de données n'ont pas été utilisées.

### ***Le Registre canadien du cancer (RCC) de 1992 à 2010***

Le RCC concerne tous les résidents du Canada ayant été diagnostiqués d'un cancer depuis 1992. Ces informations sont transmises par les registres provinciaux et territoriaux du cancer à la Division de la statistique de la santé de Statistique Canada. À titre de remplacement au Système national de déclaration des cas de cancer, le RCC est une base de données dynamique axée sur la personne plutôt que les événements et visant à recueillir les informations liées au type et à l'incidence de tumeurs primaires diagnostiquées chez une personne jusqu'à son décès. Nous n'avons pas eu recours au RCC pour notre projet de mémoire.

## ***Le Fichier de données fiscales sommaires historiques (FDFSH) de 1984 à 2011 et le Fichier maître des particuliers T1 (FMP T1) de 1981 à 1983***

Le FDFSH comprend les informations liées aux déclarations de revenus par année d'imposition. Cette base de données permet l'obtention de certaines variables comme le nom, le code postal, l'état matrimonial, le code postal lié aux déclarations de revenus, la date de décès s'il y a lieu, dont plusieurs ont facilité le couplage probabiliste avec la BCDM, la BCDC et le RCC. Dans le cadre de ce mémoire, ce fichier de données s'est avéré fort utile dans l'identification des émigrants, compte tenu des informations fournies concernant la dernière année pour laquelle une déclaration de revenus a été produite.

### **2.1.1.2 Méthodologie de jumelage**

Statistique Canada a procédé à une méthode de couplage d'enregistrement décrite dans le Guide de l'utilisateur de la DAS. Les grandes lignes sont présentées dans cette section.

En sachant que le fichier électronique des données du recensement ne comportait aucun nom et que les noms étaient nécessaires au couplage des enregistrements du recensement avec les données sur la mortalité et sur le cancer, un premier couplage des données du recensement avec les fichiers de données fiscales sommaires (FDFSH) s'est imposé. Comme première étape ayant pour but de diminuer l'apparition d'appariements faussement positifs, un couplage déterministe a été effectué entre l'univers total de la population ayant répondu au recensement de 1991 et la liste complète des Canadiens ayant produit une déclaration de revenus pour les années 1990 *ou* 1991 à n'importe quel moment entre le mois d'avril 1991 et de mars 1992.

Les dates de naissance et les codes postaux communs constituaient les premiers critères d'appariement, mis en commun avec les informations des conjoints (incluant les conjoints de fait). Étant donné que le questionnaire abrégé du recensement ne renfermait pas les codes postaux, ceux-ci ont été imputés à partir du code postal des voisins les plus proches (connus à partir du secteur de dénombrement). Les enregistrements des FDFSH, soit les fichiers de noms, correspondant le mieux aux enregistrements des individus ayant répondu aux questionnaires abrégés ont alors été retranchés. Cette étape a permis de minimiser l'occurrence d'appariements faussement positifs des individus n'étant pas éligibles à la cohorte suivie pour la mortalité. Des méthodes probabilistes ont alors été effectuées pour coupler les enregistrements des FDFSH

restants avec les enregistrements du questionnaire détaillé. Parmi les 3 576 487 individus canadiens âgés de 25 ans et plus ayant répondu au questionnaire détaillé, 80% d'entre eux (n = 2 860 244) ont été correctement appariés<sup>14</sup> avec les fichiers de nom du FDFSH (Wilkins et al., 2008).

Jusqu'aux changements apportés à la base de données autrefois nommée l'*Étude canadienne de suivi de la mortalité selon le recensement de 1991*, la taille de la cohorte représentait 15% de la population canadienne âgée de 25 ans et plus (n = 2 735 152).<sup>15</sup> Comme évoqué plus haut, en 2016, la base de données a été renommée *Cohorte santé et environnement du recensement canadien de 1991*, et comporte les ajouts d'un suivi de la mortalité allant jusqu'au 31 décembre 2011 et des codes postaux datant des années fiscales 1981 à 1983. En raison d'une mise à jour du suivi de la mortalité qui a provoqué l'élimination de certains couplages, tout particulièrement au cours de la période 1991 à 2007, la taille de la cohorte a été réduite, passant de 2 735 152 à 2 644 370 individus. L'échantillon final de la cohorte représente donc les trois quarts (74%) de la population d'intérêt. Plus d'informations sur la composition de la cohorte sont disponibles plus bas dans la section sur les limites de la CSERCan.

### **2.1.1.3 Qualité des données et limites**

#### ***Qualité des données***

Selon les rapports sur la qualité des données des recensements de Statistique Canada (2015), le taux de sous-dénombrement brut pour la province du Québec a été estimé à 3,18% dans le recensement canadien de 1991, utilisé pour l'établissement de la CSERCan. Les résidents non permanents ont été dénombrés mis à part les ressortissants étrangers et les membres de leur famille rattachés à une quelconque mission diplomatique du Canada, les militaires étrangers et les membres de leur famille, ainsi que les ressortissants étrangers visitant temporairement le pays.

Au niveau de la qualité des données, la CSERCan offre un potentiel inédit en termes d'analyses approfondies de la mortalité sur la population du Canada et celle de ses provinces. La richesse

---

<sup>14</sup> Wilkins et ses collègues (2008) ont déterminé que 99% de ces appariements étaient exacts.

<sup>15</sup> Cet effectif a été réduit de façon aléatoire de 2 860 244 à 2 735 152 individus pour que la cohorte constitue un échantillon de 15% de la population canadienne.

de la CSERCan provient du regroupement des caractéristiques socio-économiques, démographiques, et géographiques d'un échantillon de grande taille représentatif de la population canadienne de 25 ans et plus, et du suivi de mortalité de ces individus dans le temps durant plus de 20 ans. Aux fins de notre mémoire, la taille de l'échantillon permet d'assurer une bonne représentativité des sous-groupes de la population québécoise, tels que les immigrants et les personnes âgées. La longue durée d'observation de la CSERCan aide à la réalisation d'études comparatives de la mortalité selon différentes caractéristiques individuelles, dans notre cas, la langue parlée le plus souvent à la maison.

### ***Limites***

Malgré la grande qualité de la CSERCan, quelques limites méritent d'être mentionnées et auxquelles nous serons attentifs dans nos analyses.

D'abord, l'évolution des caractéristiques sociodémographiques et économiques (incluant l'appartenance à un groupe linguistique) de la cohorte n'est pas connue puisque ces renseignements proviennent du recensement détaillé de 1991 exclusivement et sont donc considérés comme fixés dans le temps. Par contre, des études ont pu démontrer à partir des données intercensitaires que le nombre de transferts linguistiques *dans toutes les directions* pour l'ensemble du Québec, et pour tous les âges (incluant les individus âgés de moins de 25 ans) s'élève à environ 1 000 ou 2 000 transferts par année (Termote, 2008). En tenant compte de ces chiffres et du fait que les transferts linguistiques ont surtout lieu en bas âge, la mobilité linguistique jouerait un rôle marginal dans notre analyse de la mortalité différentielle selon l'appartenance linguistique chez les 25 ans ou plus.

Puis, l'exclusion de la population institutionnalisée dans le recensement détaillé provoque un biais au niveau de la mortalité puisque ces résidents institutionnels ont tendance à être en moins bonne santé. Ainsi, aux âges les plus élevés surtout, on peut s'attendre à ce que les membres de la cohorte affichent des taux de mortalité inférieurs à leurs homologues de la population générale. Cependant, cette limite serait moins importante que dans des études antérieures qui auraient utilisé la même base de données, mais avec une période de suivi plus courte. À titre d'exemple, si initialement, l'échantillon ne comprend pas les individus institutionnalisés au moment du recensement en 1991, celui-ci contient cependant le suivi de mortalité des individus

admis dans une institution *après* 1991. Ainsi, le biais lié à l'exclusion de la population institutionnalisée dans l'échantillon de départ, bien que toujours existant, s'amenuise au fil du temps. Nous montrons effectivement à la section 3.1.1 que les écarts observés sont plutôt faibles. Enfin, lors du jumelage des enregistrements du questionnaire détaillé avec les fichiers de nom de la FDFSH, les probabilités d'un jumelage fructueux variaient selon les caractéristiques des individus. Nous savons que seuls les individus issus de la population d'intérêt ayant pu être couplés avec leurs déclarations de revenus en 1990 et 1991 ont été retenus dans la cohorte. Or, certaines catégories d'individus possédaient une plus grande probabilité de non-appariement (en ordre d'importance) : ceux d'origine autochtone, du quintile inférieur de suffisance de revenu, ne faisant pas partie de la population active, non mariés ou ne vivant pas avec leur conjoint, âgés de 65 ans et plus, ayant déménagé au cours de la dernière année, n'ayant pas de diplôme secondaire et enfin, résidant en milieu rural (Wilkins et al., 2008). Toutes ces catégories d'individus étaient moins susceptibles soit, d'être des déclarants, ou d'être correctement appariés avec leurs fichiers de nom. Malgré cela, la CSERCan assure une représentativité satisfaisante de la population canadienne, compte tenu du nombre important d'individus suivis dans le temps (2,6 millions) et de la grande part (80%) des enregistrements qui ont été fructueusement appariés.

## **2.1.2 Base de données sur la longévité canadienne (BDLC)**

La Base de données sur la longévité canadienne (BDLC), sous la responsabilité des professeurs Robert Bourbeau et Nadine Ouellette, constitue un projet réalisé par l'Équipe de recherche sur la mortalité et la longévité du Département de démographie de l'Université de Montréal, en collaboration avec les démographes du Max Planck Institute for Demographic Research (MPIDR) à Rostock en Allemagne et le Département de démographie de la University of California, à Berkeley aux États-Unis.

La création de la BDLC a pour but de servir aux chercheurs, étudiants, journalistes, analystes de politiques, ainsi qu'à l'ensemble des personnes intéressées par la longévité humaine au Canada. La BDLC a été construite à partir de la même méthodologie que celle utilisée pour les données de la Human Mortality Database (HMD).

Ici, les données de la BDLC ont servi à la création d'une cohorte fictive correspondant approximativement à l'univers d'analyse sur lequel se base la CSERCan. Ainsi, des taux de mortalité selon le sexe et l'âge pour l'ensemble du Québec ont été calculés pour la cohorte fictive de la BDLC, et pour la CSERCan, d'abord en ignorant les sorties par émigration au sein de la CSERCan, puis en tenant compte de l'émigration (voir la section 2.2.1.2). Dans un premier temps, la comparaison des taux de mortalité par âge et sexe de la CSERCan avec ceux de la BDLC a permis de valider lesdits taux de mortalité. Dans un deuxième temps, cette comparaison a mis en lumière la pertinence de prendre en compte l'émigration dans notre étude de la mortalité de la CSERCan.

## 2.2 Méthodologie

Afin de répondre à nos objectifs de recherche, nous proposons la méthodologie avancée par Ouellette et Bourbeau (2011) pour quantifier avec précision l'espérance de vie à 25 ans ( $e_{25}$ ), à 65 ans ( $e_{65}$ ), ainsi que l'âge modal au décès ( $M$ ) et sa mesure de dispersion associée ( $SD(M+)$ ). Notre objectif principal consistera à analyser la mortalité différentielle selon le groupe linguistique à partir de ces quatre indicateurs de la mortalité et de fournir, en s'appuyant sur les profils sociodémographiques et économiques de chaque groupe, quelques éléments potentiellement à l'origine de ces écarts. Grâce à l'approche de lissage par P-splines, ces indicateurs dérivés de la distribution lissée des décès par âge, sexe et langue, permettront de résumer la tendance centrale et la dispersion de la distribution des décès aux âges adultes et aux grands âges pour chaque groupe linguistique, en faisant la différenciation par sexe. Par ailleurs, des intervalles de confiance à 95% seront calculés pour tous nos indicateurs afin de mesurer l'incertitude relative à leurs estimations.

De plus, l'originalité de ce mémoire repose non seulement sur l'utilisation d'un modèle non paramétrique souple qui permet de modéliser la distribution des décès sans recourir à une forme fonctionnelle prédéterminée, mais également sur la considération de l'émigration dans la cohorte. La prise en compte de ce phénomène démographique, auparavant ignoré dans les travaux sur la CSERCan, s'est révélée très pertinente dans le contexte de la mortalité différentielle selon l'appartenance linguistique.



Détaillons à présent la méthodologie employée pour mener nos analyses, en débutant par le traitement des données et en terminant avec l'approche par P-splines.

## 2.2.1 Traitement des données

Nous avons d'abord traité la base de données dans le but de minimiser les sources d'imprécision et de rendre nos résultats comparables avec les études antérieures. Effectivement, nous avons jugé pertinent de répartir les réponses multiples de la variable linguistique pour faciliter les comparaisons. Cette répartition a été appliquée dans l'étude de Termote (2011), notamment. Puis, à l'aide des informations tirées du Fichier de données fiscales sommaires historiques (FDFSH, section 2.1.1), nous avons pu cibler tous les individus ayant fort probablement émigré afin de les retirer, année après année, de la population à risque de décéder.

### 2.2.1.1 Catégories linguistiques

Rappelons qu'aux fins de notre étude, nous considérons trois groupes linguistiques : les francophones<sup>16</sup>, les anglophones<sup>17</sup>, et les allophones composés d'allochtones<sup>18</sup> et d'autochtones<sup>19</sup>. Contrairement aux études linguistiques menées dans le passé, nous avons divisé le groupe d'allophones en différenciant les allochtones des autochtones, puisque ce sont deux groupes diamétralement opposés au regard de la mortalité.

#### *Répartition des langues multiples*

Afin d'assurer une certaine comparabilité avec les études antérieures sur la mortalité différentielle selon la langue au Québec, nous avons réparti les groupes de bilingues et trilingues en langue unique de la même manière que Termote (2011). La répartition des déclarations de langues multiples a été traitée selon une méthode conventionnelle largement acceptée, soit moitié-moitié pour les groupes de bilingues et un tiers chacun pour les groupes de trilingues (francophone, anglophone et allochtone/autochtone). Toutefois, étant donné que nous avons

---

<sup>16</sup> Le groupe dont les membres parlent le plus souvent le français à la maison.

<sup>17</sup> Le groupe dont les membres parlent le plus souvent l'anglais à la maison.

<sup>18</sup> Le groupe dont les membres ne parlent ni le français, ni l'anglais, ni une langue autochtone le plus souvent à la maison. La liste de ces langues se retrouve dans le tableau annexe A.I.

<sup>19</sup> Le groupe dont les membres parlent le plus souvent une langue autochtone à la maison.

quatre catégories linguistiques pour le reste du Québec plutôt que trois, nous avons procédé à une répartition plus complexe pour la catégorie d’allochtones/autochtones.

Il convient de préciser que la méthode de répartition linguistique de Termote (2011) utilisée ici diffère de celle retenue par Trempe et collègues (2014). Ces derniers mentionnent que les individus bilingues français/anglais sont exclus de leur analyse, sans fournir d’explication. Or, dans un rapport sur la mortalité évitable des deux principales communautés linguistiques du Québec par les mêmes auteurs (Trempe et al., 2013), les décès classés comme « francophone » incluent les individus pour qui la langue d’usage est le français uniquement, et le français en plus d’une autre langue non officielle. Les décès dont la langue d’usage était l’anglais uniquement ou l’anglais et une autre langue non officielle ont été classés comme « anglophones ». Nous présumons que cette méthode de répartition linguistique a également été appliquée dans l’article de validation de l’information linguistique issue du Registre des décès.

### ***Répartition des langues autochtones***

Selon le livre de codes du recensement de 1991, la variable *hlnder*, qui se réfère à la langue parlée le plus souvent à la maison, contient plus de cent huit langues uniques, trois possibilités de bilinguisme (français/anglais, français/langue(s) non officielle(s), anglais/langue(s) non officielle(s)), et une seule possibilité de trilinguisme (français/anglais/langue non officielle(s)).<sup>20</sup>

Pour la catégorie des autochtones, vingt-six langues autochtones<sup>21</sup> figuraient dans la variable *hlnder* en langue unique, incluant l’Inuktitut, le Montagnais-naskapi, le Mohawk et le Cri, qui sont toutes des langues non officielles. Par ailleurs, certains individus parlant une langue autochtone sont camouflés dans les réponses multiples qui incluent une langue non officielle. Nous avons trois cas spécifiques où cette situation s’applique : les bilingues parlant le français et une langue non officielle, les bilingues parlant l’anglais et une langue non officielle, et les trilingues parlant le français, l’anglais et une langue non officielle. Nous avons également la réponse « Langues non officielles » à la variable *hlnder* que nous avons répartie entre

---

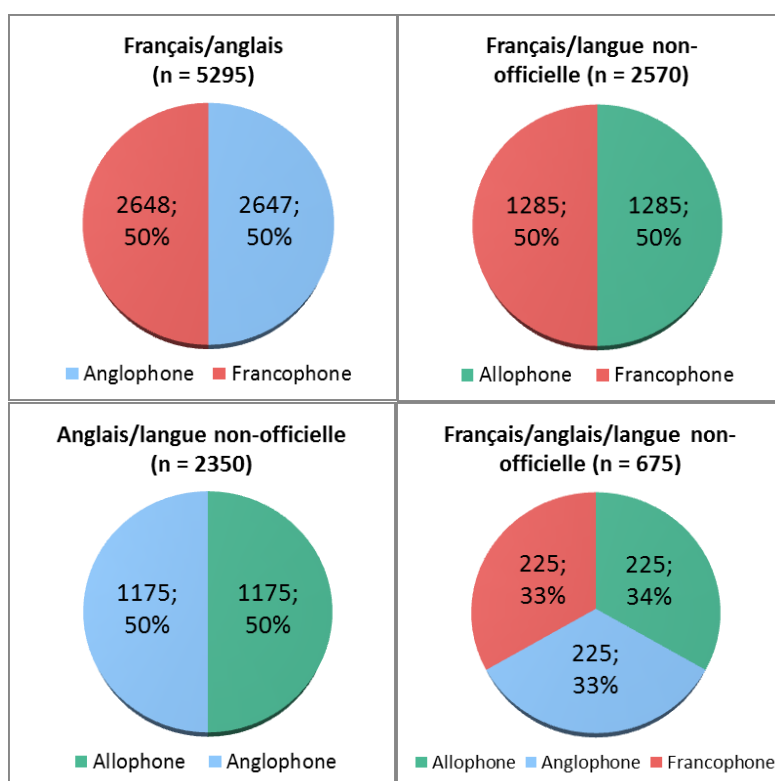
<sup>20</sup> Étant donné l’impossibilité de savoir si un individu a répondu plus d’une langue non officielle dans les possibilités de bilinguisme ou de trilinguisme, et considérant la rareté de cette éventualité, nous avons supposé que ces individus parlent une seule langue non officielle, en plus du français et/ou de l’anglais.

<sup>21</sup> La liste complète des langues autochtones se retrouve dans le tableau annexe A.II.

allochtones et autochtones (n = 120). Nous obtenons ainsi 2 805 individus bilingues, trilingues, ou ayant déclaré parler une (seule) autre langue non officielle, qui ont été catégorisés comme faisant partie du groupe d’allochtones/autochtones dans la province du Québec (partie verte de la figure 2.1). Plus précisément, 2 010 de ces individus se trouvent sur l’île de Montréal, et 795 dans le reste du Québec.

La figure 2.1 montre que comme mentionné plus haut, nous avons d’abord réparti à 50/50 tous les groupes bilingues et à 33/33/33 le groupe trilingue.

Figure 2.1 Effectifs estimés par catégorie linguistique, CSERCan, Québec, 1991 à 2011



Dans l’objectif de répartir les individus du groupe d’allophones en allochtones et autochtones le plus justement possible, nous avons décidé de procéder autrement que la règle de répartition 50/50 dans ce cas-ci, compte tenu de la rareté d’individus autochtones sur l’île de Montréal. Nous avons donc formulé trois hypothèses pour y arriver :

1. Les individus autochtones et allochtones ont autant de chance d’être bilingue ou trilingue.

2. Le nombre d'individus déclarant parler à la fois une langue autochtone et une langue allochtone est négligeable.
3. L'attribution d'une langue unique (autochtone ou allochtone) est déterminée selon les proportions de ces groupes linguistiques en réponses uniques dans le recensement de 1991, dans chacune des régions.

Sur l'île de Montréal (MTL), les individus de la CSERCan ayant répondu parler une langue autochtone est négligeable, ce qui signifie que pour MTL, l'entièreté des individus (n = 2 010) faisant partie du groupe d'allochtones/autochtones sont considérés allochtones. Pour le reste du Québec, pour les réponses uniques des langues non officielles, nous retrouvions une langue autochtone (n = 4 310) à 40% et à 60% une langue allochtone (n = 6 365). En nous basant sur ces proportions, il nous est possible de répartir le groupe d'allochtones/autochtones (n = 795) dans le RDQ de la manière suivante : 320 individus autochtones et 475 individus allophones.

Le tableau I montre les effectifs estimés par sexe, région et catégorie linguistique après répartition des réponses multiples :

Tableau I. Effectifs estimés par sexe, région et catégorie linguistique, CSERCan, Québec, 1991 à 2011

Régions	Île de Montréal				Reste du Québec			
	Femme		Homme		Femme		Homme	
Groupe linguistique	Effectifs	%	Effectifs	%	Effectifs	%	Effectifs	%
Francophone	57900	61,8%	49100	59,0%	230445	91,8%	233725	91,9%
Anglophone	22725	24,3%	20590	24,7%	14595	5,8%	15080	5,9%
Allochtone	13030	13,9%	13590	16,3%	3270	1,3%	3530	1,4%
Autochtone	N/A	N/A	N/A	N/A	2770	1,1%	1860	0,7%
Total	93655	100%	83280	100%	251080	100%	254195	100%

### 2.2.1.2 Émigration

Grâce aux données de la base canadienne de données sur la mortalité (BCDM, voir la section 2.1.1) de 1991 à 2011 jumelées avec celles du recensement de 1991, nous avons accès au suivi de la mortalité des individus de la cohorte 1991. Cependant, ces données nous fournissent uniquement l'information des individus décédés au Canada. Nous ignorons donc la

date de décès des émigrants décédés à l'extérieur du pays. Malgré le fait que l'hypothèse liée au biais du saumon a été réfutée par certains auteurs (Turra et Elo, 2008; Omariba et Vissanjée, 2014), nous avons décelé certains indices dans la CSERCan qui nous poussent à croire qu'il existe effectivement des décès non enregistrés, soit ceux ayant eu lieu à l'extérieur du pays, qui méritent d'être pris en compte dans notre analyse pour de multiples raisons.

D'abord, nous avons remarqué, en particulier aux très grands âges là où les décès se font de plus en plus rares, une part non négligeable d'individus survivant jusqu'à des âges irréalistes et, en parallèle, une baisse importante et étonnante de la trajectoire décrite par les taux de mortalité par âge (voir la figure 3.1, section 3.1). Il est devenu clair que les présumés cas de longévité exceptionnelle dans la CSERCan sont principalement des émigrants et le fait d'ignorer leur émigration mènerait à une sous-estimation notable de la mortalité aux grands âges. Le fait de ne pas tenir compte de l'émigration aux plus jeunes âges résulterait aussi en une sous-estimation des taux de mortalité. De plus, en sachant que l'émigration varie beaucoup selon l'appartenance linguistique, nous avons jugé crucial d'en tenir compte dans notre analyse afin d'éviter de sous-estimer la mortalité des allochtones. En effet, les allochtones, qui sont pour la plupart des immigrants, sont plus à même d'émigrer que les francophones ou les autochtones, par exemple. Le risque pour les individus allochtones d'atteindre des âges supposément extrêmes dans la CSERCan, simplement parce qu'ils ont en réalité quitté le Canada durant la période 1991-2011 et sont décédés à l'étranger, est donc plus élevé. Les décès d'émigrants n'étant pas comptabilisés au numérateur des taux de mortalité que nous calculons pour la CSERCan, il est judicieux de retirer les individus qui ont émigré de la population soumise au risque de décéder, soit le dénominateur des taux de mortalité, afin d'en améliorer la précision. Nous avons donc posé quelques hypothèses pour identifier les émigrants de la cohorte en nous basant sur deux variables : la dernière année de revenus déclarés (*htsf\_year*) et les décès (*dead*).

### ***Identification des émigrants***

Les hypothèses proviennent de plusieurs constatations. D'abord, l'identification des émigrants repose sur le principe qu'un individu ayant émigré ne produit plus de déclaration de

revenus au Canada.<sup>22</sup> Certaines circonstances contreviennent à cette règle simple, puisque certains individus déclarent leurs revenus avec plusieurs années de retard, ou arrêtent tout simplement de les déclarer en raison d'une conjoncture défavorable, que ce soit la maladie, l'hospitalisation ou le vieil âge. La période d'observation pour le suivi de la mortalité se terminant le 31 décembre 2011, il nous a fallu poser quelques hypothèses importantes afin d'éviter d'engendrer de « faux » émigrants, c'est-à-dire des individus toujours présents sur le territoire mais n'ayant tout simplement pas produit de déclaration de revenus pendant un certain nombre d'années.

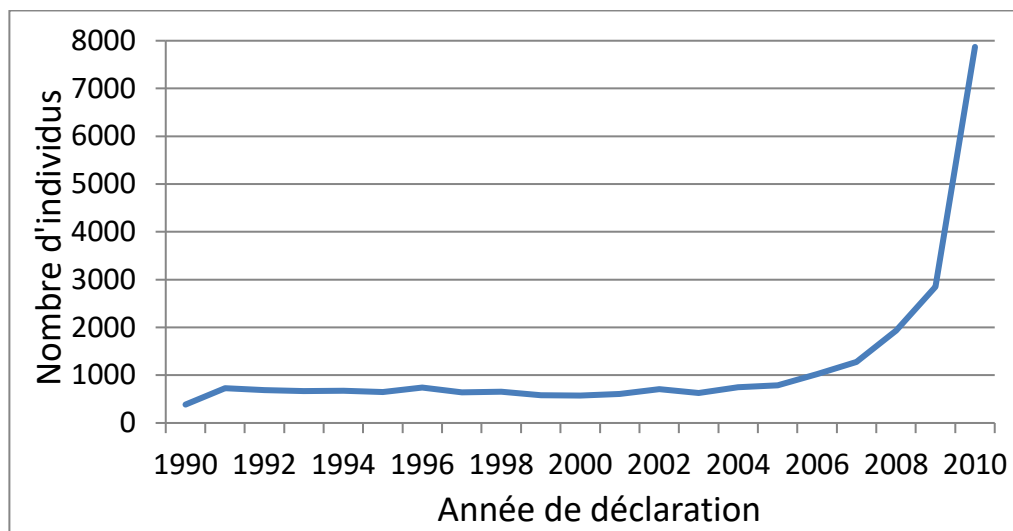
Dans un premier temps, nous avons analysé la variable *htsf\_year*, qui nous révèle la dernière année de déclaration de revenus des individus de la cohorte qui sont toujours considérés vivants à la fin de la période d'observation. La figure 2.2 montre que durant la période 1991 à 2005, le nombre d'individus varie assez peu d'une année à l'autre. Toutefois, à partir de 2006, les effectifs d'individus n'ayant pas déclaré leurs revenus augmentent drastiquement et de manière soutenue, jusqu'en 2010.

Pour interpréter correctement la courbe présentée à la figure 2.2, nous avons, dans un deuxième temps, évalué le nombre d'années écoulées depuis la dernière déclaration de revenus pour les individus déjà décédés (croisement de la variable *htsf\_year* et de l'année de décès renseignée dans la variable *dead*). Nous avons constaté que de manière générale, l'année de la dernière déclaration de revenus est fortement corrélée avec l'année de décès, et qu'il y a très peu d'individus qui ont omis de déclarer leurs revenus bien avant l'année de leur décès. La très grande majorité (97% à 99%) des individus décèdent au cours des six années suivant leur dernière déclaration de revenus.

---

<sup>22</sup> Rappelons que les individus appartenant à la CSERCan ont tous produit une déclaration de revenus pour les années 1990 ou 1991.

Figure 2.2 Répartition des membres de la CSERCan selon leur dernière année de déclaration de revenus, Québec, 1991 à 2010



Note : Le graphique s'arrête à 2010, car l'effectif pour 2011 contient tous les individus vivants ayant déclaré leurs revenus à l'année maximale (2011).

En écartant la possibilité qu'il y ait eu une grande vague d'émigration de 2006 à 2011 dû à l'absence d'événements économiques ou politiques explicatifs au Québec durant cette période, et en sachant qu'une large majorité des individus décédés semblent compléter leur déclaration de revenus dans les 6 années précédant leur décès, l'augmentation soudaine de non-déclarants observée durant la période 2006-2011 dans la figure 2.2 semble plutôt provenir d'un grand nombre de retardataires. Cette explication est la plus plausible pour expliquer la correspondance entre la fin de la période d'observation et le niveau de plus en plus élevé d'individus n'ayant pas produit de déclarations de revenus.

### ***Hypothèses pour l'émigration***

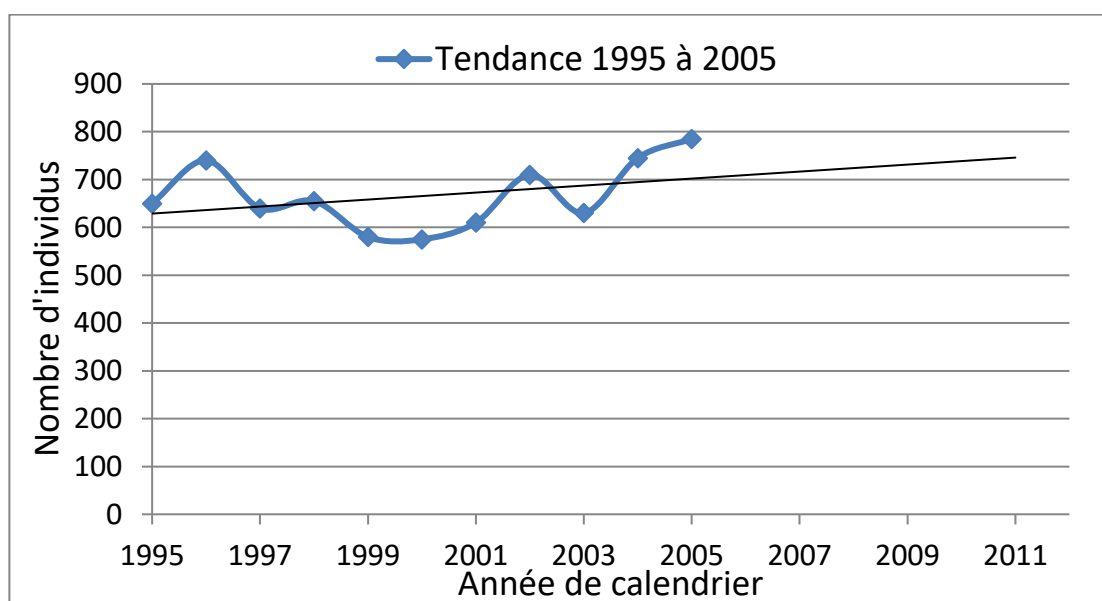
C'est dans ce contexte que nous avons posé les hypothèses suivantes pour l'émigration. Tout d'abord, une personne considérée vivante dans la base de données de la CSERCan n'ayant pas déclaré ses revenus depuis plus de 6 ans est automatiquement considérée émigrante. Ainsi, tous les individus présumés non décédés qui n'ont pas produit de déclaration de revenus depuis 2005 (inclus) ou avant 2005 sont considérés émigrants. Nous supposons que ces individus quittent le Canada en moyenne au milieu de la dernière année de déclaration de revenus, soit le 1<sup>er</sup> juillet, à part pour les individus ayant l'année 1990 comme dernière année de déclaration.

Ceux-ci étaient présents au moment du recensement le 4 juin 1991, et sont plutôt considérés partis au milieu de l'année 1991.

Ensuite, dans le but d'estimer les effectifs d'émigrants durant la période 2006-2011, nous avons analysé l'évolution des données observées durant la période 1995-2005 et simplement extrapolé la tendance linéaire qui se dégageait jusqu'en 2011 (figure 2.3). L'attribution du statut d'émigrant s'est ensuite faite de manière aléatoire auprès des individus concernés. Ainsi, parmi les individus ayant déclaré leurs revenus pour la dernière fois entre 2006 et 2011, un certain nombre d'entre eux ont été identifié « émigrants ». Par exemple, parmi les individus considérés vivants et dont la dernière déclaration de revenus remonte à 2006, 710 d'entre eux ont été sélectionnés aléatoirement à l'aide de la fonction *rannum* de Stata (StataCorp, 2012). Le tableau annexe B.I présente les nombres estimés d'émigrants par année.

À l'aide de la date de sortie imputée au 1<sup>er</sup> juillet de la dernière année de déclaration et la date de naissance, nous avons pu déterminer l'âge au départ des émigrants, ce qui nous a permis de les retirer de la population soumise au risque de décéder aux âges appropriés. Le calcul des différents cas possibles pour calculer l'exposition au risque de décéder par âge est expliqué plus en détail dans la section 2.2.2.1.

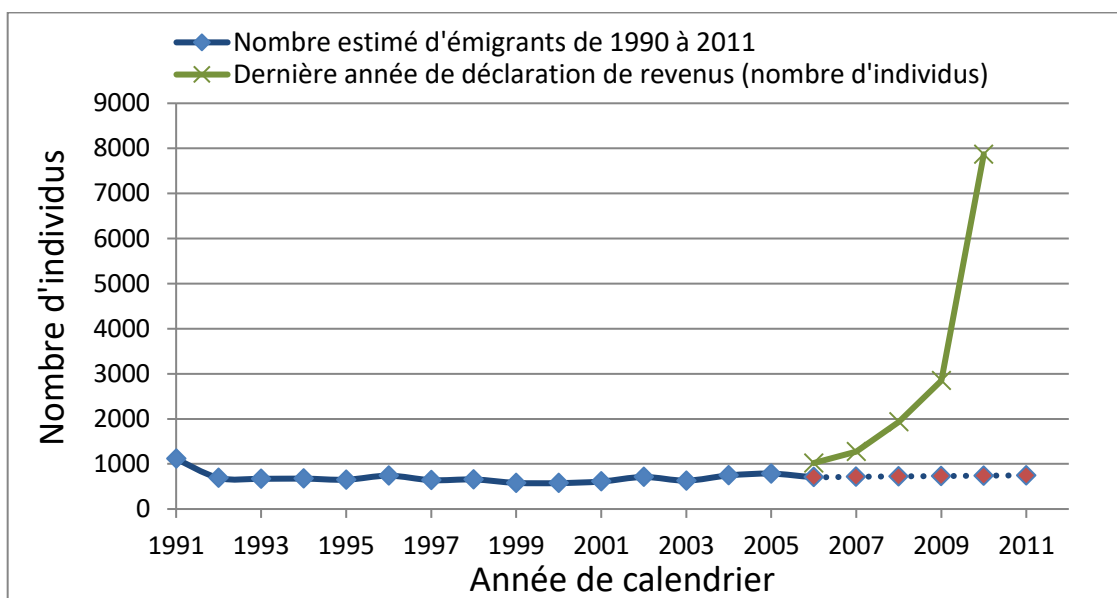
Figure 2.3 Nombre annuel d'émigrants observé (1995-2005) et projeté (2006-2011), CSERCan, Québec





La figure 2.4 présente les estimations relatives au nombre d'émigrants par année, de 1991 à 2011 (voir aussi le tableau B.I). Cette figure illustre le contraste entre le nombre de non-déclarants observés (ligne bleue et verte) et les estimations du nombre annuel d'émigrants observé pour la période 1995-2005 (losanges bleus) et projeté de 2006 à 2011 (losanges rouges).

Figure 2.4 Estimations du nombre annuel d'émigrants, CSERCan, Québec, 1991 à 2011



## 2.2.2 Taux de mortalité et force de mortalité

Par définition, les taux spécifiques de mortalité pour les intervalles d'âge  $[x, x+1[$  sont calculés en rapportant le nombre de décès observés par âge et de l'effectif de la population exposée au risque de décéder du même âge :

$$m_x = \frac{D_x}{E_x}. \quad (2.1)$$

Pour sa part, le taux de mortalité instantané ou force de mortalité, noté  $\mu_x$ , est tel que

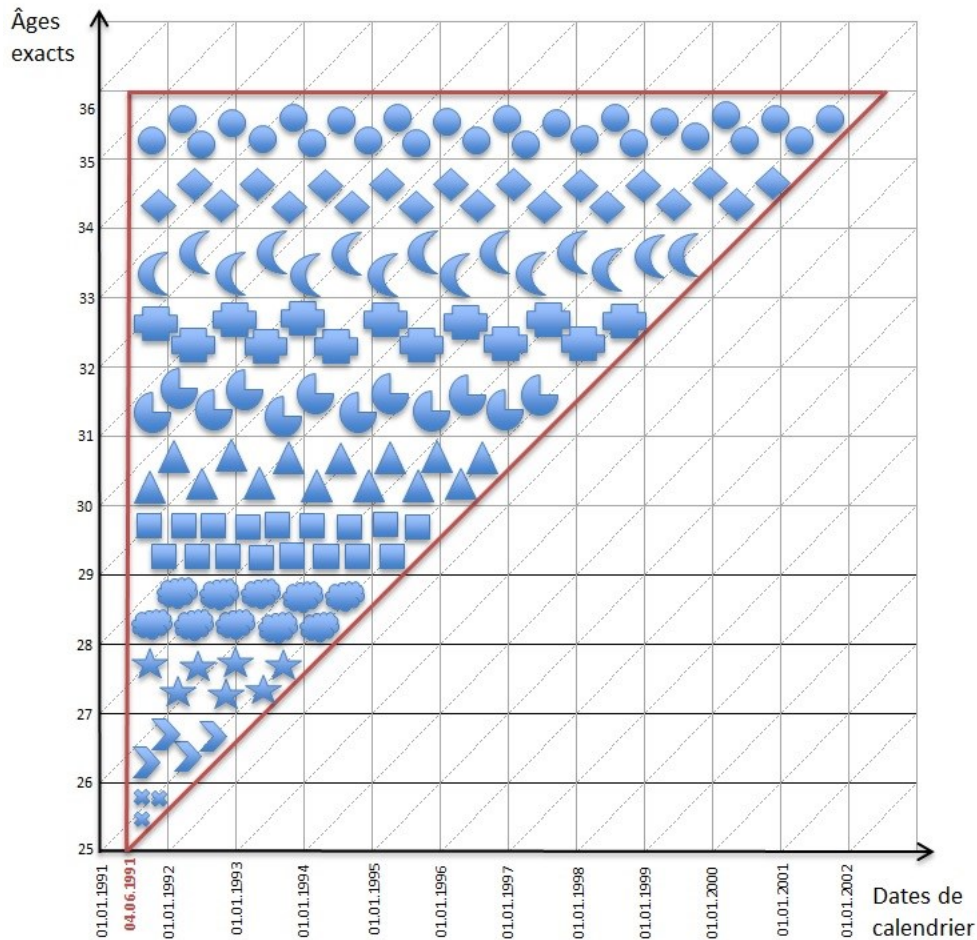
$$m_x \cong \mu_{x+\frac{1}{2}} \text{ (Thatcher et al., 1998).}$$

### 2.2.2.1 Calcul des taux de mortalité par âge pour la CSERCan

Décrivons maintenant la méthode de calculs des taux de mortalité par âge pour la CSERCan, à appliquer séparément par sexe, groupe linguistique et région. Le diagramme de

Lexis présenté à la figure 2.5 montre un extrait de l'univers d'analyse des données de la CSERCan, délimité par les lignes rouges. Les différents symboles (carrés, triangles, losanges, etc.) dépeignent les configurations décrites par les données de la CSERCan pour nos calculs de taux de mortalité par âge.

Figure 2.5 Extrait de l'univers d'analyse des données de la CSERCan, Québec, 1991 à 2011



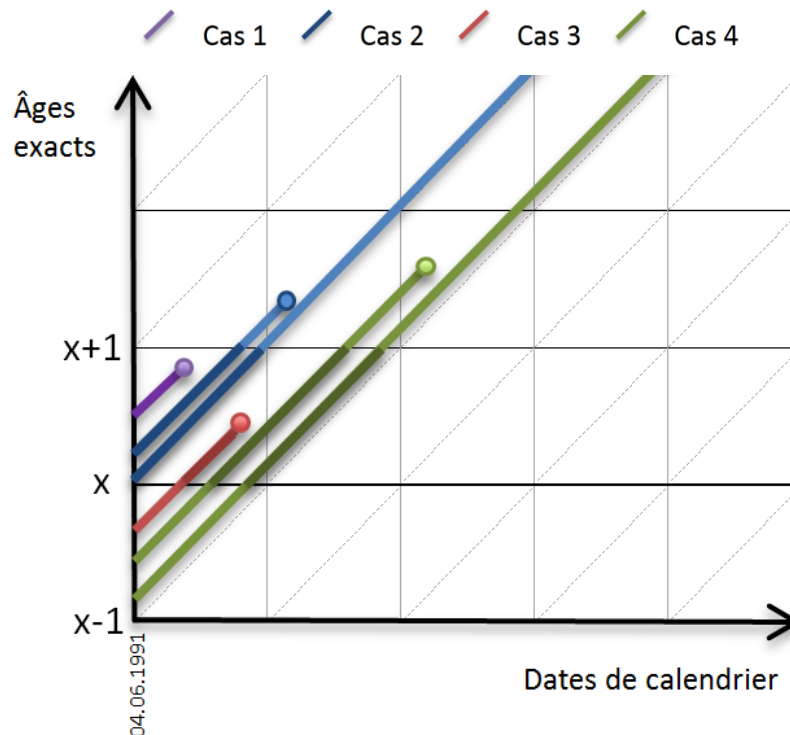
Nous avons d'abord additionné le nombre d'individus décédés dans chaque intervalle d'âge  $[x, x+1[$  pour  $x > 25$  durant la période d'intérêt pour trouver les décès par année d'âge qui représentent le numérateur de l'équation 2.1.

Compte tenu de la configuration des données de la CSERCan (figure 2.5), le calcul des personnes soumises au risque de décéder dans chaque intervalle d'âge, soit le dénominateur de

l'équation 2.1, doit prendre en compte tous les cas possibles de trajectoires individuelles. Nous dénombrons quatre possibilités, illustrées à la figure 2.6 et décrites ci-dessous :

1. La personne entre en observation à l'âge de  $x$  ans révolus, et **décède** ou **émigre** dans l'intervalle d'âge  $[x, x+1[$ . La personne contribue donc à l'exposition au risque pour la différence entre l'âge au décès et l'âge à l'entrée (âge en date du 4 juin 1991).
2. La personne entre à l'âge de  $x$  ans révolus, et soit elle **décède** ou **émigre** à l'âge de  $x+1$  ans, soit elle survit jusqu'à la fin de l'étude, le 31 décembre 2011. L'exposition au risque dans l'intervalle d'âge  $[x, x+1[$  est égale à la différence entre  $x+1$  et l'âge d'entrée à l'étude.
3. La personne entre à un âge inférieur à l'âge de  $x$  ans exacts, et **décède** ou **émigre** dans l'intervalle d'âge  $[x, x+1[$ . Son exposition au risque de décéder est égale à la différence entre l'âge au décès et l'âge exact  $x$ .
4. La personne entre à un âge inférieur à l'âge de  $x$  ans exacts, et soit elle **décède** ou **émigre** à l'âge de  $[x+1, \omega[$ , soit elle est encore en vie à la fin de l'étude. L'exposition au risque de décéder est égale à 1 an.

Figure 2.6 Calcul des années vécues selon les quatre cas possibles



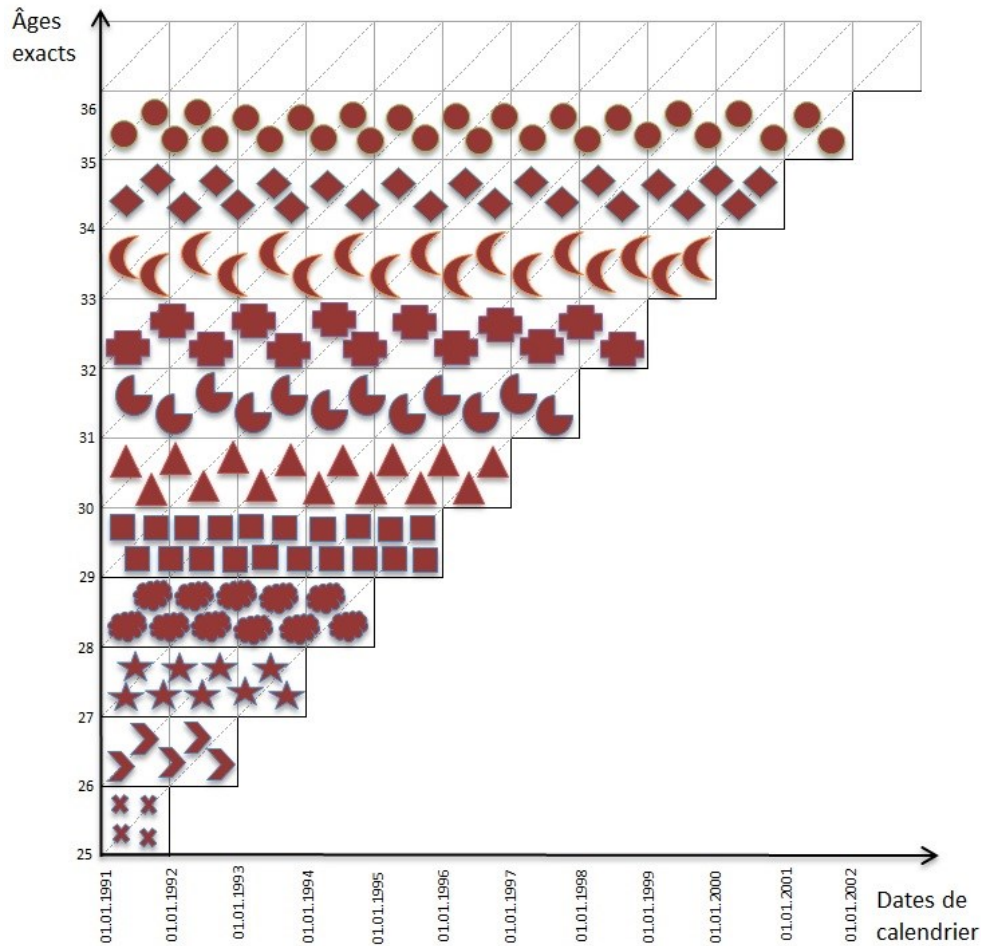
Note : Les lignes font référence aux trajectoires de vie des individus. Les cercles retrouvés au bout de certaines lignes signifient que la trajectoire de vie s'est terminée par le décès ou l'émigration. La partie plus foncée sur chaque ligne désigne l'exposition au risque de décéder pour l'âge  $x$ .

Les résultats de ces quatre cas possibles, évalués à chaque âge  $x$ , prennent la forme d'une matrice où les âges sont portés en ordonnée, et les quatre cas en abscisse. L'exposition au risque de décéder par âge est le résultat de chaque ligne de la matrice (soit l'addition des quatre cas possibles pour chaque âge  $x$ ). Ayant obtenu les nombres de décès et de personnes soumises au risque de décéder par année d'âge, nous appliquons l'équation 2.1 pour obtenir la série des taux de mortalité par âge pour la CSERCan.

### 2.2.2.2 Taux de mortalité par âge pour la cohorte BDLC

Les taux de mortalité par âge et sexe calculés dans cette sous-section concernent ceux de la cohorte fictive créée à partir des données de la BDLC, avec l'objectif de les comparer avec les taux obtenus pour la CSERCan pour valider ces derniers. La figure 2.7 fournit un aperçu de la configuration des données utilisées dans la BDLC pour reproduire aussi fidèlement que possible la CSERCan, en prenant le nombre de décès et les effectifs de la population soumise au risque par âge et sexe pour la province du Québec de 1991 à 2011.

Figure 2.7 Extrait de l'univers d'analyse des données de la cohorte BDLC, Québec, 1991 à 2011



Pour déterminer le nombre de décès par âge, soit le numérateur de l'équation 2.1, nous additionnons le nombre d'individus décédés entre  $x$  et  $x+1$  ans exacts pour les années d'intérêt. Par exemple, pour obtenir le nombre de décès à l'âge de 29 ans révolus pour la cohorte BDLC (rectangle de Lexis identifié par des carrés rouges à la figure 2.7), nous avons fait le calcul suivant :

$$D_{29} = D_{29}^{1991} + D_{29}^{1992} + D_{29}^{1993} + D_{29}^{1994} + D_{29}^{1995}.$$

Pour trouver les effectifs de la population soumis au risque de décéder par âge, soit le dénominateur de l'équation 2.1, nous avons additionné les individus ayant  $x$  ans révolus durant la période d'intérêt. À titre d'exemple, pour trouver l'effectif de la population exposée au risque de décéder à l'âge de 29 ans révolus dans la cohorte BDLC, nous avons fait le calcul suivant :

$$E_{29} = \mathcal{E}_{29}^{1991} + \mathcal{E}_{29}^{1992} + \mathcal{E}_{29}^{1993} + \mathcal{E}_{29}^{1994} + \mathcal{E}_{29}^{1995}.$$

Ayant obtenu les nombres de décès et de personnes soumises au risque de décéder par année d'âge pour la cohorte BDLC, nous pouvons alors calculer les taux de mortalité par âge à partir de l'équation 2.1.

### 2.2.3 Modèle de régression de Poisson

Sous l'hypothèse d'une force de mortalité ( $\mu_x$ ) constante dans chaque intervalle d'âge  $[x, x+1[$ , les décès observés ( $D$ ) peuvent être considérés comme des réalisations d'une loi de Poisson de moyenne  $E_x * \mu_x$ :

$$D_x \sim \text{Poisson}(E_x * \mu_x),$$

où  $E_x$  représente l'effectif de la population exposée au risque de décéder à l'âge  $x$ .

Pour estimer la force de mortalité à partir des données sur les décès observés, nous appliquons le modèle de régression de Poisson décrit par :

$$\ln(\mathbb{E}[D]) = \ln(E * \mu) = \ln(E) + \ln(\mu),$$

où les vecteurs  $D$ ,  $E$  et  $\mu$  comprennent respectivement l'information sur les décès observés, les effectifs de la population exposée au risque de décéder, et les forces de mortalité, selon l'âge.

À partir de l'approche non paramétrique de lissage par P-splines, les paramètres du modèle de Poisson sont estimés à l'aide de l'équation suivante :

$$\ln(\widehat{\mathbb{E}}[D]) = \ln(E * \hat{\mu}) = \ln(E) + \ln(\hat{\mu}) = \ln(E) + B\hat{\alpha}, \quad (2.2)$$

où  $B$  représente la matrice associée à la base B-splines évaluée aux âges  $x$ , perçus comme des régresseurs, et où le vecteur  $\hat{\alpha}$  renferme les estimations des paramètres associés aux B-splines se trouvant dans la base  $B$ . Ces différents concepts sont expliqués plus en détail dans les prochaines sous-sections.

## 2.2.4 Lissage par P-splines

Les propriétés, les attributs et l'efficacité du lissage par P-splines dans le contexte de la mortalité sont largement documentés dans l'article de Ouellette et Bourbeau (2011) et nous anticipons que cette approche nous procurera une précision augmentée dans la détection des disparités de mortalité selon l'appartenance linguistique.

Brièvement, l'usage des P-splines assure une grande souplesse en n'imposant aucune forme fonctionnelle prédéterminée lors de la modélisation de la distribution des décès par âge. Néanmoins, il est recommandé de ne pas prolonger le lissage jusqu'aux âges les plus extrêmes, en raison des petits effectifs de décès et de population soumise au risque. À ces âges, la courbe de mortalité (et par conséquent la distribution des décès) lissée par P-splines pourrait présenter des tendances irréalistes provenant de fluctuations aléatoires importantes. Selon la règle de base proposée par Horiuchi et collègues (2013), le lissage par P-splines exige plus de 10 décès par intervalle d'âge afin de prévenir des tendances indésirables dans la courbe de mortalité lissée, ce que nous avons suivi. À partir de la distribution des décès lissée correspondante (plutôt qu'avec les taux de mortalité observés par âge), nous dérivons les indicateurs de mortalité qui nous intéressent. De cette façon, il n'est pas requis de construire une table de mortalité, qui repose pour sa part sur un certain nombre d'hypothèses et d'ajustements, et cela nous offre une structure plus directe et proche des données observées.

## 2.2.5 Les B-splines et les P-splines

L'idée de l'approche par P-splines d'Eilers et Marx (1996) consiste d'une part à utiliser un très grand nombre de B-splines, formés de bouts de polynômes joints par des nœuds répartis uniformément sur le domaine des données observées, de manière à se prévaloir d'une souplesse telle que la courbe résultante puisse pratiquement passer par tous les points à modéliser. D'autre part, afin de garantir l'aspect lisse de la courbe, il est essentiel d'introduire une pénalité dans la fonction de vraisemblance, qui agit sur les coefficients des B-splines, de telle manière que les variations des coefficients estimés de B-splines adjacents soient modiques.

À partir de l'équation 2.2, nous pouvons estimer la force de mortalité lissée  $\hat{\mu}(x)$  comme suit :

$$\hat{\mu}(x) = \exp(\mathbf{B}(x)\hat{\alpha}). \quad (2.3)$$

## 2.2.6 Fonctions de survie et de densité

En nous basant sur la force de mortalité lissée  $\hat{\mu}(x)$ , nous obtenons la fonction de survie lissée avec la formule suivante :

$$\hat{S}(x) = \exp\left(-\int_0^x \hat{\mu}(t)dt\right),$$

et recourrons à une méthode d'intégration numérique simple pour résoudre l'intégrale. Nous obtenons alors la fonction de densité lissée  $\hat{f}(x)$ , décrivant la répartition des décès selon l'âge, et donnée par le produit de la force de mortalité et la fonction de survie lissées :

$$\hat{f}(x) = \hat{\mu}(x)\hat{S}(x).$$

### 2.2.6.1 Les espérances de vie et l'âge modal au décès ainsi que sa mesure de dispersion associée

Pour estimer l'espérance de vie à 25 ans, nous utilisons la relation suivante :

$$\hat{e}_{25} = \frac{\int_{25}^y (x-25)\hat{f}(x)dx + \left(y + \frac{1}{\hat{\mu}(y)} - 25\right)\hat{S}(y)}{\hat{S}(25)}. \quad (2.4)$$

Dans l'équation précédente,  $y$  correspond à l'âge maximal auquel sont observés 10 décès ou plus et la valeur de  $y$  varie donc selon le groupe linguistique, le sexe et la région.

Similairement, pour estimer l'espérance de vie à 65 ans pour chaque sexe, il suffit de remplacer 25 par 65 dans l'équation (2.4).

L'âge modal au décès estimé découle directement de la fonction de densité lissée  $\hat{f}(x)$  :

$$\hat{M} = \max_x \hat{f}(x),$$

et l'équation suivante permet ensuite d'estimer l'écart type des âges au décès au-delà du mode :

$$SD(\hat{M}+) = \sqrt{\frac{\int_{\hat{M}}^y (x-\hat{M})^2 \hat{f}(x)dx + \left(y + \frac{1}{\hat{\mu}(y)} - \hat{M}\right)^2 \hat{S}(y)}{\hat{S}(25)}}.$$



## 2.2.7 Construction des intervalles de confiance

Dans le but de déterminer s'il existe des différences significatives en matière d'espérance de vie (à 25 et à 65 ans) et d'âge modal au décès (M) entre les groupes linguistiques, les sexes et les régions à l'étude, nous avons construit des intervalles de confiance pour  $\hat{e}_{25}$ ,  $\hat{e}_{65}$  et  $\hat{M}$ . Une méthode bootstrap, initialement développée par Koissi et collègues (2006) et adaptée aux P-splines par Ouellette et collègues (2013), a été privilégiée car la loi de probabilité et la variance de la statistique M sont inconnues.

À partir de l'équation (2.2) et de la force de mortalité lissée  $\hat{\mu}$  décrite par l'équation (2.3), nous obtenons la matrice des décès lissés telle que  $\hat{d} = e * \hat{\mu}$ , où  $\hat{D} = \text{vec}(\hat{d})$  ce qui nous permet de calculer le vecteur des résidus basés sur la déviance du modèle de Poisson (McCullagh et Nelder, 1989, Sect. 2.4), notés  $r$ , comme suit :

$$r = \text{sign}(D - \hat{D}) * \sqrt{2[D \ln(\frac{D}{\hat{D}}) - D + \hat{D}]}. \quad (2.5)$$

Nous générons ensuite un nouvel ensemble de résidus bootstrap, notés  $r_b^*$ , qui sont le résultat de tirages aléatoires avec remise. La substitution des résidus du modèle Poisson  $r$  par les résidus bootstrap  $r_b^*$  dans l'équation (2.5) permet de la réorganiser de la façon suivante :

$$\hat{D} - D \ln(\hat{D}) + r_b^{*2} + D - D \ln(D) = 0. \quad (2.6)$$

La solution numérique à l'équation (2.6) nous fournit une nouvelle matrice de décès bootstrap, notés  $\hat{D}_b^*$ . Nous procédons subséquemment au lissage par P-splines à partir de  $\hat{D}_b^*$  et des effectifs de la population soumise au risque de décéder,  $E$ . Nous en dégageons une fonction de densité lissée bootstrap, de laquelle découlent nos indices de mortalité  $\hat{e}_{25}^*$ ,  $\hat{e}_{65}^*$ ,  $\hat{M}^*$  et  $SD(\widehat{M+})^*$ . Cette procédure a été répétée 5 000 fois afin de générer 5 000 valeurs bootstrap pour chacun de nos indices de mortalité, nous permettant ainsi de construire un intervalle de confiance de type bootstrap à 95% autour de chaque valeur estimée, en retenant les percentiles 0,025 et 0,975 comme limites inférieure et supérieure.

# Chapitre 3 : Présentation et interprétation des résultats

Ce troisième chapitre a pour objectif de présenter les principaux résultats de notre mémoire. Dans la première section, nous procédons à une triple validation de nos résultats. D’abord, nous évaluons si la CSERCan offre un portrait réaliste de la population québécoise. Puis, nous confirmons l’importance de tenir compte de l’émigration observée au sein de la CSERCan dans le calcul de nos taux spécifiques de mortalité. Enfin, nous validons le produit du lissage par P-splines sur la distribution des décès par âge, sexe et langue.

La deuxième section comporte les résultats de notre analyse sur la mortalité différentielle selon le groupe linguistique et le sexe, pour l’ensemble de la province. Nous déterminons s’il existe des différences significatives de mortalité aux *âges adultes* et aux *grands âges* selon l’appartenance linguistique au Québec, à partir de nos quatre indices de mortalité estimés :  $\hat{e}_{25}$ ,  $\hat{e}_{65}$ ,  $\hat{M}$  et  $SD(\hat{M}+)$ . Nous y présentons également les intervalles de confiance bootstrap à 95% pour toutes les valeurs estimées.

La troisième section est consacrée aux inégalités de mortalité *selon la région*. Nous évaluons si les différences de mortalité selon l’appartenance linguistique entre l’île de Montréal et le reste du Québec sont les mêmes, en nous basant aussi sur  $\hat{e}_{25}$ ,  $\hat{e}_{65}$ ,  $\hat{M}$  et  $SD(\hat{M}+)$ . De plus, nous discutons du groupe d’allophones dans le reste du Québec, composé de deux groupes bien distincts : les allochtones et les autochtones.

## 3.1 Validation de la méthodologie

Avant de présenter nos résultats, il nous importe de valider les taux de mortalité par âge et sexe obtenus pour l’ensemble de la province à partir de la CSERCan avec ceux de la cohorte BDLC correspondante. Comme détaillé dans la section 2.2.2.2, nous avons construit une cohorte fictive, élaborée à partir de la Base de données sur la longévité canadienne (BDLC), qui s’étend du 1<sup>er</sup> janvier 1991 au 31 décembre 2011. Nous avons ainsi pu comparer, pour chaque sexe, l’évolution des taux de mortalité par âge pour la CSERCan et la cohorte BDLC. Deux séries de

taux de mortalité cohérents confirmeraient que : (1) notre méthode pour calculer l'exposition au risque de décéder dans la CSERCan est valide ; (2) l'allure des deux distributions des décès qui en découlent<sup>23</sup> le sont aussi ; (3) la prise en compte de l'émigration dans la CSERCan est pertinente. Enfin, nous validons le résultat du lissage par P-splines sur la distribution des décès par âge, sexe et langue en comparant notre fonction de densité lissée, qui décrit cette dernière, avec les proportions des décès à chaque âge, obtenus à partir de la construction d'une table de mortalité.

### 3.1.1 Les taux spécifiques de mortalité

La figure 3.1 illustre l'allure similaire des courbes de taux de mortalité par âge de la cohorte BDLC et de la CSERCan (avec et sans prise en compte de l'émigration), pour les femmes et les hommes séparément. Comme on pouvait l'anticiper, les taux de mortalité sont légèrement inférieurs dans la CSERCan par rapport à la cohorte BDLC, et ce, à tous les âges et pour les deux sexes, entre autres parce que les pensionnaires sont exclus.<sup>24</sup> Aux âges avancés, en particulier pour les femmes, les taux de mortalité de la CSERCan calculés en *ignorant* le phénomène d'émigration sont visiblement inférieurs aux taux de la BDLC, et à ceux pour la CSERCan qui tiennent compte de l'émigration.

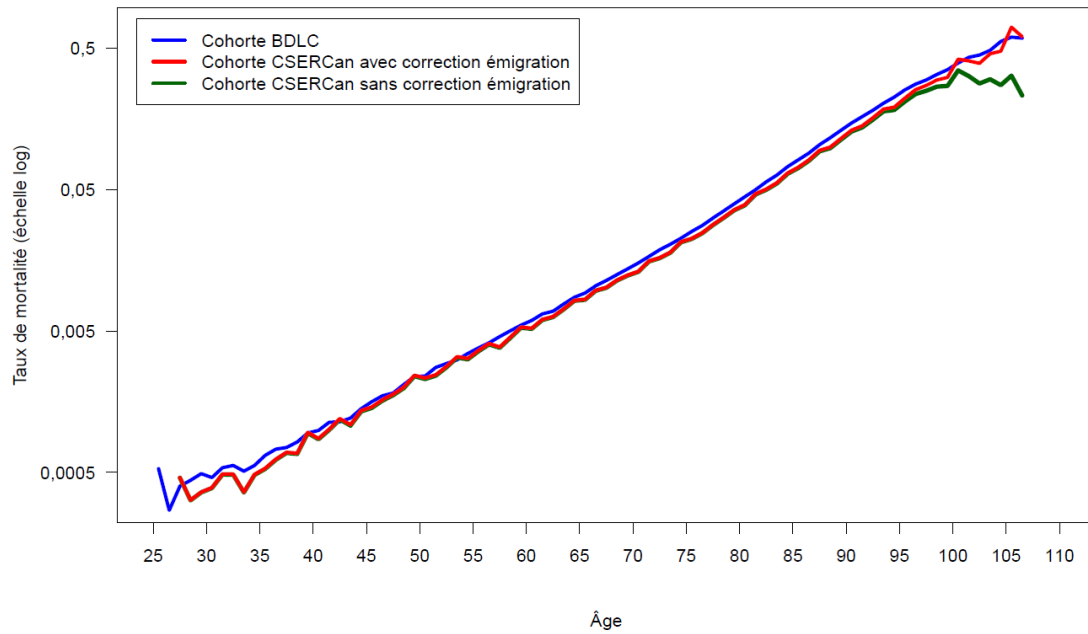
---

<sup>23</sup> Les graphiques des distributions des décès, par sexe, pour la CSERCan et la cohorte BDLC sont disponibles dans la figure annexe C.1.

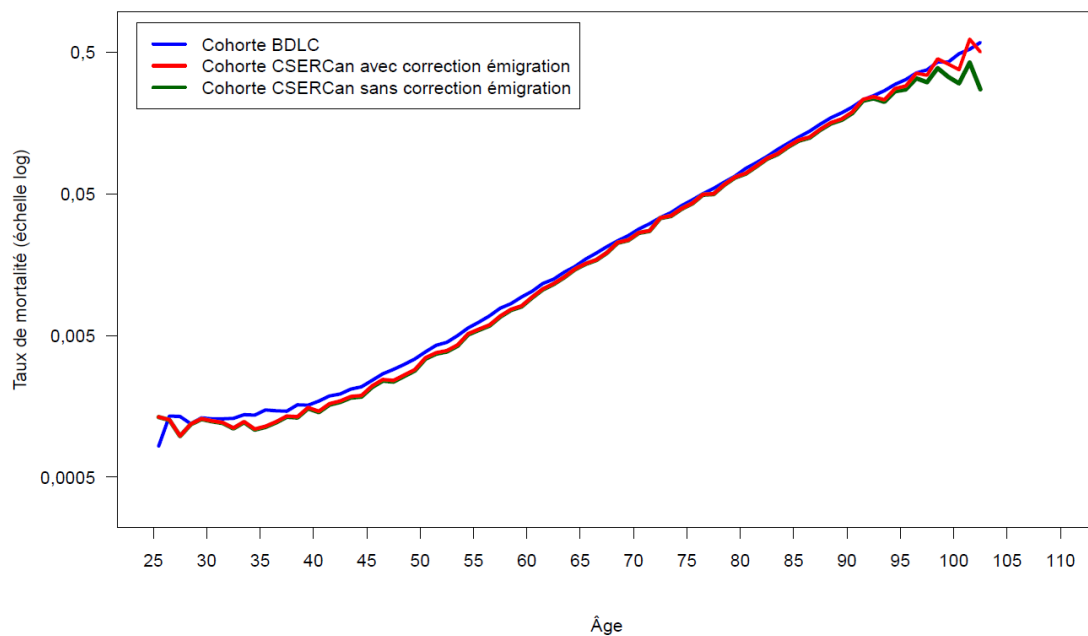
<sup>24</sup> Voir les limites de la base de données (section 2.1.1.3) pour plus de détails sur les autres facteurs pouvant expliquer l'écart entre les deux séries de taux.

Figure 3.1 Comparaison des taux de mortalité selon l'âge et le sexe des cohortes BDLC et CSERCan avec et sans émigration, Québec, 1991 à 2011

**a. Femmes**



**b. Hommes**



Chez les femmes (figure 3.1a), les taux obtenus en tenant compte de l'émigration dans la CSERCan se rapprochent des taux de la cohorte BDLC à tous les âges. Vers l'âge de 95 ans cependant, les taux CSERCan qui s'appuient sur des expositions au risque de décéder exemptes

d'émigration s'éloignent considérablement des taux BDLC. Cette divergence s'accroît avec l'âge car les effectifs de décès (numérateur du taux) vont en diminuant, alors que les individus ayant émigré, qui continuent d'être comptabilisés dans les expositions au risque de décéder (dénominateur), vont en augmentant. Comme démontré dans la section 2.2.1.2, les individus ayant émigré du Canada sont considérés comme étant « immortels » dans la base de données de la CSERCan, ce qui contribue à biaiser les taux de mortalité vers le bas. Ainsi, leur fausse contribution dans les expositions au risque de décéder aux âges suivant leur émigration gonfle artificiellement ces expositions, affectant particulièrement les taux aux âges avancés où les effectifs de décès sont de plus en plus petits.

Chez les hommes (figure 3.1b), le même phénomène s'observe, mais avec moins d'acuité. Les taux de la CSERCan corrigés par la prise en compte de l'émigration se situent tout de même plus près des taux de la cohorte BDLC que les taux non corrigés à tous les âges, et de façon plus évidente à partir de 93 ans.

En somme, la correction de l'émigration sur la CSERCan aide énormément à minimiser le biais aux âges avancés. La correction nous évite par ailleurs de censurer systématiquement les données à un âge fixe auquel le biais deviendrait criant. Comme nous le verrons dans les prochaines figures illustrant les fonctions de densité et de survie de chaque groupe linguistique suite au lissage par P-splines, nous avons pu utiliser les données jusqu'au dernier âge où nous retrouvions plus de 10 décès, comme le proposent Horiuchi et collègues (2013).<sup>25</sup>

### **3.1.2 Validation du lissage par P-splines**

Nous examinons ici la validité du lissage par P-splines. La figure 3.2 illustre les distributions de l'âge au décès obtenues pour la CSERCan avec la construction d'une table de mortalité<sup>26</sup> et le lissage par P-splines, selon le groupe linguistique<sup>27</sup> et le sexe pour l'ensemble de la province. Les deux courbes tiennent compte des expositions au risque par âge, ce qui permet une comparaison appropriée par âge et sexe. Nous observons que le modèle de régression

---

<sup>25</sup> Pour plus de détails, consulter la section 2.2.4.

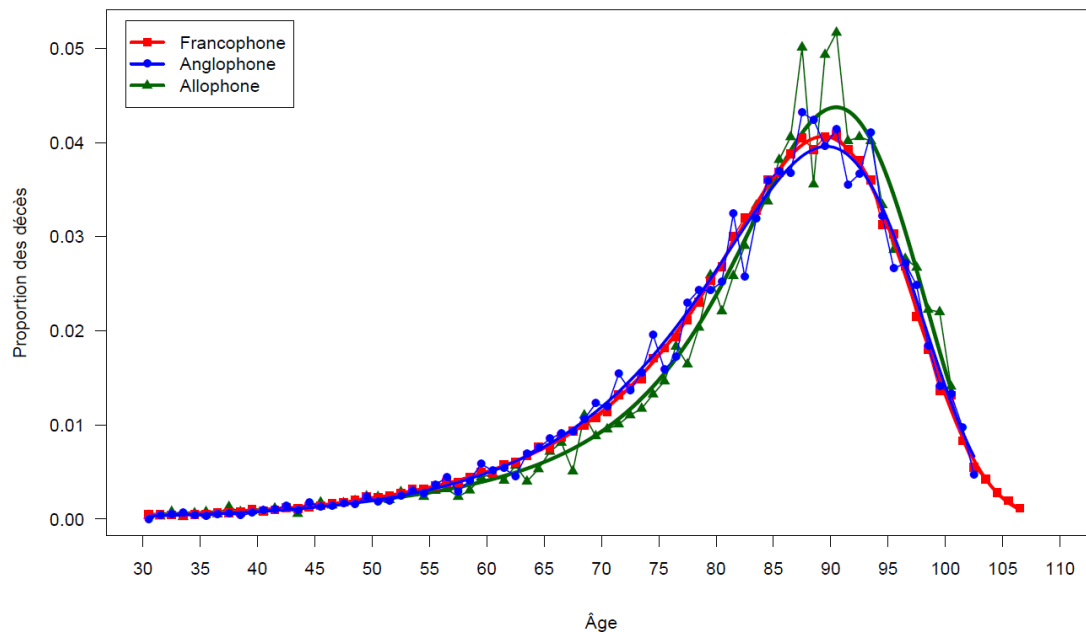
<sup>26</sup> La méthode actuarielle a été utilisée pour obtenir les quotients de la table de mortalité.

<sup>27</sup> Le groupe d'allophones comporte le groupe d'allochtones et d'autochtones dans la région du reste du Québec. La comparaison entre les proportions de décès de la table et les fonctions de densité lissée par P-splines de ces deux sous-groupes est disponible dans la figure annexe D.1.

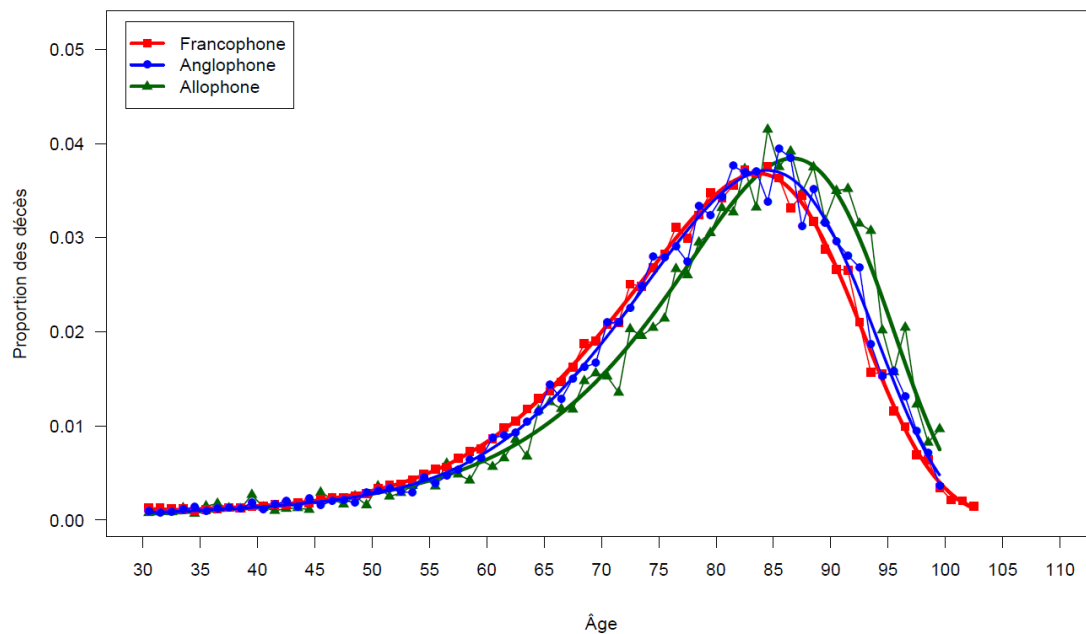
de Poisson estimé par P-splines correspond à une représentation à la fois fidèle et plus régulière des données de la table de mortalité.

Figure 3.2 Comparaison entre les proportions des décès tirées de la table de mortalité (points et lignes superposés) et les fonctions de densité lissées (lignes pleines) résultant de l'approche par P-splines, selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011

### a. Femmes



### b. Hommes



## 3.2 Mortalité différentielle au Québec

Cette deuxième section comporte les réponses à certaines de nos questions de recherche, soit celles relatives à la présence de différentiels de mortalité significatifs selon l'appartenance linguistique au Québec aux âges adultes et aux grands âges, et la comparaison des dispersions des durées de vie aux grands âges.

Rappelons que pour la CSERCan, les décès et les expositions au risque de décéder selon le groupe linguistique proviennent du jumelage des données du recensement de 1991, qui comporte l'information linguistique, de la Base canadienne de données sur la mortalité (BCDM) et du Fichier de données fiscales sommaires historiques (FDFSH), allant de 1991 à 2011. Afin de reproduire des résultats comparables avec les études antérieures, nous avons retenu trois grandes catégories linguistiques dans cette section, soit les francophones, les anglophones et les allophones. Pour l'ensemble du Québec, la CSERCan est composée de 84% de francophones, 11% d'anglophones et 5% d'allophones. Cette distribution linguistique<sup>28</sup> est la même chez les deux sexes.

Nous n'avons pas consacré de section spécifique pour le groupe d'autochtones du Québec en raison de leurs petits effectifs et de leur quasi-absence dans l'échantillon de l'île de Montréal. Nous apportons néanmoins quelques commentaires concernant leur présence importante dans le groupe d'allophones du RDQ (section 3.3). De plus, nos estimations des indicateurs de mortalité aux âges adultes et avancés pour les groupes d'allochtones et d'autochtones pour le Québec sont présentées dans l'annexe D. Pour mesurer les différentiels de mortalité, rappelons que les indices de mortalité retenus ( $\hat{e}_{25}$ ,  $\hat{e}_{65}$ ,  $\hat{M}$  et  $SD(\hat{M}+)$ ) sont dérivés des fonctions de densité lissée présentées ci-dessous. Les fonctions de survie lissées correspondantes sont également présentées, par sexe et groupe linguistique.

Dans les prochaines figures, nous constatons, chez les deux sexes, des proportions plus élevées de survivants allophones à l'âge modal au décès (figure 3.3) dont les courbes de survie sont plus rectangulaires (figure 3.4) que celles des deux autres groupes. Les différences dans la distribution des décès des femmes francophones et anglophones sont quant à elles à peine

---

<sup>28</sup> Plus de détails concernant la méthode de répartition des langues multiples sont fournis dans la section 2.2.1.1.

perceptibles à l'œil nu, ce qui porte à croire qu'il n'y a probablement pas d'écart de mortalité statistiquement significatif entre ces deux groupes (figure 3.3a). Par contre, chez les hommes, les proportions de survivants anglophones semblent légèrement supérieures à celles des francophones (figure 3.3b), entraînant ainsi une plus grande rectangularisation de leur courbe de survie (figure 3.4b).

Figure 3.3 Fonction de densité lissée selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011

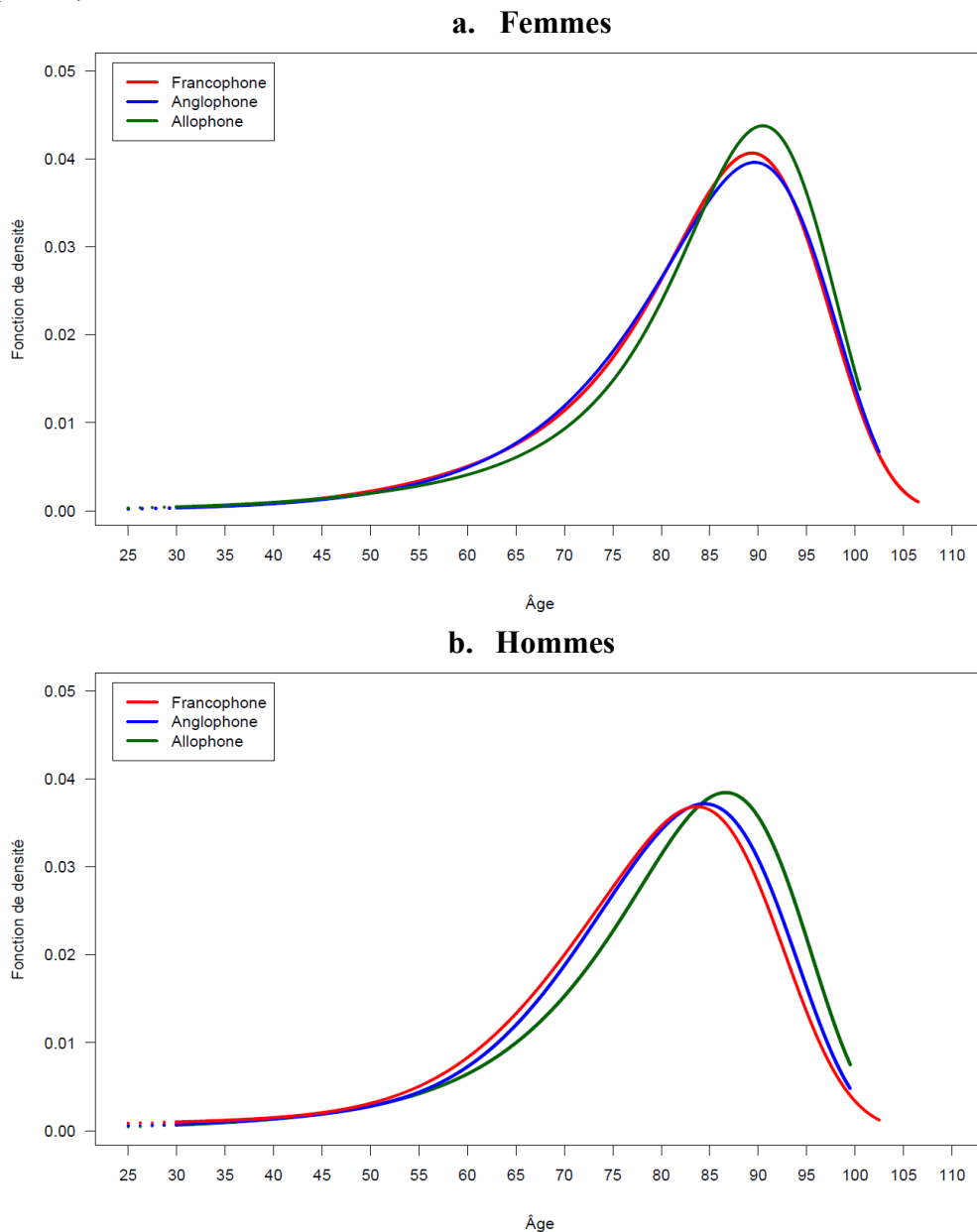
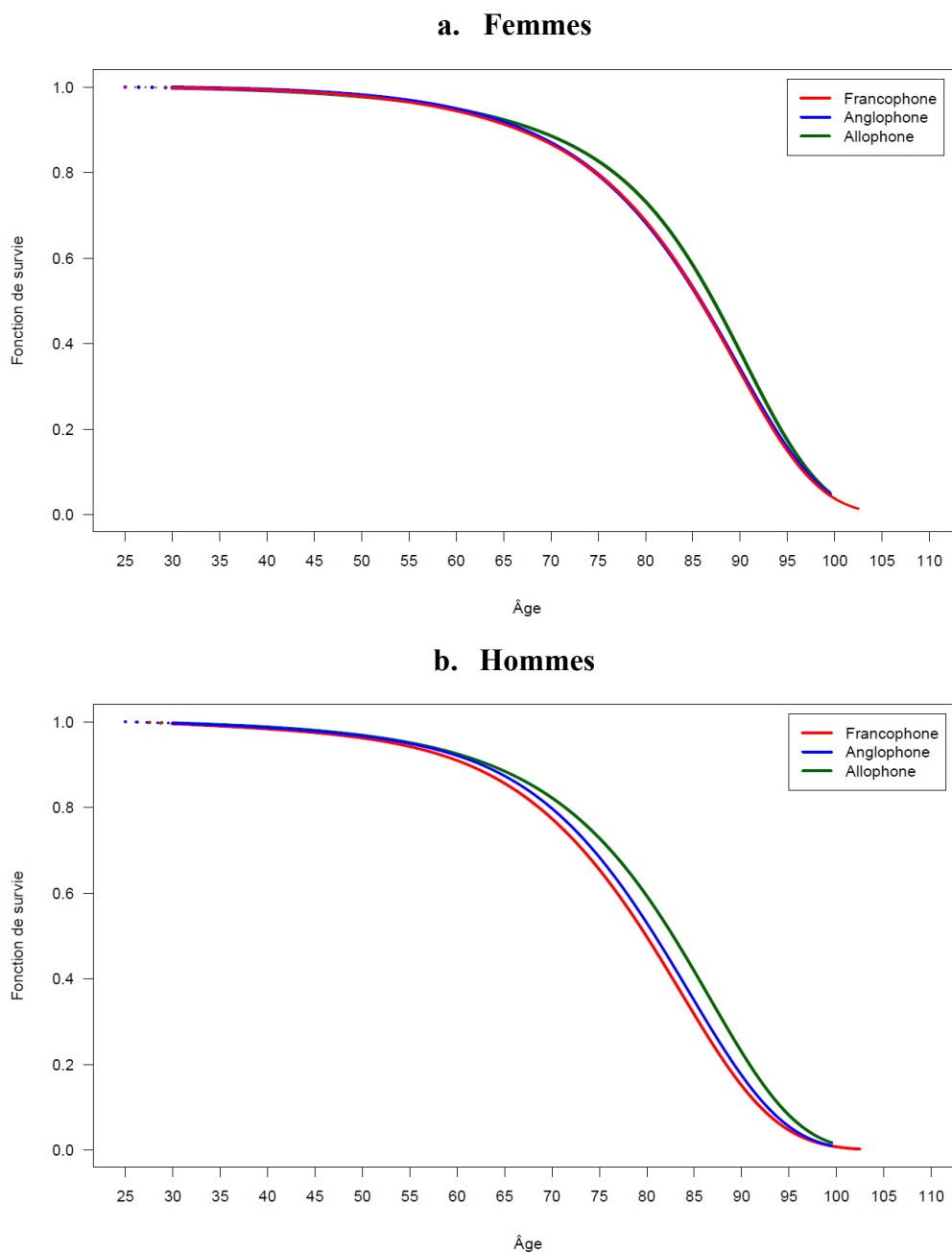




Figure 3.4 Fonction de survie lissée selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011



### 3.2.1 Mortalité différentielle aux âges adultes

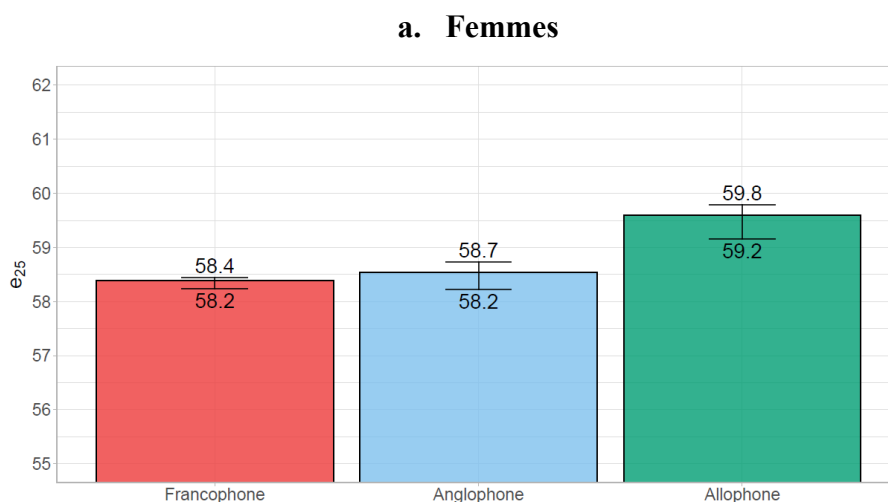
Dans le but d'étudier les différentiels de mortalité aux *âges adultes* selon l'appartenance linguistique au Québec, nous avons obtenu les valeurs pour  $\hat{e}_{25}$  à partir de la fonction de densité

lissée (figure 3.3) et calculé l'intervalle de confiance correspondant pour chaque sexe et groupe linguistique (figure 3.5).

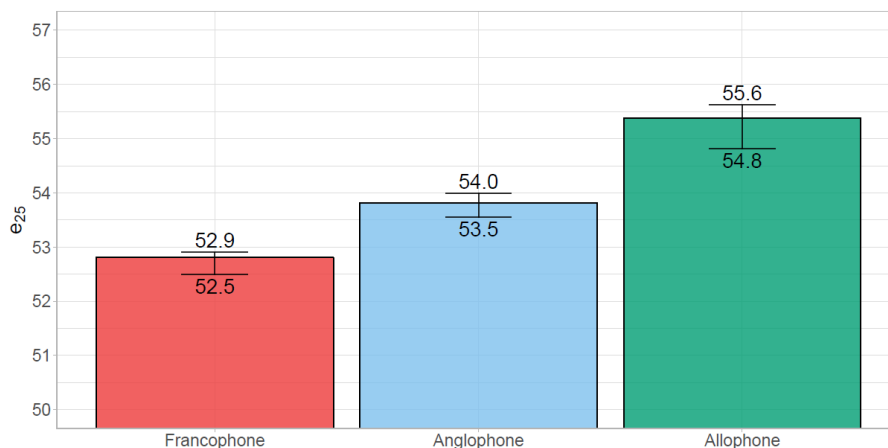
### 3.2.1.1 Espérance de vie à 25 ans

Chez les femmes, les espérances de vie à 25 ans s'élèvent à 58,4 ans pour les francophones, 58,5 ans pour les anglophones et 59,6 ans pour les allophones. La construction des intervalles de confiance autour de  $\hat{e}_{25}$  pour chaque groupe linguistique, tel qu'illustrés dans la figure 3.5a, nous permet de confirmer deux résultats probants chez les Québécoises. D'abord, la mortalité aux âges adultes des francophones et des anglophones est significativement plus élevée que celle des allophones. En effet, l'espérance de vie féminine des francophones ( $\hat{e}_{25} = 58,4$ ) et des anglophones ( $\hat{e}_{25} = 58,5$ ), respectivement, est de 1,2 et 1,1 an inférieure à celle des allophones ( $\hat{e}_{25} = 59,6$ ). Puis, la mortalité des anglophones n'est pas significativement différente de celle des francophones.

Figure 3.5 Estimations de l'espérance de vie à 25 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011



## b. Hommes



Chez les hommes, les différences d' $\hat{e}_{25}$  sont encore plus prononcées, puisque  $\hat{e}_{25}$  vaut 52,8 ans pour les francophones, 53,8 ans pour les anglophones et 55,4 ans pour les allophones (figure 3.5b). Les intervalles de confiance confirment que les avantages de 2,6 ans et de 1,6 an des allophones sur les francophones et anglophones respectivement, en plus de celui de 1,0 an des anglophones sur les francophones, sont tous significatifs au plan statistique. Ainsi, comme documenté dans la littérature, nous nous retrouvons en présence d'une surmortalité des francophones et d'une sous-mortalité des allophones, avec les anglophones qui vivent en moyenne plus longtemps que les francophones, chez les hommes de la province.

### 3.2.2 Mortalité différentielle aux grands âges

Dans notre recherche, nous abordons pour la première fois les disparités de mortalité selon la langue au Québec aux *âges avancés*. Nous avons choisi deux indicateurs de mortalité, qui sont  $\hat{e}_{65}$  et  $\hat{M}$ .

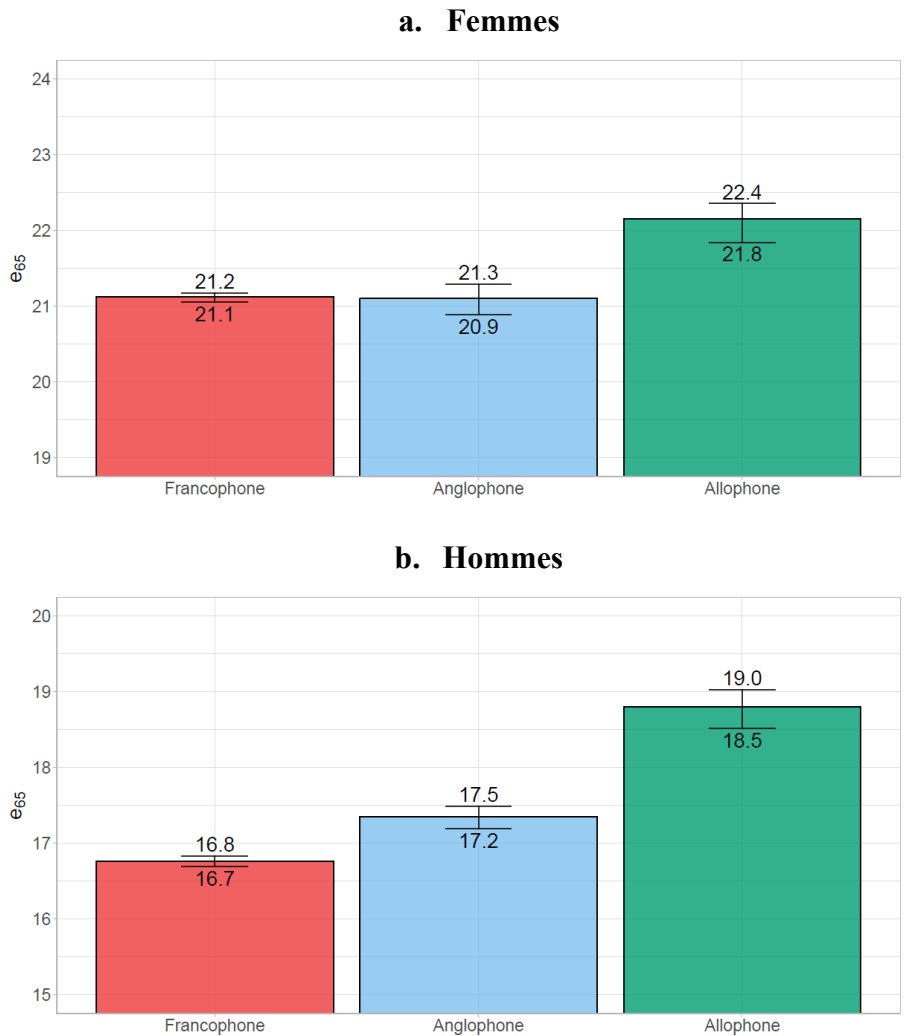
#### 3.2.2.1 Espérance de vie à 65 ans

La figure 3.6 présente  $\hat{e}_{65}$  ainsi que les intervalles de confiance par sexe et par groupe linguistique. Pour les femmes du Québec (figure 3.6a), les valeurs de  $\hat{e}_{65}$  sont de 21,1 ans pour les francophones et les anglophones, et de 22,2 ans pour les allophones. Les différences de mortalité des femmes aux âges avancés sont très similaires à celles notées aux âges adultes, c'est-à-dire que nous observons des différences de mortalité significatives entre les femmes

allophones et les francophones ou anglophones seulement, et l'avantage des allophones s'élève à 1,0 an.

Pour les hommes, les différentiels de mortalité observés aux âges adultes persistent aux âges avancés, selon nos estimations de  $\hat{e}_{65}$  (figure 3.6b). L'espérance de vie à 65 ans des hommes francophones ( $\hat{e}_{65} = 16,8$ ) est de 0,5 an et de 2,0 an inférieure à celle des hommes anglophones et allophones, respectivement. Quant aux hommes anglophones ( $\hat{e}_{65} = 17,3$ ), ils peuvent espérer vivre 1,5 an de moins à l'âge de 65 ans, par rapport aux allophones ( $\hat{e}_{65} = 18,8$ ). Tout comme pour  $\hat{e}_{25}$ , les différences sont significatives entre chaque groupe linguistique.

Figure 3.6 Estimations de l'espérance de vie à 65 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011

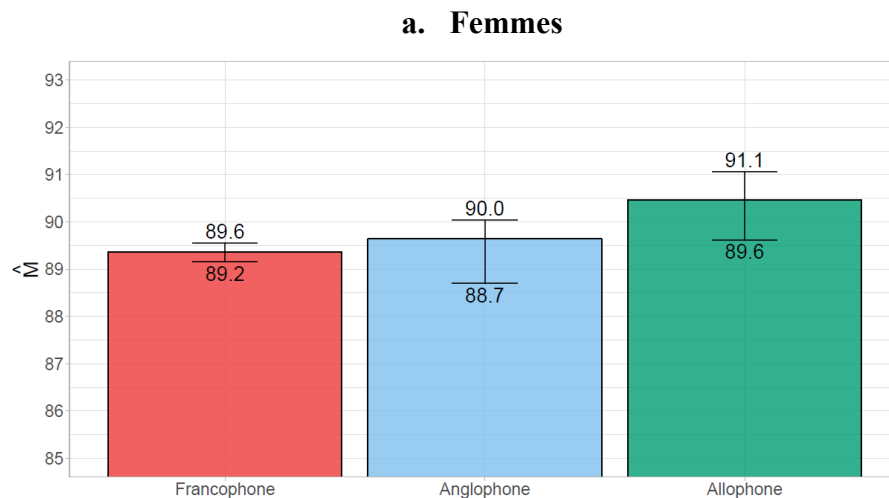


### 3.2.2.2 Âge modal au décès

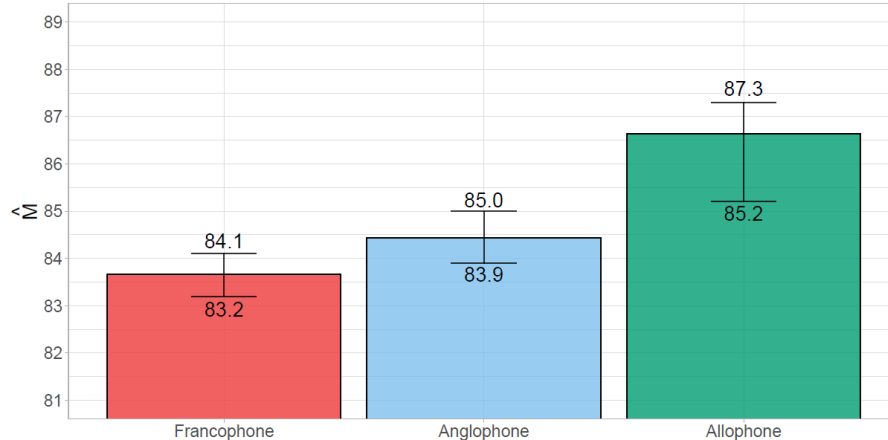
L'âge modal au décès ( $\hat{M}$ ) des groupes linguistiques ainsi que les intervalles de confiance sont présentés dans la figure 3.7, pour les femmes et les hommes au Québec. De prime abord, nous observons chez les deux sexes des résultats qui sont similaires à ceux pour  $\hat{e}_{65}$ , soit que la mortalité semble être la plus élevée chez les francophones, et la moins élevée chez les allophones. Toutefois, d'après les intervalles de confiance, ces différentiels de mortalité ne sont pas tous statistiquement significatifs.

Pour les femmes (figure 3.7a), l'âge modal au décès est pratiquement le même chez les anglophones ( $\hat{M} = 89,6$ ) et chez les francophones ( $\hat{M} = 89,4$ ), ce qui ressemble à ce que nous avons trouvé pour  $\hat{e}_{65}$ . Les femmes allophones, quant à elles, détiennent un âge modal au décès qui est de 1,1 an et 0,9 an plus élevé que celui des francophones et anglophones, sans toutefois que ces écarts soient significatifs, contrairement à ce qui avait été constaté pour  $\hat{e}_{65}$ . Ainsi, en nous basant sur l'âge modal au décès, il n'existe pas de différences de mortalité significatives selon le groupe linguistique chez les femmes du Québec.

Figure 3.7 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011



### b. Hommes



Chez les hommes, l'écart de mortalité aux grands âges mesuré par  $\widehat{M}$  demeure significatif entre le groupe d'allophones et les groupes de francophones et anglophones (figure 3.7b). L'âge modal au décès estimé pour les hommes allophones ( $\widehat{M} = 86,6$ ) surpasse de 2,9 ans et 2,1 ans respectivement, ceux estimés pour les francophones ( $\widehat{M} = 83,7$ ) et les anglophones ( $\widehat{M} = 84,5$ ). Toutefois, contrairement aux résultats pour  $\hat{e}_{65}$ , l'âge modal au décès des hommes francophones n'est pas significativement différent au plan statistique de celui des anglophones.

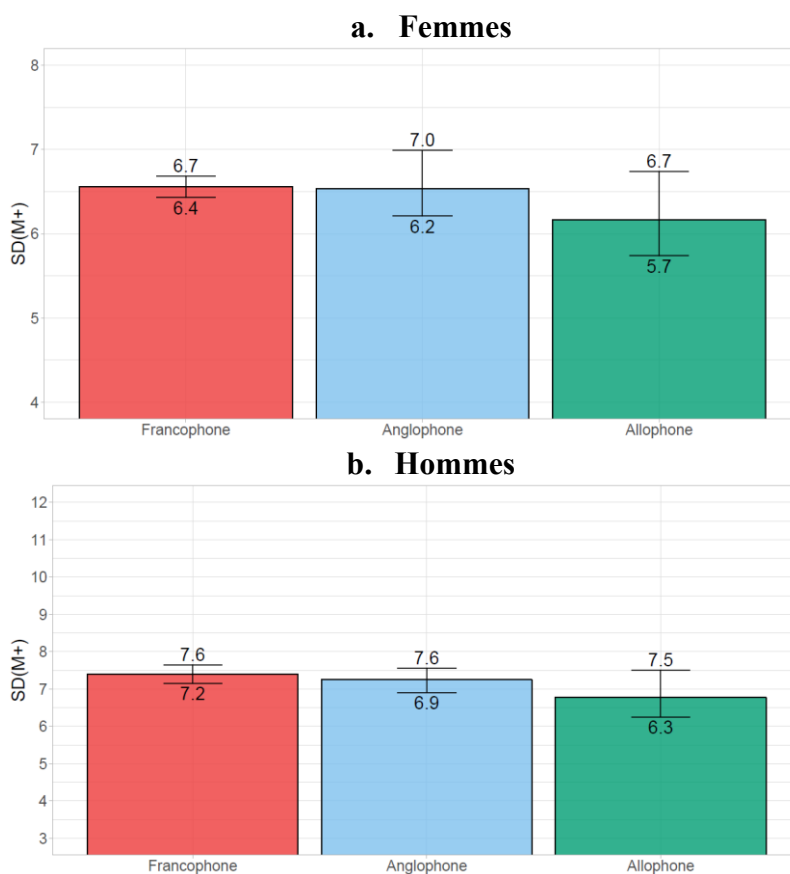
#### 3.2.2.3 Écart-type au-delà du mode

Afin de décrire le degré d'hétérogénéité des durées de vie aux âges avancés selon le groupe linguistique, nous avons estimé l'écart-type des durées de vie au-delà de l'âge modal au décès,  $SD(\widehat{M} +)$ , qui est particulièrement utile pour effectuer des comparaisons de la dispersion des décès aux grands âges entre divers sous-groupes de la population. Ainsi, une valeur plus élevée de  $SD(\widehat{M} +)$  est associée à une plus grande variabilité dans la distribution des âges au décès et au contraire, une plus petite valeur signifie qu'une plus forte proportion d'individus décédés se concentre autour de la durée de vie la plus commune.

L'examen des figures 3.3 et 3.4 illustrant les fonctions de densité et de survie lissées au début de la section suggérait un plus haut niveau de compression de la mortalité chez les allophones que chez les francophones ou anglophones. Cette observation, visible pour les deux sexes, se basait sur les proportions plus élevées de survivants allophones à l'âge modal au décès, ainsi que sur une plus grande rectangularisation de leurs courbes de survie. Dans la figure 3.8, nous

remarquons justement que  $SD(\widehat{M} +)$  est le moins élevé parmi les femmes et les hommes allophones, ce qui confirme que les décès des personnes âgées allophones se concentrent dans un intervalle d'âge plus court que chez les francophones ou les anglophones. Néanmoins, au plan statistique, les valeurs de  $SD(\widehat{M} +)$  pour chacun des groupes linguistiques et chaque sexe ne sont pas significativement différentes l'une de l'autre au seuil de 5%.

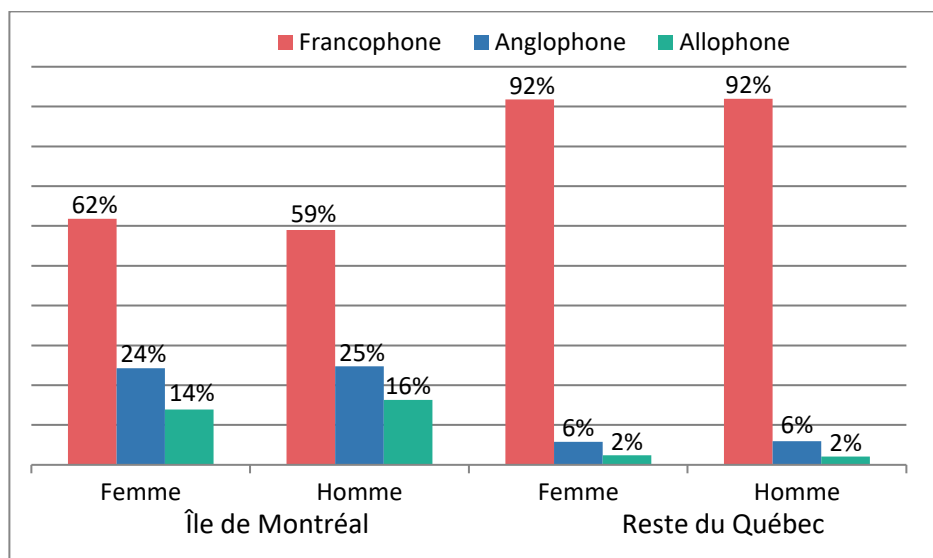
Figure 3.8 Estimations de l'écart-type au-delà du mode et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011



### 3.3 Mortalité différentielle selon la région

Cette troisième section vise à répondre à notre dernière question de recherche, à savoir si les différences de mortalité selon l'appartenance linguistique sont les mêmes sur l'île de Montréal (MTL) et dans le reste du Québec (RDQ).<sup>29</sup> L'intérêt de cette question provient de la composition différente des groupes linguistiques entre MTL et le RDQ, qui dévoile des dynamiques et des comportements démolinguistiques distincts et propres à chaque région. La figure 3.9 illustre la distribution linguistique par sexe de la cohorte pour MTL et le RDQ. La population échantillonnée de MTL est composée, pour les femmes et les hommes respectivement, de 62% et 59% de francophones, de 24% et 25% d'anglophones et de 14% et 16% d'allophones. Pour le RDQ, la composition linguistique est la même pour les deux sexes, avec 92% de francophones, 6% d'anglophones et 2% d'allophones.

Figure 3.9 Distribution (en %) des membres de la CSERCan par sexe, groupe linguistique et région, Île de Montréal et Reste du Québec, 1991 à 2011



Dans le but de comparer les trois groupes linguistiques entre les deux régions, nous avons utilisé le terme « allophone » pour décrire un groupe qui est en réalité, considérablement différent dans chaque région. Le groupe d'allophones dans la région de MTL est composé uniquement

<sup>29</sup> Notez que l'ensemble des résultats pour chaque région est disponible dans l'annexe E.

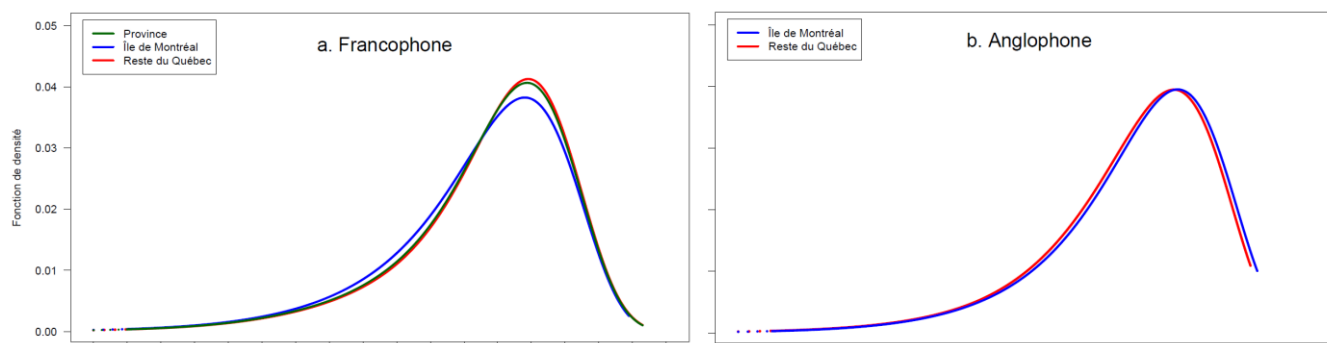


d'allochtones<sup>30</sup>, alors que dans le RDQ, les allophones sont composés de 54% d'allochtones et 46% d'autochtones chez les femmes, et de 65% d'allochtones et 35% d'autochtones chez les hommes. Également, nous dévoilons d'emblée qu'aucune différence significative régionale n'est notée dans la mortalité différentielle des allochtones, et ce, chez les deux sexes. Ainsi, cette section vise principalement à souligner les différences de mortalité entre les deux régions, pour chaque groupe linguistique. Nous débutons par commenter les disparités observées chez les femmes, et poursuivons avec les hommes.

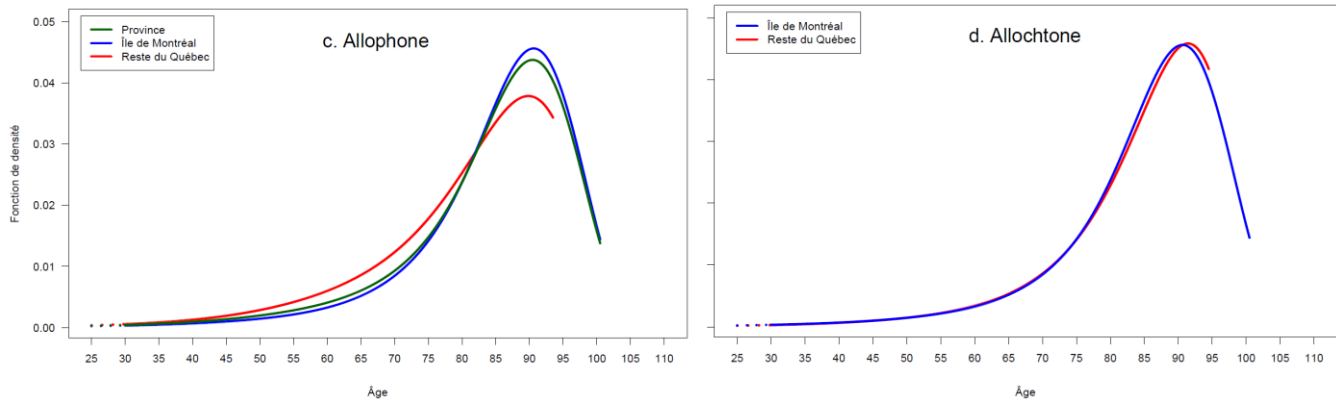
### 3.3.1 Différences régionales chez les femmes

Les fonctions de densité lissée selon la région et le groupe linguistique, qui décrivent les diverses distributions des décès par âge, permettent de bien visualiser les différences de mortalité régionales chez les femmes (figure 3.10). Les courbes révèlent de possibles différences régionales entre les femmes francophones et allophones de MTL et du RDQ. Effectivement, aux âges adultes, les francophones de MTL (ligne bleue) et les allophones du RDQ (ligne rouge) possèdent une proportion plus élevée de décès par âge que leurs homologues de la région adverse, jusqu'à l'âge de 80 ans environ. Cela porte à croire qu'il y aurait une plus grande hétérogénéité des décès pour ces sous-groupes de la population. En revanche, les différences régionales chez les femmes anglophones et allochtones paraissent négligeables.

Figure 3.10 Fonction de densité lissée pour les femmes selon la région et le groupe linguistique, CSERCan, 1991 à 2011



<sup>30</sup> Nous y avons exclu le très petit effectif d'individus autochtones inclus dans l'échantillon de MTL.

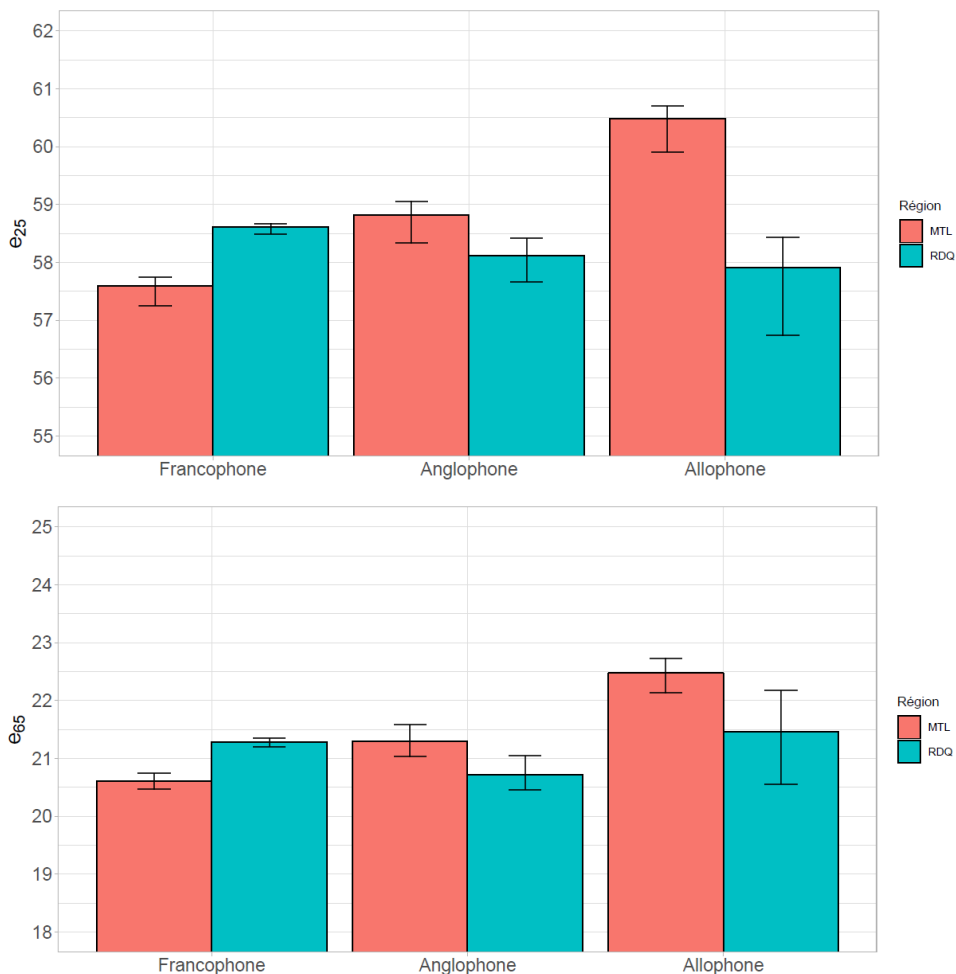


La figure 3.11 présente les espérances de vie estimées à 25 ans et à 65 ans ainsi que leurs intervalles de confiance, par groupe linguistique et sexe, pour MTL et le RDQ. L'espérance de vie à 25 ans ( $\hat{e}_{25} = 57,6$  [57,2 ; 57,7]) et à 65 ans ( $\hat{e}_{65} = 20,6$ , [20,5 ; 20,7]) des femmes francophones vivant dans MTL est significativement plus courte que celles vivant dans le RDQ ( $\hat{e}_{25} = 58,6$  [58,5 ; 58,7];  $\hat{e}_{65} = 21,3$  [21,2 ; 21,3]). Chez les allophones, la différence est significative pour  $\hat{e}_{25}$  uniquement, avec une plus forte mortalité enregistrée dans le RDQ ( $\hat{e}_{25} = 57,9$  [56,7 ; 58,4]) que dans MTL ( $\hat{e}_{25} = 60,5$  [59,9 ; 60,7]). L'écart important d' $\hat{e}_{25}$  entre les deux régions (2,6 ans) s'explique par la présence importante d'individus autochtones, dont la mortalité est connue pour être nettement supérieure à la moyenne, dans le groupe d'allophones du RDQ (figure 3.11a). À 65 ans et plus cependant, l'écart d' $\hat{e}_{65}$  n'est pas significatif (figure 3.11b). Nous croyons que la forte mortalité des autochtones aux âges adultes contribue à faire baisser  $\hat{e}_{25}$  de façon plus importante que pour  $\hat{e}_{65}$ . Ce constat fait en sorte qu'aux âges plus avancés, la part de femmes allochtones dans le groupe d'allophones du RDQ augmente et contribue davantage à hausser la moyenne des durées de vie du groupe.

Pour revenir à la figure 3.11a, il est intéressant de noter que dans le RDQ, l'espérance de vie à 25 ans des femmes francophones est plus élevée que celle des anglophones et des allophones. Cette différence est tout juste significative entre les francophones ( $\hat{e}_{25} = 58,6$  [58,5 ; 58,7]), et les allophones ( $\hat{e}_{25} = 57,9$  [56,7 ; 58,4]) et les anglophones ( $\hat{e}_{25} = 58,1$  [57,7 ; 58,4]). Il convient de mentionner que la présence d'autochtones et d'allochtones dans le groupe d'allophones, dont les tendances de mortalité sont opposées, explique le plus grand intervalle de confiance chez les femmes allophones du RDQ. La mortalité significativement plus élevée de celles-ci comparées aux francophones dans le RDQ contredit la croyance que les allophones

au Québec possèdent une mortalité généralement plus faible. Toutefois, la figure 3.11b montre que cette tendance n'est pas maintenue pour  $\hat{e}_{65}$ , pour les raisons évoquées au paragraphe précédent.

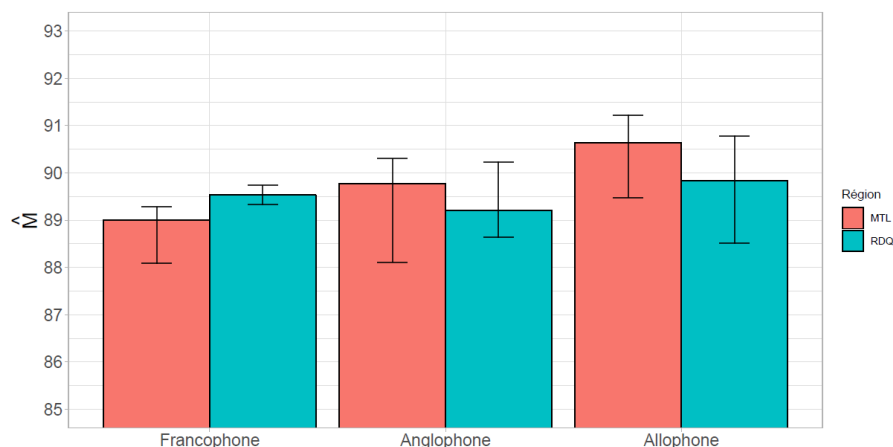
Figure 3.11 Estimations de l'espérance de vie à 25 ans (a) et 65 ans (b), et des intervalles de confiance bootstrap à 95% pour les femmes selon la région et le groupe linguistique, CSERCan, 1991 à 2011



Nos estimations de l'âge modal au décès,  $\hat{M}$ , ne révèlent pas non plus de différences notables aux âges avancés chez les femmes entre les régions (figure 3.12). Seul l'écart de 0,5 an entre les francophones dans le RDQ ( $\hat{M} = 89,5 [89,3 ; 89,7]$ ) et celles dans MTL ( $\hat{M} =$

89,0 [88,1 ; 89,3]) est tout près d'être significatif. Enfin, nous ne constatons aucun écart important entre les deux régions en matière de  $SD(\widehat{M} +)$ .<sup>31</sup>

Figure 3.12 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% pour les femmes selon la région et le groupe linguistique, CSERCan, 1991 à 2011

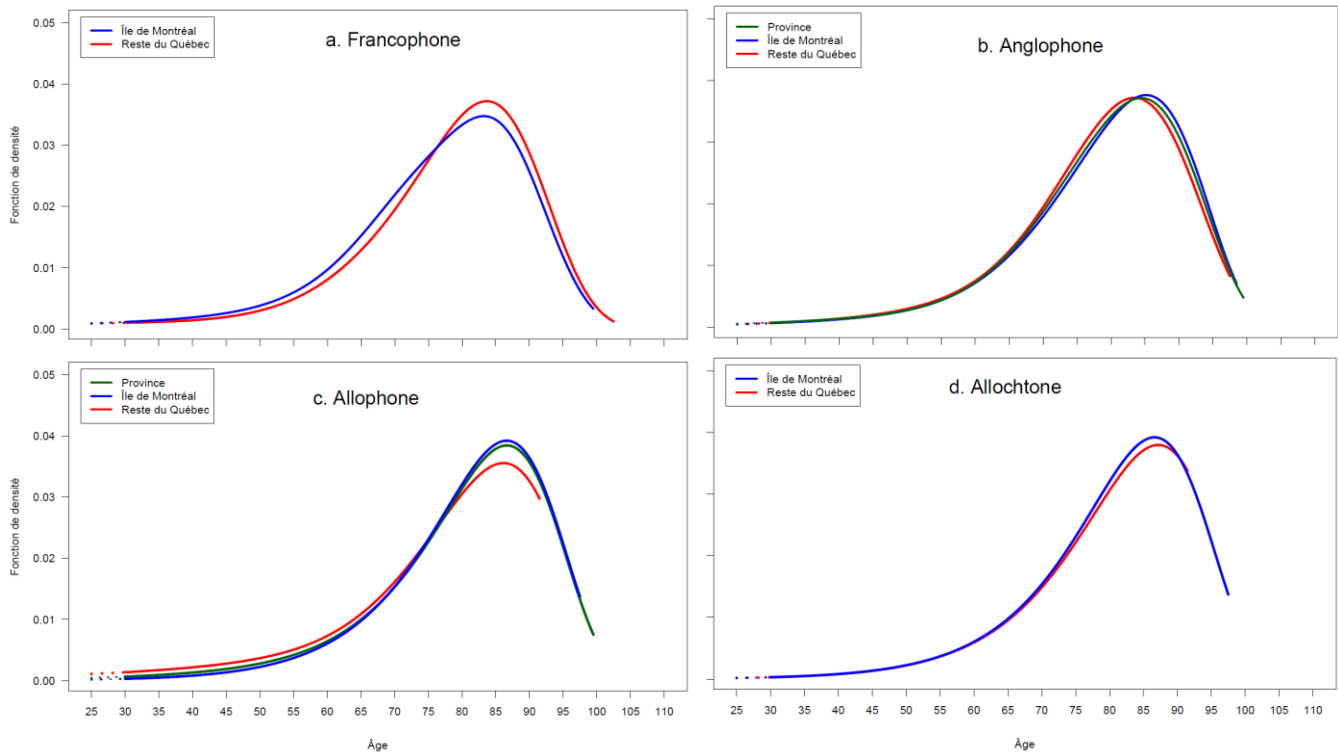


### 3.3.2 Différences régionales chez les hommes

La figure 3.13 présente les fonctions de densité lissées selon la région et le groupe linguistique, chez les hommes. Parmi les francophones, nous voyons d'abord que ceux de MTL (ligne bleue) sont plus susceptibles de décéder avant l'âge modal au décès par rapport à ceux du RDQ (ligne rouge), d'où la raison pour laquelle nous observons moins de survivants à  $\widehat{M}$  et de plus grandes proportions de décès par âge jusqu'aux alentours de 75 ans dans MTL. Chez les anglophones, ce sont plutôt ceux dans le RDQ (ligne rouge) qui semblent détenir une mortalité plus élevée aux âges adultes et avancés. Enfin, les allophones du RDQ (ligne rouge) comparés à ceux de MTL (ligne bleue) sont désavantagés sur le plan de la mortalité, notamment en raison de la présence importante d'autochtones. En effet, les différences régionales sont minimales pour les hommes allochtones, c'est-à-dire lorsque nous excluons les autochtones parmi les allophones.

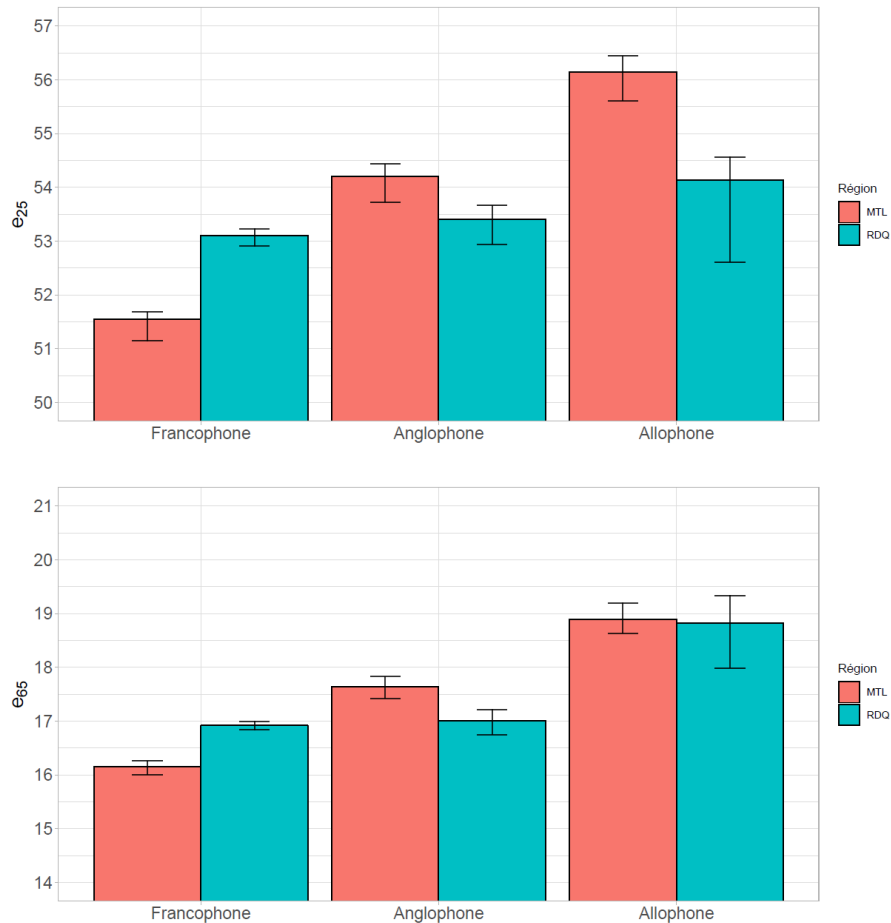
<sup>31</sup> Le graphique des différences régionales pour  $SD(\widehat{M} +)$  est disponible à la figure annexe E.5.

Figure 3.13 Fonction de densité lissée pour les hommes selon la région et le groupe linguistique, CSERCan, 1991 à 2011



La figure 3.14 présente les différences régionales pour les hommes francophones, anglophones et allophones en ce qui concerne  $\hat{e}_{25}$  et  $\hat{e}_{65}$ . En premier lieu, les francophones de MTL ( $\hat{e}_{25} = 51,5$  [51,1 ; 51,7];  $\hat{e}_{65} = 16,1$  [16,0 ; 16,3]) sont désavantagés, autant pour  $\hat{e}_{25}$  que pour  $\hat{e}_{65}$ , avec des indicateurs de 1,6 et 0,8 an inférieurs à ceux du RDQ ( $\hat{e}_{25} = 53,1$  [52,9 ; 53,2];  $\hat{e}_{65} = 16,9$  [16,8 ; 17,0]). Dans le groupe allophone, les hommes du RDQ ( $\hat{e}_{25} = 54,1$  [52,6 ; 54,6]) possèdent une espérance de vie à 25 ans significativement plus faible que ceux de MTL ( $\hat{e}_{25} = 56,1$  [55,6 ; 56,4]). Cependant, parmi les hommes de 65 ans et plus, les écarts d' $\hat{e}_{65}$  entre les régions ne sont pas significatifs, tout comme c'était le cas pour les femmes allophones et pour les mêmes raisons (section 3.3.1). Enfin, chez les hommes anglophones, un écart significatif est constaté pour  $\hat{e}_{65}$  seulement: ceux du RDQ accusent un retard de 0,6 an sur leurs homologues de MTL.

Figure 3.14 Estimations de l'espérance de vie à 25 ans (a) et 65 ans (b), et des intervalles de confiance bootstrap à 95% pour les hommes selon la région et le groupe linguistique, CSERCan, 1991 à 2011



En matière d'âge modal au décès, tout comme chez les femmes, nous n'observons aucune différence régionale chez les hommes (figure 3.15). Même le désavantage de 1,8 an des anglophones dans le RDQ ( $\hat{M} = 83,4 [82,8 ; 84,2]$ ) comparé à MTL ( $\hat{M} = 85,2 [84,2 ; 86,0]$ ) n'est pas significatif. À noter, la largeur importante de l'intervalle de confiance autour de  $\hat{M}$  pour les hommes allophones du RDQ dépeint à la fois une grande hétérogénéité des tendances de mortalité aux âges avancés au sein du groupe, et des petits effectifs de décès et de la population soumise au risque de décéder. En effet, conformément aux résultats présentés à la figure 3.15, la figure 3.16 montre que l'âge modal au décès des hommes allophones du RDQ (ligne verte) est bel est bien plus élevé (décalage vers la droite) que celui des francophones et des anglophones, mais que leur niveau de compression des durées de vie est moins important

(plus faible concentration des décès autour de  $\widehat{M}$ ). La valeur prise par  $SD(\widehat{M} +)$  est effectivement supérieure chez les allophones, sans qu'il n'y ait d'écart significatif entre les deux régions.<sup>32</sup>

Figure 3.15 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% pour les hommes selon la région et le groupe linguistique, CSERCan, 1991 à 2011

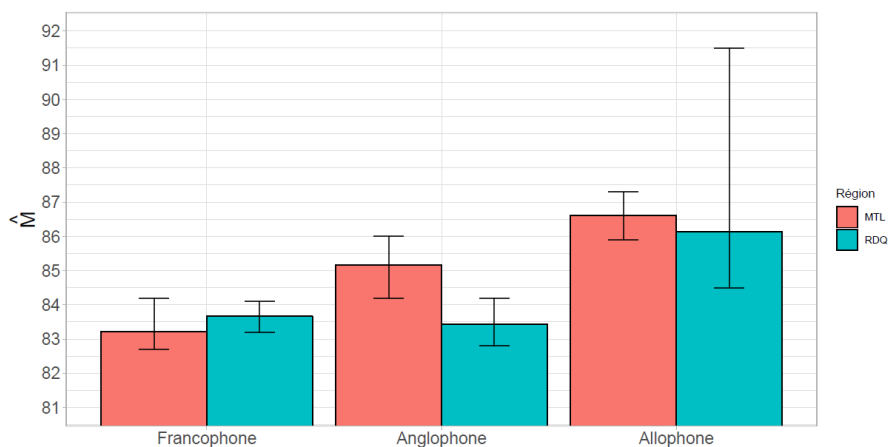
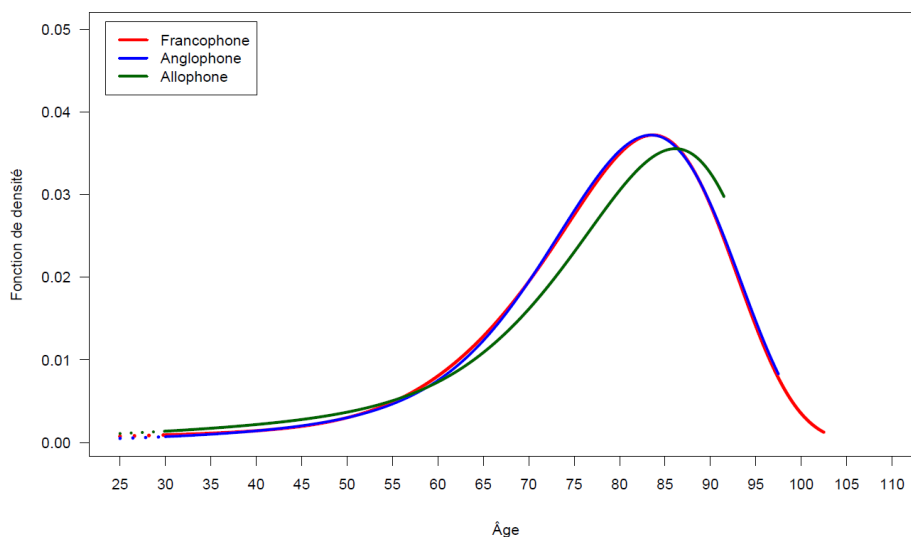


Figure 3.16 Fonction de densité lissée pour les hommes selon le groupe linguistique, CSERCan, Reste du Québec, 1991 à 2011



<sup>32</sup> Le graphique des différences régionales pour  $SD(\widehat{M} +)$  est disponible à la figure annexe E.5.

# Discussion et conclusion

Cette dernière partie du mémoire contient une interprétation plus approfondie des résultats qui découlent de notre étude, en faisant ressortir nos questions de recherche. Nous entamons avec un bilan comparatif de l'espérance de vie à 25 ans selon la langue, issue du présent mémoire et de la littérature. Nous présentons ensuite les inégalités de mortalité aux âges avancés suivies des profils sociodémographiques et économiques dans chaque région. Les limites et les futures pistes de recherche sont enfin décrites.

## Validation et actualisation de la recherche

Un des objectifs de cette recherche consistait à actualiser les observations sur la mortalité différentielle aux âges adultes selon le groupe linguistique, à partir d'une source de données fiable, soit la CSERCan. Nous dressons ici un bilan comparatif de nos valeurs pour  $\hat{e}_{25}$  avec celles publiées dans l'article de Trempe et collègues (2014).<sup>33</sup> Rappelons que leurs travaux concluaient que l'information sur la langue d'usage à la maison inscrite sur les certificats de décès convenait suffisamment pour dresser un portrait réaliste de la mortalité par groupe linguistique au Québec. Cette conclusion se basait sur des écarts de mortalité similaires entre les groupes linguistiques, entre leurs estimations et celles issues d'une compilation spéciale de l'*Étude canadienne de suivi de la mortalité selon le recensement 1991-2001* (BDJ-SC), qui leur avait été fournie sur demande par Statistique Canada. En raison des multiples sources de biais liées à la validité de l'information sur les certificats de décès ainsi qu'à l'utilisation de deux variables linguistiques différentes<sup>34</sup>, nous tenions à vérifier la solidité de leurs résultats à partir de la CSERCan, dans laquelle la variable linguistique est nettement plus fiable. Or, nos doutes se sont avérés pertinents puisque nos résultats diffèrent largement des leurs.

Dans le tableau II, nous présentons les valeurs pour  $\hat{e}_{25}$  chez les hommes et les femmes, selon la source de données et le groupe linguistique.<sup>35</sup> Nous constatons que chez les hommes, les valeurs d' $\hat{e}_{25}$  issues de la BDJ-SC et de la CSERCan sont plutôt similaires. Bien que les écarts

---

<sup>33</sup> Plus de détails dans la section 1.2.4.

<sup>34</sup> La langue maternelle est utilisée pour estimer  $\hat{e}_{25}$  dans la compilation spéciale de Statistique Canada.

<sup>35</sup> Les graphiques à barre avec intervalle de confiance de ces valeurs sont présentés dans la figure annexe F.1.



observés entre les groupes linguistiques soient plus petits que ceux calculés par l’INSPQ, ceux-ci demeurent significatifs dans la CSERCan comme dans la BDJ-SC.

Cependant, le même constat ne tient pas chez les femmes, pour lesquelles nous observons des différences importantes d’ $\hat{e}_{25}$  selon la source de données. En effet, les valeurs pour  $\hat{e}_{25}$  obtenues dans le présent mémoire à l’aide de la CSERCan se situent bien en-dessous de celles issues de la BDJ-SC, ce qui est très curieux compte tenu du fait que la BDJ-SC est une version antérieure de la CSERCan. Nos valeurs sont inférieures de 1,4 an, 2,5 ans et 4,3 ans respectivement pour les femmes francophones, anglophones et allophones. Nous nous sommes donc penchés sur ces résultats pour tenter d’expliquer ce qui semble être, soit une sous-estimation de la mortalité féminine dans la BDJ-SC, soit une surestimation dans la CSERCan. Nous avons ainsi produit de nouvelles séries pour  $\hat{e}_{25}$  avec la CSERCan, en ajustant le traitement des données de trois façons.

Tableau II. Estimations de l’espérance de vie à l’âge de 25 ans par sexe et groupe linguistique selon trois sources de données, Québec

Groupe linguistique	Femmes			Hommes		
	INSPQ 1994-1998 <sup>a</sup>	BDJ-SC 1991-2001 <sup>b</sup>	CSERCan 1991-2011 <sup>c</sup>	INSPQ 1994-1998 <sup>a</sup>	BDJ-SC 1991-2001 <sup>b</sup>	CSERCan 1991-2011 <sup>c</sup>
Francophones	56,1 [56,1 ; 56,2]	59,8 [59,7 ; 60,0]	58,4 [58,2 ; 58,4]	50,0 [49,9 ; 50,0]	51,3 [51,2 ; 51,4]	52,8 [52,5 ; 52,9]
Anglophones	58,6 [58,4 ; 58,7]	61,0 [60,6 ; 61,5]	58,5 [58,2 ; 58,7]	53,0 [52,9 ; 53,2]	52,4 [52,0 ; 52,8]	53,8 [53,5 ; 54,0]
Allophones	63,2 [62,9 ; 63,4]	63,9 [63,4 ; 64,4]	59,6 [59,2 ; 59,8]	58,5 [58,2 ; 58,8]	55,7 [55,3 ; 56,0]	55,4 [54,8 ; 55,6]

**Notes :**

<sup>a</sup> Trempe et al. (2014, tableau 7), où la variable linguistique retenue est la langue d’usage à la maison, telle que renseignée sur les certificats de décès du Registre des décès du Québec.

<sup>b</sup> Trempe et al. (2014, tableau 7), où la variable linguistique retenue est la langue maternelle issue du recensement détaillé de 1991.

<sup>c</sup> Calculs de l’auteure pour la présente recherche où la variable linguistique retenue est la langue parlée le plus souvent à la maison issue du recensement détaillé de 1991.

**Sources de données retenues :**

INSPQ — Registre des décès du Québec et estimations de la population 1996 de l’ISQ.

BDJ-SC — Compilation spéciale de l’Étude canadienne de suivi de la mortalité selon le recensement 1991-2001.

CSERCan — Cohorte santé et environnement du recensement canadien de 1991.

Abréviations : BDJ-SC, Banque de données jumelées de Statistique Canada ; INSPQ, Institut national de santé publique du Québec ; ISQ, Institut de la statistique du Québec.

En premier lieu, nous avons calculé  $\hat{e}_{25}$  en nous servant des taux de mortalité sans prendre en compte le phénomène d'émigration dans la CSERCan. Simultanément, ces nouvelles estimations ont permis de mesurer l'effet de notre correction pour l'émigration sur  $\hat{e}_{25}$ , à savoir comment la correction a contribué à diminuer le biais lié aux sorties par émigration dans la CSERCan. Cela a eu pour résultat d'augmenter nos valeurs d'espérances de vie à 58,5, 58,9 et 60,3 ans pour les femmes francophones, anglophones et allophones, soit des différences de 0,1, 0,4 et 0,7 an respectivement. Une augmentation de  $\hat{e}_{25}$  était attendue en raison de la présence d'individus « immortels » dans les expositions au risque de décéder avant la correction, affectant particulièrement les taux de mortalité aux âges avancés à la baisse. En revanche, la correction du biais de la mortalité ne s'est pas avérée aussi importante qu'on l'aurait cru, pour deux raisons. D'abord, nous avons préalablement tronqué à droite le domaine d'âges utilisé pour la modélisation des taux afin d'éviter l'apparition de tendances indésirables aux très grands âges dans la courbe de mortalité lissée par P-splines. Puis, le poids négligeable des décès aux âges avancés, comparativement à celui des décès aux âges adultes, contribue peu à faire augmenter  $\hat{e}_{25}$ .

En deuxième lieu, nous avons utilisé la langue maternelle, plutôt que la langue d'usage à la maison, pour estimer  $\hat{e}_{25}$  afin que nos estimations ressemblent le plus possible à celle de la BDJ-SC. De plus, nous n'avons pas réparti les réponses multiples pour cette variable, ce qui correspond à la méthode de répartition de Trempe et collègues<sup>36</sup>, à défaut de connaître celle de Wilkins et collègues. Ces nouveaux traitements des données de la CSERCan n'ont pas entraîné de différences notables dans les valeurs d' $\hat{e}_{25}$ , même en utilisant les taux de mortalité sans correction pour l'émigration.

En troisième et dernier lieu, nous avons reconstruit la CSERCan en omettant le suivi de mortalité allant de 2002 à 2011, pour qu'elle corresponde à la période d'observation 1991-2001 de la BDJ-SC. En utilisant la langue maternelle comme variable linguistique, nous avons, sans surprise, obtenu des valeurs d'espérance de vie plus faibles que celles pour la période 1991-2011, en particulier chez les hommes. Ceci s'explique vraisemblablement par les gains de

---

<sup>36</sup> Les réponses multiples qui incluent une langue non officielle ne sont pas réparties mais tout simplement ajoutées aux réponses simples. Par exemple, le groupe des francophones inclut les personnes ayant répondu le français et une autre langue non officielle. Les bilingues français-anglais sont exclus de leur étude.

mortalité plus forts qu’ont réalisés les hommes par rapport aux femmes durant la période 2002-2011. Les « nouvelles » espérances de vie masculines à 25 ans pour la période 1991-2001 que nous en tirons (51,4 ans, 52,0 ans et 54,7 ans pour les francophones, anglophones et allophones respectivement) sont assez similaires à celles de la BDJ-SC et permettent d’amenuiser l’écart pour les francophones et anglophones. En revanche, chez les femmes, l’écart avec les estimations issues de la BDJ-SC s’est accentué.

Cherchant toujours à élucider la différence de mortalité inexplicée chez les femmes de la CSERCan et de la BDJ-SC, nous avons comparé les écarts d’espérance de vie à 25 ans entre les sexes. Cela s’est avéré déterminant pour trouver la source du problème. Le tableau III présente ces écarts (femmes moins hommes) selon le groupe linguistique pour les trois sources de données discutées plus haut, ainsi que pour la cohorte fictive de la BDLC.

Tableau III. Écarts d’espérance de vie à 25 ans entre les sexes selon plusieurs sources de données, Québec

Groupe linguistique	INSPQ 1994-1998 <sup>a</sup>	BDJ-SC 1991-2001 <sup>b</sup>	CSERCan 1991-2011 <sup>c</sup>	BDLC 1991-2011 <sup>d</sup>
Francophone	6,1	8,5	5,6	
Anglophone	5,6	8,6	4,7	
Allophones	4,7	8,2	4,2	
Total			5,4	5,3

**Notes :**

<sup>a</sup> Trempe et al. (2014, tableau 7), où la variable linguistique retenue est la langue d’usage à la maison, telle que renseignée sur les certificats de décès du Registre des décès du Québec.

<sup>b</sup> Trempe et al. (2014, tableau 7), où la variable linguistique retenue est la langue maternelle issue du recensement détaillé de 1991.

<sup>c</sup> Calculs de l’auteure pour la présente recherche où la variable linguistique retenue est la langue parlée le plus souvent à la maison issue du recensement détaillé de 1991.

<sup>d</sup> Calculs de l’auteure pour la présente recherche.

**Sources de données retenues :**

INSPQ — Registre des décès du Québec et estimations de la population 1996 de l’ISQ.

BDJ-SC — Compilation spéciale de l’Étude canadienne de suivi de la mortalité selon le recensement 1991-2001.

CSERCan — Cohorte santé et environnement du recensement canadien de 1991.

BDLC — Base de données sur la longévité canadienne.

Abréviations : BDJ-SC, Banque de données jumelées de Statistique Canada ; INSPQ, Institut national de santé publique du Québec ; ISQ, Institut de la statistique du Québec.

Nous remarquons dans ce tableau que les écarts pour la BDJ-SC sont considérablement plus élevés comparativement aux autres sources. En effet, pour chaque groupe linguistique, les écarts entre les sexes s’élèvent à plus de 8 ans dans la BDJ-SC, alors que ceux-ci sont autour de 4 à 6 ans ailleurs. Grâce à la cohorte fictive de la BDLC construite pour la période 1991-2011, nous

savons que pour l'ensemble de la population québécoise, l'écart est de 5,3 ans et bien proche de celui estimé avec la CSERCan 1991-2011, se chiffrant à 5,4 ans.

Suite à l'ensemble de ces vérifications, les écarts irréalistes d' $\hat{e}_{25}$  entre les sexes, en raison des espérances de vie très élevées chez les femmes dans la BDJ-SC, nous poussent à croire que les valeurs de  $\hat{e}_{25}$  pour les femmes provenant de cette source de données sont incorrectes, quel que soit le groupe linguistique. Les sources d'erreur nous sont toutefois inconnues, les calculs de Statistique Canada pour Trempe et collègues (2014) n'ayant pas été documentés. Ainsi, nous confirmons que nos estimations pour  $\hat{e}_{25}$  représentent les résultats les plus fiables et les plus à jour lorsqu'il est question de la mortalité différentielle selon la langue.

Parallèlement, en comparant nos estimations avec celles de l'INSPQ pour la période 1994-1998 à l'aide du tableau II, nous constatons deux tendances bien claires, observées chez les deux sexes : la mortalité des francophones est bel et bien surestimée, et celle des allophones, largement sous-estimée, lorsque l'information sur la langue provient du Registre des décès du Québec. La sous-estimation de la mortalité des allophones, déjà documentée par Termote (2011)<sup>37</sup>, se traduit par des espérances de vie surestimées. En effet, les valeurs d' $\hat{e}_{25}$  de l'INSPQ sont nettement supérieures aux nôtres avec la CSERCan (3,6 et 3,1 ans pour les femmes et les hommes respectivement). Ce biais provient de la forte probabilité qu'il y ait, chez les allophones en particulier, occurrence d'erreurs ou de réponses manquantes dans la langue enregistrée au décès. À titre d'exemple, il est fort possible que la langue inscrite sur les bulletins de décès soit la langue d'usage du patient défunt à l'hôpital plutôt que la langue d'usage à la maison. Cette source d'erreurs est d'autant plus importante que le groupe d'allophones est constitué d'un petit effectif d'individus. La part des décédés allophones incorrectement classés comme francophones ou anglophones peut, a priori, entraîner une surestimation de la mortalité des francophones et anglophones (Termote, 2011). Nous avons effectivement constaté que, chez les francophones, les valeurs d'espérance de vie de l'INSPQ pour les femmes et les hommes sont de 2,3 et 2,8 ans inférieures aux nôtres (tableau II). Cela n'a cependant pas été observé chez les

---

<sup>37</sup> Pour plus de détails, voir la section 1.2.2.

anglophones. Au final, les écarts linguistiques entre les valeurs pour  $\hat{e}_{25}$  dans la CSERCan, dénués de ces sources de biais, sont sensiblement plus petits que ceux publiés par l'INSPQ.

Pour donner suite à la validation de nos résultats pour  $\hat{e}_{25}$ , nous souhaitons commenter l'absence de différences significatives d' $\hat{e}_{25}$  entre les femmes francophones et anglophones. Des tendances de mortalité très similaires entre ces deux groupes linguistiques font contraste avec les observations du passé, soit que les anglophones vivent en moyenne plus longtemps que les francophones (Wilkins et al., 2008; Termote, 2011; Trempe et al., 2014). En raison de l'évolution de leur mortalité respective, l'écart de l'espérance de vie à la naissance (sexes réunis) entre les populations francophone et anglophone a diminué de manière notable, passant de 4,1 ans en 1976-1981 à 2,0 ans en 2005-2007 (Termote, 2011, tableau 4). Or, bien que l'écart de mortalité entre les hommes francophones et anglophones demeure significatif, celui-ci s'est également rétréci. Nous pouvons en déduire qu'aux âges adultes, l'amenuisement de l'écart entre ces deux groupes a persisté dans le temps. Nous suggérons que les francophones, ayant préalablement un niveau d'espérance de vie plus faible, possédaient un gain potentiel plus important comparé aux anglophones.

## **Mortalité différentielle aux grands âges**

La présente recherche avait également comme objectif d'évaluer s'il existe des différences de mortalité notables aux grands âges parmi les francophones, les anglophones et les allophones au Québec. Nous avons retenu deux indices de mortalité, soit  $\hat{e}_{65}$ , qui constitue un indicateur grandement utilisé pour mesurer la mortalité aux âges avancés, et  $\hat{M}$ , qui est libre de tout seuil d'âge et auquel les chercheurs se réfèrent de plus en plus pour suivre de près les transformations de la distribution des décès aux grands âges dans les régions à faible mortalité. Nous voulions aussi comparer  $\hat{e}_{65}$  et  $\hat{M}$ , à savoir si les différentiels de mortalité qu'on en dégage sont les mêmes. Comme préalablement mentionné dans la section 1.4, un changement de perspective s'impose lorsque nous recourons à  $\hat{e}_{65}$ , qui impose un seuil d'âge arbitraire, et  $\hat{M}$ , qui se concentre sur la portion de personnes très âgées, évitant ainsi l'essentiel de la mortalité prématurée. Chez les femmes, bien que l'espérance de vie à 65 ans soit significativement plus élevée chez les allophones que chez les francophones ou anglophones, aucune différence significative ne transparaît au niveau de l'âge modal au décès. Nous pouvons en déduire que

l'avantage de survie observé au niveau de  $\hat{e}_{65}$  chez les femmes allophones prend sa source dans les décès prématurés survenant aux *jeunes* grands âges. Chez les hommes, l'écart de mortalité est significatif au niveau de  $\hat{e}_{65}$  entre chaque groupe linguistique, mais seul l'écart entre l'âge modal au décès des allophones l'est par rapport aux francophones et anglophones. L'avantage de survie des hommes allophones semble donc persister jusqu'aux âges normaux de décès, contrairement aux hommes anglophones, chez qui nous n'avons pas observé de différence significative pour  $\hat{M}$  par rapport aux francophones.

Concernant la dispersion des décès aux grands âges mesurée par  $SD(M+)$ , nous remarquons que les décès des allophones se concentrent dans un intervalle d'âges plus court que ceux des francophones et des anglophones. Cependant, les différences ne sont pas significatives, que ce soit sur l'île de Montréal ou dans le reste du Québec.

## **Profils sociodémographiques et économiques**

Les caractéristiques sociodémographiques et économiques agissent comme leviers pour expliquer la mortalité différentielle entre les groupes linguistiques du Québec. Étant donné que nous avons uniquement ces informations en date du recensement de 1991 pour les membres de la CSERCan, nous ne pouvons ni connaître l'évolution de leurs profils au cours du temps, ni déterminer avec précision le rôle que jouent ces caractéristiques pour expliquer les différentiels de mortalité. Nous pouvons au moins présenter certaines des caractéristiques plus ou moins fixes, soit le lieu de naissance, le plus haut niveau de scolarité atteint et l'état matrimonial de la cohorte en 1991, afin d'améliorer notre compréhension de la mortalité différentielle selon la langue. Notons que l'atteinte de niveaux de scolarité élevés et le fait d'être marié sont associés à de plus faibles taux de mortalité (voir entre autres Wilkins et al. (2008) et Canon (2017)).

Dans la CSERCan, le groupe d'allophones au Québec est composé de 81% et 86% de femmes et d'hommes nés à l'extérieur du Canada, respectivement (figure annexe F.2). Les figures annexe F.3 et F.4 montrent que le groupe d'allophones du Québec est plus souvent marié ou en union libre, mais possède en revanche une part d'individus ayant atteint un niveau de scolarité secondaire, collégial ou universitaire nettement plus petite que le groupe de francophones et d'anglophones. Selon une étude qui utilise une version antérieure de la CSERCan, l'avantage de mortalité des immigrants n'est pas expliqué par des facteurs sociodémographiques ou

économiques (Omariba et al, 2014). Nous l’attribuons plutôt au biais de sélection lié à l’effet de l’immigrant en bonne santé<sup>38</sup>, ce qui est particulièrement vrai sur l’île de Montréal où 95% des individus allophones sont nés à l’étranger. En considérant que le groupe d’allophones dans le RDQ comporte une part importante d’autochtones, affichant des niveaux de mortalité considérablement plus élevés que la moyenne québécoise, explorons à présent la mortalité des allophones selon la région plus en profondeur.

Dans l’annexe E, nous présentons plusieurs figures qui comparent les indicateurs de mortalité  $\hat{e}_{25}$ ,  $\hat{e}_{65}$  et  $\hat{M}$  selon le groupe linguistique pour les régions de MTL (figures E.6 à E.8) et du RDQ (figures E.9 à E.11). Pour MTL, la mortalité des femmes et des hommes allochtones est significativement plus faible que celle des francophones et des anglophones pour les espérances de vie à 25 et à 65 ans. Toutefois, aux très grands âges, l’écart en matière de  $\hat{M}$  entre les allochtones et les anglophones n’est pas significatif chez les deux sexes. Ce résultat peut être expliqué par le fait que l’avantage de mortalité des immigrants tend à s’estomper au fur et à mesure que s’allonge leur durée de résidence au Canada (Omariba et al, 2014). Aux âges très avancés, la durée de résidence au Canada de ces immigrants est probablement plus longue que courte.

Pour le RDQ, les figures E.9 à E.11 dévoilent les valeurs pour  $\hat{e}_{25}$ ,  $\hat{e}_{65}$  et  $\hat{M}$  pour chaque groupe linguistique, de façon séparée pour le groupe d’allochtones (orange), d’autochtones (mauve) et d’allophones (vert). Ces graphiques à barres démontrent que les valeurs pour  $\hat{e}_{25}$  et  $\hat{e}_{65}$  du groupe d’allophones constituent une moyenne approximative des tendances de mortalité opposées du groupe d’allochtones (faible mortalité) et d’autochtones (forte mortalité). Pour l’âge modal au décès, la part moins importante d’individus autochtones dans le groupe d’allophones aux grands âges contribue à expliquer la valeur relativement élevée de  $\hat{M}$  chez les allophones. De plus, à ces âges avancés, ce sont uniquement les hommes allophones qui conservent leur avantage de mortalité par rapport aux francophones et anglophones dans le reste du Québec.

Pour ce qui est de la mortalité différentielle entre les anglophones et les francophones du Québec, nous croyons que les diverses caractéristiques sociodémographiques et économiques

---

<sup>38</sup> Plus de détails dans la section 1.2.1

ainsi que le biais de sélection lié à l'effet de l'immigrant en bonne santé contribuent à expliquer leurs tendances de mortalité. En effet, nous notons une part considérablement plus importante d'anglophones nés à l'extérieur du Canada, soit 28% et 30% chez ceux-ci comparativement à 3% et 4% chez les francophones, pour les femmes (figure F.2a) et les hommes (figure F.2b), respectivement. Rappelons toutefois qu'aucun écart de mortalité aux âges adultes et avancés entre les femmes francophones et anglophones du Québec ne s'est révélé statistiquement significatif. Or, à l'échelle provinciale, ces dernières possèdent des profils similaires en ce qui concerne leurs niveaux de scolarité et leur état matrimonial (figures F.3a et F.4a). Par contraste, dans MTL, nous retrouvons des écarts de mortalité significatifs entre les femmes francophones et anglophones au niveau de  $\hat{e}_{25}$  et  $\hat{e}_{65}$  (figures E.6a et E.7a). Dans cette région, les femmes anglophones de la cohorte sont effectivement plus éduquées et plus souvent mariées que les francophones (figures F.5a et F.6a). En ce qui concerne les hommes, le profil des anglophones semble également expliquer en partie leur avantage sur le plan de la mortalité par rapport aux francophones. Dans les deux régions, nous observons de plus fortes proportions d'hommes anglophones mariés et diplômés au niveau collégial ou universitaire que les hommes francophones (figures F.5b à F.8b).

## Limites

Bien que ces résultats aident à mieux situer et comprendre les inégalités de mortalité, il convient de mentionner certaines limites associées à notre base de données, la CSERCan. Premièrement, nous pouvons mentionner que les profils sociodémographiques et économiques sont influencés par la structure par âge de chaque catégorie linguistique. Cependant, considérant la distribution par âge assez similaire entre les trois groupes, ces tendances sont relativement justes. Deuxièmement, comme mentionné plus haut, nous ne connaissons pas l'évolution des caractéristiques des membres de la cohorte au cours de la période 1991-2011 puisque ces informations proviennent du recensement détaillé de 1991. Toutefois, même si la langue parlée le plus souvent à la maison peut avoir changé au cours de la période 1991-2011, la mobilité linguistique joue un rôle marginal dans notre analyse sur la mortalité des adultes et aux grands âges compte tenu du nombre négligeable de transferts linguistiques après 25 ans (Termote, 2008). Troisièmement, bien que l'échantillon soit représentatif de la population du Québec, les



membres de la CSERCan sont légèrement avantagés au niveau de la mortalité, comme noté dans leurs taux de mortalité plus faibles que ceux de la cohorte BDLC (section 3.1.1). Cet avantage provient en partie de l'exclusion des pensionnaires d'établissement dans le recensement détaillé de 1991, mais également des probabilités de jumelage qui varient selon les caractéristiques de la population. Parmi les catégories d'individus possédant une plus grande probabilité de non-appariement dans l'échantillon de la CSERCan pour le Canada, nous retrouvons entre autres : 1) les individus d'origine autochtone, 2) du quintile inférieur de suffisance de revenu, 3) ne faisant pas partie de la population active, 4) non mariés ou ne vivant pas avec leur conjoint, 5) n'ayant pas de diplôme secondaire, et 6) âgés de 65 ans et plus (Wilkins et al., 2008). Mis à part les personnes âgées de 65 ans et plus, toutes les exclusions contribuent à biaiser nos estimations de l'espérance de vie à 25 et à 65 ans en les surestimant. En revanche, avec l'exclusion d'une part importante de personnes âgées qui, forcément, décéderont à des âges avancés, nos estimations pour  $\hat{e}_{25}$  et  $\hat{e}_{65}$  s'en trouveraient sous-estimées. Sachant que la probabilité d'appartenir aux catégories 2 à 5 tend à augmenter avec l'avancement en âge, la surestimation des indicateurs  $\hat{e}_{25}$  et  $\hat{e}_{65}$  serait en partie contrebalancée par la part importante de personnes âgées dans ces catégories. Considérant le léger avantage que détient la CSERCan sur le plan de mortalité, comme démontré dans la section 3.1.1, nous comprenons que l'ensemble des caractéristiques associées à de plus grandes probabilités de non-appariement tend, au total, à biaiser les valeurs d' $\hat{e}_{25}$  et d' $\hat{e}_{65}$  vers le haut. Enfin, notre compréhension des inégalités sociales de santé entre les groupes linguistiques est limitée étant donné le manque d'information au regard des habitudes et des conditions de vie, telles que l'alimentation, la sédentarité, le tabagisme et la consommation d'alcool, qui sont propres à chaque groupe.

## Conclusion

Dans le passé, divers travaux ont démontré qu'il existe des différences de mortalité en matière d'espérance de vie entre les principaux groupes linguistiques du Québec. Toutefois, l'information sur la langue provenait des bulletins de décès dans lesquels celle-ci n'est pas toujours fournie ou exacte. La fiabilité de telles données pouvait donc être remise en question. Nous avons donc profité de l'accès à la base de données CSERCan, dans laquelle la variable linguistique provient directement du recensement de 1991.

En premier lieu, ce mémoire a permis de réévaluer et d'actualiser les résultats précédents avec une précision accrue grâce à la base de données *CSERCan*. Nous avons conclu qu'il existe bel et bien des différences significatives de mortalité aux âges adultes selon l'appartenance linguistique au Québec, à l'exception des femmes francophones et anglophones pour lesquelles les tendances de mortalité diffèrent uniquement sur l'île de Montréal. La plus grande fiabilité de la variable linguistique dans la *CSERCan* nous a permis de constater que les écarts d' $\hat{e}_{25}$  sont moins importants qu'annoncé dans la littérature.

En deuxième lieu, nous avons profité de ces données pour étudier les différentiels de mortalité selon le groupe linguistique aux grands âges spécifiquement, ce qui, jusqu'à présent, n'avait jamais été fait. Nous avons constaté que les différences de mortalité observées aux âges adultes persistent chez les 65 ans et plus, mais pas nécessairement parmi l'ensemble des personnes atteignant l'âge modal au décès. Effectivement, nous ne constatons aucun écart significatif de mortalité au niveau de  $\hat{M}$ , à l'exception des hommes allophones, qui conservent leur avantage de mortalité aux très grands âges.

En troisième lieu, nous voulions comparer les dispersions des décès au-delà de l'âge modal, toujours selon l'appartenance linguistique. À ce sujet, nous avons constaté que la durée de vie modale des allophones, qui est la plus élevée chez les deux sexes, s'accompagne d'une concentration des décès dans un intervalle d'âge plus court. Ce niveau de compression de la mortalité n'est toutefois pas significativement différent de celui des francophones ou des anglophones.

En quatrième et dernier lieu, nous tenions à évaluer si les tendances de mortalité selon la langue sont les mêmes sur l'île de Montréal et dans le reste du Québec. Des différences régionales sont

notées chez les francophones et les allophones. Dans le premier cas, les individus francophones de MTL tendent à être désavantagés par rapport à ceux du RDQ. Nous attribuons cette disparité régionale à leurs différents profils sociodémographiques et économiques, ainsi qu'aux dynamiques démolinguistiques propres à MTL, où il y a une meilleure représentativité des groupes linguistiques que dans le RDQ. En ce qui concerne le groupe d'allophones, leur désavantage de mortalité dans le RDQ s'explique par la présence importante d'autochtones dans cette région, qui affichent des niveaux de mortalité plus élevés que la moyenne québécoise. Nous ne retrouvons effectivement aucune différence régionale lorsque nous comparons les groupes d'allochtones.

Notre mémoire comprend plusieurs améliorations apportées à la recherche sur la mortalité différentielle selon le groupe linguistique au Québec, non seulement par la plus grande fiabilité de la variable linguistique dans la CSERCan, mais également au niveau du traitement des données et de l'analyse. Notons la répartition des langues multiples, la prise en compte de l'émigration dans la cohorte, l'analyse par région, la considération des autochtones dans le groupe d'allophones et l'utilisation de l'approche de lissage par P-splines pour estimer les indicateurs de mortalité. Nous avons ainsi pu parfaire et mettre à jour les observations liées aux différentiels de mortalité aux âges adultes, et évaluer si ceux-ci persistent aux grands âges parmi les francophones, les anglophones et les allophones au Québec.

Pour les prochains travaux sur ce sujet, il serait intéressant d'aborder la question d'un point de vue interdisciplinaire dans le but de mieux comprendre ce qui cause les inégalités sociales de santé. Par exemple, la recherche sur la mortalité différentielle selon le groupe linguistique profiterait d'une analyse qui se concentre sur les déterminants sociaux de la santé, dont les quatre catégories se rapportent aux habitudes de vie (alcoolisme, tabagisme, sédentarité, alimentation déséquilibrée), à l'environnement physique et social, au système de soins, et à la biologie, telle que l'épigénétique. Également, conduire une analyse axée sur les causes de décès constitue une voie intéressante, qui permettrait une meilleure compréhension de la mortalité de chacune des communautés linguistiques. Enfin, depuis tout récemment, les chercheurs du CIQSS où la présente étude a été menée disposent d'une période de suivi de la mortalité allant de 1991 à 2016 pour la CSERCan. L'utilisation de la version plus récente de la base de données permettrait d'obtenir une mise à jour de nos résultats grâce à cinq années additionnelles précieuses de suivi.

# ANNEXE A

## Catégories linguistiques

Tableau A.I : Langues incluses dans la catégorie linguistique allochtone, CSERCan

Code	Groupe allochtone	Code	Groupe allochtone	Code	Groupe allochtone
3	Néerlandais	30	Serbe	57	Lao
4	Flamand	31	Slovène	58	Chinois
5	Frison	32	Tchèque	59	Langues sino-tibétaines, n.i.a.
6	Danois	33	Macédonien	60	Vietnamien
7	Islandais	34	Polonais	61	Khmer (cambodgien)
8	Norvégien	35	Ukrainien	62	Langues austro-asiatiques, n.i.a.
9	Suédois	36	Slovaque	63	Langues asiatiques, n.i.a.
10	Yiddish	37	Langues slaves, n.i.a.	64	Indonésien (malais)
11	Allemand	38	Kurde	65	Tagalog (Pilipino)
12	Langues germaniques, n.i.a.	39	Baloutche	66	Langues malayo-polynésiennes, n.i.a.
13	Portufais	40	Persan (farsi)	67	Arabe
14	Espagnol	41	Pendjabi	68	Hébreu
15	Roumain	42	Gujarati	69	Maltais
16	Italien	43	Marathi	70	Langues sémitiques, n.i.a.
17	Langues latines, n.i.a.	44	Singhalais	72	Turc
18	Grec	45	Hindi	73	Langues turques, n.i.a.
19	Arménien	46	Ourdou	74	Estonien
20	Langues gaéliques	47	Bengali	75	Finois
21	Gallois	48	Pachto	76	Hongrois
22	Langues celtiques, n.i.a.	49	Langues indo-iraniennes, n.i.a.	77	Swahili
23	Letton	50	Malayalam	78	Langues bantoues, n.i.a.
24	Lituanien	51	Tamoul	79	Langues nigéro-congolaises, n.i.a.
25	Russe	52	Telougou	80	Langues africaines, n.i.a.
26	Biélorussien	53	Langues dravidiennes, n.i.a.	107	Langues créoles
27	Bulgare	54	Japonais	108	Autres langues
28	Serbo-croate	55	Coréen		
29	Croate	56	Thai		

Note : Ces informations sont tirées du livre de codes du recensement de 1991.

Tableau A.II : Langues incluses dans la catégorie linguistique autochtone, CSERCan

<b>Code</b>	<b>Langues autochtones</b>
81	Cri
82	Ojibway
83	Pied-noir
84	Malécite
85	Micmac
86	Montagnais-naskapi
87	Langues algonquines, n.i.a.
88	Chipewyan
89	Esclave du sud
90	Porteur
91	Chilcotin
92	Flanc-de-chien
93	Esclave du nord (Peau-de-lièvre)
94	Kutchin-gwich'in (loucheux)
95	Langues athapascanes, n.i.a.
96	Mohawk
97	Langues iroquoises, n.i.a.
98	Langues salishennes
99	Tsimshian
100	Langues wakashanes
101	Haida
102	Dakota
103	Tlingit
104	Kutenai
105	Langues amérindiennes, n.i.a.
106	Inuktitut

Note : Ces informations sont tirées du livre de codes du recensement de 1991.

# ANNEXE B

## Émigration

Tableau B.I : Estimation d'émigrants, CSERCan, Québec, 1991 à 2011

Année	Effectifs	Année	Effectifs
1991	1115	2002	710
1992	690	2003	630
1993	670	2004	745
1994	675	2005	785
1995	650	2006	710
1996	740	2007	715
1997	640	2008	725
1998	655	2009	730
1999	580	2010	740
2000	575	2011	745
2001	610		
<b>Total</b>			<b>14835</b>

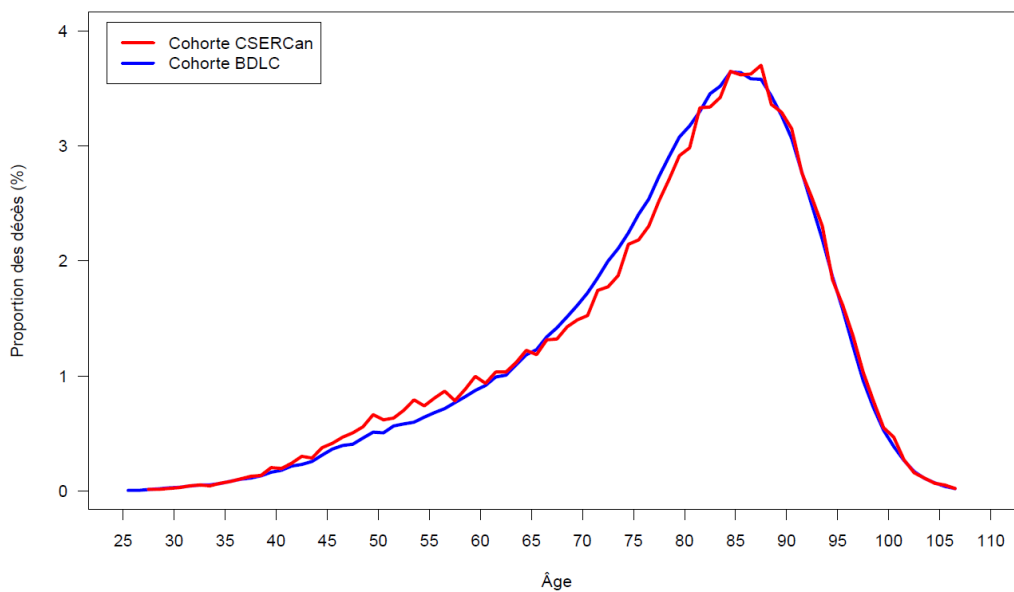
Note : Consulter la section 2.2.1.2 pour la méthodologie utilisée pour estimer ces effectifs d'émigrants.

# ANNEXE C

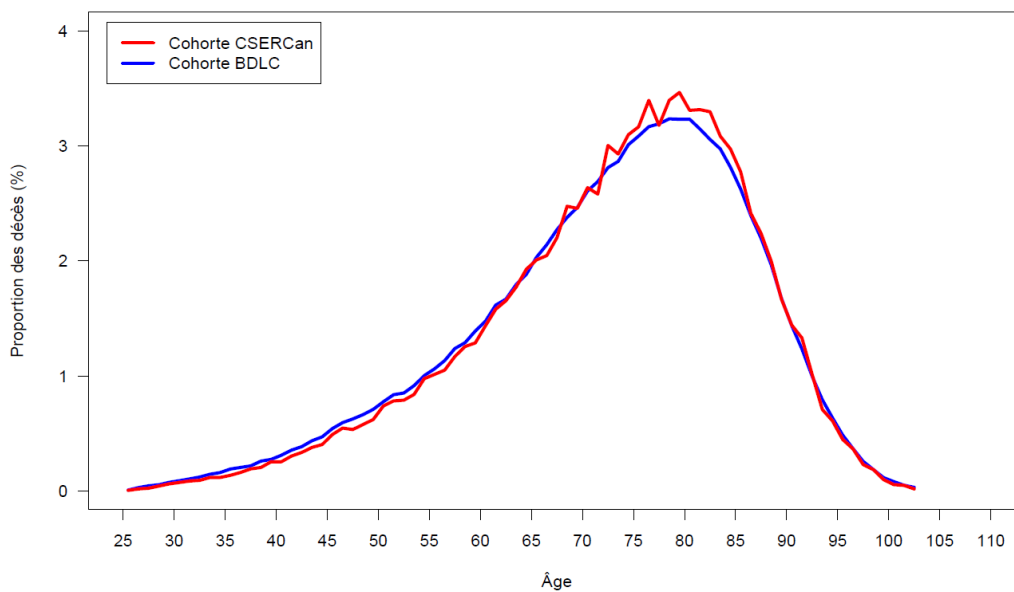
## Validation de la CSERCan avec la cohorte BDLC

Figure C.1 Proportion des décès selon l'âge et le sexe au Québec: Comparaison de la CSERCan et la cohorte BDLC, 1991 à 2011

### a. Femmes



### b. Hommes

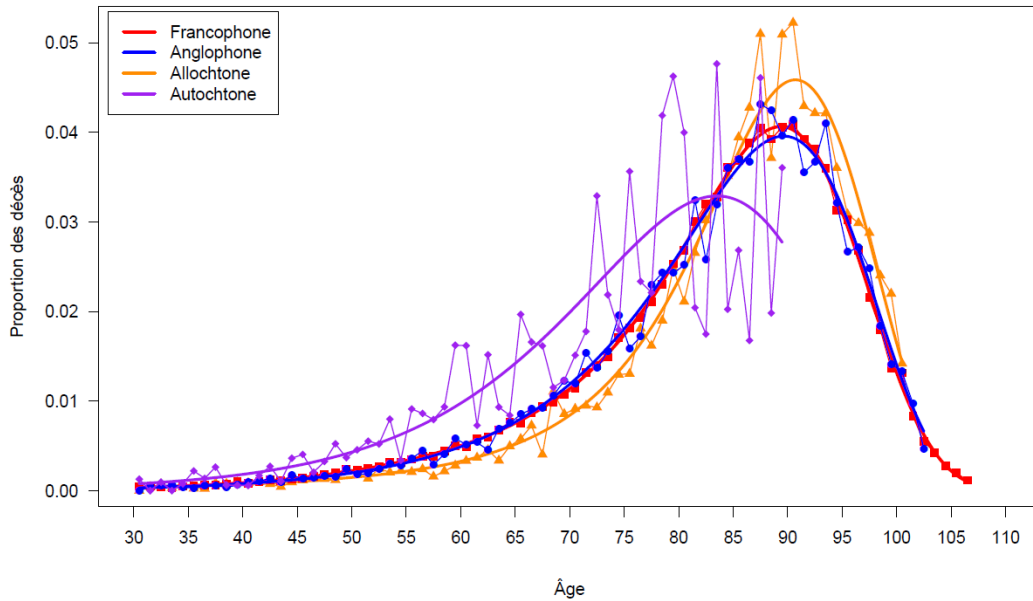


# ANNEXE D

## Résultats pour la province du Québec

Figure D.1 Comparaison entre les proportions des décès tirées de la table de mortalité et la fonction de densité lissée par P-splines selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011

### a. Femmes



### b. Hommes

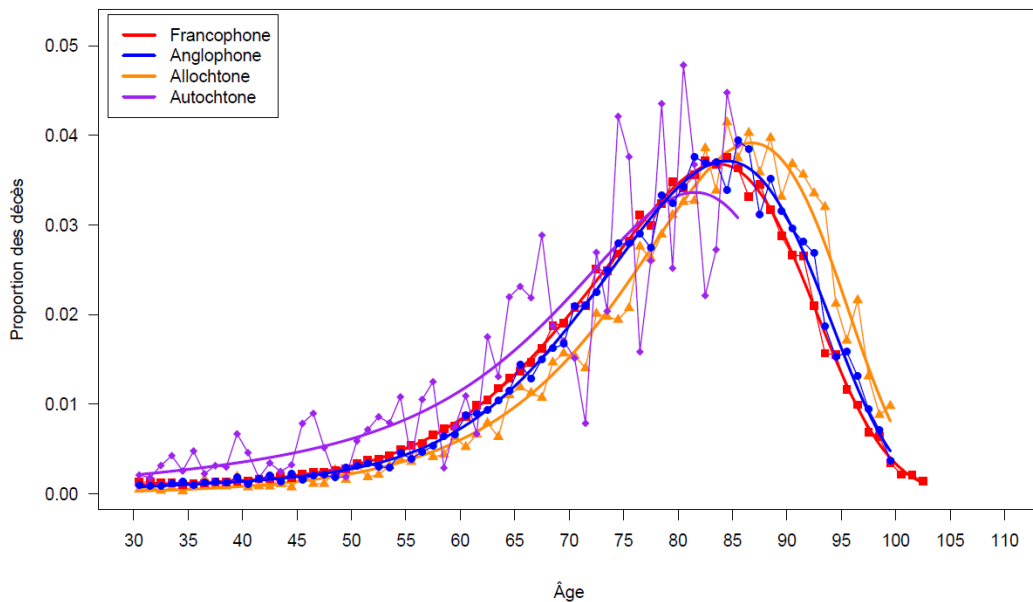
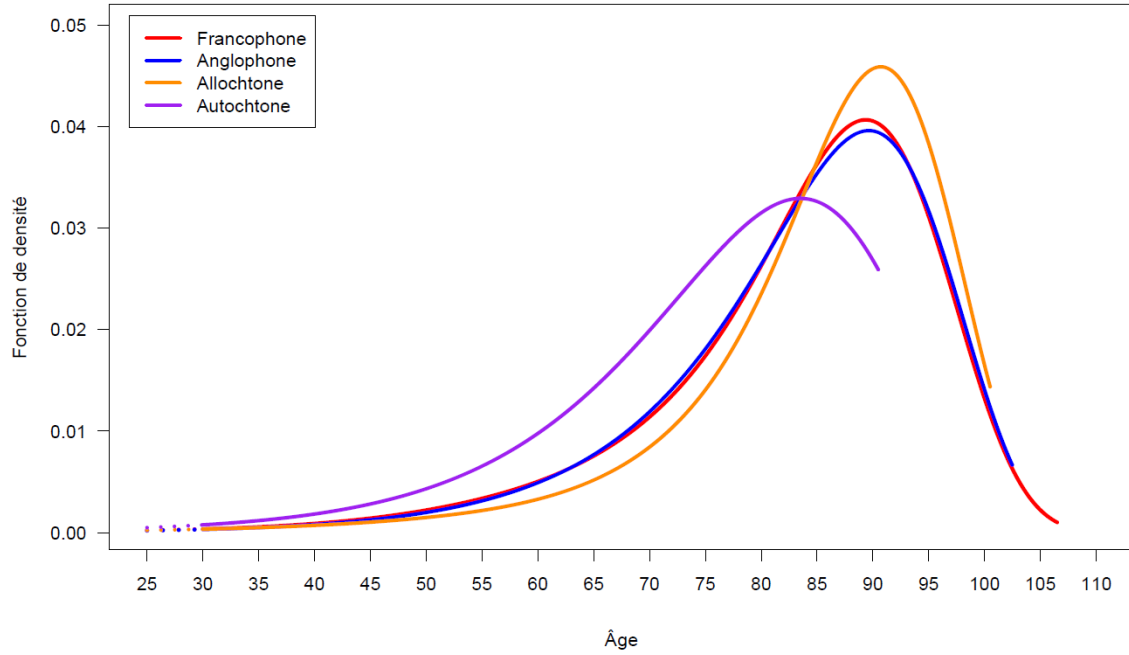




Figure D.2 Fonction de densité lissée selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011

**a. Femmes**



**b. Hommes**

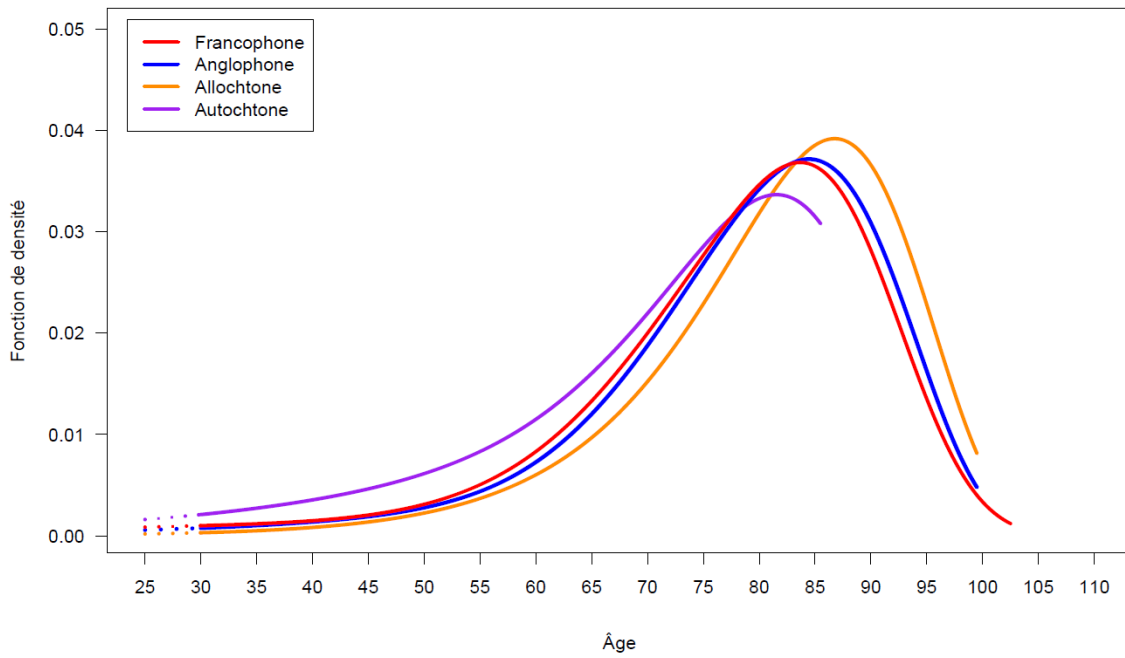
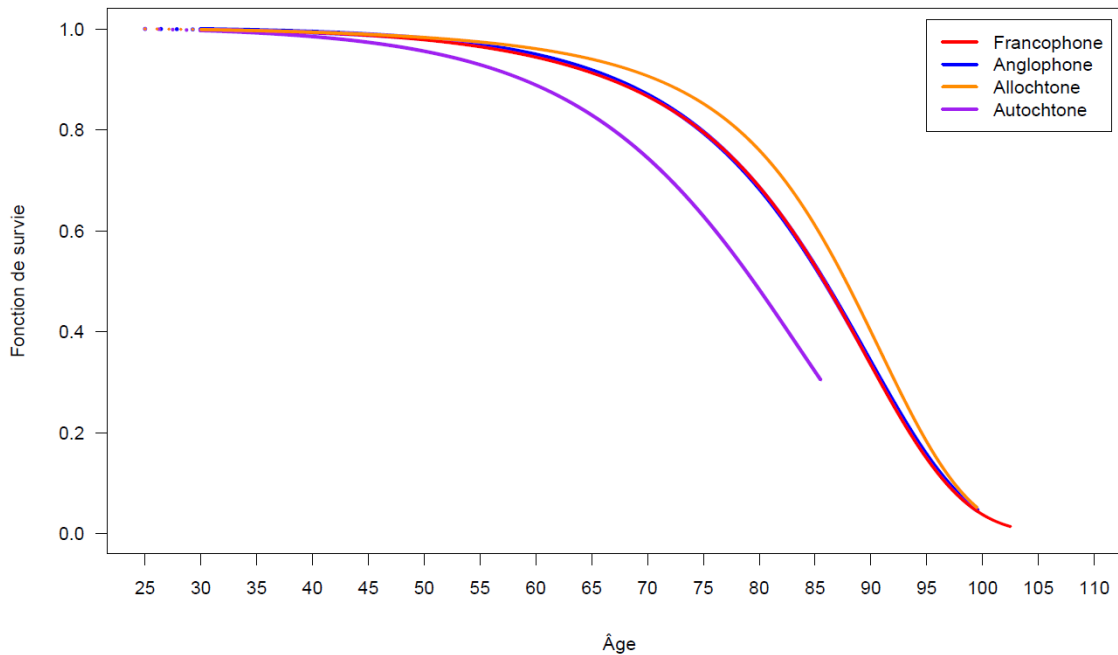


Figure D.3 Fonction de survie lissée selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011

**a. Femmes**



**b. Hommes**

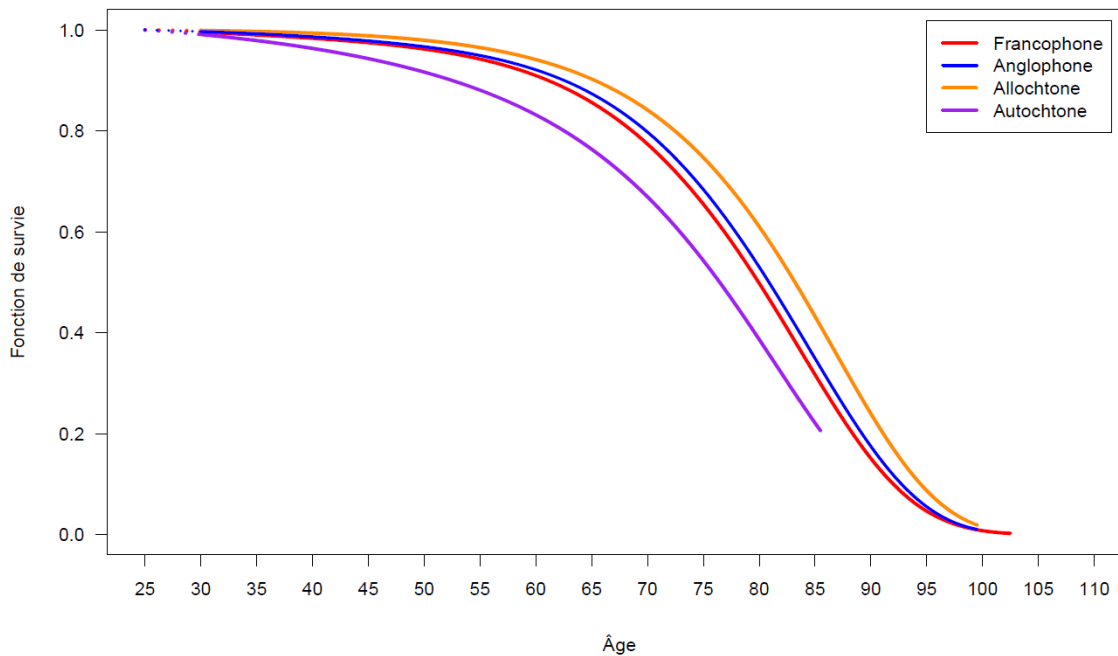


Figure D.4 Estimations de l'espérance de vie à 25 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011

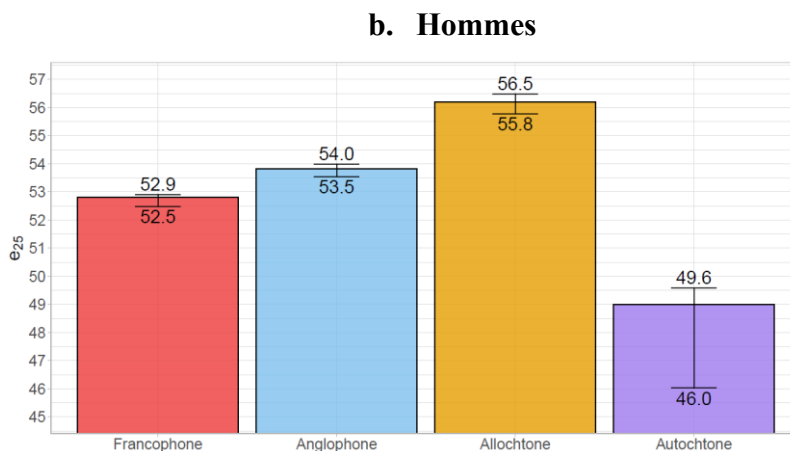
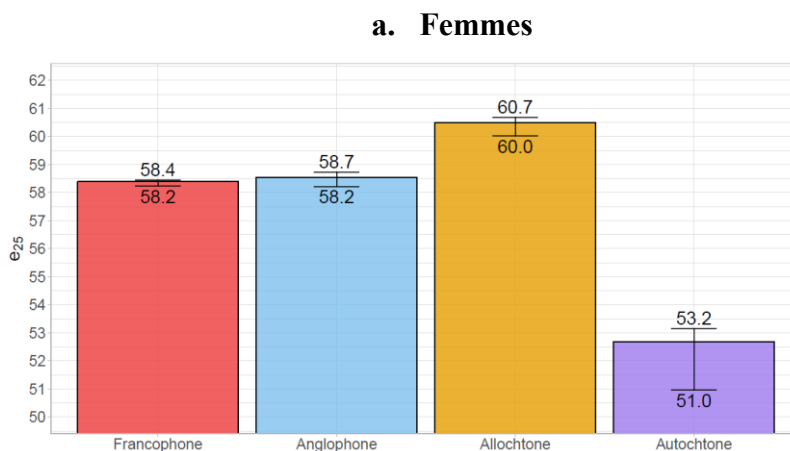
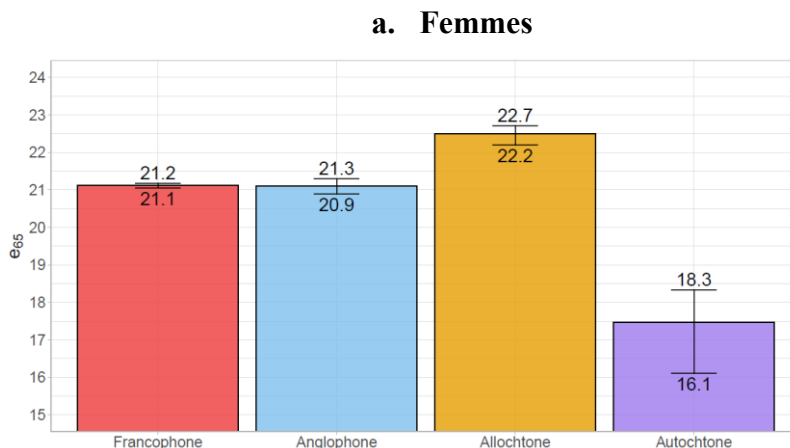


Figure D.5 Estimations de l'espérance de vie à 65 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011



### b. Hommes

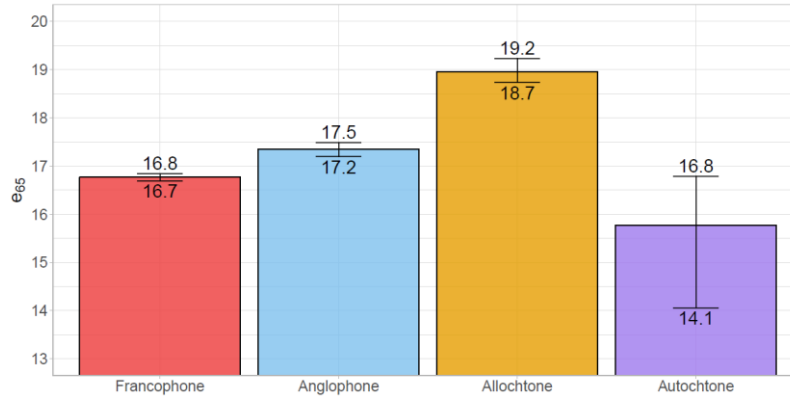
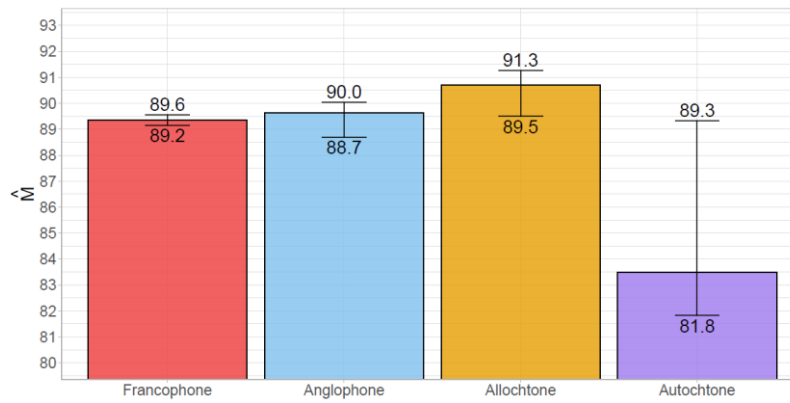


Figure D.6 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011

### a. Femmes



### b. Hommes

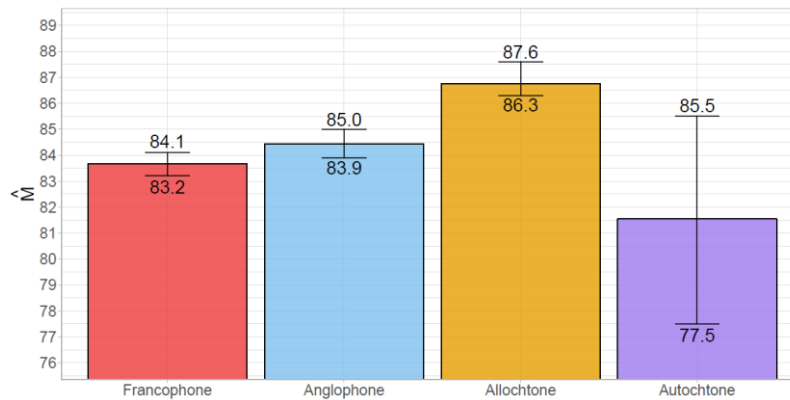
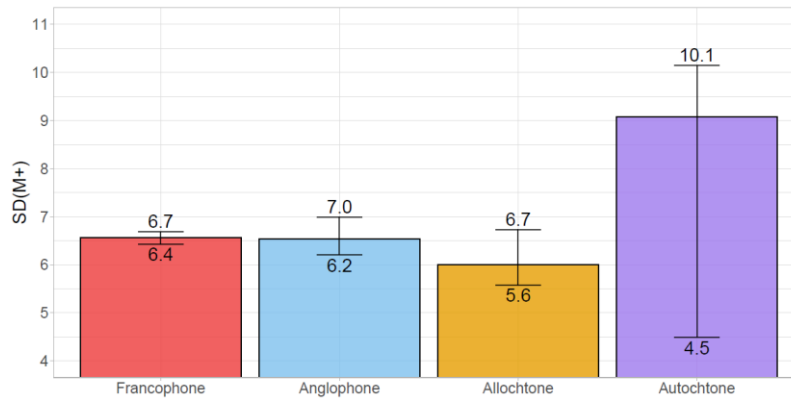
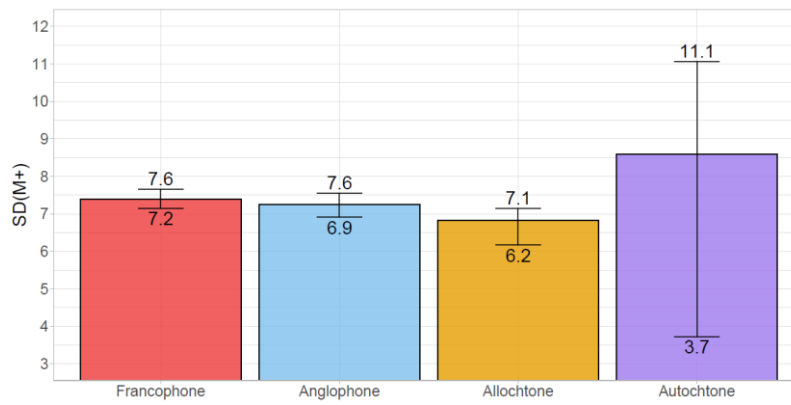


Figure D.7 Estimations de l'écart-type au-delà du mode et des intervalles de confiance bootstrap à 95% selon le groupe linguistique (4) et le sexe, CSERCan, Québec, 1991 à 2011

**a. Femmes**



**b. Hommes**

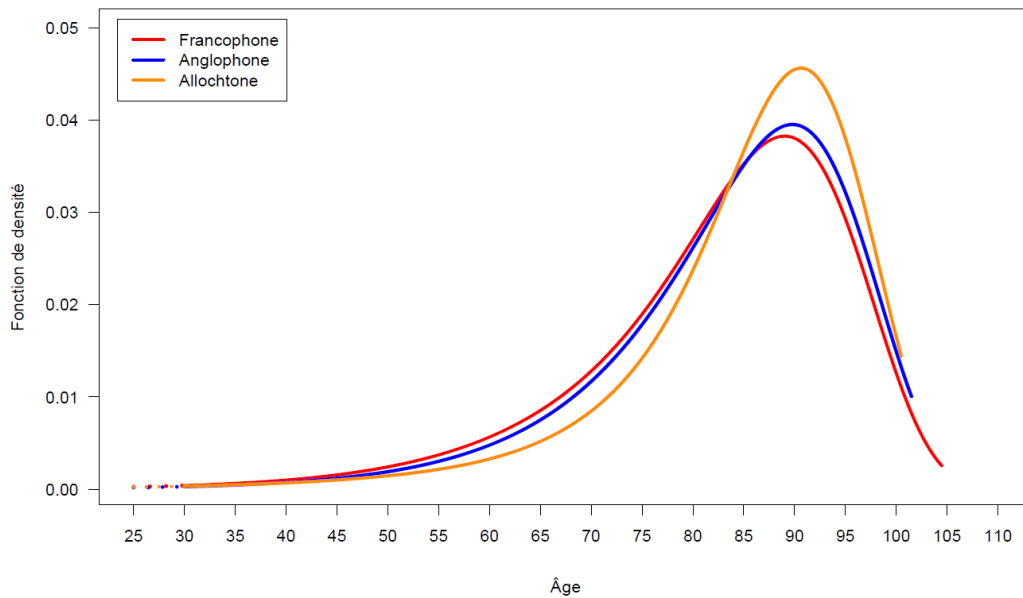


# ANNEXE E

## Résultats pour l'île de Montréal et le reste du Québec

Figure E.1 Fonction de densité lissée selon le groupe linguistique (3) et le sexe, CSERCan, Île de Montréal, 1991 à 2011

### a. Femmes



### b. Hommes

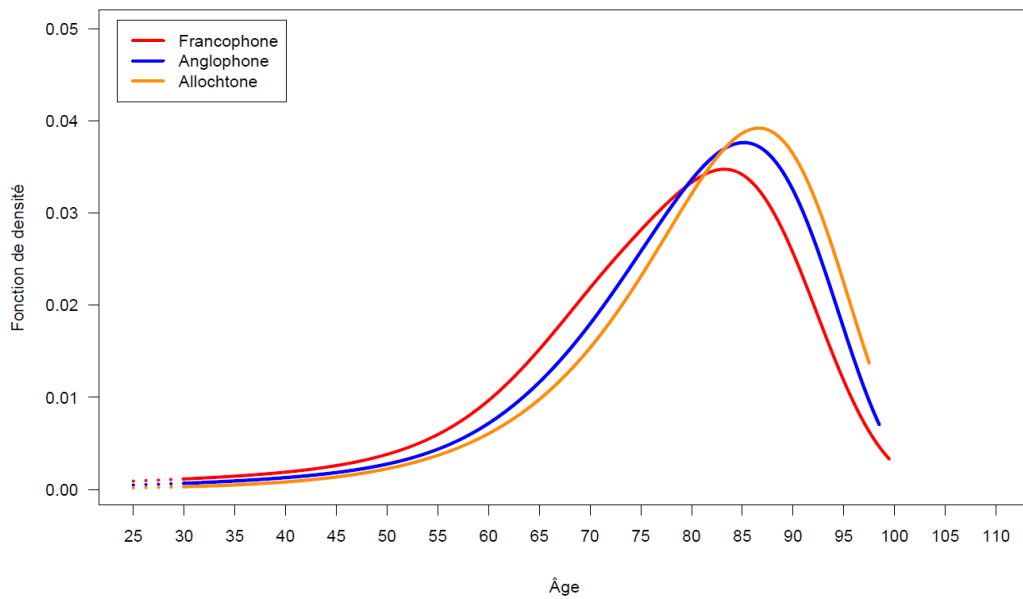
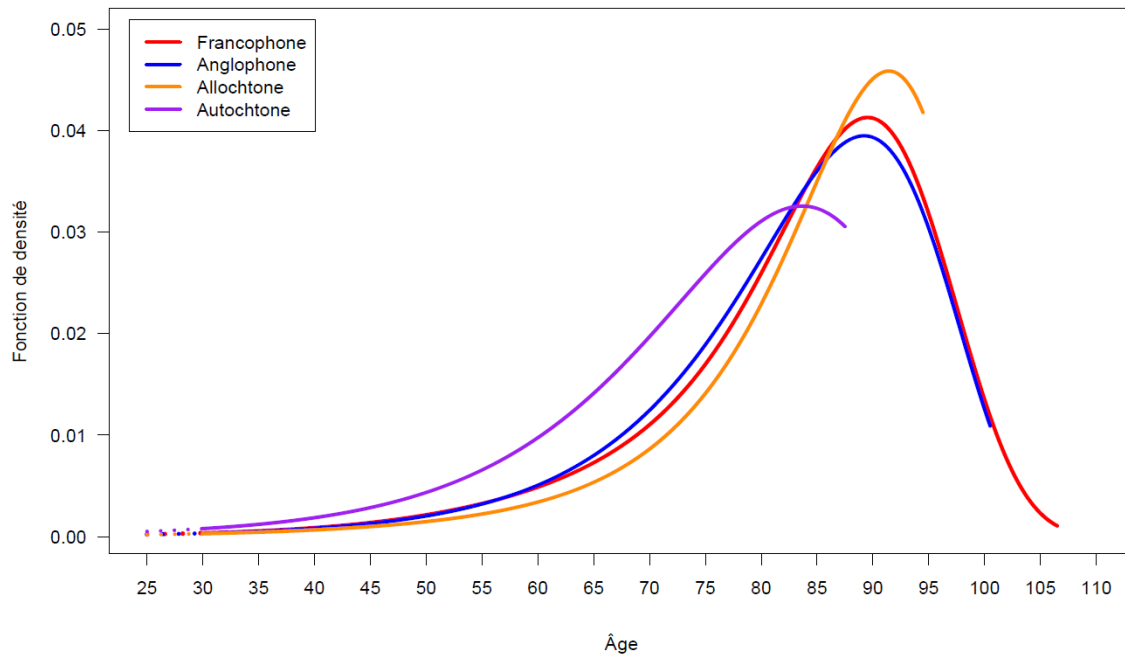


Figure E.2 Fonction de densité lissée selon le groupe linguistique (4) et le sexe, CSERCan, Reste du Québec, 1991 à 2011

**a. Femmes**



**b. Hommes**

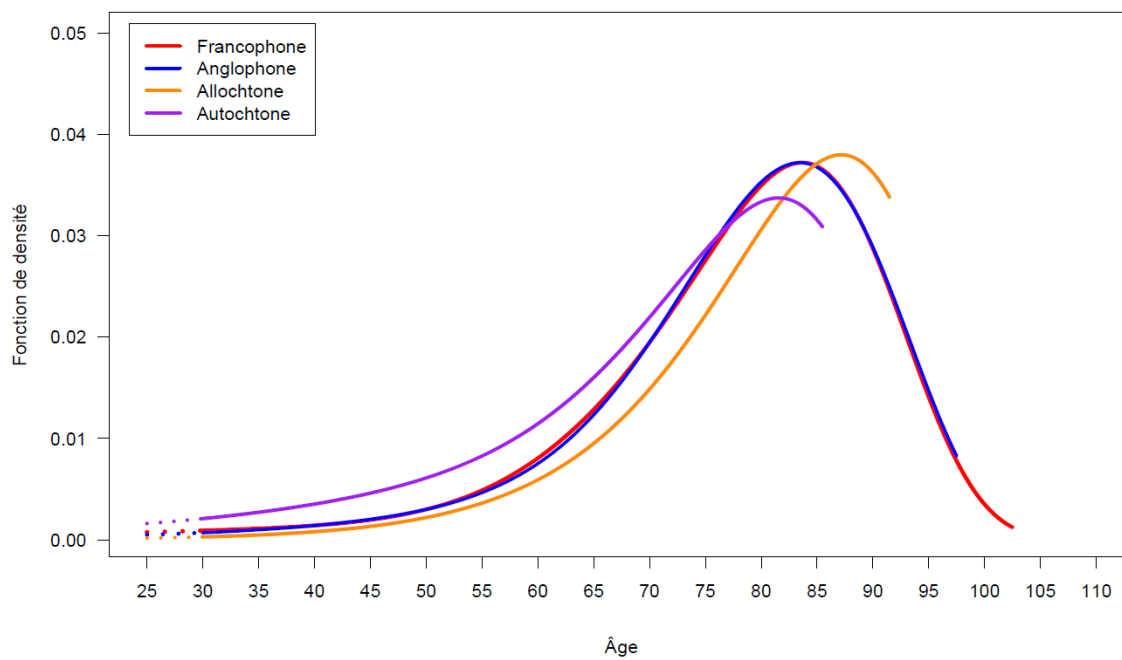
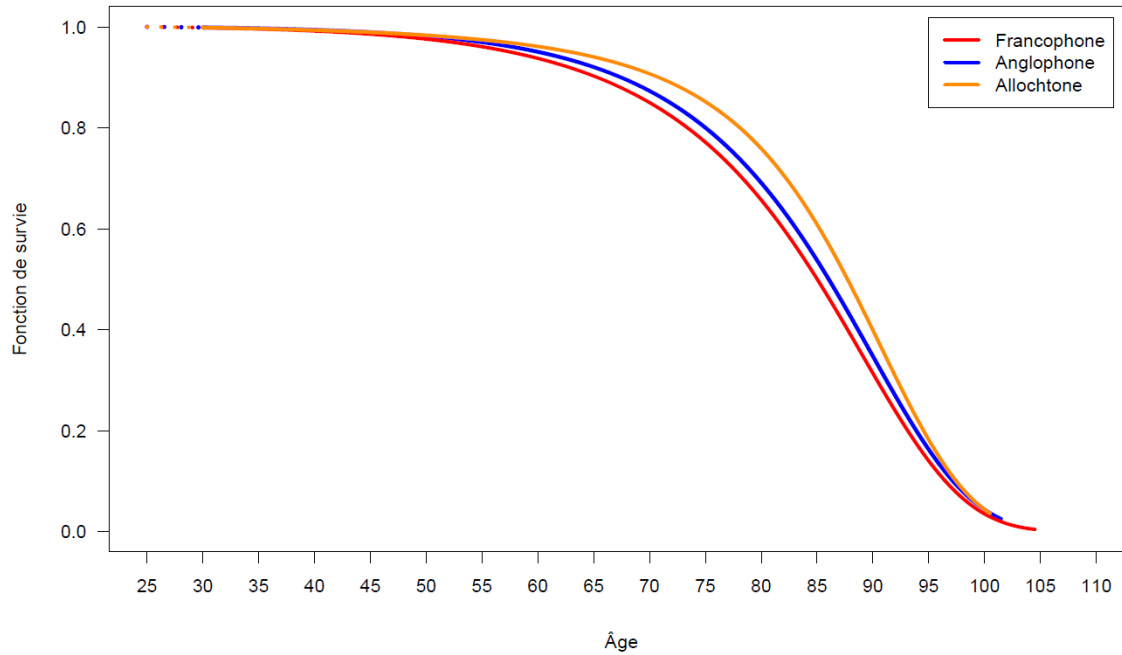


Figure E.3 Fonction de survie lissée selon le groupe linguistique (4) et le sexe, CSERCan, Île de Montréal, 1991 à 2011

**a. Femmes**



**b. Hommes**

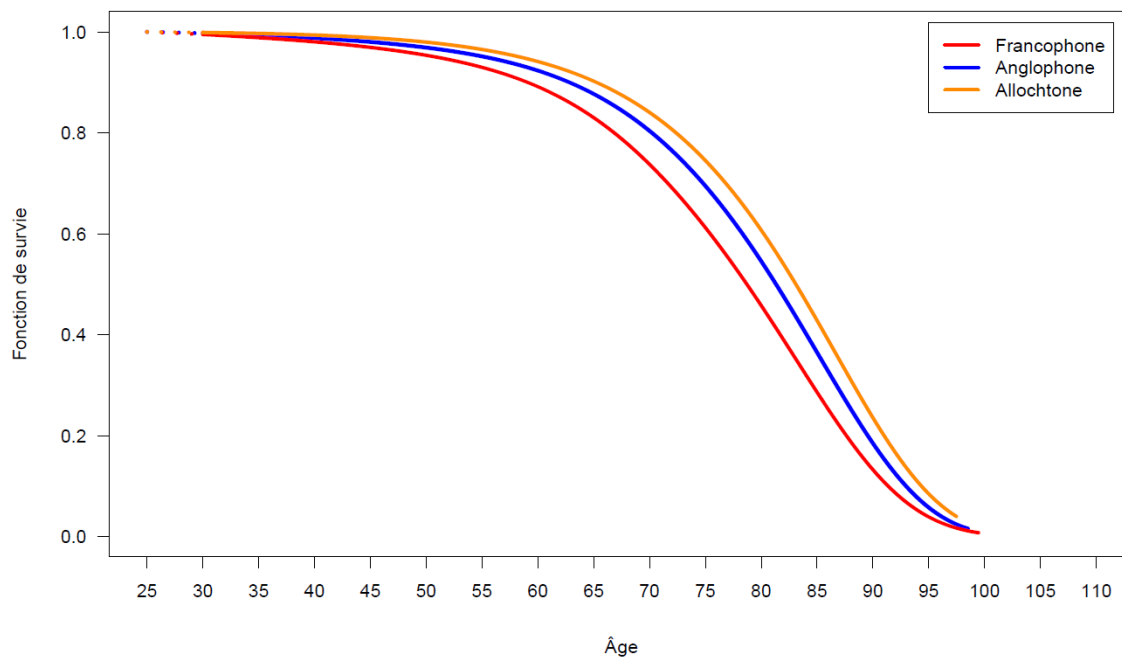
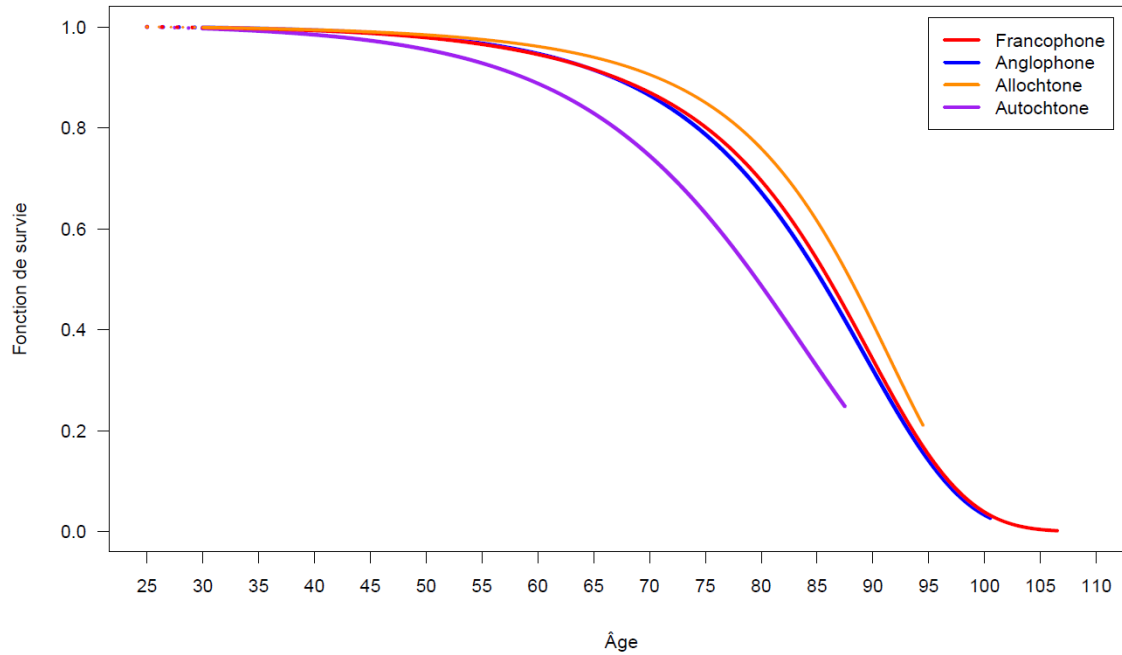




Figure E.4 Fonction de survie lissée selon le groupe linguistique (4) et le sexe, CSERCan, Reste du Québec, 1991 à 2011

**a. Femmes**



**b. Hommes**

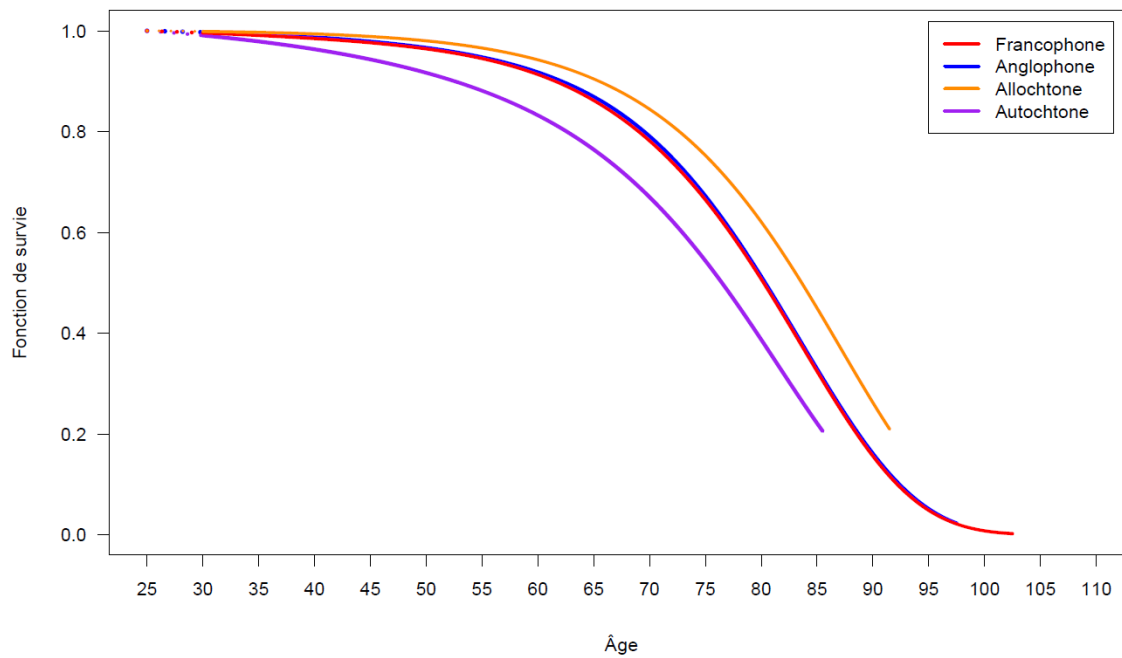


Figure E.5 Estimations de l'écart-type au-delà du mode et des intervalles de confiance bootstrap à 95% selon la région, le groupe linguistique et le sexe, CSERCan, 1991 à 2011

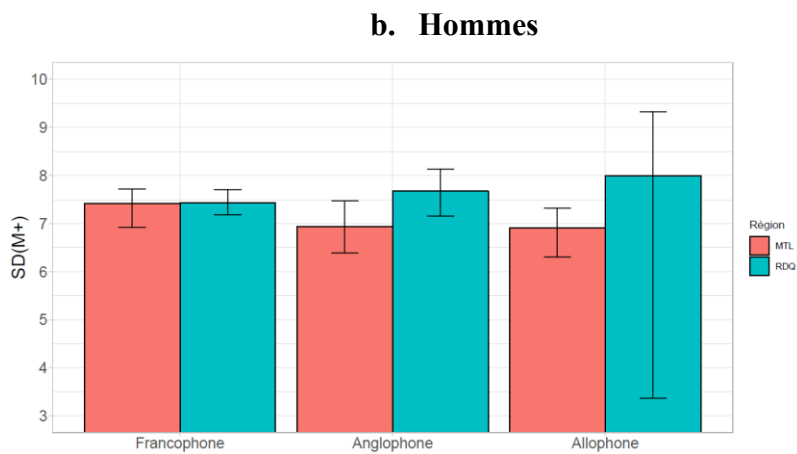
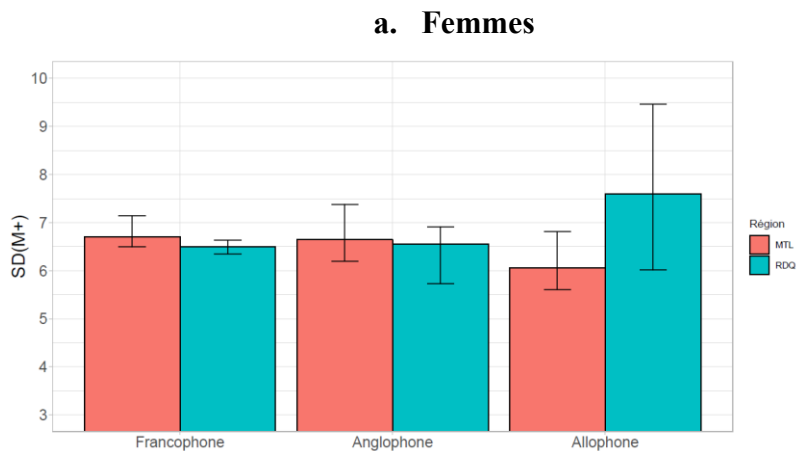
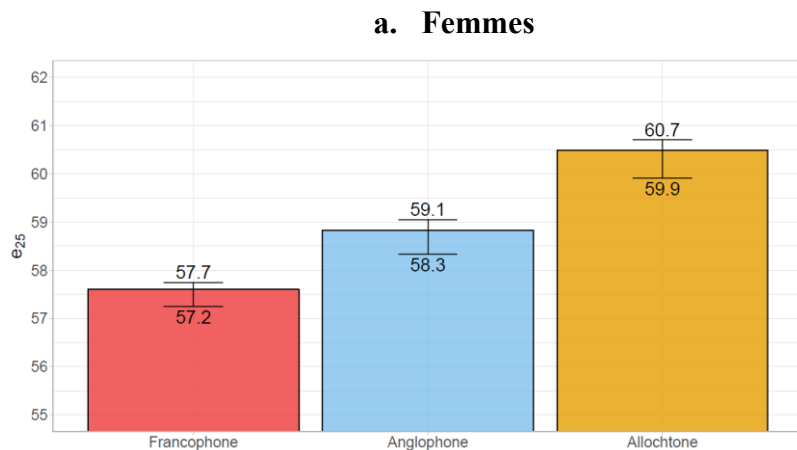


Figure E.6 Estimations de l'espérance de vie à 25 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011



### b. Homme

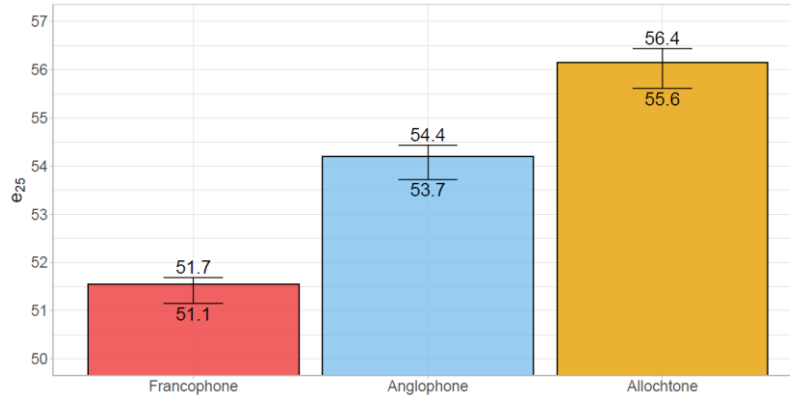
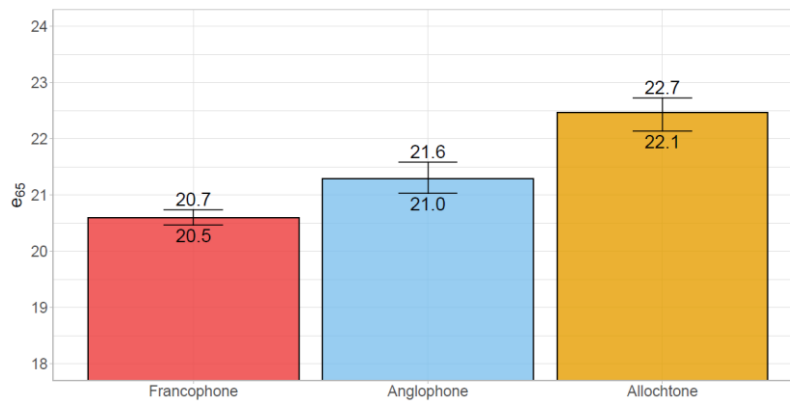


Figure E.7 Estimations de l'espérance de vie à 65 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011

### a. Femmes



### b. Hommes

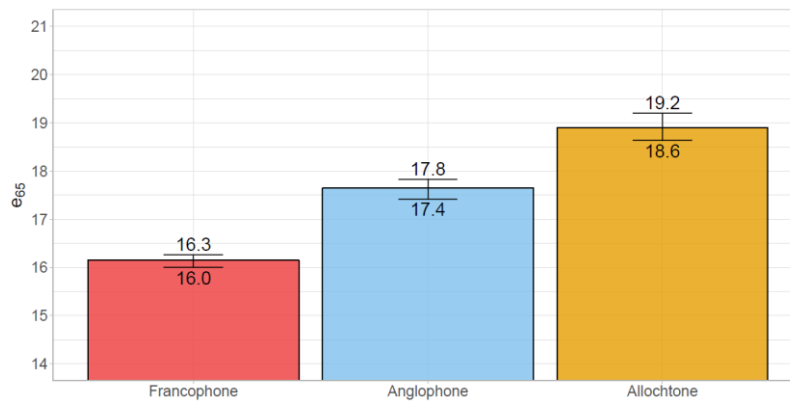
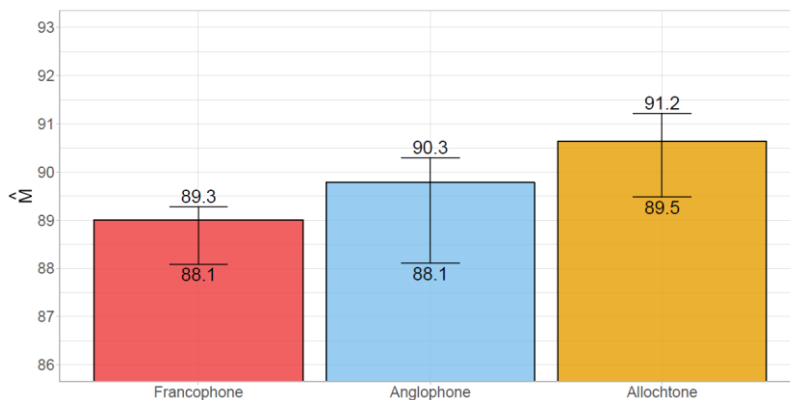


Figure E.8 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011

**a. Femmes**



**b. Hommes**

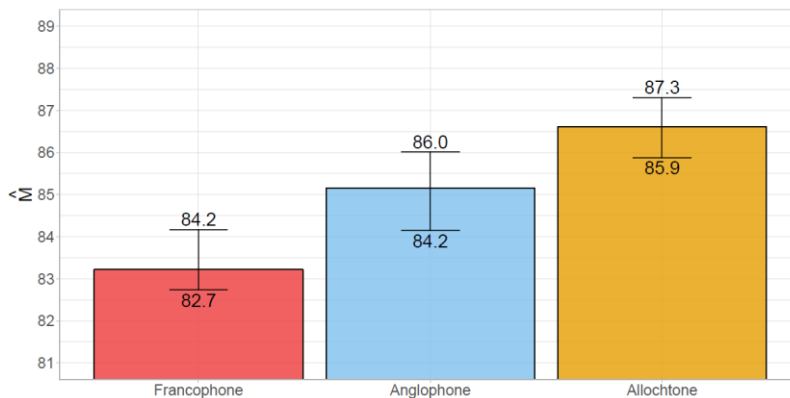
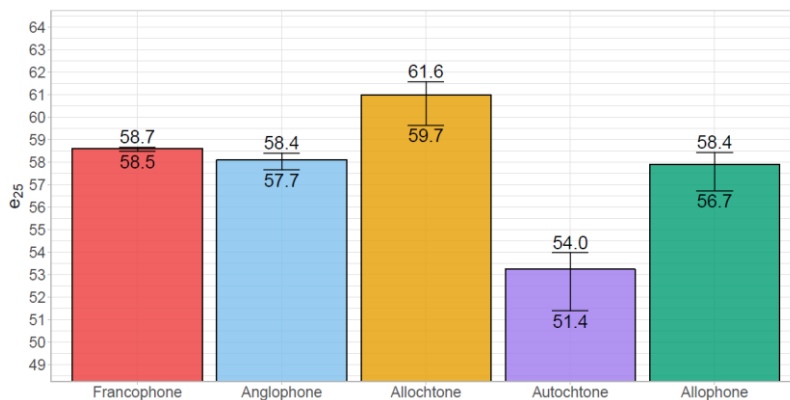


Figure E.9 Estimations de l'espérance de vie à 25 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011

**a. Femmes**



### b. Hommes

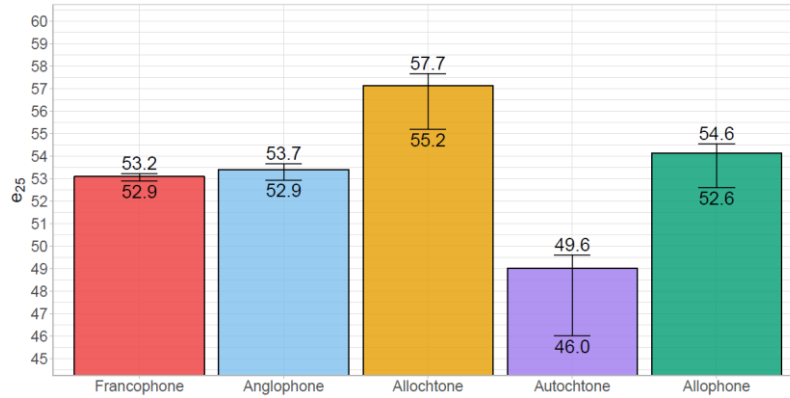
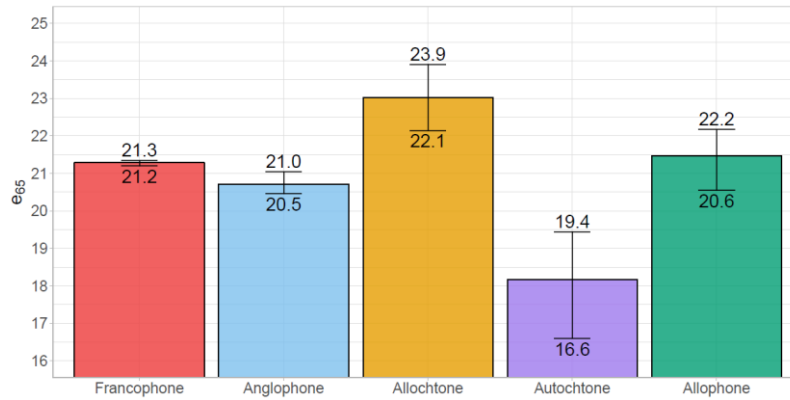


Figure E.10 Estimations de l'espérance de vie à 65 ans et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011

### a. Femmes



### b. Hommes

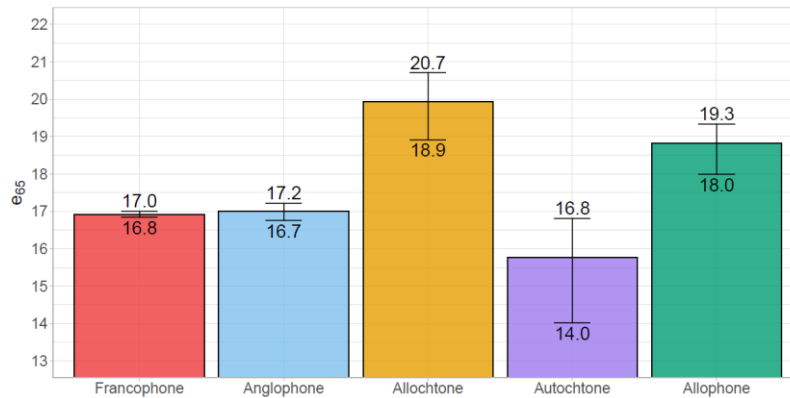
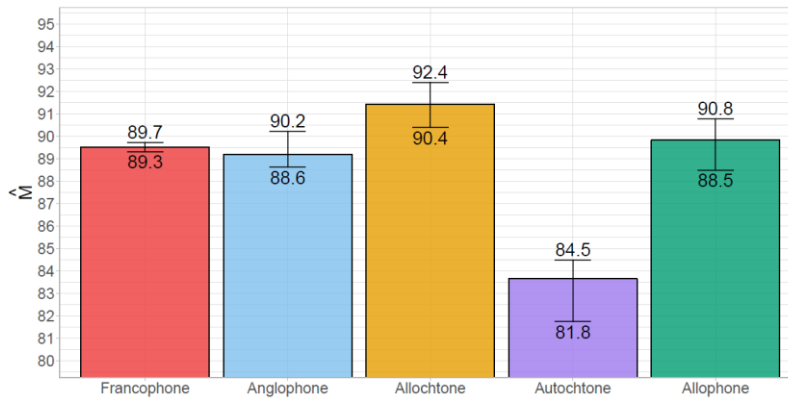
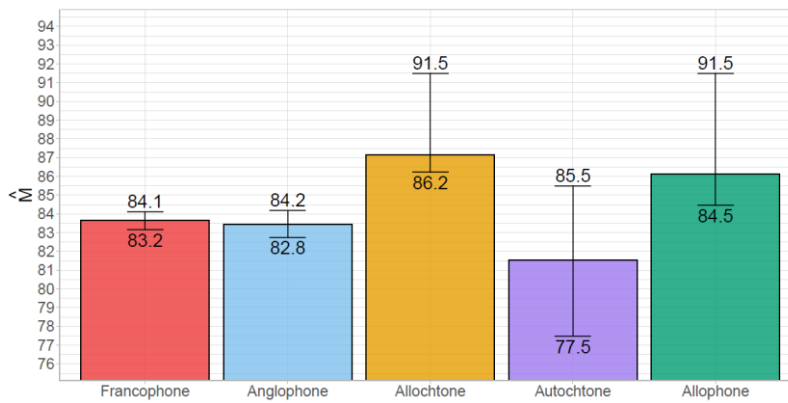


Figure E.11 Estimations de l'âge modal au décès et des intervalles de confiance bootstrap à 95% selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011

**a. Femmes**



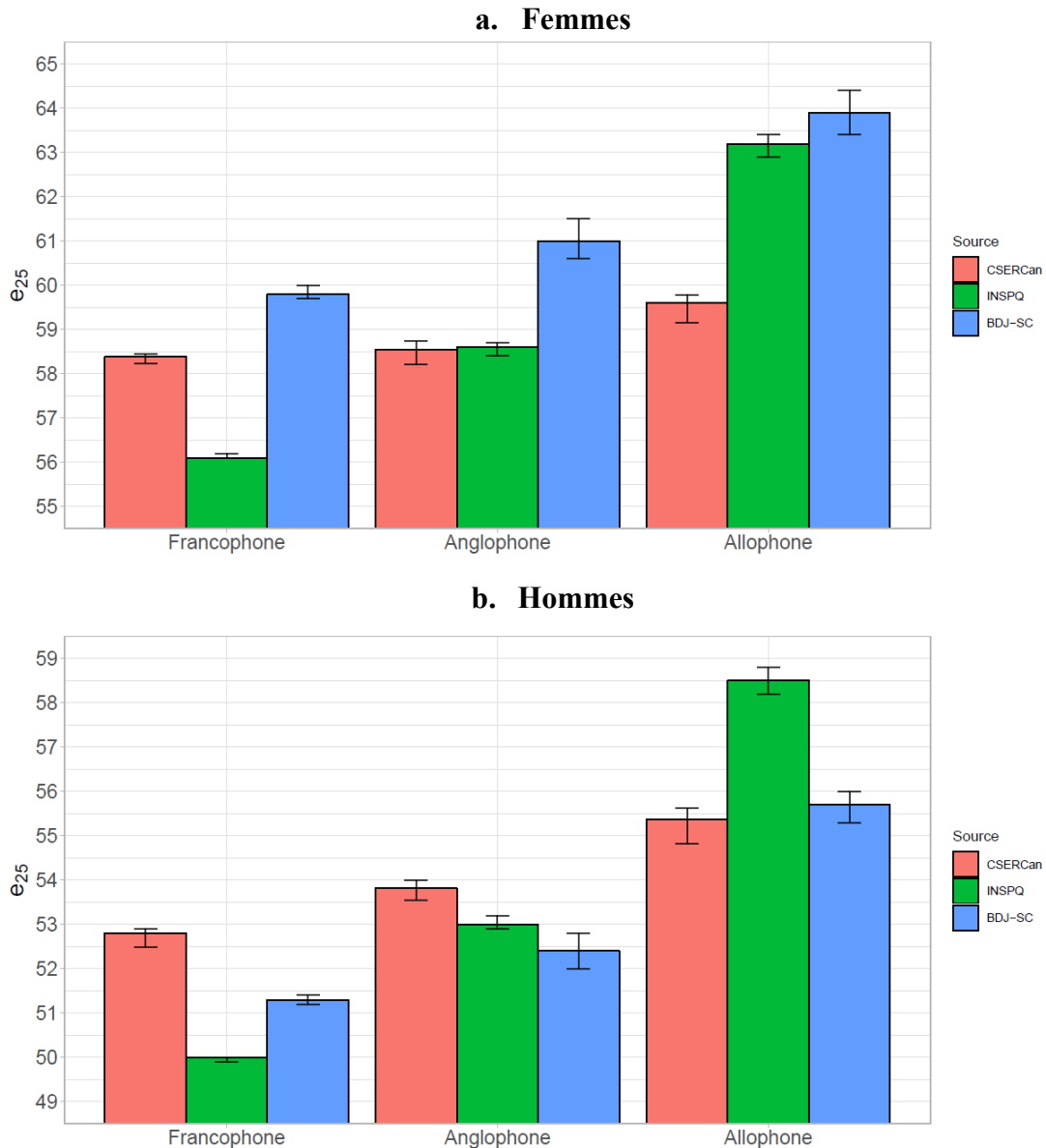
**b. Hommes**



# ANNEXE F

## Validation de $e_{25}$ et profils des groupes linguistiques

Figure F.1 Comparaison de  $e_{25}$  selon la source de données et le sexe, Québec



Sources : CSERCan — Calculs de l’auteur à partir de la Cohorte santé et environnement du recensement canadien de 1991. INSPQ — Calculs de Trempe et al. (2014) à partir du Registre des décès du Québec et des estimations de la population 1996 de l’ISQ. BDJ-SC — Calculs de Wilkins et al. (2008) à l’aide de la compilation spéciale de l’Étude canadienne de suivi de la mortalité selon le recensement 1991-2001.

Abréviations : BDJ-SC, Banque de données jumelées de Statistique Canada; INSPQ, Institut national de santé publique du Québec; ISQ, Institut de la statistique du Québec.

Figure F.2 Lieu de naissance selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011

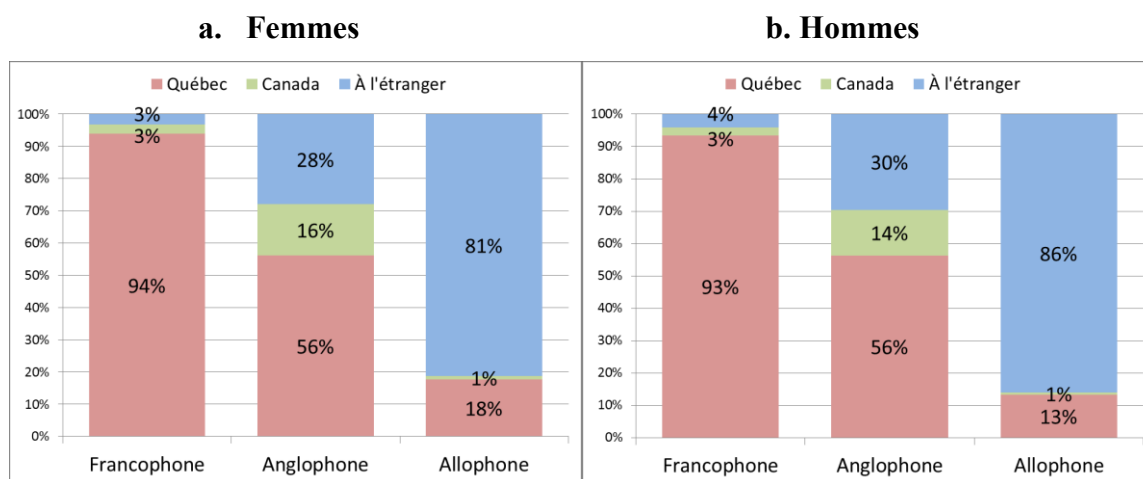
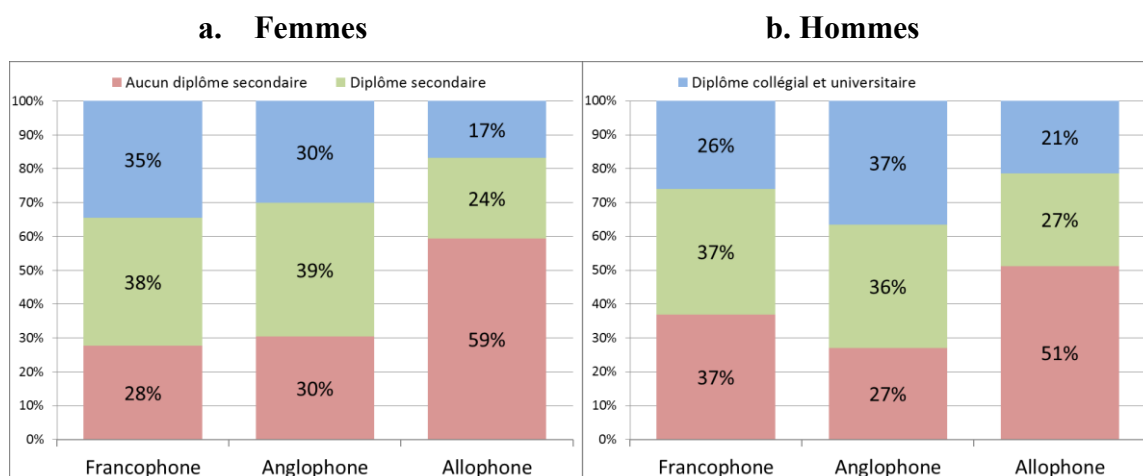


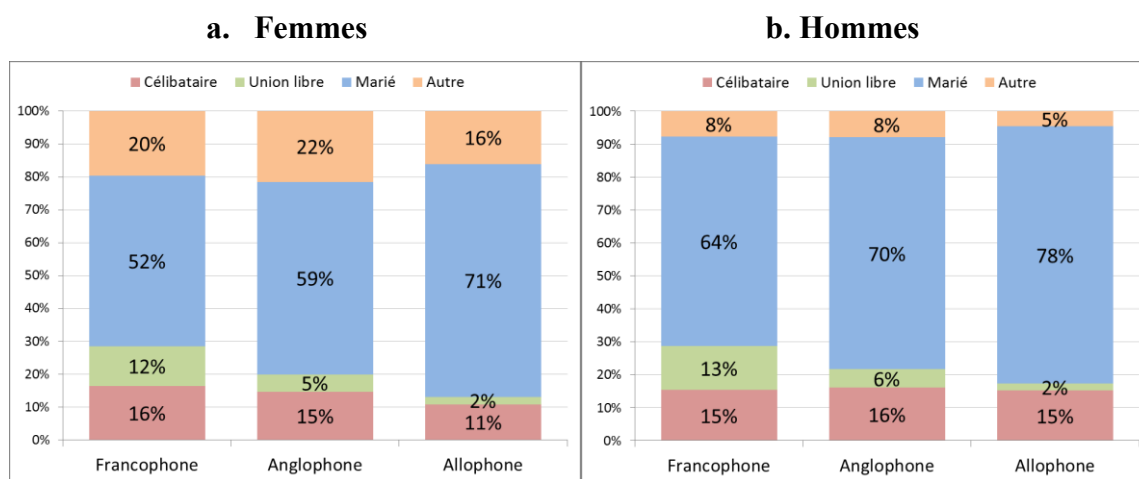
Figure F.3 Plus haut niveau de scolarité atteint selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011



Note : Nous avons réuni le diplôme collégial et universitaire pour atténuer les effets de générations; puisque les générations plus anciennes ont moins tendance à accéder à l'université que les plus récentes.

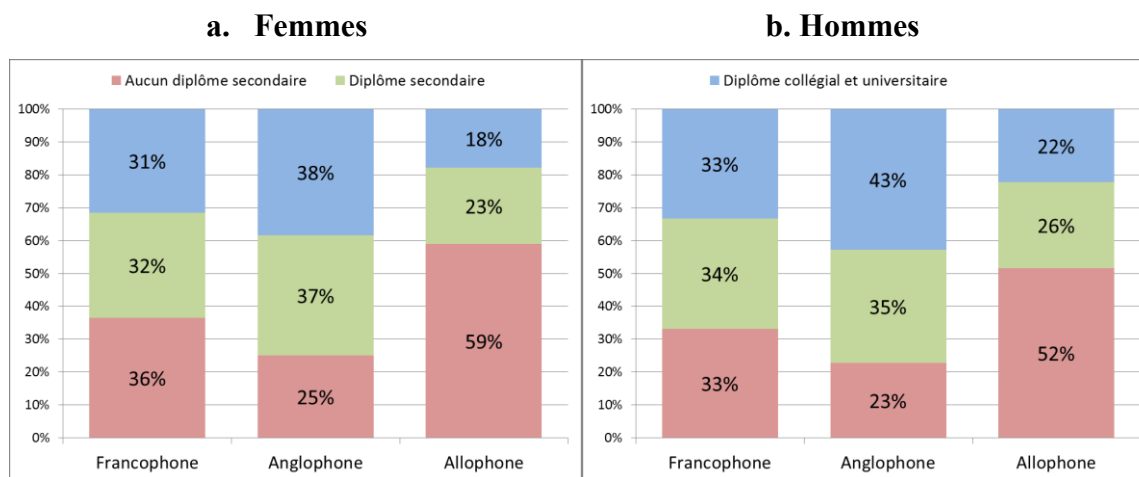


Figure F.4 État matrimonial selon le groupe linguistique et le sexe, CSERCan, Québec, 1991 à 2011



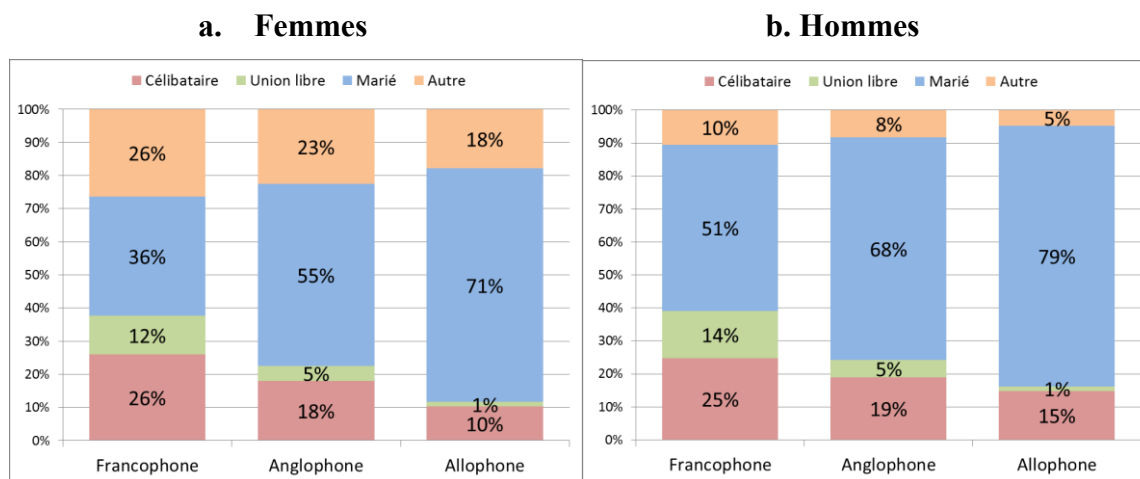
Note : Le groupe « Autre » inclut les individus séparés, divorcés et veufs.

Figure F.5 Plus haut niveau de scolarité atteint selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011



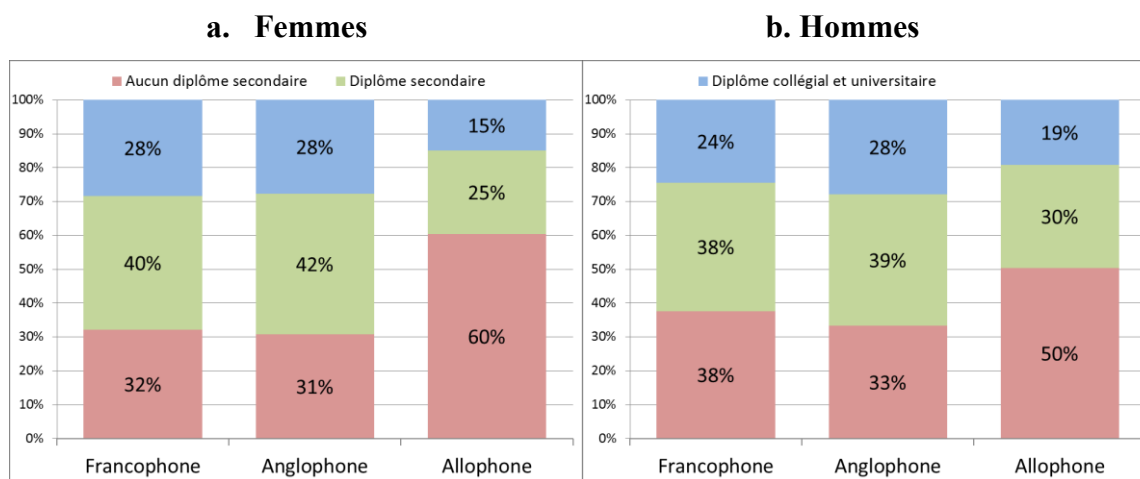
Note : Nous avons réuni le diplôme collégial et universitaire pour atténuer les effets de générations; puisque les générations plus anciennes ont moins tendance à accéder à l'université que les plus récentes.

Figure F.6 État matrimonial selon le groupe linguistique et le sexe, CSERCan, Île de Montréal, 1991 à 2011



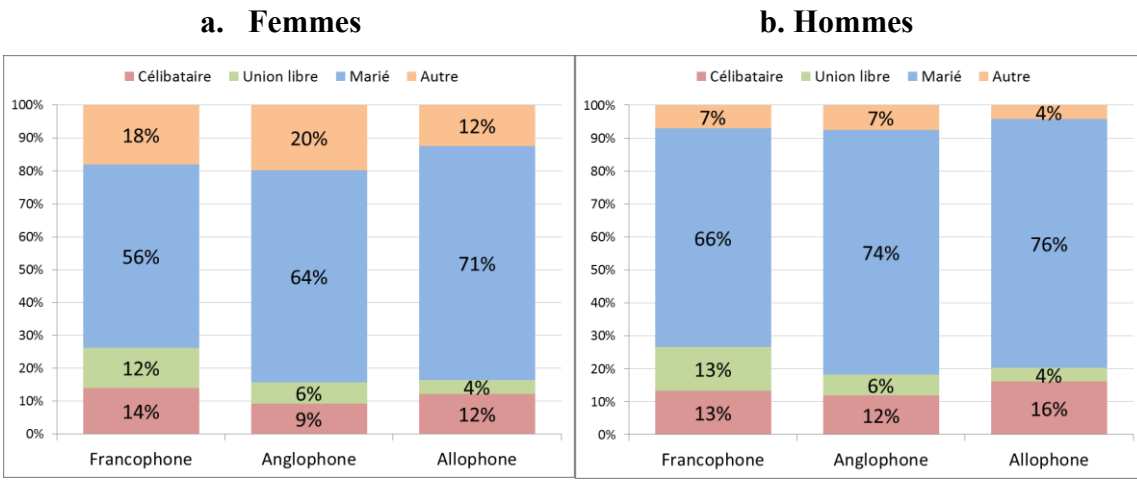
Note : Le groupe « Autre » inclut les individus séparés, divorcés et veufs.

Figure F.7 Plus haut niveau de scolarité atteint selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011



Note : Nous avons réuni le diplôme collégial et universitaire pour atténuer les effets de générations; puisque les générations plus anciennes ont moins tendance à accéder à l'université que les plus récentes.

Figure F.8 État matrimonial selon le groupe linguistique et le sexe, CSERCan, Reste du Québec, 1991 à 2011



Note : Le groupe « Autre » inclut les individus séparés, divorcés et veufs.

# Bibliographie

Base de données sur la longévité canadienne – BDLC. (2018). Université de Montréal (Canada). Disponible au <http://www.bdlc.umontreal.ca> (Données téléchargées le 16 janvier 2018).

BÉLAND, Paul. (1999). *Le français, langue d'usage public au Québec en 1997*, Rapport de recherche, Direction des études et recherches, Conseil de la langue française, Québec, 123 p.

BOURBEAU, R. et TREMBLAY, M. (1985). La mortalité et la fécondité selon le groupe linguistique au Québec, 1976 et 1981. *Cahiers québécois de démographie*, 14(1): 7–30.

BOURBEAU, R. et LEBEL, A. (2000). Mortality Statistics for the Oldest-Old: An Evaluation of Canadian Data. *Demographic Research*, 2(2).

BOURBEAU, R. (2002). L'effet de la « sélection d'immigrants en bonne santé » sur la mortalité canadienne aux grands âges. *Cahiers québécois de démographie* 31(2): 249–274.

BROWN, D. HAYWARD, M., MONTEZ, J. K., HUMMER, R., CHIU, C. T., et HIDAJAT, M. (2012). The significance of education for mortality compression in the United States. *Demography*, 49(3): 819–840.

CANON, L. (2017). *Analyse de la distribution des décès aux grands âges selon le niveau de scolarité à partir d'un suivi de la mortalité sur 20 ans au Canada*. Mémoire de maîtrise présenté à la Faculté des études supérieures de l'Université de Montréal, 75 p.

CHEUNG, S. L. K. et ROBINE, J.-M. (2007). Increase in common longevity and the compression of mortality: The case of Japan. *Population Studies*, 61(1): 85–97.

CHEUNG, S. L. K., ROBINE, J.-M., PACCAUD, F., et MARAZZI, A. (2009). Dissecting the compression of mortality in Switzerland, 1876-2005. *Demographic Research*, 21: 569–598.

EILERS, P. H. C. et MARX, B. D. (1996). Flexible smoothing with B-splines and penalties (with discussion). *Statistical Science*, 11(2): 89–102.

HORIUCHI, S., OUELLETTE, N., CHEUNG, S., et ROBINE, J. (2013). Modal age at death: Lifespan indicator in the era of longevity extension. *Vienna Yearbook of Population Research*, 1: 37-69.

KANNISTO, V. (1996). *The Advancing Frontier of Survival Life Tables for Old Age*, Monographs on Population Aging, 3. Odense University Press, Odense.

KANNISTO, V. (2000). Measuring the compression of mortality. *Demographic Research*, 3(6).

- KANNISTO, V. (2001). Mode et dispersion de la durée de vie. *Population : Biodémographie de la longévité*, 56(1-2): 183–197.
- KOISSI, M.-C., SHAPIRO, A. F. and HÖGNÄS, G. (2006). Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics*, 38: 1–20.
- LECOURS, C. (2011). *Les inégalités sociales dans la durée de vie la plus commune : la répartition des décès selon l'âge et le quintile de défavorisation au Québec en 2000-2002 et 2005-2007*. Mémoire de maîtrise présenté à la Faculté des études supérieures de l'Université de Montréal, 99 p.
- LÉGARÉ, J. (2004). Conséquences économiques, sociales et culturelles du vieillissement de la population., dans Graziella CASSELI, Jacques VALLIN et Guillaume WUNSCH (dir.), *Démographie : Analyse et synthèse VI – Population et sociétés*, Paris, Éditions de l'INED, chapitre 82: 117-135.
- LÉGARÉ, J. (2009). Le vieillissement des populations : incontournable certes... mais un plus, s'il est géré dans un esprit d'éthique intergénérationnelle. *Lien social et Politiques*, 62: 15–28.
- LEXIS, W. (1878). Sur la durée normale de la vie humaine et sur la théorie de la stabilité des rapports statistiques. *Annales de démographie internationale*, 2: 447–460.
- MCCULLAGH, P. and NELDER, J. A. (1989). *Generalized Linear Models*. Monographs on Statistics and Applied Probability. Chapman & Hall, London, deuxième édition.
- Nations Unies. (2009). *Principes et recommandation concernant les recensements de la population et des logements, deuxième révision*, [https://unstats.un.org/unsd/...social/.../Principles.../Population.../Series\\_M67Rev2-F.pdf](https://unstats.un.org/unsd/...social/.../Principles.../Population.../Series_M67Rev2-F.pdf), consulté le 4 avril 2018.
- OMARIBA, D.W.R., NG, E., et VISSANDJÉE, B. (2014). Differences between immigrants at various durations of residence and host population in all-cause mortality, Canada 1991–2006. *Population Studies* 68(3): 339–357.
- OUELLETTE N. et BOURBEAU R. (2011). Changes in the age-at-death distribution in four low mortality countries: A nonparametric approach. *Demographic Research*, 25: 595–628.
- OUELLETTE N., BOURBEAU R., et CAMARDA C. G. (2013). Regional disparities in Canadian adult mortality: A comparative study based on smoothed mortality surfaces and age-at-death distributions. *Canadian Studies in Population*, 39(3-4): 79–106.

ROBITAILLE, N., BOURBEAU R., GIRARD C., et TREMBLAY, M. (1992). L'avenir des groupes linguistiques au Québec : équilibres et options. *Cahiers québécois de démographie*, 21(2): 147–161.

STATA CORP (2012). *Stata statistical software: Release 12* [Software]. College Station, TX: StataCorp LP.

STATISTIQUE CANADA (2016). *Guide de l'utilisateur : Cohorte Santé et Environnement du Recensement du Canada de 1991 (CSERCAN)*, Division de l'analyse de la santé.

STATISTIQUE CANADA (2015). *Études de couverture du Recensement de 1971 au Recensement de 2011*, <https://www12.statcan.gc.ca/census-recensement/2011/ref/guides/98-303-x/tbl/tbl11.1-fra.cfm>, consulté le 30 juillet 2018.

TERMOTE, Marc. (1992). Le comportement démographique des groupes linguistiques à Montréal. *Cahiers québécois de démographie*, 21(2): 77–94.

TERMOTE, Marc. (1999). *Perspectives démolinguistiques du Québec et de la région de Montréal à l'aube du XXI<sup>e</sup> siècle. Implications pour le français langue d'usage public*. Québec : Conseil de la langue française, <http://www.cslf.gouv.qc.ca/publications/pub151/b151.pdf>, consulté le 7 novembre 2018.

TERMOTE, Marc. (2001). *L'évolution démolinguistique du Québec et du Canada*. Rapport pour la Commission sur l'avenir politique et constitutionnel du Québec (Bélanger-Campeau), Volume 2: 239–329.

TERMOTE, Marc. (2008). *L'intégration linguistique des immigrants au Québec. Problèmes méthodologiques, résultats empiriques et implications politiques*. Département de démographie, Université de Montréal, <https://www.uclouvain.be/cps/ucl/doc/demo/documents/Termote.pdf>, consulté le 6 juin 2018.

TERMOTE, Marc. (2011). *Perspectives démolinguistiques du Québec et de la région de Montréal (2006-2056)*. Office québécois de la langue française, 204p.

THATCHER, A., KANNISTO, V. et VAUPEL, J. W. (1998). *The force of mortality at ages 80 to 120*. Monographs on Population Aging, 5. Odense University Press, Odense.

TJEPKEMA, M., WILKINS R., SENÉCAL S. ET GUIMOND É. (2009). *Mortality of Métis and Registered Indian adults in Canada: An 11-year follow-up study*. Health Reports 2009; 20(4): 31-51.

TREMPE, N., DIOGO BARRY, A., LUSSIER, M-H. ET LO, E. (2013). La mortalité évitable des deux principales communautés linguistiques du Québec. Institut national de santé publique du Québec, 67p.

TREMPE, N., BOIVIN, M.-C., LO, E., et DIOGO BARRY, A. (2014). L'utilisation de la variable sur la langue d'usage à la maison du Registre des décès du Québec. *Cahiers québécois de démographie*, 43(1): 163–180.

TROVATO, Frank. (1985). Mortality differences among Canada's indigenous and foreign-born populations, 1951-1971, *Canadian Studies in Population*, 12, 1: 49-80.

TURRA, C.M., et ELO, I.T. (2008). The impact of salmon bias on the Hispanic mortality advantage: new evidence from social security data, *Population Research and Policy Review* 27(5): 515–530.

WILKINS, R., TJEKEMA, M., MUSTARD, C., et CHOINIÈRE, R. (2008). The Canadian census mortality follow-up study, 1991 through 2001. *Health Reports*, 19: 25–43.

WILMOTH, J. R. et HORIUCHI, S. (1999). Rectangularization revisited: Variability of age at death within human populations. *Demography*, 36(4): 475–495.

