IRM | institut du risk management

Mémoire présenté devant l'Institut du Risk Management pour la validation du cursus à la Formation d'Actuaire de l'Institut du Risk Management et l'admission à l'Institut des actuaires le

Par : Ulcinaite Agne

Titre : Prospective mortality by cause of death modelling

Confidentialité : ☑ NON ☐ OUI (Durée : ☐ 1an ☐ 2 ans)
Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

Membres présents du jury de l'Institut des actuaires :

Membres présents du jury de l'Institut du Risk Management :

Secrétariat :

Bibliothèque :

Entreprise :
Nom : SCOR
Signature et Cachet :

SCOR SE
5, avenue Kléber
F - 75795 PARIS CEDEX 16
Tél. : + 33 (0)1 58 44 70 00

Directeur de mémoire en entreprise :
Nom : TOMAS Julien
Signature :

Invité :
Nom :
Signature :

Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels
(après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise

Signature(s) du candidat(s)

Remerciements

Je tiens à remercier tout spécialement Julien Tomas pour sa disponibilité et le temps consacré à l'encadrement de mon mémoire.

Je remercie également Razvan Ionescu et Adam Shao pour avoir relu et corrigé mon mémoire. Leurs conseils de rédaction ont été très précieux.

J'exprime également ma reconnaissance envers mes collègues d'équipe BRM qui m'ont apporté leur soutien moral et intellectuel tout au long de ma démarche.

Enfin, ma gratitude va à mon mari, mon fils et mes parents, qui m'ont apporté une motivation constante tout au long de cette epreuve.

À tous ces intervenants, je présente mes remerciements, mon respect et ma gratitude.

Résumé

La vision de la mortalité future est cruciale pour les compagnies d'assurance-vie. La mortalité prédite est utilisée à la fois pour la tarification - impact direct sur la profitabilité, et pour la gestion des risques - employée pour définir la solvabilité des entreprises.

Les compagnies d'assurance utilisent des projections de mortalité future pour estimer le coût probable des risques liés à la durée de vie humaine. Les données de mortalité toutes causes sont souvent utilisées par les actuaires pour établir les hypothèses de mortalité. En revanche, l'utilisation de la modélisation de la mortalité par cause permet de construire des scénarios hypothetiques pour produire une analyse de type "what-if", qui sont constamment demandées à la fois par les régulateurs et les parties prennantes internes.

Cette thèse introduit la modélisation de la mortalité par cause de décès dans le cadre des risques concurrents et montre comment les hypothèses sur la structure de dépendance entre les causes affectent les projections de mortalité futures. Nous évaluons les impacts sur l'espérance de vie dans le scénario central et dans des scénarios hypothétiques sur la maladie d'Alzheimer et de la démence. Ces scénarios hypothétiques sont utilisés pour illustrer l'impact probable de la diversification sur les activités de mortalité et de longévité.

Mots clés: Mortalité, Longevité, Cause de décès, Risques concurrents, Copule archimédienne, Analyse de scénario.

Abstract

Mortality prediction is crucial for Life insurance companies as the future mortality prediction is used for both policy pricing - direct impact on profitability, and risk management - to set business acceptance capacity.

Insurance companies use future mortality projections to estimate probable cost of risks related to human lifespan. The all-cause mortality data is often used by the actuaries to set mortality assumptions. Otherwise, using cause-specific mortality modelling permits construction of cause-specific scenarios to produce a "what-if" type-of analysis, which are constantly requested by both regulators and internal stakeholders like risk management.

This thesis introduces cause-of-death mortality modelling within competing risk framework and shows how assumptions on cause dependency structure affect future mortality projections. We assess impacts on life expectancy in central scenario and in hypothetical scenarios on Alzheimer's and dementia cause. These hypothetical scenarios are used to illustrate probable diversification impact on mortality and longevity businesses.

Keywords: Mortality, Longevity, Cause of death, Competing risks, Archimedian copula, Scenario analysis.

Note de synthèse

Contexte

Les compagnies d'assurance sont tenues de produire des scénarios de type "what-if" à la fois par les régulateurs et les parties prenantes internes, comme les équipes de management du risque. À l'aide de scénarios de stress test, les régulateurs cherchent à valider le modèle interne et à identifier les situations de crise probables qui menaceraient la viabilité de l'entreprise. La direction de la compagnie est également intéressée par l'analyse des scénarios probables qui permet d'évaluer la résilience aux chocs, d'approuver les décisions sur l'appétit au risque et d'évaluer l'impact de la diversification du portefeuille.

Construire des scénarios de type "what-if" pour les risques de mortalité et de longévité se traduit par un besoin de créer des scénarios hypothétiques sur une ou plusieurs causes de décès. Cela permet d'élaborer une histoire autour d'un modèle mathématique et d'interpréter les résultats.

Par conséquent, le scénario sur une ou plusieurs causes de décès nécessite une modélisation de la mortalité à un niveau granulaire spécifique. Comme la mort n'est pas répétitive et est associée à une cause unique, toutes les causes se disputent la vie d'une personne. La modélisation de la mortalité par cause est soumise à un cadre de risques concurrents, en ce sens que la probabilité de chaque événement concurrent est en quelque sorte régulée par les autres événements concurrents.

Le premier chapitre de ce mémoire présente le contexte et les motivations de l'étude. Dans un deuxième temps, nous donnons un bref aperçu de l'historique et introduisons les concepts utilisés dans la modélisation de la mortalité. Différentes manières d'aborder la projection de la mortalité sont présentées dans le deuxième chapitre, en commençant par les modèles de mortalité toutes causes, puis en affinant le concept de modélisation de la mortalité par cause de décès plus granulaire. Le troisième chapitre présente les données utilisées. Les approches de modélisation de la mortalité par cause sont ensuite abordées, en supposant d'abord l'indépendance, puis différents niveaux de dépendance entre les causes à l'aide des copules Archimédiennes de survie. Le dernier chapitre présente des scénarios *footprint* sur la maladie d'Alzheimer et de la démence. Ces scénarios permettent d'illustrer un éventail de conclusions sur la diversification entre les portefeuilles de mortalité et de longévité, en s'appuyant sur la structure de dépendance supposée entre les causes.

Nous utilisons les données de la mortalité masculine U.S. par cause de décès et par

niveau de scolarité atteint. En retenant que les niveaux d'éducation atteint les plus élevés (équivalent à bac+3 et supérieur), nous proposons un proxy de la population assurée pour l'application.

Cadre d'indépendance

Lors de la modélisation de la mortalité par cause de décès, nous avons besoin d'introduire les notions d'intensité brute et nette de mortalité. Alors que la mortalité brute est directement obtenue à partir du nombre de décès et des données d'exposition, la mortalité nette estime la probabilité de mourir dans le monde hypothétique où vous ne pouvez pas mourir de causes autres que la cause d'intérêt. Il s'agit d'une construction théorique et selon l'hypothèse sur la structure de dépendance entre les causes, la relation entre la mortalité nette et brute sera différente.

La première approche de modélisation par cause presentée suppose l'indépendance entre les causes de décès, ce qui signifie que les intensités de mortalité brute et nette sont identiques. Chaque cause de décès i est modélisée séparément par le modèle de mortalité Lee and Carter (1992) classique:

$$\log(\mu_{x,t,i}) = \alpha_{x,i} + \beta_{x,i}\kappa_{t,i} + \epsilon_{x,t,i}, \tag{1}$$

ou $\alpha_{x,i}$ est une mortalité moyenne de cause i à l'âge x, $\kappa_{t,i}$ décrit l'évolution globale de la mortalité de cause i dans le temps, $\beta_{x,i}$ représente la sensibilité à la dynamique $\kappa_{t,i}$ par âge et $\epsilon_{x,t,i}$ décrit le terme d'erreur. Les projections des intensités de mortalité pour la cause i sont obtenues en extrapolant la tendance de mortalité $\kappa_{t,i}$ avec marche aléatoire avec drift comme processus stochastique :

$$\kappa_{t+1,i} = \kappa_{t,i} + \delta_i + \eta_{t,i}, \tag{2}$$

ou δ_i est un parametre de dérive pour la cause i et $\eta_{t,i} \sim N(0, \sigma^2)$.

L'approche supposant une indépendance entre les causes prévoit une stagnation de la mortalité toutes causes à long terme pour l'âge de 55 ans et amelioration de la mortalité sur toute la période de projection pour l'âge de 75 ans, voir figures 1 et 2.

Cadre de dépendance basé sur la copule

Le cadre de la copule archimédienne est utilisé pour introduire une structure de dépendance entre les causes de décès, une approche proposée par Li and Lu (2019). En spécifiant la copule de survie $\mathbf{S}$ comme copule de survie archimédienne, Li and Lu (2019) a montré que les intensités nettes par cause, par le biais de fonctions de survie nettes, peuvent être estimées de manière unique à partir des données en utilisant des

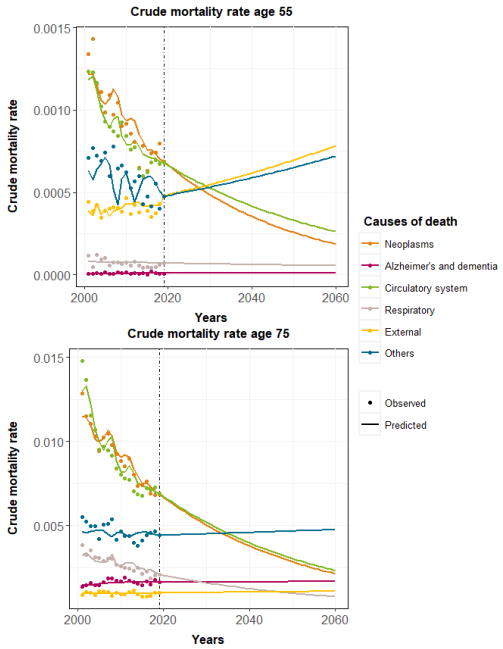

Figure 1: Mortalité par cause

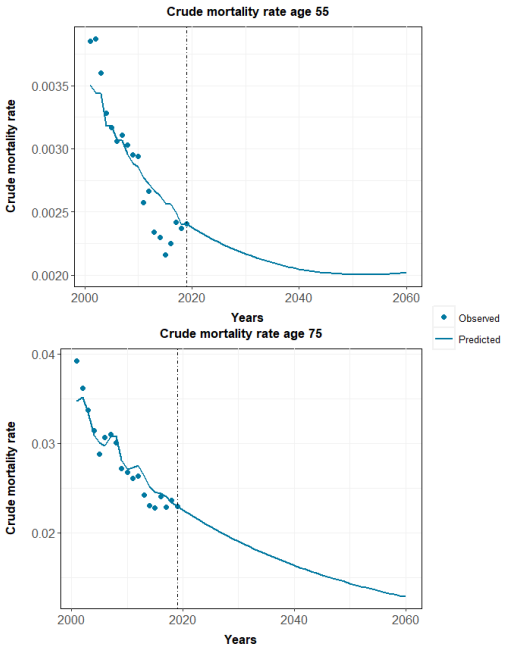

Figure 2: Mortalité toutes causes

intensités brutes et la copule :

$$S_i(t) = \psi\left[-\int_0^t \frac{exp\left(-\int_0^s \sum_{j=1}^m \mu_j(u)du\right)}{\psi' \circ \psi^{-1} \circ exp\left(-\int_0^s \sum_{j=1}^m \mu_j(u)du\right)} \mu_i(s)ds\right], \ \forall i = 1, \ldots, m. \quad (3)$$

On utilise la copule de Clayton avec générateur $\psi(t) = (1+t)^{-1/\theta}$ et on choisie le paramètre de dependence $\theta = 1$ et $\theta = 4$. Les valeurs du paramètre θ définissent l'intensité de la dépendance, et plus la valeur θ est élevée, plus la dépendance assumée entre les causes est forte.

Nous observons que les projections de mortalité par cause varient en fonction de la dépendance entre les causes supposées, voir les figures 1, 3 et 4. La mortalité toutes causes confondues qui en résulte diffère également: l'indépendance entraîne une stagnation des taux de mortalité dans la seconde moitié de la période de projection pour l'âge de 55 ans, tandis que $\theta = 1$ et $\theta = 4$ fournissent des projections de mortalité décroissantes. Inversement, la baisse des taux de mortalité à 75 ans est projetée plus élevée sous l'hypothèse d'indépendance, voir les figures 2, 5 et 6.

Les projections de mortalité obtenues donnent une espérance de vie plus faible que dans le cadre de l'indépendance, et le ralentissement de l'amélioration de la mortalité est plus fort pour une valeur θ plus élevée. Le Tableau 1 présente les espérances de vie périodiques résiduelles projetées à 55 et 75 ans sous différentes hypothèses de dépendance. Nous observons des variations importantes dans les gains projetés, pour les âges de 55 et 75 ans à la fin de la période de projection, la différence est de plus de 3 ans entre les hypothèses d'indépendance et de forte dépendance.

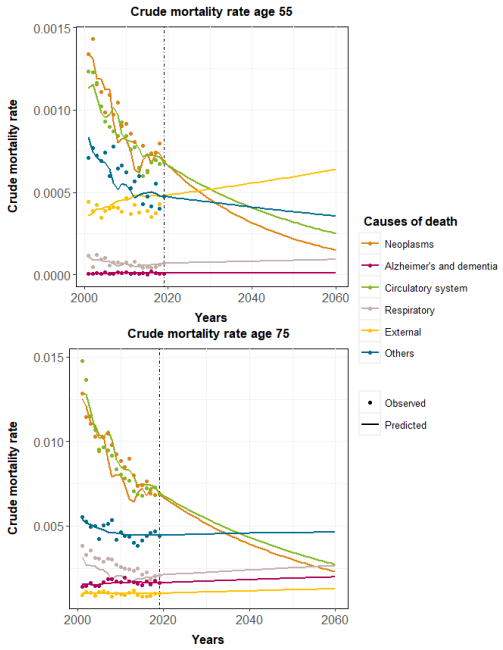

Figure 3: Mortalité nette, $\theta = 1$

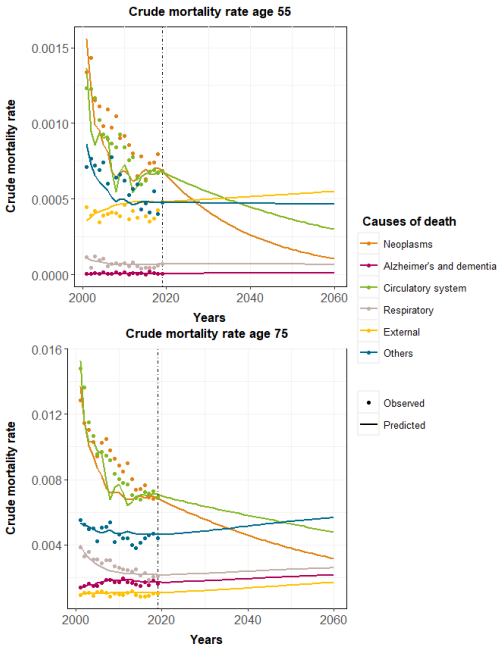

Figure 4: Mortalité nette, $\theta = 4$

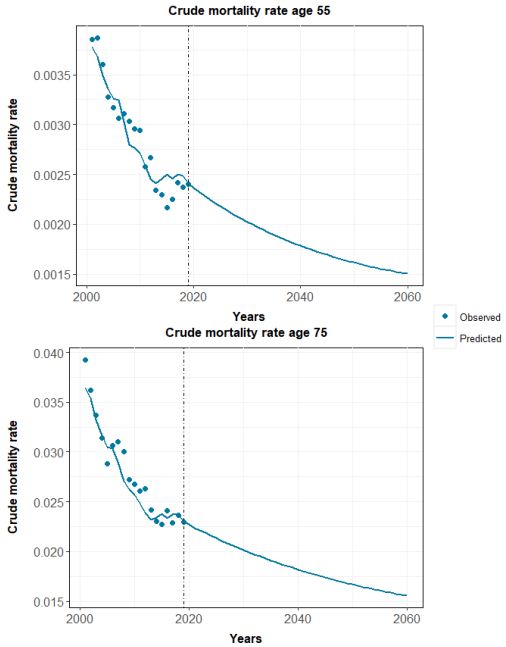

Figure 5: Mortalité globale, $\theta = 1$

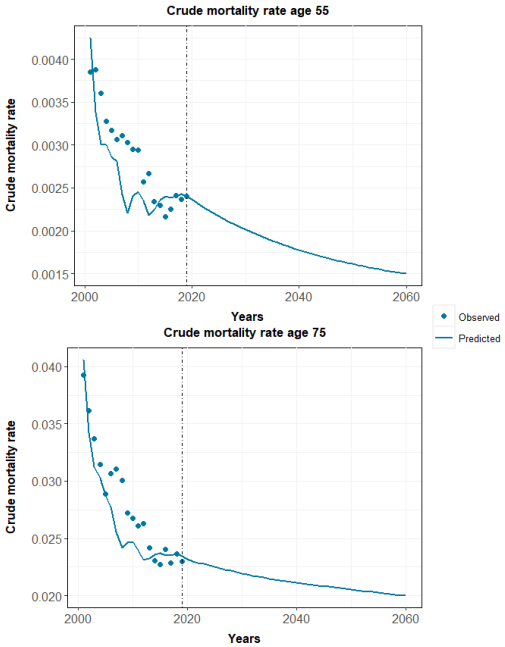

Figure 6: Mortalité globale, $\theta = 4$

Esperance de vie residuelle (années)				
Âge 55	Année 2001	Année 2019	Année 2040	Année 2060
$\theta = 0$	26	29	30.9	32.1
$\theta = 1$	26	29	30.1	30.7
$\theta = 4$	26	29	29.3	29

Âge 75	Année 2001	Année 2019	Année 2040	Année 2060
$\theta = 0$	10.1	11.8	13	13.9
$\theta = 1$	10.1	11.8	12.3	12.4
$\theta = 4$	10.1	11.8	11.6	10.8

Table 1: Espérances de vie périodiques résiduelles

Les scenarios Footprint

La modélisation de la mortalité par cause de décès permet d'intégrer des scénarios par cause sur une ou plusieurs causes d'intérêt. Nous proposons deux scénarios sur la mortalité future de la maladie d'Alzheimer et la démence et nous évaluons les impacts sur les *model points* représentant les portefeuilles de la mortalité et de la longévité et leur diversification. Les deux scénarios proposés résultent des progrès combinés des mesures de prévention et des innovations en matière de détection et de traitement des risques.

Dans le premier scénario, la mortalité due à la maladie d'Alzheimer et à la démence est réduite de 66% par rapport à la projection du scénario central sur les 15 prochaines années. Après cet horizon d'amélioration de 15 ans, la probabilité spécifique à l'âge de décès dû à la maladie d'Alzheimer et à la démence est supposé rester à 33% de sa projection pré-scénario.

Le deuxième scénario est plus extrême à la fois sur le délai aprés la percée médicale et sur l'impact significatif du scénario proposé. Nous supposons une élimination de la mortalité et de la perte d'autonomie causés par les maladies d'Alzheimer et de démence dans les 5 prochaines années de la projection. Les deux scénarios sont des vues très positives des résultats futurs possibles en ce qui concerne l'impact sur la santé, particulierement pou le second scenario.

Les figures suivantes illustrent le Scénario 1 sur la mortalité par cause et globale.

Figures 7, 8 and 9 montrent comment la diminution de la mortalité due à la maladie d'Alzheimer et à la démence n'a pas d'impact sur les autres causes sous l'hypothèse d'indépendance. En revanche, la prise en compte de la dépendance entre les causes entraîne un transfert des décès vers d'autres causes, et le transfert est plus important avec un paramètre θ plus élevé.

Les vies "sauvées" des décès dus à la maladie d'Alzheimer et à la démence transférées

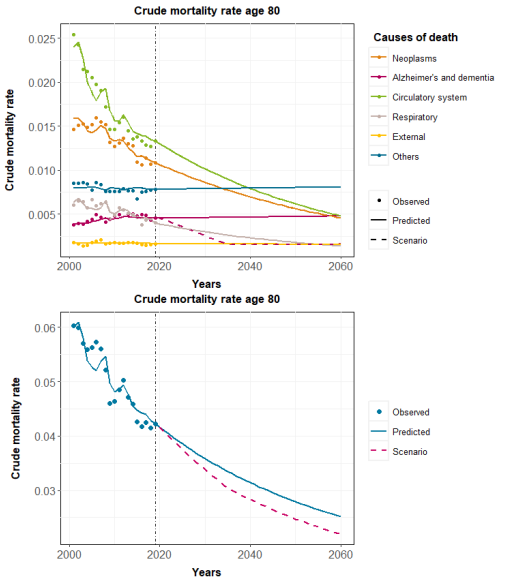

Figure 7: Scenario 1, independence

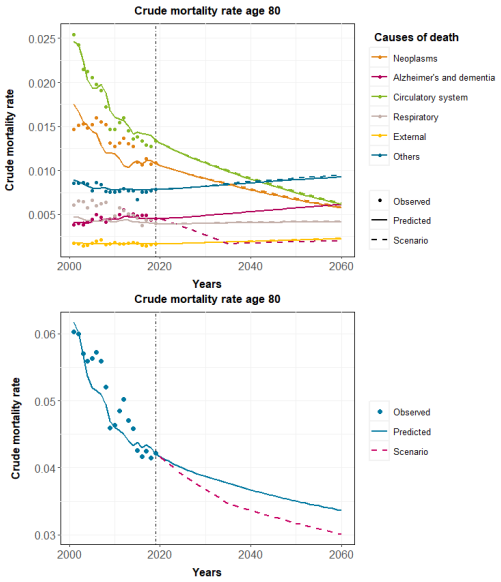

Figure 8: Scenario 1, $\theta = 1$

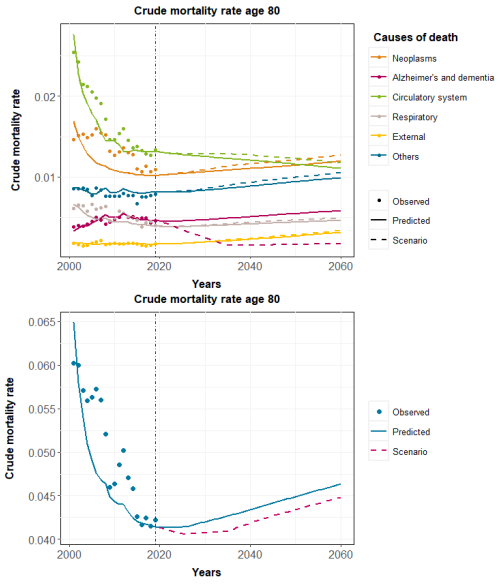

Figure 9: Scenario 1, $\theta = 4$

à d'autres causes entraînent des variations importantes des gains d'espérance de vie pour différentes hypothèses de dépendance. Le Tableau 2 décrit les gains d'espérance de vie suite à une réduction de 66 % de la mortalité due à la maladie d'Alzheimer et à la démence et nous observons des variations importantes dans les gains du scénario, en particulier entre les cas avec des causes indépendantes et en supposant une forte dépendance ($\theta = 4$).

Residual life expectancy					
Âge	Dependence	Cas	Année 2019	Année 2040	Année 2060
55	$\theta = 0$	Scénario central	29	30.9	32.1
		Scénario 1	29	31.4	32.7
		Δ (mois)	-	5.9	7.3
	$\theta = 1$	Scénario central	29	30.1	30.7
		Scénario 1	29	30.6	31.2
		Δ (mois)	-	5	6.4
	$\theta = 4$	Scénario central	29	29.3	29
		Scénario 1	29	29.5	29.1
		Δ (mois)	-	2.2	1.7
75	$\theta = 0$	Scénario central	11.8	13	13.9
		Scénario 1	11.8	13.5	14.5
		Δ (mois)	-	5.7	7
	$\theta = 1$	Scénario central	11.8	12.3	12.5
		Scénario 1	11.8	12.7	13
		Δ (mois)	-	4.7	6
	$\theta = 4$	Central scenario	11.8	11.6	10.9
		Scénario 1	11.8	11.7	11
		Δ (mois)	-	1.7	1.1

Table 2: Espérances de vie périodiques résiduelles, Scénario 1

Le deuxième scénario est plus extrême et par conséquent les gains d'espérance de vie sont plus importants. Cependant la conclusion générale est la même avec des gains d'espérance de vie offerts par le scénario avec indépendance les plus elevés et les gains les plus modestes - pour $\theta = 4$.

Pour illustrer comment les résultats précédents pourraient avoir un impact sur les résultats d'une compagnie d'assurance, nous utilisons un *model point* représentatif pour chaque portefeuille de mortalité et de longévité avec la valeur actuelle des sinistres égales dans le scénario central et nous évaluons les gains et les pertes pour le Scénario 1 et le Scénario 2, voir Tableau 3.

L'hypothèse de l'indépendance entre les causes entraîne la plus grande variation pour le montant de sinistres pour les portefeuilles de Mortalité et Longévité, ce qui était à prévoir car les gains d'espérance de vie étaient les plus élevés dans le cadre de l'indépendance, tandis que l'hypothèse d'une forte dépendance entre les causes donne lieu à des fluctuations modérées. Les deux scénarios sous l'hypothèse d'indépendance

produisent une baisse du montant total des sinistres, $\theta = 1$ donne lieu à une légère augmentation des sinistres pour le Scénario 1 et à une baisse dans le Scénario 2, et enfin, dépendance élevée génère des montants totaux de sinistres plus élevés pour les deux scénarios.

Causes indépendantes ($\theta = 0$)			
Ligne de business	Scénario central	Δ Scénario 1	Δ Scénario 2
(1) Mortalité (sinistres)	2000	-53	-83
(2) Longevité (sinistres)	2000	+33	+82
Total (1)-(2)	0	-20	-1

Copule de Clayton, $\theta = 1$			
Ligne de business	Scénario central	Δ Scénario 1	Δ Scénario 2
(1) Mortalité (sinistres)	2000	-37	-57
(2) Longevité (sinistres)	2000	+28	+68
Total (1)-(2)	0	-9	+9

Copule de Clayton, $\theta = 4$			
Ligne de business	Scénario central	Δ Scénario 1	Δ Scénario 2
(1) Mortalité (sinistres)	2000	-10	-14
(2) Longevité (sinistres)	2000	+14	+30
Total (1)-(2)	0	+4	+16

Table 3: Impact sur la diversification pour différentes valeurs θ

Nous observons que le fait de ne pas prendre en compte la structure de dépendance entre les causes pour modéliser les scénariospourraient générer des gains hypothétiques. D'autre part, travail dans un cadre de risques concurrents entraînerait des gains plus modérés de la diversification entre la mortalité et la longévité. En supposant que les causes sont indépendantes, les deux scénarios entraînent des sinistres inférieurs à ceux du scénario central, en particulier pour le Scénario 1. En passant au cas de faible dépendance ($\theta = 1$), la réduction des sinistres totaux est deux fois plus petit et nous observons une augmentation des sinistres pour le Scénario 2 (+9 vs -1). L'hypothèse d'une forte dépendance entre les causes a un impact encore plus fort et pour les deux scénarios on observe une augmentation des sinistres.

Les impacts financiers des scénarios sont sensiblement différents selon l'hypothèse de dépendance utilisée lors de la modélisation de la mortalité par cause. La variété des conclusions concernant les effets de diversification et la résilience des entreprises illustre l'importance de travailler dans un cadre de risques concurrents.

Executive summary

Context

Insurance companies are required to produce "what-if" type of scenarios by both regulators and internal stakeholders. Using stress test scenarios, regulators seek to validate internal model and to identify probable crisis situations that would threaten viability of the company. Senior management is also particularly interested in probable scenario analysis which permits assessing business resilience to shocks, supporting business acceptance and risk appetite decisions and evaluating portfolio diversification impact.

Building "what-if" type of scenarios for Mortality and Longevity risks translates into a need to create hypothetical scenarios on one on several causes of death. This allows building a story around mathematical model and interpret the results.

Consequently, scenario on one or several causes of death require mortality modelling at a granular cause-specific level. As death is not repetitive and is associated to a single cause, all causes compete with one another for the life of a person. Cause-specific mortality modelling is subject to competing risk framework, in a sense that probability of each competing event is somehow regulated by the other competing events.

This thesis first gives a global overview of mortality modelling history and existing methods. Secondly, we propose cause-specific mortality modelling approach by assuming independence between causes which is developed in the second chapter. In the following, we refine the method by taking into account dependency structure between causes within Archimedian copula framework. We assess impact of different approaches on life expectancy in central scenario and also introduce specific scenarios on Alzheimer's and dementia cause that are discussed in Chapter 4. These scenarios permit illustrating range of conclusions on diversification between mortality and longevity portfolios one might end up with, built on assumed dependency structure between causes of death.

We use U.S. male mortality data by cause of death and by educational attainment level. By retaining only most-educated (Bachelors and Graduates) we propose an insured-population proxy for application.

Independence framework

When modelling mortality by cause of death, one must introduce notions of crude and net mortality intensity. While crude mortality is directly obtained from death count and exposure data, net mortality estimates probability of dying in the hypothetical world where you cannot die of causes other than the cause of interest. This is a theoretical construct and depending on hypothesis on cause dependency structure, relation between net and crude mortality will be different.

The first introduced cause-specific modelling approach assumes independence between causes of death, meaning that crude and net mortality intensities are equal. Each cause of death i is modelled separately by classical Lee and Carter (1992) mortality model:

$$\log(\mu_{x,t,i}) = \alpha_{x,i} + \beta_{x,i}\kappa_{t,i} + \epsilon_{x,t,i}, \tag{4}$$

where $\alpha_{x,i}$ is some average cause i mortality at age x, $\kappa_{t,i}$ describes overall cause i mortality evolution through time, $\beta_{x,i}$ represents sensitivity to $\kappa_{t,i}$ dynamics by age and $\epsilon_{x,t,i}$ describes the error term. Mortality intensity projections for cause i are obtained by extrapolating mortality trend $\kappa_{t,i}$ using Random Walk with Drift as a stochastic process:

$$\kappa_{t+1,i} = \kappa_{t,i} + \delta_i + \eta_{t,i}, \tag{5}$$

where δ_i is the drift parameter for cause i and $\eta_{t,i} \sim N(0, \sigma^2)$.

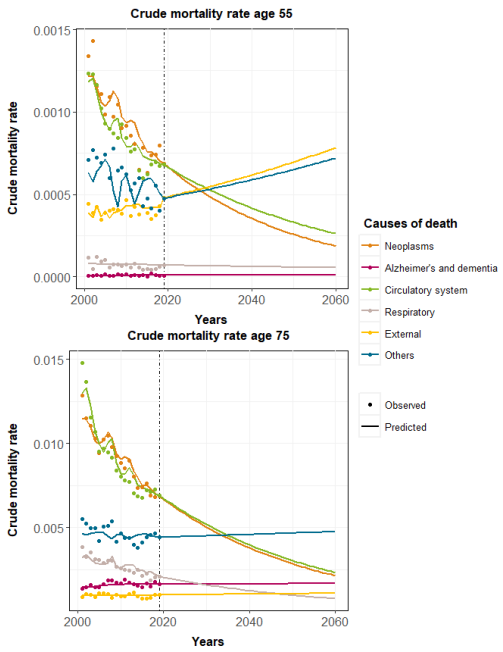

Figure 10: Cause-specific mortality

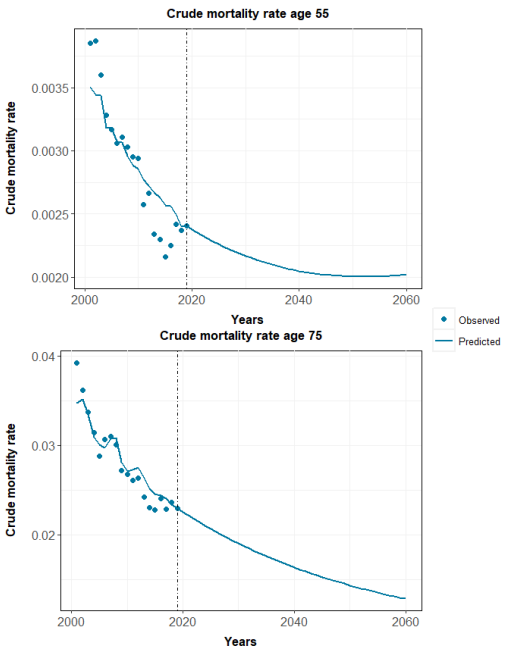

Figure 11: Aggregate mortality

Figures 10 and 11 show projected mortality stagnating in long-term for age 55 and decline in mortality rates over whole projection period for age 75.

Copula-based dependency framework

Archimedian copula framework is used to introduce dependency structure between causes of death, an approached proposed by Li and Lu (2019). By specifying the survivor copula **S** as Archimedian survivor copula, Li and Lu (2019) showed that the net cause-specific intensities, through net survival functions, can be uniquely estimated from the data by using crude-specific intensities and the copula:

$$S_i(t) = \psi \left[- \int_0^t \frac{exp\left(-\int_0^s \sum_{j=1}^m \mu_j(u)du\right)}{\psi' \circ \psi^{-1} \circ \exp\left(-\int_0^s \sum_{j=1}^m \mu_j(u)du\right)} \mu_i(s)ds \right], \ \forall i = 1, \ldots, m. \quad (6)$$

We use Clayton's copula with generator function $\psi(t) = (1+t)^{-1/\theta}$ and choose dependency parameters $\theta = 1$ and $\theta = 4$. Parameter θ values captures dependency strength, and the greater the θ value, the stronger dependence is implied between causes.

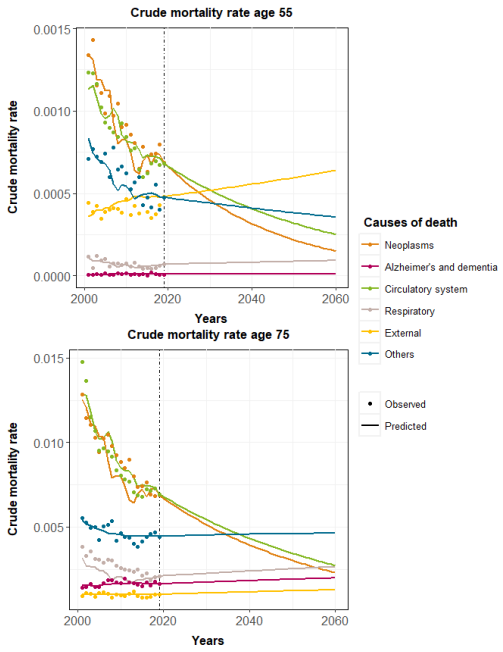

Figure 12: Net mortality, $\theta = 1$

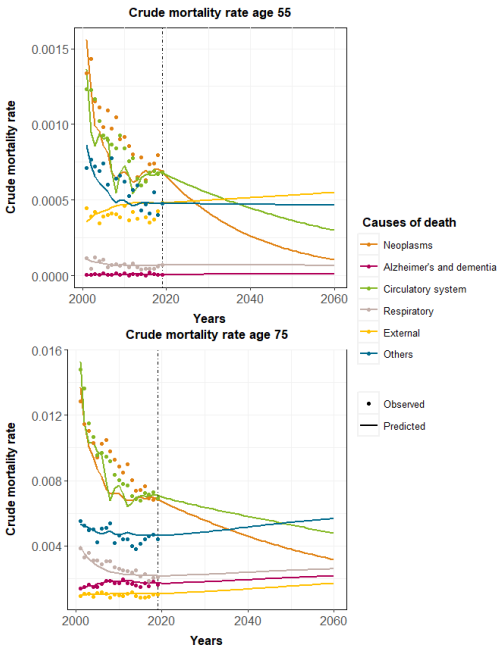

Figure 13: Net mortality, $\theta = 4$

We observe that cause-specific mortality projections vary depending on the dependency between causes assumed, see Figures 10, 12 and 13. The resulting all-cause mortality also differs - independence yields mortality rate stagnation in the second half of the projection period for age 55, while both $\theta = 1$ and $\theta = 4$ provides decreasing mortality projections. Conversely, decline in mortality rates for age 75 is projected higher under independence assumption, see Figures 11, 14 and 15.

The obtained mortality projections yield lower life expectancy than under the independence framework as we allow for transfer between causes, and slowdown in mortality improvements is stronger for higher θ value. Table 4 depicts residuals periodic

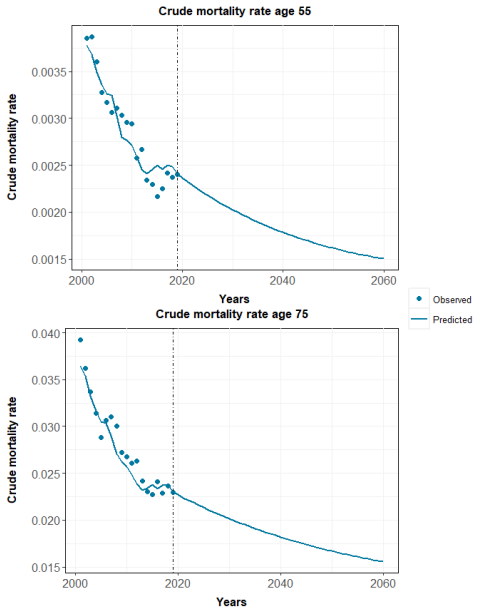

Figure 14: Aggregate mortality, $\theta = 1$

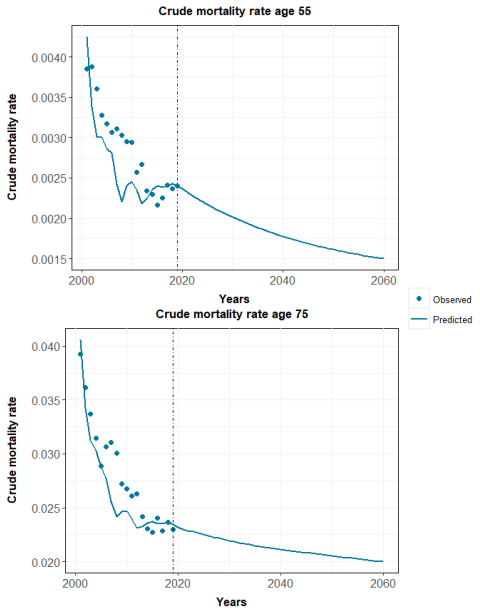

Figure 15: Aggregate mortality, $\theta = 4$

life expectancy projected at ages 55 and 75 under different dependence assumptions. We observe important variations in projected gains in life expectancy, for both ages 55 and 75. By the end of projection period the difference is more than 3 years between independence and strong dependence assumptions.

Residual life expectancy (years)				
Age 55	Year 2001	Year 2019	Year 2040	Year 2060
$\theta = 0$	26	29	30.9	32.1
$\theta = 1$	26	29	30.1	30.7
$\theta = 4$	26	29	29.3	29

Age 75	Year 2001	Year 2019	Year 2040	Year 2060
$\theta = 0$	10.1	11.8	13	13.9
$\theta = 1$	10.1	11.8	12.3	12.4
$\theta = 4$	10.1	11.8	11.6	10.8

Table 4: Residual periodic life expectancy

Footprint scenarios

Modelling mortality by cause of death permits integrating cause-specific scenarios on one or several causes of interest. We propose two scenarios on Alzheimer's and dementia cause future mortality and we assess impacts on model points representing mortality and longevity business lines and their diversification. The two proposed scenarios are results of combined progress in prevention measures and innovations in risk detection and treatments.

In the first scenario, mortality due to Alzheimer's and dementia is reduced by 66% of the central scenario projection over the next 15 years. After this 15-year horizon of improvement, the age specific probability of death due to Alzheimer's and dementia diseases is assumed to remain at 33% of its pre-scenario projection.

The second scenario is more extreme in both short timeline for the medical breakthrough to be put in place and the significant impact the scenario proposed. We assume an elimination of mortality and loss of autonomy from Alzheimer's and dementia diseases within the next 5 years of the projection.

Both scenarios are very positive views of possible future outcome with respect to health impact and the second one particularly so.

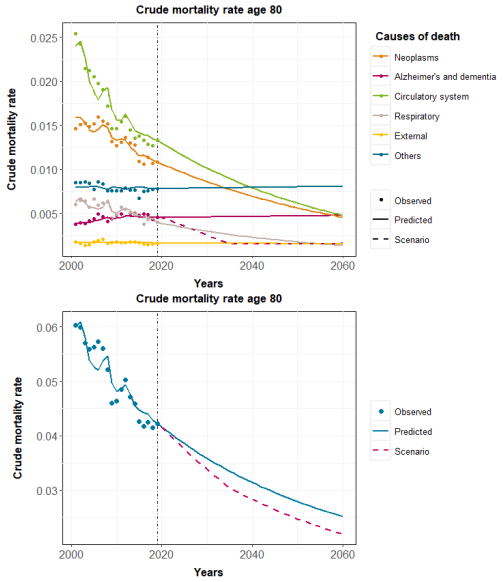

Figure 16: Scenario 1, independence

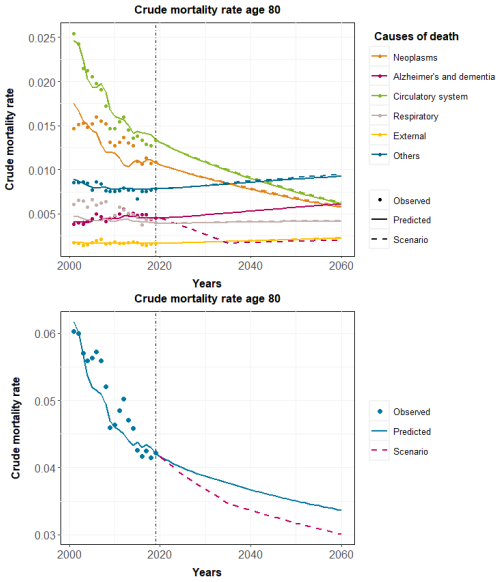

Figure 17: Scenario 1, $\theta = 1$

Figures 16, 17 and 18 illustrate Scenario 1 on cause-specific and aggregate mortality. We observe how Alzheimer's and dementia mortality rate decrease has no impact on other causes under the independence assumption, while accounting for dependence between causes results in transfer of deaths to other causes, and the transfer is more important with higher parameter θ value.

The lives "saved" from dying due to Alzheimer's and dementia cause being transferred to other causes result in important variations in life expectancy gains for different dependency hypotheses. Table 5 depicts gains in life expectancy due to 66% reduction in Alzheimer's and dementia mortality and we observe important variations in the scenario gains, particularly between cases with independent causes and assuming strong dependency ($\theta = 4$).

The second scenario is more extreme and consequently gains in life expectancy are more important, but the overall conclusion is the same with most gains in life ex-

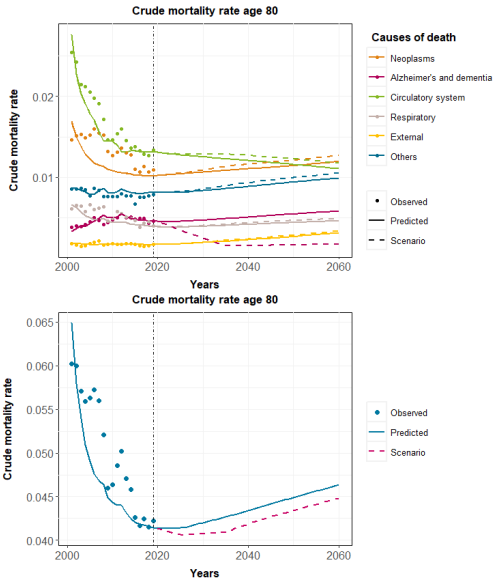

Figure 18: Scenario 1, $\theta = 4$

Residual life expectancy					
Age	Dependency	Case	Year 2019	Year 2040	Year 2060
55	$\theta = 0$	Central scenario	29	30.9	32.1
		Scenario 1	29	31.4	32.7
		Δ (months)	-	5.9	7.3
	$\theta = 1$	Central scenario	29	30.1	30.7
		Scenario 1	29	30.6	31.2
		Δ (months)	-	5	6.4
	$\theta = 4$	Central scenario	29	29.3	29
		Scenario 1	29	29.5	29.1
		Δ (months)	-	2.2	1.7
75	$\theta = 0$	Central scenario	11.8	13	13.9
		Scenario 1	11.8	13.5	14.5
		Δ (months)	-	5.7	7
	$\theta = 1$	Central scenario	11.8	12.3	12.5
		Scenario 1	11.8	12.7	13
		Δ (months)	-	4.7	6
	$\theta = 4$	Central scenario	11.8	11.6	10.9
		Scenario 1	11.8	11.7	11
		Δ (months)	-	1.7	1.1

Table 5: Scenario 1 residual periodic life expectancy

pectancy offered by scenario with independence and least gains by $\theta = 4$.

To illustrate how the previous results could impact insurance company's business, we use one representative model point for each Mortality and Longevity portfolios with equal PV claims in central scenario and assess gains and losses for both Scenario 1 and Scenario 2, see Table 6.

Assuming independence between causes results in biggest variation of claims for Mortality and Longevity business lines, which was to be expected as the life expectancy gains were the highest under the independence framework, while assuming high dependency between causes yields in moderate fluctuations. Both scenarios under independence assumption produce drop in total claims, $\theta = 1$ yields to small increase in claims for Scenario 1 and drop in Scenario 2, and lastly, considering high dependency generates higher total claim amounts for both of the scenarios.

Independent causes			
Business	Central scenario	Δ Scenario 1	Δ Scenario 2
(1) Mortality claims	2000	-53	-83
(2) Longevity claims	2000	+33	+82
Total (1)-(2)	0	-20	-1

Clayton's copula, $\theta = 1$			
Business	Central scenario	Δ Scenario 1	Δ Scenario 2
(1) Mortality claims	2000	-37	-57
(2) Longevity claims	2000	+28	+68
Total (1)-(2)	0	-9	+9

Clayton's copula, $\theta = 4$			
Business	Central scenario	Δ Scenario 1	Δ Scenario 2
(1) Mortality claims	2000	-10	-14
(2) Longevity claims	2000	+14	+30
Total (1)-(2)	0	+4	+16

Table 6: Impact on diversification for different θ values

We observe how not taking into account dependency structure between causes when modelling scenarios could yield in hypothetical gains. On the other hand, working in competing risk framework would result in more moderate gains from diversification from mortality and longevity business lines. Assuming causes being independent both scenarios results in lower claims than in central scenario, showing high diversification between Mortality and Longevity business lines, especially for Scenario 1. Moving to low dependency case ($\theta = 1$), the reduction in total claims is twice smaller and we observe and increase in claims for Scenario 2 instead of small gains (+9 vs -1). Hypothesis of strong dependency between causes has even stronger impact and for both scenarios where we observe increase in claims.

Financial scenario impacts are notably different depending on the dependency assumption used when modelling mortality at granular cause-specific level. Conclusion variety concerning diversification effects and business resilience illustrates importance of working under competing risk framework.

Contents

Introduction

Mortality prediction is crucial for Life insurance companies as the predicted future mortality is used for both policy pricing - direct impact on profitability, and risk management - used to set business acceptance capacity.

Mortality and longevity are directly opposite risks. Mortality risk refers to financial burden that an insurance company can encounter if many of their life insurance policy holders would die earlier than expected. On the other hand, longevity risk corresponds to financial difficulty the company would face due to unexpected decrease in mortality, resulting in annuity payments for longer period than expected.

Insurance companies use future mortality projections to estimate probable cost of these risks related to human lifespan. Currently most of the life insurers set future mortality assumptions by modelling all-cause mortality . This approach can be transformed by approaching mortality in more granular - cause-specific level that is put in place in this thesis.

Advantage of using cause-specific mortality modelling is adopting the framework to derive cause-specific scenarios on one on several causes of interest to produce a "what-if" type-of analysis for future mortality development. Insurance companies are requested to deliver this type of scenarios by both regulators and internal stakeholders. Regulators demand stress test scenarios to validate internal model and to identify probable crisis situations that would threaten the viability of the insurance company. Senior management is also particularly interested in probable scenario analysis which permits to assess business resilience to shocks, support business acceptance and risk appetite decisions and evaluate portfolio diversification impact.

The first chapter of this thesis presents background and motivation for the study. Secondly, we provide a brief overview on the history and introduce concepts used in the mortality modelling. Different manners to approach mortality projecting are introduced in the second chapter, starting by all-cause mortality models and then moving to a more granular cause-specific mortality modelling concept. The third chapter presents the data used. Afterwards, cause-specific mortality modelling approaches are illustrated, first assuming independence and then different levels of dependence between causes using Archimedian survival copulas. The final chapter introduces footprint scenarios on Alzheimer's and dementia causes to perform a "what-if" analysis on longevity and mortality business blocks and to assess diversification impact between the two business lines, revealing important variations in results for different dependency positions.

Methods are applied on U.S. insured male population proxy data. The outcome could be different using different country data or observation period.

SCOR uses mortality modelling by cause of death approach developed in this thesis to derive adverse footprint scenarios for both internal model validation and in the ORSA process to test the resilience of the current and expected future solvency position.

Chapter 1

Context of the study

This chapter presents background and motivation for this study. The suggested context will provide useful grounds for the modelling approach proposed by this thesis.

1.1 Reinsurance

Reinsurance is often simplistically referred to as insurance for the insurance companies. It is in fact a contract between an insurer and a reinsurer. Under this contract, an insurer - the cedent - transfers some of its risk to a reinsurance company with the latter assuming the agreed part of the insurance contracts issued by the insurance company. Reinsurance is essential tool used by the insurance companies to manage risks and the amount of capital they must hold to support those risks. Insurers may use reinsurance to achieve an optimal targeted risk profile. Reinsurance company might judge that the total risk ceded by insurance company surpasses its capacity, and in this case the reinsurance could further share the risk of the cedent with other reinsurance companies through retrocession.

We can discern two types of reinsurance: proportional (quota-share) and non proportional (excess-of-loss). In a proportional reinsurance treaties the reinsurer and the cedent company share both the premium from the policyholder and the potential losses. Consequently, they are mainly used to cover risks that occur frequently. In a non-proportional agreement, the insurance company retains a certain amount of liability for losses (known as the ceding company's retention) and pays a fee to the reinsurer for coverage above that limit amount, generally subject to a fixed upper limit. Non-proportional agreements may apply to individual policies, to an event such as a hurricane that affects many policyholders or to the insurer's aggregate losses above a certain amount, per year or per policy.

1.2 SCOR

SCOR, a leading global Tier 1 reinsurer, offers its clients a diversified and innovative range of reinsurance and insurance solutions and services to control and manage risk. Applying "The Art & Science of Risk", SCOR uses its industry-recognized expertise and cutting-edge financial solutions to serve its clients and contribute to the welfare and resilience of society. SCOR provides its clients with value-added solutions

thanks to an underwriting policy based on profitability, effective risk management and cautious investment. In terms of financial strength, independent rating agencies place SCOR among the best-ranked reinsurance companies in the world (A+ rating from S&P and Fitch, Aa3 from Moody's and A+ from AM Best). The Group's development model is driven by three Business Units, which provide a broad range of innovative reinsurance solutions: Property & Casualty, Life & Health and Investments.

Property & Casualty's broad coverage of physical, environmental and financial risks includes:

1. damage to physical assets caused by fire, natural catastrophes and other perils, as well as inherent defects in construction
2. customized risk transfer solutions for crops, livestock, forest, greenhouse and aquaculture resources
3. credit, surety and political risks, with more than 40 years' experience to draw on.

Life & Health safeguards human, social and relationship capital. With our unique expertise in biometric and health-related risks, we offer a variety of health solutions for:

1. critical illnesses such as cancer, heart attacks and stroke
2. long-term care required by conditions such as Alzheimer's disease
3. longevity risks, which are important to the security of pension systems.

Investments, together with SCOR Investment Partners – its asset management company – contributes to economic growth by helping to increase many forms of capital.

1.3 Risks related to human life

This section is inspired by the doctoral thesis by Piveteau (2021).

There are numerous risks related to the duration of human life: longevity risk, mortality risk, healthcare and long term care risks and demand of insuring those risks is at all-time high. Indeed, the Covid-19 pandemics has emphasized the need for mortality protection, rising health expenditure weight over government budgets and public pension replacement rates are constantly declining.

Longevity risk refers to the possibility that life expectancy and actual survival rates will exceed expectations or pricing assumptions and will result in financial losses due higher-than-expected cashflows required by the pension funds. Inverse from the

longevity risk, mortality risk represents financial trouble that a household meets upon premature death of one of its providing members and the insurance company faces risk that too many life insurance policy holders will die sooner than expected.

Healthcare spending and long term care risks are less directly related to human lifespan, but the incidence rates increase directly with age. Healthcare spending risk is defined as situation where a household cannot have access to needed healthcare services without experiencing undue financial hardship. And finally, the long term care risk covers financial difficulties due to autonomy loss resulting, for example, at the person integrating a nursing home.

Mortality, longevity, healthcare and long term care risk coverage has prospered in the 20th century and became almost systematic in the developed countries.

Participation of the public sector in the insurance coverage varies from one country to another leaving more or less participation to the private sector at the same time. While in France the longevity risk is covered mainly by the pension system governed by the state, in the United Kingdom pension system offers solely minimal public pension highly supplemented by mandatory occupational private pensions. Due to this reason the longevity insurance market in the UK is much more developed than in France.

The before mentioned risk coverage generalisation is supplemented by the phenomenon of the society-aging which present in all the developed countries.

Figure 19 illustrates the life expectancy increase observed during the last 60 years.

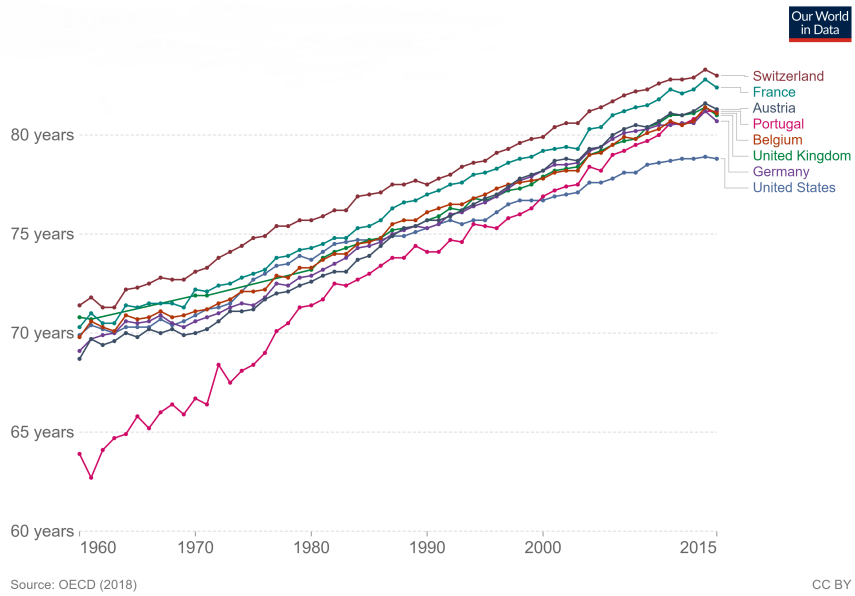

Figure 19: Life expectancy at birth 1960-2018, both sexes combined (source: OECD (2018))

Life expectancy in the developed countries now surpass or is close to 80 years, having gained 10 years approx. of life expectancy since 1960s', meaning that significant amounts are needed to counterbalance the aging society phenomenon together with the human lifespan-related risk coverage expansion.

1.3.1 Mortality

Mortality risk refers to a financial burden that an insurance company can encounter if too many of their life insurance policy holders will die earlier than what has been expected by the insurance company.

Mortality risk can be decomposed into three components as illustrated in the Figure 20:

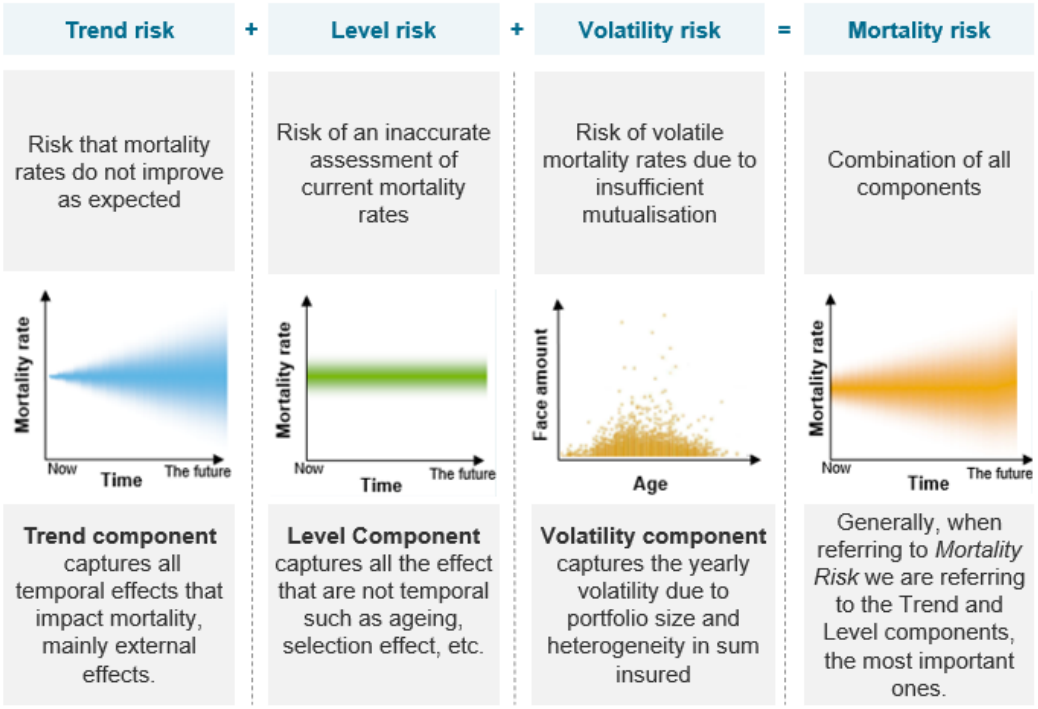

Figure 20: Decomposition of mortality risk (source: Internal)

- Level assumptions

  At SCOR, internal mortality data, also called mortality experience, allow to estimate the mortality rate - asses the mortality level. Internal data allow to capture the effect of age, gender, etc. However, there are areas where data is scarce, such as high ages, see Figure 21 .

- Trend assumptions

  In order to perform mortality forecasts, or to estimate the future mortality improvements, one needs long series of historical mortality data. Thus, to build

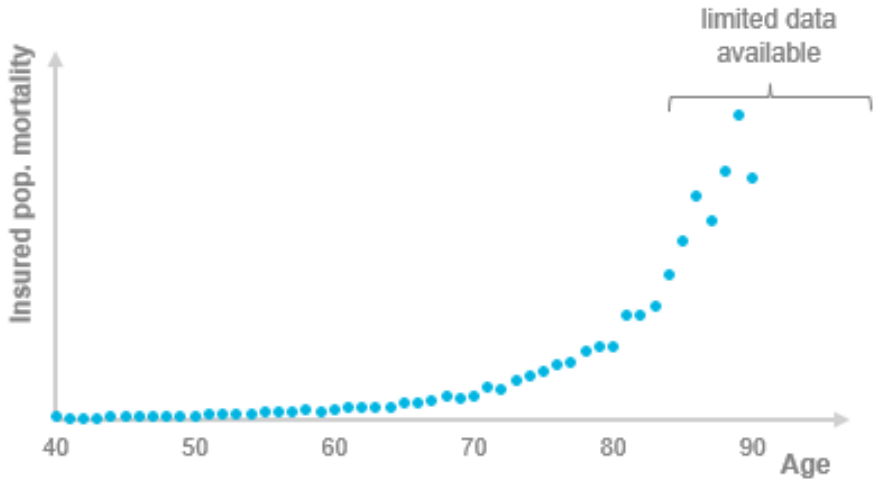

Figure 21: Mortality level (source: Internal)

mortality trend assumptions, we supplement internal mortality data with external data such as national population or educational mortality data, as a proxy of the insured population.

Different mortality datasets used to construct the mortality trend assumptions are illustrated in Figure 22:

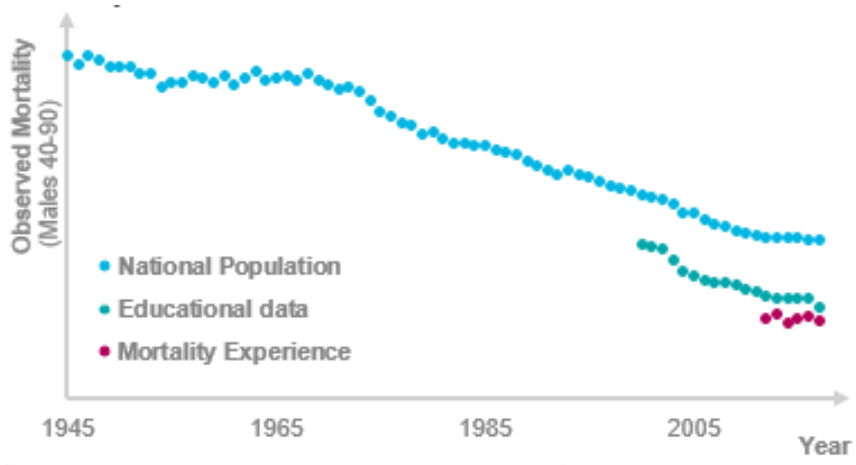

Figure 22: Mortality trend (source: Internal)

Mortality risk is one of SCOR's most significant risks. Majority of the life division's standalone economic capital covers mortality risks (mortality shock, mortality level and mortality trend) with the risk being concentrated in the United States.

While the life expectancy has been observed to increase over past decades, an example of recent Covid-19 pandemic crisis has revealed how a single event can revert the trend and put in peril life insurance companies with important life insurance policy portfolio. According to the CDC Press release (2022), "Life expectancy at birth in

the United States declined nearly a year from 2020 to 2021. That decline – 77.0 to 76.1 years – took U.S. life expectancy at birth to its lowest level since 1996. The 0.9 year drop in life expectancy in 2021, along with a 1.8 year drop in 2020, was the biggest two-year decline in life expectancy since 1921-1923." Fluctuations in life expectancy are not common and typically the life expectancy declines are quickly followed by bounce backs. In contrast to these short-term fluctuations, however, the COVID-19 pandemic induced global and severe mortality shocks in 2020 which are, to some extent, still ongoing.

1.3.2 Longevity

Longevity risk is one of the largest risks that insurance and reinsurance companies, pension plans, and the governments are exposed to. Longevity risk from the perspective of an insurance company or defined benefit plan is the burden that a company has to deal with due to unexpected decreases in mortality. This is the opposite of mortality risk, which is exposure to increases in mortality. Longevity risk has developed as experience emerges about the consistent increase in life expectancy, combined with the long term nature of many guarantees that insurance companies have written.

As already discussed, most of the developed countries have been seeing the life expectancy of their population increase due to factors that include better diet, increased access to adequate amounts of food and basic healthcare, and advances in medicine. With concurrent declines in fertility rates, many countries are witnessing a demographic shift towards a 'graying' population, where the number of people in retirement is rapidly catching up with the number of people in the workforce, see Figure 23. All across the globe this is putting strains on existing retirement systems, and leading to a shift in the risk from employers and plan sponsors to individuals.

On the other hand, some actuaries and academics argue (see National Academies of Sciences, Engineering, and Medicine (2021)) that we would soon observe the trend invert for the life expectancy increase which has been observed in the past decades. The dramatic increase of obesity rates over the last decades could be the biggest issue leading to this conclusion. The wide panel of views on future mortality trends indicate that there is a great deal of uncertainty regarding mortality improvement, leading to an ever greater need for action by the industry to understand the fundamental drivers of longevity risk.

The aging society phenomenon and expansion of insurance coverage imply considerable capital to cover the longevity risk.

Substantial assets have been accumulated in retirement savings plans to finance future pension benefits around the world. Pension assets exceeded USD 56 trillion worldwide at the end of 2020 (see OECD (2021)), a 11% increase compared to end-

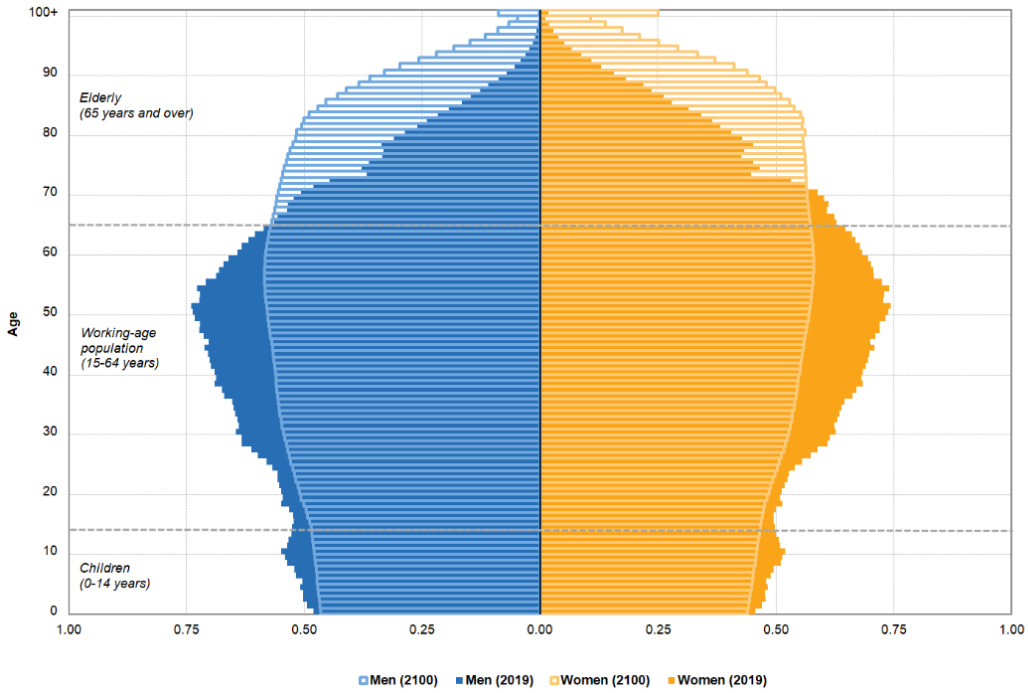

Figure 23: Population pyramids, EU-27, 2019 and 2100 (% of total population), source: Eurostat

2019 when they amounted to USD 50.6 trillion. Seven out of the 38 OECD countries held more than 90% of the total pension assets within the OECD area. The United States has the largest pension market within the OECD, with assets worth USD 35.5 trillion, representing 65.6% of the OECD area total (Figure 24). The United Kingdom recorded the second largest amount (USD 3.6 trillion, i.e. 6.6% of OECD area pension assets), followed by Canada (USD 3.1 trillion, 5.7% of OECD area pension assets), the Netherlands (USD 2.1 trillion, 3.9% of OECD area pension assets), Australia (USD 1.8 trillion, 3.3% of OECD area pension assets), Japan (USD 1.6 trillion, 2.9% of OECD area pension assets) and Switzerland (USD 1.3 trillion, 2.5% of OECD area pension assets). The 31 other OECD countries jointly hold the remaining 9.5% of pension assets in the OECD area.

Pension assets have increased faster than GDP over the last decade, highlighting the growing importance of retirement savings worldwide.

SCOR is exposed to Longevity risk which is concentrated in the UK, corresponding to dozens of billions pounds of underlying Present Value of Pension expected to be paid to annuitants. The longevity risk represents the danger that more is paid out because of increasing life expectancy. SCOR covers only the biometric risk, excluding asset risk associated with delivering pension. This type of insurance contract is called longevity swap.

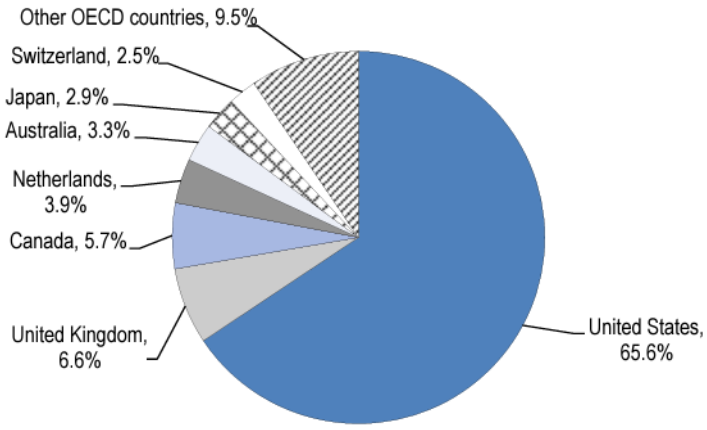

Figure 24: Geographical distribution of pension assets in the OECD area, 2020

1.4 Solvency II

The solvency of insurance and reinsurance companies is defined as the ability to meet their short, medium and long-term commitments to their clients. The solvency depends on the commitment size, guarantees and protection that are offered to the insured, and the resources put in place to meet these commitments.

Solvency II is the prudential regime in the European Union for insurance and reinsurance undertakings that has entered into force in January 2016. Solvency II sets out requirements applicable to insurance and reinsurance companies in the EU with the aim to ensure the adequate protection of policyholders and beneficiaries. Solvency II has a risk-based approach that enables to assess the "overall solvency" of insurance and reinsurance undertakings through quantitative and qualitative measures. The Solvency II regulatory framework is built on a three-pillar structure:

- Pillar I sets the quantitative requirements i.e. the assets and liabilities valuation and capital requirements. The Pillar I concentrates on the economic balance sheet, the resulting actual own funds, and the required risk-based own funds. Required own funds are determined based on a Solvency II Standard Formula, or by means of an internal model.

- Pillar II sets the qualitative requirements, including governance and risk management of the undertakings and the Own Risk and Solvency Assessment (ORSA).

- Pillar III sets the supervisory reporting and public disclosure.

The three pillars form a coherent approach that allow to understand and to manage risks across the sector.

SCOR has chosen to develop internal model to assess individual risk profile and characteristics of the company's business rather than use the standard formula approach which is proposed by default by the regulator. The internal model is subject to regulatory approval and is required to meet several requirements:

- Use test: requirement to fit the business and reflect the risk profile, and to be widely used in company's decision making and governance process.
- Statistical quality: minimum statistical standards must be met together with sound methodology justification for assumptions like management actions, risk diversification, expert judgement etc.
- Calibration standards: ensure that the SCR calibrated corresponds to 1-in-200 scenario.
- Profit and Loss attribution: review of sources and causes of P&L for each major business unit.
- Validation standards cover on-going model validation process, validation policy and sensitivity, stress and scenario testing.
- Documentation standards
- External models and data: consistency with the internal model standards and suitability to the undertaking's risk profile.

1.5 Footprint scenarios

Scenarios, characterized by deterministic events, also called *Footprint scenarios*, can be based on historical experiences or imagined situations. These scenarios permit companies performing a "what-if" type of analysis, i.e. estimating impact from events like hypothetical terrorist attack in Madison Square Garden or Great Kanto earthquake repeating itself. Scenarios can touch one or several lines of business and all the exposure-to-risk should be assessed. Metrics like crude or net losses of the company, changes in solvability ratio or rating can be used to evaluate footprint scenario impact. Footprint scenarios are easy to understand and to communicate and they contribute to risk culture development in the company. They are also complementary to probabilistic models and permit taking a step back from the models and enable assessing risks that are partly or not sufficiently modelled.

Article 242 of the delegated regulation, published in the *Official Journal of the European Union*, provides in particular that the statistical process for validating the internal model must include a reverse stress test in order to identify the most probable crisis situations which could threaten the viability of the insurance company. In addition, the ACPR notice on the internal models which specifies this provision pro-

vides that the insurance company ensures that the stress test and scenario analysis used cover the relevant risks and are monitored over time. SCOR includes scenario analyses into its validation process for the calibration of long-term mortality and longevity - most import life underwriting risks of the company.

A working group of experts established from the risk management, knowledge and business acceptance teams consider plausible but severe scenarios that would impact future mortality and determine their impact on SCOR's portfolios. The scenarios were chosen to be both severe but also plausible and also to test whether the modelled negative correlation between mortality and longevity could fail in certain extreme scenarios.

Scenario about future trends of Alzheimer's and dementia diseases has been introduced based on several criteria. First, the scenario needs to extreme but plausible - a scenario around curing Alzheimer's disease would be in line with probable medical advances in the near future. Also, the goal of the scenario is to test diversification impact between mortality and longevity business lines. The scenario is likely to have important cumulative impact from both longevity and mortality books and not to totally offset one another: Alzheimer's disease affects mainly very high ages and as the longevity insured population is older than the mortality's one. An important (negative) impact could be expected on longevity business with some but not total offset from mortality due to differences in age distribution, see Figure 25.

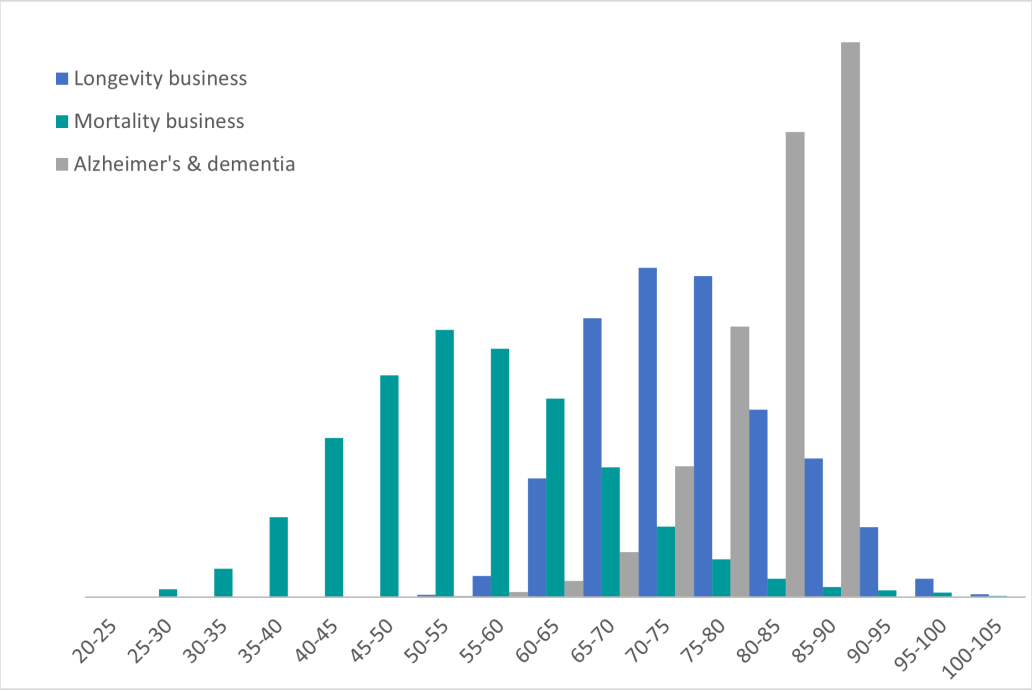

Figure 25: Age distribution in Mortality and Longevity business and US male age-at-death distribution for Alzheimer's & dementia disease

1.6 Alzheimer's and dementia

Alzheimer's disease and dementia are major public health challenges in many countries of the world, particularly in Western Europe and North America where aging population is a major issue. Population aging became a concern in countries like Latin America or China putting Alzheimer's and dementia diseases in the spotlight recently.

It is estimated that around fifty million people have dementia worldwide. About two-thirds of them have Alzheimer's, according the report by Alzheimer's International. Research on Alzheimer's and dementia has been gaining in support substantially - work on new screening, diagnostic and therapeutic applications have surged recently supported by large private and public funding. Many experts are optimistic about breakthroughs in Alzheimer's and dementia in the short to mid-term future, see Alzheimer's Association (2021b), Shekelle et al. (2020).

As discussed in Buffet et al. (2022), dementia is in fact a general term characterizing memory loss and other mental abilities extreme enough to interfere with daily life activities. It covers a wide range of specific medical conditions caused by physical changes in the brain. Alzheimer's is the most common type of dementia, but there are many others such as vascular dementia, Lewy body dementia and Huntington's disease, etc.

Dementia impacts more than 50 million people and kills about 1.5 million annually worldwide, according to Alzheimer's Association (2021b), making it the second leading cause of death in high-income countries in 2019, overtaking stroke, see World Health Organization (2020). In the U.S., more than 6 million Americans are living with Alzheimer's disease. By 2050, this number is projected to rise to nearly 13 million, according to Alzheimer's Association (2021c). In 2021, Alzheimer's and other dementias will have cost the U.S. $355 Billion. These costs are projected to balloon, hitting an estimated $1.1 Trillion by 2050.

This thesis covers an in-depth footprint scenario analysis, developed using sophisticated modeling and implementation of a shock on Alzheimer's and dementia mortality, and assesses its impact on life expectancy, portfolio diversification gains in multiple scenarios.

1.6.1 Types of dementia

Different types of dementia are associated with particular types of brain cell damage in specific regions of the brain. Figure 26 displays the most common types of dementia and their prevalence.

60-80% of dementia cases can be classified as Alzheimer's disease. In this disease,

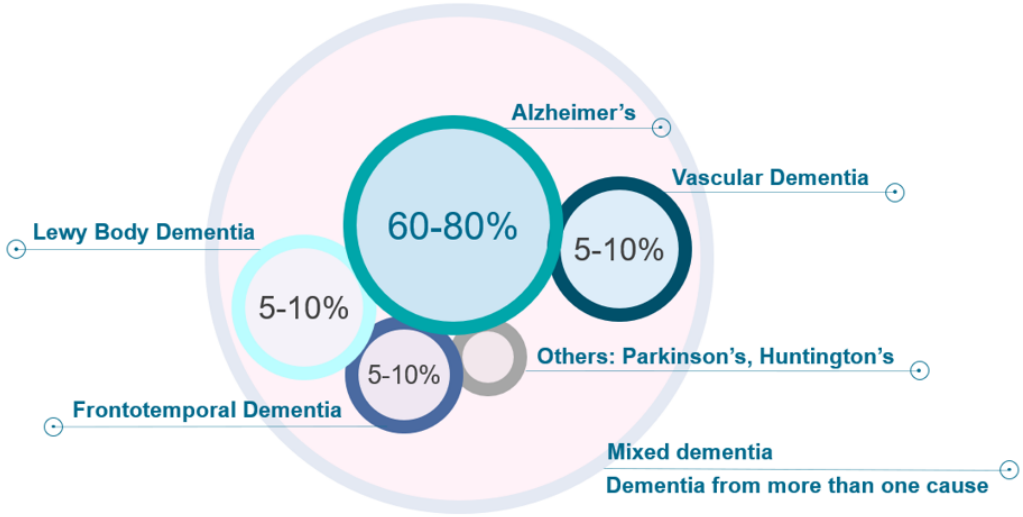

Figure 26: Types of dementia (source: Alzheimer's Association (2021a))

brain cells have trouble staying healthy and communicating due to high levels of certain proteins inside and outside brain cells.

5-10% of all people with dementia have vascular dementia, which is the second most common type of dementia. This disease develops because of microscopic bleeding and blood vessel blockage depriving various brain regions of oxygen and nutrients. Vascular dementia is common as part of mixed dementia.

Another 5 to 10% of all dementia cases fall under Lewy body dementia type. This type of dementia is caused by abnormal microscopic deposits that damage brain cells over time and lead to a decline in thinking, reasoning and independent function.

And finally, mixed dementia is diagnosed to someone simultaneously experiencing multiple types of dementia. For example, in the study carried on by James et al. (2016) involving long-term cognitive assessments followed by brain autopsy, 94% of participants who were diagnosed with dementia were diagnosed with Alzheimer's.

Coexisting pathologies were observed in 54% of autopsy cases of those diagnosed with Alzheimer's, with evidence of a vascular disease being the most common coexisting abnormality. Lewy body dementia was the second most common brain change observed.

1.6.2 Potential breakthrough and impact on the life insurance

As reviewed in Buffet et al. (2022), increasing support has been brought to clinical and pre-clinical research on Alzheimer's and dementia diseases. Following a $289 million increase in March 2022, and added to current spending, the National Institutes of Health spending is expected to total $3.5 billion on research into Alzheimer's and

dementia, according to Alzheimer's Association (2022). This unprecedent funding for Alzheimer's and dementia research will participate to speeding up investment in collaborations that accelerate discovery, scientists will be able to work at a more rapid pace to advance basic disease knowledge, explore ways to reduce risk, reveal new biomarkers for early diagnosis and drug targeting, and develop potential treatments.

Short and mid-term future for Alzheimer's and dementia breakthroughs is considered as optimistic by many experts. Shekelle et al. (2020) judge 10 breakthroughs as being at least 70% likely to occur by 2037. This optimism is also reflected in the clinical pipeline for new therapies addressing disease-modifying biologics, cognitive impairment, and neuropsychiatric symptoms with symptom-reducing agents now in Phase II and III clinical trials as reviewed by Cummings et al. (2021).

In September 2022, Lecanemab, an experimental drug was shown to alter rate of memory decline and thinking in patients with early Alzheimer's disease. This achievement has been called as a "historic moment" for dementia treatment in the research community, see Eisai (2022).

Another important discovery was made in October 2022, see Swaddiwudhipong et al. (2022). Scientists have discovered that as early as nine years before patients are diagnosed for one of dementia-related diseases, it could be possible to detect signs of brain impairment in the patient. This prominent result leads to believing that in the future, screening could be proposed to individuals at risk to help select those who would benefit from treatment to decrease the risk of developing one of the conditions.

- Progress in risk detection

  Research in the accurate and widespread early detection will be one of the most attractive areas for scientists working in the Alzheimer's and dementia field as most of the dementia diagnoses happen once the patient starts experiencing memory loss. One of the most promising areas of early detection-related research is Neuroimaging. Deep learning and other artificial intelligence methods will participate in speeding up the risk identification process. Another valuable risk assessment could be proven to be tool genetic profiling. To date, around 70 genetic variants have been reported to be associated with Alzheimer's disease risk. And finally, biomarkers are believed to offer one of the most promising paths. Researchers are currently investigating whether Alzheimer's and dementia cause consistent and measurable changes in blood levels of tau, beta-amyloid or other biomarkers before symptoms appear. Furthermore, researchers are exploring whether early detectable changes can appear elsewhere in the body such as ocular changes by neuroretinal exams.

- Progress in risk reduction and prevention

  Modifiable risk factors will play an important role to prevent, delay the onset and slow down the disease progression as result of interventions that seek to target and modify these risk factors. It is believed that over one third of the dementia cases could be prevented by addressing the risk factors in the next 25 years. More and more evidence confirm nine previously identified risk factors for dementia - low education, hypertension, hearing impairment, smoking, obesity, depression, physical inactivity, diabetes, and poor social contact. Three new factors, excessive alcohol use, traumatic brain injury, and air pollution, were also recently added to the list.

All the progress in treatment as described above would decrease the rate of progression of the disease and could modify its long-term trajectory. Furthermore, future treatment would target the disease in its earliest stages, before irreversible brain damage or mental decline occurred.

As of December 2021, more than 160 agents with diverse mechanisms of action have been registered in the clinical trials for Alzheimer's disease, see U.S. National Library of Medicine (Accessed December 16, 2021). There are currently 113 clinical studies in phase 2/3 and 48 in phase 3 clinical trials. These studies are evaluating the safety and efficacy of new molecules and repurposed drugs. Experts predict that future Alzheimer's treatments could include a combination of medications, similar to how treatments for many cancers or HIV/AIDS include more than a single drug.

Alzheimer's and dementia diseases are covered in many existing insurance policies, including life insurance, long term care and critical illness. Considering the aging population, the growth of proportion of those who suffer from Alzheimer's and dementia diseases will be significant. Without even considering impact on mortality, the cost associated with the care and treatment for patients of which insurers need to share the burden can be significant.

According to the 2020 Alzheimer's Disease Facts and Figures report by the Alzheimer's Association (2020), over 19 million family members and other unpaid caregivers provided an estimated 18.6 billion hours of care to patients with Alzheimer's or other dementias. In the United States, beneficiaries of age 65 and older receive Medicare payments that are three times higher for someone with dementias than for someone without the condition, while Medicaid payments are more than 23 times as great. Total payments in 2020 for health care, long-term care and hospice services for people age 65 and older with dementia are estimated to be as much as $305 billion. Moreover, non-monetary negative impact on caregivers and family members such as mental and physical health will be considerably high.

If this negative trend could be turned around with the advancement in cure and

treatment combined with increased prevention efforts as discussed above, it would be a tremendously positive effect for both insurance industry and the society. Insurance and reinsurance companies have a massive stake in this global effort to fight against Alzheimer's and Dementia. SCOR has committed to take a proactive part with several initiatives including internal R&D initiatives such as this footprint scenario analysis and collaborative research and partnerships with external medical experts and research institutions.

Chapter 2

Mortality modelling

This second chapter will provide a brief overview on the history and introduce concepts used in the mortality modelling. Different manners to approach mortality projecting are introduced, starting by all-cause mortality models and then moving to a more granular cause-specific mortality modelling concept.

2.1 History of mortality modelling

This section gives overview of mortality modelling history as described in Piveteau (2021) doctoral thesis.

Modelling human mortality has a very long history with demographers, insurance mathematics and academicians trying to tackle the subject for several centuries already. By aiming to understand factors influencing mortality, one addresses modelling lifespan as random variable.

Oldest known attempt to describe and model mortality dates back to the 17th century when Graunt (1973) first examined mortality in London by using public statistics provided in the London Bills of Mortality. Amongst other important results, he showed that while individual lifespan was unpredictable, mortality pattern could be easier estimated by working by causes of death or groups by age or socio-economic environment. On the turn of the century, Halley (1693) proposed a method to build reliable mortality tables using birth and death statistics. He even proposed a method to perform annuity calculation based on this mortality table. Work by DeMoivre (1725) can be qualified as a breakthrough in mortality modelling as he was the first to model survival function as a continuous function with respect to age.

Gompertz (1825) published his law of mortality and proposed the notion of force of mortality μ_x and assigned it a parametric expression. Force of mortality (also referred to as mortality intensity) at age x is defined as:

$$\mu(x) = \lim_{u \to 0} \frac{\mathbb{P}(T \leq x + u | T > x)}{u}.$$

The resulting survival function $S(x) = \mathbb{P}(X > x)$ can be written as:

$$S(x) = e^{-\int_0^x \mu_s ds}.$$

After observing adult mortality, Gompertz concluded that the force of mortality seems to have exponential growth by age and so proposed the following expression:

$$\mu_x = ae^{bx}.$$

Makeham (1860) refined the model by adding a parameter modelling general mortality unrelated to age:

$$\mu_x = ae^{bx} + c.$$

The Gompertz-Makeham model is still widely used for adult mortality modelling, see Gavrilov and Gavrilova (2011) and Kirkwood (2015). This piece of work was great impetus to parametric modeling of an average lifespan. Over the years, the method gained in sophistication to reflect observation that infantile mortality seems to have different dynamics than the exponential form observed in the adult mortality. By the end of the 20th century, the mathematical model structure became excessively complex. Heligman and Pollard (1980) suggest an eight-factor model that fits every age:

$$\mu_x = A^{(x+B)^C} + De^{-E(\log x - \log F)^2} + GH^x.$$

Parameters are divided into three groups to distinctively describe infantile, adult and old-age mortality. Concerning the old age mortality, Thatcher et al. (1998) delivered extensive studies using several sources of reliable data and fitting different models on ages 80 and above. The model evaluation of study revealed the logistic model being the best mathematical model for human mortality, putting aside the widely used Gompertz and Makeham model. The logistic model assumes the force of mortality μ_x being a logistic function of age x.

All of the models discussed above model human mortality at a fixed point in time or over a very short period of time, while actually mortality is stochastic and constantly evolving. Accordingly, to account for the mortality stochasticity, the model parameters would need to be fitted periodically as by construction they are static. With the life expectancy constantly increasing since the 19th century in the developed countries, mortality estimations made solely from the observed past became obsolete. Including mortality dynamics with respect to time becomes a necessity to obtain reliable future mortality projections.

2.2 Prospective mortality modelling

A work by Swedish astronomer Hugo Gyldén in 1875 can be called first known prospective mortality study, see Cramér and Wold (1935). Gyldén used Swedish

population mortality data over 1750-1870 to fit a straight line and proposed this linear function as probable future mortality forecast. This first method proposed was of course very primitive as the proposed projection introduced unique estimation for all ages.

Another attempt to assess mortality dynamics date to end of 19th century as Cannan (1895) integrated a cohort component to England and Wales mortality description.

In 1912, another Swedish astronomer and mathematician Andres Lindstedt developed pensioner mortality table by using Swedish mortality data up to 1907 and extrapolating annual probabilities of death. This extrapolated table was widely used in Sweden and described well the experience observed in 1911-1915. When explaining his method, Lindstedt emphasized the importance for the projections to be based on past mortality change extrapolations.

Efforts to project force of mortality were rather rudimentary up to 1970s when statistical time series methods were published by Box and Jenkins (1970). Few years later, first application of the methods were carried out in demography by Saboia (1974). Time series presented with numerous advantages, one of the most important being the probabilistic framework that would permit not only quantifying future mortality, but also it's distribution, enabling confidence intervals construction.

Up until 1990s, mortality modeling was not any different than adopted for any time series: after estimating life expectancy or force of mortality on the observed historical data, model parameters were projected as usual time series. In the 90s, stochastic models started appearing simplifying estimation process and providing relevant results. Amongst these newly-appeared model the most popular is by far the Lee-Carter (LC) developed by Lee and CarterLee and Carter (1992). We will provide detailed description of the Lee-Carter model here following as the this model will be repeatedly referred to and used in this study.

Let x be age taking values from 1 to x_{max} with time spanning from 1 to T. Let $\mu_{x,t}$ note the force of mortality for age x and year t and so the Lee-Carter model is described as follows:

$$\log(\mu_{x,t}) = \alpha_x + \beta_x \kappa_t + \epsilon_{x,t}, \tag{7}$$

where α_x is some average mortality at age x, κ_t describes overall mortality evolution through time, β_x represents sensitivity to κ_t dynamics by age and $\epsilon_{x,t}$ describes the error term.

In the original version of the Lee-Carter model, the error terms are independent and identically distributed (iid) - $\epsilon_{x,t} \sim N(0, \sigma^2)$ and the κ_t dynamics are modeled using

Random Walk with Drift (RWD) as a stochastic process:

$$\kappa_{t+1} = \kappa_t + \delta + \eta_t, \tag{8}$$

where δ is the drift parameter and $\eta_t \sim N(0, \sigma^2)$.

Model is calibrated in two steps.

First, parameters α_x, β_x and κ_t are estimated by minimising least squared errors. As the authors identified in the Lee-Carter introductory article Lee and Carter (1992), without adding constraints to the parameters the model is not identifiable. In fact, let the parameter set α_x, β_x and κ_t be one that minimizes the sum of quadratic errors of the L-C model and C be a constant not equal to zero. In this case, the parameters:

$$\kappa_t^* = \kappa_t + C,$$

$$\alpha_x^* = \alpha_x - C\beta_x,$$

$$\beta_x^* = \beta_x$$

will also minimize the sum of quadratic errors. So a constraint on the parameters is required to have identifiable model. The most conventional constraint is setting $\sum_x \beta_x = 1$ and $\sum_t \kappa_t = 0$. This is only one of possible constraints to obtain identifiability of the model and further in this thesis we will see how other than this particular constraint can prove to be more of interest for specific cases. And finally, as the κ_t dynamics are modeled by Random Walk with Drift, the δ parameter estimator is given by $\hat{\delta} = \frac{\kappa_T - \kappa_1}{T}$.

The work of Lee and Carter has been widely acknowledged and cited, and was also used as a benchmark model by the U.S. Bureau of the Census for long-run forecasts of age-specific mortality rates, see Hollmann et al. (1999).

Success of the Lee-Carter model generated variants of the latter proposed over the recent years. Multiple papers published tried to improve the Lee-Carter model by adding more principal components, cohort effect or other similar statistical quantities.

A log-bilinear model by Brouhns et al. (2002) is a Lee-Carter model extension considering deaths as a random variable of Poisson distribution. Let us denote $D_{x,t}$ the number of deaths and the $E_{x,t}$ the exposure at risk of age x at year t, and so the Lee-Carter Poisson model is described as:

$$\mu_{x,t} = exp(\alpha_x + \beta_x + \kappa_t),$$

$$D_{x,t} \sim Poisson(E_{x,t} * \mu_{x,t}) = \frac{(E_{x,t} * \mu_{x,t})^{D_{x,t}} e^{-E_{x,t} * \mu_{x,t}}}{D_{x,t}!}$$

with constraints similar to the original Lee-Carter model.

Li and Lee (2005) extended the Lee-Carter model to account for dynamics in several population mortality simultaneously, while Booth et al. (2006) proposed optimal time period choice of the time-dynamics parameter κ_t and compared 5 Lee-Carter model variants for 10 different populations. Renshaw and Haberman (2006) adds a cohort effect to the Lee-Carter model by incorporating the year of birth variable, permitting to account for the commonly observed generational differences.

Other models that are not a direct version of the famous Lee-Carter model are also introduced, such as the CBD model presented by Cairns et al. (2006a).

Even though the mentioned models are rather efficient at measuring and predicting general mortality, they are not always appropriate when one seeks to take into account mortality's heterogeneity phenomenon.

2.3 Mortality by cause of death

When speaking about heterogeneity of the mortality one can have different aspects in mind. First, we can distinguish differences between mortality on a socio-economical level - variables affecting mortality are unequally distributed in the population causing mortality risk differences. Individuals tend to make different lifestyle choices, like diet, tobacco consumption, etc. that impact their lifespan. At the same time, depending on the social system and insurance cover of the person, he or she will have variable access to medical care that will also influence their health and so mortality.

In the U.S., a group of researchers, led by Stanford University economist Raj Chetty, analyzed income data for the US population from 1.4 billion tax records between 1999 and 2014, see Chetty et al. (2016). They discovered that, from 2001 to 2014, the 25% richest Americans gained about three years of longevity, while life expectancy for the poor increased only by one year, see Figure 27.

These mortality inequalities due to personal wealth and education level can be resolved by data choice, to some extent. We are interested in insured population mortality, and for that we create an insured population data proxy. We detail more in Chapter 3 the motivation for data choice that is related to mortality heterogeneity in this aspect.

At the same time, mortality heterogeneity can represent differences in causes resulting in death. If we can distinguish mortality by different causes, one can observe variations in both death probabilities by age and their dynamics. Figure 28 depicts US male national population mortality at age 65 since 1980s' by with causes of death grouped into "Accidents", "Cancer", "Cardiovascular diseases", "Respiratory diseases" and "Other". While the dynamics seem similar for "Accidents", "Respiratory" and

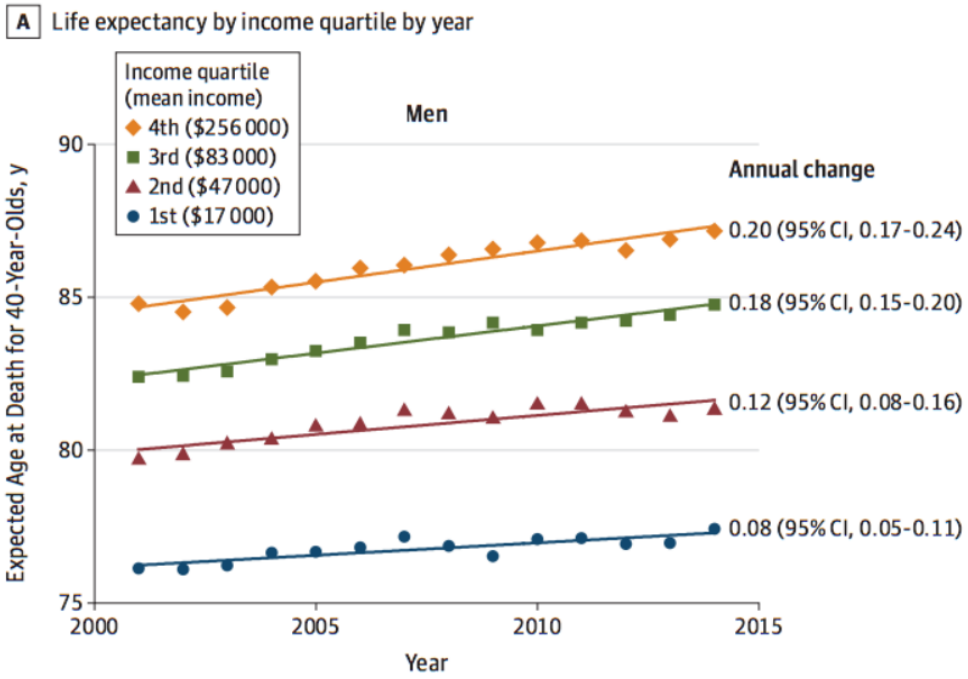

Figure 27: Changes in life expectancy by income group, 2001 to 2014 (source: Chetty et al. (2016))

"Other" causes over the observed time period, the mortality due to cardiovascular diseases has significantly different trend from the rest of the causes. Advances in prevention, decline in cigarette smoking, improved hypertension treatment etc. - the facts that contributed to this considerable drop in cardiovascular disease mortality, that are observed in majority of the developed countries, are well known. However, the trend shows signs of reversal starting around year 2010, and opinions have been surging if the cardiovascular disease mortality has not reached its plateau (see Mensah et al. (2017)). As it can be observed from Figure 29, the all-cause mortality trend resembles quite well the cardiovascular disease mortality's trajectory, as the cardiovascular mortality drop was the main reason behind the all-cause mortality decline. If we referred uniquely to all-cause mortality modelling, future mortality projections would continue the cardiovascular disease mortality declining trend into the future. Modelling mortality by cause of death permits getting into more granular level, and so expert judgement could be applied on the cardiovascular disease mortality to "correct" the future trend and take into account the slowdown of mortality decline.

2.3.1 Challenges

Cause of death is the primary pathology due to which the death occurred, recorded by the medical personnel on the death certificate. As simple as this may seem, multiple issues are related to this analysis by cause. The following challenges were recognized and in-depth covered in the 2022 Human Mortality Database conference, see Arnold

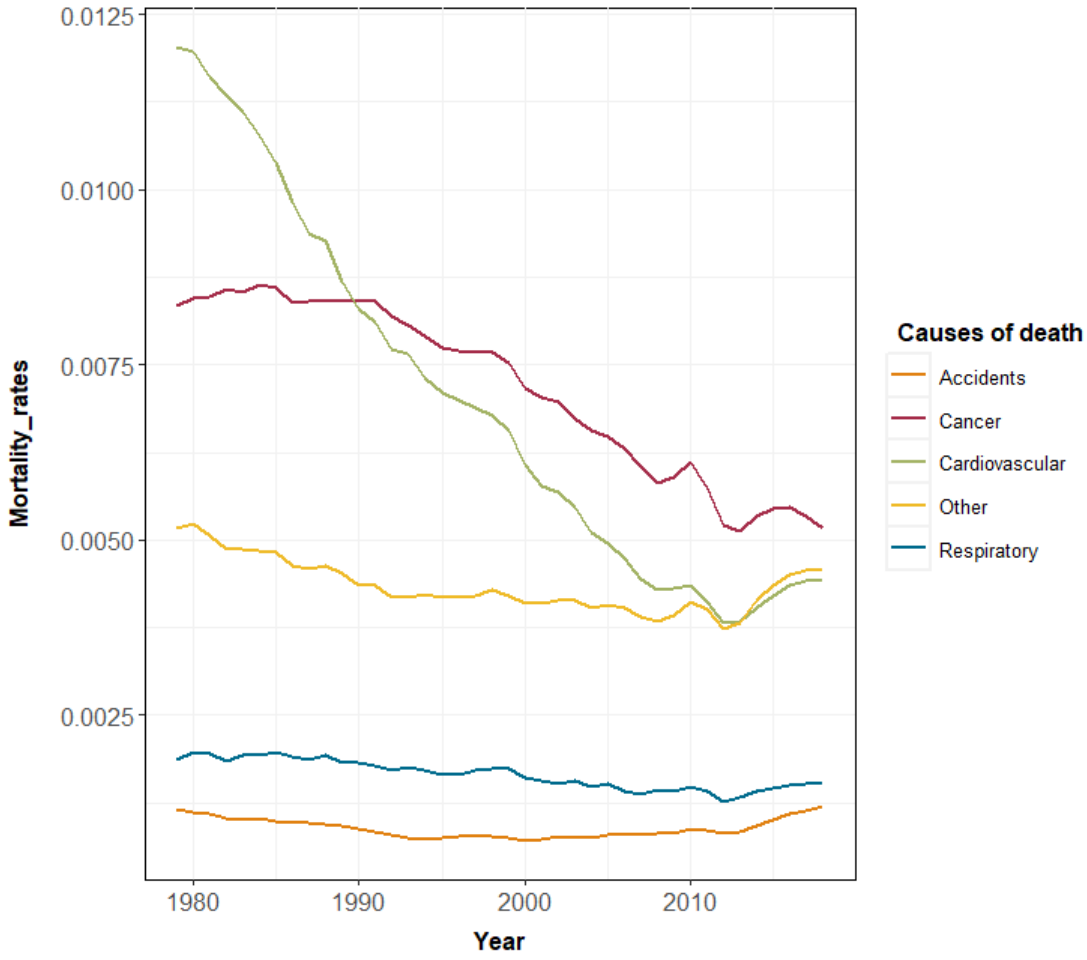

Figure 28: US male cause-specific mortality rates at age 65

(2022).

- Selection rules for the underlying cause of death might be arbitrary: medical personnel could base their decision for the underlying death cause on many human factors that we cannot control for.
- Within a cause-of-death group, we have several diseases: even if the cause-of death information might be available at a very granular level, to have sufficient data one will need to regroup small causes into larger groups.
- Changes within the classification: pathology classifications evolve through time with some pathologies emerging and some being refined in the classifications.
- Differences in interpretation of international rules, in coding practices and in training of physicians across countries and medical personnel might not consider different medical conditions throughout the years. Some conditions that were not identified earlier as causes of death and gained in importance with time would alter the cause of death distribution and would obstruct analysis

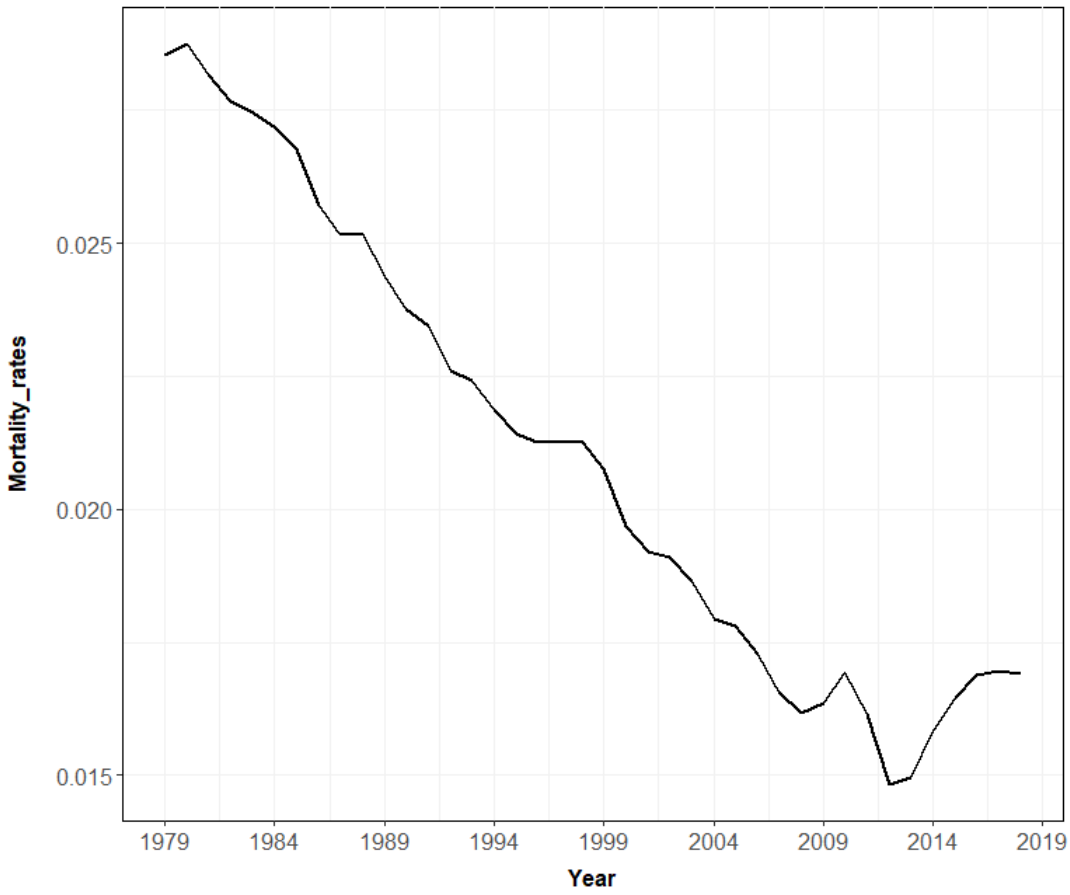

Figure 29: US male all-cause mortality rates at age 65

over long period.

- Cause-of-death reporting are less reliable at older ages where most of the deaths occur due to the frailty of elderly and often multiple conditions being present at time of death.
- Different causes impact different age-groups.
- Available time series are often very short and make plausible future mortality forecasts complicated to obtain.
- Multiple causes of death: in death certificates, several causes can be identified, with explicitly enumerating those by importance. It is difficult to clearly distinguish the principal cause of death when the person was subject to multiple conditions leading to death or if the conditions were cumulative to the death outcome.
- Miss-classifications of deaths by cause can also be an issue due to human-error presence.
- Interdependence between the causes of death (competing risks) add a layer of

complexity for modelling compared to aggregate mortality analysis. This thesis will attempt to propose specific approach to tackle this topic.

- Data limitations for disaggregated data are also important subject as time series analysis and model parameters for causes of death are often less stable than parameters in aggregate mortality model.

- Changes in diagnosis of causes over the years represent also an important issue that can prevent from having sufficiently long series of data.

- New causes of death and changes in the relative importance of known causes of death that will result to modifications and changes in classification system and again will obstruct data stability through time.

- Different causes have very different patterns and using unique model for all causes might not be appropriate.

- In every dataset there will some deaths that will be labeled as "Unknown" or "Unclassified". How should we treat deaths not ascribed to a specific cause?

- Cause-of-death mortality is linked to socio-economic groups and analyzing national population mortality might be very different than looking into, for example 10% of the wealthiest persons of the same population.

To control some of these issues and not to alter cause of death analysis, several choices can be introduced. First, we can regroup causes into sufficiently large groups that would be less or not at all sensible to treating conditions differently through time. Also, one can exclude very high ages from the analysis, and this for several reasons. Individuals who survive to high ages have very often several conditions without one explicitly main and the primary cause of death will often depend on the expert judgment of the person signing death certificate. Besides, the data become rather scarce in high ages already when dealing with all cause mortality, so splitting the numbers into more granular count by cause of death seems unreasonable and problematic. We also choose to approximate insured portfolio to account for mortality difference between national and insured populations.

2.3.2 Competing risks

The analysis of competing risks data has a long history, starting when Bernoulli (1760) published his work on smallpox impact on mortality and advantage of inoculation in 18th century Europe.

Competing risks are very important aspect to take into account when dealing with mortality by cause of death. Every human is continuously exposed to many risks of death, such as cancer, heart disease, and accidents. As death event is not repetitive

and will be in majority of cases associated to one single cause of death, these risks will compete with one another for the life of a person. If we are interested to study cardiovascular diseases as cause of death, for example, some persons will die from other causes during the time period observed. These persons will not die anymore from cardiovascular disease but also they will not be observed to the end of observation period. As a result, we call these events *competing risks*, in a sense that the probability of each competing event is somehow regulated by the other competing events. This leads to the survival process determined by multiple types of event.

2.3.3 Cause-specific mortality modelling

Concept of competing risks demands introducing additional concepts than what are required to study all-cause mortality. Hereafter we introduce notions of crude ant net mortality followed by assumptions regarding dependency structure between the survival variables.

Let us assume a homogeneous population in which every individual is exposed to m mutually exclusive causes of death and each of these individuals may succumb to a unique cause out of m available. As death may occur only once and due to a unique cause, the actual time of death of an individual can be expressed as the minimum of m cause-specific death times:

$$T = \min(T_1, \ldots, T_m).$$

The only observed cause of death is the one corresponding to the T, that is assumed to be unique.

Aggregate (all-cause) force of mortality (also referred to as mortality intensity) is the instantaneous probability of death for an individual already survived up to time t before time $(t + u)$ when u is relatively small time interval:

$$\mu(t) = \lim_{u \to 0} \frac{\mathbb{P}(T \leq t + u | T > t)}{u}.$$

For a given actual time of death T, $\Pi_i(t) := \mathbb{P}[I = i | T = t]$ describes probability that the observed cause was i.

The crude mortality intensity $\mu_i(t)$ is the instantaneous probability of death assuming that only one cause i exists for an individual who survived t years:

$$\mu_i(t) = \mu(t)\Pi_i(t) = \lim_{u \to 0} \frac{\mathbb{P}(T \leq t + u, I = i | T > t)}{u}.$$

The observed (crude) survival function is given by:

$$S(t) = \exp\left(-\int_0^t \mu(s)ds\right) = \exp\left(-\int_0^t \sum_{i=1}^{m} \mu_i(s)ds\right).$$

Despite the fact that there is only one (actual) observable time of death and the corresponding survival time, it might be interesting to study scenarios where we remove one or several causes of death. In this case, the corresponding net survival function of cause i is defined as:

$$S_i(t) = \mathbb{P}[T_1 > 0, \ldots, T_i > t, \ldots, T_m > 0] = \exp\left(-\int_0^t \lambda_i(s)ds\right),$$

where $\lambda_i(t)$ is the net cause-specific intensity of cause i. If studying a hypothetical scenario cause i, the net-cause specific intensities $\lambda_i(t)$ can be modified to reflect expert judgement on the future mortality increase or deterioration that result from adverse events modelled or medical breakthroughs:

$$\lambda_i(t) = \lim_{u \to 0} \frac{\mathbb{P}(T_i \leq t + u | T_i > t)}{u} = -\frac{d}{dt} \log S_i(t).$$

Unfortunately, as only the actual time of death and the related cause are observed, only the $\mathbb{P}(T \leq t + u, I = i | T > t)$ can be estimated from t he data, and not the $\mathbb{P}(T_i \leq t + u | T_i > t)$. To this end, the joint distribution of $(T_1, \ldots, T_m)$ needs to be introduced:

$$S(t_1, \ldots, t_m) = \mathbb{P}[T_1 > t_1, \ldots, T_m > t_m].$$

The crude cause-specific mortality intensities $\mu_i(t)$ are related to the joint distribution $S(t_1, \ldots, t_m)$ through:

$$\mu_i(t) = -\frac{\partial}{\partial t_i} \log \mathbb{P}[T_1 > t_1, \ldots, T_m > t_m]|_{t_1 = \cdots = t_m = t}.$$

Additional assumptions need to be made on the dependency structure between the causes as due to competing risks, it is impossible to observe and so to identify the exact structure, see Tsiatis (1975).

Chiang (1968) proposed to consider the causes being independent, greatly simplifying the modelling which turns out to be very popular, see Prentice et al. (1978), Wilmoth (1995), Putter et al. (2007)) and Boumezoued et al. (2019), Boumezoued et al. (2018) among others. If we assume the that the survival times are independent, net and crude cause-specific mortality intensities are equal but this "may have no resemblance to reality" as quoted by Tsiatis (1975).

2.3.4 Dependency structure using Archimedian copula

To assess coherently the expected mortality patterns in hypothetical conditions when some causes of death are modified or eliminated, one needs to assume other dependency structure than complete independence. The following of this thesis, assumes dependency structure captured by Archimedean survivor copula as suggested by Li and Lu (2019).

The survival times $(T_1, \ldots, T_m)$ has a joint Archimedian survivor copula , if the joint distribution satisfies:

$$\mathbb{P}[T_1 > t_1, \ldots, T_m > t_m] = \psi(\psi^{-1} \circ S_1(t_1) + \cdots + \psi^{-1} \circ S_m(t_m)),\ \forall t_1, \ldots, t_m > 0,$$

where the $\circ$ symbol stands for the composition of functions and ψ - the copula generator function. The Li and Lu (2019) paper proposes Clayton and Frank formula for the application, and for this thesis Clayton copula was chosen arbitrary. The Clayton copula is obtained by letting $\psi(t) = (1+t)^{-1/\theta}$ with parameter θ describing strength of dependency. The higher the θ value, the stronger dependency is assumed between survival times of different causes. When θ tends to zero, the copula reduces to independence copula. The Clayton survivor copula, denoted by $\mathbf{S}$, which is the cumulative distribution function (CDF) of the net survival functions $S_1(T_1), \ldots, S_m(T_m)$:

$$\mathbf{S}(u_1, \ldots, u_m) = \mathbb{P}[S_1(T_1) < u_1, \ldots, S_m(t_m) < u_m],\ \forall u_1, \ldots, u_m \in [0, 1],$$

$$\mathbf{S}(u_1, \ldots, u_m) = [u_1^{-\theta} + \cdots + u_m^{-\theta} - m + 1]^{-1/\theta},\ \forall u_1, \ldots, u_m \in [0, 1].$$

By specifying the survivor copula $\mathbf{S}$ as Archimedian survivor copula, Li and Lu (2019) showed that the net cause-specific intensities, through net survival functions, can be uniquely estimated from the data by using crude-specific intensities and the copula:

$$S_i(t) = \psi\left[-\int_0^t \frac{\exp\left(-\int_0^s \sum_{j=1}^m \mu_j(u)du\right)}{\psi' \circ \psi^{-1} \circ \exp\left(-\int_0^s \sum_{j=1}^m \mu_j(u)du\right)} \mu_i(s)ds\right],\ \forall i = 1, \ldots, m. \quad (9)$$

Crude mortality intensities are observable and can be calculated directly as:

$$\mu_{x,t,i} = \frac{D_{x,t,i}}{E_{x,t}}$$

where $D_{x,t,i}$ refers to number of deaths observed due to cause i of individuals of age x in year t, and $E_{x,t}$ is so called exposure-to-risk - total number of individuals aged x last birthday in year t.

Total, or aggregated, mortality rates are direct sum of mortality rates by cause:

$$\mu_{x,t} = \frac{D_{x,t}}{E_{x,t}} = \frac{\sum_i D_{x,t,i}}{E_{x,t}} = \sum_i \mu_{x,t,i},\ i = 1, \ldots m.$$

Using crude mortality intensities that are calculated directly from the data, together with specified copula and its' generator function ψ we can obtain the net cause-specific intensities $S_i(t)$ in closed form.

2.3.5 Literature review for cause-specific mortality modelling

There are multiple available methods to model and forecast mortality rates by cause. First, and maybe the simplest approach, is to maintain the use of the classical stochastic mortality models originating from Lee and Carter (1992). Alai et al. (2018) used cause specific mortality in the multinomial form to capture the underlying competing risk nature between causes, and Boumezoued et al. (2019) relied on a multivariate Lee-Carter framework to capture the joint (stochastic) dynamics of the cause-specific death rates, while accounting for a detailed treatment of cause-specific historical breakpoints.

Inspired by Aitchison (1982), compositional data framework involves projecting aggregated mortality jointly with by-cause proportions of deaths, so that the sum of the proportions of cause-specific deaths over all causes is forced to always sum to one. J. Oeppen et al (2008) first applied CoDa on the multiple-decrement mortality model to the Japanese population. Alternatively to modelling life table deaths distribution, Piveteau and Tomas (2018) introduce a constraint on aggregate mortality forecast and propose modelling proportions of each cause instead.

Another approach for net and crude force of mortality modelling is based on copulas to account for the dependence between competing risks. Kaishev et al. (2007) investigated partial and complete cause elimination by extending the results of Carriere (1994) to include up to four causes. Dimitrova et al. (2013) generalized the copula approach to cause-elimination by removing and ignoring causes.

In order to account for the cause-specific mortality dynamics in a coherent manner, Li and Lu (2019) combine crude cause-specific mortality projections to Archimedian copula-based modelling approach. The proposed concept accounts for both between-cohort mortality improvements and the dependence between competing risks within-cohort. Authors first derive net mortality intensities by using copula method which models the within-cohort dependence. Then the net intensities obtained are grouped together for all cohorts and fitted to the Lee-Carter model. Finally, future mortality intensities are projected for all causes and cohorts and all-cause future mortality is deduced with corresponding life expectancy forecast.

Our study will be based on the Li and Lu (2019) approach and bench-marked to Boumezoued et al. (2019) to asses the impact of accounting for dependency between causes.

Chapter 3

From independence to dependence cause-specific mortality modelling

This chapter will first present the data that are used for the mortality by cause modelling. Afterwards, cause-specific mortality modelling approaches are illustrated, first assuming independence and then different levels of dependence between causes.

3.1 Presentation of the data

3.1.1 Construction of insured population proxy

SCOR is majorly exposed to mortality risk in the North America's continent, therefore the analysis will be performed on the US mortality data.

Number of deaths in the U.S. national population are originating from the Centers for Disease Control and Prevention (2022). It includes data by gender, age, causes of death, educational attainment, and other fields. Census (2010) data provided exposure information. While the census is performed only once every ten years, the American Community Survey annually provides an estimated update annually, see ACS (2020). Educational attainment by sex and age range is also reported. By using this information along with other data sources, such as population estimates and proportion of educational attainment provided by the U.S. Census Bureau (2022), mortality dynamics by cause of death can be studied by educational attainment status. For our illustrations, highest educational attainment levels (Graduates and bachelors) covering the period 2001-2019 are used to derive an insured population proxy data.

It is well known that mortality rates are highly influenced by individual's socio-economic status, and the United States is recognized to be rather heterogenous country in this aspect: in the US, in 1992 income explained only 3 percent of mortality inequality, but by 2016 state-level income explained 58 percent, see (Couillard et al., 2021).

The US insured population is heavily weighted towards higher economic status, suggesting that using national population mortality data to set mortality assumptions for the insured might results in over-estimation of the mortality. The differences ap-

pear not only regarding mortality level, cause of death mortality proportions also vary between national and insured populations. Due to underwriting process (blood tests, IMC measures etc.), we observe fewer deaths from cardiovascular diseases in the insured population than in the national population, for at least 5 first years after policy inception. Figure 30 illustrates how causes have different proportions in insured and national populations by comparing deaths observed over several years in national population and company's claims in life insurance. For example, Drug-related causes (highlighted in red) are noticeably more important amongst males of national population than what is observed in claim data. Another difference observed due to heart attack (highlighted in blue) is mainly induced by the underwriting process before buying insurance policy. We can also depict cancers as having higher proportions in the insured population - the underwritting process captures well cardiovascular diseases and so greater proportions of cancers are observed.

It would be preferable to directly use internal insured population mortality data to set mortality assumptions, although in practice that is unreasonable due to time period of available observations being too limited to derive medium and long-term future mortality rates and data too scarce at some ages. For this reason insurance companies usually seek to replicate insured population combining national population data and various indicators that would replicate company's policyholders.

Socio-economic status indicators that are associated with health and mortality sequel are wealth, income, occupation, education and race. However, US mortality data by cause of death is only easily available for education and for county-level (geographical) segmentation.

The county-level data can be quite heterogeneous in their socioeconomic characteristics and the observed differences would be dampened by averaging across geographical entity.

The education data, however, is available at an individual person level. This allows for better segmentation of mortality rates and mortality improvement across the US population.

The U.S. insured population is weighted towards higher socioeconomic status which translates to some extent to higher educational attainment levels. Significant differences in mortality rates can be observed for distinct groups of education, as depicted by studies: "Life expectancy at birth between males with less than 12 years of education and those with more than 16 rose from 13.4 years in 1990 to 14.2 years in 2008", see Olshansky et al. (2012). Figure 31 illustrates the evolution of the U.S. male mortality by educational attainment for 5 age groups. Different educational levels have significantly different both mortality levels and trends, and we can observe decreasing mortality patterns persisting over all age groups for highest educational levels,

Figure 30: Density of causes of death for U.S. national and insured male and female populations

while mortality increase or stagnation can be detected for lower educational levels.

By all means, this does not state that there are no policyholders with lower educational level in the insured population, only the fact that the proportions of different educational level will be different between insured and national populations.

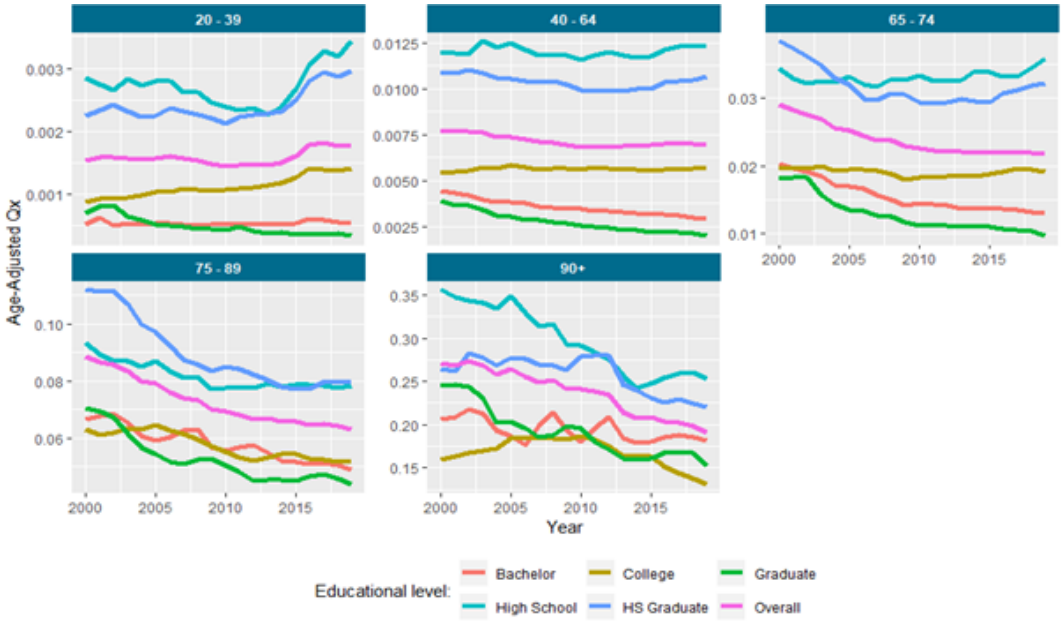

Figure 31: U.S. male age-adjusted mortality rates by educational level

Using mortality data weighted towards higher education levels will not cover entirely the mortality differences between insured and national populations and insurance companies can choose to supplement the developed assumptions with the internal claim experience or other available data. Nevertheless, for this thesis we will construct an insured population proxy by selecting mortality of highest educational level for U.S. individuals. For illustration purposes the results will be presented using the male population to avoid repetition, without loss of generality.

3.1.2 Cause of death grouping

All human deaths are assigned to a single (main) cause using the The International Classification of Diseases (ICD) in the death certificate, which ensures common cause of definition in different countries and regions. The ICD - medical classification list by World Health Organization (WHO) - has different levels of granularity available. Also, different revisions of ICD were used over time which accounts for some diseases disappearing and emerging of new diseases. The ICD-10 has been in use since around 1994 and can provide over 25 years of unified cause of death classification.

We will seek to forecast mortality rates, and we need to ensure reasonably stable historical data to build the future projections. Minimal death count must be guaranteed to obtain statistically reliable results, therefore causes of death should be classified to

reasonably large groups. Five large groups of causes are retained for our application, assessing only Alzheimer's and dementia as "small cause" as specific scenario will be applied in the last chapter. The retained grouping is presented in Table 7.

Causes of death	ICD-10 codes
Alzheimer's and dementia	F01, F03, G20-G21, G30
Neoplasms	C00-C97
Circulatory system diseases	I00-I99
Respiratory diseases	J00-J98, U04
External causes	U01, V01-Y84
Other	All other causes not included elsewhere

Table 7: Cause of death grouping and ICD-10 codes

3.2 Historical observations

Cause-specific mortality rates (or more precisely force of mortality) that are used for the further work are directly obtained from the data for all 6 retained cause of death groups between 2001 and 2019.

It is important to understand past trends and examine those through multiple axis to set assumptions for future projections precisely. At this stage we analyze data for ages between 20 and 89 to assess past trends, but for the modelling and future projections limited age range will be chosen to reflect ages at risk for the insured population in mortality and longevity business lines.

Figure 32 depicts cause-specific mortality trends over different age groups for the male population. Several observations can be deduced from the figures:

- For the youngest ages, External causes are the predominant cause of death for the male population. Age band of 20:39 seem to have steady increase in mortality due to External causes up to 2017 with a possible change in trend over the two last years observed. For other age bands, External causes follow rather linear trend over the whole observation period and mortality due to External causes reduces to negligible for ages 60 and above.

- The decreasing trend in Circulatory disease mortality is observed for ages 40 and above. This decrease was the main reason for all-cause mortality decline. Slowdown in mortality decrease is visible for ages 40:59 in most recent years, and ages 60:74 exhibit increase in Circulatory disease mortality starting around year 2012. For the eldest, Circulatory system diseases is the most important cause of death and the downwards trend appears to be continuing, only at slower pace.

- Alzheimer's and dementia is insignificant cause of death at younger ages, which is expected as these conditions affect individuals at high ages. Mortality due to

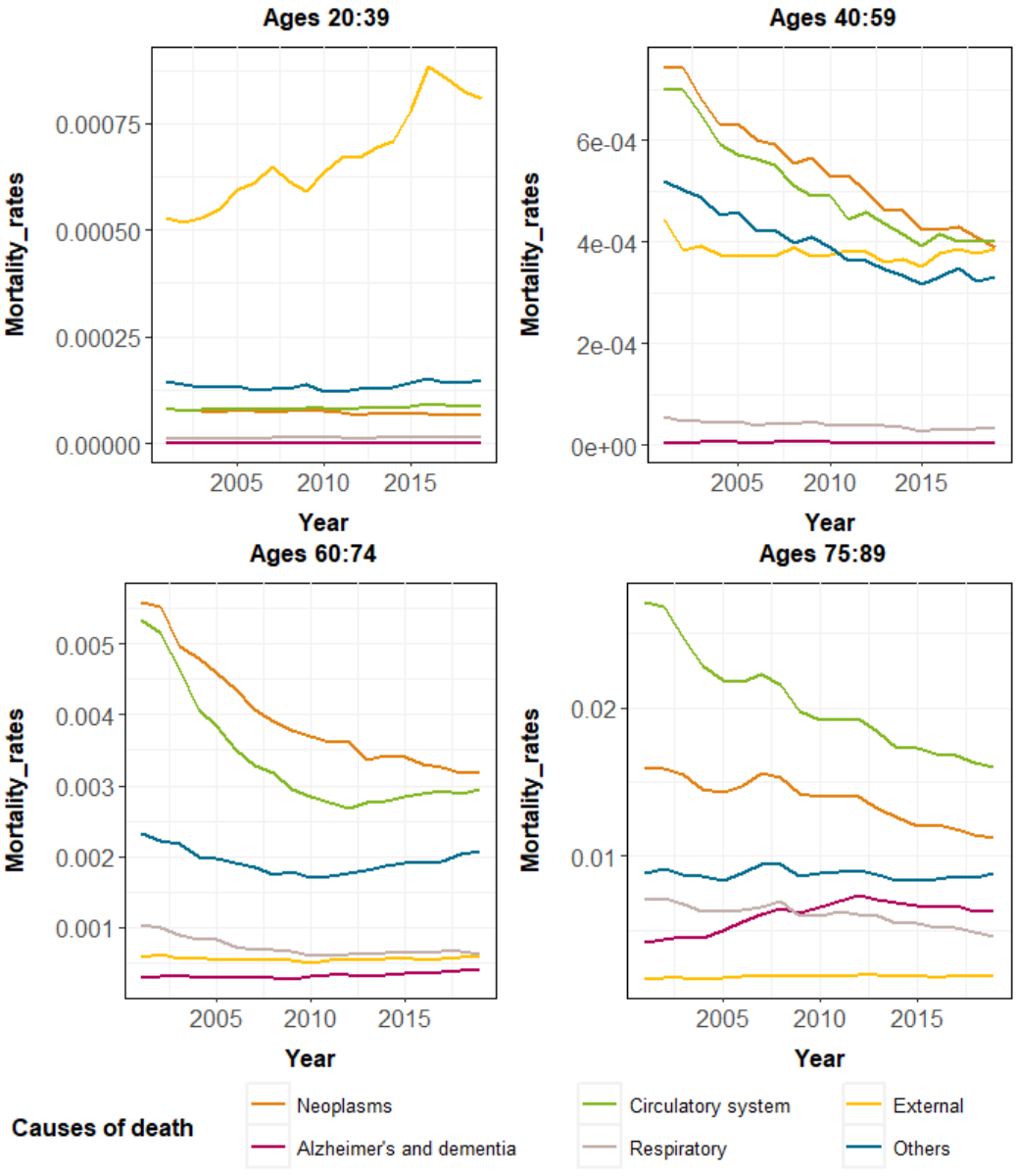

Figure 32: Male cause-specific mortality rates by age band

Alzheimer's and dementia diseases started increasing around 2005 and appears to stagnate as of year 2012.

The following Figure 33 illustrates cause-specific mortality evolution by age (in log scale) in year 2010. One can observe that age structure is not consistent for different causes. First, External cause mortality demonstrates constant linear pattern with a sudden level jump around age 35 up to age 75 approx. increasing afterwards. Alzheimer's and dementia mortality exhibits the most substantial increase with age.

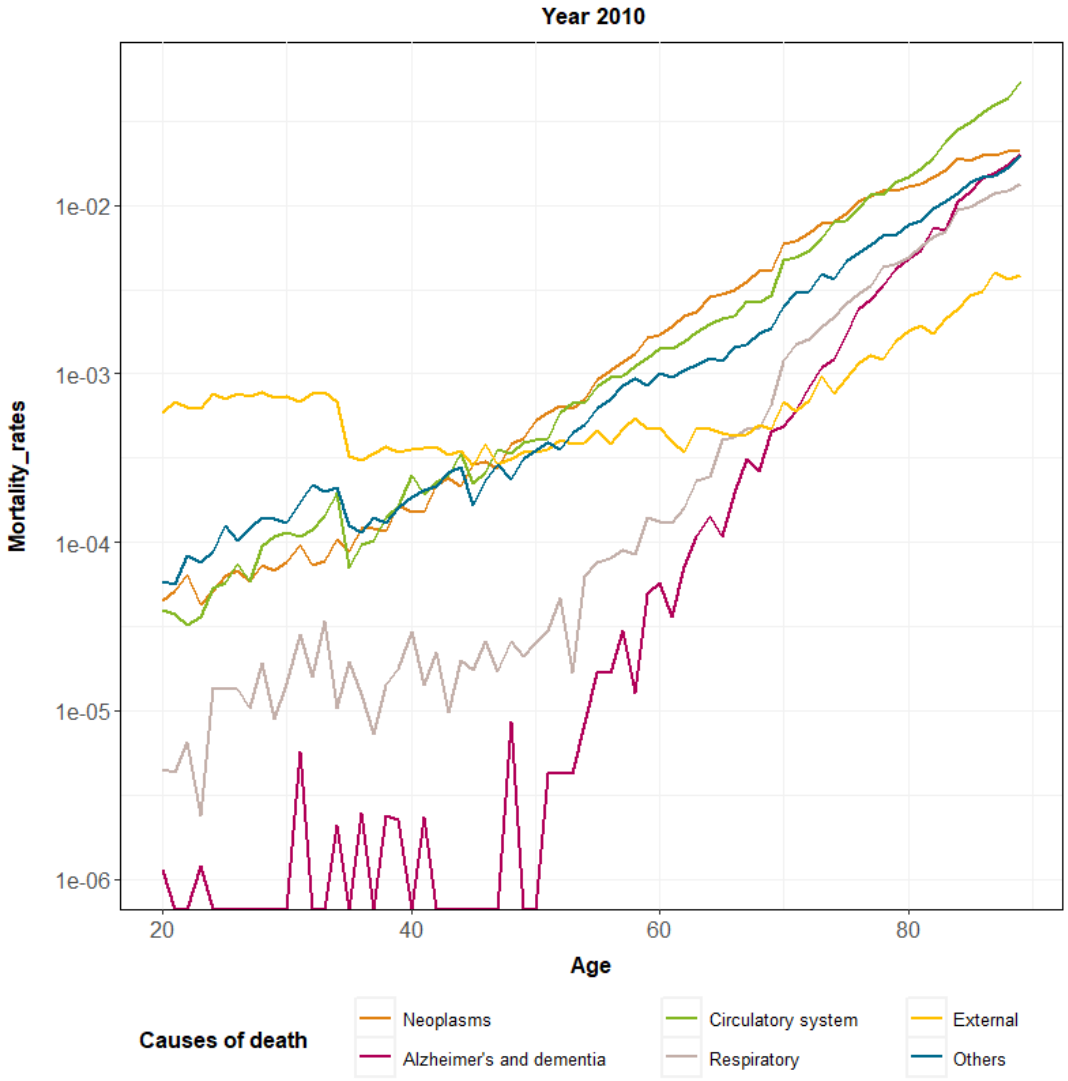

Figure 33: Cause-specific male mortality in 2010 (log scale)

Figure 34 depicts proportions of different causes composing aggregated mortality over the age palette for the last year in the observation period - 2019. External causes explain almost all the mortality of young males with the proportion decreasing with age and reaching negligibly small proportion around age 60. Mortality due to Neoplasms reaches it's peak around age 60 and it's importance start decreasing around age 75. Circulatory disease-related mortality seems to be stable for ages 50

to 89, and the Alzheimer's and dementia disease mortality comes into picture around age 60 and steadily increases to highest ages where it covers over 15% of the aggregate mortality.

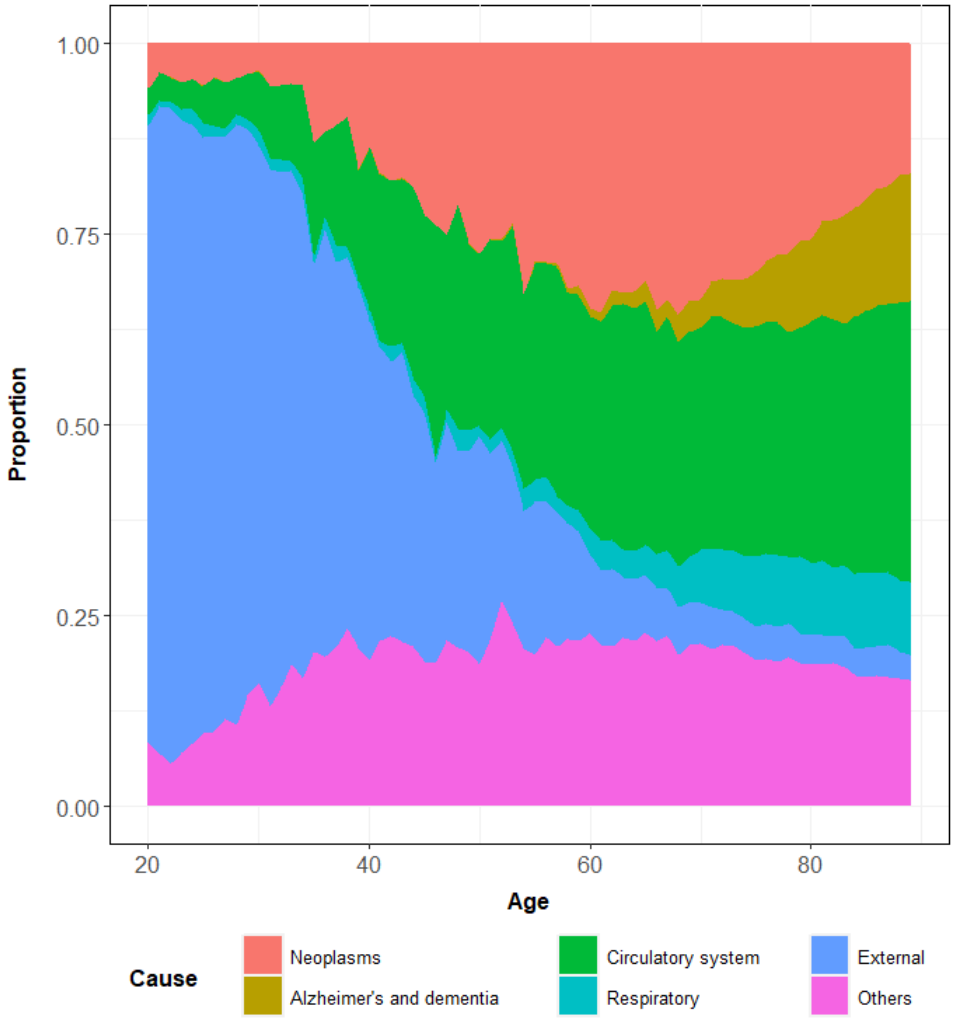

Figure 34: Cause-of-death proportions in 2019 by age

As the aggregate mortality illustrated in the Figure 29 depicts change in trend around the year 2009, at least for age 65, let us look at life expectancy gains over the whole period split into two intervals: 2001-2009 and 2010-2019, see Figure 35.

The observed changes in life expectancy are decomposed over the two periods showing the contribution of each cause at each age between 20 and 89, see Vaupel and Romo (2003) for more details. Values below zero represent a negative contribution of the specific cause to changes in life expectancy (loss), while positive values represent a positive contribution of the cause to changes in life expectancy (gain). Bars over all causes and all ages sum up to total life expectancy gain over the period. Males in the U.S. gained 2.42 years of life expectancy over 2001-2019 than can be broken down into gain of 1.77 years in the period 2001-2009 and only 0.54 for the following period

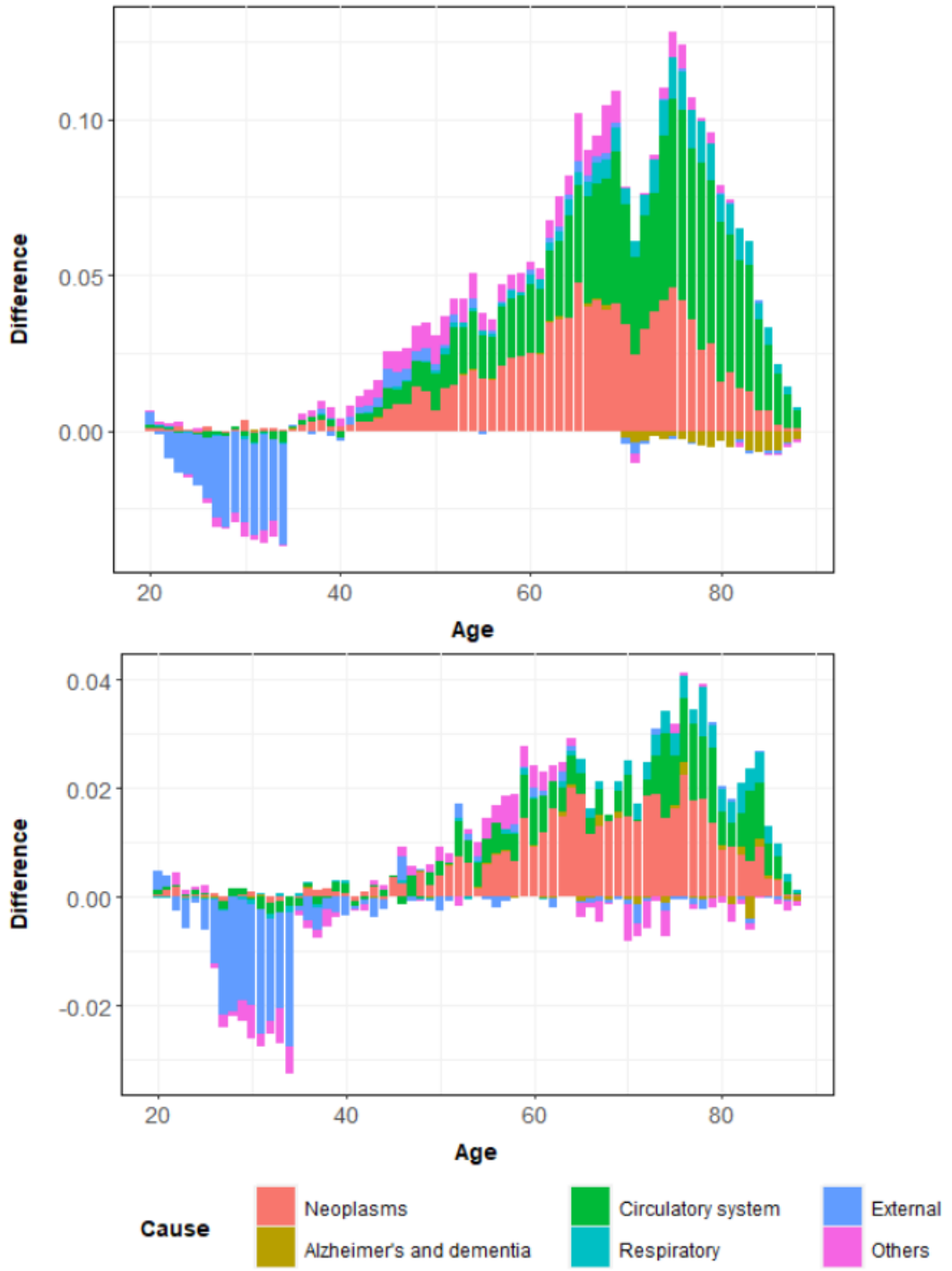

Figure 35: Gains in life expectancy between 2001 and 2009 (top) and between 2010 and 2019 (bottom) for the U.S. male population

2010-2019. Over the first time interval, important gains came from cardiovascular disease mortality, while over the second period they were much lower. Important life expectancy losses occurred due to increased external cause mortality in young ages, specially in the second part of the study period. Alzheimer's and dementia diseases participated mainly in decreasing life expectancy over 2001-2009. Overall, a general trend of decreasing gains in life expectancy can be observed over all causes rather than losses in life expectancy.

3.3 Break-point algorithm

Due to important changes in cause-specific trends observed over the observation period, simply calibrating mortality model on the whole set of data could lead to unreasonable future projections. Instead, quantitative algorithm could be applied on each cause time series data to detect significant changes in mortality trends, see (Berkum et al., 2014). The time component of the Lee-Carter model (introduced in section 2.2) can be then calibrated only on the latest retained trend if a change in trend is detected by the break-point algorithm.

Let us assume to have at our disposal period effect $\kappa_t^{(i)}$ for every cause i once the Lee-Carter model is fit on the whole mortality data series. Let us define the first-order differences by $\Delta\kappa_t^{(i)} = \kappa_t^{(i)} - \kappa_{t-1}^{(i)}$ for $t = 2, \dots T$. A random walk with piece-wise constant drift can be estimated:

$$\kappa_t^{(i)} = \begin{cases} \beta_1 + \epsilon_t, & t \leq t_1 \\ \dots & \\ \beta_j + \epsilon_j, & t_{j-1} < t \leq t_j \\ \dots & \\ \beta_{m+1} + \epsilon_t, & t_m < t \end{cases}$$

where $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ are independent over time. The model is estimated using ordinary least squares (OLS), hence, the sum of squared residuals (SSR) is minimized:

$$\text{SSR}(t_1, \dots, t_m) = \sum_{j=1}^{m+1} \sum_{t=t_{j-1}+1}^{t_j} [\Delta\kappa_t^{(i)} - \beta_j]^2$$

where $t_0 = 1$ and $t_{m+1} = T$. We distinguish m break points that divide the time series into $m + 1$ periods with different drifts. Both the number of break points and the dates of the break points are unknown.

Let $\beta(T_m)$ estimates $\{\beta_1, \dots, \beta_m\}$ based on a given m-partition $(t_1, \dots, t_m)$ denoted T_m. If we substitute these parameter estimates $\beta(T_m)$ to κ_t^i formula then the esti-

mated break points $(\hat{t_1}, \ldots, \hat{t_m})$ are such that

$$(\hat{t}_1, \ldots, \hat{t}_m) = \text{argmin}_{(t_1,\ldots,t_m)} \text{SSR}(t_1, \ldots, t_m)$$

where the minimisation is taken over all partitions $(t_1, \ldots, t_m)$ for which $t_j - t_{j-1} \geq h$. The parameter h corresponds to the minimum period that the regime should last and can be chosen upwards.

The break-point algorithm is implemented in the R package `strucchange` (see Zeileis et al. (2022) for more details).

Nevertheless, for this thesis assessing break-point problematic through the proposed algorithm remains complicated due to different quantities on which the algorithm will be applied. The approach under independence assumption will fit and project crude (which are equal to net in this case) mortality intensities, while once the dependency between causes is taken into account, net intensities are used, and in addition, different copula's θ parameter values will produce distinctive net mortality intensity values. For comparison purpose, we will use identical periods for all the approaches which are varying by cause and are set arbitrary by visual inspection to produce plausible future trend projections. Model choice can be refined by applying plausibility concept to future trend evolution. Cairns et al. (2006b) introduced the concept of biological reasonableness, drawing on the concept of economic reasonableness from interest-rate modelling, and proposed to extend it to coherent mortality trend forecast evaluation. The method is not based on any scientific, biological or medical model, but is rather a subjective way to assess qualitatively future projections by establishing opinion on biological factors, medical breakthroughs and environmental changes.

The retained calibration periods for each group of causes are presented in Table 8:

Group of causes	Period
Neoplasms	2001 - 2019
Alzheimer's and dementia	2010 - 2019
Circulatory system diseases	2005 - 2019
Respiratory diseases	2010 - 2019
External causes	2010 - 2019
Others	2008 - 2019

Table 8: Periods retained for calibration

3.4 Application

The following section presents the estimation procedure and results for mortality forecasting assuming independence between causes and compares it to copula-based method where dependence between causes is accounted for. We will proceed with data for ages 55 to 89 to have better representation of the age mix in longevity and

mortality business which will also harmonize the projections as the observed negative trends for younger ages will be excluded from observations used.

3.4.1 Independent cause-specific mortality model

The Lee-Carter model is independently calibrated for each of the causes using crude mortality intensities that are obtained directly from the data. As independence is assumed between times of death for different causes, crude and net mortality intensities are equal. The projections of mortality intensities for cause i are obtained by extrapolating mortality trend $\kappa_{i,t}$. Authors (Li and Lu, 2019) propose the following constraints as replacement to the original constraints proposed by (Lee and Carter, 1992):

$$\sum_t \kappa_{i,t} = 0, \text{ and } \frac{\kappa_{i,T} - \kappa_{i,1}}{T} = d, \text{ for } i = 1, \ldots, m, \tag{10}$$

where d is cause-invariant constant, which will facilitate comparison between mortality patterns for different causes. Under this normalization constraint, the time parameters $\kappa_{i,t}$'s have the same average trend for all causes, and so to compare different cause mortality improvements will restrain to comparing only $\beta_{i,x}$'s. The larger the $\beta_{i,x}$, the faster the improvement. We adopt the constraint (Li and Lu, 2019) proposed $d = -0.2$. With this constraint, the $\kappa_{i,t}$'s will be decreasing for all causes. Accordingly, a positive $\beta_{i,x}$ indicates that he mortality intensity of age x due to cause i is decreasing over the observed period.

β parameters in the Figure 36 indicate that Neoplasms, Circulatory system and Respiratory diseases were the causes where mortality decrease was observed for all ages 55 to 89 over the period. We observe that the decline in mortality was similar between Neoplasms and Circulatory system up to age 70, but for higher ages the decline of Circulatory system diseases was faster. Mortality deterioration due to Alzheimer's and dementia causes affected all ages over 70. β_x's for the External causes fluctuate around zero over the age range meaning that mortality mainly stagnated due to this cause. Similar observation can be made for the Other causes for ages above 70 while mortality due to this cause decreases for ages 55 to 70.

Even though if κ parameters have the same average trend due to the constraint applied (10), trends are still somewhat different - κ of Other causes exhibit increased volatility, end of the observation period had strong improvements for External causes and Alzheimer's and dementia disease mortality experienced increasing trend since 2012.

Future mortality rates are projected by extrapolating the period parameters $\kappa_{t,i}$'s following Random Walk with Drift (see Prospective mortality modelling) that are calibrated corresponding to the retained periods for each cause presented in the Table 8. We project 41 years into the future, 2060 being the final projection year.

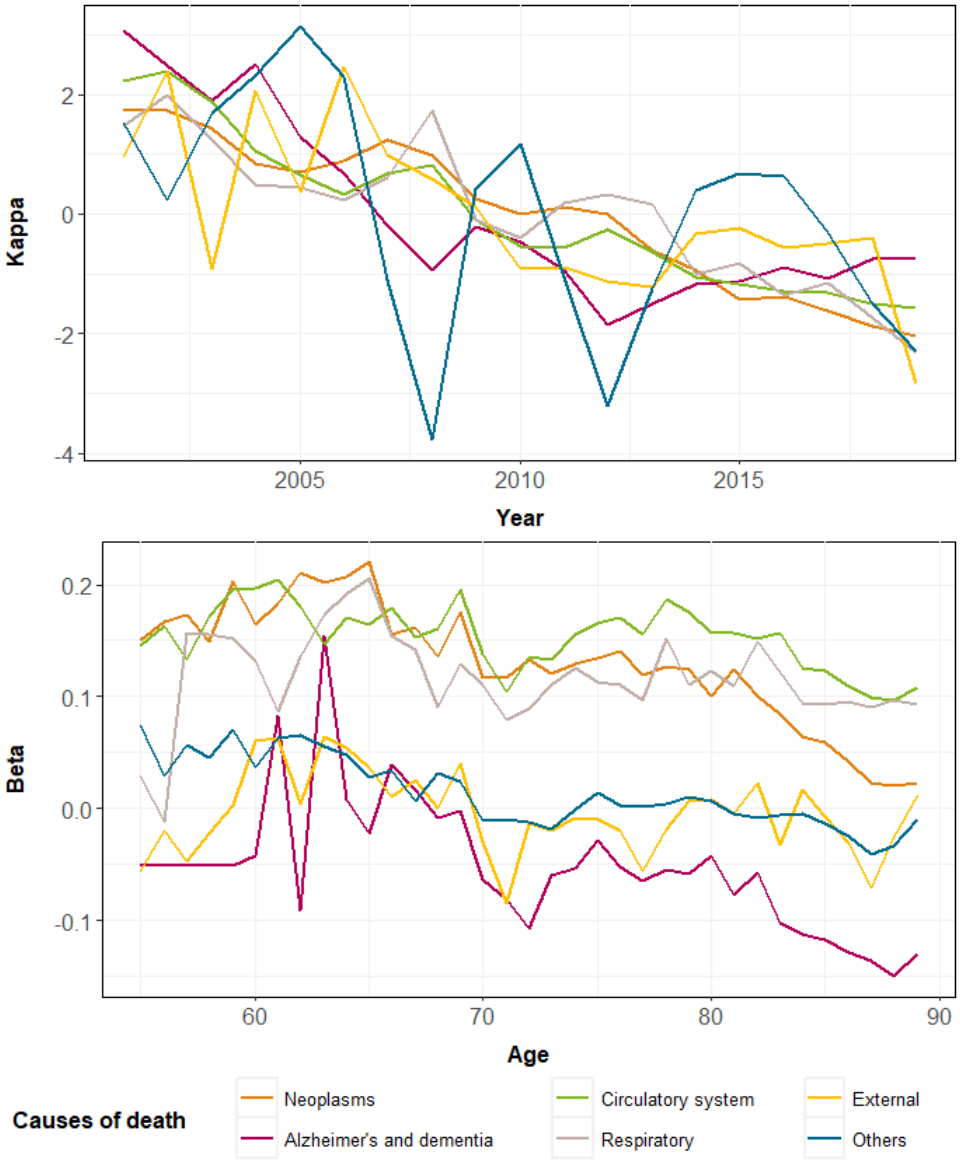

Figure 36: κ and β parameters by cause of death under independence assumption

We obtain future mortality projections for the same age range as the initial data - 50 to 89 years from the Lee-Carter model. To obtain cause of death mortality in advances ages, we use approach proposed by Piveteau (2021). The method consists first extrapolating the all-cause mortality using standard mortality table closure methods, we use Denuit and Goderniaux (2005). In the second step, contribution by age of each cause to the all-cause mortality is projected for older ages, multinomial model with P-splines is used for this purpose. The product of the contributions and all-cause mortality results in the extrapolated mortality by cause of death and we set maximum death age 120.

Figure 37 depicts projected mortality intensities for ages 55 and 75. Age 55 will be close to a representative age of mortality business while average age for longevity inforce is 75. At age 55, External and Other causes are projected with an increasing trend, and an important decrease for the Cardiovascular and Neoplasms mortality. At age 75 we depict strong decrease in mortality rates due to Circulatory system diseases and Neoplasms with slight increase in mortality due to External, Other causes and Alzheimer's and dementia. Respiratory disease mortality intensity is marginally declining for all ages.

Cause-specific mortality projections allows us to better understand and validate (or not) future hypothesis, apply expert judgement if desired and produce scenarios on one or several causes which will be the proposal of the following chapter. Nevertheless for actuarial use one needs to assess aggregated mortality rates and see the corresponding life expectancy that is forecasted. Aggregated mortality intensity is obtained directly by summing up the cause-specific intensities for both past and future projections and presented in the Figure 38.

For age 55, modest decrease in aggregated mortality is projected for approximately ten years, then stagnating and even slight increase by the end of the forecasting period, which is explained by External and Other cause future mortality trends projected that are not offset by other causes over the whole period. Age 75 aggregated mortality are expected to decline in the future.

The following table represents residual periodic life expectancy projected and observed for 55 and 75 year-old males.

Residual life expectancy				
Age	Year 2001	Year 2019	Year 2040	Year 2060
55	26	29	30.8	32.1
75	10.1	11.8	13	14

Table 9: Periodic life expectancy under independent cause assumption

We note that the life expectancy gains are projected to slow down for both ages. For age 55 we observed a gain of 3 years between 2001 and 2019, while the gain projected

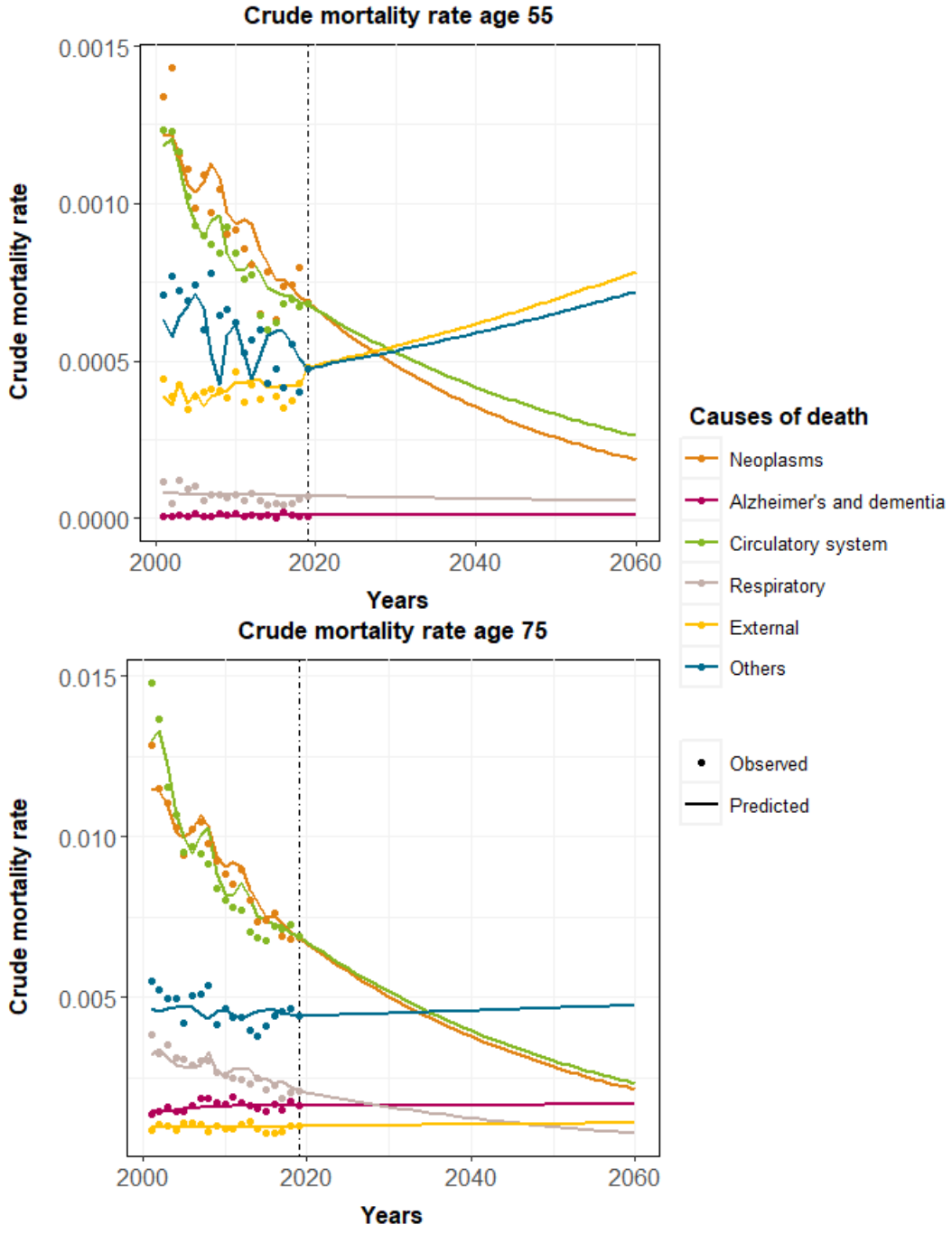

Figure 37: Cause-specific (crude=net) mortality intensity projections under independence assumption

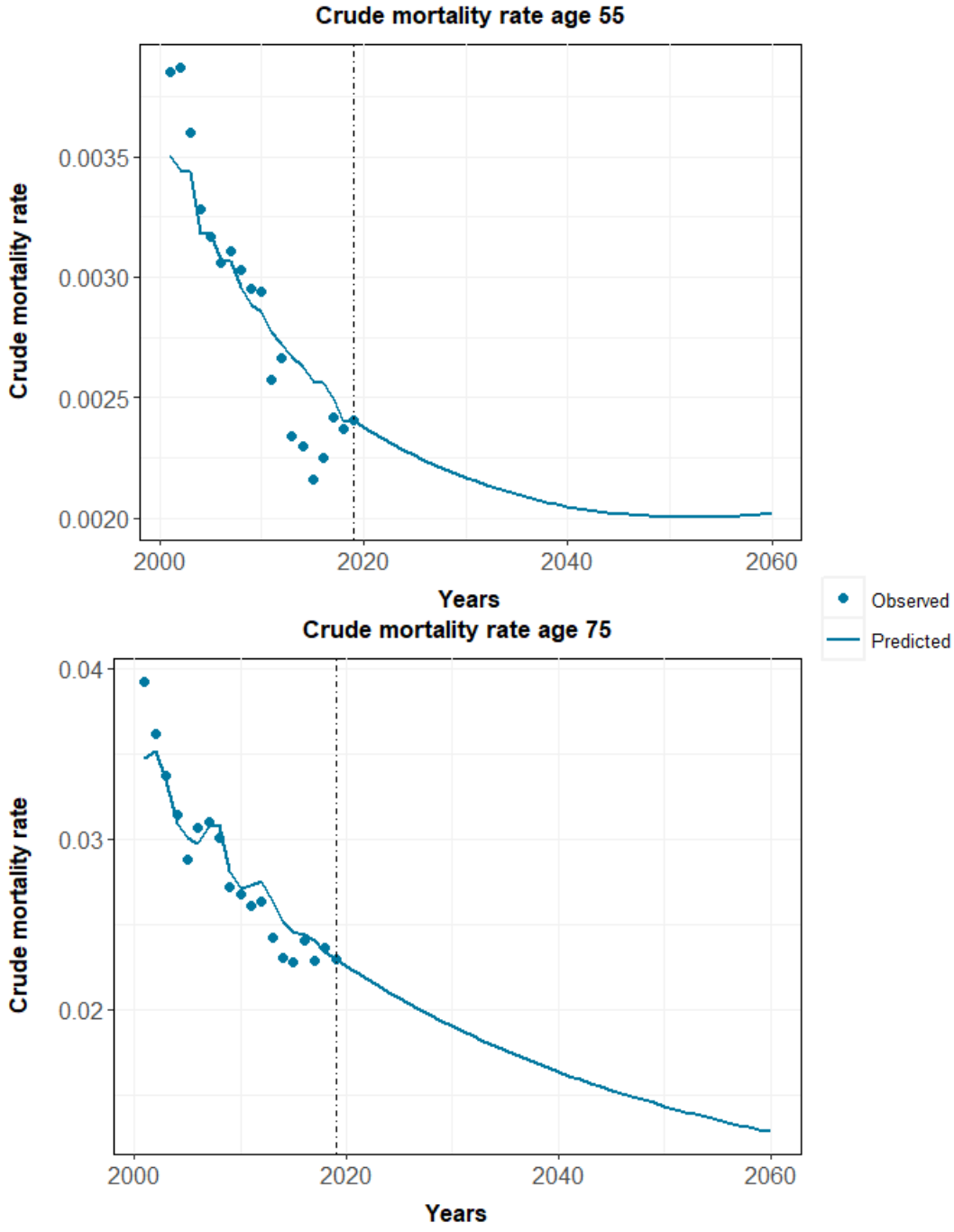

Figure 38: Aggregated mortality projections under cause independence assumption

over the next 20 years is only of 1.8 years, and the following 20 years are predicted to increase life expectancy of only 1.3 years. Similar results are obtained for age 75 (gain of 1.7 between years 2001 to 2019 and only 1 year between 2040 and 2060).

3.4.2 Cause-specific mortality modelling under copula-based dependency framework

In this part, we move away from the assumption of causes being independent to the approach introduced in section 2.3.4 that accounts for dependency between causes within the Archimedian copula framework.

We work with Clayton copula as described in the section 2.3.4. Due to copula being not identifiable, the results will be presented with two values of the θ parameter that is driving the dependence intensity. As θ goes to zero, the copula reduces to independence copula by so making net and crude intensities being equal which is exactly the approach implemented in the previous section. To the contrary, the greater the θ value, the stronger dependence is implied between causes. We choose $\theta = 1$ and $\theta = 4$ similarly to the range of values presented in the paper of Li and Lu (2019) for the application.

Here below we first introduce the estimation procedure for net mortality intensities and obtaining aggregate mortality projections:

- Crude mortality intensities $\mu_{i,c,t}$ are obtained directly from the death count and exposure data for each cause of death i, cohort c and year t:

$$\mu_{i,c,t} = \frac{D_{i,c,t,}}{E_{c,t}}, \ \forall i = 1, \ldots, m.$$

- Using the proposed net survival function expression (9), crude intensity of each cause can be transformed to the net survival function $S_{i,c}(t)$. The net mortality intensities $\lambda_{i,c,t}$ can be now derived from t he net survival functions by:

$$\lambda_{i,c,t} = -\log \frac{S_{i,c}(t)}{S_{i,c}(t-1)}.$$

- The (Lee and Carter, 1992) model is used to forecast net intensities for each cause separately. To obtain comparable results, we use identical periods by cause for the period parameter κ_t projections.

- After projecting the net intensities the reverse reasoning is applied to recover the corresponding crude intensities using equation (9). The latter are then used to obtain the aggregate future mortality and corresponding life expectancy projected at different time frames.

For comparison, we use the same Lee-Carter model constraints as in the model with independence assumption, see (10) to fit and project net mortality intensities.

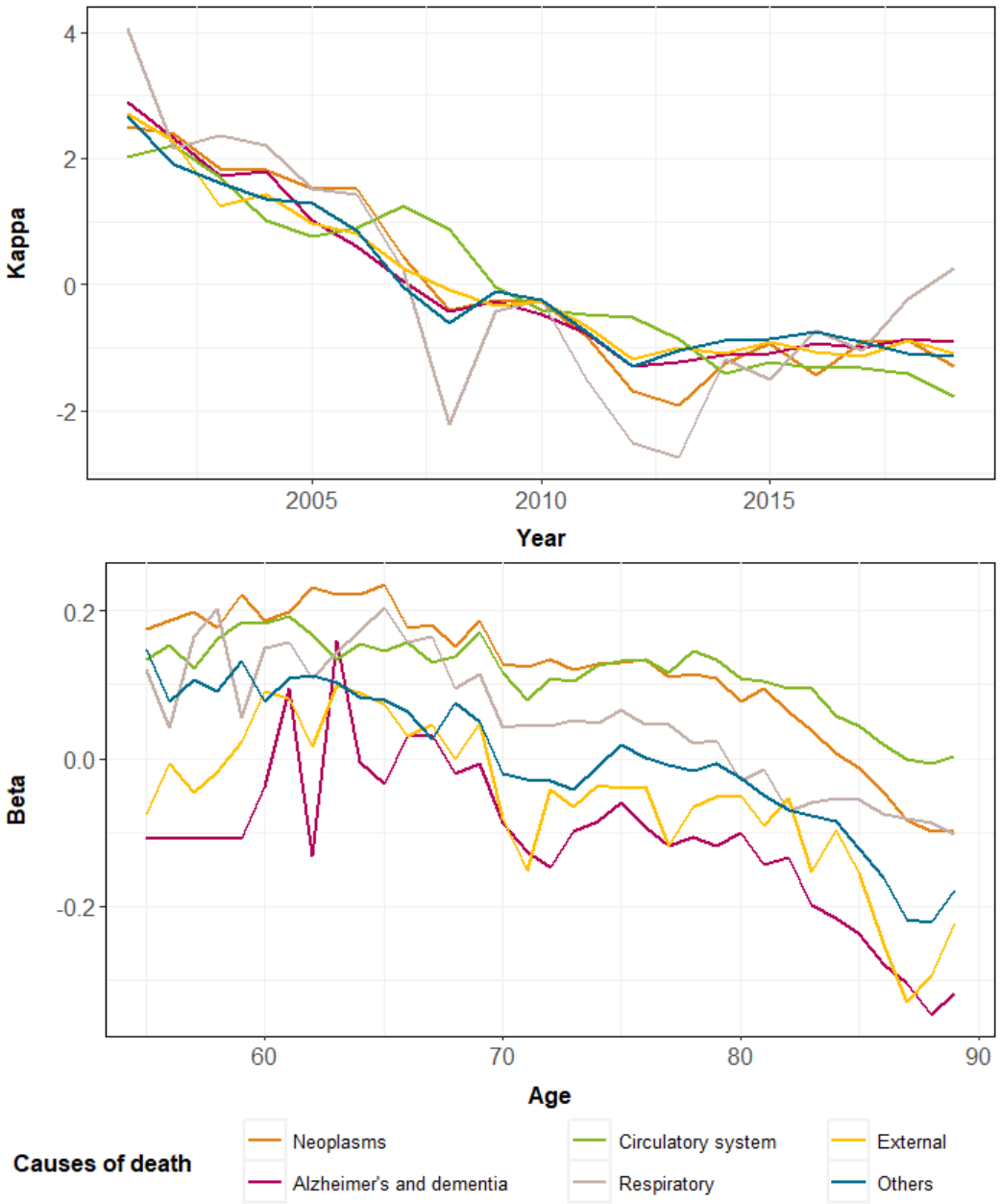

Figure 39: Net κ and β parameters with Clayton copula, $\theta = 1$

From Figure 39 and 40 we notice how the cause-specific trends are similar for different parameter values, and introducing positive dependence structure between Neoplasms and Circulatory system diseases reduces the difference between the two causes - β curves are very close for $\theta = 4$ and start diverging around age 77. In the independent case (see Figure 36) Circulatory system and Neoplasms β were positive over the whole age range, while introducing dependency structure β of net mortalities become negative for ages over 85 for Neoplasms for $\theta = 1$ and even sooner, for ages 80 and above, for $\theta = 4$ for both Neoplasms and Circulatory diseases.

We can also observe the κ values for both θ parameters plateau'ing around year 2012, with higher θ value introducing additional volatility for the Circulatory system cause and case of $\theta = 1$ - similar irregular patterns for the Respiratory diseases. The net intensity trends of fast decline in mortality up to approx. 2012 with stagnation or even increasing trend afterwards is well captured.

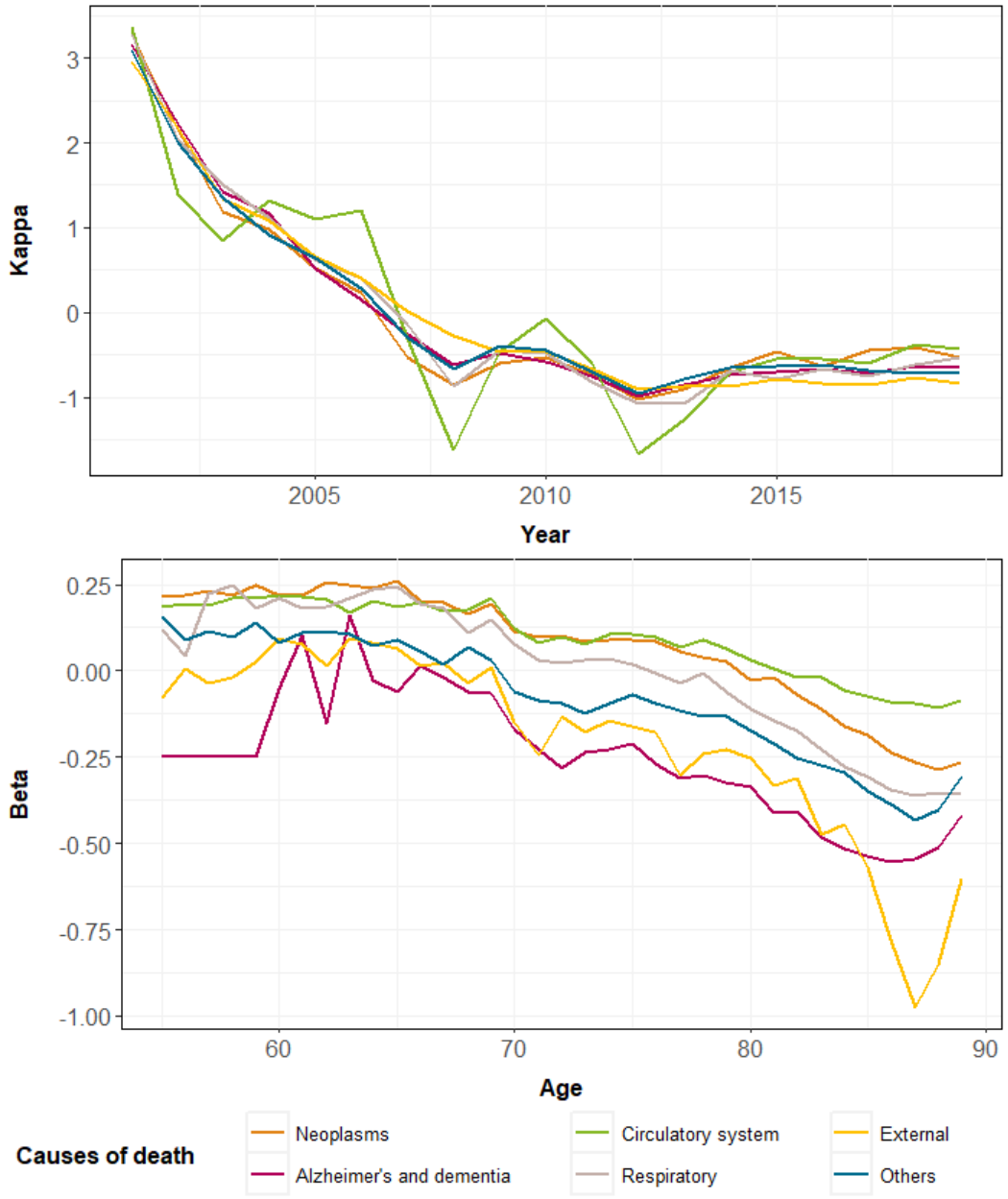

Figure 40: Net κ and β parameters with Clayton copula, $\theta = 4$

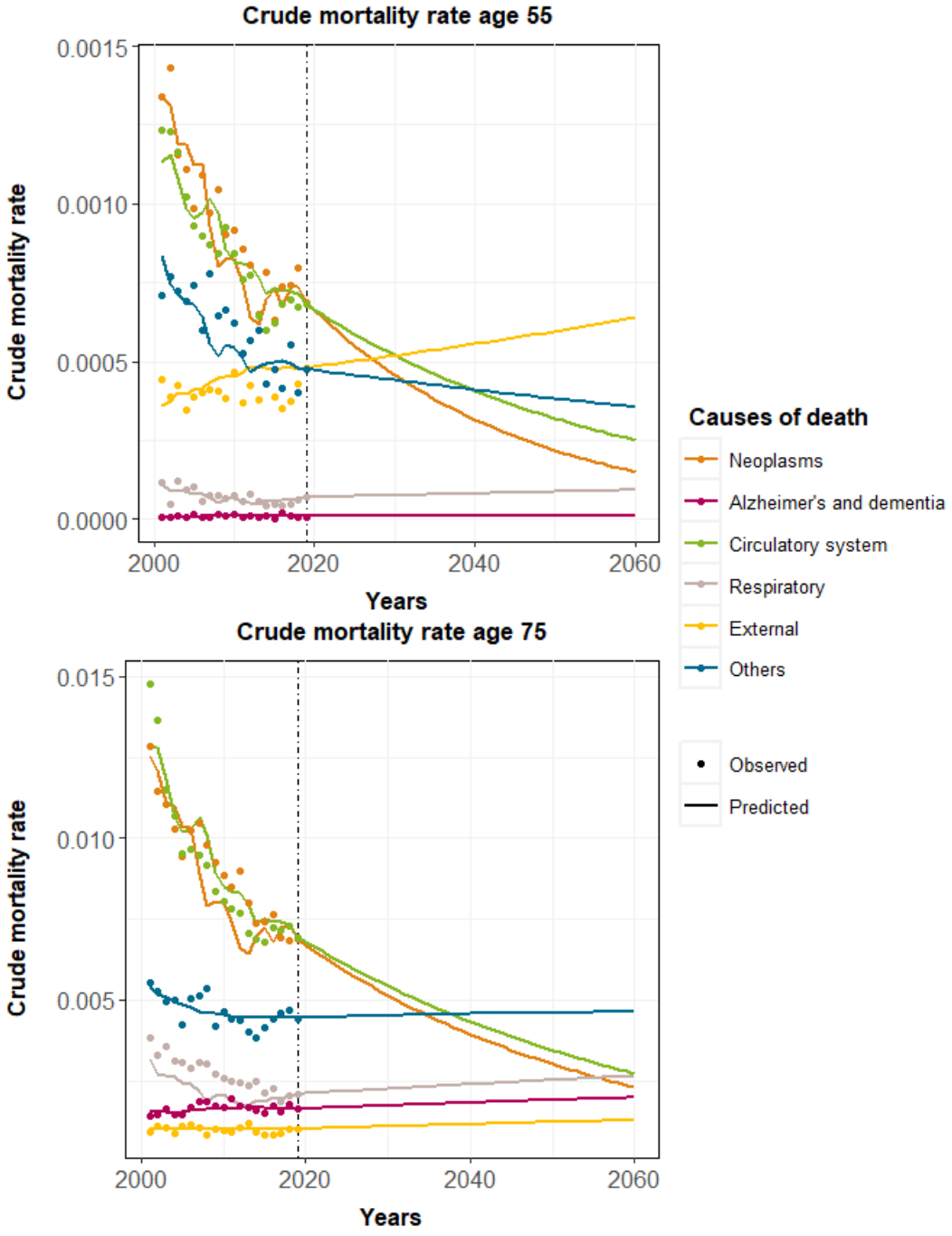

Figure 41: Crude mortality intensities, $\theta = 1$

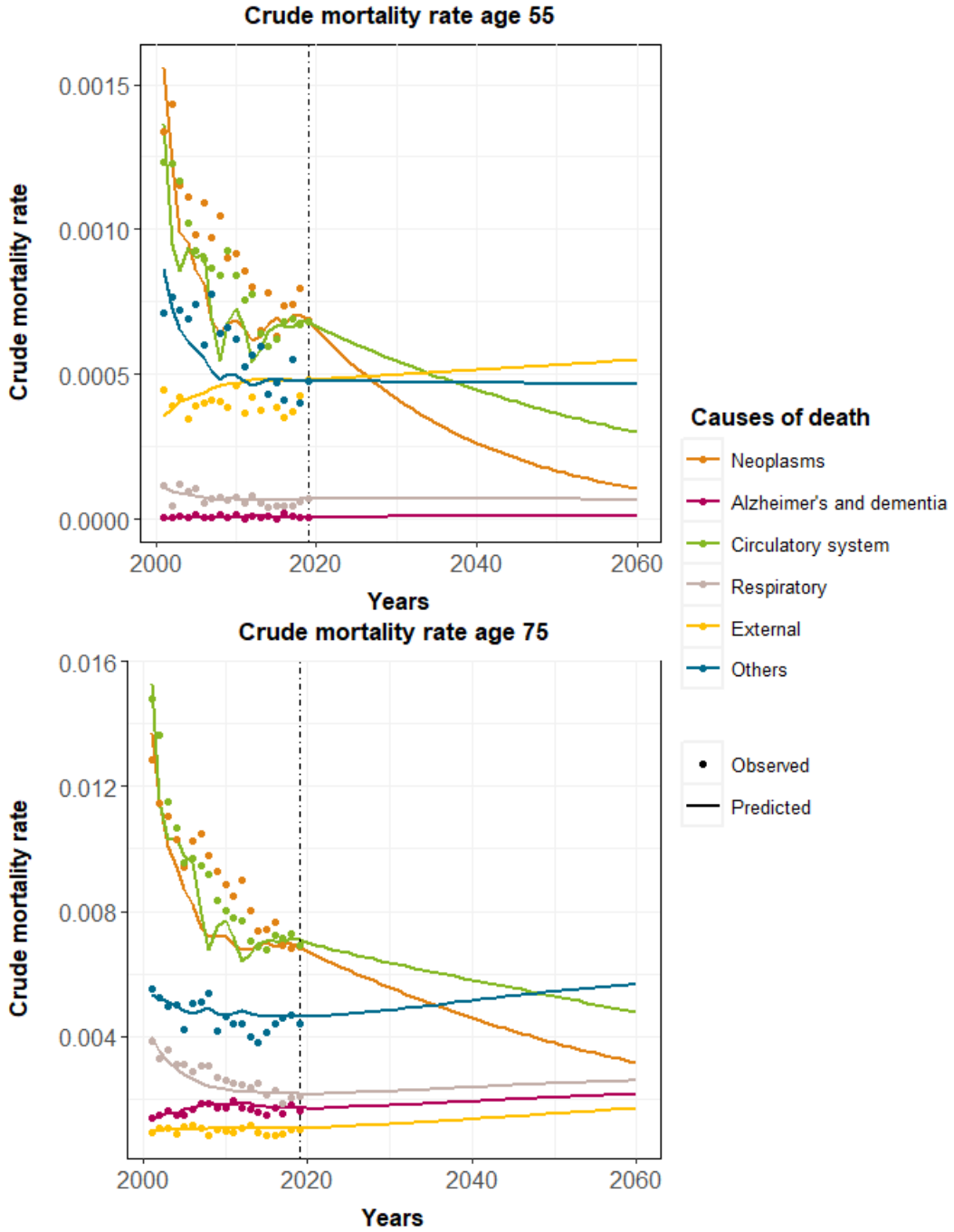

Figure 42: Crude mortality intensities, $\theta = 4$

After fitting the net mortality intensities using Lee-Carter model and projecting κ parameters using Random Walk with Drift, we obtain net mortality intensity projections that are used to recover the corresponding forecast of the crude intensities, see Figure 41 for $\theta = 1$ and Figure 42 for $\theta = 4$.

The crude mortality intensities for External and Other caused are obtained at age 55 are lower compared to the independent case and the ones for $\theta = 4$ are lower than for case where $\theta = 4$. On the other hand, for age 75 the increase in mortality intensities for External, Other, Alzheimer's and dementia and Respiratory causes is more important for larger θ values. As we assume stronger dependency between causes, so stronger common underlying factors, the trends tend to move "similarly".

The resulting all-cause mortality intensity confirm these observations and predict more significant mortality decrease for age 55 when assuming dependency (see Figures 38 and 42), while for higher ages mortality decrease is faster when assuming less or no dependency, see Figures 43 and 44. Table 10 depicts periodic life expectancy for the independence approach ($\theta = 0$) and two dependency intensity levels:

Residual life expectancy (years)				
Age 55	Year 2001	Year 2019	Year 2040	Year 2060
$\theta = 0$	26	29	30.9	32.1
$\theta = 1$	26	29	30.1	30.7
$\theta = 4$	26	29	29.3	29

Age 75	Year 2001	Year 2019	Year 2040	Year 2060
$\theta = 0$	10.1	11.8	13	13.9
$\theta = 1$	10.1	11.8	12.3	12.4
$\theta = 4$	10.1	11.8	11.6	10.8

Table 10: Periodic life expectancy under independent cause assumption

Assuming causes being independent generates highest life expectancy, the variation for age 55 going up to 2.1 years by year 2060. As we saw, the mortality projected for younger ages was decreasing with θ values, and inverse was observed for higher ages. Supposing high dependency between causes yields to mortality stagnation and even slight increase.

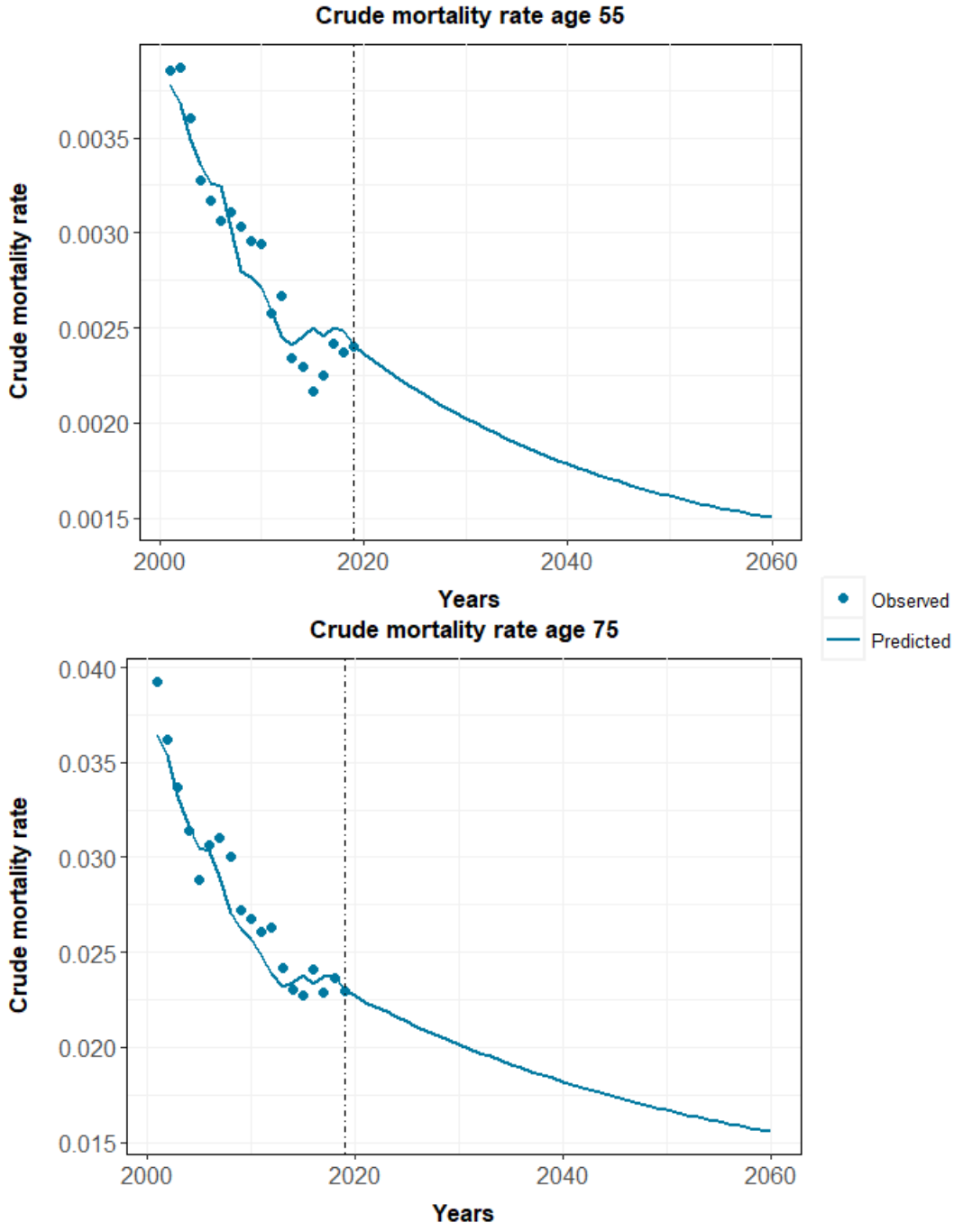

Figure 43: Crude aggregated mortality intensity, $\theta = 1$

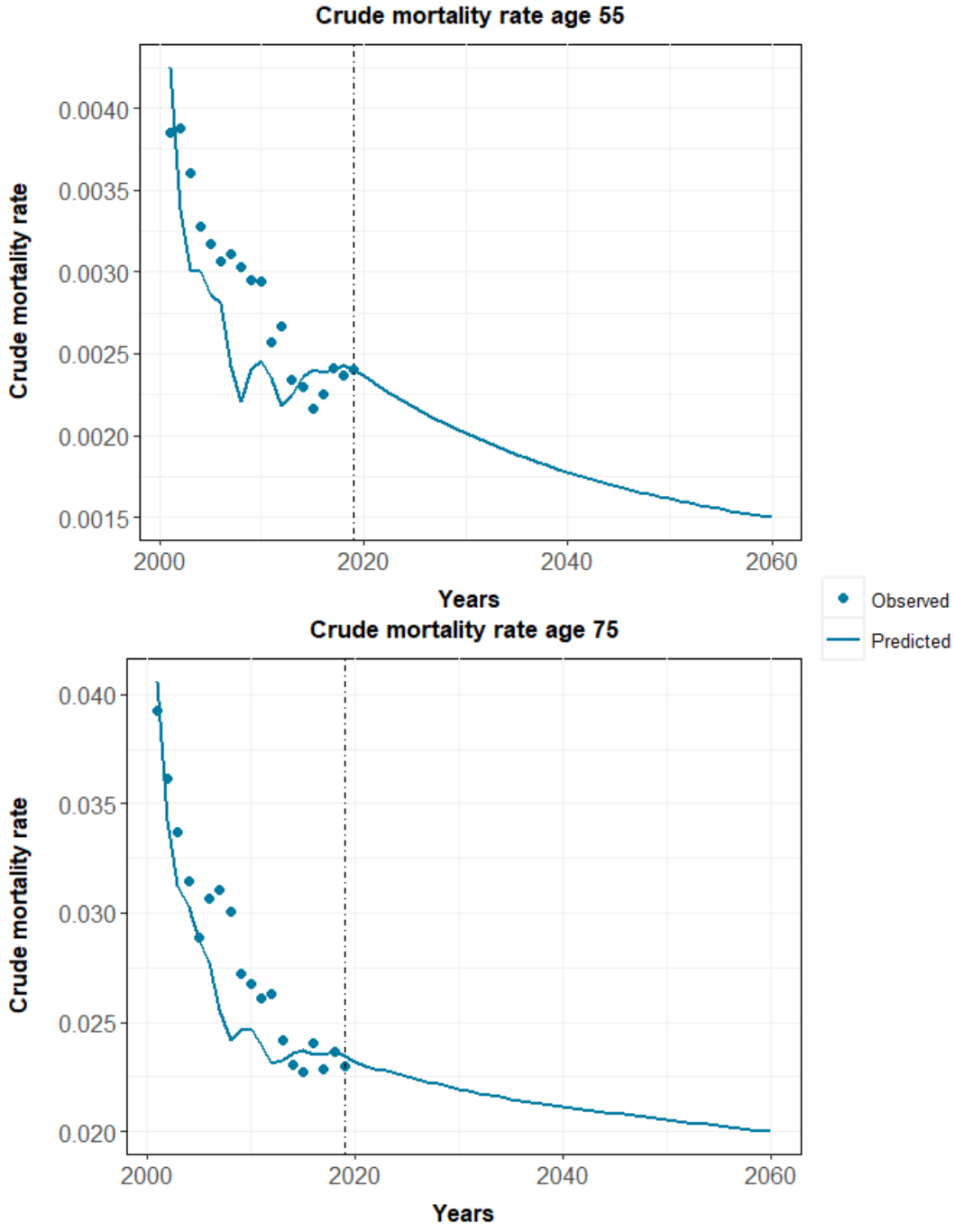

Figure 44: Crude aggregated mortality intensity, $\theta = 4$

Chapter 4

Scenario modelling for cause-specific mortality

This final chapter will introduce footprint scenarios on Alzheimer's and dementia causes to perform a "what-if" analysis on longevity and mortality business blocks and to assess diversification impact between the two business lines. We compare the approach containing dependency structure via Archimedian copula framework with the approach assuming independence between different causes.

4.1 Scenario overview

Footprint scenarios enable setting hypothetical or historical events in today's perspective to achieve a "what-if" type of analysis and assess impacts on multiple lines of business and company's resistance to shocks. By assuming different levels of progress in prevention measures and medical innovations as well as their developments in the future, we can directly study how this impacts the life expectancy, compared to a benchmark scenario.

The working group at SCOR derived two distinct scenarios: one considering slowdown of the Alzheimer's and dementia symptoms and another considering complete disease elimination in rather near future. Both of the scenarios are results of combined progress in prevention measures, innovations in risk detection and treatments, see Buffet et al. (2022).

Majority of the diagnoses are made to patients late in life once the first signs of disease start appearing, like memory losses. Recent discovery by Swaddiwudhipong et al. (2022) shows that as early as nine years before patients are diagnosed for one of dementia-related diseases, it could be possible to detect signs of brain impairment in the patient. Early detection alone will not result in better patient outcome but is expected to result in better care for persons at higher risk for the disease and help select those who would benefit from treatment to decrease the risk of developing one of the conditions.

Neuroimaging, applications of deep learning and other AI methods are expected to speed up risk identification. In addition, genetic profiling, identification of new

biomarkers and improving identification of functional and cognitive performance will help diagnose the disease in its earliest stages.

Delaying the onset of the disease combined with slowing down the progression once diagnosed is another important axis of possible breakthrough. These could be achieved through progress in prevention measures such as interventions enhancing or maintaining cognitive reserve and targeting modifiable risk factors for dementia.

And eventually, to modify the long-term trajectory of disease, therefore to decrease the progression rate of the disease, progress could be achieved through advances in treatments such as anti-neuroinflammatory drugs, antioxidants, stem cell therapies and drugs repositioning and re-purposing. These future possible treatments would target the disease in its earliest stages before irreversible brain damage or mental decline occurred.

The two hypothetical footprint scenarios for the Alzheimer's and dementia disease future developments are:

- a reduction in Alzheimer's and dementia mortality due to success in delaying onset and slowing deterioration
- an elimination of Alzheimer's and dementia as a cause of loss of autonomy and mortality.

First of the two scenarios delays the onset of moderate to severe symptoms while the second one assumes a modification of the long-term disease trajectory up to risk elimination. One needs to have in mind that in these scenarios do not correspond to a best estimate vision developed by the medical experts. Both are very positive views of possible future outcome with respect to health impact ans the second scenario is particularly so.

4.2 Scenario 1

Figure 45 illustrates the disease evolution for the first of the two scenarios where mortality decline is acquired through delaying onset of Alzheimer's and dementia conditions. On x-axis we represent individual lifetime. On y-axis the disease progression is represented, with two thresholds: first one is the diagnosis and the second one is death.

Today, once a person is diagnosed with the disease, progressive onset of symptoms is expected up to his or her death. In Scenario 1, due to early risk identification followed by reduced risk, disease onset is delayed during the pre-clinical stage of the disease providing postponed disease settling. Additionally, individuals with higher risk to the disease can be identified as a result of developed detection techniques which allowed proposing them guided treatment. After diagnosis, treatment decreases the rate of

progression of the disease. The difference between the time at death as observed today and expected by Scenario 1 results in a gain in life expectancy.

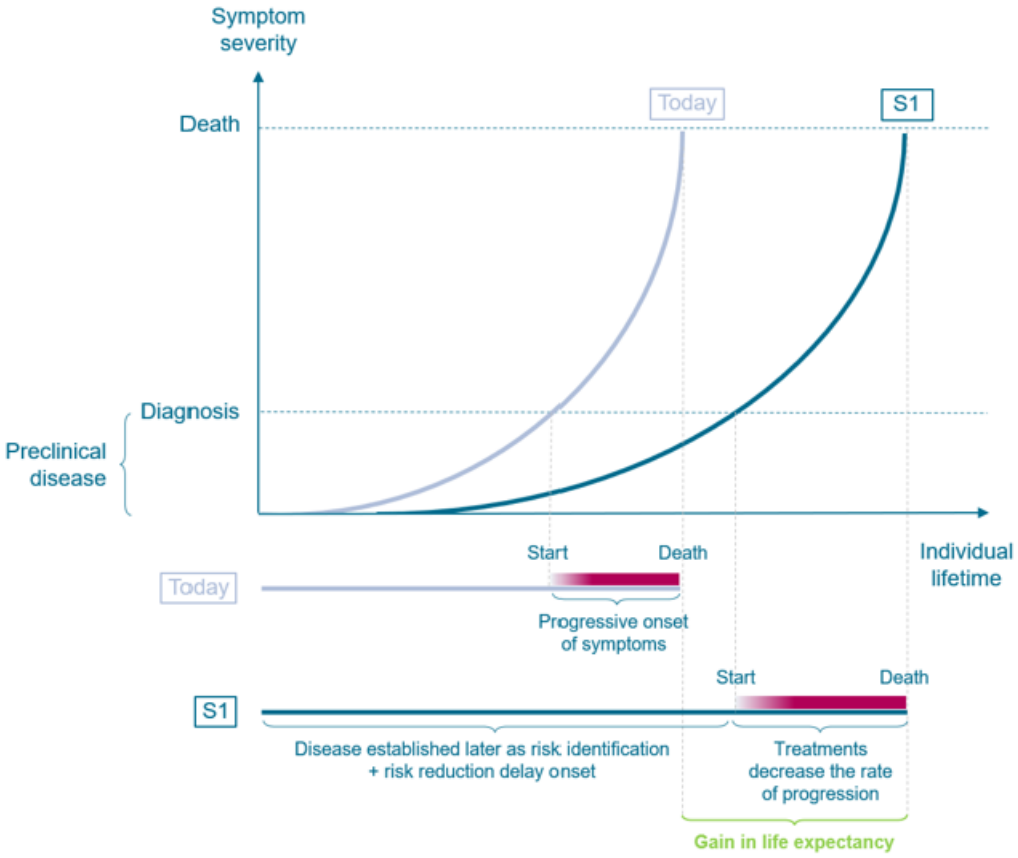

Figure 45: Illustration of Scenario 1

Scenario 1 is implemented assuming that advances in risk detection and prevention measures spread over the next 15 years. Furthermore, treatments decreasing the rate of progression of the disease lead to a reduction of Alzheimer's and dementia mortality by 66% by 2035, as 2019 is the last year of observation in the data and 2020 is the first year projected. After this 15-year horizon of improvement, the age specific probability of death due to Alzheimer's and dementia diseases is assumed to remain at 33% of its pre-scenario projection.

We apply reduction of the Alzheimer's and Dementia mortality on the future net mortality intensity projections to attain 33% of the projection in 2035. The scenario is applied first under independence assumption between different causes framework and then compared to the approach where dependency is accounted for using Archimedian copula structure. We provide visual result illustrations of different approaches by presenting impacts for an 80-year old U.S. male mortality as the Alzheimer's and dementia mortality becomes important cause of death at relatively high ages.

4.2.1 Independence framework

Under the independence assumption between causes, crude and net mortality intensities are equal. Due to this independence assumption between causes, specific scenario on Alzheimer's and dementia will not affect other cause projections and so final impact on life expectancy will come solely from this specific cause mortality rate reduction.

Figure 46 illustrates impact on crude (equal to net in this case) mortality intensities of applying the 66% reduction of Alzheimer's and dementia mortality intensity reached by year 2035 for an 80 year-old male. Cause-specific period of calibration is identical to the one used in the previous chapter, see Table 8.

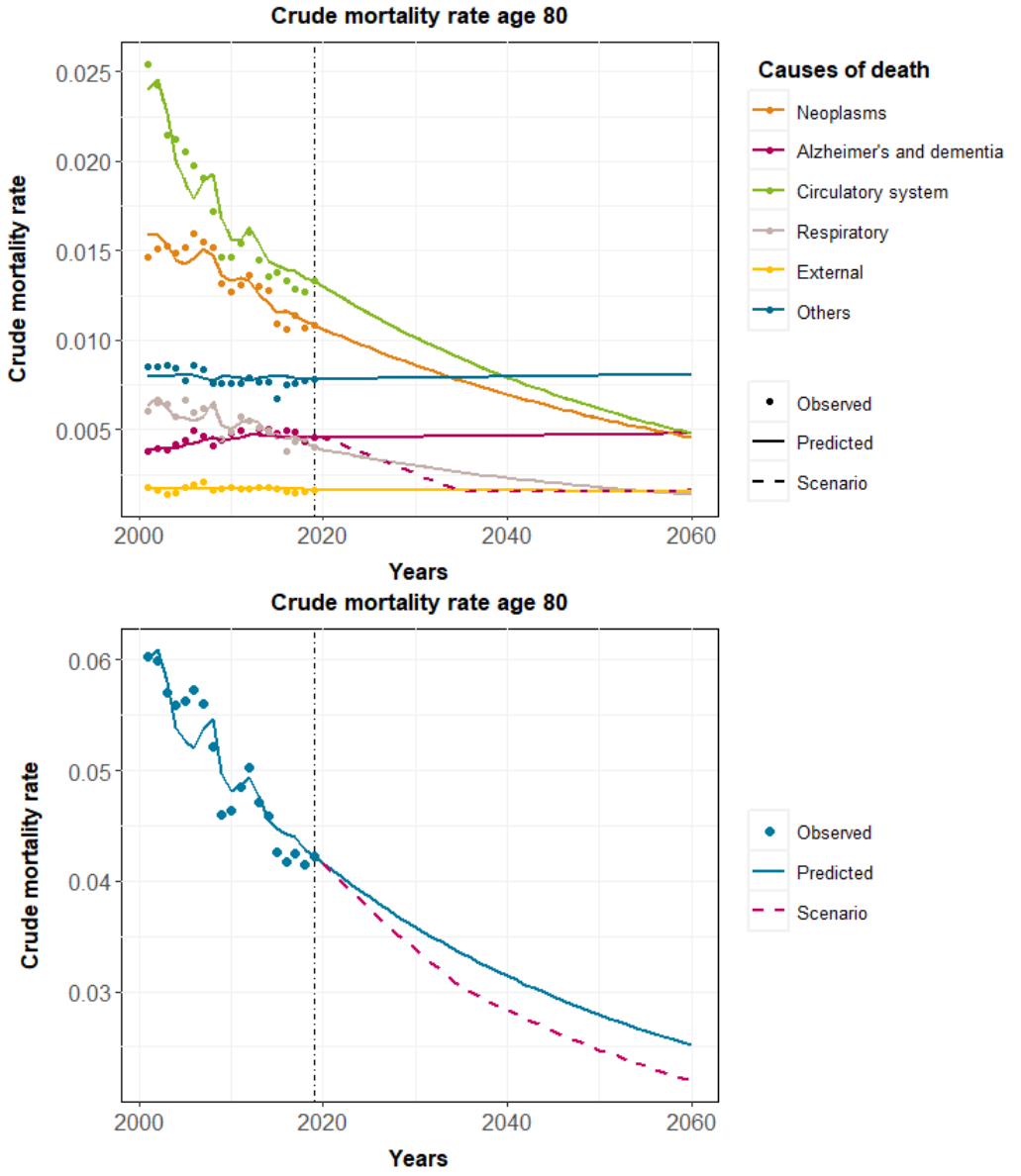

Figure 46: Scenario 1 impact on 80-year old male cause-specific (top) and aggregate (bottom) mortality intensity under independence assumption

The impact from decreasing Alzheimer's and dementia mortality by 66% over the next 15 years has visible impact due to this cause's importance at age 75, on both cause-specific and aggregate mortality. Due to assumption of independence between causes, there is no offset in mortality due to causes other than Alzheimer's and dementia, which will not be the case in the dependency framework.

The following Table 11 depicts gains in residual periodic life expectancy for ages 55 and 75. Projected gains in life expectancy are similar for both ages in 20 and 40 years - almost 6 months over the first 20 years of projection and a bit over 7 months over the whole projection period. There are very little gains between ages 55 and 75 as there are very few deaths due to Alzheimer's and dementia in younger ages, even if the relative gains are more important at age 75.

Residual life expectancy				
Age	Case	Year 2019	Year 2040	Year 2060
55	Central scenario	29	30.9	32.1
	Scenario 1	29	31.4	32.7
	Δ (months)	-	5.9	7.3
75	Central scenario	11.8	13	13.9
	Scenario 1	11.8	13.5	14.5
	Δ (months)	-	5.7	7

Table 11: Scenario 1 periodic life expectancy under independent cause assumption

4.2.2 Copula-based dependence framework

To account for dependency between causes of death in Scenario 1, as in the previous chapter, we use Clayton's copula with $\theta = 1$ and $\theta = 4$.

Dependency structure between causes in the scenario context results in modifying projections of causes other than Alzheimer's and dementia, i.e. individuals that are saved from dying of Alzheimer's and dementia have higher chances to die from other causes, proceeding by increase in forecasted mortality intensity for these alternative causes. The proportion of deaths transferred depends on the dependency parameter θ - the larger the parameter value, the more transfer between causes is observed.

Figures 47 and 48 illustrate Scenario 1 impact on crude mortality rates on cause-specific and aggregate mortality levels for the two parameter θ values. We observe little transfer to other causes for $\theta = 1$ compared to $\theta = 4$ case. Also, $\theta = 4$ case for age 80 exhibit unreasonable patterns for both historical fit and future projections even producing increase in aggregate mortality. Indeed, imposing strong dependency between causes twists the mortality patterns for all causes. Visual inspection can be helpful to refine the set of reasonable values of the dependence parameter θ.

Table 12 presents residual periodic life expectancy gains for the Scenario 1. The life expectancy gains for both ages 55 and 75 are lower than projected under the independence framework due to introduced transfer between causes. Moving from $\theta = 1$ to $\theta = 4$ deteriorates projected gains in life expectancy by more than two times by year 2040, and by the end of the projection period the gap becomes even higher.

Residual life expectancy					
Age	Dependency	Case	Year 2019	Year 2040	Year 2060
55	$\theta = 1$	Central scenario	29	30.1	30.7
		Scenario 1	29	30.6	31.2
		Δ	-	5	6.4
	$\theta = 4$	Central scenario	29	29.3	29
		Scenario 1	29	29.5	29.1
		Δ	-	2.2	1.7
75	$\theta = 1$	Central scenario	11.8	12.3	12.5
		Scenario 1	11.8	12.7	13
		Δ	-	4.7	6
	$\theta = 4$	Central scenario	11.8	11.6	10.9
		Scenario 1	11.8	11.7	11
		Δ	-	1.7	1.1

Table 12: Scenario 1 residual periodic life expectancy for $\theta = 1$ and $\theta = 4$

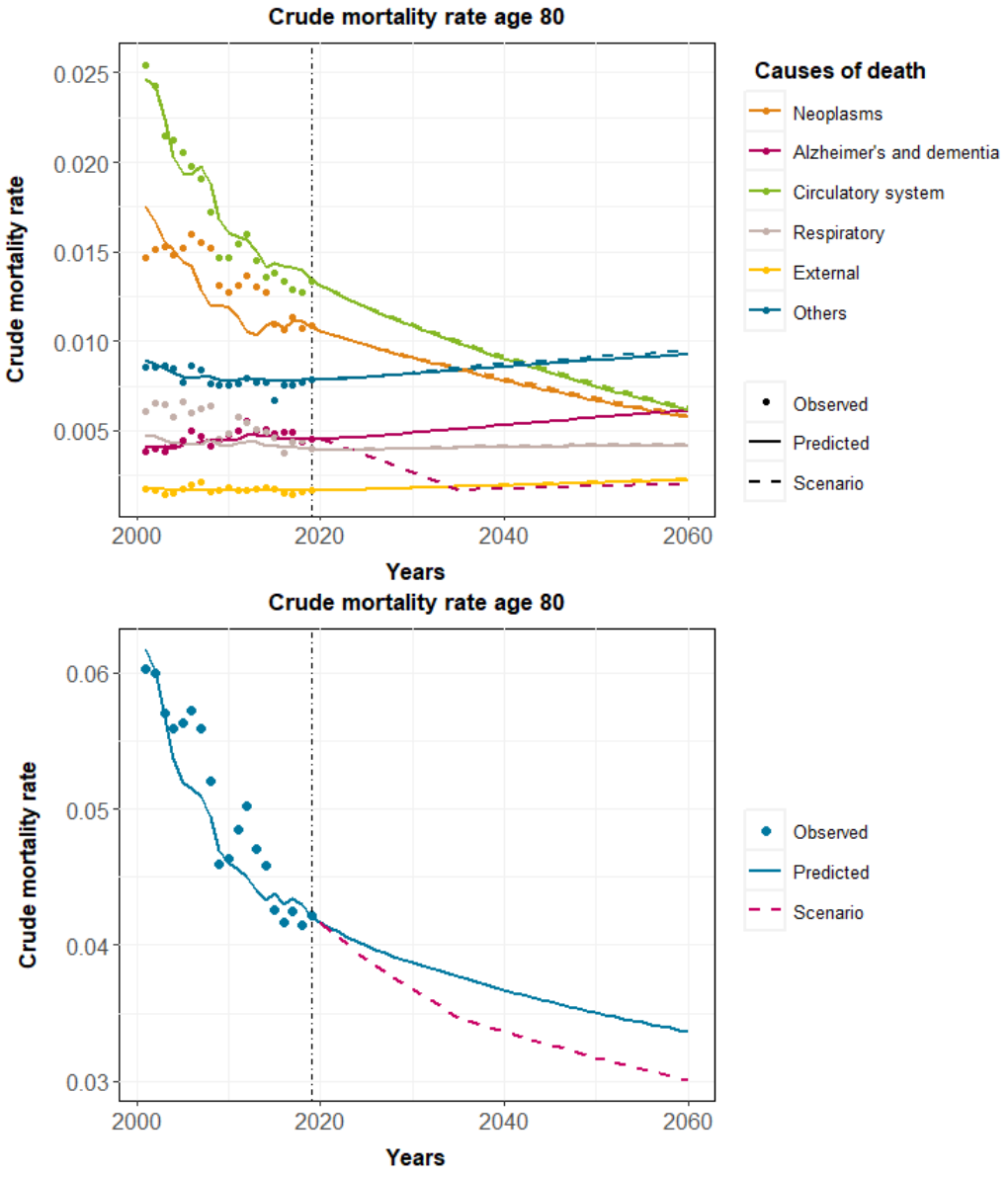

Figure 47: Scenario 1 impact on 80-year old male cause-specific (top) and aggregate (bottom) mortality intensity, Clayton copula $\theta = 1$

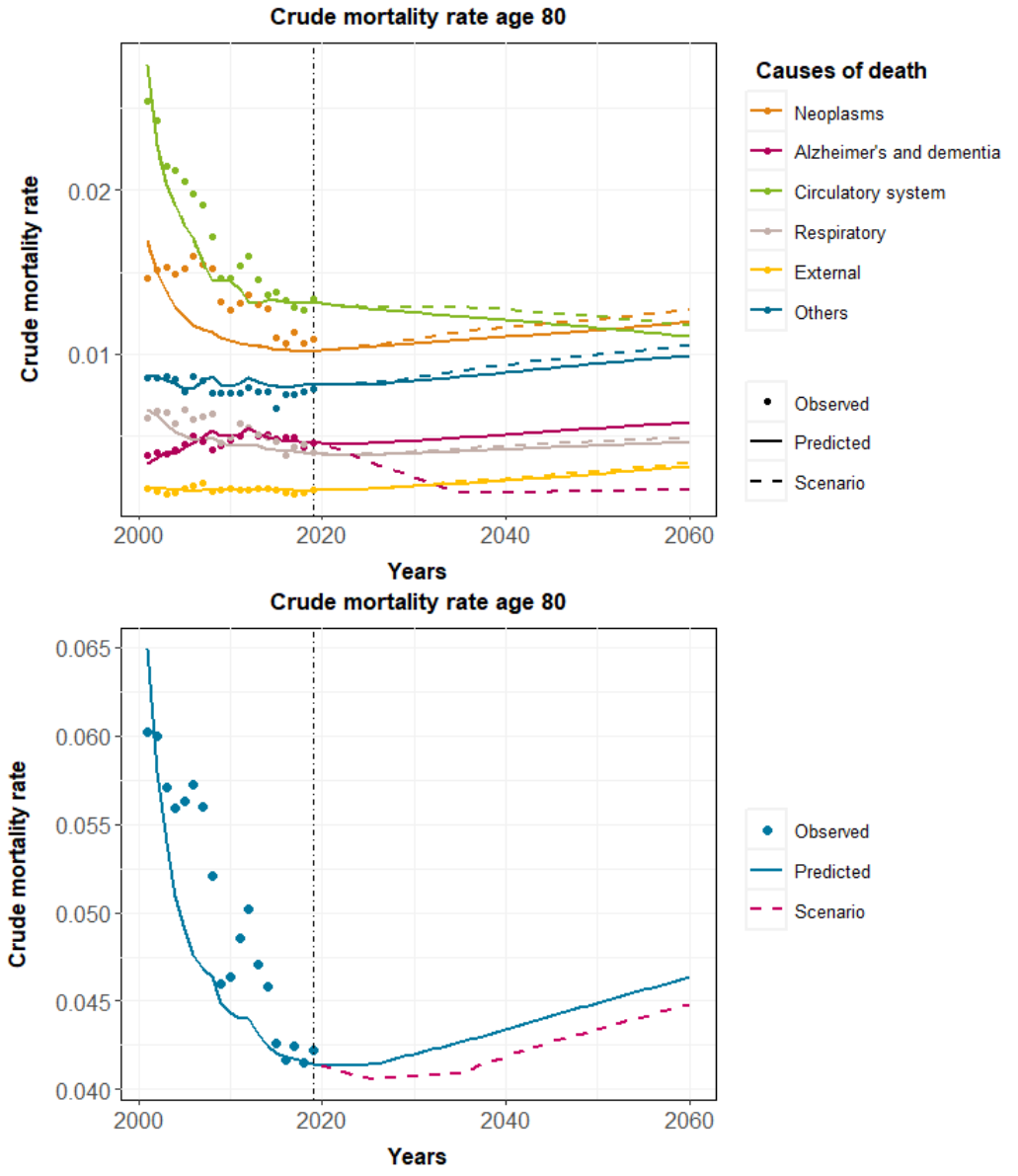

Figure 48: Scenario 1 impact on 80-year old male cause-specific (top) and aggregate (bottom) mortality intensity, Clayton copula $\theta = 4$

4.3 Scenario 2

The second scenario on the Alzheimer's and dementia disease is more extreme in both short timeline for the medical breakthrough to be put in place and the significant impact the scenario proposed - we assume an elimination of mortality and loss of autonomy from Alzheimer's and dementia diseases within the next 5 years. The scenario applies to new cases and individuals in pre-clinical stage as well as individuals already diagnosed with Alzheimer's and dementia diseases.

In the scenario setting, the incidence rates of new cases drop to zero in the near future due to progress in risk detection, disease onset delay and modified trajectory of the condition. This important combined effect is achieved through efficient individual-at-risk identification followed by guided treatment, important risk reduction and prevention measures and improved treatment.

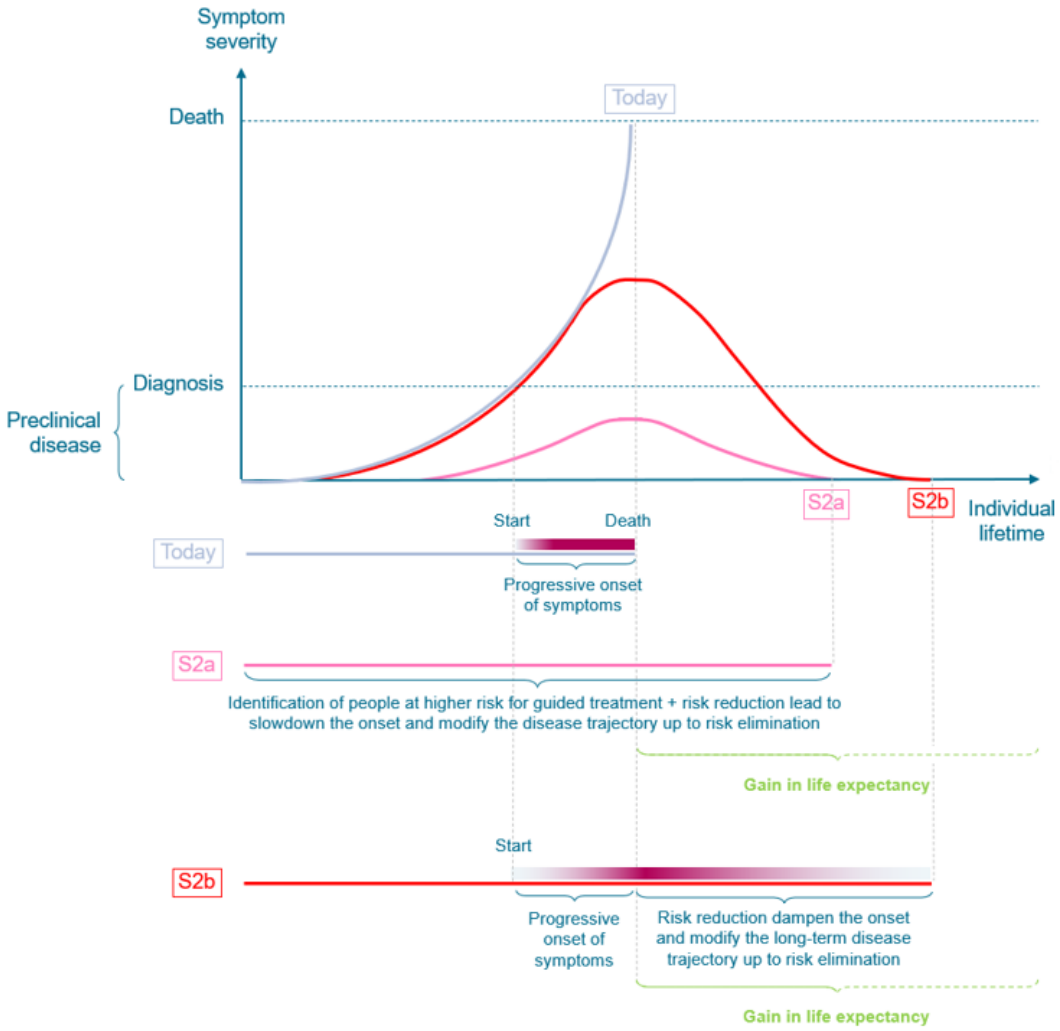

Figure 49: Illustration of Scenario 2

Scenario 2 for new cases and individuals in pre-clinical phase of the disease is il-

lustrated by scenario S2a in Figure 49. The scenario pattern for individuals already diagnosed with the condition is illustrated in scenario S2b of Figure 49 where treatments target the disease before irreversible brain damage or mental decline occurs allowing a modification of the long-term trajectory of the disease up to risk elimination.

Scenario 2 assumes complete elimination of the disease as cause of death over the next 5 years. The scenario proposed is highly hypothetical and is surely situated in the tail of the probable scenario distribution.

4.3.1 Independence framework

Similarly to the Scenario 1, under the independence assumption between causes, crude and net mortality intensities are equal. Due to this independence assumption between causes, specific scenario on Alzheimer's and dementia will not affect other cause projections and so final impact on life expectancy will come solely from this specific cause mortality rate reduction.

Figure 50 depicts impacts on cause-specific and aggregate mortality from Scenario 2. The decrease in mortality rates is significantly more sudden and drastic than in Scenario 1. The gains in residual periodic life expectancy are displayed in the Table 13. Comparing the two scenarios under independence assumption (see Table 11), we observe the second scenario providing approximately 3 additional months by 2040 and 4 months over the whole projection window of 40 years of life expectancy for both ages.

Residual life expectancy				
Age	Case	Year 2019	Year 2040	Year 2060
55	Central scenario	29	30.9	32.1
	Scenario 2	29	31.6	33
	Δ	-	9	11.2
75	Central scenario	11.8	13.1	13.9
	Scenario 2	11.8	13.8	14.8
	Δ	-	8.7	10.7

Table 13: Scenario 2 periodic life expectancy under independent cause assumption

4.3.2 Copula-based dependence framework

Dependency structure between causes in the scenario context results in modifying projections of causes other than Alzheimer's and dementia, i.e. individuals that are saved from dying of Alzheimer's and dementia have higher chances to die from other causes, proceeding by increase in forecasted mortality intensity for these alternative causes. The proportion of deaths transferred depends on the parameter θ - the larger the parameter value, the more transfer between causes is observed.

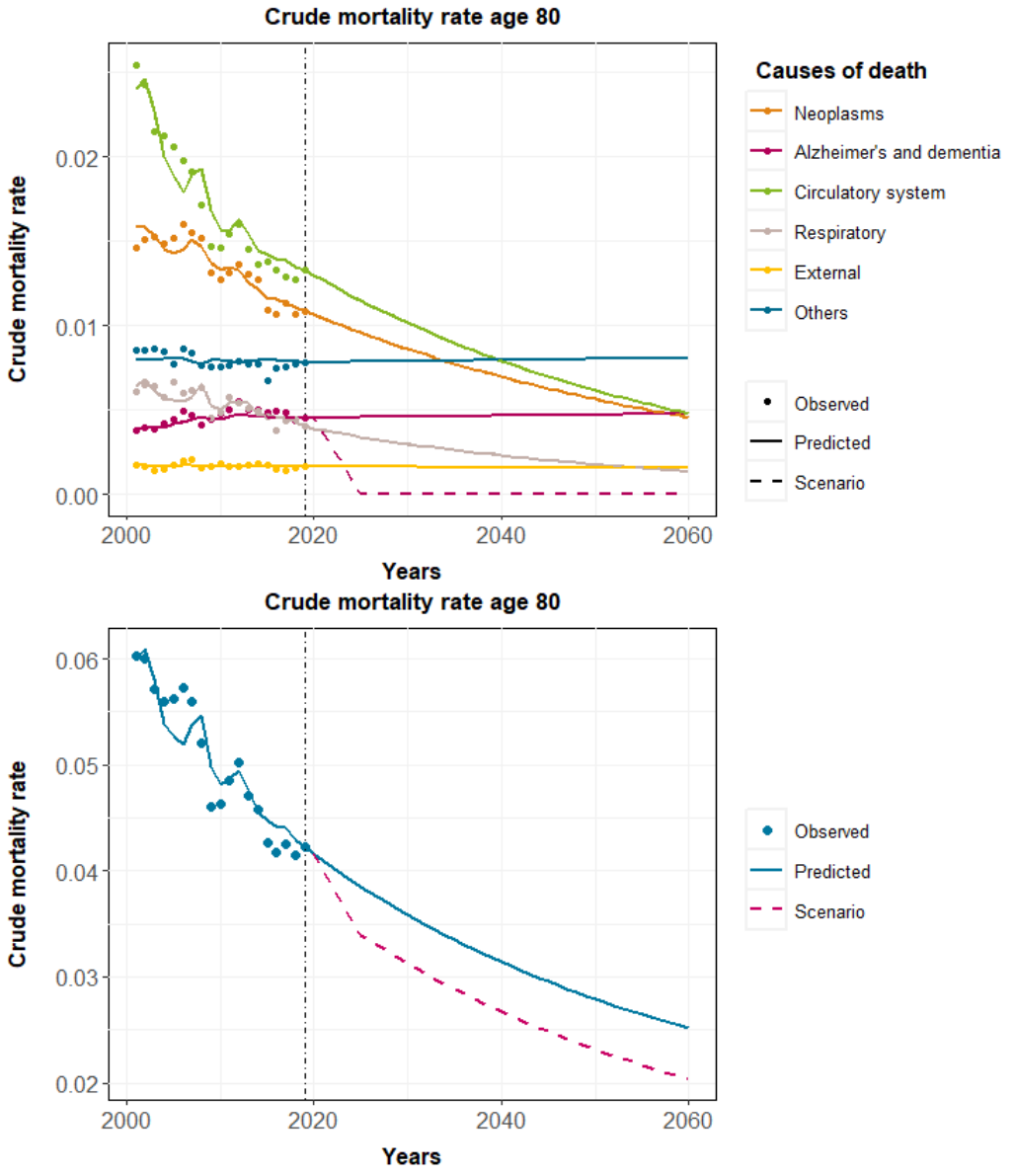

Figure 50: Scenario 2 impact on 80-year old male cause-specific (top) and aggregate (bottom) mortality intensity under independence assumption

Figures 51 and 52 illustrate Scenario 2 impact on crude mortality rates on cause-specific and aggregate mortality levels for the two parameter θ values. Again like in case of Scenario 1, we observe little transfer to other causes for $\theta = 1$ compared to $\theta = 4$ case.

Table 14 presents residual periodic life expectancy gains for the Scenario 2. The life expectancy gains for both ages 55 and 75 are lower than projected under the independence framework and the gains are substantially lower for higher θ value.

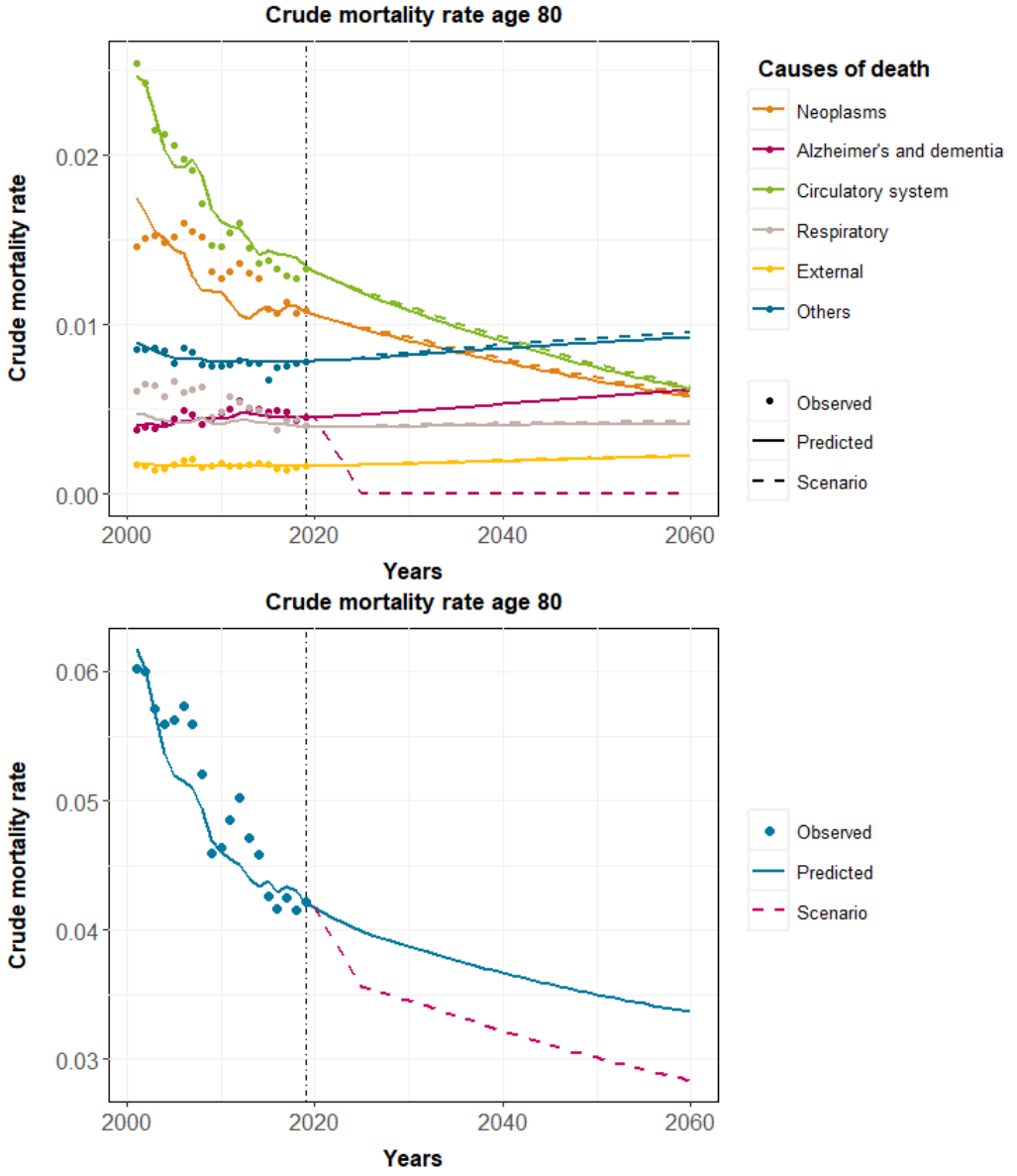

Figure 51: Scenario 2 impact on 80-year old male cause-specific (top) and aggregate (bottom) mortality intensity, Clayton copula $\theta = 1$

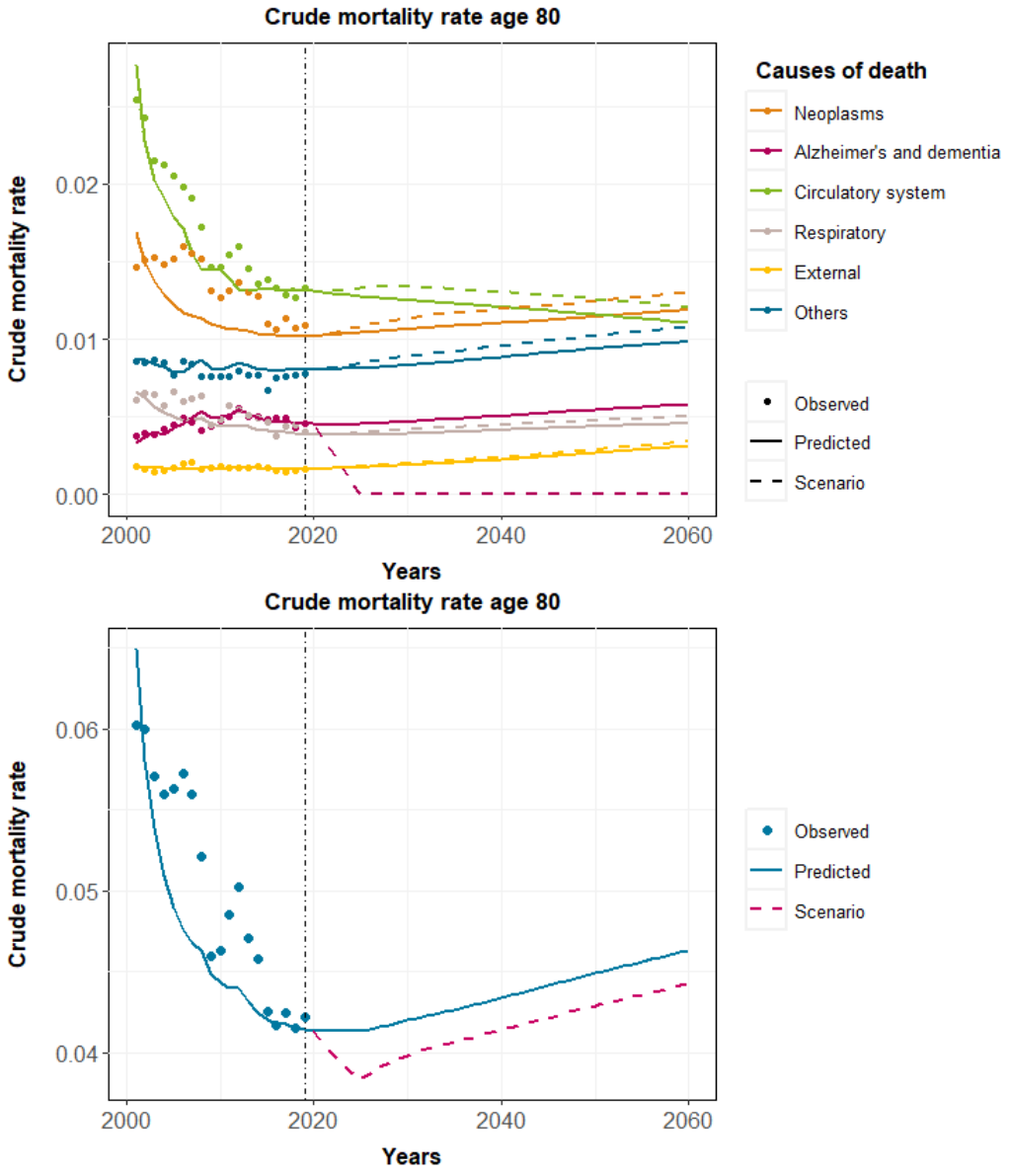

Figure 52: Scenario 1 impact on 80-year old male cause-specific (top) and aggregate (bottom) mortality intensity, Clayton copula $\theta = 4$

Residual life expectancy					
Age	Dependency	Case	Year 2019	Year 2040	Year 2060
55	$\theta = 1$	Central scenario	29	30.1	30.7
		Scenario 2	29	30.8	31.5
		Δ	-	7.5	9.6
	$\theta = 4$	Central scenario	29	29.3	29
		Scenario 2	29	29.6	29.2
		Δ	-	2.9	2.5
75	$\theta = 1$	Central scenario	11.8	12.3	12.5
		Scenario 2	11.8	12.9	13.3
		Δ	-	7.1	9
	$\theta = 4$	Central scenario	11.8	11.6	10.9
		Scenario 2	11.8	11.8	11
		Δ	-	2.2	1.6

Table 14: Scenario 2 residual periodic life expectancy for $\theta = 1$ and $\theta = 4$

4.4 Additional assumptions affecting the results

The two scenarios on Alzheimer's and dementia cause mortality and the time frame proposed for the scenario development were set by discussions with medical experts internally at SCOR. In addition to theses scenario assumptions, there are other elements that have influence on the final results.

- The within-cohort dependence among the causes of death in the copula framework is another parameter set by expert judgement. The greater the parameter used, the more transfer to other causes is made resulting in larger offset in aggregate mortality due to increased mortality of causes othet than Alzheimer's and dementia.

- The central scenario of Alzheimer's and dementia mortality forecast at high ages is also influencing the outcome. By choosing specific period to model every cause, we first look for visually reasonable future mortality projection, nevertheless it gives the modeller control over the projection results. In a general manner, the larger the projected mortality, the higher the number of deaths saved from dying of Alzheimer's and dementia, and the larger the potential impact of an improvement scenario.

- The central scenario mortality projection of causes other than Alzheimer's and dementia at high ages also affects the result. If other causes, like Neoplasms of Circulatory system diseases, have very high projected mortality, individuals would die of neoplasms or cardiovascular diseases shortly after being saved from Alzheimer's and dementia.

- And eventually, results are affected by mortality shape of at very high ages, i.e. the completion assumption of the mortality table. The method and age

limit chosen will influence the outcome as it defines the survival time of individuals saved from dying of Alzheimer's and dementia, though the effect of this assumption is minimal once comparing different approaches with the same mortality table completion method.

4.5 Financial impact

We observed how scenarios of different extremity have contrasting impacts on life expectancy. Both of the scenarios are very positive views on future mortality and so result in higher life expectancy than what was projected under the same dependency assumption central scenario.

As discussed in sections Mortality and Longevity, Mortality and Longevity risks are directly opposite from the insurer's perspective - the former is related to policy holders dying earlier than expected and the latter is related to annuitants living longer than expected and the insurer paying pensions beyond the period predicted.

Nevertheless, SCOR's Mortality and Longevity business lines do not entirely offset one another. First, age mix in the two lines of business is different - Mortality business covers mainly working population with of average age of 55, while Longevity portfolio consists of the annuitants in payment with average age of 75. Additionally, Mortality business in concentrated in the U.S. whereas Longevity covered by SCOR is situated in the U.K. And finally, size of the portfolio plays important role - for optimal diversification effect the two lines of business should be of similar size by Present Value (PV) of Claims or other metric.

To illustrate diversification repercussion on the two scenarios in a simplified way, we neglect the geographical placing and portfolio size characteristics and use only average age to differentiate between the two risks. Considering the Mortality business, changes in mortality rate assumptions have impact on outgoing claim payments at the same time affecting the premium rates. For our illustration we look uniquely to impact on claim amounts while ignoring the probable premium decrease, as the premium decrease would be marginal compared to changes in claim amounts.

Both Longevity and Mortality claims are calculated using one representative model point for each line of business, with fixed interest rate of 1.5% over the whole 40 year projection period. The Mortality claim calculations takes into account lapse assumption observed in SCOR's portfolio for a 10-year duration. Longevity pensions are assumed not to be indexed which is also a slight simplification of the reality.

For the three different dependency frameworks presented in this thesis, the framework-specific Central scenario is compared to Scenario 1 and Scenario 2 cumulative impacts - gains from Mortality book minus the losses on the Longevity side. We set sum insured for Mortality and pension amount for Longevity so that they offset one another

entirely in Central scenario and are equal to 2000.

Independent causes ($\theta = 0$)			
Business	Central scenario	Δ Scenario 1	Δ Scenario 2
(1) Mortality claims	2000	-53	-83
(2) Longevity claims	2000	+33	+82
Total (1)-(2)	0	-20	-1

Clayton's copula, $\theta = 1$			
Business	Central scenario	Δ Scenario 1	Δ Scenario 2
(1) Mortality claims	2000	-37	-57
(2) Longevity claims	2000	+28	+68
Total (1)-(2)	0	-9	+9

Clayton's copula, $\theta = 4$			
Business	Central scenario	Δ Scenario 1	Δ Scenario 2
(1) Mortality claims	2000	-10	-14
(2) Longevity claims	2000	+14	+30
Total (1)-(2)	0	+4	+16

Table 15: Impact on diversification for different θ values

Assuming independence between causes results in biggest variation of claims for Mortality and Longevity business lines, which was to be expected as the life expectancy gains were the highest under the independence framework, while assuming high dependency between causes yields in moderate fluctuations. Both scenarios under independence assumption produce drop in total claims, $\theta = 1$ yields to small increase in claims for Scenario 1 and drop in Scenario 2, and lastly, considering high dependency generates higher claim amounts for both of the scenarios.

We observe how not taking into account dependency structure between causes when modelling scenarios could yield in hypothetical gains when working in competing risk framework would result in more moderate gains from diversification from mortality and longevity business lines. Assuming causes being independent both scenarios results in lower claims than in central scenario, showing high diversification between Mortality and Longevity business lines, especially for Scenario 1. Moving to low dependency case ($\theta = 1$), the reduction in total claims is twice smaller and we observe and increase in claims for Scenario 2 instead of small gains (+9 vs -1). Hypothesis of strong dependency between causes has even stronger impact and for both scenarios we observe increase in claims.

Financial scenario impacts are notably different depending on the dependency assumption used when modelling mortality at granular cause-specific level. Conclusion variety concerning diversification effects and business resilience illustrates importance of working under competing risk framework.

Conclusion

This thesis investigates mortality modelling by cause of death under several dependency frameworks - assuming independence between causes and introducing dependency structure through Archimedian copula configuration. We compare the approaches adopting central scenario and footprint scenarios on Alzheimer's and dementia cause application basis. Results vary significantly in terms of both residual gains in life expectancy and financial impact between the proposed methods.

On one hand, by assuming independence structure between causes, we do not take into account competing risks framework. The method projects each cause of death independently with no allowance for transfers of death between causes. On the other hand, the alternative method introduced tries to capture dependency between the competing causes using copula model. We use Clayton's copula with several dependence parameters to illustrate importance of allowing for dependency between competing causes.

Application on U.S. male insured population proxy mortality data shows how dependency assumption affects residual life expectancy - the aggregate mortality projected is the most "optimistic" when assuming no dependency between causes, while high dependent parameter values can result in mortality rate increase for some ages. Visual projection inspection permits to establish reasonable values for the dependency parameter θ.

The footprint scenarios introduced propose several hypothetical future developments of the Alzheimer's and dementia disease and illustrate how cause dependence structure alters results in scenario framework. Using model points to represent mortality and longevity business lines, we observe how not taking into account dependency structure between causes when modelling scenarios could yield in hypothetical gains. On the other hand, working in competing risk framework would result in more moderate gains from diversification from mortality and longevity business lines.

While the introduced method is an innovative approach in cause of death modelling, it presents several limitations that need to be disclosed.

First of all, the proposed structure imposes unique strength of dependency between causes. We understand that Neoplasms and Circulatory system disease causes share multiple risk factors and strong transfer of deaths between the two causes can be assumed. On the other hand, External causes that regroup accidents and similar smaller causes, share less risk factors with other causes. Provided that, adjustment

in the dependency structure could enhance the method by introducing hierarchical structure between causes, allowing for grouping causes having stronger correlations vs. causes having lower correlations. The method enhancement is already proposed by Li and Lu (2019).

Additionally, cause-specific mortality trends often turn out to be more volatile over-time than aggregate mortality trends. Using Lee-Carter forecasting model for all causes may prove to be over-simplification and more complicated modelling could be introduced to account for dynamic differences. State-space time series models offer a natural extension of the Lee-Carter model both in terms of the alternative coherent fitting procedure and the possibility to flex the model by introducing additional parameters to ensure a better fit, see Piveteau (2021) or Gylys (2021).

Modelling mortality by cause of death requires in-depth analysis and additional assumptions compared to all-cause mortality models. However, the approach proves to be an advantageous tool to obtain valuable observations on mortality evolution, dissect aggregate mortality trends to cause-specific drivers and perform a "what-if" scenario analysis.

References

ACS. Census update estimation data, accessed in march 2022. 2020. URL https://www.census.gov/programs-surveys/acs/news/data-releases.2020.html.

J. Aitchison. The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2):139–177, 1982.

D. H. Alai, S. Arnold, B. Madhavi, and A. Villegas. Mind the gap: A study of cause-specific mortality by socioeconomic circumstances. *North American Actuarial Journal*, 22(2):161–181, 2018.

Alzheimer's Association. Alzheimer's disease facts and figures. *Alzheimer's Dement.*, 16:391–460, 2020. URL https://doi.org/10.1002/alz.12068.

Alzheimer's Association. Alzheimer's dementia: Global resources. 2021a. URL https://alz.org/global/.

Alzheimer's Association. Alzheimer's disease facts and figures. 17(3):1–108, 2021b.

Alzheimer's Association. Fiscal Year 2021 Alzheimer's Research Funding. *Chicago, IL*, 2021c.

Alzheimer's Association. Fiscal Year 2023 Alzheimer's Research Funding. *Chicago, IL*, 2022.

Alzheimer's International. World Alzheimer Report 2018.

S. Arnold. Challenges with cause-of-death mortality data. *2022 HMD Conference*, 2022.

F. Berkum, K. Antonio, and M. Vellekoop. The impact of multiple structural changes on mortality predictions. 2014.

D. Bernoulli. Essai d'une nouvelle analyse de la mortalite causee par la petite verole et des avantages de l'inoculation pour la prevenir. *Historie avec les Memoires, Paris: Acad. Sci.*, pages 1–45, 1760.

H. Booth, L. Hyndman, and P. De Jong. Lee-Carter mortality forecasting : A multi-country comparison of variants and extensions. *Demographic Research*, 15:289–310, 2006.

A. Boumezoued, H. Labit Hardy, N. El Karoui, and S. Arnold. Cause-of-death mortality: What can be learned from population dynamics? *Insurance: Mathematics and Economics*, 18:301–315, 2018.

A. Boumezoued, J. B. Coulomb, A. Klein, D. Louvet, and E. Titon. Modeling and forecasting cause-of-death mortality. *Society of actuaries*

(SOA), 2019. URL https://www.soa.org/resources/research-reports/2019/cod-mortality-forecasting/.

G. E. P. Box and G. Jenkins. Time series analysis forecasting and control. *XF2006175558*, 3, 1970.

N. Brouhns, M. Denuit, and J. Vermunt. A poisson log-bilinear regression approach to the construction of projected life tables. *Insurance : Mathematics and Economics*, 31:373–393, 2002.

G. Buffet, M. Fysekidis, J. Tomas, and A. Ulcinaite. Alzheimer's disease and dementia footprint scenario: A research report series volume 1 – introduction. *Assessing Risk for Alzheimer's Disease and Dementia | SCOR*, 1, 2022.

A. Cairns, D. Blake, and K. Dowd. A two-factor model for stochastic mortality with parameter uncertainty : Theory and calibration. *Journal of Risk Insurance*, 73 (4):687–718, 2006a.

A. Cairns, D. Blake, and K. Dowd. Pricing death: Frameworks for the valuation and securitization of mortality risk. *ASTIN Bulletin*, 36:79–120, 2006b.

E. Cannan. The probability of a cessation of the growth of population in England and Wales during the next century. *The Economic Journal*, 5(20):505–515, 1895.

J. Carriere. Dependent decrement theory (with discussion). *Transactions of the Society of Actuaries*, 46:45–74, 1994.

CDC Press release. Life expectancy in the U.S. dropped for the second year in a row in 2021. 2022. URL https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2022/20220831.htm.

Census. Census data. 2010. URL https://www.census.gov/programs-surveys/decennial-census/data/datasets.2010.html.

Centers for Disease Control and Prevention. Death count data by different variables, accessed in march 2022. 2022. URL https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm.

R. Chetty, M. Stepner, and S. Abraham. The association between income and life expectancy in the United States, 2001-2014. *JAMA*, 315(16):1750–1766, 2016.

C. L. Chiang. Introduction to stochastic processes in biostatistics. *John Wiley and Sons, New York, 1st edition*, 1968.

B. K. Couillard, C. L. Foote, K. Gandhi, E. Meara, and J. Skinner. Rising geographic disparities in US mortality. *J Econ Perspect*, 35(4):123—-146, 2021.

H. Cramér and H. Wold. Mortality variations in Sweden: a study in graduation and forecasting. *Scandinavian Actuarial Journal*, pages 161—-241, 1935.

J. Cummings, G. Lee, K. Zhong, J. Fonseca, and K. Taghva. Alzheimer's disease drug development pipeline: 2021. *Translational Research and Clinical Interventions*, 7 (1), 2021. URL `https://doi.org/10.1002/trc2.12179`.

A. DeMoivre. Annuities on lives : Or, the valuation of annuities upon any number of lives as also of reversions. *William Person, London*, 1725.

M. Denuit and A. C. Goderniaux. Closing and projecting lifetables using log-linear models. *Bulletin de l'Association Suisse des Actuaires, sous presse*, 2005.

D. Dimitrova, S. Haberman, and V. Kaishev. Dependent competing risks : Cause elimination and its impact on survival. *Insurance : Mathematics and Economics*, 53:464–477, 2013.

Eisai. Lecanemab confirmatory phase 3 clarity AD study met primary endpoint, showing highly statistically significant reduction of clinical decline in large global clinical study of 1,795 participants with early Alzheimer's disease. *Press release 28 Sept*, 2022. URL `https://www.eisai.com/news/2022/news202271.html`.

L. Gavrilov and N. Gavrilova. Mortality measurement at advanced ages : A study of the social security administration death master file. *North American actuarial journal : NAAJ*, 15:432–447, 2011.

B. Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 115:513–583, 1825.

J. Graunt. Natural and political observations mentioned in a following index, and made upon the bills of mortality. *London: John Martyn and James Allestry*, 1973.

R. Gylys. Application of the Lee-Carter mortality projection model and its modifications in the modelling of an insurance company's solvency capital. *Doctoral thesis, Vilnius University, Faculty of mathematics and informatics*, 2021.

E. Halley. An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives. *Philosophical Transactions of the Royal Society of London*, 17(196):596–610, 1693.

L. Heligman and J. Pollard. The age pattern of mortality. *Journal of the Institute of Actuaries*, 107, 1980.

F. Hollmann, T. Mulder, and J. Kallan. Methodology assumptions for the population projections of the United States: 1999 to 2010. *US Department of Commerce, Bureau of the Census, Population Division, Population Projections Branch*, 1999.

J. Oeppen et al. Coherent forecasting of multiple-decrement life tables: a test using Japanese cause of death data. 2008.

B. D. James, R. S. Wilson, P. A. Boyle, J. Q. Trojanowski, D. A. Bennett, and J. A. Schneider. TDP-43 stage, mixed pathologies, and clinical Alzheimer's-type dementia. *Brain*, 139(11):2983–2993, 2016.

V. Kaishev, D. Dimitrova, and S. Haberman. Modelling the joint distribution of competing risks survival times using copula functions. *Insurance : Mathematics and Economics*, 41:339–361, 2007.

T. Kirkwood. Deciphering death : A commentary on Gompertz (1825) 'On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies'. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 370, 2015.

R. Lee and L. Carter. Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87(419):659–671, 1992.

H. Li and Y. Lu. Modeling cause-of-death mortality using hierarchical Archimedean copula. *Scandinavian Actuarial Journal*, 3:247–272, 2019.

N. Li and R. Lee. Coherent mortality forecasts for a group of population : An extension of the Lee–Carter method. *Demography*, 42:575–594, 2005.

W. M. Makeham. On the law of mortality. *Journal of the Institute of Actuaries (1866-1867)*, 13(6):325–358, 1860.

G. A. Mensah, G. S. Wei, and P. D. Sorlie. Decline in cardiovascular mortality: Possible causes and implications. *Circ Res*, 120(2):366–380, 2017.

National Academies of Sciences, Engineering, and Medicine. High and rising mortality rates among working-age adults. *Washington (DC): National Academies Press (US)*, 2021. URL `https://www.ncbi.nlm.nih.gov/books/NBK571940/`.

OECD. 2018. URL `https://ourworldindata.org/life-expectancy`.

OECD. Pension markets in focus 2021. 2021. URL `https://www.oecd.org/daf/fin/private-pensions/Pension-Markets-in-Focus-2021.pdf`.

S. J. Olshansky, T. Antonucci, L. Berkman, R. Binstock, A. Boersch-Supan, J. Cacioppo, B. Carnes, L. Carstensen, L. Fried, D. Goldman, J. Jackson, M. Kohli, J. Rother, Y. Zheng, , and J. Rowe. Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Affairs*, 31(8):1803–1813, 2012.

S. Piveteau. Modélisation de la mortalité par cause de décès. *Thèse de doctorat en Sciences de gestion, dans le cadre de École doctorale Sciences économiques et gestion (Lyon)*, 2021.

S. Piveteau and J. Tomas. Mortality forecasting by cause of death and basis risk modelling with compositional data. *Presented at the Longevity 14 Conference, Amsterdam*, 2018. URL `https://www.bayes.city.ac.uk/__data/assets/pdf_file/0019/437140/TOMAS-Julian.pdf`.

R. Prentice, J. Kalbfleisch, A. V. Peterson, N. Flournoy, V. Farewell, and N. Breslow. The analysis of failure times in the presence of competing risks. *Biometrics*, 34 (4):541—-554, 1978.

H. Putter, M. Fiocco, and R. Geskus. Tutorial in biostatistics : competing risks and multi-state models. *Statistics in medicine*, 26(11):2389—-2430, 2007.

A. E. Renshaw and S. Haberman. A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38 (3):556–570, 2006.

J. Saboia. Modeling and forecasting populations by time series: The Swedish case. *Demography*, 11(3):483–492, 1974.

P. G. Shekelle, O. S. Aisen, H. J. Federoff, C. E. Finch, F. F. Hefti, J. Kirkland, F. Laferla, A. Motala, J. Hlavka, and D. P. Goldman. Prospects for future advances in Alzheimer's disease. *Leonard D. Schaeffer Center for Health Policy and Economics*, pages 1–5, 2020.

N. Swaddiwudhipong, J. D. Whiteside, F. H. Hezemans, D. Street, J. B. Rowe, and T. Rittman. Pre-diagnostic cognitive and functional impairment in multiple sporadic neurodegenerative diseases. *Alzheimer's Dementia. The journal of the Alzheimer's Association*, 2022.

A. Thatcher, J. Vaupel, and V. Kannisto. The force of mortality at ages 80 to 120. *Odense University Press, Odense, Denmark*, 1998.

A. Tsiatis. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences of the United States of America*, 72(1):20–22, 1975.

U.S. Census Bureau. Education attainement data, accessed in march 2022. 2022. URL `https://www.census.gov/topics/education/educational-attainment.html`.

U.S. National Library of Medicine. ClinicalTrials.gov database. U.S. National Institutes of Health. Accessed December 16, 2021. URL `https://clinicaltrials.gov/ct2/home`.

J. Vaupel and V. C. Romo. Decomposing change in life expectancy: A bouquet of formulas in honor of Nathan Keyfitz's 90th birthday. *Demography*, 40(2):201—-216, 2003.

J. Wilmoth. Are mortality projections always more pessimistic when disaggregated by cause of death ? *Mathematical Population Studies*, 5(4):293–319, 1995.

World Health Organization. Risk reduction of cognitive decline and dementia: Who guidelines. 2020. URL `https://www.who.int/publications/i/item/risk-reduction-of-cognitive-decline-and-dementia`.

A. Zeileis, F. Leisch, K. Hornik, C. Kleiber, B. Hansen, E. Merkle, and N. Umlauf. Package 'strucchange'. 2022. URL `extension://efaidnbmnnnibpcajpcglclefindmkaj/https://cran.r-project.org/web/packages/strucchange/strucchange.pdf`.