



Sciences Economiques et Sociales de la Santé
& Traitement de l'Information Médicale

sesstim.univ-amu.fr

Robert Darlin MBA

*PhD Student Aix Marseille Université,
INSERM, IRD, SESSTIM, Marseille, France*

**Analysis of excess mortality by cancer:
adjusting the lack of additional covariates in life tables**

juin 2018



[Cliquez ici pour voir l'intégralité des ressources associées à ce document](#)

Analyse de la mortalité en excès par cancer :
modélisations corrigeant le manque de variables
supplémentaires dans les tables de mortalité
(Travail en cours)

Robert Darlin Mba¹, Roch Giorgi*²

¹Aix Marseille Univ, INSERM, IRD, SESSTIM, Marseille, France

²Aix Marseille Univ, APHM, INSERM, IRD, SESSTIM, Hop Timone, BioSTIC, Marseille, France

*Directeur de thèse

email : darlin.mba@univ-amu.fr

Survie nette (\neq survie globale ou brute)

Survie nette (\neq survie globale ou brute)

- ➡ Survie qui serait observée si la seule cause de décès possible était la pathologie d'intérêt (e.g., cancer étudié)

Survie nette (\neq survie globale ou brute)

- Survie qui serait observée si la seule cause de décès possible était la pathologie d'intérêt (e.g., cancer étudié)

- Cadre hypothétique

Survie nette (\neq survie globale ou brute)

- ☞ Survie qui serait observée si la seule cause de décès possible était la pathologie d'intérêt (e.g., cancer étudié)
- ☞ Cadre hypothétique
- ☞ Registres de cancer (cause de décès non disponible)

Survie nette (\neq survie globale ou brute)

- ☞ Survie qui serait observée si la seule cause de décès possible était la pathologie d'intérêt (e.g., cancer étudié)
- ☞ Cadre hypothétique
- ☞ Registres de cancer (cause de décès non disponible)
- ☞ Affranchissement des différences de mortalité dues aux autres causes (comparaisons notamment entre pays et périodes)

Survie nette (\neq survie globale ou brute)

- ☞ Survie qui serait observée si la seule cause de décès possible était la pathologie d'intérêt (e.g., cancer étudié)
- ☞ Cadre hypothétique
- ☞ Registres de cancer (cause de décès non disponible)
- ☞ Affranchissement des différences de mortalité dues aux autres causes (comparaisons notamment entre pays et périodes)
- ☞ Son estimation repose sur l'utilisation des tables de mortalité (approche paramétrique ou non paramétrique)

Table de mortalité

Hommes	
Age	Prob. de décès
⋮	⋮
30-31	0.0020520
31-32	0.0021522
32-33	0.0022424
33-34	0.0023426
34-35	0.0024529
⋮	⋮
85-86	0.1321354
86-87	0.1441446
87-88	0.1573526
88-89	0.1716155
89-90	0.1870248
⋮	⋮

Femmes	
Age	Prob. de décès
⋮	⋮
30-31	0.0007502
31-32	0.0008003
32-33	0.0008503
33-34	0.0009003
34-35	0.0009504
⋮	⋮
85-86	0.0882394
86-87	0.0984490
87-88	0.1095782
88-89	0.1215099
89-90	0.1345577
⋮	⋮

Table de mortalité

Hommes	
Age	Prob. de décès
⋮	⋮
30-31	0.0020520
31-32	0.0021522
32-33	0.0022424
33-34	0.0023426
34-35	0.0024529
⋮	⋮
85-86	0.1321354
86-87	0.1441446
87-88	0.1573526
88-89	0.1716155
89-90	0.1870248
⋮	⋮

Femmes	
Age	Prob. de décès
⋮	⋮
30-31	0.0007502
31-32	0.0008003
32-33	0.0008503
33-34	0.0009003
34-35	0.0009504
⋮	⋮
85-86	0.0882394
86-87	0.0984490
87-88	0.1095782
88-89	0.1215099
89-90	0.1345577
⋮	⋮

Table de mortalité

Hommes	
Age	Prob. de décès
⋮	⋮
30-31	0.0020520
31-32	0.0021522
32-33	0.0022424
33-34	0.0023426
34-35	0.0024529
⋮	⋮
85-86	0.1321354
86-87	0.1441446
87-88	0.1573526
88-89	0.1716155
89-90	0.1870248
⋮	⋮

Femmes	
Age	Prob. de décès
⋮	⋮
30-31	0.0007502
31-32	0.0008003
32-33	0.0008503
33-34	0.0009003
34-35	0.0009504
⋮	⋮
85-86	0.0882394
86-87	0.0984490
87-88	0.1095782
88-89	0.1215099
89-90	0.1345577
⋮	⋮

Table de mortalité

Hommes	
Age	Prob. de décès
⋮	⋮
30-31	0.0020520
31-32	0.0021522
32-33	0.0022424
33-34	0.0023426
34-35	0.0024529
⋮	⋮
85-86	0.1321354
86-87	0.1441446
87-88	0.1573526
88-89	0.1716155
89-90	0.1870248
⋮	⋮

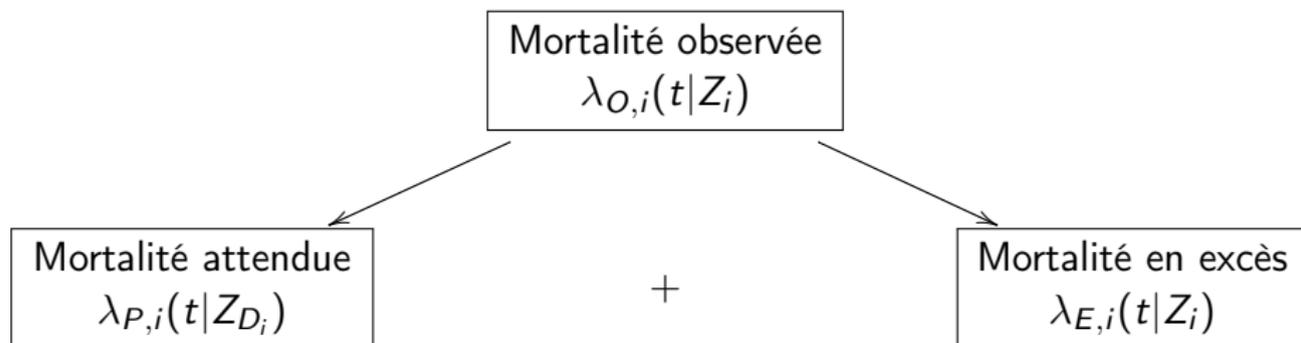
Femmes	
Age	Prob. de décès
⋮	⋮
30-31	0.0007502
31-32	0.0008003
32-33	0.0008503
33-34	0.0009003
34-35	0.0009504
⋮	⋮
85-86	0.0882394
86-87	0.0984490
87-88	0.1095782
88-89	0.1215099
89-90	0.1345577
⋮	⋮

Table de mortalité

Hommes		Femmes	
Age	Prob. de décès	Age	Prob. de décès
⋮	⋮	⋮	⋮
30-31	0.0020520	30-31	0.0007502
31-32	0.0021522	31-32	0.0008003
32-33	0.0022424	32-33	0.0008503
33-34	0.0023426	33-34	0.0009003
34-35	0.0024529	34-35	0.0009504
⋮	⋮	⋮	⋮
85-86	0.1321354	85-86	0.0882394
86-87	0.1441446	86-87	0.0984490
87-88	0.1573526	87-88	0.1095782
88-89	0.1716155	88-89	0.1215099
89-90	0.1870248	89-90	0.1345577
⋮	⋮	⋮	⋮

- 👉 En France, Table[année, sexe, âge, département]
- 👉 Aux Etats-Unis, Table[année, sexe, âge, État, [race](#)]
- 👉 En Angleterre, Table[année, sexe, âge, comté, [classe sociale](#)]

Modèle additif de mortalité



- ☞ t : temps (délai) depuis le diagnostic
- ☞ Z_i : vecteur des covariables d'intérêt du patient i
- ☞ Z_{D_i} : vecteur de variables démographiques du patient i , avec $Z_{D_i} \subset Z_i$
- ☞ $\lambda_{O,i}$: mortalité observée du patient i
- ☞ $\lambda_{E,i}$: mortalité en excès du patient i (imputable au cancer)
- ☞ $\lambda_{P,i}$: mortalité attendue dans la population générale du patient i

Problème

Table de mortalité standard (appropriée) ?

Problème

Table de mortalité standard (appropriée) ?

☞ Dans certains registres de cancer spécialisés en France

Cancer étudié
(registre de cancer,
mortalité observée) : Z

- sexe
- année
- âge
- **gradient social**

Table de mortalité
(population générale,
mortalité attendue) : Z_D

- sexe
- année
- âge
- ~~gradient social~~

Problème

Table de mortalité standard (appropriée) ?

👉 Dans certains registres de cancer spécialisés en France

Cancer étudié
(registre de cancer,
mortalité observée) : Z

- sexe
- année
- âge
- **gradient social**

Table de mortalité
(population générale,
mortalité attendue) : Z_D

- sexe
- année
- âge
- ~~gradient social~~

👉 $gradient \in Z$, $gradient \notin Z_D$ (table de mortalité non stratifiée sur $gradient$)

Problème

Table de mortalité standard (appropriée) ?

☞ Dans certains registres de cancer spécialisés en France

Cancer étudié
(registre de cancer,
mortalité observée) : Z

- sexe
- année
- âge
- **gradient social**

Table de mortalité
(population générale,
mortalité attendue) : Z_D

- sexe
- année
- âge
- ~~gradient social~~

☞ $gradient \in Z$, $gradient \notin Z_D$ (table de mortalité non stratifiée sur $gradient$)

☞ Effet de $gradient$ sur la mortalité en excès $\lambda_E(t|Z)$ biaisé

☞ Effets des autres covariables biaisés

Pour résoudre ce problème :

☞ Modèle de régression multivarié

$$\lambda_{O,i}(t|Z_i) = \alpha_j \lambda_{P,i}(t|Z_{D_i}) + \lambda_{E,i}(t|Z_i)$$

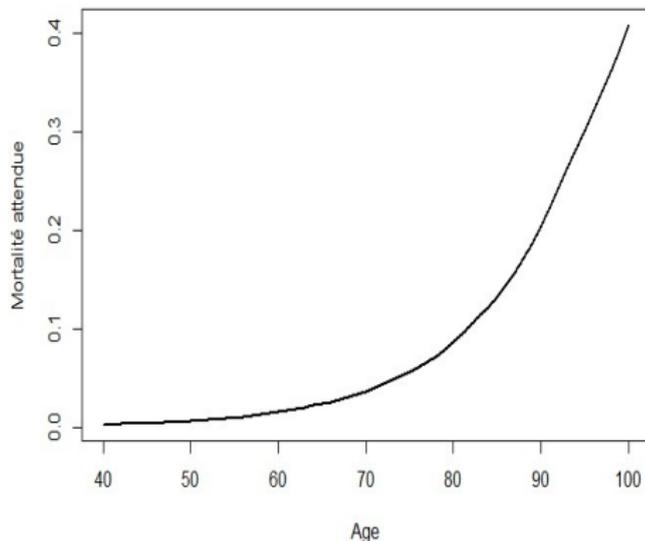
$j = 1, \dots, J$: modalité de la variable supplémentaire

Pour résoudre ce problème :

👉 Modèle de régression multivarié

$$\lambda_{O,i}(t|Z_i) = \alpha_j \lambda_{P,i}(t|Z_{D_i}) + \lambda_{E,i}(t|Z_i)$$

$j = 1, \dots, J$: modalité de la variable supplémentaire



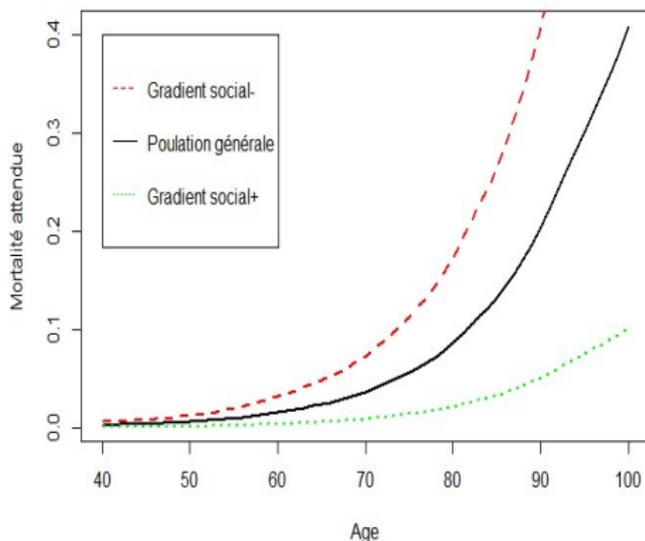
Pour résoudre ce problème :

👉 Modèle de régression multivarié

$$\lambda_{O,i}(t|Z_i) = \alpha_j \lambda_{P,i}(t|Z_{D_i}) + \lambda_{E,i}(t|Z_i)$$

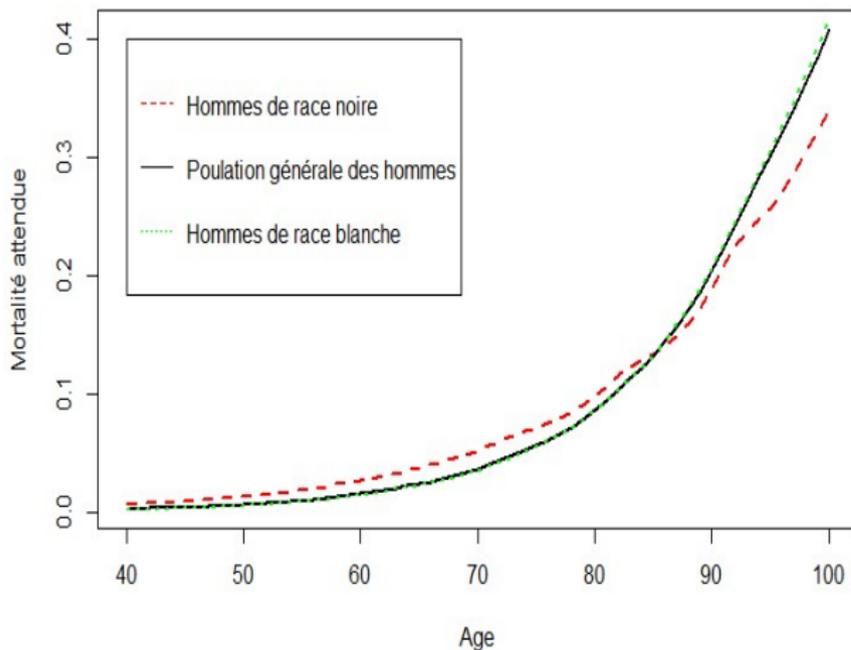
$j = 1, \dots, J$: modalité de la variable supplémentaire

👉 Pour $j \in \{\text{gradient social-}, \text{gradient social+}\}$



☞ Limite : hypothèse (forte) de proportionnalité entre les risques instantanés des différentes modalités de la variable supplémentaire

- 👉 Limite : hypothèse (forte) de proportionnalité entre les risques instantanés des différentes modalités de la variable supplémentaire
- 👉 Dans certaines situations, ce n'est toujours pas le cas



Objectifs

- ☞ Montrer par simulations certaines limites de modèles existants
- ☞ Présenter un modèle de régression dont la proportionnalité est fonction du temps (*travail préliminaire*)

Modèles existants

Modèle 1 : Estève et al.

$$\lambda_{P,i}(t|age_i) + \lambda_{E,i}(t|age_i, x_i)$$

Modèles existants

Modèle 1 : Estève et al.

$$\lambda_{P,i}(t|age_i) + \lambda_{E,i}(t|age_i, x_i)$$

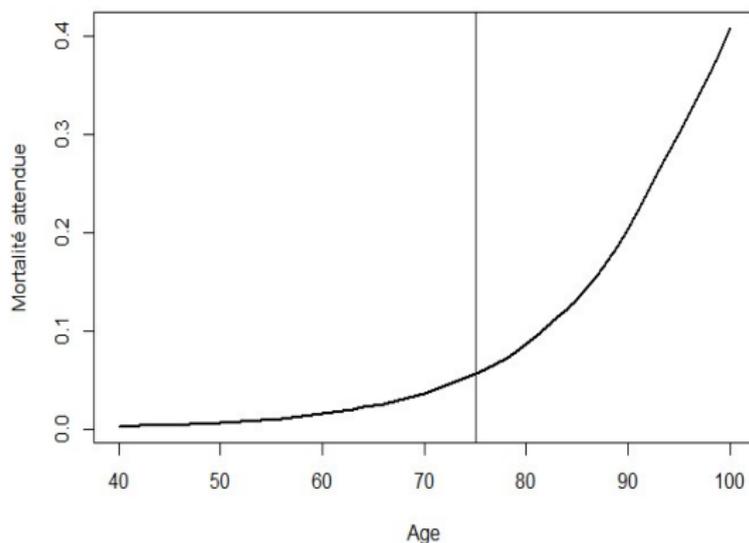
Modèle 2 : Touraine et al.

$$\alpha_j \lambda_{P,i}(t|age_i) + \lambda_{E,i}(t|age_i, x_i)$$

- ☞ 1000 jeux de données de 5000 patients chacun
- ☞ Covariables et effets sur la mortalité en excès
 - Age au diagnostic, 3 lois uniformes ($[30-65[$, $[65-75[$ et $[75-85[$ respectivement 25%, 35% et 40%)
 - $\beta_{age} = 0.3$
 - Variable (binaire) supplémentaire, $\%(x = 1) = 10$ (e.g., $x =$ gradient social)
 - $\beta_x = -0.2$
- ☞ Temps de survie
 - T_P suivant 2 scénarios

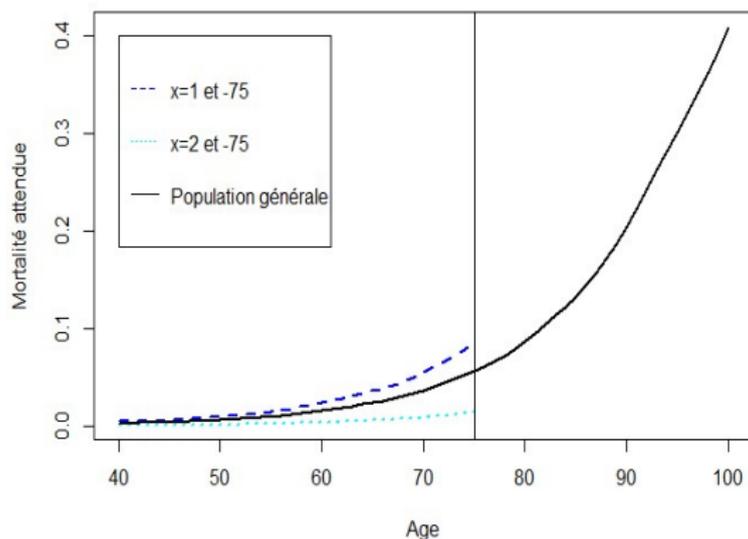
Scénario de référence

- Scénario où la mortalité attendue dans la population générale est la même que celle de la population étudiée



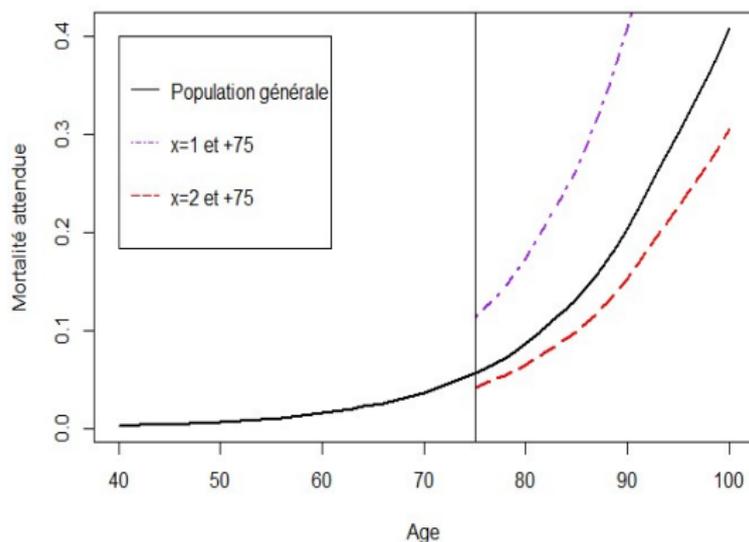
Construction d'un scénario où il n'y a pas de proportionnalité

- Scénario où la mortalité attendue dans la population générale diffère de celle de la population étudiée



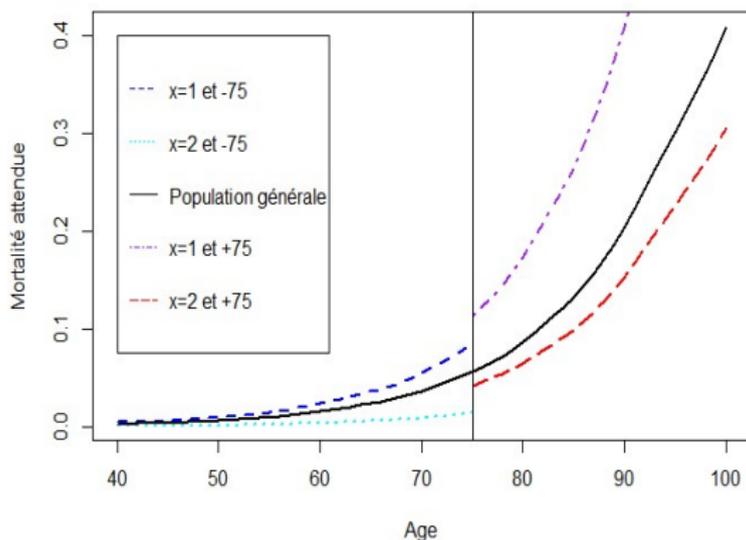
Construction d'un scénario où il n'y a pas de proportionnalité

- Scénario où la mortalité attendue dans la population générale diffère de celle de la population étudiée



Construction d'un scénario où il n'y a pas de proportionnalité

- Scénario où la mortalité attendue dans la population générale diffère de celle de la population étudiée



- 👉 1000 jeux de données de 5000 patients chacun
- 👉 Covariables et effets sur la mortalité en excès
 - Age au diagnostic, 3 lois uniformes ($[30-65[$, $[65-75[$ et $[75-85[$ respectivement 25%, 35% et 40%)
 - $\beta_{age} = 0.3$
 - Variable (binaire) supplémentaire, $\%(x = 1) = 10$ (e.g., $x = \text{gradient social}$)
 - $\beta_x = -0.2$
- 👉 Temps de survie
 - T_P suivant 2 scénarios

- 👉 1000 jeux de données de 5000 patients chacun
- 👉 Covariables et effets sur la mortalité en excès
 - Age au diagnostic, 3 lois uniformes ($[30-65[$, $[65-75[$ et $[75-85[$ respectivement 25%, 35% et 40%)
 - $\beta_{age} = 0.3$
 - Variable (binaire) supplémentaire, $\%(x = 1) = 10$ (e.g., $x =$ gradient social)
 - $\beta_x = -0.2$
- 👉 Temps de survie
 - T_P suivant 2 scénarios
 - T_E à partir d'une distribution de Weibull généralisée
 - C censure administrative à 6 ans

$\Rightarrow T_O = \min(T_P, T_E, C)$

Indicateurs de performance

Biais

Différence entre la valeur de l'estimateur et celle de son espérance

Indicateurs de performance

Biais

Différence entre la valeur de l'estimateur et celle de son espérance

EQM : Écart Quadratique Moyen

Précision de l'estimateur, il est fonction du biais et de la variabilité de l'estimateur

Indicateurs de performance

Biais

Différence entre la valeur de l'estimateur et celle de son espérance

EQM : Écart Quadratique Moyen

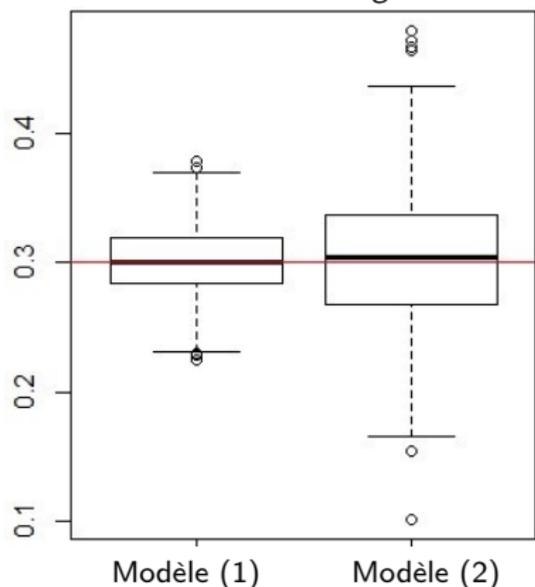
Précision de l'estimateur, il est fonction du biais et de la variabilité de l'estimateur

TCE : Taux de Couverture Empirique

Proportion d'occurrences où l'intervalle de confiance comprend la vraie valeur du paramètre

👉 Scénario de référence (table standard), mortalité attendue dans la population générale est la même que celle de la population étudiée

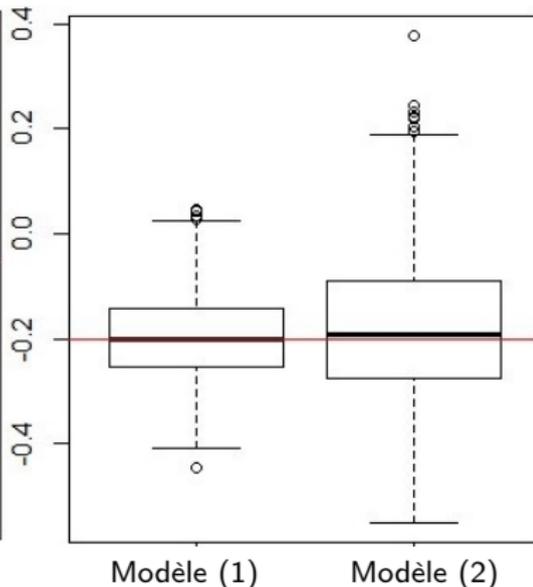
Estimation de $\beta_{age} = 0.3$



EQM : 0.0271
TCE : 95.40

0.0518
94.80

Estimation de $\beta_X = -0.2$

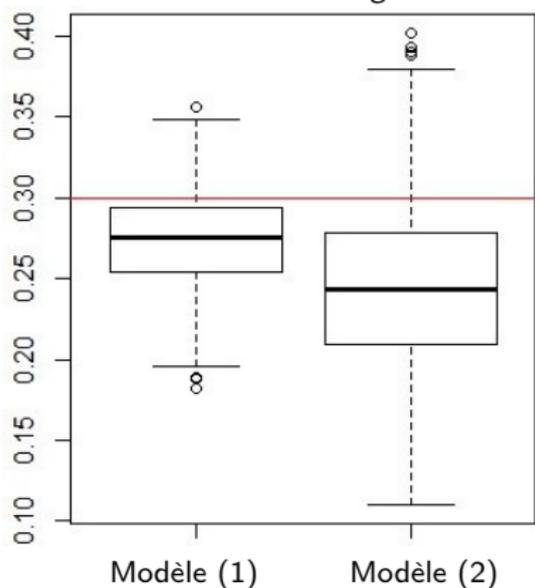


0.0826
94.10

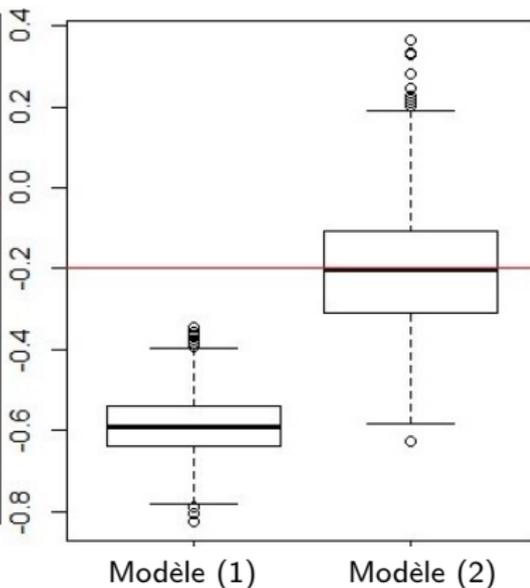
0.1398
95.90

- 👉 Scénario où la mortalité attendue dans la population générale diffère de celle de la population étudiée

Estimation de $\beta_{age} = 0.3$



Estimation de $\beta_X = -0.2$



EQM : 0.0389
TCE : 84.50

0.0756
78.90

0.3956
00.20

0.1552
95.70

 Modèle proposé

$$\lambda_{O,i}(t|Z_i) = \sum_{k=1}^K \alpha_{jk} \lambda_{P,i}(t|Z_{D_i}) I_k(\text{age}_i + t) + \lambda_{E,i}(t|Z_i)$$

$j = 1, \dots, J$: modalité de la variable supplémentaire

$k = 1, \dots, K$: classe d'âge au décès ($\text{age}_i + t$) dans la table de mortalité

☞ Modèle proposé

$$\lambda_{O,i}(t|Z_i) = \sum_{k=1}^K \alpha_{jk} \lambda_{P,i}(t|Z_{D_i}) I_k(\text{age}_i + t) + \lambda_{E,i}(t|Z_i)$$

$j = 1, \dots, J$: modalité de la variable supplémentaire

$k = 1, \dots, K$: classe d'âge au décès ($\text{age}_i + t$) dans la table de mortalité

Dans ce modèle :

☞ $x \in Z$ et $x \notin Z_D$

☞ $\alpha_{jk} : (j * k)$ paramètres multiplicatifs corrigeant la mortalité attendue des patients de la modalité j et de la classe d'âge au décès k

☞ $\lambda_{E,i}(t|Z) = \lambda_{E,0,i}(t) e^{\beta' Z_i}$ avec $\lambda_{E,0,i}(t) = \sum_{m=1}^M e^{\tau_m} I_m(t)$

☞ Modèle proposé

$$\lambda_{O,i}(t|Z_i) = \sum_{k=1}^K \alpha_{jk} \lambda_{P,i}(t|Z_{D_i}) I_k(\text{age}_i + t) + \lambda_{E,i}(t|Z_i)$$

$j = 1, \dots, J$: modalité de la variable supplémentaire

$k = 1, \dots, K$: classe d'âge au décès ($\text{age}_i + t$) dans la table de mortalité

Dans ce modèle :

☞ $x \in Z$ et $x \notin Z_D$

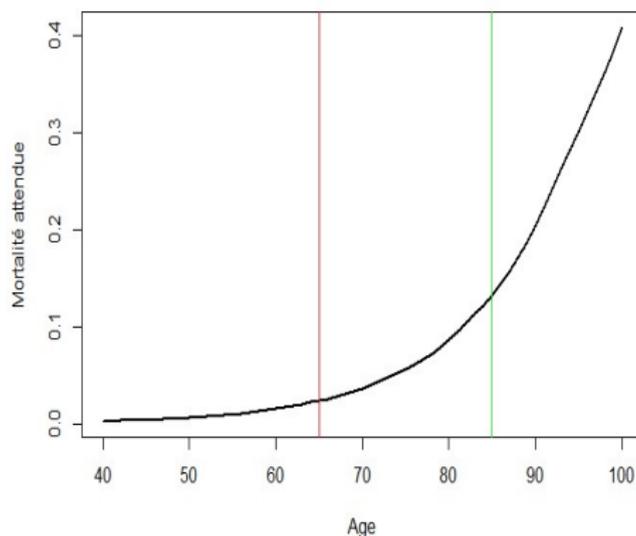
☞ $\alpha_{jk} : (j * k)$ paramètres multiplicatifs corrigeant la mortalité attendue des patients de la modalité j et de la classe d'âge au décès k

☞ $\lambda_{E,i}(t|Z) = \lambda_{E,0,i}(t) e^{\beta' Z_i}$ avec $\lambda_{E,0,i}(t) = \sum_{m=1}^M e^{\tau_m} I_m(t)$

☞ Estimation des paramètres par Maximum de Vraisemblance

Limites

- 👉 Modèle proportionnel par morceaux (on pourrait implémenter des fonctions plus souples notamment les splines)
- 👉 Cut-off (pour les intervalles d'âges)



- Obtenir les résultats des simulations en cours
- Analyse de sensibilité sur les cut-off pour les classes d'âge au décès
- Étendre le travail à k ($k > 2$) classes d'âge au décès
- Appliquer le modèle sur des données réelles du cancer colorectal avec pour variable supplémentaire Catégorie socio-professionnelle (table de mortalité non stratifiée sur cette variable)

Références bibliographiques

- Cheuvart B, Ryan L. Adjusting for age-related competing mortality in long-term cancer clinical trials. *Statistics in Medicine* 1991 ; 10 : 65-77.
- Danieli C, Remontet L, Bossard N, Roche L, Belot A. Estimating net survival : the importance of allowing for informative censoring. *Statistics in Medicine* 2012 ; 31 :775 -786.
- Estève J, Benhamou E, C, Raymond L. Relative survival and estimation of net survival : elements for further discussion. *Statistics in Medicine* 1990 ; 9 : 529-38.
- Grafféo N, Jooste V, Giorgi R. The impact of additional life-table variables on excess mortality estimates. *Statistics in Medicine* 2012 ;31(30) :4219-30.
- Touraine C, Grafféo N, Giorgi R and the CENSUR working survival group. An excess hazard model adjusting for lack of additional life table variables. 35th annual conference of the International Society for Clinical Biostatistics (2014). Viennes, Autriche.



Merci pour votre attention !

