



CONTRIBUTION À L'ANALYSE DES RISQUES DE MORTALITÉ ET DE MORBIDITÉ EN ASSURANCE À L'AIDE DE DONNÉES DE SANTÉ

Oleksandr Sorochynskyi
24^e Congrès des Actuaires – 2025-06-17



Estimation de l'espérance de vie en bonne santé avec les données PMSI



Objectif : une alternative aux enquêtes

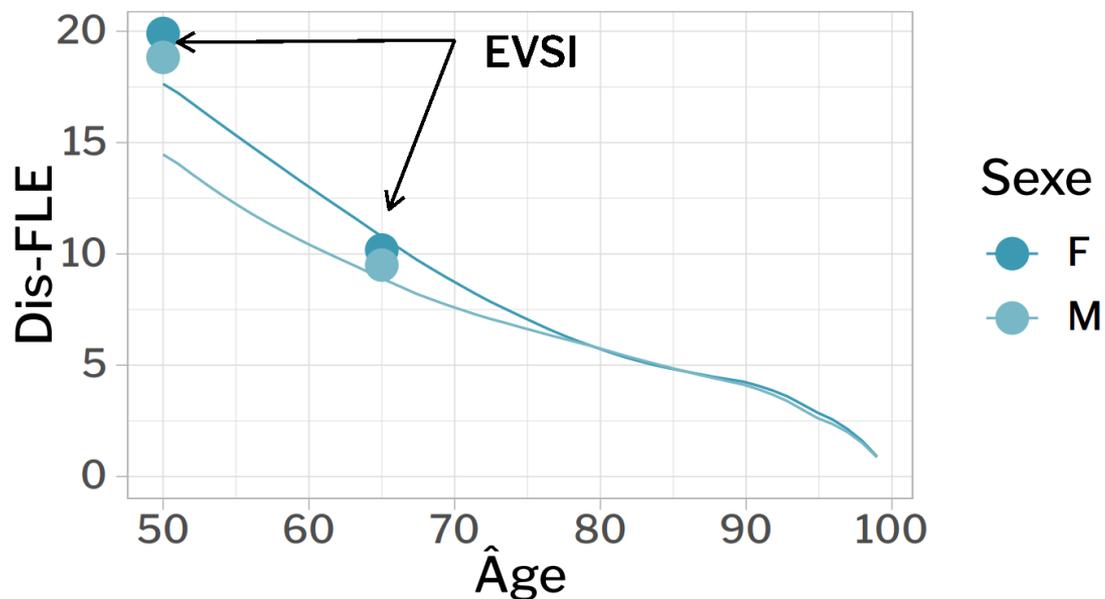
Et si on calculait l'espérance de vie en bonne santé à partir des données hospitalières ?



- Les enquêtes (HLY / EVSI) reposent sur des données autodéclarées
- Les données PMSI : exhaustives, standardisées, individuelles
- Définir la santé comme l'absence d'événement grave (37 pathologies)

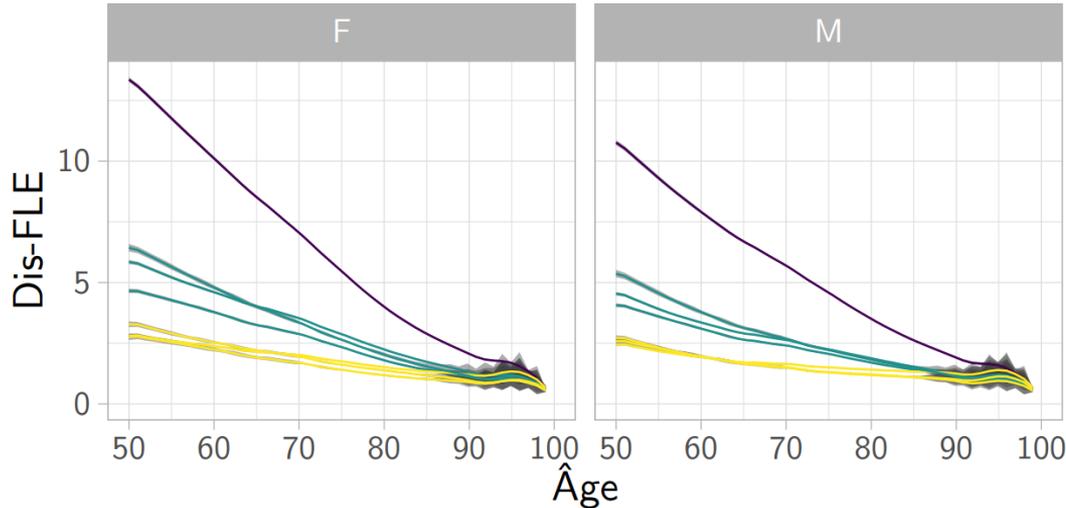
→ Un indicateur alternatif :
Disease-Free Life Expectancy (Dis-FLE)

Dis-FLE ajusté à la population générale



Nouvel indicateur comparable à l'EVS, même si des écarts existent

Dis-FLE selon les profils de risque



- Sans facteur de risque
- Un seul facteur de risque
- Facteurs de risque multiples

- Facteurs de risque : tabac, alcool, obésité
- Impact de la présence d'un facteur de risque: -10 ans sur le Dis-FLE

Pour en savoir plus, voir...



Blog «Actu d'actuaire» -

Mesure de l'espérance de vie en bonne santé en France

Oleksandr Sorochynskyi, Quentin Guibert, Frédéric Planchet et Michaël Schwarzingger (3 juin 2024). « Estimating Disease-Free Life Expectancy Based on Clinical Data from the French Hospital Discharge Database ». In : *Risks* 12.6, p. 92. issn : 2227-9091. doi : 10.3390/risks12060092.
url : <https://www.mdpi.com/2227-9091/12/6/92> (visité le 16/03/2025)



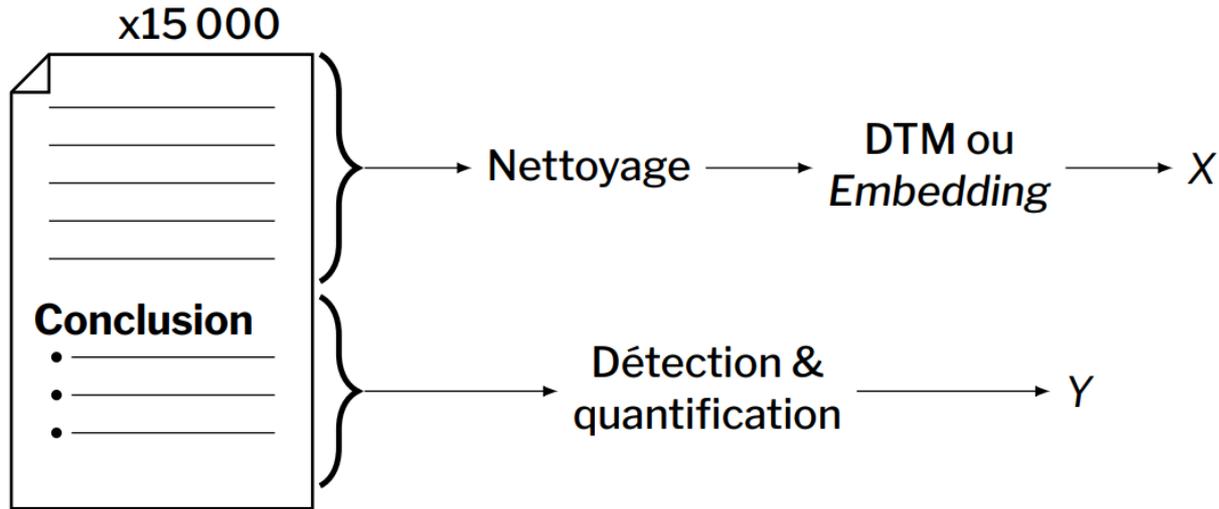
Prévision précoce des coûts des sinistres corporels

Contexte et objectif



- **Contexte**
Évaluation des **postes de préjudice** pour les sinistres corporels.
- **Objectif**
Intégrer automatiquement les documents textuels dans le **provisionnement** des sinistres corporels.

Schéma du traitement de données



$$Y \sim f(X)$$

Performances R^2 des modèles



Poste de préjudice	RF hurdle	RF regr.	LASSO hurdle	LASSO-RF hurdle
DFP	25 %	24 %	42 %	30 %
DSA-H	21 %	20 %	28 %	24 %
SE	41 %	44 %	49 %	47 %
ATP	14 %	14 %	18 %	17 %
PGPA	35 %	35 %	39 %	40 %
DFT	27 %	31 %	38 %	32 %

Limites \Rightarrow Pistes de sortie



- **Source unique : rapports médicaux**
 \Rightarrow collecter aussi les documents en amont (CR d'hospitalisation, correspondances).
- **Performances $<$ attentes**
 \Rightarrow monter une baseline humaine pour savoir si la tâche — ou le modèle — est le vrai frein.
- **Performances $<$ attentes**
 \Rightarrow explorer des LLM / modèles pré-entraînés adaptés au domaine.

Calcul des âges biologiques à partir des données NHANES

Qu'est-ce qu'un « âge biologique » ?



« Un âge qui reflète mieux l'état biologique du corps »

- immédiatement interprétable;
- intégration simple dans les outils existants.

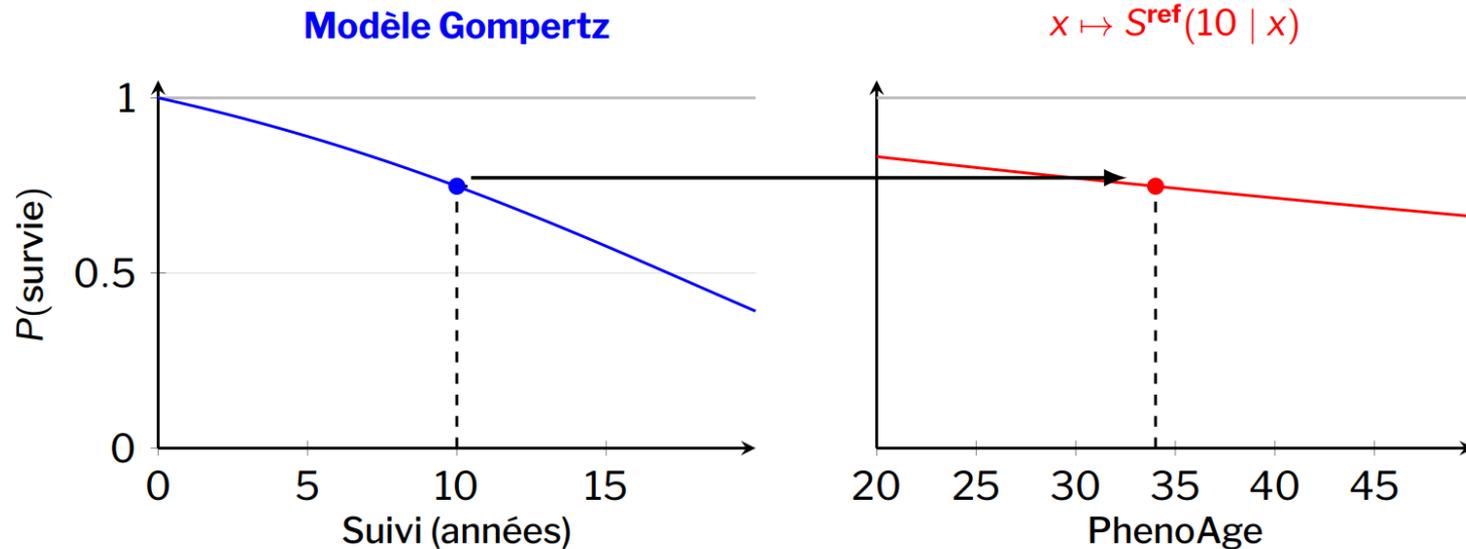
Objectifs

- Introduire les âges biologiques dans la littérature actuarielle;
- Tester une application des âges biologiques.

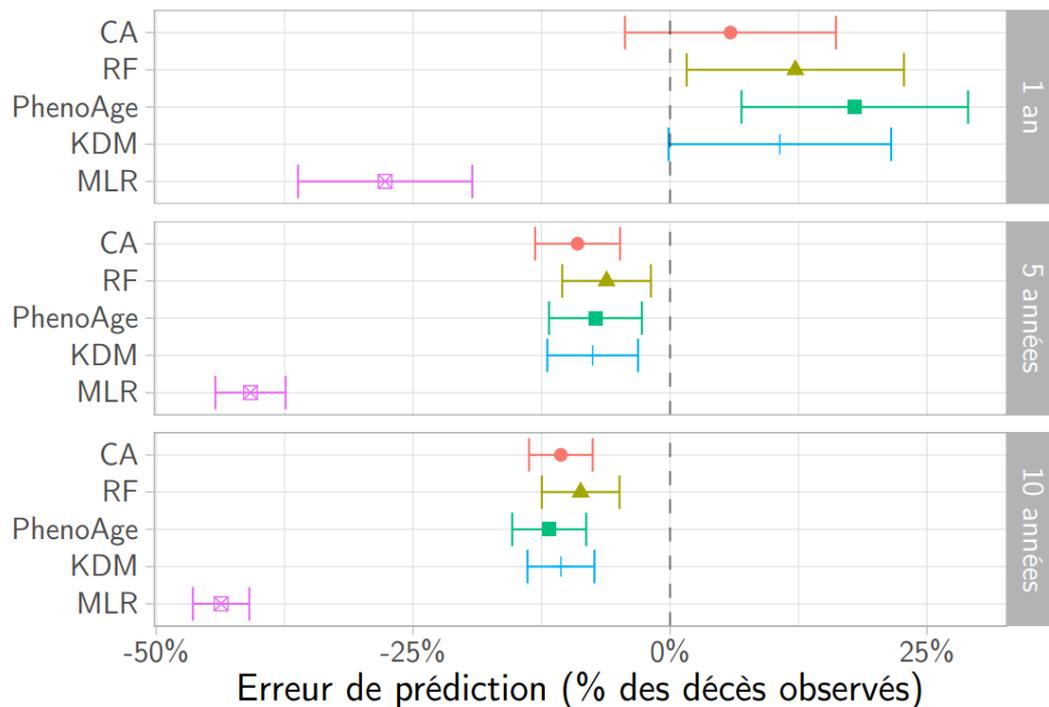
Exemple d'un âge biologique : PhenoAge

Principe : utiliser des modèles Gompertz pour

1. estimer la probabilité de survie ;
2. transformer en âge par la distribution de référence.

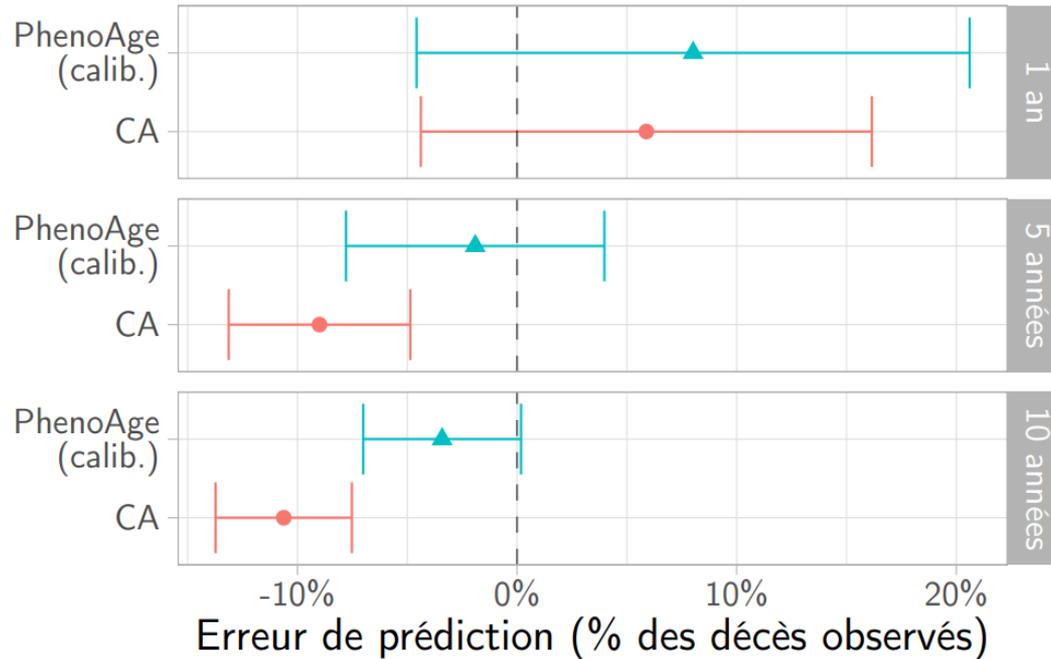


Application naïve à une table de mortalité

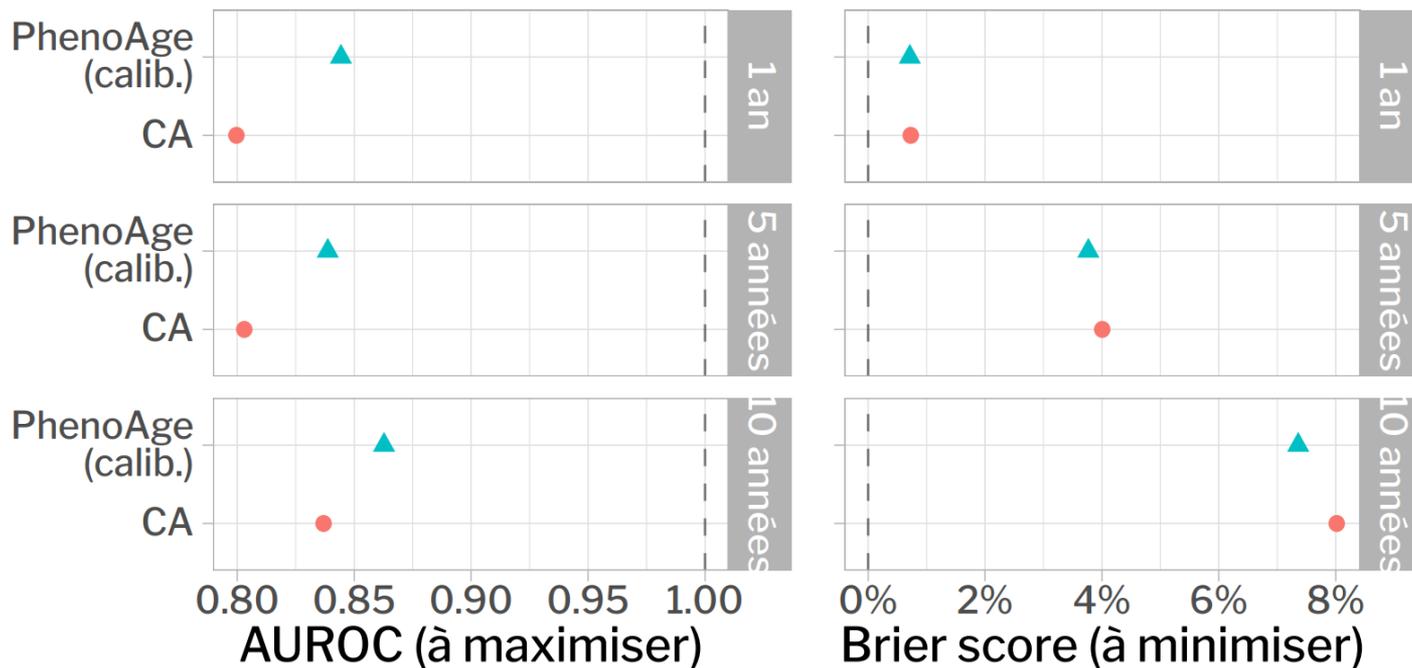


- Comparaison entre le nombre de décès observés et estimés
- L'approche naïve n'améliore pas la qualité des estimations

Application avec un Âge biologique calibré



Indicateurs de précision individuelle des prédictions



Pour en savoir plus, voir...

[Blog «Actu d'actuaire» -
Age biologique et prévention en assurance](#)

Article soumis à *European Actuarial Journal...*





Merci

Oleksandr Sorochnytskyi
oleksandr.sorochnytskyi@primact.fr

