Modélisation de l'histoire d'une exposition sur le changement d'un marqueur au cours du temps

Ariane Bercu

Institut de santé publique, d'épidémiologie et de développement (ISPED)

Mémoire de Master 2 Biostatistiques

Etudiant :
BERCU Ariane

Encadrant :
PROUST-LIMA Cécile

Modélisation de l'histoire d'une exposition sur le changement d'un marqueur au cours du temps

au sein de

Bordeaux population health (BPH)
Equipe Biostatistiques

Mars 2021 - Fin Juillet 2021

Remerciements

Je tiens à remercier grandement Cécile Proust-Lima, ma maître de stage, pour son expertise, son suivi attentif et pour sa confiance.

Je remercie chalereusement Viviane Philipps, ingénieure d'étude, pour son aide à la programmation et son écoute.

Je tiens également à remercier tous les membres de l'équipe Biostatistique pour leur accueil, leur aide et pour mon intégration à l'équipe en ce contexte sanitaire particulier.

Table des matières

Table des figures

Liste des tableaux

1 Introduction

1.1 Structure d'accueil

Le Centre de Recherche Inserm-Université de Bordeaux U1219, plus communément nommé Bordeaux population health (BPH) est une Unité Mixte de Recherche (UMR) dirigée par le professeur Christophe Tzourio. Cette unité est un laboratoire labellisé soutenu à la fois par l'Inserm et l'Université de Bordeaux [1].

L'Institut National de la Santé et de la Recherche Médicale (Inserm) a été fondé en 1964 par Raymond Marcellin, ministre de la Santé à cette période [2]. Depuis lors, l'Inserm est acteur de mission de Santé Publique en investissant dans la recherche biomédicale et appliquée. Son expertise collective se déploie dans 350 structures de recherche sur l'ensemble du territoire français, mettant l'Inserm au premier rang européen des institutions accadémiques de recherche dans le domaine biomédical. L'Inserm entretient de fortes relations de coopération européennes et internationales, permettant de dynamiser les collaborations à thématiques de Santé dans le monde. L'Inserm est donc à la fois un acteur local et international dont l'expertise est sollicitée et reconnue.

L'Université de Bordeaux est un Etablissement Public à caractère Scientifique, Culturel et Professionnel (EPSCP). La gouvernance est asssurée par le président Manuel Tunon de Lara. La recherche à l'université est organisée en 11 départements de recherche, dont celui de Santé Publique auquel est rattaché le centre BPH.

Le BPH est constitué de onze équipes labellisées sur la période 2016-2021 [3]. Ces équipes sont composées d'un grand nombre de chercheurs, ingénieurs et doctorants qui participent à une vision globale des problématiques de Santé couvrant un large champ de pathologies, expositions, méthodes et populations. Chaque équipe du centre a une thématique propre de recherche, tout en participant à des projets communs. L'équipe de Biostatistique développe des méthodes statistiques pour des données épidémiologiques ou cliniques. L'équipe est dirigée depuis 2014 par la directrice de recherche Hélène Jacqmin-Gadda. L'axe de recherche principal de l'équipe porte sur les modèles dynamiques, c'est-à-dire les modèles prenant en compte la dimension temporelle des processus et des évènements de Santé. On retrouve dans cette catégorie les modèles multivariés pour données dépendantes du temps tel que les modèles conjoints pour données longitudinales. Une des applications princpales des travaux de l'équipe est l'étude du vieillissement cérébral, notamment l'histoire naturelle et les facteurs

de risque de la maladie d'Alzheimer. D'autres domaines d'application sont le virus de l'immunodéficience humaine (VIH)), la cancérologie (e.g cancer du poumon) ou la maladie rénale chronique.

Ce stage a été encadré par la directrice de recherche Cécile Proust-Lima au sein de l'équipe de Biostatistiques du BPH.

1.2 Contexte de recherche

En épidémiologie, nous nous intéressons très souvent à l'association entre une exposition et la survenue d'un évènement de santé. Dans la majorité des études, l'exposition est considérée uniquement à un temps donné, le plus souvent l'inclusion, et l'événement de santé est évalué de façon répétée à partir de ce temps donné (marqueur répété ou temps d'événément), l'événement de santé étant un processus dynamique. Pourtant, très souvent, l'exposition est elle aussi un processus qui évolue dans le temps, et son association avec l'événement de santé peut être complexe et dépendre de la fenêtre temporelle pendant laquelle elle est évaluée. Ainsi prendre une photo à un temps donné de l'histoire de l'exposition peut nous éloigner de l'hypothèse biologique réelle liée à l'exposition.

En épidémiologie vie-entière, plusieurs hypothèses de mécanismes biologiques d'impact sur la santé ont été identifiés comme l'accumulation de l'exposition au cours du temps ou une exposition accrue sur une fenêtre d'exposition courte [4]. Cette problématique d'association complexe et dépendante du temps est particulièrement présente dans de nombreuses maladies chroniques qui sont liées à l'âge comme la démence, les maladies cardio-vasculaires ou les cancers lorsque l'on s'intéresse à des expositions environnemmentales comme la pollution [5], ou de style de vie comme le tabagisme et l'Indice de Masse Corporel (IMC) [6].

De nouvelles avancées en méthodes statistiques ont été développées, permettant d'explorer une représentation plus appropriée de l'exposition et comprendre si et comment l'histoire d'une exposition affecte la santé. L'approche la plus communément utilisée est le Cumulative Index Exposure (CIE). Elle consiste à calculer la dose cumulée à une exposition sur une fenêtre de temps de taille S. Le CIE en t est l'intégrale de l'exposition de $t - S$ à t (ou la somme des expositions dans le cas discret), c'est-à-dire l'aire sous la courbe de l'exposition. La méthode CIE est basée sur l'hypothèse que c'est l'accumulation de l'exposition qui est importante, quelle que soit la temporalité. Deux sujets qui ont un historique de l'exposition très différent auront le même risque d'événement de santé tant que l'accumulation sur la

fenêtre S est la même. Cette mesure résumée de l'histoire d'exposition repose donc sur une hypothèse forte qu'il est impératif de vérifier pour capter l'impact réel de l'exposition cumulée sur l'évènement de santé. Par exemple, nous pouvons regarder l'histoire d'Hypertension Artérielle (HTA) sur la santé cardiovasculaire des personnes âgées. Selon l'hypothèse du CIE, l'HTA à un âge plus avancé aurait le même poids que l'HTA à des âges jeunes. Cela serait en contradiction avec la littérature qui suggère un poids plus important de l'exposition à un âge jeune [6, 7].

Pour tenir compte de l'importance de la temporalité des expositions, une extension du CIE a été proposée, appelée Weighted CIE (WCIE) [8, 9, 10]. Dans l'approche WCIE, des poids dépendants du temps sont attribués à chaque mesure de l'exposition afin de caractériser son importance dans la temporalité de l'accumulation de l'exposition. Les poids peuvent être fixés d'après la littérature et l'avis d'experts ou bien, le plus souvent, en absence d'a priori, ils peuvent être directement estimés à partir des données. L'application de la méthode WCIE reste aujourd'hui très limitée dans les études de cohortes. En effet, elle nécessite des expositions mesurées de façon complète et sans erreur de mesure alors qu'en pratique, les expositions dans les études de cohortes sont mesurées par intermittence aux visites et très souvent bruitées (e.g. poids, HTA). De plus, la méthode WCIE a été développée pour deux grands types d'événement de santé, un marqueur mesuré à un temps donné (en fin de fenêtre d'exposition) ou un temps d'événement censuré au cours d'un suivi. Elle ne permet pas de traiter des marqueurs répétés au cours du temps, comme utilisé en vieillissement par exemple avec l'évolution cognitive.

Récemment, une approche proposée par Wagner et al [10] a étendu la méthodologie WCIE à des expositions mesurées par intermittence et avec erreur, et à un marqueur de santé répété au cours du temps. La méthodologie utilisée repose sur une approche landmark qui sépare totalement la fenêtre de mesure de l'exposition cumulée et la survenue de l'évènement de santé. Cette première approche permet de traiter l'exposition comme un processus dynamique mais elle ne permet pas de lier les processus de l'exposition et du marqueur en tout temps alors que les hypothèses biologiques suggéreraient plutôt que l'exposition peut impacter le marqueur jusqu'à son instant d'évaluation.

1.3 Objectif et Missions

L'objectif de ce stage est d'étendre la méthode WCIE pour étudier l'effet cumulé d'une exposition bruitée et mesurée par intermittence sur la pente instantanée d'un marqueur de maladie par une approche de modélisation conjointe. L'objectif épidémiologique est d'étudier la temporalité de l'effet de l'Indice de Masse Corporel (IMC) sur le déclin cognitif chez les personnes âgées. Les missions pour remplir l'objectif de stage sont les suivantes :

1. Appréhender la littérature sur le sujet et la méthodologie statistique envisagée ;
2. Définir le modèle conjoint, les équations du modèle et l'écriture de la vraisemblance ;
3. Implémenter l'estimation des paramètres du modèle sous forme de programme sous R ;
4. Appliquer la méthodologie sur les données de la cohorte en population des trois cités (3C).

2 Méthodologie

2.1 Modèle conjoint pour une exposition et un marqueur répétés

Nous cherchons à développer un modèle permettant de modéliser l'impact de l'histoire d'une exposition E sur un marqueur de la maladie Y, tous les deux mesurés de façon répétée au cours du temps. Nous sommes principalement intéressés par l'influence du processus de l'exposition E sur le changement ultérieur du marqueur Y. Pour ce faire, nous devons retrouver le mécanisme qui explique comment le marqueur change au cours du temps en accord avec l'approche dynamique de la causalité [**?**, 11]. Le principe est donc d'évaluer l'association dépendante du temps de l'exposition sur le changement du marqueur Y plutôt que sur le niveau ultérieur du marqueur Y.

Le schéma 1 résume le principe général de notre méthodologie. L'exposition mesurée à un temps t est notée $E_i(t)$ pour le sujet i, $i = 1, \ldots, N$. Nous notons 0_E le temps de référence ("baseline") de l'exposition tel que $t = 0$. Nous considérons dans ce travail qu'il s'agit du premier temps de mesure de l'exposition dans l'échantillon des N sujets. Notons, $Y_i(t)$ le marqueur associé pour le sujet i à un temps t. Pour permettre d'étudier l'effet de E sur Y avec une antériorité de S, les mesures de Y ne sont considérées qu'à partir de $0_Y = 0_E + S = S$.

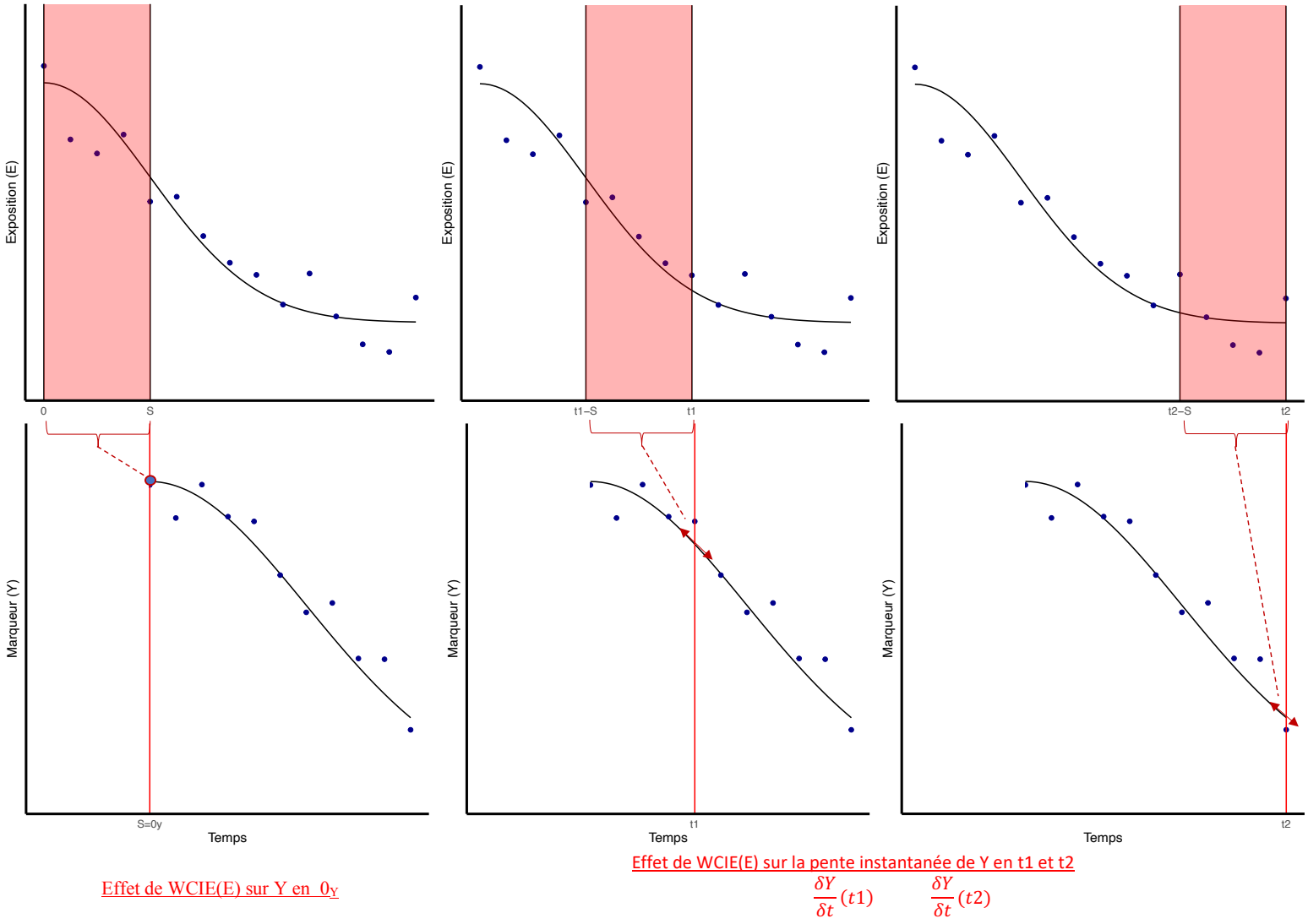

FIGURE 1: Schéma explicatif du modèle conjoint de l'histoire d'une exposition E et d'un marqueur Y pour un sujet.

Dans ce schéma, la première ligne montre la trajectoire fictive des données d'exposition d'un sujet et la deuxième ligne montre la trajectoire fictive du marqueur pour ce même sujet. La fenêtre d'histoire de l'exposition à associer au marqueur est représentée dans le rectangle rouge. Il s'agit des S unités de temps précédant l'évaluation du marqueur. Elle évolue donc avec le temps. Ainsi à $t = S$, l'histoire de l'exposition E est prise en compte seulement sur la fenêtre $[0, S]$ où S est l'amplitude fixe des fenêtres d'exposition pour quantifier son impact sur le niveau initial de Y en S. Puis pour $t > S$, l'histoire de l'exposition E sur la fenêtre $[t - S, t]$ est prise en compte pour quantifier son impact sur le changement de pente instantanée de Y en t.

Nous supposons dans ce travail que l'exposition et le marqueur sont tous deux continus et suivent une distribution multivariée normale. Pour définir le modèle, nous séparons les données observées E et Y de leur niveau sous-jacent débruité, appelé respectivement E^* et Y^*. Ce niveau sous-jacent est vu comme le "vrai" niveau de l'exposition ou du marqueur, et les relations entre les deux processus se font à leur niveau.

$$E_i(t) = E_i^*(t) + \varepsilon_i(t) \tag{1}$$

$$Y_i(t) = Y_i^*(t) + \epsilon_i(t), \tag{2}$$

avec $\varepsilon_i(t) \sim N(0, \sigma_\varepsilon^2)$ et $\epsilon_i(t) \sim N(0, \sigma_\epsilon^2)$, les erreurs de mesure $\varepsilon_i(t)$ et $\epsilon_i(t)$ sont indépendantes pour tous t, $\varepsilon_i(t_1)$ et $\varepsilon_i(t_2)$ sont indépendants $\forall i, t_1, t_2$ avec $t_1 \neq t_2$, de même pour $\epsilon_i(t_1)$ et $\epsilon_i(t_2)$.

L'évolution de l'exposition sous-jacente est supposée suivre le modèle linéaire mixte suivant pour tout i=1,. . .,N et t$\geq$ 0 :

$$E_i^*(t) = {X_{iE}}^\top \otimes F_E(t)^\top \beta + F_{ER}(t)^\top b_i \tag{3}$$

avec le vecteur d'effets aléatoires $b_i \sim N(0, B)$, b_i et $\varepsilon_i(t)$ indépendants. Nous avons X_{iE} les variables explicatives (comprenant un intercept). Les vecteurs F_E, F_{ER} incluent la base de fonctions décrivant la forme de l'évolution temporelle (ils incluent l'intercept). Nous nous focalisons dans ce travail sur des fonctions M-splines ou polynomiales du temps t.

L'évolution du marqueur sous-jacent est aussi défini par un modèle mixte. Mais contrairement à l'écriture classique, nous scindons ici le modèle mixte en deux :

— le modèle pour le niveau initial du marqueur débruité en S

— le modèle pour son changement débruité au cours du temps ;

avec les deux parties pouvant dépendre de l'exposition E^*.

Pour tout i=1,...,N et t$\geq$ S,

$$\begin{cases} Y_i^*(S) & = {X_{i0}}^\top \alpha + u_{i0} + \underbrace{\sum_{s=-S}^{0} w_0(s) E_i^*(S+s)}_{WCIE_i(0,S)} \\ \frac{\delta Y_i^*(t)}{\delta(t)} & = {X_{i1}}^\top \otimes F_Y(t)^\top \delta + F_{YR}(t)^\top u_{i1} + \underbrace{\sum_{s=-S}^{0} w_1(s) E_i^*(t+s)}_{WCIE_i(t-S,t)} , \forall\, t > S \end{cases} \tag{4}$$

avec $u_i = (u_{0i}, u_{1i}) \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, D = \begin{pmatrix} \sigma_0^2 & C^\top \\ C & C_1 \end{pmatrix}\right)$, u_i et $\epsilon_i(t)$ indépendants.

Nous avons X_{i0} et X_{i1} les variables explicatives associées au niveau initial de Y et à sa pente instantanée (elles comprennent l'intercept). Les vecteurs F_Y, F_{YR} incluent la base de fonctions décrivant la forme de l'évolution temporelle de la pente instantanée (ils incluent l'intercept). Comme pour l'exposition, nous nous focalisons dans ce travail sur des fonctions M-splines ou polynomiales du temps t.

D'après les équations précédentes, nous pouvons en déduire l'écriture de $Y_i^*(t)$ pour tout i=1,...,N et t $\geq$ S,

$$
\begin{aligned}
Y_i^*(t) &= Y_i^*(S) + \int_S^t \frac{\delta Y_i^*(l)}{\delta l} dl && (5)\\
&= {X_{i0}}^\top \alpha + u_{i0} + \sum_{s=-S}^{0} w_0(s) E_i^*(S+s) \\
&+ \int_S^t {X_{i1}}^\top \otimes F_Y(l)^\top dl\delta + \int_S^t F_{YR}(l)^\top dl u_{i1} + \int_S^t \sum_{s=-S}^{0} w_1(s) E_i^*(l+s) dl && (6)\\
&= {X_{i0}}^\top \alpha + u_{i0} + \sum_{s=-S}^{0} w_0(s) [{X_{iE}}^\top \otimes F_E(S+s)^\top \beta] \\
&+ {X_{i1}}^\top \otimes \Delta\mathbb{F}_Y(t,S)^\top \delta + \Delta\mathbb{F}_{YR}(t,S)^\top u_{i1} \\
&+ \sum_{s=-S}^{0} w_1(s) {X_{iE}}^\top \otimes \Delta\mathbb{F}_E(t+s,S+s)^\top \beta \\
&+ {b_i}^\top \Big(\sum_{s=-S}^{0} w_0(s) F_{ER}(S+s) + \sum_{s=-S}^{0} w_1(s) \Delta\mathbb{F}_{ER}(t+s,S+s) \Big) && (7)
\end{aligned}
$$

On note de façon générique $\Delta\mathbb{F}(u,v)$ la différence des primitives de F entre u et v, avec $u \geq v$.
Les équations 4 montrent comment l'écriture de Y_i^* dépend de E_i^*, puis par développement 5, 6 et 7 comment elle dépend des effets aléatoires de E soit b_i (en violet). Ainsi nous avons obtenu une formulation de la distribution de Y_i qui s'écrit comme un modèle mixte avec des combinaisons linéaires des effets aléatoires b_i et u_i. Ce modèle est un modèle non linéaire mixte car les paramètres du WCIE, $w_.(t)$, interviennent de façon non linéaire dans le modèle.

Par la suite, nous montrons que notre modèle s'écrit sous la forme classique d'un modèle multivarié non indépendant dont le calcul de vraisemblance est connu et explicite. Pour faciliter la compréhension nous présentons notre modèle sous forme vectorielle pour un individu i avec i=1,...,N à un temps t (t$\geq$0)

En posant

$$\xi_{iE1}(t) = \sum_{s=-S}^{0} w_0(s) X_{iE} \otimes F_E(0_Y + s) + \sum_{s=-S}^{0} w_1(s) X_{iE} \otimes \Delta\mathbb{F}_E(t_s)$$

et

$$\kappa_{iE}(t) = \sum_{s=-S}^{0} w_0(s) F_{ER}(0_Y + s) + \sum_{s=-S}^{0} w_1(s) \Delta\mathbb{F}_{ER}(t_s)$$

Nous pouvons définir le modèle joint pour l'exposition et le marqueur :

$$\begin{aligned}\begin{pmatrix} E_i(t) \\ Y_i(t) \end{pmatrix} = &\begin{pmatrix} {X_{iE}}^\top \otimes F_E(t)^\top & 0 & 0 \\ \xi_{iE1}(t)^\top & {X_{i0}}^\top & {X_{i1}}^\top \otimes \Delta\mathbb{F}_Y(t)^\top \end{pmatrix} \times \begin{pmatrix} \beta \\ \alpha \\ \delta \end{pmatrix} \\ &+ \begin{pmatrix} F_{ER}(t)^\top & 0 & 0 \\ \kappa_{iE}(t)^\top & 1 & \Delta\mathbb{F}_{YR}(t)^\top \end{pmatrix} \times \begin{pmatrix} b_i \\ u_{i0} \\ u_{i1} \end{pmatrix} + \begin{pmatrix} \varepsilon_i(t) \\ \epsilon_i(t) \end{pmatrix}\end{aligned} \tag{8}$$

Dans l'équation 8, nous retrouvons l'écriture classique d'un modèle linéaire mixte bivarié avec les paramètres fixes à estimer β, α, δ et $\theta_{k.}$ (en marron) et les paramètres de variance à estimer pour $b_i, u_{i0}, u_{i1}, \varepsilon_i$ et ϵ_i (en olive).

L'originalité de notre modèle mixte bivarié porte sur les termes conjoints $\xi_{iE1}(t)$ et $\kappa_{iE}(t)$ associés aux effets fixes β et aux effets aléatoires b_i. En effet, l'équation 8, est celle d'un modèle linéaire mixte bivarié uniquement si ces termes sont déterminés. Or dans notre approche, ces termes incluent eux-mêmes des paramètres à estimer avec les poids $w_0(s)$ et $w_1(s)$ qui définissent les WCIE. Dans la suite, nous modélisons ces poids par des B-spline ou N-splines, ainsi pour tout s=$-S$,...,0 et K>0, $w_0(s) = \sum_{k=1}^{K} \theta_{0k} B_k(s)$ et $w_1(s) = \sum_{k=1}^{K} \theta_{1k} B_k(s)$ avec les θ_{0k} et θ_{1k} les paramètres à estimer.

Nous utilisons la méthode du maximum de vraisemblance pour estimer les paramètres de notre modèle. La vraisemblance jointe est déduite des équations 7 et 3 qui montrent que les deux variables E et Y suivent une distribution multivariée normale.

Pour i=1,...,N,

$$\begin{aligned} Y_i|b_i &\sim N(m_i(t), V(t)) \\ E_i|b_i &\sim N({X_{iE}}^\top \otimes {F_E(t)}^\top \beta + {F_{ER}(t)}^\top b_i, \sigma_\varepsilon^2) \\ b_i &\sim N(0, B) \end{aligned}$$

$$\begin{aligned} m_i(t) &= {X_{i0}}^\top \alpha + \textstyle\sum_{s=-S}^{0} w_0(s)[{X_{iE}}^\top \otimes {F_E(S+s)}^\top \beta] \\ &+ {X_{i1}}^\top \otimes \Delta\mathbb{F}_Y(t,S)^\top \delta \\ &+ \textstyle\sum_{s=-S}^{0} w_1(s) {X_{iE}}^\top \otimes \Delta\mathbb{F}_E(t+s, S+s)^\top \beta \\ &+ {b_i}^\top (\textstyle\sum_{s=-S}^{0} w_0(s) F_{ER}(S+s) + \sum_{s=-S}^{0} w_1(s) \Delta\mathbb{F}_{ER}(t+s, S+s)) \end{aligned}$$

$$V(t) = (1, \Delta\mathbb{F}_{YR}(t,S)^\top, 1) \begin{pmatrix} \sigma_0^2 & C^\top & 0^\top \\ C & C_1 & 0^\top \\ 0 & 0 & \sigma_\epsilon^2 \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbb{F}_{YR}(t,S)^\top \\ 1 \end{pmatrix}$$

On note ψ l'ensemble des paramètres à estimer du modèle, alors la vraisemblance du modèle est la suivante :

$$\begin{aligned} L(\psi) &= \prod_{i=1}^{N} l_i(\psi) \\ l_i(\psi) &= \int_b f_{Y_i|b_i}(y_i, b) f_{E_i|b_i}(e_i, b) f_{b_i}(b) db \end{aligned}$$

Avec $f_{Y_i|b_i}(y_i, b)$,$f_{E_i|b_i}(e_i, b)$ et $f_{b_i}(b)$ des fonctions de densité de lois normales explicitées précédemment. Par calcul, nous pouvons obtenir une écriture explicite analytique de la log-vraisemblance grâce aux proporités d'intégration des lois normales.

2.2 Algorithme d'optimisation et implémentation informatique

2.2.1 Calcul de la vraisemblance

Nous avons choisi de récupérer la valeur de la log-vraisemblance de notre modèle par la fonction hlme du package lcmm qui estime des modèles linéaires mixtes par maximum de vraisemblance. Cela a nécessité de modifier préalablement la fonction pour traiter les erreurs hétéroscédastiques. En spécifiant qu'on ne souhaite pas d'optimisation, la fonction hlme renvoie la log-vraisemblance pour une valeur donnée de l'ensemble des paramètres.

2.2.2 Algorithme Marquardt-Levenberg

La maximisation de cette log-vraisemblance en fonction de l'ensemble des paramètres souhaités est faite par l'algorithme d'optimisation de Marquardt-Levenberg implémenté sous R dans le package marqlevAlg et la fonction mla [12, 13]. Nous notons $\psi^{(k+1)}$ l'estimation des paramètres à l'itération $k+1$ d'après le point de départ $\psi^{(0)}$, jusqu'à convergence nous avons,

$$\psi^{(k+1)} = \psi^{(k)} - \delta_k(\tilde{H}(\mathcal{L}(\psi^{(k)})))^{-1}\nabla(\mathcal{L}(\psi^{(k)}))$$

avec $\nabla(\mathcal{L}(\psi^{(k)}))$ le gradient de la fonction de log-vraisemblance en $\psi^{(k)}$, $\tilde{H}(\mathcal{L}(\psi^{(k)}))$ est la hessienne où les termes diagonaux sont remplacés par :

$$\tilde{H}(\mathcal{L}(\psi^{(k)}))_{ii} = H(\mathcal{L}(\psi^{(k)}))_{ii} + \lambda_k[(1-\eta_k)|H(\mathcal{L}(\psi^{(k)}))_{ii}| + \eta_k tr(H(\mathcal{L}(\psi^{(k)})))]$$

Les paramètres $\delta_k, \eta_k, \lambda_k$ sont des scalaires déterminés à chaque itération k. Le paramètre δ_k est la longueur de pas optimal localement. Les paramètres η_k et λ_k sont déterminés de manière à garantir que $\tilde{H}(\mathcal{L}(\psi^{(k)}))$ est une matrice définie positive, c'est-à-dire toutes ces valeurs propres sont positives et que $\tilde{H}(\mathcal{L}(\psi^{(k)})) \underset{\psi^{(k)}\to\widehat{\psi}}{\longrightarrow} H(\mathcal{L}(\psi^{(k)}))$

Dans des cas complexes, cet algorithme a prouvé une meilleure convergence que Newton-Raphson et a été implémenté dans R dans le package marqLevAlg. Le package nous sort des indicateurs de convergence pour m paramètres et par défaut $\epsilon_a = \epsilon_b = \epsilon_d = 10^{-4}$:

— ca, critères de convergence pour la stabilisation des paramètres :

$$\sum_{j=1}^{m}(\psi_j^{(k+1)} - \psi_j^{(k)})^2 < \epsilon_a$$

— cb, critères de convergence pour la fonction de stabilisation :

$$|\mathcal{L}(\psi^{(k+1)}) - \mathcal{L}(\psi^{(k)})| < \epsilon_b$$

— rdm, critère de convergence sur les dérivées premières et secondes :

$$\frac{\nabla(\mathcal{L}(\psi^{(k)}))(H(\mathcal{L}(\psi^{(k)})))^{-1}\nabla(\mathcal{L}(\psi^{(k)}))}{m} < \epsilon_d$$

La fonction proposée par mla repose sur trois critères de convergence, convergence en paramètres, en log-vraisemblance et en distance relative au maximum de la log-vraisemblance (rdm). Ce dernier critère permet de s'assurer que l'algorithme ne converge pas vers des maximums locaux. Des scénarios où le critère rdm ne converge pas ou avec difficultés ont été identifiés :

— un paramètre de variance des effets aléatoires proche de 0 ;

— la fonction de log-vraisemblance présente un plateau à son maximum ;

Dans le cas de modèles complexes, le maximum peut se situer sur un plateau. La convergence en log-vraisemblance et en paramètres sont atteints au plateau mais le rdm reste non statisfait. Néanmoins, les paramètres de sorties peuvent être interprétés sachant que les autres critères de convergence sont satisfaits.
Dans ce rapport nous allons donc distinguer deux cas de convergence, (1) les trois critères sont satisfaits, (2) nous avons convergence pour la log-vraisemblance et les paramètres. Nous avons au moins une convergence de type (2) dans les modèles présentés et pour un temps de calcul limité à 30 heures maximum.

Après avoir vérifié l'ensemble de ces critères, une interprétation des paramètres estimés est possible.

2.2.3 Programme implémenté sous R

Les entrées du programme

Pour commencer, l'utilisateur devra avoir recours à deux fonctions internes au programme qui ont été créés afin de faciliter l'appel du modèle.
Nous avons d'abord ftime pour spécifier les fonctions du temps F dans les équations X et Y (F_E, F_{ER}, F_Y et F_{YR}) qui prend comme paramètres :

- *time* la variable temps du modèle,
- *scale* l'échelle du temps (par défault 1)
- *type* qui prend "mSpline" ou "poly", des splines ou la fonction polynomiale,
- *df* les degrés de liberté,
- *degree* le degré (par défault à 3) des splines ou de la fonction polynomiale,
- *knots* les noeuds internes des splines,
- *Boundary.knots* noeuds externes des splines

La fonction ftime va nous rendre une liste comprenant tous ces paramètres d'entrée, au minimum l'utilisateur doit spécifier *time* et *type*. Dans le corps du programme cette fonction va nous permettre de construire les fonctions du temps.

La seconde fonction présente dans les entrées est WCIE qui permet de spécifier la forme des poids du WCIE. Les arguments de WCIE sont similaires à ceux de ftime avec *type* ("nSpline" ou "bSpline"), *df*, *degree*, *knots*, *Boundary.knots* et *theta* les coefficients initiaux associés aux splines. Cette fonction nous permettra dans le corps du programme de spécifier les poids de l'exposition cumulée sur le marqueur.

L'estimation du modèle joint se fait par l'appel à la fonction WCIEconjoint qui dépend des arguments suivants :

- *data*, la base de données utilisée en data.frame.
- *subject*, la variable d'identification des sujets, entre guillemets(").
- $fixedE$, une formule avec la variable réponse d'exposition à gauche et à droite, les covariables explicatives de l'exposition. Par défaut, un intercept est inclus et il est nécessaire d'inclure ftime.
- $fixedYI$, une formule avec la variable réponse du marqueur à gauche et à droite, les covariables explicatives du niveau intial du marqueur. Par défaut, un intercept est

inclus et il est nécessaire d'inclure la fonction WCIE .

— $fixedDY$, une formule unilatérale avec les covariables explicatives de la pente du marqueur. Par défaut, un intercept est inclus et il est nécessaire d'inclure les fonctions ftime et WCIE.
— $randomE$, une formule unilatérale avec les covariables ayant un effet aléatoire sur l'exposition. Par défaut un intercept est inclus, si aucun intercept est souhaité écrire -1 en premier et il est nécessaire d'inclure la fonction ftime.
— $randomDY$, une formule unilatérale avec les covariables ayant un effet aléatoire sur la pente du marqueur. Par défaut un intercept est inclus, si aucun intercept est souhaité écrire -1 en premier et il est nécessaire d'inclure la fonction ftime.
— S, un nombre supérieur ou égal à 0 correspond à l'amplitude de la fenêtre d'exposition désirée.
— $baselineE$, la valeur de notre temps t=0 pour l'exposition, par défault il est NULL et sera déterminé d'après la base de donnée.
— $binit$, un vecteur contenant les valeurs initiales pour le modèle conjoint, soit $(\beta_0, \beta, \alpha, \delta_0, \delta, \mathbb{V}_t, \sigma^2_\varepsilon, \sigma^2_\epsilon)^T$ où $\mathbb{V}_t$ est un vecteur composé des termes de la matrice de covariance.
— $posfix$, un vecteur composé des indices du vecteur $(\theta_0, \theta_1, \text{binit})^T$ que l'on ne souhaite pas estimer, $posfix$ est fixé à NULL par défaut.

La fonction $ftime$ est nécessaire pour certains appels, cependant si l'utilisateur ne souhaite pas de fonction du temps il lui suffit d'appeler ftime avec degree=0. Nous retrouvons classiquement les objets $fixed$ et $random$ présents dans les fonctions lme, lmer et hlme. De plus, les paramètres d'entrée de la fonction d'optimisation mla peuvent être renseignés.

Le corps du programme

La première partie du programme se focalise sur la lecture des paramètres d'entré et l'évaluation de leur type et spécificités. Il s'agit de vérifier le type des variables et de vérifier les contraintes associées. Ces vérifications sont effectuées sur les entrées de WCIEconjoint et de ftime et WCIE utilisés dans les appels. Après cette vérification, nous avons créé les tables de l'exposition et du marqueur avec respectivement les variables présentes dans leurs effets fixes et aléatoires. Puis, nous avons appliqué une sélection de données sur ces deux tables :

— un sujet est gardé dans le dataset s'il a au moins une mesure de l'exposition avant la dernière mesure du marqueur ;

— un sujet est gardé s'il a au moins une mesure du marqueur à partir de $0_Y = 0_E + S$ (déterminé en entrée ou d'après la base de données) ;
— après ces restrictions, le sujet doit être présent dans les deux tables de données sinon il est supprimé.

La seconde partie du programme consiste à créer les matrices des effets fixes, des effets random présentées dans l'équation (8). Nous calculons d'abord les termes indépendants de WCIE, nommés XE, XY, ZE et ZY et définis comme :

$$\begin{pmatrix} \underbrace{{X_{iE}}^\top \otimes F_E(t)^\top}_{XE} & 0 \\ 0^\top & \underbrace{{X_{i0}}^\top \quad {X_{i1}}^\top \otimes \Delta\mathbb{F}_Y(t)^\top}_{XY} \end{pmatrix} \qquad \begin{pmatrix} \underbrace{F_{ER}(t)^\top}_{ZE} & 0 \\ 0^\top & \underbrace{1 \quad \Delta\mathbb{F}_{YR}(t)^\top}_{ZY} \end{pmatrix}$$

Dans chacune de ces matrices XE, XY, ZE et ZY nous avons d'abord l'intercept s'il est nécessaire/spécifié, les variables explicatives, le temps, l'intéraction entre le temps et un sous-ensemble de variables explicatives. Par défaut, nous avons au minimum l'intercept présent dans ces matrices.

En suivant, la fonction `loglike_WCIE_init` prend le relai, cette fonction est effectuée systématiquement (soit indépendamment de l'entrée pour *binit*). Cette fonction prend comme paramètres d'entrée les tables de données de Y et E, ainsi que XE, XY, ZE et ZY. Grâce à ces matrices nous pouvons calculer les modèles pour l'exposition E et le marqueur Y en les considérant indépendants avec la fonction hlme classique. Cela va nous permettre d'avoir une première estimation des paramètres fixes et des paramètres de variance. La fonction va retourner les valeurs initiales des paramètres des modèles indépendants dans un vecteur et les indices dans ce même vecteur des paramètres de covariance entre E et Y à ne pas estimer (fixé à 0).

L'avant dernière étape de notre programme est de créer la fonction sur laquelle l'algorithme de Marquant-Levenberg (fonction mla) va tourner. Nous nommons cette fonction `loglike_WCIE`, elle prend en paramètres d'entrée b le vecteur contenant l'ensemble des valeurs des paramètres pour calculer la log-vraisemblance. La fonction retourne pour un b donné la log-vraisemblance du modèle associé.

Pour se faire il calcule les termes dépendants de WCIE à partir des *theta* contenus dans *b*. Cela nous permet d'avoir les matrices complètes suivantes :

$$\begin{pmatrix} XE & 0 \\ \xi_{iE1}(t)^\top & XY \end{pmatrix} \qquad \begin{pmatrix} ZE & 0 \\ \kappa_{iE}(t)^\top & ZY \end{pmatrix}$$

Dans un premier temps nous avons utilisé La fonction hlme pour erreurs hétérocédastes sur ce modèle multivarié complet sans optimisation pour retourner la log-vraisemblance correspondante. Puis par des soucis d'optimisation de temps de calcul nous avons créé une fonction spécifique loglikF qui retourne la log-vraisemblance du modèle sans optimisation.

La dernière étape est la partie centrale du programme qui appelle la fonction mla pour maximiser la log-vraisemblance loglike_WCIE avec *boptim* vecteur contenant les valeurs des paramètres à optimiser. Les sorties du programme sont basées sur les sorties *mla* :

- *cl*, le résumé de l'appel de la fonction marqLevAlg ;
- *ni*, le nombre d'itérations de marqLevAlg avant le critère d'arrêt ;
- *istop*, le statut de convergence valant 1 si la convergence est atteinte, 2 si le maximum d'itérations est atteint et 4 si problème computationel ;
- *v*, la matrice triangulaire supérieure de la matrice de covariance estimée des paramètres ;
- *grad*, le gradient au point d'arrêt ;
- $fn.value$, fonction d'évaluation au point d'arrêt (la log-vraisemblance) ;
- *b*, les paramètres estimés au point d'arrêt ;
- *ca*, critères de convergence pour la stabilité des paramètres ;
- *cb*, critères de convergence pour la stabilité de la log-vraisemblance ;
- *rdm*, critère de convergence pour les dérivées premières et secondes ;
- *time*, le temps de calcul de la fonction.

Les interprétations du modèle portent principalement sur *b* et *v* à condition que les critères de convergence *istop*, *ca*, *cb* et *rdm* soient vérifiés.

2.3 Calculs a posteriori

Pour permettre une meilleure compréhension des modèles estimés, nous avons utilisé différentes quantités calculées a posteriori à partir des paramètres estimés $\widehat{\psi}$ et leur variance $\widehat{V}(\widehat{\psi})$.

2.3.1 Trajectoires des poids

La trajectoire estimée des poids $w_0(s)$ et $w_1(s)$ est déduite des paramètres $\widehat{\theta_.}$ et leur variance $\widehat{V}(\widehat{\theta_.})$ où $\widehat{\theta_.} = (\widehat{\theta_{.0}}, \widehat{\theta_{.1}}, \ldots, \widehat{\theta_{.K}})$. Ainsi pour tout $s = 0, \ldots, -S$, la trajectoire estimée de $w_.(s)$ est :

$$\widehat{w_.(s)} = \sum_{k=0}^{K} \widehat{\theta_{.k}} B_k(s), \text{ avec } B_0(s) = 1$$

et son intervalle de confiance à 95% est obtenu par approximation normale à partir de la variance estimée :

$$\begin{aligned} \widehat{V}(\widehat{w_.(s)}) &= \widehat{V}(\sum_{k=0}^{K} \widehat{\theta_{.k}} B_k(s)) \\ &= (1, B_1, \ldots, B_K)\widehat{V}(\widehat{\theta_.})(1, B_1, \ldots, B_K)^\top \end{aligned}$$

Nous pouvons remarquer que la variance de $w_.$ évolue avec s.

2.3.2 Effet moyen de l'exposition sur la période S

Pour résumer l'association entre l'exposition et le marqueur sur l'intervalle $[-S, \ldots, 0]$, nous pouvons calculer l'effet global estimé $\widehat{\omega_.} = \sum_{s=-S}^{0} \widehat{w_.(s)}$ et sa variance $\widehat{V}(\widehat{\omega_.})$:

$$\begin{aligned} \widehat{V}(\widehat{\omega}) &= \widehat{V}(\sum_{s=-S}^{0} \widehat{w_.(s)}) \\ &= \sum_{s=-S}^{0} \widehat{V}(\sum_{k=0}^{K} \widehat{\theta_{.k}} B_k(s)) \\ &= (\sum_{s=-S}^{0} 1, \sum_{s=-S}^{0} B_1, \ldots, \sum_{s=-S}^{0} B_K)\widehat{V}(\widehat{\theta_.})(\sum_{s=-S}^{0} 1, \sum_{s=-S}^{0} B_1, \ldots, \sum_{s=-S}^{0} B_K)^\top \end{aligned}$$

2.3.3 Trajectoires prédites du marqueur

Nous désirons obtenir des trajectoires prédites de notre marqueur Y pour des profils d'évolution donnés de l'exposition E. Pour se faire, nous supposons connaître la valeur réelle de la trajectoire d'exposition notée $\tilde{E}^*$, et à partir de notre modèle nous pouvons en déduire l'espérance prédite de notre marqueur Y :

$$\begin{aligned}\widehat{E}(Y_i(t)|\tilde{E}_i^*,\widehat{\psi}) &= {X_{i0}}^\top\widehat{\alpha}+\sum_{s=-S}^{0}\widehat{w_0(s)}\tilde{E}_i^*(S+s)\\ &\quad+\int_S^t {X_{i1}}^\top\otimes F_Y(l)^\top dl\widehat{\delta}+\int_S^t\sum_{s=-S}^{0}\widehat{w_1(s)}\tilde{E}_i^*(l+s)dl\\ &= \mathcal{X}_{i\tilde{E}^*}(t)^\top\widehat{\begin{pmatrix}\alpha\\ \theta_0\\ \delta\\ \theta_1\end{pmatrix}}\end{aligned}$$

avec ${\theta_0}^\top=(\theta_{0k})_k$ et ${\theta_1}^\top=(\theta_{1k})_k$, pour k allant de 0 à K

$$\mathcal{X}_{i\tilde{E}^*}(t)^\top=(X_{i0},(\sum_{s=-S}^{0}B_k(s)\tilde{E}_i^*(S+s))_k,\int_S^t{X_{i1}}^\top\otimes F_Y(l)^\top dl,(\sum_{s=-S}^{0}B_k(s)\int_S^t\tilde{E}_i^*(l+s))_k)$$

Les scénarios considérés pour l'évolution de l'exposition $\tilde{E}^*$ sont une non-variation de $\tilde{E}^*$, une évolution linéaire positive et négative. Nous considérons également différents niveaux initiaux de l'exposition, sélectionnés au vu de sa distribution.

Il est aussi possible de contraster les trajectoires obtenues à partir de profils d'exposition différents. Nous notons E_1 et E_2 deux profils d'exposition différents et supposons que les autres variables explicatives sont fixées à la même valeur X_0 et X_1. La différence estimée de trajectoire du marqueur associée aux deux profils et sa variance sont donc :

$$\widehat{E}(Y(t)|\tilde{E}_1,\widehat{\psi})-\widehat{E}(Y(t)|\tilde{E}_2,\widehat{\psi})=(\mathcal{X}_{\tilde{E}_1}(t)^\top-\mathcal{X}_{\tilde{E}_2}(t)^\top)\widehat{\begin{pmatrix}\alpha\\ \theta_0\\ \delta\\ \theta_1\end{pmatrix}}$$

$$\widehat{V}(\widehat{E}(Y(t)|\tilde{E}_1,\widehat{\psi}) - \widehat{E}(Y(t)|\tilde{E}_2,\widehat{\psi})) = (\mathcal{X}_{\tilde{E}_1}(t)^\top - \mathcal{X}_{\tilde{E}_2}(t)^\top)\widehat{V}\begin{pmatrix}\widehat{\begin{pmatrix}\alpha\\ \theta_0\\ \delta\end{pmatrix}}\\ \theta_1\end{pmatrix}(\mathcal{X}_{\tilde{E}_1}(t)^\top - \mathcal{X}_{\tilde{E}_2}(t)^\top)^\top$$

Nous représentons les prédictions moyennes de Y pour divers scénarios via les courbes prédites de Y avec leurs intervalles de confiance à 95%. Nous évaluons aussi les différences moyennes entre ces courbes graphiquement avec leurs intervalles de confiance à 95%. En complément, nous rapportons les prédictions moyennes et leurs différences pour des temps d'analyses t particuliers.

3 Application

3.1 Données de l'étude Trois Cités

Les données sur lequelles nous appliquons notre modèle sont les données issues de la cohorte des trois cités [14]. Cette cohorte avait pour objectif l'étude des facteurs de risques vasculaires sur le vieillissement cérébral dans la population âgée de 65 ans et plus [15]. Les sujets éligibles ont été tirés au sort sur les listes électorales de Bordeaux, Dijon et Montpellier. L'étude a débuté en 1999 avec environ 10 000 sujets qui ont été suivis pendant 17 ans. Un grand nombre de données sont collectées via entretiens et examens tous les 2/3 ans.

Nous allons nous intéresser spécifiquement à la relation entre l'Indice de Masse Corporel (IMC), qui mesure d'adiposité, et l'évolution cognitive. L'IMC est un facteur modifiable déjà identifié comme associé au vieillissement cognitif. La relation entre l'IMC et l'évolution cognitive est contradictoire suivant la fenêtre temporelle étudiée [16, 17] : les premiers travaux suggèrent qu'un IMC élevé à des âges jeunes est associé à un plus grand risque de démence et de déclin cognitif, alors qu'un IMC élevé à des âges plus avancés a été retrouvé associé à un moindre déclin cognitif. L'IMC comme facteur protecteur aux âges avancés pourrait être dû à un phénomène de causalité inverse [18, 17, 10]. On dit qu'il y a causalité inverse quand une maladie sous-jacente induit un changement de comportement qui altère l'exposition, ici l'IMC. Ainsi, la temporalité, comme notre méthode permet de le faire, est essentielle à prendre en compte dans l'impact de l'histoire de l'exposition de l'IMC sur l'évolution cognitive.

Pour la mesure du niveau cognitif, nous avons choisi de nous focaliser sur le Set Test d'Isaacs (IST) [19] qui est un test de fluence verbale [19]. L'IST a montré de bonnes performances de détection précoce du processus de déclin cognitif [20, 4].
Pour cette analyse, nous travaillons sur l'âge comme base de temps et considérons une fenêtre d'exposition d'IMC de S=10 ans en amont de l'évaluation cognitive. Le choix de cette fenêtre se base sur le fait que nous avons 17 années de suivi, fixer S=10 nous permet d'avoir une antériorité substantielle tout en nous assurant 7 années de suivi cognitif au niveau individuel. Dans l'ensemble nous avons une amplitude d'environ 30 années entre 75 ans (65+10) et 105 ans pour le suivi de l'IST.

Au sein de l'étude 3C nous nous sommes concentrés sur l'échantillon répondant aux critères suivants :

- le centre de Bordeaux ;
- au moins une mesure de l'IMC avant ou au moment de la dernière mesure de l'IST ;
- au moins une mesure de l'IST après $0_E + S$, soit 65,36+10=75,36 ans ;
- variables explicatives non manquantes : âge à l'inclusion dans l'étude (en années, nommé AGE_INIT), le sexe avec pour modalités hommes ou femmes [21] (hommes en référence et variable nommé $SEXE$), le gène APOE4 avec pour modalités non ou oui (non en référence, nommé $APOE4$) [22, 23] et le niveau d'étude avec pour modalités court ou long (court en référence, nommé $EDUC$) [24] ;

Le sexe est codé 0 - hommes et 1 - femmes, respectivement 0 - non et 1 - oui pour le gène APOE4. Le niveau d'études est considéré comme court si le sujet a un niveau d'éducation de type sans étude ou primaire avec ou sans diplôme ou secondaire court, codé 0. Un niveau d'étude long correspond à un secondaire long ou à des études universitaires, codé 1.

3.2 Descriptif du centre de Bordeaux de la cohorte 3C

A partir de l'échantillon de 9 294 sujets de la cohorte 3C, le centre ayant la majorité des sujets est celui de Dijon avec 4 931 (53,1%), puis vient Montpellier avec 2 259 sujets (24,3%) et enfin Bordeaux avec 2 104 sujets (22,6%).
Notre sélection est résumée dans le flow-chart en figure 2. Après l'exclusion des sujets ayant des données manquantes pour l'âge à l'inclusion et le sexe, ainsi que ceux n'ayant aucune données pour l'IST ou l'IMC, nous nous retrouvons à 1 810 sujets dans le centre de Bordeaux. Puis nous devons prendre en compte que S=10, ainsi 189 sujets ont leurs mesures de l'IST

exclusivement avant 75,36 ans ou leurs mesures de l'IMC exclusivement après la fin de mesure de l'IST. Notre sélection est composés de 1621 sujets.

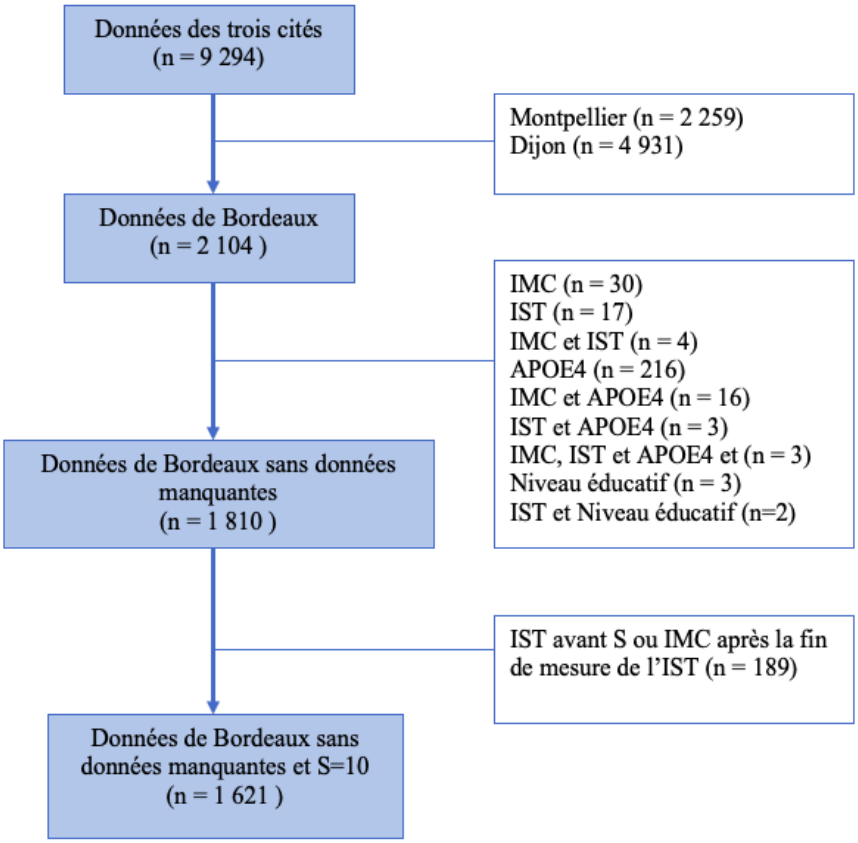

FIGURE 2: Flow-chart de la sélection des sujets de la cohorte 3C pour la modélisation conjointe.

Le tableau 1 décrit l'ensemble des caractéristiques des sujets de la cité de Bordeaux pour notre sélection et les autres centres. Nous avons un ensemble de 1 621 sujets. Les sujets comptent parmi eux 611 hommes (37,7%) et 1 010 femmes (62,3%). Parmi, l'ensemble des sujets, le plus jeune rentre dans l'étude à 65,4 ans et le plus âgé à 94,5 ans. La moyenne d'âge à l'inclusion est à 74,9 ans (sd=4,9) et 50% des sujets ont un âge à l'inclusion compris entre 71,1 et 78,5 ans. Parmi les sujets, 996 (61,4%) ont effectué des études courtes et 625 (38,6%) des études longues. La majorité des sujets ne présente pas le gène APOE4, 1 313 (81,0%) sujets.
Les sujets ont été vu au maximum 8 fois au cours des 17 années de suivi (12 808 observations).

Le score IST à 75 ans, IST initial, est en moyenne de 29,8 points (n=70 ;sd=6,4) et 50% des sujets ont un score à 75 ans compris entre 26 et 35. A la première visite, le score d'IST est en moyenne de 28,7 points (n=1621 ;sd=6,4) et 50% des sujets ont un score initial compris entre 24 et 33 points.

L'IMC à 75 ans est en moyenne de 26,4 kg/m^2 (n=371 ;sd=4,1) et 50% des sujets ont un IMC initial compris entre 23,6 et 28,5 kg/m^2. A la première visite, l'IMC est en moyenne de 26,4 kg/m^2 (n=1621 ;sd=4,1) et 50% des sujets ont un IMC initial compris entre 24 et 29 kg/m^2.
Le score d'IST et l'IMC à 75 ans sont obtenus en arrondissant à l'entier inférieur les âges observés.

Centre	Sélection Bordeaux	Bordeaux	Dijon	Montpellier	Tous centre
Nombre de sujets	1 621	2 104	4 931	2 259	9 294
Variables qualitatives n (%)					
Sexe					
Homme	611 (37,7)	816 (38,8)	1 888 (38,3)	946 (41,9)	3 650 (39,3)
Femme	1 010 (62,3)	1 288 (61,2)	3 043 (61,7)	1 313 (58,1)	5 644 (60,7)
Etudes					
courtes	996 (61,4)	1 320 (62,7)	3 162 (64,1)	1 233 (54,6)	5 715 (61,5)
longues	625 (38,6)	779 (37,1)	1 758 (35,7)	1 024 (45,4)	3 561 (38,3)
Gène APOE4					
Non	1 313 (81,0)	1 501 (71,3)	3 625 (73,5)	1 771 (78,4)	6 897 (74,2)
Oui	308 (19,0)	365 (17,3)	977 (19,8)	436 (19,3)	1 778 (19,1)
Varibales quantitatives moyenne (n ;sd)					
Age à l'inclusion (années)	74,9 (1621 ;4,9)	74,6 (2104 ;5,1)	74,6 (4931 ;5,7)	73,3 (2259 ;5,7)	74,3 (9264 ;5,6)
IST à 75 ans (points)	29,8 (70 ;6,8)	30,3 (457 ;6,6)	35,0 (1223 ;6,7)	33,3 (538 ;6,5)	33,7 (2218 ;6,9)
IST à l'inclusion (points)	28,7(16201 ;6,4)	29,1 (2035 ;6,5)	33,0 (4891 ;7,1)	30,0 (2204 ;6,8)	31,4 (9130 ;7,1)
IMC à 75 ans (kg/m^2)	26,4 (371 ;4,1)	26,3 (459 ;4,3)	25,6 (983 ;3,8)	25,3 (421 ;3,6)	25,7 (1863 ;3,9)
IMC à l'inclusion (kg/m^2)	26,4 (16201 ;4,1)	26,3 (2048 ;4,2)	25,6 (4896 ;4,1)	25,1 (2240 ;3,7)	25,7 (9184 ;4,1)

TABLE 1: Descriptif de la population de la cohorte des 3C par sous-ensembles.

L'étude 3C étant une étude de cohorte non interventionnelle nous ne nous attendons pas à avoir un effet du centre sur la santé des sujets. Cependant en nous référant au tableau 1, nous remarquons des disparités au niveau des centres de Montpellier et Dijon. Le centre de Montpellier comprend un plus grand nombre d'hommes, de sujets jeunes et de sujets ayant effectués des études longues. Le centre de Dijon a en son sein moins de personnes atteintes de démence au début du suivi que les autres centres.
Dans un souhait de représentativité de la population des trois cités et d'éviter l'effet centre, le choix de Bordeaux comme centre de notre analyse est conforté par les résultats précédents.

Pour finir, nous avons vérifié graphiquement (figure 3) la normalité de notre exposition l'IMC et de notre marqueur IST à 75 ans en vue de l'application du modèle linéaire mixte. Nous avons aussi décrit les trajectoires individuelles de l'IMC et du IST dans la figure 4. L'IST a tendance à diminuer plus le sujet avance en âge, il en va de même pour l'évolution du IMC.

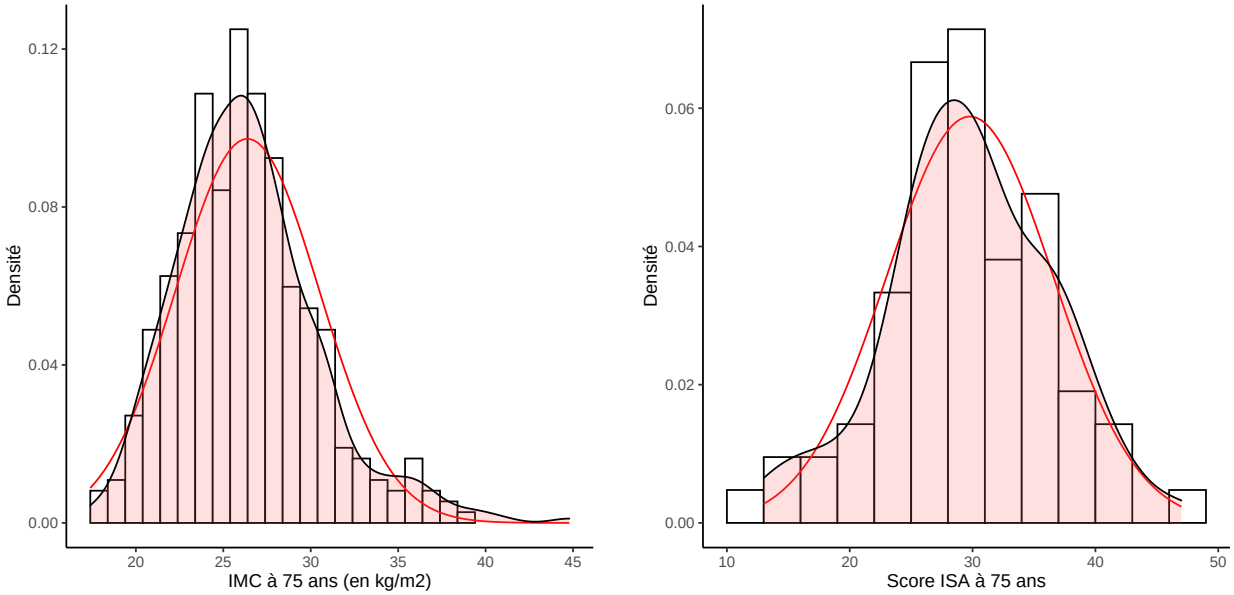

FIGURE 3: Histogramme de la distribution de l'IMC (à gauche,n=371) et du IST (à droite,n=70) pour les sujets à 75 ans de l'étude 3C dans la sélection de Bordeaux, distribution normale en rouge et distribution observée en rose pâle.

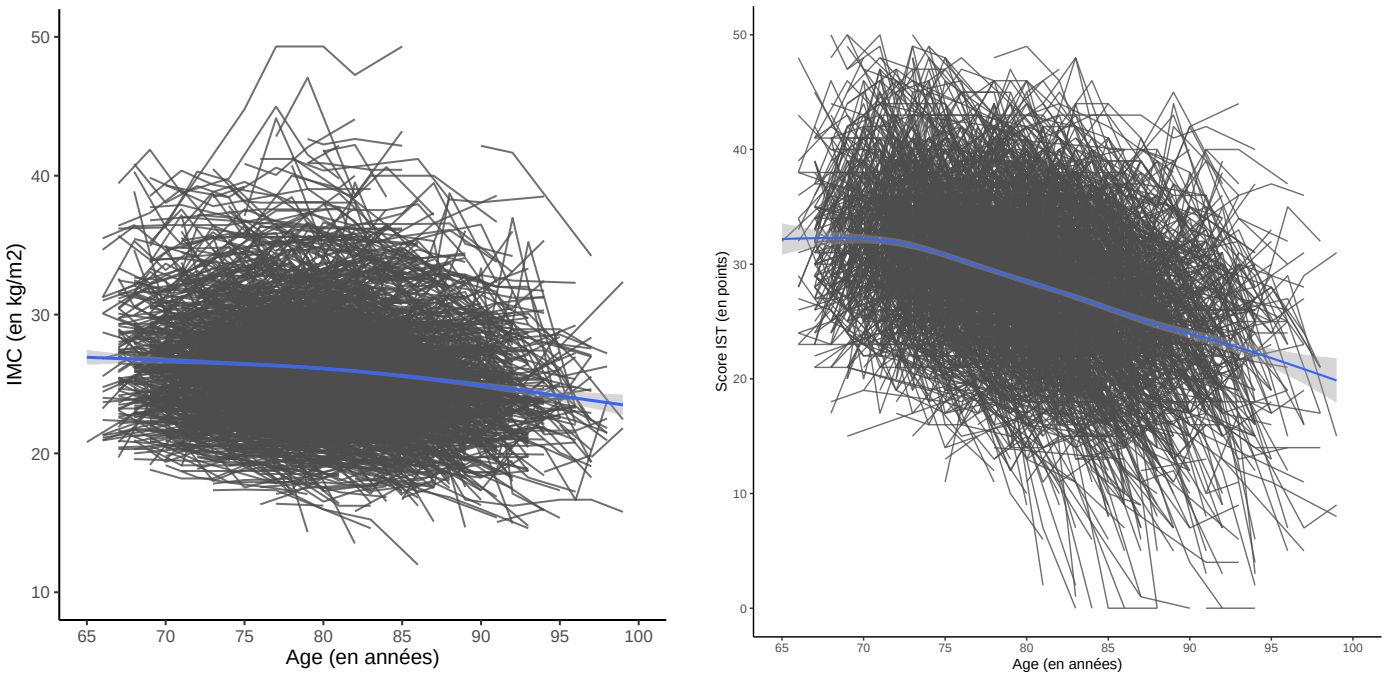

FIGURE 4: Evolution de l'IMC (à gauche) et du IST (à droite) pour les sujets de l'étude 3C dans la sélection de Bordeaux, lissage par splines cubiques (n=1 621)

3.3 Etudes préliminaires des trajectoires de l'IMC et l'IST

Dans un premier temps, nous souhaitons déterminer les modèles mixtes les plus adéquates pour l'IMC et l'IST pris séparément. Nous avons réalisé cette première étude sur l'ensemble de la cohorte des trois cités puis confirmé sur notre échantillon de Bordeaux. Cette première étape a pour visée de déterminer les fonctions du temps les plus pertinentes pour décrire l'évolution de chaque marqueur avant de poursuivre sur l'estimation de leurs associations via le modèle conjoint.

Les modèles analysés ont pour base de temps l'âge recentré à 64 ans et redimensionné en dizaine d'années (10^{-1}), avec les fonctions de l'âge suivantes :

- linéaire et quadratique ;
- Mspline avec degré allant de 1 à 3, avec 1 à 3 noeuds internes respectivement médiane (50%), terciles (33, 66%) quartiles (25,50 et 75%).

Pour notre application sur la sélection de Bordeaux avec 1 621 sujets, nous avons un ensemble de 8 410 observations pour l'IMC et respectivement 6 293 pour l'IST.
Pour comparer l'ensemble des modèles nous nous basons sur l'aboutissement à une convergence satisfaisante, la minimisation du critère d'information d'Akaike (AIC), la représentation de la fonction de l'âge et le fit individuel. Nous adoptons une méthode parcimonieuse, c'est-à-dire que si le gain entre deux modèles est faible nous allons nous orienter vers le modèle ayant le moins d'effets aléatoires à estimer.

3.3.1 Modélisation préliminaire de la trajectoire de l'IMC

Dans les modèles estimés sur la cohorte entière, l'ensemble des résultats comparatifs des modèles pour l'IMC sont présents dans le tableau 2. D'après ce tableau, dès que nous passons à des mspline de degré 1 à 2 noeuds les performances sont meilleures que celles d'un modèle quadratique. Nous observons une amélioration constante du modèle dès que le degré des msplines augmente. Cependant, les modèles avec des splines de degré 3 à 2 et 3 noeuds n'ont pas convergé correctement après 500 itérations de hlme (AIC=142 145 #param=28 ; AIC=142 113 #param 36). Alors que l'algorithme convergeait selon les paramètres et la log-vraisemblance, la Hessienne restait non-inversible. Ce problème de convergence peut s'expliquer par un très grand nombre d'effets aléatoires respectivement 6 et 7. De par leurs complexités ces deux modèles ont été exclus. Nous avons deux modèles qui présentent les meilleures performances : degré 2 avec 3 noeuds et degré 3 avec 1 noeud. Le gain d'AIC entre les deux modèles est faible en comparaison aux autres évolutions, seulement 22 points et nous avons 6 effets aléatoires en plus dans le modèle le plus performant. Les courbes de prédictions de l'IMC sont à retrouver en annexe. D'après la figure 17 en annexe, nous observons que la courbe de prédiction de l'IMC de ces deux modèles sont très similaires au cours de l'âge. Nous avons préféré rester parcimonieux et garder le modèle avec des splines de degré 3 et 1 noeud internes.

Modèle	Quadratique	Degré 1, 1 noeud	Degré 1, 2 noeuds	Degré 1 3 noeuds	Degré 2 1 noeud	Degré 2, 2 noeuds	Degré 2, 3 noeuds	**Degré 3, 1 noeud**
AIC	142 801	142 981	142 681	142 541	142 449	142 324	142 226	**142 248**
Loglike	-71 3901	-71 480	-71 325	-71 250	-71 209	-71 141	-71 085	**-71 103**
# paramètres	10	10	15	21	15	21	28	**21**
# effets aléatoires (param)	3 (6)	3 (6)	4 (10)	5 (15)	4 (10)	5 (15)	6 (21)	**5 (15)**

TABLE 2: Résumé des modèles linéaires mixtes pour l'IMC sur l'ensemble de la Cohorte des troisCités (n=9 185 ;# observations=33 824).

Dans l'échantillon de Bordeaux (n=1 621), nous reprenons les modèles précédents jusqu'au degré 2 à 2 noeuds (voir tableau 3). Le modèle le plus pertinent est le modèle avec des splines de degré 2 à 2 noeuds internes, il présente le meilleur AIC=35 723, soit une amélioration de 233 points par rapport au modèle quadratique. Ce modèle comprend 5 effets aléatoires sur la fonction mspline de l'âge soit 15 paramètres liés aux effets aléatoires.

Modèle	Quadratique	Degré 1, 1 noeud	Degré 1, 2 noeuds	Degré 1 3 noeuds	Degré 2 1 noeud	**Degré 2, 2 noeuds**
AIC	35 956	35 982	35 849	35 824	35 776	**35 723**
Loglike	-17 968	-17 981	-17 909	-17 891	-17 873	**-17 840**
# paramètres	10	10	15	21	15	21
# effets aléatoires (#param)	3 (6)	3 (6)	4 (10)	5 (15)	4 (10)	**5 (15)**

TABLE 3: Résumé des modèles linéaires mixtes pour l'IMC dans la sélection du centre de Bordeaux avec S=10 (n=1 621 ;# observations=8 410).

3.3.2 Modélisation préliminaire de la trajectoire de l'IST

Sur l'ensemble de la cohorte, les résultats comparatifs des modèles pour l'IST sont présentés dans le tableau 4. Les modèles avec des msplines de degré 1 ont des performances moindres ou égales au modèle quadratique. Parmi les modèles avec des splines de degré 2, seulement celui avec 1 noeud interne a convergé correctement après 500 itérations.
Parmi les modèles ayant convergé avec une Hessienne inversible, le modèle quadratique et mspline de degré 2 à 1 noeud interne présentent les meilleures performances. Les courbes de prédiction de l'IST sont à retrouver en annexe, dans la figure 18. La prédiction des modèles et le fit de ces modèles sont très proches, ils se distinguent par un écart de 5 paramètres et de 75 points d'AIC. Nous avons donc décidé de garder le modèle quadratique. Il est classiquement utilisé dans le cadre de la modélisation du IST et présente un très bon compromis entre AIC minimisé et nombre limité d'effets aléatoires.

Modèle	Linéaire	**Quadratique**	Degré 1, 1 noeud	Degré 1, 2 noeuds	Degré 1 3 noeuds	Degré 2 1 noeud
AIC	243 205	**242 035**	242 261	242 090	242 035	241 960
Loglike	-121 596	**-121 008**	-121 120	-121 030	-120 997	-120 965
# paramètres	6	**10**	10	15	21	15
# effets aléatoires (#param)	2 (3)	**3 (6)**	3 (6)	4 (10)	5 (15)	4 (10)

TABLE 4: Résumé des modèles linéaires mixtes pour l'IST sur l'ensemble de la Cohorte des trois cités (n=9 243 ;# observations=39 315).

Pour les modèles dans la sélection de Bordeaux (n=1 621), nous proposons de réévaluer le quadratique et le modèle de degré 2 à 1 noeud pour l'IST (voir tableau 5). Précédemment les deux modèles avaient des performances similaires avec un écart d'AIC en faveur du modèle de degré 2 à 1 noeud. Sur les données de Bordeaux et pour S=10, la différence de score d'AIC est nulle, les deux modèles ont un AIC à 38 479. Ces résultats confortent notre choix de garder le modèle quadratique pour modéliser IST dans la sélection de Bordeaux.
Par la suite, nous avons voulu voir si l'effet aléatoire sur l'âge en quadratique était pertinent dans le modèle. Ainsi, nous avons comparé le modèle précédent à un modèle où les effets aléatoires sont linéaires. Le modèle avec des effets aléatoires linéaires présente un AIC à 38 488 pour 7 paramètres dont 3 pour les effets aléatoires. Ainsi, en passant à ce modèle linéaire pour les effets aléatoires nous augmentons notre AIC de 9 points mais diminuons nos paramètres de 3. Afin de rester dans une démarche parcimonieuse, nous allons préférer ce dernier modèle pour la suite de notre modélisation.

Modèle	**Quadratique**	Degré 2 1 noeud
AIC	**38 479**	38 479
Loglike	**-19 229**	-19 224
# paramètres	**10**	15
# effets aléatoires (#param)	**3 (6)**	4 (10)

TABLE 5: Résumé des modèles pour l'IST dans le centre de Bordeaux avec S=10 (n=1 621 ;# observations=6 293).

3.4 Modèle conjoint pour estimer l'association temporelle entre l'IMC et l'IST

3.4.1 Ecriture du modèle

La partie précédente nous a permis d'identifier les modèles indépendants les plus adéquates pour spécifier la trajectoire de l'IST et de l'IMC avec l'âge. Cependant nous avons rencontré des difficultés récurrentes de convergence sur le critère rdm si bien que nous avons simplifié les modèles à présenter. Le modèle de l'IST est resté inchangé mais nous avons retenu au final le modèle quadratique pour l'IMC, ce choix permet de réduire le nombre de paramètres de 11 dont 9 paramètres d'effets aléatoires.

L'objectif du modèle conjoint est d'étudier la relation entre l'histoire d'exposition IMC et la pente du score IST, ajusté sur des covariables. Le modèle considère un effet simple du WCIE de l'IMC sur la pente du score IST, ainsi nous posons la même hypothèse pour les variables d'ajustement. Les variables d'ajustement ont été choisies d'après la littérature, soit l'âge à l'entrée dans l'étude, le gène APOE4, le sexe et le statut éducatif. Ainsi, notre modèle peut s'écrire sous la forme suivante :

Pour tous sujets i =1, ..., 1 621 à la mesure j et $AGE_{ij} \geq 0_{IMC}$,

$$\begin{aligned} IMC_{ij} &= IMC^*(AGE_{ij}) + \varepsilon(AGE_{ij}) \\ &= \beta_0 + \beta_1 AGE_{ij} + \beta_2 {AGE_{ij}}^2 + b_{i0} + b_{i1} AGE_{ij} + b_{i2} {AGE_{ij}}^2 + \varepsilon_{ij} \end{aligned}$$

avec, $b_i \sim N(0, B)$, $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$,

pour tous sujets i et âges, $b_i \perp\!\!\!\perp \varepsilon_{ij}$,

pour tous sujets i et âges tel que $AGE_{i1} \neq AGE_{i2}$, alors $\varepsilon_{i1} \perp\!\!\!\perp \varepsilon_{i2}$.

Pour tous sujets i =1, ..., 1 621 à la mesure j et $AGE_{ij} \geq 0_{IST} = 0_{IMC} + 10$,

$$IST_{ij} = IST^*(AGE_{ij}) + \epsilon_{ij},$$

$$\begin{cases} IST_i^*(0_{IST}) & = \alpha_0 + \alpha_1 AGE\_INIT_i + \alpha_2 APOE4_i + \alpha_3 SEXE_i + \alpha_4 EDUC_i \\ & \quad + u_{i0} + \underbrace{\sum_{s=-10}^{0} w_0(s) IMC_i^*(0_{IST}+s)}_{WCIE_i(0_{IST}-10;0_{IST})} \\ \frac{\delta IST^*(AGE_{ij})}{\delta(AGE_{ij})} & = \delta_0 + \delta_1 APOE4_i + \delta_2 SEXE_i + \delta_3 EDUC_i + \delta_4 AGE_{ij} \\ & \quad + u_{i1} + \underbrace{\sum_{s=-10}^{0} w_1(s) IMC^*(AGE_{ij}+s)}_{WCIE(AGE_{ij}-10;AGE_{ij})} \, , \forall \, AGE_{ij} > 0_{IST} \end{cases}$$

avec, $u_i = (u_{0i}, u_{1i}) \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, D = \begin{pmatrix} \sigma_0^2 & C^t \\ C & C_1^2 \end{pmatrix})$, $\epsilon(AGE_{ij}) \sim N(0, \sigma_\epsilon^2)$,

pour tous sujets i et âges, $u_i \perp\!\!\!\perp \epsilon_{ij}$,

pour tous sujets i et âges tel que $AGE_{i1} \neq AGE_{i2}$, alors $\epsilon_{i1} \perp\!\!\!\perp \epsilon_{i2}$.

3.4.2 Estimations jointes pour différentes spécifications de l'association

Nous souhaitons analyser notre modèle pour différentes hypothèses sur le $WCIE$. Les modèles présentés vont donc varier dans leurs hypothèses du $WCIE$ sur le niveau initial de IST et sur la pente de IST. L'objectif principal de notre modèle étant de modéliser la relation entre l'IMC et la pente de l'IST, nous allons focaliser notre attention sur $w_1(s)$ plutôt que $w_0(s)$. Les WCIE présentés sont les suivants :

- NULL, càd $w_.(s) = 0$,
- CIE, càd $w_.(s) = \theta_{.0}$,
- WCIE ns, càd $w_.(s) = \theta_{.0} + \sum_{k=1}^{K} \theta_{.k} N_k(s)$ avec $(N_k(.))$ une base de splines naturelles cubiques (ns)
- WCIE bs, càd $w_.(s) = \theta_{.0} + \sum_{k=1}^{K} \theta_{.k} B_k(s;0)$ avec $(B_k(.;d))$ une base de B-splines de degré d (bs).

Pour l'ensemble des modèles le nombre de paramètres d'effets aléatoires reste constant à 5, soit 9 paramètres d'effets aléatoires (3 pour l'IST et 6 pour l'IMC puisque les effets aléatoires de l'IST et l'IMC sont mutuellement indépendants). Le nombre d'effets fixes va augmenter avec la complexification du WCIE allant de 24 à 28 paramètres. La complexification des modèles porte sur le nombre de noeuds, pour un noeud celui-ci est fixé à -5, pour deux noeuds ils sont fixés à -7 et -3. Les fonctions splines utilisées pour modéliser les poids sont présentés dans les figures en annexe. L'inversibilité de la hessienne est indiquée pour chaque modèle, dans tableau 6.

w_0	w_1	AIC	Loglike	#param	Inversion
NULL	NULL	74 283,08	-37 117,54	24	Oui
CIE	NULL	74 283,16	-37 116,58	25	Oui
CIE	CIE	74 284,92	-37 116,46	26	Oui
CIE	WCIE, ns 0 noeud	74 274,78	-37 110,39	27	Oui
WCIE, ns 0 noeud	WCIE, ns 0 noeud	74 276,54	-37 110,27	28	Oui
CIE	WCIE, ns 1 noeud	74 277,44	-37 110,72	28	Oui
CIE	WCIE, ns 2 noeuds	74 279,28	-37 110,64	29	Non
CIE	WCIE, bs degré 0 et 3 noeuds (bs par morceaux)	74 279,00	-37 110,50	29	Non
CIE	WCIE, bs degree 1 et 2 noeuds	74 279,58	-37 110,79	29	Non

TABLE 6: Résumé des caractéristiques des modèles conjoints WCIE de IST et IMC (n=1621 ;# observations =14 703)

Pour w_0, on remarque que l'AIC augmente quand nous passons à un modèle WCIE, ns 0 noeud, respectivement 74 276,54 versus 74 274,78 pour le modèle CIE avec w_1 WCIE, ns 0 noeud. Ce résultat suggère une évolution constante des poids sur [65 ;75] ans, soit l'hypothèse du CIE est valide.
Pour $w_0(s)$ sur le niveau intial d'IST fixé à un CIE, $\forall s$, $w_0(s) = \theta_{00}$. Pour ces modèles, la valeur de θ_{00} est inchangée, nous obtenons $\theta_{00} = -0,006$ ($IC_{95\%} = [-0,015; 0,003]$). Ainsi, le niveau intial du score IST à 75 ans n'est pas statistiquement associé à l'histoire de l'IMC sur [65 ;75] ans.

Pour w_1, on note une amélioration du critère d'AIC en passant d'un modèle où w_1 CIE à w_1 WCIE ns avec hessienne inversible, respectivement 74 284,92 et 74 274,78 ou 74 277,44. Ces résultats suggèrent que le modèle WCIE pour les poids w_1 est préférable à un CIE, soit que les poids sont non constants sur les dix dernières années au cours du suivi.
Pour $w_1(s)$ considéré constant pour tous s, soit CIE, le coefficient $\theta_{10} = 0,002$ ($IC_{95\%} =$

$[-0,009; 0,013]$) n'est statistiquement différent de 0. Ainsi, la pente instantanée du score IST n'est pas statistiquement associée à l'histoire de l'IMC sur les dix dernières années en CIE.
Puis quand nous considérons $w_1(s)$ évolutif grâce à des splines naturelles cubiques (ns), $\forall s$, $w_1(s) = \theta_{10} + \sum_{k=1}^{K} \theta_{1k} N_k(s)$, au moins un coefficient θ_{1k} est statistiquement différent de 0. Il est important de signaler que quand nous passons de "0 noeud" à "1 noeud", la fonction $N_1(s)$ est différente. Nous ne rajoutons pas seulement 1 noeud mais nous faisons évoluer la base des fonctions splines, soit l'ensemble des fonctions. De ce fait le paramètre θ_{11} est tantôt significativement différent de 0 dans le modèle WCIE 0 noeud puis non significativement différent de 0 dans le modèle WCIE avec 1 noeud.

Pour une meilleure compréhension des coefficients des poids $w_1(s)$, nous allons les représenter pour $s = -10, \ldots, 0$ dans la figure 5 pour les modèles CIE, WCIE ns 0 noeud et WCIE ns avec 1 noeud. Dans ces modèles la hessienne est inversible, ainsi nous allons pouvoir représenter les intervalles de confiance à 95%.
Un modèle WCIE ns 0 noeud interne revient à considérer une évolution linéaire croissante des poids. Ainsi, dans les dix dernières années au cours du suivi, de la 10^{e} à la 6^{e} années les poids sont statistiquement inférieurs à 0, respectivement -0,196,-0,157,-0,118,-0,078 et-0,039. La 5^{e} année a un poids statistiquement non différent de 0. Puis, de la 4^{e} à l'année actuelle les poids sont statistiquement supérieurs à 0, respectivement 0,039, 0,078, 0,118, 0,157 et 0,196. Pour WCIE ns avec 1 noeud, l'évolution des poids diffère du modèle WCIE ns 0 noeud. Le poids de la 10^{e} année n'est pas statistiquement différent de 0. Puis les poids deviennent statistiquement négatifs de la 9^{e} à la 4^{e} années, respectivement -0,069, -0,109, -0,135, -0,142, -0,121 et -0,069. La 3^{e} année a un poids non statistiquement différent de 0. Enfin de la 2^{e} année à l'année actuelle, les poids sont statistiquement supérieurs à 0, respectivement 0,109, 0,221 et 0,340.

Les modèles WCIE ns permettent de détecter les deux phénomènes souhaités contrairement au CIE qui ne détecte pas d'association entre la pente du score IST et l'histoire de l'IMC. Les poids du WCIE ns montrent d'abord qu'un IMC plus élevé est un facteur de risque du déclin cognitif plus en amont de l'évaluation cognitive. Puis, un IMC bas est associé à un déclin plus important de cognition à l'approche de l'évaluation cognitive. Cela va dans le sens de la causalité inverse avec un effet sur l'IMC associé à une maladie sous-jacente.

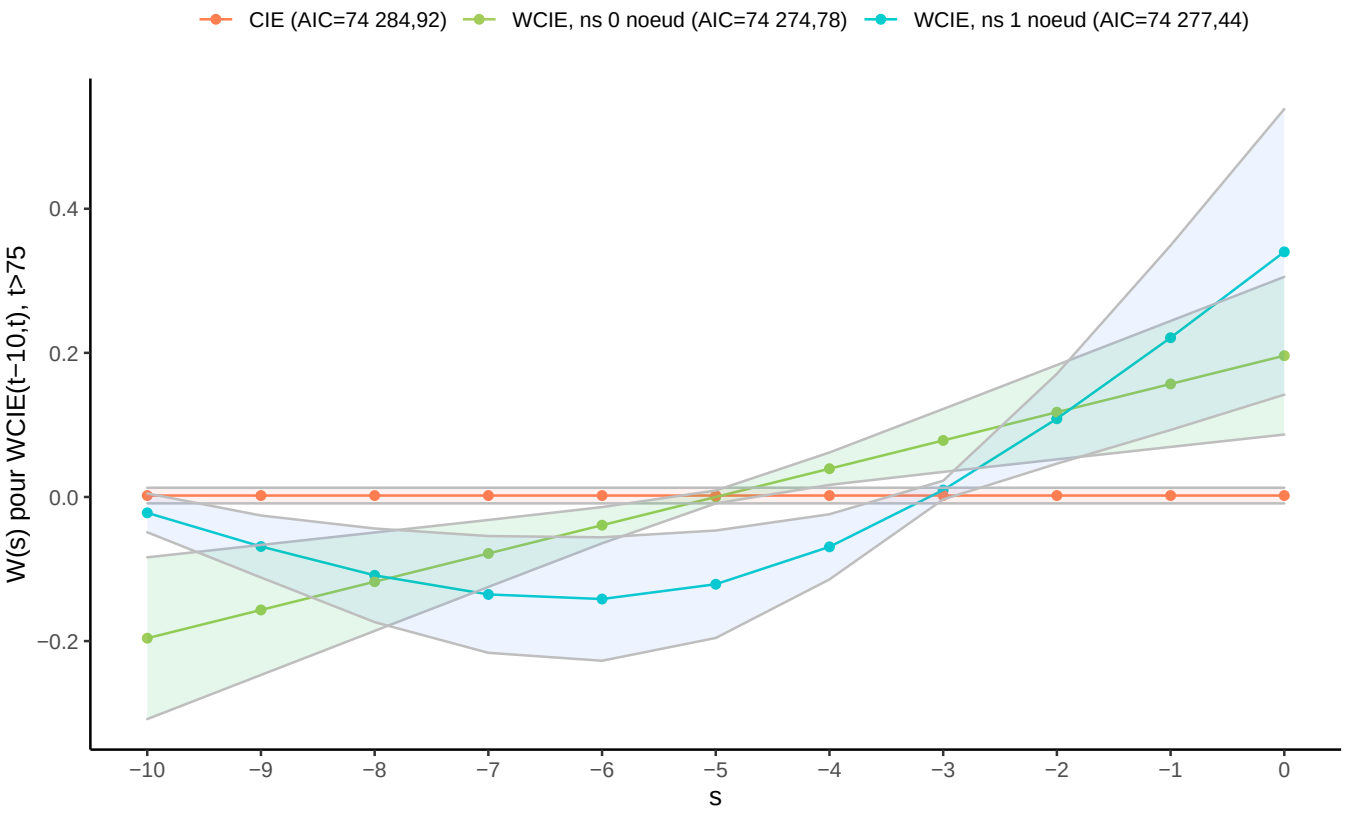

FIGURE 5: Présentation des paramètres de l'histoire de l'IMC sur la pente de l'IST, pour le CIE (rouge), le WCIE avec des nsplines à 0 noeud interne (vert) et avec 1 noeud interne (bleu) (S=10 ;n=1621 ;# observations =14 703).

Cependant, les interprétations des WCIE ns ne vont pas toutes dans le même sens :

— Le poids le plus négatif est associé à la 10^{e} année pour WCIE ns 0 noeud, versus à la 6^{e} années pour WCIE ns 1 noeud ;
— La 10^{e} année a un poids nul sur la pente du score IST de l'année actuelle dans le modèle WCIE ns 1 noeud ;
— La causalité inverse débute dès la 4^{e} année pour WCIE ns 0 noeud, versus 2^{e} pour WCIE ns avec 1 noeud ;
— L'accroissement des poids de la causalité inverse est plus important pour WCIE ns avec 1 noeud que WCIE ns 0 noeud, respectivement 0,099 et 0,039.
— Dans le cas de WCIE ns 0 noeud, la 4^{e} année est associée à la causalité inverse alors que pour WCIE ns avec 1 noeud elle est associée au facteur de risque .

Pour étayer ces divergences, nous allons pousser l'analyse vers les modèles où la hessienne n'est pas inversible. Bien que ces modèles aient convergé correctement en paramètres et en vraisemblance, le critère *rdm* n'est pas vérifié pour ces modèles, ainsi les intervalles de confiance ne pourront pas être présentés. Les modèles estimés sont des ns avec 2 noeuds, bs constant par morceaux et bs de degré 1 avec 2 noeuds. Dans le tableau 6, nous pouvons retrouver les caractéristiques principales de ces modèles et leurs courbes dans la figure 6.

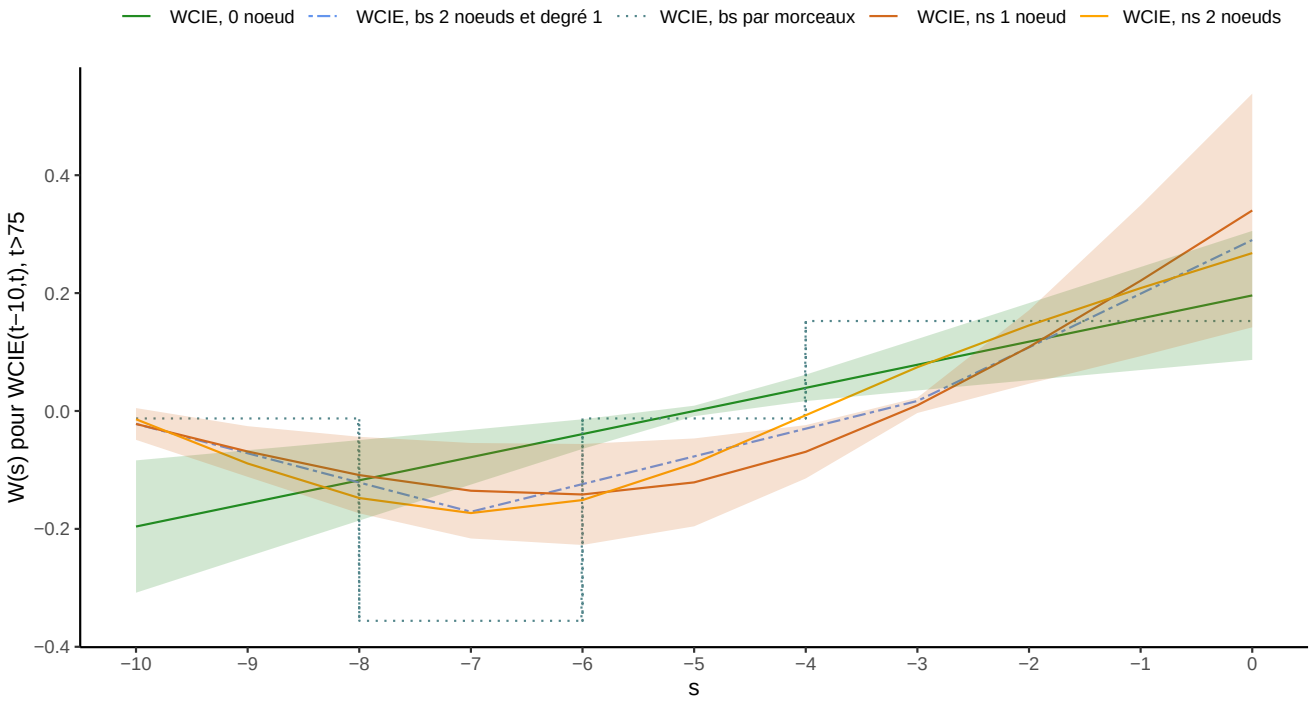

FIGURE 6: Courbes de prédiction des paramètres de l'histoire de l'IMC sur la pente instantanée de l'IST pour des modèles complémentaires (n=1621 ;# observations =14 703).

Le modèle WCIE ns 0 noeud est le seul modèle à suggérer une évolution linéaire positive sur [-10 ;0].

Les modèles WCIE ns avec 1 noeud, 2 noeuds, bs par morceaux et bs 2 noeuds degré 1 ont des courbes très similaires dans leurs tendances. Ces courbes suggèrent un poids proche de 0 pour la 10^e année, puis les poids deviennent négatifs atteignant leur minima entre [-8 ;-6]. En suivant les poids augmentent progressivement, atteignant 0 entre [-4 ;-3] et leur maximum strictement positif à l'année actuelle. Ces concordances entre les modèles soulignent la nécessité de prendre des poids non linéaires sur [-10 ;0]. Ces résultats montrent que les prédictions de $w_1(s)$ du modèle ns 0 noeud ne permettent pas de détecter un phénomène évolutif non linéaire. Ainsi les prédictions des poids proposées par le modèle ns avec 1 noeud interne semblent plus pertinentes.

3.5 Prédictions issues du modèle conjoint

3.5.1 Prédiction de l'effet global de l'IMC sur la pente instantanée de l'IST

Pour débuter nous allons nous intéresser à la prédiction de l'effet global de $w_1(s)$ sur la fenêtre S, noté ω, dans les modèles WCIE ns pour lesquels la Hessienne était inversible. Nous retrouvons dans le tableau 7 les estimations de l'effet global pour les modèles WCIE ns à 0 noeud interne et avec un noeud interne.
Dans les deux modèles l'effet global est statistiquement non différent de 0, respectivement <0,001 pour le modèle WCIE 0 noeud et 0,013 pour WCIE avec 1 noeud. Ces résultats ne sont pas surprenants au vu des courbes $w_1(s)$ dans la figure 5. Pour le modèle WCIE 0 noeud, la symétrie autour de 0 des valeurs de $w_1(s)$ et de ses intervalles de confiance explique cette non significativité. Pour le modèle WCIE avec 1 noeud, $w_1(s)$ prend des valeurs négatives de -9 à -4 proches de -0,1 puis positives de -2 à 0 allant de 0,1 à presque 0,4, ces valeurs vont se contrebalancer aboutissant à un effet global très proche de 0.

w_0	w_1	ω ($IC_{95\%}$)
CIE	WCIE, ns 0 noeud	0,000 ([-0,098 ;0,099])
CIE	WCIE, ns 1 noeud	0,013 ([-0,080 ;0,106])

TABLE 7: Effet global des paramètres liés à l'histoire d'IMC des 10 ans précédant l'évaluation de la pente instantanée de l'IST (n=1621 ;# observations =14 703)

3.5.2 Prédiction de la trajectoire d'IST via des scénarios d'évolution de l'IMC

Par la suite, nous nous sommes intéressés à la prédiction de la trajectoire moyenne d'IST en fonction de divers profils d'évolution de l'IMC. Les prédictions sont présentées pour un sujet homme, ayant fait des études courtes, ne présentant pas le gène APOE4 et rentrant dans l'étude à 64 ans. Nous avons considéré un IMC initial à 20, 25 et 30 kg/m^2, puis une évolution constante, linéaire positive ou négative. Pour l'évolution linéaire nous avons considéré un coefficient de +/-0,25, soit pour +0,25 ; toutes les dix années le sujet va prendre +/-2,5 kg/m^2 en IMC.

Dans la figure 7 ci-dessous, nous avons en haut les valeurs simulées de l'IMC selon le profil choisi, soit un ensemble de 9 courbes. Au bas de la figure, nous avons les prédictions correspondantes obtenues pour l'IST moyen. Nous pouvons remarquer que les profils ayant la même forme d'évolution ont des courbes d'IST très proches, avec un score IST meilleur dès que l'IMC initial est plus bas. Au début des prédictions de l'IST, les courbes des profils sont très proches voire confondues, puis à partir de 80 ans, l'écart commence à se creuser, les courbes présentant le meilleur score d'IST sont celles avec une évolution croissante puis constante et enfin décroissante. Pour étayer ces analyses, nous allons calculer les différences moyennes attendues entre les profils et les intervalles de confiance à 95% associées.

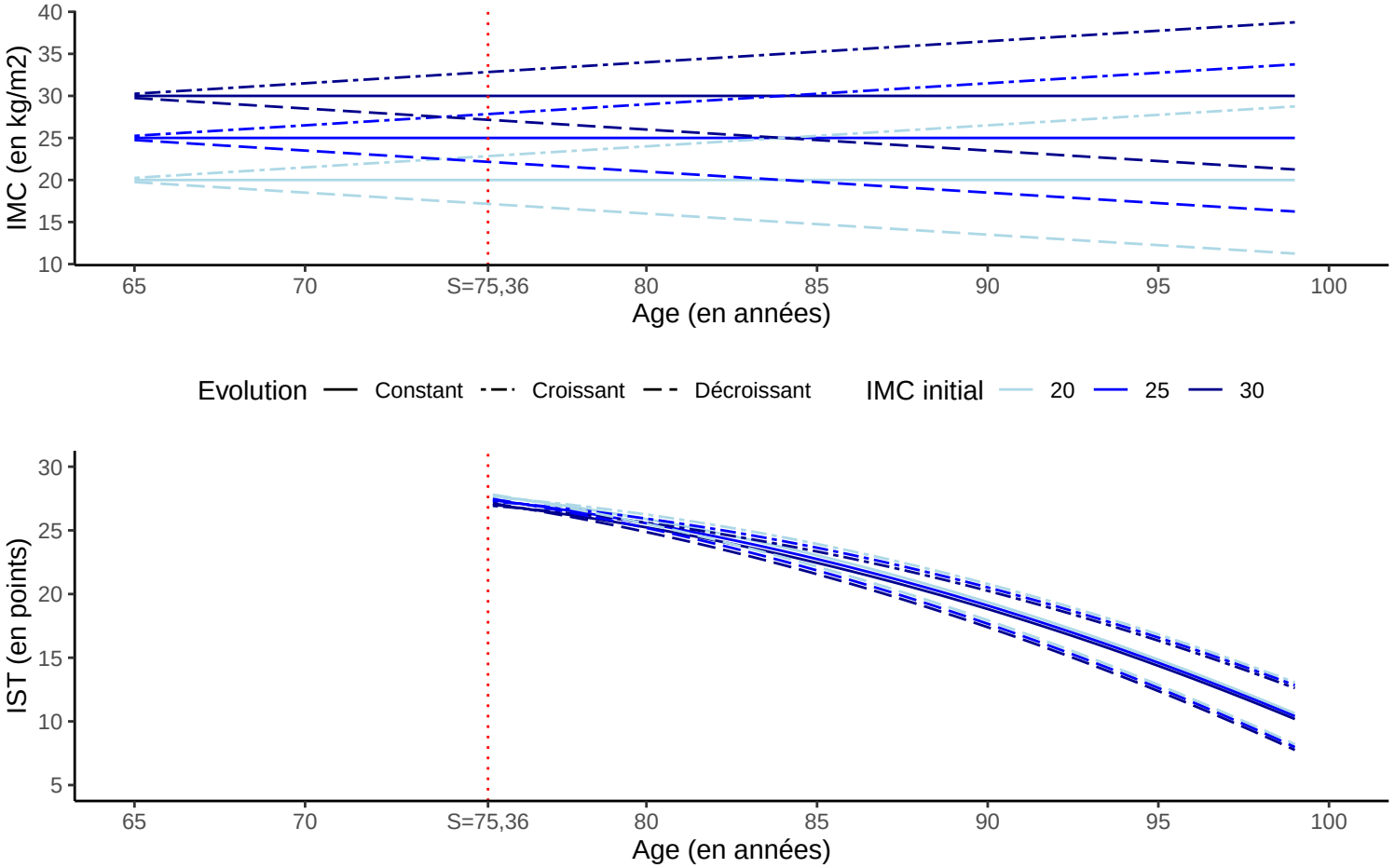

FIGURE 7: Courbes de prédiction de l'IST pour des profils d'évolution de l'IMC chez un homme, sans gène APOE4, ayant fait des études courtes et rentrant dans l'étude à 64 ans.

Nous allons regarder les valeurs de l'IST et les différences attendues pour un même niveau intial et pour une même évolution, les résultats pouvant être transposés aux autres niveaux initiaux et aux autres évolutions.

Pour un même niveau intial fixé à 30 kg/m^2, nous avons dans la figure 8 les courbes simulées de l'IMC selon l'évolution et les prédictions de l'IST avec leurs intervalles de confiance. Nous pouvons remarquer que les trois intervalles de confiance se chevauchent pour tous les âges compris entre 75 et 100 ans. Néanmoins, nous observons que les intervalles de confiance s'écartent les uns des autres plus l'âge augmente. L'âge avançant, l'incertitude sur la valeur prédite de l'IST augmente, les intervalles de confiance sont plus larges.

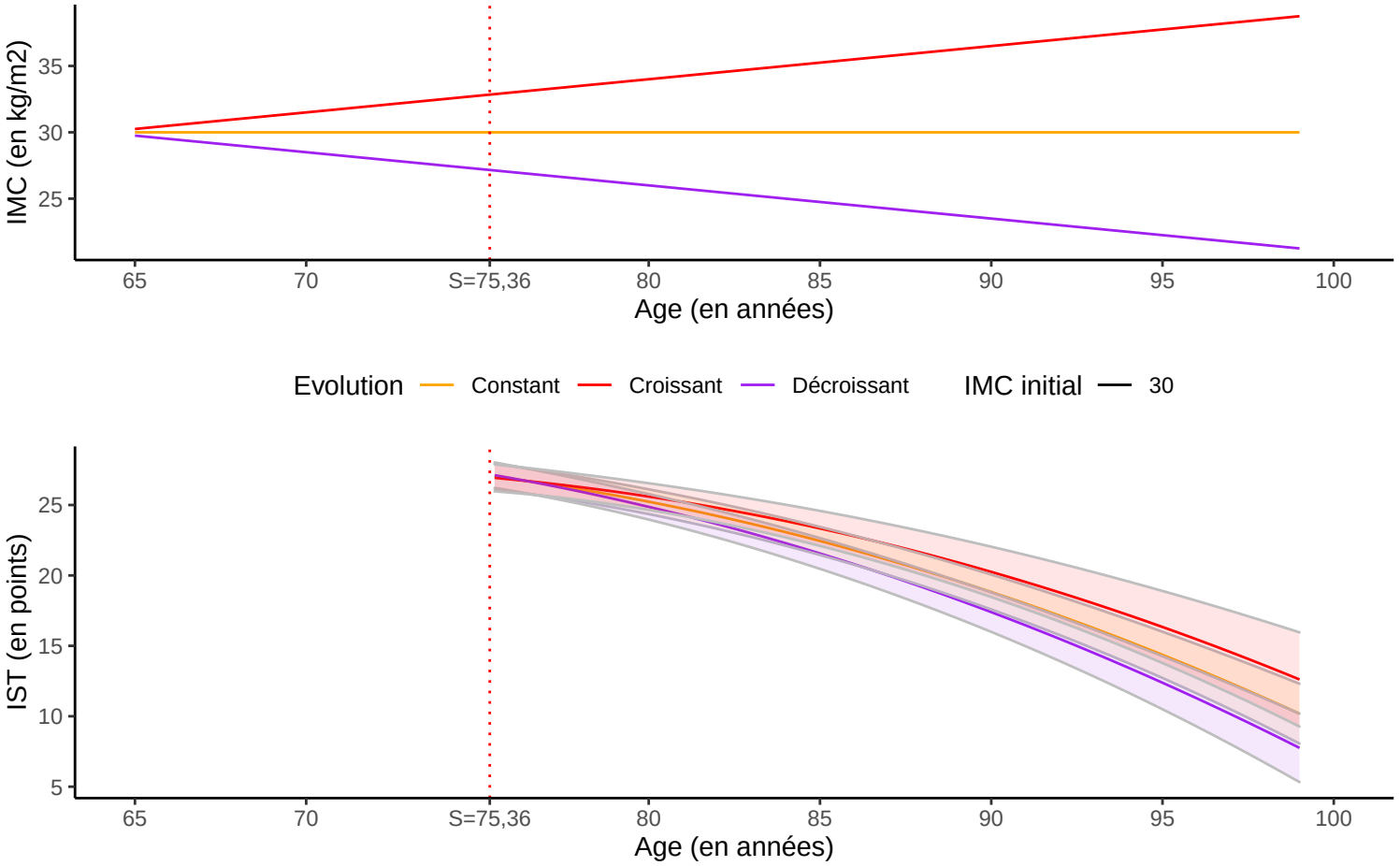

FIGURE 8: Courbes de prédiction l'IST pour un IMC initial à 30 kg/m^2 avec divers profils d'évolution de l'IMC chez un homme, sans gène APOE4, ayant fait des études courtes et rentrant dans l'étude à 64 ans.

Pour une évolution constante, les informations sont présentes dans la figure 9. Nous remarquons que les valeurs prédites de l'IST sont très proches avec des intervalles de confiance confondus pour tous les âges. De même, l'âge avançant, l'incertitude sur la valeur prédite de l'IST augmente et les intervalles restent confondus.

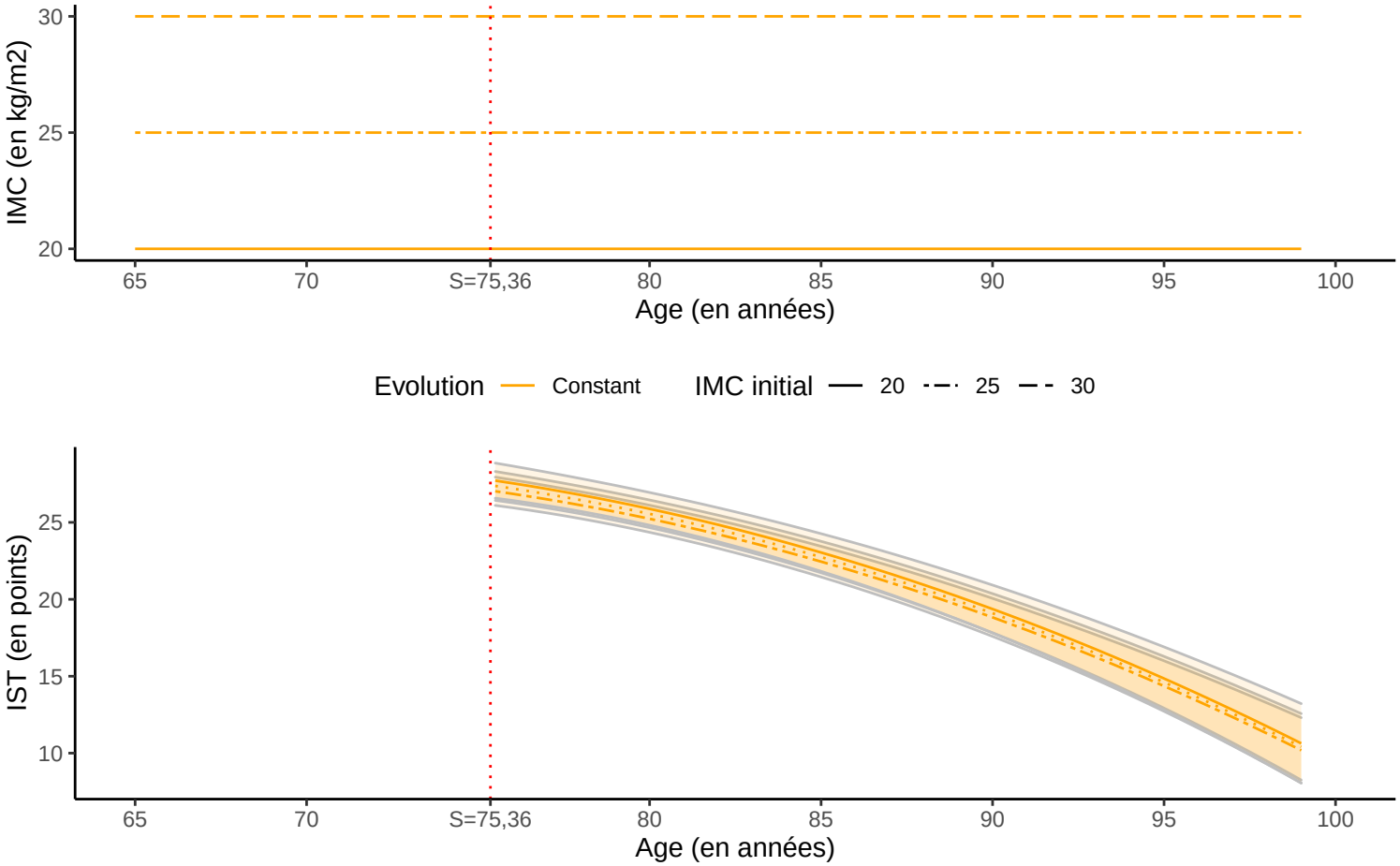

FIGURE 9: Courbes de prédiction de l'IST pour une évolution constante de l'IMC avec divers profils initiaux de l'IMC chez un homme, sans gène APOE4, ayant fait des études courtes et rentrant dans l'étude à 64 ans.

Dans notre contexte, les profils sont très liés les uns aux autres, il est préférable et nécessaire de regarder plutôt les intervalles de confiance des différences afin de statuer sur les différences de prédiction de l'IST. En effet, si deux espérances sont fortement corrélées, la variance de la différence de ces espérances est nettement inférieure à la variance de l'une ou l'autre des espérances. Ainsi, nous pouvons nous attendre à des différences prédites statistiquement différentes de 0, pour certains âges fixés.

Dans les figures 10 et 11, nous avons obtenus les différences moyennes prédites respectivement pour un même niveau intial à 30 kg/m^2 et pour une évolution constante. Ces résultats sont transposables aux autres évolutions et niveaux initiaux.

Dans la figure 10, les différences moyennes prédites pour le profil croissant versus constant et constant versus décroissant sont similaires, ainsi leurs courbes et leurs intervalles de confiance se superposent. La différence moyenne pour croissant versus décroissant est plus importante que les autres différences et croît de manière plus marquée au cours des âges. Pour toutes les différences, l'incertitude concernant la valeur prédite augmente avec l'âge, ainsi l'ampleur des intervalles de confiance augmente. Dans l'ensemble, dès 79 ans, les différences moyennes prédites sont statistiquement différentes de 0 et avec une valeur positive. Pour un sujet ayant une évolution croissante de son IMC alors son score d'IST est statistiquement supérieur à celui d'un sujet ayant une évolution décroissante ou constante de son IMC, dès 79 ans, toutes choses égales par ailleurs. Pour un sujet ayant une évolution constante de son IMC alors son score d'IST est statistiquement supérieur à celui d'un sujet ayant une évolution décroissante de son IMC, dès 79 ans, toutes choses égales par ailleurs.

Dans le tableau 8, nous avons les valeurs prédites de ces différences par types et âges. Pour deux sujets ayant 76 ans, le sujet ayant une évolution croissante de l'IMC aura en moyenne un score d'IST inférieur de 0,05 ($IC_{95\%} = [-0,18; 0,09]$) points par rapport au score d'IST du sujet ayant une évolution constante de l'IMC, toutes choses égales par ailleurs. Cette différence est statistiquement non différente de 0, de même pour constant versus décroissant. Cette différence va croître avec l'âge, à 90 ans elle sera de 1,43 ($IC_{95\%} = [0,40; 2,45]$) points, statistiquement différente de 0. On observe la même tendance pour la différence croissante versus décroissante avec une pente plus croissante au cours de l'âge. A 76 ans, la différence moyenne prédite est de -0,09 ($IC_{95\%} = [-0,37; 0,19]$) points, statistiquement non différent de 0. Puis pour deux sujets ayant 90 ans, le sujet ayant une évolution croissante de l'IMC aura en moyenne un score d'IST supérieur de 2,86 ($IC_{95\%} = [0,80; 4,91]$) points par rapport au score d'IST du sujet ayant une évolution décroissantes de l'IMC, toutes choses égales par ailleurs.

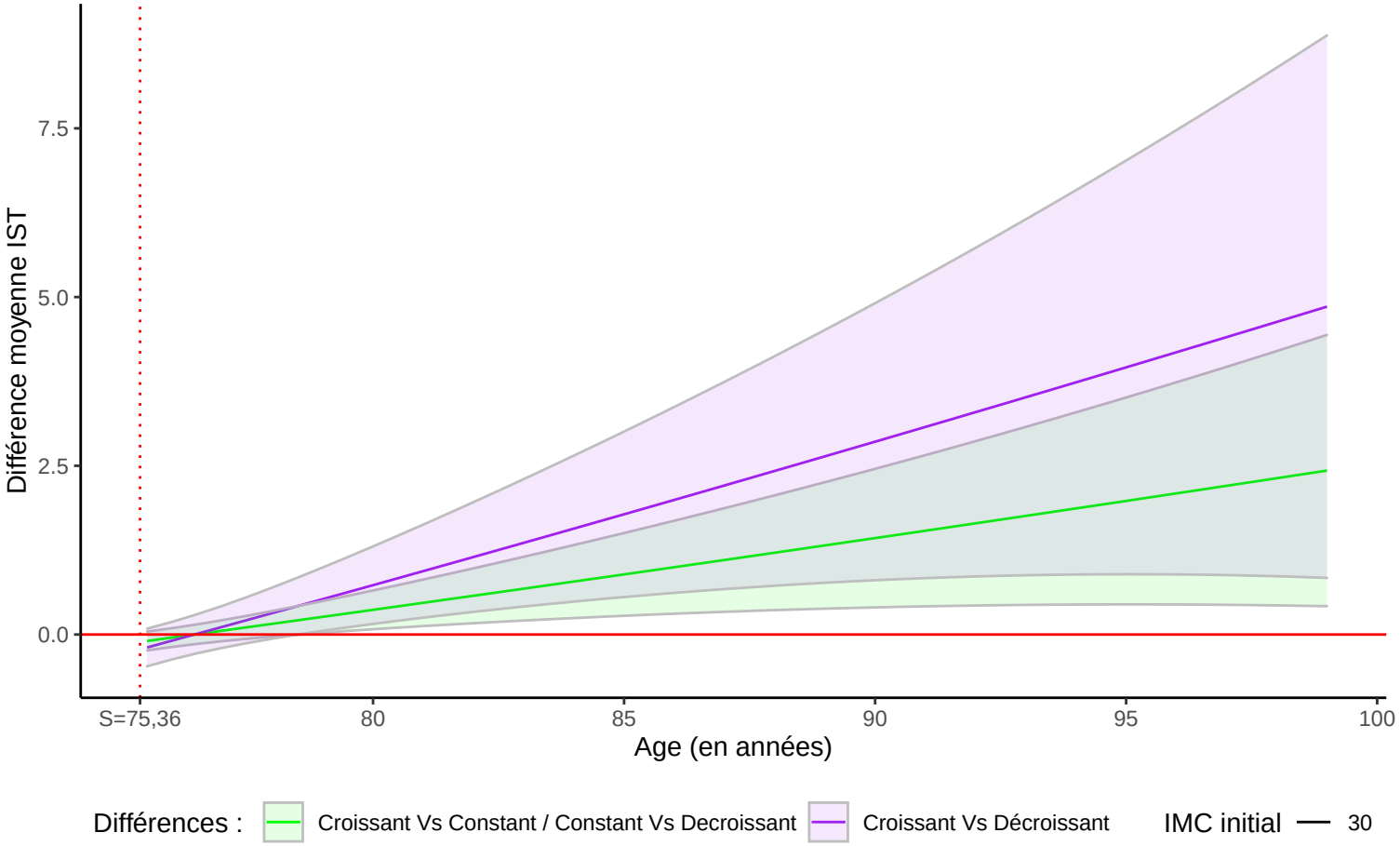

FIGURE 10: Courbes des différences prédites d'IST pour un IMC initial à 30 kg/m^2 avec divers profils d'évolution de l'IMC chez un homme, sans gène APOE4, ayant fait des études courtes et rentrant dans l'étude à 64 ans.

Ages (en années)	Croissant\Décroissant Vs Constant	Croissant Vs Décroissant
76	-0,05 ([-0,18 ;0,09])	-0,09 ([-0,37 ;0,19])
78	0,16 ([-0,03 ;0,35])	0,32 ([-0,06 ;0,70])
80	0,37 ([0,08 ;0,65])	0,73 ([0,16 ;1,31])
82	0,57 ([0,17 ;0,98])	1,15 ([0,33 ;1,96])
84	0,78 ([0,24 ;1,33])	1,57 ([0,49 ;2,65])
86	1,00 ([0,31 ;1,69])	1,99 ([0,62 ;3,37])
88	1,21 ([0,36 ;2,06])	2,42 ([0,72 ;4,12])
90	1,43 ([0,40 ;2,45])	2,86 ([0,80 ;4,91])

TABLE 8: Différence moyenne prédite d'IST pour un IMC initial à 30 kg/m^2 avec divers profils d'évolution de l'IMC chez un homme, sans gène APOE4, ayant fait des études courtes et rentrant dans l'étude à 64 ans.

Dans la figure 11, nous retrouvons une concordance avec les résultats de la figure 9, soit pour tous âges il n'y a pas de différence statistiquement significative ($\alpha = 0,05$) entre un IMC initial à 20, 25 ou 30 kg/m^2 pour une évolution constante. L'incertitude concernant la valeur prédite de la différence augmente avec l'âge, ainsi l'ampleur des intervalles de confiance augmente. De plus nous observons que les courbes de 20 versus 25 et 25 versus 30 sont similaires ainsi que leurs intervalles de confiance. Pour deux sujets ayant 76 ans, le sujet ayant un IMC initial à 20 kg/m^2 aura en moyenne un score d'IST supérieur de 0,34 ($IC_{95\%} = [-0,09; 0,78]$) points par rapport au score d'IST du sujet ayant un IMC initial à 25 kg/m^2, toutes choses égales par ailleurs. Puis deux sujets ayant 90 ans, le sujet ayant un IMC initial à 20 kg/m^2 aura en moyenne un score d'IST supérieur de 0,27 ($IC_{95\%} = [-0,33; 0,87]$) points par rapport au score d'IST du sujet ayant un IMC initial à 25 kg/m^2, toutes choses égales par ailleurs. Nous avons les mêmes résultats pour un sujet ayant un IMC initial à 25 kg/m^2 par rapport à un sujet ayant un IMC initial à 30 kg/m^2. Pour le cas 20 versus 30, la différence moyenne prédite est plus importante, même si elle demeure non signifiative. Nous avons aussi une tendance à la décroissance, à 76 ans 0,69 ($IC_{95\%} = [-0,17; 1,57]$) puis à 90 ans 0,54 ($IC_{95\%} = [-0,66; 1,75]$).

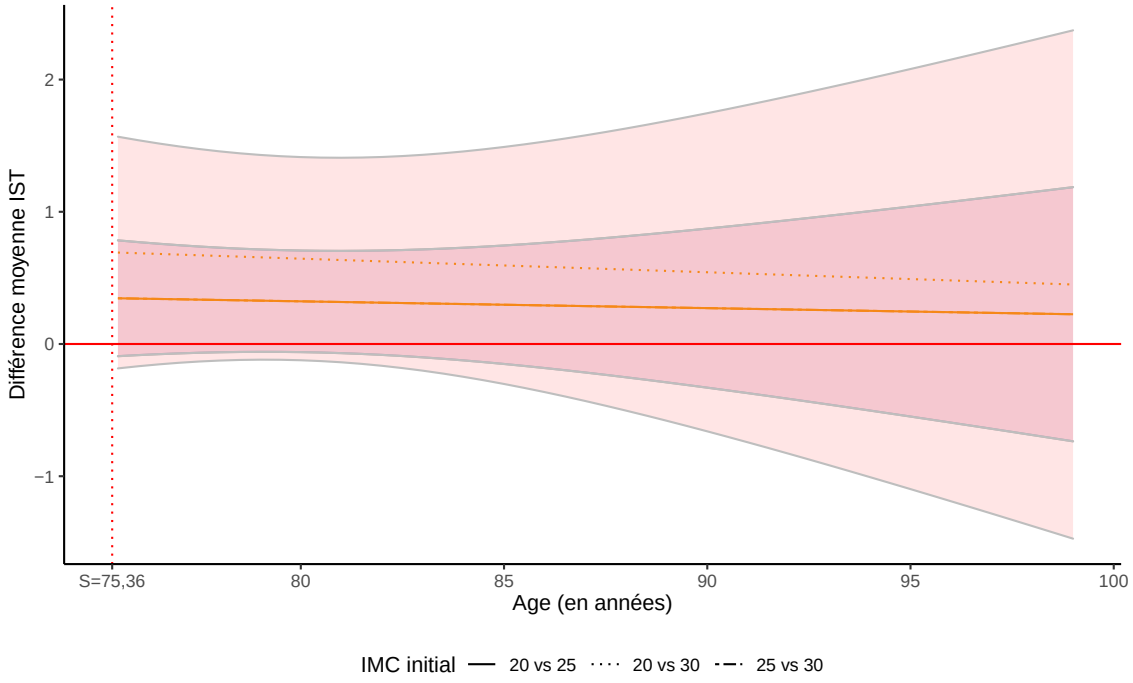

FIGURE 11: Courbes des différences prédites d'IST pour une évolution constante de l'IMC avec divers profils initiaux de l'IMC chez un homme, sans gène APOE4, ayant fait des études courtes et rentrant dans l'étude à 64 ans.

4 Discussion

4.1 Apports de ce travail

Le vieillissement de la population et l'absence de traitement curatif a fait des maladies neurodégénératives un défi majeur pour le système de santé et la recherche scientifique. Dans le but d'une politique de prévention, la compréhension du lien entre des facteurs modifiables et le vieillissement cognitif est essentielle [6, 17, 25]. Mais les méthodes statistiques actuelles ne permettent pas d'évaluer des associations complexes et dépendantes du temps entre ces facteurs modifiables tout au long de la vie et le vieillissement cognitif qui apparait à des âges avancés.

Nous avons proposé une nouvelle méthodologie statistique qui permet d'évaluer la trajectoire d'association entre l'histoire d'une exposition dépendante du temps et un marqueur dépendant du temps. Cette méthode se démarque des travaux précédents de plusieurs façons.

Premièrement, nous considérons que l'exposition est une variable endogène, c'est à dire un processus dynamique spécifique au sujet qui peut aussi être bruité et évalué à des temps discrets, et nous modélisons sa trajectoire via un modèle mixte. Dans les travaux précédents, l'exposition était soit mesurée à un temps donné fixe [26, 27] ou vu comme un processus non bruité mesuré en temps continu [28, 29, 30].

Deuxièmement, nous considérons un outcome répété modélisé via un modéle mixte. Les travaux antérieurs portent souvent sur des outcomes binaires ou de type temps de survie censurés par des modèles de logistique ou de survie. [21, 29, 30].

Troisièmement, nous modélisons les deux processus dans un modèle conjoint pour correctement prendre en compte leur inter-dépendance.

Enfin, nous nous focalisons sur l'effet de l'histoire de l'exposition sur le changement instantané de l'outcome pour être en accord avec l'approche dynamique de la causalité [25, 11]. Notre hypothèse est que l'effet de l'exposition est évalué sur une fenêtre d'amplitude fixe S et mouvante dans le temps pour toujours correspondre à la même antériorité par rapport au point d'évaluation du changement instantané. Cette approche nous permet de nous rapprocher des hypothèses développementales par rapport à une approche qui scinderait les fenêtres d'étude de l'exposition et du marqueur comme c'est fait dans l'approche landmark [10].

Cette modélisation conjointe via un WCIE a été entiérement développée au cours de ce stage, en commençant par les calculs mathématiques jusqu'à l'implémentation informatique à vocation d'être finalisée en package et l'application aux données de la cohorte 3C.

Dans notre application aux sujets du centre bordelais de la cohorte 3C, nous avons étudié la relation entre l'histoire de l'IMC et l'évolution cognitive via le test de fluence verbale IST. Pour se faire, nous avons considéré une fenêtre mouvante d'exposition de dix années pour l'IMC. Notre approche a permis d'infirmer l'hypothèse d'une accumulation progressive et constante de l'IMC sur le changement cognitive (CIE). Nous avons retrouvé un changement de sens de l'effet de l'IMC au cours des dix dernières années de suivi. D'abord, un IMC élevé à distance de l'évaluation cognitive a un effet délétère sur la pente instantanée de l'IST, cela va dans le sens de l'adiposité comme facteur de risque long terme. Puis un IMC élevé à proximité de l'évaluation cognitive a un effet protecteur sur le déclin de l'IST, probablement dû à la causalité inverse. Ces conclusions sont en concordance avec les observations de travaux précédents [10, 17, 16, 31]. Nous avons aussi observé que pour un même niveau initial d'IMC, une évolution différente de l'IMC impactait significativement le déclin de l'IST alors que des niveaux initiaux différents d'IMC avec une même évolution ne semblaient pas être associés à des évolutions d'IST très différentes.

4.2 Points forts de l'approche statistique

Le point fort de ce travail est que la méthodologie a été pensée et réfléchie pour toutes expositions et tous marqueurs tant que les hypothèses des modèles mixtes sont vérifiées. Ainsi, au delà de l'application présentée, cette méthode pourra permettre d'analyser de nombreuses relations exposition/marqueur, telle que l'histoire de la tension artérielle ou du cholestérol sur la santé cardiovasculaire.

Nous avons considéré un marqueur continu avec un modèle linéaire mixte. Mais cette méthode est facilement adaptable à des marqueurs qualitatifs par le biais de modèle mixte généralisé. De plus, nous avons défini le WCIE comme la somme des expositions annuelles pondérées sur $[-S, 0]$, mais il est simple de considérer plutôt l'intégrale sur cette période.

4.3 Limites du travail

La principale limite de cette approche est que cette méthode nécessite un grand nombre d'années de suivi. En fonction de nos hypothèses biologiques sur le processus dynamique de l'exposition, l'amplitude pour mesurer l'histoire va varier. Dans le cas d'une exposition où le sens de l'effet change au cours du temps, l'amplitude pour mesurer l'histoire doit être importante. Pour l'application, nous avons choisi une amplitude de dix années, il aurait été préférable d'augmenter cette fenêtre à vingt années et d'étudier l'IMC plus tôt dans la vie mais le design de la cohorte ne le permettait pas avec une entrée dans l'étude après 65 ans et un maximum de suivi à 17 ans. Cette méthode serait particulièrement intéressante avec des cohortes vie-entière qui suivent les individus de l'âge adulte à l'âge avancé. Si l'histoire de l'exposition encore plus antérieure est associée à l'événement, les années non prises en compte pourraient agir comme facteur de confusion non observé car associé à la fois à l'exposition considérée et à l'événement. Cela pourrait biaiser les résultats de l'analyse, comme montré en annexe 5.1 avec une histoire d'exposition uniquement sur l'année précédant l'évaluation cognitive.

Nous avons vu que dans le cas de modèles complexes l'inversion de la hessienne peut poser des difficultés. Dans notre application, la complexité du modèle a rendu difficiles certaines estimations et nous a amené à devoir simplifier le modèle pour l'IMC. Il serait intéressant d'explorer les raisons de ces problèmes de convergence, et la robustesse de nos résultats à des spécifications différentes du modèle pour l'IST.

Les modèles mixtes sont dit robustes aux données manquantes si elles sont manquantes aléatoirement (MAR). Ainsi, notre modèle est de même valide et robuste sous cette hypothèse restrictive. Dans notre application, les données manquantes liées à la sortie d'étude sont dûes principalement au décès et à la survenue de démence, et sont possiblement informatives. Prendre en compte ces sorties d'étude informatives serait possible en remplaçant le modèle mixte pour l'IST par un modèle conjoint pour l'IST et les causes de sortie d'étude. L'interprétation du WCIE ne changerait pas.

4.4 Note sur le temps de calcul

Dans les modèles pour données répétées, le temps de calcul peut augmenter exponentiellement en fonction du nombre d'effets aléatoires. Ce temps de calcul augmente d'autant plus que nous sommes dans le cas de modèles conjoints pour deux marqueurs répétés dans le temps. Ainsi, un des objectifs de ce travail était d'optimiser le temps de calcul afin d'obtenir des estimations dans un temps de calcul raisonnable.

Tout d'abord, la fonctions *mla* permet de travailler sur plusieurs coeurs à la fois. Ainsi, les modèles conjoints de ce rapport ont pu être estimés sur le serveur Curta de l'université de Bordeaux en utilisant 10 coeurs pour la fonctions mla.

De plus, *mla* est une fonction itérative qui pour chaque itération va calculer la log-vraisemblance du modèle. Tout en prenant en compte les spécificités de notre modèle nous devions pouvoir minimiser le temps de calcul de la log-vraisemblance.
Pour ce faire nous avons envisagé deux options, en premier lieu nous avons utilisé la fonction *hlme* pour variances hétérocédastes. Cette fonction est présente dans le packages lcmm où le corps du programme est codé en Fortran, C++ et R. Cette fonction réalise normalement l'estimation complète d'un modèle mais nous pouvions récupérer uniquement la valeur de vraisemblance pour un vecteur de paramètres donnés en spécifiant maxiter=0. Cependant cette approche donne des temps de calcul très importants et limitants. Pour le modèle $w_0 = CIE$ et $w_1 = CIE$, le temps de calcul est de 2,8 heures pour 11 itérations, soit environ 15 minutes pour une itération. Pour le modèle $w_0 = CIE$ et $w_1 = WCIE$, ns 0 noeud, le temps de calcul est de 6,2 heures pour 12 itérations, soit environ 30 minutes pour une itération. Ce rapport temps de calcul itérations dépend du chemin itératif pris par *mla* et de la complexité du modèle, en général il augmente plus le modèle est complet.
En second lieu nous avions envisagé d'utiliser seulement la fonction de calcul de la log-vraisemblance de hlme. Ainsi nous avons un code en Fortran qui calcule la log-vraisemblance et nous l'appellons dans code R avec la fonction $loglikF$. Cette seconde méthode permet de se dispenser des autres calculs et sorties faits par *hlme*, ainsi nous nous limitons à un code Fortran. Pour le modèle $w_0 = CIE$ et $w_1 = CIE$, le temps de calcul est de 3,57 heures pour 54 itérations, soit environ 4 minutes pour une itération. Pour le modèle $w_0 = CIE$ et $w_1 = WCIE$, ns 0 noeud, le temps de calcul est de 0,72 heures pour 17 itérations, soit environ 2,5 minutes pour une itération.

Nous remarquons clairement un gain de temps de calcul avec la seconde méthode, $loglikF$. Par la suite, il sera proposé aux utilisateurs d'utiliser cette méthode qui est plus performante.

4.5 Perspectives

Les perspectives de ce travail sont vastes. Tout d'abord, un travail de simulation a été entamé (voir annexe 5.6) mais a dû être mis en suspens faute de temps. Il serait utile et pertinent de faire tourner des scénarios pour évaluer les performances du programme d'estimation en terme de biais d'estimation et taux de couverture.

Par la suite une analyse de sensibilité du modèle à l'hypothèse des données manquantes aléatoires serait pertinente en considérant conjointement la survenue de démence et décès [32].

Enfin, cette méthodologie pourrait être approfondie sur l'étude 3C. Les autres tests cognitifs pourraient être analysés en tant que marqueurs de l'évolution cognitive. La base de sujets incluse dans l'analyse pourrait être élargie à d'autres centres voir à tous les sujets en prenant en compte l'effet centre.

Références

[1] Garcia V. Livret d'accueil Centre de recherche Inserm U1219 ; 2017. Available from : https ://issuu.com/valeriegarcia1/docs/livret$_a ccueil_u$1219 − oct2017 − 5.

[2] Lévy Y. Inserm Plaquette ; 2017. Available from : https ://www.inserm.fr/sites/default/files/2018-01/Inserm$_P laquette_2$01706$_F R_0 .pdf$.

[3] BPH. Le Centre | Bordeaux population Health ; 2021. Available from : https ://www.bordeaux-population-health.center/le-centre/.

[4] Wang M, Liao X, Laden F, Spiegelman D. Quantifying risk over the life course - latency, age-related susceptibility, and other time-varying exposure metrics. Statistics in Medicine. 2016 Jun ;35(13) :2283–2295. Available from : http ://-doi.wiley.com/10.1002/sim.6864.

[5] Baldi I, Cordier S, Coumoul X, Elbaz A, Gamet-Payrastre L, Lebailly P, et al. Pesticides : Effets sur la santé. 101 rue de Tolbiac, 75013 Paris : Inserm ; 2019.

[6] Yaffe K, Vittinghoff E, Hoang T, Matthews K, Golden SH, Zeki Al Hazzouri A. Cardiovascular Risk Factors Across the Life Course and Cognitive Decline : A Pooled Cohort Study. Neurology. 2021 Apr ;96(17) :e2212–e2219. Available from : http ://www.neurology.org/lookup/doi/10.1212/WNL.0000000000011747.

[7] Walker KA, Sharrett AR, Wu A. Association of Midlife to Late-Life Blood Pressure Patterns With Incident Dementia ;6(322) :535–545. Available from : https ://pubmed.ncbi.nlm.nih.gov/31408138/.

[8] Sylvestre MP, Abrahamowicz M. Modélisation flexible des effets cumulatifs des expositions dépendantes du temps sur le risque. Statistics in Medicine. 2009 Aug ;28(27) :3437–3453. Available from : https ://onlinelibrary.wiley.com/doi/abs/10.1002/sim.3701.

[9] Mauff K, Steyerberg EW, Nijpels G, van der Heijden AAWA, Rizopoulos D. Extension of the association structure in joint models to include weighted cumulative effects. Statistics in Medicine. 2017 Jul ;36(23) :3746–3759. Available from : http ://-doi.wiley.com/10.1002/sim.7385.

[10] Wagner M, Grodstein F, Leffondre K, Samieri C, Proust-Lima C. Time-varying associations between an exposure history and a subsequent health outcome : a landmark approach to identify critical windows ;p. 41.

[11] Voelkle MC, Gische C, Driver CC, Lindenberger U. The Role of Time in the Quest for Understanding Psychological Mechanisms. Multiva-

riate Behavioral Research. 2018 Nov ;53(6) :782–805. Available from : https ://www.tandfonline.com/doi/full/10.1080/00273171.2018.1496813.

[12] Philipps V, Hejblum BP, Prague M, Commenges D, Proust-Lima C. Robust and Efficient Optimization Using a Marquardt-Levenberg Algorithm with R Package marqLevAlg. arXiv :200903840 [stat]. 2020 Sep ;ArXiv : 2009.03840. Available from : http ://arxiv.org/abs/2009.03840.

[13] Prague M, Diakite A, Commenges D. Package 'marqLevAlg' - Algorithme de Marquardt-Levenberg en R : Une alternative à 'optimx' pour des problèmes de minimisation ; 2012. Available from : https ://hal.archives-ouvertes.fr/hal-00717566/file/2$_{P}rague.pdf$.

[14] INSERM, de La Salpêtrière H, Pasteur I, UMR CC, Cyceron G. Vascular Factors and Risk of Dementia : Design of the Three-City Study and Baseline Characteristics of the Study Population ;22(6) :316–325. Available from : https ://www.karger.com/Article/FullText/72920.

[15] INSERM. Etude des 3 cités ;. Available from : http ://www.three-city-study.com/l-etude-des-trois-cites-3c-historique.php.

[16] Singh-Manoux A, Dugravot A, Shipley M, Brunner EJ, Elbaz A, Sabia S, et al. Obesity trajectories and risk of dementia : 28 years of follow-up in the Whitehall II Study. Alzheimers Dement. 2018 Feb ;14(2) :178–186. Available from : https ://alz-journals.onlinelibrary.wiley.com/doi/epdf/10.1016/j.jalz.2017.06.2637.

[17] Wagner M, Grodstein F, Proust-Lima C, Samieri C. Long-Term Trajectories of Body Weight, Diet, and Physical Activity From Midlife Through Late Life and Subsequent Cognitive Decline in Women. American Journal of Epidemiology. 2020 Apr ;189(4) :305–313. Available from : https ://academic.oup.com/aje/article/189/4/305/5645465.

[18] Wagner M, Helmer C, Tzourio C, Berr C, Proust-Lima C, Samieri C. Evaluation of the Concurrent Trajectories of Cardiometabolic Risk Factors in the 14 Years Before Dementia. JAMA Psychiatry. 2018 Oct ;75(10) :1033–1042. Available from : http ://archpsyc.jamanetwork.com/article.aspx ?doi=10.1001/jamapsychiatry.2018.2004.

[19] Issac B, Kennie AT. The Set Test as an Aid to the Detection of Dementia in Old People. The British Journal of Psychiatry. 1973 Oct ;123(575) :467–470. Available from : https ://www.cambridge.org/core/journals/the-british-journal-of-psychiatry/article/abs/set-test-as-an-aid-to-the-detection-of-dementia-in-old-people/3200078F4CC3AD0C3FD938978F4C0141.

[20] Lechevallier-Michel N, Fabrigoule C, Lafont S, Letenneur L, Dartigues JF. Normes pour le MMSE, le test de rétention visuelle de Benton, le set test d'Isaacs, le sous-test des codes de la WAIS et le test de barrage de Zazzo chez des sujets âgés de 70 ans et plus : données de la cohorte PAQUID. Revue Neurologique. 2004 Nov ;160(11) :1059–1070. Available from : https ://linkinghub.elsevier.com/retrieve/pii/S0035378704711431.

[21] Artero S, Ancelin ML, Portet F, Dupuy A, Berr C, Dartigues JF, et al. Risk profiles for mild cognitive impairment and progression to dementia are gender specific. Journal of Neurology, Neurosurgery & Psychiatry. 2008 Sep ;79(9) :979–984. Available from : https ://jnnp.bmj.com/lookup/doi/10.1136/jnnp.2007.136903.

[22] Dufouil, Richard, Flévet, Dartigues, Ritchie, Tzourio, et al. APOE genotype, cholesterol level, lipid-lowering treatment, and dementia ;64(9). Available from : https ://n.neurology.org/content/64/9/1531.

[23] Campion D, Brice A, Hannequin D, Frebourg T, Martinez M, Agid Y, et al. Les facteurs génétiques dans l'étiologie de la maladie d'Alzheimer. Médecine/sciences. 1996 Jul ;12(6–7) :723–731. Available from : https ://www.ipubli.inserm.fr/bitstream/handle/10608/813/$MS_1996_6 - 7_723.pdf$.

[24] Glymour MM, Tzourio C, Dufouil C. Is Cognitive Aging Predicted by One's Own or One's Parents' Educational Level ? Results From the Three-City Study. American Journal of Epidemiology. 2012 Apr ;175(8) :750–759. Available from : https ://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwr509.

[25] Taddé BO, Jacqmin-Gadda H, Dartigues J, Commenges D, Proust-Lima C. Dynamic modeling of multivariate dimensions and their temporal relationships using latent processes : Application to Alzheimer's disease. Biometrics. 2019 Oct ;76(3) :886–899. Available from : https ://onlinelibrary.wiley.com/doi/10.1111/biom.13168.

[26] Samieri C, Jutand MA, Féart C, Capuron L, Letenneur L, Barberger-Gateau P. Dietary Patterns Derived by Hybrid Clustering Method in Older People : Association with Cognition, Mood, and Self-Rated Health. Journal of the American Dietetic Association. 2008 Sep ;108(9) :1461–1471. Available from : https ://linkinghub.elsevier.com/retrieve/pii/S0002822308012674.

[27] Godin O, Dufouil C, Ritchie K, Dartigues JF, Tzourio C, Pérès K, et al. Depressive Symptoms, Major Depressive Episode and Cognition in the Elderly : The Three-City Study. Neuroepidemiology. 2007 Apr ;28(2) :101–108. Available from : https ://www.karger.com/Article/FullText/101508.

[28] Danieli C, Sheppard T, Costello R, Dixon WG, Abrahamowicz M. Modeling of cumulative effects of time-varying drug exposures on within-subject changes in a continuous outcome. Statistical Methods in Medical Research. 2020 Sep ;29(9) :2554–2568. Available from : http ://journals.sagepub.com/doi/10.1177/0962280220902179.

[29] Ryan J, Carrière I, Scali J, Ritchie K, Ancelin ML. Life-time estrogen exposure and cognitive functioning in later life. Psychoneuroendocrinology. 2009 Feb ;34(2) :287–298. Available from : https ://linkinghub.elsevier.com/retrieve/pii/S0306453008002503.

[30] Ritchie K, Jaussent I, Stewart R, Dupuy AM, Courtet P, Malafosse A, et al. Adverse childhood environment and late-life cognitive functioning : Childhood adversity and late-life cognition. International Journal of Geriatric Psychiatry. 2010 Oct ;26(5) :503–510. Available from : https ://onlinelibrary.wiley.com/doi/10.1002/gps.2553.

[31] Albanese E, Launer LJ, Egger M, Prince MJ, Giannakopoulos P, Wolters FJ, et al. Body mass index in midlife and dementia : Systematic review and meta-regression analysis of 589,649 men and women followed in longitudinal studies. Alzheimers Dement. 2017 Jan ;8(1) :165–178. Available from : https ://onlinelibrary.wiley.com/doi/abs/10.1016/j.dadm.2017.05.007.

[32] Roderick J, Little A, Donald A, Rubin B. Statistical Analysis With Missing Data. Journal of Educational Statistics. 1991 ;16(2) :150–155. Available from : https ://www.jstor.org/stable/1165119.

5 Annexes

5.1 Modélisation des trajectoires d'effet des années actuelle et précédente de l'IMC sur la pente instantanée de l'IST

Dans cette partie nous souhaitons déterminer la relation entre l'IMC actuel et précédent (S=1) et la pente instantanée de l'IST. Par la suite nous souhaitons comparer ces résultats à ceux obtenus avec S=10, ainsi nous allons garder les mêmes sujets que lors de la prédiction avec S=10.

Le poids pour le modèle $w_1(s)$ a été modélisé avec un nspline 0 noeud sur $[-1;0]$, les résultats de ce modèle seront juxtaposés à ceux du modèle avec nspline 0 noeud sur $[-10;0]$. Pour S=1, le modèle a convergé en log-vraisemblance, en paramètres et en rdm, nous obtenons une log-vraisemblance à -37 113,28 avec 27 paramètres soit AIC=74 280,56, $w_1(s) = \theta_{10} + \theta_{11} N_1(s)$, avec :

— $\theta_{10} = -1,657$, $IC_{95\%} = [-3,133; -0,180]$

— $\theta_{11} = 4,154$, $IC_{95\%} = [0,508; 7,801]$

Pour l'année précédente le poids $w_1(-1) = -1,657$ ($IC_{95\%} = [-3,133; -0,181]$) et pour l'année courante $w_1(0) = 1,674$ ($IC_{95\%} = [0,225; 3,122]$), l'ensemble des poids sont statistiquement significativement différents de 0 pour α=0,05. L'effet global est de 0,017 avec $IC_{95\%} = [-0,07; 0,108]$.
En comparaison, si nous regardons les poids $w_1(-1)$ et $w_1(0)$ pour S=10, nous obtenons respectivement 0,157 ($IC_{95\%} = [0,069; 0,244]$) et 0,196 ($IC_{95\%} = [0,087; 0,305]$). L'effet global de w_1 sur la période $[-1;0]$ est de 0,353 avec $IC_{95\%} = [0,156; 0,550]$.

Les deux méthodes avec S=1 et S=10 présentent des résultats contradictoires tant sur $w_1(s)$ $s = -1$ ou 0 que pour l'effet global, ces divergences s'expliquent par un biais de confusion.
Dans le cas où nous considérons S=1, les années $[-10;-2]$ sont un facteur de confusion sur la relation de l'IMC sur $[-1;0]$ avec la pente instantanée de l'IST. Nous avons bien le niveau de l'IMC sur $[-1;0]$ qui est fortement lié à celui sur $[-10;-2]$ et d'après le modèle ns 0 noeud

S=10 les années $[-10; -2]$ de l'IMC impactent la pente instantanée de l'IST. Ainsi, même si nous nous intéressons seulement à l'année actuelle et précédente, nous devons simuler avec S ≥ 10 afin d'ajuster sur ce facteur de confusion pour obtenir des résultats valides pour $w_1(s)$ et $w_0(s)$.

5.2 Interprétation des paramètres fixes du modèle conjoint

Dans le tableau 9, nous pouvons observer l'évolution des paramètres fixes de l'IMC et l'IST en fonction du WCIE ns de w_1 pour les modèles où l'inversion de la hessienne est possible. Nous pouvons remarquer que la plus grande variation de ces effets fixes est observée quand nous passons d'un modèle indépendant à un modèle CIE ou WCIE ns. Les effets fixes restent relativement stablse d'un modèle à l'autre, exceptés pour l'intercept de la pente instantanée du IST qui a la plus grande amplitute de variation.

Nous allons vous présenter les résultats de ces effets fixes pour WCIE ns avec 1 noeud.

Pour un sujet rentrant dans l'étude à 64 ans, son niveau initial d'IMC est en moyenne de 26,73 kg/m^2 ($IC_{95\%} = [26, 39; 27, 07]$).

Pour deux sujets ayant un écart de 5 ans, le plus âgé ayant 75 ans, ce dernier aura en moyenne un IMC inférieur de 5,75 kg/m^2 par rapport au plus jeune ($IC_{95\%} = [-7, 09; -4, 38]$).

Pour deux sujets ayant un âge d'inclusion différent de 5 ans, le plus âgé aura en moyenne un score d'IST initial supérieur de 0,53 points par rapport au plus jeune ($IC_{95\%} = [0, 24; 0, 82]$), toutes choses égales par ailleurs.

Pour deux sujets ayant un statut gène APOE4 différent, le sujet ayant le gène APOE4 aura en moyenne un score d'IST initial inférieur de -0,03 points par rapport à l'autre sujet, cette différence n'est pas statistiquement significative pour α=0.05, toutes choses égales par ailleurs. Le sujet ayant le gène APOE4 aura en moyenne un coefficient de pente instantanée inférieur de 1,33 points par rapport à l'autre sujet ($IC_{95\%} = [-2, 52; -0.14]$), toutes choses égales par ailleurs.

Une femme a en moyenne un score d'IST intial supérieur de 0,77 points ($IC_{95\%} = [0, 09; 1, 45]$) par rapport à un homme, toutes choses égales par ailleurs. La femme aura en moyenne un coefficient de pente instantanée supérieur de 0,14 points par rapport à un homme, cette différence n'est pas statistiquement significative α=0,05, toutes choses égales par ailleurs.

Un sujet ayant fait des études longues aura en moyenne un score d'IST initial supérieur de 3,65 points par rapport à un sujet ayant fait des études courtes ($IC_{95\%} = [2,98; 4,31]$), toutes choses égales par ailleurs. Le sujet ayant fait des études longues aura en moyenne un coefficient de pente instantanée inférieure de 0,17 points par rapport à un sujet ayant fait des études courtes, cette différences n'est pas statistiquement significative α=0,05, toutes choses égales par ailleurs.

Pour deux sujets ayant un écart d'âge de 5 ans, le sujet le plus âgé aura en moyenne un coefficient de pente instantanée inférieure de -1,68 points de coefficient par rapport au plus jeune ($IC_{95\%} = [-2,15; -1,21]$), toutes choses égales par ailleurs.

Variable	Indépendant	CIE	WCIE, ns 0 noeuds	WCIE, ns 1 noeud
IMC				
Intercept	$26,74^{***}$	$26,74^{***}$	$26,73^{***}$	$26,73^{***}$
Age	0,29	0,30	0,31	0,32
Age^2	$-0,44^{***}$	$-0,44^{***}$	$-0,45^{***}$	$-0,45^{***}$
Niveau initial du IST				
Intercept	$27,28^{***}$	$29,06^{***}$	$29,17^{***}$	$29,15^{***}$
Age à l'inclusion	$0,10^{**}$	$0,10^{**}$	$0,11^{**}$	$0,11^{**}$
APOE4 (ref=non)	-0,01	-0,04	-0,02	-0.03
Sexe (ref=homme)	$0,81^{*}$	$0,77^{*}$	$0,77^{*}$	$0,77^{*}$
Etudes (ref=courtes)	$3,72^{***}$	$3,64^{***}$	$3,65^{***}$	$3,65^{***}$
w_0				
θ_{00}		-0,006	-0,006	-0,006
Pente instantané du IST				
Intercept	0,81	0,16	0,61	0,33
APOE4 (ref=non)	$-1,33^{*}$	$-1,32^{*}$	$-1,33*$	$-1,33*$
Sexe (ref=homme)	0,11	0,13	0,14	0,14
Etudes (ref=courtes)	-0,23	-0,19	-0,17	-0,17
Age	$-3,64^{***}$	$-3,34^{***}$	$-3,34^{***}$	$-3,36^{***}$
w_1				
θ_{10}		0,002	$-0,196^{**}$	-0,022
θ_{11}			$0,489^{**}$	0,000
θ_{12}				$0,470^{***}$

TABLE 9: Tableau comparatif des modèles sur les effets fixes de l'IST et IMC (n=1 621 ;# observations=14 703).

* : pvaleur $<$ 0.05 ** : pvaleur $<$ 0.001 *** : pvaleur $<$ 0.0001

gris : non estimé

L'âge à l'inclusion est centré à 64 et l'âge est centré à 64 ans et réduit à l'échelle de 0.1 : $\frac{Age-64}{10}$

5.3 Adéquation du modèle conjoint

Dans cette partie nous allons évaluer l'adéquation du modèle conjoint de notre méthodologie. L'adéquation est vérifiée graphiquement en regardant la valeur observée et prédite avec son intervalle de confiance à 95% pour tous âges, pour l'IST et l'IMC. L'adéquation du modèle a été vérifiée pour les modèles ayant leur hessienne inversible. Nous retrouvons les graphiques correspondant dans la figure 12 pour l'IMC et 13 et 14 pour l'IST ci-dessous.

Nous pouvons observer une bonne adéquation du modèle pour l'IMC, les valeurs prédites et observées étant très similaires. Nous retrouvons bien que pour tous les modèles dans la figure 12 la prédiction est la même, en effet les modèles ont la même structure de prédiction pour l'IMC.

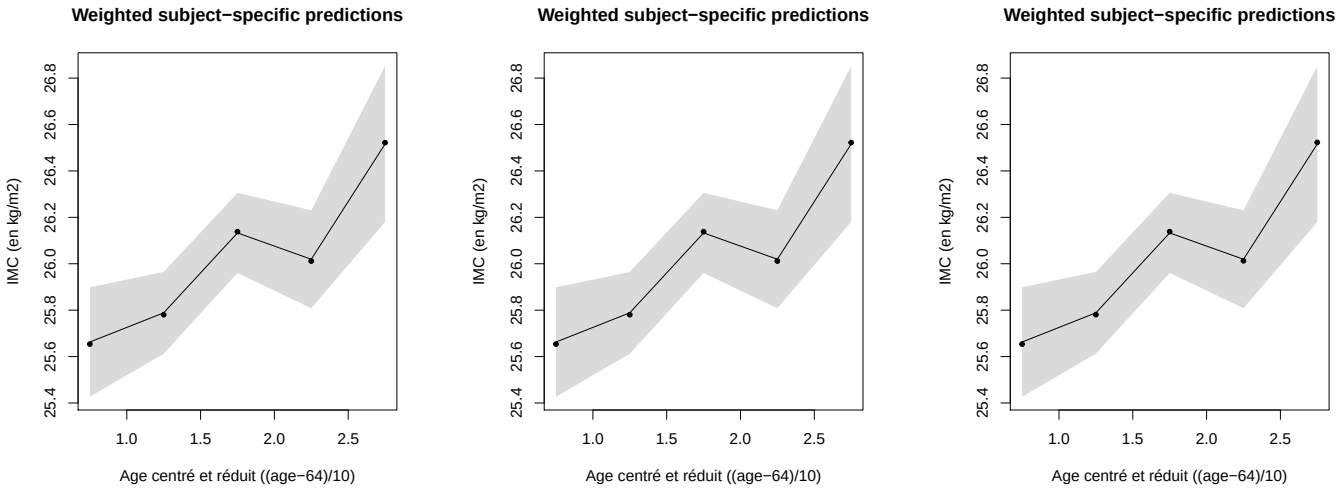

FIGURE 12: Graphique comparatif des valeurs prédites et observées pour l'IMC dans les modèles WCIE ns 0 noeud (à gauche) et 1 noeud (à droite) (n=1 621 ; # observations =14 703) .

Pour l'IST nous remarquons que le modèle prédictif avec CIE présente de moins bonnes performances d'adéquation. Les valeurs prédites sont statisfaisantes dans le modèle CIE (voir figure 14), cependant nous remarquons une amélioration de la véracité de ces prédictions quand nous passons à un modèle WCIE dans la figure 13. Parmi les modèles WCIE ns, l'adéquation des modèles WCIE ns 0 noeud et 1 noeud est similaire.

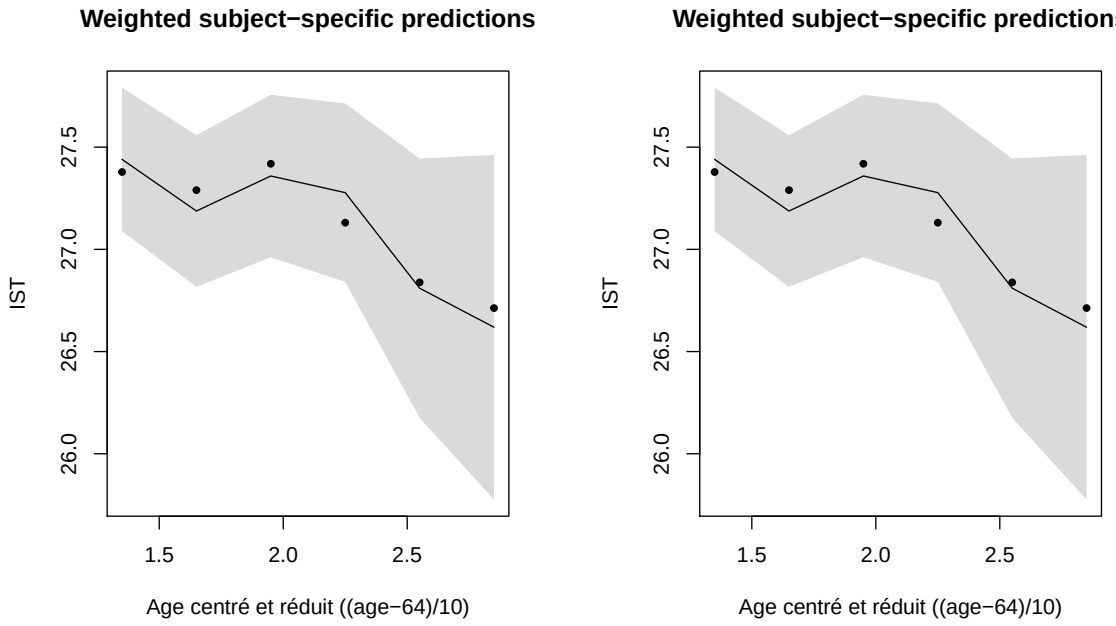

FIGURE 13: Graphique comparatif des valeurs prédites et observées pour l'IST dans les modèles WCIE ns 0 noeud (à gauche) et 1 noeud (à droite) (n=1 621 ; # observations =14 703) .

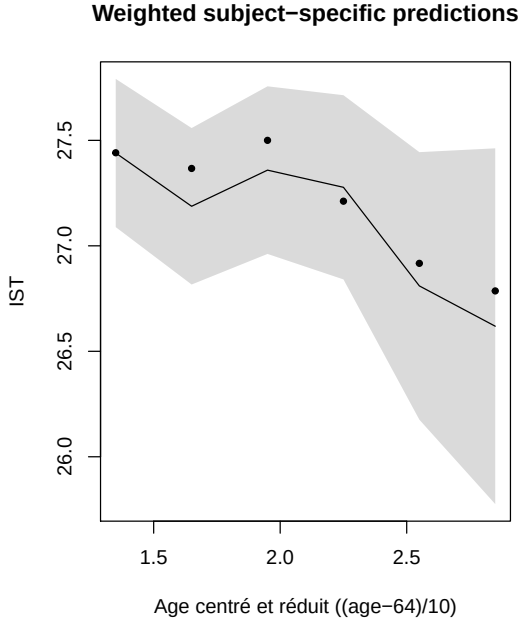

FIGURE 14: Graphique comparatif des valeurs prédites et observées pour l'IST dans le modèle CIE (n=1 621 ; # observations =14 703) .

En complément, nous avons vérifié la spécification du modèle via :

— la comparaison des résidus spécifiques aux sujets et des prédictions spécifiques aux sujets ;

— Q-Q plot des résidus spécifiques aux sujets.

Nous mettons dans cette annexe seulement les résultats pour le modèle WCIE ns avec 1 noeud dans la figure 15 pour l'IMC et la figure 16 pour l'IST. On observe bien que les résidus sont centrés en 0 et distribués également autour de cet axe, tant pour l'IMC que l'IST. Le Q-Q plot nous permet de vérifier que nos résidus spécifiques aux sujets suivent bien une loi normale, cette hypothèse est vérifiée ici tant pour l'IMC que l'IST.
Les autres modèles présentes des résultats similaires.

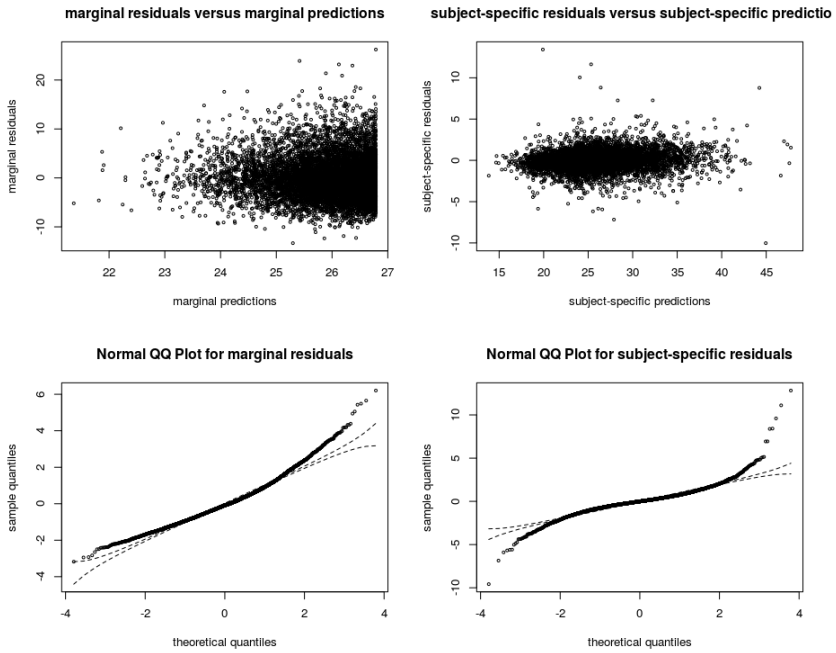

FIGURE 15: Distribution des résidus spécifiques aux sujets dans le modèle WCIE ns avec 1 noeud pour l'IMC (n=1 621 ; # observations =14 703) .

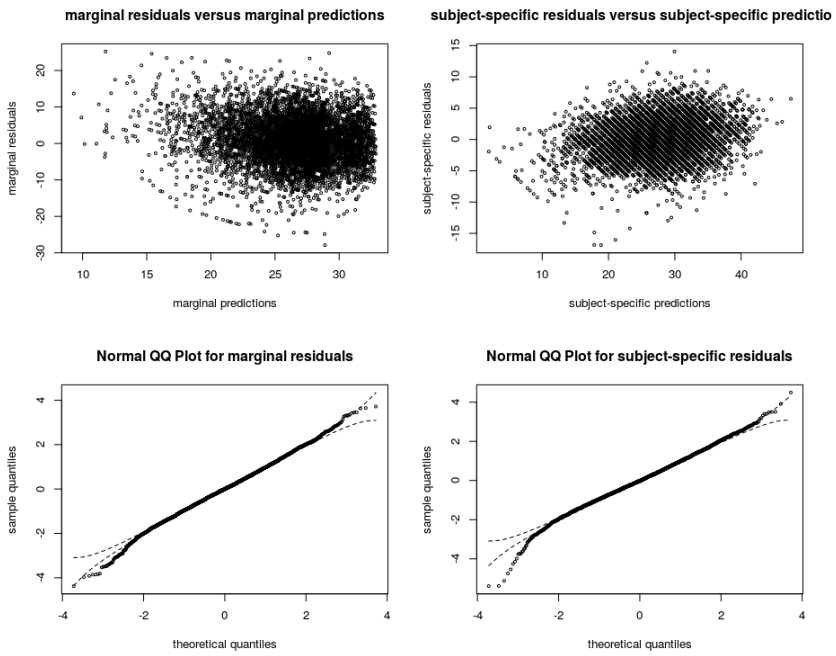

FIGURE 16: Distribution des résidus spécifiques aux sujets dans le modèle WCIE ns avec 1 noeud pour l'IST (n=1 621 ; # observations =14 703) .

5.4 Représentation des prédictions séparées de l'IMC et l'IST

Dans cette partie nous avons représenté les prédictions pour des âges donnés de l'IMC (figure 17) et l'IST (figure 18) dans les modèles indépendants. Ces prédictions ont été faites sur l'ensemble du jeu de données des trois cités, leurs représentations graphique nous permettent de regarder les convergences ou divergences entre les fonctions du temps utilisées dans les modèles.

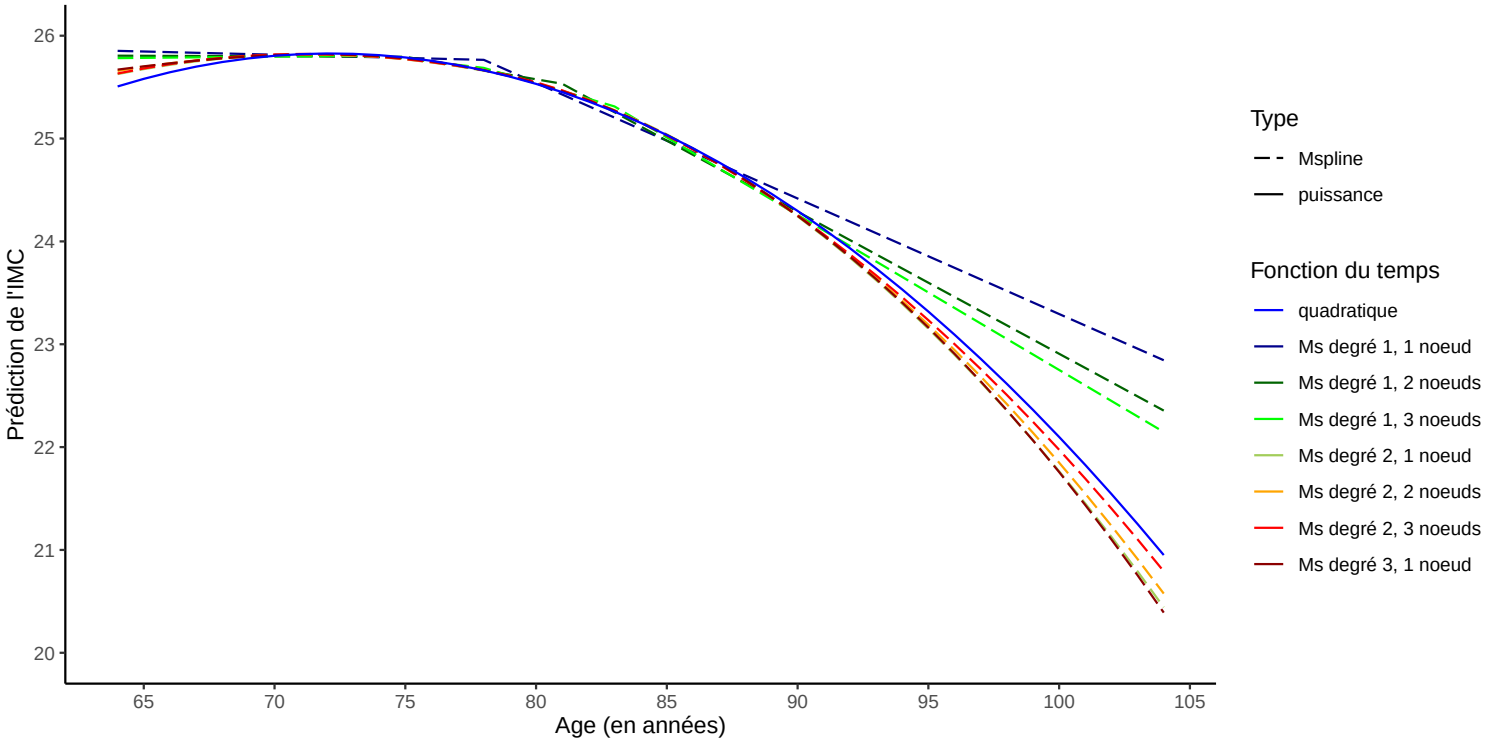

FIGURE 17: Prédictions des modèles mixtes pour l'IMC en fonction de la fonction du temps choisie (n=9 294# observations = 33 824).

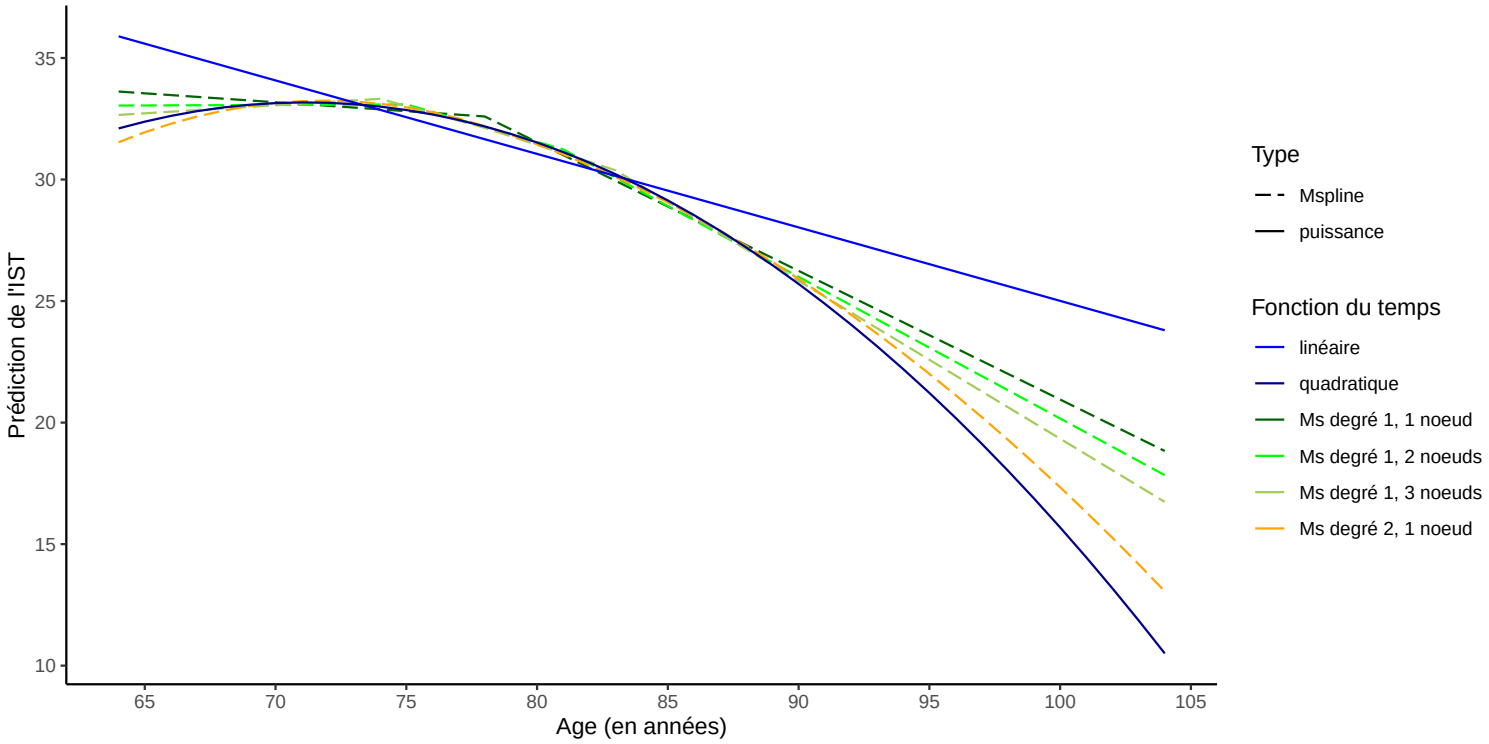

FIGURE 18: Prédictions des modèles mixtes pour l'IST en fonction de la fonction du temps choisie (n=9 294# observations = 39 315).

5.5 Représentation des fonctions splines

Dans cette partie nous avons choisi de représenter les ns et bsplines utilisés pour la modélisation des poids dans les modèles WCIE.

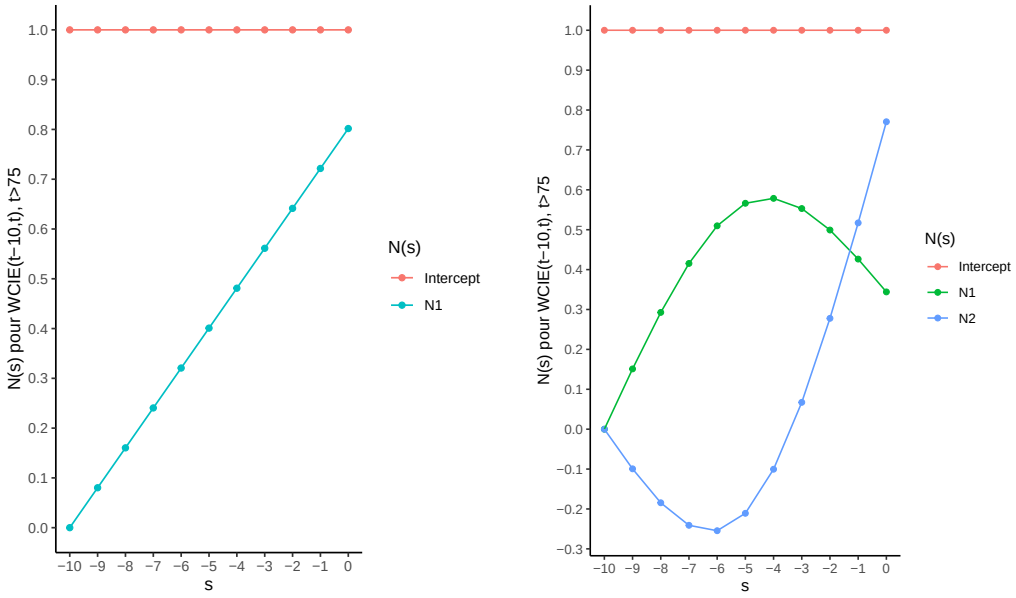

FIGURE 19: Présentation des fonctions nspline du WCIE de l'IMC sur la pente instantanée de l'IST, à gauche 0 noeud et à droite avec 1 noeud (S=10).

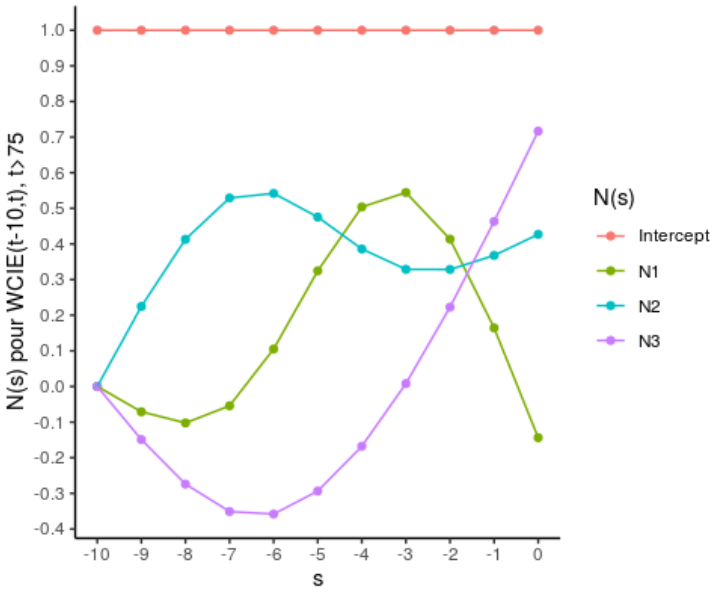

FIGURE 20: Présentation des fonctions nspline du WCIE de l'IMC sur la pente instantanée de l'IST avec 2 noeud (S=10).

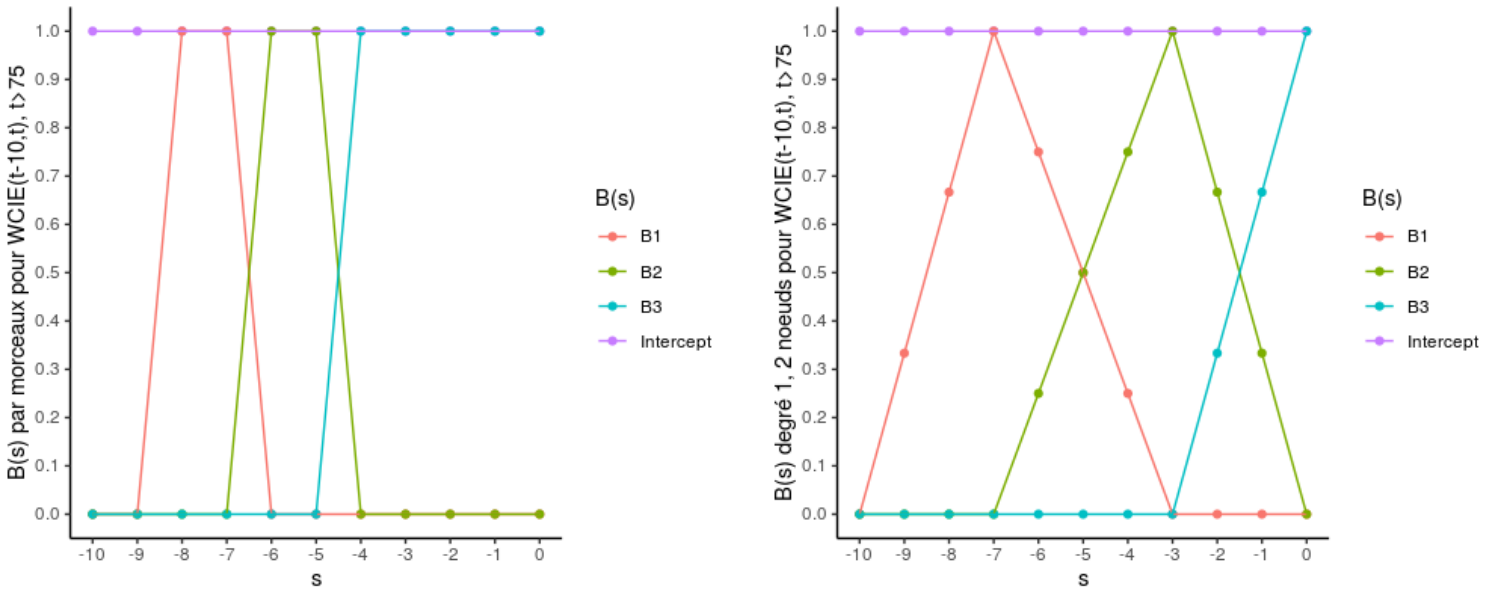

FIGURE 21: Présentation des fonctions bspline du WCIE de l'IMC sur la pente instantanée de l'IST, à gauche par morceaux et à droite de degré 1 à 2 noeuds (S=10).

5.6 Données de simulation

Les données de simuation ont pour visée de valider notre méthodologie d'estimation et d'implémentation informatique. Nous avons généré des données de cohort basées sur les caractéristiques de la cohorte des trois cités. Nous allons générer des temps de visite spécifique au sujet en utilisant une distribution uniforme sur [-6,6] mois autour de la visite théorique toutes les deux années du début de suivi 0 à 30 années de suivi maximum. A chaque visite nous simulons la valeur observée de l'exposition E et celle du marqueur Y d'après le modèle conjoint. Nous considérons le scénario suivant :

- E est explicité par une fonction linéaire du temps et un intercept aléatoire ;
- La valeur initiale de Y, $Y(S)$, est explicitée par une variable explicative vexp, l'histoire de l'exposition sur les dix premières années de suivi $[0, S]$ et un intercept aléatoire ;
- La pente de Y en t est explicitée par la variable explicative vexp et l'histoire de l'exposition sur les dix années précédentes $[t - S, t]$;

Vexp est modélisée par une binomiale de probabilité de succès de 0.5 pour chaque sujet et sans donnée manquante (par la suite nous pourrons en introduire). Nous considérons une proportion de données manquantes de l'ordre de dix pourcents pour l'exposition et le marqueur. Nous supprimons les valeurs du marqueur antérieur à 0+S, soit les données de Y pour les dix premières années. Les poids $w.(s)$ sont modélisés en fonction de deux scénarios :

- Scénario A, une association négative constante pour toutes les valeurs de s (CIE) ;
- Scénario B, une association négative évolutive, qui se rapproche de la valeur 0 quand s se rapproche de 0, une normale tronquée est utilisée pour modéliser $w.(s)$ avec comme moyenne 0 et écart-type 2, retranché de 1 et divisé par 10 ;

Le modèle est implémenté pour 500 répliques avec des échantillions de 500 participants. Pour chaque scénario nous allons regarder le biais d'estimation et le taux de couverture à 95 pourcents pour chaque paramètre fixe et erreur de mesure.