



**HAL**  
open science

# **Modélisation conjointe de la dynamique de biomarqueurs et de données de survie à risques compétitifs : développements et applications aux maladies infectieuses sévères**

Alexandra Lavalley-Morelle

## ► To cite this version:

Alexandra Lavalley-Morelle. Modélisation conjointe de la dynamique de biomarqueurs et de données de survie à risques compétitifs : développements et applications aux maladies infectieuses sévères. Statistiques [math.ST]. Université Paris Cité, 2023. Français. ⟨NNT : 2023UNIP5111⟩. ⟨tel-04902080⟩

**HAL Id: tel-04902080**

**<https://theses.hal.science/tel-04902080v1>**

Submitted on 20 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## Université Paris Cité

ÉCOLE DOCTORALE PIERRE LOUIS DE SANTÉ PUBLIQUE : ÉPIDÉMIOLOGIE ET  
SCIENCES DE L'INFORMATION BIOMÉDICALE (ED 393)

**Unité de recherche :** Infection, Anti-microbien, Modélisation, Évolution (UMR-S 1137)

**Équipe :** Modélisation biostatistique, pharmacométrie et investigation clinique dans les  
maladies infectieuses

# MODÉLISATION CONJOINTE DE LA DYNAMIQUE DE BIOMARQUEURS ET DE DONNÉES DE SURVIE À RISQUES COMPÉTITIFS : DÉVELOPPEMENTS ET APPLICATIONS AUX MALADIES INFECTIEUSES SÉVÈRES

Par Alexandra LAVALLEY-MORELLE

Thèse de doctorat de biostatistiques et biomathématiques

Dirigée par France MENTRÉ et co-encadrée par Jimmy MULLAERT

Présentée et soutenue publiquement le 13 novembre 2023

Devant un jury composé de :

Cécile PROUST-LIMA, DR	Université de Bordeaux	Rapporteur
Étienne DANTAN, MCU-HDR	Université de Nantes	Rapporteur
Raphaël PORCHER, PU-PH	Université Paris Cité	Examineur
Maud DELATTRE, CR	INRAE	Examineur
Jean-Jacques PARIENTI, PU-PH	Université de Caen Normandie	Examineur
France MENTRÉ, PU-PH	Université Paris Cité	Directrice de thèse
Jimmy MULLAERT, MCU-PH	Université de Versailles- Saint-Quentin-en-Yvelines	Membre invité

# Remerciements

## Résumé en Français

**Titre : Modélisation conjointe de la dynamique de biomarqueurs et de données de survie à risques compétitifs : développements et applications aux maladies infectieuses sévères**

Aujourd'hui, les hôpitaux sont équipés de systèmes d'information qui enregistrent quotidiennement les résultats d'analyses biologiques prescrites par les médecins. Ces observations biologiques longitudinales représentent une grande quantité d'information qui peut être utilisée dans un modèle joint pour prédire le pronostic du patient de façon dynamique, et ainsi guider les décisions thérapeutiques ou prévoir les besoins des hôpitaux. Les modèles joints proposés dans la littérature impliquent généralement un modèle linéaire à effets mixtes décrivant l'évolution d'un unique biomarqueur, et un modèle de survie décrivant le risque instantané de décès, en considérant la sortie de l'hôpital comme événement en compétition. Cependant, les biomarqueurs ont souvent une dynamique non-linéaire qui n'est pas optimalement capturée par des modèles linéaires.

Pour répondre à cette question, le premier projet de cette thèse s'est intéressé au pronostic des patients admis en réanimation pour sepsis, dont le score Sequential Organ Failure Assessment (SOFA) était évalué chaque jour jusqu'au décès ou à la sortie de réanimation. Un modèle non-linéaire à effets mixtes décrivant l'évolution du SOFA est combiné avec un modèle de survie à risques compétitifs pour modéliser le risque de décès à l'hôpital, en prenant en compte la sortie comme événement en compétition. Dans le modèle de survie, une approche par risques de sous-distribution, communément distinguée de l'approche cause-spécifique, est employée. Nous avons montré la capacité de l'algorithme Stochastic Approximation Expectation Maximization (SAEM) implémenté dans le logiciel Monolix à fournir des estimations non biaisées des paramètres d'un tel modèle.

D'autre part, le fait de passer d'un seul biomarqueur à des centaines pourrait certainement améliorer les performances prédictives, mais poserait également des problèmes statistiques. Certains auteurs ont décrit des problèmes computationnels et d'identifiabilité dus au nombre élevé d'effets aléatoires lorsque de multiples modèles longitudinaux sont

estimés conjointement avec un modèle de survie. Cela explique probablement pourquoi les modèles publiés se limitent souvent à des modèles longitudinaux comportant au maximum deux biomarqueurs et souligne la nécessité de disposer de méthodes de sélection dans ce contexte. C'est dans ce cadre que le deuxième projet de cette thèse s'est intéressé au pronostic des patients hospitalisés pour infection au SARS-CoV-2, avec un suivi de 59 biomarqueurs jusqu'au décès ou à la sortie d'hospitalisation. Une stratégie de sélection a été appliquée pour construire un modèle joint multivarié prédictif du risque de décès des patients. L'algorithme SAEM de Monolix et la stratégie de sélection ont démontré de bonnes propriétés dans ce cadre plus complexe. Dans ces deux projets, nous avons quantifié, par des approches de prédictions dynamiques individuelles, le bénéfice du suivi des biomarqueurs pour améliorer le pronostic des patients.

Enfin, dans un dernier projet, nous avons fait le constat que les outils disponibles pour estimer des modèles joints non-linéaires étaient presque exclusivement disponibles dans des logiciels issus de la pharmacométrie. Pour diffuser l'utilisation de ces méthodes, nous avons étendu le package R `saemix`, initialement dédié aux modèles non-linéaires à effets mixtes, au cas des modèles à réponses multiples et joints. Des modèles joints très spécifiques peuvent être estimés, l'utilisateur définissant lui-même la vraisemblance du modèle. Une étude de simulation a démontré de bonnes performances d'estimation des paramètres et de leur incertitude. Ces travaux ouvrent la voie au développement de modèles conjoints sophistiqués.

**Mots-clés : Modélisation conjointe, Modèles non-linéaires à effets-mixtes, Analyse de survie, Risques compétitifs, Algorithme SAEM, Prédictions dynamiques, Maladies infectieuses sévères**

# Résumé en Anglais

**Title : Joint modeling of biomarker dynamics and survival with competing risks : developments and applications on severe infectious diseases**

Nowadays, hospitals are equipped with laboratory information systems that routinely gather all the results of biological analyses prescribed by the physicians. These longitudinal biological observations represent a huge amount of information that can be used in a joint model to provide individual dynamic predictions of patient prognosis, and thus guide therapeutic management or forecast hospital needs. Joint models proposed in the literature usually involve a linear mixed-effects model describing the evolution of a single biomarker, and a survival model describing the instantaneous risk of death, considering hospital discharge as a competing event. However, biomarkers often follow nonlinear dynamics that are not optimally captured by linear models.

To address this question, the first project of this thesis focused on the prognosis of patients admitted in Intensive Care Unit (ICU) for sepsis, whose Sequential Organ Failure Assessment (SOFA) score was daily assessed until death or discharge. A nonlinear mixed-effects model to describe the evolution of the SOFA score is combined with a competing risk survival model to model the risk of in-hospital death, considering ICU discharge as a competing event. In the survival model, a subdistribution risk approach, commonly distinguished from the cause-specific approach, is considered. We showed the ability of the Stochastic Approximation Expectation Maximization (SAEM) algorithm implemented in Monolix software to provide unbiased estimates of the model parameters.

On the other hand, shifting the modeling from a single biomarker to potentially hundreds could certainly improve predictive performances, but also comes with statistical issues. Some authors described computational and identifiability issues due to the high number of random effects when multiple longitudinal models are jointly estimated with a survival model. This probably explains why published models are often limited to longitudinal models with at most two biomarkers, and highlights the need for selection methods in this context. Within this framework, the second project of this thesis focused on the

prognosis of patients hospitalized for SARS-CoV-2 infection, with the follow-up of 59 biomarkers until death or discharge. A selection strategy was applied to build a multivariate joint model predicting the risk of patient death. Both the SAEM algorithm implemented on Monolix and the selection strategy demonstrated good properties in this more complex setting. In these two projects, we used individual dynamic prediction approaches to quantify the benefits of biomarker monitoring for improving patient prognosis.

Finally, in a last project, we emphasized that the tools available for nonlinear joint model estimation were almost exclusively available in pharmacometric software. To spread the use of these methods, we extended the R package `saemix`, initially designed to fit nonlinear mixed-effects models, to the case of multiple-response and joint models. A wide range of joint models can be estimated, as the user defines himself the model's likelihood. A simulation study demonstrated good performances for parameters and standard errors estimation. This work paves the way for the development of sophisticated joint models.

**Key words : Joint modeling, Nonlinear mixed-effects models, Survival analysis, Competing risks, SAEM algorithm, Dynamic predictions, Infectious diseases**

---

## Table des matières

---

<b>1</b>	<b>Introduction</b>	<b>13</b>
1.1	Modélisation conjointe de données longitudinales et de données de survie .	14
1.1.1	Modèles à effets mixtes . . . . .	14
1.1.2	Modèles de survie à risques compétitifs . . . . .	15
1.1.3	Modélisation conjointe . . . . .	20
1.1.4	Prédictions individuelles dynamiques . . . . .	23
1.2	Inférence dans le cadre fréquentiste . . . . .	27
1.2.1	Estimation des paramètres . . . . .	27
1.2.2	Estimation des erreurs standard . . . . .	30
1.2.3	Logiciels disponibles . . . . .	33
1.3	Applications aux maladies infectieuses sévères . . . . .	36
1.3.1	Le sepsis . . . . .	36
1.3.2	La COVID-19 . . . . .	39
1.4	Objectifs de la thèse . . . . .	45
1.4.1	Modélisation conjointe univariée . . . . .	45
1.4.2	Modélisation conjointe multivariée . . . . .	47
1.4.3	Dissémination . . . . .	47

<b>2</b>	<b>Modélisation conjointe avec risques compétitifs : application à des prédictions de survie de patients admis en réanimation pour sepsis, avec un suivi journalier du score SOFA</b>	<b>49</b>
2.1	Résumé . . . . .	49
2.2	Article 1 (publié) . . . . .	51
<b>3</b>	<b>Modèle joint multivarié avec risques compétitifs pour prédire la survie des patients hospitalisés pour infection au SARS-CoV-2</b>	<b>83</b>
3.1	Résumé . . . . .	83
3.2	Article 2 (accepté) . . . . .	85
<b>4</b>	<b>Extension du code du package libre <i>saemix</i> pour l’ajustement de modèles conjoints</b>	<b>129</b>
4.1	Résumé . . . . .	129
4.2	Article 3 (soumis) . . . . .	131
<b>5</b>	<b>Discussion et conclusion</b>	<b>167</b>
5.1	Discussion . . . . .	168
5.2	Perspectives . . . . .	172
5.2.1	Conclusion générale . . . . .	175

## Publications de thèse

- **LAVALLEY-MORELLE A.**, TIMSIT JF., MENTRÉ F., MULLAERT J, THE OUTCOMEREA NETWORK, Joint modeling under competing risks : Application to survival prediction in patients admitted in Intensive Care Unit for sepsis with daily Sequential Organ Failure Assessment score assessments, *CPT : Pharmacometrics & Systems Pharmacology*. 2022;11(11) :1472-1484. doi : 10.1002/psp4.12856
- **LAVALLEY-MORELLE A.**, PEIFFER-SMADJA N., GRESSENS S., SOUHAIL B., LAHENS A., BOUNHIOL A., LESCURE FX., MENTRÉ F., MULLAERT J., Multivariate joint model under competing risks to predict death of hospitalized patients for SARS-CoV-2 infection, *Accepté sous conditions\* pour publication dans Biometrical Journal le 26 juillet 2023 (\* conditionnellement au fait que l'équipe de reproduction valide le code fournit pour reproduire les résultats)*
- **LAVALLEY-MORELLE A.**, MENTRÉ F., COMETS E.\*, MULLAERT J.\*, Extending the code in the open-source *saemix* package to fit joint models, *Soumis dans Computer Methods and Programs in Biomedicine le 5 juillet 2023*

\* Co-derniers auteurs

## Communications orales

- **Journées du GDR Statistiques et Santé** - 1<sup>er</sup> octobre 2020, "Prédictions individuelles dynamiques du risque de décès à l'aide d'un modèle conjoint : application aux patients admis en soins intensifs pour sepsis."
- **29th Population Approach Group in Europe (PAGE) conference** - 7 septembre 2021, "Individual dynamic predictions and their evaluation in joint analysis with non-linear longitudinal model and parametric competing risks model : application on sepsis patients"
- **30th Population Approach Group in Europe (PAGE) conference** - 30 juin 2022, "Longitudinal biomarkers predicting death of hospitalized patients for SARS-CoV-2 infection : a joint analysis with competing risks"
- **43rd Annual Conference of the International Society for Clinical Biostatistics (ISCB)** - 24 août 2022, "Marker selection in joint analysis with competing risks : application to hospitalized patients for SARS-CoV-2 infection"
- **Journées de Biostatistiques** - 17 novembre 2022, "Modèle joint multivarié avec risques compétitifs pour prédire le risque de décès des patients hospitalisés pour infection au SARS-CoV-2"
- **Séminaire invité - Anticipe U1086 Caen** - 26 juin 2023, "Longitudinal biomarkers predicting death of hospitalized patients for SARS-CoV-2 infection : a joint analysis with competing risks"
- **31st Population Approach Group in Europe (PAGE) conference, Lewis Sheiner student session award** - 29 juin 2023, "Joint modeling of biomarkers dynamics and survival with competing risks to predict the prognostic of patients hospitalized with severe infectious diseases"

## Communications affichées

- **42nd Annual Conference of the International Society for Clinical Biostatistics (ISCB)** - Juillet 2021, "Individual dynamic predictions and their evaluation in joint analysis of nonlinear longitudinal model and parametric competing risk model : application on sepsis patients"
- **15èmes conférences francophones d'Epidémiologie clinique (EPICLIN)** - Mai 2022, "Individual dynamic predictions in joint analysis of nonlinear longitudinal model and parametric competing risk model : application on patients admitted in ICU for sepsis"

## Prix scientifiques

- **Prix du second meilleur poster, 42nd Annual Conference of the International Society for Clinical Biostatistics (ISCB)**, 19 juillet 2021
- **Prix étudiant Lewis Sheiner, 31st Population Approach Group in Europe (PAGE) conference**, 29 juin 2023, <https://www.page-meeting.org/?abstract=10394>



# CHAPITRE 1

---

## Introduction

---

Ce travail de thèse a été initié afin d'étudier des méthodes de modélisation conjointe de données longitudinales et de survie lorsque plusieurs événements sont en compétition, situation fréquemment rencontrée dans les applications cliniques. En particulier, dans le cadre de maladies infectieuses sévères, ce travail s'est intéressé à des modèles conjoints possiblement non-linéaires afin de caractériser le lien entre la dynamique de biomarqueur(s) et du risque de décès et de sortie d'hospitalisation. Cette démarche contribue à la médecine personnalisée en fournissant des outils pronostiques performants comme des prédictions d'évènements individualisées et mises à jour au fur et à mesure que de nouvelles observations longitudinales sont récoltées.

Dans cette introduction, le principe de la modélisation conjointe de données longitudinales et de données de survie, ainsi que leur utilisation à des fins prédictives seront d'abord introduits. Puis, les méthodes d'estimation existantes seront exposées, en insistant sur les logiciels disponibles et leurs limites. Les deux domaines d'applications cliniques de la thèse, centrés sur des maladies infectieuses sévères seront ensuite présentés. Finalement, ce chapitre se clôturera en énonçant les trois objectifs de la thèse.

# 1.1 Modélisation conjointe de données longitudinales et de données de survie

## 1.1.1 Modèles à effets mixtes

Généralement, lors d'études de cohortes ou d'essais cliniques, de nombreuses mesures sont récoltées de façon répétée au cours du suivi du patient. Ces mesures répétées font souvent référence au suivi longitudinal d'un biomarqueur clinique. Pour analyser correctement ces informations, trois principaux facteurs sont à prendre en compte : l'analyse simultanée de plusieurs patients différents, l'erreur de mesure existante sur chaque observation, et la corrélation entre les mesures d'un même patient. Les modèles à effets-mixtes représentent un outil particulièrement adapté à l'analyse de ces données (Brown, 2021). Dans le cadre de cette approche,  $y = (y_i, 1 \leq i \leq N)$  dénote le vecteur des observations longitudinales, avec  $N$  le nombre de patients analysés. Le vecteur  $y_i = (y_{ij}, 1 \leq j \leq n_i)$  réfère aux observations du patient  $i$ , avec  $n_i$  son nombre total de mesures. Ainsi,  $y_{ij}$  dénote la  $j$ -ème mesure du biomarqueur pour le patient  $i$  au temps  $t_{ij}$ . Pour  $i \in \{1, \dots, N\}$  et  $j \in \{1, \dots, n_i\}$ , un modèle à effets mixtes peut s'écrire de la façon suivante :

$$y_{ij} = m(t_{ij}, \psi_i) + g(t_{ij}, \psi_i, \sigma) \epsilon_{ij} \quad (1.1)$$

La fonction  $m$  représente le modèle structurel qui décrit l'évolution individuelle du biomarqueur, en l'absence d'erreur de mesure. Cette fonction dépend du temps et de paramètres individuels  $\psi_i$ . Ces paramètres individuels dépendent eux d'un vecteur de paramètres propres à la population, appelés effets fixes et notés  $\mu$ , d'effets aléatoires  $\eta_i$  propres au patient  $i$ , et d'un éventuel vecteur de covariables  $Z$ . Ce sont ces effets aléatoires qui permettent de prendre en compte la corrélation des mesures d'un même patient, et de quantifier la manière dont les individus diffèrent du comportement moyen (de population). On suppose également que  $\eta_i \sim \mathcal{N}(0, \Omega)$ ,  $\Omega$  représentant la matrice de variance-covariance des effets aléatoires. Les termes d'erreurs  $\epsilon_{ij}$  sont supposés indépendants et identique-

ment distribués selon une loi normale  $\mathcal{N}(0, 1)$  et la fonction  $g$  représente la forme du modèle d'erreur. Des formes très utilisées pour  $g$  sont le modèle d'erreur combiné, où  $g(t_{ij}, \psi_i, \sigma) = \sigma_a + \sigma_b m(t_{ij}, \psi_i)$ , proportionnel (lorsque  $\sigma_a = 0$ ) ou encore additif (lorsque  $\sigma_b = 0$ ).

Le vecteur des paramètres à estimer est noté  $\theta^{(l)}$  et se compose des effets fixes  $\mu$ , des éventuels effets des covariables  $\beta$ , de la matrice de variance-covariance des effets aléatoires  $\Omega$ , et des paramètres du modèle d'erreur  $\sigma$ .

### 1.1.2 Modèles de survie à risques compétitifs

L'analyse de données de survie classique repose sur la modélisation du temps écoulé jusqu'à l'observation d'un évènement d'intérêt. Ces données sont souvent incomplètes par le biais du phénomène de censure (Leung et al., 1997). Nous considérons dans ce travail un processus de censure à droite, c'est à dire lorsque le patient n'a éventuellement pas fait l'évènement à sa dernière observation. Une hypothèse majeure dans l'analyse de survie est que la censure est non-informative, c'est à dire que les individus censurés ont le même risque futur d'évènement que les individus encore observés (on parle d'indépendance conditionnelle). Cependant, dans certaines études, il arrive qu'un autre évènement vienne perturber l'observation de l'évènement d'intérêt. C'est ce qu'on appelle un évènement en compétition. Par exemple, si on souhaite modéliser le temps jusqu'au décès des patients admis en réanimation pour une infection sévère, la sortie de réanimation vient empêcher d'observer le décès. Censurer les patients sortis violerait l'hypothèse de censure non-informative. En effet, la sortie de réanimation apporte une information sur l'état de santé du patient : il sort de réanimation et donc n'est plus observé dans l'analyse parce que son état s'est amélioré.

Les modèles de survie classiquement utilisés (type modèle de Cox) pour l'étude d'un unique évènement ne sont plus adaptés. C'est pour cette raison que les modèles de survie à risques compétitifs ont été introduits. Dans le cadre de cette thèse, uniquement deux évènements en compétition sont considérés. Cependant, les principes qui suivent peuvent

être facilement étendus au cas où plus de deux évènements sont en compétition. Soit  $T$  la variable aléatoire qui représente les temps d'évènements et  $C$  celle qui représente les temps de censure (non-informative) à droite. Les données de survie observées sont le couple de vecteurs  $(\tilde{T}, \delta)$  avec  $\tilde{T} = (\min(T_1, C_1), \dots, \min(T_N, C_N))$  et  $\delta = (\delta_1, \dots, \delta_N)$  le vecteur des indicatrices d'évènements. Pour un patient  $i \in \{1, \dots, N\}$ ,  $\delta_i = 1$  (respectivement  $\delta_i = 2$ ) si le patient a fait l'évènement 1 (respectivement l'évènement 2), et  $\delta_i = 0$  si le patient est censuré. L'incidence cumulée spécifique à la cause  $\delta = e$  ( $e = \{1, 2\}$ ) est définie par :

$$F_e(t) = \mathbb{P}(T \leq t, \delta = e) \quad (1.2)$$

Plusieurs approches ont été décrites pour modéliser cette quantité (Mozumder et al., 2018). La première représente l'extension naturelle du modèle de Cox et est appelée l'approche cause-spécifique (Prentice et al., 1978). Cette dernière s'intéresse au risque instantané d'évènement  $\delta = e$  sachant que le patient n'a fait aucun des évènements. En d'autres termes, les risques instantanés cause-spécifique sont définis de la manière suivante :

$$h_e^{CS}(t) = \lim_{dt \rightarrow 0} \frac{\mathbb{P}(t < T < t + dt, \delta = e \mid T > t)}{dt} \quad (1.3)$$

Les patients ayant fait un évènement  $\delta \neq e$  sont censurés pour la cause  $\delta = e$ . Pour dériver la fonction d'incidence cumulée spécifique à la cause  $\delta = e$ , la notion de survie toute cause confondue, notée  $S$ , nécessite d'être introduite :

$$S(t) = \exp \left( - \int_0^t \sum_{e=1}^2 h_e^{CS}(u) du \right) \quad (1.4)$$

La fonction d'incidence cumulée est ensuite obtenue par l'équation suivante :

$$F_e(t) = \int_0^t S(u) h_e^{CS}(u) du \quad (1.5)$$

Sous l'approche cause-spécifique, le lien entre risque instantané et fonction d'incidence cumulée spécifique à un évènement est donc indirect, puisqu'il fait intervenir le risque

instantané de l'autre évènement, par le biais de la survie toute cause confondue. En tant que conséquence directe de cette relation, une variable peut être associée d'un sens (positivement ou négativement) au risque instantané d'évènement, mais associée dans le sens inverse à la fonction d'incidence cumulée (donc à la probabilité d'évènement), ceci dû à l'interaction avec le risque instantané de l'autre évènement.

Pour contourner ce problème, l'approche sous-distribution a été introduite (Fine and Gray, 1999). Celle-ci s'intéresse au risque instantané d'évènement  $\delta = e$  sachant que le patient n'a pas fait l'évènement en question, ou a fait un des évènements en compétition. Les risques instantanés de sous-distribution sont définis par :

$$h_e^{SD}(t) = \lim_{dt \rightarrow 0} \frac{\mathbb{P}(t < T < t + dt, \delta = e \mid T > t \cup (T \leq t \cap \delta \neq e))}{dt} \quad (1.6)$$

Ainsi, les patients ayant fait un évènement  $\delta \neq e$  sont toujours considérés comme à risque pour la cause  $\delta = e$ . Dans ce cas, le lien entre le risque instantané et la fonction d'incidence cumulée est direct :

$$F_e(t) = 1 - \exp\left(-\int_0^t h_e^{SD}(u) du\right) \quad (1.7)$$

Même si les deux approches sont essentiellement deux paramétrisations du même processus aléatoire, des travaux ont suggéré que l'approche cause-spécifique était préférable pour des études étiologiques, alors que l'approche sous-distribution était plus propice à des fins prédictives (Putter et al., 2007; Lau et al., 2009). En effet, pour cette dernière approche, les covariables associées au risque instantané de sous-distribution sont directement liées à l'incidence cumulée. Enfin, d'importants travaux ont montré qu'il était préférable de fournir les estimations sous les deux approches (Latouche et al., 2013). Quelque soit l'approche considérée, le risque instantané lié à l'évènement  $\delta = e$  peut être décrit à l'aide d'une fonction paramétrique de type  $\lambda_e(t) = h_e(t) \exp(\beta_e W)$  où  $h_e(t)$  est une fonction de risque de base commune à tous les patients,  $W$  le vecteur des covariables individuelles, et  $\beta_e$  le paramètre mesurant la force d'association entre la covariable et le risque instantané d'évènement  $\delta = e$ . Une hypothèse de proportionnalité des risques est sous-jacente aux

modèles, mais ne peut pas être vraie simultanément pour les deux approches.

Un point important à souligner est l'interprétation des paramètres sous ces deux approches. La suite fait référence à l'exemple tiré de l'article de Austin and Fine (2017), regroupant des patients hospitalisés pour insuffisance cardiaque entre 1995 et 2009 et inclus dans l'étude EFFECT (Tu et al., 2009). Les patients sont suivis 5 ans depuis l'hospitalisation et leur temps jusqu'au décès ainsi que la cause du décès sont reportés. Les auteurs évaluent d'abord l'effet du sexe sur le décès toute cause confondue 5 ans après hospitalisation, avec un modèle de Cox classique. Le hazard ratio estimé et son intervalle de confiance à 95% est de 1.18 [1.14,1.23] pour les hommes (les femmes étant la modalité de référence). Sous l'hypothèse des risques proportionnels, on peut alors conclure que le risque instantané de décéder dans les 5 ans après une hospitalisation pour insuffisance cardiaque est 18% plus grand chez les hommes que chez les femmes. Comme le hazard ratio est supérieur à 1, on peut également conclure que les hommes ont une plus grande probabilité de décéder que les femmes. Les auteurs considèrent ensuite le cadre des risques compétitifs, où ils évaluent le décès pour cause cardiovasculaire dans les 5 ans après hospitalisation, en considérant toute autre cause de décès comme évènement en compétition. En considérant l'approche cause-spécifique, le hazard ratio associé au sexe sur le décès pour cause cardiovasculaire est de 1.20 [1.14,1.27]. Le risque instantané de décès pour cause cardiovasculaire est ainsi 20% plus grand chez les hommes toujours vivants que chez les femmes toujours vivantes. Sous cette approche, il n'est pas possible de tirer des conclusions de l'effet du sexe sur l'incidence cumulée de décès pour cause cardiovasculaire, donc sur la probabilité d'évènement. En effet, comme évoqué plus haut, le risque instantané d'une cause donnée et son incidence cumulée ne sont pas directement liés. En considérant l'approche sous-distribution, le hazard ratio associé au sexe sur le décès pour cause cardiovasculaire est de 1.16 [1.10,1.22]. Le risque instantané de décès cardiovasculaire est 16% plus grand chez les hommes (implicitement, non décédés pour cause cardiovasculaire) que chez les femmes (idem). Sous cette approche et contrairement à l'approche cause-spécifique, la correspondance directe entre l'effet d'une covariable sur

le risque instantané d'évènement et son effet sur l'incidence cumulée permet de retrouver l'interprétation du modèle de Cox. On peut ainsi conclure que les hommes ont une plus forte probabilité de décéder pour cause cardiovasculaire dans les 5 ans suivants une hospitalisation pour insuffisance cardiaque que les femmes.

Modéliser les risques instantanés de sous-distribution définis en équation 1.6 en semi-paramétrique est l'approche initialement proposée par Fine et Gray (Fine and Gray, 1999). Cette approche comporte deux principales limites. La première réside dans l'hypothèse de proportionnalité des risques de sous-distribution, difficile à apprécier en pratique. La deuxième limite est l'absence de contrainte pour la borne supérieure de l'incidence cumulée toute cause confondue (qui théoriquement ne peut pas excéder 100%, soit 1). En effet, avec l'approche sous-distribution et contrairement à l'approche cause-spécifique, la somme des incidences cumulées associées aux évènements peut dépasser 1 pour certaines valeurs de covariables. Pour illustrer ce propos, nous définissons deux risques instantanés cause-spécifiques  $\lambda_1^{CS}(t)$  et  $\lambda_2^{CS}(t)$ , et deux risques instantanés de sous-distribution  $\lambda_1^{SD}(t)$ ,  $\lambda_2^{SD}(t)$ . Notons  $\Lambda$  le risque instantané cumulé. Pour chaque cause d'évènement  $e \in \{1, 2\}$  :

$$F_e(t) = \int_0^t \lambda_e^{CS} \exp(-\Lambda_1^{CS} + \Lambda_2^{CS}), \text{ pour l'approche cause spécifique,}$$

$$F_e(t) = \int_0^t \lambda_e^{SD} \exp(-\Lambda_e^{SD}), \text{ pour l'approche sous-distribution.}$$

$\forall t \in \mathbb{R}^+$  la somme des incidences cumulées de chaque évènement vaut :

$$\begin{aligned} F_1(t) + F_2(t) &= \int_0^t (\lambda_1^{CS} + \lambda_2^{CS}) \exp(-\Lambda_1^{CS} - \Lambda_2^{CS}) \\ &= 1 - \exp(-\Lambda_1^{CS}(t) - \Lambda_2^{CS}(t)) < 1 \end{aligned} \quad (1.8)$$

$$\begin{aligned} F_1(t) + F_2(t) &= 1 - \exp(-\Lambda_1^{SD}(t)) + 1 - \exp(-\Lambda_2^{SD}(t)) \\ &= 2 - \exp(-\Lambda_1^{SD}(t)) - \exp(-\Lambda_2^{SD}(t)) < 2 \end{aligned} \quad (1.9)$$

Sous l'approche cause-spécifique (équation 1.8), la somme des incidences cumulées est par nature inférieure ou égale à 1, ce qui n'est pas nécessairement vrai pour l'approche sous-distribution (équation 1.9). Des précautions sont nécessaires quant à l'utilisation de

cette dernière approche.

Au cours de cette thèse, nous nous sommes intéressés à l'approche sous-distribution, peu considérée dans le cadre des modèles conjoints en dépit de ses propriétés intéressantes dans le cadre de prédictions individuelles.

### 1.1.3 Modélisation conjointe

Dans de nombreuses applications, des données longitudinales (suivi de biomarqueurs) et des données de survie sont récoltées simultanément. Dans ce cadre, modéliser le processus longitudinal par un modèle à effets mixtes peut engendrer des biais dans l'estimation des paramètres, car la présence d'évènements comme le décès censure la partie longitudinale de façon informative. Plus le lien entre le processus longitudinal et de survie est fort, plus ce phénomène est accentué (Desmée et al., 2015). Prenons l'exemple de patients hospitalisés en réanimation pour une infection donnée. Supposons qu'on s'intéresse à un biomarqueur à la fois lié au décès en réanimation (association positive) et à la sortie de réanimation (association négative). Plus le temps passe, moins les données longitudinales seront représentatives car les patients ayant de grandes ou de faibles valeurs de biomarqueurs seront sortis de l'étude pour cause de décès ou de sortie de réanimation. Des méthodes spécifiques permettent d'éviter ce biais dans l'estimation des paramètres du modèle à effets mixtes.

D'autre part, analyser les données de survie en incluant la dynamique du biomarqueur en tant que covariable temps-dépendante n'apparaît pas adapté pour trois raisons principales. Premièrement, cette stratégie néglige l'erreur de mesure du biomarqueur. Deuxièmement, elle suppose de façon irréaliste que les valeurs du biomarqueur sont constantes par morceaux entre deux temps de mesures. Finalement, en tant que variable endogène (Rizopoulos, 2011) (c'est à dire que l'existence de la variable est directement liée à l'observation de l'évènement), la trajectoire complète du biomarqueur n'est pas disponible.

La modélisation conjointe des processus longitudinaux et de survie est une façon adaptée de répondre aux limites méthodologiques exposées ci-dessus. Dans le cadre de cette

thèse, nous nous intéressons aux modèles joints à effets aléatoires partagés, qui consistent à modéliser les risques d'évènements comme des fonctions dépendantes de l'évolution individuelle du biomarqueur (Wu et al., 2011) et donc dépendant des effets aléatoires du modèle longitudinal. Il existe aussi des modèles joints à classes latentes (Proust-Lima et al., 2016) qui supposent que des sous groupes d'individus appartenant à une même classe latente partagent une évolution du biomarqueur et des risques d'évènements similaires. Le cadre de ces modèles est semblable à celui des modèles joints à effets aléatoires partagés, à la différence que la distribution des effets aléatoires est discrète et non continue. Les modèles joints à classes latentes ne sont pas traités dans le cadre de cette thèse.

Un modèle conjoint à effets aléatoires partagés dans le cadre de données de survie à deux risques compétitifs décrits par des fonctions de sous-distribution s'écrit de la façon suivante :

$$\begin{aligned}
y_{ij} &= m(t_{ij}, \psi_i) + g(t_{ij}, \psi_i, \sigma) \epsilon_{ij} \\
\lambda_{1i}(t, \psi_i, W_i) &= h_1(t) \exp(\alpha_1 f(t, \psi_i) + \beta_1 \cdot W_i) \\
\lambda_{2i}(t, \psi_i, W_i) &= h_2(t) \exp(\alpha_2 f(t, \psi_i) + \beta_2 \cdot W_i)
\end{aligned} \tag{1.10}$$

où  $f$  décrit l'association entre le processus longitudinal et de survie,  $\alpha_1$  et  $\alpha_2$  mesurant la force de l'association avec l'évènement  $\delta = 1$  et  $\delta = 2$  respectivement. La fonction  $f$  peut prendre différentes formes pour représenter la valeur courante ( $f(t, \psi_i) = m(t, \psi_i)$ ), la pente ( $f(t, \psi_i) = \partial_t m(t, \psi_i)$ ), ou encore l'aire sous la courbe ( $f(t, \psi_i) = \int_0^t m(u, \psi_i) du$ ) du biomarqueur (Desmée et al., 2017; Keriou et al., 2022). Comme dans le cas de données de survie classique exposé en Section 1.1.2, l'incidence cumulée associée à une cause donnée (disons l'évènement d'intérêt) est obtenue par  $F_{1i}(t \mid \psi_i, W_i; \theta) = 1 - \exp\left(-\int_0^t \lambda_{1i}(u \mid \psi_i, W_i; \theta) du\right)$ .

Il arrive fréquemment que dans de nombreux cas, non pas un mais plusieurs biomarqueurs soient suivis de façon longitudinale. Dans ce contexte, la modélisation conjointe peut être facilement étendue au cas multivarié. Supposons  $K$  biomarqueurs disponibles,

le modèle conjoint multivarié s'écrit comme suit :

$$\begin{aligned}
y_{ij1} &= m_1(t_{ij1}, \psi_{i1}) + g_1(t_{ij1}, \psi_{i1}, \sigma_1) \epsilon_{ij} \\
&\dots \\
y_{ijK} &= m_K(t_{ijK}, \psi_{iK}) + g_K(t_{ijK}, \psi_{iK}, \sigma_K) \epsilon_{ij} \\
\lambda_{1i}(t, \psi_i, W_i) &= h_1 \exp \left( \sum_{k=1}^K (\alpha_{1k} f_k(t, \psi_{ik})) + \beta_1 \cdot W_i \right) \\
\lambda_{2i}(t, \psi_i, W_i) &= h_2 \exp \left( \sum_{k=1}^K (\alpha_{2k} f_k(t, \psi_{ik})) + \beta_2 \cdot W_i \right) \tag{1.11}
\end{aligned}$$

où  $y_{ijk}$  représente la  $j^{\text{ème}}$  mesure du biomarqueur  $k$  pour le patient  $i$ . Les  $K$  biomarqueurs sont liés au processus de survie par les coefficients de liens  $\alpha_{1k}$  et  $\alpha_{2k}$ ,  $\forall k \in \{1, \dots, K\}$ .

Lorsque le nombre de modèles longitudinaux est important, des limites statistiques se présentent. En effet, certains auteurs ont décrit des problèmes computationnels et d'identifiabilité dus au nombre élevé d'effets aléatoires lorsque de multiples modèles longitudinaux sont estimés conjointement avec un modèle de survie. Cela explique probablement pourquoi les modèles publiés se limitent souvent à des modèles longitudinaux comportant au maximum deux biomarqueurs (Andrinopoulou et al., 2017; Long and Mills, 2018; Rajeswaran et al., 2018) et souligne la nécessité de disposer de méthodes de sélection dans ce contexte.

Dans un contexte de données en grande dimension où les biomarqueurs correspondent à des niveaux d'expression de gènes, des auteurs (Liu et al., 2019) ont proposé un modèle d'analyse factorielle pour résumer la co-évolution des expressions géniques. Partant d'un modèle avec 55 gènes, les auteurs ont retenu un sous-modèle longitudinal où 3 composantes principales résumaient l'information génique. Cette approche rend l'estimation calculable, mais ne propose pas de sélection à proprement parler car tous les gènes ont contribué, avec différents poids strictement positifs, à construire les composantes principales. Récemment, une autre étude s'est adressée à cette question dans le contexte des patients souffrant de

maladies cardiaques chroniques, où 92 biomarqueurs longitudinaux étaient disponibles (Klimczak-Tomaniak et al., 2022). Les auteurs ont considéré un modèle de Cox avec covariables temps-dépendantes (correspondant au 92 biomarqueurs) et ont appliqué une sélection de type LASSO. Cette procédure leur a permis de sélectionner 9 biomarqueurs pour estimer ensuite un modèle joint multivarié. Parmi ces 9 biomarqueurs, seulement 2 avaient un coefficient de lien significativement associé à leur évènement d'intérêt. La limite de cette méthodologie réside dans le fait que les biomarqueurs ont été sélectionnés à l'aide d'un modèle ne tenant pas compte des censures informatives causées par les évènements, ce qui peut mener à biaiser les résultats et exclure des biomarqueurs potentiellement intéressants.

Alors que les procédures de sélection des variables sont connues pour les modèles de régression standard, il n'existe pas, à notre connaissance, d'applications aux modèles conjoints surtout en présence de risques compétitifs.

#### 1.1.4 Prédiction individuelle dynamique

Ces dernières années, la notion de médecine personnalisée est devenue centrale. Cette dernière a l'objectif de proposer le traitement le plus adapté à chaque patient et ainsi d'améliorer leur prise en charge. Pour répondre à cela, il apparaît primordial de disposer d'outils puissants utilisant non seulement les caractéristiques individuelles du patient, mais aussi toutes les données collectées au cours de son suivi. La modélisation conjointe entre un processus longitudinal et un processus de survie permet d'avoir accès à la fois aux prédictions de l'évolution du biomarqueur et aux prédictions des risques d'évènements. Pour améliorer la précision de la prédiction, ces prédictions peuvent être actualisées à chaque nouvelle observation disponible, et sont ainsi communément appelées les prédictions dynamiques (Rizopoulos, 2011).

Le principe repose sur l'utilisation des données disponibles chez un patient jusqu'à un temps  $l$ , appelé temps landmark, afin de prédire son évolution future du biomarqueur et ses risques d'évènements associés à un temps ultérieur  $l + t$ , où  $t$  est appelé fenêtre de

prédiction. Cette méthode s'applique évidemment sur les patients qui n'ont pas connu d'évènement avant la date  $l$ . De façon plus formelle, supposons un modèle conjoint entre un processus longitudinal et un processus de survie à risques compétitifs, dont le vrai vecteur de paramètres est noté  $\theta$ . Dans le cadre de l'approche sous-distribution, les variables  $T_1 = T \times \mathbb{1}_{\delta=1} + \infty \times \mathbb{1}_{\delta=2}$  et  $T_2 = T \times \mathbb{1}_{\delta=2} + \infty \times \mathbb{1}_{\delta=1}$ , représentant les délais jusqu'à l'évènement d'intérêt et en compétition respectivement, sont introduites. Pour un nouveau patient  $i$  dont les observations longitudinales jusqu'au temps landmark  $l$  sont notées  $Y_i(l)$ , nous cherchons à prédire ses valeurs futures de biomarqueurs, et surtout, son risque d'évènement d'intérêt à un temps  $l+t$ . On s'intéresse donc à la probabilité conditionnelle de faire l'évènement d'intérêt entre  $l$  et  $l+t$  sachant que le patient n'a pas fait d'évènement avant la date  $l$  :

$$\pi_{1i}(l+t) = \mathbb{P}(T_{1i} < l+t \mid T_i > l, Y_i(l), W_i; \theta) \quad (1.12)$$

Par soucis de représentation graphique, la probabilité  $s_{1i}(l+t) = 1 - \pi_{1i}(l+t)$  est souvent préférée et se décompose à l'aide de la formule de Bayes de la manière suivante :

$$\begin{aligned} s_{1i}(l+t) &= \mathbb{P}(T_{1i} > l+t \mid T_i > l, Y_i(l), W_i; \theta) \\ &= \frac{(\mathbb{P}(T_i > l+t, \delta_i = 1) + \mathbb{P}(\delta_i = 2)) \cap \mathbb{P}(T_i > l)}{\mathbb{P}(T_i > l)} \\ &= \frac{(\mathbb{P}(T_i > l+t, \delta_i = 1) + \mathbb{P}(T_i > l, \delta_i = 2))}{\mathbb{P}(T_i > l, \delta_i = 1) + \mathbb{P}(T_i > l, \delta_i = 2)} \end{aligned} \quad (1.13)$$

L'estimation de cette probabilité nécessite de calculer un estimateur  $\hat{\theta}$  de  $\theta$  sur un jeu de données, puis d'estimer la distribution conditionnelle des effets aléatoires  $\eta_i$  d'un nouveau patient  $i$  sans évènement au temps  $l$ , sachant ses données longitudinales  $Y_i(l)$ . Avec les notations simplifiées  $F_{1i}(t) = F_{1i}(t \mid \psi_i, W_i; \theta)$  (idem pour  $F_{2i}(t)$ ),  $s_{1i}$  se calcule comme suit :

$$\hat{s}_{1i}(l+t) = \frac{(F_{1i}(\infty) - F_{1i}(l+t)) + (F_{2i}(\infty) - F_{2i}(l))}{(F_{1i}(\infty) - F_{1i}(l)) + (F_{2i}(\infty) - F_{2i}(l))} \quad (1.14)$$

Si le modèle décrit par des risques de sous-distribution est correctement spécifié,

$F_{1i}(\infty) + F_{2i}(\infty) = 1, \forall i$  et l'expression devient :

$$\hat{s}_{1i}(l+t) = \frac{1 + F_{1i}(l+t) - F_{2i}(l)}{1 - F_{1i}(l) - F_{2i}(l)} \quad (1.15)$$

Or, comme évoqué dans la section 1.1.2, il se peut que l'incidence cumulée toute cause confondue ( $F_1 + F_2$ ) dépasse 1 pour certaines valeurs de covariables, lorsqu'une approche par risques de sous-distribution est considérée. Pour ne pas obtenir de résultats erronés, nous proposons d'utiliser l'équation 1.14 pour le calcul des probabilités individuelles.

Pour évaluer les performances de ces prédictions individuelles, plusieurs indicateurs peuvent être considérés. Les plus utilisés sont l'AUC (Blanche et al., 2013), mesurant la capacité de discrimination du modèle, et le score de Brier (Gerds and Schumacher, 2006), mesurant à la fois la discrimination et la calibration du modèle. La discrimination représente la capacité du modèle à séparer les patients en fonction de leur évènement, et la calibration mesure la précision des prédictions fournies. L'indicateur de l'AUC a initialement été développé pour le cas d'une réponse binaire (Pepe, 2003) (par exemple, évaluer la capacité d'un test à distinguer les patients malades des non malades), puis a été étendu aux données de survie (Heagerty and Zheng, 2005), et aux risques compétitifs (Blanche et al., 2013). Notamment, il permet d'évaluer la capacité d'un biomarqueur  $Bm$ , mesuré à un temps  $t$  bien précis, à distinguer les sujets qui vont expérimenter un évènement à court terme des sujets qui ne l'expérimenteront qu'à long terme ou jamais. Dans ce cadre, il faut définir les concepts de cas et de contrôles (Heagerty and Zheng, 2005). Plusieurs définitions existent, et dans ce travail de thèse, nous considérons qu'un cas est un individu ayant fait l'évènement d'intérêt dans une fenêtre de temps définie, et un contrôle, un individu n'ayant fait aucun des évènements ou ayant fait l'évènement en compétition, dans cette même fenêtre. Cette définition fait référence à l'approche cumulative dynamique définie dans le papier de Heagerty and Zheng (2005). Suivant cette définition, l'AUC est définie de la façon suivante :

$$AUC(t) = \mathbb{P}(Bm_i > Bm_j \mid T_i \leq t, \delta_i = 1, \{T_j > t\} \cup \{T_j \leq t, \delta_j \neq 1\}) \quad (1.16)$$

Dans le cadre de modèles conjoints et pour prendre en compte l'historique du patient jusqu'à un temps landmark  $l$  défini,  $Bm_i$  est remplacé par la probabilité  $\hat{\pi}_{1i}(l+t)$  de faire l'évènement d'intérêt entre  $l$  et  $l+t$ , et l'AUC dépend à la fois du temps landmark  $l$  et de la fenêtre de prédiction  $t$ . Pour faciliter les notations, nous définissons  $D_i(l, t) = \mathbb{1}_{l < T_i \leq l+t, \delta_i=1}$  c'est à dire que  $D_i(l, t) = 1$  si l'individu a fait l'évènement d'intérêt entre  $l$  et  $l+t$ , et  $D_i(l, t) = 0$  si l'individu a fait l'évènement en compétition entre  $l$  et  $l+t$  ou n'a fait aucun des évènements dans cette fenêtre de temps. Ainsi, l'AUC peut être réécrite de la manière suivante :

$$AUC(l, t) = \mathbb{P}(\pi_{1i}(l+t) > \pi_{1j}(l+t) \mid D_i(l, t) = 1, D_j(l, t) = 0, T_i > l, T_j > l) \quad (1.17)$$

L'AUC est généralement comprise entre 0.5 (modèle qui prédit aléatoirement l'ordre entre les prédictions et la réalisation des évènements pour tous les couples de patients de l'échantillon) et 1 (modèle qui prédit parfaitement).

Le score de Brier est quant à lui une forme d'erreur quadratique pour évaluer la distance entre l'observation (par exemple, l'indicatrice d'évènement chez un patient) et la prédiction (probabilité de faire l'évènement donnée par le modèle). Les travaux de Paul Blanche (Blanche et al., 2015) ont défini le score de Brier dynamique adapté aux risques compétitifs. Pour une probabilité  $\pi_1(l+t)$  de faire l'évènement d'intérêt entre  $l$  et  $l+t$ , et  $D(l, t) = \mathbb{1}_{l < T \leq l+t, \delta=1}$  :

$$BS(l, t) = \mathbb{E} [(D(l, t) - \pi_1(l+t))^2 \mid T > l] \quad (1.18)$$

Contrairement à l'AUC, des faibles valeurs de score de Brier sont associées à de bonnes capacités prédictives.

L'estimation de ces indicateurs repose sur la capacité à classer chaque individu en cas ou contrôle. Or, en présence de censure, il n'est plus possible de faire ce classement pour les individus concernés. C'est pourquoi des estimateurs prenant en compte les données censurées ont été proposés, notamment les estimateurs IPCW (Inverse Probability of

Censoring Weighting) décrits dans d'excellents travaux (Blanche et al., 2013).

## 1.2 Inférence dans le cadre fréquentiste

L'inférence des modèles joints peut-être menée dans le cadre fréquentiste ou bayésien. Dans cette thèse, nous évoquerons uniquement le cadre fréquentiste.

### 1.2.1 Estimation des paramètres

Dans le cadre de modèles joints entre un processus longitudinal et un processus de survie, l'objectif est d'estimer le vecteur des paramètres de population  $\theta = \{\theta^{(l)}, \theta^{(s)}\}$  où  $\theta^{(l)}$  réfère aux paramètres du modèle longitudinal, et  $\theta^{(s)}$  aux paramètres du modèle de survie. Nous raisonnons par maximum de vraisemblance. Deux approches principales ont été décrites dans la littérature pour estimer  $\theta$  dans ce cadre. La première repose sur une approche en deux étapes (Tsiatis et al., 1995), où les paramètres  $\theta^{(l)}$  du modèle longitudinal définis dans l'équation 1.1.1 sont d'abord estimés indépendamment du processus de survie. Les modèles à effets mixtes dépendant de paramètres individuels  $\psi = (\psi_i, 1 \leq i \leq N)$  non observés, on cherche ainsi à maximiser la vraisemblance marginale du modèle notée  $l(y; \theta^{(l)})$ , obtenue par intégration sur les paramètres individuels :

$$l(y; \theta^{(l)}) = \prod_{i=1}^N \int_{\mathcal{D}_{\psi_i}} p(y_i | \psi_i; \theta^{(l)}) p(\psi_i; \theta^{(l)}) d\psi_i \quad (1.19)$$

où  $p(y_i | \psi_i; \theta^{(l)})$  représente la densité de probabilité des observations longitudinales conditionnellement aux paramètres individuels, et  $p(\psi_i; \theta^{(l)})$  la densité de probabilité des paramètres individuels. Lorsque le modèle est linéaire et homoscédastique, il existe une forme explicite de l'intégrale, et la vraisemblance peut être maximisée directement. Lorsque le modèle est non-linéaire, il n'existe pas de forme explicite. Des algorithmes ont ainsi été développés pour permettre un calcul approché ou exact de la vraisemblance. Les premiers algorithmes développés proposent d'approcher le modèle structurel sous l'intégrale par des développements de Taylor d'ordre 1, dans le but de le linéariser. Parmi eux, on

trouve la méthode FO (linéarisation autour de l'espérance des effets aléatoires) (Sheiner et al., 1972) et FOCE (linéarisation autour de l'estimation individuelle des effets aléatoires) (Lindstrom and Bates, 1990). Puis des auteurs ont proposé d'approcher l'intégrale par une approximation de Laplace (Wolfinger, 1993). Ces méthodes ne sont pas optimales car elles reposent sur une forme approchée de la vraisemblance et non pas sur sa forme exacte, pouvant ainsi mener à des estimations imprécises.

D'autre part, des méthodes approximant la forme exacte de la vraisemblance peuvent être employées : l'intégration numérique et l'intégration stochastique. La quadrature de Gauss (Davidian and Gallant, 1992) ou plus tard la quadrature de Gauss adaptative (Pinheiro and Bates, 1995) sont généralement utilisées pour approcher numériquement l'intégrale. L'algorithme Stochastic Approximation Expectation Maximization (SAEM) (Delyon et al., 1999) est, quant à lui, le plus utilisé pour approcher l'intégrale de façon stochastique et représente une version stochastique de l'algorithme Expectation Maximization (EM). Dans le cadre de modèles linéaires, les algorithmes itératifs tels que l'algorithme EM fonctionnent en plusieurs étapes. A l'itération  $k$ , on complète les données en tirant des paramètres individuels  $\psi^{(k)}$  dans leur distribution supposée connue. Le couple  $(y, \psi^{(k)})$  caractérise ainsi les données complètes. L'étape E calcule l'espérance conditionnelle de la log vraisemblance complète, notée  $Q_k(\theta^{(l)}) = \mathbb{E}_\psi \left( \log l(y, \psi; \theta^{(l)}) \mid y; \theta_{k-1}^{(l)} \right)$ , et pour finir l'étape M calcule la valeur  $\theta_k^{(l)}$  qui maximise  $Q_k(\theta^{(l)})$ . Dans ce contexte, la vraisemblance complète du modèle  $l(y, \psi; \theta^{(l)})$  peut être décrite par :

$$l(y, \psi; \theta^{(l)}) = \prod_{i=1}^N p(y_i \mid \psi_i; \theta^{(l)}) p(\psi_i; \theta^{(l)}) \quad (1.20)$$

Dans le cadre de modèles non linéaires, l'étape E ne peut pas être explicitement calculée. Une alternative est d'utiliser une version stochastique de l'algorithme EM. Comme l'algorithme EM, il se divise en 3 étapes. A l'itération  $k$  de l'algorithme :

1. *Etape de simulation* : Tirage de paramètres individuels  $\psi^{(k)}$  dans la distribution conditionnelle  $p(\cdot \mid y; \theta_k^{(l)})$ .

2. *Approximation stochastique* : Mise à jour de  $Q_k(\theta^{(l)})$  :

$$Q_k(\theta^{(l)}) = (1 - \gamma_k) Q_{k-1}(\theta^{(l)}) + \gamma_k (\log l(y, \psi^{(k)}; \theta^{(l)})) \quad (1.21)$$

où  $(\gamma_k)$  est une séquence décroissante de nombres positifs avec  $\gamma_1 = 1$ .

3. *Étape de maximisation* : Mise à jour de  $\theta_k^{(l)}$  :

$$\theta_{k+1}^{(l)} = \text{Arg max}_{\theta^{(l)}} Q_k(\theta^{(l)})$$

Quelque soit la méthode utilisée et une fois le maximum de vraisemblance atteint, les paramètres individuels peuvent être estimés par des estimations bayésiennes empiriques (EBEs) et sont notés  $\hat{\psi} = (\hat{\psi}_i, 1 \leq i \leq N)$ . Dans une deuxième étape, ces paramètres individuels sont utilisés pour estimer les fonctions de risque :

$$\begin{aligned} \lambda_{1i}(t \mid \hat{\psi}_i, W_i; \theta^{(s)}) &= h_1(t) \exp \left( \alpha_1 f(t, \hat{\psi}_i) + \beta_1 \cdot W_i \right) \\ \lambda_{2i}(t \mid \hat{\psi}_i, W_i; \theta^{(s)}) &= h_2(t) \exp \left( \alpha_2 f(t, \hat{\psi}_i) + \beta_2 \cdot W_i \right) \end{aligned} \quad (1.22)$$

Les paramètres du modèle de survie  $\theta^{(s)}$  sont à leur tour estimés en maximisant la vraisemblance qui, dans le cadre de l'approche sous-distribution, s'écrit sous la forme suivante :

$$\begin{aligned} l(\tilde{T}, \delta \mid \hat{\psi}; \theta^{(s)}) &= \prod_i \left[ \lambda_{1i}(\tilde{T}_i \mid \hat{\psi}_i, W_i; \theta^{(s)}) \times (1 - F_{1i}(\tilde{T}_i \mid \hat{\psi}_i, W_i; \theta^{(s)})) \right]^{\mathbb{1}_{\delta_i=1}} \\ &\quad \times \left[ \lambda_{2i}(\tilde{T}_i \mid \hat{\psi}_i, W_i; \theta^{(s)}) \times (1 - F_{2i}(\tilde{T}_i \mid \hat{\psi}_i, W_i; \theta^{(s)})) \right]^{\mathbb{1}_{\delta_i=2}} \\ &\quad \times \left[ 1 - F_{1i}(\tilde{T}_i \mid \hat{\psi}_i, W_i; \theta^{(s)}) - F_{2i}(\tilde{T}_i \mid \hat{\psi}_i, W_i; \theta^{(s)}) \right]^{\mathbb{1}_{\delta_i=0}} \end{aligned} \quad (1.23)$$

Des travaux ont montré que cette approche en deux étapes conduit à des estimations biaisées des paramètres (Desmée et al., 2015), et une approche par maximisation de la vraisemblance conjointe est nécessaire pour s'affranchir de ce biais Tsiatis et al. (1995); Rizopoulos (2012); Desmée et al. (2015).

Dans une approche de maximisation de la vraisemblance conjointe, on suppose que

conditionnellement aux paramètres individuels  $\psi$ , les processus longitudinaux et de survie sont indépendants. On cherche ainsi à maximiser la vraisemblance marginale conjointe suivante :

$$l(y, \tilde{T}, \delta; \theta) = \prod_{i=1}^N \int_{\mathcal{D}_{\psi_i}} p(y_i | \psi_i; \theta) p(\tilde{T}_i, \delta_i | \psi_i; \theta) p(\psi_i; \theta) d\psi_i \quad (1.24)$$

où  $p(\tilde{T}_i, \delta_i | \psi_i; \theta)$  représente la densité de probabilité des observations de survie (explicitée dans l'équation 1.23 dans le cadre de l'approche sous-distribution). Cette fonction de vraisemblance n'a une nouvelle fois, pas de forme explicite. Elle peut cependant être approchée numériquement à l'aide de quadratures telles que Gauss (Guedj et al., 2011) ou Laplace (Rizopoulos et al., 2009), puis maximisée à l'aide d'algorithmes d'optimisation de type Newton ou EM (Wulfsohn and Tsiatis, 1997; Rizopoulos, 2012). D'autres alternatives sont les méthodes de Monte Carlo et d'approximations stochastiques, développées pour approximer la vraisemblance en étant intégrées à l'algorithme EM. On trouve notamment l'algorithme MCEM (Wei and Tanner, 1990) et SAEM (Delyon et al., 1999). Les approches numériques sont généralement plus complexes et plus coûteuses en termes de temps de calculs que les approches stochastiques, car elles reposent sur la forme marginale de la vraisemblance impliquant une intégration sur les effets aléatoires. Dans le cadre d'intégration stochastique via l'algorithme SAEM, les 3 étapes décrites précédemment restent les mêmes. Toutefois, la vraisemblance complète comporte un terme additionnel référant aux contributions des données de survie, et est notée  $l(y, \tilde{T}, \delta, \psi; \theta)$ .

## 1.2.2 Estimation des erreurs standard

Dans toute procédure d'estimation, il est important de fournir une incertitude sur l'estimation des paramètres du modèle et donc de fournir des erreurs standard des estimateurs. Dans le cadre d'approches par maximum de vraisemblance, les erreurs standard sont usuellement dérivées de la Matrice d'Information de Fisher (FIM, pour Fisher Information Matrix) observée (Efron and Hinkley, 1978).

Cette dernière peut se définir, par la hessienne (basée sur les dérivées secondes, lors-

qu'elles existent) ou par le produit des scores (basés sur les dérivées premières) de la log vraisemblance :

$$I_{hess}(\theta) = -\partial_{\theta}^2 \log l(y, \tilde{T}, \delta; \theta) \quad (1.25)$$

$$I_{sco}(\theta) = \left( \partial_{\theta} \log l(y, \tilde{T}, \delta; \theta) \right) \left( \partial_{\theta} \log l(y, \tilde{T}, \delta; \theta) \right)^t \quad (1.26)$$

L'utilisation d'algorithmes itératifs complétant les données ont l'avantage de reposer sur la forme complète de la vraisemblance (cf Section 1.2.1), mais ne donnent pas accès directement au calcul de la FIM observée, qui repose sur sa forme marginale. Des travaux ont donc été menés pour fournir une estimation de la FIM observée dans ces approches. En utilisant le principe de décomposition de Louis (Louis, 1982), la FIM observée  $I_{hess}(\theta)$  peut se décomposer en une fonction d'espérances conditionnelles basée sur les dérivées premières et secondes de la log vraisemblance complète :

$$I_{hess}(\theta) = - \left( \mathbb{E}_{\psi|y, \tilde{T}, \delta; \theta} \left[ \partial_{\theta}^2 \log l(y, \tilde{T}, \delta, \psi; \theta) \right] + \text{Var}_{\psi|y, \tilde{T}, \delta; \theta} \left( \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \right) \right) \quad (1.27)$$

où :

$$\begin{aligned} \text{Var}_{\psi|y, \tilde{T}, \delta; \theta} \left( \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \right) = \\ \mathbb{E}_{\psi|y, \tilde{T}, \delta; \theta} \left[ \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \left( \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \right)^t \right] \\ - \mathbb{E}_{\psi|y, \tilde{T}, \delta; \theta} \left[ \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \right] \times \mathbb{E}_{\psi|y, \tilde{T}, \delta; \theta} \left[ \left( \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \right)^t \right] \end{aligned}$$

Des algorithmes stochastiques permettent ensuite d'approcher  $I_{hess}(\theta)$  (Kuhn and La-vielle, 2005; Samson et al., 2007). En particulier, deux quantités  $\Delta_0$  et  $G_0$  sont initiali-sées arbitrairement. Au sein d'une itération  $k$  de l'algorithme SAEM, chaque espérance conditionnelle peut être estimée à l'aide de méthodes de Monte-Carlo qui génèrent des réalisations  $\psi^{(k)}$  des paramètres individuels, sachant le vecteur des données  $(y, \tilde{T}, \delta)$  et le vecteur des paramètres de population  $\theta_k$ . Puis, les quantités  $\Delta_k$  et  $G_k$  sont mises à jour,

pour une itération  $k$  donnée :

$$\begin{aligned}
\Delta_k &= (1 - \gamma_k) \Delta_{k-1} + \gamma_k (\partial_\theta \log l(y, \tilde{T}, \delta, \psi^{(k)}; \theta_k)) \\
G_k &= (1 - \gamma_k) G_{k-1} + \\
&\quad \gamma_k \left( \partial_\theta^2 \log l(y, \tilde{T}, \delta, \psi^{(k)}; \theta_k) + \partial_\theta \log l(y, \tilde{T}, \delta, \psi^{(k)}; \theta_k) \left( \partial_\theta \log l(y, \tilde{T}, \delta, \psi^{(k)}; \theta_k) \right)^t \right)
\end{aligned} \tag{1.28}$$

Lorsque l'algorithme SAEM a convergé vers  $\hat{\theta}$ , le maximum de vraisemblance de  $\theta$  (disons à l'itération  $K$ ), la FIM est estimée par

$$I_{hess}(\hat{\theta}) = - (G_K - \Delta_K (\Delta_K)^T) \tag{1.29}$$

La matrice de variance-covariance  $\hat{\Sigma}$  de  $\hat{\theta}$  est finalement dérivée par :  $\hat{\Sigma}(\hat{\theta}) = I_{hess}(\hat{\theta})^{-1}$ .

Plus tard, des développeurs basés chez Lixoft (<https://lixoft.com/>), ont adapté l'algorithme stochastique décrit ci-dessus pour améliorer notamment la convergence. Entre autre, les algorithmes estimant le vecteur des paramètres  $\theta$  et la FIM observée  $I_{hess}(\theta)$  ne sont plus itérés simultanément. Une fois l'algorithme SAEM ayant convergé vers  $\hat{\theta}$ , l'estimateur du maximum de vraisemblance de  $\theta$ ,  $K'$  réalisations de paramètres individuels  $\psi_i^{(k)}$  sont générées, puis les quantités suivantes sont calculées :

$$\begin{aligned}
D_{K'} &= \frac{1}{K'} \sum_{k=1}^{K'} \partial_\theta^2 \log l(y, \tilde{T}, \delta, \psi^{(k)}; \hat{\theta}) \\
G_{K'} &= \frac{1}{K'} \sum_{k=1}^{K'} \left( \partial_\theta \log l(y, \tilde{T}, \delta, \psi^{(k)}; \hat{\theta}) \right) \left( \partial_\theta \log l(y, \tilde{T}, \delta, \psi^{(k)}; \hat{\theta}) \right)^t \\
\Delta_{K'} &= \frac{1}{K'} \sum_{k=1}^{K'} \partial_\theta \log l(y, \tilde{T}, \delta, \psi^{(k)}; \hat{\theta})
\end{aligned} \tag{1.30}$$

Et la FIM observée est estimée par :

$$I_{hess}(\hat{\theta}) = - (D_{K'} + G_{K'} - \Delta_{K'} (\Delta_{K'})^t) \tag{1.31}$$

Lorsque  $I_{hess}(\hat{\theta})$  ne peut pas être inversée (notamment quand l'estimation est complexe ou que la matrice est mal conditionnée), des réalisations supplémentaires de paramètres individuels peuvent être générées, jusqu'à ce que la matrice soit correctement estimée.

Dans le même esprit que le principe de Louis, des travaux récents (Delattre and Kuhn, 2023) ont proposé une décomposition de  $I_{sco}(\theta)$  à partir de la formule suivante :

$$I_{sco}(\theta) = \mathbb{E}_{\psi|y, \tilde{T}, \delta; \theta} \left[ \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \right] \mathbb{E}_{\psi|y, \tilde{T}, \delta; \theta} \left[ \left( \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \right)^t \right] \quad (1.32)$$

Ces derniers ont également proposé un algorithme stochastique pour estimer  $I_{sco}(\theta)$ . La quantité  $\Delta_0$  est initialisée arbitrairement. Puis, au sein d'une itération  $k$  de l'algorithme SAEM,  $\Delta_k$  est mise à jour comme suit :

$$\Delta_k = (1 - \gamma_k) \Delta_{k-1} + \gamma_k (\partial_{\theta} \log l(y, \tilde{T}, \delta, \psi^{(k)}; \theta_k)) \quad (1.33)$$

Lorsque la convergence de l'algorithme SAEM est assurée, disons à l'itération  $K$  :

$$I_{sco}(\hat{\theta}) = \Delta_K (\Delta_K)^T \quad (1.34)$$

La matrice de variance-covariance de  $\hat{\theta}$  est obtenue une nouvelle fois par inversion de  $I_{sco}(\hat{\theta})$ . Cet estimateur a l'avantage de reposer sur les dérivées premières de la vraisemblance, plus intéressant d'un point de vue computationnel. Cet algorithme a démontré de bonnes performances dans le cadre de modèles à variables latentes (réponse unique). Cependant, il n'a pas encore été évalué aux cas des modèles joints.

### 1.2.3 Logiciels disponibles

Les logiciels pour l'estimation de modèles conjoints sont disponibles depuis de nombreuses années. Dans la suite, nous distinguons deux grandes familles de logiciels : les logiciels dits "généralistes" comme R (R Core Team, 2022), ou SAS (Institute, 1985),

	R			SAS		
	<i>JM</i>	<i>joineR</i>	<i>joineRML</i>	<i>frailtypack</i>	<i>%JM</i>	<i>%JMfit</i>
<b>Sous-modèle longitudinal</b>						
MLEM	•	•	•	•	•	•
MNLEM				•		
Splines	•	•	•	•	•	
Modèles multiples			•			
<b>Sous-modèle de survie</b>						
<b>Risque de base</b>						
Non spécifié (Cox)	•	•	•			
Constant	•			•	•	•
Weibull	•			•		
Splines	•			•	•	•
<b>Extensions</b>						
Risques compétitifs (CS)	•	•	•		•	
Risques compétitifs (SD)						
Évènements récurrents				•		
<b>Estimation</b>						
<b>Approximation des intégrales</b>						
Quadratures de Gauss	•	•		•	•	•
Laplace	•					
Monte Carlo			•			
<b>Algorithme d'optimisation</b>						
De type EM	•		•			
De type Newton	•	•		•	•	•
Hybrides (EM puis Quasi Newton)	•					

TABLE 1.1 – Principaux outils et leurs caractéristiques disponibles sur les logiciels R et SAS pour l’inférence conjointe. Notes : Modèles multiples = multiples sous-modèles longitudinaux autorisés, CS = cause-spécifique, SD = sous-distribution.

et les logiciels utilisés en pharmacométrie comme Monolix (Traynard et al., 2020) ou NONMEM (Bauer, 2019).

Nous restons dans le cadre de l’inférence fréquentiste et des modèles joints à effets aléatoires partagés. Dans le cadre des logiciels généralistes, de nombreux outils sont disponibles pour l’estimation des modèles joints et les principaux sont présentés dans la table 1.1.

La plupart des outils disponibles sur les logiciels généralistes permettent d’estimer des modèles joints considérant uniquement une dynamique linéaire du biomarqueur. Notons cependant que la plupart de ces packages donnent la possibilité d’utiliser des combinaisons

linéaires de splines pour donner plus de flexibilité au modèle. Toutefois, lorsque la dynamique du biomarqueur nécessite d’être modélisée par des modèles mécanistiques non-linéaires ou des systèmes d’équations différentielles non linéaires pour lesquels il n’existe pas de forme explicite des solutions, ces packages ne représentent pas des choix optimaux. Seul le package R *frailtypack* (Rondeau et al., 2019a) permet d’estimer un modèle joint avec un modèle mécanistique non-linéaire (modèle de dynamique tumorale) et un modèle de survie avec évènement terminal ou répété. Cependant, seulement une fonction définissant le modèle structurel longitudinal est à disposition (voir la documentation du package (Rondeau et al., 2019b) pour la description de la fonction), et aucune autre fonction ne peut être considérée. Les logiciels généralistes sont pourtant fournis de nombreux outils permettant l’estimation de MNLEM classiques. Par exemple, le logiciel R est doté des packages *nlme* (Pinheiro et al., 2023), *nlmixr* (Fidler et al., 2022), ou encore *saemix* (Comets et al., 2017). Aucun de ces packages n’a été étendu au cas des réponses multiples ou modèles joints. A noter que le package *saemix* a été étendu en 2022 pour l’estimation des réponses non gaussiennes (binaires, catégorielles, comptage, survie), disponible dans la version 3.0 déposée sur le CRAN (<https://github.com/iame-researchCenter/saemix/blob/7638e1b09ccb01cdff173068e01c266e906f76eb/docsaem.pdf>).

Beaucoup d’outils destinés aux modèles conjoints ont été étendus aux risques compétitifs (les packages R *JM* (Rizopoulos, 2010), *joiner* (Philipson et al., 2018), *joinerML* (Hickey et al., 2018), ou encore la macro SAS *%JM* (Garcia-Hernandez and Rizopoulos, 2018)), mais ces derniers offrent uniquement la possibilité d’utiliser une approche cause-spécifique en dépit de l’approche sous-distribution.

Pour estimer des modèles joints non-linéaires, des logiciels issus de la pharmacométrie représentent une solide alternative. Les deux plus utilisés sont Monolix et NONMEM. Dans ces deux logiciels, l’algorithme SAEM est implémenté et permet l’estimation de modèles joints paramétriques flexibles. En effet, l’utilisateur spécifie le modèle de son choix pour la dynamique du biomarqueur (modèle à effets mixtes, équations différentielles, etc...) et pour le modèle de survie. La plupart des études publiées utilisant ces logiciels

modélisent l'évolution d'un biomarqueur par des modèles non-linéaires à effets mixtes en lien avec l'évolution du risque instantané d'un évènement terminal (Desmée et al., 2015; Zhudenkov et al., 2022; Kerioui et al., 2022). Une étude a utilisé le logiciel NONMEM pour l'estimation d'un modèle conjoint non-linéaire avec plusieurs évènements modélisés par un modèle multi-états, en considérant l'approche cause-spécifique pour modéliser les risques d'évènements (Krishnan et al., 2021). Cependant, aucune évaluation des performances d'estimation n'a été proposée.

## 1.3 Applications aux maladies infectieuses sévères

### 1.3.1 Le sepsis

Depuis 2016 et la conférence de consensus Sepsis-3, le sepsis est défini comme une dysfonction d'organe menaçant le pronostic vital et causé par une réponse inappropriée de l'hôte à une infection (Singer et al., 2016). L'infection causant le sepsis peut être tout aussi bien d'origine bactérienne, virale, fongique que parasitaire. Il touche particulièrement les personnes ayant un système immunitaire affaibli, déficient, ou des personnes déjà fragilisées par une ou plusieurs comorbidités. Les nouveaux-nés et les personnes âgées sont également plus fréquemment touchés. Le sepsis s'accompagne d'une production exacerbée de médiateurs inflammatoires, notamment des cytokines (molécules qui assurent les communications entre les cellules du système immunitaire). Ce phénomène affecte le fonctionnement des organes vitaux de manière aiguë et provoque des lésions dans les tissus, ce qui peut entraîner de graves séquelles fonctionnelles voire le décès. L'Organisation Mondiale de la Santé (OMS) estime à près de 30 millions le nombre de personnes atteintes de sepsis chaque année, avec un taux de mortalité d'environ 25% (World Health Assembly, 2017). Lorsque le sepsis est caractérisé par une défaillance circulatoire et une souffrance cellulaire majeure, on parle de choc septique. Dans ce cas, le taux de mortalité peut s'élever jusqu'à 50% (Cohen et al., 2015; Rudd et al., 2020).

Des critères ont été définis pour caractériser un sepsis. On diagnostique un sepsis en

Système	SCORE SOFA				
	0	1	2	3	4
<b>Respiration</b>					
PaO <sub>2</sub> /FiO <sub>2</sub> (mmHg)	≥ 400	<400	<300	< 200 avec soutien ventilatoire	<100 avec soutien ventilatoire
<b>Coagulation</b>					
Plaquettes (10 <sup>3</sup> /μl)	≥ 150	<150	<100	< 50	<20
<b>Foie</b>					
Bilirubine (μmol/l)	<1,2	[1,2,2[	[2,6[	[6,12[	>12
<b>Cardiovasculaire</b>					
	PAM ≥ 70 mmHg	PAM < 70 mmHg	Dop. < 5 ou dobutamine	Dop. [5,15] ou adr. ou noradr. ≤ 0.1	Dop. >15 ou adr. ou noradr. > 0.1
<b>Système nerveux central</b>					
Echelle de Glasgow	15	13-14	10-12	6-9	<6
<b>Rénal</b>					
Créatinine (μmol/l)	<1.2	[1.2,2.0[	[2.0,3.5[	[3.5,5[	> 5
Diurèse (ml/j)				< 500	< 200

TABLE 1.2 – Composantes du score Sequential Organ Failure Assessment (SOFA). Abréviations : Dop. = Dopamine, adr. = adrénaline, noradr. = noradrénaline.

présence d’une infection et lorsque le score SOFA (Sepsis Organ Function Assessment) est au moins égal à 2. Ce score est utilisé en soins intensifs pour déterminer et suivre l’état d’un patient en défaillance d’organe(s) (Vincent et al., 1996; Timsit et al., 2002). Il s’appuie sur six critères, correspondant au système respiratoire, neurologique, cardiovasculaire, hépatique, rénal et à la coagulation. Il varie de 0 à 24, 24 étant le scénario le plus défavorable pour le patient. La table 1.2 détaille le calcul de ce score.

Le score SOFA est devenu un outil prédictif de la mortalité en soins intensifs. Les premières études ont montré le caractère pronostique du score SOFA mesuré dans les 24 premières heures d’hospitalisation, du score SOFA moyen ou le plus haut au cours du suivi en réanimation (Vincent et al., 1998; Ferreira et al., 2001; Timsit et al., 2002). Certains travaux ont également évalué la différence de SOFA entre deux temps distincts (par exemple, entre  $J_0$  et  $J_3$ ) (Karakike et al., 2019), ce qui constitue un embryon de la modélisation dynamique. Pour prendre en compte correctement l’ensemble de la dynamique du score SOFA, des études plus récentes ont considéré son suivi longitudinal par le biais d’une modélisation conjointe, et ont démontré le caractère pronostique de l’évolution du SOFA sur le risque de décès en soins intensifs (Deslandes and Chevret, 2010; Musoro

et al., 2018).

Le premier travail appliqué de cette thèse s’inscrit dans cette lignée. Nous souhaitons utiliser l’information disponible dans le suivi du score SOFA pour construire un outil prédictif du pronostic du patient. Pour cela, nous travaillons sur l’entrepôt de données OUTCOMEREA (<https://outcomerea.fr/>) qui regroupe des patients admis en réanimation dans des centres Français depuis Janvier 1997. Le recueil de données continue encore aujourd’hui. Ce registre est conforme à la législation française relative aux recherches biomédicales. Les patients ou leur famille ont donné leur autorisation pour la collecte, la conservation et l’utilisation de leurs données personnelles anonymisées. L’investigateur principal de ce registre est le Pr Jean-François Timsit, chef de service de réanimation médicale à l’Hôpital Bichat AP-HP, Paris.

Nous nous sommes intéressés à un sous-ensemble de cette base de données qui regroupe les patients majeurs admis en réanimation pour sepsis ou choc septique de 1997 à 2015. Les ré-admissions, admissions pour soins palliatifs, morts cérébrales, ou les décisions précoces de renoncer à des interventions de réanimation pour le maintien en vie ont été exclus de l’analyse. Au total,  $N = 6046$  patients ont été inclus et aléatoirement répartis en un échantillon d’apprentissage ( $N_{app} = 4050$ ) et de validation ( $N_{val} = 1996$ ). Pour chacun des patients, un suivi journalier du score SOFA jusqu’au temps de décès ou de sortie de réanimation est disponible. Le pronostic (décédé/sorti de réanimation) est également rapporté. De nombreuses variables caractéristiques des patients sont à disposition, comme l’âge, le sexe, les comorbidités (diabète, insuffisance cardiaque chronique, maladies rénales chroniques, cirrhose hépatique, maladies pulmonaires obstructives chroniques et immunodéficience), etc. La table 1.3 rassemble les principales caractéristiques des patients.

La Figure 1.1 A montre l’évolution du score SOFA de 100 patients de la base d’apprentissage, tirés aléatoirement. Les patients décédés (courbes rouges) ont généralement connu des trajectoires montantes et élevées de SOFA. A l’inverse, les patients sortis de réanimation (courbes bleues) ont plutôt des trajectoires descendantes et avec de faibles

Variable	Données d'apprentissage	Données de validation
	$N_{app}=4050$	$N_{val}=1996$
<b>Age - no. (%)</b>		
< 60	1565 (39%)	782 (39%)
[60,75[	1333 (33%)	682 (34%)
$\geq 75$	1152 (28%)	532 (27%)
<b>Sexe - no. (%)</b>		
Homme	2491 (61%)	1272 (64%)
Femme	1559 (39%)	726 (36%)
<b>Comorbidités</b>		
0	1988 (49%)	985 (49%)
1	1340 (33%)	682 (34%)
>1	722 (18%)	329 (17%)
<b>SOFA à l'admission - med [q1,q3]</b>	6 [4,9]	6 [4,9]
<b>Mortalité à 30 jours - no. (%)</b>	781 (19%)	362 (18%)
<b>Sorties à 30 jours - no. (%)</b>	2921 (72%)	1426 (71%)

TABLE 1.3 – Principales caractéristiques des patients à l'admission en réanimation

valeurs de SOFA. Ces trajectoires ne sont pas observées en totalité car elles sont censurées par le processus de survie. L'intérêt de la modélisation conjointe apparaît donc clairement.

### 1.3.2 La COVID-19

La COVID-19 est une maladie causée par un virus nommé SARS-CoV-2. Il a été identifié pour la première fois en Chine, en novembre 2019 et a rapidement infecté des millions de personnes menant ainsi à une pandémie mondiale. Aujourd'hui en France, la pandémie s'est traduite par des vagues épidémiques successives avec des chiffres d'incidence et de mortalité très variables. Par exemple, lors de la première vague (printemps 2020), la Direction de la Recherche, des Études, de l'Évaluation et des Statistiques (DREES) rapporte un taux de mortalité hospitalier de 19% avec une forte hétérogénéité en fonction de l'âge (Clanché, 2022). Dans les récentes vagues de l'épidémie, il est difficile de donner un taux de mortalité intra-hospitalier mais ce dernier a considérablement chuté (Clanché, 2022). Ceci peut s'expliquer notamment par de meilleures connaissances de la maladie, une prise en charge médicale optimisée, la stratégie de vaccination mise en place, et l'arrivée de

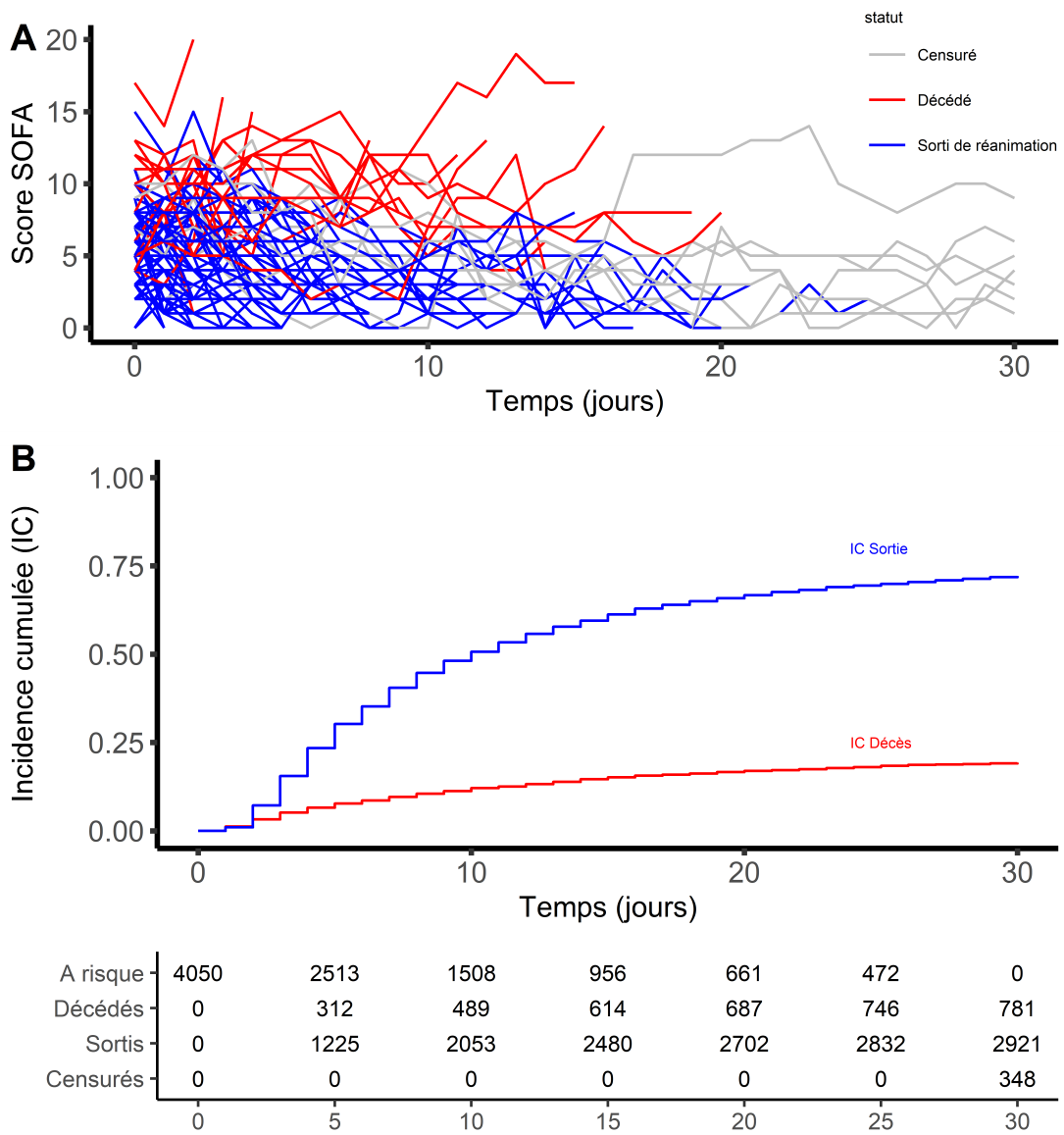


FIGURE 1.1 – Évolution du score SOFA chez 100 patients tirés aléatoirement parmi les patients du jeu de données d'apprentissage (A) et fonctions d'incidences cumulées du décès et de la sortie de réanimation calculées sur les données d'apprentissage (B)

nouveaux variants moins agressifs.

Les symptômes de la COVID-19 sont assez similaires à ceux d'autres infections respiratoires aiguës telles que la grippe saisonnière. On recense notamment des symptômes généraux : la fatigue, la fièvre, la toux et des symptômes liés à l'atteinte pulmonaire notamment la dyspnée, une sensation d'étouffement. La plupart des personnes en bonne santé présentent des symptômes légers à modérés et ne nécessitent pas de prise en charge spécifique pour éliminer le virus. Toutefois, certaines personnes notamment les personnes fragiles, immunodéprimées, âgées ou présentant certaines comorbidités ont plus de risques de présenter une forme sévère de la maladie. Dans les cas extrêmes, le virus est responsable d'un syndrome de détresse respiratoire aiguë qui peut être fatal pour l'individu infecté.

Les systèmes hospitaliers dans le monde entier ont été mis à rude épreuve, du fait de la propagation très rapide et alarmante de la pandémie. Beaucoup de patients dont le pronostic vital était engagé ont dû être pris en charge dans des hôpitaux provisoires, souvent par un personnel moins expérimenté. Dans ce genre de situation, il apparaît important de disposer de systèmes de notation clinique relativement simples pour identifier les patients les plus à risques et les diriger vers la structure et la prise en charge adaptée le plus rapidement possible. C'est dans ce cadre que dès la première vague de l'épidémie en 2020, des scores cliniques ont été développés pour prédire le pronostic des patients hospitalisés pour infection au SARS-CoV-2. Les développements de scores ont été très nombreux, dans des populations et contextes cliniques très variés (Häger et al., 2022), mais reposent généralement sur des informations relatives au patient recueillies à l'admission à l'hôpital (par exemple, des données démographiques, cliniques et des résultats de laboratoire) pour estimer la probabilité de développer un événement clinique (par exemple, le passage en unités de soins intensifs, le décès, etc) (Knight et al., 2020; Shang et al., 2020; Armiñanzas et al., 2021).

Plusieurs revues systématiques ont été publiées pour valider et comparer des scores prédictifs (Lombardi et al., 2021; Vedovati et al., 2022; de Jong et al., 2022; Hassan et al., 2022). Le score 4C apparaît avoir les meilleures performances prédictives du décès intra-

hospitalier dans les 28 jours suivant l'hospitalisation dans la plupart des études. Le détail de son calcul est donné dans la table 1.4 ci-dessous.

<b>Age (années)</b>	<b>&lt;50</b>	0
	<b>50-59</b>	+2
	<b>60-69</b>	+4
	<b>70-79</b>	+6
	<b>≥ 80</b>	+7
<b>Sexe à la naissance</b>	<b>Female</b>	0
	<b>Male</b>	+1
<b>Nombre de comorbidités</b>	<b>0</b>	0
Les comorbidités incluent : maladies cardiaques chroniques, maladies respiratoires chroniques (sauf l'asthme), maladies rénales chroniques, maladies hépatiques moyennes ou sévères, démence, maladies neurologiques chroniques, maladies des tissus conjonctifs, diabètes, VIH ou SIDA et tumeurs	<b>1</b>	+1
	<b>≥ 2</b>	+2
<b>Fréquence respiratoire, cycles/min</b>	<b>&lt;20</b>	0
	<b>20-29</b>	+1
	<b>≥30</b>	+2
<b>Saturation en oxygène</b>	<b>≥92%</b>	0
	<b>&lt;92%</b>	+2
<b>Échelle de coma de Glasgow</b>	<b>15</b>	0
	<b>&lt;15</b>	+2
<b>Urée à l'admission (mmol/L)</b>	<b>&lt;7</b>	0
	<b>7-14</b>	+1
	<b>&gt;7</b>	+2
<b>Protéine C-réactive à l'admission (mg/L)</b>	<b>&lt;50</b>	0
	<b>50-100</b>	+1
	<b>&gt;100</b>	+2

TABLE 1.4 – Système de notation du score 4C. Les poids associés à chaque modalité des 8 composantes sont donnés en gris.

Les scores pronostiques développés ont permis de fournir des outils prédictifs faciles d'utilisation, mais leur principale limite réside dans le fait qu'ils ne tiennent pas compte du côté évolutif et dynamique de la maladie. Or, l'évolution clinique au jour le jour des patients atteints de COVID-19 est un élément essentiel du suivi permettant d'identifier des signes précoces d'aggravation et d'avoir une idée précise de l'évolution de la maladie.

Une étude (Tong-Minh et al., 2022) a considéré le suivi longitudinal de certains biomarqueurs immunologiques pour prédire le risque de décès des patients admis en soins intensifs pour COVID-19 sévère par le biais d'une modélisation conjointe. Cette étude a montré des résultats prometteurs en soulignant que la dynamique de la procalcitonine (PCT), de l'interleukine 6 (IL-6) et du récepteur soluble de l'activateur plasminogène d'urokinase (suPAR) étaient significativement associés au risque de décès en réanimation. Les auteurs ont cependant modélisé le décès en réanimation comme unique évènement terminal, en censurant les patients sortis de soins intensifs. Cette approche va à l'encontre de l'hypothèse de censure non-informative et peut amener à des biais. De plus, ces derniers n'ont évalué que 4 marqueurs immunologiques. De nombreux autres marqueurs restent encore à évaluer.

Le deuxième travail de cette thèse s'articule autour de cette problématique. Nous souhaitons évaluer le suivi de multiples biomarqueurs longitudinaux pour prédire le pronostic des patients hospitalisés pour infection au SARS-CoV-2 dans les 30 jours après l'admission à l'hôpital. Pour se faire, nous travaillons sur une cohorte rétrospective nommée RisCoV dont l'investigateur principal est le Pr François-Xavier Lescure, du Service des Maladies Infectieuses et Tropicales (SMIT) de l'Hôpital Bichat, AP-HP, Paris. Au total, 327 patients hospitalisés de janvier à juillet 2020 dans le SMIT de l'Hôpital Bichat AP-HP à Paris sont inclus dans l'analyse. Une extraction de l'entrepôt de données de l'AP-HP a été réalisée par les internes du service pour regrouper l'ensemble des résultats des analyses biologiques prescrites par les praticiens. Ainsi, 59 biomarqueurs ont été suivis de façon longitudinale et sont présentés dans la table 1.5. Une collecte manuelle a également permis d'obtenir des informations démographiques et cliniques relatives aux patients. Le score 4C à l'admission à l'hôpital du patient a été reconstruit grâce aux données recueillies.

La table 1.6 présente les principales caractéristiques des patients, et la Figure 1.2 les courbes d'incidences cumulées du décès intra-hospitalier et de la sortie de l'hôpital calculées sur l'ensemble des patients du jeu de données. Trente jours après l'admission, environ 14% des patients sont décédés et 73% sortis d'hospitalisation.

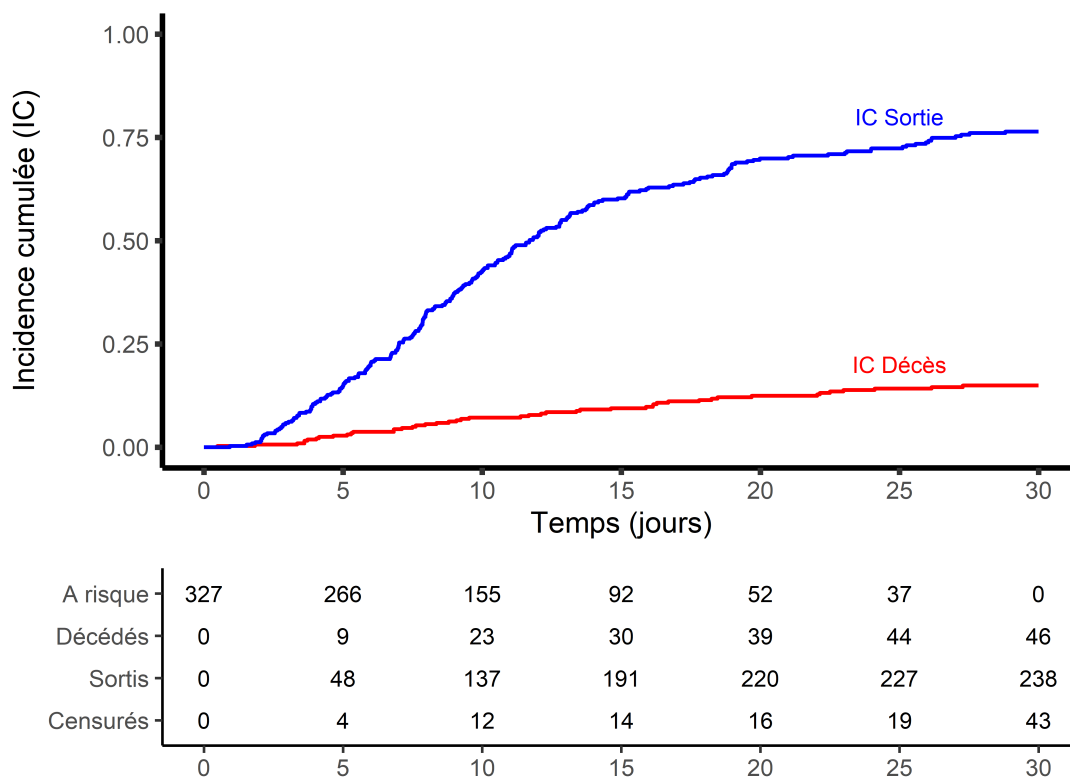


FIGURE 1.2 – Fonctions d’incidences cumulées du décès et de la sortie de réanimation calculées sur l’ensemble des patients de la cohorte RisCoV

Biomarqueur	Unité	N*	n **	Biomarqueur	Unité	N*	n **
<b>Numération formule sanguine</b>				<b>Fonctions rénales / Lyse cellulaire</b>			
Erythrocytes	$\times 10^{12}/L$	326	6.6	Natrémie	mmol/L	327	7.1
Volume globulaire moyen (VGM)	fL	326	6.6	Kaliémie	mmol/L	325	7.1
Hémoglobine	g/dL	326	6.6	Chlorémie	mmol/L	300	5.7
Hématocrite	%	326	6.6	Calcémie	mmol/L	315	5.3
Réticulocytes	$\times 10^9/L$	170	4.4	Phosphates	mmol/L	281	4.9
Leucocytes	$\times 10^9/L$	326	6.6	Magnésium	mmol/L	260	5.2
Polynucléaires basophiles	$\times 10^9/L$	322	6.4	Trou anionique	mmol/L	280	4.8
Polynucléaires éosinophiles	$\times 10^9/L$	322	6.5	Bicarbonates	mmol/L	295	5.7
Polynucléaires neutrophiles	$\times 10^9/L$	326	6.5	Urémie	mmol/L	325	6.9
Taux de granulocytes immatures	%	139	3.7	Protidémie	g/L	290	5.6
Lymphocytes	$\times 10^9/L$	326	6.5	Créatininémie	$\mu\text{mol/L}$	327	6.5
Monocytes	$\times 10^9/L$	326	6.5	Lactate deshydrogénase (LDH)	U/L	297	4.2
Plaquettes	$\times 10^9/L$	326	6.6	Créatine phosphokinase (CPK)	U/L	309	4.3
<b>Coagulation</b>				Phosphatases alcalines	U/L	313	5.3
Taux de prothrombine (TP)	%	303	4.9	<b>Fonctions hépatiques/pancréatiques</b>			
Temps de céphaline activée (TCA)	IU	297	4.6	Alanine aminotransférase (ALAT)	U/L	321	5.6
Factor II	%	153	4.5	Aspartate aminotransférase (ASAT)	U/L	321	5.6
Factor V	%	155	4.5	Gamma GT	U/L	314	5.1
Fibrinogène	g/L	294	4.4	Bilirubine totale	$\mu\text{mol/L}$	322	5.5
Fibrines monomères	$\mu\text{g/mL}$	112	5.3	Lipasémie	U/L	158	4.3
D-Dimères	ng/mL	218	3.9	Albuminémie	g/L	295	3.8
<b>Fonctions pulmonaires</b>				<b>Marqueurs de l'inflammation</b>			
pH artériel	IU	246	8.3	Protéine C-réactive (CRP)	log(mg/L)	318	5.5
pO2a	mm Hg	246	8.3	Procalcitonine (PCT)	$\mu\text{g/L}$	229	4
pCO2a	mm Hg	246	8.3	Orosomucoïde	g/L	101	4.2
Lactate artériel	mmol/L	231	8.1	Ferritine	$\mu\text{g/L}$	214	3.3
Taux d'oxyhémoglobine	%	245	8	Haptoglobine	g/L	116	4.7
<b>Échantillons urinaires</b>				<b>Marqueurs cardiaques</b>			
Natriurèse	mmol/L	132	3	Troponine Ic ultra-sensible	$\mu\text{g/L}$	253	3.6
Kaliurèse	mmol/L	132	3	NT-proBNP	ng/L	262	3.5
Chlorure urinaire	mmol/L	62	3.6				
Protéinurie	g/L	109	2.9				
Urée urinaire	mmol/L	119	3				
Créatininurie	mmol/L	134	3				
Glycosurie	mmol/L	76	3.2				

\*Nombre de patients avec au moins une observation entre l'admission (J0) et J30

\*\*Nombre moyen d'observations pour les patients ayant au moins une observation entre J0 et J30

TABLE 1.5 – Biomarqueurs considérés dans l'analyse, classifiés selon 8 catégories. Pour chaque biomarqueur, le nombre de patients avec au moins une observation entre l'admission (J0) et J30 ainsi que le nombre moyen d'observations sont donnés.

## 1.4 Objectifs de la thèse

### 1.4.1 Modélisation conjointe univariée

Le premier objectif de cette thèse est d'étudier la modélisation conjointe de données longitudinales et de survie dans le cas où un biomarqueur est modélisé par un modèle non-linéaire à effets mixtes (MNLEM), et où plusieurs événements sont en compétition et modélisés par une approche par risques de sous-distribution. Pour développer un tel mo-

<b>Nombres de patients</b>	<b>327</b>
<b>Age - no. (%)</b>	
<50	73 (22)
[50-59[	74 (23)
[60-69[	71 (22)
[70-79[	62 (19)
$\geq 80$	47 (14)
<b>Sexe - Homme - no. (%)</b>	<b>198 (61)</b>
<b>Comorbidités - no. (%)</b>	
0	139 (43)
1	93 (28)
$\geq 2$	95 (29)
<b>Mortalité à 30 jours - no. (%)</b>	<b>46 (14%)</b>
<b>Sorties à 30 jours - no. (%)</b>	<b>238 (73%)</b>

TABLE 1.6 – Principales caractéristiques des 327 patients recueillies à l’admission

dèle, nous nous placerons dans le contexte des patients admis en réanimation pour sepsis issus de l’entrepôt de données OUTCOMEREA (cf Section 1.3.1), avec un suivi journalier du score SOFA. Nous développerons un modèle conjoint non-linéaire avec risques compétitifs pour prédire le risque de décès en réanimation des patients, en considérant la sortie de réanimation comme évènement en compétition. Les paramètres seront estimés à l’aide de l’algorithme SAEM disponible sur le logiciel Monolix. Nous proposerons aussi des prédictions dynamiques à partir de ce modèle, et comparerons leurs performances prédictives par rapport à celles d’un modèle qui n’utilise que l’information disponible à l’admission du patient. Pour valider une telle approche, nous proposerons une étude de simulation dans le but d’évaluer les capacités de l’algorithme SAEM de Monolix à fournir des estimations précises et non biaisées des paramètres du modèle. Le Chapitre 2 fait référence à ces premiers travaux.

## 1.4.2 Modélisation conjointe multivariée

Dans un second temps, nous étendrons la modélisation précédente aux cas où plusieurs processus longitudinaux sont modélisés. Nous travaillerons sur les patients hospitalisés pour infection au SARS-CoV-2 de la cohorte RisCoV présentée en Section 1.3.2, avec des mesures de plusieurs biomarqueurs longitudinaux. L'objectif sera de construire un modèle joint multivarié, en proposant une stratégie de sélection des biomarqueurs, pour prédire le risque de décès intra-hospitalier des patients en considérant la sortie de l'hôpital comme évènement en compétition. Monolix sera utilisé une nouvelle fois pour l'estimation des paramètres du modèle. Nous évaluerons également les bénéfices d'une telle modélisation par rapport à une modélisation qui ne tient compte que de l'information disponible à l'admission, ceci à l'aide de l'AUC temps-dépendante. Dans le but d'évaluer les performances d'estimation des paramètres, la stratégie de sélection proposée, et les propriétés de l'indicateur de l'AUC, nous proposerons une étude de simulation. Le Chapitre 3 fait référence à ces travaux.

## 1.4.3 Dissémination

Dans le troisième travail de thèse, nous souhaitons étendre l'utilisation des modèles conjoints non-linéaires aux logiciels généralistes, où très peu de développements ont été implémentés. Pour se faire, nous étendrons le code du package *saemix* (version 3.1 disponible sur le CRAN) initialement dédié à l'estimation de MNLEM au cas des réponses multiples et modèles joints. Nous implémenterons également l'algorithme stochastique récemment développé par Delattre and Kuhn (2023) pour le calcul des erreurs standard, possédant des propriétés computationnelles intéressantes. Pour l'évaluation de ces développements, nous considérerons quatre modèles joints différents et proposerons une étude de simulation pour valider (i) l'estimation des paramètres, (ii) l'estimation des erreurs standard, (iii) l'erreur de type-I lors du test d'indépendance entre les deux sous-modèles. Ce travail est développé dans le Chapitre 4.



---

Modélisation conjointe avec risques compétitifs : application à des prédictions de survie de patients admis en réanimation pour sepsis, avec un suivi journalier du score SOFA

---

### 2.1 Résumé

Dans ce travail, nous avons proposé un cadre pour l'estimation d'un modèle conjoint composé d'un modèle non-linéaire à effets mixtes pour décrire l'évolution d'un biomarqueur et d'un modèle de survie à risques compétitifs pour décrire les risques instantanés d'évènements par des fonctions de sous-distribution. Pour cela, nous nous sommes placés dans le contexte des patients admis en réanimation pour sepsis, issus de la base de données OUTCOMEREA, dont le score SOFA était évalué chaque jour jusqu'au décès ou à la sortie de réanimation.

Comme évoqué en Section 1.3.1, les données ont été aléatoirement réparties en un échantillon d'apprentissage et en un échantillon de validation. La construction du modèle s'est réalisée sur les données d'apprentissage. Pour le sous-modèle longitudinal, un

modèle non-linéaire à effets mixtes reposant sur une fonction composée d'une double exponentielle a été considérée pour modéliser les dynamiques individuelles du score SOFA. Cette fonction autorise aussi bien les évolutions monotones que non-monotones et est composée de 5 paramètres comprenant tous des effets aléatoires pour prendre en compte la variabilité inter-individuelle. Le modèle de survie à risques compétitifs s'inscrit dans l'approche sous-distribution en vertu de ses propriétés intéressantes pour la prédiction. En particulier il est composé de deux fonctions de sous-distribution décrivant les risques instantanés de décès et de sortie de réanimation, impliquant un risque de base constant, la valeur courante du SOFA et l'âge du patient. L'estimation est réalisée par maximisation de la vraisemblance via l'algorithme SAEM implémenté dans le logiciel Monolix (version 2018R2). Similairement aux études publiées (Deslandes and Chevret, 2010; Musoro et al., 2018), nous avons démontré qu'une hausse du score SOFA est associée à un risque de décès accru et à un risque de sortie diminué. Pour chaque patient de l'échantillon de validation, des prédictions dynamiques de l'évolution du score SOFA et du risque de décès en réanimation ont été calculées, en considérant plusieurs temps landmark (J0, J3, J6, J9) et plusieurs horizons (jusqu'à J30). Les performances ont été évaluées grâce à l'AUC et au score de Brier temps-dépendants, et comparées à celles d'un modèle qui n'utilise que le score SOFA disponible à l'admission (appelé le modèle baseline). La modélisation conjointe a montré sa supériorité en termes de performances prédictives lorsque la quantité d'information disponible est suffisante (par exemple pour un temps landmark = J6 et un horizon = J30, AUC [IC 95%] = 0.63 [0.58,0.68] et 0.77 [0.74,0.80] pour le modèle baseline et joint respectivement,  $p < 10^{-8}$ ).

Pour évaluer les performances de l'estimation, une étude de simulation a été menée. Pour cela, 200 jeux de données ont été simulés à partir du modèle joint proposé ci-dessus, en considérant 1000 patients avec une mesure du score SOFA disponible chaque jour jusqu'au décès ou à la sortie de réanimation. Les valeurs utilisées pour simuler les données longitudinales et de survie ont été choisies en s'inspirant des estimations trouvées sur le vrai jeu de données OUTCOMEREA. Pour chaque jeu de données simulé, les

performances d'estimation sont évaluées en calculant les biais relatifs (RB), les erreurs quadratiques moyennes relatives (RRMSE) ainsi que les erreurs d'estimation relatives (REE) des paramètres du modèle. Les simulations ont montré que l'algorithme SAEM était capable de fournir des estimations non biaisées et précises des paramètres d'un tel modèle. Le coefficient de lien, qui représente un intérêt particulier dans le cadre de la modélisation conjointe, est estimé avec un biais relatif autour de 2% et une RRMSE d'environ 9%, ce qui souligne de très bonnes performances.

Ce travail a permis de démontrer que le logiciel Monolix, et notamment son algorithme SAEM, possède de bonnes propriétés pour l'estimation de modèles joints non-linéaires avec risques compétitifs décrits par une approche sous-distribution. Le modèle proposé peut être employé ou adapté dans d'autres applications où des données longitudinales et de survie sont récoltées simultanément, grâce à la souplesse du logiciel Monolix. Ce travail pourra donc ouvrir le champs d'utilisation des modèles conjoints non-linéaires lorsque plusieurs évènements sont en compétition.

Ces premier travaux ont fait l'objet d'un article publié dans *CPT : Pharmacometrics & Systems Pharmacology* le 6 octobre 2022.

## 2.2 Article 1 (publié)



## ARTICLE

# Joint modeling under competing risks: Application to survival prediction in patients admitted in Intensive Care Unit for sepsis with daily Sequential Organ Failure Assessment score assessments

Alexandra Lavalley-Morelle<sup>1</sup> | Jean-François Timsit<sup>1,2</sup> | France Mentré<sup>1,3</sup>  | Jimmy Mullaert<sup>1,3</sup> | The OUTCOMEREA network

<sup>1</sup>Université Paris Cité, IAME, INSERM, Paris, France

<sup>2</sup>Service de Réanimation Médicale et Infectieuse, AP-HP, Hôpital Bichat, Paris, France

<sup>3</sup>Département Epidémiologie Biostatistiques et Recherche Clinique, AP-HP, Hôpital Bichat, Paris, France

## Correspondence

Alexandra Lavalley-Morelle, Inserm UMR 1137 IAME, UFR de Médecine Bichat, 16, rue Henri Huchard, 75018 Paris, France.

Email: [alexandra.lavalley-morelle@inserm.fr](mailto:alexandra.lavalley-morelle@inserm.fr)

## Funding information

Innovative Medicines Initiative Joint Undertaking, Grant/Award Number: 115737-2

## Abstract

Joint models of longitudinal process and time-to-event data have recently gained attention, notably to provide individualized dynamic predictions. In the presence of competing risks, models published mostly involve cause-specific hazard functions jointly estimated with a linear or generalized linear model. Here we propose to extend the modeling to full parametric joint estimation of a nonlinear mixed-effects model and a subdistribution hazard model. We apply this approach on 6046 patients admitted in intensive care unit (ICU) for sepsis with daily Sequential Organ Failure Assessment (SOFA) score measurements. The joint model is built on a randomly selected training set of two thirds of patients and links the current predicted SOFA measurement to the instantaneous risks of ICU death and discharge from ICU, both adjusted on the patient age. Stochastic Approximation Expectation Maximization algorithm in Monolix is used for estimation. SOFA evolution is significantly associated with both risks: 0.37, 95% confidence interval (CI) = [0.35, 0.39] for the risk of death and  $-0.38$ , 95% CI =  $[-0.39, -0.36]$  for the risk of discharge. A simulation study, inspired from the real data, shows the good estimation properties of the parameters. We assess on the validation set the added value of modeling the longitudinal SOFA follow-up for the prediction of death compared with a model that includes only SOFA at baseline. Time-dependent receiver operating characteristic area under the curve and Brier scores show that when enough longitudinal individual information is available, joint modeling provides better predictions. The methodology can easily be applied to other clinical applications because of the general form of the model.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *CPT: Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics.

## Study Highlights

### WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Joint modeling under competing risk setting has been rarely described in the literature, especially when dealing with nonlinear mixed-effects models (NLMEMs). Most published models considered a linear biomarker evolution, possibly with spline functions to deal with nonlinearity. One considers an NLMEM model jointly estimated with a cause-specific multistate model using NONMEM software.

### WHAT QUESTION DID THIS STUDY ADDRESS?

We propose an approach to model joint estimation of NLMEM and competing risks using a parametric subdistribution hazard model. We apply this new approach on a large data set of patients admitted in intensive care unit (ICU) for sepsis, with daily Sequential Organ Failure Assessment (SOFA) score assessments, to predict the risk of death with the competing risk being discharge from ICU.

### WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

The Stochastic Approximation Expectation Maximization algorithm can be used to provide unbiased estimates of such model parameters. In the real case application, we showed that the joint modeling improves the quality of predictions when sufficient SOFA follow-up is considered compared with a model that only uses baseline information.

### HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

The methodology can be easily extended to other clinical applications of individualized predictions, which is useful because a competing risk setting arises in many biomedical applications.

## INTRODUCTION

A major ambition of precision medicine is to provide every patient with a personalized and adapted medical treatment. Individual dynamic predictions of a clinical outcome based on discrete longitudinal observations of a biomarker fully enter the scope of precision medicine. To get individual predictions that can be updated as more observations become available, a natural approach is to use the fit of a joint model combining (i) the longitudinal model for the evolution of the biomarker and (ii) the time-to-event process describing the risk of the clinical outcome in consideration.

This joint modeling approach has recently gained attention,<sup>1-4</sup> mainly for two reasons. First, a much simpler strategy where biomarker observations are included as a time-varying covariate in a survival model imposes to make a hypothesis about the marker evolution between measurements, neglects measurement errors, and is known to be incorrect because of the endogenous nature of the covariate (i.e., whose existence is directly related to failure status).<sup>5,6</sup> Second, one could consider fitting a first model for the biomarker evolution and then plugging the predicted evolution as a time-dependent covariate in the survival model, resulting in a so-called two-stage

procedure. Unfortunately, the dropout process induced by the survival model, that is, the onset of an event that precludes any further biomarker measurements, generates missing data for the biomarker that can be not at random (missing not at random [MNAR]). This leads to possible bias in the estimations as described in a published study.<sup>7</sup> The joint modeling approach genuinely overcomes these two methodological issues. A recent publication<sup>8</sup> proposes a workflow for a joint model of longitudinal and time-to-event data.

Competing risks are frequent in biomedical applications, where patients are at risk of multiple and mutually exclusive events (e.g., death and relapse in oncology or death and discharge in a hospital ward). When dealing with competing events, two distinct risks can be modeled: the cause-specific hazard (CSH)<sup>9</sup> and the subdistribution hazard (SDH).<sup>10</sup> Although both approaches are essentially two different parametrizations of the same random process, some authors argue that the CSH parameters are more suitable for etiologic studies and SDH parameters for risk prediction.<sup>11,12</sup> We also emphasize that the hypothesis of proportionality of risks on CSH and SDH are generally not compatible.<sup>13</sup> Although the cause-specific approach was the most often used in the context of joint modeling,<sup>3,14-16</sup> we decided here to analyze the data assuming

that the hypothesis of proportional hazards holds on the SDHs.

The joint estimation with competing risks complicates the analysis, especially when the biomarker evolution is associated with the risk of both events. In the literature, linear<sup>17–19</sup> and nonlinear<sup>1,4,20</sup> joint models have been extensively described. However, when dealing with competing risks, linear mixed-effects submodels, possibly involving a combination of spline functions to deal with nonlinearity (using R-INLA<sup>21</sup> or JMBayes<sup>15</sup>) are predominant. To our knowledge, only one study<sup>22</sup> considered a nonlinear mixed-effects model (NLMEM) jointly estimated with a cause-specific multistate model using NONMEM software (ICON plc).

A popular application when treating of joint models is to derive individual dynamic predictions of both marker evolution and associated risks of event. It consists of considering a so-called landmark time and predicting, for one individual, the future evolution of the marker and the cumulative incidences of the competing events conditionally to the overall survival up to that landmark time.<sup>1,23</sup> In that case, the competing risk framework imposes to carefully define the at-risk population for predictions, in particular when assessing their performances.

In this work, we propose to use Monolix software (Lixoft) to jointly estimate a nonlinear mixed-effects submodel for modeling the longitudinal trajectory of a biomarker and a parametric SDH submodel for modeling multiple risks of events. This approach is then applied to a subset of OUTCOMEREA database in patients admitted in intensive care unit (ICU) for sepsis, a life-threatening condition where severe organ dysfunctions are caused by the infection of the organism by a pathogen.<sup>24</sup> We also propose a simulation study to assess performances of the estimation in terms of relative bias and relative root mean square error.

## METHODS

### General framework

#### Joint Model

##### *Nonlinear mixed-effects models*

Let  $N$  be the number of subjects and  $\{y_{i1}, \dots, y_{in_i}\}$  the vector of longitudinal observations of subject  $i$  (for  $i = 1, \dots, N$ ). Observation  $y_{ij}$  denotes the  $j$ th measurement of patient  $i$  at time  $t_{ij}$  (for  $j = 0, \dots, n_i$ ). Let  $m$  be the known nonlinear function describing the structural model and  $g$  the known function describing the error model. The NLMEM writes as follows:

$$y_{ij} = m(t_{ij}, \psi_i) + g[m(t_{ij}, \psi_i), \sigma] \epsilon_{ij} \quad (1)$$

$\psi_i = f(\mu, \eta_i, Z_i)$  denotes the individual parameters expressed as a function of population parameters, noted  $\mu$ ; individual random effects, noted  $\eta_i$ ; and an optional vector of individual covariates  $Z_i$ . The random effects are assumed to be normally distributed with mean 0 and variance–covariance matrix  $\Omega$ , and independent of the residual Gaussian error, noted  $\epsilon_{ij}$ , of mean 0 and variance 1.  $\sigma$  denotes the vector of the error model parameters.

##### *SDH models*

Let the variable  $T$  describe the time-to-event distribution,  $C$  the noninformative censoring time, and  $K$  denotes the event type that occurs. For the sake of simplicity, we consider only two events  $k \in \{1, 2\}$ . We set  $\delta = 1_{T < C} \times K$  as the event indicator, thus  $\delta = K$  in case of failure and  $\delta = 0$  in case of censoring. The couple  $(\tilde{T}, \delta)$  with  $\tilde{T} = \min(T, C)$  is observed for each individual. The cumulative incidence  $\mathbb{P}(T < t, K = k)$  on event  $k$  is modeled with an SDH  $\lambda_k$  associated with the variable  $T_k$  defined as:  $T_k = T \times 1_{K=k} + \infty \times 1_{K \neq k}$ . To model  $T_1$  and  $T_2$  and derive cumulative incidences, we introduce two parametric SDH functions  $\lambda_1$  and  $\lambda_2$ .

##### *Likelihood and estimation*

Longitudinal and survival parts are supposed to be linked as described in the following equation:

$$\begin{aligned} y_{ij} &= m(t_{ij}, \psi_i) + g[m(t_{ij}, \psi_i), \sigma] \epsilon_{ij} \\ \lambda_{1i}(t|\psi_i, W_i) &= h_1(t, \chi_1) \exp[\alpha_1 \times m(t, \psi_i) + \beta_1 \cdot W_i] \\ \lambda_{2i}(t|\psi_i, W_i) &= h_2(t, \chi_2) \exp[\alpha_2 \times m(t, \psi_i) + \beta_2 \cdot W_i] \end{aligned} \quad (2)$$

where  $h_1$  and  $h_2$  are the baseline hazard functions for the risk of event 1 and the risk of event 2, and which depend on vectors of parameters  $\chi_1$  and  $\chi_2$ , respectively.  $\alpha_1$  and  $\alpha_2$  are the coefficients that link the current predicted value of the marker with the instantaneous risks of event 1 and event 2, respectively. For the sake of clarity, we only considered a linear link between the longitudinal process and the subdistribution risks. Depending on the modeled data, more complex link functions could also be considered (e.g., involving splines or integration).  $W_i$  is the optional vector of covariates with coefficient  $\beta_1$  or  $\beta_2$ .

Both longitudinal and survival submodels are linked by shared random effects, and the joint likelihood writes:

$$L(\theta) = \prod_{i=1}^N \int_{\eta_i} \left[ \left( \prod_{j=0}^{n_i} p(y_{ij} | \eta_i, Z_i; \theta) \right) p(\tilde{T}_i, \delta_i | \eta_i; \theta) p(\eta_i; \theta) \right] d\eta_i \quad (3)$$

where  $\theta = \{\mu, \Omega, \sigma, \chi_1, \chi_2, \alpha_1, \alpha_2, \beta_1, \beta_2\}$  is the vector of parameters to estimate.  $p(y_{ij}|\eta_i, Z_i; \theta)$  is the density function of longitudinal process,  $p(\eta_i; \theta)$  the density function of random effects, and  $p(t, \delta|\eta_i; \theta)$  is the density function of the survival process given by:

$$p(t, \delta|\eta_i; \theta) = [\lambda_{1i}(t|\psi_i, W_i; \theta) \times (1 - F_{1i}(t|\psi_i, W_i; \theta))]^{1_{\delta=1}} \times [\lambda_{2i}(t|\psi_i, W_i; \theta) \times (1 - F_{2i}(t|\psi_i, W_i; \theta))]^{1_{\delta=2}} \times (1 - F_{1i}(t|\psi_i, W_i; \theta) - F_{2i}(t|\psi_i, W_i; \theta))^{1_{\delta=0}} \quad (4)$$

where, for  $k = \{1, 2\}$ ,  $F_{ki}(t|\psi_i, W_i; \theta) = 1 - \exp(-\int_0^t \lambda_{ki}(u|\psi_i, W_i; \theta) du)$  is the individual cumulative incidence function for event  $k$  at time  $t$ .

The population parameter  $\theta$  can be estimated by the maximization of the likelihood with the Stochastic Approximation Expectation Maximization (SAEM) algorithm implemented in Monolix (<http://lixoft.com>). This algorithm has been shown to be powerful, having good theoretical properties and accurately estimating parameters.<sup>25</sup> Variance-covariance matrix  $\Sigma$  of the population parameters is estimated with a stochastic approximation.<sup>26</sup> We note  $\hat{\theta}$  and  $\hat{\Sigma}$  their estimates.

## Individual dynamic predictions

A popular application of joint models is to provide individual dynamic predictions for the biomarker evolution and for survival outcome(s). For a new patient  $i$  who has longitudinal observations until a landmark time  $l$ ,  $Y_i(l) = \{y_{ij}; 0 \leq t_{ij} \leq l\}$ , we aim at predicting his future marker value at time  $l+t$ , denoted  $m_i(l+t|l)$ , and his event 1 survival probability  $s_i(l+t|l)$ . The time  $t$  is called horizon time. Because dynamic predictions are applied to patients who are event free before time  $l$ , we focus on the conditional probability of not experiencing the event of interest expressed as:

$$s_i(l+t|l) = \mathbb{P}(T_{1i} > l+t | T_i > l, Y_i(l), Z_i, W_i, \theta) \quad (5)$$

Thus, for each landmark time  $l$ , the marker measurements of patient  $i$  up to time  $l$  and the information provided by the population parameters are used to compute the a posteriori distribution of the individual parameters and infer the desired quantities. But as the a posteriori distribution has no closed form, Markov chain simulation is needed. In particular, some authors<sup>27</sup> compared different software and computing methods to draw a posteriori values and showed that Monolix software and its Metropolis-Hastings (MH) algorithm are able to provide those quantities with good properties. Then, the Monte Carlo process described in Rizopoulos<sup>23</sup> can be used to derive  $R$  samples of individual parameters. For a repetition  $r \in [1, R]$ ,

1. draw  $\theta^{(r)} \sim \mathcal{N}(\hat{\theta}, \hat{\Sigma})$  to take into account estimation uncertainty on population parameters
2. draw a realization of  $\eta_i^{(r)}$  on the a posteriori distribution:  $\eta_i^{(r)} \sim \{\eta_i | T_i > l, Y_i(l), Z_i, W_i, \theta^{(r)}\}$  and infer  $\psi_i^{(r)} = f(\mu^{(r)}, \eta_i^{(r)}, Z_i)$
3. compute  $m_i^{(r)}(l+t|\psi_i^{(r)})$  and  $s_i^{(r)}(l+t|\psi_i^{(r)}) = \frac{F_{1i}^{(r)}(\infty) + F_{2i}^{(r)}(\infty) - F_{1i}^{(r)}(l+t) - F_{2i}^{(r)}(l)}{F_{1i}^{(r)}(\infty) + F_{2i}^{(r)}(\infty) - F_{1i}^{(r)}(l) - F_{2i}^{(r)}(l)}$  with  $F_{ki}^{(r)}(t) = F_{ki}(t|\psi_i^{(r)}, W_i; \theta^{(r)})$ .

Estimates of  $m_i(l+t|l)$  and  $\pi_i(l+t|l)$  are derived with

$$\begin{aligned} \hat{m}_i(l+t|l) &= \text{median} \left\{ m_i^{(r)}(l+t|\psi_i^{(r)}), r = 1, \dots, R \right\} \\ \hat{s}_i(l+t|l) &= \text{median} \left\{ s_i^{(r)}(l+t|\psi_i^{(r)}), r = 1, \dots, R \right\} \end{aligned} \quad (6)$$

Prediction intervals can be derived by reporting the corresponding percentiles.

Performances are assessed in terms of discrimination with the time-dependent receiver operating characteristic area under the curve (ROC AUC)<sup>28</sup> and calibration with the time-dependent Brier score.<sup>28</sup> Details on the definition and computation of those indicators can be found in Supplementary Information S1.

## Application framework

### Data and objectives

#### Description of the data

The data are a subset of the OUTCOMEREA database composed of patients admitted in the ICU in 30 French centers between 1997 and 2015. Ethical and legal aspects of the database are described in Supplementary Information S2. In this application, we focus on patients included for sepsis or septic shock aged older than 18 years. Exclusion criteria were readmissions, admissions for palliative care, brain deaths, or early decisions not to forego life-sustaining therapy. The patient set was randomly divided into two subsets to build a training data set composed of two thirds of the data and a validation data set composed of the remaining one third. Daily assessment of the sepsis-related Sequential Organ Failure Assessment (SOFA) score was available for each patient stay. The SOFA score quantifies the number and the severity of organ dysfunctions between 0 (no organ dysfunction) and 24 (very severe organ dysfunctions).<sup>29-31</sup> This score is measured daily from ICU admission to discharge from ICU, but not after. ICU duration stay; status (alive/dead) at the end of the stay; and baseline covariates such as age, sex, and presence of comorbidities (e.g., diabetes, chronic heart failure, chronic kidney diseases,

liver cirrhosis, chronic obstructive pulmonary diseases and immunocompromised status) were also available.

### Objectives

The objective is to estimate jointly the individual SOFA score evolutions and the associated risk of ICU death and discharge within 30 days after ICU admission. It is worth noting that death and discharge are competing events because the occurrence of discharge precludes the possibility to observe the ICU death. It is also reasonable to presume that discharged patients do not have the same risk of death than other patients still hospitalized in the ICU. In this application, SOFA score evolution is modeled by a parametric NLMEM, death and discharge are modeled by a parametric competing risk sub-distribution model with ICU death as the event of interest and discharge from ICU as the competing event. The added value of modeling the longitudinal evolution of the SOFA score is assessed by comparing prediction performances of a baseline model (model that only includes baseline SOFA measurement and other baseline covariates) versus the joint model.

## Baseline and joint models

### Baseline model

SDHs are parametric functions with a constant baseline hazard noted  $h_1$  for the risk of death and  $h_2$  for the risk of discharge. Both instantaneous risks include the baseline SOFA measurement. Covariate selection for age, sex, and number of comorbidities is based on Bayesian information criterion (BIC). All combinations are tested, and the one that minimizes the BIC is added in the baseline model. Briefly, the model writes:

$$\begin{aligned}\lambda_{1i}(t|\psi_i, W_i) &= h_1 \exp\left[\beta_{S_1} SOFA_0 + \beta_1 \cdot W_i\right] \\ \lambda_{2i}(t|\psi_i, W_i) &= h_2 \exp\left[\beta_{S_2} SOFA_0 + \beta_2 \cdot W_i\right]\end{aligned}\quad (7)$$

### Joint model

To allow monotonous or nonmonotonous evolution (e.g., initial increase of SOFA followed by a decrease or vice versa), the structural model  $m$  is defined as:

$$m(t_{ij}, \psi_i) = \phi\left[b_{0i} + a_i \times (\exp(b_{1i} \times (t_{ij} - t_{lag,i})) - \exp(b_{2i} \times (t_{ij} - t_{lag,i})))\right] \quad (8)$$

where  $\phi$  is a capping function used to enforce the predicted SOFA score to remain between 0 and 24:

$$\phi(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x \leq 24 \\ 24 & \text{if } x > 24 \end{cases}$$

and we have:

$$\begin{cases} b_{0i} = \mu_0 + \eta_{0i} \\ b_{1i} = \mu_1 + \eta_{1i} \\ b_{2i} = \mu_2 + \eta_{2i} \\ a_i = \mu_a e^{\eta_{ai}} \\ t_{lag,i} = \mu_{lag} + \eta_{t_{lag},i} \end{cases} \quad \eta_{pi} \sim \mathcal{N}(0, \omega_p)$$

The parameter  $t_{lag}$  stands for the delay between symptom onset and the first SOFA evaluation. Details of coefficient interpretations can be found in Supplementary Information S3.

We used a combined error based on a model selection procedure using the BIC criteria. Thus,  $g[m(t_{ij}, \psi_i), \sigma] = \sigma_a + \sigma_b \times m(t_{ij}, \psi_i)$ , with  $\sigma = \{\sigma_a, \sigma_b\}$ . Details referring to error model selection are presented in Table S1 of the Supplementary Materials. In our joint model, the survival part includes baseline covariate(s) to model the dependence between prognosis and a patient's characteristics. In accordance with physicians, we made the hypothesis that the covariates only affect survival, for a given SOFA trajectory, and not the SOFA process itself. Therefore, it was not tested in the longitudinal sub-model. Of note, exploring and testing various covariates on each of the five parameters of the structural model would require a careful handling of multiple testing issues.

SDHs are the same parametric functions as presented in the baseline model but including a coefficient linking the longitudinal process and the respective instantaneous risk.

We used Monolix 2018R2 with a maximum of 2000 iterations in the exploratory phase and in the stochastic approximation for both baseline and joint model estimations.

Details and help for readers on practical model implementation can be found in Supplementary Information S4.

## Individual dynamic predictions

For both baseline and joint modeling, four landmark times are considered at 0, 3, 6, and 9 days after patient admission. Multiple horizon times are also considered, from landmark time plus 1 day to 30 days after admission.

To assess the performances of the models and according to the methods described, at a landmark time  $l$  and an horizon time  $t$ , a case is a patient who died in the ICU between  $l$  and  $l+t$ , and a control is a patient still hospitalized in the ICU or who was discharged between  $l$  and  $l+t$ . We note that,

in our context, no censoring occurs between landmark and horizon times, and there is only an administrative censoring at 30 days. Thus, estimators of the AUC and Brier score in the absence of censoring presented in Supplementary Information S1 are used (Equations 2 and 4).

For the settings of the MH algorithm, we set one chain with three proposal functions that are used in turn: the population distribution, a unidimensional Gaussian random walk, or a multidimensional Gaussian random walk. For the random walks, the variance of the Gaussian is automatically adapted to reach an optimal acceptance ratio of 0.3. The algorithm stops when, for all parameters, the average conditional means and standard deviations of the last 50 iterations do not deviate by more than 5%.

### Evaluation of the estimation performances by simulation

We conducted a simulation study to confirm that the SAEM algorithm implemented on Monolix 2018R2 provides good estimation performances of such a joint model. We assessed the estimation accuracy in terms of relative bias and evaluated the incertitude of estimation using relative root mean square errors (RRMSEs).

#### Simulation setting

We simulated  $M = 200$  data sets of  $N = 1000$  patients assuming that SOFA score measurements were available every day for at most 30 days after the ICU admission. SOFA measurements were reported for each patient until time of ICU death or ICU discharge or also end of the first 30 days of the stay, with no other censoring process.

The longitudinal data were simulated according to the model proposed in the application methods. Thus, this model is used:

$$y_{ij} = m(\psi_i, t_{ij}) + [\sigma_a + \sigma_b \times m(\psi_i, t_{ij})] \epsilon_{i,j}$$

$$m(t_{ij}, \psi_i) = \phi[b_{0i} + a_i \times (\exp(b_{1i} \times (t_{ij} - t_{lag,i})) - \exp(b_{2i} \times (t_{ij} - t_{lag,i})))]$$
(9)

with

$$\phi(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x \leq 24 \\ 24 & \text{if } x > 24 \end{cases} \quad \begin{cases} b_{0i} = \mu_0 + \eta_{0i} \\ b_{1i} = \mu_1 + \eta_{1i} \\ b_{2i} = \mu_2 + \eta_{2i} \\ a_i = \mu_a e^{\eta_{ai}} \\ t_{lag,i} = \mu_{lag} + \eta_{lag,i} \end{cases} \quad \eta_{pi} \sim \mathcal{N}(0, \omega_p)$$

For the survival part, we considered here two competing events and one categorical covariate  $W$  with three

modalities. We adapted methods already described<sup>32,33</sup> to generate failure time data. The subdistribution for the event of interest is given by:

$$\mathbb{P}(T_{1i} < t | \psi_i, W_i; \theta) = 1 - \exp\left(-\int_0^t \frac{p_1 g_1 \exp(-g_1 \times s) \times \lambda(s)}{1 - p_1 (1 - \exp(-g_1 \times s))} ds\right)$$
(10)

with  $\lambda(s) = \exp(\alpha_1 \times m(\psi_i, t_{ij}) + \beta_1 \times W_i)$ . Parameters  $p_1$  and  $g_1$  are used to control the number of event 1 at infinite time and the speed of onset. The subdistribution for the competing risk was then obtained using an exponential distribution with rate  $\frac{t}{b}$ , where  $b$  controls the speed of event 2 onset:

$$\mathbb{P}(T_{2i} < t | \psi_i, W_i; \theta) = (1 - F_1(\infty)) \times \left(1 - \exp\left(-\frac{t}{b}\right)\right)$$
(11)

#### Estimation setting

For each data set  $m \in \{1, \dots, M\}$ , we estimated the following joint model:

$$y_{ij} = m(t_{ij}, \psi_i) + [\sigma_a + \sigma_b \times m(\psi_i, t_{ij})] \epsilon_{i,j}$$

$$\lambda_{1i}(t | \psi_i, W_i) = \frac{p_1 g_1 \exp(-g_1 \times t)}{1 - p_1 (1 - \exp(-g_1 \times t))} \exp[\alpha_1 \times m(t, \psi_i) + \beta_1 \cdot W_i]$$

$$\lambda_{2i}(t | \psi_i, W_i) = \frac{p_2 g_2 \exp(-g_2 \times t)}{1 - p_2 (1 - \exp(-g_2 \times t))} \exp[\alpha_2 \times m(t, \psi_i) + \beta_2 \cdot W_i]$$
(12)

Estimation was performed by the maximization of the likelihood using the SAEM algorithm implemented on Monolix 2018 software with the same settings as explained in the application.

#### Evaluation criteria

We aimed to assess the performances of the estimation on a subset of the vector of parameters, the longitudinal parameters, and the survival parameters for the event of interest (death). We defined  $\rho = \{\mu, \omega, \alpha_1, \beta_1\}$  and  $\hat{\rho}$  its estimate. The model specification for data simulation does not allow  $\alpha_2$  and  $\beta_2$  to be easily identified, which is why the estimation evaluation was based on vector  $\hat{\rho}$ . We evaluated the relative bias and the RRMSE of  $\hat{\rho}$  defined as:

$$\text{Relative bias} = \frac{\mathbb{E}[\hat{\rho}] - \rho}{\rho} \times 100$$

and

$$\text{RRMSE} = \sqrt{\frac{\mathbb{E}[(\hat{\rho} - \rho)^2]}{\rho}} \times 100$$
(13)

We also provided violin plots of relative estimation errors (REE) defined for each data set  $m \in \{1, \dots, M\}$  as:

$$\text{REE} = \frac{\hat{\rho}^m - \rho}{\rho} \times 100 \quad (14)$$

## RESULTS

### Data description

A total of 6046 patients were considered, with 4050 patients included in the training set and 1996 in the validation set. The data flowchart is detailed in [Figure 1](#).

Most of the patients were men (62%), and the median age at admission was 65.1 years ([Q1–Q3]: [52.7–76.2]). About half of the patients had no comorbidity at admission, one third had one, and one third had more than one. Median admission SOFA was 6 and 30 days after admission, about 19% of the patients died, whereas 72% were discharged. [Table 1](#) summarizes the main patient characteristics at admission in both data sets.

[Figure 2a](#) shows the longitudinal SOFA trajectories of 100 random patients of the training data set. [Figure 2b](#) shows the cumulative incidences of both competing events (death and discharge) on the training set, truncated at day 30.

### Baseline model fit

The baseline model includes the baseline SOFA measurement and the patient age in three categories with cutoffs at 60 and 75 years (see [Table S2](#) of the Supplementary Materials). The baseline SOFA already has a strong effect on the instantaneous risks of death and discharge: at a constant age category, a patient with a one-point higher baseline SOFA has his instantaneous risk of death multiplied by 1.22 with a 95% confidence interval

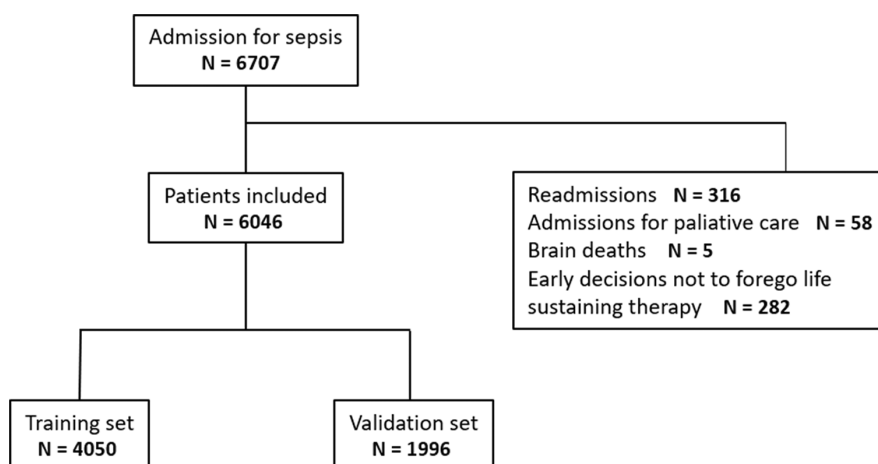
(CI) of [1.20, 1.24] and his instantaneous risk of discharge divided by 1.13 with a 95% CI of [1.12, 1.14]. Age also has an effect on both risks: for the same baseline SOFA, patients aged [60, 75[years have an instantaneous risk of death multiplied by 1.40 (with 95% CI: [1.18, 1.68]), and patients aged 75 years or older have their risk multiplied by 1.84 (with 95% CI: [1.55, 2.19]), compared with patients aged <60 years. For the risk of discharge from ICU, the effect is more moderate: patients aged [60, 75[years have an instantaneous risk of discharge divided by 1.23 (with 95% CI: [1.14, 1.35]), and patients aged 75 years or older have their risk divided by 1.30 (with 95% CI: [1.19, 1.43]). All parameter estimates can be found in [Table S3](#) of the Supplementary Materials.

### Joint model fit

The parameter estimates of the joint model can be found in [Table 2](#).

The  $p$  values from the Wald test are reported for the link coefficients and those associated with covariate effects. As expected, we showed a significant association between an increase of the SOFA score and a higher risk of death ( $\alpha_1 = 0.371$  with 95% CI: [0.356, 0.386]) and a lower risk of discharge ( $\alpha_2 = -0.375$  with 95% CI: [-0.390, -0.360]). We also reported the significant effect of age class on both death and discharge risk. Age has a more important impact on the risk of death, similar to the baseline model: for the same SOFA evolution, patients aged [60, 75[years have an instantaneous risk of death multiplied by 1.42 (with 95% CI: [1.20, 1.68]), and patients aged 75 years or older have their risk multiplied by 2.15 (with 95% CI: [1.70, 2.72]), compared with the reference class age (<60 years).

Some diagnostic plots and indicators that assessed the quality of the longitudinal model and the joint model, respectively, can be found in [Supplementary Information S5](#).

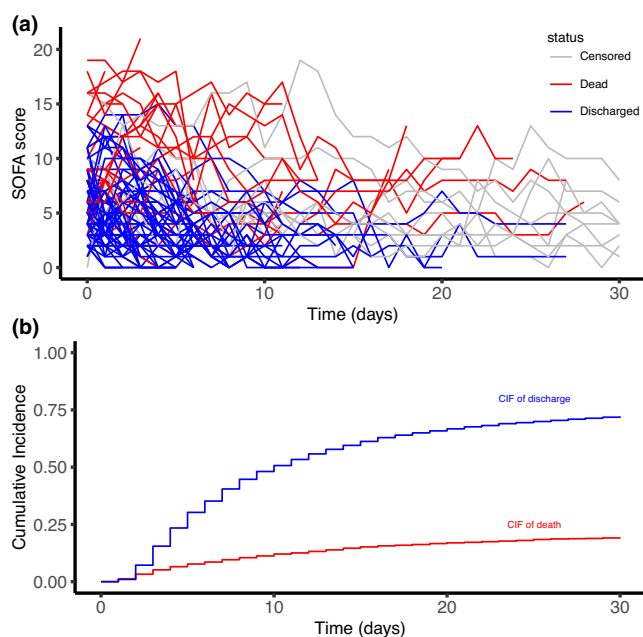


**FIGURE 1** Data flowchart. A total of 6046 patients were included in the analysis, with 4050 in the training set and 1996 in the validation set.

**TABLE 1** Main patient characteristics at intensive care unit admission

Variable	Training set ( <i>n</i> = 4050)	Validation set ( <i>n</i> = 1996)
Age, years, <i>n</i> (%)		
<60	1565 (39)	782 (39)
[60, 75]	1333 (33)	682 (34)
≥75	1152 (28)	532 (27)
Sex, <i>n</i> (%)		
Male	2491 (61)	1272 (64)
Female	1559 (39)	726 (36)
Comorbidities, <i>n</i> (%)		
0	1988 (49)	985 (49)
1	1340 (33)	682 (34)
>1	722 (18)	329 (17)
Admission SOFA, median [Q1–Q3]	6 [4–9]	6 [4–9]
30-day mortality, <i>n</i> (%)	781 (19)	362 (18)
30-day discharge, <i>n</i> (%)	2921 (72)	1426 (71)

Abbreviation: SOFA, Sequential Organ Failure Assessment.

**FIGURE 2** Exploratory plots of longitudinal and competing risk intensive care unit (ICU) data of the training set: (a) the Sequential Organ Failure Assessment (SOFA) score evolution of 100 random patients and (b) the cumulative incidences of ICU death and discharge from ICU (no censoring occurred during the follow-up). CIF, cumulative incidence function.

## Dynamic predictions

Figure 3 illustrates dynamic predictions for two patients (a and b) of the validation set using the joint model. It

shows the observed and predicted SOFA evolution on the top and predicted individual survival probability on the bottom, depending on landmark time. Predicted SOFA evolution and predicted survival are updated at each landmark time. Patient a of the figure shows an initial clinical deterioration, followed by an improvement of his clinical state. The more landmark time increases, better is his predicted survival. The SOFA score of Patient b monotonously increases over time, resulting in his predicted survival at 30 days getting worse as the landmark increases. He finally died on Day 26.

To visualize the covariate effect on predictions, in Supplementary Information S6 we added the median [Q1–Q3] survival predictions in the validation population for the three age categories and for baseline SOFA splits in three categories for various landmark times.

To assess the predictive performances of the baseline and joint models, we used the 1996 patients of the validation set. Time-dependent ROC AUC and the Brier score including 95% CIs within the joint model are presented in Figure 4. The ones for the baseline model are presented in Figure S1 of the Supplementary Materials. The joint modeling approach allows to improve the quality of prediction when the follow-up is sufficient. For instance, when considering landmark time = 6 and horizon time = 30, AUC = 0.63 (95% CI: [0.58, 0.68]) and AUC = 0.77 (95% CI: [0.74, 0.80]) for the baseline and the joint model, respectively. However, when considering an early landmark where only one SOFA observation is available, the joint model does not perform better than the baseline model (for landmark time = 0 and horizon time = 30, AUC = 0.73, 95% CI = [0.70, 0.77] and AUC = 0.69 with 95% CI = [0.66, 0.72] for the baseline and joint models, respectively).

## Results of the simulation study

Values of parameters (except  $p_1$ ,  $g_1$  and  $b$ ) used to simulate data were based on estimates obtained by maximization of the likelihood using this model on the OUTCOMEREA database presented previously. The details of the values can be found in Table 3.  $p_1$ ,  $g_1$  were set to 0.01, 0.05, respectively, and  $b$  was fixed to 6 to obtain about 20% Cause 1 failures, 78% Cause 2 failures, and 2% administrative censoring at time 30, similar to the application.

Table 3 provides the relative biases and RRMSEs expressed in percentage. Relative biases are low for all parameters, which reflects good accuracy. The parameter that links the longitudinal and the survival process ( $\alpha_1$ ) was of particular interest. The simulation showed that it is

**TABLE 2** Joint model parameter estimation

Parameter	Value	SE	RSE (%)	<i>p</i> Value <sup>a</sup>
Longitudinal submodel				
Fixed effects				
$\mu_0$	1.69	0.0547	3.23	
$\mu_1$	-0.262	0.0051	1.93	
$\mu_2$	-1.13	0.0188	1.66	
$\mu_a$	10.4	0.135	1.29	
$\mu_{t_{lag}}$	-1.48	0.0286	1.94	
Random effects				
$\omega_0$	2.56	0.0424	1.66	
$\omega_1$	0.194	0.00411	2.12	
$\omega_2$	0.494	0.013	2.63	
$\omega_a$	0.459	0.00962	2.10	
$\omega_{t_{lag}}$	0.669	0.02	3.00	
Error parameters				
$\sigma_a$	0.887	0.0123	1.39	
$\sigma_b$	0.127	0.00265	2.08	
Survival submodel				
Death				
$h_1$	0.000476	$6.52 \times 10^{-5}$	13.7	
$\alpha_1$	0.371	0.00788	2.12	<10 <sup>-4</sup>
$\beta_{1,Age1}(ref)$	0			
$\beta_{1,Age2}$	0.348	0.0876	25.1	<10 <sup>-4</sup>
$\beta_{1,Age3}$	0.766	0.119	15.6	<10 <sup>-4</sup>
Discharge				
$h_2$	0.325	0.00901	2.77	
$\alpha_2$	-0.375	0.00768	2.05	<10 <sup>-4</sup>
$\beta_{2,Age1}(ref)$	0			
$\beta_{2,Age2}$	-0.195	0.0187	9.58	<10 <sup>-4</sup>
$\beta_{2,Age3}$	-0.225	0.0302	13.4	<10 <sup>-4</sup>

Abbreviation: RSE, relative standard error.

<sup>a</sup>Wald test *p* values are reported for link and covariate coefficients.

The parameters are those from (8).

accurately estimated with a relative bias of about 2% and also precisely estimated with a RRMSE about 9%. Overall, the parameters are precisely estimated except parameters associated with the baseline covariate, which have high RRMSEs.

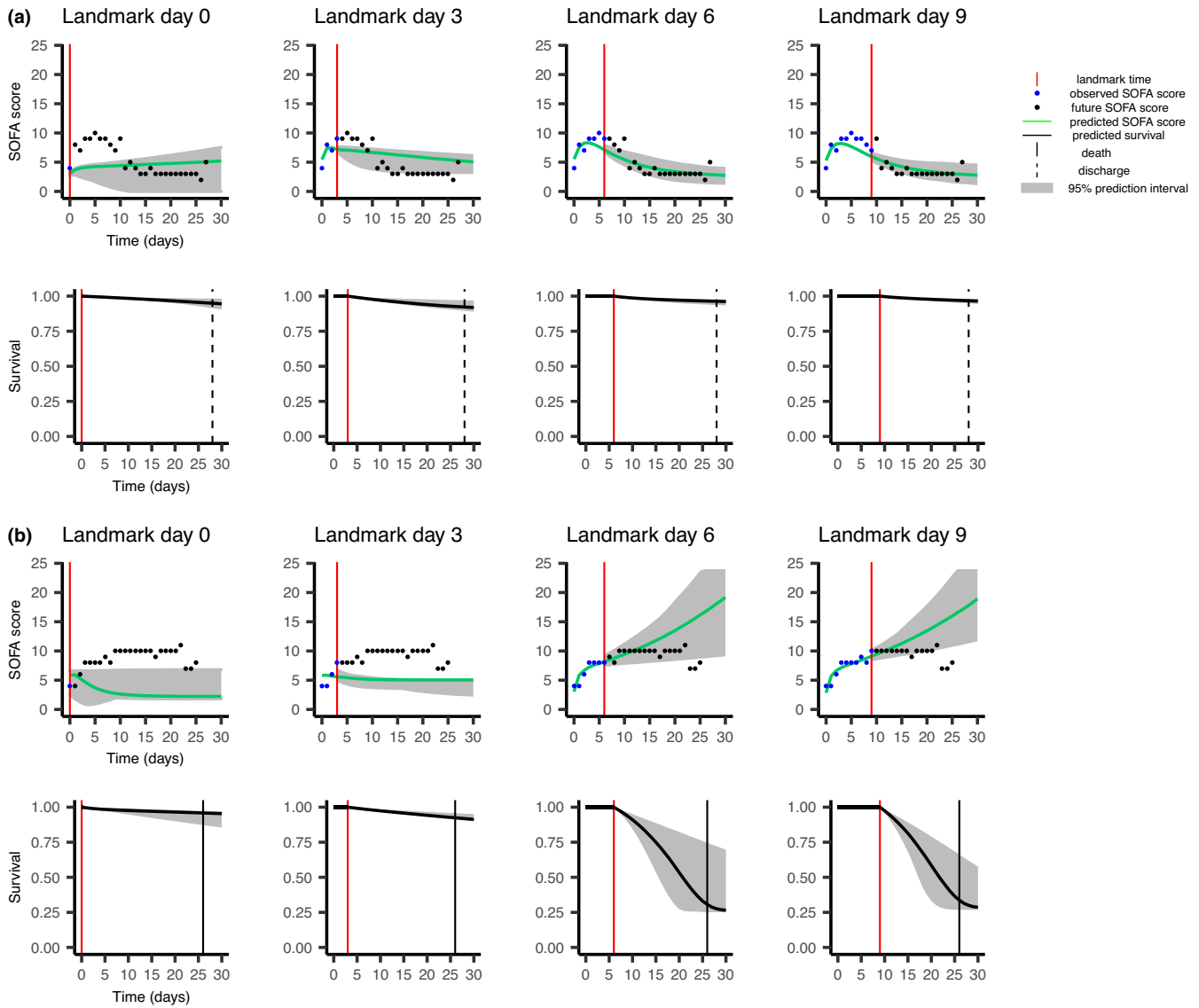
Violin plots can be found in Figure S2A of the Supplementary Materials and provide the same conclusion.

## DISCUSSION

In many biomedical applications, the patients are not only at risk of a single event, but of several. Joint modeling

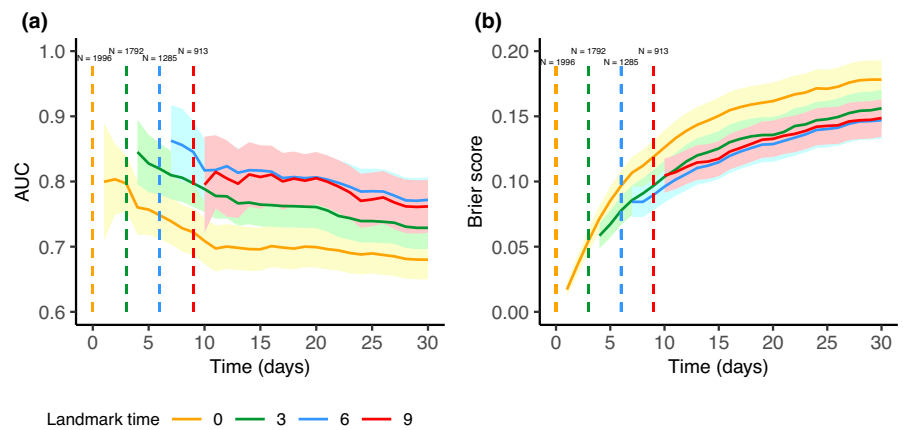
under competing risks is seldom explored, particularly in a full parametric setting.

In our application, a significant association was found between the longitudinal evolution of the SOFA and each outcome (death and discharge): an increase of the SOFA score is associated with a higher risk of death and a lower risk of discharge. In the joint model, higher uncertainty was found in the estimates of the baseline age effect when adjusted on the SOFA evolution. Of note, simulations according to the baseline model (see Figure S2B of the Supplementary Materials), which is adjusted on the baseline SOFA value, also reported the same trend: high uncertainty in the estimates of the baseline age effect, which highlights the difficulty



**FIGURE 3** Dynamic predictions for two random patients (a and b) of the validation data set. For each patient, predictions of the Sequential Organ Failure Assessment (SOFA) evolution are on the top, and predictions of the risk of death are on the bottom. Patient a, who is older than 75years, was discharged 28days after his intensive care unit admission, whereas Patient b, also older than 75years, died 26days after his admission.

**FIGURE 4** Time-dependent receiver operating characteristic area under the curve (ROC AUC) (a) and Brier score (b) computed on the validation set depending on landmark time. The 95% confidence intervals are represented by the corresponding colored intervals. Number of at-risk patients at each landmark time is specified on the top.



**TABLE 3** Estimated relative bias and root mean square errors based on 200 simulations

Parameter	True value	Relative bias (%)	RRMSE (%)
Longitudinal part			
Fixed effects			
$\mu_0$	2	2	11
$\mu_1$	-0.25	5	10
$\mu_2$	-1	4	8
$\mu_a$	10	-0.06	5
$\mu_{t_{lag}}$	-1	6	12
Random effects			
$\omega_0$	3	6	9
$\omega_1$	0.2	15	18
$\omega_2$	0.5	-0.5	13
$\omega_a$	0.5	-14	16
$\omega_{t_{lag}}$	0.8	6	11
Error parameters			
$\sigma_a$	0.9	-0.3	2
$\sigma_b$	0.1	3	5
Survival part			
$\alpha_1$	0.35	2	9
$\beta_{1,Age2}$	0.34	-12	85
$\beta_{1,Age3}$	0.76	7	39

Abbreviation: RRMSE, relative root mean square error.

The parameters are those from (12).

in estimating age effect when the SOFA score is in the model.

For landmark time 3 and more, as sufficient information becomes available, individual predictions from the joint model are significantly improved and outperform those from the baseline model. At landmark time 0, no added value of the joint model was shown as there is no longitudinal information available. Individual predictions at any landmark time, using all available SOFA scores at that time, can be very useful for clinicians in routine medical practice as it may help clinical decisions such as therapeutic escalation or limitation, as indeed those decisions arise during the follow-up and not at admission.

The proposed joint model could probably be improved on some points. First, we could test for the covariate effects in the parameters of the longitudinal submodel. Second, we did not assess graphically the quality of the survival model fit because usual diagnostic plots are, to date, unavailable for this kind of model. This is because these plots rely on simulations, which are currently not available for SDH formulation and would require additional

developments that are beyond the scope of this article. Finally, we considered the subdistribution approach to treat the competing risk setting, and sensitivity analyses can be done with a cause-specific approach.

Our approach with joint modeling is justified when missing data for the longitudinal biomarker are MNAR. However, depending of the richness of the design, missing data can be considered missing at random (MAR), and some authors<sup>34</sup> showed that treating MAR data with a joint modeling approach can lead to bias in estimates.

The simulation study showed that our estimation procedure has good statistical properties with no major bias in longitudinal parameters and in the coefficient that links the longitudinal process with the survival process. However, we cannot evaluate the bias on event 2 because of the model specification in the simulation process. Further developments are thus needed to fully investigate the quality of estimation on both events as well as the type I error and power of statistical tests involving the coefficients.

In conclusion, we showed in this work that Monolix software and its SAEM algorithm can be used for joint modeling with competing events and provide accurate parameter estimates. We also presented a real-case application with dynamic predictions. The methodology is easy to reproduce and can be easily extended to other clinical applications because of the general form of the model. The longitudinal model can also be extended to ordinary differential equations in a more pharmacometric application.

## AUTHOR CONTRIBUTIONS

A.L.-M. wrote the manuscript. J.-F.T., F.M., and J.M. designed the research. A.L.-M. and J.M. performed the research. A.L.-M. analyzed the data.

## ACKNOWLEDGMENTS

The members of the OUTCOMEREA network are detailed in Supplementary Information S7.

## FUNDING INFORMATION

Dr. Timsit was supported by the Innovative Medicines Initiative under grant agreement no. 115737-2-COMBACTE-MAGNET.

## CONFLICT OF INTEREST

The authors declared no competing interests for this work. As Editor-in-Chief of *CPT: Pharmacometrics & Systems Pharmacology*, France Mentre was not involved in the review or decision process for this paper.

## ORCID

France Mentre  <https://orcid.org/0000-0002-7045-1275>

## REFERENCES

1. Desmée S, Mentré F, Veyrat-Follet C, et al. Nonlinear joint models for individual dynamic prediction of risk of death using Hamiltonian Monte Carlo: application to metastatic prostate cancer. *BMC Med Res Methodol.* 2017;17(1):105. doi:10.1208/s12248-015-9745-5
2. Musoro JZ, Zwinderman AH, Abu-Hanna A, Bosman R, Geskus RB. Dynamic prediction of mortality among patients in intensive care using the sequential organ failure assessment (SOFA) score: a joint competing risk survival and longitudinal modeling approach. *Stat Neerl.* 2018;72(1):34-47. doi:10.1111/stan.12114
3. Ferrer L, Putter H, Proust-Lima C. Individual dynamic predictions using landmarking and joint modelling: validation of estimators and robustness assessment. *Stat Methods Med Res.* 2019;28(12):3649-3666.
4. Kerioui M, Mercier F, Bertrand J, et al. Bayesian inference using Hamiltonian Monte-Carlo algorithm for nonlinear joint modeling in the context of cancer immunotherapy. *Stat Med.* 2020;39(30):4853-4868.
5. Latouche A, Porcher R, Chevret S. Note on including time-dependent covariate in regression model for competing risks data. *Biom J.* 2005;47(6):807-814.
6. Rizopoulos D. *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R.* Chapman & Hall/CRC; 2012.
7. Saha C, Jones MP. Asymptotic bias in the linear mixed effects model under non-ignorable missing data mechanisms. *J R Stat Soc Series B Stat Methodol.* 2005;67(1):167-182.
8. Zhudenkov K, Gavrilov S, Sofronova A, et al. A workflow for the joint modeling of longitudinal and event data in the development of therapeutics: tools, statistical methods, and diagnostics. *CPT Pharmacometrics Syst Pharmacol.* 2022;11(4):425-437.
9. Prentice RL, Kalbfleisch JD, Peterson AV, et al. The analysis of failure times in the presence of competing risks. *Biometrics.* 1978;34(4):541-554.
10. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc.* 1999;94(446):496-509.
11. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med.* 2007;26(11):2389-2430. doi:10.1002/sim.2712
12. Lau B, Cole S, Gange S. Competing risk regression models for epidemiologic data. *Am J Epidemiol.* 2009;170(2):244-256. doi:10.1093/aje/kwp107
13. Grambauer N, Schumacher M, Beyersmann J. Proportional subdistribution hazards modeling offers a summary analysis, even if misspecified. *Stat Med.* 2010;29(7-8):875-884.
14. Proust-Lima C, Dartigues JF, Jacqmin-Gadda H. Joint modeling of repeated multivariate cognitive measures and competing risks of dementia and death: a latent process and latent class approach. *Stat Med.* 2016;35(3):382-398.
15. Andrinopoulou ER, Rizopoulos D, Takkenberg JJ, et al. Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. *Stat Methods Med Res.* 2017;26(4):1787-1801.
16. Rué M, Andrinopoulou ER, Alvares D, Armero C, Forte A, Blanch L. Bayesian joint modeling of bivariate longitudinal and competing risks data: an application to study patient-ventilator asynchronies in critical care patients. *Biome J.* 2017;59(6):1184-1203.
17. Sweeting MJ, Thompson SG. Joint modelling of longitudinal and time-to-event data with application to predicting abdominal aortic aneurysm growth and rupture. *Biom J.* 2011;53(5):750-763. doi:10.1002/bimj.201100052
18. Baghfalaki T, Ganjali MA. Bayesian approach for joint modeling of skew-normal longitudinal measurements and time to event data. *Revstat.* 2015;13(2):169-191.
19. Borges A, Sousa I, and Castro L. Joint modelling of longitudinal CEA tumour marker progression and survival data on breast cancer. *AIP Conf Proc* 2017; 1836 (1):020043.
20. Tardivon C, Desmée S, Kerioui M, et al. Association between tumor size kinetics and survival in patients with urothelial carcinoma treated with Atezolizumab: implication for patient follow-up. *Clin Pharmacol Ther.* 2019;106(4):810-820. doi:10.1002/cpt.1450
21. van Niekerk J, Bakka H, Rue H. Competing risks joint models using R-INLA. *Stat Model.* 2021;21(1-2):56-71.
22. Krishnan S, Friberg L, Bruno R, Beyer U, Jin JY, Karlsson MO. Multistate model for Pharmacometric analyses of overall survival in HER2-negative breast cancer patients treated with docetaxel. *CPT Pharmacometrics Syst Pharmacol.* 2021;10(10):1255-1266. doi:10.1002/PSP4.12693
23. Rizopoulos D. Dynamic predictions and prospective accuracy in JointModels for longitudinal and time-to-event data. *Biometrics.* 2011;67(3):819-829.
24. Soo A, Zuege DJ, Fick GH, et al. Describing organ dysfunction in the intensive care unit: a cohort study of 20,000 patients. *Crit Care.* 2019;23(1):186.
25. Delyon B, Lavielle M, Moulines EJ. Convergence of a stochastic approximation version of EM algorithm. *Ann Stat.* 1999;27(1):94-128.
26. Kuhn E, Lavielle M. Maximum likelihood estimation in nonlinear mixed effects model. *Comput Stat Data Anal.* 2005;49:1020-1038.
27. Riglet F, Mentré F, Veyrat-Follet C, Bertrand J. Bayesian individual dynamic predictions with uncertainty of longitudinal biomarkers and risks of survival events in a joint modelling framework: a comparison between Stan, Monolix, and NONMEM. *AAPS J.* 2020;22(2):50.
28. Blanche P, Proust-Lima C, Loubère L, Berr C, Dartigues JF, Jacqmin-Gadda H. Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics.* 2015;71(1):102-113.
29. Vincent JL, Moreno R, Takala J, et al. The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. *Intensive Care Med.* 1996;22(7):707-710.
30. Vincent JL, de Mendonça A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine. *Crit Care Med.* 1998;26(11):1793-1800.
31. Timsit JF, Fosse JP, Troché G, et al. Calibration and discrimination by daily logistic organ dysfunction scoring comparatively with daily sequential organ failure assessment scoring for predicting hospital mortality in critically ill patients. *Crit Care Med.* 2002;30(9):2003-2013.
32. Fine JP. Regression modeling of competing crude failure probabilities. *Biostatistics.* 2001;2(1):85-97.

33. Deslandes E, Chevret S. Joint modeling of multivariate longitudinal data and the dropout process in a competing risk setting: application to ICU data. *BMC Med Res Methodol.* 2010;10:69.
34. Thomadakis C, Meligkotsidou L, Pantazis N, Touloumi G. Longitudinal and time-to-drop-out joint models can lead to seriously biased estimates when the drop-out mechanism is at random. *Biometrics.* 2019;75(1):58-68.

### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Lavalley-Morelle A, Timsit J-F, Mentré F, Mullaert J, The OUTCOMEREA network. Joint modeling under competing risks: Application to survival prediction in patients admitted in Intensive Care Unit for sepsis with daily Sequential Organ Failure Assessment score assessments. *CPT Pharmacometrics Syst Pharmacol.* 2022;11:1472-1484. doi:[10.1002/psp4.12856](https://doi.org/10.1002/psp4.12856)

# Supplementary Information S1 - Time-dependant ROC AUC and Brier score indicators

The predictive performances of the model can be assessed in terms of discrimination (i.e. ability to separate patients according to their prognosis) with time-dependent Area Under the Curve (AUC), and calibration (i.e. accuracy of predictions) with time-dependent Brier score. We note  $\pi_i(l+t|l) = 1 - s_i(l+t|l)$ , the probability for patient  $i$  to experience event 1 between  $l$  and  $l+t$ . At a landmark time  $l$  and an horizon time  $t$ , we define a case as a patient who experienced the event of interest (event 1) between  $l$  and  $l+t$ , and a control as a patient who did not experience any event or experienced the competing event (the event 2) between  $l$  and  $l+t$ .

In the following, we present definitions and estimates of AUC and Brier score according to Blanche et al. works (2013) [1]. Estimates are provided without accounting for censoring. If data contains censored observations, refined methods to handle censoring based on inverse probability of censoring weighting are preferable, also described in Blanche et al. (2013)[1].

The AUC for a landmark time  $l$  and a prediction horizon  $t$  is defined as:

$$AUC(l, t) = \mathbb{P}(\pi_i(l+t|l) > \pi_j(l+t|l) \mid D_i(l, t) = 1, D_j(l, t) = 0, T_i > l, T_j > l) \quad (1)$$

where  $D_i(l, t) = \mathbb{1}_{(l < T_i \leq l+t, K_i=1)}$ ; that is to say,  $D_i(l, t) = 1$  if the subject experiences the event of interest between time  $l$  and  $l+t$ , and  $D_i(l, t) = 1$  if the subject experiences the competing event between time  $l$  and  $l+t$  or is event-free at time  $l+t$ . AUC is estimated by:

$$\widehat{AUC}(l, t) = \frac{\sum_{i=1}^{N_l} \sum_{j=1}^{N_l} \mathbb{1}_{\hat{\pi}_i(l+t|l) > \hat{\pi}_j(l+t|l)} D_i(l, t) (1 - D_j(l, t))}{\sum_{i=1}^{N_l} \sum_{j=1}^{N_l} D_i(l, t) (1 - D_j(l, t))} \quad (2)$$

where  $N_l$  denotes the number of at-risk patients at time  $l$ .

The Brier score for a landmark time  $l$  and a prediction horizon  $t$  is defined as:

$$BS(l, t) = \mathbb{E} \left[ (D(l, t) - \pi(l+t|l))^2 \mid T > l \right] \quad (3)$$

and estimated by:

$$\widehat{BS}(l, t) = \frac{1}{N_l} \sum_{i=1}^{N_l} (D_i(l, t) - \hat{\pi}_i(l+t|l))^2 \quad (4)$$

A method for the derivation of asymptotic confidence intervals is proposed in the literature [1].

## References

- [1] Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat. Med.* 2013; 32: 5381-5397

## **Supplementary Information S2 - Ethical considerations**

The database is in accordance with French legislation concerning biomedical research. Patients or their family gave authorization for collection, conservation, and use of their personal anonymized data. Authorizations were obtained from the French Informatics and Liberty Commission (CNIL, registration no. 8999262), the French Advisory Committee for Data Processing in Health Research (CCTIRS). The database protocol was submitted to the Institutional Review Board of the Clermont-Ferrand University hospital (Clermont-Ferrand, France), who waived the need for informed consent (IRB no. 5891).

## Supplementary Information S3 : Longitudinal submodel

The non-linear structural function  $m$  is defined as :

$$m(t_{ij}, \psi_i) = \phi \left[ b_{0i} + a_i \times \left( e^{b_{1i} \times (t_{ij} - t_{lag_i})} - e^{b_{2i} \times (t_{ij} - t_{lag_i})} \right) \right]$$

$b_0$  is the SOFA value at time  $t_{lag}$ , for example if  $t_{lag} = 0$ ,  $b_0$  is the SOFA value at admission, and if  $t_{lag} = -1$ ,  $b_0$  is the SOFA value one day after patient admission. Parameter  $a$  plays a role on the slope of the SOFA evolution. The figure below illustrates the possible and relevant clinical evolution depending on parameters  $b_1$ ,  $b_2$ , and  $a$  and when  $t_{lag} = 0$ . If  $t_{lag} > 0$  the curves shift to the right, and if  $t_{lag} < 0$ , to the left. The dotted line corresponds to a doubling of parameter  $a$ . For an initial SOFA value at  $b_0$ , plot A represents the best scenario for the patient with a monotonous decrease of the SOFA over time. Plot B is the worst scenario with a monotonous increase of the SOFA over time. Plots C and D presents when the patient knows a clinical state change during the ICU stay, for example an increase of the SOFA during the first days and then a decrease (C) or vice-versa (D).

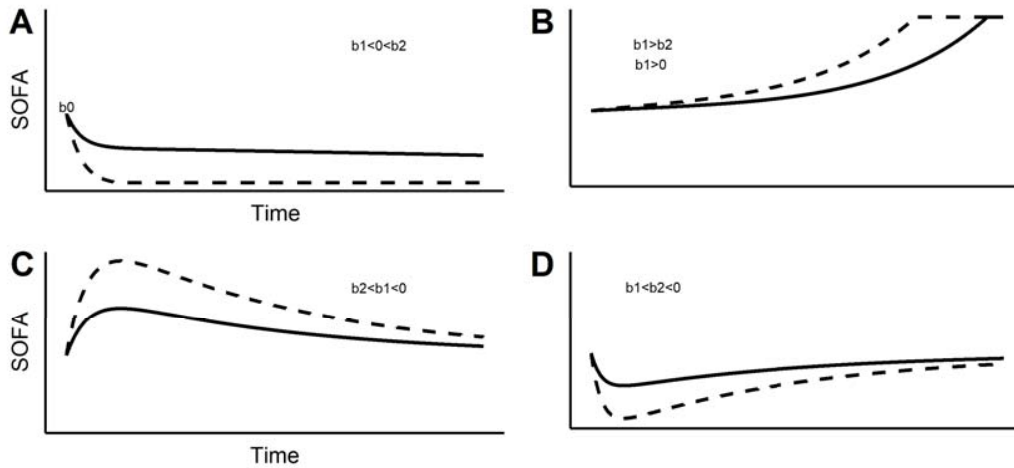


Figure: Possible shapes of the non-linear function depending on parameters  $\beta_1$ ,  $\beta_2$ , and  $a$ . The dotted curve represents a doubling of parameter  $a$ .

# Supplementary Information S4 - Implementation of the joint model on Monolix software

## Creating data set

The data set is composed of three parts: the longitudinal one and the two survival ones. We denote *Obs\_id* the observation id, 1 for longitudinal data, 2 for time-to-death data, and 3 for time-to-discharge data. We suppose that patient A was discharged from ICU at day 7. For this patient, the data has the following form:

Patient	Time	Obs	Obs_id	COV
A	0	11	1	2
A	1	11	1	2
A	2	12	1	2
A	3	9	1	2
A	4	8	1	2
A	5	7	1	2
A	6	8	1	2
A	7	7	1	2
A	0	0	2	2
A	30	0	2	2
A	0	0	3	2
A	7	1	3	2

Because we are in a subdistribution approach, the assumption is that the patient will never experience the other event (death in ICU), that is the reason why the event with observation id 2 was not observed until the end of the study (at time 30).

## Model file

If we consider the model presented in Method section, the model file is written as:

```
DESCRIPTION : Joint Model

[LONGITUDINAL]
input={beta0,b1,b2,a,tlag,h1,h2,alpha1,alpha2,h_cov1,h_cov2}

EQUATION :

SOFA = max(min(beta0 + a*(exp(b1*(t-tlag))-exp(b2*(t-tlag))),24),0)

haz1 = h1*exp(alpha1*SOFA + h_cov1)
haz2 = h2*exp(alpha2*SOFA + h_cov2)

DEFINITION:
death = {type=event, maxEventNumber= 1, hazard=haz1}
discharge = {type=event, maxEventNumber=1, hazard=haz2}

OUTPUT:
output={SOFA, death, discharge}
```

## Mlxtran file

```
<DATAFILE>

[FILEINFO]
file = './data.txt'
delimiter = space
header = {id, Time, Obs, Obs_id, age}

[CONTENT]
id = {use=identifier}
Time = {use=time}
Obs = {use=observation, name={y1, y2, y3}, yname={'1', '2', '3'}, type={continuous, event, event}}
Obs_id = {use=observationtype}
age = {use=covariate, type=categorical}

<MODEL>

[COVARIATE]
input = age

age = {type=categorical, categories={1, 2, 3}}

[INDIVIDUAL]
input = {beta0_pop, omega_beta0, b1_pop, omega_b1, b2_pop, omega_b2, a_pop, omega_a, tlag_pop, omega_tlag, h1_pop, h2_pop,
        alpha1_pop, alpha2_pop, h_cov1_pop, h_cov2_pop, age, beta_h_cov1_age_2, beta_h_cov1_age_3, beta_h_cov2_age_2, beta_h_cov2_age_3}

age = {type=categorical, categories={1, 2, 3}}

DEFINITION:
beta0 = {distribution=normal, typical=beta0_pop, sd=omega_beta0}
b1 = {distribution=normal, typical=b1_pop, sd=omega_b1}
b2 = {distribution=normal, typical=b2_pop, sd=omega_b2}
a = {distribution=logNormal, typical=a_pop, sd=omega_a}
tlag = {distribution=normal, typical=tlag_pop, sd=omega_tlag}
h1 = {distribution=logNormal, typical=h1_pop, no-variability}
h2 = {distribution=logNormal, typical=h2_pop, no-variability}
alpha1 = {distribution=normal, typical=alpha1_pop, no-variability}
alpha2 = {distribution=normal, typical=alpha2_pop, no-variability}
h_cov1 = {distribution=normal, typical=h_cov1_pop, covariate=age, coefficient={0, beta_h_cov1_age_2, beta_h_cov1_age_3}, no-variability}
h_cov2 = {distribution=normal, typical=h_cov2_pop, covariate=age, coefficient={0, beta_h_cov2_age_2, beta_h_cov2_age_3}, no-variability}

[LONGITUDINAL]
input = {sigma_a, sigma_b}

file = './model.txt'

DEFINITION:
y1 = {distribution=normal, prediction=SOFA, errorModel=combined1(sigma_a, sigma_b)}

<FIT>
data = {y2, y3, y1}
model = {deces, sortie, y1}

<PARAMETER>
sigma_a = {value=1, method=MLE}
a_pop = {value=1, method=MLE}
alpha1_pop = {value=0, method=MLE}
alpha2_pop = {value=0, method=MLE}
sigma_b = {value=1, method=MLE}
b1_pop = {value=1, method=MLE}
b2_pop = {value=1, method=MLE}
beta0_pop = {value=1, method=MLE}
beta_h_cov1_age_2 = {value=0, method=MLE}
beta_h_cov1_age_3 = {value=0, method=MLE}
beta_h_cov2_age_2 = {value=0, method=MLE}
beta_h_cov2_age_3 = {value=0, method=MLE}
h1_pop = {value=1, method=MLE}
h2_pop = {value=1, method=MLE}
h_cov1_pop = {value=0, method=FIXED}
h_cov2_pop = {value=0, method=FIXED}
omega_a = {value=1, method=MLE}
omega_b1 = {value=1, method=MLE}
omega_b2 = {value=1, method=MLE}
omega_beta0 = {value=1, method=MLE}
omega_tlag = {value=1, method=MLE}
tlag_pop = {value=1, method=MLE}

<MONOLIX>
```

```
[TASKS]
populationParameters()
individualParameters(method = {conditionalMean, conditionalMode })
fim(method = StochasticApproximation)
logLikelihood(method = ImportanceSampling)
plotResult(method = {outputplot, indfits, obspred, parameterdistribution, covariatemodeldiagnosis, randomeffects,
                    {covariancemodeldiagnosis, saemresults }})

[SETTINGS]
GLOBAL:
exportpath = './'

POPULATION:
smoothingiterations = 2000
exploratoryiterations = 2000

INDIVIDUAL:
optimizationiterations = 2000

FIM:
maxiterations = 2000
```

## Supplementary Information S5 - Model assessment

### Goodness of fit plots for the longitudinal model

The following goodness of fit plots showed the adequacy of the longitudinal model.

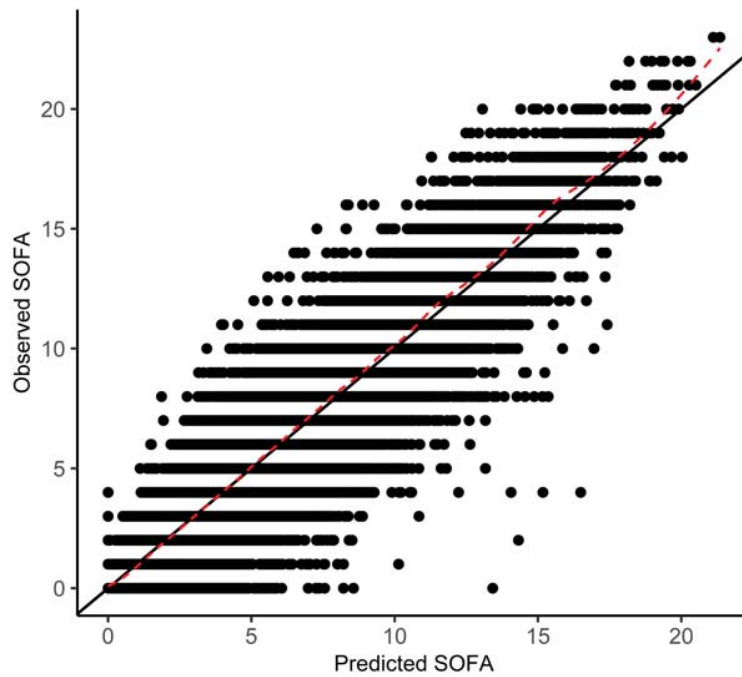


Figure A: Observations versus individual predictions of the SOFA score, on the training data set. The solid black line represents the identity line and the dashed red line the spline line.

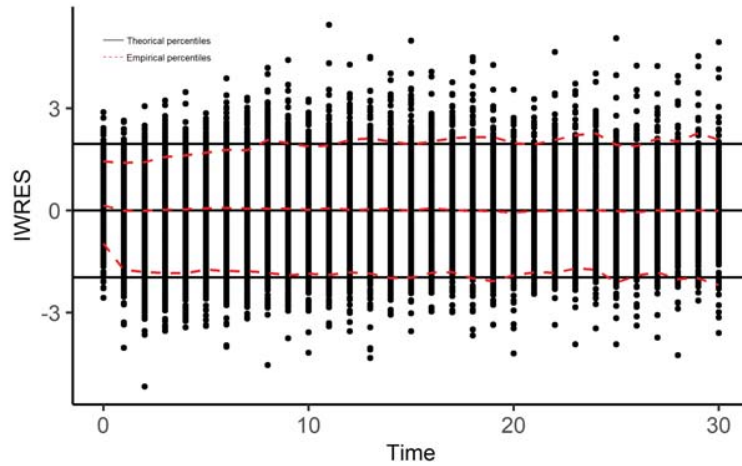


Figure B: Individual Weighted RESiduals (IWRES) versus time for SOFA score, on the training data set. The solid black lines denote respectively the 2.5th, 50th and 97.5th theoretical percentiles and the dashed red lines the corresponding empirical percentiles.

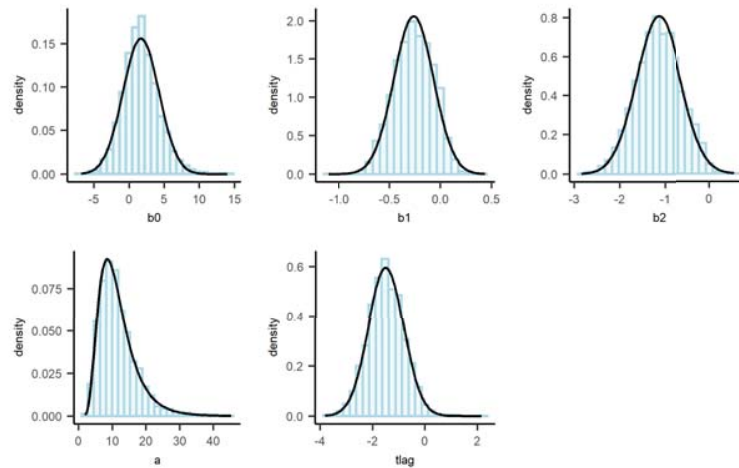


Figure C: Distribution of longitudinal individual parameters ( $b_0, b_1, b_2, a, t_{lag}$ ) of the training data set. Blue bars correspond to the histogram of the estimated individual parameters and the black curve to the distribution predicted using estimated means and variances.

### Adequacy of the joint model

To assess the adequacy of the joint model on the training data set, we computed two indicators: ROC AUC and Brier score at day 30 for both event (death and discharge), using all the information available up to day 30. We obtained very good performances: ROC AUC = 0.96 (with 95% CI [0.94,0.97]) and Brier score = 0.059 (with 95% CI [0.054,0.064]) for predicting the risk of ICU death and ROC AUC = 0.98 (with 95% CI [0.95,1.00]) and Brier score = 0.046 (with 95% CI [0.042,0.051]) for the risk of discharge.

## Supplementary Information S6 - Marginal trends

We computed median (Q1-Q3) survival predictions in the validation population for the 3 age categories and for baseline SOFA splits in 3 categories, for landmark times 0 and 6. Baseline SOFA cutoffs were chosen to have initially about one third of the patients in each category.

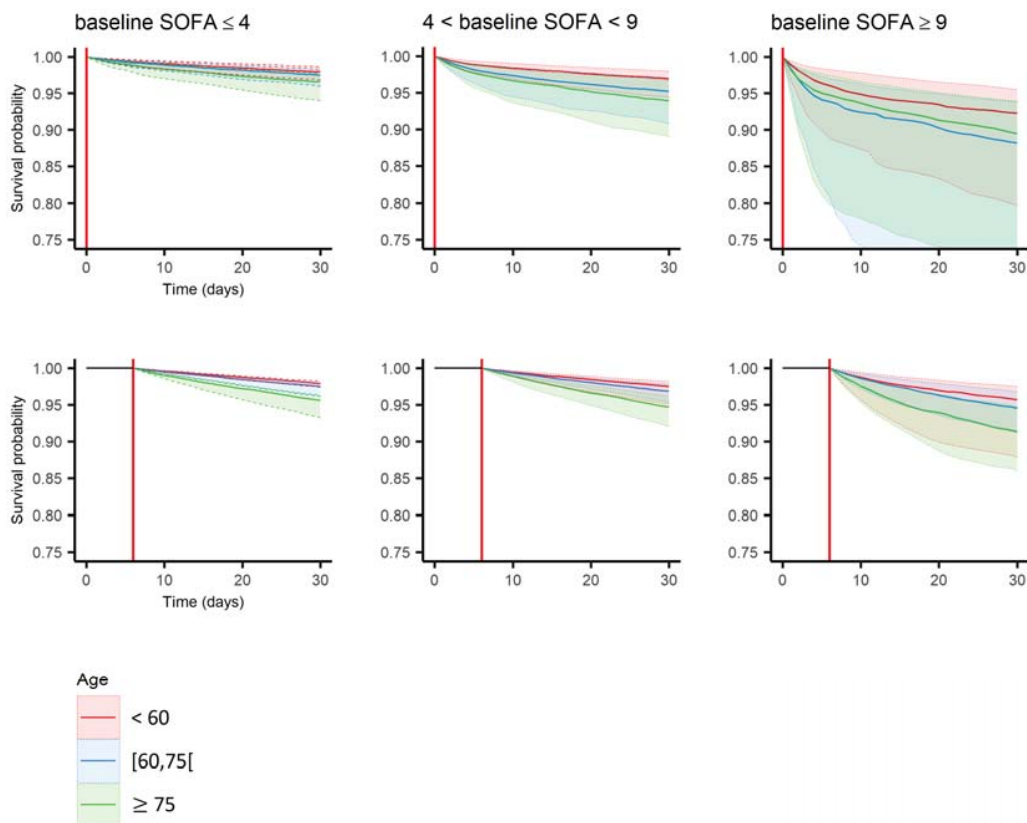


Figure: Median (line) and Q1-Q3 (bands) population survival predictions given different age categories and different baseline SOFA. Survival predictions were computed at landmark time 0 (top) and landmark time 6 (bottom) for each individual of the validation data set (number of subjects at risk in each group is given in table below).

For landmark time 0, the uncertainty on predictions is more important than for landmark time 6. For both landmark times and all the initial SOFA categories, older patients tend to have a lower survival probability than younger ones. This difference is exacerbated for high initial SOFA.

Baseline SOFA	Age (y)	Landmark 0	Landmark 6
		N	N
low ( $\leq 4$ )	<60	300	154
	$60 \leq \text{age} \leq 75$	222	131
	$\geq 75$	154	96
medium ( $4 < \text{SOFA} < 9$ )	<60	285	186
	$60 \leq \text{age} \leq 75$	241	166
	$\geq 75$	228	161
high ( $\geq 9$ )	<60	197	139
	$60 \leq \text{age} \leq 75$	219	159
	$\geq 75$	150	93
Total		1996	1285

Table: Number of at-risk patients in the validation set for landmark 0 and 6, for different age categories and different baseline SOFA.

## Supplementary information S7 - Members of the Outcome rea network

### Members of the OUTCOMEREA Study Group:

*Scientific Committee:* Jean-François Timsit (Medical and Infectious Diseases ICU, Bichat-Claude Bernard Hospital, Paris, France; UMR 1137 Inserm –Paris Diderot university IAME, F75018, Paris); Elie Azoulay (Medical ICU, Saint Louis Hospital, Paris, France); Maïté Garrouste-Orgeas (Paliative care, Institut franco Britanique, Paris, France); Jean-Ralph Zahar (Infection Control Unit, Avicenne Hospital, Bobigny, France); Bruno Mourvillier (Medical ICU, CHU Reims, France); Michael Darmon (Medical ICU, APHP Saint Louis hospital Paris, France). Biostatistical and information system expertise: Jean-François Timsit (Medical and Infectious Diseases ICU, Bichat-Claude Bernard Hospital, Paris, France; UMR 1137 Inserm –Paris Diderot University IAME, F75018, Paris); Corinne Alberti (Medical Computer Sciences and Biostatistics Department, Robert Debré Hospital, Paris, France); Stephane Ruckly (OUTCOMEREA Organization and Inserm UMR 1137 IAME, F75018, Paris); Sébastien Bailly (Grenoble University Hospital Inserm UMR 1137 IAME, F75018, Paris) and Aurélien Vannieuwenhuyze (Tourcoing, France).

*Investigators of the OUTCOMEREA database:* Christophe Adrie (ICU, CH Melun, and Physiology, Cochin Hospital, Paris, France); Carole Agasse (Medical ICU, University Hospital Nantes, France); Bernard Allaouchiche (ICU, Hospices civils de Lyon, Lyon sud, Lyon, France); Olivier Andremont (ICU, Bichat Hospital, Paris, France); Pascal Andreu (CHU Dijon, Dijon, France); Laurent Argaud (Medical ICU, Hospices Civils de Lyon, Lyon, France); Claire Arah-Somohano (Medical ICU, University Hospital, Grenoble, France); Elie Azoulay (Medical ICU, Saint Louis Hospital, Paris, France); Francois Barbier (Medical-Surgical ICU, Orleans, France); Jean-Pierre Bedos (ICU, Versailles Hospital, Le Chesnay, France); Thomas Baudry (Medial ICU, Edouard Herriot Hospital, Lyon France); Jérôme Bedel (ICU, Versailles Hospital, Le Chesnay, France); Julien Bohé (ICU, Hôpital Pierre Benite, Lyon France); Lila Bouadma (ICU, Bichat Hospital, Paris, France); Jeremy Bourenne (Réanimation des Urgences, Timone-2; APHM, Marseille, France); Noel Brule (Medical ICU, University Hospital Nantes, France); Frank Chemouni (Polyvalent ICU, Grand Hôpital de l'Est Francilien, Jossigny, France); Julien Carvelli (Réanimation des Urgences, Timone-2; APHM, Marseille, France); Elisabeth Coupez (ICU, G Montpied Hospital, Clermont-Ferrand, France); Martin Cour (Medial ICU, Edouard Heriot Hospital, Lyon France); Michael Darmon (ICU, APHP St louis, Paris France); Claire Dupuis (ICU, G Montpied Hospital, Clermont-Ferrand, France); Etienne de Montmollin (ICU, Bichat Hospital, Paris, France); Loa Dopeux (ICU, G Montpied Hospital, Clermont-Ferrand, France); Anne-Sylvie Dumenil (Antoine Béclère Hospital, Clamart, France); Jean-Marc Forel (AP HM, Medical ICU, Hôpital Nord Marseille); Marc Gainnier (Réanimation des Urgences, Timone-2; APHM, Marseille, France); Charlotte Garret (Medical ICU, University Hospital Nantes, France); Dany Goldgran-Tonedano (CH le Raincy-Montfermeil; France); Steven Grangé (ICU, CHU Rouen, France); Antoine Gros (ICU, Versailles Hospital, Versailles, France); Hédia Hamed (CH le Raincy-Montfermeil); Akim Haouache (Surgical ICU, H Mondor Hospital, Creteil, France); Tarik Hissem (ICU, Eaubonne, France); Vivien Hong Tuan Ha (ICU, CH Meaux, France); Sébastien Jochmans (ICU, CH Melun); Jean-Baptiste Joffredo (ICU, G Montpied Hospital, Clermont-Ferrand, France); Hatem Kallel (ICU, Cayenne General Hospital, Cayenne, France); Guillaume Lacave (ICU, Versailles Hospital, Le Chesnay, France); Virgine Laurent (ICU, Versailles Hospital, Le Chesnay, France); Alexandre Lautrette (ICU, G Montpied Hospital, Clermont-Ferrand, France); Clément Le Bihan (ICU, Bichat Hospital, Paris, France); Virgine Lemiale (Medical ICU, Saint Louis Hospital, Paris, France); David Luis (Médecine Intensive et Réanimation, CH Simone Veil, Beauvais, France); Guillaume Marcotte (Surgical ICU, Hospices Civils de Lyon, Lyon, France); Jordane Lebut (ICU, Bichat Hospital, Paris, France); Bruno Mourvillier (ICU, CHU Reims, Reims, France); Mathild Neuville (ICU, Foch Hospital, Paris, France); Laurent Nicolet (Medical ICU, University Hospital Nantes, France); Johanna Oziel (Medico-Surgical ICU, Hôpital Avicenne APHP, Bobigny, France); Laurent Papazian (Hopital Nord, Marseille, France); Juliette Patrier (ICU, Bichat Hospital, Paris, France); Benjamin Planquette (Pulmonology ICU, George Pompidou Hospital Hospital, Paris, France); Aguila Radjou (ICU, Bichat Hospital, Paris, France); Marie Simon (Medial ICU, Edouard Heriot Hospital, Lyon France); Romain Sonneviller (ICU, Bichat Hospital, Paris, France); Jean Reignier (Medical ICU, University Hospital Nantes, France); Bertrand Souweine (ICU, G Montpied Hospital, Clermont-Ferrand, France); Carole Schwebel (ICU, A Michallon Hospital, Grenoble, France); Shidasp Siami (ICU, Etampes Hospital, France); Nicolas Terzi (ICU, A Michallon Hospital, Grenoble, France); Gilles Troché (ICU, André Mignot Hospital, Le Chesnay, France); Fabrice Thiollieres (ICU, Hospices civils de Lyon, Lyon sud, Lyon, France); Guillaume Thierry (ICU, St Etienne, France); Guillaume Van Der Meersch (Medical-Surgical ICU, University Hospital Avicenne); Marion Venot (Medical ICU, Saint Louis Hospital,

Paris, France); Florent Wallet (ICU, Hospices Civils de Lyon, Lyon sud, Lyon, France); Sondes Yaacoubi (CH le Raincy-Montfermeil); Olivier Zambon (Medical ICU, University Hospital Nantes, France); Jonathan Zarka (Reanimation Polyvalente, Centre Hospitalier de Marne la Vallée, France).

*Study Monitors:* Mireille Adda, Vanessa Vindrieux, Marion Provent, Sylvie de la Salle, Pauline Enguerrand, Vincent Gobert, Stéphane Guessens, Helene Merle, Nadira Kaddour, Boris Berthe, Samir Bekkhouche, Kaouttar Mellouk, Méline Lebrazic, Carole Ouisse, Diane Maugars, Christelle Aparicio, Igor Theodose, Manal Nouacer, Veronique Deiler, Fariza Lamara, Myriam Moussa, Atika Mouaci, Nassima Viguiet.

## Supplementary Materials - Predictive performances of the baseline model

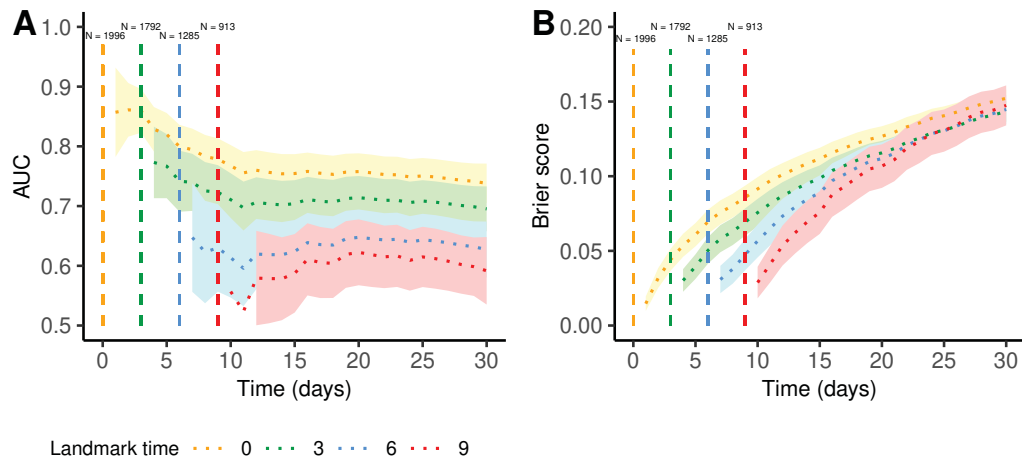


Figure S1: Time dependant AUC (A) and Brier score (B) within the baseline model computed on the validation set depending on landmark time. 95% confidence intervals are represented by the corresponding coloured intervals. Number of at-risk patients at each landmark time is specified on the top.

# Supplementary Materials - Simulation study: relative errors on parameter estimation

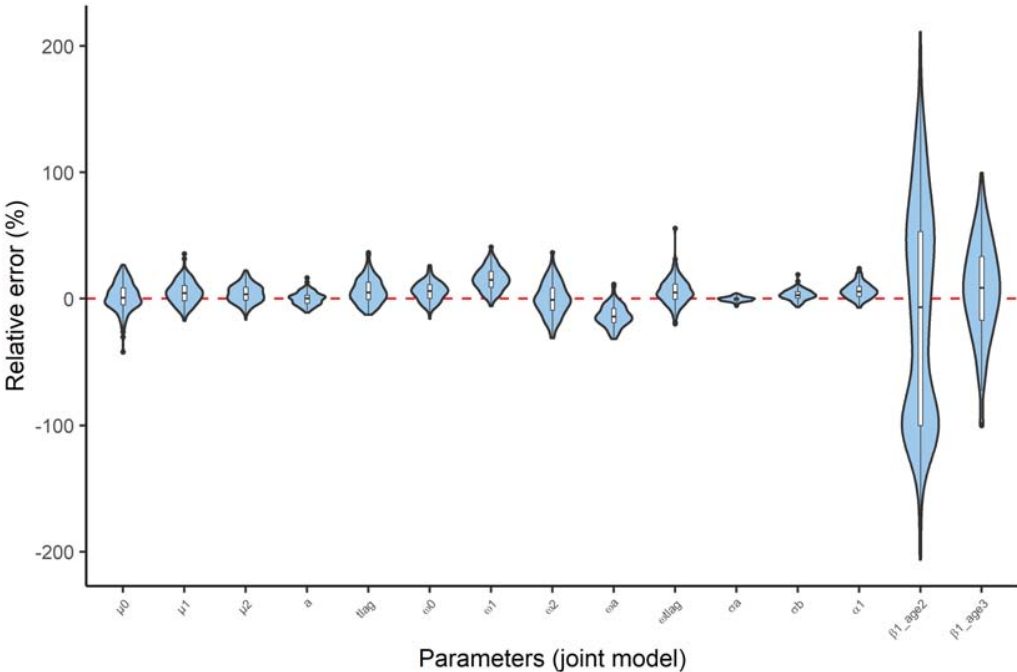


Figure S2 A: Distribution of relative errors of parameter estimates of the joint model (%)

The violin plots demonstrate good properties of estimators, in particular for the coefficient which links the longitudinal process and the survival process. All longitudinal parameters are accurate and precisely estimated too. However, the variability on parameters associated with baseline covariate is high, although the bias is closed to zero.

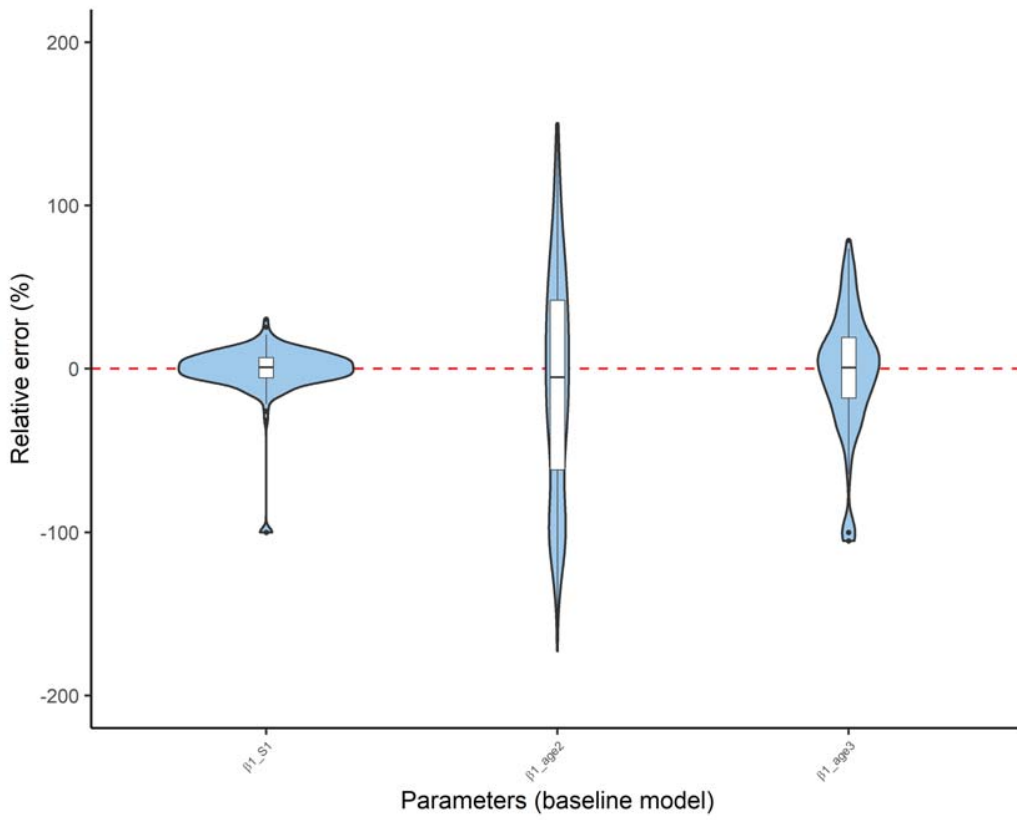


Figure S2 B: Distribution of relative errors of parameter estimates of the baseline model (%)

## Supplementary Materials - BIC for different error models (joint model)

Table S1: Bayesian information criteria (BIC) for different error models tested within the joint model

<b>Error model</b>	<b>BIC*</b>
constant	217 596
proportional	217 703
combined	216 819

\* Computed on the training data set

## Supplementary Materials - BIC for different models including baseline SOFA score and some covariates

Table S2 : Bayesian information criteria (BIC) for different joint models including the baseline SOFA and some baseline covariates in the survival part.

<b>Covariates included</b>	<b>BIC*</b>
$\emptyset$	31 327
Age	31 273
Gender	31 338
Comorbidities	31 332
Age + Gender	31 284
Age + Comorbidities	31 293
Gender + Comorbidities	31 344
Age + Gender + Comorbidities	31 304

\*Computed on the training data set.

## Supplementary Materials - Baseline model fit

Table S3 : Parameter estimates of the baseline model (with baseline SOFA and age at admission)

Parameter	Value	S.E.	RSE (%)	p-value
<i>Death</i>				
$h_1$	0.0012	0.00013	11	
$\beta_{S_1}$	0.20	0.0093	5	$< 10^{-16}$
$\beta_{1, Age_{[60,75]}}$	0.34	0.091	27	0.0002
$\beta_{1, Age_{\geq 75}}$	0.61	0.089	15	$7 \times 10^{-12}$
<i>Discharge</i>				
$h_2$	0.13	0.0053	4	
$\beta_{S_2}$	-0.12	0.0052	4	$< 10^{-16}$
$\beta_{2, Age_{[60,75]}}$	-0.21	0.044	20	$2 \times 10^{-6}$
$\beta_{2, Age_{\geq 75}}$	-0.26	0.046	18	$2 \times 10^{-8}$

---

### Modèle joint multivarié avec risques compétitifs pour prédire la survie des patients hospitalisés pour infection au SARS-CoV-2

---

#### 3.1 Résumé

Dans ce travail, nous avons étendu la méthodologie développée au Chapitre 2 au cas multivarié où plusieurs biomarqueurs longitudinaux sont mesurés. Pour cela, nous avons considéré les données de la cohorte RisCoV présentée en Section 1.3.2 où 327 patients sont hospitalisés pour infection au SARS-CoV-2 avec des mesures irrégulières de plusieurs biomarqueurs. L'objectif est de proposer puis de valider une méthode de sélection de biomarqueurs pour construire un modèle joint multivarié avec risques compétitifs pour prédire la survie des patients hospitalisés.

Dans une étape univariée, chaque biomarqueur est modélisé par un modèle joint composé d'un modèle linéaire ou non-linéaire à effets mixtes et d'un modèle de survie avec risques compétitifs impliquant des fonctions de sous-distribution (pour le risque instantané de décès et de sortie d'hôpital). Le modèle non-linéaire est semblable à celui proposé dans le Chapitre 2. Les risques de sous-distribution impliquent la valeur courante du bio-

marqueur et sont ajustés sur la valeur du score 4C. Les biomarqueurs sont ensuite choisis en fonction de leur qualité de modélisation et de la significativité de leur lien entre les parties longitudinale et de survie. Sept biomarqueurs ont ainsi été sélectionnés pour une étape multivariée, où une stratégie de sélection backward basée sur la significativité du coefficient de lien a été appliquée. L'estimation a été réalisée avec le logiciel Monolix. Nous avons notamment quantifié, à l'aide de l'AUC temps-dépendante, les performances pronostiques de la combinaison de trois biomarqueurs sélectionnés par la méthode énoncée ci-dessus : la numération des polynucléaires neutrophiles sanguins, le pH, et la protéine C-réactive. Nous avons comparé ces performances avec celles d'un modèle qui ne considère aucune information longitudinale, en modélisant uniquement le score 4C disponible à l'admission. Ce modèle est appelé le modèle baseline, en analogie au précédent travail. En particulier, la modélisation conjointe témoigne de meilleures performances lorsque suffisamment de mesures sont disponibles (pour un temps landmark = J6 et un horizon = J30, AUC [IC 95%] = 0.73 [0.65,0.81] et 0.81 [0.73,0.89] pour le modèle baseline et joint multivarié respectivement,  $p = 0.04$ ).

Nous avons conduit une étude de simulation pour valider (i) les capacités de l'algorithme SAEM de Monolix à fournir des estimations non biaisées des paramètres dans ce contexte, (ii) la stratégie de sélection proposée, et (iii) les propriétés de l'indicateur de l'AUC. Pour cela, 100 jeux de données de 300 patients ont été simulés, avec un suivi de 7 biomarqueurs ( $bm_1$  à  $bm_7$ ). Les valeurs utilisées pour simuler les données se sont inspirées des estimations trouvées sur le vrai jeu de données. Les biomarqueurs  $bm_1$ ,  $bm_2$ ,  $bm_3$  ont été associés au risque de décès,  $bm_4$  et  $bm_5$  n'ont pas été associés mais avaient la pente de leur modèle longitudinal corrélée à celle de  $bm_2$  et  $bm_3$  respectivement, et  $bm_6$  et  $bm_7$  n'ont été ni associés ni corrélés. Pour valider les performances d'estimation, le vrai modèle (incluant  $bm_1$ ,  $bm_2$  et  $bm_3$ ) a été estimé sur les jeux de données simulés et les biais relatifs (RB), les erreurs quadratiques moyennes relatives (RRMSE) et les erreurs relatives d'estimation (REE) ont été calculés. Pour valider la stratégie de sélection, la stratégie backward a été appliquée et l'ensemble final de biomarqueurs a été reporté pour chaque

jeu de données simulé. Finalement, pour évaluer l'indicateur de l'AUC, les performances prédictives (évaluées par l'AUC pour un temps landmark et un horizon donnés) obtenues à partir des estimations des vrais modèles ont été comparées à celles obtenues à partir des estimations des modèles finaux reportés après la backward.

L'algorithme SAEM de Monolix a ainsi été montré comme performant pour l'estimation de ces modèles, avec toutefois des estimateurs moins précis que dans le cadre de données riches comme dans le précédent travail. La stratégie backward a montré de bonnes capacités à sélectionner les biomarqueurs réellement associés (modèles finaux incluant au moins  $bm_1$ ,  $bm_2$  et  $bm_3$  dans 75% des cas). Enfin, l'indicateur de l'AUC rapporte des performances prédictives légèrement réduites lorsque le modèle final ne correspond pas au modèle ayant servi à simuler les données (vrai modèle).

La méthode de sélection proposée peut être appliquée dans de nombreuses autres applications où plusieurs biomarqueurs sont mesurés au cours du suivi. Toutefois, une tendance exponentielle des temps de calcul en fonction du nombre de biomarqueurs inclus dans le modèle joint multivarié a été soulignée. Cette tendance est majoritairement portée par l'algorithme estimant les erreurs standard des paramètres, très coûteux d'un point de vue computationnel lorsque le nombre de sous-modèles longitudinaux est important. La modélisation pourrait également être étendue pour considérer une sélection de type LASSO, connue pour avoir de meilleures propriétés que les sélections de type forward/backward.

Ce travail a fait l'objet d'un article accepté pour publication dans *Biometrical Journal* le 26 juillet 2023, sous condition que le code fourni soit validé par l'équipe de reproduction scientifique.

## 3.2 Article 2 (accepté)

**ARTICLE TYPE**

# Multivariate joint model under competing risks to predict death of hospitalized patients for SARS-CoV-2 infection

Alexandra Lavalley-Morelle\*<sup>1</sup> | Nathan Peiffer-Smadja<sup>1,2</sup> | Simon B. Gressens<sup>2</sup> | Bérénice Souhail<sup>2</sup> | Alexandre Lahens<sup>2</sup> | Agathe Bounhiol<sup>2</sup> | François-Xavier Lescure<sup>1,2</sup> | France Mentré<sup>1,3</sup> | Jimmy Mullaert<sup>1,3</sup>

<sup>1</sup>Université Paris Cité, INSERM, IAME, F-75018 Paris, France

<sup>2</sup>Department of Infectious and Tropical Diseases, AP-HP, Bichat-Claude Bernard University Hospital, F-75018 Paris, France

<sup>3</sup>Department of Epidemiology, Biostatistics and Clinical Research, AP-HP, Bichat-Claude Bernard University Hospital, F-75018 Paris, France

**Correspondence**

\*Alexandra Lavalley-Morelle,  
Corresponding address. Email:  
alexandra.lavalley-morelle@inserm.fr

**Abstract**

During the COVID-19 pandemic, several clinical prognostic scores have been proposed and evaluated in hospitalized patients, relying on variables available at admission. However, capturing data collected from the longitudinal follow-up of patients during hospitalization may improve prediction accuracy of a clinical outcome. To answer this question, 327 patients diagnosed with COVID-19 and hospitalized in an academic French hospital between January and July 2020 are included in the analysis. Up to 59 biomarkers were measured from the patient admission to the time to death or discharge from hospital. We consider a joint model with multiple linear or non-linear mixed-effects models for biomarkers evolution, and a competing risks model involving subdistribution hazard functions for the risks of death and discharge. The links are modeled by shared random effects, and the selection of the biomarkers is mainly based on the significance of the link between the longitudinal and the survival part. Three biomarkers are retained: the blood neutrophil counts, the arterial pH and the C-reactive protein. The predictive performances of the model are evaluated with the time-dependent AUC for different landmark and horizon times, and compared with those obtained from a baseline model, that only consider information available at admission. The joint modeling approach helps to improve predictions when sufficient information is available. For landmark 6 days and horizon of 30 days, we obtain AUC [95% CI] 0.64 [0.55, 0.74] and 0.84 [0.75,0.93] for the baseline and the joint model respectively ( $p=0.04$ ). Statistical inference is validated through a simulation study.

**KEYWORDS:**

biomarkers, competing risks, COVID-19, dynamic predictions, multivariate joint model

## 1 | INTRODUCTION

The pandemic of coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), rapidly infected millions of people around the world, causing saturation of intensive care units and emergency

<sup>0</sup>**Abbreviations:** ROC AUC, receiver operator characteristic area under the curve; RSE, relative standard error

departments. In this context, personalized predictions of the survival of hospitalized patients can be useful to guide therapeutic management (e.g. escalation or limitation of care) and to forecast hospital needs (beds, staff,...).

Since 2020, many prognostic scores (Knight et al., 2020; Liang et al., 2020; Myrstad et al., 2020) using information available at hospital admission (e.g. from medical history, clinical presentation or biological measurements) have been developed to quantify the risk of in-hospital mortality for COVID-19 patients (Tjendra et al., 2020; Zhang et al., 2020; Zhu et al., 2020). These scores ignore the biological evolution during hospitalization, and attempts to use this information remain scarce. Some studies quantified the prognostic effect of an early change in biomarker measurements (e.g. between two days (Lavillegrand et al., 2021; Li et al., 2021; Mueller et al., 2020)). However, when biomarkers evolve over time, incorporating the full follow-up in a statistical model may be useful to better predict disease prognostic. This can be achieved by jointly estimating a (possibly nonlinear) longitudinal mixed model describing the biomarkers evolution at individual level, and a survival model describing hazard evolution. These models are usually linked by shared random effects that models the association between the individual biomarkers evolution and their translation into hazard (Rizopoulos, 2011). Joint modeling also allows to derive individual dynamic predictions and recent examples of application are numerous (Desmée et al., 2017; Keroui et al., 2020; Lavalley-Morelle et al., 2022).

Nowadays, most hospitals are equipped with laboratory information systems that routinely gather all results of biological analyses prescribed by the physicians. During a hospital stay, consecutive clinical observations and biological results represent a huge amount of information that can be used in a joint model to provide individual dynamic predictions. Of note, when studying in-hospital death as a time-to-event outcome, patient discharge should be taken into account as a competing event. Shifting the modeling from one single biomarker to potentially hundreds could certainly improve prediction accuracy, but also comes with statistical issues. First, Shen et al. (Shen & Li, 2021) well described computational and identifiability issues due to the high number of random effects when multiple correlated longitudinal models are jointly estimated with a survival model. This probably explains why published models are mostly limited to longitudinal models with two biomarkers (Andrinopoulou et al., 2017; Long & Mills, 2018; Rajeswaran et al., 2018), and highlights the need of methods for variable selection in this context. Recently, a study (Tong-Minh et al., 2022) evaluated the follow-up of four immune biomarkers (C-reactive protein, procalcitonin, interleukin-6, and soluble urokinase-type plasminogen activator receptor) to predict the risk of in-hospital death in COVID-19 patients. This study provided promising results, but the involved biomarkers were selected based on clinical considerations and no statistical selection procedure is described. In a high dimensional setting where biomarkers correspond to gene expression levels, Liu et al (Liu et al., 2019) proposed a factor analysis model to summarize the co-evolution of gene expressions. Again, this approach makes the estimation computationally tractable, but does not perform marker selection since the loading matrix is full. In order to efficiently leverage information from hospital information systems, it appears that variable selection represents a second methodological challenge. While variable selection procedures are known for standard regression models, applications in joint models in the presence of competing risks are, to our knowledge, lacking.

The objective of the present work is twofold. Taking advantage on real data of patients hospitalized for SARS-CoV-2 infection, the first objective is to propose an original selection strategy to build a multivariate joint model with a subset of biomarkers most associated with the risk of in-hospital death and discharge from hospital. The second objective is to compare the predictive performances of such a model against a model that only consider baseline information to show the added value of the proposed strategy. This article is organized as follows: section 2 introduces the analyzed data; section 3 presents methodological aspects about uni- and multivariate joint modeling; section 4 details the methodology to perform dynamic predictions and the performances assessment; section 5 shows the results and section 6 provides a simulation study to support the methodology. Finally, section 7 is devoted to the discussion.

## 2 | DATA

The study is based on real-life hospital data, obtained from  $N = 327$  patients hospitalized in the Department of Infectious and Tropical Diseases in an academic hospital (Bichat, France) during the first wave of the COVID-19 pandemic (January to July 2020). The project has been approved by the local ethic committee (IRB number 00006477). A manual data collection allowed to collect clinical variables relative to the patients: baseline characteristics (age, gender, ...). The primary outcome is the time until in-hospital death within 30 days. Discharge is considered as a competing event. Figure 1 displays the cumulative incidence functions for the in-hospital death and the discharge from hospital. Thirty days after admission, about 14% of the patient died, while 72% were discharged. The censored observations correspond to hospital transfer during the stay or the end of the observation study (at day 30).

An automated extraction from the hospital data warehouse allowed to obtain all results of 59 biological exams prescribed during the hospitalization, up to day 30 or to the time to death or discharge from hospital. The availability and the frequency of the biomarker measurements differ between patients. These biomarkers were classified into 8 categories: complete blood count, coagulation, pulmonary functions, urine samples, kidney functions/cellular lysis, liver/pancreatic functions, markers of inflammation and cardiac markers. Supplementary Table S1 shows the complete list description of these biomarkers.

In this study, we consider a baseline score adapted from the 4C-score (Knight et al., 2020). This score was developed in 2020 to predict the in-hospital mortality for COVID-19 patients and combines 8 variables (age, sex at birth, number of comorbidities, respiratory rate, peripheral oxygen saturation on air room, Glasgow coma scale, urea and C-reactive protein (CRP) at admission) with different weights (see Supplementary Table S2). However, since the collected data does not include respiratory rate, peripheral oxygen saturation on room air and Glasgow coma scale, we set to 0 the corresponding weights. Indeed, because the patients are hospitalized in a medical ward and not in intensive care unit (ICU), these items are likely to be closed to 0. For baseline urea and CRP values, we used the first available result within 48 hours. The remaining missing values are imputed with the most frequently encountered value knowing the other characteristics of the score (age, sex and number of comorbidities). Actually, it affects few individuals (4% and 8% of the patients for the urea and CRP respectively). Table 1 details the score components of the analyzed patients. The median [Q1-Q3] score was 6 [4-9] at admission.

### 3 | BUILDING OF A JOINT MODEL WITH MULTIPLE BIOMARKERS

The main idea of this building process is to model each biomarker individually with a joint model, select the biomarkers properly fitted and among them, keep the ones which evolution is significantly associated with the risk of in-hospital death. Then, the multivariate analysis is performed including the selected biomarkers.

#### 3.1 | Univariate modeling

##### 3.1.1 | General univariate model

Let  $T$  the variable describing the time-to-event distribution,  $C$  the non-informative censoring distribution and  $\delta$  the event indicator (1 for in-hospital death, 2 for discharge and 0 for censoring). The couple  $(\tilde{T}, \delta)$  is observed with  $\tilde{T} = \min(T, C)$ . We consider the subdistribution hazard framework to model the competing risks. In that case, we define  $T_e = T \times \mathbb{1}_{\delta=e} + \infty \times \mathbb{1}_{\delta \neq e}$  for  $e = \{1, 2\}$ .

Let  $N$  be the total number of analyzed patients, and  $K$  the number of biomarkers. Let  $N_k$  be the number of subjects who have at least one measurement for biomarker  $k$  and  $\{y_{i1k}, \dots, y_{in_{ik}k}\}$  the vector of longitudinal observations of subject  $i$  (for  $i = 1, \dots, N_k$ ) for this biomarker. Hence, observation  $y_{ijk}$  denotes the  $j^{\text{th}}$  measurement of patient  $i$  for biomarker  $k$  at time  $t_{ijk}$  (for  $j = 0, \dots, n_{ik}$ ). Let  $m_k$  be the function describing the structural model of biomarker  $k$ . Let also  $\lambda_{1ik}$  and  $\lambda_{2ik}$  be the instantaneous risks of in-hospital death and discharge respectively, associated to individual  $i$  and biomarker  $k$ . We model each biomarker  $k$ , with a joint model defined as follows:

$$\begin{aligned} y_{ijk} &= m_k(t_{ijk}, \psi_{ik}) + g[m_k(t_{ijk}, \psi_{ik}), \sigma_k] \epsilon_{ij} \\ \lambda_{1ik}(t|\psi_{ik}, \text{Score}_i) &= h_{1k} \exp \left[ \alpha_{1k} \times \left( m_k(t, \psi_{ik}) - \text{med}_k \right) + \beta_{1k} \cdot \text{Score}_i \right] \\ \lambda_{2ik}(t|\psi_{ik}, \text{Score}_i) &= h_{2k} \exp \left[ \alpha_{2k} \times \left( m_k(t, \psi_{ik}) - \text{med}_k \right) + \beta_{2k} \cdot \text{Score}_i \right] \end{aligned} \quad (1)$$

where  $\psi_{ik} = f(\mu_k, \eta_{ik})$  denotes the individual parameters associated to biomarker  $k$  expressed as a function of fixed effects, noted  $\mu_k$  and individual random effects, noted  $\eta_{ik}$ . The random effects are assumed to be normally distributed with mean 0 and variance-covariance matrix  $\Omega_k = \text{diag}(\omega_k^2)$ , and independent of the residual Gaussian error, noted  $\epsilon_{ij}$ , of mean 0 and variance 1.  $\sigma_k$  denotes the vector of the error model parameters.  $\text{Score}_i$  denotes the baseline score associated to patient  $i$ .  $h_{1k}$  and  $h_{2k}$  are the baseline constant hazards for the risk of in-hospital death and the risk of discharge.  $\alpha_{1k}$  and  $\alpha_{2k}$  are the coefficients which link the current predicted value of the biomarker  $k$  with the instantaneous risks of events. We only considered a linear link between the longitudinal process and the subdistribution risks. The median value of all observations for biomarker  $k$  is denoted  $\text{med}_k$  and appears in the model in order to avoid numerical issues during estimation. Finally,  $\beta_{1k}$  and  $\beta_{2k}$  are the coefficients associated to the baseline score effects on both instantaneous risks. We emphasize that under a joint modeling approach, it is not necessary

to know the value of the biomarker after the occurrence of the competing event since only the predictions given the structural model are linked with both hazards.

The structural model  $m_k$  describing the marker observations can be (1) a linear mixed-effects model or (2) a nonlinear mixed-effects model which is more flexible for describing non-monotonic trend, but also less parsimonious. In the case of linear joint models, the function  $m_k$  is defined as follows :

$$m_k(t_{ijk}, \psi_{ik}) = b_{0ik} + b_{1ik} \times t_{ijk} \quad (2)$$

and for nonlinear joint models, the function  $m$  writes :

$$m_k(t_{ijk}, \psi_{ik}) = b_{0ik} + a_{ik} \times \left( \exp \left( b_{1ik} \times (t_{ijk} - t_{lag_{ik}}) \right) - \exp \left( b_{2ik} \times (t_{ijk} - t_{lag_{ik}}) \right) \right) \quad (3)$$

And we have,

$$\begin{cases} b_{0ik} = \mu_{0k} + \eta_{0ik} \\ b_{1ik} = \mu_{1k} + \eta_{1ik} \\ b_{2ik} = \mu_{2k} + \eta_{2ik} \\ a_{ik} = \mu_{ak} \exp(\eta_{aik}) \\ t_{lag_{ik}} = \mu_{lag_k} + \eta_{t_{lag_{ik}}} \end{cases}$$

For a joint model associated to biomarker  $k$ , the joint likelihood writes:

$$L(\theta_k) = \prod_{i=1}^N \int_{\eta_{ik}} \left[ \left( \prod_{j=0}^{n_{ik}} p(y_{ijk} | \eta_{ik}, Score_i; \theta_k) \right) p(\tilde{T}_i, \delta_i | \eta_{ik}; \theta_k) p(\eta_{ik}; \theta_k) \right] d\eta_{ik} \quad (4)$$

where  $\theta_k = \{\mu_k, \Omega_k, \sigma_k, h_{1k}, h_{2k}, \alpha_{1k}, \alpha_{2k}, \beta_{1k}, \beta_{2k}\}$  is the vector of parameters to estimate.  $p(y_{ijk} | \eta_{ik}, Score_i; \theta_k)$  is the density function of longitudinal process associated to biomarker  $k$ ,  $p(\eta_{ik}; \theta_k)$  the density function of random effects and  $p(t, \delta | \eta_{ik}; \theta_k)$  is the density function of the survival process given by (Jeong & Fine, 2006):

$$\begin{aligned} p(t, \delta | \eta_{ik}; \theta_k) &= [\lambda_{1ik}(t | \psi_{ik}, Score_i; \theta_k) \times (1 - F_{1ik}(t | \psi_{ik}, Score_i; \theta_k))]^{\mathbb{1}_{\delta=1}} \\ &\times [\lambda_{2ik}(t | \psi_{ik}, Score_i; \theta_k) \times (1 - F_{2ik}(t | \psi_{ik}, Score_i; \theta_k))]^{\mathbb{1}_{\delta=2}} \\ &\times (1 - F_{1ik}(t | \psi_{ik}, W_i; \theta_k) - F_{2ik}(t | \psi_{ik}, Score_i; \theta_k))^{\mathbb{1}_{\delta=0}} \end{aligned} \quad (5)$$

where, for  $e = \{1, 2\}$  :

$F_{eik}(t | \psi_{ik}, Score_i; \theta_k) = 1 - \exp \left( - \int_0^t \lambda_{eik}(u | \psi_{ik}, Score_i; \theta_k) du \right)$  is the individual cumulative incidence function for event  $e$  at time  $t$ .

The vector of population parameters  $\theta$  can be estimated by maximization of the likelihood. However, to the best of our knowledge, no available R package can handle this likelihood expression. Therefore, we used the SAEM algorithm (Delyon et al., 1999) implemented on Monolix software version 2018R2, as in Lavalley-Morelle et al., 2022. Its variance-covariance matrix  $\Sigma_k$  is estimated by inversion of the Fisher Information Matrix (FIM), computed with a stochastic approximation using the Louis method (Louis, 1982). To ensure better estimation properties, the number of SAEM algorithm chains is fixed at 10 and the maximum number of iterations for both exploratory phase and stochastic approximation is set at 1000, keeping the default auto-stop criteria when convergence is reached. A previous work (Lavalley-Morelle et al., 2022) showed that Monolix software is able to provide unbiased and accurate estimates of parameters of a competing risk joint model in this context.

### 3.1.2 | Selection of biomarkers after univariate modeling

#### Linear submodel evaluation

We started by modeling each biomarker individually with the linear joint model for sake of simplicity. The initial values for the algorithm are set to the estimates obtained by modeling a "no-link" model (ie, with the link coefficients  $\alpha_{1k}$  and  $\alpha_{2k}$  set to 0). A combined error is considered at this step (ie,  $g[m_k(t_{ijk}, \psi_{ik}), \sigma_k] = \sigma_{ak} + \sigma_{bk} \times m_k(t_{ijk}, \psi_{ik})$ ). After fitting  $K = 59$  linear joint models, we obtain  $\hat{\theta}_k$  the estimate of  $\theta_k$ .

#### Error model selection

The following approach was used to select the error model associated to biomarker  $k$ : if  $\hat{\sigma}_{bk} \leq 10^{-4}$ , an additive error is considered for the biomarker ( $g[m_k(t_{ijk}, \psi_{ik}), \sigma_k] = \sigma_{ak}$ ), if  $\hat{\sigma}_{ak} \leq 10^{-4}$ , a proportional error is considered for the biomarker

$(g[m_k(t_{ijk}, \psi_{ik}), \sigma_k]) = \sigma_{bk} \times m_k(t_{ijk}, \psi_{ik})$ ). Otherwise, the error model which minimizes the BIC criteria is considered. The models are then re-fitted with the correct error function, and we define:

$$e\hat{r}_k = \begin{cases} \frac{\hat{\sigma}_{ak}}{med_k} & \text{if the biomarker } k \text{ has an additive error} \\ \hat{\sigma}_{bk} & \text{if the biomarker } k \text{ has a proportional error} \\ \frac{\hat{\sigma}_{ak}}{med_k} + \hat{\sigma}_{bk} & \text{if the biomarker } k \text{ has a combined error} \end{cases}$$

### Fit criteria

Biomarkers with a small measurement error ( $e\hat{r}_k \leq 25\%$ ) and with all relative standard errors (RSE) relative to the longitudinal parameters under 100% are kept with the linear joint model for the rest of the analysis. Biomarkers with a small measurement error ( $e\hat{r}_k \leq 25\%$ ) but with at least one RSE of a longitudinal parameter over 100%, are excluded from the analysis. Finally, biomarkers with a larger measurement error ( $e\hat{r}_k > 25\%$ ) are evaluated using a nonlinear longitudinal submodel.

### Nonlinear submodel

After the linear fit, some biomarkers were tested with the nonlinear submodel. First, due to potential identifiability issues,  $\mu_{t_{lagk}}$  and  $\omega_{t_{lagk}}$  are fixed to 0 if the RSE associated to parameter  $t_{lag}$  is over 100%. In fact,  $t_{lag}$  is an optional parameter which give more flexibility to the structural function, however, it can be easily removed to simplify the model. Then, same strategy is conducted for the selection and exclusion of the biomarkers but by being more flexible on the threshold for the measurement error. Hence, biomarkers with a measurement error  $e\hat{r}_k \leq 50\%$  and with all RSE of longitudinal parameters under 100% are kept with the nonlinear joint model for the rest of the analysis. Biomarkers with a measurement error  $e\hat{r}_k > 50\%$  or with at least one RSE of a longitudinal parameter over 100%, are excluded from the analysis.

### Survival submodel evaluation

For each biomarker selected either with a linear or a nonlinear longitudinal submodel, the last step was to evaluate the survival part of the model. Two criteria are considered. The first involves the good quality of survival parameter estimation. The biomarkers for which the RSE of survival parameters are below 100% are selected for the next. The others are excluded from the analysis. The second criteria involves the link between the longitudinal process and the instantaneous risk of death. We want to select biomarkers which longitudinal evolution is significantly associated with the instantaneous risk of death. To do that, we perform a Wald test on parameter  $\alpha_{1k}$  and we note  $p$  the corresponding p-value. If  $p \leq 0.05$ , the biomarker  $k$  is selected for the next; otherwise, the biomarker is excluded from the analysis.

A sensibility analysis on the thresholds was conducted. Two alternative scenarios were considered: a stringent scenario and a flexible scenario. In the stringent one, the thresholds for the measurement errors and the RSE of parameters were reduced by 40% of their values. In the flexible one, they were increased by 40% of their values. The different cases are summarized in Supplementary Table S3a.

## 3.2 | Multivariate modeling

The biomarkers selected at the univariate stage are grouped following the initial classification (given in Supplementary Table S1). For each group, the biomarker with the most significant p-value for  $\alpha_{1k}$  and which there is at least one measurement in half of the patients is considered for inclusion in the multivariate model. Thus, at most 8 biomarkers are selected for the first step of the multivariate model which contains all that biomarkers. Then a first backward selection process including the selected biomarkers is performed. At each iteration we remove the biomarker with the highest Wald test p-value for  $\alpha_{1k}$ . We stop when all p-values for the coefficients  $\alpha_{1k}$  are under 5%. If exists a biomarker for which the standard error of  $\alpha_{1k}$  has not been estimated for convergence reasons, it is also removed. Finally, a second backward selection is performed by removing the biomarker with the highest Wald test p-value for  $\alpha_{2k}$  and stop, as previously, when all p-values for the coefficients  $\alpha_{2k}$  are significant. At the end of this step,  $K'$  biomarkers which evolution are significantly associated with both instantaneous risk of death and discharge are selected.

For a given iteration where  $K''$  biomarkers are involved in the modeling, the multivariate joint model writes:

$$\begin{aligned}
y_{ij1} &= m_1(t_{ij1}, \psi_{i1}) + g[m_1(t_{ij1}, \psi_{i1}), \sigma_1] \epsilon_{ij} \\
&\dots \\
y_{ijK''} &= m_{K''}(t_{ijK''}, \psi_{iK''}) + g[m_{K''}(t_{ijK''}, \psi_{iK''}), \sigma_{K''}] \epsilon_{ij}
\end{aligned} \tag{6}$$

$$\begin{aligned}
\lambda_{1i}(t|\psi_i, Score_i) &= h_1 \exp \left[ \sum_{k=1}^{K''} \left( \alpha_{1k} \times \left( m_k(t, \psi_{ik}) - med_k \right) \right) + \beta_1 \cdot Score_i \right] \\
\lambda_{2i}(t|\psi_i, Score_i) &= h_2 \exp \left[ \sum_{k=1}^{K''} \left( \alpha_{2k} \times \left( m_k(t, \psi_{ik}) - med_k \right) \right) + \beta_2 \cdot Score_i \right]
\end{aligned}$$

with  $\psi_i = (\psi_{i1}^T, \dots, \psi_{iK''}^T)^T$ . As for the univariate model,  $\lambda_{1i}, \lambda_{2i}$  are the subdistribution hazards for the risk of in-hospital death and discharge from hospital,  $h_1, h_2$  the constant baseline risks, and  $\beta_1, \beta_2$  the baseline score effects.

Here, the likelihood writes:

$$L(\theta) = \prod_{i=1}^N \int_{\eta_i} \left[ \left( \prod_{k=1}^{K''} \prod_{j=0}^{n_{ik}} p(y_{ijk} | \eta_{ik}, Score_i; \theta) \right) p(\tilde{T}_i, \delta_i | \eta_i; \theta) p(\eta_i; \theta) \right] d\eta_i \tag{7}$$

where  $\theta = \{\mu, \Omega, \sigma, h_1, h_2, \alpha_1, \alpha_2, \beta_1, \beta_2\}$  is the vector of parameters to estimate, with  $\mu = (\mu_1^T, \dots, \mu_{K''}^T)^T$ ,  $\Omega = \text{diag}(\omega_1^2, \dots, \omega_{K''}^2)$ ,  $\sigma = (\sigma_1^T, \dots, \sigma_{K''}^T)^T$ ,  $\alpha_1 = (\alpha_{11}, \dots, \alpha_{1K''})^T$ ,  $\alpha_2 = (\alpha_{21}, \dots, \alpha_{2K''})^T$ .

Same definitions of  $p(\eta_i; \theta)$  and  $p(t, \delta | \eta_i; \theta)$  as presented in section 3.1.1 are used. Similarly to the univariate joint models, the multivariate joint models are estimated by the SAEM algorithm on Monolix software, with the same settings as previously described.

## 4 | DYNAMIC PREDICTIONS AND PERFORMANCES COMPARISON

### 4.1 | Dynamic predictions

For patient  $i$  who has longitudinal observations for biomarker  $k$  until a landmark time  $l$ :  $Y_{ik}(l) = \{y_{ijk}; 0 \leq t_{ij} \leq l\}$ , we aim at predicting its future biomarker  $k$  values  $m_{ik}(l+t|l)$  and his associated survival probability  $s_i(l+t|l)$  where  $t$  is the horizon time. Since the dynamic predictions are applied for patients who are still hospitalized at time  $l$ , we focus on the following conditional probability:

$$s_i(l+t|l) = \mathbb{P}(T_{1i} > l+t | T_i > l, Y_{ik}(l), Score_i, \theta)$$

For each landmark time  $l$ , the biomarker measurements of patient  $i$  up to time  $l$  and the information provided by the population parameters estimated on the whole data are used to compute the a posteriori distribution of the individual parameters and infer the desired quantities. Monte-Carlo process as described in literature (Lavalley-Morelle et al., 2022) is adapted to derive 200 samples of individual parameters. Briefly, the probability of not experiencing the death between  $l$  and  $l+t$ , for a replicate  $r$  and an individual  $i$ , is computed as follows :

$$s_i^{(r)}(l+t|\psi_i^{(r)}) = \frac{F_{1i}^{(r)}(\infty) + F_{2i}^{(r)}(\infty) - F_{1i}^{(r)}(l+t) - F_{2i}^{(r)}(l)}{F_{1i}^{(r)}(\infty) + F_{2i}^{(r)}(\infty) - F_{1i}^{(r)}(l) - F_{2i}^{(r)}(l)} \tag{8}$$

and estimates of  $m_{ik}(l+t|l)$ ,  $k = 1, \dots, K'$ , and  $s_i(l+t|l)$  are obtained by taking the median over 200 replicates. Prediction intervals are obtained by reporting the 2.5<sup>th</sup> and the 97.5<sup>th</sup> percentiles of the distribution.

## 4.2 | Comparison with the baseline model

For the baseline model, two subdistribution hazard functions are considered (for the risk of in-hospital death and discharge from hospital). Both are adjusted on the value of the baseline score.

$$\begin{aligned}\lambda_{1i}(t) &= h_1 \exp(\beta_1 \cdot \text{Score}_i) \\ \lambda_{2i}(t) &= h_2 \exp(\beta_2 \cdot \text{Score}_i)\end{aligned}\quad (9)$$

For both baseline and multivariate joint modeling, three landmark times are considered at 3, 6 and 9 days after patient admission. Multiple horizon times are also considered, from landmark time plus one day to 30 days after admission. The predictive performances are assessed and compared by computing the time-dependent ROC AUC for both the multivariate joint model and the baseline model. Details on computation are given in Supplementary Information S1. We consider the cumulative-dynamic definition described in previous works (Blanche et al., 2015) using the *timeROC* (Blanche et al., 2013) package available in *R* software. Finally, tests for comparing the time-dependent ROC AUC of both models is also computed using the *timeROC* package.

## 5 | RESULTS

### 5.1 | Baseline model

Table 2 presents parameter estimates for the baseline model. A significant effect of the baseline score on both risks is highlighted: a patient with a one point higher baseline score has his risk of death multiplied by 1.43 (95% CI [1.30,1.58]) and his risk of discharge divided by 1.15 (95% CI [1.11,1.20]).

### 5.2 | Uni- and multivariate joint models

Figure 2 details the selection process of the 59 biomarkers. Based on the value of the residual error of the longitudinal submodels, 31 biomarkers are considered with a linear joint model and 22 with a nonlinear joint model. For the 31 linear biomarkers, 2 are excluded because at least one RSE of a survival parameter is over 100%, and 7 excluded because the link between the longitudinal and the survival part is not significant. For the nonlinear biomarkers, 19 are excluded because at least one RSE of a longitudinal or a survival parameter is over 100%. Hence, 22 biomarkers are selected with a linear model and 3 with a nonlinear model.

Supplementary Table S4 presents those 25 biomarkers grouped following the initial classification and ranked by increasing p-value Wald test on  $\alpha_{1k}$ . Seven biomarkers are selected for the multivariate analysis: blood neutrophil counts, D-dimers, arterial pH, lactate dehydrogenase (LDH), albuminemia, C-reactive protein (CRP, in logarithmic scale) and NT-proBNP.

After backward selection, we selected 3 biomarkers significantly associated with both risks of in-hospital death and discharge: the blood neutrophil counts, the arterial pH and the CRP. Figure 3 shows the longitudinal evolution of each biomarker. Parameter estimates for the multivariate joint model are shown in Table 3. An increase of the blood neutrophil counts or the C-reactive protein, or a decrease of arterial pH is associated with a higher risk of death ( $\alpha_{1n} = 0.14$ , 95%CI [0.07, 0.20],  $\alpha_{1c} = 0.63$ , 95%CI [0.20, 1.06],  $\alpha_{1p} = -11.4$ , 95%CI [-16.93, -5.87]) and a lower risk of discharge ( $\alpha_{2n} = -0.14$ , 95%CI [-0.25,-0.03],  $\alpha_{2c} = -1.09$ , 95%CI [-1.44, -0.74],  $\alpha_{2p} = 25.2$ , 95%CI [12.75, 37.65]). Even after adjustment on the biomarker levels, the baseline score also has a significant effect on both risks: a higher baseline score is associated with a higher risk of death ( $\beta_1 = 0.33$ , 95%CI [0.21,0.45]) and a lower risk of discharge ( $\beta_2 = -0.12$ , 95%CI [-0.18,-0.05]). Some diagnostic plots that assessed the quality of the longitudinal submodels can be found in Supplementary Figure S1.

The results of uni and multivariate selection under stringent and flexible scenarios of the sensibility analysis are described in Supplementary Table S3b and Supplementary Figures S2a and S2b. When stringent thresholds are used, the interest of the backward strategy is limited because only 3 biomarkers entered in the multivariate analysis and remained in the final model. Of note, those 3 biomarkers are not selected in the current scenario. On the contrary when thresholds are too flexible, more biomarkers are involved and the final multivariate joint model retains 2 biomarkers that are also selected in the current scenario (pH and CRP). Of note, under the flexible scenario, blood neutrophil counts are modeled linearly and are removed during the backward process.

Finally, some details about computation time are given in Supplementary Information S2. We evaluated through simulations the CPU times needed to estimate parameters as well as standard errors. Estimation of both parameters and standard errors

is scalable according to the number of patients included (the estimation process will take twice as long if there are twice as many subjects). However, the relation between the CPU time spent to estimate standard errors and the number of biomarkers involved is almost exponential. Considering a too large set of biomarkers in the backward process would be inappropriate with this approach.

### 5.3 | Dynamic predictions and performances comparison

First, to illustrate the dynamic predictions, we consider two patients. Patient A, who has a baseline score of 8, was discharged at day 24. Patient B, who has a baseline score of 6, died at day 22. We first computed survival probabilities from the baseline model for these two patients at horizon time 30. We predicted a high survival probability for both patients: 90% (95%CI [87,93]) and 95% (95%CI [93,97]) for patient A and B respectively. Of note, patient B has a greater probability of discharge because of a lower baseline score. Figure 4 shows the dynamic longitudinal and survival predictions from the full multivariate joint model for the two patients, for different landmark and horizon times. For patient A, all the predictions are good, in light with the result of the baseline model and the actual outcome for this patient. For patient B, we see that his survival probability is getting worse as the landmark time increases. He finally died at time 22. In this case, the use of the full information available up to the landmark time helps to improve predictions.

To assess the prediction performances of the model, time-dependent ROC AUC are computed for the baseline model as well as the multivariate joint model. Figure 5 shows the results of this estimation. Given the small sample of failures for the times close to the landmark time, we are interested in predictions for late horizon times. The multivariate joint model has higher ROC AUC for late horizon times, compared to the baseline model. For example, for landmark time 9 and horizon time 30, AUC [95% CI] is 0.64 [0.55, 0.74], and 0.84 [0.75,0.93] for the baseline and the joint model respectively (Table 4 ). There is a significant difference in ROC AUC between both models for landmark times 6 and 9. In other words, when sufficient information is available, the multivariate joint model performs better compared to the baseline model and the benefit is higher as more information becomes available.

Of note, we also computed ROC AUC for multivariable joint model under stringent and flexible scenarios of the sensibility analysis. ROC AUC [95% CI] for landmark = D9 and horizon = D30 are respectively 0.76 [0.67,0.85] and 0.81 [0.72,0.90] for the stringent and the flexible scenario. This reduction, compared with our current scenario, is not significant (p-values are respectively 0.05 and 0.09). Details are given in Supplementary Table S3b.

Finally, to show the interest of biomarker selection, we present the results of the multivariate joint model estimation including the 7 biomarkers considered in the multivariate analysis in Supplementary Table S5. Some biomarker effects disappear after adjustments on all others and the complexity of estimation reflects a high degree of uncertainty particularly in the estimates of link coefficients. The ROC AUC is not significantly improved with the 7 biomarkers (ROC AUC [95% CI] = 0.85 [0.77, 0.93],  $p = 0.50$ ).

## 6 | SIMULATION STUDY

We performed a simulation study to answer three points. The first is to show that the SAEM algorithm implemented on Monolix software provides good estimation performances of such multivariate joint model. The second is to assess the performances of the selection process at the multivariate stage (ie, the backward process). And finally, we will assess the ROC AUC indicator under different scenarios because some authors described a possible overestimation of this indicator (Schmid et al., 2013).

### 6.1 | Simulation settings

We simulated  $M = 100$  datasets of  $N = 300$  patients assuming that  $K = 7$  biomarkers ( $bm_1$  to  $bm_7$ ) were available at most 30 days after the hospital admission. The biomarkers were simulated according to the design of the application, that is to say considering the patterns and the estimated parameters of the 7 biomarkers included in the multivariate analysis. Biomarker measurements were reported for each patient at different frequencies until time to in-hospital death or discharge from hospital or also end of the first 30 days of the stay, with no other censoring process. The table 5 summarizes those characteristics. Longitudinal submodels refer to equation 2 and 3 with  $t_{lag} = 0$ .

For the survival part, we considered two competing events. We supposed that  $bm_1$ ,  $bm_2$  and  $bm_3$  are truly associated with the event 1 (risk of in-hospital death),  $bm_4$  and  $bm_5$  are not associated but have their slopes correlated with those of  $bm_2$  and  $bm_3$  respectively, and  $bm_6$  and  $bm_7$  are not associated and not correlated with other biomarkers. A scenario with stronger correlations was also considered. In this one, we simulated  $bm_4$  and  $bm_5$  having a correlation on intercept and slope parameters of 0.95 with the ones of  $bm_2$  and  $bm_3$  respectively. We followed methods already described (Fine, 2001; Lavalley-Morelle et al., 2022) to generate failure time data. The subdistribution for the risk of in-hospital death is given by :

$$\mathbb{P}(T_{1i} < t | \psi_i, W_i; \theta) = 1 - \exp\left(-\int_0^t \frac{p_1 g_1 \exp(-g_1 \times s) \times \lambda(s)}{1 - p_1(1 - \exp(-g_1 \times s))} ds\right) \quad (10)$$

with  $\lambda(s) = \exp\left[\alpha_{11} \times (m_1(\psi_{i1}, t_{ij1}) - med_1) + \alpha_{12} \times (m_2(\psi_{i2}, t_{ij2}) - med_2) + \alpha_{13} \times (m_3(\psi_{i3}, t_{ij3}) - med_3)\right]$ . Parameters  $p_1$  and  $g_1$  are used to control the number of event 1 at infinite time and the speed of onset.

The subdistribution for the competing risk was then obtained using an exponential distribution with rate  $\frac{t}{b}$  where  $b$  controls the speed of event 2 onset :

$$\mathbb{P}(T_{2i} < t | \psi_i, W_i; \theta) = (1 - F_1(\infty)) \times \left(1 - \exp\left(-\frac{t}{b}\right)\right) \quad (11)$$

$p_1$ ,  $g_1$  were set to 0.05, 0.1 respectively and  $b$  was fixed to 10 to obtain about 17% cause 1 failures, 69% cause 2 failures, and 14% administrative censoring at time 30, similarly to the application. Further details about true values for model parameters are given in Supplementary Information S3.

## 6.2 | Evaluation of the estimation performances, the selection process and ROC AUC indicator

### 6.2.1 | Evaluation of the estimation performances

For each simulated dataset, the true model is estimated and writes as follows, considering  $K'' = 3$ :

$$\begin{aligned} y_{ij1} &= m_1(t_{ij1}, \psi_{i1}) + g[m_1(t_{ij1}, \psi_{i1}), \sigma_1] \epsilon_{ij} \\ &\dots \\ y_{ijK''} &= m_{K''}(t_{ijK''}, \psi_{iK''}) + g[m_{K''}(t_{ijK''}, \psi_{iK''}), \sigma_{K''}] \epsilon_{ij} \end{aligned} \quad (12)$$

$$\begin{aligned} \lambda_{1i}(t | \psi_i) &= \frac{p_1 g_1 \exp(-g_1 \times t)}{1 - p_1(1 - \exp(-g_1 \times t))} \exp\left[\sum_{k=1}^{K''} \left(\alpha_{1k} \times (m_k(t, \psi_{ik}) - med_k)\right)\right] \\ \lambda_{2i}(t | \psi_i) &= \frac{p_2 g_2 \exp(-g_2 \times t)}{1 - p_2(1 - \exp(-g_2 \times t))} \exp\left[\sum_{k=1}^{K''} \left(\alpha_{2k} \times (m_k(t, \psi_{ik}) - med_k)\right)\right] \end{aligned}$$

The model is estimated by maximization of the likelihood using the SAEM algorithm implemented on Monolix 2018 software with the same settings as explained in the application.

We aim to assess the performances of the estimation on the longitudinal parameters and the survival parameters for the event 1. We define  $\theta = \{\mu, \omega, p_1, g_1, \alpha_{11}, \alpha_{12}, \alpha_{13}\}$  and  $\hat{\theta}$  its estimate. The model specification for data simulation does not allow to easily identify  $p_2$ ,  $g_2$ ,  $\alpha_{21}$ ,  $\alpha_{22}$  and  $\alpha_{23}$  that is why estimation evaluation was based on vector  $\hat{\theta}$ . We evaluated the empirical relative bias and relative root mean square errors (RRMSE) of  $\hat{\theta}$  defined as :

$$\text{Relative bias} = \frac{1}{M} \sum_{m=1}^M \frac{\hat{\theta}^m - \theta}{\theta} \times 100 \quad \text{and} \quad \text{RRMSE} = \sqrt{\frac{1}{M} \sum_{m=1}^M \left(\frac{\hat{\theta}^m - \theta}{\theta} \times 100\right)^2} \quad (13)$$

We also provided violin plots of relative estimation errors (REE) defined for each dataset  $m \in \{1, \dots, M\}$  as :

$$\text{REE}^m = \frac{\hat{\theta}^m - \theta}{\theta} \times 100 \quad (14)$$

### 6.2.2 | Evaluation of the backward process

The objective is to assess the ability of the backward process to find the "true" model, that is to say the final set of biomarkers truly associated with the risk of in-hospital death. The first iteration of the process starts with the 7 biomarkers included in the multivariate joint model, and at each iteration, the biomarker with the highest p-value on  $\alpha_{1k}$  is removed. We stop when all the p-values for  $\alpha_{1k} < 5\%$ . For each dataset  $m \in \{1, \dots, M\}$  and for a given iteration of the backward process involving  $K''$  biomarkers, we estimated the joint model defined in equation 12.

We finally report the final set of biomarkers after the backward process for each simulation and each scenario of correlations.

### 6.2.3 | Evaluation of the ROC AUC indicator

To answer this objective, we computed the ROC AUC for the landmark time D9 and horizon time D30 for both true model and final model provided at the end of the backward process, for each simulated data set. We compared both distributions using paired t-test.

## 6.3 | Results

### 6.3.1 | Evaluation of the estimation performances

Table 6 provides relative bias and relative root mean square errors expressed in percentage. Relative bias and RRMSE are low for most of parameters. Longitudinal parameters are accurately estimated (while more uncertainty for some random effects, with higher relative bias and RRMSE). The parameters which link the longitudinal and the survival processes ( $\alpha_{11}$ ,  $\alpha_{12}$ ,  $\alpha_{13}$ ) are of particular interest and are estimated with a relative bias around 13% and a RRMSE around 30%. The design of the simulation may cause this moderate uncertainty of estimation for those parameters. In a previous work where only one biomarker was handled in a richer design and with a higher number of patients (Lavalley-Morelle et al., 2022), the RRMSE were reduced. Overall, this suggests good estimation performances given the size of the dataset and in particular, the small number of event 1 (deaths). Violin plots can be found in Supplementary Figure S3 and provide same conclusion.

### 6.3.2 | Evaluation of the backward process

The upset plot on Figure 6 is an efficient way to visualize the proportion of replicates finding the truly associated biomarkers. In the first scenario of correlations (correlation of 0.8 on slope parameters), 75% of the replicates contains the three correct biomarkers ( $bm_1$ ,  $bm_2$  and  $bm_3$ ): 64% only those three, 8% with one additional biomarker and 2% with two additional biomarkers. In 21% of cases, the process found only two of the true associated biomarkers. For the remaining cases, the process only found one associated biomarker (2%) or  $bm_4$  (correlated with  $bm_2$ ) was selected instead of  $bm_2$  (2%). We emphasize that all final models should have similar predictive performances.

For the scenario of stronger correlations, the upset plot is given in Supplementary Figure S4. Similarly to the previous scenario of correlations, 72% of the replicates found a correct set of biomarkers: 43% the true set of biomarkers, 15% with a very correlated biomarker instead of the associated one, 13% with one additional biomarker and 1% with three additional biomarkers. In 20% of cases the backward selected only two associated biomarkers, and in 4%, only one. For the remaining cases (4%), the process retains one or two associated biomarkers with an additional one.

Overall, the backward strategy demonstrates good performances even in the presence of strong correlations.

### 6.3.3 | Evaluation of the ROC AUC indicator

Figure 7 provides the distribution of ROC AUC for landmark = D9 and horizon time = D30 under the true and the final model, for the first scenario of correlations. Of note, 64% of ROC AUC were equal for both models, as the backward strategy leads to the true set of biomarkers in those cases. The figure shows that ROC AUC, in most cases, tends to be lower when a wrong model is selected, and more rarely, overestimated. In mean [95% CI], the ROC AUC is reduced by 0.010 [0.003, 0.017] ( $p = 0.007$ ) under the final model, which, although significant, is very small.

## 7 | DISCUSSION

In this work, we proposed to evaluate the added-value of modeling the full trend of various biomarkers to predict the risk of in-hospital death for COVID-19 patients, compared to using only information available at hospital admission. We developed a strategy to select between various biomarkers most associated with prognostic. The proposed methodology has the advantage to select a subset of most relevant biomarkers based on statistical methodology and not only on clinical considerations. We explored more information than an entire clinical-based approach that would directly considered a small set of biomarkers in the final model. However, we emphasize that building a predictive tool according to statistical criteria or only clinical knowledge may produce similar results.

We found that the dynamics of the blood neutrophil counts, arterial pH and C-reactive protein were significantly associated with the risk of in-hospital death and discharge from hospital: an increase of the blood neutrophil counts or the CRP, or a decrease of the pH is associated with a higher risk of death and a lower risk of discharge. We showed that the individual predictions from the multivariate joint model were significantly improved and outperformed those from the baseline model when sufficient information was available. Hence, such a tool can be useful for clinicians in routine medical practice as it may support clinical decisions such as therapeutic escalation or limitation of care. These decisions usually occur during the follow-up, and considering up-to-date longitudinal information to derive survival probabilities seems crucial. More generally, our methodology could be straightforwardly extended to other outcomes or medical conditions, and be easily implemented into hospital information systems to provide clinicians with dynamically updated prognostic scores and their estimated uncertainty.

Our real-case study also has some limitations. First, we recall that the baseline score referring to the 4C-Score was computed for each patient with some missing data for CRP and urea components. While it affects few individuals, the single imputation of missing data, given the other variables of the score, can introduce a small bias in the modeling, damped by the fact that the entire score slightly depends only slightly on those components. Second, one can argue that complex biomarker trends are not accounted for in the choice of the modeling. However, in such case, we think that none parametric model can approximate the evolution, and non-parametric models are out of the scope of this paper. Moreover, it worth noting that biomarkers which evolution are hard to model are generally non-informative biomarkers that have no added-value in terms of predictions. Then, two different thresholds for the residual error were used for the selection of linear and nonlinear models during univariate analysis. This relies on subjective appreciations and others thresholds could be proposed as well. Some biological measurements exhibit erratic behavior and are hard to model: the residual error should not be too high, but some flexibility is needed. We showed in the sensibility analysis that the choice of the thresholds in the univariate analysis impacts the set of biomarkers selected in the final model. However, all the scenarios considered lead to predictive performances that do not differ significantly. We also considered a set of biomarkers frequently prescribed to avoid at most selective missingness. However, it is worth noting that some biomarkers may have been more often measured in severe patients. In an extreme case, the population parameters estimated for such biomarker can be biased and not reflecting the dynamic in the whole population. In the survival part, the baseline risk functions considered may seem restrictive. However, the objective of the work is to provide a powerful predictive tool and we showed through the application that the predictions performed well with this parametrization. That is why no more complex function was considered. Then, in the backward processes, we chose to select biomarkers that are associated with both competing risks. As an alternative approach and for other applications, we could allow for different biomarkers being associated with one or both risks. Finally, our model should also be validated on an external data set before being used in practice. Due to the sample size and the number of events, an internal validation by splitting data was found inappropriate.

In a methodological point of view, we demonstrated well estimation performances of the SAEM algorithm. We evaluated the estimation in a complex design with 3 longitudinal submodels (with one nonlinear) and 2 competing risks with a small number of event 1. Bias and RRMSE were relatively low for most of parameters. However, as we are in a subdistribution approach, estimation performances relative to survival parameters associated with event 2 are impossible to evaluate. It may be an interesting future research. Moreover, although backward selection procedure has known drawbacks, it has never been proposed in the context of joint modeling with competing risks and we found here interesting performances in the simulation study even when introducing correlations between biomarkers. We emphasize that strong correlation structures were considered in the simulations. In reality, correlations strongly depend on measurement errors and are probably lower than those considered. Other selection strategies as the forward or forward-backward stepwise ones may be considered, although they may be more time-consuming because of the number of combinations to test at each iteration of the process. We acknowledge that the backward strategy may be computationally expensive depending on the machine hardware and software configuration. We showed through simulations that this strategy may be extended for a larger set of patients but not for a too large set of biomarkers. In fact,

the stochastic algorithm used to compute the FIM shows an exponential growth in terms of CPU time spent as the number of biomarkers increases. This algorithm is implemented according to the Louis method and has the disadvantage of being run once parameters have been estimated (see Kuhn and Lavielle, 2005 for more details). The whole estimation process (parameters and standard errors) is thus, time consuming. Recently, Delattre and Kuhn (Delattre & Kuhn, 2023) developed an interesting alternative: a stochastic algorithm to compute the FIM as the SAEM algorithm iterates. It offers the potential to be less time consuming, however, it has not been evaluated for multiple responses and joint models so future developments are needed to consider this alternative. Finally, we note that the *compare* function used for testing and comparing time-dependent AUC between the baseline and the multivariate joint model was not designed for nested models. It would imply a loss of power and not a rise of the type-I error. Hence, there is no a negative impact on our results.

A possible extension of this work could be the implementation of a selection procedure based on penalized likelihood such as LASSO penalization, because forward/backward selection on regression models have been shown to be outperformed by this type of methods in a high-dimensional context.

## ACKNOWLEDGEMENTS

Authors thank Dr Arthur Mageau, AP-HP, Bichat Claude Bernard University Hospital, F-75018 Paris, France, for his clinical expertise and help.

## FUNDING INFORMATION

This work has received no external funding.

## CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

## DATA AVAILABILITY STATEMENT

Research data are not shared.

## References

- Andrinopoulou, E.-R., Rizopoulos, D., Takkenberg, J. J., & Lesaffre, E. (2017). Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. *Statistical Methods in Medical Research*, 26(4), 1787–1801. <https://doi.org/10.1177/0962280215588340>
- Blanche, P., Dartigues, J., & Jacqmin-Gadda, H. (2013). Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Statistics in Medicine*, 32(30), 5381–5397. <https://doi.org/10.1002/sim.5958>
- Blanche, P., Proust-Lima, C., Loubère, L., Berr, C., Dartigues, J., & Jacqmin-Gadda, H. (2015). Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics*, 71(1), 102–113. <https://doi.org/10.1111/biom.12232>
- Delattre, M., & Kuhn, E. (2023). Estimating Fisher Information Matrix in Latent Variable Models based on the Score Function. <https://hal.science/hal-02285712>
- Delyon, B., Lavielle, M., & Moulines, E. (1999). Convergence of a stochastic approximation version of EM algorithm. *The Annals of Statistics*, 27, 94.

- Desmée, S., Mentré, F., Veyrat-Follet, C., Sébastien, B., & Guedj, J. (2017). Nonlinear joint models for individual dynamic prediction of risk of death using Hamiltonian Monte Carlo: Application to metastatic prostate cancer. *BMC Medical Research Methodology*, *17*(1), 105. <https://doi.org/10.1186/s12874-017-0382-9>
- Fine, J. P. (2001). Regression modeling of competing crude failure probabilities. *Biostatistics (Oxford, England)*, *2*(1), 85–97. <https://doi.org/10.1093/biostatistics/2.1.85>
- Jeong, J.-H., & Fine, J. (2006). Direct Parametric Inference for the Cumulative Incidence Function [Publisher: [Wiley, Royal Statistical Society]]. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *55*(2), 187–200. Retrieved November 8, 2021, from <https://www.jstor.org/stable/3592662>
- Keroui, M., Mercier, F., Bertrand, J., Tardivon, C., Bruno, R., Guedj, J., & Desmée, S. (2020). Bayesian inference using Hamiltonian Monte-Carlo algorithm for nonlinear joint modeling in the context of cancer immunotherapy. *Statistics in Medicine*, *39*(30), 4853–4868. <https://doi.org/https://doi.org/10.1002/sim.8756>
- Knight, S. R., Ho, A., Pius, R., Buchan, I., Carson, G., Drake, T. M., Dunning, J., Fairfield, C. J., Gamble, C., Green, C. A., Gupta, R., Halpin, S., Hardwick, H. E., Holden, K. A., Horby, P. W., Jackson, C., Mclean, K. A., Merson, L., Nguyen-Van-Tam, J. S., ... Young, P. (2020). Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Development and validation of the 4C Mortality Score. *The BMJ*, *370*, m3339. <https://doi.org/10.1136/bmj.m3339>
- Kuhn, E., & Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis*, *49*(4), 1020–1038. <https://doi.org/10.1016/j.csda.2004.07.002>
- Lavalley-Morelle, A., Timsit, J.-F., Mentré, F., Mullaert, J., & Network, T. O. (2022). Joint modeling under competing risks: Application to survival prediction in patients admitted in Intensive Care Unit for sepsis with daily Sequential Organ Failure Assessment score assessments [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/psp4.12856>]. *CPT: Pharmacometrics & Systems Pharmacology*, *11*(11), 1472–1484. <https://doi.org/10.1002/psp4.12856>
- Lavillegrand, J.-R., Garnier, M., Spaeth, A., Mario, N., Hariri, G., Pilon, A., Berti, E., Fieux, F., Thietart, S., Urbina, T., Turpin, M., Darrivere, L., Fartoukh, M., Verdonk, F., Dumas, G., Tedgui, A., Guidet, B., Maury, E., Chantran, Y., ... Ait-Oufella, H. (2021). Elevated plasma IL-6 and CRP levels are associated with adverse clinical outcomes and death in critically ill SARS-CoV-2 patients: Inflammatory response of SARS-CoV-2 patients. *Annals of Intensive Care*, *11*, 9. <https://doi.org/10.1186/s13613-020-00798-x>
- Li, T., Wang, X., Zhuang, X., Wang, H., Li, A., Huang, L., Zhang, X., Xue, Y., Wei, F., & Ma, C. (2021). Baseline characteristics and changes of biomarkers in disease course predict prognosis of patients with COVID-19. *Internal and Emergency Medicine*, *16*(5), 1165–1172. <https://doi.org/10.1007/s11739-020-02560-4>
- Liang, W., Liang, H., Ou, L., Chen, B., Chen, A., Li, C., Li, Y., Guan, W., Sang, L., Lu, J., Xu, Y., Chen, G., Guo, H., Guo, J., Chen, Z., Zhao, Y., Li, S., Zhang, N., Zhong, N., & He, J. (2020). Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. *JAMA Internal Medicine*, *180*(8), 1–9. <https://doi.org/10.1001/jamainternmed.2020.2033>
- Liu, M., Sun, J., Herazo-Maya, J. D., Kaminski, N., & Zhao, H. (2019). Joint Models for Time-to-Event Data and Longitudinal Biomarkers of High Dimension. *Statistics in biosciences*, *11*(3), 614–629. <https://doi.org/10.1007/s12561-019-09256-0>
- Long, J. D., & Mills, J. A. (2018). Joint modeling of multivariate longitudinal data and survival data in several observational studies of Huntington’s disease. *BMC Medical Research Methodology*, *18*(1), 138. <https://doi.org/10.1186/s12874-018-0592-9>
- Louis, T. A. (1982). Finding the Observed Information Matrix When Using the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *44*(2), 226–233. <https://doi.org/10.1111/j.2517-6161.1982.tb01203.x>
- Mueller, A. A., Tamura, T., Crowley, C. P., DeGrado, J. R., Haider, H., Jezmir, J. L., Keras, G., Penn, E. H., Massaro, A. F., & Kim, E. Y. (2020). Inflammatory Biomarker Trends Predict Respiratory Decline in COVID-19 Patients. *Cell Reports. Medicine*, *1*(8), 100144. <https://doi.org/10.1016/j.xcrm.2020.100144>
- Myrstad, M., Ihle-Hansen, H., Tveita, A. A., Andersen, E. L., Nygård, S., Tveit, A., & Berge, T. (2020). National Early Warning Score 2 (NEWS2) on admission predicts severe disease and in-hospital mortality from Covid-19 - a prospective cohort study. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, *28*(1), 66. <https://doi.org/10.1186/s13049-020-00764-3>
- Rajeswaran, J., Blackstone, E. H., & Barnard, J. (2018). Joint Modeling of Multivariate Longitudinal Data and Competing Risks Using Multiphase Sub-models. *Statistics in Biosciences*, *10*(3), 651–685. <https://doi.org/10.1007/s12561-018-9223-6>

- Rizopoulos, D. (2011). Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics*, *67*(3), 819–829. <https://doi.org/10.1111/j.1541-0420.2010.01546.x>
- Schmid, M., Kestler, H. A., & Potapov, S. (2013). On the validity of time-dependent AUC estimators. *Briefings in Bioinformatics*, *16*(1), 153–168. <https://doi.org/10.1093/bib/bbt059>
- Shen, F., & Li, L. (2021). Backward joint model and dynamic prediction of survival with multivariate longitudinal data. *Statistics in Medicine*, *40*(20), 4395–4409. <https://doi.org/10.1002/sim.9037>
- Tjendra, Y., Al Mana, A. F., Espejo, A. P., Akgun, Y., Millan, N. C., Gomez-Fernandez, C., & Cray, C. (2020). Predicting Disease Severity and Outcome in COVID-19 Patients: A Review of Multiple Biomarkers. *Archives of Pathology & Laboratory Medicine*, *144*(12), 1465–1474. <https://doi.org/10.5858/arpa.2020-0471-SA>
- Tong-Minh, K., van der Does, Y., van Rosmalen, J., Ramakers, C., Gommers, D., van Gorp, E., Rizopoulos, D., & Endeman, H. (2022). Joint Modeling of Repeated Measurements of Different Biomarkers Predicts Mortality in COVID-19 Patients in the Intensive Care Unit. *Biomarker Insights*, *17*, 11772719221112370. <https://doi.org/10.1177/11772719221112370>
- Zhang, L., Yan, X., Fan, Q., Liu, H., Liu, X., Liu, Z., & Zhang, Z. (2020). D-dimer levels on admission to predict in-hospital mortality in patients with Covid-19. *Journal of Thrombosis and Haemostasis*, *18*(6), 1324–1329. <https://doi.org/10.1111/jth.14859>
- Zhu, Z., Cai, T., Fan, L., Lou, K., Hua, X., Huang, Z., & Gao, G. (2020). The potential role of serum angiotensin-converting enzyme in coronavirus disease 2019. *BMC Infectious Diseases*, *20*(1), 883. <https://doi.org/10.1186/s12879-020-05619-x>

## TABLES

**TABLE 1** Baseline characteristics (components of the baseline score) for the 327 patients of the data

<b>Number of patients</b>	327
<b>Age - Number (%)</b>	
<50	73 (22)
[50-59[	74 (23)
[60-69[	71 (22)
[70-79[	62 (19)
≥ 80	47 (14)
<b>Gender - Male - Number (%)</b>	198 (61)
<b>Comorbidities - Number (%)</b>	
0	139 (43)
1	93 (28)
≥ 2	95 (29)
<b>Urea (mmol/L) - med [Q1-Q3]</b>	5.6 [4.1 - 8.1]
NA - Number (%)	2 (0.6)
1st measurement > 48h - Number (%)	11 (3)
<b>CRP (mg/L) - med [Q1-Q3]</b>	67.5 [30.3 - 120.8]
NA - Number (%)	9 (3)
1st measurement > 48h - Number (%)	15 (5)
<b>Score - med [Q1-Q3]</b>	6 [4 - 9]

**TABLE 2** Parameter estimates for the baseline model. P-value of Wald tests are provided for the coefficients associated to the baseline score effect.

<b>Parameter (unit)</b>	<b>Value</b>	<b>SE</b>	<b>RSE (%)</b>	<b>P-value</b>
<b>Death</b>				
$h_1(d^{-1})$	0.0003	0.00016	53	
$\beta_1$	0.36	0.05	14	$<10^{-5}$
<b>Discharge</b>				
$h_2(d^{-1})$	0.13	0.016	13	
$\beta_2$	-0.14	0.02	13	$<10^{-5}$

**TABLE 3** Parameter estimates of the multivariate joint model. P-value Wald of tests are reported for the link coefficients and the coefficients associated with the baseline score effect.

Parameter (unit)	Value	SE	RSE (%)	p-value
<b>Longitudinal submodel</b>				
<b>Blood neutrophil counts</b>				
$\mu_{0n}$ ( $10^9.L^{-1}$ )	4.59	0.16	3.6	
$\mu_{1n}$ ( $10^9.L^{-1}.d^{-1}$ )	-0.15	0.024	16.0	
$\mu_{2n}$ ( $10^9.L^{-1}.d^{-1}$ )	-0.16	0.016	10.0	
$\mu_{an}$ ( $10^9.L^{-1}$ )	5.30	1.58	29.8	
$\omega_{0n}$ ( $10^9.L^{-1}$ )	2.10	0.15	7.1	
$\omega_{1n}$ ( $10^9.L^{-1}.d^{-1}$ )	0.13	0.017	13.0	
$\omega_{2n}$ ( $10^9.L^{-1}.d^{-1}$ )	0.076	0.018	24.2	
$\omega_{an}$ ( $10^9.L^{-1}$ )	0.83	0.21	24.8	
$\sigma_{bn}$	0.32	0.0070	2.2	
<b>arterial pH</b>				
$\mu_{0p}$	7.44	0.0035	0.05	
$\mu_{1p}$ ( $d^{-1}$ )	0.0027	0.00079	29.1	
$\omega_{0p}$	0.039	0.0027	7.1	
$\omega_{1p}$ ( $d^{-1}$ )	0.0053	0.00061	11.6	
$\sigma_{ap}$	0.055	0.00092	1.7	
<b>C-reactive protein</b>				
$\mu_{0c}$ ( $\log(mg.L^{-1})$ )	4.18	0.061	1.5	
$\mu_{1c}$ ( $\log(mg.L^{-1}).d^{-1}$ )	-0.16	0.012	7.3	
$\omega_{0c}$ ( $\log(mg.L^{-1})$ )	0.93	0.048	5.2	
$\omega_{1c}$ ( $\log(mg.L^{-1}).d^{-1}$ )	0.15	0.011	7.4	
$\sigma_{ac}$ ( $\log(mg.L^{-1})$ )	0.71	0.015	2.1	
<b>Survival submodel</b>				
<b>Death</b>				
$h_1$ ( $d^{-1}$ )	0.00037	0.00024	65.3	
$\alpha_{1n}$ ( $L.10^{-9}$ )	0.14	0.033	24.2	$< 10^{-5}$
$\alpha_{1p}$	-11.4	2.82	24.9	$< 10^{-5}$
$\alpha_{1c}$ ( $-\log(mg.L^{-1})$ )	0.63	0.22	34.7	0.004
$\beta_1$	0.33	0.060	18.1	$< 10^{-5}$
<b>Discharge</b>				
$h_2$ ( $d^{-1}$ )	0.014	0.0069	51.4	
$\alpha_{2n}$ ( $L.10^{-9}$ )	-0.14	0.054	39.1	0.01
$\alpha_{2p}$	25.2	6.35	25.2	$< 10^{-5}$
$\alpha_{2c}$ ( $-\log(mg.L^{-1})$ )	-1.09	0.18	16.2	$< 10^{-5}$
$\beta_2$	-0.12	0.032	27.3	0.0002

**TABLE 4** Tests for comparing ROC AUC at horizon time 30 between the baseline and the joint model

<b>Horizon time = 30</b>	<b>Landmark day 3</b>	<b>Landmark day 6</b>	<b>Landmark day 9</b>
<b>AUC [95% CI] – baseline model</b>	0.79 [0.65,0.81]	0.73 [0.65,0.81]	0.64 [0.55, 0.74]
<b>AUC [95% CI] – joint model</b>	0.82 [0.75,0.89]	0.81 [0.73,0.89]	0.84 [0.75,0.93]
<b>p-value</b>	0.37	0.04	$< 10^{-5}$

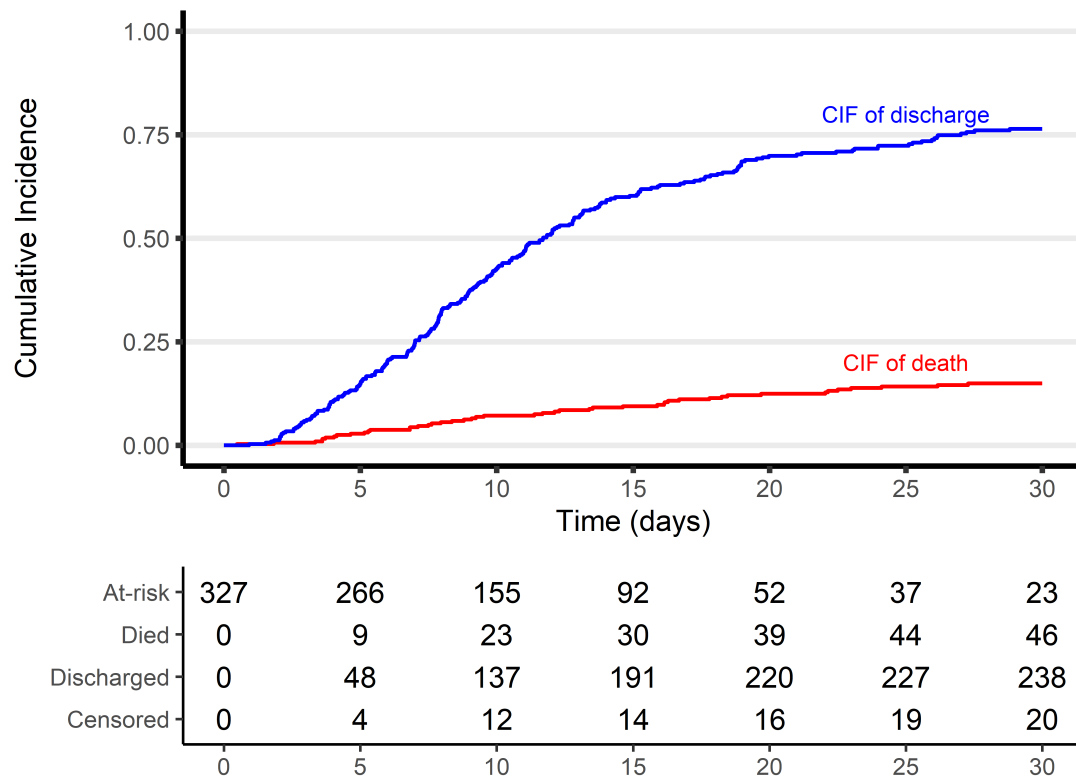
<b>Biomarker</b>	<b>Correspondance</b>	<b>Longitudinal submodel</b>	<b>Error model</b>	<b>Measurement frequency</b>
<i>bm</i> <sub>1</sub>	Blood neutrophil counts	Nonlinear	Proportional	Every 2 days
<i>bm</i> <sub>2</sub>	Arterial pH	Linear	Additive	Every 1.5 days
<i>bm</i> <sub>3</sub>	CRP	Linear	Additive	Every 2 days
<i>bm</i> <sub>4</sub>	NT-proBNB	Linear	Additive	Every 3 days
<i>bm</i> <sub>5</sub>	D-dimers	Linear	Proportional	Every 3 days
<i>bm</i> <sub>6</sub>	Albuminemia	Linear	Proportional	Every 3 days
<i>bm</i> <sub>7</sub>	LDH	Linear	Proportional	Every 3 days

**TABLE 5** Characteristics of the 7 simulated biomarkers

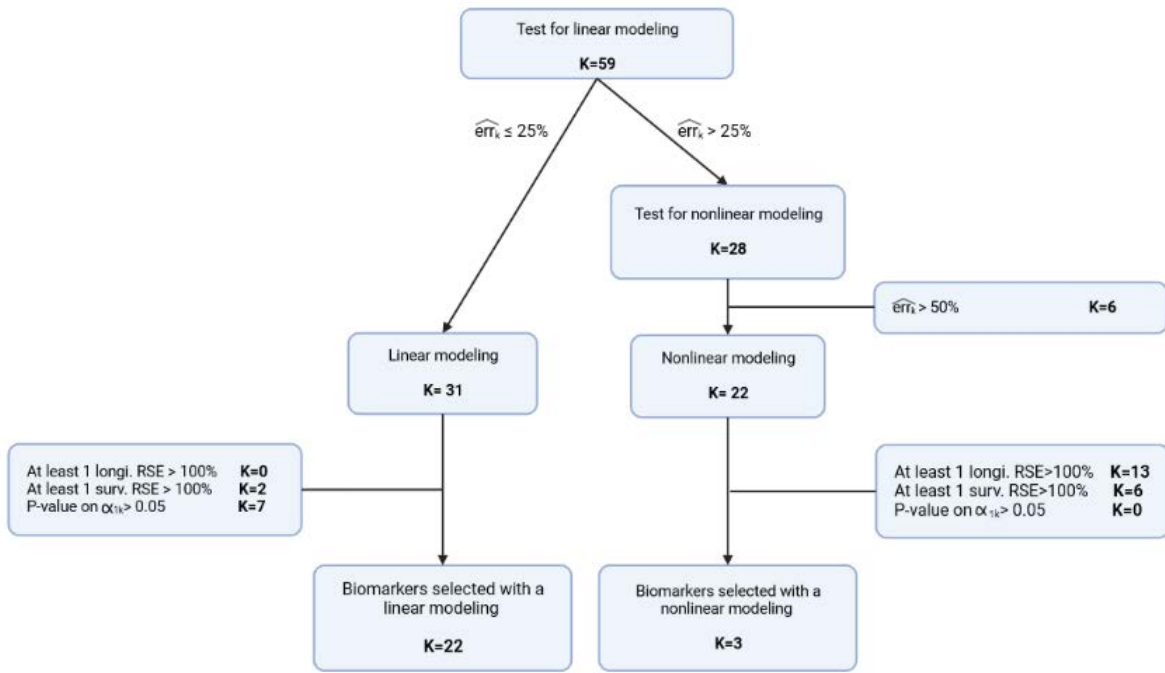
**TABLE 6** Estimated relative bias and Root Mean Square Errors (RMSE) based on 100 simulations

<b>Parameter</b>	<b>True value</b>	<b>Relative Bias (%)</b>	<b>RRMSE (%)</b>
<b><i>Longitudinal part - biomarker 1</i></b>			
<i>Fixed effects</i>			
$\mu_{01}$	4.6	0.4	3
$\mu_{11}$	-0.15	-8	15
$\mu_{21}$	-0.16	-7	14
$\mu_{a1}$	5.3	-9	17
<i>Random effects</i>			
$\omega_{01}$	2.0	-2	11
$\omega_{11}$	0.10	14	28
$\omega_{21}$	0.07	-8	36
$\omega_{a1}$	0.80	25	46
<i>Error parameter</i>			
$\sigma_{b1}$	0.30	2	3
<b><i>Longitudinal part - biomarker 2</i></b>			
<i>Fixed effects</i>			
$\mu_{02}$	7.4	$3^{-5}$	0.04
$\mu_{12}$	0.003	-6	18
<i>Random effects</i>			
$\omega_{02}$	0.04	2	11
$\omega_{12}$	0.005	6	17
<i>Error parameter</i>			
$\sigma_{a2}$	0.05	-0.1	1
<b><i>Longitudinal part - biomarker 3</i></b>			
<i>Fixed effects</i>			
$\mu_{03}$	4.2	-0.08	1
$\mu_{13}$	-0.16	-5	9
<i>Random effects</i>			
$\omega_{03}$	0.90	-0.9	11
$\omega_{13}$	0.15	7	15
<i>Error parameter</i>			
$\sigma_{a3}$	0.70	-0.1	2
<b><i>Survival part</i></b>			
$p_1$	0.05	-23	37
$g_1$	0.10	16	32
$\alpha_{11}$	0.14	15	35
$\alpha_{12}$	-11	13	32
$\alpha_{13}$	0.60	11	23

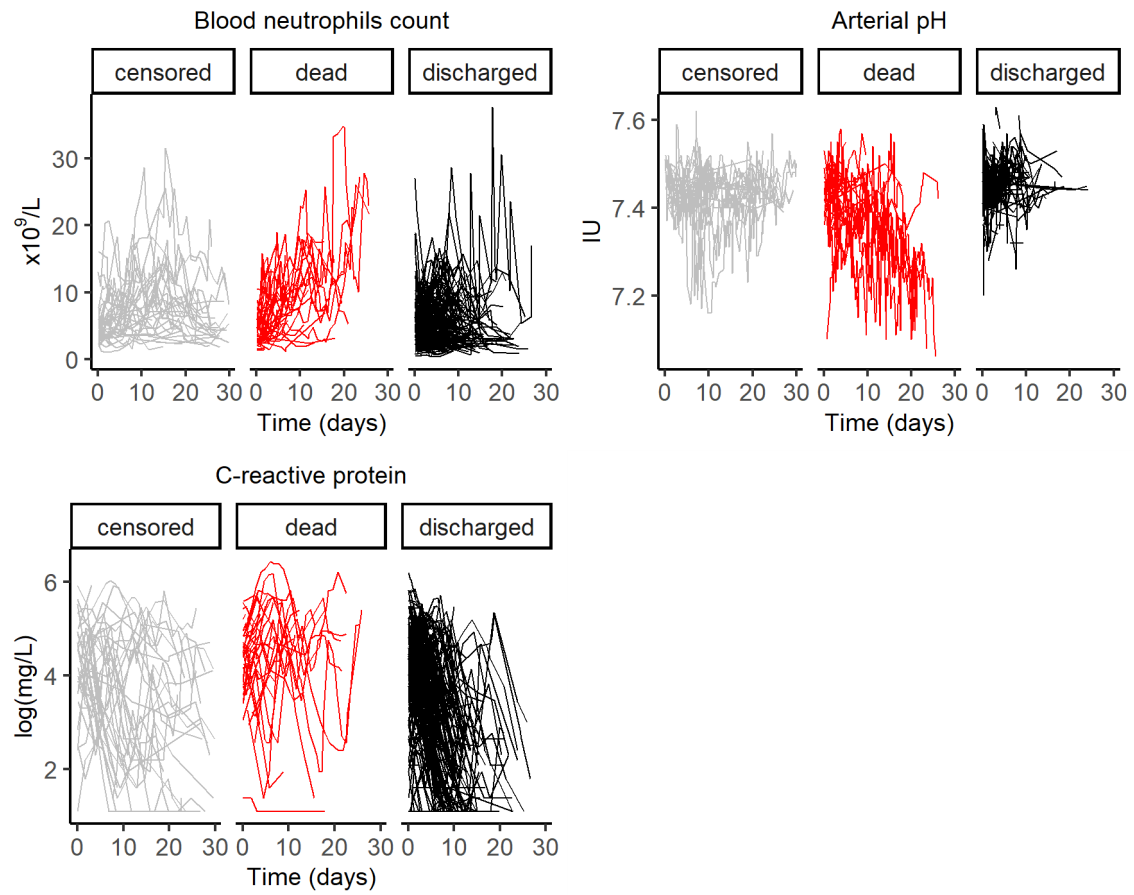
## FIGURES



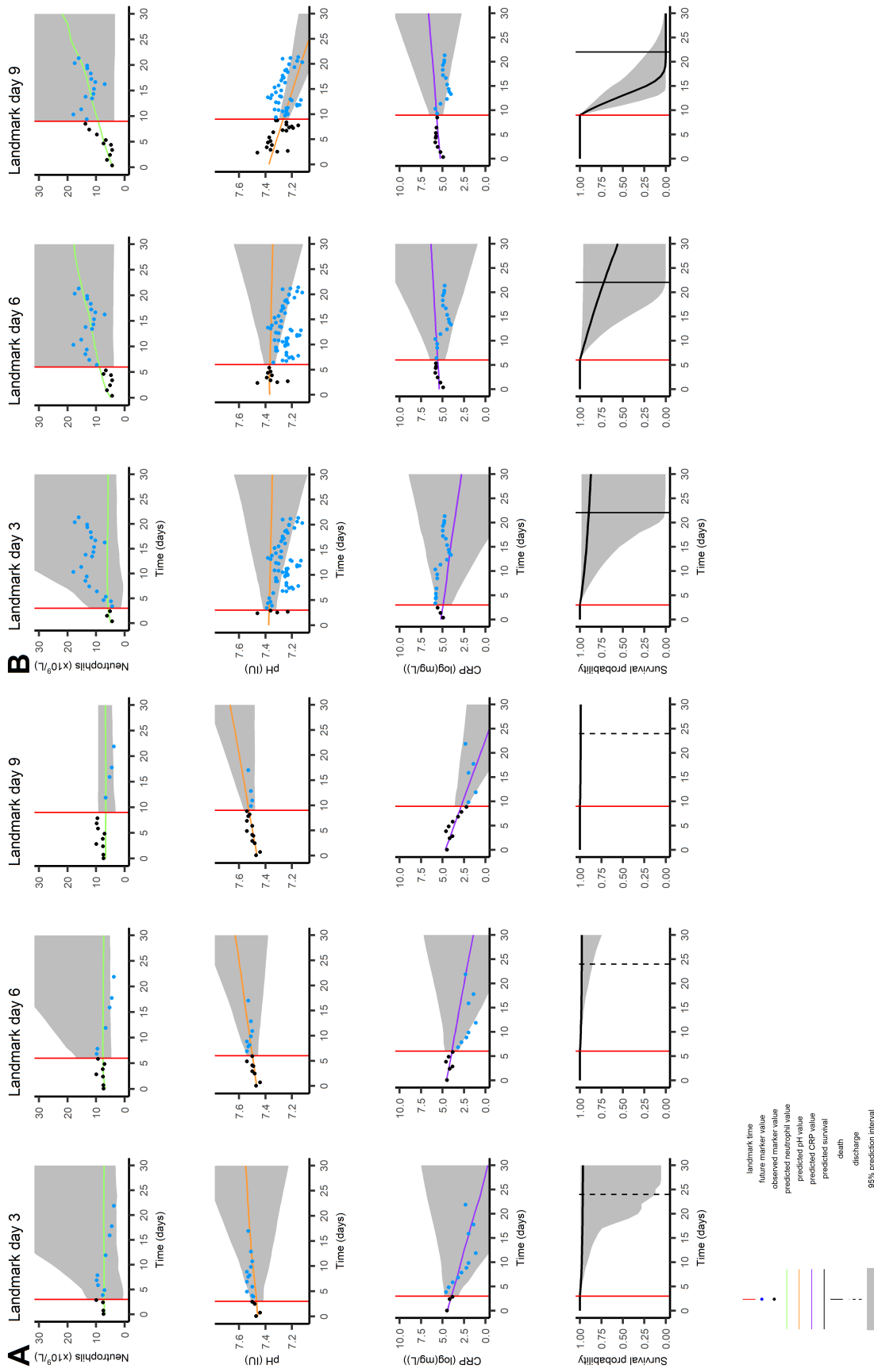
**FIGURE 1** Cumulative incidence functions (CIF) for in-hospital death (red) and hospital discharge (blue). The number of events and at-risk patients is given on the bottom.



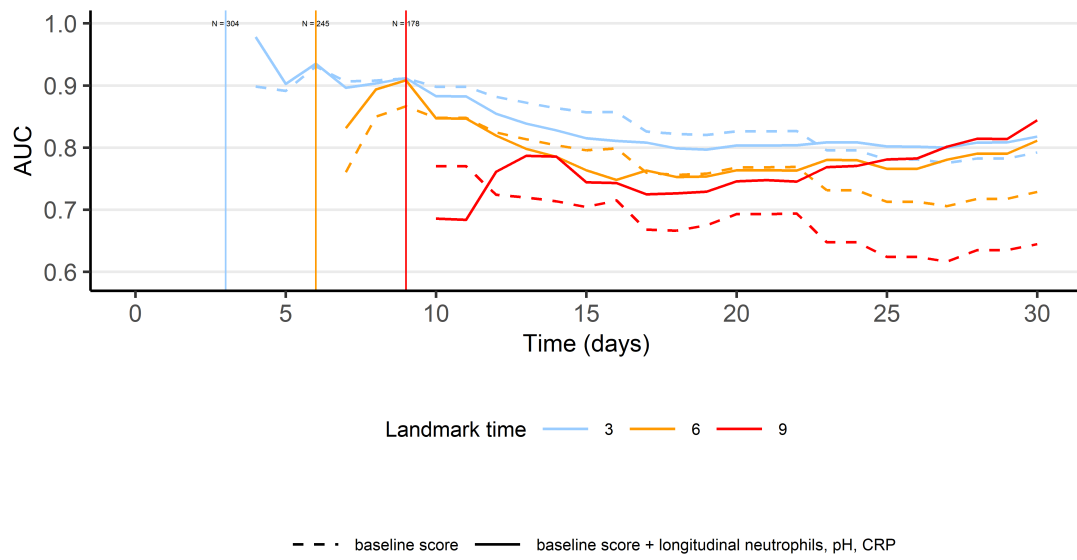
**FIGURE 2** Flowchart for the univariate selection



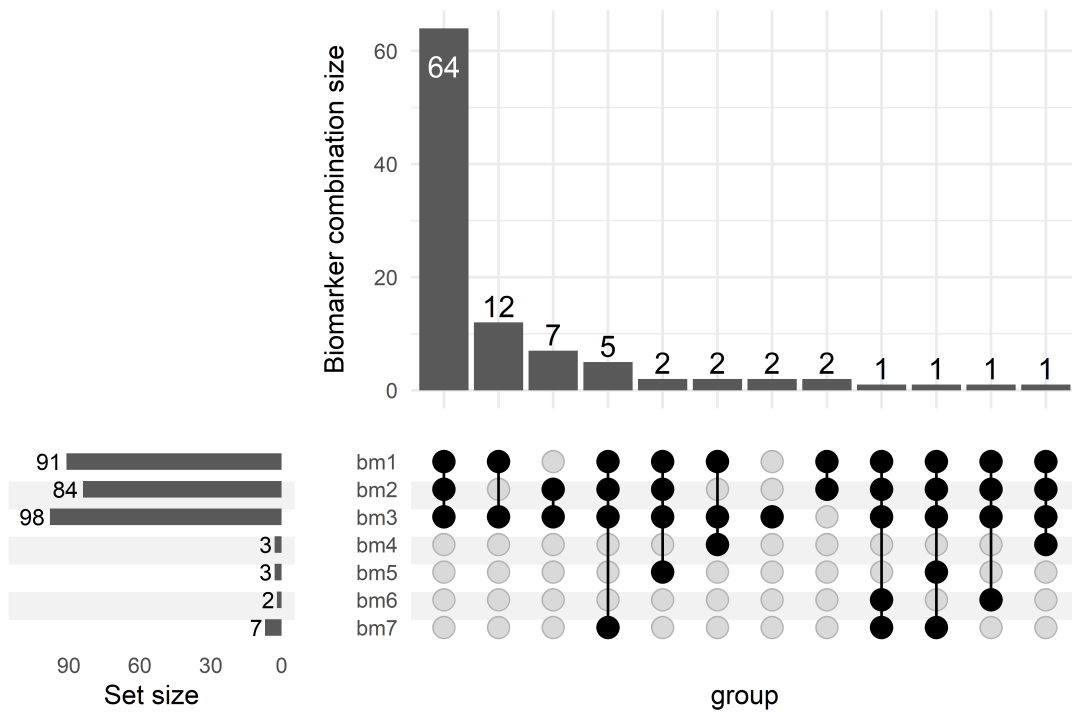
**FIGURE 3** Individual longitudinal evolution of the selected biomarkers, stratified by clinical outcome



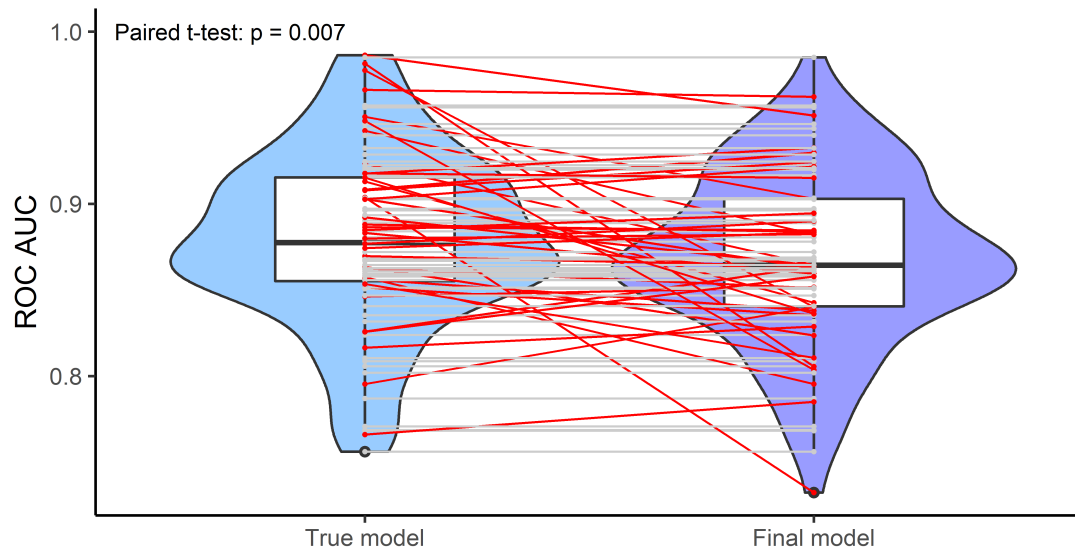
**FIGURE 4** Dynamic predictions and 95% prediction intervals for patient A (left) and patient B (right). Patient A was discharged at time 24 and patient B died at time 22. Predictions for the biomarkers evolution are on the first 3 rows (blood neutrophils count, arterial pH and CRP respectively), and predictions for the survival on the last row.



**FIGURE 5** Time-dependent AUC computed on the data set for various landmark and horizon times. The number of at-risk patients at each landmark is given on the top.



**FIGURE 6** Final set of biomarkers selected after the backward process (simulation study)



**FIGURE 7** Distributions of ROC AUC of the true model and the final model for landmark time = D9 and horizon time = D30. True model corresponds to the multivariate joint model with bm1, bm2 and bm3 and the final model the one with the biomarkers selected after the backward process. Grey lines represent situations when the final model is equal to the true model and red lines the other cases.

**How to cite this article:** Lavalley-Morelle A., N. Peiffer-Smadja, S B. Gressens, B. Souhail, A. Bounhiol, FX. Lescure, F. Mentré ,and J. Mullaert (2022), Multivariate joint model under competing risks to predict death of hospitalized patients for SARS-CoV-2 infection, *Biometrical Journal*, .

# **Supplementary Materials - Multivariate joint model under competing risks to predict death of hospitalized patients for SARS-CoV-2 infection**

## **1 Supplementary Tables**

**TABLE S1** Biomarkers considered in the analysis, classified in 8 categories. For each biomarker, the number of patients who have at least one observations before day 30 as well as the mean number of observations are given.

Biomarker	Unit	N*	n**	Biomarker	Unit	N*	n**
<b>Complete blood count</b>				<b>Kidney functions / cellular lysis</b>			
Erythrocytes	$\times 10^{12}/L$	326	6.6	Natremia	mmol/L	327	7.1
Mean corpuscular volume (MCV)	fL	326	6.6	Kalemia	mmol/L	325	7.1
Hemoglobin	g/dL	326	6.6	Chloremia	mmol/L	300	5.7
Hematocrit	%	326	6.6	Calcemia	mmol/L	315	5.3
Reticulocytes	$\times 10^9/L$	170	4.4	Phosphates	mmol/L	281	4.9
Leukocytes	$\times 10^9/L$	326	6.6	Magnesium	mmol/L	260	5.2
Basophil polynuclear cells	$\times 10^9/L$	322	6.4	Anion gap	mmol/L	280	4.8
Eosinophil polynuclear cells	$\times 10^9/L$	322	6.5	Bicarbonates	mmol/L	295	5.7
Neutrophil polynuclear cells	$\times 10^9/L$	326	6.5	Uremia	mmol/L	325	6.9
Immature granulocytes ratio	%	139	3.7	Protidemia	g/L	290	5.6
Lymphocytes	$\times 10^9/L$	326	6.5	Creatininemia	$\mu\text{mol/L}$	327	6.5
Monocytes	$\times 10^9/L$	326	6.5	Lactate dehydrogenase (LDH)	U/L	297	4.2
Platelets	$\times 10^9/L$	326	6.6	Creatine phosphokinase (CPK)	U/L	309	4.3
				Alkaline phosphatases	U/L	313	5.3
<b>Coagulation</b>				<b>Liver/pancreatic functions</b>			
Prothrombin Ratio (PR)	%	303	4.9	Alamine amino transferase (ALAT)	U/L	321	5.6
Activated partial thromboplastin time (aPTT)	IU	297	4.6	Aspartate amino transferase (ASAT)	U/L	321	5.6
Activated factor II	%	153	4.5	Gamma GT	U/L	314	5.1
Activated factor V	%	155	4.5	Total bilirubin	$\mu\text{mol/L}$	322	5.5
Fibrinogen	g/L	294	4.4	Lipasemia	U/L	158	4.3
Fibrin monomers	$\mu\text{g/mL}$	112	5.3	Albuminemia	g/L	295	3.8
D-Dimers	ng/mL	218	3.9				
<b>Pulmonary functions</b>				<b>Markers of inflammation</b>			
Arterial pH	IU	246	8.3	C-reactive protein (CRP)	log(mg/L)	318	5.5
pO2a	mm Hg	246	8.3	Procalcitonin (PCT)	$\mu\text{g/L}$	229	4
pCO2a	mm Hg	246	8.3	Orosomuroid	g/L	101	4.2
Arterial lactate	mmol/L	231	8.1	Ferritin	$\mu\text{g/L}$	214	3.3
Oxyhemoglobin ratio	%	245	8	Haptoglobin	g/L	116	4.7
<b>Urine samples</b>				<b>Cardiac markers</b>			
Natriuresis	mmol/L	132	3	Ultrasensitive troponin I	$\mu\text{g/L}$	253	3.6
kaliuresis	mmol/L	132	3	NT pro BNP	ng/L	262	3.5
Urinary chloride	mmol/L	62	3.6				
Proteinuria	g/L	109	2.9				
Urinary urea	mmol/L	119	3				
Creatinyuria	mmol/L	134	3				
Glycosuria	mmol/L	76	3.2				

\*Number of patients with at least observation between admission and day 30

\*\*Mean number of observations for the patients having at least one between admission and day 30

**TABLE S2** 4C-score components. The items of the 8 components are given in black and the associated weights, in grey.

<b>Age (years)</b>	<b>&lt;50</b>	0
	<b>50-59</b>	+2
	<b>60-69</b>	+4
	<b>70-79</b>	+6
	<b>≥ 80</b>	+7
<b>Sex at birth</b>	<b>Female</b>	0
	<b>Male</b>	+1
<b>Number of comorbidities</b> Comorbidities include chronic cardiac disease, chronic respiratory disease (excluding asthma), chronic renal disease, mild to severe liver disease, dementia, chronic neurological conditions, connective tissue disease, diabetes mellitus, HIV or AIDS and malignancy	<b>0</b>	0
	<b>1</b>	+1
	<b>≥ 2</b>	+2
<b>Respiratory rate, breaths/min</b>	<b>&lt;20</b>	0
	<b>20-29</b>	+1
	<b>≥30</b>	+2
<b>Peripheral oxygen saturation on air room</b>	<b>≥92%</b>	0
	<b>&lt;92%</b>	+2
<b>Glasgow Coma Scale</b>	<b>15</b>	0
	<b>&lt;15</b>	+2
<b>Urea at admission (mmol/L)</b>	<b>&lt;7</b>	0
	<b>7-14</b>	+1
	<b>&gt;14</b>	+2
<b>C-reactive protein at admission (mg/L)</b>	<b>&lt;50</b>	0
	<b>50-100</b>	+1
	<b>&gt;100</b>	+2

**TABLE S3a** Thresholds considered for the sensibility analysis in the univariable part.

	<b>Stringent scenario</b>	<b>Current scenario</b>	<b>Flexible scenario</b>
Measurement error for linear biomarkers (%)	15	25	35
Measurement error for nonlinear biomarkers (%)	30	50	70
RSE for model parameters (%)	60	100	140

**TABLE S3b** Results of the univariable and multivariable analysis for the depending on the scenario (stringent, current or flexible one)

	<b>Stringent scenario</b>	<b>Current scenario</b>	<b>Flexible scenario</b>
Number of linear biomarkers (univariable analysis)	8	22	26
Number of nonlinear biomarkers (univariable analysis)	0	3	1
Number of biomarkers selected for the multivariable analysis	3	7	8
Number of biomarkers included in the final model	3	3	2
Biomarkers included in the final model	Erythrocytes, Magnesium, Albuminemia	Blood neutrophil counts, arterial pH, CRP	arterial pH, CRP
ROC AUC [95%] CI (Imk=D9, horizon = D30)	0.76 [0.67,0.85]	0.84 [0.75,0.93]	0.81 [0.72,0.90]
p-value comparing ROC AUC with the current scenario	0.05	-	0.09

TABLE S4 Biomarkers selected after the univariate stage, ranked by increasing p-value Wald test on  $\alpha_{1k}$

Biomarker	N*	n**	$\alpha_{1k}$	RSE for $\alpha_{1k}$ (%)	$-\log_{10}$ p-value for $\alpha_{1k}$	longitudinal submodel
<b>Complete blood count</b>						
<b>Blood neutrophil counts</b>	<b>326</b>	<b>6.5</b>	<b>0.24</b>	<b>16.17</b>	<b>9.20</b>	<b>nonlinear</b>
Platelets	326	6.6	-0.004	27.66	3.52	linear
Erythrocytes	326	6.6	-0.44	45.28	1.57	linear
Hemoglobin	326	6.6	-0.14	49.06	1.38	linear
<b>Coagulation</b>						
<b>D-Dimers</b>	<b>218</b>	<b>3.9</b>	<b>1.08</b>	<b>14.86</b>	<b>10.78</b>	<b>linear</b>
Activated factor V	155	4.5	0.04	18.40	7.26	linear
aPTT	297	4.6	1.50	20.00	6.24	linear
Fibrinogen	294	4.4	0.70	22.10	5.22	linear
Activated factor II	153	4.5	-0.02	45.62	1.55	linear
<b>Pulmonary functions</b>						
<b>pHa</b>	<b>246</b>	<b>8.3</b>	<b>-20.61</b>	<b>11.18</b>	<b>18.42</b>	<b>linear</b>
pCO2a	246	8.3	0.19	12.48	14.95	linear
<b>Urine samples</b>						
<b>Kidney functions / cellular lysis</b>						
<b>LDH</b>	<b>297</b>	<b>4.2</b>	<b>0.01</b>	<b>12.97</b>	<b>13.90</b>	<b>linear</b>
Uremia	325	6.9	0.07	18.11	7.48	nonlinear
Kaliuresis	132	3.0	0.10	19.19	6.73	nonlinear
Magnesium	260	5.2	6.65	23.96	4.52	linear
Calcemia	315	5.3	-6.00	25.27	4.12	linear
Creatininemia	327	6.5	0.003	32.27	2.71	linear
Phosphates	281	4.9	1.92	39.10	1.98	linear
Kalemia	325	7.1	0.99	44.01	1.64	linear
<b>Liver/pancreatic functions</b>						
<b>Albuminemia</b>	<b>295</b>	<b>3.8</b>	<b>-0.11</b>	<b>27.43</b>	<b>3.57</b>	<b>linear</b>
Lipasemia	158	4.3	0.88	17.22	8.19	linear
<b>Markers of inflammation</b>						
<b>CRP</b>	<b>318</b>	<b>5.5</b>	<b>1.25</b>	<b>17.63</b>	<b>7.85</b>	<b>linear</b>
Haptoglobin	116	4.7	0.47	19.03	6.83	linear
Orosomuroid	101	4.2	1.85	19.87	6.32	linear
<b>Cardiac markers</b>						
<b>NT-proBNP</b>	<b>267</b>	<b>3.5</b>	<b>0.48</b>	<b>23.01</b>	<b>4.86</b>	<b>linear</b>

**TABLE S5** Parameter estimates of the multivariate joint model. P-value Wald of tests are reported for the link coefficients and the coefficients associated with the baseline score effect.

Parameter (unit)	Value	SE	RSE (%)	Parameter (unit)	Value	SE	RSE (%)	p-value
<b>Longitudinal submodel</b>				<b>Survival submodel</b>				
<b>Blood neutrophil counts</b>				<b>Death</b>				
$\mu_{0n}$ ( $10^9.L^{-1}$ )	4.56	0.18	3.9	$h_1$ ( $d^{-1}$ )	0.00031	0.00031	101	
$\mu_{1n}$ ( $10^9.L^{-1}.d^{-1}$ )	-0.14	0.030	22.3	$\alpha_{1n}$ ( $L.10^{-9}$ )	0.14	NA	NA	NA
$\mu_{2n}$ ( $10^9.L^{-1}.d^{-1}$ )	-0.15	0.019	12.1	$\alpha_{1p}$	-8.28	NA	NA	NA
$\mu_{an}$ ( $10^9.L^{-1}$ )	5.60	2.32	41.4	$\alpha_{1c}$ ( $-\log(mg.L^{-1})$ )	0.86	0.25	28.6	0.0006
$\omega_{0n}$ ( $10^9.L^{-1}$ )	2.06	0.16	7.8	$\alpha_{1d}$ ( $mL.ng^{-1}$ )	0.34	0.43	127	0.43
$\omega_{1n}$ ( $10^9.L^{-1}.d^{-1}$ )	0.12	0.018	15.0	$\alpha_{1a}$ ( $L.g^{-1}$ )	0.085	0.064	75.0	0.18
$\omega_{2n}$ ( $10^9.L^{-1}.d^{-1}$ )	0.071	0.018	24.6	$\alpha_{1l}$ ( $LU^{-1}$ )	$7.8^{-6}$	0.00037	$4.7^3$	0.98
$\omega_{an}$ ( $10^9.L^{-1}$ )	0.81	0.29	36.0	$\alpha_{1NT}$ ( $L.ng^{-1}$ )	0.34	0.17	50.9	0.05
$\sigma_{bn}$	0.32	0.0070	2.2	$\beta_1$	0.29	0.065	22.5	$< 10^{-5}$
<b>arterial pH</b>				<b>Discharge</b>				
$\mu_{0p}$	7.44	0.0035	0.05	$h_2$ ( $d^{-1}$ )	0.013	0.0078	60.6	
$\mu_{1p}$ ( $d^{-1}$ )	0.0028	0.00080	28.8	$\alpha_{2n}$ ( $L.10^{-9}$ )	-0.14	0.058	42.2	0.02
$\omega_{0p}$	0.038	0.0027	7.0	$\alpha_{2p}$	25.2	6.67	26.4	0.0002
$\omega_{1p}$ ( $d^{-1}$ )	0.0052	0.00062	12.0	$\alpha_{2c}$ ( $-\log(mg.L^{-1})$ )	-1.10	0.18	16.3	$< 10^{-5}$
$\sigma_{ap}$	0.055	0.00092	1.7	$\alpha_{2d}$ ( $mL.ng^{-1}$ )	-0.00052	0.0053	1010	0.92
<b>C-reactive protein</b>				$\alpha_{2a}$ ( $L.g^{-1}$ )				
$\mu_{0c}$ ( $\log(mg.L^{-1})$ )	4.18	0.061	1.5	$\alpha_{2l}$ ( $LU^{-1}$ )	$-1.8^{-5}$	$-6.9^{-5}$	381	0.79
$\mu_{1c}$ ( $\log(mg.L^{-1}).d^{-1}$ )	-0.16	0.011	7.3	$\alpha_{2NT}$ ( $L.ng^{-1}$ )	-0.0010	0.005	451	0.84
$\omega_{0c}$ ( $\log(mg.L^{-1})$ )	0.94	0.049	5.2	$\beta_2$	-0.12	0.041	35.4	0.0034
$\omega_{1c}$ ( $\log(mg.L^{-1}).d^{-1}$ )	0.15	0.011	7.4					
$\sigma_{bc}$ ( $\log(mg.L^{-1})$ )	0.71	0.015	2.1					
<b>D-Dimers</b>								
$\mu_{0d}$ ( $ng.mL^{-1}$ )	6.75	0.072	1.1					
$\mu_{1d}$ ( $ng.mL^{-1}.d^{-1}$ )	0.033	0.011	33.5					
$\omega_{0d}$ ( $ng.mL^{-1}$ )	0.89	0.055	6.2					
$\omega_{1d}$ ( $ng.mL^{-1}.d^{-1}$ )	0.081	0.0092	11.4					
$\sigma_{bd}$	0.079	0.0025	3.1					
<b>Albuminemia</b>								
$\mu_{0a}$ ( $g.L^{-1}$ )	31.4	0.34	1.1					
$\mu_{1a}$ ( $g.L^{-1}.d^{-1}$ )	-0.16	0.040	25.5					
$\omega_{0a}$ ( $g.L^{-1}$ )	5.01	0.25	5.0					
$\omega_{1a}$ ( $g.L^{-1}.d^{-1}$ )	0.33	0.037	11.3					
$\sigma_{aa}$ ( $g.L^{-1}$ )	3.05	0.081	2.7					
<b>LDH</b>								
$\mu_{0l}$ ( $U.L^{-1}$ )	332	7.34	2.2					
$\mu_{1l}$ ( $U.L^{-1}.d^{-1}$ )	-2.36	0.96	40.7					
$\omega_{0l}$ ( $U.L^{-1}$ )	107	5.71	5.4					
$\omega_{1l}$ ( $U.L^{-1}.d^{-1}$ )	8.13	0.83	10.2					
$\sigma_{bl}$	0.22	0.0057	2.6					
<b>NTpro-BNP</b>								
$\mu_{0NT}$ ( $ng.L^{-1}$ )	5.21	0.11	2.2					
$\mu_{1NT}$ ( $ng.L^{-1}.d^{-1}$ )	0.061	0.012	20.1					
$\omega_{0NT}$ ( $ng.L^{-1}$ )	1.68	0.084	5.0					
$\omega_{1NT}$ ( $ng.L^{-1}.d^{-1}$ )	0.078	0.012	15.5					
$\sigma_{aNT}$ ( $ng.L^{-1}$ )	0.77	0.023	2.9					

## 2 Supplementary Figures

The following goodness of fit plots showed the adequacy of each longitudinal submodels.

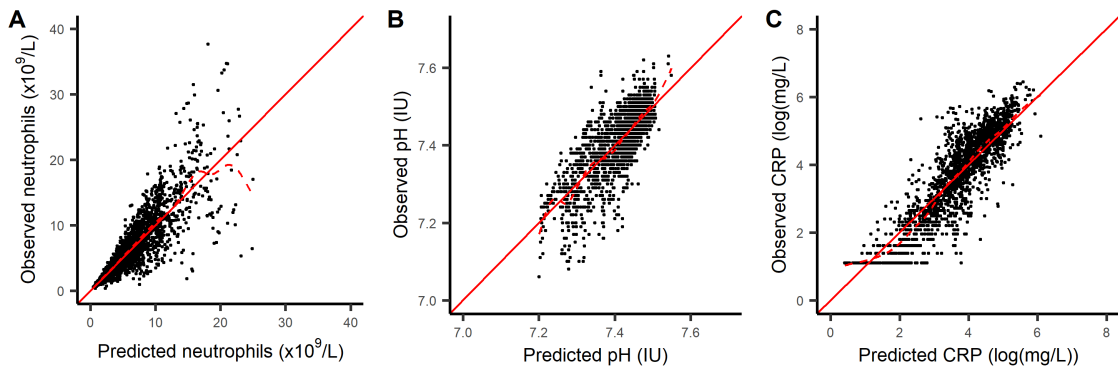


Figure S1a: Observations versus individual predictions of the different biomarkers: **A**: blood neutrophil counts, **B**: arterial pH and **C**: C-reactive protein. The solid red line represents the identity line and the dashed red line the spline line.

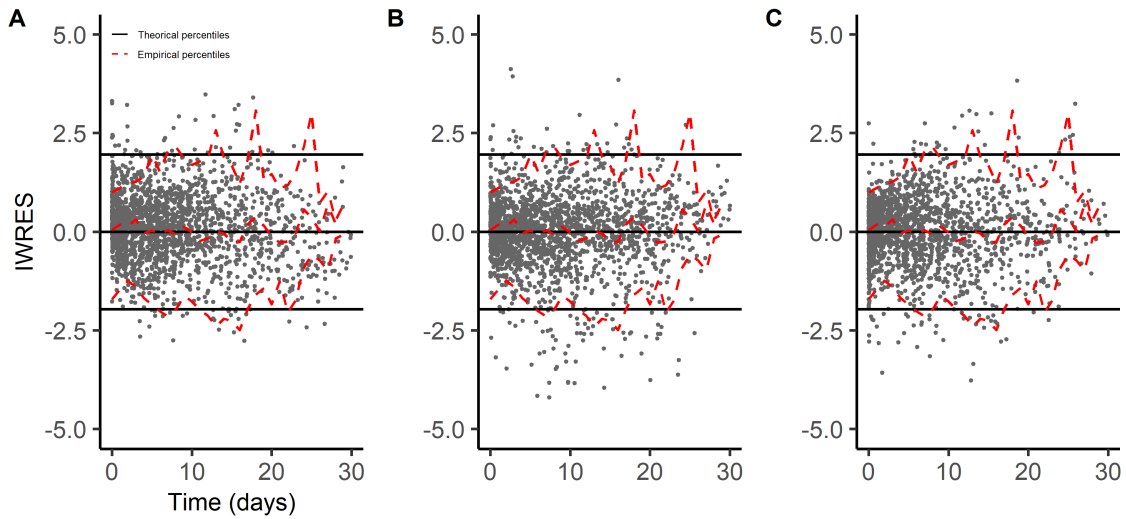


Figure S1b: Individual Weighted RESiduals (IWRES) versus time for the different biomarkers: **A**: blood neutrophil counts, **B**: arterial pH and **C**: C-reactive protein. The solid black lines denote respectively the 2.5th, 50th and 97.5th theoretical percentiles and the dashed red lines the corresponding empirical percentiles.

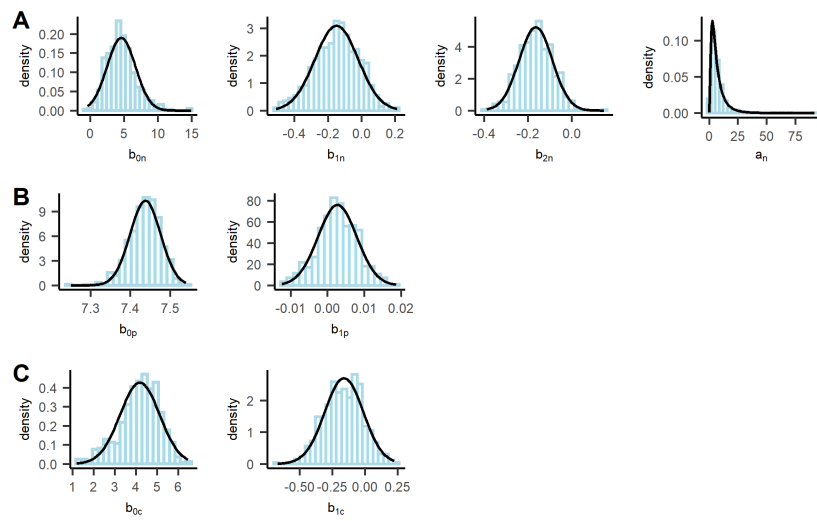


Figure S1c: Distribution of longitudinal individual parameters. **A:** ( $b_{0n}, b_{1n}, b_{2n}, a_n$ ) for blood neutrophil counts, **B:**  $b_{0p}, b_{1p}$  for arterial pH and **C:**  $b_{0c}, b_{1c}$  for C-reactive protein). Blue bars correspond to the histogram of the estimated individual parameters and the black curve to the distribution predicted using estimated means and variances. Units of parameters are the same as in Table 3

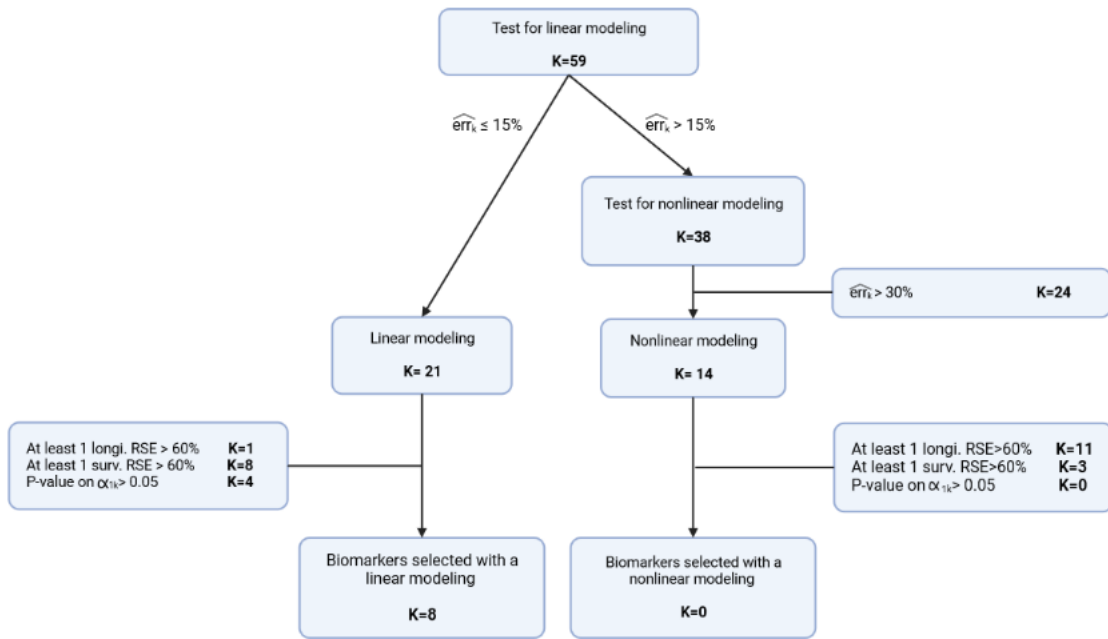


Figure S2a Flowchart under the stringent scenario

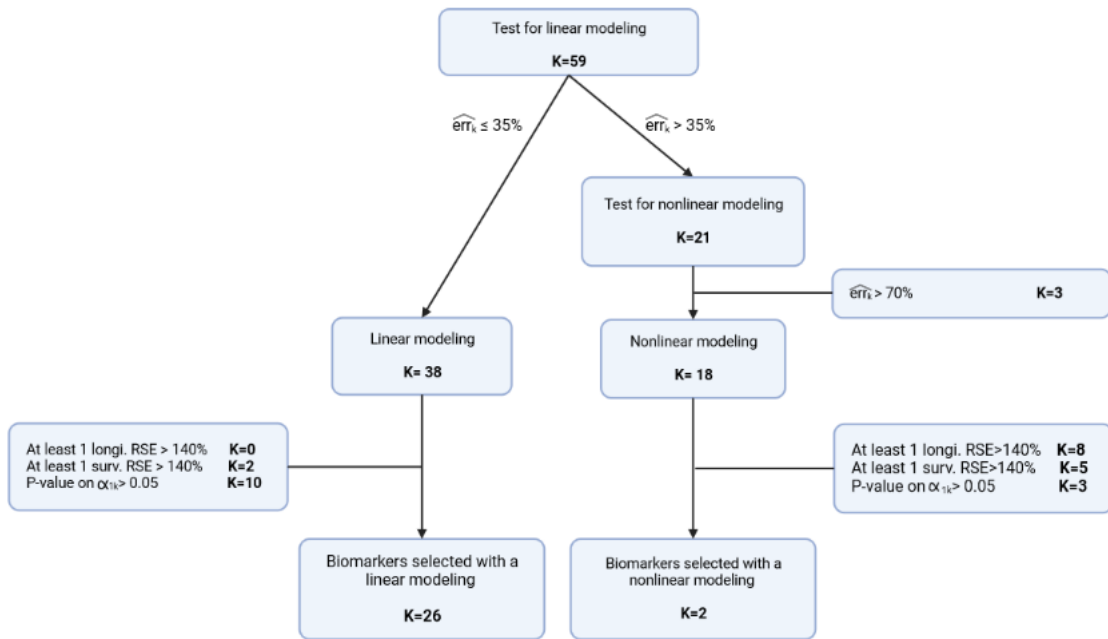


Figure S2b Flowchart under the flexible scenario

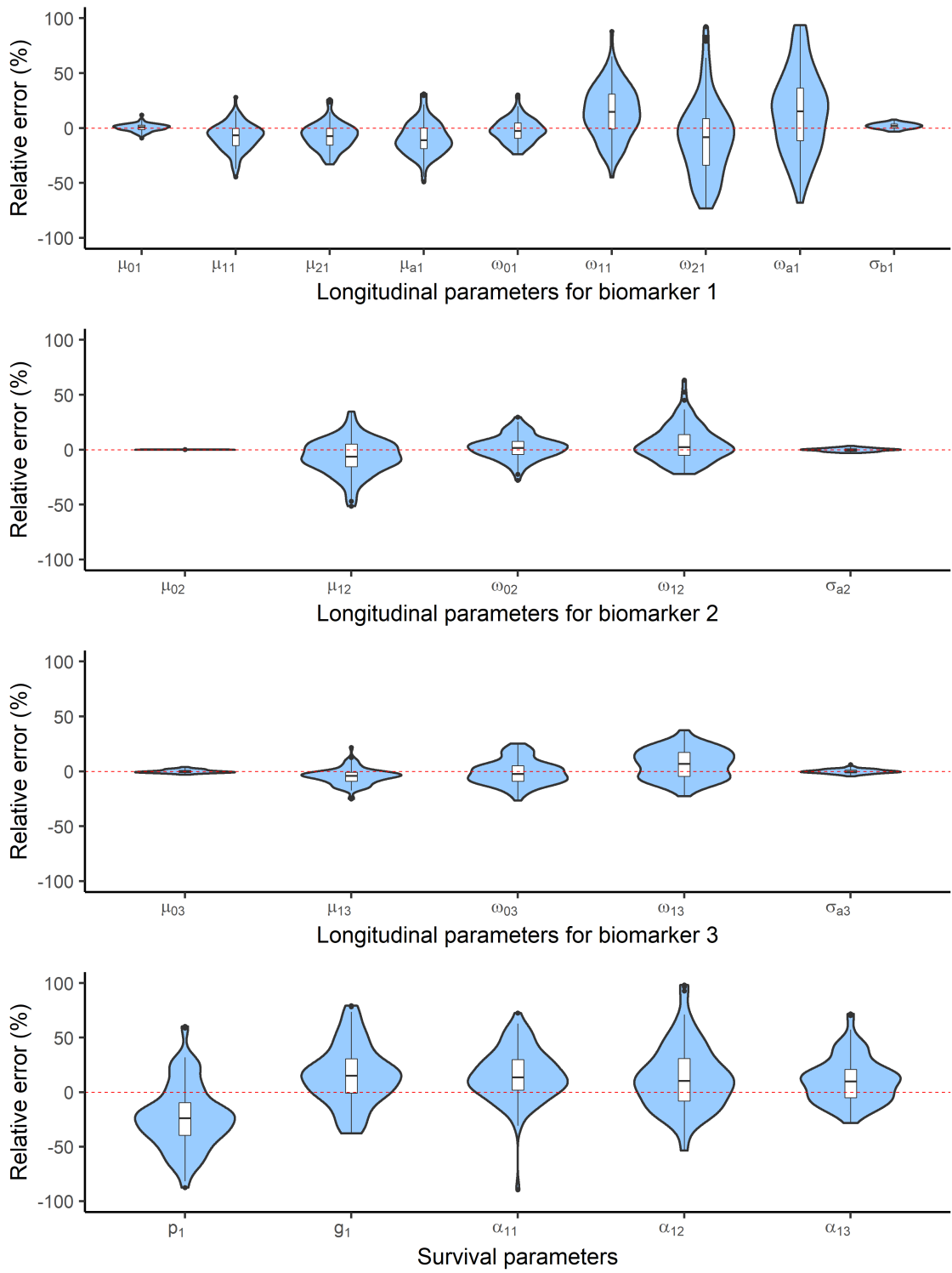


Figure S3: Distribution of relative errors of parameter estimates of the true multivariate joint model (%)

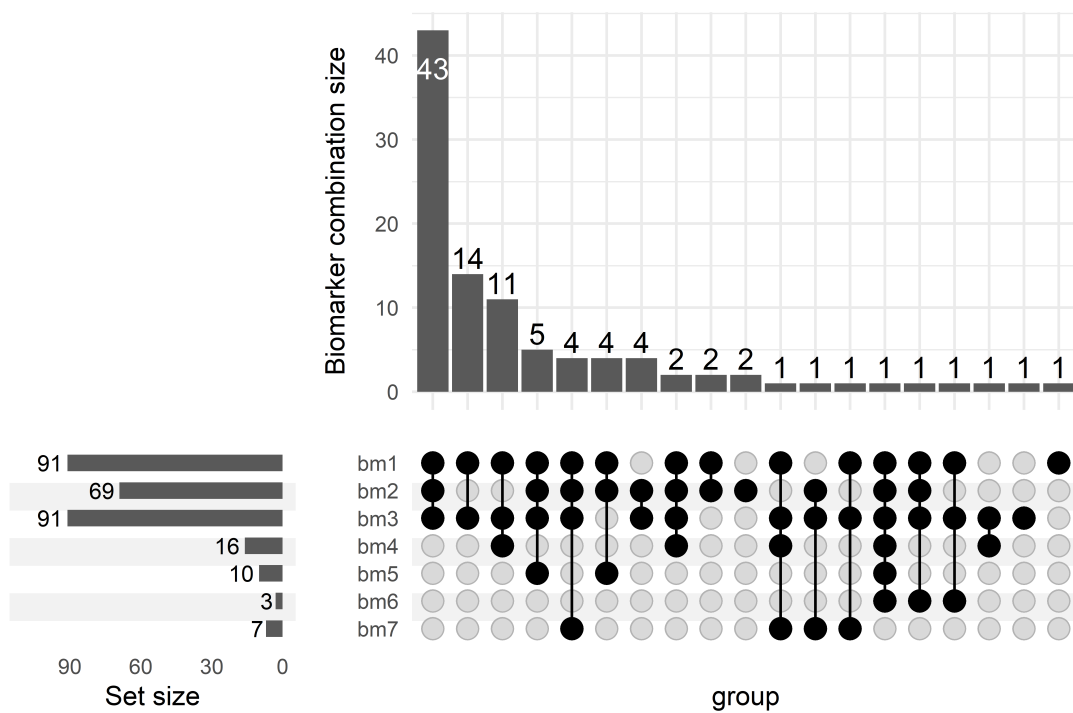


Figure S4: Upset plot for the scenario of stronger correlations.  $bm_1$ ,  $bm_2$  and  $bm_3$  are truly associated,  $bm_4$ ,  $bm_5$  are not associated but have their intercept and slope parameters correlated with  $bm_2$  and  $bm_3$  respectively, and  $bm_6$  and  $bm_7$  are not associated and uncorrelated.

### 3 Supplementary Information

#### 3.1 Supplementary Information S1 - ROC AUC indicator

The ability to separate patients according to their prognosis is assessed with time-dependent Area Under the Curve (AUC). We note  $\pi_i(l+t|l) = 1 - s_i(l+t|l)$ , the probability for patient  $i$  to die between  $l$  and  $l+t$ . At a landmark time  $l$  and an horizon time  $t$ , we define a case as a patient who died between  $l$  and  $l+t$ , and a control as a patient who did not experience any event or who was discharged between  $l$  and  $l+t$ .

In the following, we present definition and estimate of time-dependent ROC AUC according to Blanche et al. works (2013) [1].

The AUC for a landmark time  $l$  and a prediction horizon  $t$  is defined as:

$$AUC(l, t) = \mathbb{P}(\pi_i(l+t|l) > \pi_j(l+t|l) \mid D_i(l, t) = 1, D_j(l, t) = 0, T_i > l, T_j > l) \quad (1)$$

where  $D_i(l, t) = \mathbb{1}_{(l < T_i \leq l+t, K_i=1)}$ ; that is to say,  $D_i(l, t) = 1$  if the subject die between time  $l$  and  $l+t$ , and  $D_i(l, t) = 0$  if the subject is discharged between time  $l$  and  $l+t$  or is event-free at time  $l+t$ . AUC is estimated by:

$$\widehat{AUC}(l, t) = \frac{\sum_{i=1}^{N_l} \sum_{j=1}^{N_l} \mathbb{1}_{\hat{\pi}_i(l+t|l) > \hat{\pi}_j(l+t|l)} D_i(l, t) (1 - D_j(l, t)) \hat{W}_i(l, t) \hat{W}_j(l, t)}{\sum_{i=1}^{N_l} \sum_{j=1}^{N_l} D_i(l, t) (1 - D_j(l, t)) \hat{W}_i(l, t) \hat{W}_j(l, t)} \quad (2)$$

where  $N_l$  denotes the number of at-risk patients at time  $l$  and  $\hat{W}_i(l, t)$  are the weights for censored observations defined as:

$$\hat{W}_i(l, t) = \frac{\mathbb{1}_{T_i > l+t}}{\hat{G}(l+t|l)} + \frac{\mathbb{1}_{l < T_i \leq l+t} \Delta_i}{\hat{G}(T_i|l)} \quad (3)$$

$\hat{G}(x)$  is the Kaplan–Meier estimator of survival function of the censoring time at  $x$  and  $\forall x > s$ ,  $\hat{G}(x|l) = \frac{\hat{G}(x)}{\hat{G}(l)}$  estimates the conditional probability of not being censored at time  $x$  conditionally on being uncensored at time  $s$  [1].

A method for the derivation of asymptotic confidence intervals is proposed in the literature [1].

### References

- [1] Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat. Med.* 2013; 32: 5381-5397

### 3.2 Supplementary Information S2 - Details about computation time

We fitted models on a shared computing center including 8 heterogeneous machines. The less performant one includes 24 CPUs clocked at 2.2GHz (Intel Xeon E5-2430) and 96GB of RAM memory while the most performant one, includes 16 CPUs clocked at 3.8GHz (Intel Xeon Gold 5222) and 320GB of RAM memory.

To see the scalability of the algorithms according to the number of patients and biomarkers we performed two simulations. In the first one, we simulated 3 biomarkers (corresponding to bm1, bm2 and bm3 of the simulation study) with  $N=100, 200, 300, 400, 800$  or  $1600$  subjects. We reported the CPU time needed to estimate parameters, to estimate standard errors and the total of both estimations. Of note, all the following results were obtained using the most performant machine (with 16 CPUs clocked at 3.8GHz (Intel Xeon Gold 5222) and 320GB of RAM memory). As shown in Figure S5a, the CPU time needed to estimate both parameters and standard errors is almost proportional to the number of subjects  $N$ . In fact, the estimation process will take twice as long if there are twice as many subjects. The CPU time spent is driven by the estimation of parameters because the time spent to estimate standard errors seems constant depending on the number of subjects involved.

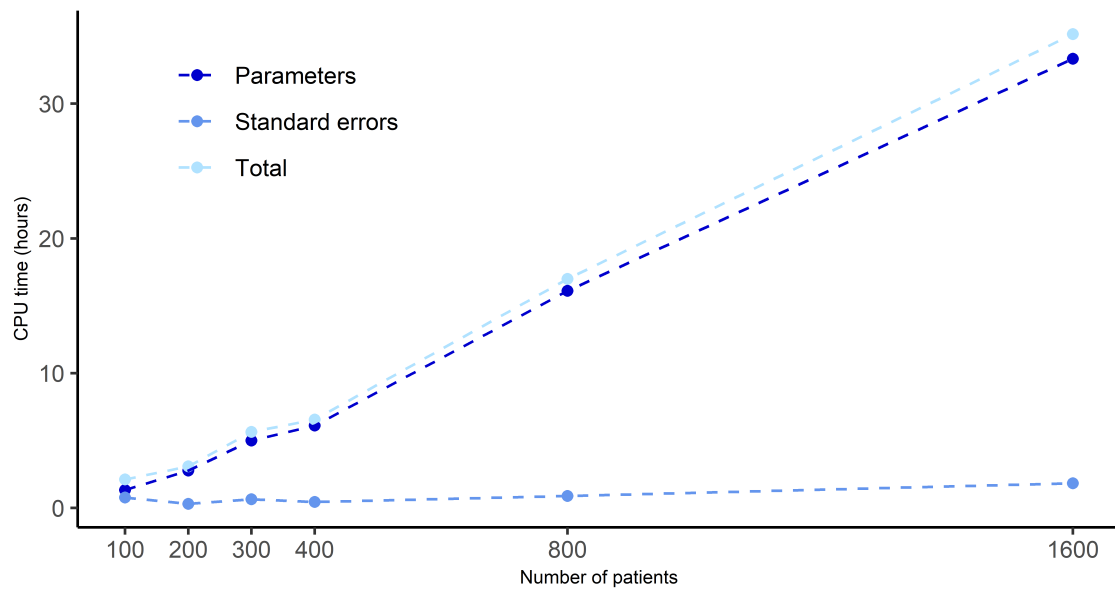


Figure S5a: CPU time spent to estimate parameters and standard errors of a multivariate joint model involving 3 biomarkers, depending on the number of subjects included

In the second simulation, we simulated joint models including  $N=300$  patients with  $K=1$  to  $K=7$  biomarkers (bm1 to bm7 of the simulation study) involved. For each joint model, we reported the CPU time needed to estimate parameters, to estimate standard errors and the total of both estimations. Figure S5b shows that the CPU time spent to estimate parameters is almost scalable according to the number of biomarkers included in the joint model. However, the CPU time spent to estimate standard errors is growing faster than the number of biomarkers (close to an exponential distribution). Hence, considering a too large set of biomarkers in the backward process would be inappropriate with this approach.

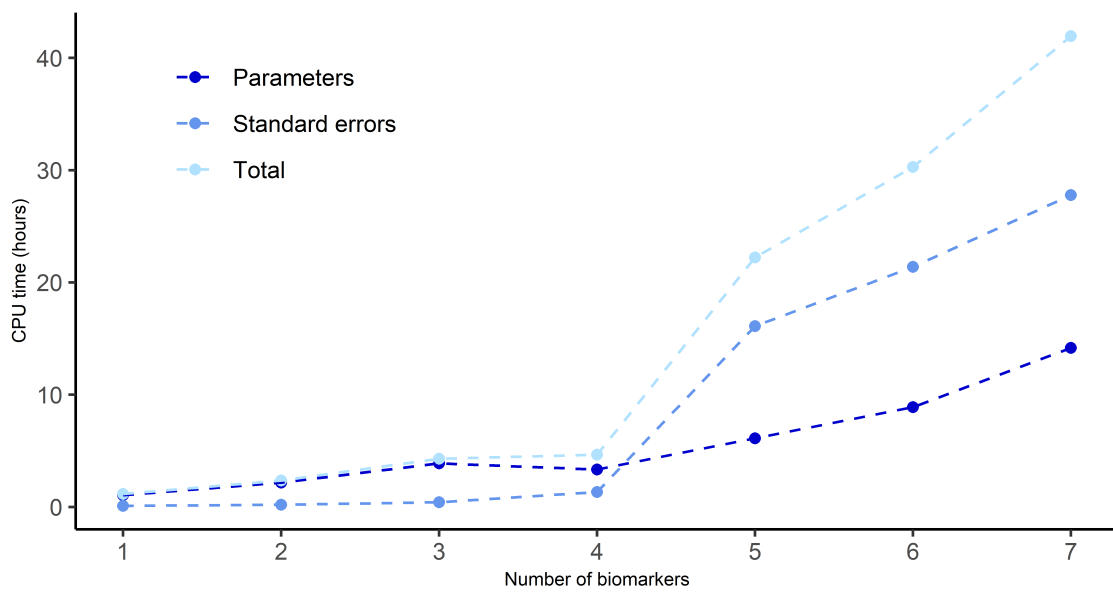


Figure S5b: CPU time spent to estimate parameters and standard errors of a multivariate joint model including 300 patients, depending on the number of biomarkers involved in the modeling

### 3.3 Supplementary Information S3 - True parameter values used in the simulation study

To recall the model considered writes as follows:

**Longitudinal submodels:**

$$y_{ij1} = m_1(t_{ij1}, \psi_{i1}) + \sigma_{b1} m_1(t_{ij1}, \psi_{i1}) \epsilon_{ij}$$

for  $k$  in {2,3,4}:

$$y_{ijk} = m_k(t_{ijk}, \psi_{ik}) + \sigma_{ak} \epsilon_{ij}$$

for  $k$  in {5,6,7}:

$$y_{ijk} = m_k(t_{ijk}, \psi_{ik}) + \sigma_{ak} m_k(t_{ijk}, \psi_{ik}) \epsilon_{ij}$$

with:

$$m_1(t_{ij1}, \psi_{i1}) = (\mu_{01} + \eta_{01i}) + (\mu_{a1} \exp(\eta_{a1i})) \times [\exp((\mu_{11} + \eta_{11i}) t_{ij1}) - \exp((\mu_{21} + \eta_{21i}) t_{ij1})]$$

for  $k$  in {2,...,7}:

$$m_k(t_{ijk}, \psi_{ik}) = (\mu_{0k} + \eta_{0ki}) + (\mu_{1k} + \eta_{1ki}) \times t_{ijk}$$

**Survival submodel:**

$$\lambda_{1i}(t|\psi_i) = \frac{p_1 g_1 \exp(-g_1 \times t)}{1 - p_1(1 - \exp(-g_1 \times t))} \exp \left[ \sum_{k=1}^3 (\alpha_{1k} \times (m_k(t, \psi_{ik}) - med_k)) \right]$$

$$\lambda_{2i}(t|\psi_i) = \frac{p_2 g_2 \exp(-g_2 \times t)}{1 - p_2(1 - \exp(-g_2 \times t))} \exp \left[ \sum_{k=1}^3 (\alpha_{2k} \times (m_k(t, \psi_{ik}) - med_k)) \right]$$

#### Longitudinal submodels

*Fixed effects*

$\mu_{01}$  4.6

$\mu_{11}$  -0.15

$\mu_{21}$  -0.16

$\mu_{a1}$  5.3

$\sigma_{b1}$  0.3

$\mu_{02}$  7.4

$\mu_{12}$  0.03

$\sigma_{a2}$  0.05

$\mu_{03}$  4.2

$\mu_{13}$  -0.16

$\sigma_{a3}$  0.7

$\mu_{04}$  5.2

$\mu_{14}$  0.02

$\sigma_{a4}$  0.7

$\mu_{05}$  7

$\mu_{15}$  -0.05

$\sigma_{b5}$  0.08

$\mu_{06}$  5.2

$\mu_{16}$  -0.07

$\sigma_{b6}$  0.1

$\mu_{07}$  338

$\mu_{17}$  -5

$\sigma_{b7}$  0.2

*Random effects*

$\omega_{01}$  2

$\omega_{11}$  0.1

$\omega_{21}$  0.07

$\omega_{a1}$  0.8

$\omega_{02}$  0.04

$\omega_{12}$  0.005

$\omega_{03}$  0.9

$\omega_{13}$  0.15

$\omega_{04}$  1.7

$\omega_{14}$  0.08

$\omega_{05}$  0.8

$\omega_{15}$  0.1

$\omega_{06}$  0.7

$\omega_{16}$  0.09

$\omega_{07}$  107

$\omega_{17}$  5

#### Survival submodel

$p_1$  0.05

$g_1$  0.1

$\alpha_{11}$  0.14

$\alpha_{12}$  -11

$\alpha_{13}$  0.6

#### Correlation (random effects)

$\rho(\eta_{12}, \eta_{14})$  0.8

$\rho(\eta_{13}, \eta_{15})$  0.8



---

# Extension du code du package libre *saemix* pour l'ajustement de modèles conjoints

---

## 4.1 Résumé

Comme présenté dans le Chapitre 1, l'objectif de ce troisième travail est de diffuser l'utilisation des modèles conjoints non-linéaires, principalement disponibles dans des logiciels issus de la pharmacométrie, aux logiciels généralistes.

Nous nous sommes intéressés à cette question en considérant le package *saemix* (Comets et al., 2017) disponible sur le CRAN (version 3.1), initialement dédié à l'estimation de modèles non-linéaires à effets mixtes. Nous avons utilisé la version en développement disponible sur Github et étendu les principales fonctions aux cas des réponses multiples et des modèles joints. L'algorithme stochastique récemment développé par Delattre and Kuhn (2023), présenté en Section 1.2.2, a été intégré pour estimer les erreurs standard des paramètres. Cet algorithme a démontré de bonnes propriétés notamment en termes computationnels mais n'a été évalué que sur des modèles à réponse unique (modèles à effets mixtes). Nous proposons une étude de simulation considérant quatre modèles joints,

composés d'une combinaison entre un modèle linéaire ou non-linéaire à effets mixtes, et un modèle de survie paramétrique avec un évènement terminal ou deux risques compétitifs. Le but est d'évaluer (i) les performances d'estimation des paramètres, (ii) l'estimation des erreurs standard, et (iii) l'erreur de type-I lors du test d'indépendance entre les sous-modèles longitudinaux et de survie. Pour chacun des quatre modèles proposés, 100 jeux de données de 100 patients ont été simulés, avec une mesure de biomarqueur reportée chaque jour jusqu'à l'évènement. L'estimation des paramètres est évaluée à l'aide des biais relatifs (RB), des erreurs quadratiques moyennes relatives (RRMSE) et des erreurs d'estimation relatives (REE). L'estimation des erreurs standard est évaluée en comparant la distribution des erreurs standard obtenues par l'algorithme stochastique, avec les erreurs standard empiriques. Pour le troisième objectif, un test de Wald est effectué pour tester la significativité du coefficient de lien, et l'erreur de type-I de ce test est reportée.

Cette étude de simulation a démontré de bonnes performances en termes d'estimation des paramètres et des erreurs standard, malgré une plus grande incertitude sur les paramètres associés aux variances des effets aléatoires dans le cadre de modèles joints non-linéaires. Cependant, cette tendance avait été observée dans les précédents travaux lorsque l'estimation était effectuée sur le logiciel Monolix. L'erreur de type-I lors du test d'indépendance entre les deux sous-modèles est également contrôlée pour chacun des modèles joints.

Ces travaux ouvrent la voie à l'estimation de modèles conjoints sophistiqués et personnalisés car l'utilisateur définit lui même la vraisemblance du modèle. L'ensemble du code et des fonctions étendues sont disponibles sur Github à l'adresse suivante : <https://github.com/saemixdevelopment/saemixextension/tree/master/joint>. Des travaux restent à mener pour étendre l'ensemble du package, notamment, développer des outils diagnostiques du modèle. Il serait également intéressant de mener des études comparatives, entre Monolix et la version étendue de *saemix* en termes de temps de calcul et de précision d'estimation, notamment des erreurs standard.

Ce travail a donné lieu à un article soumis dans *Computer Methods and Programs in*

*Biomedicine*, le 5 juillet 2023.

## **4.2 Article 3 (soumis)**

## Extending the code in the open-source *saemix* package to fit joint models

Alexandra Lavalley-Morelle<sup>a,\*</sup>, France Mentré<sup>a,b</sup>, Emmanuelle Comets<sup>a,c,1</sup>, Jimmy Mullaert<sup>a,b,1</sup>

<sup>a</sup> Université Paris Cité, INSERM, IAME, F-75018 Paris, France

<sup>b</sup> Department of Epidemiology, Biostatistics and Clinical Research, AP-HP, Bichat-Claude Bernard University Hospital, F-75018 Paris

<sup>c</sup> Université de Rennes, Inserm, EHESP, Irset - UMRS 1085, F-35000 Rennes, France

---

### Abstract

**Background and Objective:** Joint modeling of longitudinal and time-to-event data has received considerable attention over recent years. Developments about joint models have been extensive, with complex parametrizations such as nonlinear models for longitudinal outcomes and flexible time-to-event models for survival outcomes, possibly involving competing risks. However, in most popular software such as R, only limited possibilities in terms of modeling are available, (e.g. linear longitudinal model or pre-defined survival models). Hence, the global objective of this work is to extend the code from the *saemix* package (version 3.1 on CRAN) to allow fitting a wide range of parametric joint models fully defined by the user.

**Methods:** We used the *saemix* package, designed to fit nonlinear mixed-effects models, and extended the main functions in the Stochastic Approximation Expectation Maximization (SAEM) algorithm to joint model estimation. To compute standard errors of parameter estimates, we implemented a recently developed stochastic algorithm. A simulation study is proposed to assess (i) the performances of parameter estimation, (ii) the standard error computation and (iii) the type I error when testing independence between the two submodels. Four joint models are considered in the simulation study, combining a linear or nonlinear mixed-effects model for the longitudinal submodel, with a single terminal event or a competing risk model.

**Results:** For all simulation scenarios, parameters are precisely and accurately estimated with low bias and uncertainty. The empirical standard errors of parameters obtained over all simulations are very close to those computed with the stochastic algorithm. For more complex joint models (involving a nonlinear longitudinal model), some estimates of random effects variances have higher uncertainty and their standard errors are moderately under-estimated. Finally, the type-I error is controlled for each joint model.

**Conclusions:** *saemix* is a flexible open-source package that we tweaked to extend it to fit various parametric joint models. Further developments would be to extend functions used to derive diagnostic tools.

**Key words:** competing risks, joint modeling, mixed-effects model, SAEM algorithm, *saemix*, time-to-event

---

\*Corresponding author: alexandra.lavalley-morelle@inserm.fr

<sup>1</sup>Both authors contributed equally to the work

## 1. Introduction

In many clinical applications, both longitudinal and survival outcomes are observed. Different approaches exist to analyse separately or jointly both outcomes [1, 2]. A well established one is the joint modeling approach [2], which consists in estimating simultaneously the parameters of both models linked by shared random effects. This approach has been extensively studied over the last decade [3, 4, 5]. The first developments for joint models involved a linear mixed-effects model describing a single biomarker evolution and a time-to-event model (Cox proportional hazard model) describing the instantaneous risk of event [6]. Nowadays, many extensions have been proposed including nonlinear mixed-effects models [3], multiple longitudinal models [7], categorical/count data [8] or also competing risks [5].

Softwares for the joint analysis of longitudinal and event time data have been available for many years. In what follows, we present a non exhaustive list of tools dedicated to joint modeling in a frequentist framework, focusing on more popular softwares such as R, SAS, and Stata. With respect to R software, a number of packages are available. First, the *joineR* package [9] can be considered for the joint analysis of a continuous longitudinal outcome with a time-to-event outcome. Only a linear modeling and a Cox proportional hazard parametrization (semi-parametric specification) are handled for the longitudinal and the survival outcome respectively. This package also allows for the competing risk setting under cause-specific approach (with unspecified baseline risks). To allow for a parametric specification on the time-to-event model, the *JM* package [10] can be used with different baseline hazard parametrizations. It also allows for competing risk setting assuming relative risk model for each competing event (with piecewise proportional hazards model). The continuous longitudinal outcome is handled by a linear modeling (possibly involving splines). To handle multiple continuous longitudinal outcomes, the package *joineRML* [11] can be used to fit joint models with multivariate longitudinal data and a single time-to-event outcome using Cox parametrization. An other class of joint models allows to stratify the population into classes of patients that are homogeneous both with respect to the evolution of the biomarker(s) and to the occurrence of the event(s). Those joint latent class models are available using *lcmm* package [12]) which allows for a linear (possibly using splines) modeling of a single or multiple continuous longitudinal outcomes and different parametrizations of the survival outcome (semi-parametric or parametric form, possibly including competing risks). All the previous mentioned packages only allow for a linear modeling of the biomarker dynamics, possibly including linear combination of splines functions. To the best of our knowledge, only the package *frailtypack* [13] allows for a joint model with a mechanistic nonlinear model and possibly recurrent and/or terminal events for the survival model. However, only one expression for the nonlinear longitudinal model is currently implemented and no other expressions can be considered. With respect to SAS software, two macros may be used. The macros *%JM* [14] and *%JMfit* [15] are designed to fit joint model of a single longitudinal outcome (continuous, binary, categorical or count data) and a single time-to-event outcome using parametric parametrization. A similar tool is available on Stata software, with the package *stjm* [16] but only allowing for a continuous longitudinal outcome.

All of these tools reach limitations when it comes to the description of biomarker dynamics because mainly linear functions or pre-defined ODE can be used. We acknowledge that some previous mentioned packages allow for the use of a linear combination of spline functions that give more flexibility on the modeling. However, if the biomarker evolution need to be investigated with various mechanistic models, these packages may not represent ideal choices. In that case, pharmacometrics software such as NONMEM

[17] and Monolix [18] may be a good alternative. They allow to make very flexible nonlinear joint models and recent studies also used it under complex survival settings such as competing risks [19, 5].

An important step in parameter estimation is to provide an associated uncertainty. The observed Fisher information matrix (FIM) is usually computed to obtain the variance-covariance matrix associated with maximum-likelihood estimates. Several methods have been described in the literature to compute exactly the FIM or to approximate it [20, 21, 22]. In the context of nonlinear joint models involving survival data, the most common approach is to use a stochastic approximation algorithm using Louis decomposition principle [20]. This approach may be not optimal in a computational point of view in particular because the algorithm involves the computation of second derivative of the complete data likelihood (see [23] for more details). Recently, Delattre and Kuhn proposed a new numerical method to get a stochastic approximation of the observed FIM in latent variables model [24]. This method has interesting properties, in particular it only requires the calculation of the first derivatives of the complete log-likelihood and it is fully integrated within the SAEM algorithm. While it has been evaluated for nonlinear mixed-effects models and mixture models [24], this is not the case for joint models.

In this work, we extended the main functions of the *saemix* package, an open source implementation in R of the SAEM [25] algorithm which provides a flexible and fast algorithm to estimate population parameters in nonlinear mixed-effects models, to the case of multiple responses and joint models. We integrated the stochastic algorithm recently developed by Delattre and Kuhn [24] to compute standard errors of parameter estimates. This extension allows us to leverage a tool in a widely used software to fit a large range of joint models (linear/nonlinear continuous outcome(s) with parametric survival model (including competing risks or not) fully defined by the user. We present a simulation study to show the good properties of estimation of both parameters and standard errors, and a small real-case study based on data available in R software.

This paper is organized as follows: Section 2 describes the methods referring to the *saemix* extension, with the parameters and standard errors estimation procedures, and then the simulation study. In Section 3, we present the results of the simulation study, and also the results of an application based on real data. We close the paper with a discussion in Section 4.

## 2. Methods

### 2.1. Description of the SAEM algorithm for standard mixed-effects models

A mixed-effects model describing the vector of observed data  $y = (y_{ij}; 1 \leq i \leq N, 1 \leq j \leq n_i)$  is commonly defined by the following equation:

$$y_{ij} = m(t_{ij}, \psi_i) + g(t_{ij}, \psi_i, \sigma) \epsilon_{ij} \quad (1)$$

where  $m$  is the structural function describing the observations,  $\psi = (\psi_i; 1 \leq i \leq N)$  the vector of unobserved individual parameters composed of fixed and individual random effects, and  $\epsilon_{ij}$  the gaussian residual error with mean 0 and variance 1 (independent and identically distributed). The objective is to estimate the vector of parameters  $\theta = (\mu, \Omega, \sigma)$ , with  $\mu$  the fixed-effects parameters,  $\Omega$  the variance-covariance matrix of random effects, and  $\sigma$  the residual error model parameters by maximizing the likelihood defined

as an integral over the random effects and noted  $l(y; \theta)$ :

$$l_y(y; \theta) = \prod_i \int_{\mathcal{D}_{\psi_i}} \prod_j p(y_{ij} | \psi_i; \theta) p(\psi_i; \theta) d\psi_i \quad (2)$$

In the case of linear mixed-effects models (LMEM), the estimation can be treated with the regular iterative EM algorithm [26]. At iteration  $k$ , data are completed by drawing individual parameters realizations, the E-step computes the conditional expectation of the complete log-likelihood  $Q_k(\theta) = \mathbb{E}(\log l(y, \psi; \theta) | y; \theta_{k-1})$ , and the M-step computes the value  $\theta_k$  which maximize  $Q_k(\theta)$ . Note, the complete data of the model  $(y, \psi)$  is composed of the vector of observations and unobserved individual parameters and the complete likelihood can be expressed as:

$$l(y, \psi; \theta) = \prod_{i,j} p(y_{ij} | \psi_i; \theta) \prod_i p(\psi_i; \theta) \quad (3)$$

In the cases of nonlinear mixed-effects model (NLMEM), where the structural function does not linearly depend on the random effects, the E-step cannot be computed in a closed-form. An alternative is to use a stochastic version of the EM algorithm, the SAEM algorithm, divided in three steps. At iteration  $k$  of the algorithm:

1. *Simulation-step* : draw  $\psi^{(k)}$  from the conditional distribution  $p(\cdot | y; \theta_k)$ .
2. *Stochastic approximation* : update  $Q_k(\theta)$  according to

$$Q_k(\theta) = (1 - \gamma_k) Q_{k-1}(\theta) + \gamma_k (\log l(y, \psi^{(k)}; \theta)) \quad (4)$$

where  $(\gamma_k)$  is a decreasing sequence of positive numbers with  $\gamma_1 = 1$ .

3. *Maximization-step* : update  $\theta_k$  according to

$$\theta_{k+1} = \text{Arg max}_{\theta} Q_k(\theta)$$

In practice, for the mixed-effects model defined in 1, we have:

$$p(y_{ij} | \psi_i; \theta) = \frac{1}{\sqrt{2\pi}g(t_{ij}, \psi_i, \sigma)} \exp\left(-\frac{(y_{ij} - m(t_{ij}, \psi_i))^2}{2g(t_{ij}, \psi_i, \sigma)^2}\right),$$

while the density of individual parameters can be expressed in the case of general  $\Omega$  (non-diagonal) as:

$$p(\psi_i; \theta) = \frac{1}{(2\pi)^{D/2} |\Omega|^{1/2}} \exp\left(-\frac{1}{2} (\psi_i - \mu)^T \Omega^{-1} (\psi_i - \mu)\right)$$

where  $D$  is the number of random effects, and  $|\cdot|$  refers to the matrix determinant.

For simple forms of  $g$  including the constant and proportional residual error model, the complete likelihood therefore belongs to the curved exponential family and can be written as:

$$l(y, \psi; \theta) = \exp(-\Psi(\theta) + \langle S(y, \psi), \phi(\theta) \rangle) \quad (5)$$

The stochastic approximation (step 2) of the algorithm reduces to updating the sufficient statistics (at iteration  $k$ ):

$$\begin{aligned} S_{1,k} &= (1 - \gamma_k) S_{1,k-1} + \gamma_k \sum_{i,j} (y_{ij} - m(t_{ij}, \psi_i^{(k)}))^2 \\ S_{2,k} &= (1 - \gamma_k) S_{2,k-1} + \gamma_k \sum_i \psi_i^{(k)} \\ S_{3,k} &= (1 - \gamma_k) S_{3,k-1} + \gamma_k \sum_i \psi_i^{(k)} \psi_i^{(k)T} \end{aligned}$$

An example of the derivation of sufficient statistics for a gaussian mixed-effects model is given in Appendix A.1.

For a general residual variance model  $g = \sigma_a + \sigma_b m^{\sigma_c}$ , the complete model does not belong to the exponential family and an additional optimisation step is required to estimate the variance of error parameters.

## 2.2. Extension of the algorithm for joint models

We now suppose that we also observe survival data. Let  $T$  the variable describing the time-to-event distribution,  $C$  the non-informative censoring distribution and  $\delta$  the event indicator. The couple  $(\tilde{T}, \delta)$  is observed with  $\tilde{T} = \min(T, C)$ . In *saemix*, we only consider parametric survival models. To introduce joint model principle, we first present the case of a single terminal event, but more complex frameworks such as competing risks may also be considered. In that case, a joint model describing simultaneously longitudinal data by a mixed-effects model and survival data by parametric function, is usually defined by:

$$h(t | \psi_i) = h_0(t) \times \exp(\alpha m(t, \psi_i)) \quad (6)$$

along with the definition of  $m(t, \psi_i)$  in equation 1.  $h(t | \psi_i)$  is the instantaneous risk of event with  $h_0(t)$  defining the baseline hazard function with vector of parameters  $h_b$ , and  $\alpha$  being the link coefficient. Here, the vector of population parameters to estimate becomes  $\theta = (\mu, \Omega, \sigma, h_b, \alpha)$ . The complete data is now  $(y, \psi, T, \delta)$  and its likelihood defined by:

$$l(y, T, \delta, \psi; \theta) = \prod_{i,j} p(y_{ij} | \psi_i; \theta) \prod_i p(T_i, \delta_i | \psi_i; \theta) \prod_i p(\psi_i; \theta) \quad (7)$$

The 3 steps of the SAEM algorithm remain as described in the previous subsection. However, the log likelihood changed and has an additional term corresponding to the contribution of the survival part (see equation 7). Survival parameters usually only have a population value and are free from random effects. In *saemix*, we however use the exponentiation approach proposed by Kuhn and Lavielle [23]. This trick consists in adding artificial decreasing variability to these parameters in the first  $K_1$  iterations of the algorithm (exploratory phase) and forcing it to decrease progressively to 0 before the final smoothing phase. Using the exponentiation trick, the complete joint likelihood can be expressed as a function belonging to the curved exponential family. This adaptive method demonstrates well convergence properties [27]. An example of the derivation of sufficient statistics for a joint model is given in Appendix A.2. During the smoothing phase (after iteration  $K_1$ ), an optimization algorithm is used at each iteration (Nelder-Mead simplex algorithm [28]) to get estimated for survival parameters.

### 2.3. Computation of standard errors

Standard errors of parameter estimates are derived from the inverse of the observed Fisher Information Matrix (FIM), which is defined in the following equation:

$$\begin{aligned} I(\theta) &= -\partial_{\theta}^2 \log l_{y, \tilde{T}, \delta}(y, \tilde{T}, \delta; \theta) \\ &= \left( \partial_{\theta} \log l_{y, \tilde{T}, \delta}(y, \tilde{T}, \delta; \theta) \right) \left( \partial_{\theta} \log l_{y, \tilde{T}, \delta}(y, \tilde{T}, \delta; \theta) \right)^t \end{aligned} \quad (8)$$

As the equation 8 involves the marginal log likelihood, not available using the SAEM algorithm, a recent work [24] proposed a decomposition in the same spirit as the Louis principle [20], based on the score functions of the complete log likelihood:

$$I_{sco}(\theta) = \mathbb{E}_{\psi|y, \tilde{T}, \delta; \theta} \left[ \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \right] \mathbb{E}_{\psi|y, \tilde{T}, \delta; \theta} \left[ \left( \partial_{\theta} \log l(y, \tilde{T}, \delta, \psi; \theta) \right)^t \right] \quad (9)$$

This decomposition only requires the calculation of the first derivatives of the complete log likelihood and is therefore interesting in a computational point of view. They also proposed a stochastic approximation method to evaluate this quantity that can be easily integrated within the SAEM algorithm as follows:

Initialize  $\Delta_i^0 = O_D$  (zero matrix of dimension  $D$ ), then

1. *Simulation-step* : draw  $\psi^{(k)}$  from the conditional distribution  $p(\cdot|y; \theta_k)$ .
2. *Stochastic approximations* : update  $Q_k(\theta)$  and  $\Delta^k$  according to

$$\begin{aligned} Q_k(\theta) &= (1 - \gamma_k) Q_{k-1}(\theta) + \gamma_k (\log l(y, \tilde{T}, \delta, \psi^{(k)}; \theta)) \\ \Delta_k &= (1 - \gamma_k) \Delta_{k-1} + \gamma_k (\partial_{\theta} \log l(y, \tilde{T}, \delta, \psi^{(k)}; \theta_k)) \end{aligned}$$

3. *Maximization-step* : update  $\theta_k$  according to

$$\theta_{k+1} = \text{Arg max}_{\theta} Q_k(\theta)$$

When convergence is reached (iteration  $K$ ), we obtain  $\hat{\theta}$  the maximum likelihood estimator of  $\theta$  and the approximation of the FIM  $I_{sco}(\hat{\theta})$  based on the score functions is given by:

$$I_{sco}(\hat{\theta}) = \Delta_K \Delta_K^t \quad (10)$$

We implemented this stochastic procedure to evaluate the observed FIM and then derive the variance-covariance matrix of  $\hat{\theta}$  defined by  $\hat{\Sigma} = I_{sco}(\hat{\theta})^{-1}$ .

### 2.4. Extending the code in saemix to joint models

In this work, we used the development version of *saemix* available on Github, based on version 3.1 available on CRAN, and modified the main estimation functions to simultaneously fit several outcomes, and handle joint models. We also implemented

the algorithm developed by Delattre and Kuhn [24] to estimate standard errors of parameters. As users can themselves define the likelihood of the model, *saemix* allows considerable modeling flexibility in parametric framework.

Our modified code along with examples to use our extensions are available on Github at the following address: <https://github.com/saemixdevelopment/saemixextension/tree/master/joint>. To guide the user, an example for joint model estimation using the code is presented in Appendix C and on Github (file *test\_realdataJM.R*).

In the following, we evaluate the performances of our extension for computing parameter estimates as well as standard errors estimates. This evaluation is based on a simulation study covering different scenarios including competing risk settings.

## 2.5. Simulation study

### 2.5.1. Aims

The objective of the simulation is threefold. Firstly, we aim to assess the new *saemix* extension for the estimation of six different models: a linear mixed-effects model (LMEM), a nonlinear mixed-effect model (NLMEM), a joint model with a linear mixed-effects model and a single time-to-event model (JM LMEM-TTE) or a competing risks model (JM LMEM-CR), and a joint model with a nonlinear mixed-effects model and a single time-to-event model (JM NLMEM-TTE) or a competing risks model (JM NLMEM-CR). Secondly, we aim to assess the performances of the standard errors computation for these models. Finally, in the cases of joint models, we aim to assess the type-I error when testing independence between submodels. It is of particular interest in the context of joint modeling to use a Wald test to assess the effect of the biomarker evolution on the risk of event. We used simulations to determine whether the type-I error of this test is controlled.

### 2.5.2. Data-generating mechanism

For each of the 6 models presented above, we simulate  $M = 100$  datasets of  $N = 100$  patients. We assume a rich design and parameters are chosen in order to have about 50% of failures for single event models; about 45% of failures from event 1 and 45% of failures from event 2 for competing risks models. All details on parameter values are summarized in Table 1. Code for generating datasets is provided in Appendix B.1. We now detail the equations of each model.

#### LMEM and NLMEM

We assume one biomarker daily measured from  $t=0$  to  $t=30$ , with  $\{y_{i1}, \dots, y_{i30}\}$  the vector of longitudinal observations of subject  $i$  (for  $i = 1, \dots, N$ ). In the case of the LMEM, we define:

$$\begin{aligned} y_{ij} &= m_i(t_{ij}, \psi_i) + g[m_i(t_{ij}, \psi_i), \sigma] \epsilon_{ij} \\ &= \psi_{i0} + \psi_{i1} \times t_{ij} + \sigma \epsilon_{ij} \\ &= (\mu_0 + \eta_{0i}) + (\mu_1 + \eta_{1i}) \times t_{ij} + \sigma \epsilon_{ij} \end{aligned} \quad (11)$$

$$\text{with } \begin{pmatrix} \eta_{i0} \\ \eta_{i1} \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \omega_0^2 & 0 \\ 0 & \omega_1^2 \end{pmatrix} \right)$$

In the case of the NLMEM, we define:

$$\begin{aligned}
 y_{ij} &= m_{nl}(t_{ij}, \psi_i) + g[m_{nl}(t_{ij}, \psi_i), \sigma] \epsilon_{ij} \\
 &= \psi_{i0} + \psi_{ia} \times [\exp(-\psi_{i1} t_{ij}) - \exp(-\psi_{i2} t_{ij})] + \sigma \epsilon_{ij} \\
 &= (\mu_0 + \eta_{0i}) + (\mu_a \exp(\eta_{ai})) \times [\exp(-(\mu_1 \exp(\eta_{1i})) t_{ij}) - \exp(-(\mu_2 \exp(\eta_{2i})) t_{ij})] + \sigma \epsilon_{ij}
 \end{aligned} \tag{12}$$

$$\text{with } \begin{pmatrix} \eta_{i0} \\ \eta_{i1} \\ \eta_{i2} \\ \eta_{ia} \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \omega_0^2 & 0 & 0 & 0 \\ 0 & \omega_1^2 & 0 & 0 \\ 0 & 0 & \omega_2^2 & 0 \\ 0 & 0 & 0 & \omega_a^2 \end{pmatrix} \right)$$

#### *JM with a LMEM/NLMEM and a time-to-event model*

We suppose here that biomarker measurements are also available every day. Measurements are reported until time to event or end of the study period ( $t=30$ ). Longitudinal data are simulated according to equation (11) or (12) (LMEM or NLMEM). Time-to-event data are simulated given the cumulative incidence function defined as follows:

$$F_i(t) = \mathbb{P}(T_i < t | \psi_i; \theta) = 1 - \exp\left(-\int_0^t h(s, \psi_i) ds\right) \tag{13}$$

$$h(t, \psi_i) = \begin{cases} h_0 \times \exp(\alpha m_i(t, \psi_i)) & \text{for the linear model} \\ h_0 \times \exp(\alpha m_{nl}(t, \psi_i)) & \text{for the nonlinear model} \end{cases}$$

where  $h_0$  is the baseline constant risk and  $\alpha$  the coefficient linking the current predicted biomarker value to the instantaneous risk of event. For each simulated patients, longitudinal observations after the simulated event times are dropped.

#### *JM with a LMEM/NLMEM and a competing risks model (subdistribution approach)*

As previously, biomarker measurements are available every day, for at most  $t=30$ . We now assume that measurements are reported until time to event 1 (event of interest), time to event 2 (competing event), or end of the study period ( $t=30$ ). We consider the subdistribution hazard framework to model the competing risks. In that case, we define  $T_e = T \times \mathbb{1}_{\delta=e} + \infty \times \mathbb{1}_{\delta \neq e}$  for  $e = \{1, 2\}$ . The cumulative incidence function for the event 1 is given by:

$$F_1(t) = \mathbb{P}(T_{1i} < t | \psi_i; \theta) = 1 - \exp\left(-\int_0^t h_1(s, \psi_i) ds\right) \tag{14}$$

where

$$h_{1i}(t, \psi_i; \theta) = \frac{p_1 g_1 \exp(-g_1 t)}{1 - p_1 (1 - \exp(-g_1 t))} \lambda(t) \tag{15}$$

$$\lambda(t) = \begin{cases} \exp(\alpha \times m_l(t, \psi_i)) & \text{for the linear model} \\ \exp(\alpha \times m_{nl}(t, \psi_i)) & \text{for the nonlinear model} \end{cases}$$

The subdistribution for the competing risk is then obtained using an exponential distribution with rate  $\frac{t}{b}$ :

$$\mathbb{P}(T_{2i} < t | \psi_i; \theta) = (1 - F_1(\infty)) \times \left(1 - \exp\left(-\frac{t}{b}\right)\right) \tag{16}$$

and the instantaneous hazard is derived from the previous equation:

$$h_{2i}(t, \psi_i; \theta) = \frac{1}{b} \frac{(1 - F_1(\infty)) \exp(-t/b)}{1 - (1 - F_1(\infty))(1 - \exp(-t/b))} \tag{17}$$

	LMEM	NLMEM	JM LMEM-TTE	JM NLMEM-TTE	JM LMEM-CR	JM NLMEM-CR
<i>Fixed effects</i>			<i>Survival model</i>			
$\mu_0$	4	4	$h_0$	0.005	-	-
$\mu_1$	0.20	0.15	$p_1$	-	0.15	0.15
$\mu_2$	-	0.20	$g_1$	-	0.10	0.10
$\mu_a$	-	10	$\alpha$	0.20	0.20	0.30
			$b$	-	15	15
<i>Random effects</i>						
$\omega_0^2$	4	4				
$\omega_1^2$	0.10	0.10				
$\omega_2^2$	-	0.10				
$\omega_a^2$	-	0.50				
<i>Error model</i>						
$\sigma$	1	1				

Table 1: Values used to simulate longitudinal data (left) and survival data (right).

### JM with no-link

In order to evaluate the type-I error of the Wald test on the link function, we simulated  $M = 1000$  datasets considering models presented in section 2.5.2, but with  $\alpha = 0$ . We also changed  $h_0$  and  $p_1$  values to keep the same proportion of failures for each model. Hence,  $h_0$  was set to 0.025 and  $p_1$  to 0.50. For each of the replicates, the tested hypothesis are:  $H_0 : \alpha = 0$  vs  $H_1 : \alpha \neq 0$ .

### 2.5.3. Estimands

For each model and simulated dataset, we estimate  $\theta$ , the vector of population parameters:

- $\theta = \{\mu, \Omega, \sigma\}$  for LMEM/NLMEM;
- $\theta = \{\mu, \Omega, \sigma, h_0, \alpha\}$  for JM with LMEM/NLMEM and TTE;
- $\theta = \{\mu, \Omega, \sigma, p_1, g_1, \alpha, b\}$  for JM with LMEM/NLMEM and CR

and we also compute  $\hat{\Sigma}$ , the variance-covariance matrix of  $\hat{\theta}$  according to the method described in section 2.3.

#### 2.5.4. Performance measures

In the following,  $\theta$  denotes the true set of parameters used to simulate data. For the first objective of assessing estimation performances, we report Relative Bias (RB), Relative Root Mean Square Errors (RRMSE) and Relative Estimation Errors (REE).

$$\text{RB}(\hat{\theta}) = \frac{1}{M} \sum_{m=1}^M \frac{\hat{\theta}^{(m)} - \theta}{\theta} \times 100 \quad \text{RRMSE}(\hat{\theta}) = \sqrt{\frac{1}{M} \sum_{m=1}^M \left( \frac{\hat{\theta}^{(m)} - \theta}{\theta} \times 100 \right)^2} \quad \text{REE}^{(m)}(\hat{\theta}) = \frac{\hat{\theta}^{(m)} - \theta}{\theta} \times 100$$

For the second objective of assessing the standard errors, we report the relative standard errors (RSE) computed using the stochastic algorithm and compare them to the relative empirical standard errors.

$$\text{SE}_{emp}(\hat{\theta}) = \sqrt{\frac{\sum_m (\hat{\theta}^{(m)} - \bar{\hat{\theta}})^2}{m-1}} \quad \text{RSE}_{emp}(\hat{\theta}) = \frac{\text{SE}_{emp}(\hat{\theta})}{\bar{\hat{\theta}}}$$

where  $\bar{\hat{\theta}}$  is the mean estimates of  $\theta$ .

Finally, for the third objective of assessing the type-I error when testing independence between the two submodels, we computed the Wald test statistic:  $z^{(m)} = \frac{\hat{\theta}^{(m)}}{SE_{stoc}^{(m)}(\hat{\theta}^{(m)})}$ , where  $SE_{stoc}^{(m)}$  refers to the stochastic standard errors computed in simulation  $m$ .

#### 2.5.5. Implementation

For each model and each simulated dataset, parameters are estimated by maximization of the likelihood using the extension of *saemix* available on Github. The algorithm is composed of an exploratory and a smoothing phases: in the first  $K_1$  iterations, the algorithm explore the parameter space without memory (that is to say  $\gamma_k = 1$ ) to converge quickly to a neighborhood of the maximum likelihood estimator. In the following  $K_2$  iterations, the stochastic approximation is performed and  $\gamma_k = 1/(k - K_1 + 1)$ . To achieve faster convergence, we used the "true" values presented in Table 1 as initial guesses for parameter estimates. The default simulated annealing option is kept during the first  $K_1/2$  iterations (it constrains the variance of the random effects and the residual error parameters to decrease by maximum 3%). Parameters without variability are estimated with an artificial variability during the simulated annealing phase, progressively forced to 0. We initialized their artificial variances to 1% of their values. In order to provide more stability, we put the number of chains of the algorithm to 3. We also raised the number of iterations to ensure good convergence: we fixed  $K_1 = 1000$  iterations in the exploratory phase (including 500 for the simulated annealing phase) and  $K_2 = 500$  for the smoothing phase.

#### 2.6. Real case study

We consider the R data set *prothro* available in *JM* package, where 488 liver cirrhosis patients are followed at most 12 days with prothrombin measurements. Patients have a median of 6 (Q1-Q3=[3,9]) measurements. Status (dead/censored) at the end of the follow-up is available for each patient. About 60% of patients died during the study. We consider the joint model presented in equation 13 with a linear mixed-effects model to model the individual prothrombin evolution and a parametric time-to-event model to model the instantaneous risk of death in liver cirrhosis patients.

The linear joint model writes:

$$\begin{aligned}
 y_{ij} &= m(t_{ij}, \psi_i, z_i) + \sigma \epsilon_{ij} \\
 &= (\mu_0 + \eta_{0i}) + (\mu_1 + \eta_{1i}) t_{ij} + \sigma \epsilon_{ij} \\
 h_i(t, \psi_i) &= h_0 \exp(\alpha m(t, \psi_i))
 \end{aligned} \tag{18}$$

Details on code implementation are given in Appendix C. For the estimation, the default settings are kept (1 chain, 300/100 iterations in the exploratory/smoothing phases).

### 3. Results

#### 3.1. Results of the simulation study

##### 3.1.1. Parameter and standard errors estimation

Table 2 presents the RB and RRMSE for the joint models presented in the Methods section. The link coefficients, which are of particular interest in the context of joint modeling, are precisely and accurately estimated. For those parameters, the RB and RRMSE are low (< 7% for RB and < 25% for RRMSE) for all joint models. Overall, joint model parameters are very well estimated, while more uncertainty in estimates for some variances of random effects in particular for nonlinear modeling. In that case, the RRMSE are greater than 50% but the RB remains moderate. Similar parameter estimates for the corresponding parameters were obtained when fitting only the continuous biomarker alone (LMEM and NLMEM) and are shown in Appendix D.

	JM LMEM-TTE		JM NLMEM-TTE		JM LMEM-CR		JM NLMEM-CR	
	RB (%)	RRMSE (%)	RB (%)	RRMSE (%)	RB (%)	RRMSE (%)	RB (%)	RRMSE (%)
<i>Longitudinal model</i>								
<i>Fixed effects</i>								
$\mu_0$	0.20	5.4	0.3	5.1	-0.20	5.6	-1.5	5.6
$\mu_1$	0.13	15.2	1.6	8.8	-0.49	19.1	1.4	12.3
$\mu_2$	-	-	-1.3	7.0	-	-	0.69	11.9
$\mu_a$	-	-	21.2	43.6	-	-	25.6	56.6
<i>Random effects</i>								
$\omega_0^2$	1.3	14.0	-2.4	15.1	-3.50	16.0	0.33	15.0
$\omega_1^2$	1.6	15.7	-17.9	60.2	-2.8	18.3	-1.24	84.6
$\omega_2^2$	-	-	-18.3	47.5	-	-	-0.34	73.8
$\omega_a^2$	-	-	2.0	51.2	-	-	-22.0	52.6
<i>Error model</i>								
$\sigma$	0.39	1.5	0.28	1.7	0.30	2.1	0.10	2.5
<i>Survival model</i>								
$h_0$	-1.6	33.0	-3.8	30.5	-	-	-	-
$p_1$	-	-	-	-	-3.2	41.5	1.6	32.4
$g_1$	-	-	-	-	1.4	20.4	-1.5	19.7
$\alpha$	3.3	15.0	2.9	18.5	6.1	23.6	1.6	18.3
$b$	-	-	-	-	2.8	21.1	-0.56	17.2

Table 2: Relative Bias (RB) and Relative Root Mean Square Error (RRMSE) obtained on 100 simulations, for the four joint models

The distribution of REE in following figures (Figure 1 to 4), provide same conclusion. It is an efficient and simple way to see the good properties of parameter estimation. The higher uncertainty on random effect variances clearly appears for nonlinear joint models. Of note, we obtain same results in parameter estimated for the NLMEM alone (see Appendix E).

The figures also provide results for relative standard errors estimation. For each joint model, standard errors estimates are very close to the empirical ones. That suggests good estimation properties. Note however RSE for random effects variances are under-estimated for nonlinear modeling. We also emphasize that similar results were obtained for NLMEM (see Appendix E), highlighting an issue inherent to the model and not due to the joint modeling. This result may be related to the higher uncertainty of REE for those same parameters. This issue with the variances however does not affect the standard errors of the link coefficients.

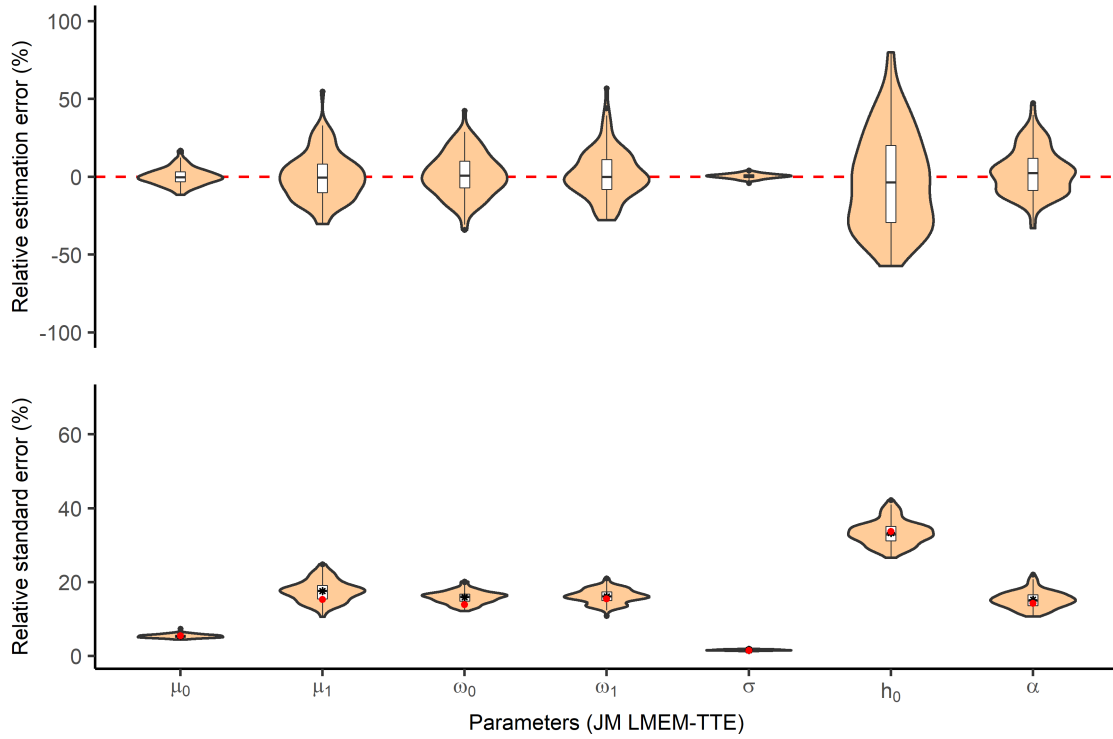


Figure 1: Distribution of the relative estimation errors (top) and relative standard errors (RSE) (bottom) for joint models with a linear mixed-effects model and a time-to-event model. Stars correspond to the mean distribution of the RSE. Red points correspond to the empirical RSE obtained over all the simulations.

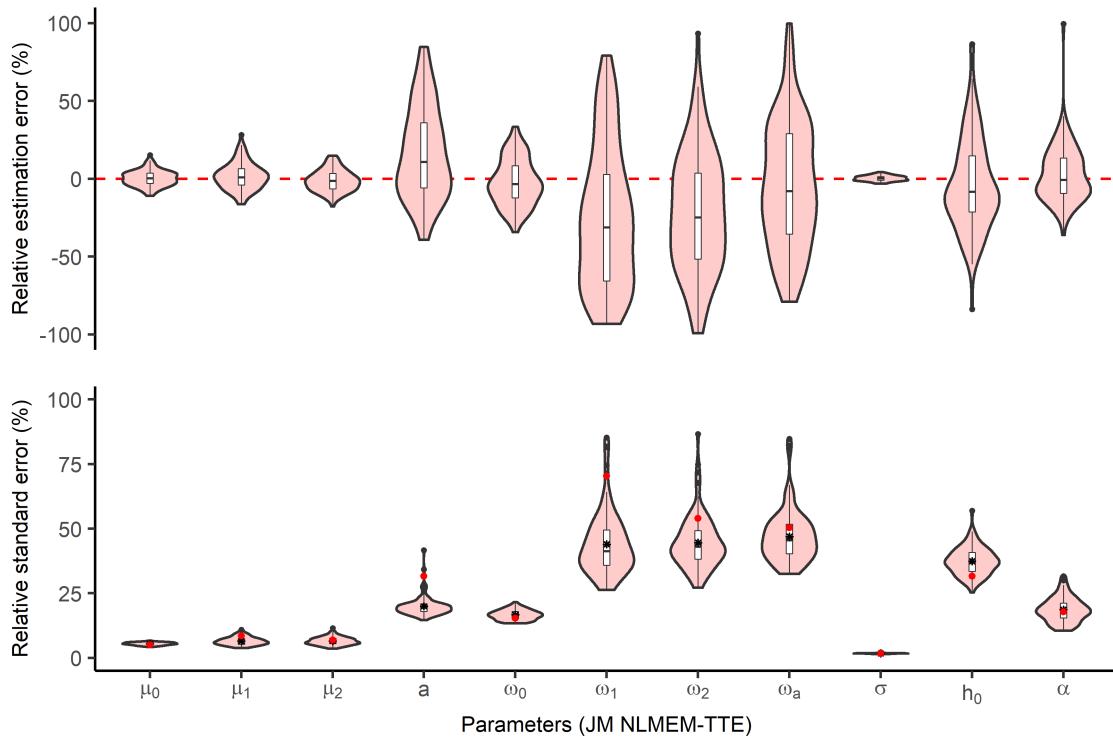


Figure 2: Distribution of the relative estimation errors (top) and relative standard errors (RSE) (bottom) for joint models with a nonlinear mixed-effects model and a time-to-event model. Stars correspond to the mean distribution of the RSE. Red points correspond to the empirical RSE obtained over all the simulations.

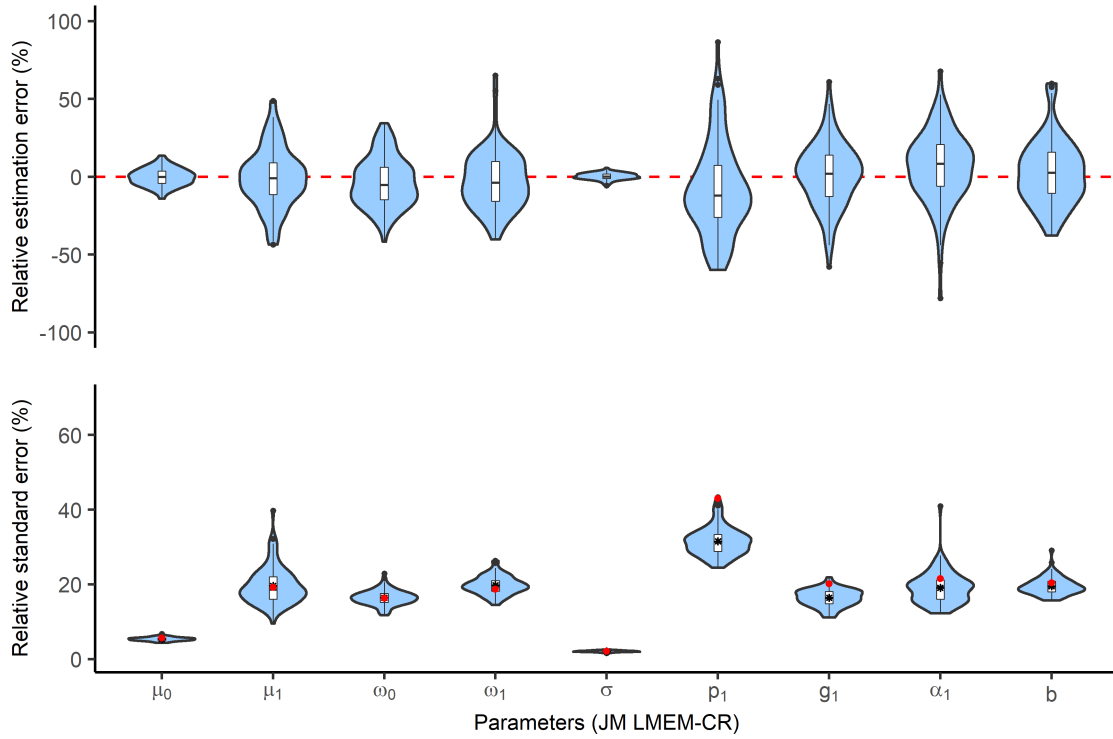


Figure 3: Distribution of the relative estimation errors (top) and relative standard errors (RSE) (bottom) for joint models with a linear mixed-effects model and a competing risks model. Stars correspond to the mean distribution of the RSE. Red points correspond to the empirical RSE obtained over all the simulations.

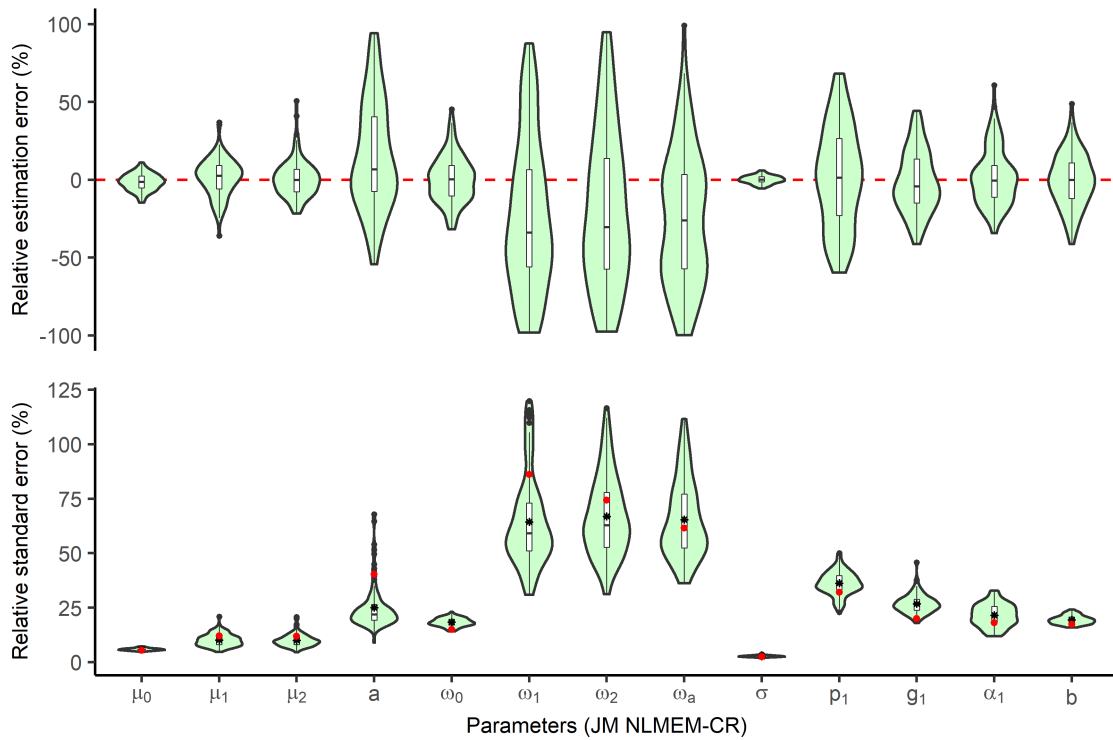


Figure 4: Distribution of the relative estimation errors (top) and relative standard errors (RSE) (bottom) for joint models with a nonlinear mixed-effects model and a competing risks model. Stars correspond to the mean distribution of the RSE. Red points correspond to the empirical RSE obtained over all the simulations.

### 3.1.2. Type-I error of Wald test on link coefficients

Figure 5 (plot A) shows the empirical type-1 errors computed over the 1000 simulated data sets (with  $\alpha = 0$ ). Type-I error is controlled for each joint model. With the most complex scenario (JM NLMEM-CR) the Wald test is conservative (empirical type-1 error = 3.2% with 95%CI = [2.2, 4.5]). On plot B, distribution of empirical versus theoretical p-values (on the  $-\log_{10}$ ) yields the same conclusion. We observe a deviation to the right, for JM NLMEM-CR around p-values between 1 and 5%. In particular, observed p-values are lower than the expected ones which corresponds to a conservative test for this range of p-values.

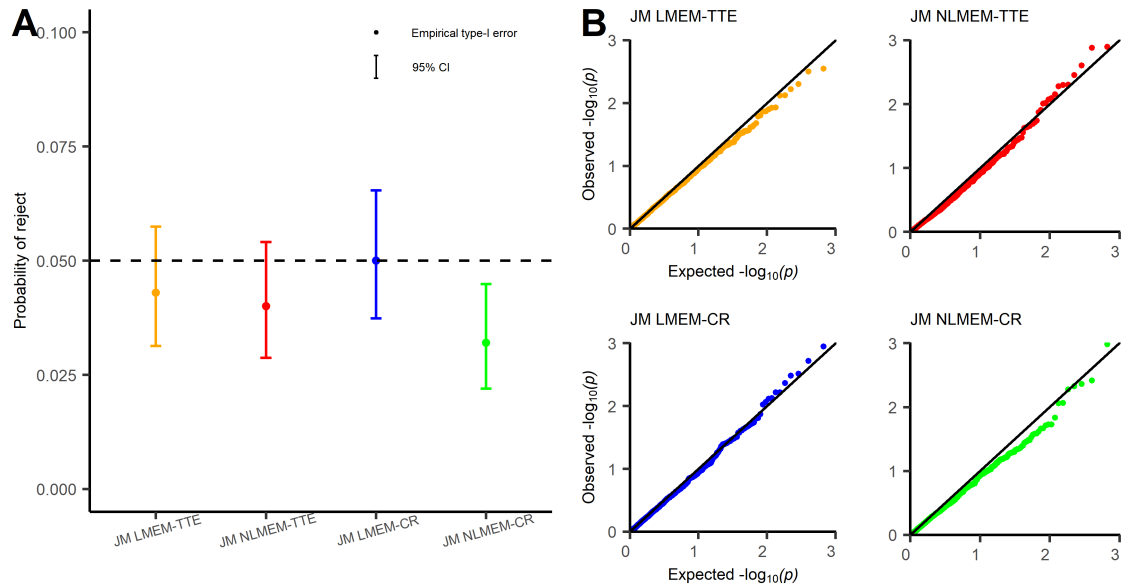


Figure 5: Type-I error with 95% CI for each joint model (A) and empirical  $-\log_{10}$  p-values of the Wald statistics versus theoretical ones (B)

### 3.2. Real-case application

Table 3 presents parameter estimates for the joint model applied to the *prothro* data. Parameters are well estimated with small RSE ( $< 30\%$  except for the slope of the linear structural function). The dynamic of prothrombin is predictive of death in liver cirrhosis patients, as decrease of prothrombin over time is associated with a higher risk of death ( $\alpha = -0.039$  with 95%CI =  $[-0.045, -0.033]$ ,  $p < 10^{-5}$ ). As a quality check, we also fitted the model using the *JM* package (*JointModel* function) and obtained similar parameter and RSE estimates (see Appendix F).

## 4. Conclusion

In this work, we modified the code from the R package *saemix* to fit joint models. This tool allows to fit a wide range of parametric joint models with an explicit expression of the likelihood. As a non exhaustive list of examples, we applied it to linear joint models with a single event model, and to more complex joint models including NLMEM and competing risks. Multiple longitudinal models can be easily considered, if various biomarkers are available. Other error model functions (proportional or combined) can also be considered, as well as baseline covariates (not presented in this article but an example is proposed on Github). In comparison with other packages available for joint model estimation where modeling is mainly limited to linear forms

Parameters	Estimates	S.E.	R.S.E (%)
<i>Longitudinal model</i>			
<i>Fixed effects</i>			
$\mu_0$	73.3	1.00	1.3
$\mu_1$	0.57	0.33	57.9
<i>Random effects</i>			
$\omega_0^2$	369	31.4	8.5
$\omega_1^2$	16	1.99	12.4
<i>Error model</i>			
$\sigma$	17	0.17	1.0
<i>Survival model</i>			
$h_0$	3.1	0.21	6.8
$\alpha$	-0.039	0.003	7.7

Table 3: Parameter estimates on the real data application

for the longitudinal response and pre-defined forms for the survival model [9, 10, 11, 13], the advantage is that users can fit very specific joint models by defining a personalized likelihood function.

In terms of standard error calculation, standard algorithms used in the framework of Gaussian responses based on linearization of the likelihood such as FO or FOCE [21] cannot be used in the framework of joint models with survival responses. The current version of the *saemix* offers several bootstrap approaches but has not yet implemented numerical or stochastic computation of the FIM for discrete data and survival data models. We opted for a recently developed stochastic algorithm [24] to estimate the FIM of the model. The advantage is that it integrates perfectly with the SAEM algorithm, and that it only involves first derivatives. Alternatively, the Monolix software uses Louis' decomposition principle combined with a stochastic algorithm to estimate the FIM of a joint model, but it is more computationally expensive because it involves second derivatives and it is run after the SAEM algorithm.

We showed through the simulation study good performances for both the estimation of the parameters in the joint model and for the computation of the standard errors. We reported a higher uncertainty in some estimates of random effects variances, for complex models (nonlinear modeling) and also a small under-estimation of relative standard errors for the same parameters. Of note, similar results are obtained using Monolix software, which highlights difficulties inherent to the proposed model [5]. However, the link coefficients which are of particular interest in a joint modeling context, are accurately and precisely estimated, as are their standard errors. Moreover, we showed that the type-I error when assessing the significance of those coefficients (Wald test) is controlled for each proposed model.

We applied our approach to data collected in liver cirrhosis patients. The joint model is easily implemented and shows that a decrease of the prothrombin over time is associated with a higher risk of death. Although the objective is not to compare the results with the existing *JM* package, very similar results were obtained using this previous one. The advantage is that more complex longitudinal and survival models can be explored with the *saemix* extension compared to currently available packages in R.

We provide code that is *ad hoc* but available and editable by any user. In the model definition, the technical part remains the writing the likelihood function. It has the advantage of being flexible in the sense that any parametric joint model can be considered, but the disadvantage of being written explicitly so the user has to perfectly know the model expression. In the presence of survival

data, the likelihood involves integrals that may have no close form depending on the model complexity. This issue can be overcome by using numerical integration, which is simple to implement but also makes parameter estimation much more time consuming.

We emphasize that the work presents some limits. In the simulation study, we assessed the performances of the developed algorithm only in a rich design. Furthermore, parameters were initialized to their true values. It may be interesting to assess the performances of the algorithm in a sparser design, and test the sensitivity to the values of initial estimates. Note, version 3.1 of *saemix* was evaluated with different initial parameters [29]. Performances can be degraded when the initial parameters are far from the true ones, but this issue can be corrected by increasing the number of iterations, and/or the number of chains. A similar effect could be expected in this context. Moreover, the code was not yet extended to allow for correlation between random effects and to provide diagnostic tools. However, users have access to estimated individual parameters and individual predictions and hence, can build their own diagnostic plots such as observations vs predictions plots, normalised prediction distribution errors (NPDE) [30, 31], or individual residuals for the longitudinal part [31], and martingale or Cox-snell residuals for the survival part [32].

All the extended functions of the R package *saemix* are available on Github at the following address: <https://github.com/saemixdevelopment/saemixextension/tree/master/joint>.

## Acknowledgments

The authors gratefully acknowledge Dr Maud Delattre (UMR MIA-Paris, AgroParisTech, INRAE, Université Paris-Saclay, 75005, Paris, France) for her kindness, availability and help with the FIM computation.

## References

- [1] Magdalena Murawska, Dimitris Rizopoulos, and Emmanuel Lesaffre. A Two-Stage Joint Model for Nonlinear Longitudinal Response and a Time-to-Event with Application in Transplantation Studies. *Journal of Probability and Statistics*, 2012:e194194, 2012.
- [2] Dimitris Rizopoulos. *Joint models for longitudinal and time-to-event data: with applications in R*. Chapman & Hall / CRC biostatistics series. Boca Raton, Fla., 2012.
- [3] Solène Desmée, France Mentré, Christine Veyrat-Follet, and Jérémie Guedj. Nonlinear Mixed-effect Models for Prostate-specific Antigen Kinetics and Link with Survival in the Context of Metastatic Prostate Cancer: A Comparison by Simulation of Two-stage and Joint Approaches. *The AAPS Journal*, 17(3):691–699, 2015.
- [4] Marion Keroui, François Mercier, Julie Bertrand, Coralie Tardivon, René Bruno, Jérémie Guedj, and Solène Desmée. Bayesian inference using Hamiltonian Monte-Carlo algorithm for nonlinear joint modeling in the context of cancer immunotherapy. *Statistics in Medicine*, 39(30):4853–4868, 2020.
- [5] Alexandra Lavalley-Morelle, Jean-François Timsit, France Mentré, Jimmy Mullaert, and The OUTCOMEREA Network. Joint modeling under competing risks: Application to survival prediction in patients admitted in Intensive Care Unit for sepsis with daily Sequential Organ Failure Assessment score assessments. *CPT: Pharmacometrics & Systems Pharmacology*, 11(1472–1484), 2022.
- [6] Michael S. Wulfsohn and Anastasios A. Tsiatis. A Joint Model for Survival and Longitudinal Data Measured with Error. *Biometrics*, 53(1):330–339, 1997.
- [7] Jeffrey D. Long and James A. Mills. Joint modeling of multivariate longitudinal data and survival data in several observational studies of Huntington’s disease. *BMC Medical Research Methodology*, 18(1):138, 2018.
- [8] Jaewan Choi, Jianwen Cai, Donglin Zeng, and Andrew F. Olshan. Joint Analysis of Survival Time and Longitudinal Categorical Outcomes. *Statistics in Biosciences*, 7(1):19–47, 2015.
- [9] Pete Philipson, Ines Sousa, Peter J. Diggle, Paula Williamson, Ruwanthi Kolamunnage-Dona, Robin Henderson, and Graeme L. Hickey. *joiner: Joint Modelling of Repeated Measurements and Time-to-Event Data*, 2018.

- [10] Dimitris Rizopoulos. JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software*, 35(9):1–33, 2010.
- [11] Graeme L. Hickey, Pete Philipson, and Andrea Jorgensen and. joinerml: a joint model and software package for time-to-event and multivariate longitudinal outcomes. *BMC Medical Research Methodology*, 18(1), 2018.
- [12] Cécile Proust-Lima, Viviane Philipps, and Benoit Liqueur. Estimation of extended mixed models using latent classes and latent processes: The R package lmm. *Journal of Statistical Software*, 78(2):1–56, 2017.
- [13] Virginie Rondeau, Yassin Mazroui, and Juan R. Gonzalez. frailtypack: An R package for the analysis of correlated survival data with frailty models using penalized likelihood estimation or parametrical estimation. *Journal of Statistical Software*, 47(4):1–28, 2012.
- [14] Alberto Garcia-Hernandez and Dimitris Rizopoulos. %JM: A SAS Macro to Fit Jointly Generalized Mixed Models for Longitudinal Data and Time-to-Event Responses. *Journal of Statistical Software*, 84:1–29, 2018.
- [15] Danjie Zhang, Ming-Hui Chen, Joseph G. Ibrahim, Mark E. Boye, and Wei Shen. JMFIt: A SAS Macro for Joint Models of Longitudinal and Survival Data. *Journal of Statistical Software*, 71:1–24, 2016.
- [16] Michael J. Crowther. STJM: Stata module to fit shared parameter joint models of longitudinal and survival data. *Statistical Software Components*, 2013.
- [17] Robert J. Bauer. NONMEM Tutorial Part I: Description of Commands and Options, With Simple Examples of Population Analysis. *CPT: Pharmacometrics & Systems Pharmacology*, 8(8):525–537, 2019.
- [18] Pauline Traynard, Géraldine Ayrat, Monika Twarogowska, and Jonathan Chauvin. Efficient Pharmacokinetic Modeling Workflow With the MonolixSuite: A Case Study of Remifentanyl. *CPT: Pharmacometrics & Systems Pharmacology*, 9(4):198–210, 2020.
- [19] Sreenath M. Krishnan, Lena E. Friberg, René Bruno, Ulrich Beyer, Jin Y. Jin, and Mats O. Karlsson. Multistate model for pharmacometric analyses of overall survival in HER2-negative breast cancer patients treated with docetaxel. *CPT: Pharmacometrics & Systems Pharmacology*, 10(10):1255–1266, 2021.
- [20] Thomas A. Louis. Finding the Observed Information Matrix When Using the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2):226–233, 1982.
- [21] Mary J. Lindstrom and Douglas M. Bates. Nonlinear Mixed Effects Models for Repeated Measures Data. *Biometrics*, 46(3):673–687, 1990.
- [22] Marie-Karelle Riviere, Sebastian Ueckert, and France Mentré. An MCMC method for the evaluation of the Fisher information matrix for non-linear mixed effect models. *Biostatistics*, 17(4):737–750, 2016.
- [23] Estelle Kuhn and Marc Lavielle. Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis*, 49(4):1020–1038, 2005.
- [24] Maud Delattre and Estelle Kuhn. Estimating Fisher Information Matrix in Latent Variable Models based on the Score Function, 2023.
- [25] Bernard Delyon, Marc Lavielle, and Eric Moulines. Convergence of a stochastic approximation version of EM algorithm. *The Annals of Statistics*, 27:94, 1999.
- [26] Arthur P. Dempster, Nan M. Laird, and Donald B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.
- [27] Vianney Debavelaere and Stéphanie Allasonnière. On the curved exponential family in the Stochastic Approximation Expectation Maximization Algorithm. *ESAIM: Probability and Statistics*, 25:408–432, 2021.
- [28] John A. Nelder and Roger Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, 1965.
- [29] Emmanuelle Comets, Audrey Lavenu, and Marc Lavielle. Parameter Estimation in Nonlinear Mixed Effect Models Using saemix, an R Implementation of the SAEM Algorithm. *Journal of Statistical Software*, 80:1–41, 2017.
- [30] Emmanuelle Comets, Karl Brendel, and France Mentré. Computing normalised prediction distribution errors to evaluate nonlinear mixed-effect models: the npde add-on package for R. *Computer Methods and Programs in Biomedicine*, 90(2):154–166, 2008.
- [31] Emmanuelle Comets and France Mentré. Developing tools to evaluate non-linear mixed effect models: 20 years on the npde adventure. *AAPS Journal*, 23(4):75, 2021.
- [32] David. Cox and E. Joyce Snell. A General Definition of Residuals. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30(2):248–275, 1968.

## Appendix A. Derivation of Sufficient Statistics for a gaussian mixed-effects model and for a joint model

### Appendix A.1. Mixed-effects model with homoscedastic variance

We consider the mixed-effects model with an additive error defined as:

$$y_{ij} = m(t_{ij}, \psi_i) + \sigma \epsilon_{ij} \quad (\text{A.1})$$

with  $\psi_i$  the individual parameters ( $i = 1, \dots, N$ ), a function of fixed-effects  $\mu_d$  and random effects  $\eta_{d,i}$  ( $1 \leq d \leq D$ ). The vectors of random effects are assumed to be normally distributed with diagonal variance covariance matrix  $\Omega$ .  $\epsilon_{ij}$  is the residual error following a normal distribution with mean 0 and variance 1. For simplicity in the rest of this section, we will assume that the distribution of the individual parameters is normal:

$$\psi_{d,i} = \mu_d + \eta_{d,i}$$

but the equations below generalize to models where each component  $\psi_{d,i}$  can be transformed to a linear function of  $\mu_d$ ,  $\eta_{d,i}$ , and optionally covariates  $z_{cov,i}$  and associated coefficients  $\beta_{d,cov}$ .

The complete likelihood is:

$$\begin{aligned} l(y, \psi; \theta) &= \prod_{ij} p(y_{ij} | \psi_i; \theta) \prod_i p(\psi_i; \theta) \\ &= \prod_{ij} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_{ij} - m(t_{ij}, \psi_i))^2}{2\sigma^2}\right) \prod_i \frac{1}{(2\pi)^{D/2} |\Omega|^{1/2}} \exp\left(-\frac{1}{2} (\psi_i - \mu)^T \Omega^{-1} (\psi_i - \mu)\right) \\ &= \prod_{ij} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_{ij} - m(t_{ij}, \psi_i))^2}{2\sigma^2}\right) \prod_i \prod_d \frac{1}{(2\pi)^{D/2} \omega_d} \exp\left(-\frac{(\psi_{d,i} - \mu_d)^2}{2\omega_d^2}\right) \end{aligned}$$

Taking the logarithm of previous equation, we can write:

$$\begin{aligned} \ln l(y, \psi; \theta) &= \sum_{i,j} -\ln \sqrt{2\pi}\sigma - \sum_{i,j} \frac{(y_{ij} - m(t_{ij}, \psi_i))^2}{2\sigma^2} \\ &\quad + N \sum_d -\ln((2\pi)^{D/2} \omega_d) - \sum_i \sum_d \frac{(\psi_{d,i})^2}{2\omega_d^2} + \sum_i \sum_d \frac{(\mu_d \psi_{d,i})}{\omega_d^2} - N \sum_d \frac{(\mu_d)^2}{2\omega_d^2} \end{aligned} \quad (\text{A.2})$$

The terms in red depend only on the population parameters, while the terms in blue depend on the data and the individual and population parameters. In the framework of the exponential family, the red terms correspond to  $\Psi(\theta)$  in equation 5, so that:

$$\Psi(\theta) = \sum_{i,j} \ln(\sqrt{2\pi}\sigma) + N \sum_d \ln(\sqrt{2\pi}\omega_d^2) + N \sum_d \frac{(\mu_d)^2}{2\omega_d^2}$$

The blue terms can now be factorised as the scalar product of two terms which conform to equation (5). We define:

$$S(y, \phi) = \begin{pmatrix} 1/2(y_{1,1} - m(t_{1,1}, \psi_1))^2 \\ 1/2(y_{1,2} - m(t_{1,2}, \psi_1))^2 \\ \dots \\ 1/2(y_{1,n_1} - m(t_{1,n_1}, \psi_1))^2 \\ 1/2(y_{2,1} - m(t_{2,1}, \psi_2))^2 \\ \dots \\ 1/2(y_{2,n_2} - m(t_{2,n_2}, \psi_2))^2 \\ \dots \\ 1/2(y_{N,1} - m(t_{N,1}, \psi_N))^2 \\ \dots \\ 1/2(y_{N,n_N} - m(t_{N,n_N}, \psi_N))^2 \\ \\ \psi_{11} \\ \dots \\ \psi_{1N} \\ \dots \\ \psi_{D1} \\ \dots \\ \psi_{DN} \\ \\ 1/2(\psi_{11})^2 \\ \dots \\ 1/2(\psi_{1N})^2 \\ \dots \\ 1/2(\psi_{D1})^2 \\ \dots \\ 1/2(\psi_{DN})^2 \end{pmatrix}$$

composed of 3 stacked vectors

$$S_y = \left( \frac{1}{2}(y_{i,j} - m(t_{i,j}, \psi_i))^2 \right)_{i=1, \dots, N, j=1, \dots, n_i} \quad S_\psi = (\psi_{di})_{i=1, \dots, N, d=1, \dots, D} \quad S_{\psi^2} = \left( \frac{1}{2}(\psi_{di})^2 \right)_{i=1, \dots, N, d=1, \dots, D}$$

We then define the function  $\phi(\theta)$  as a similar stacked vector composed of 3 parts of the same sizes as  $S_y$ ,  $S_\psi$  and  $S_{\psi^2}$  respectively:

$$\phi_1(\theta) = \begin{pmatrix} -\frac{1}{\sigma^2} \\ \dots \\ -\frac{1}{\sigma^2} \end{pmatrix}_{\sum_i n_i \text{times}}$$

$$\phi_2(\theta) = \begin{pmatrix} \frac{\mu_1}{\omega_1^2} \\ \dots \\ \frac{\mu_D}{\omega_D^2} \\ \dots \\ \frac{\mu_D}{\omega_D^2} \end{pmatrix}_{N \text{times}}$$

$$\phi_3(\theta) = \begin{pmatrix} -\frac{1}{\omega_1^2} \\ \dots \\ -\frac{1}{\omega_1^2} \\ \dots \\ -\frac{1}{\omega_D^2} \\ \dots \\ -\frac{1}{\omega_D^2} \end{pmatrix}_{N \text{times}}$$

We can see that

$$\langle S_y, \phi_1 \rangle = - \sum_{i,j} \frac{(y_{ij} - m(x_{ij}, \psi_i))^2}{2 \sigma^2}$$

$$\langle S_\psi, \phi_2 \rangle = \sum_d \frac{(\mu_d \psi_d)}{\omega_d^2}$$

and

$$\langle S_{\psi^2}, \phi_3 \rangle = - \sum_d \frac{(\psi_d)^2}{2\omega_d^2}$$

so that we have isolated the scalar product  $\langle S(y, \psi), \phi(\theta) \rangle$  through the definition of the three sufficient statistics  $S_y$ ,  $S_\psi$  and  $S_{\psi^2}$ .

In turn, these statistics are sufficient because we can write the population parameters as a function of them only:

$$\mu_d = \frac{1}{N} \sum_i \psi_{di}$$

$$\omega_d^2 = \sqrt{\frac{1}{N} \sum_i (\psi_{di})^2}$$

$$\sigma^2 = \sqrt{\frac{1}{n_{obs}} \sum_{ij} (y_{ij} - m(t_{ij}, \psi_i))^2}$$

### Appendix A.2. Joint model with linear mixed-effects model and single event

We consider the joint model defined as:

$$\begin{aligned} y_{ij} &= m(t_{ij}, \psi_i) + \sigma \epsilon_{ij} \\ &= \psi_{0i} + \psi_{1i} t_{ij} + \sigma \epsilon_{ij} \\ h_i(t, \psi_i) &= h_0 \exp(\alpha m(t, \psi_i)) \end{aligned} \quad (\text{A.3})$$

For longitudinal parameters, we assume same distributions as in equation A.1.  $h_i(t, \psi)$  is the instantaneous risk of event for individual  $i$  at time  $t$  with  $h_0$  the baseline risk and  $\alpha$  the link coefficient. We introduce  $S_i(t, \psi_i) = \exp\left(-\int_0^t h_i(s, \psi_i) ds\right)$ .

In the first  $K_1$  iterations of the algorithm, where artificial variability is added on parameters  $h_0$  and  $\alpha$ , the vectors  $\mu$  and  $\eta$  contains 2 additional parameters referring to  $h_0$  and  $\alpha$ . Hence,  $\mu = (\mu_0, \mu_1, \mu_{h_0}, \mu_\alpha)$  and  $\eta = (\eta_{0i}, \eta_{1i}, \eta_{h_0i}, \eta_{\alpha i})$ . In that case, the joint likelihood of the complete data  $(y, T, \delta, \psi)$  can be written as:

$$\begin{aligned} l(y_i, T_i, \delta_i, \psi_i) &= \prod_{ij} p(y_{ij}|\psi_i; \theta) \prod_i p(T_i, \delta_i|\psi_i; \theta) \prod_i p(\psi_i; \theta) \\ &= \prod_{ij} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_{ij} - m(t_{ij}, \psi_i))^2}{2\sigma^2}\right) \prod_i h_i(T_i, \psi_i)^{\delta_i} S_i(T_i, \psi_i) \prod_i \prod_d \frac{1}{(2\pi)^{D/2} \omega_d} \exp\left(-\frac{(\psi_{di} - \mu_d)^2}{2\omega_d^2}\right) \end{aligned}$$

and the log-likelihood as:

$$\begin{aligned} \ln l(y, \psi; \theta) &= \sum_{i,j} -\ln \sqrt{2\pi}\sigma - \sum_{i,j} \frac{(y_{ij} - m(t_{ij}, \psi_i))^2}{2\sigma^2} \\ &\quad + \delta_i (\log(\psi_{ih_0}) + \psi_{i\alpha} m(T_i, \psi_i)) - \frac{\psi_{ih_0}}{\psi_{i\alpha} \psi_{i1}} (\exp(\psi_{i\alpha} m(T_i, \psi_i)) - \exp(\psi_{i\alpha} \psi_{i0})) \\ &\quad + N \sum_d -\ln((2\pi)^{D/2} \omega_d) - \sum_i \sum_d \frac{(\psi_{di})^2}{2\omega_d^2} + \sum_i \sum_d \frac{(\mu_d \psi_{di})}{\omega_d^2} - N \sum_d \frac{(\mu_d)^2}{2\omega_d^2} \end{aligned}$$

The log-likelihood again decomposes in the general form:

$$\log l(y, \psi; \theta) = -\Psi(\theta) + \langle S(y, \psi), \phi(\theta) \rangle + A(T, \delta, \psi) \quad (\text{A.4})$$

$A(T, \delta, \psi)$  depends only the complete survival data  $(T, \delta, \psi)$  and is ignored in the computation of sufficient statistics, while the red and blue terms are the same as in equation A.2 and factorise similarly.

After the  $K_1$  first iterations, the variability on  $h_0$  and  $\alpha$  is set to 0. The vector  $\mu$  and  $\eta$  are now only composed of longitudinal parameters and are derived through the sufficient statistics. An optimization algorithm is used for  $h_0$  and  $\alpha$ . Note, This approach is also used to obtain estimates of the residual error parameters when using a combined error model for the continuous outcome, in which case the complete log-likelihood is strictly speaking no longer in the curved exponential family.

## Appendix B. Code

### Appendix B.1. Datasets simulation

The following code is used to simulate the datasets in R. It does not require the *saemix* code or its extensions to run.

```

1 ##### SIMULATION OF DATASETS #####
2 ##### 6 models: LMEM      NLMEM      LMEM/NLMEM – TTE      LMEM/NLMEM – CR
3
4
5 ##### LMEM #####
6
7 set.seed(1996)
8 N = 100
9
10 longi = data.frame(id=NA, obs=NA, time=NA)
11
12 # Values used for simulation
13 b0=4
14 b1 = 0.2
15 omega_b0 = sqrt(4)
16 omega_b1 = sqrt(0.1)
17 sigma_a = 1
18
19 for (i in 1:100){
20   r1 = rnorm(N, mean=b0, sd=omega_b0) # vector of random intercepts
21   r2 = rnorm(N, mean=b1, sd=omega_b1) # vector of random slopes
22   tt = 0:30 # design for longitudinal observations
23   longi = data.frame(id = as.vector(sapply(1:N, function(x) rep(x, length(tt)))),
24                     time = rep(tt, N))
25   longi$obs = r1[longi$id] + r2[longi$id]*longi$time +
26             rnorm(nrow(longi), sd=sigma_a)
27
28   write.table(longi, paste0(".../data", i, ".txt"), row.names = F)
29 }
30
31
32 ##### NLMEM #####
33
34 set.seed(1996)
35 N = 100
36 longi = data.frame(id=NA, obs=NA, time=NA)
37
38 b0<-4
39 b1<- 0.15

```

```

40 b2 <- 0.2
41 a = 10
42 omega_b0<-sqrt(4)
43 omega_b1<-sqrt(0.1)
44 omega_b2<-sqrt(0.1)
45 omega_a<-sqrt(0.5)
46 sigma_a<-1
47
48 for (i in 1:100){
49   r1 = rnorm(N,mean=b0 ,sd=omega_b0)
50   r2 = b1*exp(rnorm(N,mean=0 ,sd=omega_b1))
51   r3 = b2*exp(rnorm(N,mean=0, sd=omega_b2))
52   r4 = a*exp(rnorm(N,mean=0, sd=omega_a))
53   tt = 0:30
54   longi = data.frame(id = as.vector(sapply(1:N,function(x) rep(x, length(tt))),
55                                     time = rep(tt,N))
56   longi$obs = r1[longi$id] + r4[longi$id] *
57               (exp(-r2[longi$id]*longi$time)-exp(-r3[longi$id]*longi$time)) +
58               rnorm(nrow(longi),sd=sigma_a)
59
60   write.table(longi,paste0(".../data",i,".txt"),row.names = F)
61 }
62
63
64
65 ##### LMEM - TTE #####
66
67 set.seed(1996)
68 N=100
69
70 longi = data.frame(id=NA,obs=NA,time=NA)
71
72 b0=4
73 b1 = 0.2
74 omega_b0 = sqrt(4)
75 omega_b1 = sqrt(0.1)
76 sigma_a = 1
77 h0=0.005
78 alpha=0.2
79
80 for (i in 1:100){
81   r1 = rnorm(N,mean=b0 ,sd=omega_b0)
82   r2 = rnorm(N,mean=b1 ,sd=omega_b1)

```

```

83 tt = 0:30
84 longi = data.frame(id = as.vector(sapply(1:N, function(x) rep(x, length(tt)))),
85                               time = rep(tt, N))
86 longi$obs = r1[longi$id] + r2[longi$id]*longi$time +
87             rnorm(nrow(longi), sd=sigma_a)
88
89 # function that gives an event time given a random intercept r1 and a random
90 # slope r2
91 gettimes = function(r1, r2){
92   x = rexp(1)
93   H = function(tt) h0*exp(alpha*r1)/alpha/r2*(exp(alpha*r2*tt)-1) - x
94   z = try(uniroot(H, interval=c(0,30))$root, silent=T)
95   ifelse(class(z)=="try-error", Inf, z)
96 }
97 #survival data with a censoring time = D30
98 tte.data = data.frame(id=1:N, time = mapply(gettimes, rint, reff), obs = 1)
99 tte.data$obs[tte.data$time>30]=0
100 tte.data$time[tte.data$time>30]=30
101 # removal of post-event longitudinal data
102 longi = longi[longi$time<=sapply(longi$id,
103                               function(x) tte.data$time[tte.data$id==x]),]
104
105 longi$ytype=1
106 tte.data$ytype=2
107
108 alldata_saem = rbind(longi, tte.data)
109
110 write.table(alldata_saem, paste0("../data", i, ".txt"), row.names = F)
111 }
112
113
114 ##### NLMEM - TTE #####
115
116 set.seed(1996)
117 N=100
118
119 longi = data.frame(id=NA, obs=NA, time=NA)
120
121 b0<-4
122 b1<- 0.15
123 b2 <- 0.2
124 a = 10
125 omega_b0<-sqrt(4)

```

```

126 omega_b1<-sqrt(0.1)
127 omega_b2<-sqrt(0.1)
128 omega_a<-sqrt(0.5)
129 sigma_a<-1
130 h0=0.005
131 alpha=0.3
132
133 for (i in 1:100){
134   r1 = rnorm(N,mean=b0 ,sd=omega_b0)
135   r2 = b1*exp(rnorm(N,mean=0, sd=omega_b1))
136   r3 = b2*exp(rnorm(N,mean=0, sd=omega_b2))
137   r4 = a*exp(rnorm(N,mean=0, sd=omega_a))
138   tt = 0:30
139   longi = data.frame(id = as.vector(sapply(1:N,function(x) rep(x, length(tt))),
140                                     time = rep(tt,N))
141   longi$obs = r1[longi$id] + r4[longi$id] *
142             (exp(-r2[longi$id]*(longi$time))-exp(-r3[longi$id]*(longi$time)))+
143             rnorm(nrow(longi),sd=sigma_a)
144
145   gettimes = function(r1,r2,r3,r4){
146     x = rexp(1)
147     t = seq(0,100,length.out=1000)
148     pas = t[2]-t[1]
149     h = h0*exp(alpha*(r1+r4*(exp(-r2*(t))-exp(-r3*(t)))))
150     H = cumsum(h)*pas
151     z=(which(H>x)[1])*pas
152     z=ifelse(is.na(z)==T,Inf,z)
153   }
154   tte.data = data.frame(id=1:N,time = mapply(gettimes,r1,r2,r3,r4), obs = 1)
155   tte.data$obs[tte.data$time>30]=0
156   tte.data$time[tte.data$time>30]=30
157
158   longi = longi[longi$time<=sapply(longi$id,
159                                   function(x) tte.data$time[tte.data$id==x]),]
160
161   longi$ytype=1
162   tte.data$ytype=2
163
164   alldata_saem = rbind(longi,tte.data)
165
166   write.table(alldata_saem,paste0(".../data",i,".txt"),row.names = F)
167 }
168

```

```

169
170 ##### IMEM - CR #####
171
172 set.seed(1996)
173 N = 100
174
175 b0.pop<-4
176 b1.pop<- 0.2
177 omega_b0<-sqrt(4)
178 omega_b1<-sqrt(0.1)
179 sigma_a<-1
180 p1=0.15
181 g1 = 0.1
182 alpha1 = 0.2
183
184 for (i in 1:100){
185   rint = rnorm(N,mean=b0.pop ,sd=omega_b0)
186   reff = rnorm(N,mean=b1.pop ,sd=omega_b1)
187   tt = 0:30
188   longi = data.frame(id = as.vector(sapply(1:N, function(x) rep(x, length(tt)))),
189                       time = rep(tt,N))
190   longi$obs = rint[longi$id] + reff[longi$id]*longi$time +
191             rnorm(nrow(longi),sd=sigma_a)
192
193   gettimes = function(r1,r2){
194     t = seq(0,1000,length.out=10000)
195     pas = t[2]-t[1]
196     h = (p1*g1*exp(-g1*t)/(1-p1*(1-exp(-g1*t))))*exp(alpha1*(r1+r2*t))
197     F1 = 1-exp(-cumsum(h)*pas)
198     F2 = (1-rev(F1)[1])*(1-exp(-t/15))
199
200     p1 = rev(F1)[1]
201     p2 = 1-p1
202
203     u = runif(1)
204     event = ifelse(u<p1,1,2)
205
206     if(event==1) z=(which(F1/p1>runif(1))[1])*pas
207     if(event==2) z=(which(F2/p2>runif(1))[1])*pas
208     c(z,event)
209   }
210
211 a=mapply(gettimes,rint,reff)

```

```

212
213 tte.data = data.frame(id=1:N, time = a[1,], obs = a[2,])
214 tte.data$obs[tte.data$time > 30] = 0
215 tte.data$time[tte.data$time > 30] = 30
216
217 longi = longi[longi$time <= sapply(longi$id,
218                                 function(x) tte.data$time[tte.data$id == x]),]
219 longi$ytype = 1
220 tte.data$ytype = 2
221
222 alldata_saem = rbind(longi, tte.data)
223
224 write.table(alldata_saem, paste0(".../data", i, ".txt"), row.names = F)
225
226 }
227
228
229 ##### NLMEM - CR #####
230
231 set.seed(1996)
232 N = 100
233
234 longi = data.frame(id = NA, obs = NA, time = NA)
235 b0 <- -4
236 b1 <- 0.15
237 b2 <- 0.2
238 a = 10
239 omega_b0 <- sqrt(4)
240 omega_b1 <- sqrt(0.1)
241 omega_b2 <- sqrt(0.1)
242 omega_a <- sqrt(0.5)
243 sigma_a <- 1
244 p1 = 0.15
245 g1 = 0.1
246 alpha1 = 0.3
247
248 for (i in 1:100){
249   r1 = rnorm(N, mean=b0, sd=omega_b0)
250   r2 = b1 * exp(rnorm(N, mean=0, sd=omega_b1))
251   r3 = b2 * exp(rnorm(N, mean=0, sd=omega_b2))
252   r4 = a * exp(rnorm(N, mean=0, sd=omega_a))
253
254   tt = 0:30

```

```

255 longi = data.frame(id = as.vector(sapply(1:N, function(x) rep(x, length(tt))),
256                               time = rep(tt, N))
257 longi$obs = r1[longi$id] + r4[longi$id] *
258             (exp(-r2[longi$id]*(longi$time))-exp(-r3[longi$id]*(longi$time)))+
259             rnorm(nrow(longi), sd=sigma_a)
260
261 gettimes = function(r1, r2, r3, r4){
262   t = seq(0, 1000, length.out=10000)
263   pas = t[2]-t[1]
264   h = (p1*g1*exp(-g1*t)/(1-p1*(1-exp(-g1*t))))*
265       exp(alpha1*(r1+r4*(exp(-r2*(t))-exp(-r3*(t)))))
266   F1 = 1-exp(-cumsum(h)*pas)
267   F2 = (1-rev(F1)[1])*(1-exp(-t/15))
268
269   p1 = rev(F1)[1]
270   p2 = 1-p1
271
272   u = runif(1)
273   event = ifelse(u<p1, 1, 2)
274
275   if(event==1) z=(which(F1/p1>runif(1))[1])*pas
276   if(event==2) z=(which(F2/p2>runif(1))[1])*pas
277   c(z, event)
278 }
279 aa=mapply(gettimes, r1, r2, r3, r4)
280
281 tte.data = data.frame(id=1:N, time = aa[1,], obs = aa[2,])
282 tte.data$obs[tte.data$time>30]=0
283 tte.data$time[tte.data$time>30]=30
284 longi = longi[longi$time<=sapply(longi$id,
285                               function(x) tte.data$time[tte.data$id==x]),]
286
287 longi$ytype=1
288 tte.data$ytype=2
289
290 alldata_saem = rbind(longi, tte.data)
291
292 write.table(alldata_saem, paste0(".../data", i, ".txt"), row.names = F)
293 }

```

### Appendix C. Prothrombin example - Using code extension

```

1 ##### Joint model fit: liver cirrhosis patients
2
3 library(JM)
4 library(ggplot2)
5 library("viridis")
6 library(rlang)
7
8 # Loading of the data
9 data("prothro")
10 data("prothros")
11
12 gp = ggplot(data=prothro[prothro$id %in% 1:20,], aes(x=time, y=pro, group = id))+
13   geom_point(lwd=1.5)+geom_line(col="#CC0033",lwd=0.8)+theme_classic()+
14   ylab("Marker observations")+xlab("Days")
15 gp
16
17
18 ##### estimation with saemix
19
20 # Set directory
21 saemixDir <- "../GitHub/saemixextension"
22 workDir <- file.path(saemixDir, "joint")
23 setwd(workDir)
24
25 # Loading of original functions of the library
26 progDir<-file.path(saemixDir, "R")
27 source(file.path(progDir, "aaa_generics.R"))
28 #source(file.path(progDir, "global.R"))
29 source(file.path(progDir, "SaemixData.R"))
30 source(file.path(progDir, "SaemixRes.R"))
31 source(file.path(progDir, "SaemixModel.R"))
32 source(file.path(progDir, "SaemixObject.R"))
33 source(file.path(progDir, "func_plots.R"))
34
35 # Loading of new functions
36 source(file.path(workDir, "multi_aux2.R"))
37 source(file.path(workDir, "multi_initializeMainAlgo.R"))
38 source(file.path(workDir, "multi_estep.R"))
39 source(file.path(workDir, "multi_mstep.R"))
40 source(file.path(workDir, "multi_main.R"))
41 source(file.path(workDir, "multi_map.R"))

```

```

42 source( file . path( workDir , "compute_LL_multi.R" ) )
43
44 # Formatting data with column id, observations, time, ytype
45 d1 = prothro[,c(1,2,3)]
46 d1$ytype=1
47 colnames(d1)[2] = "obs"
48 d2 = prothros[,c(1,3,2)]
49 d2$ytype = 2
50 colnames(d2)[2] = "obs"
51 colnames(d2)[3] = "time"
52 data_joint = rbind(d1,d2)
53
54 # creating saemixData object
55 dataJM<-saemixData(name.data=data_joint, name.group=c("id"),
56                   name.predictors=c("time","obs"),name.response="obs",
57                   name.ytype = "ytype")
58
59 # Initialization of parameters
60 param<-c(73,1.25,0.6,0.0001)
61 omega.sim<-c(18, 3, 0.05, 0.01)
62 sigma.sim <- 17
63
64 # model
65 JMmodel<-function(psi, id, xidep) {
66   ytype<-xidep$ytype
67   b0 <- psi[id,1]
68   b1 <- psi[id,2]
69   h0 <- psi[id,3]
70   alpha <- psi[id,4]
71
72   ypred <- b0+b1*xidep[,1] # longitudinal predictions
73
74   T<-xidep[ytype==2,1]
75   Nj <- length(T)
76   ev = xidep$obs[ytype==2]
77   cens<-which(ev==0)
78   ind <- which(ev==1)
79   b0b = b0[ytype==2] # vectors of same length than T
80   b1b = b1[ytype==2]
81   h0b = h0[ytype==2]
82   alphab = alpha[ytype==2]
83
84   haz <- h0b*exp(alphab*(b0b+b1b*T)) # instantaneous hazard

```

```

85 H <- (h0b/(alphab*b1b))*exp((b0b+b1b*T)*alphab)-
86     (h0b/(alphab*b1b))*exp(alphab*b0b) # cumulative hazard
87
88 logpdf <- rep(0,Nj)
89 logpdf[cens] <- -H[cens]
90 logpdf[ind] <- -H[ind] + log(haz[ind])
91
92 ypred[ytype==2] = logpdf # survival predictions
93 return(ypred)
94 }
95
96 # creating saemixModel object
97 # log normal distribution for h0, normal for other parameters
98 # no variability on survival parameters (h0 and alpha)
99 # homoscedastic error
100 jointTTE<-saemixModel(model=JMmodel,description="JM lin longi one tte",
101                       modeltype=c("structural","likelihood"),
102                       psi0=matrix(param,ncol=4,byrow=TRUE,
103                                   dimnames=list(NULL,c("b0","b1","h0","alpha"))),
104                       transform.par=c(0,0,1,0),covariance.model=diag(c(1,1,0,0)),
105                       fixed.estim = c(1,1,1,1),error.model = "constant",
106                       omega.init = diag(omega.sim))
107
108 saemix.data<-dataJM
109 saemix.model<-jointTTE
110 saemix.options<-saemixControl(seed=12345, map=F, fim=T, ll.is=T)
111
112 yfit <- saemix.multi(saemix.model, saemix.data, saemix.options)
113 summary(yfit) # parameter estimates + likelihood
114 yfit@results@fim # inverse of the Fisher Information Matrix
115 sqrt(diag(yfit@results@fim)) # SE of parameter estimates
116
117
118
119 ##### with JM
120
121 modlin = lme(pro~time, random = ~1+time|id, data = prothro)
122 modsurv = coxph(Surv(Time, death) ~ 1, data = prothros, x = TRUE)
123
124 jm = jointModel(modlin, modsurv, timeVar = "time", method = "piecewise-PH-GH",
125                control = list(lng.in.kn=1))
126 summary(jm)

```

**Appendix D. Relative bias and relative root mean square errors obtained on 100 simulations, for the LMEM and NLMEM**

	LMEM		NLMEM	
	RB (%)	RRMSE (%)	RB (%)	RRMSE (%)
<i>Longitudinal model</i>				
<i>Fixed effects</i>				
$\mu_0$	-0.61	5.2	-0.1	5.0
$\mu_1$	3.5	18.6	0.08	6.1
$\mu_2$	-	-	-0.9	6.0
$\mu_a$	-	-	10.9	27.4
<i>Random effects</i>				
$\omega_0^2$	-2.5	15.0	-0.26	14.7
$\omega_1^2$	1.2	14.6	-7.9	39.3
$\omega_2^2$	-	-	-7.4	38.0
$\omega_a^2$	-	-	-13.1	36.2
<i>Error model</i>				
$\sigma$	0.15	1.4	0.19	1.4

Table D.4: Relative Bias (RB) and Relative Root Mean Square Error (RRMSE) obtained on 100 simulations, for the LMEM and NLMEM

## Appendix E. Distribution of relative estimation errors and relative standard errors for LMEM and NLMEM

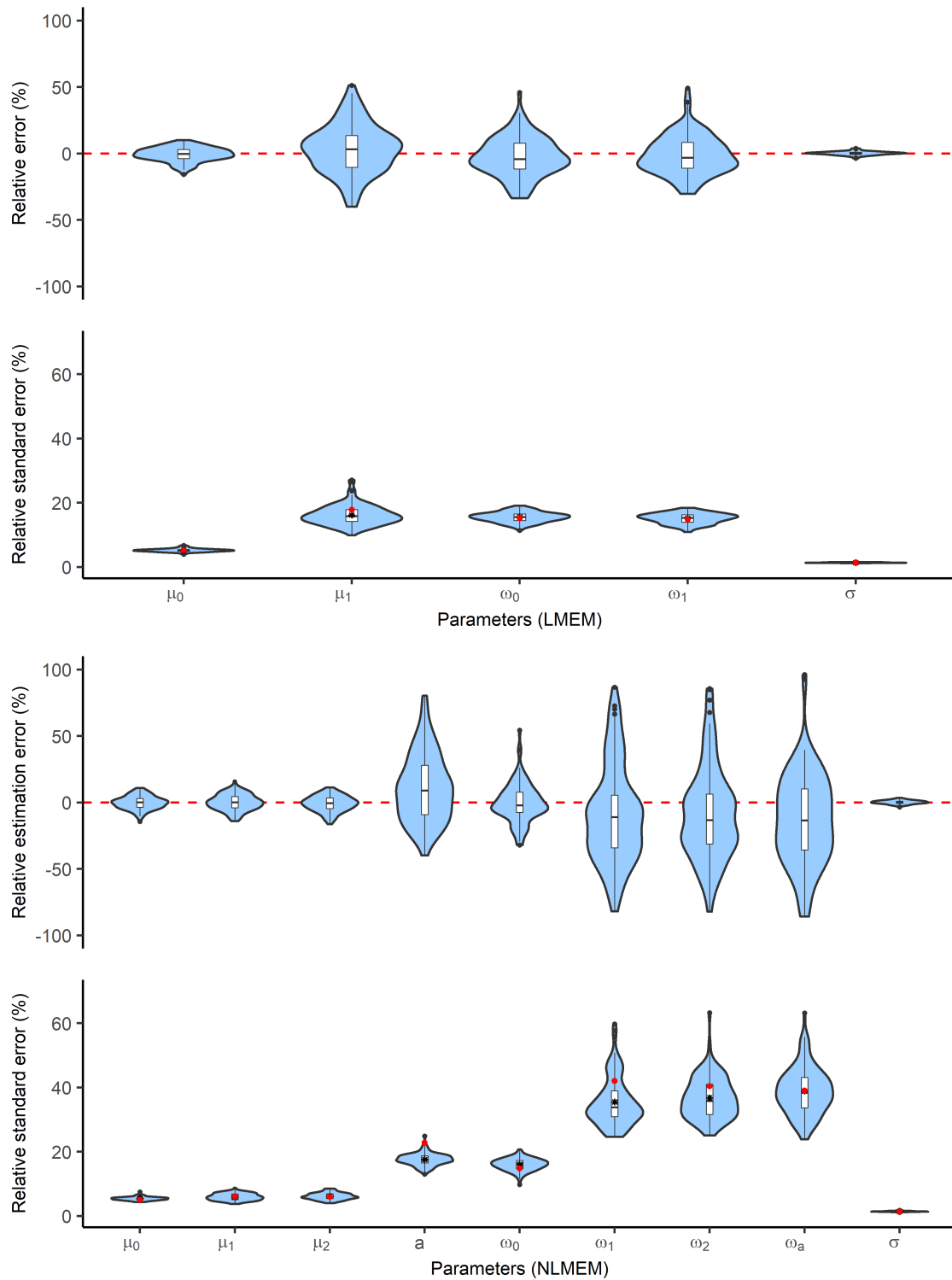


Figure E.6: Distribution of the relative estimation errors (top) and relative standard errors (bottom) for linear and nonlinear mixed-effects models. Stars correspond to the mean distribution of the RSE. Red points correspond to the empirical RSE obtained over all the simulations.

## Appendix F. Prothrombin example - Fit using *JM* package

Of note, *JM* allows to specify a piecewise constant baseline risk with at least 1 internal knot that we put at the median of the observed event times. We denote by  $h_0^{(1)}$  and  $h_0^{(2)}$  the two baseline constant risks.

Parameters	Estimates	S.E.	R.S.E (%)
<i>Longitudinal model</i>			
<i>Fixed effects</i>			
$\mu_0$	73.4	0.87	1.2
$\mu_1$	0.73	0.28	38.4
<i>Random effects</i>			
$\omega_0^2$	384	0.04	0.01
$\omega_1^2$	15	0.37	2.5
<i>Error model</i>			
$\sigma$	17.2	0.015	0.09
<i>Survival model</i>			
$h_0^{(1)}$	3.01	0.24	8.0
$h_0^{(2)}$	3.08	0.27	8.8
$\alpha$	-0.039	0.003	7.7

Table F.5: Parameter estimates on the real data application

## CHAPITRE 5

---

### Discussion et conclusion

---

L'objectif principal de cette thèse s'adresse au développement de modèles conjoints non-linéaires avec risques compétitifs, en proposant des applications aux maladies infectieuses sévères. Dans un premier temps, nous avons montré qu'il est possible d'estimer les paramètres d'un modèle joint composé d'un MNLEM et de deux événements en compétition décrits par leur risque de sous-distribution à l'aide d'un logiciel habituellement utilisé en pharmacométrie. En effet, nous avons montré les capacités de l'algorithme SAEM implémenté dans le logiciel Monolix à fournir des estimations précises des paramètres d'un tel modèle. Une application réelle basée sur des patients admis en réanimation pour sepsis avec un suivi quotidien du score SOFA a permis d'illustrer la démarche. Puis nous avons étendu la modélisation au cas multivarié, dans le cadre d'une cohorte où plusieurs biomarqueurs longitudinaux étaient mesurés chez des patients hospitalisés pour infection au SARS-CoV-2. Dans cet exemple, le nombre de points par biomarqueur est considérablement plus pauvre que dans l'exemple des patients sepsis. Nous avons proposé une stratégie de sélection des biomarqueurs puis démontré les bonnes performances d'estimation de l'algorithme SAEM et de cette stratégie de sélection. Dans un dernier projet, nous

avons pour objectif de diffuser l'utilisation de ces modèles principalement modélisables dans des logiciels issus de la pharmacométrie, aux logiciels généralistes comme R. Pour cela, nous avons étendu le code du package *saemix* initialement dédié à l'estimation de MNLEM au cas des réponses multiples et modèles joints.

## 5.1 Discussion

Dans le cadre de risques compétitifs, situation rencontrée dans de nombreuses applications médicales, les développements en termes de modélisation conjointe se limitaient généralement à décrire la dynamique du biomarqueur par des modèles linéaires mixtes. En effet, l'utilisation de MNLEM mécanistiques entraîne une complexité numérique plus importante. Au début de cette thèse, nous n'avions trouvé aucune étude qui s'adressait à cette question. Entre temps, (Krishnan et al., 2021) ont proposé un modèle conjoint non-linéaire en lien avec plusieurs évènements modélisés par un modèle multi-états. Les paramètres de leur modèle ont été estimés via l'algorithme SAEM implémenté dans le logiciel NONMEM, mais aucune étude de simulation n'a été proposée pour étudier les performances d'estimation. Dans les deux premiers travaux de cette thèse, nous avons montré que l'estimation de modèles conjoints non-linéaires avec risques compétitifs décrits par des fonctions de sous-distribution était réalisable à l'aide du logiciel Monolix. Bien qu'aucune fonction prédéfinie ne soit disponible par défaut pour estimer un tel modèle, et que le code ne soit pas en source ouverte, nous avons modifié le jeu de données pour rendre l'estimation calculable. Ainsi, nous avons montré que l'algorithme SAEM implémenté dans ce logiciel fournissait des estimations précises et non biaisées des paramètres, et ceci dans différents cas de figure. En effet, dans le premier travail, les données et le design étaient riches : plus de 6000 patients ayant une mesure quotidienne du marqueur. Dans le deuxième travail, le design était plus pauvre, avec 327 patients ayant plusieurs biomarqueurs mesurés à intervalles irréguliers. Toutefois, nous n'avons pas évalué de manière précise l'influence du nombre de patients, du nombre de mesures par patient ou encore du nombre d'évènements sur la qualité de l'estimation des paramètres. Nous avons

observé plus d'incertitude sur l'estimation lorsque les données étaient moins riches, mais une évaluation rigoureuse de plans optimaux (connus sous le nom de "optimal design" en anglais) serait une extension intéressante. De plus, les valeurs initiales des paramètres ont été fixées aux vraies valeurs utilisées pour simuler les données. Nous n'avons pas évalué la sensibilité de l'algorithme à ces valeurs initiales. D'autres études dans le contexte des MNLEM ont souligné que les valeurs initiales éloignées pouvaient mener à des estimations moins précises, mais que ce problème pouvait être corrigé en augmentant notamment le nombre de chaînes et/ou d'itérations de l'algorithme (Makowski and Lavielle, 2006).

Dans l'ensemble de ces travaux, nous avons modélisé les risques compétitifs par des fonctions de sous-distribution pour leur avantage d'être directement liées aux incidences cumulées des évènements. Comme évoqué en introduction et à la différence des risques cause-spécifique, les risques de sous-distribution font l'objet d'une mauvaise spécification. En effet, déjà dans une démarche de régression avec covariables (sans modélisation conjointe), pour certaines valeurs de covariables et certains temps définis, la somme des incidences cumulées des deux évènements pour un sujet donné peut dépasser 1. Ce phénomène est clairement démontré dans d'importants travaux menés récemment (Austin et al., 2021). Dans le cadre de patients victimes d'un infarctus du myocarde aigu, les auteurs modélisent le risque de décès pour cause cardiovasculaire versus le risque de décès pour cause autre que cardiovasculaire. Ils évaluent l'effet de l'âge et du sexe sur ces risques et montrent que, pour certaines combinaisons d'âge et de sexe, la probabilité estimée de décès toute cause confondue excède 1. Pour les patients concernés (5% dans leur cas), les auteurs proposent d'utiliser les risques cause-spécifique pour retrouver les bonnes propriétés. Dans ce travail de thèse, nous avons proposé d'utiliser une alternative dans le calcul des probabilités. Notre méthode repose sur l'utilisation de la valeur de la somme des incidences cumulées des deux évènements à un temps infini. Cette approche n'a pas été évaluée en détail par des approches de simulation, mais a cependant permis d'obtenir des probabilités d'évènement entre 0 et 1 pour tous les sujets, y compris les sujets problématiques. En outre, le problème de mauvaise spécification ne permet pas d'évaluer, dans

des démarches de simulation, les paramètres utilisés pour estimer la fonction de sous-distribution de l'évènement en compétition. Un célèbre aphorisme énoncé par Gorges Box dit, "all models are wrong, but some are useful" (Box, 1979). Cette citation trouve un grand intérêt ici. Certes, la mauvaise spécification des risques est inhérente à l'approche sous-distribution, mais celle-ci présente l'avantage de fournir un lien direct entre le risque instantané et l'incidence cumulée, ce qui n'est pas le cas avec l'approche cause-spécifique. Pour cette raison, l'approche par risques de sous-distribution est une approche "facile à interpréter" pour modéliser les données en présence de risques compétitifs, surtout lorsque le but concerne la prédiction des évènements.

Dans les deux applications cliniques proposées, nous avons montré l'importance de considérer le suivi longitudinal de biomarqueurs pour améliorer les prédictions du pronostic des patients. Cependant, lorsque peu de mesures de biomarqueurs sont disponibles, l'utilisation d'une modélisation complexe n'est plus adaptée, et les modèles dits baseline, reposant sur des données disponibles à l'admission, sont préférés. En effet, dans le cadre de l'application des patients sepsis tirés de l'entrepôt OUTCOMEREA, le modèle incluant uniquement le SOFA à l'admission témoigne de meilleures performances prédictives que le modèle joint lorsque seulement cette valeur est disponible. De même, dans le cadre de l'application des patients COVID-19 de la cohorte RISCOV, le modèle joint n'était pas significativement meilleur que le modèle incluant uniquement le score 4C à l'admission, lorsque seulement les mesures des neutrophiles, du pH et de la CRP disponibles jusqu'à J3 étaient modélisées. Pour des temps landmark plus lointains, nous avons montré les bénéfices de la modélisation conjointe.

Dans le premier travail, nous avons traité le SOFA en tant que variable continue, comme dans la plupart des autres études (Deslandes and Chevret, 2010; Musoro et al., 2018). Comme le SOFA est une somme de sous-scores (par organe) discrets (catégoriels ordonnés) la modélisation IRT (pour Item Response Theory) aurait pu être considérée. Cette modélisation a déjà été appliquée dans le cadre des modèles à effets mixtes (Buatois et al., 2017), puis dans le cadre des modèles joints (Saulnier et al., 2022).

Dans le contexte des modèles joints et lorsque beaucoup de biomarqueurs sont disponibles, la première difficulté consiste à les modéliser correctement. Dans le cadre de patients hospitalisés pour infection au SARS-CoV-2 ayant de multiples biomarqueurs suivis au cours du temps, nous avons proposé deux modèles pour décrire la dynamique des biomarqueurs, qui arbitrent différemment entre complexité et flexibilité. Le premier, le modèle linéaire, est peu flexible mais a l'avantage d'être parcimonieux et facilite ainsi l'estimation des paramètres. Le deuxième, le modèle non-linéaire composé d'une double exponentielle, est beaucoup plus flexible puisqu'il autorise des évolutions monotones et non-monotones mais complexifie l'estimation. Nous reconnaissons que des évolutions complexes de biomarqueurs ne sont pas autorisées avec ces deux modèles. Nous avons établi une stratégie pour assigner un des deux modèles à chaque biomarqueur puis pour exclure les biomarqueurs non significativement associés au décès. Une analyse de sensibilité a été proposée et a révélé la robustesse de la stratégie, en soulignant toutefois que des critères trop stricts pourraient limiter les performances en excluant trop de biomarqueurs. Selon les données récoltées et le design de l'étude, ces critères sont vraisemblablement à adapter. Dans une étape multivariée, nous avons établi une sélection basée sur une approche backward, ce qui nous a permis d'arriver à un modèle final parcimonieux mais ayant de bonnes capacités prédictives. Nous avons insisté sur le fait de ne pas considérer un ensemble trop important de biomarqueurs au début de l'étape multivariée, pouvant mener à des limites computationnelles en termes de temps de calcul et de stabilité de l'algorithme. Nous reconnaissons que les approches de régression backward sont connues pour être surpassées par des approches de régression dites pénalisées (Tsiatis et al., 1995). Cependant, ces concepts n'ont jamais été appliqués à la problématique de sélection de sous-modèles longitudinaux dans un modèle conjoint et nécessitent de nouveaux développements. Ce point sera développé dans les perspectives en Section 5.2. Les stratégies de sélection proposées dans ce travail sont certes *ad hoc* mais originales et relativement facilement applicables dans d'autres contextes cliniques.

Dans un dernier projet, nous avons fait le constat que très peu d'outils en dehors

des logiciels utilisés en pharmacométrie étaient disponibles pour l'estimation de modèles joints lorsque la dynamique du biomarqueur était modélisée par un modèle non-linéaire. En effet, si on considère le logiciel R, très largement utilisé par la communauté scientifique, seul le package *frailtypack* (Rondeau et al., 2019a) permet d'estimer un modèle conjoint non-linéaire avec évènement terminal ou récurrent. Toutefois, une seule forme possible pour le modèle structurel non-linéaire est autorisée (modèle de croissance tumorale). Nous avons donc étendu le code du package *saemix* (Comets et al., 2017) initialement dédié à l'estimation de modèles non-linéaires à effets mixtes, au cas des réponses multiples et modèles joints dans un cadre totalement paramétrique. Nous avons également intégré un algorithme stochastique récemment développé (Delattre and Kuhn, 2023) pour calculer les erreurs standard des paramètres. Une étude de simulation nous a permis de démontrer les bonnes performances d'estimation des paramètres et de leur incertitude. Cette extension a l'avantage de fournir aux utilisateurs un outil flexible dans le sens où n'importe quel modèle joint paramétrique peut être estimé, car l'utilisateur définit lui-même la vraisemblance du modèle. Ce dernier doit donc être en mesure d'écrire explicitement la vraisemblance, ce qui n'est pas tout le temps possible. Dans ce cas, des techniques d'intégration numérique peuvent être employées mais allongent considérablement le temps de calcul.

## 5.2 Perspectives

Plusieurs perspectives s'offrent à ce travail de thèse, mêlant méthodologie et applications cliniques.

Que ce soit dans l'application des patients admis en soins intensifs pour sepsis ou des patients hospitalisés pour infection au SARS-CoV-2, les traitements reçus tout au long du séjour à l'hôpital n'ont pas été pris en compte dans l'analyse. La modélisation proposée pourrait toutefois considérer des éventuels effets de traitements sur les risques de décès et de sortie d'hospitalisation et mener à des évaluations thérapeutiques comme déjà proposées dans de nombreuses applications (Desmée et al., 2015; van Oudenhoven et al., 2020;

Kerioui et al., 2023). Le calcul des prédictions individualisées et dynamiques proposé dans ces travaux pourrait également faire l'objet de développements d'applications sur tablettes ou smartphones à destination des cliniciens. Les prédictions dynamiques évoquées dans ces travaux reposent sur le calcul de lois conditionnelles et sont ainsi faciles à implémenter. Ces outils représentent des aides à la décision, mais ne remplacent en aucun cas les connaissances et intuitions des praticiens. Des outils similaires ont déjà été proposés dans d'autres contextes (Choo et al., 2019; Naseer Qureshi et al., 2020).

D'un point de vue méthodologique et comme évoqué à plusieurs reprises dans ce manuscrit, nous avons proposé une méthode pour contourner le problème de la mauvaise spécification de l'approche sous-distribution afin de calculer des prédictions individuelles dynamiques. A ce jour, nous n'avons pas évalué cette méthode et d'importants développements seraient nécessaires pour la valider. Ce point représente un sujet de recherche novateur et intéressant car très peu d'études ont traité de cette question, pourtant fondamentale dans ce genre d'approche.

Lorsque nous avons travaillé sur la question de la sélection de sous-modèles longitudinaux dans des modèles joints multivariés, nous avons proposé une stratégie certes *ad hoc* mais performante dans le cadre de notre application. Cette stratégie repose notamment sur une procédure backward où à chaque itération, le biomarqueur ayant le lien le moins significatif sort de l'analyse. Pour rappel, lorsqu'un biomarqueur est exclu, il est exclu des deux risques de sous-distribution. Des analyses de sensibilité permettraient d'éclairer sur l'impact de ce choix en termes de performances prédictives. En effet, nous pourrions considérer des scénarios alternatifs où le biomarqueur est exclu uniquement pour le risque de sous-distribution en question. Le principe de la stratégie évoquée repose sur l'utilisation du test de Wald sur les coefficients de liens, et donc de l'estimation des erreurs standard des paramètres. Bien que le logiciel Monolix ait été montré performant pour estimer les erreurs standard dans le contexte des MNLEM (Samson et al., 2007) puis des modèles joints avec un évènement terminal (Mbogning et al., 2015), il serait particulièrement intéressant de valider l'estimation pour le cas de l'approche avec risques compétitifs. Dans

la discussion, nous avons mentionné que notre stratégie pouvait être surpassée en utilisant des techniques de régressions dites pénalisées comme la régression LASSO. Inclure une sélection LASSO au sein d'un modèle joint pour sélectionner les sous-modèles longitudinaux n'a encore jamais été proposé et représenterait une contribution majeure. Ces développements nécessitent de modifier l'algorithme SAEM et donc d'avoir accès au code source. L'extension de *saemix* pour le cas des réponses multiples et modèles joints proposée dans le troisième travail de cette thèse pourrait ainsi servir de base à ces nouveaux développements.

Par ailleurs, dans ce travail d'extension de *saemix*, nous avons inclus un algorithme stochastique pour estimer la FIM observée et dériver les erreurs standard des estimations des paramètres. Dans un futur travail, nous pourrions implémenter également la méthode basée sur la décomposition de Louis et utiliser les algorithmes stochastiques décrits en Section 1.2.2. A titre informatif, l'algorithme stochastique décrit dans l'équation 1.30 est utilisé par le logiciel Monolix. Il serait intéressant de faire des analyses pour comparer les trois approches notamment en termes de précision d'estimation et de temps de calcul. Des comparaisons directes avec le logiciel Monolix pourraient également être menées, mais des facteurs seraient à prendre en compte car des critères d'auto-stop sont implémentés dans Monolix, à la différence de *saemix*.

Ce dernier projet de thèse offre de nombreuses autres possibilités de développement. Le code est disponible sur Github et des exemples sont présentés afin de l'utiliser facilement. Pour information, la version courante de *saemix* est la version 3.2. L'extension du code que nous avons proposé basé sur la version 3.1 est utilisable depuis la version courante. La suite de ce travail serait d'intégrer le code dans le package *saemix* directement utilisable depuis R. Pour cela, des développements restent à prévoir notamment pour autoriser la corrélation entre effets aléatoires ou pour fournir des graphes diagnostiques adaptés. Cette future extension sera idéalement programmée une fois les versions *saemix 3* finalisées et formera éventuellement une version 4.

### 5.2.1 Conclusion générale

Les modèles conjoints pour données longitudinales et de survie à risques compétitifs ont connu une utilisation importante ces dernières années. Souvent limités à des modèles linéaires pour décrire le processus longitudinal et des approches cause-spécifique pour décrire les risques instantanés d'évènements, leur adaptation aux applications où les biomarqueurs ont des dynamiques mécanistiques non-linéaires et où les risques instantanés se veulent être décrits par des approches sous-distribution est non sans intérêt. Dans ce contexte, ces travaux de thèse proposent des développements et offrent des outils adaptés à ce genre de modélisation. D'un point de vue clinique, ces travaux soutiennent le développement de la médecine personnalisée et peuvent aussi être utilisés dans des domaines autres que les maladies infectieuses, afin d'optimiser l'évaluation thérapeutique et la prédiction d'évènements au sein d'une population. D'un point de vue méthodologique, cette thèse ouvre la voie à de nombreuses perspectives de développements pour améliorer notamment la prise en compte de données en grande dimension au sein d'une modélisation conjointe et pour diffuser l'utilisation de ces méthodes aux logiciels généralistes.



---

## Bibliographie

---

- Andrinopoulou, E.-R., Rizopoulos, D., Takkenberg, J. J., and Lesaffre, E. (2017). Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. *Statistical Methods in Medical Research*, 26(4) :1787–1801.
- Armiñanzas, C., Arnaiz de las Revillas, F., Gutiérrez Cuadra, M., Arnaiz, A., Fernández Sampedro, M., González-Rico, C., Ferrer, D., Mora, V., Suberviola, B., Latorre, M., Calvo, J., Olmos, J. M., Cifrián, J. M., and Fariñas, M. C. (2021). Usefulness of the COVID-GRAM and CURB-65 scores for predicting severity in patients with COVID-19. *International Journal of Infectious Diseases*, 108 :282–288.
- Austin, P. C. and Fine, J. P. (2017). Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Statistics in Medicine*, 36(27) :4391–4400.
- Austin, P. C., Steyerberg, E. W., and Putter, H. (2021). Fine-Gray subdistribution hazard models to simultaneously estimate the absolute risk of different event types : Cumulative total failure probability may exceed 1. *Statistics in Medicine*, 40(19) :4200–4212.
- Bauer, R. J. (2019). NONMEM Tutorial Part I : Description of Commands and Options, With Simple Examples of Population Analysis. *CPT : Pharmacometrics & Systems Pharmacology*, 8(8) :525–537.

- Blanche, P., Dartigues, J.-F., and Jacqmin-Gadda, H. (2013). Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Statistics in Medicine*, 32(30) :5381–5397.
- Blanche, P., Proust-Lima, C., Loubère, L., Berr, C., Dartigues, J.-F., and Jacqmin-Gadda, H. (2015). Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics*, 71(1) :102–113.
- Box, G. E. P. (1979). Robustness in the Strategy of Scientific Model Building. In Launer, R. L. and Wilkinson, G. N., editors, *Robustness in Statistics*, pages 201–236. Academic Press.
- Brown, V. A. (2021). An Introduction to Linear Mixed-Effects Modeling in R. *Advances in Methods and Practices in Psychological Science*, 4(1) :2515245920960351.
- Buatois, S., Retout, S., Frey, N., and Ueckert, S. (2017). Item Response Theory as an Efficient Tool to Describe a Heterogeneous Clinical Rating Scale in De Novo Idiopathic Parkinson’s Disease Patients. *Pharmaceutical Research*, 34(10) :2109–2118.
- Choo, C. C., Devakaran, B., Chew, P. K. H., and Zhang, M. W. B. (2019). Smartphone Application in Postgraduate Clinical Psychology Training : Trainees’ Perspectives. *International Journal of Environmental Research and Public Health*, 16(21) :4206.
- Clanché, F. (2022). Covid-19 et mortalité : les chiffres disponibles et leurs usages. *DREES MÉTHODES*.
- Cohen, J., Vincent, J.-L., Adhikari, N. K. J., Machado, F. R., Angus, D. C., Calandra, T., Jaton, K., Giulieri, S., Delaloye, J., Opal, S., Tracey, K., van der Poll, T., and Pelfrene, E. (2015). Sepsis : a roadmap for future research. *The Lancet. Infectious Diseases*, 15(5) :581–614.
- Comets, E., Lavenu, A., and Lavielle, M. (2017). Parameter Estimation in Nonlinear

- Mixed Effect Models Using saemix, an R Implementation of the SAEM Algorithm. *Journal of Statistical Software*, 80 :1–41.
- Davidian, M. and Gallant, A. R. (1992). Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with application to quinidine. *Journal of Pharmacokinetics and Biopharmaceutics*, 20(5) :529–556.
- de Jong, V. M. T., Rousset, R. Z., Antonio-Villa, N. E., Buenen, A. G., Van Calster, B., Bello-Chavolla, O. Y., Brunskill, N. J., Curcin, V., Damen, J. A. A., Fermín-Martínez, C. A., Fernández-Chirino, L., Ferrari, D., Free, R. C., Gupta, R. K., Haldar, P., Hedberg, P., Korang, S. K., Kurstjens, S., Kusters, R., Major, R. W., Maxwell, L., Nair, R., Naucler, P., Nguyen, T.-L., Noursadeghi, M., Rosa, R., Soares, F., Takada, T., van Royen, F. S., van Smeden, M., Wynants, L., Modrák, M., CovidRetro collaboration, Asselbergs, F. W., Linschoten, M., CAPACITY-COVID consortium, Moons, K. G. M., and Debray, T. P. A. (2022). Clinical prediction models for mortality in patients with covid-19 : external validation and individual participant data meta-analysis. *British Medical Journal (Clinical research ed.)*, 378 :e069881.
- Delattre, M. and Kuhn, E. (2023). Estimating Fisher Information Matrix in Latent Variable Models based on the Score Function. hal-02285712v2.
- Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a stochastic approximation version of EM algorithm. *The Annals of Statistics*, 27 :94.
- Deslandes, E. and Chevret, S. (2010). Joint modeling of multivariate longitudinal data and the dropout process in a competing risk setting : application to ICU data. *BMC Medical Research Methodology*, 10 :69.
- Desmée, S., Mentré, F., Veyrat-Follet, C., and Guedj, J. (2015). Nonlinear Mixed-Effect Models for Prostate-Specific Antigen Kinetics and Link with Survival in the Context of Metastatic Prostate Cancer : a Comparison by Simulation of Two-Stage and Joint Approaches. *The AAPS Journal*, 17(3) :691–699.

- Desmée, S., Mentré, F., Veyrat-Follet, C., Sébastien, B., and Guedj, J. (2017). Nonlinear joint models for individual dynamic prediction of risk of death using Hamiltonian Monte Carlo : application to metastatic prostate cancer. *BMC Medical Research Methodology*, 17(1) :105.
- Efron, B. and Hinkley, D. V. (1978). Assessing the accuracy of the maximum likelihood estimator : Observed versus expected Fisher information. *Biometrika*, 65(3) :457–483.
- Ferreira, F. L., Bota, D. P., Bross, A., Mélot, C., and Vincent, J.-L. (2001). Serial Evaluation of the SOFA Score to Predict Outcome in Critically Ill Patients. *JAMA*, 286(14) :1754–1758.
- Fidler, M., Xiong, Y., Schoemaker, R., Wilkins, J., Trame, M., Hooijmaijers, R., Post, T., and Wang, W. (2022). *nlmixr : Nonlinear Mixed Effects Models in Population Pharmacokinetics and Pharmacodynamics*. R package version 2.0.7.
- Fine, J. P. and Gray, R. J. (1999). A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*, 94(446) :496–509.
- Garcia-Hernandez, A. and Rizopoulos, D. (2018). %JM : A SAS Macro to Fit Jointly Generalized Mixed Models for Longitudinal Data and Time-to-Event Responses. *Journal of Statistical Software*, 84 :1–29.
- Gerds, T. A. and Schumacher, M. (2006). Consistent estimation of the expected Brier score in general survival models with right-censored event times. *Biometrical Journal*, 48(6) :1029–1040.
- Guedj, J., Thiébaud, R., and Commenges, D. (2011). Joint modeling of the clinical progression and of the biomarkers’ dynamics using a mechanistic model. *Biometrics*, 67(1) :59–66.
- Hassan, S., Ramspek, C. L., Ferrari, B., Diepen, M. v., Rossio, R., Knevel, R., Mura, V. l., Artoni, A., Martinelli, I., Bandera, A., Nobili, A., Gori, A., Blasi, F., Canetta,

- C., Montano, N., Rosendaal, F. R., and Peyvandi, F. (2022). External validation of risk scores to predict in-hospital mortality in patients hospitalized due to coronavirus disease 2019. *European Journal of Internal Medicine*, 102 :63–71.
- Heagerty, P. J. and Zheng, Y. (2005). Survival model predictive accuracy and ROC curves. *Biometrics*, 61(1) :92–105.
- Hickey, G. L., Philipson, P., and and, A. J. (2018). joinerml : a joint model and software package for time-to-event and multivariate longitudinal outcomes. *BMC Medical Research Methodology*, 18(1).
- Häger, L., Wendland, P., Biergans, S., Lederer, S., de Arruda Botelho Herr, M., Erhardt, C., Schmauder, K., Kschischo, M., Malek, N. P., Bunk, S., Bitzer, M., Gladstone, B. P., and Göpel, S. (2022). External Validation of COVID-19 Risk Scores during Three Waves of Pandemic in a German Cohort—A Retrospective Study. *Journal of Personalized Medicine*, 12(11) :1775.
- Institute, S. (1985). *SAS user's guide : Statistics*, volume 2. Sas Inst.
- Karakike, E., Kyriazopoulou, E., Tsangaris, I., Routsis, C., Vincent, J.-L., and Giamarellos-Bourboulis, E. J. (2019). The early change of SOFA score as a prognostic marker of 28-day sepsis mortality : analysis through a derivation and a validation cohort. *Critical Care*, 23(1).
- Kerioui, M., Bertrand, J., Bruno, R., Mercier, F., Guedj, J., and Desmée, S. (2022). Modelling the association between biomarkers and clinical outcome : An introduction to nonlinear joint models. *British Journal of Clinical Pharmacology*, 88(4) :1452–1463.
- Kerioui, M., Bertrand, J., Desmée, S., Le Tourneau, C., Mercier, F., Bruno, R., and Guedj, J. (2023). Assessing the Increased Variability in Individual Lesion Kinetics During Immunotherapy : Does It Exist, and Does It Matter ? *JCO Precision Oncology*, (7) :e2200368. Publisher : Wolters Kluwer.

Klimczak-Tomaniak, D., de Bakker, M., Bouwens, E., Akkerhuis, K. M., Baart, S., Rizopoulos, D., Mouthaan, H., van Ramshorst, J., Germans, T., Constantinescu, A., Manintveld, O., Umans, V., Boersma, E., and Kardys, I. (2022). Dynamic personalized risk prediction in chronic heart failure patients : a longitudinal, clinical investigation of 92 biomarkers (Bio-SHiFT study). *Scientific Reports*, 12 :2795.

Knight, S. R., Ho, A., Pius, R., Buchan, I., Carson, G., Drake, T. M., Dunning, J., Fairfield, C. J., Gamble, C., Green, C. A., Gupta, R., Halpin, S., Hardwick, H. E., Holden, K. A., Horby, P. W., Jackson, C., Mclean, K. A., Merson, L., Nguyen-Van-Tam, J. S., Norman, L., Noursadeghi, M., Olliaro, P. L., Pritchard, M. G., Russell, C. D., Shaw, C. A., Sheikh, A., Solomon, T., Sudlow, C., Swann, O. V., Turtle, L. C., Openshaw, P. J., Baillie, J. K., Semple, M. G., Docherty, A. B., Harrison, E. M., Baillie, J. K., Semple, M. G., Openshaw, P. J., Carson, G., Alex, B., Bach, B., Barclay, W. S., Bogaert, D., Chand, M., Cooke, G. S., Docherty, A. B., Dunning, J., da Silva Filipe, A., Fletcher, T., Green, C. A., Harrison, E. M., Hiscox, J. A., Ho, A. Y. W., Horby, P. W., Ijaz, S., Khoo, S., Klenerman, P., Law, A., Lim, W. S., Mentzer, A. J., Merson, L., Meynert, A. M., Noursadeghi, M., Moore, S. C., Palmarini, M., Paxton, W. A., Pollakakis, G., Price, N., Rambaut, A., Robertson, D. L., Russell, C. D., Sancho-Shimizu, V., Scott, J. T., Sigfrid, L., Solomon, T., Sriskandan, S., Stuart, D., Summers, C., Tedder, R. S., Thomson, E. C., Thwaites, R. S., Turtle, L. C., Zambon, M., Hardwick, H., Donohue, C., Ewins, J., Oosthuyzen, W., Griffiths, F., Norman, L., Pius, R., Drake, T. M., Fairfield, C. J., Knight, S., Mclean, K. A., Murphy, D., Shaw, C. A., Dalton, J., Girvan, M., Saviciute, E., Roberts, S., Harrison, J., Marsh, L., Connor, M., Halpin, S., Jackson, C., Gamble, C., Leeming, G., Law, A., Hendry, R., Scott-Brown, J., Greenhalf, W., Shaw, V., McDonald, S., Ahmed, K. A., Armstrong, J. A., Ashworth, M., Asimwe, I. G., Bakshi, S., Barlow, S. L., Booth, L., Brennan, B., Bullock, K., Catterall, B. W., Clark, J. J., Clarke, E. A., Cole, S., Cooper, L., Cox, H., Davis, C., Dincarslan, O., Dunn, C., Dyer, P., Elliott, A., Evans, A., Fisher, L. W., Foster, T., Garcia-Dorival, I., Greenhalf, W., Gunning, P., Hartley, C., Ho, A., Jensen, R. L., Jones, C. B., Jones,

T. R., Khandaker, S., King, K., Kiy, R. T., Koukorava, C., Lake, A., Lant, S., Latawiec, D., Lavelle-Langham, L., Lefteri, D., Lett, L., Livoti, L. A., Mancini, M., McDonald, S., McEvoy, L., McLauchlan, J., Metelmann, S., Miah, N. S., Middleton, J., Mitchell, J., Moore, S. C., Murphy, E. G., Penrice-Randal, R., Pilgrim, J., Prince, T., Reynolds, W., Ridley, P. M., Sales, D., Shaw, V. E., Shears, R. K., Small, B., Subramaniam, K. S., Szemiel, A., Taggart, A., Tanianis-Hughes, J., Thomas, J., Trochu, E., van Tonder, L., Wilcock, E., Zhang, J. E., Adeniji, K., Agranoff, D., Agwuh, K., Ail, D., Alegria, A., Angus, B., Ashish, A., Atkinson, D., Bari, S., Barlow, G., Barnass, S., Barrett, N., Bassford, C., Baxter, D., Beadsworth, M., Bernatoniene, J., Berridge, J., Best, N., Bothma, P., Brealey, D., Brittain-Long, R., Bulteel, N., Burden, T., Burtenshaw, A., Caruth, V., Chadwick, D., Chamblor, D., Chee, N., Child, J., Chukkambotla, S., Clark, T., Collini, P., Cosgrove, C., Cupitt, J., Cutino-Moguel, M.-T., Dark, P., Dawson, C., Dervisevic, S., Donnison, P., Douthwaite, S., DuRand, I., Dushianthan, A., Dyer, T., Evans, C., Eziefula, C., Fegan, C., Finn, A., Fullerton, D., Garg, S., Garg, S., Garg, A., Godden, J., Goldsmith, A., Graham, C., Hardy, E., Hartshorn, S., Harvey, D., Havalda, P., Hawcutt, D. B., Hobrok, M., Hodgson, L., Holme, A., Hormis, A., Jacobs, M., Jain, S., Jennings, P., Kaliappan, A., Kasipandian, V., Kegg, S., Kelsey, M., Kendall, J., Kerrison, C., Kerslake, I., Koch, O., Koduri, G., Koshy, G., Laha, S., Larkin, S., Leiner, T., Lillie, P., Limb, J., Linnett, V., Little, J., MacMahon, M., MacNaughton, E., Mankregod, R., Masson, H., Matovu, E., McCullough, K., McEwen, R., Meda, M., Mills, G., Minton, J., Mirfenderesky, M., Mohandas, K., Mok, Q., Moon, J., Moore, E., Morgan, P., Morris, C., Mortimore, K., Moses, S., Mpenge, M., Mulla, R., Murphy, M., Nagel, M., Nagarajan, T., Nelson, M., Otahal, I., Pais, M., Panchatsharam, S., Paraiso, H., Patel, B., Pepperell, J., Peters, M., Phull, M., Pintus, S., Pooni, J. S., Post, F., Price, D., Prout, R., Rae, N., Reschreiter, H., Reynolds, T., Richardson, N., Roberts, M., Roberts, D., Rose, A., Rousseau, G., Ryan, B., Saluja, T., Shah, A., Shanmuga, P., Sharma, A., Shawcross, A., Sizer, J., Smith, R., Snelson, C., Spittle, N., Staines, N., Stambach, T., Stewart, R., Subudhi, P., Szakmany, T., Tatham, K., Thomas, J.,

- Thompson, C., Thompson, R., Tridente, A., Tupper-Carey, D., Twagira, M., Ustianowski, A., Vallotton, N., Vincent-Smith, L., Visuvanathan, S., Vuylsteke, A., Waddy, S., Wake, R., Walden, A., Welters, I., Whitehouse, T., Whittaker, P., Whittington, A., Wijesinghe, M., Williams, M., Wilson, L., Wilson, S., Winchester, S., Wiselka, M., Wolverson, A., Wooton, D. G., Workman, A., Yates, B., and Young, P. (2020). Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol : development and validation of the 4C Mortality Score. *British Medical Journal*, 370 :m3339.
- Krishnan, S. M., Friberg, L. E., Bruno, R., Beyer, U., Jin, J. Y., and Karlsson, M. O. (2021). Multistate model for pharmacometric analyses of overall survival in HER2-negative breast cancer patients treated with docetaxel. *CPT : pharmacometrics & systems pharmacology*, 10(10) :1255–1266.
- Kuhn, E. and Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects model. *Computational Statistics & Data Analysis*, 49 :1020–1038.
- Latouche, A., Allignol, A., Beyersmann, J., Labopin, M., and Fine, J. P. (2013). A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *Journal of Clinical Epidemiology*, 66(6) :648–653.
- Lau, B., Cole, S. R., and Gange, S. J. (2009). Competing Risk Regression Models for Epidemiologic Data. *American Journal of Epidemiology*, 170(2) :244–256.
- Leung, K.-M., Elashoff, R. M., and Afifi, A. A. (1997). Censoring Issues in Survival Analysis. *Annual Review of Public Health*, 18(1) :83–104.
- Lindstrom, M. J. and Bates, D. M. (1990). Nonlinear Mixed Effects Models for Repeated Measures Data. *Biometrics*, 46(3) :673–687.
- Liu, M., Sun, J., Herazo-Maya, J. D., Kaminski, N., and Zhao, H. (2019). Joint Models for Time-to-Event Data and Longitudinal Biomarkers of High Dimension. *Statistics in Biosciences*, 11(3) :614–629.

- Lombardi, Y., Azoyan, L., Szychowiak, P., Bellamine, A., Lemaitre, G., Bernaux, M., Daniel, C., Leblanc, J., Riller, Q., and Steichen, O. (2021). External validation of prognostic scores for COVID-19 : a multicenter cohort study of patients hospitalized in Greater Paris University Hospitals. *Intensive Care Medicine*, 47(12) :1426–1439.
- Long, J. D. and Mills, J. A. (2018). Joint modeling of multivariate longitudinal data and survival data in several observational studies of Huntington’s disease. *BMC Medical Research Methodology*, 18(1) :138.
- Louis, T. A. (1982). Finding the Observed Information Matrix When Using the EM Algorithm. *Journal of the Royal Statistical Society : Series B (Methodological)*, 44(2) :226–233.
- Makowski, D. and Lavielle, M. (2006). Using SAEM to estimate parameters of models of response to applied fertilizer. *Journal of Agricultural, Biological, and Environmental Statistics*, 11(1) :45–60.
- Mbogning, C., Bleakley, K., and Lavielle, M. (2015). Joint modelling of longitudinal and repeated time-to-event data using nonlinear mixed-effects models and the stochastic approximation expectation–maximization algorithm. *Journal of Statistical Computation and Simulation*, 85(8) :1512–1528.
- Mozumder, S. I., Rutherford, M. J., and Lambert, P. C. (2018). Direct likelihood inference on the cause-specific cumulative incidence function : a flexible parametric regression modelling approach. *Statistics in medicine*, 37(1) :82–97.
- Musoro, J. Z., Zwinderman, A. H., Abu-Hanna, A., Bosman, R., and Geskus, R. B. (2018). Dynamic prediction of mortality among patients in intensive care using the sequential organ failure assessment (SOFA) score : a joint competing risk survival and longitudinal modeling approach. *Statistica Neerlandica*, 72(1) :34–47.
- Naseer Qureshi, K., Din, S., Jeon, G., and Piccialli, F. (2020). An accurate and dynamic

- predictive model for a smart M-Health system using machine learning. *Information Sciences*, 538 :486–502.
- Pepe, M. S. (2003). *The Statistical Evaluation of Medical Tests for Classification and Prediction*. OUP Oxford.
- Philipson, P., Sousa, I., Diggle, P. J., Williamson, P., Kolamunnage-Dona, R., Henderson, R., and Hickey, G. L. (2018). *joiner : Joint Modelling of Repeated Measurements and Time-to-Event Data*.
- Pinheiro, J., Bates, D., and R Core Team (2023). *nlme : Linear and Nonlinear Mixed Effects Models*. R package version 3.1-162.
- Pinheiro, J. C. and Bates, D. M. (1995). Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model. *Journal of Computational and Graphical Statistics*, 4(1) :12–35.
- Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The Analysis of Failure Times in the Presence of Competing Risks. *Biometrics*, 34(4) :541–554.
- Proust-Lima, C., Dartigues, J.-F., and Jacqmin-Gadda, H. (2016). Joint modeling of repeated multivariate cognitive measures and competing risks of dementia and death : a latent process and latent class approach. *Statistics in Medicine*, 35(3) :382–398.
- Putter, H., Fiocco, M., and Geskus, R. B. (2007). Tutorial in biostatistics : competing risks and multi-state models. *Statistics in Medicine*, 26(11) :2389–2430.
- R Core Team (2022). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rajeswaran, J., Blackstone, E. H., and Barnard, J. (2018). Joint Modeling of Multivariate Longitudinal Data and Competing Risks Using Multiphase Sub-models. *Statistics in Biosciences*, 10(3) :651–685.

- Rizopoulos, D. (2010). JM : An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software*, 35(9) :1–33.
- Rizopoulos, D. (2011). Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics*, 67(3) :819–829.
- Rizopoulos, D. (2012). *Joint models for longitudinal and time-to-event data : with applications in R*. Chapman & Hall / CRC biostatistics series. Boca Raton, Fla.
- Rizopoulos, D., Verbeke, G., and Lesaffre, E. (2009). Fully Exponential Laplace Approximations for the Joint Modelling of Survival and Longitudinal Data. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 71(3) :637–654.
- Rondeau, V., Gonzalez, J. R., Mazroui, Y., Mauguen, A., Diakite, A., Laurent, A., Lopez, M., Król, A., and Sofeu, C. L. (2019a). *frailtypack : General Frailty Models : Shared, Joint and Nested Frailty Models with Prediction ; Evaluation of Failure-Time Surrogate Endpoints*. R package version 3.0.3.
- Rondeau, V., Gonzalez, J. R., Mazroui, Y., Mauguen, A., Diakite, A., Laurent, A., Lopez, M., Król, A., and Sofeu, C. L. (2019b). *frailtypack : General Frailty Models : Shared, Joint and Nested Frailty Models with Prediction ; Evaluation of Failure-Time Surrogate Endpoints*. R package version 3.0.3.
- Rudd, K. E., Johnson, S. C., Agesa, K. M., Shackelford, K. A., Tsoi, D., Kievlan, D. R., Colombara, D. V., Ikuta, K. S., Kissoon, N., Finfer, S., Fleischmann-Struzek, C., Machado, F. R., Reinhart, K. K., Rowan, K., Seymour, C. W., Watson, R. S., West, T. E., Marinho, F., Hay, S. I., Lozano, R., Lopez, A. D., Angus, D. C., Murray, C. J. L., and Naghavi, M. (2020). Global, regional, and national sepsis incidence and mortality, 1990–2017 : analysis for the Global Burden of Disease Study. *The Lancet*, 395(10219) :200–211.
- Samson, A., Lavielle, M., and Mentré, F. (2007). The SAEM algorithm for group com-

- parison tests in longitudinal data analysis based on non-linear mixed-effects model. *Statistics in Medicine*, 26(27) :4860–4875.
- Saulnier, T., Philipps, V., Meissner, W. G., Rascol, O., Pavy-Le Traon, A., Foubert-Samier, A., and Proust-Lima, C. (2022). Joint models for the longitudinal analysis of measurement scales in the presence of informative dropout. *Methods*, 203 :142–151.
- Shang, Y., Liu, T., Wei, Y., Li, J., Shao, L., Liu, M., Zhang, Y., Zhao, Z., Xu, H., Peng, Z., Wang, X., and Zhou, F. (2020). Scoring systems for predicting mortality for severe patients with COVID-19. *eClinicalMedicine*, 24 :100426.
- Sheiner, L. B., Rosenberg, B., and Melmon, K. L. (1972). Modelling of individual pharmacokinetics for computer-aided drug dosage. *Computers and Biomedical Research, an International Journal*, 5(5) :411–459.
- Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., Bellomo, R., Bernard, G. R., Chiche, J.-D., Coopersmith, C. M., Hotchkiss, R. S., Levy, M. M., Marshall, J. C., Martin, G. S., Opal, S. M., Rubenfeld, G. D., van der Poll, T., Vincent, J.-L., and Angus, D. C. (2016). The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*, 315(8) :801–810.
- Timsit, J.-F., Fosse, J.-P., Troché, G., De Lassence, A., Alberti, C., Garrouste-Orgeas, M., Bornstain, C., Adrie, C., Cheval, C., Chevret, S., and OUTCOMEREA Study Group, France (2002). Calibration and discrimination by daily Logistic Organ Dysfunction scoring comparatively with daily Sequential Organ Failure Assessment scoring for predicting hospital mortality in critically ill patients. *Critical Care Medicine*, 30(9) :2003–2013.
- Tong-Minh, K., van der Does, Y., van Rosmalen, J., Ramakers, C., Gommers, D., van Gorp, E., Rizopoulos, D., and Endeman, H. (2022). Joint Modeling of Repeated Measurements of Different Biomarkers Predicts Mortality in COVID-19 Patients in the Intensive Care Unit. *Biomarker Insights*, 17 :11772719221112370.

- Traynard, P., Ayrat, G., Twarogowska, M., and Chauvin, J. (2020). Efficient Pharmacokinetic Modeling Workflow With the MonolixSuite : A Case Study of Remifentanyl. *CPT : Pharmacometrics & Systems Pharmacology*, 9(4) :198–210.
- Tsiatis, A. A., Degruittola, V., and Wulfsohn, M. S. (1995). Modeling the Relationship of Survival to Longitudinal Data Measured with Error. Applications to Survival and CD4 Counts in Patients with AIDS. *Journal of the American Statistical Association*, 90(429) :27–37.
- Tu, J., Donovan, L., Lee, D., Wang, J., Austin, P., Alter, D., and Ko, D. (2009). Effectiveness of Public Report Cards for Improving the Quality of Cardiac Care : The EFFECT Study : A Randomized Trial. *JAMA : the journal of the American Medical Association*, 302 :2330–7.
- van Oudenhoven, F. M., Swinkels, S. H. N., Ibrahim, J. G., and Rizopoulos, D. (2020). A marginal estimate for the overall treatment effect on a survival outcome within the joint modeling framework. *Statistics in Medicine*, 39(28) :4120–4132.
- Vedovati, M. C., Barbieri, G., Urbini, C., D’Agostini, E., Vanni, S., Papalini, C., Pucci, G., Cimini, L. A., Valentino, A., Ghiadoni, L., and Becattini, C. (2022). Clinical prediction models in hospitalized patients with COVID-19 : A multicenter cohort study. *Respiratory Medicine*, 202 :106954.
- Vincent, J.-L., de Mendonça, A., Cantraine, F., Moreno, R., Takala, J., Suter, P. M., Sprung, C. L., Colardyn, F., and Blecher, S. (1998). Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units : results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine. *Critical Care Medicine*, 26(11) :1793–1800.
- Vincent, J.-L., Moreno, R., Takala, J., Willatts, S., De Mendonça, A., Bruining, H., Reinhart, C. K., Suter, P. M., and Thijs, L. G. (1996). The SOFA (Sepsis-related Organ

- Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Medicine*, 22(7) :707–710.
- Wei, G. C. G. and Tanner, M. A. (1990). A Monte Carlo Implementation of the EM Algorithm and the Poor Man’s Data Augmentation Algorithms. *Journal of the American Statistical Association*, 85(411) :699–704.
- Wolfinger, R. (1993). Covariance structure selection in general mixed models. *Communications in Statistics - Simulation and Computation*, 22(4) :1079–1106.
- World Health Assembly, . (2017). Improving the prevention, diagnosis and clinical management of sepsis. Technical Report WHA70.7, World Health Organization.
- Wu, L., Liu, W., Yi, G. Y., and Huang, Y. (2011). Analysis of Longitudinal and Survival Data : Joint Modeling, Inference Methods, and Issues. *Journal of Probability and Statistics*, 2012 :e640153.
- Wulfsohn, M. S. and Tsiatis, A. A. (1997). A Joint Model for Survival and Longitudinal Data Measured with Error. *Biometrics*, 53(1) :330–339.
- Zhudenkov, K., Gavrilov, S., Sofronova, A., Stepanov, O., Kudryashova, N., Helmlinger, G., and Peskov, K. (2022). A workflow for the joint modeling of longitudinal and event data in the development of therapeutics : Tools, statistical methods, and diagnostics. *CPT : Pharmacometrics & Systems Pharmacology*, 11(4) :425–437.

# Modélisation conjointe de la dynamique de biomarqueurs et de données de survie à risques compétitifs : développements et applications aux maladies infectieuses sévères

Au cours des dernières années, l'estimation jointe d'un modèle décrivant l'évolution longitudinale d'un biomarqueur et d'un modèle de survie a largement contribué au développement de la médecine personnalisée. Cependant, en présence de risques compétitifs, la plupart des modèles joints incluent un modèle linéaire à effets mixtes pour décrire la dynamique du biomarqueur, souvent peu flexible. Dans un premier travail, nous avons développé un modèle conjoint non-linéaire avec risques compétitifs pour prédire la survie des patients admis en réanimation pour sepsis, avec un suivi journalier du score SOFA. Les risques compétitifs sont décrits par une approche sous-distribution, souvent moins considérée que l'approche cause-spécifique en dépit de ses vertus intéressantes pour la prédiction. Nous avons montré la capacité de l'algorithme SAEM implémenté dans le logiciel Monolix à fournir des estimations non biaisées des paramètres d'un tel modèle. Dans un second travail où plusieurs biomarqueurs longitudinaux étaient disponibles, nous avons développé un modèle joint multivarié avec risques compétitifs et une stratégie pour sélectionner un sous-ensemble de biomarqueurs les plus associés au pronostic des patients hospitalisés pour infection au SARS-CoV-2. L'algorithme SAEM de Monolix et la stratégie de sélection ont démontré de bonnes propriétés dans ce cadre plus complexe. Dans ces deux travaux, nous avons quantifié, par des approches de prédictions dynamiques individuelles, le bénéfice du suivi des biomarqueurs pour améliorer le pronostic des patients. Finalement dans un dernier projet, nous avons souligné que les outils disponibles pour l'estimation des modèles joints non-linéaires étaient principalement disponibles dans des logiciels issus de la pharmacométrie. C'est pourquoi, nous avons étendu le package R *saemix*, initialement dédié aux modèles non linéaires à effets mixtes, au cas des modèles à réponse multiple et joints. Un large spectre de modèles joints peut être estimé, l'utilisateur définissant lui-même la vraisemblance du modèle. Une étude de simulation a démontré de bonnes performances d'estimation des paramètres et de leur incertitude. Ces travaux ouvrent la voie au développement de modèles conjoints sophistiqués et personnalisés.

**Mots-clés :** *Modélisation conjointe, Modèles non-linéaires à effets-mixtes, Analyse de survie, Risques compétitifs, Algorithme SAEM, Prédications dynamiques, Maladies infectieuses sévères*

## Joint modeling of biomarker dynamics and survival with competing risks : developments and application on severe infectious diseases

In recent years, the joint estimation of a model describing the longitudinal evolution of a biomarker and a survival model has greatly contributed to the development of personalized medicine. However, in the presence of competing risks, most joint models include a linear mixed-effects model describing a biomarker dynamic, often few flexible. In a first work, we developed a nonlinear joint model with competing risks to predict the survival of patients admitted in intensive care unit for sepsis, with daily monitoring of the SOFA score. Competing risks are described using a subdistribution approach, often less considered than the cause-specific one despite its interesting predictive properties. We demonstrated the ability of the SAEM algorithm implemented on Monolix software to provide unbiased estimates of the model parameters. In a second work where several longitudinal biomarkers were available, we developed a multivariate joint model with competing risks and a strategy to select a subset of biomarkers most associated with the prognosis of patients hospitalized for SARS-CoV-2 infection. Both the SAEM algorithm implemented on Monolix and the selection strategy demonstrated good properties in this more complex setting. In these two works, we used individual dynamic prediction approaches to quantify the benefits of biomarker monitoring for improving patient prognosis. Finally, in a last project, we pointed out that the tools available for nonlinear joint model estimation were mainly available in pharmacometric software. For this reason, we extended the R package *saemix*, initially designed to fit nonlinear mixed-effects models, to the case of multiple-response and joint models. A wide range of joint models can be estimated, as the user defines himself the likelihood of the model. A simulation study demonstrated good performances of parameter and standard error estimation. This work paves the way for the development of sophisticated and customized joint models.

**Key-words :** *Joint modelling, Nonlinear mixed-effects models, Survival analysis, Competing risks, SAEM algorithm, Dynamic predictions, Infectious diseases*