



HAL
open science

Modélisation conjointe de données longitudinales non-linéaires et de données de survie : Application au cancer de la prostate métastatique

Solène Desmée

► **To cite this version:**

Solène Desmée. Modélisation conjointe de données longitudinales non-linéaires et de données de survie : Application au cancer de la prostate métastatique. Statistiques [stat]. Université Paris Diderot (Paris 7) Sorbonne Paris Cité, 2016. Français. NNT : . tel-01484089

HAL Id: tel-01484089

<https://theses.hal.science/tel-01484089v1>

Submitted on 6 Mar 2017

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.

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

Unité de recherche : UMR 1137 - Infection, Antimicrobiens, Modélisation, Evolution
Équipe : Modélisation biostatistique, pharmacométrie et investigation clinique dans les
maladies infectieuses

DOCTORAT

Spécialité : Biostatistique

Solène DESMÉE

MODÉLISATION CONJOINTE DE DONNÉES LONGITUDINALES
NON-LINÉAIRES ET DE DONNÉES DE SURVIE : APPLICATION AU CANCER
DE LA PROSTATE MÉTASTATIQUE

Thèse dirigée par le Pr. France Mentré et le Dr. Jérémie Guedj

Soutenue le 21 juin 2016

JURY

Madame Virginie RONDEAU	Rapporteur
Monsieur Michel TOD	Rapporteur
Monsieur René BRUNO	Examineur
Monsieur Aurélien LATOUCHE	Examineur
Monsieur Matthieu RESCHE-RIGON	Examineur
Monsieur Bernard SÉBASTIEN	Examineur
Madame France MENTRÉ	Directeur de thèse
Monsieur Jérémie GUEDJ	Co-directeur de thèse

Merci !

Merci à toutes les personnes que j'ai rencontrées au cours de cette thèse et qui m'ont encouragée

À France de m'avoir encadrée durant ces années de thèse et de m'avoir proposé ce sujet qui s'inscrivait dans la continuité de mes expériences passées. Je tiens à te remercier de transmettre tes connaissances scientifiques et ta rigueur et, par ton dynamisme et tes qualités humaines, de permettre à chacun de s'épanouir au sein de l'équipe BIPID et au-delà.

À Jérémie d'avoir accepté de co-encadrer cette thèse. Le partage de tes connaissances en modélisation des maladies infectieuses m'a beaucoup apporté, et ce même si j'étais un "outlier" ! Merci pour nos discussions scientifiques qui ont éveillé mon esprit critique, ta disponibilité et ton soutien au quotidien.

Aux membres du jury. Merci à mes rapporteurs, Virginie Rondeau et Michel Tod, de me faire l'honneur d'évaluer ce travail. Vos remarques constructives ont contribué à l'amélioration de ce manuscrit. Merci à René Bruno, Aurélien Latouche et Matthieu Resche-Rigon d'avoir accepté de participer à mon jury de thèse.

À Bernard. Merci d'être là pour cette étape importante dans ma carrière professionnelle. Je te suis très reconnaissante de m'avoir fait découvrir la modélisation il y a maintenant 5 ans lors de mon stage à Sanofi, de tes encouragements constants depuis et tes conseils bienveillants.

À Christine de m'avoir fait confiance pour mener ces projets et avoir toujours été présente pour moi même depuis que tu es aux US. Je suis ravie que notre collaboration continue.

Aux équipes DSAR et biostatistique de Sanofi et au consortium Biointelligence qui ont financé cette thèse. Merci à l'ensemble des collaborateurs que j'ai rencontrés depuis mon stage en 2011 et qui ont contribué à ce travail. En particulier, je tiens à remercier chaleureusement Claire, Christine, Dorothee, Guilhem, Jean-Baptiste, Jeffrey Barrett, Kimiko, Laurent NGuyen, Marion et Thu d'avoir notamment participé à de nombreux échanges scientifiques. Mais aussi les statisticiens, Antoine, Astrid, Emmanuelle, Franck, Pierre Mancini, Sandrine et Sylvain Nicolas pour votre accueil et vos conseils. Merci également à Michel pour avoir partagé avec moi tes connaissances et pour nos discussions, et à Philippe dont la thèse a été mon premier aperçu de cette aventure.

À l'ensemble de mes collègues de l'unité IAMÉ, passés et présents : Adrien, Andjelika, Anne, Blaise, Camille, Cédric, Charles, Claire, Cyrielle, Dieynaba, Dorothee, Elodie, Emmanuelle C., Emmanuelle P., Florence, François, Giulia, Guillaume, Jean, Jimmy, Josselin, Julie, Julien, Karim, Lidia, Lucille, Marc, Marie, Marie-Karelle, Marie-Paule, Mèril, Minerva, Minh, Pauline, Philippine, Rebecca, Sabrina, Sarah, Sebastian, Simon, Steven, Tiphane, Thu Thuy, Tram, Vincent,... Merci à tous de créer cette bonne ambiance de travail. Merci à François et Hervé pour votre patience informatique. Et un grand merci à Houda pour ton aide précieuse.

À mes anciens collègues du Registre National du Cancer de l'Enfant. Et en particulier à mes collègues de bureau, Claire et Alicia, de ne pas m'avoir découragée de faire une thèse !

À Taraneh pour m'avoir permis de faire mes premiers pas en biostatistique à l'Agence de la biomédecine.

Un grand merci à mes courageux relecteurs !

À mon amie de toujours Cindy et à Vivien. Depuis toutes ces années, c'est toujours un plaisir de vous voir et de passer du temps ensemble aussi naturellement.

À Fiona, Flavie, Marie-Caroline et Émeline. On a chacune pris des chemins différents mais vous retrouver est toujours une bulle d'oxygène.

À mes amis des années lycée et Buck & Co : Clément et Gaëlle (et Estéban !), Jonathan et Marilou, Jules, Julie et Jérôme (et Laureen et Estelle), Laura et Nico, Luc et Mathilde, Mathilde, Morgane et Erwan, Pierre et Fred (et Mahin !), Vincent et Laure, pour votre présence joyeuse et philosophique depuis toutes ces années. Merci à Gaël et Caro, Gwen et Romain, James et Angie, Laurent et Caro et Thérèse pour les soirées d'ici et d'ailleurs.

À Vincent pour tes conseils avisés de docteur, tes encouragements permanents et pour me changer les idées quand il le faut, et merci à Valérie. Merci à Charlotte et Sébastien, que vous êtes beaux avec le petit Raphaël !

À Agathe pour nos indispensables soirées filles !

À la famille Civilise pour votre accueil chaleureux et pour m'avoir fait découvrir la Bretagne, Cologne et Cherbourg.

À Marie-Thérèse et André pour toutes vos pensées positives et bienveillantes. À Martine et Serge pour vos gentilles attentions. J'ai de la chance de vous avoir comme marraine et parrain !

À mes frères et sœurs, Thierry et Christel, Aurore et Yann, toujours présents auprès de moi. Je suis très contente Thierry de t'avoir donné envie de reprendre tes études à 40 ans :-) Merci ma grande sœur d'être toi...

À ma nièce, Camille, et à mes neveux, Elliott, Nathan, Hugo et Bastien, mon filleul, pour les week-ends pleins de vie !

À mes grands-parents qui m'ont tant appris, bien au-delà d'une thèse... Papi, tu auras tout juste sû que je commençais cette aventure, mais je pense que tu aurais été fier d'avoir une petite fille docteur. Mamie Gisèle, merci pour nos bavardages téléphoniques et de rester comme tu es ! Mamie Do, en écrivant ces quelques lignes, mon cœur se serre, 5 saisons que tu me manques... Merci pour tout.

À mes parents, sans qui je n'aurais pu accomplir ce chemin. Merci de m'accompagner depuis toujours de vos petits et grands conseils et de m'avoir soutenue, dans les bons et mauvais moments, même lorsque vous ne compreniez pas tout !

À Alex... Merci pour ton grand soutien pendant ces 3 années, même à distance... Cette soutenance marque enfin le début d'une vie à 2...

Table des matières

1	Introduction	9
1.1	Le cancer de la prostate	9
1.1.1	Définitions	9
1.1.2	Épidémiologie	10
1.1.3	Évolution et traitements	10
	Diagnostic	10
	Stades	11
	Traitements	12
	Docetaxel	14
1.1.4	Biomarqueurs d'évolution sous traitement	14
	Taille de la tumeur	14
	Antigène spécifique de la prostate (PSA) et dérivés	15
1.1.5	VENICE, un essai clinique de phase III	17
	Caractéristiques de l'étude	17
	Description des données	18
1.2	Analyse de données longitudinales	21
1.2.1	Modèles mathématiques de la cinétique du PSA	21
1.2.2	Modèles Non-Linéaires à Effets Mixtes (MNLEM)	22
1.2.3	Méthodes d'estimation	23
1.3	Données de survie	25
1.3.1	Temps d'évènement et données censurées	25
1.3.2	Modèles à risques proportionnels	27
1.4	Modèles conjoints	28
1.4.1	Principe et spécification du modèle dans le contexte des MNLEM	28
1.4.2	Méthodes d'estimation	30
	Approche en deux étapes	30
	Vraisemblance conjointe	30
	Approche "conjointe séquentielle"	32
1.4.3	Tests statistiques	33
1.4.4	Outils diagnostiques	34
	Évaluation du sous-modèle longitudinal	34
	Évaluation du sous-modèle de survie	34
1.4.5	Modèles conjoints et cancer de la prostate	35

1.5	Prédictions individuelles dynamiques	36
1.5.1	Utilisation de la modélisation conjointe	36
1.5.2	Évaluation des modèles prédictifs	38
	Aire sous la courbe ROC	38
	Score de Brier	41
	Prise en compte de la censure : les estimateurs IPCW	42
	Autres métriques	43
1.6	Objectifs de la thèse	43
2	Modèles non-linéaires à effets mixtes et lien avec la survie : Comparaison par simulation des approches conjointes et en deux étapes	45
2.1	Résumé	45
2.2	Article 1 (publié)	47
3	Développement d'un modèle conjoint mécanistique caractérisant la relation entre la cinétique non-linéaire du PSA et la survie dans le cancer de la prostate métastatique	63
3.1	Résumé	63
3.2	Article 2 (publié)	65
4	Prédictions dynamiques individuelles et évaluation des capacités prédictives d'un modèle conjoint non-linéaire dans le cancer de la prostate métastatique	83
4.1	Résumé	83
4.2	Article 3 (en préparation)	85
5	Discussion et conclusion	107
5.1	Discussion	107
5.2	Perspectives	111
5.3	Conclusion générale	114

Productions scientifiques liées à la thèse

Article publié

DESMÉE S., MENTRÉ F., VEYRAT-FOLLET C., GUEDJ J., 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 American Association of Pharmaceutical Scientists Journal*, 2015 ; 17 :691–9. DOI :10.1208/s12248-015-9745-5.

DESMÉE S., MENTRÉ F., VEYRAT-FOLLET C., SÉBASTIEN B., GUEDJ J., Using the SAEM Algorithm for Mechanistic Joint Models Characterizing the Relationship between Nonlinear PSA Kinetics and Survival in Prostate Cancer Patients. *Biometrics*, 2016. DOI : 10.1111/biom.12537.

Article en préparation

DESMÉE S., MENTRÉ F., RIVIÈRE MK., VEYRAT-FOLLET C., GUEDJ J., Dynamic predictive performances of individual survival using nonlinear joint models : application to PSA kinetics in metastatic prostate cancer.

Communications orales

DESMÉE S., MENTRÉ F., VEYRAT-FOLLET C., SÉBASTIEN B. GUEDJ J., Joint modelling for nonlinear longitudinal PSA kinetics and survival data in metastatic prostate cancer patients. *25th Population Approach Group in Europe (PAGE) conference*, Lisbonne, Portugal, Juin 2016.

DESMÉE S., MENTRÉ F., VEYRAT-FOLLET C., GUEDJ J., Joint modelling of PSA kinetics and survival in metastatic prostate cancer patients. *Séminaire hebdomadaire du service de biostatistique et d'épidémiologie de l'Institut Gustave Roussy*, Villejuif, France, Septembre 2015 (invitée).

DESMÉE S., MENTRÉ F., VEYRAT-FOLLET C., GUEDJ J., Joint modelling of nonlinear longitudinal and survival data : simulation study and application in metastatic prostate cancer. *36th Annual Conference of the International Society for Clinical Biostatistics (ISCB)*, Utrecht, Pays-Bas, Août 2015.

DESMÉE S., MENTRÉ F., VEYRAT-FOLLET C., GUEDJ J., Joint modelling of survival and PSA kinetics in metastatic castration-resistant prostate cancer patients. *Journées 2015 du Groupe de Recherche (GDR) Statistique & Santé*, Paris, France, Juin 2015.

DESMÉE S., GUEDJ J., VEYRAT-FOLLET C., MENTRÉ F., Modélisation conjointe de données longitudinales non-linéaires et de survie dans le contexte du cancer de la prostate métastatique et hormono-résistant. 47^{èmes} *Journées de Statistique*, Lille, France, Juin 2015.

Communication affichée

DESMÉE S., GUEDJ J., VEYRAT-FOLLET C., MENTRÉ F., Simulation study of joint model of survival and PSA kinetics using non-linear mixed-effect models and the stochastic approximation expectation-maximization algorithm. 2nd *Workshop of Statistical Analysis of Multi-Outcome Data (SAM)*, Cambridge, Royaume-Unis, Juin 2014.

Récompense scientifique

Prix étudiant Lewis Sheiner, 25th *Population Approach Group in Europe (PAGE) conference*, Lisbonne, Portugal, Juin 2016, <http://www.page-meeting.org/default.asp?abstract=5972>.

Chapitre 1

Introduction

Ce travail de thèse a été initié afin d'étudier des méthodes de modélisation de données longitudinales et de temps d'évènement, très fréquemment collectés dans le milieu médical. En particulier, nous nous sommes intéressés à des modèles conjoints mécanistiques afin de caractériser le lien entre la cinétique non-linéaire d'un biomarqueur et un évènement d'intérêt, dans le contexte d'un essai clinique de phase III sur le cancer de la prostate métastatique.

Dans cette introduction, nous allons donc dans un premier temps présenter le contexte médical et les données disponibles. Puis nous exposerons les méthodes usuelles pour l'analyse de données longitudinales et de temps d'évènement, afin d'introduire la modélisation conjointe de ces données. Le principe des prédictions individuelles dynamiques sera ensuite exposé, ainsi que les méthodes d'évaluation d'un modèle prédictif utilisées dans cette thèse. Les objectifs de la thèse seront présentés à la fin de ce chapitre.

1.1 Le cancer de la prostate

1.1.1 Définitions

La prostate est une glande de l'appareil génital masculin située sous la vessie, en avant du rectum. Sa fonction principale est de sécréter une partie du liquide séminal, un des constituants du sperme, et de le stocker. Pour fonctionner correctement, la prostate a besoin d'androgènes, hormones masculines, dont la plus importante est la testostérone.

Le cancer est caractérisé par une prolifération anormalement importante de cellules. Les cellules ainsi produites forment une tumeur. Pour vivre et continuer à se multiplier, ces cellules annexent les ressources des cellules saines. Elles stimulent notamment la création de nouveaux vaisseaux sanguins à partir de vaisseaux pré-existants, cela s'appelle l'angiogenèse. Au cours de l'évolution de la maladie, certaines cellules tumorales peuvent migrer et former des métastases.

1.1.2 Épidémiologie

Avec 1,1 million d'hommes diagnostiqués en 2012 dans le monde, le cancer de la prostate représente 15% des cancers diagnostiqués chez les hommes et est le 2^{ème} type de cancer le plus fréquent chez les hommes dans le monde, derrière le cancer du poumon (FERLAY et al., 2015). Néanmoins, il existe de fortes disparités en fonction des régions du monde avec des taux d'incidence standardisés sur l'âge variant d'un facteur 25. Ainsi, 70% des cas sont diagnostiqués dans les pays dits développés (FERLAY et al., 2013). Cela est généralement imputé au fait que, dans ces régions, le dépistage par dosage de l'antigène spécifique de la prostate (*Prostate-Specific Antigen* noté PSA) suivi d'une biopsie en cas de résultat suspect est devenu courant (FERLAY et al., 2010).

Avec 307 000 décès estimés en 2012, le cancer de la prostate est la 5^{ème} cause de décès par cancer chez les hommes dans le monde (FERLAY et al., 2015). Les différences géographiques de taux de mortalité sont moins importantes que pour le taux d'incidence, avec des variations d'un facteur 10. Cependant, le taux de mortalité est plus important dans les pays en développement que dans les pays développés.

En France métropolitaine en 2012, le cancer de la prostate est le type de cancer le plus fréquemment diagnostiqué chez les hommes avec 98 nouveaux cas pour 100 000 personnes-années (standardisé sur l'âge dans le monde) et la 3^{ème} cause de décès par cancer chez les hommes avec 10 décès pour 100 000 personnes-années (FERLAY et al., 2013).

1.1.3 Évolution et traitements

Diagnostic

Trois facteurs de risque du cancer de la prostate sont communément admis : l'âge, les antécédents familiaux et l'origine ethnique et géographique (JEMAL et al., 2011; LEITZMANN et ROHRMANN, 2012; HEIDENREICH et al., 2014a).

Le cancer de la prostate n'a souvent aucun symptôme précoce, ce qui rend primordial son dépistage par des tests médicaux afin d'éviter l'évolution de la maladie vers une forme avancée et métastasée et ainsi réduire la mortalité. Il existe trois principaux outils pour diagnostiquer le cancer de la prostate : le toucher rectal, le dosage du PSA et la biopsie (HEIDENREICH et al., 2014a).

Le toucher rectal consiste à palper la prostate à travers la paroi du rectum. Lors de cet examen, la taille de la prostate est évaluée, et un nodule, c'est-à-dire un petit renflement anormal, peut être détecté. Cependant, lorsqu'une anomalie est révélée via cet examen, le cancer de la prostate a généralement déjà métastasé. C'est pourquoi il est recommandé de l'associer à un dosage du PSA (GOSSELAAR et al., 2008).

Chez un homme sain, le PSA est une protéine sécrétée exclusivement par les cellules prostatiques. Sa fonction est d'améliorer la fertilité en liquéfiant le sperme pour faciliter la

circulation des spermatozoïdes. Le PSA est présent naturellement dans le sang de tous les hommes et augmente avec l'âge (LUBOLDT et al., 2007). En cas de cancer de la prostate, les cellules tumorales conservent les mêmes fonctionnalités que les cellules prostatiques et on observe donc une augmentation du PSA dans le sang. Il n'existe pas de réel consensus concernant un seuil de PSA discriminant correctement les individus sains des individus malades (SHARIAT et al., 2011), même si en routine le seuil de 4 ng/mL est souvent utilisé pour détecter les cas suspects, celui de 2.6 ng/mL peut aussi être utilisé (KRUMHOLTZ et al., 2002).

D'autres outils diagnostiques sont étudiés mais ne sont pas utilisés en routine. Lorsque l'on dose le PSA, c'est le PSA total qui est mesuré. Il est en effet la somme de deux composantes : le PSA dit "libre" et le PSA conjugué à des protéines. En cas de cancer, la proportion de PSA libre diminue. Le ratio PSA libre sur PSA total permet alors de distinguer les augmentations du PSA dues à un cancer et donc nécessitant une biopsie, de celles dues à d'autres maladies de la prostate, telle que l'hypertrophie bénigne de la prostate (KOBORI et al., 2008). Récemment a été découvert un gène spécifique au cancer de la prostate, le PCA3, qui peut être mesuré dans l'urine après un massage prostatique (SHAPPELL et al., 2009). Il n'est produit que par les cellules cancéreuses et permet donc de distinguer le cancer de la prostate des autres maladies augmentant le PSA dans le sang. C'est un outil prometteur car non invasif, mais qui nécessite encore d'être validé.

Le diagnostic définitif repose sur les données de biopsies prostatiques transrectales échoguidées. Une biopsie est réalisée si le PSA est élevé et le toucher rectal suspect, en fonction de l'âge du patient et des possibles comorbidités (HEIDENREICH et al., 2014a).

Actuellement, l'efficacité des programmes de dépistage généralisé dans la population est controversée (ILIC et al., 2007). Même s'ils réduisent la mortalité (SCHRÖDER et al., 2012), ils sont également associés à un fort risque de sur-diagnostic (ANDRIOLE et al., 2009; SCHRÖDER et al., 2009). De plus, il a été montré que ces méthodes de dépistage à grande échelle n'étaient pas coût-efficaces, notamment à cause du coût des traitements en excès (HEIJNSDIJK et al., 2009). Les dernières recommandations sont donc de mieux cibler les individus à risque et de tester régulièrement ces personnes (HEIDENREICH et al., 2014a). Ainsi, l'*Association Européenne d'Urologie* suggère de mesurer le PSA vers l'âge de 40-45 ans afin d'avoir une valeur de référence (SCHRÖDER et al., 2009). Ensuite, un dosage tous les 2 à 4 ans chez les hommes dont la valeur de PSA initiale est supérieure à 1 ng/mL est recommandé, tandis qu'un dosage tous les 8 ans chez les autres suffirait (HEIDENREICH et al., 2013). Chez les hommes âgés de plus de 75 ans, avec un PSA de référence inférieur à 3 ng/mL, continuer les dosages n'est pas nécessaire car le risque de décéder d'un cancer de la prostate est alors très faible (morbidité compétitive importante).

Stades

Une fois le cancer diagnostiqué, il est nécessaire de le caractériser. La tumeur est classée d'après les résultats des premiers examens, selon une classification appelée "TNM" pour *Tumor/Nodes/Metastases* (SOBIN et al., 2011). Cette classification est réalisée à l'aide de 3 critères :

- T suivi d'un nombre de 0 à 4 indique la taille de la tumeur primitive ainsi que son invasion, évaluées par imagerie par résonance magnétique (IRM)
 - * T0 : absence de tumeur
 - * T1 : tumeur non palpable au toucher rectal et non visible par imagerie
 - * T2 : tumeur limitée à la prostate, palpable au toucher rectal
 - * T3 : extension de la tumeur au-delà de la capsule prostatique
 - * T4 : extension aux organes adjacents (sphincter, rectum, muscle releveur de l'anus, paroi pelvienne) ou tumeur fixée
- N décrit la propagation dans les ganglions lymphatiques régionaux
 - * N0 : absence de métastase ganglionnaire régionale
 - * N1 : atteinte(s) ganglionnaire(s) régionale(s)
- M indique si le cancer s'est propagé à d'autres parties du corps (métastases à distance)
 - * M0 : absence de métastases à distance
 - * M1 : métastases à distance

Au sein des cancers non métastasés qui peuvent être localisés (T1/T2, N0 et M0) ou localement avancés (T3/T4, N0/N1 et M0), D'AMICO et al. (1998) ont défini 3 groupes de risque (risque faible, intermédiaire et haut) basés sur le taux de PSA, le stade TNM et le score de Gleason qui mesure l'agressivité des cellules cancéreuses d'après la biopsie. Dans le cancer de la prostate, la métastase la plus fréquente est osseuse (observée chez 90% des patients ayant des métastases) (HESS et al., 2006).

Au moment du diagnostic, aux Etats-Unis, environ 80% des malades ont un cancer localisé, 12% un cancer localement avancé et 4% présentent déjà des métastases (SIEGEL et al., 2016). La survie à 5 ans lorsque le cancer est diagnostiqué précocément est très élevée (>99%) alors qu'elle n'est que de 28% lorsque la tumeur a métastasé.

L'évaluation du stade de la maladie est donc primordiale pour proposer au patient le traitement le plus adapté pour lui.

Traitements

Le traitement va dépendre du stade de la maladie, mais également de l'âge du patient et de son état de santé. Le choix du traitement va donc être fait par une équipe médicale pluridisciplinaire et en prenant en compte les préférences du patient.

- Une *surveillance active* est proposée aux patients ayant un très faible risque de

décès par cancer de la prostate (DALL'ERA et al., 2012). Elle consiste à ne pas traiter le patient mais à le suivre régulièrement afin d'initier un traitement curatif en cas de progression de la maladie. Cette méthode a pour but de diminuer les risques de traitement excessif.

- La *prostatectomie radicale* consiste en l'ablation de la prostate et des vésicules séminales par une intervention chirurgicale lourde réalisée sous anesthésie générale. Cette opération a montré un bénéfice en terme de survie par rapport à la surveillance active, en particulier chez les patients de moins de 65 ans et ayant un risque intermédiaire à fort de progression (BILL-AXELSON et al., 2014; HEIDENREICH et al., 2014a). Les effets secondaires principaux sont l'impuissance et l'incontinence.
- Deux alternatives à la prostatectomie sont actuellement en cours d'évaluation : la *cryochirurgie* et les *ultrasons focalisés de haute intensité (HIFU pour High Intensity Focused Ultrasound)* (JÁCOME-PITA et al., 2014). La cryothérapie détruit la tumeur en congelant la zone traitée à l'aide de neige carbonique ou d'azote liquide. Au contraire, le traitement par HIFU détruit la prostate et la tumeur par la chaleur. Ces deux méthodes entraînent l'apoptose des cellules ciblées.
- La *radiothérapie* consiste à envoyer des radiations ciblées sur la prostate. Ces rayons vont créer des lésions dans l'ADN des cellules afin de les détruire. La radiothérapie peut être *externe*, c'est-à-dire que les rayons sont envoyés sur les cellules depuis l'extérieur du corps, ou par implantation de composants radioactifs directement dans la prostate (appelée *curiethérapie* ou *brachythérapie*). Les avantages de ces méthodes sont qu'elles sont indolores et ciblées, donc permettent de préserver autant que possible les tissus sains alentour et ont des effets secondaires moins importants. La radiothérapie est recommandée pour les cancers à faible risque, et en combinaison avec de l'hormonothérapie pour les cancers à risques intermédiaire et fort (MOTTET et al., 2014).
- L'*hormonothérapie* repose sur le caractère hormono-dépendant du cancer de la prostate puisque le développement des cellules prostatiques, leur fonctionnalité et leur multiplication nécessitent de l'androgène. Un traitement empêchant l'activité de l'androgène (et en particulier de la testostérone) entraîne la mort des cellules (appelée apoptose) prostatiques, dont font partie les cellules tumorales. L'hormonothérapie est proposée aux patients non éligibles aux traitements locaux décrits précédemment et aux patients dont le cancer s'étend au-delà de la capsule prostatique et présentant des métastases (SALOMON et al., 2013; HEIDENREICH et al., 2014b). La castration peut être médicamenteuse ou plus rarement chirurgicale.
- La *chimiothérapie* est utilisée pour traiter les patients ayant un cancer de la prostate métastatique non répondeur à l'hormonothérapie (on parle de cancer

hormono-résistant). Son but est de diminuer la croissance tumorale, voire de la stopper, et de soulager les douleurs liées au cancer. Le docetaxel (Taxotere[®]) est la chimiothérapie de référence en première ligne (TANNOCK et al., 2004; BERTHOLD et al., 2008). En cas d'échec avec le docetaxel, une seconde ligne de chimiothérapie peut être proposée, le cabazitaxel (Jevtana[®]) (DE BONO et al., 2010; POUESSEL et al., 2012; LHEUREUX et JOLY, 2012). Ces traitements sont associés à de la prednisone, un anti-inflammatoire qui permet de limiter les effets indésirables.

- Des *soins palliatifs* sont proposés dans les stades avancés du cancer afin d'aider le malade à supporter la douleur due aux métastases osseuses.

Les traitements peuvent être successifs ou combinés. Dans tous les cas, le patient sera ensuite suivi très régulièrement pour évaluer la progression de la maladie.

Docetaxel

Dans cette thèse, nous allons tout particulièrement nous intéresser au cancer de la prostate métastatique hormono-résistant, noté CPmHR (en anglais : mCRPC pour *metastatic Castration-Resistant Prostate Cancer*). Cela correspond à un stade très avancé de la maladie, les patients ayant déjà une longue histoire thérapeutique derrière eux, avec généralement une prostatectomie, puis une hormonothérapie ayant échoué. Dans ce cas, les recommandations européennes sont d'initier une chimiothérapie par docetaxel associé à de la prednisone (HEIDENREICH et al., 2014b).

Le docetaxel appartient à la famille des taxanes. Il est obtenu par transformation chimique d'une molécule naturelle (hémi-synthèse) extraite des aiguilles de l'if européen. Le docetaxel a plusieurs mécanismes d'action (HERBST et KHURI, 2003). Tout d'abord, c'est un agent antimicrotubule. Les microtubules étant impliqués dans la division cellulaire (mitose), le docetaxel inhibe la prolifération des cellules. D'autre part, le docetaxel stimule la mort des cellules (apoptose). Enfin, en inhibant l'angiogenèse (néovascularisation) des cellules tumorales, il empêche un processus vital pour les cellules tumorales.

Mais même avec cette chimiothérapie, docetaxel associé à prednisone, la survie médiane des patients ayant un CPmHR est de 18.9 mois (intervalle de confiance à 95% IC95=[17.0 - 21.2]) (TANNOCK et al., 2004).

1.1.4 Biomarqueurs d'évolution sous traitement

Nous allons nous intéresser aux biomarqueurs utilisés pour suivre l'évolution de la maladie après l'initiation d'un traitement. Dans le cas du CPmHR, le critère principal étant la survie, un biomarqueur sera utile s'il est prédictif de la survie.

Taille de la tumeur

Les marqueurs les plus évidents de l'évolution d'un cancer solide sont liés aux caractéristiques de la tumeur elle-même : volume, poids, nombre de cellules ou densité. Des modèles animaux ont ainsi tout d'abord été développés car il est plus facile d'avoir des données fréquentes et précises (à la fin de l'étude l'animal est euthanasié et sa tumeur disséquée) (SIMEONI et al., 2004; RIBBA et al., 2011).

Chez l'homme, lors des essais cliniques, la taille de la tumeur peut être mesurée régulièrement par des techniques d'imagerie médicale. Lors du premier examen, plusieurs lésions (anomalies du tissu biologique dues à la tumeur) sont sélectionnées pour constituer les "lésions cibles" qui seront tout particulièrement mesurées tout au long du suivi. Il existe alors plusieurs critères, dont les plus connus sont le critère *RECIST* (*Response Evaluation Criteria in Solid Tumors*) qui n'utilise qu'une seule dimension de la taille de la tumeur, en faisant la somme des plus longs diamètres des lésions cibles mesurées sur un nombre limité d'organes, et le critère *WHO* (*World Health Organization*) qui utilise deux dimensions de la tumeur, en faisant la somme des produits des plus longs diamètres perpendiculaires (CHOI et al., 2005). Le critère *RECIST* v 1.1, utilisé dans l'essai clinique VENICE (voir paragraphe 1.1.5), transforme une mesure continue en une variable catégorielle classant la réponse au traitement en 4 classes (THERASSE et al., 2000; EISENHAUER et al., 2009) :

- *Réponse complète* : disparition de toutes les lésions cibles
- *Réponse partielle* : diminution d'au moins 30% de la somme des plus longs diamètres des lésions cibles par rapport à la valeur de référence mesurée à l'initiation du traitement
- *Progression de la maladie* : augmentation d'au moins 20% de la somme des plus longs diamètres des lésions cibles par rapport à la plus petite valeur mesurée depuis l'initiation du traitement ou apparition d'une ou plusieurs nouvelles lésions
- *Maladie stable* : ni réponse partielle, ni progression de la maladie

Mais ces critères sont critiqués pour plusieurs raisons. Tout d'abord, leur reproductibilité inter- et intra- observateur n'est pas évidente, d'autre part, en catégorisant une variable continue, on perd de l'information. Enfin, la dynamique de la croissance tumorale n'est pas prise en compte puisque les mesures sont réalisées à des temps discrets (SHARMA et al., 2012).

Antigène spécifique de la prostate (PSA) et dérivés

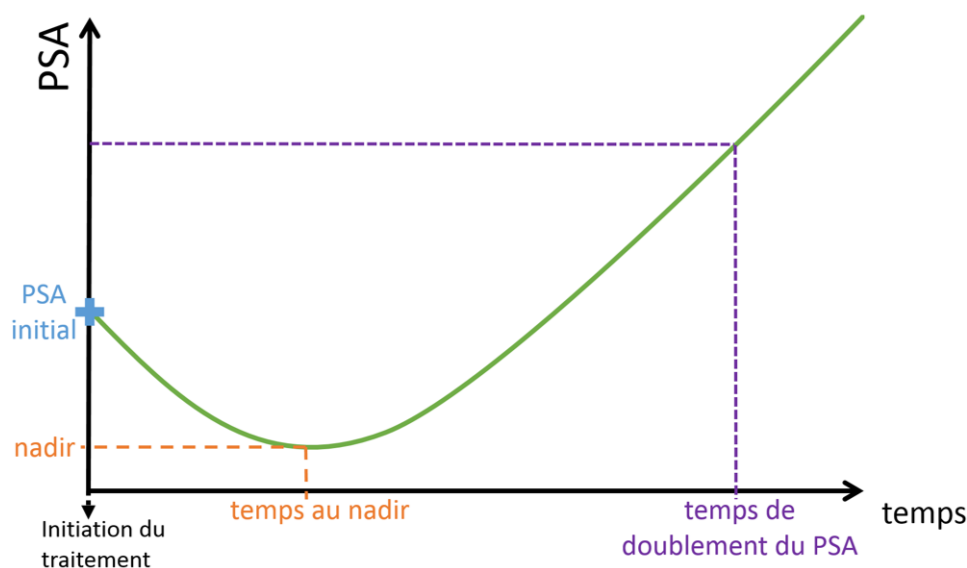
L'inconvénient de l'évaluation de la taille de la tumeur par imagerie médicale est qu'elle est coûteuse et ne peut pas être réalisée fréquemment, contrairement au PSA qui nécessite une simple prise de sang.

Le PSA apparaît comme un "substitut" à la taille de la tumeur puisqu'il est produit par les cellules prostatiques dont les cellules tumorales font partie. S'il est controversé

en tant qu'outil diagnostique, le PSA est un outil pronostique utilisé pour le suivi des patients après un traitement (STEPHENSON et al., 2006; ROACH et al., 2006).

Après initiation d'une chimiothérapie chez un patient ayant un CPmHR, on s'attend à observer une diminution du PSA, voire à atteindre un seuil indétectable (de 0.1 ng/mL). Cependant, les patients développent généralement une résistance au traitement (SERUGA et al., 2011) ce qui entraîne dans un second temps un rebond du taux de PSA comme illustré par la figure 1.1.

FIGURE 1.1 – Evolution naturelle du PSA après mise sous traitement (radiothérapie, hormonothérapie ou chimiothérapie) chez un patient qui rechute



Néanmoins, il n'existe pas de consensus sur la pertinence d'une valeur au dessus de laquelle le patient pourrait être considéré comme ayant un cancer qui progresse (STEPHENSON et al., 2006). C'est pourquoi des marqueurs dérivés de la cinétique du PSA ont été proposés. Parmi eux :

- Le *temps de doublement du PSA* (noté PSA-DT pour *PSA doubling time*) est le temps écoulé entre l'initiation du traitement et le moment où le taux de PSA initial a doublé (figure 1.1). Il est souvent utilisé comme facteur pronostique du risque de décès par cancer après radiothérapie (FITZPATRICK et al., 2009; ARMSTRONG et al., 2012).
- Le *nadir* est la plus petite valeur de PSA mesurée après l'initiation du traitement (figure 1.1). Le *temps au nadir* (*time-to-nadir*) est le temps écoulé entre l'initiation du traitement et le moment où le nadir est atteint. Si ces deux valeurs sont utilisées pour évaluer l'efficacité des traitements dans le cas de cancers moins avancés, elles n'ont pas montré d'association avec le risque de décès dans le cas d'une maladie avancée (LILJA et al., 2008).

- Chez les patients ayant un PSA initial supérieur à 10 ng/mL, une *réponse PSA* est définie comme une diminution de 50% du PSA. Il a été établi qu'un traitement qui n'entraînait pas de diminution du PSA n'était pas efficace, mais qu'une diminution du PSA ne garantit pas forcément l'efficacité du traitement (FLEMING et al., 2006; PETRYLAK et al., 2006).
- Une *progression PSA* est définie, chez les patients non-répondeurs PSA, comme une augmentation de 25% du PSA par rapport au PSA à l'initiation du traitement, et chez les patients répondeurs PSA, comme une augmentation d'au moins 25% du PSA par rapport au nadir (à condition que ces augmentations soient d'au moins 2 ng/mL) (SCHER et al., 2008).

Ces variables présentent toutes l'inconvénient de ne pas toujours être observées puisque le PSA est mesuré à des temps discrets, rendant leur méthode d'évaluation sujette à caution, et surtout elles ne prennent pas en compte toute la cinétique du PSA. Elles ne reflètent donc qu'une petite fraction de l'information collectée tout au long du suivi du patient.

1.1.5 VENICE, un essai clinique de phase III

Caractéristiques de l'étude

Le cancer de la prostate étant un problème de santé publique majeur et touchant les hommes âgés, de plus en plus nombreux avec le vieillissement de la population des pays dits développés, la recherche de nouveaux traitements est un défi pour l'industrie pharmaceutique. Dans cette thèse, nous avons analysé des données provenant d'une étude clinique de phase III menée par SANOFI et appelée VENICE (TANNOCK et al., 2013).

Cette étude multicentrique, randomisée en double aveugle avec un bras contrôle, examinait l'utilisation en première ligne de traitement du CPmHR de la combinaison aflibercept et docetaxel avec prednisone. L'objectif principal était de démontrer l'efficacité, en plus du traitement de référence docetaxel associé à prednisone, d'aflibercept versus "*placebo*" en terme de survie globale chez des patients ayant un CPmHR.

Les objectifs secondaires étaient d'évaluer l'efficacité d'aflibercept comparé au *placebo* en terme de réponse PSA, réaction à la douleur, temps de survenue d'évènements osseux, survie sans progression de la maladie (i.e. sans progression tumorale selon RECIST, sans progression PSA, sans progression de la douleur et sans évènement osseux), réponse tumorale (en utilisant RECIST), survie sans progression PSA, survie sans progression de la douleur et qualité de vie, et d'évaluer la sûreté des deux bras de traitement.

Pour ce faire, 1224 patients ayant un CPmHR et n'ayant jamais été traités par chimiothérapie ont été randomisés entre août 2007 et février 2010, menant à 612 patients par bras de traitement en intention de traiter. Les patients du bras aflibercept devaient recevoir de l'aflibercept 6 mg/kg par voie intra-veineuse toutes les 3 semaines, en plus du

docetaxel 75 mg/m² par voie intra-veineuse toutes les 3 semaines en combinaison avec de la prednisone 5 mg pris par voie orale deux fois par jour. Ceux du bras *placebo* recevaient un placebo à la place de l'affibercept par voie intra-veineuse toutes les 3 semaines, en plus du docetaxel 75 mg/m² par voie intra-veineuse toutes les 3 semaines en combinaison avec de la prednisone 5 mg pris par voie orale deux fois par jour. Huit patients du bras *placebo* n'ont finalement pas été traités contre 7 dans le bras affibercept. De plus, 6 patients inclus dans le bras *placebo* reçurent au moins une dose d'affibercept par erreur (TANNOCK et al., 2013).

Comme décrit dans le protocole, chaque patient est traité toutes les 3 semaines jusqu'à progression de la maladie (i.e. progression tumorale selon RECIST, progression PSA, progression de la douleur ou évènement osseux), toxicité intolérable ou demande du patient. Les résultats des examens cliniques, des tests en laboratoire et les effets indésirables sont collectés avant l'initiation du traitement, durant tous les cycles de chimiothérapie et jusqu'à 30 jours après la dernière administration. En particulier, le PSA doit être mesuré juste avant l'initiation du traitement, puis toutes les 3 semaines (juste avant administration de la chimiothérapie), à la fin du traitement, puis toutes les 12 semaines jusqu'au décès du patient ou son retrait de l'étude. Les tailles de tumeur sont également documentées selon le critère RECIST (somme des plus longs diamètres des lésions cibles) à l'initiation du traitement, puis toutes les 12 semaines jusqu'au décès du patient ou son retrait de l'étude.

Finalement, cette étude n'a pas montré d'amélioration de la survie globale chez les hommes ayant un CPmHR et traités en première ligne de chimiothérapie par affibercept en combinaison avec docetaxel et prednisone versus docetaxel seul et prednisone (rapport des risques de 0.94 avec IC95%=[0.82-1.08]). De plus, une augmentation de la toxicité a été observée dans le bras affibercept (TANNOCK et al., 2013). Le docetaxel est donc resté le traitement de référence en première ligne du CPmHR.

Description des données

Dans cette thèse, nous nous sommes intéressés aux 598 patients du bras *placebo* qui recevaient du docetaxel 75 mg/m² par voie intra-veineuse toutes les 3 semaines en combinaison avec de la prednisone 5 mg pris par voie orale deux fois par jour.

Au final, nos analyses se concentreront sur les 596 patients ayant un CRmHR traités par la chimiothérapie de référence, docetaxel et prednisone, et ayant au moins une mesure de PSA au cours de l'essai clinique (2 patients n'avaient aucune mesure de PSA). Ces patients ont été divisés par tirage aléatoire en deux jeux de données :

- un jeu de données dit "*d'apprentissage*" de $N = 400$ patients ("*training dataset*")
- un jeu de données dit "*de validation*" comportant les $N' = 196$ patients restants ("*validation dataset*")

La table 1.1 décrit ces 596 patients. Les deux jeux de données ainsi constitués sont

TABLE 1.1 – Caractéristiques des patients et des données disponibles des jeux de données d’apprentissage et de validation

	Jeu de données d’apprentissage N=400	Jeu de données de validation N'=196
Age médian (min-max) (années)	68 (41-87)	68 (40-84)
BSA médian (min-max) (m ²)	2.0 (1.4-2.6)	2.0 (1.4-2.5)
Taille médiane (min-max) de la tumeur initiale* (mm)	35 (11-239)	38 (10-189)
Nombre total de mesures de PSA (chez # patients)	6627 (400)	3185 (196)
Nombre total de mesures de PSA pré-traitement (chez # patients)	1385 (400)	690 (196)
Nombre total de mesures de PSA sous traitement (chez # patients)	3934 (391)	1893 (193)
Nombre total de mesures de PSA après traitement (chez # patients)	1308 (295)	602 (151)
Nombre médian (min-max) de mesures de PSA par patient	15 (3-57)	15 (5-59)
Nombre médian (min-max) de mesures de PSA pré-traitement par patient	4 (1-4)	4 (1-4)
Nombre médian (min-max) de mesures de PSA sous traitement par patient	9 (0-48)	9 (0-50)
Nombre médian (min-max) de mesures de PSA après traitement par patient	2 (0-26)	2 (0-21)
PSA médian (min-max) à la 1 ^{ère} mesure avant traitement (ng/mL)	22.7 (0.1**-1354.0)	37.6 (0.1**-1197.0)
PSA médian (min-max) à l’initiation du traitement (ng/mL)	80.1 (0.1**-3821.0)	126.6 (0.1**-2383.0)
Temps médian (min-max) écoulé entre la 1 ^{ère} valeur de PSA et l’initiation du traitement (jours)	104 (2-1195)	101 (0-357)
Estimateur de Kaplan-Meier de la survie médiane [IC95%] (jours)	656 [598-741]	598 [547-732]

* chez les patients dont la taille de la tumeur est mesurable à l’initiation du traitement, soit 236 et 118 patients dans les jeux de données d’apprentissage et de validation respectivement

** limite de quantification du PSA = 0.1 ng/mL

très comparables en termes d’âge, de surface corporelle (notée *BSA* pour *body surface area*) et de taille initiale de la tumeur. Les mesures de PSA sont très nombreuses, et en particulier, on dispose de mesures sous traitement, mais également avant le traitement (du fait que le patient doit avoir une maladie qui progresse documentée) et souvent après traitement. De ce fait, nous distinguons la 1^{ère} valeur de PSA mesurée et la valeur du PSA à l’initiation du traitement. Comme la maladie progresse, le PSA augmente avant l’initiation du traitement (figure 1.2). On observe bien la cinétique biphasique typique de

patient initiant un traitement : le PSA va dans un premier temps diminuer, puis rebondir, entraînant l'arrêt de la chimiothérapie puisque l'un des critères d'arrêt du traitement est la progression PSA. Le PSA peut atteindre des valeurs très importantes (le maximum mesuré au cours du suivi d'un patient était 11 000 ng/mL), signe d'une maladie très avancée. Vu cette forte variabilité des valeurs de PSA, il est généralement log-transformé : nous travaillerons avec $\log(\text{PSA}+1)$ afin d'assurer une distribution normale et des valeurs positives.

La survie est estimée à partir de l'initiation du traitement. La figure 1.3 montre la courbe de Kaplan-Meier (voir définition dans le paragraphe 1.3.1) pour les 400 patients du jeu de données d'apprentissage dont la survie médiane est de 656 jours (IC95%=[598-741]), soit environ 22 mois après l'initiation de la chimiothérapie. C'est plus que dans l'étude de TANNOCK et al. (2004) qui a établi l'efficacité du docetaxel et qui montrait une survie médiane de 19 mois. Cela semble dû au fait que les hommes inclus dans VENICE initiaient une chimiothérapie plus précocement (TANNOCK et al., 2013). La figure 1.3 montre également que la survie dans le jeu de données de validation, dont le PSA médian initial est plus élevé que dans le jeu de données d'apprentissage, évolue très similairement, avec tout de même une survie médiane légèrement plus faible (598 jours avec un IC95%=[547-732]).

FIGURE 1.2 – Evolution du PSA chez 60 patients tirés au sort parmi les patients du jeu de données d'apprentissage

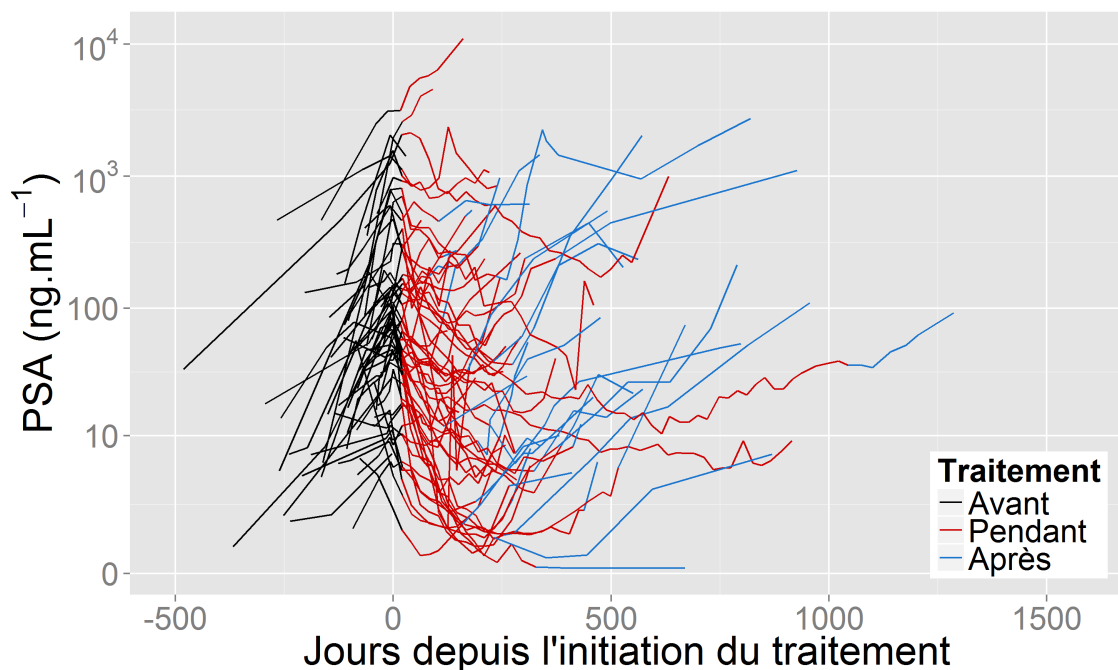
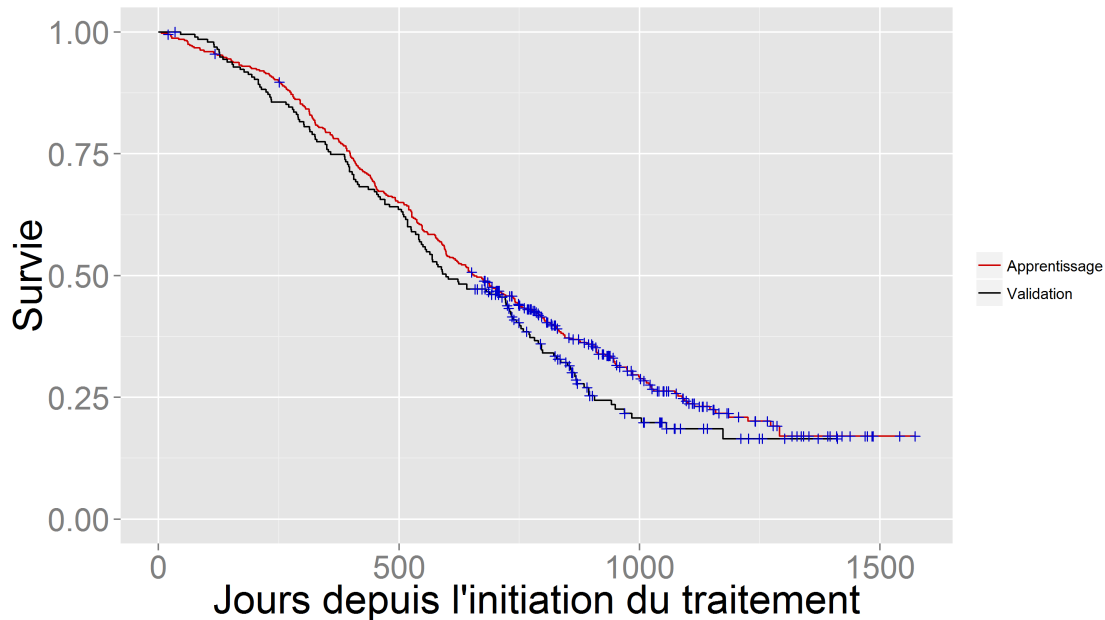


FIGURE 1.3 – Courbes de Kaplan-Meier chez les 400 patients du jeu de données d’apprentissage et chez les 196 patients du jeu de données de validation. Les croix bleues correspondent aux patients censurés.



1.2 Analyse de données longitudinales

Au cours d’un essai clinique, de nombreuses mesures sont récoltées de manière répétée chez tous les patients. Pour exploiter toute l’information fournie par ces données, il faut tenir compte des corrélations entre les mesures réalisées chez un même patient, de l’erreur de mesure (qu’elle soit due à l’appareil de mesure ou aux variations biologiques intrapatient) et du fait qu’on analyse tous les patients simultanément. Les modèles (non-linéaires) à effets mixtes (MNLEM) sont adaptés à l’analyse simultanée des données longitudinales de l’ensemble des sujets.

Dans cette thèse, nous nous sommes intéressés à l’évolution (ou cinétique) du PSA chez des patients ayant un CPmHR et initiant une chimiothérapie.

Tout d’abord, il faut définir le modèle mathématique sous-jacent, appelé modèle structurel, décrivant l’évolution du biomarqueur au cours du temps (section 1.2.1), puis le modèle statistique utilisé, le MNLEM (section 1.2.2). Enfin, nous présenterons les différentes méthodes d’estimation disponibles pour procéder à ce type d’analyses (section 1.2.3).

1.2.1 Modèles mathématiques de la cinétique du PSA

Le modèle le plus courant de cinétique du PSA à l’initiation d’un traitement est un modèle biexponentiel (KAPLAN et al., 1991; ZAGARS et POLLACK, 1997; STEIN et al.,

2008; SUBTIL et RABILLOUD, 2010) s'écrivant de manière simple :

$$PSA(t) = A_1 \exp(-d \times t) + A_2 \exp(c \times t)$$

où d , le taux de décroissance, quantifie la phase de décroissance au début du traitement, tandis que c , le taux de croissance, amorce le rebond du PSA correspondant à l'apparition d'une résistance au traitement. PROUST-LIMA et al. (2008) proposent également un modèle linéaire de l'évolution de $\log(PSA + 0.1)$ décomposé en 3 parties : le niveau de PSA après radiothérapie, l'évolution à court terme dans l'année suivant le traitement et l'évolution à long-terme. Ce modèle capture convenablement la non-linéarité biphasique de la cinétique du PSA. Au vu de la cinétique du PSA, ces modèles mathématiques vont correctement ajuster les données longitudinales de PSA. Mais un inconvénient est que les paramètres ne permettent pas une interprétation physiologique. Ce sont des modèles descriptifs, utiles dans un contexte épidémiologique en particulier. Il a ainsi été montré qu'un modèle biexponentiel décrivait moins bien des données après prostatectomie qu'un modèle utilisant un paramètre biologique de l'évolution du PSA, la clairance (vitesse d'élimination du PSA dans un volume de sang par unité de temps) (YOU et al., 2008).

La cinétique du PSA peut également être décrite par un modèle reposant sur les processus biologiques sous-jacents décrits par des équations différentielles (KRONIK et al., 2010; DE CHARRY et al., 2015). Ces modèles sont appelés *modèles mécanistiques*. Le système d'équations différentielles décrit les dynamiques de production et d'élimination de plusieurs quantités biologiques, considérées comme des compartiments, et les corrélations entre ces compartiments. Ils peuvent ainsi faciliter l'utilisation simultanée de plusieurs biomarqueurs de l'évolution de la maladie. Les modèles mécanistiques permettent de prendre en compte le vrai mécanisme de production du PSA ainsi que les effets du traitement, rendant possible la simulation de différentes prises en charge thérapeutiques dans le développement clinique de nouvelles molécules.

Ce type de modèle est très utilisé dans les maladies infectieuses avec des modèles de cinétique virale (NGUYEN et GUEDJ, 2015) ou bactérienne (NGUYEN et al., 2014) mais plus rarement en oncologie où les processus biologiques sont complexes et encore mal connus. Notons tout de même le développement de modèles de la croissance tumorale, appelée "*tumor growth model*", pour lesquels RIBBA et al. (2014b) ont proposé une librairie de modèles.

À notre connaissance, aucun modèle mécanistique de la cinétique du PSA n'a été proposé dans le cadre du CPmHR.

Ces modèles mécanistiques reposant sur des équations différentielles complexes, l'évolution du PSA devra alors être considérée comme un processus non-linéaire.

1.2.2 Modèles Non-Linéaires à Effets Mixtes (MNLEM)

Une fois le modèle structurel défini, l'analyse des données peut être réalisée à l'aide des MNLEM. Cette méthode permet d'estimer les caractéristiques de population, c'est-à-dire la moyenne et la variabilité entre les sujets (appelée variabilité inter-sujet), et les paramètres de chaque sujet. Un avantage de cette méthode est qu'elle peut être utilisée dans le cas de données individuelles éparées.

On note N le nombre de sujets dans l'étude. Pour un individu i ($i = 1, \dots, N$), la $j^{\text{ème}}$ observation du biomarqueur mesurée au temps t_{ij} et notée y_{ij} , est supposée obéir au modèle statistique suivant :

$$y_{ij} = b(t_{ij}, \psi_i, w_i) + g(t_{ij}, \psi_i, w_i, \Sigma)e_{ij}$$

avec les notations :

- i : individu disposant de n_i observations du biomarqueur ($j = 1, \dots, n_i$) aux temps t_{ij} . On notera $y_i = \{y_{i1}, \dots, y_{in_i}\}$ le vecteur des n_i observations du sujet i .
- b : modèle structurel non-linéaire choisi pour décrire l'évolution du biomarqueur, commun à tous les sujets.
- ψ_i : vecteur des paramètres individuels du sujet i .
- w_i : vecteur des covariables du sujet i .
- g : fonction représentant la forme du modèle d'erreur.
- e_{ij} : erreurs résiduelles, supposées indépendantes et identiquement distribuées selon une distribution gaussienne centrée-réduite ($e_{ij} \sim \mathcal{N}(0, 1)$).
- Σ : paramètres expliquant la variabilité des erreurs résiduelles.

Les paramètres individuels peuvent être décomposés en effets fixes, notés μ et représentant un effet médian "typique" de la population, et en effets aléatoires, notés η_i , spécifiques de chaque sujet et caractérisant leur écart par rapport à ce comportement typique. La variabilité inter-sujet peut en partie être expliquée par les covariables de l'individu w_i (par exemple l'âge, le sexe, le poids, le génotype ou une co-médication). Soit τ la transformation de la $q^{\text{ème}}$ composante du vecteur des paramètres individuels, $\psi_{q,i}$, alors l'effet de la covariable peut être inclus dans le modèle comme suit :

$$\tau(\psi_{q,i}) = \tau(\mu_q) + \eta_{q,i} + \nu'w_i \text{ où } \nu \text{ est le vecteur d'effet des covariables.}$$

Par exemple, la distribution de la $q^{\text{ème}}$ composante du vecteur des paramètres individuels est souvent supposée normale (τ est la fonction identité) : $\psi_{q,i} = \mu_q + \eta_{q,i} + \nu'w_i$ ou log-normale (τ est la fonction logarithme) : $\log(\psi_{q,i}) = \log(\mu_q) + \eta_{q,i} + \nu'w_i$. Cette seconde transformation est souvent utilisée car elle contraint les paramètres biologiques à être positifs. La distribution peut également être supposée logit-normale (τ est alors la fonction logit) : $\text{logit}(\psi_{q,i}) = \text{logit}(\mu_q) + \eta_{q,i} + \nu'w_i$, afin de considérer des paramètres individuels contraints entre 0 et 1. Nous supposons généralement que $\eta_{ij} \sim \mathcal{N}(0, \Omega)$ avec Ω la matrice de variance-covariance des effets aléatoires. Il est également supposé que les $e_i|\eta_i$ sont

indépendants d'un sujet à l'autre, mais aussi que pour un même sujet e_i et η_i sont indépendants.

Un modèle d'erreur combiné est souvent utilisé : $g(t_{ij}, \psi_i, w_i, \Sigma) = \sigma + \sigma' \times b(t_{ij}, \psi_i, w_i)$ où $\Sigma = \{\sigma, \sigma'\}$ sont les paramètres à estimer. Deux cas particuliers sont fréquemment rencontrés, celui d'une erreur constante (ou homoscédastique) de variance $g^2 = \sigma^2$ ($\sigma' = 0$) et celui d'une erreur proportionnelle à la valeur prédite par le modèle, de coefficient de variation σ' ($\sigma = 0$).

Le vecteur des paramètres de population à estimer, noté θ_l , contient donc : les effets fixes μ , le vecteur d'effet des covariables ν si le modèle en contient, les éléments de la matrice de variance-covariance des effets aléatoires Ω et les paramètres Σ expliquant la variabilité des erreurs résiduelles.

Suite à l'estimation des paramètres de population, les MNLEM permettent l'estimation des effets aléatoires $\hat{\eta}_i$ et donc des paramètres individuels $\hat{\psi}_i$. Des estimations bayésiennes empiriques (EBEs) des paramètres individuels peuvent être obtenues à partir de la distribution *a priori* des paramètres et de la vraisemblance *a posteriori* pour ce sujet.

1.2.3 Méthodes d'estimation

L'estimation des paramètres de population dans les MNLEM nécessite des méthodes d'estimation complexes. La plupart des méthodes d'estimation utilisées pour les MNLEM reposent sur le maximum de vraisemblance (DARTOIS et al., 2007). Cependant, l'intégrale dans la fonction de vraisemblance n'ayant pas une forme analytique du fait de la non-linéarité du modèle structurel en ses paramètres, les MNLEM requièrent des algorithmes complexes.

Les premières méthodes développées proposent de faire une approximation de la vraisemblance reposant sur un développement de Taylor d'ordre 1 du modèle structurel b afin de linéariser le modèle et permettre un calcul explicite de la vraisemblance. La méthode FO (*pour First-Order*, SHEINER et al. (1972)) linéarise le modèle autour de l'espérance des effets aléatoires (i.e. 0) tandis que la méthode FOCE (*pour First-Order Conditional Estimate*, LINDSTROM et BATES (1990)) le linéarise autour des estimations individuelles des effets aléatoires (i.e. $\hat{\eta}_i$). Pour estimer les paramètres de population, on peut alors utiliser un algorithme itératif tel que l'algorithme de minimisation de Newton-Raphson. Mais ces algorithmes ont tendance à produire des estimateurs biaisés et peu précis lorsque le nombre d'observations par sujet croît moins vite que le nombre de sujets (VONESH, 1996) ou lorsque la variance des effets aléatoires est grande (GE et al., 2004). Par la suite, une méthode d'estimation où la vraisemblance est calculée par une approximation de Laplace a été proposée (WOLFINGER, 1993). Ces méthodes sont toutes implémentées dans le logiciel NONMEM (SHEINER et BEAL, 1998) largement utilisé en pharmacométrie dans l'industrie pharmaceutique. La méthode FO et l'approximation de

Laplace sont disponibles dans la `proc NLMIXED` du logiciel SAS tandis que l'algorithme FOCE et l'approximation de Laplace sont proposés dans les packages `nlme` et `lme4` du logiciel R (PINHEIRO et BATES, 2000).

Le problème majeur de ces méthodes réside dans leur principe même : elles maximisent une vraisemblance approchée, et non la vraisemblance exacte. L'amélioration des capacités informatiques a permis de considérer des méthodes d'estimation dites "exactes", plus précises mais aussi plus intensives en temps de calcul. Deux approches peuvent être utilisées afin de calculer exactement la vraisemblance : l'intégration numérique et l'intégration stochastique.

La quadrature de Gauss consiste en une approximation de la valeur numérique de l'intégrale par une somme pondérée prise en un certain nombre de points du domaine d'intégration (GOLUB et WELSCH, 1969). Les noeuds et les poids sont déterminés afin de mieux couvrir la distribution des effets aléatoires. Cette méthode nécessite un nombre suffisamment élevé de noeuds de quadrature, ce qui implique une convergence lente et un nombre limité d'effets aléatoires dans le MNLEM. Une version améliorée de cette méthode est la quadrature de Gauss adaptative (PINHEIRO et BATES, 1995) dont les noeuds ne sont pas centrés sur l'espérance des effets aléatoires mais sur les modes conditionnels estimés de ces effets aléatoires. La quadrature de Gauss adaptative requiert ainsi moins de noeuds que la quadrature de Gauss standard pour une estimation plus précise et moins biaisée des paramètres. L'approximation de Laplace est un cas particulier de la quadrature de Gauss adaptative avec un seul noeud. La quadrature de Gauss adaptative est implémentée dans la `proc NLMIXED` de SAS.

Plus récemment, l'algorithme *Stochastic Approximation Expectation Maximisation* (SAEM) a été développé par DELYON et al. (1999) puis KUHN et LAVIELLE (2005) ont proposé une version de cet algorithme couplé à une procédure de chaînes de Markov de Monte Carlo (MCMC). Cette méthode est une version stochastique de l'algorithme EM (Expectation Maximisation) où les effets aléatoires individuels sont considérés comme des valeurs manquantes (DEMPSTER et al., 1977). Les algorithmes EM sont un processus itératif où chaque itération est divisée en une étape de calcul de l'espérance de la vraisemblance complète conditionnellement aux observations (étape E), et une étape de maximisation de cette quantité (étape M). Dans le cas de MNLEM, l'espérance de la vraisemblance ne peut pas être calculée. L'algorithme SAEM va alors diviser l'étape E en deux étapes : une simulation des effets aléatoires individuels en utilisant une méthode MCMC (étape S), puis le calcul de l'espérance de la vraisemblance grâce à une approximation stochastique (étape A). L'algorithme SAEM a été implémenté dans le logiciel Monolix (<http://www.lixoft.eu>), puis dans le logiciel NONMEM depuis la version 7 et dans le package R `saemix` (www.saemix.biostat.fr, COMETS et al. (2011)).

PLAN et al. (2012) ont récemment comparé les performances des principaux algorithmes d'estimation des paramètres d'un MNLEM dans le contexte de modèles dose-

réponse simples. Les conclusions de cette étude sont que les approches FOCE et Laplace sont rapides et robustes, l'algorithme SAEM de Monolix donne de bons résultats mais est sensible aux paramètres initiaux fournis, enfin la quadrature de Gauss adaptative donne des estimations non biaisées dans les différents scénarios envisagés mais nécessite les plus longs temps de calcul.

Au cours de cette thèse, nous avons utilisé l'algorithme SAEM implémenté dans Monolix.

1.3 Données de survie

1.3.1 Temps d'évènement et données censurées

L'analyse de données de survie est l'étude du temps écoulé jusqu'à la survenue d'un évènement d'intérêt. Dans notre contexte de cancer très avancé, le critère principal d'efficacité d'un traitement est la survie globale, c'est-à-dire le temps écoulé entre l'initiation du traitement et le décès pour toutes causes confondues.

Le temps de survie X est une variable aléatoire positive et continue qui désigne le temps écoulé entre une origine (ici l'initiation du traitement) et la survenue de l'évènement d'intérêt (ici le décès). La distribution du temps de survie peut être décrite par plusieurs fonctions de façon équivalente (KLEIN et MOESCHBERGER, 2005) :

- La fonction de survie est la probabilité de survivre au moins jusqu'au temps $t \geq 0$:

$$S(t) = \mathbb{P}(X > t)$$

C'est une fonction décroissante qui vaut 1 au temps d'origine et tend vers 0 quand t tend vers l'infini.

- La fonction de répartition est la probabilité de décéder avant le temps $t \geq 0$:
 $F(t) = \mathbb{P}(X \leq t) = 1 - S(t)$.
- La densité de probabilité f est une fonction positive définie telle que pour tout $t \geq 0$, $F(t) = \int_0^t f(u)du$. Si F est dérivable au temps t alors $f(t) = \frac{dF(t)}{dt}$. Elle peut être vue comme le risque de mourir dans un petit intervalle de temps après t .
- Le risque instantané, noté h (pour *hazard function*), est la probabilité de décéder dans un petit intervalle de temps après t , conditionnellement au fait d'avoir survécu jusqu'au temps t :

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{P}(t \leq X < t + \Delta t | X \geq t)}{\Delta t} = \frac{f(t)}{S(t)} = -\frac{d \ln(S(t))}{dt}$$

- Le risque instantané cumulé est l'intégrale du risque instantané :

$$H(t) = \int_0^t h(u)du = -\ln(S(t))$$

On en déduit également la relation : $S(t) = \exp(-H(t)) = \exp(-\int_0^t h(u)du)$.

Les données de survie ont comme caractéristique que l'évènement d'intérêt ne peut pas toujours être observé, on dit alors que la durée de vie est censurée. On se placera dans le cas d'une censure à droite qui peut être due à un patient considéré comme "perdu de vue" ou au fait que l'étude s'arrête avant qu'on ait observé l'évènement chez certains patients. Chez ces patients censurés, on dispose quand même d'une information dont on pourra tenir compte : ils ont survécu au moins jusqu'à la date de dernière visite.

On note N le nombre de sujets dans l'étude. Pour un patient $i \in \{1, \dots, N\}$, X_i désigne son temps de survie et C_i son temps de censure. Les données de survie observées sont alors (T_i, δ_i) avec $T_i = \min(X_i, C_i)$ et $\delta_i = \mathbf{1}_{\{X_i \leq C_i\}}$ l'indicatrice du statut vital : $\delta_i = 1$ si le patient i a eu l'évènement au temps T_i (i.e. s'il est décédé au temps T_i) et $\delta_i = 0$ s'il a été censuré (i.e. s'il est vivant à T_i). Si on fait l'hypothèse que la censure est non informative, la contribution du sujet i à la vraisemblance peut s'écrire :

$$L_i(\theta) = h(T_i|\theta)^{\delta_i} S(T_i|\theta) \tag{1.1}$$

La fonction de survie peut être estimée par un estimateur non paramétrique, l'estimateur de Kaplan-Meier (KAPLAN et MEIER, 1958). Soit Y_i le nombre d'individus à risque au temps T_i (i.e. les sujets vivants et non censurés à ce temps) et d_i le nombre d'évènements au temps T_i . L'estimateur de Kaplan-Meier s'écrit alors :

$$\hat{S}_{KM}(t) = \prod_{i; T_i \leq t} \left(1 - \frac{d_i}{Y_i}\right)$$

$\hat{S}_{KM}(t)$ est une fonction en escalier décroissante, continue à droite (voir par exemple la figure 1.3). Pour construire un intervalle de confiance, la variance de l'estimateur de Kaplan-Meier peut être estimée en utilisant l'estimateur de Greenwood qui s'écrit :

$$\widehat{Var}(\hat{S}_{KM}(t)) = \hat{S}_{KM}(t)^2 \sum_{i; T_i \leq t} \frac{d_i}{Y_i(Y_i - d_i)}$$

Il est également possible de définir des modèles de survie paramétriques. Pour cela, on spécifie généralement la forme du risque instantané : constant (loi exponentielle), monotone croissant ou décroissant (loi de Weibull ou loi Gamma) et en forme de \cup ou de \cap (loi de Weibull généralisée, log-logistique ou log-normale). Dans ce document, nous utiliserons :

- Le risque instantané constant, défini par une loi exponentielle qui ne dépend que d'un paramètre $\lambda > 0$: $h(t) = \frac{1}{\lambda}$.
- Un risque instantané monotone, défini par une loi de Weibull qui dépend de deux paramètres : $\lambda > 0$, le paramètre d'échelle (*scale*) et k , le paramètre de forme de la distribution (*shape*). Le risque instantané s'écrit alors : $h(t) = \frac{k}{\lambda} \left(\frac{t}{\lambda}\right)^{(k-1)}$.

1.3.2 Modèles à risques proportionnels

Les modèles à risques proportionnels sont utiles pour étudier l'effet d'une covariable sur la durée de survie. Ils sont définis à partir de la fonction de risque instantané h et ils supposent un effet constant des covariables au cours du temps sur la survie. Pour un sujet i ($i = 1, \dots, N$) ayant pour vecteur de covariables w_i , ces modèles sont définis pour tout $t > 0$ par la fonction de risque instantané :

$$h_i(t|w_i) = h_0(t)g(\gamma'w_i)$$

où h_0 , appelée fonction de risque instantané de base, commune à tous les sujets, représente le risque instantané sans influence d'aucune covariable et g est une fonction à définir qui dépend de γ , le vecteur de coefficients de régression associé à w_i .

On remarque que pour tout couple d'individus i et j qui ont pour covariables w_i et w_j respectivement, le rapport des risques instantanés est constant au cours du temps : $\frac{h_i(t|w_i)}{h_j(t|w_j)} = \frac{g(\gamma'w_i)}{g(\gamma'w_j)}$. Cela explique pourquoi ce type de modèle est appelé "à risques proportionnels". Mais cette appellation autorise également à avoir des covariables dépendant du temps $w_i(t)$, du moment que leur effet sur la survie est constant (i.e. γ ne dépend pas du temps).

Parmi les modèles à risques proportionnels, le modèle semi-paramétrique de Cox est le plus courant en analyse de données de survie (COX, 1972). Il suppose que la fonction g soit la fonction exponentielle et h_0 une fonction indéterminée :

$$h_i(t|w_i) = h_0(t) \exp(\gamma'w_i)$$

Il est également possible d'avoir des modèles paramétriques en spécifiant le risque instantané de base h_0 à l'aide d'une fonction paramétrique vue section 1.3.1.

Le modèle de Cox a été étendu aux covariables dépendant du temps (ANDERSEN et GILL, 1982), tout comme les modèles paramétriques. Ces modèles supposent une évolution du biomarqueur entièrement spécifiée par une fonction en escalier, l'absence d'erreur de mesure et que l'existence de la covariable ne dépend pas de l'évènement.

Il faut alors distinguer deux type de covariables dépendant du temps (RIZOPOULOS, 2012b) : les covariables exogènes dont l'évolution n'est pas impactée par la survenue de l'évènement (par exemple, l'évènement "tentative de suicide" n'affecte pas la covariable dépendant du temps "saison") et les covariables endogènes (par exemple, l'évènement "rechute" affecte l'évolution du biomarqueur PSA). Les extensions du modèle de Cox et des modèles paramétriques sont valides pour des covariables exogènes, mais pas pour des covariables endogènes car leur évolution complète n'est pas disponible et elles sont mesurées avec une erreur. Or en recherche clinique, un biomarqueur est une covariable dépendant du temps endogène et donc analyser son évolution comme si c'était une covariable exogène mènerait à des résultats erronés. Il est donc nécessaire de s'intéresser à des méthodes adaptées à ce type de données longitudinales.

1.4 Modèles conjoints

1.4.1 Principe et spécification du modèle dans le contexte des MNLEM

Les modèles conjoints ont été introduits à la fin des années 1990, avec pour première application l'étude du lien entre évolution des CD4 chez des patients séropositifs et délai de progression vers le sida (FAUCETT et THOMAS, 1996; WULFSOHN et TSIATIS, 1997; TSIATIS et DAVIDIAN, 2004). Les modèles conjoints sont utiles non seulement dans le cas d'un modèle de survie avec une covariable dépendant du temps mesurée avec une erreur, mais aussi lorsque des données répétées d'un biomarqueur sont censurées par des sorties d'études ("*dropouts*") informatives (WU et al., 2011).

Le principe est de modéliser les données longitudinales du biomarqueur par un modèle à effets mixtes (cf paragraphe 1.2.2), de modéliser le risque de décès par un modèle à risques proportionnels (cf paragraphe 1.3.2) et de lier les deux en utilisant une structure latente commune. Cette dernière doit capturer l'association entre les deux processus afin qu'ils soient indépendants conditionnellement à cette structure latente. On rencontre deux types de modélisation conjointe : les modèles conjoints "*à effets aléatoires partagés*" et les modèles conjoints "*à classes latentes*". Les modèles conjoints à effets aléatoires partagés supposent une structure latente continue (à travers les effets aléatoires). Ils consistent à modéliser le risque d'évènement comme une fonction dépendant de l'évolution individuelle du biomarqueur (WU et al., 2011) et donc dépendant des effets aléatoires du modèle à effets mixtes utilisé pour décrire les données longitudinales, d'où le nom de modèles conjoints à "effets aléatoires partagés". Quant aux modèles conjoints à classes latentes, ils supposent une structure latente discrète (à travers des classes latentes) impliquant l'existence de sous-groupes d'individus qui partagent la même évolution du biomarqueur et le même risque d'évènement (LIN et al., 2002; PROUST-LIMA et al., 2014). Dans cette thèse, nous nous intéressons uniquement aux modèles conjoints à effets aléatoires partagés.

Un modèle conjoint est constitué de deux sous-modèles, l'un décrivant le processus longitudinal et l'autre le délai de survenue de l'évènement. En utilisant les mêmes notations que dans les paragraphes 1.2.2 et 1.3.2, pour un sujet i ($i = 1, \dots, N$) chez lequel on observe $\{y_i, T_i, \delta_i\}$ où $y_i = \{y_{i1}, \dots, y_{in_i}\}$ est le vecteur des n_i observations du sujet i ($j = 1, \dots, n_i$), $T_i = \min(X_i, C_i)$ avec X_i le temps de survie et C_i son temps de censure, et $\delta_i = \mathbf{1}_{\{X_i \leq C_i\}}$ l'indicatrice du statut vital, les deux sous-modèles s'écrivent de façon générale :

$$y_{ij} = b(t_{ij}, \psi_i, w_i) + g(t_{ij}, \psi_i, w_i, \Sigma)e_{ij} \quad (1.2)$$

$$h_i(t|\mathbf{B}_i(t)) = h_0(t) \exp(\gamma'w_i + \beta f(t, \psi_i, w_i)) \quad (1.3)$$

Où $\mathbf{B}_i(t) = \{b(u, \psi_i, w_i), 0 \leq u \leq t\}$ décrit le vrai processus longitudinal non-linéaire

du biomarqueur jusqu'au temps $t > 0$ et le vecteur de paramètres β mesure l'effet d'une fonction f dépendant des vraies valeurs du biomarqueur $b(t, \psi_i, w_i)$. Si $\beta = 0$, cela signifie que les deux processus sont indépendants.

Différentes paramétrisations pour la fonction de lien f entre l'évolution du biomarqueur et la survie sont possibles. La plus fréquente est celle supposant que c'est la vraie valeur courante du biomarqueur qui impacte le risque d'évènement, $f(t, \psi_i, w_i) = b(t, \psi_i, w_i)$ (FAUCETT et THOMAS, 1996; WULFSOHN et TSIATIS, 1997; IBRAHIM et al., 2010). Mais ce peut être également des fonctions de l'évolution du biomarqueur, comme la dérivée de $b(t, \psi_i, w_i)$, son intégrale ou encore une combinaison de ces différentes caractéristiques de l'évolution du biomarqueur (YE et al., 2008; SÈNE et al., 2014a). Il est également possible de considérer que l'effet du biomarqueur est retardé dans le temps avec $f(t, \psi_i, w_i) = b(t - c, \psi_i, w_i)$ où c est le décalage temporel (RIZOPOULOS, 2012b). Enfin, le vecteur des effets aléatoires η_i peut être introduit directement dans la fonction de risque instantané avec $f(t, \psi_i, w_i) = \eta_i$ (JACQMIN-GADDA et al., 2004; LIU et HUANG, 2009). Selon la dimension de la fonction f , β est un scalaire ou un vecteur.

Le risque instantané de base $h_0(t)$ peut être non spécifié, comme dans un modèle de Cox, mais il a été montré que cela pouvait entraîner une sous-estimation des erreurs standards des estimations des paramètres (HSIEH et al., 2006). Pour éviter ce problème, on peut expliciter $h_0(t)$ en utilisant une fonction de risque correspondant à une distribution paramétrique, telle qu'une loi de Weibull, log-normale ou de Gamma (JACQMIN-GADDA et al., 2004; HU et al., 2011; ANDRINOPOULOU et al., 2015). Des fonctions plus flexibles ont également été proposées dans la littérature pour $h_0(t)$, telles que des fonctions constantes par morceaux (RIZOPOULOS, 2012a; PARK et QIU, 2014) ou des fonctions splines (RIZOPOULOS et al., 2009; SÈNE et al., 2014a).

Le vecteur de paramètres relatifs à la survie θ_s contient donc γ , β et éventuellement les paramètres de la fonction $h_0(t)$ si le risque instantané de base est paramétrique.

De nombreuses extensions des modèles conjoints ont été proposées. Certaines permettent la modélisation conjointe de données longitudinales et de temps d'évènement de différents types, appelés risques compétitifs (par exemple la rechute et le décès). Cela permet d'étudier l'association entre une covariable dépendant du temps endogène et le risque pour l'une des causes d'intérêt (ELASHOFF et al., 2008; HU et al., 2009; HUANG et al., 2011). Des modèles conjoints de données longitudinales et d'évènements récurrents tels que des hospitalisations ou des fractures osseuses ont également été proposés (LIU et HUANG, 2009; VIGAN, 2014; KRÓL et al., 2016). Enfin, plusieurs biomarqueurs peuvent être introduits dans la fonction de risque instantané (CHI et IBRAHIM, 2006; RIZOPOULOS et GHOSH, 2011; ANDRINOPOULOU et al., 2014).

1.4.2 Méthodes d'estimation

Le but est d'estimer $\theta = \{\theta_l, \theta_s\}$ où $\theta_l = \{\mu, \nu, \Omega, \Sigma\}$ sont les paramètres relatifs au MNLEM et $\theta_s = \{\gamma, \beta, \lambda, k\}$ les paramètres relatifs à la survie (si h_0 risque instantané de base est une fonction paramétrique de Weibull par exemple). Deux méthodes d'estimation sont principalement rencontrées dans la littérature : une approche en deux étapes et une approche basée sur la vraisemblance conjointe.

Approche en deux étapes

Au vu de la complexité numérique du calcul de la vraisemblance conjointe (voir équation 1.4), les premières approches ont consisté à diviser le problème en deux étapes successives. Cette approche est simple, intuitive et facile à implémenter dans les logiciels standards de statistique (TSIATIS et al., 1995; PROUST-LIMA et al., 2008) :

1. Les paramètres de population de la cinétique du biomarqueur sont tout d'abord estimés, ignorant le processus de censure dû au décès, et les paramètres individuels sont estimés par des estimations bayésiennes empiriques (EBEs), notées $\hat{\psi}_i$.
2. Puis les paramètres relatifs à la survie sont estimés en maximisant la vraisemblance usuelle (définie à partir de l'équation 1.1) avec pour fonction de risque instantané $h_i(t|\hat{\psi}_i, w_i) = h_0(t) \exp(\gamma'w_i + \beta f(t, \hat{\psi}_i, w_i))$.

Dans la seconde étape, on peut également introduire directement les paramètres individuels comme covariables dans la fonction de risque instantané (MURAWSKA et al., 2012), ou encore ajouter la pente de croissance du biomarqueur (PROUST-LIMA et al., 2008). Cette méthode peut être utilisée lorsque les données longitudinales sont collectées avant le début du suivi pour la survie car alors le biomarqueur est une covariable exogène (MURAWSKA et al., 2012).

S'il a été montré qu'une approche en deux étapes diminuait le biais obtenu avec un modèle de Cox avec des covariables dépendant du temps (DAFNI et TSIATIS, 1998; YE et al., 2008), les résultats restent biaisés et, comme cela a été étudié dans le contexte de modèles linéaires à effets mixtes, une approche par vraisemblance conjointe est nécessaire pour éliminer ce biais (IBRAHIM et al., 2010; SWEETING et THOMPSON, 2011).

Vraisemblance conjointe

Les modèles conjoints à effets aléatoires partagés supposent que, conditionnellement aux effets aléatoires η_i , le processus longitudinal et les temps d'évènement sont indépendants. La contribution d'un sujet i à la log-vraisemblance conjointe s'écrit (RIZOPOULOS, 2012b) :

$$LL_i(\theta) = \log \int p(y_i|\eta_i; \theta)p(T_i, \delta_i|\eta_i; \theta)p(\eta_i; \theta)d\eta_i \quad (1.4)$$

où $\theta = \{\theta_l, \theta_s\}$ est le vecteur des paramètres à estimer, relatifs aux sous-modèles longitudinal et de survie, $p(y_i|\eta_i; \theta)$ est la fonction de densité de probabilité des observations longitudinales conditionnellement aux effets aléatoires, $p(T_i, \delta_i|\eta_i; \theta) = h_i(T_i|\mathbf{B}_i(T_i); \theta)^{\delta_i} S_i(T_i|\mathbf{B}_i(T_i); \theta)$ est la vraisemblance du modèle de survie, avec $S_i(t|\mathbf{B}_i(t); \theta) = \exp(-\int_0^t h_i(s|\mathbf{B}_i(s); \theta) ds)$ et $p(\eta_i; \theta)$ la fonction de densité de probabilité des effets aléatoires.

Déjà lorsque le processus longitudinal est défini par un modèle linéaire à effets mixtes, l'intégrale présente dans la fonction de survie $S_i(t|\mathbf{B}_i(t); \theta)$ et celle présente dans l'équation (1.4) n'ont pas de solution analytique. Elles doivent donc être approchées par des techniques d'intégration numérique, telles que la quadrature de Gauss, la méthode de Monte-Carlo (WULFSOHN et TSIATIS, 1997; HENDERSON et al., 2000; SONG et al., 2002), ou par approximation de Laplace lorsque le nombre d'effets aléatoires est grand (RIZOPOULOS et al., 2009). Plus récemment, la quadrature de Gauss-Hermite pseudo-adaptative a été proposée comme alternative et a montré de bonnes propriétés numériques (RIZOPOULOS, 2012a).

L'estimation des paramètres des deux sous-modèles (1.2) et (1.3) peut alors se faire simultanément en maximisant la log-vraisemblance conjointe $LL(\theta) = \sum_i LL_i(\theta)$. Plusieurs approches peuvent être utilisées pour l'estimation : l'algorithme Expectation-Maximization (EM) qui considère les effets aléatoires comme des données manquantes, l'algorithme de type Newton-Raphson ou une combinaison des deux (LIN et al., 2002; RIZOPOULOS et al., 2009). Il est également possible d'estimer les paramètres par une approche bayésienne utilisant des techniques MCMC (WANG et TAYLOR, 2001; CHI et IBRAHIM, 2006).

Les logiciels de statistique standards proposent des outils d'estimation de modèles conjoints à effets aléatoires partagés. Notamment, le package JM de R (RIZOPOULOS, 2010) est souvent utilisé. Il permet de choisir le type de fonction de risque instantané de base (Cox, Weibull, constante par morceaux, B-spline) et la fonction de lien entre le processus longitudinal et celui de survie (impact de l'évolution du biomarqueur sur la survie à travers sa valeur courante, sa dérivée, son intégrale, ses effets aléatoires, ou une combinaison de ces fonctions). On peut également y spécifier la méthode d'intégration numérique (quadrature de Gauss-Hermite standard ou pseudo-adaptative, approximation de Laplace) et l'algorithme d'estimation (EM, Newton-Raphson ou la combinaison des deux). Le package `frailtypack` de RONDEAU et al. (2012) permet également, entre autre, la modélisation conjointe de données longitudinales et de survie, en utilisant des quadratures de Gauss-Hermite pour approcher les intégrales et l'algorithme de Marquardt pour maximiser la vraisemblance conjointe. Le modèle peut être paramétrique ou semi-paramétrique et la fonction de lien entre le processus longitudinal et celui de survie peut être le vecteur des effets aléatoires ou la valeur courante du biomarqueur. Des macros SAS ont également été développées (GUO et CARLIN, 2004; GARCIA-HERNANDEZ et

RIZOPOULOS, 2015). Cependant, une limite de ces outils est qu'ils ne proposent de modéliser l'évolution du biomarqueur qu'avec une fonction linéaire ou, en cas d'une évolution non-linéaire, avec des fonctions splines. Cela ne permet pas l'utilisation de modèles physiologiques dont les paramètres ont une interprétation biologique et qui, par essence, sont non-linéaires.

L'algorithme SAEM implémenté dans le logiciel Monolix initialement dédié aux MNLEM, a récemment été étendu à l'estimation de modèles conjoints. Une première évaluation a permis de montrer qu'il fournissait des estimations sans biais des paramètres longitudinaux et de survie, même lorsque la partie longitudinale était décrite par un MNLEM (MBOGNING et al., 2015). Le logiciel NONMEM permet également l'estimation des paramètres d'un modèle conjoint dont le processus longitudinal est non-linéaire, utilisant l'approximation de Laplace (HU et SALE, 2003; HOLFORD, 2005; BJÖRNSSON et al., 2015).

Comme pour les MNLEM, après l'estimation des paramètres de population, on peut estimer les effets aléatoires $\hat{\eta}_i$. Les estimations bayésiennes empiriques (EBEs) des paramètres individuels peuvent être obtenues comme le mode de la distribution conditionnelle $p(\eta_i|T_i, \delta_i, y_i; \hat{\theta})$.

Approche "conjointe séquentielle"

Une approche intermédiaire inspirée de ce qui est réalisé en pharmacocinétique / pharmacodynamique (PKPD), où les paramètres de population pharmacocinétiques (PK) sont tout d'abord estimés, puis fixés dans le modèle PKPD (ZHANG et al., 2003), peut être envisagée. Nous l'appelons modélisation "conjointe séquentielle". Elle consiste en :

1. Une première étape, identique à la première étape de la modélisation en deux étapes décrite précédemment, permet d'estimer les paramètres de population relatifs au processus longitudinal.
2. Une seconde étape consiste à fixer les paramètres de population longitudinaux aux valeurs estimées dans l'étape 1 et à estimer les paramètres de population relatifs à la survie en maximisant la vraisemblance conjointe (1.4.2).

La différence entre la modélisation en deux étapes et la modélisation conjointe séquentielle réside uniquement dans la seconde étape, puisque les paramètres de population longitudinaux sont estimés de la même façon dans l'étape 1 des deux méthodes. Dans l'approche en deux étapes, les paramètres individuels longitudinaux sont fixés et utilisés comme covariables dans la seconde étape, tandis que dans l'approche conjointe séquentielle, les paramètres individuels longitudinaux sont réestimés pour le calcul de la vraisemblance conjointe.

Mais si en PKPD le processus PK peut être supposé indépendant de la pharmacodynamique (PD), ce n'est pas le cas avec des données longitudinales et de

survie, où l'estimation des paramètres longitudinaux peut être biaisée si on ne prend pas en compte la censure due aux décès ou aux *dropouts* (BONATE et SUTTLE, 2013).

1.4.3 Tests statistiques

Comme les paramètres des modèles conjoints sont estimés par maximum de vraisemblance, il est possible d'utiliser les tests asymptotiques standards basés sur la fonction de vraisemblance (RIZOPOULOS, 2012b). Dans le cas simple mais fréquent où on s'intéresse à tester si le processus longitudinal a un effet sur la survie, l'hypothèse nulle s'écrit : $H_0 : \beta = 0$ et l'hypothèse alternative : $H_1 : \beta \neq 0$.

On pourra utiliser :

- Le test de Wald dont la statistique de test est définie par : $W = \frac{\hat{\beta}}{s.e.(\hat{\beta})}$ où $\hat{\beta}$ est l'estimation de β et $s.e.(\hat{\beta})$ son erreur standard estimée. Sous l'hypothèse nulle, la distribution asymptotique de W est une loi normale centrée réduite. On rejettera donc H_0 au risque α si $|W| > \phi^{-1}(1 - \alpha/2)$ où ϕ^{-1} est la fonction quantile de la loi normale centrée réduite.
- Le test du rapport de vraisemblance (noté LRT pour *Likelihood ratio test*) dont la statistique de test est définie par : $LRT = -2[LL(\hat{\theta}_0) - LL(\hat{\theta})]$ où LL désigne la fonction de log-vraisemblance conjointe (définie page 31) et $\hat{\theta}_0$ et $\hat{\theta}$ les estimateurs du maximum de vraisemblance sous l'hypothèse nulle et alternative respectivement. Sous l'hypothèse nulle, la distribution asymptotique du LRT est une loi du χ^2 à un degré de liberté.

Notons que le test de Wald ne nécessite qu'une seule estimation des paramètres par maximum de vraisemblance, sous l'hypothèse alternative, alors que le LRT nécessite d'ajuster les données sous les deux modèles et est donc par conséquent plus coûteux en temps de calcul. De plus, dans Monolix, le calcul de la vraisemblance est réalisé par des méthodes d'échantillonnage préférentiel (Important Sampling), et une erreur standard de la log-vraisemblance est également estimée.

Ainsi, le risque de première espèce et la puissance de ces tests à détecter l'effet du processus longitudinal sur la survie peuvent être évaluées lors d'études de simulation. Sous l'hypothèse nulle, le risque de première espèce sera la proportion de tests rejetant à tort H_0 , tandis que sous l'hypothèse alternative, la puissance sera la proportion de tests rejetant H_0 à raison.

Test de Wald et LRT ne peuvent être utilisés que pour comparer deux modèles emboîtés. Si on souhaite comparer deux modèles qui ne sont pas emboîtés, des critères d'informations sont généralement utilisés. Le principe est de comparer la vraisemblance des modèles pénalisée par le nombre de paramètres du modèle. Deux critères d'information sont fréquemment utilisés :

- L'AIC (pour *Akaike's Information Criteria*) : $AIC = -2LL(\hat{\theta}) + 2n_{par}$ où n_{par} est

le nombre de paramètres du modèle.

- Le BIC (pour *Bayesian Information Criteria*) : $BIC = -2LL(\hat{\theta}) + n_{par} \log(N)$ où N est le nombre total de sujets.

On sélectionnera le modèle fournissant la plus petite valeur d'AIC ou de BIC. Le BIC a tendance à sélectionner des modèles plus parcimonieux que l'AIC.

1.4.4 Outils diagnostiques

L'évaluation des modèles conjoints consiste principalement en l'étude séparée des deux sous-modèles, à l'aide de résidus (erreur de prédiction) spécifiques aux MNLEM et aux modèles de survie. Dans cette section, nous présentons les outils diagnostiques graphiques qui ont été utilisés au cours de cette thèse.

Évaluation du sous-modèle longitudinal

Un premier examen visuel est celui des ajustements individuels. Pour chaque patient, les EBEs des paramètres individuels (voir section 1.2.2), notées $\hat{\psi}_i$, sont calculées et il est alors possible de tracer l'évolution individuelle du biomarqueur obtenue avec ces EBEs, et de la superposer aux observations du biomarqueur chez ce patient. On s'attend alors à avoir une évolution prédite proche des points d'observation pour chaque sujet.

On peut également s'intéresser aux IWRES (pour *individual weighted residuals*) qui sont des résidus individuels (différence entre observation et prédiction) pondérés. Ils sont calculés pour tous les individus i à tous les temps t_{ij} :

$$IWRES_{ij} = \frac{y_{ij} - b(t_{ij}, \hat{\psi}_i, w_i)}{g(t_{ij}, \hat{\psi}_i, w_i, \hat{\Sigma})}$$

Dans le cas d'une erreur résiduelle constante, la pondération $g(t_{ij}, \hat{\psi}_i, w_i, \hat{\Sigma}) = \hat{\sigma}$ est l'écart-type estimé de l'erreur résiduelle. Les IWRES sont alors tracés versus le temps et versus les valeurs prédites du biomarqueur, afin de détecter d'éventuelles tendances dans l'erreur de prédiction (LAVIELLE, 2014).

Évaluation du sous-modèle de survie

Les résidus de Cox-Snell permettent de vérifier si le modèle de survie ajuste correctement les données. Pour chaque individu, ils sont calculés comme la valeur estimée du risque instantané cumulé au temps T_i :

$$r_i^{CS} = \int_0^{T_i} h_i(s | \hat{\mathbf{B}}_i(T_i); \hat{\theta}) ds$$

où $\hat{\mathbf{B}}_i(T_i)$ est l'évolution du biomarqueur prédite par le modèle.

Si le modèle ajuste bien les données, on s'attend à ce que la probabilité de survivre jusqu'au temps t , $S(t) = \mathbb{P}(X_i > t)$, suive une loi uniforme sur $[0, 1]$. Cela revient à avoir un risque instantané cumulé, $H(t) = -\log(S(t))$ qui suit une loi exponentielle de paramètre 1. Il faut alors vérifier que les résidus de Cox-Snell r_i^{CS} sont distribués selon une loi exponentielle de paramètre 1. Mais il ne faut pas omettre la possibilité de censure

de X_i et donc de r_i^{CS} . Pour cela, on compare la fonction de survie des r_i^{CS} estimée par Kaplan-Meier à la fonction de survie de la distribution exponentielle de paramètre 1, $\exp(-t)$ (RIZOPOULOS, 2012b; SÈNE et al., 2014a).

Les résidus de martingale sont définis comme la différence entre l'événement observé et le risque cumulé d'événement au temps d'événement, c'est-à-dire :

$$r_i^M = \delta_i - r_i^{CS}$$

Ces résidus permettent ainsi d'évaluer la capacité du modèle à prédire l'événement. On les trace en fonction des valeurs prédites par le modèle du biomarqueur introduit dans le modèle (RIZOPOULOS, 2012b; SÈNE et al., 2014a).

1.4.5 Modèles conjoints et cancer de la prostate

Le cancer de la prostate est un domaine dans lequel la modélisation conjointe a particulièrement été utilisée, du fait de la facilité à mesurer le PSA, supposé lié à l'apparition et à l'évolution de la maladie. Ainsi, des modèles diagnostiques ont été étudiés (LIN et al., 2002), mais ce sont essentiellement des modèles conjoints pour données répétées de PSA et survenue d'une rechute après radiothérapie qui ont été développés. En 2002, un premier modèle conjoint où le PSA était décrit par une fonction linéaire en deux morceaux a été proposé avec une estimation réalisée avec une procédure MCMC (PAULER et FINKELSTEIN, 2002). Parallèlement, LAW et al. (2002) et YU et al. (2004) développaient un modèle conjoint de guérison dans lequel chaque patient est supposé être soit guéri par la radiothérapie, soit susceptible de rechuter. Pour cela, l'évolution du PSA était modélisée par un modèle à effets mixtes hiérarchique dont la fonction structurelle était biexponentielle. Le premier proposait une estimation par maximum de vraisemblance, utilisant les algorithmes EM et Newton-Raphson, tandis que le second proposait une approche bayésienne. En 2008, un modèle où la cinétique du PSA est définie par des splines utilise une méthode d'estimation en deux étapes (YE et al., 2008). La même année, PROUST-LIMA et al. (2008) développent un modèle où le $\log(PSA(t) + 0.1)$ est décrit par un modèle linéaire à effets mixtes, l'estimation des paramètres y sera également réalisée par une approche en deux étapes. Ce modèle sera largement réutilisé et adapté car il capture bien la cinétique du PSA, tout d'abord toujours avec une méthode d'estimation en deux étapes (KENNEDY et al., 2010; TAYLOR et al., 2014), puis avec une estimation par maximum de vraisemblance (SÈNE et al., 2014a). Le modèle a été adapté pour une modélisation conjointe à classes latentes (PROUST-LIMA et al., 2014). Enfin, toujours à partir de ce modèle, un modèle conjoint multi-état a été développé (FERRER et al., 2015), décrivant l'évolution du PSA et les différentes transitions entre les événements cliniques de la progression de la maladie (par exemple une rechute locale après radiothérapie, suivie d'une hormono-thérapie, puis d'une rechute à distance métastatique et le décès), les différents processus étant liés par des effets aléatoires partagés.

Plusieurs de ces modèles ont permis de faire de la prédiction individuelle de l'évolution du PSA et du risque de rechute après radiothérapie (TAYLOR et al., 2005; YU et al., 2008; PROUST-LIMA et TAYLOR, 2009; SÈNE et al., 2014b). TAYLOR et al. (2013) mettent même à disposition un *calculateur* (<https://psacalc.sph.umich.edu/>) permettant à des patients ayant un cancer de la prostate local et n'ayant pas subi de chirurgie, d'estimer leur risque de rechute après radiothérapie en fournissant simplement quelques caractéristiques à l'initiation du traitement et les mesures de PSA réalisées depuis.

A notre connaissance, aucun modèle conjoint pour données de PSA et de survie n'a été proposé pour des patients ayant un cancer de la prostate très avancé, tel un cancer de la prostate métastatique hormono-résistant.

1.5 Prédictions individuelles dynamiques

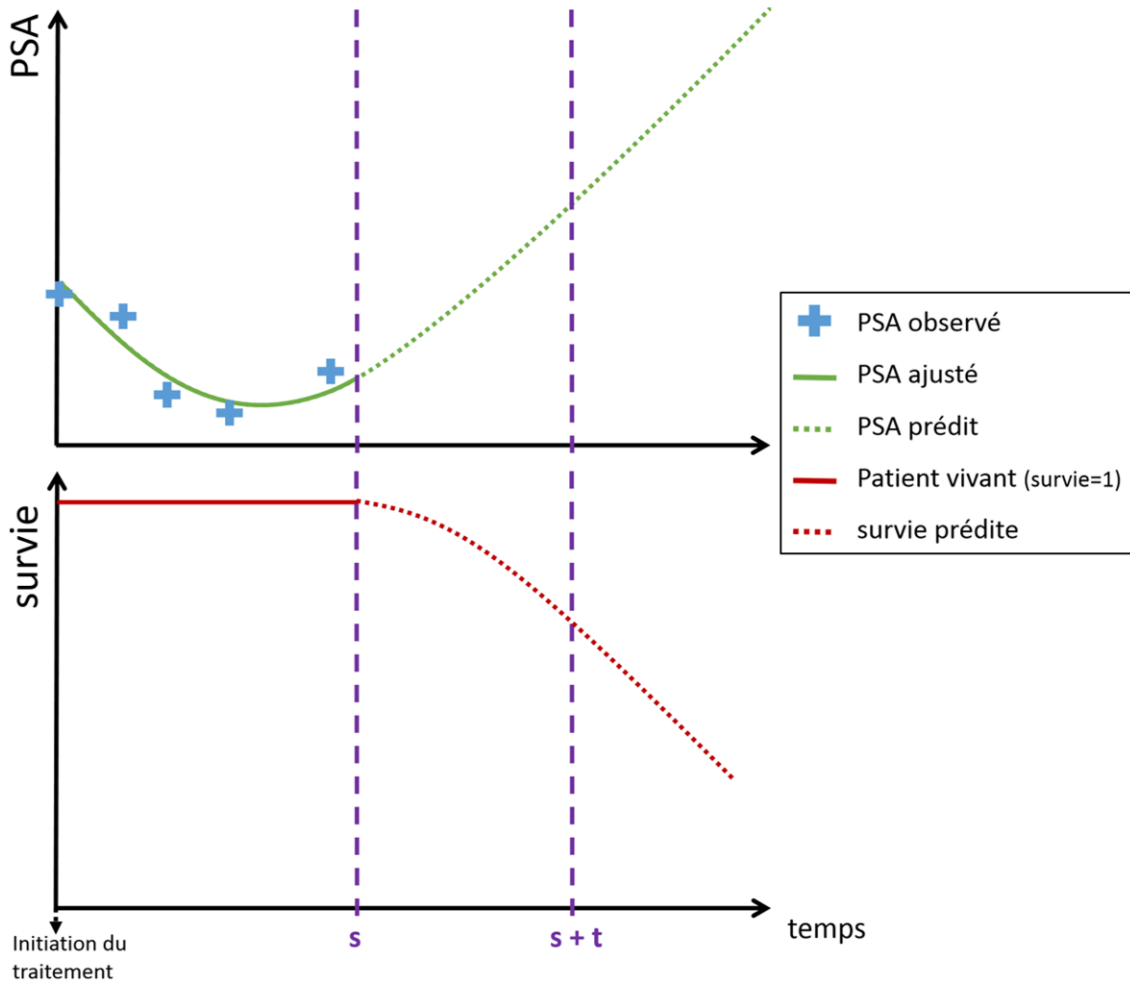
En pleine expansion, la médecine personnalisée vise à proposer la prise en charge la mieux adaptée à chaque individu. Pour cela, il apparaît nécessaire de prendre en compte la meilleure information individuelle disponible, non seulement les caractéristiques générales du patient, mais aussi toute l'information récoltée au cours de son suivi. L'objectif est alors d'utiliser toutes ces données récoltées de façon dynamique chez un patient pour prédire la survenue de l'évènement d'intérêt chez ce patient. On parle de prédiction individuelle dynamique car les prédictions peuvent être actualisées dès que de nouvelles mesures sont collectées, permettant d'étudier l'évolution de la maladie, d'améliorer la précision de la prédiction et d'adapter le traitement.

Le principe de la prédiction individuelle dynamique est d'utiliser toutes les données disponibles chez un patient à un temps s appelé temps *landmark*, afin de prédire la future évolution du biomarqueur et le risque de décès au temps $s + t$, où $t > 0$ est appelé fenêtre de prédiction, sachant que le patient est vivant au temps landmark s (voir figure 1.4).

La médecine personnalisée encourage donc la recherche d'outils méthodologiques pour faire de la prédiction individuelle. Dans ce cadre, la modélisation conjointe de données longitudinales et de survie apparaît comme un outil particulièrement adapté pour fournir des prédictions individuelles dynamiques de l'évènement d'intérêt à l'aide des mesures longitudinales (PROUST-LIMA et TAYLOR, 2009; RIZOPOULOS, 2011). Avoir une compréhension mécanistique de la maladie et du traitement pouvant améliorer les prédictions, il paraît nécessaire d'étendre les outils développés pour des modèles conjoints dont le processus longitudinal est décrit par un modèle linéaire à effets mixtes, à des modèles conjoints non-linéaires.

Dans cette section, nous verrons donc comment un modèle conjoint peut être utilisé pour fournir des prédictions individuelles dynamiques, puis nous introduirons des mesures de discrimination et de calibration permettant d'évaluer la prédictibilité de tels modèles.

FIGURE 1.4 – Principe de la prédiction individuelle dynamique de l'évolution du biomarqueur et de la survie



1.5.1 Utilisation de la modélisation conjointe

De façon plus formelle, en reprenant les notations des précédentes sections, soit un modèle conjoint dont les vraies valeurs des paramètres de population sont notées θ . L'objectif est de prédire la future évolution du biomarqueur et surtout le risque de décès chez un nouveau patient i pour lequel on dispose de mesures longitudinales du biomarqueur jusqu'au temps landmark s : $Y_i(s) = \{y_{ij}; 0 \leq t_{ij} \leq s\}$. On cherche donc à estimer en particulier la probabilité conditionnelle de décéder avant un temps $s+t$ sachant que le patient est vivant au temps s ($X_i > s$) :

$$\pi_i(s+t|s) = \mathbb{P}(X_i < s+t | X_i > s, Y_i(s); \theta) \quad (1.5)$$

Dans cette définition, on note la nature dynamique de la prédiction puisque, si une nouvelle mesure du biomarqueur est collectée à un temps $s' > s$, on peut actualiser la prédiction en calculant $\pi_i(s+t|s')$.

L'estimation de cette probabilité nécessite dans un premier temps d'estimer θ sur un jeu de données, on note cet estimateur $\hat{\theta}$, puis de calculer les effets aléatoires individuels du nouveau patient i à partir du modèle conjoint et des données longitudinales $Y_i(s)$, sachant que le patient est vivant au temps s , notés $\eta_i(s)$. Les EBEs de $\eta_i(s)$, notés $\hat{\eta}_i(s)$, peuvent être utilisés, fournissant l'estimateur de $\pi_i(s+t|s)$ suivant :

$$\hat{\pi}_i(s+t|s) = \frac{S_i(s+t|\hat{\eta}_i(s); \hat{\theta})}{S_i(s|\hat{\eta}_i(s); \hat{\theta})} + \mathcal{O}\left(\frac{1}{n_i(s)}\right) \quad (1.6)$$

où $\hat{\eta}_i(s)$ est le mode de la distribution conditionnelle $p(\eta_i(s)|Y_i(s); \hat{\theta})$, S_i est la fonction de survie du patient i (définie page 31) et $n_i(s)$ est le nombre de mesures du biomarqueur chez le patient i au temps *landmark* s . RIZOPOULOS (2011) a montré par simulation que cet estimateur était utile en pratique. Cependant, il est difficile d'en calculer l'erreur type et donc de fournir un intervalle de confiance. Pour ce faire, une méthode de Monte Carlo peut être utilisée (RIZOPOULOS, 2011), permettant de prendre en compte l'incertitude sur l'estimation des paramètres de population et individuels dans la prédiction et de proposer une prédiction accompagnée d'un intervalle de prédiction. Dans cette thèse, nous utiliserons Stan (STAN DEVELOPMENT TEAM, 2015), un logiciel de programmation spécialement conçu pour la statistique bayésienne, pour tirer au sort des valeurs dans la distribution des effets aléatoires individuels.

1.5.2 Évaluation des modèles prédictifs

L'examen des performances prédictives d'un modèle repose sur plusieurs notions (GERDS et al., 2008), dont les plus utilisées sont la discrimination et la calibration. La discrimination est la capacité d'un modèle à distinguer les sujets qui vont expérimenter l'évènement des sujets qui ne l'expérimenteront pas dans la fenêtre de temps considérée. La calibration est la capacité du modèle à fournir une prédiction proche de la réalité.

Dans cette section, nous présenterons en détail les métriques utilisées dans la thèse : l'AUC (BLANCHE et al., 2013a) et le score de Brier (GERDS et SCHUMACHER, 2006), puis d'autres métriques plus brièvement.

Aire sous la courbe ROC

Dans ce paragraphe, nous présentons tout d'abord l'AUC (pour *Area Under the ROC Curve*) dans le cas simple d'une réponse binaire, puis son adaptation à une réponse de survie.

AUC pour une réponse binaire

A l'origine, la notion de discrimination a été introduite afin d'évaluer les capacités diagnostiques d'un nouveau test, généralement basé sur une mesure à un temps donné d'un biomarqueur continu, notée B_i pour le sujet i . La discrimination est alors la capacité du biomarqueur à distinguer les sujets malades des sujets non malades. Notons M_i la variable binaire, indicatrice de la maladie chez le sujet i : $M_i = 1$ si i est malade (i est alors appelé 'cas'), $M_i = 0$ sinon ('contrôle'). Supposons sans perte de généralité qu'une valeur élevée de B_i indique un risque plus important de maladie qu'une faible valeur. C'est le cas par exemple du biomarqueur PSA dans le diagnostic du cancer de la prostate.

On va alors chercher quel est le seuil, c , qui permet au mieux le diagnostic de la maladie, c'est-à-dire d'obtenir un test positif ($B_i > c$) pour les sujets effectivement malades, et un test négatif ($B_i \leq c$) pour les sujets sains. La sensibilité ($Se(c)$) reflète alors le taux de vrais positifs et la spécificité ($Sp(c)$) le taux de vrais négatifs :

$$Se(c) = \mathbb{P}(B_i > c | M_i = 1)$$

et

$$Sp(c) = \mathbb{P}(B_i \leq c | M_i = 0)$$

L'objectif est de maximiser à la fois la sensibilité et la spécificité. On peut tracer l'ensemble des points $Se(c)$ versus $1 - Sp(c)$: c'est la courbe *ROC (Receiver Operating Characteristic)*. Par définition, la courbe ROC est comprise dans le carré $[0, 1] \times [0, 1]$ et va du point $(0; 0)$ au point $(1; 1)$. Elle est confondue avec la première bissectrice si le biomarqueur n'a aucun pouvoir discriminant entre les sujets malades et les sujets sains. Le point $(0; 1)$ correspond à une sensibilité et à une spécificité égales à 100%, c'est donc le point correspondant à une discrimination parfaite.

On peut s'intéresser à la capacité de discrimination du biomarqueur de manière plus générale, en utilisant une mesure résumée de la courbe ROC. On examine alors l'aire sous la courbe ROC (AUC) (PEPE, 2003) :

$$AUC = \int_0^1 ROC(u) du$$

Plus elle est grande, meilleure est la discrimination du biomarqueur. L'AUC a également une formulation plus intuitive : c'est la probabilité qu'un sujet malade ait une mesure de biomarqueur supérieure à celle d'un sujet sain (HANLEY et MCNEIL, 1982) :

$$AUC = \mathbb{P}(B_i > B_j | M_i = 1, M_j = 0)$$

où i et j sont deux sujets indépendants. Pour chaque paire $\{i, j\}$ de sujets "comparables", c'est-à-dire dont le statut de la maladie M est différent, l'AUC est la probabilité que l'ordre de leurs niveaux de biomarqueur soit concordant avec leurs statuts de la maladie. Une AUC de 1 indique une discrimination maximale, tandis qu'un AUC de 0.5 indique une discrimination aléatoire.

AUC pour un temps d'évènement

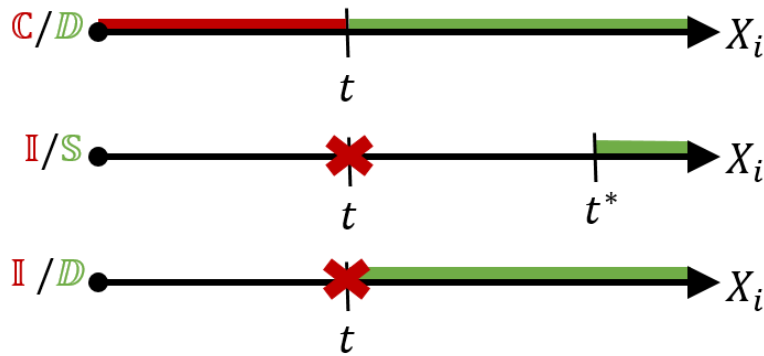
La définition de l'AUC ci-dessus est la définition d'origine, proposée dans le cadre de tests diagnostiques, avec un biomarqueur continu et une variable binaire caractérisant la présence/absence de la maladie. Depuis, elle a été étendue au cadre des tests pronostiques, afin d'évaluer la capacité d'un biomarqueur, toujours noté B et mesuré au temps $t = 0$, à distinguer les sujets qui vont expérimenter un évènement (apparition de la maladie, rechute ou décès par exemple) à court terme des sujets qui ne l'expérimenteront qu'à long terme. La variable binaire indiquant la maladie est alors remplacée par un temps de survie noté X .

Pour étendre les notions de sensibilité et spécificité à des données de survie, les concepts de 'cas' et 'contrôle' doivent être redéfinis (HEAGERTY et ZHENG, 2005). Pour un temps t donné, on appellera le sujet i :

- 'cas incident' si $X_i = t$
- 'cas cumulatif' si $X_i \leq t$
- 'contrôle statique' si $X_i > t^*$ pour un t^* fixé
- 'contrôle dynamique' si $X_i > t$

A partir de ces définitions, trois couples $\{Se, Sp\}$ sont principalement utilisés et sont illustrés par la figure 1.5 :

FIGURE 1.5 – Représentation schématique des 3 grandes familles de couples cas (rouge)/contrôle (vert)



1. Sensibilité cumulative et spécificité dynamique, C/D (HEAGERTY et al., 2000) :

$$Se^C(c, t) = \mathbb{P}(B_i > c | X_i \leq t)$$

$$Sp^D(c, t) = \mathbb{P}(B_i \leq c | X_i > t)$$

Sous ces définitions, à chaque temps t , tous les sujets de la population sont classés soit en 'cas', soit en 'contrôle' en fonction de leur statut au temps t . Un patient i bascule du groupe 'contrôle' au groupe 'cas' au temps X_i .

2. Sensibilité incidente et spécificité statique, I/S (SLATE et TURNBULL, 2000) :

$$Se^{\mathbb{I}}(c, t) = \mathbb{P}(B_i > c | X_i = t)$$

$$Sp^{\mathbb{S}}(c, t) = \mathbb{P}(B_i \leq c | X_i > t^*)$$

Ces définitions supposent que les sujets ne changent pas de groupe : ils sont soit ‘cas’, soit ‘contrôle’. Les cas sont stratifiés en fonction de leur temps d’évènement, les contrôles sont les sujets qui n’expérimentent pas l’évènement avant un temps tardif t^* fixé.

3. Sensibilité incidente et spécificité dynamique, \mathbb{I}/\mathbb{D} (HEAGERTY et ZHENG, 2005) :

$$Se^{\mathbb{I}}(c, t) = \mathbb{P}(B_i > c | X_i = t)$$

$$Sp^{\mathbb{D}}(c, t) = \mathbb{P}(B_i \leq c | X_i > t)$$

Avec ces définitions, un sujet i joue le rôle de ‘contrôle’ pour tous les temps $t < X_i$ et joue le rôle de ‘cas’ uniquement lorsque $t = X_i$.

Ces différentes définitions sont dépendantes du temps : les capacités du biomarqueur (mesuré au temps $t = 0$) à discriminer les ‘cas’ des ‘contrôles’ vont changer avec le temps d’horizon de la prédiction. Ainsi la courbe ROC et l’AUC associées vont également dépendre du temps auquel on souhaite prédire et refléter comment la discrimination du biomarqueur évolue au cours du temps.

Dans le contexte de prédictions dynamiques, l’utilisation de la définition \mathbb{C}/\mathbb{D} apparaît être la plus pertinente car on recherche à prédire un évènement dans une fenêtre de temps et non à un temps précis (ce qui est fait avec une sensibilité incidente) et à distinguer les patients n’ayant pas expérimenté l’évènement à la fin de cette fenêtre de temps plutôt qu’à un temps prédéfini t^* (ce qui est fait avec une spécificité statique) (BLANCHE et al., 2013a). Pour cette définition, l’AUC utilisant un biomarqueur mesuré au temps 0 s’écrit :

$$AUC^{\mathbb{C}/\mathbb{D}}(t) = \mathbb{P}(B_i > B_j | X_i \leq t, X_j > t)$$

Dans cette thèse, B_i sera remplacé par la probabilité de décéder avant le temps $s + t$ sachant que le patient est vivant au temps s estimée en utilisant le modèle conjoint, $\hat{\pi}_i(s + t | s)$ (voir équation 1.6). L’AUC dépendra donc à la fois du temps landmark s (durée au cours de laquelle le biomarqueur est mesuré à plusieurs reprises) et de la fenêtre de prédiction t (horizon pour lequel on souhaite prédire l’évènement) :

$$AUC(s, t) = \mathbb{P}(\pi_i(s + t | s) > \pi_j(s + t | s) | s < X_i < s + t, X_j > s + t)$$

Score de Brier

L’erreur de prédiction mesure la distance entre l’observation (le statut vital d’un patient à un temps t donné) et la prédiction (la probabilité de l’évènement prédite par le modèle au temps t , notée $\hat{\pi}(t)$). Plusieurs erreurs de prédiction ont été proposées dans la littérature (KORN et SIMON, 1990), dont l’erreur de prédiction quadratique, également appelée score de Brier (BS). Ce score a été initialement proposé par BRIER

(1950) pour évaluer la précision des prévisions météorologiques. Il a plus tard été étendu aux données de survie (GRAF et al., 1999) et aux événements censurés à droite (GERDS et SCHUMACHER, 2006). Pour une probabilité de décéder avant un horizon t , $\pi(t)$, le score de Brier est défini par :

$$BS(t) = \mathbb{E}[(\mathbf{1}_{\{X < t\}} - \pi(t))^2]$$

De bonnes capacités prédictives sont associées à un faible BS. Le score de Brier peut être compris entre 0 et 1. Cependant, puisqu'un modèle non-informatif qui attribuerait une probabilité de décès de 0.5 à tous les sujets conduit à un BS de 0.25, on s'attend à avoir des valeurs inférieures à 0.25.

Une propriété intéressante du score de Brier est qu'il s'écrit en deux composantes : un terme de "vraie variation" et un terme d'erreur du modèle due à une mauvaise spécification du modèle (VAN HOUWELINGEN et PUTTER, 2011). Le BS mesure donc les performances prédictives globales, capturant à la fois les capacités de discrimination et de calibration du modèle (STEYERBERG et al., 2010).

Néanmoins l'interprétation du Brier score peut apparaître compliquée, surtout dans le cas de prédictions dynamiques car il dépend de l'incidence cumulée de l'évènement. Pour pallier cette difficulté, on peut normaliser le score de Brier et utiliser une métrique qu'on appellera *scaled Brier score* (sBS) (souvent appelée R^2 dans la littérature). Elle est définie par GRAF et al. (1999) :

$$sBS(t) = 1 - \frac{BS(t)}{BS_{KM}(t)}$$

où BS_{KM} est le score de Brier obtenu en utilisant l'estimateur de Kaplan-Meier comme prédiction de la survie pour tous les sujets. L'interprétation de cette mesure est plus simple, elle permet de comparer les performances d'un modèle à celles d'un modèle n'apportant aucune information individuelle. On cherche alors à avoir un grand sBS.

Comme pour l'AUC, dans un contexte de prédictions dynamiques, BS et sBS dépendront à la fois de s et de t , $\hat{\pi}(t)$ étant remplacé par $\hat{\pi}(s+t|s)$ qui utilise toutes les données du patient i jusqu'au temps s pour prédire la survenue de l'évènement à l'horizon $s+t$. On a alors :

$$BS(s, t) = \mathbb{E}[(\mathbf{1}_{\{s < X < s+t\}} - \pi(s+t|s))^2 | X > s]$$

Prise en compte de la censure : les estimateurs IPCW

Pour présenter l'AUC et le BS, nous avons supposé que l'évènement d'intérêt X était observé. Or, nous avons vu précédemment que la principale difficulté à surmonter lorsqu'on utilise des données de survie est de prendre en compte le fait que les observations peuvent être censurées. L'estimation des mesures de discrimination et de calibration doit donc être adaptée à la possibilité de censure.

Dans ce but, une famille d'estimateurs a été proposée, les estimateurs non

paramétriques *IPCW* (pour *Inverse Probability of Censoring Weighting*), disponibles pour l'AUC et le BS (GRAF et al., 1999; HUNG et CHIANG, 2010). Ils prennent en compte la censure en pondérant par l'estimateur de Kaplan-Meier de la distribution de la censure C (BLANCHE et al., 2015). Un avantage de ces estimateurs est qu'il sont "libres de tout modèle", c'est-à-dire qu'il suffit de fournir les prédictions obtenues avec un modèle $\hat{\pi}(s + t|s)$ pour les calculer, sans hypothèse supplémentaire.

Autres métriques

D'autres métriques de discrimination et calibration existent.

Tout d'abord, le C-index, indice de concordance (HARRELL et al., 1996), est défini par : $\mathbb{P}(B_i > B_j | X_i < X_j)$ où B_i est la valeur du biomarqueur du sujet i et X_i son temps de décès. Il peut être adapté au cas de données de survie censurées (UNO et al., 2011). Contrairement à l'AUC qui dépend du temps d'horizon, le C-index est une mesure globale qui ne dépend pas d'un horizon spécifique t . Ici le C-index nous apparaissait donc moins pertinent pour évaluer les capacités prédictives dans un horizon de temps donné.

PENCINA et al. (2008) ont proposé deux indices permettant de comparer la discrimination de deux biomarqueurs : l'IDI pour *Integrated Discrimination Improvement* et le NRI pour *Net Reclassification Index*. Cependant HILDEN et GERDS (2014) montrent que ces nouvelles mesures peuvent mener à des conclusions fausses.

Enfin, une métrique récente de la précision de la calibration est l'entropie croisée pronostique (EPOCE), une mesure basée sur la vraisemblance et adaptée aux données censurées à droite (COMMENGES et al., 2012). Peu utilisée, elle présente pourtant de bonnes propriétés. Néanmoins, comme pour le C-index, l'EPOCE est une mesure des capacités pronostiques globales du modèle et ne permet donc pas de s'intéresser aux capacités à prédire à un horizon particulier.

1.6 Objectifs de la thèse

L'objectif principal de cette thèse est d'étudier la modélisation conjointe de données longitudinales et de survie dans le cas où la cinétique du biomarqueur est décrite par un modèle non-linéaire à effets mixtes (MNLEM). Nous nous placerons dans le contexte d'un essai clinique sur le cancer de la prostate métastatique hormono-résistant (CPmHR) et les analyses seront effectuées pour des patients de l'étude clinique VENICE recevant la chimiothérapie de première ligne de référence, docetaxel associé à de la prednisone.

Dans un premier temps, l'estimation par l'algorithme SAEM des paramètres de modèles conjoints où les données longitudinales sont modélisées par un MNLEM, nécessite d'être évaluée. Pour cela, des données longitudinales et de survie seront simulées, considérant différents liens entre les deux processus. L'étude de simulation

permettra d'évaluer les performances de l'algorithme SAEM de Monolix à fournir des estimations précises des paramètres de population, sous différents scénarios, en utilisant un modèle conjoint non-linéaire (Chapitre 2). La modélisation conjointe sera comparée à une approche en deux étapes et à une approche "conjointe séquentielle", en termes d'estimation des paramètres, de risque de première espèce et de puissance à détecter un lien entre les deux processus.

Dans un second temps, un modèle conjoint mécanistique sera développé chez les patients du jeu de données d'apprentissage (décrits dans le paragraphe 1.1.5), afin de caractériser le lien entre l'évolution non-linéaire du PSA et la survie chez des patients ayant un CPmHR et traités par docetaxel associé à de la prednisone. Pour le sous-modèle longitudinal, le modèle structurel du MNLEM devra décrire le mécanisme de production du PSA et le mécanisme d'action du docetaxel, nécessitant l'utilisation d'équations différentielles. Plusieurs fonctions de lien entre processus longitudinal et de survie seront testées et le modèle conjoint sélectionné sera évalué en utilisant à la fois le jeu de données d'apprentissage et le jeu de données de validation (Chapitre 3).

Enfin, nous évaluerons l'utilisation d'un modèle conjoint non-linéaire pour fournir des prédictions individuelles dynamiques accompagnées d'intervalles de prédiction, chez des patients ayant un CPmHR, étendant des outils développés dans le contexte linéaire. Les performances prédictives du modèle seront étudiées à l'aide de métriques de discrimination et de calibration, l'AUC, le score de Brier (BS) et le BS normalisé (sBS), dépendant du temps (Chapitre 4).

Dans la suite de ce manuscrit, nous présentons donc ces trois travaux de thèse et leurs articles associés. Pour conclure, nous discuterons de l'ensemble de ces travaux et des perspectives qu'offre cette thèse dans le Chapitre 5.

Chapitre 2

Modèles non-linéaires à effets mixtes et lien avec la survie : Comparaison par simulation des approches conjointes et en deux étapes

2.1 Résumé

Nous avons réalisé une étude de simulation afin de comparer les performances de la modélisation conjointe à celles de la modélisation en deux étapes et de la modélisation "conjointe séquentielle" (voir section 1.4.2) en termes d'estimation des paramètres de population, de risque de première espèce et de puissance à détecter un lien entre processus longitudinal et processus de survie, lorsque les données longitudinales sont décrites par un MNLEM. Nous nous sommes placés dans un contexte d'essai clinique sur le cancer de la prostate métastatique, considérant un modèle conjoint pour des données longitudinales de PSA et de temps de décès.

Pour le modèle longitudinal, nous avons développé un modèle d'évolution non-linéaire du PSA à partir d'équations différentielles, supposant un mécanisme de production du PSA simplifié. Sous l'hypothèse d'un traitement ayant un effet constant ε sur la prolifération des cellules prostatiques (cellules de la prostate et cellules cancéreuses) jusqu'à un temps $T_{esc} > 0$, puis de l'apparition d'une résistance complète au traitement au temps T_{esc} (l'effet du traitement $e(t) = \varepsilon$ si $0 \leq t \leq T_{esc}$ et $e(t) = 0$ si $t > T_{esc}$), le système d'équations différentielles aboutit à une expression analytique de l'évolution du PSA, $PSA(t, \psi_i)$ où ψ_i sont les paramètres de l'individu i (équation (3) de l'article 1). Le modèle de survie utilisé est le modèle à risques proportionnels paramétrique suivant :

$$h_i(t|PSA(t, \psi_i)) = h_0(t) \exp(\beta PSA(t, \psi_i))$$

avec $h_0(t) = \frac{k}{\lambda} (\frac{t}{\lambda})^{(k-1)}$ un risque instantané de base de Weibull. Le modèle de survie suppose donc un lien entre la valeur courante du PSA et la survie, quantifié par le paramètre β .

Les valeurs utilisées pour simuler les patients ont été choisies en s'inspirant des vraies données du bras *placebo* de l'essai VENICE décrit dans le paragraphe 1.1.5. Trois scénarios de simulation ont été étudiés en faisant varier la force du lien $\beta \in \{0, 0.005, 0.02\}$ et en choisissant $k = 1.5$ et λ de façon à conserver une survie de 25% à 735 jours (fin de l'étude) pour un "patient médian" (i.e. ayant pour paramètre relatif au PSA les paramètres médians). Un quatrième scénario considère un lien fort $\beta = 0.02$, et k et λ aux valeurs utilisées pour le scénario sans lien ($\beta = 0$), entraînant une survie courte pour un patient médian (figure 2 de l'article 1). $M = 100$ jeux de données de $N = 500$ patients ont été simulés sous chaque scénario.

L'estimation des paramètres relatifs au modèle d'évolution du PSA et relatifs au modèle de survie a été réalisée pour chaque jeu de données et chaque scénario par modélisations en deux étapes, "conjointe séquentielle" et conjointe, en utilisant l'algorithme SAEM implémenté dans Monolix. Le modèle utilisé pour ces analyses était celui utilisé pour les simulations, et les valeurs initiales des paramètres de population étaient celles du scénario sans lien ($\beta = 0$). Les critères de convergence pour l'estimation des paramètres étaient ceux par défaut de l'algorithme SAEM de Monolix (PLAN et al., 2012). Pour chaque modèle et chaque scénario, les erreurs d'estimation relatives (REE) des paramètres ont été calculées pour chacun des M jeux de données. Le risque de première espèce et la puissance à détecter un effet du PSA sur la survie (hypothèse nulle $H_0 : \beta = 0$) ont également été estimés en utilisant le test du rapport de vraisemblance (LRT). Un LRT corrigé a été proposé afin de pallier le fait que Monolix estimait la log-vraisemblance avec une erreur standard non négligeable.

Comme attendu, aucun biais n'a été observé lorsqu'il n'y avait aucun lien entre PSA et survie, quelle que soit la méthode d'estimation des paramètres. En présence d'un lien entre les deux processus ($\beta \neq 0$), cette étude de simulation a montré que la modélisation conjointe était la seule méthode permettant d'obtenir des estimations non biaisées à la fois des paramètres du modèle du biomarqueur et des paramètres du modèle de survie. Concernant les paramètres de cinétique du PSA, de faibles biais apparaissaient en utilisant les approches en deux étapes et "conjointe séquentielle", notamment sur le paramètre représentant le taux de prolifération des cellules prostatiques r , guidant le rebond du PSA. De plus, l'approche en deux étapes conduisait à systématiquement sous-estimer le paramètre de lien β , tandis que l'approche "conjointe séquentielle" parvenait à corriger ce biais. Enfin, si la puissance à détecter un effet de la cinétique du PSA sur la survie était de 100% pour les 3 méthodes, le risque de première espèce était lui de 4, 3 et 12 % pour la modélisation conjointe, "conjointe séquentielle" et en deux étapes respectivement.

Cette étude montre ainsi que modéliser conjointement des données longitudinales non-linéaires et de survie est possible en utilisant l'algorithme SAEM et que les estimations fournies par un tel modèle sont meilleures que celles obtenues avec des méthodes plus simples. Cela ouvre la voie à l'utilisation de modèles conjoints capturant de façon plus

physiologique l'évolution d'un biomarqueur et son impact sur un temps d'évènement, ce qui peut être utile dans l'évaluation d'un médicament et dans la prédiction d'un évènement.

Ce travail a fait l'objet d'un article publié dans *The American Association of Pharmaceutical Scientists (AAPS) Journal* en 2015.

2.2 Article 1 (publié)

Research Article

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

Solène Desmée,^{1,2,4} France Mentré,^{1,2} Christine Veyrat-Follet,³ and Jérémie Guedj^{1,2}

Received 5 November 2014; accepted 20 February 2015; published online 5 March 2015

Abstract. In metastatic castration-resistant prostate cancer (mCRPC) clinical trials, the assessment of treatment efficacy essentially relies on the time to death and the kinetics of prostate-specific antigen (PSA). Joint modeling has been increasingly used to characterize the relationship between a time to event and a biomarker kinetics, but numerical difficulties often limit this approach to linear models. Here, we evaluated by simulation the capability of a new feature of the Stochastic Approximation Expectation-Maximization algorithm in Monolix to estimate the parameters of a joint model where PSA kinetics was defined by a mechanistic nonlinear mixed-effect model. The design of the study and the parameter values were inspired from one arm of a clinical trial. Increasingly high levels of association between PSA and survival were considered, and results were compared with those found using two simplified alternatives to joint model, a two-stage and a joint sequential model. We found that joint model allowed for a precise estimation of all longitudinal and survival parameters. In particular, the effect of PSA kinetics on survival could be precisely estimated, regardless of the strength of the association. In contrast, both simplified approaches led to bias on longitudinal parameters, and two-stage model systematically underestimated the effect of PSA kinetics on survival. In summary, we showed that joint model can be used to characterize the relationship between a nonlinear kinetics and survival. This opens the way for the use of more complex and physiological models to improve treatment evaluation and prediction in oncology.

KEY WORDS: joint model; metastatic prostate cancer; NLMEM; PSA; SAEM.

INTRODUCTION

Prostate cancer is the second most frequently diagnosed cancer in men and is responsible for about 300,000 deaths worldwide every year (1). Although treatment can be effective, a number of factors such as resistance or delayed treatment (4% of cancer have metastasized at diagnosis (2)) considerably worsen the treatment outcome. In the case of metastatic castration-resistant prostate cancer (mCRPC), the evaluation of chemotherapy efficacy primarily relies on the overall survival (3) and is complemented by the analysis of the prostate-specific antigen (PSA). Although countless studies have explored the relationship between different PSA kinetic parameters (such as doubling time or time to

nadir) and survival, the choice of a clearly defined parameter remains controversial.

The lack of consensus on how to use PSA kinetics is exacerbated by the difficulty to properly handle PSA kinetics and the time to event (time to death or dropout) in statistical models. Several methods have been proposed in the literature. The simplest one is to plug the individual PSA kinetic parameters into a survival model (4). However, the fact that these parameters are often not directly observed from the data makes this approach error prone. A second approach is to use a model to describe the entire PSA kinetics using mixed-effect models to precisely account for between-subjects variability, and then to plug individual model predictions as covariates in a survival model (5,6). However, this method, called in the following as “two-stage” approach, is prone to bias because it does not take into account the relationship between marker’s kinetic and the time to event and the uncertainty in the individual model predictions (7). In order to eliminate the bias found in the two-stage approach, one can use models which simultaneously, or “jointly,” handle longitudinal and time-to-event data (7–13). For the latter, one can either aim to estimate all parameters simultaneously (“joint” model) or in a sequential manner (“joint sequential” model), as it has been suggested in the pharmacometric field (14).

Electronic supplementary material The online version of this article (doi:10.1208/s12248-015-9745-5) contains supplementary material, which is available to authorized users.

¹ INSERM, IAME, UMR 1137, F-75018, Paris, France.

² Université Paris Diderot, IAME, UMR 1137, Sorbonne Paris Cité, F-75018, Paris, France.

³ Drug Disposition, Disposition Safety and Animal Research Department, Sanofi, Chilly-Mazarin, France.

⁴ To whom correspondence should be addressed. (e-mail: solene.desmee@insERM.fr)

The main challenge in using joint model is the numerical complexity involved by the calculation and the maximization of the likelihood. This difficulty can be circumvented by using a linear model for the PSA kinetics as implemented in the large majority of published models (7,15) or softwares (16). However, this precludes the use of physiologically more accurate models for PSA kinetics which are in essence nonlinear.

In the last years, pharmacometric softwares have addressed the need for joint model when the longitudinal model is defined by a nonlinear mixed-effect model (NLMEM). This approach was firstly implemented in NONMEM using a Laplacian approximation of the likelihood and was essentially used to account for informative dropouts (17,18). Recently, the Stochastic Approximation Expectation-Maximization (SAEM) algorithm, a method relying on an “exact” calculation of the likelihood, was extended to include time-to-event data (19,20) in Monolix and NONMEM.

Here, we evaluated the benefit of joint models using the SAEM algorithm for characterizing the relationship between survival and a biomarker having a nonlinear kinetics. We compared by simulation the precision and the type 1 error of longitudinal and survival parameters obtained using a joint, a joint sequential, and a two-stage model in the context of a clinical study in mCRPC according to the strength of the association between PSA kinetics and survival.

MODELS AND NOTATIONS

A Mechanistic Model for PSA Kinetics

In the absence of treatment, we assume that prostatic cells, C (mL^{-1}), proliferate with rate r (day^{-1}) and are eliminated with rate d (day^{-1}). PSA (ng.mL^{-1}) is secreted with a production rate p (ng.day^{-1}) and cleared from the blood with rate δ (day^{-1}). We suppose that a chemotherapy for mCRPC acts by blocking cell proliferation with time-varying effectiveness, $e(t)$, and

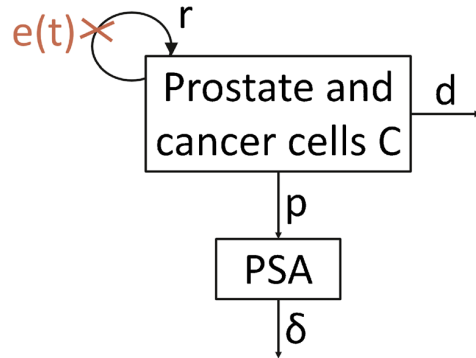


Fig. 1. Schema of the secretion of PSA by prostate and cancer cells. PSA is expressed in ng.mL^{-1} and C in mL^{-1} . r is the rate of prostatic cells proliferation in absence of treatment (day^{-1}), d the rate of prostatic cells elimination (day^{-1}), p the rate of PSA secretion by C (ng.day^{-1}), δ the rate of PSA elimination (day^{-1}), and $e(t)$ the time-dependent treatment effect

hence, the proliferation rate under treatment is given by $r'=r(1-e(t))$ with $0 \leq e(t) \leq 1$ (Fig. 1):

$$\begin{cases} \frac{dC}{dt} = r(1-e(t))C(t) - dC(t) \\ \frac{d\text{PSA}}{dt} = pC(t) - \delta\text{PSA}(t) \end{cases} \quad (1)$$

Treatment is initiated at $t=0$; $\text{PSA}(0)=\text{PSA}_0$ and $C(0)=C_0$ are PSA value and the number of prostatic cells at treatment initiation, respectively. Here, treatment is assumed to be constantly effective until a certain time, T_{esc} , after what tumor escapes and treatment has no longer an effect:

$$e(t) = \begin{cases} \varepsilon & \text{if } t \leq T_{\text{esc}} \\ 0 & \text{if } t > T_{\text{esc}} \end{cases} \quad (2)$$

Further, we made the assumption of quasi steady-state at treatment initiation, and thus, $C_0 = \frac{\delta \times \text{PSA}_0}{p}$. With this piecewise constant treatment effect (Eq. 2), the model (Eq. 1) has an analytical solution given by:

$$\text{PSA}(t, \psi) = \begin{cases} \frac{\delta \text{PSA}_0}{r(1-\varepsilon)-d+\delta} e^{(r(1-\varepsilon)-d)t} + \left[\text{PSA}_0 - \frac{\delta \text{PSA}_0}{r(1-\varepsilon)-d+\delta} \right] e^{-\delta t} & \text{if } t \leq T_{\text{esc}} \\ \frac{\delta \text{PSA}_0}{r-d+\delta} e^{(r-d)t-r\varepsilon T_{\text{esc}}} + \left[\text{PSA}(T_{\text{esc}}) - \frac{\delta \text{PSA}_0 e^{(r(1-\varepsilon)-d)T_{\text{esc}}}}{r-d+\delta} \right] e^{-\delta(t-T_{\text{esc}})} & \text{if } t > T_{\text{esc}} \end{cases} \quad (3)$$

Because only four parameters can be identified from Eq. 3, we fixed d to 0.046 day^{-1} , corresponding to a half-life of tumor cells of 15 days, consistent with an apoptotic index of 5% in metastatic prostate cancer (21). Moreover, we fixed δ to 0.23 day^{-1} , corresponding to a PSA half-life in blood of about 3 days (22). Thus, PSA kinetics was defined by the vector parameter $\psi = \{r, \text{PSA}_0, \varepsilon, T_{\text{esc}}\}$.

Statistical Model for PSA Measurements

Let y_{ij} denote the j th longitudinal measurement of $\log(\text{PSA}+1)$ for the individual i at time t_{ij} , where $i=1, \dots, N$, $j=1, \dots, n_i$, N is the number of subjects, and n_i the number of measurements in subject i . The nonlinear mixed-effect model (NLMEM) for PSA is defined as follows:

$$y_{ij} = \log(\text{PSA}(t_{ij}, \psi_i) + 1) + \sigma e_{ij} \quad (4)$$

where PSA is given by the formula (Eq. 3), ψ_i is the vector of the individual parameters, and e_{ij} is the residual error which follows a standard normal distribution with mean 0 and variance 1. The vector of the individual parameters ψ_i is decomposed as fixed effects $\mu = \{r, \text{PSA}_0, \varepsilon, T_{\text{esc}}\}$ representing median effects of the population and random effects η_i specific for each individual. It is assumed that $\eta_i \sim N(0, \Omega)$ with Ω the variance-covariance matrix. In this work, Ω is supposed to be diagonal. Each diagonal element ω_q^2 represents the variance of the q th component of the random effect vector η_i . We assumed log-normal distribution for r , PSA_0 , and T_{esc} :

$$\log(\psi_{q,i}) = \log(\mu_q) + \eta_{q,i}$$

and logit-normal distribution for ε :

$$\logit(\varepsilon_i) = \logit(\mu_\varepsilon) + \eta_{\varepsilon,i} \text{ with } \logit(x) = \log\left(\frac{x}{1-x}\right) \text{ for } 0 < x < 1.$$

The population parameter vector of PSA noted θ_1 is composed of $\{\mu, \Omega, \sigma\}$.

Statistical Model for Survival

Let X_i denote the time to event and C_i the censoring time for the i th subject. Survival data (T_i, δ_i) are observed in all individuals, where $T_i = \min(X_i, C_i)$ and $\delta_i = 1$, if $X_i \leq C_i$, 0 otherwise. For the event process, we used a hazard function of the form:

$$h_i(t | \text{PSA}(t, \psi_i)) = h_0(t) \exp(\beta \text{PSA}(t, \psi_i)) \quad (5)$$

where the baseline hazard function, h_0 , is a Weibull hazard function $h_0(t) = \frac{k}{\lambda} \left(\frac{t}{\lambda}\right)^{k-1}$. The parameter β measures the strength of the association between the PSA kinetics and the risk of death. If $\beta=0$, the survival process is independent on the PSA evolution, and survival data can be fitted by a Weibull model without adjusting for PSA. If $\beta \neq 0$, the survival process depends on the PSA kinetics. The survival parameters to estimate are $\theta_s = \{\lambda, k, \beta\}$.

Joint Model

Joint models assume conditional independence: given the random effects η_i , longitudinal measurements and survival events are independent. All parameters $\theta = \{\theta_1, \theta_s\}$ are simultaneously estimated. Thus, the joint log-likelihood for the subject i can be written as follows (23):

$$l_i(\theta) = \log \int p(y_i | \eta_i; \theta) \left\{ h_i(T_i | \eta_i; \theta)^{\delta_i} S_i(T_i | \eta_i; \theta) \right\} p(\eta_i; \theta) d\eta_i \quad (6)$$

where $S_i(t | \eta_i; \theta) = \exp\left(-\int_0^t h_0(s | \eta_i; \theta) \exp(\beta \text{PSA}(s, \psi_i)) ds\right)$ is the survival function conditionally on the random effects, $p(y_i | \eta_i; \theta)$ the density of the longitudinal observations

conditionally on the random effects, and $p(\eta_i; \theta)$ the density of the random effects.

Two-Stage Model

In order to simplify the heavy calculations involved by Eq. 6, one may also use a two-stage approach. In the first step, PSA kinetics parameters (θ_1) are estimated assuming complete independence of PSA kinetics and survival and the Empirical Bayes Estimates (EBEs), defined as the mode of the conditional distribution $p(\psi_i | y_i, \hat{\theta}_1)$, are used to provide individual parameters, noted $\hat{\psi}_i$. In the second step, the survival parameters (θ_s) are estimated maximizing the usual log-likelihood $l(\theta_s) = \sum_{i=1}^n \log \left\{ h_i(T_i | \text{PSA}(t, \hat{\psi}_i); \theta_s)^{\delta_i} S_i(T_i | \text{PSA}(t, \hat{\psi}_i); \theta_s) \right\}$.

This method is analogous to the sequential method called ‘‘Individual Pharmacokinetic Parameters’’ (IPP) used in pharmacometric field to handle pharmacokinetic/pharmacodynamic (PKPD) data (14).

Joint Sequential Model

In order to reduce the number of parameters to estimate in Eq. 6, an alternative consists of first estimating population PSA kinetic parameters, as done in the two-stage model, and then estimating parameters related to survival (θ_s) fixing the all population PSA parameters (θ_1) in Eq. 6 to the values obtained at the previous stage. This method, inspired from the ‘‘Population PK parameter and individual PK data’’ for the combined analysis of PKPD data, has been shown to limit the bias compared to two-stage approach (14).

SIMULATION STUDY

Design

Simulation setting was inspired by one arm of a phase III study of clinical development of chemotherapy for metastatic prostate cancer (3). $M=100$ datasets with $N=500$ patients were simulated with PSA measurements every 3 weeks for 2 years (i.e., the last possible measurement time was at $t=735$ days), leading to a maximum of 36 observations per patient (Table I). Follow-up was censored at $t=735$, and no other mechanism than death was considered for dropout. For the simulation of the time to death, k was fixed to 1.5, and three values for β were considered: 0, 0.005, and 0.02,

Table I. Values of the Population PSA Parameters Used for the Simulations in all Scenarios

	Fixed effects	Transformation	Inter-individual standard deviation (ω)
r (day^{-1})	0.05	log-normal	0.1
PSA_0 (ng.mL^{-1})	80	log-normal	0.6
ε	0.3	logit-normal	1.5
T_{esc} (day)	140	log-normal	0.6
σ	0.36	-	-

Table II. Values of the Population Survival Parameters Used for the Simulations of the Four Scenarios

	Scenario No link	Scenario Low link	Scenario High link	Scenario Short survival
β	0	0.005	0.02	0.02
λ (day)	580	765	2150	580
k	1.5	1.5	1.5	1.5

corresponding to “no link”, “low link” and “high link” between PSA kinetics and survival, respectively. In order to maintain a comparable amount of PSA measurements across scenarios, λ was determined in each scenario such that the probability of survival at the end of the follow-up (i.e., 735 days) was equal to 25% with the median PSA kinetic parameters (Table I), leading to $\lambda=580, 765,$ and 2150 in scenarios “No link”, “Low link” and “High link”, respectively (Table II). Lastly, we evaluated the effects of having both a large baseline risk and a strong effect of PSA kinetics by setting $\beta=0.02$ and $\lambda=580$ in the scenario called “Short survival”. Figure 2 shows the survival functions for these four scenarios for the “median patient”, i.e., a virtual patient having PSA kinetic parameters equal to the median values in the population.

In order to take into account the effect of withdrawals from PSA follow-up, we also explored additional scenarios where PSA and/or vital status were censored in case of PSA progression defined as an increase of 25% above the nadir and of at least 2 ng.mL compared to nadir (see Supplementary File 1).

All simulations were carried out with the R software version 3.0.1 (24).

Parameter Estimation

The estimation of two-stage, joint sequential, and joint models was performed for each dataset and each scenario using the Stochastic Approximation Expectation-Maximization (SAEM) algorithm in the software Monolix version 4.2.2.

The initial values of the parameters were those used for the scenario “No link.” Minus twice log-likelihood ($-2LL$) was calculated by Importance Sampling, fixing the Monte-Carlo sizes to achieve a sufficient level of precision for Likelihood Ratio Test (LRT) to 100,000 Monte-Carlo sizes for the scenario “No link,” 20,000 for the other scenarios in joint and joint sequential models, and 2,000 in two-stage model.

SAEM was run using one chain, and we verified that using three chains, as suggested in a related context (19), gave similar results. Central Processing Unit (CPU) times for parameter estimations and $-2LL$ estimations were recorded on a i7 64 bits 3.33 GHz.

Evaluation Criteria

We used Relative Estimation Errors defined by:

$REE(\hat{\theta}_m) = \frac{\hat{\theta}_m - \theta^*}{\theta^*} \times 100$, where θ^* and $\hat{\theta}_m$ are the true and estimated parameters, respectively, for dataset $m, m=1 \dots M$. Boxplots of the REEs with the 10 and 90% percentiles were plotted. When $\beta=0$ (scenario “No link”), the boxplot of the absolute estimation error, $\hat{\beta}_k$, with the 10 and 90% percentiles was plotted.

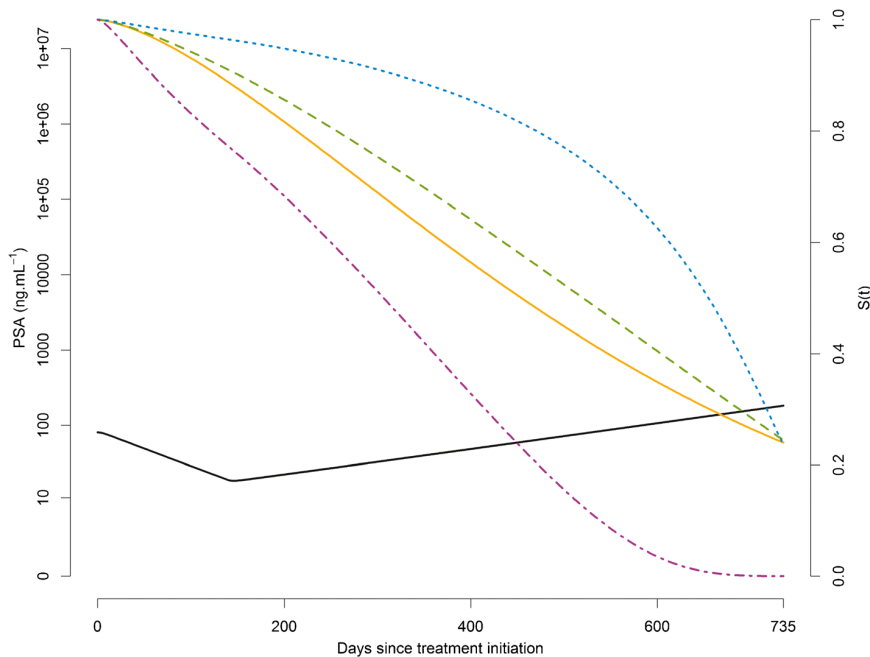


Fig. 2. Typical evolution of $PSA(t)$ (solid black) and survival functions in the typical patient (who have the fixed effects of the Table I as PSA parameters) for the scenarios “No link” ($\beta=0, \lambda=580$) (solid orange), “Low link” ($\beta=0.005, \lambda=765$) (dashed green), “High link” ($\beta=0.02, \lambda=2150$) (dotted blue), and “Short survival” ($\beta=0.02, \lambda=580$) (dotdash purple)

The type 1 error and power were calculated as the proportion among the M datasets for which LRT (called in the following “uncorrected LRT”) led to reject the null hypothesis $H_0: \beta=0$. The type 1 error was evaluated in the scenario “No link,” and the power was calculated for the three other scenarios. The significance level of the tests α for the observed type 1 error was fixed to 0.05, leading to a 95% prediction interval for 100 replicates equal to [0.7–9.3%].

In some cases, the estimation of $-2LL$, which relies on a stochastic approximation, was associated with a non-negligible standard error. Because this can lead to an inflation of the type 1 error, we also evaluated a “corrected LRT.” In

this test, H_0 was rejected if $-2LL(H_0)+2LL(H_1)$ belonged to $\left(\chi_1^2 + z_\alpha \sqrt{se_{-2LL}(H_0)^2 + se_{-2LL}(H_1)^2}; \infty\right)$ where $-se_{-2LL}(H_0)$ and $se_{-2LL}(H_1)$ are the estimated standard errors of $-2LL$ under H_0 and H_1 , respectively, whereas χ_1^2 and z_α are the chi-square value with one degree of freedom and the α th quantile of the standard normal distribution, respectively. $\sqrt{se_{-2LL}(H_0)^2 + se_{-2LL}(H_1)^2}$ is the uncertainty for the sum of the two estimated $-2LL$. We used a corrected LRT if $z_\alpha \sqrt{se_{-2LL}(H_0)^2 + se_{-2LL}(H_1)^2}$ was non-negligible compared to χ_1^2 , i.e., for a significance level of 5%, $se_{-2LL}(H_0)^2$ or $se_{-2LL}(H_1)^2$ non-negligible compared to 2.

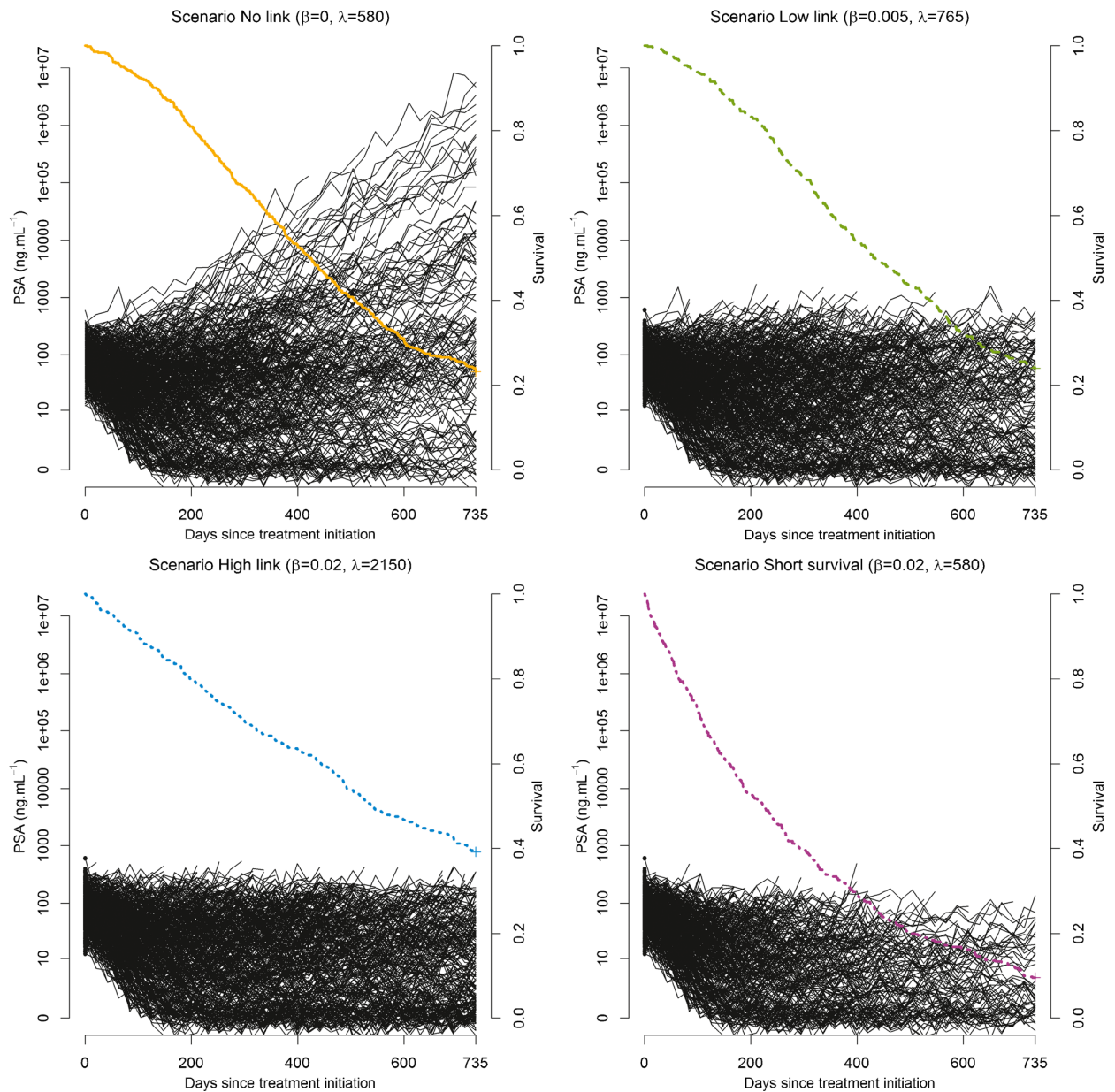


Fig. 3. Spaghetti plots (black lines) and estimated Kaplan-Meier curves (colored lines) for one typical dataset ($N=500$) for each of the four scenarios

RESULTS

Simulated Data

Figure 3 shows the spaghetti plots along with the Kaplan-Meier curves of one typical dataset for each of the four scenarios. Because the scenario “No link” assumes that death does not depend on PSA kinetics, PSA rebound after loss of treatment efficacy (at time T_{esc}) was more frequently observed than in the three other scenarios. As expected (see Methods), the numbers of measurements in the first three scenarios were largely similar (Table III). In the last scenario where both the baseline hazard function and the effect of PSA kinetics were large, early events frequently occurred and the total amount of PSA data was much smaller.

Estimation Performance

No bias was observed in the scenario “No link” regardless of the approach. In the two-stage approach, increasing effect of PSA on survival led to higher levels of bias (Figs. 4 and 5). In particular, there was a systematical underestimation of the PSA effect on survival with median (Q1;Q3) REE for β which equals to -8.4% (-12.6; -4.1), -14.8% (-18.9; -10.2), and -9.1% (-16.7; -5.7) in scenarios “Low link”, “High link” and “Short survival,” respectively. The bias was corrected by using joint sequential models or joint models, with median (Q1;Q3) REE for β of 0.005% (-3.6; 5.3), -0.6% (-3.3; 2.9), and 0.03% (-4.8; 5.7) in scenarios “Low link”, “High link” and “Short survival,” respectively, for this latter method. Although the three approaches led to low REEs for the almost all parameters of PSA kinetics (Fig. 4), a bias was observed for the proliferation rate of tumor cells, r , with median (Q1;Q3) REE which equals to -0.4% (-0.7; -0.05), -0.8% (-1.0; -0.5), and -1.2% (-1.8; -0.7) in scenarios “Low link”, “High link” and “Short survival,” respectively, when using the two-stage or joint sequential approach. Here as well, the bias was largely reduced when using a joint model, with median (Q1;Q3) REE for r which equals to 0.01% (-0.4; 0.3), -0.004% (-0.3; 0.2), and -0.08% (-0.6; 0.5), in scenarios “Low link”, “High link” and “Short survival,” respectively. Of note, large REE was found for the parameters λ and k (|REE|>30%) in the scenario “High link” when using joint and joint sequential models, due to the presence of a local extremum of the likelihood function.

Table III. Number of PSA Measurements Per Patient and Median Survival in the Total Number of Simulated (50,000) Patients for the Four Simulated Scenarios

Number of PSA measurements	Scenario No link	Scenario Low link	Scenario High link	Scenario Short survival
1-5	7%	7%	8%	29%
6-10	12%	13%	10%	20%
11-20	26%	31%	22%	26%
21-35	30%	29%	21%	17%
36	24%	21%	39%	8%
Median survival (day)	457	414	552	217

Lastly, we also considered additional scenarios where PSA data and/or vital status was censored in case of PSA increase (see Supplementary File 1). Although the precision of the parameter estimates was deteriorated due to the reduction in the amount of data available, no substantial bias in the PSA kinetic parameters was found using joint model. However, bias was found in survival-related parameters, in particular when both PSA and vital status were censored after disease progression and this bias systematically led to an overestimation of the hazard function.

Test Performance

The uncorrected LRT led to a type 1 error of 21, 9, and 12% for the joint model, the joint sequential model, and the two-stage model, respectively, i.e., outside the 95% prediction interval ([0.7-9.3%] for 100 replicates). The use of a corrected LRT (see methods) led to a smaller type 1 error of 4, 3 and 12% for the joint model, the joint sequential model, and the two-stage model, respectively. The reason why the type 1 errors for the two-stage model with or without correction were similar is due to the fact that the standard error of -2LL was negligible with this model ($<10^{-4}$). For the scenarios with $\beta \neq 0$, the power was 100% with all three models, regardless of the correction.

Computation Time

Mean CPU times for the simultaneous estimation of the 12 parameters using joint model were about five times larger than the total CPU times using two-stage model (70 vs 14 min) and about 1.2 times larger than the total CPU times using joint sequential model (70 vs 59 min), ignoring the time for specific data manipulation required for setting the two-stage and joint sequential approaches. Mean CPU times for the -2LL estimation using joint model (respectively, joint sequential model) and 100,000 chains were about 3.1 (respectively, 3.6) times larger than when using the joint approach and 20,000 chains (264 vs 86 min (respectively, 207 vs 57 min)) but led to a mean $se_{-2LL}(H_1)$ of 1.96 vs 5.14 (respectively, 2.40 vs 5.90).

DISCUSSION

Numerical complexities have long limited the use of joint models to longitudinal processes defined by linear mixed-effect models (7,10,15). Here, we evaluated by simulation, in the context of PSA and survival in metastatic prostate cancer, the capability of a new feature of the Stochastic Approximation Expectation-Maximization algorithm in Monolix to estimate the parameters of a joint model where the longitudinal process was defined by a nonlinear model. And we compared the results to two simplified approaches, two-stage and joint sequential models.

We found that joint model provided unbiased parameters of both longitudinal and survival processes. In contrast, the use of a two-stage model (5,25) led to large biases when PSA kinetics and survival were linked. In particular, the impact of the biomarker kinetics on the survival, as measured by the link parameter β , was systematically underestimated, consistent with results found in linear mixed-effect model (8,9). Beside survival parameters, a bias on the tumor proliferation

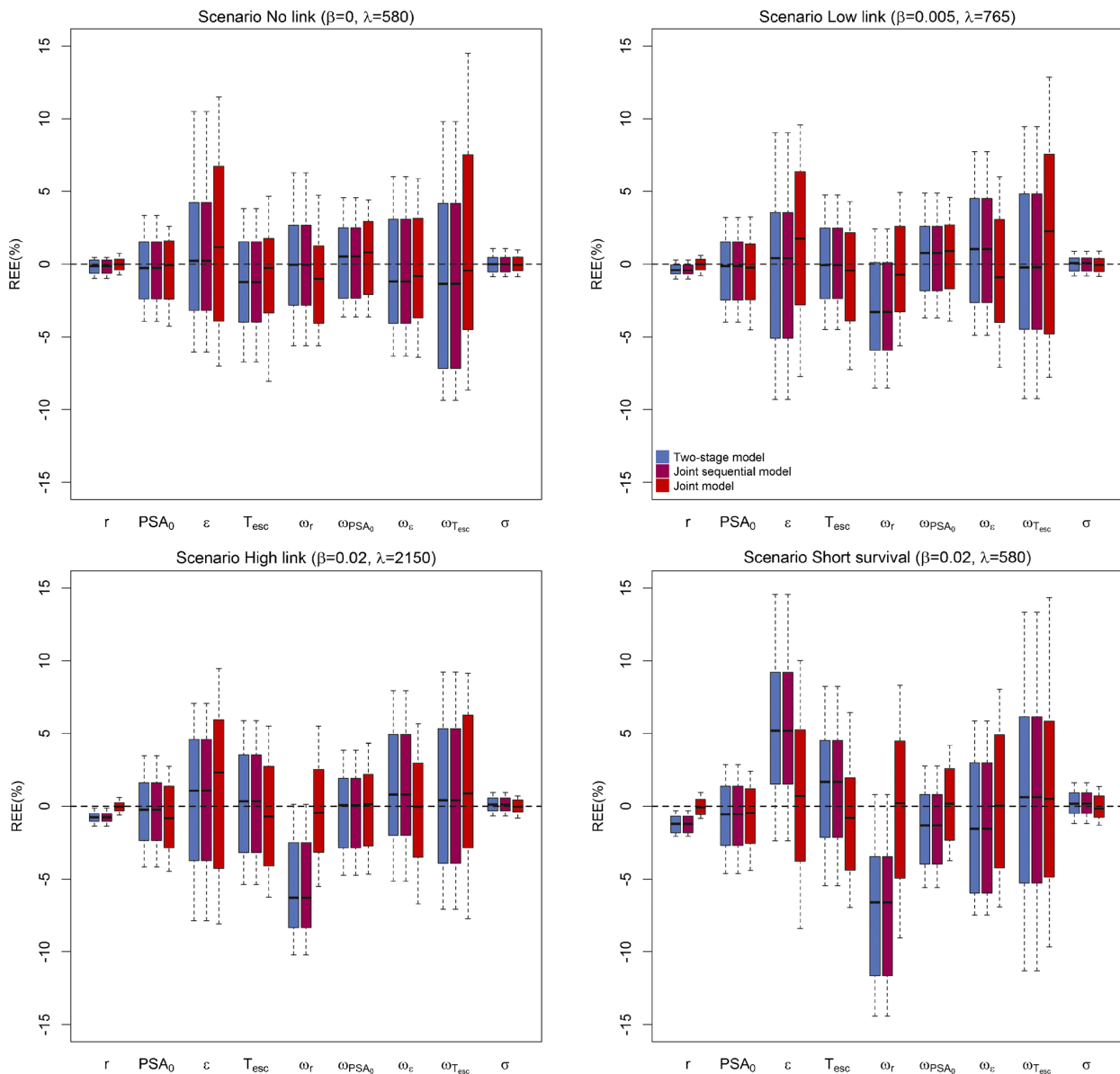


Fig. 4. Boxplots of the relative estimation errors for parameters related to PSA for two-stage model (blue), joint sequential model (purple), and joint model (red), for the four scenarios. *Top left: scenario No link ($\beta=0, \lambda=580$), top right: scenario Low link ($\beta=0.005, \lambda=765$), bottom left: scenario High link ($\beta=0.02, \lambda=2150$) and bottom right: scenario Short survival ($\beta=0.02, \lambda=580$)*

rate, r , which is the driving force for the increase in PSA, was also observed in scenarios with $\beta \neq 0$. The fact that no such bias occurred when using joint model shows that the simultaneous estimation of the hazard function also improved the estimation of the longitudinal parameters. Moreover, a two-stage approach led to an inflation of type 1 error (i.e., concluding to an effect of PSA on survival while there is none) which could be explained by the shrinkage of the EBEs in patients with short survival (26). By construction, the joint sequential model led to the same biases on PSA longitudinal parameters than the two-stage model, but the estimation of survival parameters was close to that obtained with the joint model. This, therefore, suggests that joint sequential model

could be a relevant approach when joint model cannot be performed.

In spite of the increasing numerical capability, the likelihood of joint models remains particularly complex to calculate. Here, we reported that the likelihood was estimated with a relatively large uncertainty. Increasing the Monte-Carlo sizes led to smaller standard errors of the likelihood but considerably increased the computation time. The impact of this error on type 1 error was in part accounted by using a corrected likelihood ratio test (LRT). However, more studies will be needed to precisely determine when this correction is necessary and whether it can be improved. Here, for instance, the corrected rejection region did not take into account the

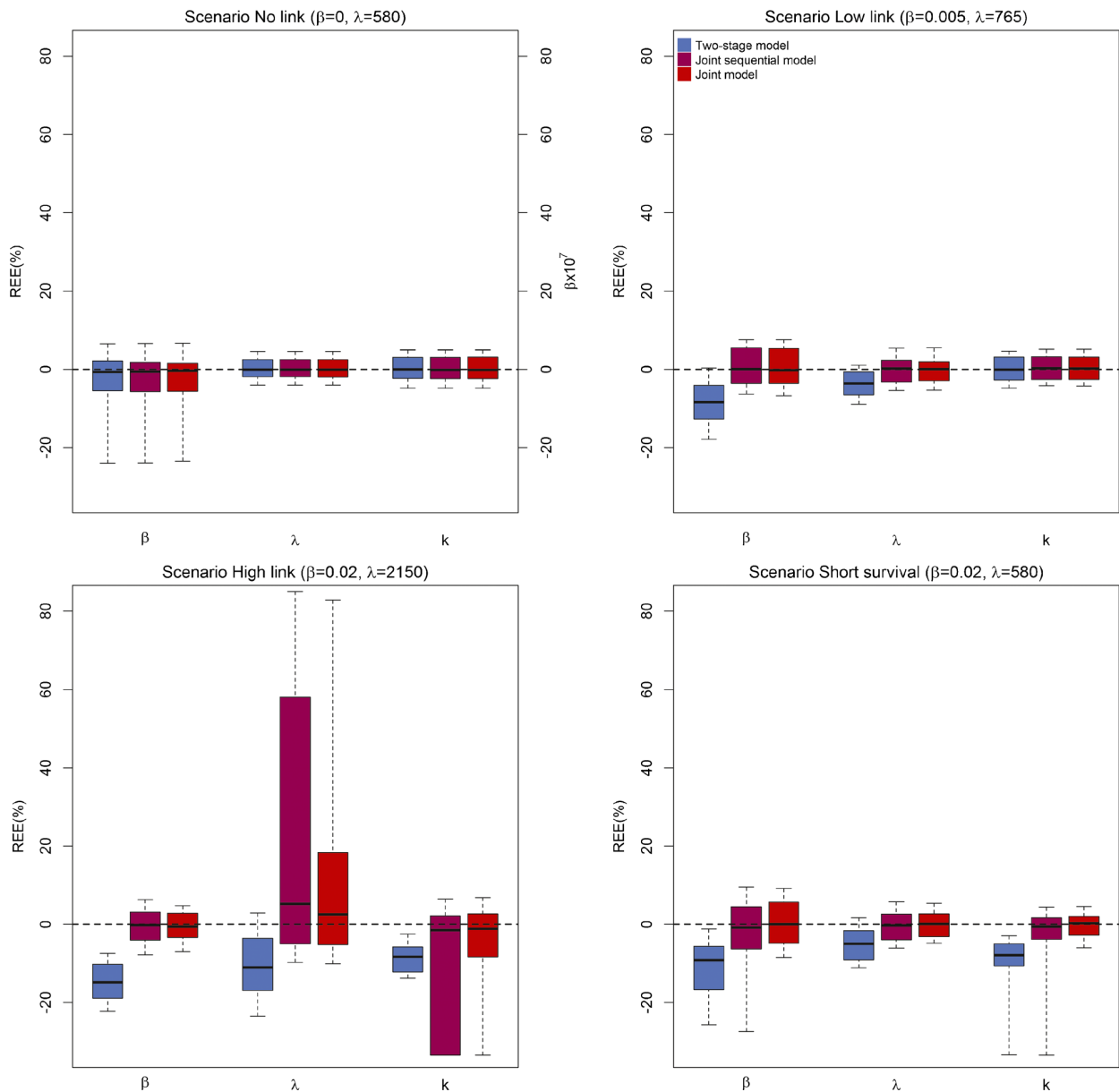


Fig. 5. Boxplots of the relative estimation errors for parameters related to survival (in % except for β of the scenario *No link* for which estimated values are represented) for two-stage model (blue), joint sequential model (purple), and joint model (red), for the four scenarios. *Top left: scenario No link ($\beta=0, \lambda=580$), top right: scenario Low link ($\beta=0.005, \lambda=765$), bottom left: scenario High link ($\beta=0.02, \lambda=2150$), and bottom right: scenario Short survival ($\beta=0.02, \lambda=580$)*

covariance between the likelihood calculated under the null and alternative hypotheses. Although the Wald test may be an alternative, we found here that standard errors of the estimates tended to be smaller than the root mean-squared errors (RMSE), indicating a potential underestimation of the standard errors. Furthermore, and in spite of the stochastic algorithm, the estimation of parameters related to survival was complicated in some cases by the existence of local extremum of the likelihood function. This stresses the need, in practice, to perform several runs of likelihood maximization using different initial values.

The main advantage of nonlinear models is the possibility to develop physiological models based on nonlinear differential equations, which naturally integrate the correlations between the different biomarkers. In this study, we used a rather simple model where the treatment effectiveness was piecewise constant, which allowed us to derive a biexponential analytical solution for the PSA kinetics. However, this model may clearly be oversimplistic, and more complex models will be needed that rely on several markers, such as tumor size or drug pharmacokinetics. The facilitated use of these models via joint models holds the promise that the determinants of survival may be much better characterized (20).

Regarding the survival model, we restricted here to a rather simple fully parametric model, where the baseline hazard function was a Weibull model (27), and the hazard function was related to the current PSA values. In practice, complex survival models could be evaluated, and standard tools for model selection (e.g., AIC or BIC) could be used to evaluate the effect of various transformations of the biomarker, such as the derivative or the cumulative value of PSA. Lastly, with a fully parametric model, the prediction and the simulation of the individual hazard function can easily be performed, making possible to guide treatment adaptation in a dynamic manner (28).

CONCLUSION

SAEM algorithm implemented in Monolix was shown to provide precise estimates for joint models where the longitudinal model was defined by a nonlinear mixed-effect model. This opens the way for a more systematic use of joint models and a better understanding of the relationship between biomarker kinetics and survival, especially in the field of metastatic cancer where survival and nonlinear biomarker kinetics are intrinsically related.

ACKNOWLEDGMENTS

The authors would like to thank Drug Disposition Department, Sanofi, Paris which supported Solène Desmée by a research grant during this work. They also thank Hervé Le Nagard for the use of the computer cluster services hosted on the “Centre de Biomodélisation UMR1137”.

REFERENCES

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase [Internet]. Lyon, France: International Agency for Research on Cancer. 2013. Report No.: 11. Available from: <http://globocan.iarc.fr>.
2. Howlader N, Noone A, Krapcho M, Garshell J, Miller D, Altekruse S, et al. SEER Cancer Statistics Review, 1975–2011 [Internet]. National Cancer Institute. 2013. Available from: http://seer.cancer.gov/csr/1975_2011/.
3. Tannock IF, Fizazi K, Ivanov S, Karlsson CT, Fléchon A, Skoneczna I, et al. Afibercept versus placebo in combination with docetaxel and prednisone for treatment of men with metastatic castration-resistant prostate cancer (VENICE): a phase 3, double-blind randomised trial. *Lancet Oncol*. 2013;14:760–8.
4. Petrylak DP, Ankerst DP, Jiang CS, Tangen CM, Hussain MHA, Lara PN, et al. Evaluation of prostate-specific antigen declines for surrogacy in patients treated on SWOG 99–16. *J Natl Cancer Inst*. 2006;98:516–21.
5. Proust-Lima C, Taylor JMG, Williams SG, Ankerst DP, Liu N, Kestin LL, et al. Determinants of change in prostate-specific antigen over time and its association with recurrence after external beam radiation therapy for prostate cancer in five large cohorts. *Int J Radiat Oncol Biol Phys*. 2008;72:782–91.
6. Murawska M, Rizopoulos D, Lesaffre E. A two-stage joint model for nonlinear longitudinal response and a time-to-event with application in transplantation studies. *J Probab Stat [Internet]*. 2012 [cited 2014 Apr 23];2012. Available from: <http://www.hindawi.com/journals/jps/2012/194194/abs/>.
7. Wu L, Liu W, Yi GY, Huang Y. Analysis of longitudinal and survival data: joint modeling, inference methods, and issues. *J Probab Stat [Internet]*. 2011 [cited 2013 Jul 10];2012. Available from: <http://www.hindawi.com/journals/jps/2012/640153/abs/>.
8. 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;53:750–63.
9. Ibrahim JG, Chu H, Chen LM. Basic concepts and methods for joint models of longitudinal and survival data. *J Clin Oncol*. 2010;28:2796–801.
10. Taylor JM, Wang Y. Surrogate markers and joint models for longitudinal and survival data. *Control Clin Trials*. 2002;23:626–34.
11. Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics*. 1997;53:330.
12. Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Stat Sin*. 2004;14:809–34.
13. Hsieh F, Tseng Y-K, Wang J-L. Joint modeling of survival and longitudinal data: likelihood approach revisited. *Biometrics*. 2006;62:1037–43.
14. Zhang L, Beal SL, Sheiner LB. Simultaneous vs. sequential analysis for population PK/PD data I: best-case performance. *J Pharmacokinet Pharmacodyn*. 2003;30:387–404.
15. Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics*. 2000;1:465–80.
16. Rizopoulos D. Joint models for longitudinal and time-to-event data: with applications in R. CRC Press. 2012. p. 275
17. Hu C, Sale ME. A joint model for nonlinear longitudinal data with informative dropout. *J Pharmacokinet Pharmacodyn*. 2003;30:83–103.
18. Björnsson MA, Friberg LE, Simonsson USH. Performance of nonlinear mixed effects models in the presence of informative dropout. *AAPS J*. 2014;17:245–55.
19. Vigan M, Stirnemann J, Mentré F. Evaluation of estimation methods and power of tests of discrete covariates in repeated time-to-event parametric models: application to Gaucher patients treated by imiglucerase. *AAPS J*. 2014;16:415–23.
20. Mbogning C, Bleakley K, Lavielle M. Joint modelling of longitudinal and repeated time-to-event data using nonlinear mixed-effects models and the stochastic approximation expectation–maximization algorithm. *J Stat Comput Simul*. 2015;85:1512–28.
21. Tu H, Jacobs SC, Borkowski A, Kyprianou N. Incidence of apoptosis and cell proliferation in prostate cancer: relationship with TGF- β 1 and bcl-2 expression. *Int J Cancer*. 1996;69:357–63.
22. Polascik TJ, Oesterling JE, Partin AW. Prostate specific antigen: a decade of discovery—what we have learned and where we are going. *J Urol*. 1999;162:293–306.
23. Rizopoulos D, Verbeke G, Lesaffre E. Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. *J R Stat Soc Ser B Stat Methodol*. 2009;71:637–54.
24. R Development Core Team. R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>. R Foundation for statistical computing, Vienna, Austria.
25. Taylor JMG, Yu M, Sandler HM. Individualized predictions of disease progression following radiation therapy for prostate cancer. *J Clin Oncol*. 2005;23:816–25.
26. Ribba B, Holford N, Mentré F. The use of model-based tumor-size metrics to predict survival. *Clin Pharmacol Ther*. 2014;96:133–5.
27. Yu M, Law NJ, Taylor JMG, Sandler HM. Joint longitudinal survival cure models and their application to prostate cancer. *Stat Sin*. 2004;14:835–62.
28. Ribba B, Holford NH, Magni P, Trocóniz I, Gueorguieva I, Girard P, et al. A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis. *CPT Pharmacomet Syst Pharmacol*. 2014;3:e113.

Supplementary file 1: Impact of withdrawal from PSA follow-up on joint estimation

It is frequent in clinical trial protocols that treatment is stopped when a disease progression (increase in PSA or tumor size, adverse events) is observed. When treatment is stopped, the patient may drop out of the study. Although the vital status may continue to be collected, the PSA measurements are no longer recorded in the study.

Simulation framework

For the sake of simplicity we consider here that a disease progression was only due to a PSA progression defined as *an increase of 25% above the nadir and of at least 2 ng/ml compared to nadir*. PSA data were removed accordingly in all datasets of the scenarios 'No link', 'High link' and 'Short survival' (see main analysis), leading to a dramatic decrease in the number of PSA measurements (Table S1).

In order to assess the impact of withdrawal data, two cases were considered:

- The vital status is known at the end of the study (scenarios 'Withdrawal')
- The vital status is censored at the time of the disease progression (scenarios 'Withdrawal + censor')

Thus, in the first case the number of observed deaths is equal to that observed in the main analysis, and in the second case the number of observed deaths is much smaller than observed in the main analysis (scenarios 'No withdrawal'). This results in much larger confidence interval for the Kaplan-Meier curves (Figure S1).

Results

The variability of the parameter estimates increases when patients withdrawal from PSA follow-up (Figures S2 and S3) which was expected because there is a smaller number of PSA measurements. PSA kinetic parameters (Figure S2) were not affected by withdrawal, with or without censor of vital status. However all three parameters related to survival were estimated with a bias in case of censor of survival (Figure S3): β and k tend to be overestimated while λ

tends to be underestimated. These trends led to an overestimation of the hazard function and hence an underestimation of the survival function.

Conclusion

Drop out caused by predefined levels of PSA progression does not affect the estimation of the parameters associated with PSA kinetics.

Regarding survival parameters, a bias towards an overestimation of the hazard function may occur, in particular when not only PSA but also the vital status is not collected after disease progression.

Table S1: Number of PSA measurements per dataset and per patient and mean number of events per dataset in the total number of simulated datasets

	Scenario No link			Scenario High link			Scenario Short survival			
	No withdrawal	Withdrawal	Withdrawal + censor	No withdrawal	Withdrawal	Withdrawal + censor	No withdrawal	Withdrawal	Withdrawal + censor	
Number of PSA measurements per dataset	11 104	3 062	3 062	12102	3488	3488	6 834	2 804	2 804	
Number of PSA measurements per patient	1-5	7%	65%	65%	8%	64%	64%	29%	70%	70%
	6-10	12%	22%	22%	10%	21%	21%	20%	19%	19%
	11-20	26%	7%	7%	22%	7%	7%	26%	7%	7%
	21-35	30%	3%	3%	21%	3%	3%	17%	3%	3%
	36	24%	1%	1%	39%	4%	4%	8%	1%	1%
Mean number of events per dataset	379	379	66	305	305	36	462	462	145	

Figure S1: Spaghetti-plots (black lines) and estimated Kaplan-Meier curves (colored solid lines) with their 95% confidence interval (colored dashed lines) for one typical dataset (N=500) for each scenario, without withdrawal (red curves), with withdrawal after PSA progression (pink curves) and with withdrawal and censor after PSA progression (grey curves).

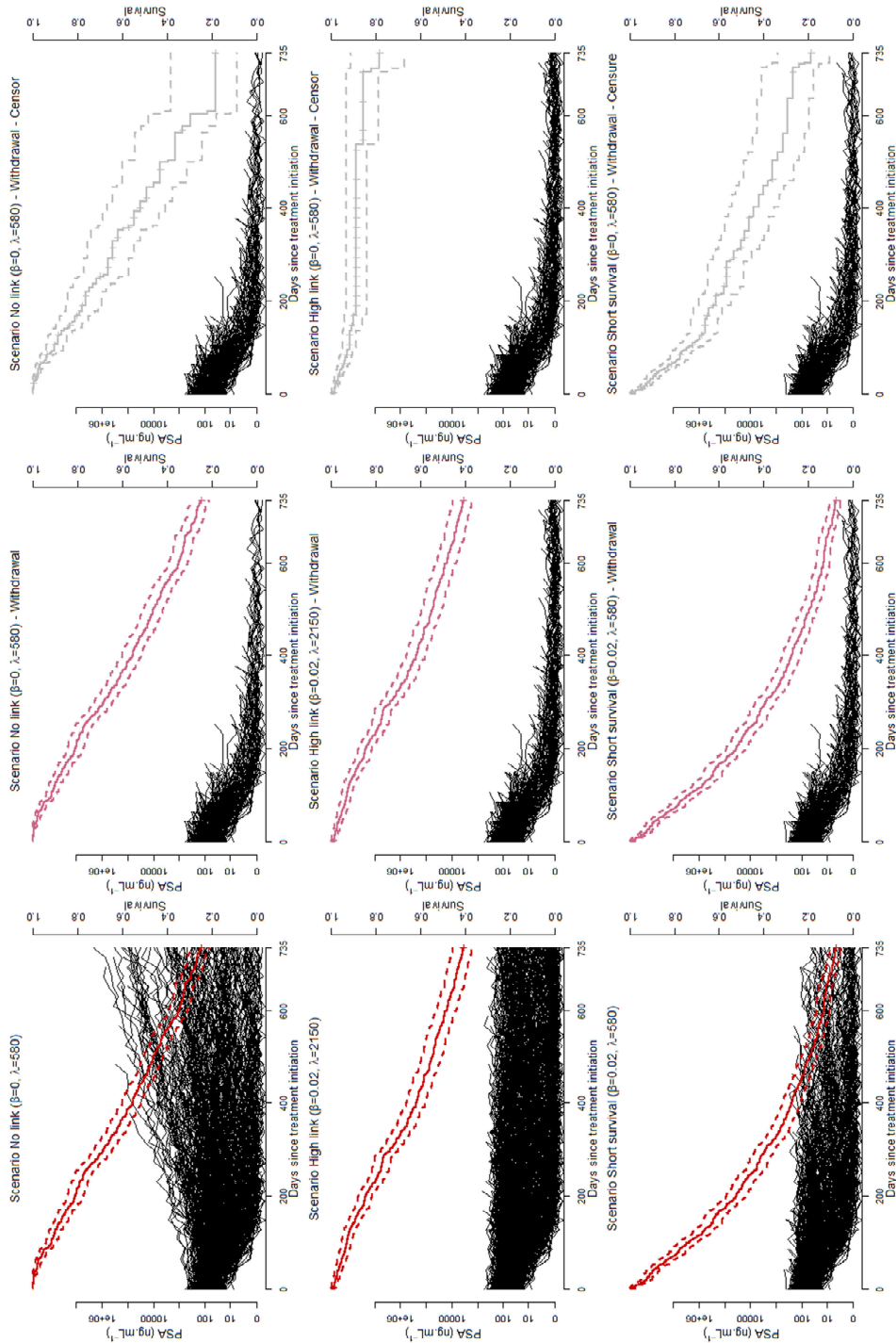


Figure S2: Boxplots of the relative estimation errors for the parameters related to PSA for joint model without withdrawal (red), with withdrawal (pink) and with withdrawal and censor (gray) and for the 3 scenarios. Top: scenario No link ($\beta=0, \lambda=580$), middle: scenario High link ($\beta=0.02, \lambda=2150$) and bottom: scenario Short survival ($\beta=0.02, \lambda=580$).

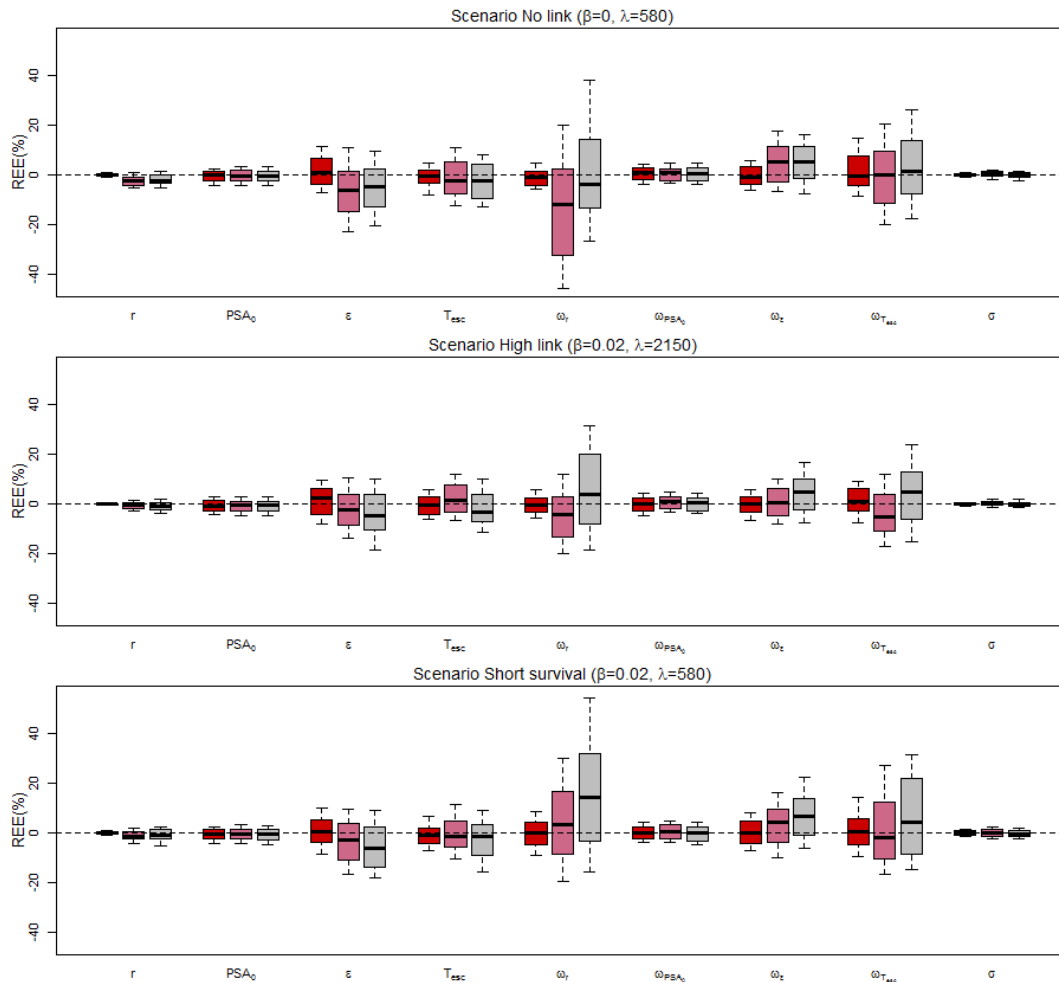
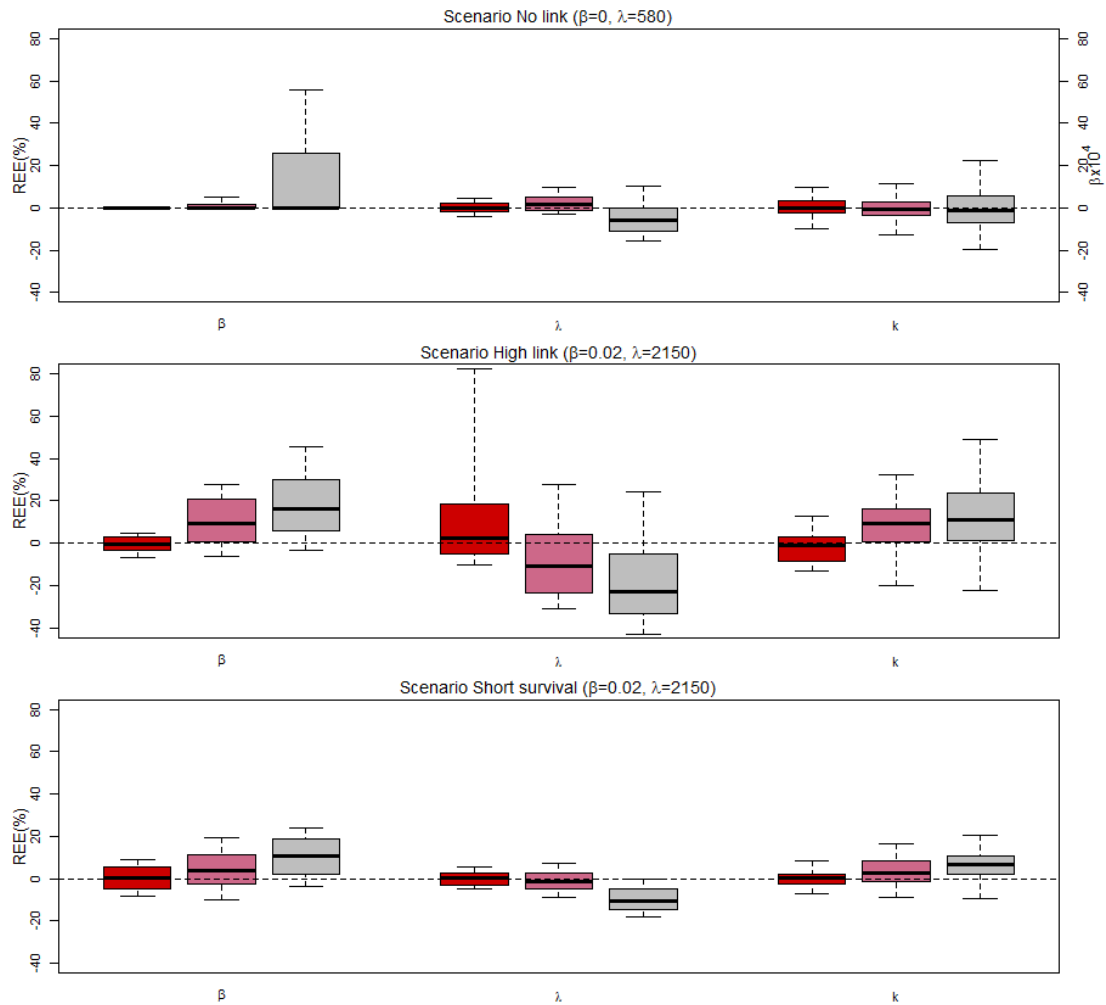


Figure S3: Boxplots of the relative estimation errors for the parameters related to survival (in % except for β of the scenario No link for which estimated values are represented) for joint model without withdrawal (red), with withdrawal (pink) and with withdrawal and censor (gray) and for the 3 scenarios. Top: scenario No link ($\beta=0, \lambda=580$), middle: scenario High link ($\beta=0.02, \lambda=2150$) and bottom: scenario Short survival ($\beta=0.02, \lambda=580$).



Chapitre 3

Développement d'un modèle conjoint mécanistique caractérisant la relation entre la cinétique non-linéaire du PSA et la survie dans le cancer de la prostate métastatique

3.1 Résumé

La précédente étude de simulation a permis d'établir qu'avec l'algorithme SAEM de Monolix, nous disposons d'un outil fiable dans l'estimation conjointe des paramètres longitudinaux et de survie, dans le cadre de MNLEM (DESMÉE et al., 2015), confirmant les résultats de MBOGNING et al. (2015) pour des modèles plus simples. Dans ce second travail, nous avons développé un modèle conjoint non linéaire mécanistique pour les données longitudinales de PSA et de survie de patients ayant un CPmHR appartenant au bras *placebo* de VENICE présentés dans le paragraphe 1.1.5, et donc traités par la chimiothérapie de référence, docetaxel associé à de la prednisone.

L'hypothèse principale pour développer la partie longitudinale de notre modèle repose sur le mécanisme d'action du docetaxel qui engendre de la résistance chez un grand nombre de patients (SERUGA et al., 2011). Ainsi nous avons supposé que le PSA est produit par deux types de cellules : des cellules sensibles au docetaxel et des cellules résistantes au docetaxel. Ces deux types de cellules peuvent être vues comme des compartiments non-observés (figure 1 de l'article 2) qui permettent de décrire toute la cinétique du PSA : l'augmentation du PSA pré-traitement (les cellules productrices de PSA se multiplient, en particulier les cellules sensibles qui sont majoritaires), puis la diminution du PSA à l'initiation du traitement (le docetaxel est supposé inhiber la prolifération des cellules sensibles avec un effet constant ε) et enfin, le rebond du PSA, dû à l'apparition d'une résistance (les cellules résistantes deviennent prédominantes, de part leur avantage évolutif, et se multiplient). Un système de 3 équations différentielles (ODEs)

résulte de ce mécanisme de production du PSA (équation (1) de l'article 2). Un MNLEM (voir section 1.2.2) reposant sur ces ODEs avec une erreur résiduelle constante sur le $\log(PSA+1)$ décrit la partie longitudinale du modèle conjoint. Les données longitudinales de PSA sous la limite de quantification (0.1 ng/mL) ont été considérées comme des données censurées à gauche (SAMSON et al., 2006).

Le temps de décès est modélisé par un modèle de risque proportionnel paramétrique avec comme risque instantané de base une fonction de Weibull. Les covariables disponibles (surface corporelle, âge, race, temps écoulé depuis l'hormono-thérapie) n'ont été introduites dans le modèle de survie que si elles avaient un effet significatif (au seuil de 20%) lors d'une analyse préliminaire univariée. Plusieurs fonctions de lien ont été envisagées pour caractériser la relation entre la cinétique du PSA et la survenue du décès (voir section 1.4.1). La plupart reposaient classiquement sur l'évolution du PSA elle-même (valeur courante du PSA, pente courante du PSA, intégrale du PSA). Mais nous avons également exploité la nature mécanistique du modèle longitudinal, utilisant la cinétique des quantités non-observées que sont le nombre courant de cellules sensibles et le nombre courant de cellules résistantes.

Les paramètres θ des différents modèles conjoints ont été estimés sur les données des 400 patients du jeu de données d'apprentissage avec l'algorithme SAEM. Le modèle fournissant le plus petit BIC a été sélectionné et évalué à l'aide de résidus pour la partie longitudinale (ajustements individuels par les EBEs et analyse des résidus pondérés individuels notés *IWRES*) et pour la partie survie (résidus de Cox-Snell et résidus de Martingale) (voir section 1.4.4). Pour évaluer la prédiction de la survie, nous avons également comparé la courbe de la fonction de survie estimée par Kaplan-Meier à la courbe de survie moyenne, c'est-à-dire la moyenne des fonctions de survie individuelles estimées par le modèle en fixant les paramètres de population à $\hat{\theta}$ et en utilisant uniquement les données longitudinales de PSA pour calculer les estimations bayésiennes empiriques (EBEs) individuelles $\hat{\psi}_i : \hat{S}(t) = \frac{1}{N} \sum_{i=1}^N S_i(t|PSA(t, \hat{\psi}_i); \hat{\theta})$. Cette dernière évaluation a également été menée sur les 196 patients du jeu de données de validation afin d'étudier la capacité du modèle à prédire la survie dans un jeu de données différent.

Finalement, aucune covariable n'a été introduite dans les modèles conjoints. Le modèle conjoint final montre que ce sont le nombre courant de cellules sensibles et le nombre courant de cellules résistantes qui sont associés à la survie. Ce modèle fournissait de bons ajustements individuels de l'évolution du PSA, capturant l'entière cinétique du biomarqueur (avant, sous et après traitement) chez des patients aux profils variés. L'examen des résidus n'a pas révélé d'erreur de spécification du modèle longitudinal. De même, les résidus pour la survie apparaissaient satisfaisants. Enfin, la courbe de survie moyenne s'inscrivait bien dans l'intervalle de confiance à 95% de la fonction de survie estimée par Kaplan-Meier, pour les deux jeux de données.

Pour conclure, ce travail propose le premier modèle conjoint prenant en compte le

mécanisme de production du PSA chez des patients ayant un CPmHR et l'action du docetaxel. Il met en exergue le rôle prédominant des cellules résistantes au docetaxel, non observées, sur la survie, et suggère l'utilité de la modélisation conjointe mécanistique pour améliorer l'évaluation thérapeutique et prédire la survenue d'un évènement à l'aide de mesures longitudinales de biomarqueurs.

Ce travail a fait l'objet d'un article accepté dans *Biometrics* en mars 2016. Le jeu de données d'apprentissage, ainsi que les codes MLXTRAN pour estimer les paramètres du modèle conjoint final avec Monolix sont disponibles en matériel supplémentaire. Certaines figures de l'article peuvent être obtenues directement avec Monolix (les prédictions individuelle et les IWRES), mais d'autres ne peuvent l'être (les résidus de Cox-Snell et de Martingale et la courbe de survie moyenne). Ainsi, nous fournissons un script R permettant de reproduire toutes les figures de l'article 2 afin de faciliter le développement de modèles conjoints mécanistiques et de promouvoir leur utilisation.

3.2 Article 2 (publié)

Using the SAEM Algorithm for Mechanistic Joint Models Characterizing the Relationship between Nonlinear PSA Kinetics and Survival in Prostate Cancer Patients

Solène Desmée,^{1,2,*} France Mentré,^{1,2} Christine Veyrat-Follet,³ Bernard Sébastien,⁴ and
Jérémie Guedj^{1,2}

¹INSERM, IAME, UMR 1137, F-75018 Paris, France

²Université Paris Diderot, IAME, UMR 1137, Sorbonne Paris Cité, F-75018 Paris, France

³Drug Disposition, Disposition Safety and Animal Research Department, Sanofi, Alfortville, France

⁴Biostatistics and Programming, Sanofi, Chilly-Mazarin, France

**email*: solene.desmee@inserm.fr

SUMMARY. Joint modeling is increasingly popular for investigating the relationship between longitudinal and time-to-event data. However, numerical complexity often restricts this approach to linear models for the longitudinal part. Here, we use a novel development of the Stochastic-Approximation Expectation Maximization algorithm that allows joint models defined by nonlinear mixed-effect models. In the context of chemotherapy in metastatic prostate cancer, we show that a variety of patterns for the Prostate Specific Antigen (PSA) kinetics can be captured by using a mechanistic model defined by nonlinear ordinary differential equations. The use of a mechanistic model predicts that biological quantities that cannot be observed, such as treatment-sensitive and treatment-resistant cells, may have a larger impact than PSA value on survival. This suggests that mechanistic joint models could constitute a relevant approach to evaluate the efficacy of treatment and to improve the prediction of survival in patients.

KEY WORDS: Joint model; Metastatic prostate cancer; Nonlinear mixed effect model; Prostate specific antigen; SAEM algorithm; Survival.

1. Introduction

The field of joint modeling, which aims to characterize the relationship between longitudinal biomarkers and a time-to-event, has received a lot of attention from biostatisticians in the last decade (Tsiatis and Davidian, 2004; Rizopoulos, Verbeke, and Lesaffre, 2009; Wu et al., 2011). Moreover, large efforts have been made to reach out beyond the academic community and packages in R or codes and macro in SAS are now available (Guo and Carlin, 2004; Rizopoulos, 2010; Garcia-Hernandez and Rizopoulos, 2015). However, a major limitation of these tools is that they essentially rely on linear models for the longitudinal part (Tsiatis and Davidian, 2004; Rizopoulos, 2012; Asar et al., 2015). Although these models are often sufficient to describe sparse longitudinal data, such as those obtained in epidemiological studies, they may not be appropriate to describe the kinetics of frequently sampled markers such as those obtained in clinical trials for treatment evaluation. In this context, nonlinear mechanistic models, for instance based on ordinary differential equations (ODEs), may be necessary to characterize the dynamic changes in the longitudinal markers but the difficulty to calculate the likelihood in nonlinear mixed effect models (NLMEM) complicates considerably the use of these models. In the last years, several stochastic methods have been developed to propose efficient algorithms for inference in NLMEM (Kuhn and Lavielle, 2005; Plan et al., 2012). One of them is the Stochastic Approxima-

tion Expectation-Maximization (SAEM) algorithm (Delyon, Lavielle, and Moulines, 1999). As in other EM algorithms, the algorithm is an iterative process where each iteration is divided into a step where the complete likelihood conditional on observations is calculated (E-step), and a step where the complete likelihood is maximized (M-step). In addition, in the SAEM algorithm, the E-step is divided into two parts: a simulation of individual parameters using a Markov Chain Monte Carlo (MCMC) algorithm (S-step), and then a calculation of the expected likelihood using a stochastic approximation (A-step). Recently, the SAEM algorithm implemented in Monolix (www.lixoft.eu) has been expanded to the context of joint models and was shown by simulation to provide precise estimates when the longitudinal marker was defined by a NLMEM (Desmée et al., 2015; Mbogning, Bleakley, and Lavielle, 2015).

Here, we applied this approach to characterize, on real data, the relationship between the kinetics of the Prostate-Specific Antigen (PSA) and survival in metastatic Castration-Resistant Prostate Cancer (mCRPC) using data from a phase 3 clinical trial (Tannock et al., 2013). In this context of patients with advanced disease, the incidence of death is high and the PSA kinetics is closely monitored after treatment initiation to rapidly detect a breakthrough in PSA and propose rescue strategies. Thus, unlike studies where long-term PSA levels can be described by linear models (Proust-Lima et al.,

2008) here, the rapid changes in PSA levels require dynamical nonlinear models. Similarly to what is done to characterize the effect of an anti-viral treatment on a pathogen (Perelson and Guedj, 2015), the PSA kinetics can be schematically viewed as a dynamic interaction between the chemotherapy and the process of production and elimination of cancer cells (Seruga, Ocana, and Tannock, 2011). This interaction can naturally be modeled using a system of nonlinear ODEs where both parameters and processes have a biological interpretation. Here, we aimed to show how a mechanistic joint approach can be used to characterize the relationship between the nonlinear kinetics of a biomarker and the time-to-death in the context of treatment evaluation.

The outline is as follows: in Section 2, we introduce the clinical data and the methods used to characterize the relationship between PSA kinetics and survival. In Section 3, we show how to construct and evaluate a joint model in the context of a mCRPC clinical trial. Lastly, in section 4, we conclude by pointing out the advantages and the limits of this approach.

2. Material and Methods

2.1. Description of the Data

We analyzed the data of 598 men with metastatic Castration-Resistant Prostate Cancer (mCRPC) treated with docetaxel and prednisone, the first-line reference chemotherapy, which constituted the control arm of a phase 3 clinical trial (Tannock et al., 2013). In the protocol, PSA had to be measured within 8 days before treatment initiation, every 21 days during treatment, and then every 84 days after treatment. Baseline available covariates were body size, age, race, and time elapsed since hormone therapy. Two patients had no PSA measurement and were not included in the analysis.

For the sake of internal validation, the dataset was randomly split into a training and a validation dataset of 400 and 196 patients, respectively. In the training dataset, 6627 PSA measurements were collected, among which 1385 (20.9%) were pretreatment and 3934 (59.4%) were on treatment. The median [minimum; maximum] duration between the first PSA measurement and the treatment initiation was 104 days [2; 1195]. The median number of measurements per patient was 15 [3; 57] with a median number of pre-, on-, and post-treatment measurements of 4 [1; 14], 9 [0; 48], and 2 [0; 26], respectively. The limit of quantification (LoQ) was 0.1 ng.mL⁻¹ and 165 observations (2.5%) from 21 patients (5.3%) were below LoQ. In the validation dataset, the median numbers of PSA measurements per patient was also equal to 15 and the total number of PSA measurements was 3185. In the training dataset 286 patients deceased (71.5%), leading to a median survival (Kaplan–Meier 95% confidence interval) of 656 days [598; 741]. In the validation dataset, 145 patients deceased (74.0%), leading to a median survival of 598 days [547; 732]. Unlike it is stated otherwise, all results given below are obtained using the training dataset only.

2.2. Mechanistic PSA Kinetic Model

In this model (Figure 1), PSA is produced by two types of cells, namely treatment-sensitive cells (S) and treatment-resistant cells (R) (Seruga et al., 2011). In absence of

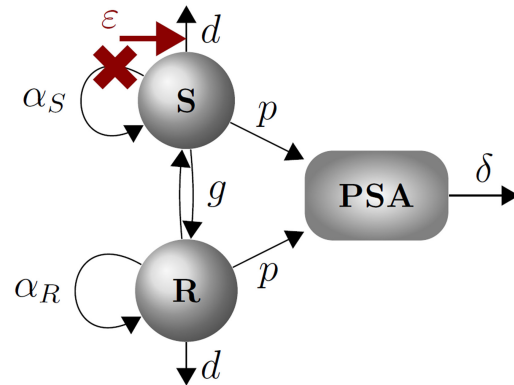


Figure 1. Schema of the secretion of PSA by sensitive (S) and resistant (R) cells. PSA is expressed in ng.mL⁻¹ and S and R in mL⁻¹. In absence of treatment, α_S and α_R are the rates of S and R proliferation (day⁻¹), respectively, d the rate of S and R elimination (day⁻¹), g is the mutation rate (day⁻¹), p the rate of PSA secretion by S and R (ng.day⁻¹), and δ the rate of PSA elimination (day⁻¹). Treatment can inhibit S cells proliferation (big cross) or stimulate S cells elimination (big arrow) with the constant effectiveness ϵ . This figure appears in color in the electronic version of this article.

treatment, sensitive and resistant cells proliferate with rates α_S and α_R , respectively, and are eliminated with a similar rate d . The total number of sensitive and resistant cells is limited by a saturation term, noted N_{\max} . Mutations from S to R and from R to S occur with an identical rate, noted g . PSA is secreted by both treatment-sensitive and -resistant cells with a rate p and is cleared from the blood with a rate δ .

The associated system of ordinary differential equations (ODEs) is as follows:

$$\begin{cases} \frac{dS}{dt} = \alpha_S \left(1 - \frac{S+R}{N_{\max}}\right) S + g(R-S) - dS \\ \frac{dR}{dt} = \alpha_R \left(1 - \frac{S+R}{N_{\max}}\right) R + g(S-R) - dR \\ \frac{dPSA}{dt} = pS + pR - \delta PSA. \end{cases} \quad (1)$$

Let PSA_b , S_b , and R_b be the initial values of PSA, sensitive, and resistant cell counts, respectively, that is, their values at first PSA measurement. In order to determine their values, we used the fact that by definition the treatment-resistant cells are less fit to grow than the treatment-sensitive cells ($RF = \alpha_R/\alpha_S < 1$). Therefore, treatment-sensitive cells should be largely predominant at baseline ($S_b \gg R_b$) and thus quasi-steady state approximations can be made to derive baseline conditions:

$$S_b = \frac{\delta}{p} PSA_b.$$

Then, one can obtain after few calculations on equations (1):

$$R_b = \frac{g}{d - RF \times (g + d)} \times \frac{\delta}{p} PSA_b.$$

Time $t = 0$ indicates the beginning of treatment that has a constant and non-null effectiveness against the treatment-sensitive cells ($\varepsilon > 0$) while it has no efficacy against resistant cells. Two mechanisms of actions for docetaxel were considered (Herbst and Khuri, 2003; Petrylak, 2005): it can either inhibit angiogenesis (i.e., decreases the cancer cells proliferation from α_S before treatment initiation to $\alpha_S \times (1 - \varepsilon)$ afterward) or increase cell apoptosis (i.e., stimulates the cancer cells elimination from d before treatment initiation to $d \times (1 + \varepsilon)$ afterward).

The following reparameterizations were done to improve the model identifiability. First, since before treatment initiation PSA increased, the proliferation rate of sensitive cells is necessarily larger than the apoptosis rate ($\alpha_S > d$). Second, after treatment initiation, no patient achieved a sustained low level of PSA, which means that the proliferation rate of resistant cells is also larger than the apoptosis rate ($\alpha_R > d$). Thus, $\alpha_S > \alpha_R > d$ and we parameterized d as $d = RE \times RF \times \alpha_S$ with $0 < RE < 1$. For the sake of parameter identifiability, we fixed δ to 0.23 day^{-1} , corresponding to a PSA half-life in blood of about 3 days (Ruffion, Rebillard, and Grima, 2005) while p and g were determined by a sensitivity analysis (see Web Appendix A).

Finally, the mathematical model for PSA kinetics was defined by the vector parameter: $\boldsymbol{\mu} = (\alpha_S, RF, RE, \varepsilon, PSA_b, N_{\max})$.

2.3. Statistical Model for PSA Measurements

Nonlinear mixed-effect models (NLMEM) were used to analyze all the longitudinal PSA measurements (before, during, and after treatment).

Let N be the number of patients and $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})$ the vector of observations in patient i , where y_{ij} is the observed Naperian logarithm of PSA + 1 for the patient i , $i = 1, \dots, N$, at time t_{ij} , $j = 1, \dots, n_i$. A constant error model is assumed on the logarithm of PSA + 1:

$$y_{ij} = \log(PSA(t_{ij}, \boldsymbol{\psi}_i) + 1) + e_{ij}, \quad (2)$$

where $\boldsymbol{\psi}_i$ is the vector of the individual parameters, $PSA(t_{ij}, \boldsymbol{\psi}_i)$ is given by the system of ODEs (1) and e_{ij} is the residual Gaussian error of mean 0 and variance σ^2 . $\boldsymbol{\psi}_i$ is decomposed as a vector of fixed effects $\boldsymbol{\mu}$ representing median parameters of the population and random effects $\boldsymbol{\eta}_i$ specific for each individual. It is assumed that $\boldsymbol{\eta}_i \sim \mathcal{N}(0, \boldsymbol{\Omega}^2)$ with $\boldsymbol{\Omega} = \text{diag}(\omega_{\alpha_S}, \omega_{RF}, \omega_{RE}, \omega_{\varepsilon}, \omega_{PSA_b}, \omega_{N_{\max}})$. We assumed log-normal distribution for α_S , PSA_b , and N_{\max} and logit-normal distribution for RF and RE . Regarding ε , we used a logit-normal distribution when an effect in blocking cell proliferation was evaluated and a log-normal distribution when an effect of treatment in enhancing cell death was evaluated (see Section 2.2).

2.4. Characterization of the Relationship between PSA Kinetics and Survival

Let X_i and C_i denote the time-to-death and the censoring time, respectively, for patient i . $T_i = \min(X_i, C_i)$ and $\delta_i = \mathbb{1}_{\{X_i \leq C_i\}}$ are observed. The time-to-death was modeled using a parametric risk-proportional model assuming a Weibull func-

tion for the baseline hazard function: for $t \geq 0$,

$$h_i(t|\mathbf{PSA}(t, \boldsymbol{\psi}_i)) = h_0(t) \exp(\boldsymbol{\gamma}^T \mathbf{w}_i + \boldsymbol{\beta}^T f(t, \boldsymbol{\psi}_i)), \quad (3)$$

where $\mathbf{PSA}(t, \boldsymbol{\psi}_i) = \{PSA(s, \boldsymbol{\psi}_i); 0 \leq s < t\}$ denotes the history of the true unobserved longitudinal process up to t , h_0 is the Weibull baseline hazard function $h_0(t) = \frac{k}{\lambda} (\frac{t}{\lambda})^{k-1}$, $\boldsymbol{\gamma}$ is the vector of coefficients associated with the vector of baseline covariates \mathbf{w}_i and $\boldsymbol{\beta}$ is the vector of coefficients associated with the ODE model outputs $f(t, \boldsymbol{\psi}_i)$ (equation 1).

The following models for the relationship between survival and PSA kinetics were considered:

- No link: $f(t, \boldsymbol{\psi}_i) = 0$;
- Initial PSA: $f(t, \boldsymbol{\psi}_i) = \log(PSA(0, \boldsymbol{\psi}_i) + 1)$;
- Current PSA: $f(t, \boldsymbol{\psi}_i) = \log(PSA(t, \boldsymbol{\psi}_i) + 1)$;
- Current PSA slope: $f(t, \boldsymbol{\psi}_i) = \frac{d \log(PSA(t, \boldsymbol{\psi}_i) + 1)}{dt}$;
- Area under PSA: $f(t, \boldsymbol{\psi}_i) = \int_0^t \log(PSA(u, \boldsymbol{\psi}_i) + 1) du$;
- Current S: $f(t, \boldsymbol{\psi}_i) = \log(S(t, \boldsymbol{\psi}_i))$;
- Current R: $f(t, \boldsymbol{\psi}_i) = \log(R(t, \boldsymbol{\psi}_i))$;
- Current S and R: $f(t, \boldsymbol{\psi}_i) = (\log(S(t, \boldsymbol{\psi}_i)), \log(R(t, \boldsymbol{\psi}_i)))$. In this model $\boldsymbol{\beta}$ has two components, $\boldsymbol{\beta} = (\boldsymbol{\beta}, \boldsymbol{\beta}')$.

Of note the last three models exploit the mechanistic nature of the model as they relate to quantities that are not observed.

2.5. Maximum-Likelihood Estimation

The log-likelihood for subject i is given by the following:

$$LL_i(\boldsymbol{\theta}) = \log \int p(\mathbf{y}_i | \boldsymbol{\eta}_i; \boldsymbol{\theta}) p(T_i, \delta_i | \boldsymbol{\eta}_i; \boldsymbol{\theta}) p(\boldsymbol{\eta}_i; \boldsymbol{\theta}) d\boldsymbol{\eta}_i, \quad (4)$$

where $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\Omega}, \sigma, \lambda, k, \boldsymbol{\gamma}, \boldsymbol{\beta})$ is the vector of parameters to estimate, $p(\mathbf{y}_i | \boldsymbol{\eta}_i; \boldsymbol{\theta})$ is the probability density function of the longitudinal observations conditionally on the random effects $\boldsymbol{\eta}_i$, $p(T_i, \delta_i | \boldsymbol{\eta}_i; \boldsymbol{\theta}) = h_i(T_i | \mathbf{PSA}(T_i, \boldsymbol{\psi}_i); \boldsymbol{\theta})^{\delta_i} S_i(T_i | \mathbf{PSA}(T_i, \boldsymbol{\psi}_i); \boldsymbol{\theta})$ is the likelihood of the survival part, with $S_i(t | \mathbf{PSA}(t, \boldsymbol{\psi}_i); \boldsymbol{\theta}) = \exp(-\int_0^t h_i(s | \mathbf{PSA}(s, \boldsymbol{\psi}_i); \boldsymbol{\theta}) ds)$ and $p(\boldsymbol{\eta}_i; \boldsymbol{\theta})$ is the probability density function of the random effects.

The likelihood was maximized using the SAEM algorithm (Delyon et al., 1999; Kuhn and Lavielle, 2005) implemented in Monolix version 4.3.2. The likelihood was estimated by Importance Sampling with a Monte-Carlo size of 200,000 to ensure efficient precision. SAEM was run using one chain after checking that using three chains gave similar results (Vigan, Stirnemann, and Mentré, 2014). All calculations were performed in a i7 64bits 3.33 GHz.

2.6. Model Selection and Evaluation

We first selected the model for the longitudinal data using only longitudinal data. For each couple of parameters (p, g) explored in the sensitivity analysis (Web Appendix A), the two possible mechanisms of action of treatment were tested and only the model giving the lowest Bayesian Information Criterion (BIC) was retained.

Next, the selection of the joint model was based on the BIC (Park and Qiu, 2014). Of note baseline covariates were tested in the joint model only if they were significant in univariate analysis using a Weibull survival model and a likelihood

Table 1

Goodness-of-fit statistics (standard error) and parameters estimates (relative standard error (%)) of PSA kinetics and survival in the 400 patients of the training dataset for the 8 joint models.

Models	No link	Initial PSA	Current PSA	PSA slope	Area under PSA	Current S	Current R	Current S and R
-2LL	14508 (7)	14486 (7)	14350 (7)	14485 (7)	14479 (7)	14471 (7)	14334 (7)	14326 (8)
BIC	14598 (7)	14582 (7)	14446 (7)	14581 (7)	14575 (7)	14567 (7)	14430 (7)	14421 (8)
α_S (day ⁻¹)	0.066 (3)	0.060 (3)	0.078 (3)	0.078 (3)	0.061 (3)	0.062 (3)	0.068 (3)	0.067 (3)
RF	0.9997 (0)	0.9996 (0)	0.9998 (0)	0.9998 (0)	0.9997 (0)	0.9997 (0)	0.9998 (0)	0.9998 (0)
RE	0.81 (1)	0.79 (1)	0.84 (1)	0.84 (0)	0.79 (1)	0.81 (1)	0.81 (1)	0.82 (1)
ε	0.42 (4)	0.46 (4)	0.35 (4)	0.35 (5)	0.47 (4)	0.45 (4)	0.43 (3)	0.43 (3)
PSA_b (ng.mL ⁻¹)	22.2 (8)	22.2 (8)	22.0 (8)	22.5 (8)	22.2 (8)	22.3 (8)	21.7 (8)	21.9 (8)
N_{max} (mL ⁻¹)	56 (4)	57 (4)	81 (4)	77 (4)	57 (4)	71 (4)	65 (4)	120 (4)
λ (day)	885 (4)	1615 (8)	4259 (15)	920 (4)	1435 (7)	675 (5)	877 (6)	906 (7)
k	1.52 (5)	1.53 (3)	1.28 (2)	1.48 (2)	1.19 (2)	1.73 (5)	1.01 (0)	1 -
β	-	0.21 (12)	0.40 (7)	17 (17)	0.00023 (8)	0.067 (15)	0.38 (6)	0.00032 (21)
β'	-	-	-	-	-	-	-	0.39 (7)
ω_{α_S}	0.48 (5)	0.53 (5)	0.49 (5)	0.45 (6)	0.52 (5)	0.49 (4)	0.50 (5)	0.48 (5)
ω_{RF}	2.68 (4)	2.72 (4)	2.55 (4)	2.51 (4)	2.66 (4)	2.59 (4)	2.6 (4)	2.55 (4)
ω_{RE}	0.63 (4)	0.64 (4)	0.58 (5)	0.60 (5)	0.67 (4)	0.66 (4)	0.64 (5)	0.59 (5)
ω_{ε}	1.14 (6)	1.21 (6)	1.10 (6)	1.04 (7)	1.15 (6)	1.21 (6)	0.90 (6)	0.95 (6)
ω_{PSA_b}	1.63 (4)	1.63 (4)	1.66 (4)	1.62 (4)	1.62 (4)	1.63 (4)	1.64 (4)	1.64 (4)
$\omega_{N_{max}}$	2.21 (6)	2.22 (6)	2.30 (6)	2.14 (6)	2.17 (6)	2.22 (6)	2.27 (6)	2.57 (7)
σ	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)

ratio test with a significance level of 20%. Then, the model evaluation relied on the analysis of the model individual predictions and residuals in the training dataset, as well as the analysis of the survival predictions in the validation dataset. The residuals for the longitudinal part were assessed using the Individual Weighted Residuals, noted IWRES, defined by $IWRES_{ij} = \frac{y_{ij} - \log(PSA(t_j, \hat{\psi}_i))}{\hat{\sigma}}$ where $\hat{\sigma}$ is the estimated standard deviation of the residual error and $\hat{\psi}_i$ is the vector of the estimated individual parameters, that is, the Empirical Bayes Estimates (EBEs) defined as the mode of the conditional distribution $p(\psi_i | y_i; \theta)$ with θ the estimation of the population parameters θ . IWRES were plotted versus time and versus predicted PSA values.

Because it is difficult to interpret the individual predicted hazard functions, the evaluation of the survival part of the model essentially relied on the residuals. Cox-Snell and Martingale residuals, noted r_i^{CS} and r_i^M , respectively, are defined by $r_i^{CS} = \int_0^{T_i} h_i(s | \mathbf{PSA}(T_i, \hat{\psi}_i)) ds$ and $r_i^M = \delta_i - r_i^{CS}$, $i = 1, \dots, n$. The Kaplan-Meier curve for the r^{CS} was compared to the survival curve of the unit exponential distribution $\exp(-t)$, while the Martingale residuals were displayed versus the predicted values inferred from the PSA kinetics and introduced in the final joint model (Rizopoulos, 2012; Sène, Bellera, and Proust-Lima, 2014). In addition to the residuals and in order to evaluate the overall prediction for the survival, the mean survival curve, defined by $S(t) = \frac{1}{N} \sum_{i=1}^N S_i(t | \mathbf{PSA}(t, \hat{\psi}_i); \hat{\theta})$ was calculated and compared to the Kaplan-Meier curve.

In order to evaluate the ability of the model to predict the survival in a different dataset, the mean survival curve was also calculated in the validation dataset. For that purpose, population parameters were fixed to the values found in the training dataset (e.g., $\hat{\theta}$) and individual parameters ψ_i were

estimated from the EBEs. Of note the time-to-death in the validation dataset was not used to estimate the mean survival curve of this dataset.

Lastly, the parametric assumption for the baseline hazard function was relaxed and spline functions for the baseline hazard h_0 , namely piecewise constant, linear, and restricted cubic, were tested. More details can be found in the Web Appendix B.

3. Results

3.1. Model Selection

A model assuming an effect of docetaxel on cell angiogenesis systematically provided a better fit to the longitudinal data than a model assuming an effect on cell apoptosis. Therefore, only the results of the model with an effect on cell angiogenesis will be discussed below. The values for the PSA production rate and for the mutation rate, noted p and g , respectively, were fixed to 20 ng.day⁻¹ and 10⁻⁷ day⁻¹ after the sensitivity analysis study (see Web Appendix A). None of the baseline covariates had a p-value lower than 0.2 in univariate analysis and therefore none of them were tested in the joint model ($\gamma = 0$).

Parameter estimates obtained with the eight candidate joint models are summarized in Table 1. PSA kinetic parameters were largely insensitive to the choice of the survival part of the model. In all cases, they were precisely estimated with relative standard error smaller than 8% for both fixed effects and variance components. In particular, the treatment effect in blocking angiogenesis, ε , was consistently estimated to about 43% ($p < 10^{-15}$ by likelihood ratio test) and the fitness of resistant cells was close to that of sensitive cells ($RF = 99.98\%$). The model using the current PSA value

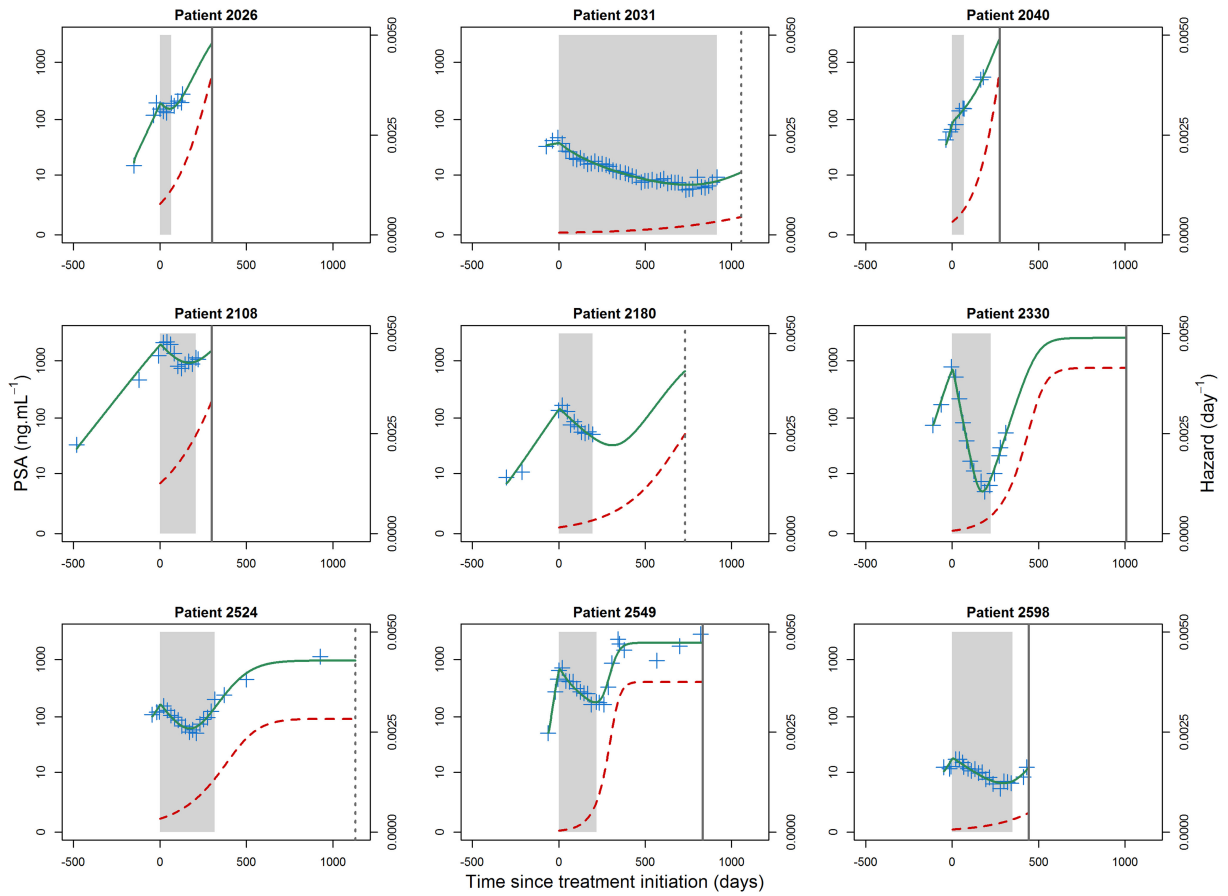


Figure 2. Individual fits of PSA kinetics and hazard function in patients with various PSA profiles. Crosses denote PSA measurements and solid lines are the model predictions using the final joint model. The vertical lines indicate the vital status (solid = death, dotted = censor) at the last recorded time and dashed curves are the individual hazard functions predicted by the final joint model. The gray area indicates the treatment period. This figure appears in color in the electronic version of this article.

outperformed all models relying on PSA in terms of BIC. Surprisingly the models based on the current PSA slope and on the cumulative PSA (area under PSA), did not lead to a large improvement in the BIC compared to the No Link model, that is, the parametric Weibull model with no effect of PSA on survival.

One of the advantages of mechanistic models is that unobserved quantities can also be evaluated. Interestingly a model relying on the current number of resistant cells, R , led to a BIC smaller to what was obtained with the current PSA model. Further in the model relying on both the current number of treatment-sensitive and treatment-resistant cells (Current S and R), the shape parameter, k , was not significantly different from 1, suggesting that the inclusion of R and S could capture most of the time-dependent change in the hazard function. This model led to the lowest BIC and therefore was retained as the final joint model for evaluation since no further combination was found to improve the BIC. Using this model and a Monte Carlo size of 200,000, the CPU times for parameter

estimation and likelihood estimation were 8.72×10^3 seconds and 2.13×10^5 seconds, respectively.

Further, we studied the effect of relaxing the assumption of parametric baseline hazard function. However, there was no reduction in BIC using piecewise constant, linear or restricted cubic splines on the final model (see Web Appendix B).

3.2. Model Evaluation

As shown in Figure 2, various kinetic patterns of PSA before, during, and after treatment could be well captured by the model. The analysis of the IWRES did not suggest any model misspecification (Figure 3) in the fitting across time or PSA values.

For the survival part, the Kaplan–Meier curve of the Cox–Snell residuals (Figure 3) was close to the theoretical survival curve (unit exponential distribution). However, there was a slight overestimation of the survival for small values of the residuals, indicating a possible underestimation of the risk of death in the patients with a small risk of death. Martingale

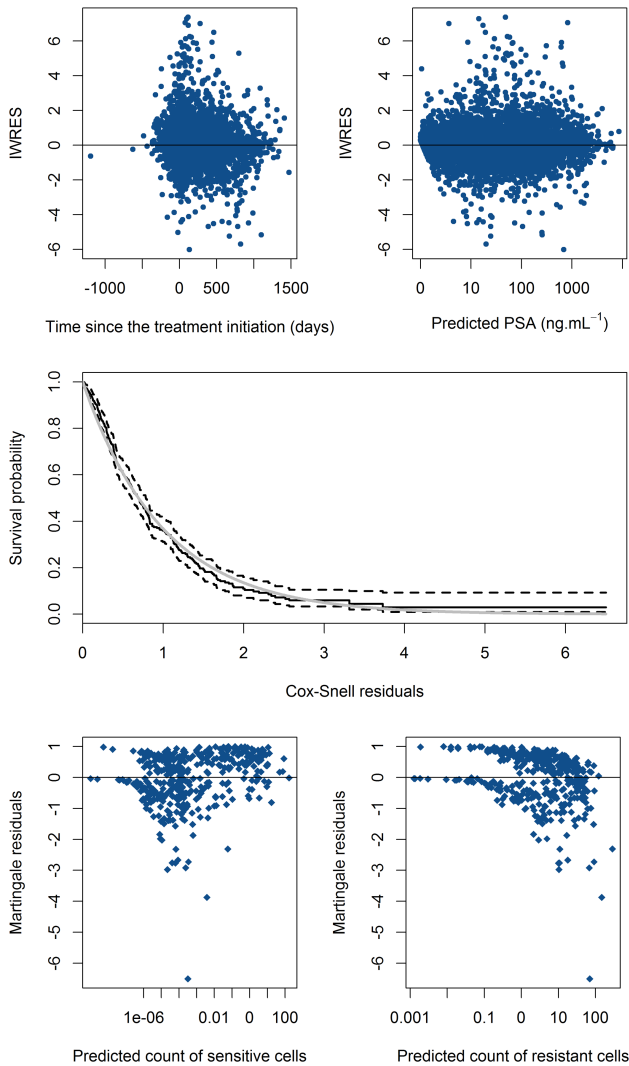


Figure 3. Residuals for the longitudinal and survival parts for the 400 patients included in the training dataset: *Top*: Individual weighted residuals (IWRES) of the longitudinal PSA model versus time (*left*) and versus predicted PSA values (*right*). *Middle*: Kaplan–Meier estimates of the Cox–Snell residuals (solid black line) and its confidence interval (dashed black lines) and the survival function of the unit exponential distribution (solid gray line). *Bottom*: Martingale residuals versus predicted count of sensitive cells (*left*) and versus predicted count of resistant cells (*right*). This figure appears in color in the electronic version of this article.

residuals (Figure 3) were satisfactory according to the number of sensitive cells but tended to underestimate the survival in patients with large number of resistant cells. Concerning the survival prediction, the mean survival curve was close to the Kaplan–Meier curve (Figure 4). In order to evaluate the capability of the model to capture the fact that PSA value at treatment initiation is associated with survival, we also com-

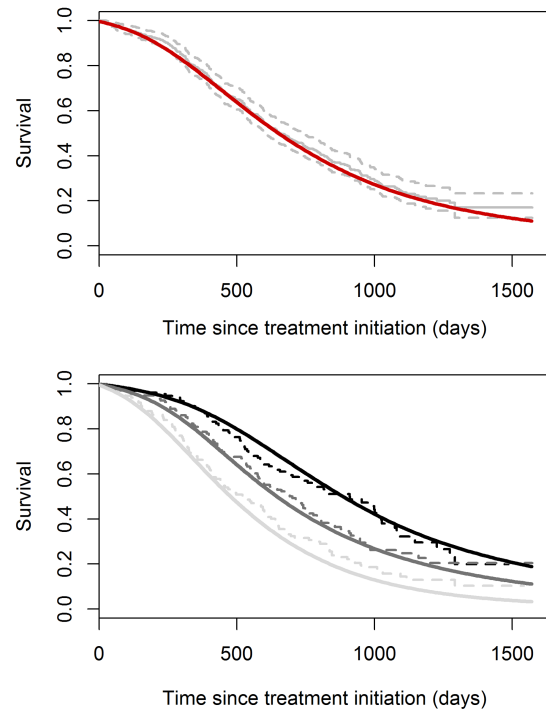


Figure 4. For the 400 patients included in the training dataset: *Top*: Estimated Kaplan–Meier survival curve (gray solid line) and its confidence interval (gray-dashed line) and mean survival curve (solid line). *Bottom*: Estimated Kaplan–Meier survival curves (dotted lines) and mean survival curves (solid lines) stratified by observed PSA values at treatment initiation (black: $PSA_0 \leq 40$, dark gray: $40 < PSA_0 \leq 140$, and light gray: $PSA_0 \geq 140$). This figure appears in color in the electronic version of this article.

pared the mean survival curves and the Kaplan–Meier curves according to the initial PSA value, PSA_0 . Interestingly the model still well captured the survival function in the patients with $PSA_0 \leq 40$ or $40 < PSA_0 \leq 140$. However, we note that the model tended to underestimate the survival in patients with $PSA_0 \geq 140$, that is, patients with very high baseline PSA values, but the mean survival curve remained included in the Kaplan–Meier 95% confidence interval.

Finally, we applied the joint model on the 196 mCRPC patients from the validation dataset. For that purpose, we fixed the population parameters to the values found in the training dataset and we calculated the EBEs for each patient (see Section 2.6). In other words, the individual trajectories for both PSA and the hazard function in a patient were calculated using only the observed PSA measurements of that patient and the information on the vital status in this dataset was not used. As can be seen in Figure 5, the Kaplan–Meier curve in the training and in the validation dataset did not exactly overlay and there was a 58 days difference in the median survival between the two datasets. In spite of this discrepancy, the mean survival curve well fitted the Kaplan–Meier curve of the validation dataset, showing that

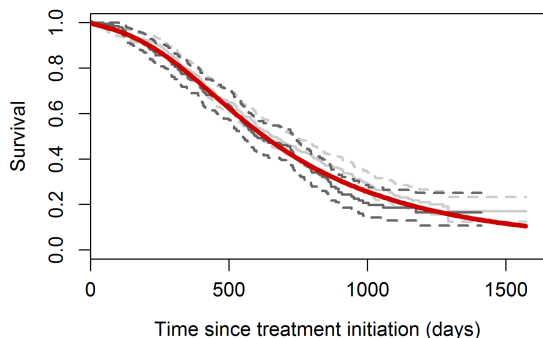


Figure 5. Estimated Kaplan–Meier survival curves (thin solid lines) and their confidence intervals (dashed lines) for the 196 patients included in the validation dataset (dark gray) and for the 400 patients included in the training dataset (light gray) and mean survival curve for the 196 patients included in the validation dataset (thick solid line). This figure appears in color in the electronic version of this article.

the inclusion of the individual PSA kinetics was sufficient to correctly predict the survival in this dataset.

All data and codes needed to reproduce the results, evaluate the residuals, and generate the figures are available as Supplementary Materials with a README file.

4. Discussion

A mechanistic joint model was proposed to study the relationship between survival and longitudinal data described by a system of nonlinear ODEs. Unlike what is done in virology and bacteriology, the field of mechanistic model for treatment evaluation in oncology is still in its infancy. Here, following the efforts of other groups that aimed to develop cancer models that can be used for clinical purposes (Ribba et al., 2014), we showed that mechanistic joint models could be used for prediction of time-to-event in clinical trials. Our model allowed to capture a variety of patterns in PSA kinetics observed before, during, and after chemotherapy. Further the model predicted that two quantities, namely docetaxel-sensitive and resistant cancer cells, may have a large impact on survival. Of note these two quantities were not observed, and the only observed quantity was the PSA value, which was assumed to be proportional to the total number of cancer cells. The kinetics of docetaxel-resistant cells had a larger impact on survival than the kinetics of treatment-sensitive cells. Since the kinetics of resistant cells drives the increase in PSA levels on the long-run, our prediction is consistent and expands the observation that the final tumor growth rate is highly predictive of the time-to-death (Stein et al., 2008). The relevance of the model was further reinforced by the fact that it could be used to predict the survival curve of the 196 patients that had not been used for model building. By including only the information on PSA and ignoring the information on survival in these patients, the model prediction well matched with the Kaplan–Meier curve observed in these patients.

One of the main advantages of mechanistic joint model is that parameters have a biological interpretation. Therefore, putative scenarios can be performed to anticipate the impact

of changes in experimental settings. Here, for instance, one could evaluate by simulation the effect of a drug that could affect the resistant cells, either by blocking the proliferation of resistant cells or by having a complementary mechanism of action. Further, this approach could be useful to guide and design cancer clinical trials by using model predictions (such as here S and R) as surrogate markers and thus possibly reduce the follow-up time or the number of patients needed to assess treatment efficacy or to compare different treatment strategies. However, this will require to pursue the development of these models and to integrate data that are rarely collected or analyzed longitudinally such as the count of circulating tumor cells (CTC) (Wilbaux et al., 2015), the size of the tumors and/or drug pharmacokinetics. ODE models are particularly suitable for that purpose since correlations between the biomarkers can be naturally taken into account through their mechanistic interactions. Obviously, these models will need to be validated internally but also externally, which was not done here, to assess their predictive ability in a different context.

However, the complexity of these models comes with a cost. In terms of feasibility, these models require frequent sampling longitudinal measurements to characterize the kinetics. This limits their use in the context of epidemiological studies where usually large cohorts of patients are followed with sparse measurements. Although SAEM algorithm, unlike other methods such as adaptive Gaussian quadratures (Guedj, Thiébaud, and Commenges, 2011; Prague et al., 2013), can deal with a large number of random effects (Lavielle et al., 2011), it remains time-consuming. Further, in our case and in spite of this computation time, the calculation of the likelihood was still associated with a non-negligible standard errors, making the use of these models not reliable to identify small effects. Lastly, it remains to be demonstrated that such mechanistic joint models can be used in practice for dynamic predictions, that is, to improve the prediction of an individual patient followed prospectively (Rizopoulos, 2011).

In conclusion, we used a novel feature of SAEM algorithm of Monolix that allows joint modeling with a longitudinal process described by ODEs to develop a mechanistic joint model for PSA kinetics and time-to-death. This model sheds a new light on the relationship between PSA and death in mCRPC patients and this approach opens the way for the use of more complex and physiological models to improve treatment evaluation and prediction.

5. Supplementary Materials

Web Appendices referenced in Sections 2 and 3, as well as training dataset, MLXTRAN code, R codes, and a README file are available with this article at the *Biometrics* website on Wiley Online Library.

ACKNOWLEDGEMENTS

The authors would like to thank the Drug Disposition Department, Sanofi, Paris, which supported Solène Desmée by a research grant during this work. They also thank Guilhem Darche for the data management and Hervé Le Nagard for the use of the computer cluster services hosted on the “Centre de biomodélisation UMR1137.”

REFERENCES

- Asar, Ö., Ritchie, J., Kalra, P. A., and Diggle, P. J. (2015). Joint modelling of repeated measurement and time-to-event data: An introductory tutorial. *International Journal of Epidemiology* **44**, 334–344.
- Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a stochastic approximation version of the em algorithm. *Annals of Statistics* **27**, 94–128.
- 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**, 691–699.
- Garcia-Hernandez, A. and Rizopoulos, D. (2015). %JM Macro v2.01 Reference Manual. Available at: www.jm-macro.com
- 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**, 59–66.
- Guo, X. and Carlin, B. P. (2004). Separate and joint modeling of longitudinal and event time data using standard computer packages. *The American Statistician* **58**, 16–24.
- Herbst, R. S. and Khuri, F. R. (2003). Mode of action of docetaxel a basis for combination with novel anticancer agents. *Cancer Treatment Reviews* **29**, 407–415.
- Kuhn, E. and Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis* **49**, 1020–1038.
- Lavielle, M., Samson, A., Karina Fermin, A., and Mentré, F. (2011). Maximum likelihood estimation of long-term HIV dynamic models and antiviral response. *Biometrics* **67**, 250–259.
- 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**, 1512–1528.
- Park, K. Y. and Qiu, P. (2014). Model selection and diagnostics for joint modeling of survival and longitudinal data with crossing hazard rate functions. *Statistics in Medicine* **33**, 4532–4546.
- Perelson, A. S. and Guedj, J. (2015). Modelling hepatitis c therapy—predicting effects of treatment. *Nature Reviews Gastroenterology & Hepatology* **12**, 437–445.
- Petrylak, D. (2005). Therapeutic options in androgen-independent prostate cancer: Building on docetaxel. *BJU International* **96**, 41–46.
- Plan, E. L., Maloney, A., Mentré, F., Karlsson, M. O., and Bertrand, J. (2012). Performance comparison of various maximum likelihood nonlinear mixed-effects estimation methods for dose-response models. *The AAPS journal* **14**, 420–432.
- Prague, M., Commenges, D., Guedj, J., Drylewicz, J., and Thiébaud, R. (2013). Nimrod: A program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations. *Computer Methods and Programs in Biomedicine* **111**, 447–458.
- Proust-Lima, C., Taylor, J. M. G., Williams, S. G., Ankerst, D. P., Liu, N., Kestin, L. L., et al. (2008). Determinants of change in prostate-specific antigen over time and its association with recurrence after external beam radiation therapy for prostate cancer in five large cohorts. *International Journal of Radiation Oncology, Biology, Physics* **72**, 782–791.
- Ribba, B., Holford, N. H., Magni, P., Trocóniz, I., Gueorguieva, I., Girard, P., et al. (2014). A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis. *CPT: Pharmacometrics & Systems Pharmacology* **3**, e113.
- Rizopoulos, D. (2010). JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software* **35**, 1–33.
- Rizopoulos, D. (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* **67**, 819–829.
- Rizopoulos, D. (2012). *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*. Boca Raton: Chapman & Hall/CRC.
- 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**, 637–654.
- Ruffion, A., Rebillard, X., and Grima, F. (2005). Psa doubling time and method of calculation. *Progres en Urologie: Journal de l'Association Francaise d'Urologie et de la Societe Francaise d'Urologie* **15**, 1035–1041.
- Sène, M., Bellera, C. A., and Proust-Lima, C. (2014). Shared random-effect models for the joint analysis of longitudinal and time-to-event data: Application to the prediction of prostate cancer recurrence. *Journal de la Société Française de Statistique* **155**, 134–155.
- Seruga, B., Ocana, A., and Tannock, I. F. (2011). Drug resistance in metastatic castration-resistant prostate cancer. *Nature Reviews. Clinical Oncology* **8**, 12–23.
- Stein, W. D., Figg, W. D., Dahut, W., Stein, A. D., Hoshen, M. B., Price, D., et al. (2008). Tumor growth rates derived from data for patients in a clinical trial correlate strongly with patient survival: A novel strategy for evaluation of clinical trial data. *The Oncologist* **13**, 1046–1054.
- Tannock, I. F., Fizazi, K., Ivanov, S., Karlsson, C. T., Fléchon, A., Skoneczna, I., et al. (2013). Afibercept versus placebo in combination with docetaxel and prednisone for treatment of men with metastatic castration-resistant prostate cancer (Venice): A phase 3, double-blind randomised trial. *The lancet Oncology* **14**, 760–768.
- Tsiatis, A. A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: An overview. *Statistica Sinica* **14**, 809–834.
- Vigan, M., Stirnemann, J., and Mentré, F. (2014). Evaluation of estimation methods and power of tests of discrete covariates in repeated time-to-event parametric models: Application to gaucher patients treated by imiglucerase. *The AAPS Journal* **16**, 415–423.
- Wilbaux, M., Tod, M., De Bono, J., Lorente, D., Mateo, J., Freyer, G., et al. (2015). A joint model for the kinetics of ctc count and psa concentration during treatment in metastatic castration-resistant prostate cancer. *CPT: Pharmacometrics & Systems Pharmacology* **4**, 277–285.
- 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**, Volume 2012 (2012), Article ID 640153, 17 pages.

Received September 2015. Revised February 2016.
Accepted March 2016.

Web-based Supplementary Materials for Using the SAEM algorithm for mechanistic joint models characterizing the relationship between nonlinear PSA kinetics and survival in prostate cancer patients

Solène Desmée^{1,2}, France Mentré^{1,2}, Christine Veyrat-Follet³,
Bernard Sébastien⁴ and Jérémie Guedj^{1,2}

¹INSERM, IAME, UMR 1137, F-75018 Paris, France

²Université Paris Diderot, IAME, UMR 1137, Sorbonne Paris Cité, F-75018 Paris, France

³Drug Disposition, Disposition Safety and Animal Research Department, Sanofi, Alfortville, France

⁴Biostatistics and Programming, Sanofi, Chilly-Mazarin, France

Web Appendix A: Preliminary sensitivity analysis

We present below the results of the sensitivity analysis for p and g using the longitudinal PSA data and the model which assumed a treatment effect on cell angiogenesis. Four orders of magnitude for p and g were explored and selection was made according to the BIC (Web Table 1).

The parameters values were largely insensitive to the change of p and g values, except N_{max} , the maximum number of PSA-producing cells, which, as expected, decreased when p , the rate of PSA production, increases (Web Tables 2 and 3).

Web Table 1: Estimated BIC (s.e.) according to the values of p and g using only the PSA measurements in the 400 patients of the training dataset

	p				
	0.2	2	20	200	
g	10^{-5}	10488 (6)	10586 (5)	10492 (7)	10491 (7)
	10^{-6}	10184 (8)	10208 (7)	10238 (7)	10207 (8)
	10^{-7}	10166 (8)	10186 (8)	10155 (7)	10194 (7)
	10^{-8}	10188 (8)	10165 (7)	10181 (8)	10182 (7)

Web Table 2: Parameters estimates (relative standard error (%)) of PSA kinetics in the 400 patients of the training dataset for the 16 combinations of p and g values used in the sensitivity analysis (first part).

	$p = 0.2$		$p = 0.2$		$p = 0.2$		$p = 0.2$		$p = 0.2$		$p = 0.2$		$p = 0.2$	
	$g = 10^{-5}$	$g = 10^{-6}$	$g = 10^{-7}$	$g = 10^{-8}$	$g = 10^{-5}$	$g = 10^{-6}$	$g = 10^{-7}$	$g = 10^{-8}$	$g = 10^{-5}$	$g = 10^{-6}$	$g = 10^{-7}$	$g = 10^{-8}$	$g = 10^{-5}$	$g = 10^{-6}$
α_S (day $^{-1}$)	0.055 (3)	0.064 (3)	0.059 (3)	0.059 (3)	0.055 (3)	0.061 (3)	0.061 (3)	0.061 (3)	0.055 (3)	0.061 (3)	0.055 (3)	0.061 (3)	0.067 (3)	0.067 (3)
RF	0.9813 (0)	0.9976 (0)	0.9997 (0)	0.9999 (0)	0.9806 (0)	0.9977 (0)	0.9977 (0)	0.9977 (0)	0.9996 (0)	0.9996 (0)	0.9996 (0)	0.9999 (0)	0.9999 (0)	0.9999 (0)
RE	0.80 (1)	0.81 (1)	0.78 (1)	0.78 (1)	0.81 (1)	0.81 (1)	0.81 (1)	0.81 (1)	0.77 (1)	0.77 (1)	0.77 (1)	0.81 (1)	0.81 (1)	0.81 (1)
ε	0.56 (4)	0.44 (4)	0.49 (4)	0.47 (4)	0.55 (4)	0.47 (4)	0.47 (4)	0.47 (4)	0.53 (4)	0.53 (4)	0.53 (4)	0.41 (4)	0.41 (4)	0.41 (4)
PSA_b (ng.mL $^{-1}$)	21.6 (8)	22.0 (8)	22.1 (8)	22.2 (8)	21.4 (8)	22.1 (8)	22.1 (8)	22.1 (8)	22.2 (8)	22.2 (8)	22.2 (8)	22.3 (8)	22.3 (8)	22.3 (8)
N_{max} (mL $^{-1}$)	4403 (2)	6248 (2)	4817 (2)	4492 (2)	518 (3)	639 (3)	639 (3)	639 (3)	508 (3)	508 (3)	508 (3)	518 (3)	518 (3)	518 (3)
ω_{α_S}	0.51 (5)	0.53 (4)	0.49 (4)	0.52 (5)	0.54 (5)	0.50 (5)	0.50 (5)	0.50 (5)	0.51 (4)	0.51 (4)	0.51 (4)	0.47 (5)	0.47 (5)	0.47 (5)
ω_{RF}	1.89 (4)	2.28 (4)	2.57 (4)	2.84 (4)	1.92 (4)	2.36 (4)	2.36 (4)	2.36 (4)	2.58 (4)	2.58 (4)	2.58 (4)	2.73 (4)	2.73 (4)	2.73 (4)
ω_{RE}	0.72 (5)	0.69 (4)	0.72 (4)	0.67 (5)	0.75 (4)	0.66 (5)	0.66 (5)	0.66 (5)	0.69 (4)	0.69 (4)	0.69 (4)	0.64 (4)	0.64 (4)	0.64 (4)
ω_{ε}	1.30 (6)	1.19 (6)	1.26 (6)	1.28 (6)	1.38 (7)	1.22 (6)	1.22 (6)	1.22 (6)	1.32 (6)	1.32 (6)	1.32 (6)	1.03 (5)	1.03 (5)	1.03 (5)
ω_{PSA_b}	1.63 (4)	1.63 (4)	1.63 (4)	1.64 (4)	1.62 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.64 (4)	1.64 (4)	1.64 (4)
$\omega_{N_{max}}$	2.09 (5)	2.19 (6)	2.14 (5)	2.23 (6)	2.07 (6)	2.25 (6)	2.25 (6)	2.25 (6)	2.32 (6)	2.32 (6)	2.32 (6)	2.31 (6)	2.31 (6)	2.31 (6)
σ	0.33 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.33 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.32 (1)

Web Table 3: Parameters estimates (relative standard error (%)) of PSA kinetics in the 400 patients of the training dataset for the 16 combinations of p and g values used in the sensitivity analysis (second part).

	$p = 20$		$p = 20$		$p = 20$		$p = 20$		$p = 20$		$p = 20$		$p = 20$	
	$g = 10^{-5}$	$g = 10^{-6}$	$g = 10^{-7}$	$g = 10^{-8}$	$g = 10^{-5}$	$g = 10^{-6}$	$g = 10^{-7}$	$g = 10^{-8}$	$g = 10^{-5}$	$g = 10^{-6}$	$g = 10^{-7}$	$g = 10^{-8}$	$g = 10^{-5}$	$g = 10^{-6}$
α_S (day $^{-1}$)	0.065 (3)	0.059 (3)	0.070 (3)	0.063 (3)	0.078 (3)	0.063 (3)	0.068 (3)	0.063 (3)	0.078 (3)	0.063 (3)	0.068 (3)	0.063 (3)	0.078 (3)	0.063 (3)
RF	0.9842 (0)	0.9968 (0)	0.9998 (0)	0.9999 (0)	0.9869 (0)	0.9974 (0)	0.9997 (0)	0.9974 (0)	0.9869 (0)	0.9974 (0)	0.9997 (0)	0.9974 (0)	0.9869 (0)	0.9999 (0)
RE	0.83 (1)	0.79 (1)	0.83 (1)	0.80 (1)	0.87 (1)	0.81 (1)	0.82 (1)	0.87 (1)	0.87 (1)	0.81 (1)	0.82 (1)	0.87 (1)	0.87 (1)	0.80 (1)
ε	0.45 (4)	0.48 (4)	0.39 (4)	0.44 (4)	0.37 (4)	0.45 (4)	0.41 (4)	0.37 (4)	0.37 (4)	0.45 (4)	0.41 (4)	0.37 (4)	0.37 (4)	0.43 (4)
PSA_b (ng.mL $^{-1}$)	21.6 (8)	22.1 (8)	22.2 (8)	22.2 (8)	21.5 (8)	21.8 (8)	22.2 (8)	21.5 (8)	21.5 (8)	21.8 (8)	22.2 (8)	21.5 (8)	21.5 (8)	22.5 (8)
N_{max} (mL $^{-1}$)	72 (4)	48 (4)	62 (4)	50 (4)	8 (4)	6 (4)	6 (4)	8 (4)	8 (4)	6 (4)	6 (4)	8 (4)	6 (4)	6 (4)
ω_{α_S}	0.52 (5)	0.49 (5)	0.45 (4)	0.51 (5)	0.49 (5)	0.51 (5)	0.50 (5)	0.49 (5)	0.49 (5)	0.51 (5)	0.50 (5)	0.49 (5)	0.49 (5)	0.48 (5)
ω_{RF}	1.84 (4)	2.35 (4)	2.62 (4)	2.79 (4)	1.93 (4)	2.38 (4)	2.71 (4)	1.93 (4)	1.93 (4)	2.38 (4)	2.71 (4)	1.93 (4)	1.93 (4)	2.66 (4)
ω_{RE}	0.68 (5)	0.61 (4)	0.64 (5)	0.60 (4)	0.59 (4)	0.66 (5)	0.57 (4)	0.59 (4)	0.59 (4)	0.66 (5)	0.57 (4)	0.59 (4)	0.59 (4)	0.61 (4)
ω_{ε}	1.14 (6)	1.38 (6)	1.07 (6)	1.25 (6)	1.06 (5)	1.20 (6)	1.16 (6)	1.06 (5)	1.06 (5)	1.20 (6)	1.16 (6)	1.06 (5)	1.06 (5)	1.22 (6)
ω_{PSA_b}	1.63 (4)	1.63 (4)	1.64 (4)	1.65 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.63 (4)	1.64 (4)
$\omega_{N_{max}}$	2.26 (6)	2.17 (6)	2.10 (6)	2.08 (5)	2.15 (5)	2.29 (6)	2.12 (5)	2.15 (5)	2.15 (5)	2.29 (6)	2.12 (5)	2.15 (5)	2.15 (5)	2.27 (6)
σ	0.33 (1)	0.32 (1)	0.32 (1)	0.32 (1)	0.33 (1)	0.32 (1)	0.32 (1)	0.33 (1)	0.33 (1)	0.32 (1)	0.32 (1)	0.33 (1)	0.32 (1)	0.32 (1)

Web Appendix B: Sensitivity analysis for the baseline hazard function

The assumption of parametric baseline hazard function was relaxed and we evaluated the effect of using piecewise constant, linear and restricted cubic splines, for the final joint model where the baseline hazard function was simply exponential. In each case, spline functions were tested using 3 knots corresponding to the 10^e, 50^e and 90^e percentiles of the observed times-to-death, i.e., $t_1 = 192$, $t_2 = 520$ and $t_3 = 947$ days, respectively (Harrell, 2001).

- The piecewise constant function was defined as followed:

$$h_0(t) = a_1 + a_2 \mathbb{1}_{[t_1, t_2[} + a_3 \mathbb{1}_{[t_2, t_3[} + a_4 \mathbb{1}_{[t_3, \infty[}.$$

- The linear spline function was defined as followed:

$$\log(h_0(t)) = a_1 + a_2 t + b_1(t - t_1)_+ + b_2(t - t_2)_+ + b_3(t - t_3)_+$$

$$\text{where } (t - t_1)_+ = \begin{cases} 0 & \text{if } t < t_1 \\ t - t_1 & \text{if } t > t_1 \end{cases}$$

- The restricted cubic spline function was defined as followed:

$$\log(h_0(t)) = a_1 + a_2 t + a_3 \left[(t - t_1)_+^3 - (t - t_2)_+^3 \frac{t_3 - t_1}{t_3 - t_2} + (t - t_3)_+^3 \frac{t_2 - t_1}{t_3 - t_2} \right]$$

As can be seen in the Web Table 4, none of these models reduced the BIC over the exponential model.

Web Table 4: Parameters estimates (relative standard error (%)) of PSA kinetics and survival in the 400 patients of the training dataset for the 3 spline functions considered for h_0 compared to the final joint model (exponential h_0)

h_0	Exponential		Piecewise constant		Linear spline		Restricted cubic spline	
-2LL	14326	(8)	14317	(7)	14325	(7)	14333	(7)
BIC	14421	(8)	14431	(7)	14445	(7)	14441	(7)
α_S	0.067	(3)	0.067	(3)	0.081	(3)	0.060	(3)
RF	0.9998	(0)	0.9998	(0)	0.9999	(0)	0.9998	(0)
RE	0.82	(1)	0.79	(1)	0.84	(0)	0.77	(1)
ε	0.43	(3)	0.43	(4)	0.33	(4)	0.48	(4)
PSA_b	21.9	(8)	21.8	(8)	22.0	(8)	22.0	(8)
N_{max}	120	(4)	80	(4)	166	(5)	64	(4)
λ	906	(7)	-	-	-	-	-	-
a_1	-	-	0.000913	(3)	-6.79	(1)	-7.09	(1)
a_2	-	-	0.00041	(4)	7.6×10^{-6}	(15)	0.0014	(13)
a_3	-	-	0.00029	(8)	-	-	-3.7×10^{-9}	(18)
a_4	-	-	5.0×10^{-6}	(136)	-	-	-	-
b_1	-	-	-	-	7.7×10^{-6}	(86)	-	-
b_2	-	-	-	-	7.1×10^{-6}	(129)	-	-
b_3	-	-	-	-	2.1×10^{-5}	(472)	-	-
β	0.00032	(21)	0.0031	(5)	0.0029	(52)	0.011	(10)
β'	0.39	(7)	0.38	(5)	0.39	(5)	0.36	(11)
ω_{α_S}	0.48	(5)	0.51	(5)	0.51	(5)	0.49	(5)
ω_{RF}	2.55	(4)	2.61	(4)	2.38	(4)	2.47	(4)
ω_{RE}	0.59	(5)	0.10	(5)	0.07	(5)	0.11	(5)
ω_{ε}	0.95	(6)	0.97	(6)	0.91	(6)	1.12	(6)
ω_{PSA_b}	1.64	(4)	1.64	(4)	1.64	(4)	1.64	(4)
$\omega_{N_{max}}$	2.57	(7)	2.35	(6)	2.51	(7)	2.37	(6)
σ	0.32	(1)	0.32	(1)	0.32	(1)	0.32	(1)

Guideline for data and codes

Using the SAEM algorithm for mechanistic joint models characterizing the relationship between nonlinear PSA kinetics and survival in prostate cancer patients

Solène Desmée, France Mentré, Christine Veyrat-Follet,
Bernard Sébastien and Jérémie Guedj

Contents

1	Data	1
2	MLXTRAN code and Monolix use	2
2.1	Main code	2
2.2	Structural model	3
2.3	Running Monolix	3
3	R code	4

1 Data

The training dataset is named JM400.txt. Column names are:

- TIME: time from treatment initiation (negative before treatment initiation)
- ID: patient's identifiers
- $Y = \log(\text{PSA}+1)$ if YTYPE equals to 1
= vital status (1 if death, 0 if censored) if YTYPE equals to 2
- CENS = 1 if PSA is below the limit of quantification (0.1 ng.ml^{-1}) and 0 otherwise
- YTYPE: type of observation, 1 for PSA, 2 for vital status
- TIME_ENDtr: covariate that indicates the day of end of treatment

2 MLXTRAN code and Monolix use

2.1 Main code

We used Monolix version 4.3.2 on Matlab R2013a. We provide a folder `Project_JM.mlxtran` which specifies the project features and the different paths. It is structured as followed:

- **DESCRIPTION:** comments on the project.
- **DATA:** information from the dataset `JM400.txt`. The column names of the present dataset are already keywords for Monolix. The path needs to be set.
- **INDIVIDUAL:** specifies the probability distribution of the individual parameters and the inter-individual variability.
- **STRUCTURAL_MODEL:** indicates the external file which contains the structural model, here `Structural_model_finalJM.txt` (see details in the paragraph 2.2). You have to change the path.
- **OBSERVATIONS:** specifies the statistical model of the observation model. `y1` corresponds to the longitudinal $\log(\text{PSA}+1)$ observations and the error model is constant. `survival` corresponds to the time-to-event.
- **TASKS:** specifies the algorithms settings and the tasks to perform.
 - All algorithms settings are specified in the available `Project_JM_algorithms.xmlx` file. The main settings to examine here are:
 - * The seed can be changed (`<algorithms seed="123456">` line 2).
 - * The number of iterations in the E- and M steps of SAEM can be fixed (`<iop_Kauto value="0,0"/>` line 25 and `<vna value="500,200"/>` line 22) or automatic (`<iop_Kauto value="1,1"/>` and `<vna value="500,200"/>` indicates that the maximum number of iterations is 500 for the E-step and 200 for the M-step). We recommend to successively run several population parameters estimation in order to be sure to not have a local extremum. Moreover SAEM convergence can be visually checked throughout the iterations.
 - * We choose 1 Markov chain (`<nmc value="1"/>` line 28) after trying 3 chains without change in results.
 - * To obtain precise estimation of the log-likelihood (small s.e.), we increased the Monte-Carlo size for LL to 200,000 (`<M_is value="200000"/>` line 4).
 - Results folder can be chosen. Here `%MLXPROJECT%` indicates that result folder is the same than the `mlxtran` file.
 - The tasks are:
 - * Estimation of the population parameters: Initializations of fixed effects (`pop_`), standard deviation of the random effects (`omega_`) and residual error (`a_`) parameters have to be specified. SAEM is not very sensitive to initial values. But for illustration and rapidity, we proposed here values closed to the final values.
 - * Estimation of the Fisher Information matrix (and standard errors): in case of joint modelling, Fisher Information Matrix has to be computed by stochastic approximation.

- * Estimation of the individual parameters: we use the conditional modes (i.e., maximization of the conditional probabilities $p(\Phi_i|y_i; \hat{\theta})$). The file "indiv_parameters.txt" is created at this moment and will be useful to plot the figures presented in the article and whose R codes are described in the section 3.
- * Estimation of the log-likelihood: in case of joint modelling, likelihood has to be estimated by important sampling.

2.2 Structural model

The `Structural_model_finalJM.txt` is structured by blocks as followed:

- **DESCRIPTION:** comments on the model
- **INPUT:** vector of parameters to estimate
- **EQUATION:** description of the mathematical model
 t is the keyword for the time regressor.
`odeType=stiff` indicates that Monolix have to use the solver for stiff ODE system. Monolix will use the CVODE solver for stiff ODEs (Cohen and Hindmarsh, 1996).
Here we can fix parameters, specify constraints and transformations.
`PSA_0`, `S_0` and `R_0` are the initial values of the ODE system, i.e., the values of PSA, S and R respectively at first time of PSA observation.
`ddt_S`, `ddt_R` and `ddt_PSA` are the derivative of S, R and PSA respectively with respect to time t .
- **OBSERVATION:** The probability distribution of the event data is defined with the hazard function.
`type=event` specify that the data is of type "event".
- **OUTPUT:** vector of the output of the model

2.3 Running Monolix

The project can be directly run using the following commands in Matlab:

```
monolix;  
aMonolixProject = MonolixProject('Path/to/Project_JM.mlxtran');  
aMonolixProject.load('Path/to/Project_JM.mlxtran');  
aMonolixProject.run()
```

Once the project has been created, one can directly open the project from the Monolix interface in order to change the model, the data or the algorithm calibration in a user-friendly manner.

3 R code

After the Monolix code has been run the R code can be used to generate the figures of the manuscript. The following needs to be done beforehand:

- install the `survival` and `deSolve` packages
- change the working directory (i.e., where Monolix outputs were saved) in `setwd()` (line 18) and the directory to save figures in `way` (line 21)
- change the values of the parameters to the value found in the file "estimated.txt" generated by Monolix (lines 25-27 and 29)
- import the individual parameters provided by Monolix (line 36)
- import the data (line 43)

RCodes.R provides the R codes to reproduce the Figures:

2. Individual fits of PSA kinetics and hazard function (line 116). Other samples of patients can be chosen (line 121).
3. Residuals for the longitudinal part (IWRES) and the survival part (Cox-Snell and Martingale residuals) (line 234)
4. Estimated Kaplan-Meier survival curve and its confidence interval and mean survival curve (line 453)

Chapitre 4

Prédictions dynamiques individuelles et évaluation des capacités prédictives d'un modèle conjoint non-linéaire dans le cancer de la prostate métastatique

4.1 Résumé

Précédemment, nous avons montré comment développer et évaluer un modèle conjoint mécanistique. La modélisation conjointe étant une méthode particulièrement adaptée à la prédiction de la survenue d'un évènement chez un patient en utilisant ses mesures longitudinales (RIZOPOULOS et al., 2013), il est nécessaire d'évaluer ses performances prédictives également dans le cas où le processus longitudinal est décrit par un MNLEM. L'objectif de ce projet est donc d'évaluer la modélisation conjointe non-linéaire pour faire de la prédiction dynamique individuelle, à l'aide de métriques de discrimination et de calibration dépendant du temps (voir paragraphe 1.5.2).

En reprenant le modèle longitudinal développé pour le premier projet de cette thèse (Chapitre 2) qui a l'avantage d'avoir une solution analytique, nous avons estimé les paramètres de population θ d'un modèle conjoint liant le $\log(PSA + 1)$ à la survie chez les $N = 400$ patients du "jeu de données d'apprentissage" ayant un CPmHR et traités par docetaxel et prednisone. A partir de là, nous avons supposé connus le vrai modèle conjoint et le vecteur des paramètres de population associé θ .

$N_{sim} = 200$ patients ont été simulés sous ce modèle conjoint avec des mesures du biomarqueur toutes les 3 semaines jusqu'au temps de décès simulé. Les prédictions individuelles de l'évolution du biomarqueur et du risque de décès ont été estimées, utilisant toutes les mesures du biomarqueur disponibles jusqu'au temps *landmark* $s \geq 0$, notées $Y_i(s)$, ainsi que le fait que l'individu considéré était vivant au temps s ($X_i > s$). Pour tenir compte de l'incertitude sur l'estimation des paramètres individuels, nous avons tiré au sort dans la distribution des effets aléatoires individuels, η_i , $L = 200$ échantillons de Monte-Carlo en utilisant le logiciel de statistique bayésienne Stan version

2.8 (STAN DEVELOPMENT TEAM, 2015). Nous obtenons ainsi des prédictions individuelles dynamiques accompagnées d'intervalles de prédiction à 95%, aux temps $s + t$ où $t > 0$ est l'horizon de la prédiction.

Pour l'évaluation de la discrimination et de la calibration du modèle, les estimateurs IPCW (voir page 42) de l'aire sous la courbe ROC ($AUC(s, t)$), du score de Brier ($BS(s, t)$) et du score de Brier normalisé ($sBS(s, t)$) ont été calculés pour des temps *landmark* $s \in \{0, 6, 12, 18\}$ mois et des horizons $t > 2$ mois à l'aide du package `timeROC` (BLANCHE, 2013). Pour chaque s , l'évolution des performances prédictives du modèle en fonction de l'horizon t peut ainsi être étudiée.

Cette étude de simulation montre tout d'abord que les probabilités de couverture des intervalles de prédiction à 95% de Monte-Carlo pour l'évolution non-linéaire du biomarqueur et du risque de décès ($\forall t > 0$) sont satisfaisantes, puisqu'elles valent environ 95% pour $s = \{0, 6, 12\}$ mois. Néanmoins, pour $s = 18$ mois, les probabilités de couverture sont légèrement inférieures à 95%, ce qui peut être expliquées par des intervalles de prédiction étroits. Le temps *landmark* $s = 0$ ne suffit pas pour fournir une prédiction de la survie discriminante ($AUC(0, t)$ vaut environ 0.6 $\forall t > 0$) et bien calibrée ($sBS(0, t)$ est proche de 0 $\forall t > 0$). Augmenter s améliore discrimination (par exemple, pour $s = 12$ mois, $AUC(12, 6) = 0.80$) et calibration ($sBS(12, 6) = 0.25$). Notons qu'un sBS positif traduit le fait que le modèle conjoint permet d'obtenir des prédictions mieux calibrées qu'avec un estimateur de Kaplan-Meier. Cela reflète qu'avec un plus long suivi longitudinal, l'estimation des paramètres individuels est meilleure, il y a moins de *shrinkage* (i.e. régression vers la moyenne). Ces résultats incitent à envisager d'autres études de simulation (par exemple en considérant un lien entre les processus longitudinaux et de survie, β , plus fort).

Ensuite, nous avons procédé de même pour les $N' = 196$ patients du "jeu de données de validation" pour obtenir leurs prédictions dynamiques individuelles puis calculer les métriques $AUC(s, t)$, $BS(s, t)$ et $sBS(s, t)$. Sur ces vraies données, les prédictions individuelles dynamiques se précisent lorsque la durée de collecte de mesures de PSA augmente. Les comportements des AUC, BS et sBS sont similaires à ceux de l'étude de simulation. Cependant, pour tous les temps *landmark* $s \geq 6$ mois, les $AUC(s, t)$ stagnent à la valeur de 0.75 environ. Cette différence peut s'expliquer par le fait qu'il y ait moins de mesures de PSA par patient quand le suivi s'allonge que dans la simulation. Concernant le BS, le gain par rapport à l'utilisation d'un estimateur de Kaplan-Meier est modéré puisque sBS vaut au mieux 0.25 (atteint pour $s = 18$ et $t = 12$ mois). Au final, l'analyse de ces vraies données suggèrent qu'il est nécessaire d'avoir au moins 12 mois de suivi de PSA pour fournir des prédictions individuelles de la survie ayant des propriétés de discrimination et de calibration satisfaisantes jusqu'à un horizon de 18 mois.

Ce travail montre comment évaluer les performances prédictives de modèles conjoints non-linéaires, en utilisant des méthodes initialement développées dans un contexte

linéaire. Il illustre également l'importance d'une durée minimum de collecte de mesures longitudinales, s , suffisante pour prédire correctement le risque de survenue d'un évènement à un temps d'horizon donné. Une future étape sera d'appliquer ces méthodes en utilisant le modèle conjoint mécanistique développé dans le chapitre 3.

L'écriture d'un article sur ce projet est actuellement en préparation.

4.2 Article 3 (en préparation)

Dynamic predictive performances of individual survival using nonlinear joint models: application to PSA kinetics in metastatic prostate cancer

Solène Desmée^{1,2*}, France Mentré^{1,2}, Marie-Karelle Rivière^{1,2}, Christine Veyrat-Follet³, and Jérémie Guedj^{1,2}

*Correspondence: solene.desmee@inserm.fr

¹ INSERM, IAME, UMR 1137, F-75018, Paris, France

² Université Paris Diderot, IAME, UMR 1137, Sorbonne Paris Cité, F-75018, Paris, France

³ Drug disposition, Disposition Safety and Animal Research Department, Sanofi, Bridgewater, USA

Abstract

Background: The development of nonlinear joint model which simultaneously analyzes nonlinear longitudinal and survival data opens the possibility to carry out individual dynamic predictions of the risk of death, as shown using linear mixed-effects models. Here, in a context of nonlinear mixed-effects models (NLMEM), the objective is to provide individual dynamic predictions and assess predictive performances of joint model.

Methods: Assuming that the joint model has been previously built on a large population and using the estimated population parameters as a prior, a Bayesian approach using the software Stan is proposed to calculate the a posteriori distribution of the risk of death for a new patient, knowing his longitudinal measurements until a given landmark time. Time-dependent measures of model discrimination (AUC) and calibration (Brier score) are also computed in that NLMEM context. Methods are applied on metastatic prostate cancer patients treated with chemotherapy for which prostate-specific antigen (PSA) was frequently collected for long period of time. It is first applied on *in silico* patients, then illustrated on real data.

Results: Individual dynamic predictions were computed with their 95% Monte Carlo prediction intervals. Prediction intervals in simulated patients offered satisfying coverage probabilities (close to 95%) for both longitudinal and survival processes. Discrimination and calibration metrics were plotted for several landmark times, showing that a sufficiently long biomarker follow-up improves predictions of the risk of death. In particular, on real data, 12 months of PSA follow-up from treatment initiation appeared necessary to provide discriminant and well-calibrated individual predictions of survival beyond 18 months with AUC around 0.75 and Brier score smaller than 0.21.

Conclusions: We showed how methods used to prospectively predict patient's survival in linear models and to assess predictive performances can be expanded to NLMEM. Because nonlinear models can be used to develop physiological model incorporating several biomarkers, this approach may be valuable to optimize information collected during clinical trials to early detect treatment effectiveness and adapt patient's treatment.

Key-words: Calibration, Discrimination, Individual dynamic prediction, Metastatic prostate cancer, Monte Carlo sample, Nonlinear joint model, PSA, Survival

Background

Personalized medicine requests effective methodological tools to adequately guide treatment strategy and patient care. Often biomarkers measured all along the patient's history are assumed to be link to a time-to-event of interest. Nevertheless most prognostic models use only a small fraction of the available longitudinal information [1], not the entire biomarker kinetics. However it appears relevant to use all longitudinal measurements to predict the individual risk of event. This is possible using joint modelling, which consists in simultaneously handling longitudinal and time-to-event data by maximizing the joint distribution of both processes [2–4]. In this approach, the longitudinal part is described by a mixed-effect model and the survival part is described by a parametric or semi-parametric hazard function which depends on the true, unobserved, biomarker kinetics. Thus, once a joint model is developed, it can be used to make “dynamic predictions” [5–7], i.e. individualized predictions that can be updated over the patient's follow-up. Nevertheless the numerical difficulties have long limited the use of joint models to linear models for the longitudinal process. Recently, we and others have shown how statistical softwares used for nonlinear mixed-effect models (NLMEM) can now be expanded to nonlinear joint models [8,9]. Here we now aim to use this approach for individual dynamic predictions, as well as expanding tools to assess the predictive performance of these models taking into account the dynamic nature of the predictions through time-dependent Area under the Receiver Operating Characteristic (ROC) curve (AUC) and Brier score [6,10,11].

The methods are illustrated on data from metastatic castration-resistant prostate cancer (mCRPC) patients included in a phase III clinical trial and receiving the reference chemotherapy [12]. In that context of advanced disease, the prostate-specific antigen (PSA), which is correlated to the risk of death [13], is frequently measured.

The outline of the paper is as follows; in the section “Methods” we present the joint model approach, the Bayesian approach used for dynamic individual predictions and the method for computation of discrimination and calibration metrics accompanied. In the section “Results”, we show individual dynamic predictions and time-dependent metrics for discrimination and calibration in simulated and real mCRPC data. The sections “Discussion” and “Conclusion” point out the advantages and the limits of this approach.

Methods

Joint modelling for nonlinear longitudinal and survival data

Let N the number of patients and $y_i = \{y_{i1}, \dots, y_{in_i}\}$ the vector of the longitudinal observations in patient i , where y_{ij} denotes the j^{th} measurement of the biomarker for the individual i at time t_{ij} , where $i = 1, \dots, N$, $j = 1, \dots, n_i$, and n_i is the number of measurements in subject i . Here the kinetics of the biomarker is assumed to be nonlinear. Thus NonLinear Mixed-Effects Models (NLMEM) are used to analyze the longitudinal evolution:

$$y_{ij} = b(t_{ij}, \psi_i) + e_{ij} \quad (1),$$

where $b(t, \psi_i)$ is the true and unobserved value of the biomarker at time point t , defined by a nonlinear structural model through the function b , ψ_i is the vector of the individual parameters and e_{ij} the residual Gaussian error of mean 0 and variance σ^2 . ψ_i is decomposed as a vector of fixed effects μ representing median parameters of the population and random effects η_i specific for each individual. The random effects are assumed to be normally distributed with mean zero and variance-covariance matrix Ω ($\eta_i \sim \mathcal{N}(0, \Omega)$), and are assumed independent of the residual error e_{ij} .

Let X_i denote the time-to-event and C_i the censoring time for the i^{th} subject. $T_i = \min(X_i, C_i)$ is observed and we note $\delta_i = 1$ if $X_i \leq C_i$, 0 otherwise. The hazard function, h , can be written as follows:

$$h_i(t | \mathbf{B}(t, \psi_i)) = h_0(t) \exp(\beta f(t, \psi_i) + \gamma^T w_i) \quad (2),$$

where $\mathbf{B}(t, \psi_i) = \{b(u, \psi_i); 0 \leq u < t\}$ denotes the history of the true unobserved longitudinal process up to t , h_0 is a parametric baseline hazard function and γ is the vector of coefficients associated with a vector of baseline covariates w_i . The link function f depends on the true biomarker kinetics $b(t, \psi_i)$. The parameter β quantifies the strength of the association between the link function and the risk of death.

Dynamic individual predictions

Let us assume that the vector of population parameters θ is known, from a training dataset. Now we are interested in a new subject i with longitudinal biomarker measurements available until a landmark time $s \geq 0$: $Y_i(s) = \{y_{ij}; 0 \leq t_{ij} \leq s\}$, and for whom we aim to predict his risk of death before a time $s + t$, where $t > 0$ is called the

horizon time. Because the patient is alive at time s , we focused on the conditional probability of death between s and the prediction horizon $s + t$:

$$\pi_i(s + t|s) = \Pr(X_i < s + t | X_i > s, Y_i(s); \theta) \quad (4)$$

For each landmark time s , we used only the biomarker measurements of i up to time s to compute the distribution of the individual parameters and infer the survival probability with a prediction interval taking into account the uncertainty of the individual parameter estimation. Monte Carlo samples of $\pi_i(s + t|s)$ can be obtained using the approach proposed by Rizopoulos (2011) [5]. For $l = 1, \dots, L$ where L is the number of Monte Carlo samples:

1. Draw $\eta_i^{(l)} \sim \{\eta_i | X_i > s, Y_i(s), \theta\}$ and infer $\psi_i^{(l)}$
2. Compute: $y_i^{(l)}(u) = b(u, \psi_i^{(l)})$ with $u > 0$

$$\text{And: } \pi_i^{(l)}(s + t|s) = \frac{S_i(s|\eta_i^{(l)}; \theta) - S_i(s + t|\eta_i^{(l)}; \theta)}{S_i(s|\eta_i^{(l)}; \theta)}$$

where $S_i(t|\eta_i; \theta) = \Pr(X_i > t | \mathbf{B}(t, \psi_i)) = \exp\left(-\int_0^t h_i(s|\mathbf{B}(s, \psi_i); \theta) ds\right)$ is the survival function conditionally on the random effects.

The difficulty here is the first step since the posterior distribution of the random effects given the observed data is of non-standard form (we know that a priori $\eta_i \sim \mathcal{N}(0, \Omega)$ but the form of the a posteriori distribution for the specific subject i is unknown for a nonlinear model). Therefore we used the Stan software version 2.8 [14], a language for Bayesian statistical models, and its R interface. Stan makes possible to specify the a priori distribution of the random effects and, given the longitudinal data and the fact that patient is alive at s , provides values drawn in the a posteriori distribution. To take into account the vital status of each subject i at time s , we used a Gauss-Legendre quadrature of order 8 to integrate the hazard function.

The L realizations $\{y_i^{(l)}(u), l = 1, \dots, L\}$ and $\{\pi_i^{(l)}(s + t|s), l = 1, \dots, L\}$ can be used to derive estimates of the logarithm of PSA+1 and estimates of $\pi_i(s + t|s)$ as:

$$\hat{y}_i(u) = \text{median}\{y_i^{(l)}(u), l = 1, \dots, L\}$$

$$\text{And: } \hat{\pi}_i(s + t|s) = \text{median}\{\pi_i^{(l)}(s + t|s), l = 1, \dots, L\} \quad (5)$$

And 95% prediction intervals for the individual predictions are obtained using the 2.5% and 97.5% Monte Carlo sample percentiles.

Model discrimination and calibration

We are interested in discrimination, i.e., the capacity of the model to discriminate patients of low and high risk of death, and in calibration, i.e., the capacity of the model to predict future events. We used the definition of time-dependent AUC corresponding to a cumulative sensitivity and a dynamic specificity [15,16] and the definition of time-dependent BS presented in Schoop et al. [17]. At landmark time s and for a prediction horizon t these metrics are defined as follows:

$$AUC(s, t) = \Pr(\pi_i(s + t|s) > \pi_j(s + t|s) \mid s < X_i < s + t, X_j > s + t)$$

$$BS(s, t) = \mathbb{E} \left[\left(\mathbb{1}_{\{s < X < s+t\}} - \pi(s + t|s) \right)^2 \mid X > s \right]$$

AUC and BS are complementary measures of the dynamic predictive accuracy. AUC measures the capacity of the function $\pi_i(s + t|s)$ to distinguish between patients of low and high risk of dying in the horizon time t . With these definitions, a dummy model such that $\pi_i(s + t|s) = 0.5$ for all i, s and t will lead to $AUC=0.5$ and $BS=0.25$. BS measures the average discrepancies between true vital status and estimated predictions. It has the advantage to quantify both discrimination and calibration. . Note that AUC does not depend on the disease prevalence while the Brier score is not free of disease prevalence. Therefore, BS corresponding to different values of s , and hence associated with a different number of events, cannot be directly compared. So BS must be carefully interpreted because its scaling changes with s [10]. To make possible comparisons of BS over different landmark times s , scaled Brier score were introduced [11,18]: $sBS(s, t) = \frac{BS_{KM}(s,t) - BS(s,t)}{BS_{KM}(s,t)}$ where $BS_{KM}(s, t)$ is the Brier score for reference predictions, here obtained with Kaplan-Meier estimates of the survival function. The higher sBS, the better it the prediction. It can be interpreted as how much the BS is decreased using the model compared to a simpler prediction.

To estimate these metrics, we used weighted estimators to account for right censoring using Inverse Probability of Censoring Weights (IPCW) [19,20]. Thus the IPCW estimators are:

$$\widehat{AUC}(s, t) = \frac{\sum_{i=1}^N \sum_{j=1}^N \mathbb{1}_{\{\hat{\pi}_i(s+t|s) > \hat{\pi}_j(s+t|s)\}} \tilde{D}_i(s, t)(1 - \tilde{D}_j(s, t)) \widehat{W}_i(s, t) \widehat{W}_j(s, t)}{\sum_{i=1}^N \sum_{j=1}^N \tilde{D}_i(s, t)(1 - \tilde{D}_j(s, t)) \widehat{W}_i(s, t) \widehat{W}_j(s, t)}$$

And:

$$\widehat{BS}(s, t) = \frac{1}{\sum_{i=1}^N \mathbb{1}_{\{T_i > s\}}} \sum_{i=1}^N \widehat{W}_i(s, t) (\tilde{D}_i(s, t) - \hat{\pi}_i(s+t|s))^2$$

Where $\tilde{D}_i(s, t) = \mathbb{1}_{\{s < T_i \leq s+t\}}$ and the weights $\widehat{W}_i(s, t) = \frac{\mathbb{1}_{\{T_i > s+t\}}}{\widehat{G}(s+t|s)} + \frac{\tilde{D}_i(s, t) \delta_i}{\widehat{G}(T_i|s)}$ take into account censor, with $\widehat{G}(u)$ the Kaplan-Meier estimator of survival function of the censoring time at time u , i.e., $\Pr(C > u)$ and $\forall u > s$, $\widehat{G}(u|s) = \frac{\widehat{G}(u)}{\widehat{G}(s)}$ and $\hat{\pi}_i(s+t|s)$ as defined in the formula (5). Thus, once $\hat{\pi}_i(s+t|s)$ has been obtained, as AUC and BS are model free, they can be calculated using packages developed in the context of linear models, such as the `timeROC` R package [21].

The scaled Brier score was obtained using the estimated Brier score: $s\widehat{BS}(s, t) = 1 - \frac{\widehat{BS}(s, t)}{\widehat{BS}_{KM}(s, t)}$ where $\widehat{BS}_{KM}(s, t)$ is the Brier Score obtained using the Kaplan-Meier estimate at $s+t$ in the sample on which the joint model was developed, the training dataset.

Building a reference nonlinear joint model

We aimed to develop a reference nonlinear joint model. For that purpose, we used data from metastatic Castration-Resistant Prostate Cancer (mCRPC) patients of the control arm of a phase 3 clinical trial [12] that included 598 men treated with the first-line reference chemotherapy: docetaxel in combination with prednisone. Only PSA measurements at baseline (i.e., measured within 8 days before treatment initiation) and after treatment initiation were used. 596 patients had at least one PSA measurement and were kept for the analyses. In the protocol, PSA had to be measured every 3 weeks during treatment and then every 12 weeks after treatment. The date of death or last visit was precisely registered. Data were randomly split into two datasets as in the paper [13]: a training dataset containing $N = 400$ patients to develop the reference joint model and to estimate the population parameters, and a validation dataset containing the $N' = 196$ remaining patients to provide individual dynamic predictions and assess the predictability of the model on real data.

Thus in the application, the observed biomarker y_{ji} is the j^{th} longitudinal measurement of $\log(\text{PSA} + 1)$ for the individual i at time t_{ij} . In the following, $b(t, \psi)$ of the Eq. (1) is in fact $\log(\text{PSA}(t, \psi) + 1)$ where $\text{PSA}(t, \psi)$ is defined by the nonlinear structural model already described in a previous study [9]. In brief, PSA evolution over time, t , was described by the following biexponential function:

$$\text{PSA}(t, \psi) = \begin{cases} \frac{\delta \text{PSA}_0}{r(1-\varepsilon)-d+\delta} e^{(r(1-\varepsilon)-d)t} + \left[\text{PSA}_0 - \frac{\delta \text{PSA}_0}{r(1-\varepsilon)-d+\delta} \right] e^{-\delta t} & \text{if } t \leq T_{esc} \\ \frac{\delta \text{PSA}_0}{r-d+\delta} e^{(r-d)t-r\varepsilon T_{esc}} + \left[\text{PSA}(T_{esc}) - \frac{\delta \text{PSA}_0 e^{(r(1-\varepsilon)-d)T_{esc}}}{r-d+\delta} \right] e^{-\delta(t-T_{esc})} & \text{if } t > T_{esc} \end{cases} \quad (5)$$

Where PSA_0 is the PSA value at treatment initiation, δ (day^{-1}) is the rate of PSA elimination, r (day^{-1}) the rate of prostatic cells proliferation in absence of treatment, d (day^{-1}) the rate of prostatic cells elimination. ε is a constant treatment effect and T_{esc} the time at which tumor escapes and treatment has no longer an effect.

Because only 4 parameters can be identified from Eq. 5, we fixed d to 0.046 day^{-1} , corresponding to a half-life of tumor cells of 15 days, consistent with an apoptotic index of 5% in metastatic prostate cancer [22]. Moreover we fixed δ to 0.23 day^{-1} , corresponding to a PSA half-life in blood of about 3 days [23]. Thus PSA kinetics was defined by the vector parameter $\psi = \{r, \text{PSA}_0, \varepsilon, T_{esc}\}$.

In the NLMEM it is assumed that Ω is a diagonal variance-covariance matrix of the random effects. Each diagonal element ω_q^2 represents the variance of the q^{th} component of the random effect vector η_i . We assume log-normal distribution for r , PSA_0 and T_{esc} and logit-normal distribution for ε .

For the survival process, we assumed no covariate effect ($\gamma = 0$) and the 5 following link functions f were considered:

- No link: $f(t, \psi_i) = 0$
- Initial PSA: $f(t, \psi_i) = \log(\text{PSA}(0, \psi_i) + 1)$
- Current PSA: $f(t, \psi_i) = \log(\text{PSA}(t, \psi_i) + 1)$
- Current PSA slope : $f(t, \psi_i) = \frac{d \log(\text{PSA}(t, \psi_i) + 1)}{dt}$
- Area under PSA: $f(t, \psi_i) = \int_0^t \log(\text{PSA}(u, \psi_i) + 1) du$

The $N = 400$ patients of the training dataset were used to select the best joint model among the 5 above. All the available PSA measurements and vital status in the N patients were used, from the treatment initiation to the end of follow-up. We selected the joint model that provided the smaller Bayesian Information Criterion (BIC) [24] using the software Monolix version 4.3.2.

Simulations

$N_{sim} = 200$ patients were simulated with PSA measurements every 3 weeks for 3 years or until death using the link function of the selected joint model and associated population parameters (see Results and Table 1). No other mechanism than death was considered for dropout. Dynamic individual predictions were done as explained in the paragraph “Dynamic individual predictions” with $L = 200$. For each landmark and each horizon time, we calculated the coverage probabilities of the 95% prediction intervals, i.e., the proportion of simulated patients for whom the true value of the marker (either PSA value or the risk of death) was contained in the 95% prediction interval of the marker. Then, we estimated time-dependent AUC, BS and sBS for several landmark time s and horizon time t . For sBS estimation, as in this simulation study we did not have another reference dataset, $BS_{KM}(s, t)$ was obtained using the Kaplan-Meier estimates computed in the $N_{sim} = 200$ simulated patients themselves.

Application to real data

Individual dynamic predictions were applied in the $N' = 196$ patients from the validation dataset, i.e. patients not included for joint model development. $L = 200$ Monte-Carlo samples of individual parameters for each patient were provided at several landmark times s using the Stan software version 2.8 in order to supply predictions. Lastly, time-dependent discrimination and calibration metrics were computed as explained above.

Results

Joint model selection

In the training dataset including the $N = 400$ patients, 5,710 PSA measurements were collected and the median [minimum; maximum] number of measurements per patient was 13 [1 ; 55]. Regarding survival, 286 deaths occurred among the 400 patients (71.5%), leading to a median survival [Kaplan-Meier 95% confidence interval] of 656

days [598 ; 741]. All these data were used to estimate parameters using the 5 candidate joint models (Table 1). PSA kinetic parameters were largely insensitive to the choice of the survival part of the model. In all cases they were precisely estimated with relative standard error smaller than 8% for both fixed effects and variance components. In particular the treatment effect ε was consistently estimated to about 35% and the median time at which patient escapes the treatment, T_{esc} , was about 140 days.

When comparing the BIC of the different joint models, the model relying on the current PSA clearly outperformed all the other models. Thus this model was selected for carrying out the individual predictions and θ was fixed to values estimated with this model.

Simulation study

Data were simulated according to the model linking the current PSA to time-to-death (Table 1). 200 new patients were generated and for each patient alive at various landmark times s , we obtained PSA kinetics and survival probabilities for $t > 0$ accompanied by their Monte Carlo 95% prediction intervals. The coverage probabilities of these prediction intervals over the set of horizon time t were closed to 95% for both PSA evolution and risk of death for $s = \{0, 6, 12\}$ months (Figure 1). For $s = 18$ months, coverage probabilities were smaller than 95%, associated with narrow prediction intervals (for $t \leq 12$ months, mean width of the individual 95% prediction intervals for PSA was smaller than 13.6 ng.mL^{-1} and for risk of death smaller than 13.2%).

For $s = 0$, AUC values were around the low level of 0.6 (Figure 2). When landmark time increases, AUC values increased and reached satisfactory levels: for example, with $s = 12$ months, we found AUC value of 0.80 for horizon $t = 6$ months, 0.82 for $t = 12$ months; with $s = 18$ months, 0.79 and 0.90 for $t = 6$ and $t = 12$ months respectively. This reflects the fact that a long follow-up of the biomarker allows to improve individual parameter estimation and the good properties of AUC as a discrimination measurement.

BS and sBS estimations were provided in Figure 3. BS for $s = 0$ quickly deteriorates until the level of 0.25, which corresponds to a dummy prediction. This means that an initial value is not enough for individual survival prediction. For a given landmark time s , BS increased when the horizon time t increases, meaning that the model well

calibrates in the short term, then calibration is deteriorated. Moreover, BS improved (i.e. was smaller) when landmark times s increases, which indicates that having more and more longitudinal measurements improves precision of survival predictions. Especially for $s = 12$ or $s = 18$, BS are smaller than 0.16 and 0.15 respectively, regardless the horizon time. sBS (Figure 3) complements this interpretation by comparing BS obtained with the joint model to BS_{KM} obtained with the Kaplan-Meier estimates. sBS for $s = 0$ was about 0 regardless t , confirming that baseline value cannot improve prediction of survival by itself. For $s > 0$, sBS increase when t increases, but never exceed 0.5, which means that for long-time horizon, our joint model predicts moderately better survival than Kaplan-Meier estimates.

Application to real data

In the validation dataset including $N' = 196$ new patients, a total of 2,720 PSA measurements were collected and the median number of measurements per patient was 13 [2 ; 57]. 145 deaths occurred (74.0%), leading to a median survival of 598 days [547 ; 732].

Figure 4 illustrates dynamic predictions for 3 typical patients of the validation dataset. When the landmark increased, the median prediction of PSA was increasingly close to the future PSA observations and precise with shrinking 95% prediction intervals. In these 3 patients, we observed that as soon as the PSA nadir after treatment initiation was reached (which corresponds to T_{esc} was observed), PSA predictions were more accurate, because of improvement of the Bayesian estimations. Concerning survival function, predictions were accompanied by a large 95% prediction interval making impossible to give a firm opinion on risk of death for a specific patient (50% was included in the prediction interval at the death or censor time) when the number of PSA observations available were not enough (landmark times $s < 18$ for patients 2073 and 2466 and landmark times $s < 12$ for patient 2558).

Similarly to the simulation study, for $s = 0$, AUC quickly decreased to about 0.5 (Figure 5). For $s = \{6, 12, 18\}$ months, AUC values remained stable close to 0.75 regardless of the horizon time t . Of note, contrary to the simulations, in real data the amount of PSA measurements stagnated, which can explained this stabilization.

BS evolution (Figure 6) is very similar to those of the simulations. BS for $s = 0$ quickly deteriorates until it is slightly greater than 0.25. For a given landmark time s , BS

increased when the horizon time t increases. On the other hand, as previously, sBS for $s = 0$ was around zero for all horizon t . For $s = 6$ months, sBS increased when the horizon time t increases (Figure 6), then decreased, meaning that the joint model calibrates better than a Kaplan-Meier estimates for medium-term predictions. For $s \in \{12, 18\}$ months, we only observed sBS increase when t increases. Nevertheless, relative to the prediction with the Kaplan-Meier estimate for patients in the training dataset, there is only a moderate gain of accuracy since sBS reached 0.29 at best.

On these mCRPC patients, it seems that a sufficient longitudinal follow-up duration ($s = 12$) is necessary to provide survival prediction in a horizon time of 18 months with relatively satisfactory discrimination and calibration, even if, using this joint model, the gain is low compared to Kaplan-Meier estimates of the survival.

Discussion

We showed how methods for individual dynamic predictions and model discrimination and calibration assessment developed in a linear context can be expanded to individual dynamic predictions using nonlinear joint modelling. When longitudinal and survival processes are linked, the good performances of joint modelling for individual predictions in terms of both discrimination and accuracy, obtained when the longitudinal part of the joint model was described by a linear mixed effect model [6], are here highlighted in a nonlinear context.

Monolix could be used to jointly fit nonlinear longitudinal and survival data [9], and could also provide empirical Bayesian estimates of the individual parameters. Nevertheless, it does not allow to draw realizations in the a posteriori distribution of the individual random effects, hence the use of a software for Bayesian statistics, Stan. The coverage probabilities of the prediction intervals were satisfying and further the use of joint models for predict risk of death from longitudinal measurements. Of note, the smaller coverage probabilities for long follow-up ($s = 18$ months) reflect an optimistic precision of the predictions. Here for the individual parameter estimation, we assumed that population parameters θ were exactly estimated because the relative standard errors were small. But a more rigorous approach would have been to take also into account this uncertainty [6].

The AUC and BS metrics used here are model-free and thus very convenient for expansion to the nonlinear context. The predictions becomes more discriminant and well calibrated when the landmark time increases, corresponding to a longer follow-up duration. In the simulation study, the improvement of AUC and BS when the landmark time increases illustrates the shrinkage of the individual parameter estimates when longitudinal measurements are sparse. Of note the number of biomarker observations per patient at each landmark in the simulation study is always the maximum number of measurements scheduled by the protocol (i.e. every 3 weeks), whereas in the real patients, PSA measurements become less frequent per patient. That could explained the fact that AUC values were similar for the 3 positive landmark times in the real data, because the amount of PSA information did not really increase. Moreover, the application to real data is here ideal since both datasets (for model selection and prediction assessment) belong to the same clinical trial, so in future this approach should be externally validated before to be possibly routinely used.

The main advantage of nonlinear joint model is the possibility to develop physiological models. In this study which aimed to propose a methodology to provide and assess individual dynamic predictions using a nonlinear joint model, the longitudinal part of the model is a simplification of the real mechanism of PSA production and will require to be improved, using ordinary differential equations (ODEs) [13]. This should be possible very soon with the development of Stan for stiff ODEs and opens the way for more accurate predictions.

Conclusion

The present work shows how to use nonlinear joint model for individual dynamic predictions and how assess the predictive capacities of such model. The application on mCRPC patients illustrates how results of such methodology can be interpreted in a clinical context. We used a Bayesian approach to take into account the uncertainty in the individual parameter estimation in individual predictions. To conclude, nonlinear joint model opens the way for the use of all available biomarker measurements in personalized medicine to early detect treatment effect and adapt patient treatment in cancer but also in other therapeutic areas.

References

1. Armstrong AJ, Eisenberger MA, Halabi S, Oudard S, Nanus DM, Petrylak DP, et al. Biomarkers in the Management and Treatment of Men with Metastatic Castration-Resistant Prostate Cancer. *Eur Urol.* 2012;61:549–59. doi:10.1016/j.eururo.2011.11.009.
2. Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Stat Sin.* 2004;14:809–834.
3. Rizopoulos D, Verbeke G, Lesaffre E. Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. *J R Stat Soc Ser B Stat Methodol.* 2009;71:637–54. doi:10.1111/j.1467-9868.2008.00704.x.
4. Wu L, Liu W, Yi GY, Huang Y. Analysis of Longitudinal and Survival Data: Joint Modeling, Inference Methods, and Issues. *J Probab Stat.* 2011;2012. doi:10.1155/2012/640153.
5. Rizopoulos D. Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics.* 2011;67:819–829. doi:10.1111/j.1541-0420.2010.01546.x.
6. Rizopoulos D, Murawska M, Andrinopoulou E-R, Molenberghs G, Takkenberg JJM, Lesaffre E. Dynamic Predictions with Time-Dependent Covariates in Survival Analysis using Joint Modeling and Landmarking. *ArXiv13066479 Stat.* 2013.
7. Proust-Lima C, Taylor JMG. Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: a joint modeling approach. *Biostat Oxf Engl.* 2009;10:535–49. doi:10.1093/biostatistics/kxp009.
8. Mbogning C, Bleakley K, Lavielle M. Joint modelling of longitudinal and repeated time-to-event data using nonlinear mixed-effects models and the stochastic approximation expectation–maximization algorithm. *J Stat Comput Simul.* 2015;85:1512–28. doi:10.1080/00949655.2013.878938.
9. Desmée S, Mentré F, Veyrat-Follet C, Guedj J. 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. *AAPS J.* 2015;17:691–9. doi:10.1208/s12248-015-9745-5.
10. Blanche P, Proust-Lima C, Loubère L, Berr C, Dartigues J-F, 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.* 2014. doi:10.1111/biom.12232.
11. Mauguen A, Rachet B, Mathoulin-Pélissier S, Lawrence GM, Siesling S, MacGrogan G, et al. Validation of death prediction after breast cancer relapses using joint models. *BMC Med Res Methodol.* 2015;15:27. doi:10.1186/s12874-015-0018-x.
12. Tannock IF, Fizazi K, Ivanov S, Karlsson CT, Fléchon A, Skoneczna I, et al. Aflibercept versus placebo in combination with docetaxel and prednisone for treatment of men with metastatic castration-resistant prostate cancer (VENICE): a phase 3, double-blind randomised trial. *Lancet Oncol.* 2013;14:760–8. doi:10.1016/S1470-2045(13)70184-0.
13. Desmée S, Mentré F, Veyrat-Follet C, Sébastien B, Guedj J. Using the SAEM algorithm for mechanistic joint models characterizing the relationship between nonlinear PSA kinetics and survival in prostate cancer patients. *Biometrics.* 2016:(in press).
14. Stan Development Team. *Stan Modeling Language Users Guide and Reference Manual, Version 2.8.0.* 2015.

15. Heagerty PJ, Zheng Y. Survival Model Predictive Accuracy and ROC Curves. *Biometrics*. 2005;61:92–105. doi:10.1111/j.0006-341X.2005.030814.x.
16. Parast L, Cheng S-C, Cai T. Landmark Prediction of Long-Term Survival Incorporating Short-Term Event Time Information. *J Am Stat Assoc*. 2012;107:1492–501. doi:10.1080/01621459.2012.721281.
17. Schoop R, Graf E, Schumacher M. Quantifying the Predictive Performance of Prognostic Models for Censored Survival Data with Time-Dependent Covariates. *Biometrics*. 2008;64:603–10. doi:10.1111/j.1541-0420.2007.00889.x.
18. Graf E, Schmoor C, Sauerbrei W, Schumacher M. Assessment and comparison of prognostic classification schemes for survival data. *Stat Med*. 1999;18:2529–45. doi:10.1002/(SICI)1097-0258(19990915/30)18:17/18<2529::AID-SIM274>3.0.CO;2-5.
19. Blanche P, Latouche A, Viallon V. Time-Dependent AUC with Right-Censored Data: A Survey. In: Lee M-LT, Gail M, Pfeiffer R, Satten G, Cai T, Gandy A, editors. *Risk Assess. Eval. Predict.*, Springer New York; 2013, p. 239–51.
20. Gerds TA, Schumacher M. Consistent Estimation of the Expected Brier Score in General Survival Models with Right-Censored Event Times. *Biom J*. 2006;48:1029–40. doi:10.1002/bimj.200610301.
21. Blanche P. Package “timeROC”: time-dependent ROC curve and AUC for censored survival data. Vienna Austria R Found Stat Comput. 2013.
22. Tu H, Jacobs SC, Borkowski A, Kyprianou N. Incidence of apoptosis and cell proliferation in prostate cancer: Relationship with TGF- β 1 and bcl-2 expression. *Int J Cancer*. 1996;69:357–363. doi:10.1002/(SICI)1097-0215(19961021)69:5<357::AID-IJC1>3.0.CO;2-4.
23. Polascik TJ, Oesterling JE, Partin AW. Prostate specific antigen: a decade of discovery--what we have learned and where we are going. *J Urol*. 1999;162:293–306.
24. Park KY, Qiu P. Model selection and diagnostics for joint modeling of survival and longitudinal data with crossing hazard rate functions. *Stat Med*. 2014;33:4532–46. doi:10.1002/sim.6259.

Table 1 Goodness-of-fit statistic and parameters estimates (relative standard error (%)) of PSA kinetics and survival in the $N = 400$ patients for the 5 joint models

Models	No link	Initial PSA	Current PSA	Current PSA slope	Area under PSA
BIC	14350	14331	14192	14291	14327
r (day ⁻¹)	0.054 (1)	0.054 (1)	0.054 (1)	0.055 (1)	0.054 (1)
PSA_0 (ng.mL ⁻¹)	74.6 (8)	74.8 (8)	73.9 (8)	73.4 (8)	74.9 (8)
ε	0.35 (5)	0.35 (5)	0.34 (5)	0.35 (5)	0.35 (5)
T_{esc} (day)	138 (4)	137 (4)	138 (4)	142 (4)	136 (4)
λ	885 (4)	1620 (13)	3800 (9)	1500 (9)	1410 (13)
k	1.52 (3)	1.54 (5)	1.19 (1)	1.33 (9)	1.15 (7)
β	-	0.21 (19)	0.32 (4)	100 (10)	0.00025 (20)
ω_r	0.098 (5)	0.10 (5)	0.098 (4)	0.11 (5)	0.10 (5)
ω_{PSA_0}	1.57 (4)	1.58 (4)	1.57 (4)	1.55 (4)	1.56 (4)
ω_ε	1.35 (5)	1.28 (5)	1.34 (5)	1.22 (5)	1.36 (6)
$\omega_{T_{esc}}$	0.68 (5)	0.66 (5)	0.64 (5)	0.63 (5)	0.66 (5)
σ	0.38 (1)	0.38 (1)	0.38 (1)	0.38 (1)	0.38 (1)

Figure 1 Coverage probabilities of the 95% prediction intervals for PSA values (green) and risk of death (red) according to the landmark time s and the horizon time $t > 0$ in the simulated patients. The grey areas are the prediction intervals of 95%. The number of patients at risk in the simulated dataset is indicated at bottom for each landmark time.

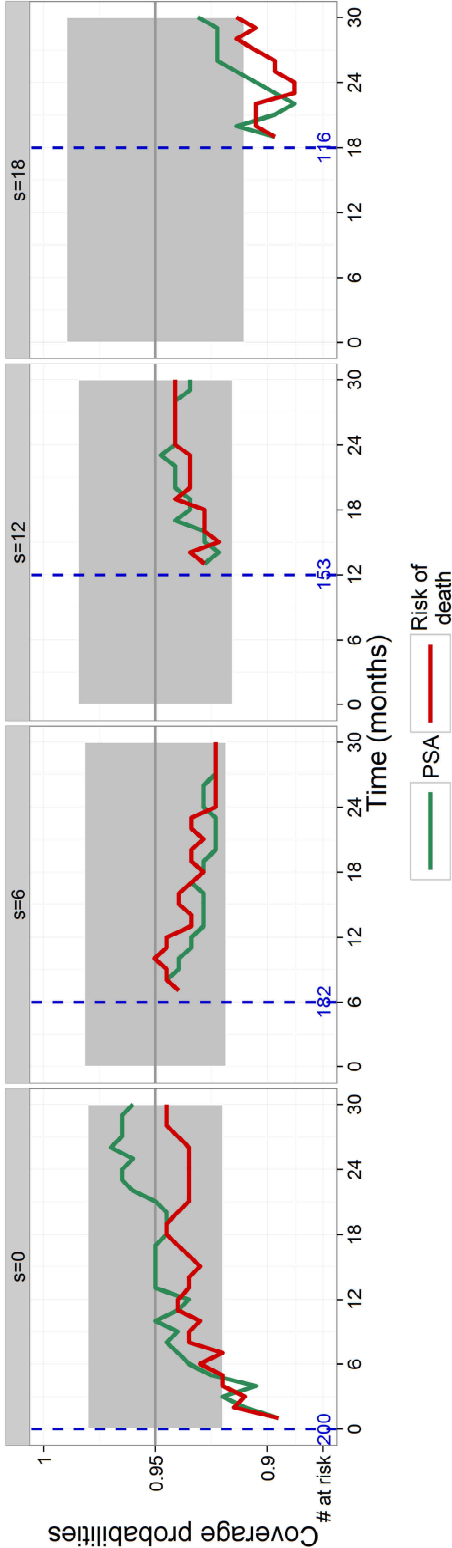


Figure 2 Time-dependent AUC in the simulated dataset of 200 patients for 4 values of landmark (months) and increasing values of horizon ($t > 2$ months). The number of patients at risk in the simulated dataset is indicated at bottom, as well as the median number [minimum-maximum] of PSA observations per patient at risk.

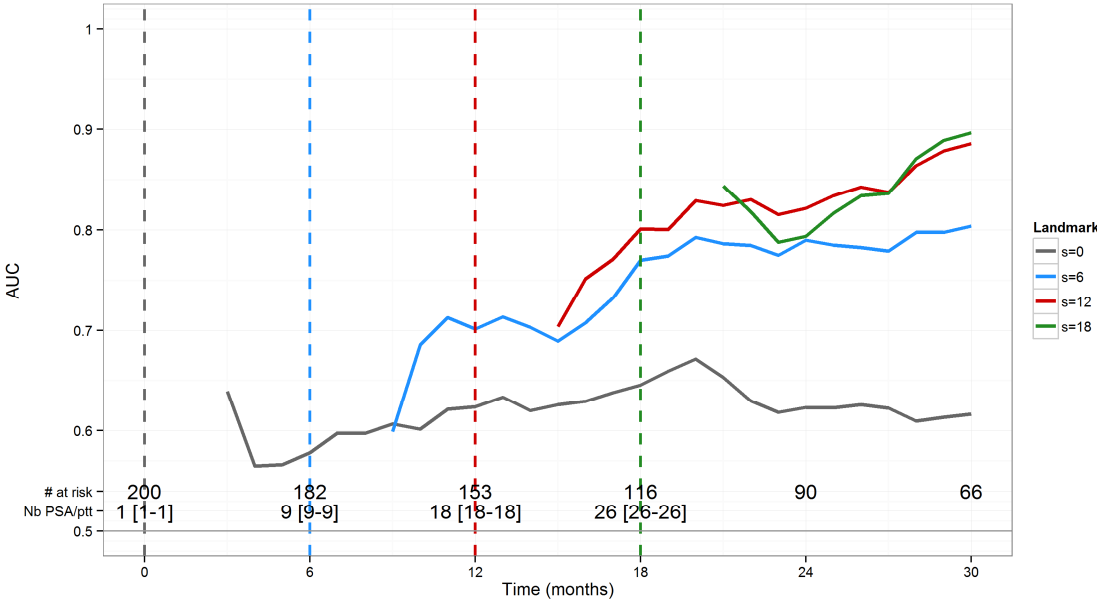


Figure 3 Time-dependent Brier Scores (top) and sBS (bottom) in the simulated dataset of 200 patients for 4 values of landmark time and increasing values of horizon ($t > 2$ months).

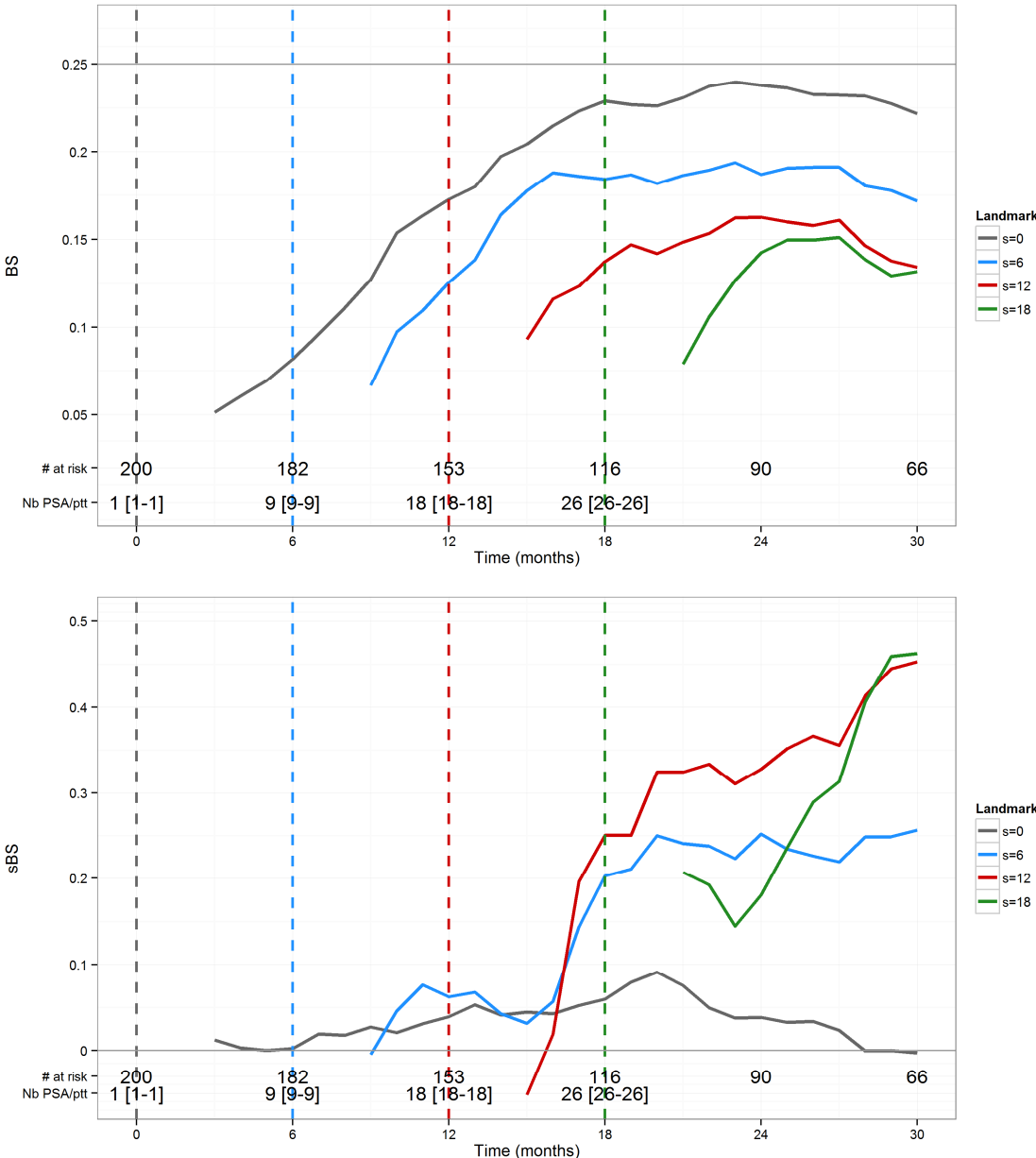


Figure 4 Dynamic individual predictions of PSA evolution and survival for 3 typical patients of validation dataset. The blue vertical lines indicate the landmark time s : left blue crosses denotes PSA measurements observed before s and used for predictions, the right black crosses denotes PSA measurements observed after s and not included for predictions. The green lines are the model prediction for PSA, the red lines the model prediction for survival function. The grey areas are the 95% prediction intervals. The black vertical lines indicate the vital status (solid=dead, dotted=censored).

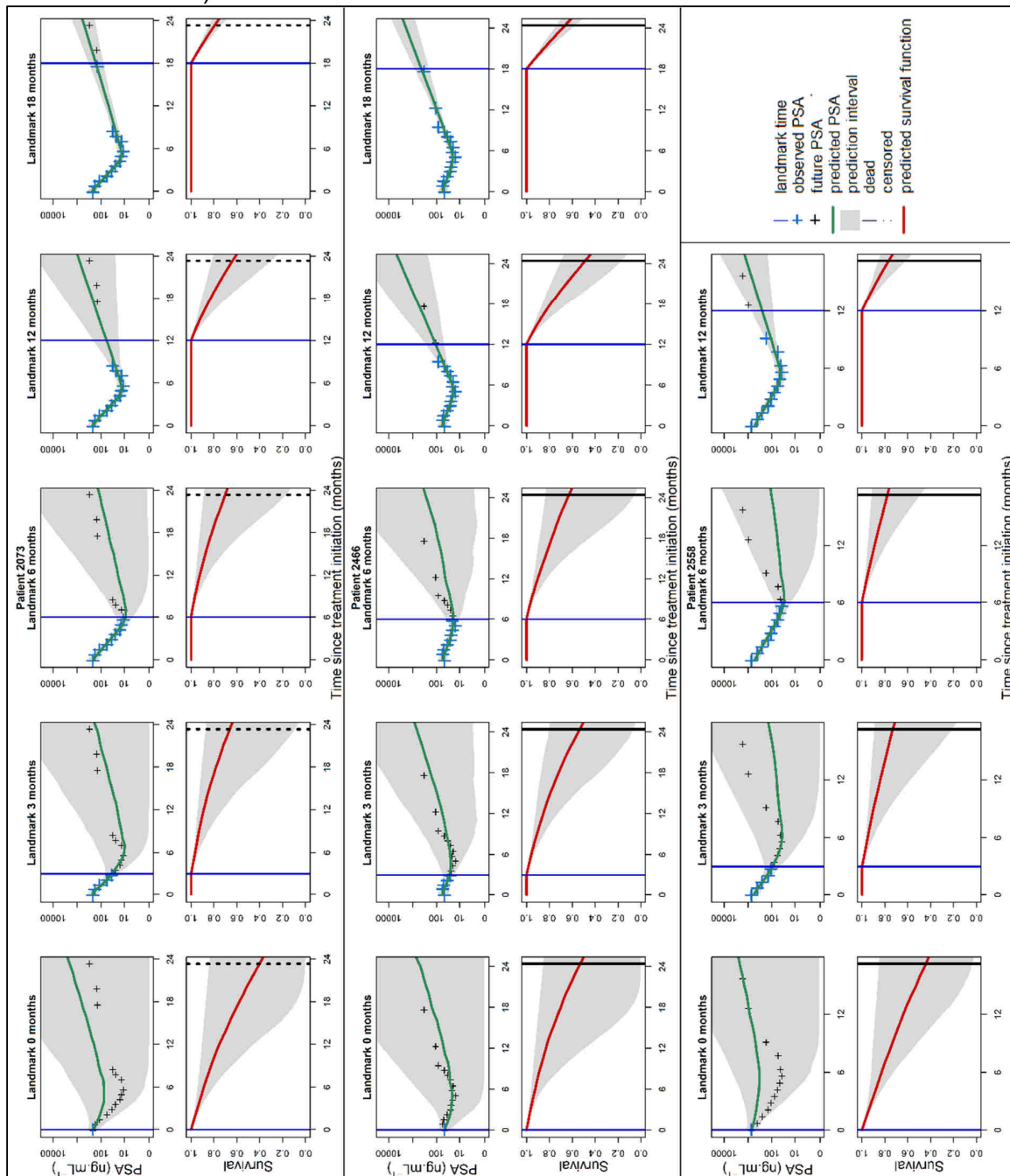


Figure 5 Time-dependent AUC obtained with the real data from N'=196 patients of the validation dataset for 4 values of landmark and increasing values of horizon. The number of patients at risk is indicated bottom, as well as the median number [minimum-maximum] of PSA observations per patient at risk.

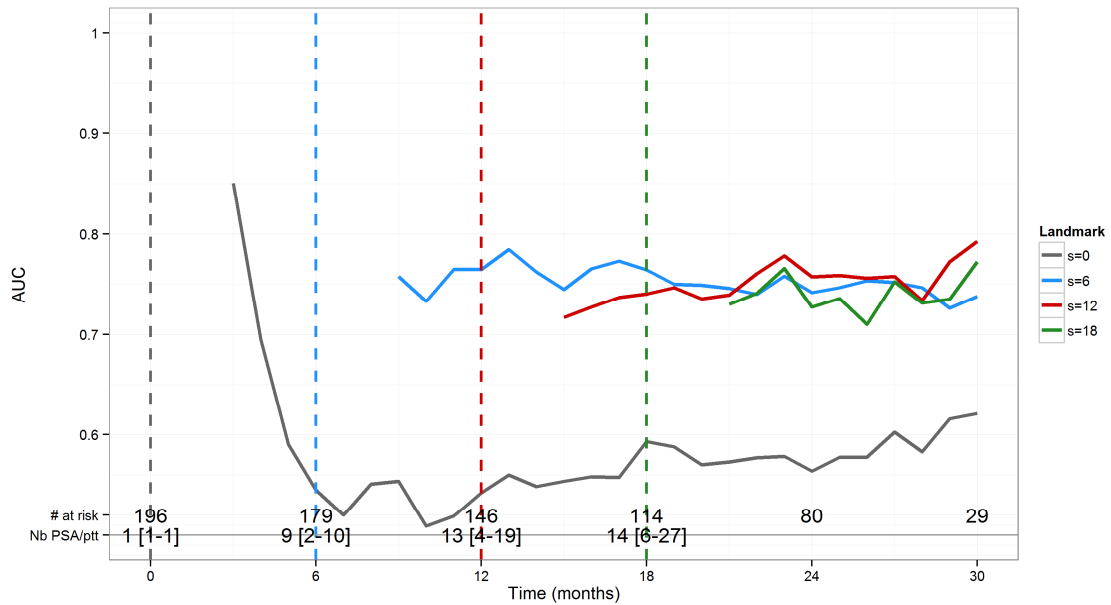
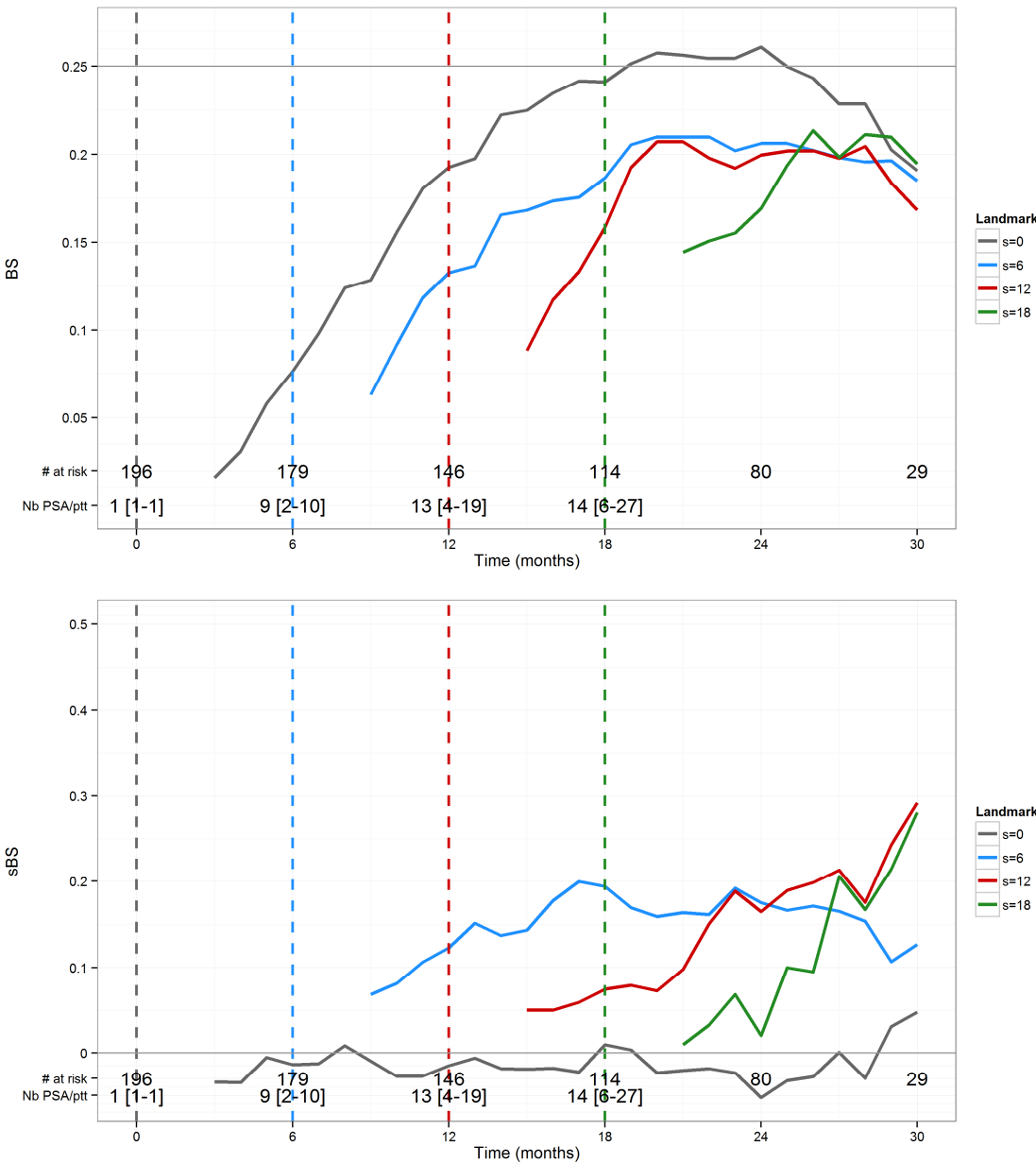


Figure 6 Time-dependent Brier Scores (top) and sBS (bottom) obtained with the real data from N'=196 patients of the validation dataset for 4 values of landmark time and increasing values of horizon.



Chapitre 5

Discussion et conclusion

L'objectif principal de ces travaux de thèse était d'étudier l'utilisation de modèles non-linéaires à effets mixtes dans l'analyse conjointe de données longitudinales et de données de survie. Dans un premier temps, nous avons montré que l'algorithme SAEM implémenté dans Monolix fournissait des estimations non biaisées des paramètres d'un modèle conjoint dont les données longitudinales étaient décrites par un MNLEM. Une fois cet outil évalué, nous avons pu développer un modèle conjoint mécanistique, reposant sur un système d'équations différentielles décrivant le mécanisme de production du PSA et son lien avec la survie chez des patients ayant un CPmHR et traités par la chimiothérapie de référence, docetaxel associé à de la prednisone. Enfin, nous avons montré comment faire de la prédiction dynamique individuelle et évaluer les performances prédictives des modèles conjoints non-linéaires, étendant des méthodes initialement développées dans un contexte de modèle linéaire à effets mixtes.

Ces trois projets contribuent au développement de modèles conjoints pour données longitudinales non-linéaires et données de survie, et offrent de nouvelles perspectives pour leur utilisation.

5.1 Discussion

Jusqu'à présent, la modélisation conjointe se limitait généralement à des processus longitudinaux décrits par des modèles linéaires à effets mixtes, du fait de la complexité numérique liée aux MNLEM. Le premier projet a donc été d'évaluer par simulations les propriétés statistiques des nouvelles fonctionnalités de l'algorithme SAEM de Monolix permettant l'estimation conjointe de paramètres relatifs à un MNLEM et à une fonction de survie. Cette étude de simulation considérait plusieurs scénarios dont le design était inspiré par les données réelles, avec différentes associations entre la valeur courante du biomarqueur et la survie. Le MNLEM avait pour modèle structurel la solution analytique d'un système d'ODEs décrivant de manière simplifiée le mécanisme de production du PSA, correspondant à une fonction biexponentielle mais dont les paramètres avaient une interprétation biologique. Les résultats obtenus par modélisation conjointe ont été comparés à ceux obtenus par les approches en deux étapes et "conjointe séquentielle", dont

l'idée d'origine vient du domaine des analyses PKPD (ZHANG et al., 2003) et pourrait attirer les pharmacométriciens pour des raisons techniques et de temps de calculs. La modélisation conjointe fournissait des estimations sans biais à la fois pour les paramètres relatifs au processus longitudinal et pour les paramètres relatifs à la fonction de survie. Au contraire, l'approche en deux étapes menait à des biais, en particulier, sous-estimait systématiquement β , la force du lien entre les deux processus. Ces résultats établis dans un contexte de MNLEM sont cohérents avec les conclusions d'études qui comparaient modèles conjoints et modèles en deux étapes dans le cadre de modèles linéaires à effets mixtes (IBRAHIM et al., 2010; SWEETING et THOMPSON, 2011). De plus, modèle conjoint et modèle "conjoint séquentiel" étaient satisfaisants en termes de risque de première espèce et de puissance à détecter un effet de l'évolution du biomarqueur sur la survie. Cela n'était pas le cas de la modélisation en deux étapes qui conduisait à une inflation du risque de première espèce pouvant s'expliquer par le fait que, dans la première étape, un phénomène de *shrinkage* (i.e. régression vers la moyenne) des EBEs individuels apparaît pour les patients décédant précocement (RIBBA et al., 2014a). Le modèle conjoint améliore également les estimations du processus longitudinal par rapport à l'approche en deux étapes, mais aussi par conséquent, par rapport à l'approche "conjointe séquentielle". Cela corrobore la nécessité d'avoir recours à des modèles conjoints également pour modéliser des données longitudinales en présence de *dropouts* (WU et al., 2011).

À l'initiation de cette thèse, la modélisation conjointe avec le logiciel Monolix n'avait pas encore été étudiée. Entre temps, MBOGNING et al. (2015) ont mené des études de simulations évaluant les performances de l'algorithme SAEM avec des exemples standards en pharmacométrie de modèles conjoints. De même, VIGAN (2014) a étudié par simulations les propriétés de l'algorithme SAEM dans le cadre d'évènements répétés. Ces études ont des conclusions similaires à nos travaux concernant les bonnes propriétés de l'algorithme SAEM pour estimer les paramètres de modèles conjoints non-linéaires.

Néanmoins, nous avons également noté certaines limites à l'utilisation de l'algorithme SAEM pour de la modélisation conjointe. L'estimation de la log-vraisemblance reste coûteuse en temps de calcul et, malgré cela, est fournie avec une erreur type (s.e.) non-négligeable, que ce soit dans notre étude de simulation ou sur les vraies données du second projet. Les méthodes reposant sur la vraisemblance semblent alors peu fiables pour identifier de petits effets sur la survie. Le test de Wald pourrait être une alternative, cependant l'étude de simulation a suggéré une tendance à sous-estimer les s.e. car elles étaient plus petites que les erreurs quadratiques moyennes (RMSE pour *root mean squared errors*). L'interprétation de ces tests doit donc rester prudente.

L'avantage principal de l'utilisation de MNLEM pour la modélisation conjointe est la possibilité d'utiliser des modèles plus physiologiques pour décrire l'évolution de biomarqueur en prenant en compte le mécanisme de la maladie et de l'éventuel traitement

à l'aide d'ODEs. Ainsi, dans cette thèse, nous avons développé un modèle conjoint mécanistique décrivant le lien entre la cinétique du PSA et la survie chez des patients ayant un CpmHR et traités par docetaxel. Le MNLEM repose sur un système d'ODEs établi d'après l'hypothèse de l'existence de cellules résistantes au docetaxel et de cellules sensibles productrices de PSA (SERUGA et al., 2011). D'autres modèles mécanistiques de la cinétique du PSA existent, comme celui de DE CHARRY et al. (2015) qui repose sur deux compartiments produisant du PSA : l'un composé de cellules saines, l'autre composé de cellules cancéreuses, mais utilisant des données PSA avant intervention chirurgicale, donc sans besoin de prendre en compte l'effet d'un traitement. Le modèle longitudinal proposé ici capture une grande variété de profils d'évolution du PSA, avant traitement, sous traitement et même après traitement, qui constitue l'originalité et la richesse de nos données.

Le phénomène d'émergence d'une résistance à la chimiothérapie est connu en oncologie mais rarement modélisé alors qu'il a un rôle important dans l'évolution de la maladie et dans l'évaluation de traitements anti-tumoraux car il réduit les chances de survie du patient (GATENBY, 2009; ENRIQUEZ-NAVAS et al., 2015). L'hypothèse de la présence de cellules résistantes au traitement avant même la mise sous traitement est reconnue en oncologie, du fait de nombreuses mutations des cellules tumorales (EASWARAN et al., 2014). Puis, nous avons appliqué les principes d'évolution darwiniens aux populations de cellules sensibles et résistantes. Notamment, les cellules résistantes acquérant un avantage évolutif en présence du traitement prolifèrent. La théorie de la dynamique darwinienne paraît intéressante à modéliser afin de proposer à l'avenir des schémas thérapeutiques innovants, reposant par exemple sur une stabilisation de la taille de la tumeur par équilibre du nombre de cellules sensibles et résistantes, plutôt que sur une diminution de la tumeur, souvent suivie d'un rebond dû à la résistance (ENRIQUEZ-NAVAS et al., 2015).

Même si la cinétique du PSA est correctement décrite par un modèle longitudinal dont les paramètres ont une interprétation mécanistique, certaines limites demeurent. Tout d'abord, pour des raisons d'identifiabilité, certains paramètres ont dû être fixés. Même si une étude de sensibilité (matériel supplémentaire A de l'article 2) et la littérature ont permis de fixer ces valeurs, il est toujours préférable d'estimer l'ensemble des paramètres sur les données. Tous les paramètres sont supposés constants au cours du temps, or il est possible que l'évolution de la maladie entraîne des variations de certaines fonctionnalités. De plus, un seul mécanisme d'action du docetaxel a pu être pris en compte, également pour des raisons d'identifiabilité, tandis que plusieurs effets concomitants du docetaxel sont connus (HERBST et KHURI, 2003) : le modèle sélectionné sur le critère BIC considère un effet inhibiteur sur la prolifération des cellules tumorales tandis qu'un autre effet possible testé du docetaxel est la stimulation de l'apoptose de ces cellules. Enfin, notre modèle reste un modèle mécanistique et non "biologique", dans le sens où on considère le corps comme un ensemble de compartiments interagissant, et non à une échelle microscopique

comme celle des cellules (ANDERSON et QUARANTA, 2008).

Notre modèle conjoint final exploite la nature mécanistique de la partie longitudinale, établissant un lien entre le nombre de cellules résistantes et sensibles, non-observées, et la survie chez les patients ayant un CPMHR. Le rôle des cellules résistantes est prépondérant, ce qui est cohérent avec un taux de croissance tumorale fortement prédictif du risque de décès (STEIN et al., 2011). La nature paramétrique du modèle final avec, en particulier, un risque instantané de base constant est également sujette à questions. Cette hypothèse paramétrique a été assouplie en considérant des fonctions splines simples (matériel supplémentaire B de l'article 2), mais aucune amélioration du critère BIC n'a été obtenue, suggérant qu'un modèle simple convenait pour ajuster nos données de survie. De plus, la pertinence du modèle de survie a été renforcée par le fait que les prédictions de la survie obtenues en utilisant uniquement les mesures de PSA des 196 patients du jeu de données de validation (non inclus pour la construction du modèle), se superposaient à la courbe de Kaplan-Meier observée chez ces patients.

Enfin, nous avons étudié la capacité des modèles conjoints non-linéaires à réaliser de la prédiction individuelle dynamique, exploitant des outils initialement développés dans le cadre linéaire. Afin de fournir non seulement une prédiction mais surtout un intervalle de prédiction, nous avons tenu compte de l'incertitude de l'estimation des paramètres individuels. Pour cela, supposant que le modèle conjoint développé sur un jeu de données de taille suffisante était le bon, nous avons obtenu des échantillons de Monte-Carlo des prédictions individuelles en utilisant une méthode bayésienne proposée par RIZOPOULOS (2011), sans prendre en compte toutefois l'incertitude sur l'estimation des paramètres de population. Cette approche permet, pour un nouveau patient ayant une ou plusieurs mesures de PSA, de prédire à la fois l'évolution future du biomarqueur et surtout celle de son risque de décès. Il a déjà été montré que la modélisation conjointe offrait un cadre idéal pour faire de la prédiction, et en particulier, donnait des prédictions plus précises que d'autres méthodes comme le *landmarking* (i.e. obtenir des probabilités de survivre à partir d'un modèle de Cox ajusté sur les patients du jeu de données initial vivants au temps *landmark* considéré) dans un contexte linéaire (RIZOPOULOS et al., 2013).

Les performances de ces prédictions devaient alors être évaluées en prenant en compte leur caractère dynamique. Nous avons choisi de nous concentrer sur trois métriques de discrimination et calibration du modèle, l'AUC, le score de Brier (BS) et le score de Brier normalisé (sBS). Nous avons obtenu des métriques dépendant du temps de *landmark* et de l'horizon de la prédiction, permettant de caractériser les performances des prédictions en fonction de la durée du suivi longitudinal et de l'horizon de la prédiction désirée. Les courbes d'AUC dynamiques sont relativement simples à interpréter et quantifient l'évolution de la discrimination en fonction de t . Les courbes de BS visent à décrire plus particulièrement les capacités de calibration du modèle, mais leur interprétation est moins

aisée, d'où l'introduction du sBS permettant de mesurer le gain du modèle par rapport à une prédiction plus naïve. Ces métriques ont l'avantage de ne pas dépendre du modèle de prédiction et ont donc été faciles à étendre au contexte non-linéaire. Elles nous ont montré que les performances prédictives du modèle s'amélioreraient lorsque le suivi du biomarqueur s'allongeait, du fait du phénomène de *shrinkage* lorsque les mesures sont encore peu nombreuses. Dans ce contexte, un suivi long d'au moins 12 mois pourrait être conseillé pour proposer des prédictions fiables pour un horizon allant jusqu'à 18 mois.

Nous avons ici procédé à une validation dite "interne" des capacités prédictives du modèle conjoint (tout comme dans le second projet), c'est-à-dire en utilisant 196 patients issus du même essai clinique que les 400 patients ayant servi à développer le modèle. Or, ce type de validation est critiquable. L'idéal pour être en mesure d'assurer la généralisation du modèle serait de faire une validation "externe", avec des données provenant d'un tout autre jeu de données. Une solution alternative ne nécessitant qu'un seul jeu de données et appartenant donc aux validations de type interne, aurait été de procéder à une validation croisée ("*cross-validation*"). Cela consiste à partitionner le jeu de données en K échantillons de même taille, de développer le modèle avec $(K - 1)$ échantillons et d'évaluer le modèle sur l'échantillon restant, puis de répéter cela K fois afin de calculer des métriques moyennes (AUC ou BS) (GERDS et al., 2008). Si $K = N$ avec N le nombre de patients, on parle de méthode "leave-one-out", mais cela est très long en temps de calculs. Le choix du nombre K est donc délicat et encore discuté car il faut trouver le bon compromis entre validation du modèle et temps de calculs raisonnable (ARLOT et al., 2010).

5.2 Perspectives

Plusieurs perspectives à ces travaux peuvent être envisagées, qu'elles soient méthodologiques ou appliquées.

D'un point de vue méthodologique, plusieurs études de simulation pourraient être menées afin de poursuivre l'évaluation de la modélisation conjointe avec l'algorithme SAEM. Une première consisterait à comparer les performances à estimer les paramètres d'un modèle conjoint non-linéaire de l'algorithme SAEM de Monolix avec celles de l'approximation de Laplace utilisée dans NONMEM. Cela a déjà été fait pour des MNLEM d'un côté (PLAN et al., 2012), et avec uniquement NONMEM pour des événements répétés de l'autre (KARLSSON et al., 2011), ce dernier concluant que SAEM devrait être privilégié lorsqu'il y a peu d'événements répétés. Il pourrait également être envisagé de comparer les performances de l'algorithme SAEM à celles d'approches bayésiennes telles que celle implémentée dans le programme NIMROD (PRAGUE et al., 2013).

Dans ce travail de thèse, nous nous sommes placés dans le cas d'un unique bras de

traitement pour étudier la modélisation conjointe non-linéaire et le lien entre un processus longitudinal et la survenue d'un évènement. Cependant, dans les essais cliniques, l'objectif est de comparer deux ou plusieurs bras de traitement et de détecter un effet du traitement étudié. Pour cela, il est donc nécessaire d'évaluer le risque de première espèce et la puissance à détecter un effet du traitement sur l'évolution du biomarqueur et/ou sur la survie en utilisant un modèle conjoint mécanistique, pour lequel on s'attend à mieux capturer le lien entre biomarqueur et survie. La modélisation conjointe pourrait également être utile pour diminuer le nombre de patients nécessaires pour détecter un effet du traitement sur la survie (ou l'absence d'effet), et/ou diminuer la durée de suivi des patients. Dans ce contexte d'essai clinique avec deux bras de traitement, les résultats obtenus avec une modélisation conjointe réalisée au cours de l'essai, pourront être comparés aux résultats obtenus avec une analyse classique réalisée dans l'industrie pharmaceutique à la fin de l'essai, comme un test du log-rank sur les données de survie seules (TANNOCK et al., 2013). Ensuite, des études sur le design optimal pourraient être réalisées. Le choix du protocole (nombre de patients par bras, nombre de prélèvements par patient et temps de prélèvement) ayant un impact sur la précision d'estimation des paramètres et sur la puissance des tests, il est donc nécessaire de l'optimiser, c'est-à-dire de choisir le protocole permettant l'estimation a priori la plus précise (et tenant compte des contraintes expérimentales et éthiques). Dans un contexte de modèle conjoint, on peut se demander si prendre en compte un temps d'évènement modifierait le design optimal. L'optimisation des temps de mesures (ou de doses) est possible en maximisant le déterminant de la matrice d'information de Fisher, en supposant connu le modèle, par une approche MCMC récemment développée pour des MNLEM (RIVIÈRE-JOURDAN et MENTRÉ, 2015) et des temps d'évènement (UECKERT et al., 2015). Néanmoins cette méthode n'a pas encore été évaluée avec des modèles conjoints.

Concernant les capacités prédictives des modèles conjoints mécanistiques, elles restent encore à approfondir. Les propriétés des métriques de discrimination et calibration utilisées, à savoir l'AUC, le BS et le sBS, ont déjà été souvent étudiées (BLANCHE et al., 2013a; STEYERBERG et al., 2010). En particulier, (BLANCHE et al., 2015) proposent des intervalles de confiance pour l'AUC et le BS dérivés de résultats asymptotiques, supposant une distribution normale de la métrique, et étendu à des intervalles de confiance simultanés afin de comparer plusieurs AUC ou BS (BLANCHE et al., 2015).

D'un point de vue clinique, le nombre important d'essais cliniques en cours pour la recherche de traitements pour le cancer de la prostate avancé (OUDARD, 2013) va offrir la possibilité d'approfondir les modèles proposés et de les valider. Tout d'abord, des stratégies de sélection du modèle conjoint devront être mises en place. En effet, une étude approfondie de l'influence à la fois sur l'évolution du biomarqueur et sur la survenue de l'évènement de facteurs d'ajustement, tels que l'âge ou la taille de la tumeur

à l'initiation du traitement, devra être menée. De plus, différentes formes de lien pourront être proposées, et par exemple avec un effet retardé du biomarqueur sur le risque de décès. Enfin, la partie longitudinale du modèle conjoint pourra être complexifiée afin de prendre en compte plus précisément les mécanismes de la maladie et du traitement, au travers d'équations différentielles. Pour cela, des données riches sont nécessaires, avec non seulement des mesures de PSA fréquentes chez les patients, mais, comme le PSA ne semble pas suffisant pour prédire la survie des patients ayant un CPmHR (COLLETTE et al., 2006), avec également l'apport de nouveaux biomarqueurs. Dans un tel modèle, nous pourrions bien sûr ajouter des données de tailles de la tumeur, déjà mesurée régulièrement dans le cadre de l'évaluation de l'évolution de la maladie par le critère RECIST (voir le paragraphe 1.1.4), mais également la densité de la tumeur, ou son activité, pouvant être évaluées par des techniques d'imagerie de plus en plus performantes. En effet, dans le développement de nouvelles thérapies ciblées (*targeted therapies*), il a été observé que la taille de la tumeur ne diminuait pas nécessairement immédiatement après l'initiation du traitement mais que sa densité et son activité diminuaient. Dans ce contexte, il a été montré qu'un critère composite de la taille de la tumeur, de sa densité et de sa vascularisation serait préférable au critère RECIST (CHOI, 2008). Enfin, on peut également évoquer l'ajout de données de nouveaux biomarqueurs pour améliorer le modèle comme les *CTCs* (*Circulating Tumor Cells*) ou cellules tumorales circulantes. Les CTCs sont des cellules qui se sont détachées de la tumeur et qui transitent dans le sang, elles peuvent être à l'origine du développement de métastases. Leur concentration étant très faible dans le sang, il a fallu attendre de récents progrès techniques pour mettre au point une méthode de détection (LIBERKO et al., 2013). Les premières études réalisées décrivent les CTCs comme un bon "*surrogate*" de la survie, c'est-à-dire un biomarqueur pouvant être mesuré facilement, précocement et surtout reflétant correctement le risque de décès, chez les patients ayant un CPmHR même si ces résultats doivent encore être confirmés (DE BONO et al., 2008; SCHER et al., 2015). Dans le cadre d'essais cliniques comparant plusieurs bras de traitement, des données de PK pourront être prises en compte. L'effet du traitement, au lieu d'être supposé constant comme dans notre modèle, dépendrait alors chez chaque patient de la dose et des temps d'administration.

La modélisation conjointe non-linéaire ouvre donc la voie à des modèles prenant en compte plus de biomarqueurs et pouvant potentiellement améliorer les capacités prédictives. Dans le troisième projet de cette thèse (Chapitre 4) sur la prédiction individuelle dynamique, nous sommes revenus à un modèle conjoint dont le MNLEM avait un modèle structurel sous une forme analytique, cela pour des raisons techniques (la version 2.8 du logiciel Stan qui nous a permis d'obtenir les distributions des effets aléatoires individuels ne pouvant gérer des ODEs "raides" (*stiff*)). Un futur travail sera donc d'étudier les performances prédictives d'un modèle conjoint mécanistique défini par des ODEs (par exemple celui développé dans le chapitre 3), ce qui devrait être possible

avec la version 2.9 de Stan récemment mise à disposition. Cela permettrait en plus, avec les données pré-traitement du patient, de prédire sa survie en fonction du temps d'initiation du traitement.

Enfin, le développement de modèles mécanistiques intégrant toutes les connaissances sur la progression de la maladie et les mécanismes d'action des traitements et leur lien avec la survie, permettrait de proposer et tester *in silico* des schémas thérapeutiques innovants, afin toujours d'améliorer le traitement des patients (ANDERSON et QUARANTA, 2008).

5.3 Conclusion générale

Les modèles conjoints pour données longitudinales et de survie ont connu une expansion importante ces dernières années. Souvent limités à des modèles linéaires pour décrire le processus longitudinal à cause de difficultés numériques, leur besoin se fait pourtant croissant en pharmacométrie où la cinétique des biomarqueurs est non-linéaire.

Dans ce contexte, ces travaux de thèse ouvrent la voie au développement de modèles conjoints mécanistiques, plus physiologiques, et qui permettent de tenir compte de l'influence de plusieurs biomarqueurs sur la survie, au moyen d'équations différentielles. Cette approche méthodologique offre des outils pratiques au-delà du contexte du cancer de la prostate métastatique hormono-résistant, dans d'autres aires thérapeutiques telles que le diabète de type 2 ou des maladies infectieuses où la cinétique de biomarqueurs peut être décrite par des équations différentielles. Nous pensons que la modélisation conjointe permettra de mieux comprendre le lien entre l'évolution de biomarqueurs et la survenue d'un évènement afin d'améliorer et d'optimiser l'évaluation thérapeutique et la prédiction d'évènement chez les patients pour leur proposer le traitement le plus adapté.

Bibliographie

1. ANDERSEN P.K., GILL R.D., Cox's regression model for counting processes : a large sample study. *The Annals of Statistics*, 1982, p. 1100–1120
2. ANDERSON A.R., QUARANTA V., Integrative mathematical oncology. *Nature Reviews Cancer*, 2008, vol. 8, p. 227–234
3. ANDRINOPOULOU E.R., RIZOPOULOS D., GELEIJNSE M.L., LESAFFRE E., BOGERS A.J., TAKKENBERG J.J., Dynamic prediction of outcome for patients with severe aortic stenosis : application of joint models for longitudinal and time-to-event data. *BMC Cardiovascular Disorders*, 2015, vol. 15, p. 1–8
4. ANDRINOPOULOU E.R., RIZOPOULOS D., TAKKENBERG J.J., LESAFFRE E., Joint modeling of two longitudinal outcomes and competing risk data. *Statistics in Medicine*, 2014, vol. 33, p. 3167–3178
5. ANDRIOLE G.L., CRAWFORD E.D., GRUBB III R.L., BUYS S.S., CHIA D., CHURCH T.R., FOUAD M.N., GELMANN E.P., KVALE P.A., REDING D.J., WEISSFELD J.L., YOKOCHI L.A., O'BRIEN B., CLAPP J.D., RATHMELL J.M., RILEY T.L., HAYES R.B., KRAMER B.S., IZMIRLIAN G., MILLER A.B., PINSKY P.F., PROROK P.C., GOHAGAN J.K., BERG C.D., Mortality results from a randomized prostate-cancer screening trial. *New England Journal of Medicine*, 2009, vol. 360, p. 1310–1319
6. ARLOT S., CELISSE A., et al., A survey of cross-validation procedures for model selection. *Statistics surveys*, 2010, vol. 4, p. 40–79
7. ARMSTRONG A.J., EISENBERGER M.A., HALABI S., OUDARD S., NANUS D.M., PETRYLAK D.P., SARTOR A.O., SCHER H.I., Biomarkers in the management and treatment of men with metastatic castration-resistant prostate cancer. *European Urology*, 2012, vol. 61, p. 549–559
8. ASAR Ö., RITCHIE J., KALRA P.A., DIGGLE P.J., Joint modelling of repeated measurement and time-to-event data : an introductory tutorial. *International Journal of Epidemiology*, 2015, vol. 44, p. 334–344
9. BERTHOLD D.R., POND G.R., SOBAN F., DE WIT R., EISENBERGER M., TANNOCK I.F., Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer : updated survival in the TAX 327 study. *Journal of Clinical Oncology*, 2008, vol. 26, p. 242–245
10. BILL-AXELSON A., HOLMBERG L., GARMO H., RIDER J.R., TAARI K., BUSCH C., NORDLING S., HÄGGMAN M., ANDERSSON S.O., SPÅNGBERG A., ANDREN O., PALMGREN J., STEINECK G., ADAMI H.O., JOHANSSON J.E., Radical prostatectomy or watchful waiting in early prostate cancer. *New England Journal of Medicine*, 2014, vol. 370, p. 932–942

11. BJÖRNSSON M.A., FRIBERG L.E., SIMONSSON U.S., Performance of nonlinear mixed effects models in the presence of informative dropout. *The AAPS Journal*, 2015, vol. 17, p. 245–255
12. BLANCHE P., Package 'timeROC' : time-dependent ROC curve and AUC for censored survival data. *Vienna, Austria : R Foundation for Statistical Computing*, 2013
13. BLANCHE P., DARTIGUES J.F., JACQMIN-GADDA H., Review and comparison of ROC curve estimators for a time-dependent outcome with marker-dependent censoring. *Biometrical Journal*, 2013a, vol. 55, p. 687–704
14. BLANCHE P., LATOUCHE A., VIALLOU V., Time-dependent AUC with right-censored data : a survey. In *Risk Assessment and Evaluation of Predictions*, p. 239–251, Springer, 2013b
15. BLANCHE P., PROUST-LIMA C., LOUBÈRE L., BERR C., DARTIGUES J.F., 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, vol. 71, p. 102–113
16. BONATE P.L., SUTTLE B., Effect of censoring due to progressive disease on tumor size kinetic parameter estimates. *The AAPS Journal*, 2013, vol. 15, p. 832–839
17. BRIER G.W., Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 1950, vol. 78, p. 1–3
18. CHI Y.Y., IBRAHIM J.G., Joint models for multivariate longitudinal and multivariate survival data. *Biometrics*, 2006, vol. 62, p. 432–445
19. CHOI H., Response evaluation of gastrointestinal stromal tumors. *The oncologist*, 2008, vol. 13, p. 4–7
20. CHOI J.H., AHN M.J., RHIM H.C., KIM J.W., LEE G.H., LEE Y.Y., KIM I.S., Comparison of WHO and RECIST criteria for response in metastatic colorectal carcinoma. *Cancer Research and Treatment*, 2005, vol. 37, p. 290–293
21. COLLETTE L., BURZYKOWSKI T., SCHRÖDER F.H., Prostate-specific antigen (PSA) alone is not an appropriate surrogate marker of long-term therapeutic benefit in prostate cancer trials. *European Journal of Cancer*, 2006, vol. 42, p. 1344–1350
22. COMETS E., LAVENU A., LAVIELLE M., saemix : stochastic approximation expectation maximization SAEM algorithm. *R package version 10*, 2011
23. COMMENGES D., LIQUET B., PROUST-LIMA C., Choice of prognostic estimators in joint models by estimating differences of expected conditional Kullback–Leibler risks. *Biometrics*, 2012, vol. 68, p. 380–387
24. COX D.R., Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, 1972, vol. 34, p. 187–220
25. DAFNI U.G., TSIATIS A.A., Evaluating surrogate markers of clinical outcome when measured with error. *Biometrics*, 1998, vol. 54, p. 1445–1462

26. DALL'ERA M.A., ALBERTSEN P.C., BANGMA C., CARROLL P.R., CARTER H.B., COOPERBERG M.R., FREEDLAND S.J., KLOTZ L.H., PARKER C., SOLOWAY M.S., Active surveillance for prostate cancer : a systematic review of the literature. *European Urology*, 2012, vol. 62, p. 976–983
27. D'AMICO A.V., WHITTINGTON R., MALKOWICZ S.B., SCHULTZ D., BLANK K., BRODERICK G.A., TOMASZEWSKI J.E., RENSHAW A.A., KAPLAN I., BEARD C.J., WEIN A., Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *Journal of the American Medical Association*, 1998, vol. 280, p. 969–974
28. DARTOIS C., BRENDEL K., COMETS E., LAFFONT C., LAVELLE C., TRANCHAND B., MENTRÉ F., LEMENUÉL-DIOT A., GIRARD P., Overview of model-building strategies in population PK/PD analyses : 2002–2004 literature survey. *British Journal of Clinical Pharmacology*, 2007, vol. 64, p. 603–612
29. DE BONO J.S., OUDARD S., OZGUROGLU M., HANSEN S., MACHIELS J.P., KOCAK I., GRAVIS G., BODROGI I., MACKENZIE M.J., SHEN L., ROESSNER M., GUPTA S., SARTOR A.O., Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment : a randomised open-label trial. *The Lancet*, 2010, vol. 376, p. 1147–1154
30. DE BONO J.S., SCHER H.I., MONTGOMERY R.B., PARKER C., MILLER M.C., TISSING H., DOYLE G.V., TERSTAPPEN L.W., PIANTA K.J., RAGHAVAN D., Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. *Clinical Cancer Research*, 2008, vol. 14, p. 6302–6309
31. DE CHARRY F., COLOMBAN O., YOU B., RUFFION A., PAPAREL P., WILBAUX M., TOD M., FREYER G., PERRIN P., Identification of the most aggressive carcinomas among patients diagnosed with prostate cancers using mathematical modeling of PSA growth. *Clinical Genitourinary Cancer*, 2015, (sous presse)
32. DELYON B., LAVIELLE M., MOULINES E., Convergence of a stochastic approximation version of the EM algorithm. *Annals of Statistics*, 1999, vol. 27, p. 94–128
33. DEMPSTER A.P., LAIRD N.M., RUBIN D.B., Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B (methodological)*, 1977, vol. 39, p. 1–38
34. DESMÉE S., MENTRÉ F., VEYRAT-FOLLET C., GUEDJ J., 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*, 2015, vol. 17, p. 691–699
35. DESMÉE S., MENTRÉ F., VEYRAT-FOLLET C., SÉBASTIEN B., GUEDJ J., Using the SAEM algorithm for mechanistic joint models characterizing the relationship between nonlinear PSA kinetics and survival in prostate cancer patients. *Biometrics*, 2016
36. EASWARAN H., TSAI H.C., BAYLIN S.B., Cancer epigenetics : tumor heterogeneity, plasticity of stem-like states, and drug resistance. *Molecular Cell*, 2014, vol. 54, p. 716–727

37. EISENHAUER E., THERASSE P., BOGAERTS J., SCHWARTZ L., SARGENT D., FORD R., DANCEY J., ARBUCK S., GWYTHYER S., MOONEY M., RUBINSTEIN L., SHANKAR L., DODD L., KAPLAN R., LACOMBE D., VERWEIJ J., New response evaluation criteria in solid tumours : revised RECIST guideline (version 1.1). *European Journal of Cancer*, 2009, vol. 45, p. 228–247
38. ELASHOFF R.M., LI G., LI N., A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics*, 2008, vol. 64, p. 762–771
39. ENRIQUEZ-NAVAS P.M., WOJTKOWIAK J.W., GATENBY R.A., Application of evolutionary principles to cancer therapy. *Cancer Research*, 2015, vol. 75, p. 4675–4680
40. FAUCETT C.L., THOMAS D.C., Simultaneously modelling censored survival data and repeatedly measured covariates : a Gibbs sampling approach. *Statistics in Medicine*, 1996, vol. 15, p. 1663–1685
41. FERLAY J., SHIN H.R., BRAY F., FORMAN D., MATHERS C., PARKIN D.M., Estimates of worldwide burden of cancer in 2008 : GLOBOCAN 2008. *International Journal of Cancer*, 2010, vol. 127, p. 2893–2917
42. FERLAY J., SOERJOMATARAM I., DIKSHIT R., ESER S., MATHERS C., REBELO M., PARKIN D.M., FORMAN D., BRAY F., Cancer incidence and mortality worldwide : sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, 2015, vol. 136, p. 359–386
43. FERLAY J., SOERJOMATARAM I., M E., DIKSHIT R., ESER S., MATHERS C., REBELO M., PARKIN D.M., FORMAN D., BRAY F., GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide : IARC CancerBase No. 11 [internet]. Technical report, Lyon, France : International Agency for Research on Cancer, 2013, <http://globocan.iarc.fr>
44. FERRER L., RONDEAU V., DIGNAM J.J., PICKLES T., JACQMIN-GADDA H., PROUST-LIMA C., Joint modelling of longitudinal and multi-state processes : application to clinical progressions in prostate cancer. *arXiv preprint arXiv :150607496*, 2015
45. FITZPATRICK J.M., BANU E., OUDARD S., Prostate-specific antigen kinetics in localized and advanced prostate cancer. *British Journal of Urology International*, 2009, vol. 103, p. 578–587
46. FLEMING M.T., MORRIS M.J., HELLER G., SCHER H.I., Post-therapy changes in PSA as an outcome measure in prostate cancer clinical trials. *Nature Clinical Practice Oncology*, 2006, vol. 3, p. 658–667
47. GARCIA-HERNANDEZ A., RIZOPOULOS D., %JM Macro v2.01 Reference Manual. Available at : www.jm-macro.com, 2015
48. GATENBY R.A., A change of strategy in the war on cancer. *Nature*, 2009, vol. 459, p. 508–509
49. GE Z., BICKEL P.J., RICE J.A., An approximate likelihood approach to nonlinear mixed effects models via spline approximation. *Computational Statistics & Data Analysis*, 2004, vol. 46, p. 747–776

50. GERDS T.A., CAI T., SCHUMACHER M., The performance of risk prediction models. *Biometrical Journal*, 2008, vol. 50, p. 457–479
51. GERDS T.A., SCHUMACHER M., Consistent estimation of the expected Brier score in general survival models with right-censored event times. *Biometrical Journal*, 2006, vol. 48, p. 1029–1040
52. GOLUB G.H., WELSCH J.H., Calculation of gauss quadrature rules. *Mathematics of computation*, 1969, vol. 23, p. 221–230
53. GOSSELAAR C., ROOBOL M.J., ROEMELING S., SCHRÖDER F.H., The role of the digital rectal examination in subsequent screening visits in the European randomized study of screening for prostate cancer (ERSPC), Rotterdam. *European Urology*, 2008, vol. 54, p. 581–588
54. GRAF E., SCHMOOR C., SAUERBREI W., SCHUMACHER M., Assessment and comparison of prognostic classification schemes for survival data. *Statistics in Medicine*, 1999, vol. 18, p. 2529–2545
55. GUEDJ J., THIÉBAUT R., COMMENGES D., Joint modeling of the clinical progression and of the biomarkers' dynamics using a mechanistic model. *Biometrics*, 2011, vol. 67, p. 59–66
56. GUO X., CARLIN B.P., Separate and Joint Modeling of Longitudinal and Event Time Data Using Standard Computer Packages. *The American Statistician*, 2004, vol. 58, p. 16–24
57. HANLEY J.A., MCNEIL B.J., The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 1982, vol. 143, p. 29–36
58. HARRELL F.E., LEE K.L., MARK D.B., Tutorial in biostatistics multivariable prognostic models : issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, 1996, vol. 15, p. 361–387
59. HEAGERTY P.J., LUMLEY T., PEPE M.S., Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*, 2000, vol. 56, p. 337–344
60. HEAGERTY P.J., ZHENG Y., Survival model predictive accuracy and ROC curves. *Biometrics*, 2005, vol. 61, p. 92–105
61. HEIDENREICH A., ABRAHAMSSON P.A., ARTIBANI W., CATTO J., MONTORSI F., VAN POPPEL H., WIRTH M., MOTTET N., Early detection of prostate cancer : European Association of Urology recommendation. *European Urology*, 2013, vol. 64, p. 347–354
62. HEIDENREICH A., BASTIAN P.J., BELLMUNT J., BOLLA M., JONIAU S., VAN DER KWAST T., MASON M., MATVEEV V., WIEGEL T., ZATTONI F., MOTTET N., EAU guidelines on prostate cancer. Part 1 : Screening, diagnosis, and local treatment with curative intent - update 2013. *European Urology*, 2014a, vol. 65, p. 124–137
63. HEIDENREICH A., BASTIAN P.J., BELLMUNT J., BOLLA M., JONIAU S., VAN DER KWAST T., MASON M., MATVEEV V., WIEGEL T., ZATTONI F., MOTTET N., EAU guidelines on prostate cancer. Part II : Treatment of advanced, relapsing, and castration-resistant prostate cancer. *European Urology*, 2014b, vol. 65, p. 467–479

64. HEIJNSDIJK E., DER KINDEREN A., WEVER E., DRAISMA G., ROOBOL M., DE KONING H., Overdetection, overtreatment and costs in prostate-specific antigen screening for prostate cancer. *British Journal of Cancer*, 2009, vol. 101, p. 1833–1838
65. HENDERSON R., DIGGLE P., DOBSON A., Joint modelling of longitudinal measurements and event time data. *Biostatistics*, 2000, vol. 1, p. 465–480
66. HERBST R.S., KHURI F.R., Mode of action of docetaxel – a basis for combination with novel anticancer agents. *Cancer Treatment Reviews*, 2003, vol. 29, p. 407–415
67. HESS K.R., VARADHACHARY G.R., TAYLOR S.H., WEI W., RABER M.N., LENZI R., ABBRUZZESE J.L., Metastatic patterns in adenocarcinoma. *Cancer*, 2006, vol. 106, p. 1624–1633
68. HILDEN J., GERDS T.A., A note on the evaluation of novel biomarkers : do not rely on integrated discrimination improvement and net reclassification index. *Statistics in Medicine*, 2014, vol. 33, p. 3405–3414
69. HOLFORD N., Simultaneous Modelling of Disease Progression and Time to Event with NONMEM. *PAGE meeting*, 2005, <http://www.page-meeting.org/page/page2005/PAGE2005P80.pdf>
70. HOWLANDER N., NOONE A., KRAPCHO M., GARSHELL J., MILLER D., ALTEKRUSE S., KOSARY C., YU M., RUHL J., TATALOVICH Z., MARIOTTO A., LEWIS D., CHEN H., FEUER E., CRONIN K., SEER cancer statistics review, 1975–2011. Technical report, National Cancer Institute, 2013, http://seer.cancer.gov/csr/1975_2011/
71. HSIEH F., TSENG Y.K., WANG J.L., Joint modeling of survival and longitudinal data : likelihood approach revisited. *Biometrics*, 2006, vol. 62, p. 1037–1043
72. HU C., SALE M.E., A joint model for nonlinear longitudinal data with informative dropout. *Journal of Pharmacokinetics and Pharmacodynamics*, 2003, vol. 30, p. 83–103
73. HU C., SZAPARY P.O., YEILDING N., ZHOU H., Informative dropout modeling of longitudinal ordered categorical data and model validation : application to exposure–response modeling of physician’s global assessment score for ustekinumab in patients with psoriasis. *Journal of Pharmacokinetics and Pharmacodynamics*, 2011, vol. 38, p. 237–260
74. HU W., LI G., LI N., A bayesian approach to joint analysis of longitudinal measurements and competing risks failure time data. *Statistics in Medicine*, 2009, vol. 28, p. 1601–1619
75. HUANG X., LI G., ELASHOFF R.M., PAN J., A general joint model for longitudinal measurements and competing risks survival data with heterogeneous random effects. *Lifetime Data Analysis*, 2011, vol. 17, p. 80–100
76. HUNG H., CHIANG C.T., Estimation methods for time-dependent AUC models with survival data. *Canadian Journal of Statistics*, 2010, vol. 38, p. 8–26
77. IBRAHIM J.G., CHU H., CHEN L.M., Basic concepts and methods for joint models of longitudinal and survival data. *Journal of Clinical Oncology*, 2010, vol. 28, p. 2796–2801

78. ILIC D., O'CONNOR D., GREEN S., WILT T., Screening for prostate cancer : a Cochrane systematic review. *Cancer Causes & Control*, 2007, vol. 18, p. 279–285
79. JÁCOME-PITA F., SÁNCHEZ-SALAS R., BARRET E., AMARUCH N., GONZALEZ-ENGUITA C., CATHELINÉAU X., Focal therapy in prostate cancer : the current situation. *Ecancermedicalscience*, 2014, vol. 8
80. JACQMIN-GADDA H., THIÉBAUT R., DARTIGUES J.F., Modélisation conjointe de données longitudinales quantitatives et de délais censurés : Joint modeling of quantitative longitudinal data and censored survival time. *Revue d'Épidémiologie et de Santé Publique*, 2004, vol. 52, p. 502–510
81. JEMAL A., BRAY F., CENTER M.M., FERLAY J., WARD E., FORMAN D., Global cancer statistics. *CA : A Cancer Journal for Clinicians*, 2011, vol. 61, p. 69–90
82. KAPLAN E.L., MEIER P., Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 1958, vol. 53, p. 457–481
83. KAPLAN I.D., COX R.S., BAGSHAW M.A., A model of prostatic carcinoma tumor kinetics based on prostate specific antigen levels after radiation therapy. *Cancer*, 1991, vol. 68, p. 400–405
84. KARLSSON K.E., PLAN E.L., KARLSSON M.O., Performance of three estimation methods in repeated time-to-event modeling. *The AAPS Journal*, 2011, vol. 13, p. 83–91
85. KENNEDY E.H., TAYLOR J.M., SCHAUBEL D.E., WILLIAMS S., The effect of salvage therapy on survival in a longitudinal study with treatment by indication. *Statistics in Medicine*, 2010, vol. 29, p. 2569–2580
86. KLEIN J.P., MOESCHBERGER M.L., *Survival analysis : techniques for censored and truncated data*. Springer Science & Business Media, 2005
87. KOBORI Y., KITAGAWA Y., MIZOKAMI A., KOMATSU K., NAMIKI M., Free-to-total prostate-specific antigen (PSA) ratio contributes to an increased rate of prostate cancer detection in a Japanese population screened using a PSA level of 2.1–10.0 ng/ml as a criterion. *International Journal of Clinical Oncology*, 2008, vol. 13, p. 229–232
88. KORN E.L., SIMON R., Measures of explained variation for survival data. *Statistics in Medicine*, 1990, vol. 9, p. 487–503
89. KRÓL A., FERRER L., PIGNON J.P., PROUST-LIMA C., DUCREUX M., BOUCHÉ O., MICHIELS S., RONDEAU V., Joint model for left-censored longitudinal data, recurrent events and terminal event : Predictive abilities of tumor burden for cancer evolution with application to the FFCD 2000-05 trial. *Biometrics*, 2016
90. KRONIK N., KOGAN Y., ELISHMERENI M., HALEVI-TOBIAS K., VUK-PAVLOVIĆ S., AGUR Z., Predicting outcomes of prostate cancer immunotherapy by personalized mathematical models. *PLoS One*, 2010, vol. 5, p. e15482
91. KRUMHOLTZ J.S., CARVALHAL G.F., RAMOS C.G., SMITH D.S., THORSON P., YAN Y., HUMPHREY P.A., ROEHL K.A., CATALONA W.J., Prostate-specific antigen cutoff of 2.6 ng/ml for prostate cancer screening is associated with favorable pathologic tumor features. *Urology*, 2002, vol. 60, p. 469–473

92. KUHN E., LAVIELLE M., Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis*, 2005, vol. 49, p. 1020–1038
93. LAVIELLE M., *Mixed effects models for the population approach : models, tasks, methods and tools*. CRC Press, 2014
94. LAVIELLE M., SAMSON A., KARINA FERMIN A., MENTRÉ F., Maximum Likelihood Estimation of Long-Term HIV Dynamic Models and Antiviral Response. *Biometrics*, 2011, vol. 67, p. 250–259
95. LAW N.J., TAYLOR J.M., SANDLER H., The joint modeling of a longitudinal disease progression marker and the failure time process in the presence of cure. *Biostatistics*, 2002, vol. 3, p. 547–563
96. LEITZMANN M.F., ROHRMANN S., Risk factors for the onset of prostatic cancer : age, location, and behavioral correlates. *Clinical Epidemiology*, 2012, vol. 4, p. 1–11
97. LHEUREUX S., JOLY F., Cabazitaxel après docétaxel, une nouvelle option dans le cancer métastatique de la prostate résistant à la castration. *Bulletin du Cancer*, 2012, vol. 99, p. 875–880
98. LIBERKO M., KOLOSTOVA K., BOBEK V., Essentials of circulating tumor cells for clinical research and practice. *Critical Reviews in Oncology/Hematology*, 2013, vol. 88, p. 338–356
99. LILJA H., ULMERT D., VICKERS A.J., Prostate-specific antigen and prostate cancer : prediction, detection and monitoring. *Nature Reviews Cancer*, 2008, vol. 8, p. 268–278
100. LIN H., TURNBULL B.W., MCCULLOCH C.E., SLATE E.H., Latent class models for joint analysis of longitudinal biomarker and event process data : application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association*, 2002, vol. 97, p. 53–65
101. LINDSTROM M.J., BATES D.M., Nonlinear mixed effects models for repeated measures data. *Biometrics*, 1990, p. 673–687
102. LIU L., HUANG X., Joint analysis of correlated repeated measures and recurrent events processes in the presence of death, with application to a study on acquired immune deficiency syndrome. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 2009, vol. 58, p. 65–81
103. LUBOLDT H.J., SCHINDLER J.F., RÜBBEN H., Age-specific reference ranges for prostate-specific antigen as a marker for prostate cancer. *EAU-EBU Ipdte Series*, 2007, vol. 5, p. 38–48
104. MAUGUEN A., RACHET B., MATHOULIN-PÉLISSIER S., LAWRENCE G.M., SIESLING S., MACGROGAN G., LAURENT A., RONDEAU V., Validation of death prediction after breast cancer relapses using joint models. *BMC Medical Eesearch Methodology*, 2015, vol. 15, p. 1
105. MBOGNING C., BLEAKLEY K., LAVIELLE M., 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*, 2015, vol. 85, p. 1512–1528

106. MOTTET N., BASTIAN P., BELLMUNT J., VAN DEN BERGH R., BOLLA M., VAN CASTEREN N., CORNFORD P., JONIAU S., MASON M., MATVEEV V., VAN DER KWAST H., ROUVIÈRE O., WIEGEL T., Guidelines on prostate cancer. *European Association of Urology*, 2014, vol. 65, p. 124–37
107. MURAWSKA M., RIZOPOULOS D., LESAFFRE E., 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, vol. 2012
108. NGUYEN T., GUEDJ J., HCV kinetic models and their implications in drug development. *CPT : Pharmacometrics & Systems Pharmacology*, 2015, vol. 4, p. 231–242
109. NGUYEN T.T., GUEDJ J., CHACHATY E., DE GUNZBURG J., ANDREMONT A., MENTRÉ F., Mathematical modeling of bacterial kinetics to predict the impact of antibiotic colonic exposure and treatment duration on the amount of resistant enterobacteria excreted. *PLoS Computational Biology*, 2014, vol. 10, p. e1003840
110. OUDARD S., Progress in emerging therapies for advanced prostate cancer. *Cancer Treatment Reviews*, 2013, vol. 39, p. 275–289
111. PARAST L., CHENG S.C., CAI T., Landmark prediction of long-term survival incorporating short-term event time information. *Journal of the American Statistical Association*, 2012, vol. 107, p. 1492–1501
112. PARK K.Y., QIU P., Model selection and diagnostics for joint modeling of survival and longitudinal data with crossing hazard rate functions. *Statistics in Medicine*, 2014, vol. 33, p. 4532–4546
113. PAULER D.K., FINKELSTEIN D.M., Predicting time to prostate cancer recurrence based on joint models for non-linear longitudinal biomarkers and event time outcomes. *Statistics in Medicine*, 2002, vol. 21, p. 3897–3911
114. PENCINA M.J., D'AGOSTINO R.B., VASAN R.S., Evaluating the added predictive ability of a new marker : from area under the ROC curve to reclassification and beyond. *Statistics in Medicine*, 2008, vol. 27, p. 157–172
115. PEPE M.S., *The statistical evaluation of medical tests for classification and prediction*. Oxford University Press, USA, 2003
116. PERELSON A.S., GUEDJ J., Modelling hepatitis C therapy - predicting effects of treatment. *Nature Reviews Gastroenterology & Hepatology*, 2015, vol. 12, p. 437–445
117. PETRYLAK D., Therapeutic options in androgen-independent prostate cancer : building on docetaxel. *British Journal of Urology International International*, 2005, vol. 96, p. 41–46
118. PETRYLAK D.P., ANKERST D.P., JIANG C.S., TANGEN C.M., HUSSAIN M.H., LARA P.N., JONES J.A., TAPLIN M.E., BURCH P.A., KOHLI M., BENSON M.C., SMALL E.J., RAGHAVAN D., CRAWFORD E.D., Evaluation of prostate-specific antigen declines for surrogacy in patients treated on SWOG 99-16. *Journal of the National Cancer Institute*, 2006, vol. 98, p. 516–521

119. PINHEIRO J.C., BATES D.M., Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of computational and Graphical Statistics*, 1995, vol. 4, p. 12–35
120. PINHEIRO J.C., BATES D.M., Mixed-effects models in S and S-plus. *Springer, New-York*, 2000
121. PLAN E.L., MALONEY A., MENTRÉ F., KARLSSON M.O., BERTRAND J., Performance comparison of various maximum likelihood nonlinear mixed-effects estimation methods for dose–response models. *The AAPS Journal*, 2012, vol. 14, p. 420–432
122. POLASCIK T.J., OESTERLING J.E., PARTIN A.W., Prostate specific antigen : a decade of discovery-what we have learned and where we are going. *The Journal of Urology*, 1999, vol. 162, p. 293–306
123. POUESSEL D., OUDARD S., GRAVIS G., PRIOU F., SHEN L., CULINE S., Cabazitaxel dans le cancer de la prostate métastatique résistant à la castration ayant progressé pendant ou après traitement par docétaxel : l’expérience française de l’essai TROPIC. *Bulletin du Cancer*, 2012, vol. 99, p. 731–741
124. PRAGUE M., COMMENGES D., GUEJ J., DRYLEWICZ J., THIÉBAUT R., NIMROD : A program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations. *Computer Methods and Programs in Biomedicine*, 2013, vol. 111, p. 447–458
125. PROUST-LIMA C., SÉNE M., TAYLOR J.M., JACQMIN-GADDA H., Joint latent class models for longitudinal and time-to-event data : A review. *Statistical Methods in Medical Research*, 2014, vol. 23, p. 74–90
126. PROUST-LIMA C., TAYLOR J.M., Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA : a joint modeling approach. *Biostatistics*, 2009, p. 535–549
127. PROUST-LIMA C., TAYLOR J.M.G., WILLIAMS S.G., ANKERST D.P., LIU N., KESTIN L.L., BAE K., SANDLER H.M., Determinants of change in prostate-specific antigen over time and its association with recurrence after external beam radiation therapy for prostate cancer in five large cohorts. *International Journal of Radiation Oncology, Biology, Physics*, 2008, vol. 72, p. 782–791
128. R CORE TEAM, R : A language and environment for statistical computing. Technical report, R Foundation for Statistical Computing, Vienna, Austria, 2016, <http://www.R-project.org/>
129. RIBBA B., HOLFORD N., MENTRÉ F., The use of model-based tumor-size metrics to predict survival. *Clinical Pharmacology and Therapeutics*, 2014a, vol. 96, p. 133–135
130. RIBBA B., HOLFORD N.H., MAGNI P., TROCÓNIZ I., GUEORGUEVA I., GIRARD P., SARR C., ELISHMERENI M., KLOFT C., FRIBERG L.E., A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis. *CPT : Pharmacometrics & Systems Pharmacology*, 2014b, vol. 3, p. e113

131. RIBBA B., WATKIN E., TOD M., GIRARD P., GRENIER E., YOU B., GIRAUDO E., FREYER G., A model of vascular tumour growth in mice combining longitudinal tumour size data with histological biomarkers. *European Journal of Cancer*, 2011, vol. 47, p. 479–490
132. RIVIÈRE-JOURDAN M.K., MENTRÉ F., *Package 'MIXFIM' : Evaluation of the FIM in NLMEMs using MCMC*. 2015, <https://cran.r-project.org/web/packages/MIXFIM/index.html>
133. RIZOPOULOS D., JM : An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data. *Journal of Statistical Software*, 2010, vol. 35, p. 1–33
134. RIZOPOULOS D., Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics*, 2011, vol. 67, p. 819–829
135. RIZOPOULOS D., Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive gaussian quadrature rule. *Computational Statistics & Data Analysis*, 2012a, vol. 56, p. 491–501
136. RIZOPOULOS D., *Joint Models for Longitudinal and Time-to-Event Data : With Applications in R*. CRC Press, 2012b
137. RIZOPOULOS D., GHOSH P., A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Statistics in Medicine*, 2011, vol. 30, p. 1366–1380
138. RIZOPOULOS D., MURAWSKA M., ANDRINOPOULOU E.R., MOLENBERGHS G., TAKKENBERG J.J., LESAFFRE E., Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *arXiv preprint arXiv :13066479*, 2013
139. RIZOPOULOS D., VERBEKE G., LESAFFRE E., Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 2009, vol. 71, p. 637–654
140. ROACH M., HANKS G., THAMES H., SCHELLHAMMER P., SHIPLEY W.U., SOKOL G.H., SANDLER H., Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer : recommendations of the RTOG-ASTRO Phoenix Consensus Conference. *International Journal of Radiation Oncology* Biology* Physics*, 2006, vol. 65, p. 965–974
141. RONDEAU V., MAZROUI Y., GONZALEZ J.R., 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*, 2012, vol. 47, p. 1–28, <http://www.jstatsoft.org/v47/i04/>
142. RUFFION A., REBILLARD X., GRIMA F., PSA doubling time and method of calculation. *Progrès en Urologie : Journal de l'Association Française d'Urologie et de la Société Française d'Urologie*, 2005, vol. 15, p. 1035–1041
143. SALOMON L., BASTIDE C., BEUZEBOC P., CORMIER L., FROMONT G., HENNEQUIN C., MONGIAT-ARTUS P., PEYROMAURE M., PLOUSSARD G., RENARD-PENNA R.,

- ROZET F., AZRIA D., COLOBY P., MOLINIÉ V., RAVERY V., REBILLARD X., RICHAUD P., VILLERS A., SOULIÉ M., LES MEMBRES DU CCAFU, Recommandations en onco-urologie 2013 du CCAFU : Cancer de la prostate. *Progrès en Urologie*, 2013, vol. 23, p. S69–S101
144. SAMSON A., LAVIELLE M., MENTRÉ F., Extension of the SAEM algorithm to left-censored data in nonlinear mixed-effects model : Application to HIV dynamics model. *Computational Statistics & Data Analysis*, 2006, vol. 51, p. 1562–1574
145. SCHER H.I., HALABI S., TANNOCK I., MORRIS M., STERNBERG C.N., CARDUCCI M.A., EISENBERGER M.A., HIGANO C., BUBLEY G.J., DREICER R., PETRYLAK D., KANTOFF P., BASCH E., KELLY W.K., FIGG W.D., SMALL E.J., BEER T.M., WILDING G., MARTIN A., HUSSAIN M., Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone : recommendations of the Prostate Cancer Clinical Trials Working Group. *Journal of Clinical Oncology*, 2008, vol. 26, p. 1148–1159
146. SCHER H.I., HELLER G., MOLINA A., ATTARD G., DANILA D.C., JIA X., PENG W., SANDHU S.K., OLMOS D., RIISNAES R., MCCORMACK R., BURZYKOWSKI T., KHEOK T., FLEISHER M., BUYSE M., DE BONO J.S., Circulating tumor cell biomarker panel as an individual-level surrogate for survival in metastatic castration-resistant prostate cancer. *Journal of Clinical Oncology*, 2015, vol. 33, p. 1348–1355
147. SCHOOP R., GRAF E., SCHUMACHER M., Quantifying the predictive performance of prognostic models for censored survival data with time-dependent covariates. *Biometrics*, 2008, vol. 64, p. 603–610
148. SCHRÖDER F.H., HUGOSSON J., CARLSSON S., TAMMELA T., MÄÄTTÄNEN L., AUVINEN A., KWIATKOWSKI M., RECKER F., ROOBOL M.J., Screening for prostate cancer decreases the risk of developing metastatic disease : findings from the European randomized study of screening for prostate cancer (erspc). *European Urology*, 2012, vol. 62, p. 745–752
149. SCHRÖDER F.H., HUGOSSON J., ROOBOL M.J., TAMMELA T.L., CIATTO S., NELEN V., KWIATKOWSKI M., LUJAN M., LILJA H., ZAPPA M., DENIS L.J., RECKER F., BERENQUER A., MÄÄTTÄNEN L., BANGMA C.H., AUS G., VILLERS A., REBILLARD X., VAN DER KWAST T., BLIJENBERG B.G., MOSS S.M., DE KONING H.J., AUVINEN A., Screening and prostate-cancer mortality in a randomized European study. *New England Journal of Medicine*, 2009, vol. 360, p. 1320–1328
150. SÈNE M., BELLERA C.A., PROUST-LIMA C., Shared random-effect models for the joint analysis of longitudinal and time-to-event data : application to the prediction of prostate cancer recurrence. *Journal de la Société Française de Statistique*, 2014a, vol. 155, p. 134–155
151. SÈNE M., TAYLOR J.M., DIGNAM J.J., JACQMIN-GADDA H., PROUST-LIMA C., Individualized dynamic prediction of prostate cancer recurrence with and without the initiation of a second treatment : Development and validation. *Statistical Methods in Medical Research*, 2014b, p. 1–20
152. SERUGA B., OCANA A., TANNOCK I.F., Drug resistance in metastatic castration-resistant prostate cancer. *Nature Reviews Clinical Oncology*, 2011, vol. 8, p. 12–23

BIBLIOGRAPHIE

153. SHAPPELL S.B., FULMER J., ARGUELLO D., WRIGHT B.S., OPPENHEIMER J.R., PUTZI M.J., PCA3 urine mRNA testing for prostate carcinoma : patterns of use by community urologists and assay performance in reference laboratory setting. *Urology*, 2009, vol. 73, p. 363–368
154. SHARIAT S.F., SEMJONOW A., LILJA H., SAVAGE C., VICKERS A.J., BJARTELL A., Tumor markers in prostate cancer I : blood-based markers. *Acta Oncologica*, 2011, vol. 50, p. 61–75
155. SHARMA M.R., MAITLAND M.L., RATAIN M.J., RECIST : no longer the sharpest tool in the oncology clinical trials toolbox-point. *Cancer Research*, 2012, vol. 72, p. 5145–5149
156. SHEINER L., BEAL S., NONMEM version 5.1. *University of California, NONMEM Project Group, San Francisco*, 1998
157. SHEINER L.B., ROSENBERG B., MELMON K.L., Modelling of individual pharmacokinetics for computer-aided drug dosage. *Computers and Biomedical Research*, 1972, vol. 5, p. 441–459
158. SIEGEL R.L., MILLER K.D., JEMAL A., Cancer statistics, 2016. *CA : A Cancer Journal for Clinicians*, 2016, vol. 66, p. 7–30
159. SIMEONI M., MAGNI P., CAMMIA C., DE NICOLAO G., CROCI V., PESENTI E., GERMANI M., POGGESI I., ROCCHETTI M., Predictive pharmacokinetic-pharmacodynamic modeling of tumor growth kinetics in xenograft models after administration of anticancer agents. *Cancer Research*, 2004, vol. 64, p. 1094–1101
160. SLATE E.H., TURNBULL B.W., Statistical models for longitudinal biomarkers of disease onset. *Statistics in Medicine*, 2000, vol. 19, p. 617–637
161. SOBIN L.H., GOSPODAROWICZ M.K., WITTEKIND C., *TNM classification of malignant tumours*. John Wiley & Sons, 2011
162. SONG X., DAVIDIAN M., TSIATIS A.A., A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. *Biometrics*, 2002, vol. 58, p. 742–753
163. STAN DEVELOPMENT TEAM, Stan modeling language users guide and reference manual, version 2.8.0. 2015
164. STEIN W.D., FIGG W.D., DAHUT W., STEIN A.D., HOSHEN M.B., PRICE D., BATES S.E., FOJO T., Tumor Growth Rates Derived from Data for Patients in a Clinical Trial Correlate Strongly with Patient Survival : A Novel Strategy for Evaluation of Clinical Trial Data. *The Oncologist*, 2008, vol. 13, p. 1046–1054
165. STEIN W.D., GULLEY J.L., SCHLOM J., MADAN R.A., DAHUT W., FIGG W.D., NING Y.M., ARLEN P.M., PRICE D., BATES S.E., FOJO T., Tumor regression and growth rates determined in five intramural NCI prostate cancer trials : the growth rate constant as an indicator of therapeutic efficacy. *Clinical Cancer Research*, 2011, vol. 17, p. 907–917
166. STEPHENSON A.J., KATTAN M.W., EASTHAM J.A., DOTAN Z.A., BIANCO F.J., LILJA H., SCARDINO P.T., Defining biochemical recurrence of prostate cancer after radical prostatectomy : a proposal for a standardized definition. *Journal of Clinical Oncology*, 2006, vol. 24, p. 3973–3978

167. STEYERBERG E.W., VICKERS A.J., COOK N.R., GERDS T., GONEN M., OBUCHOWSKI N., PENCINA M.J., KATTAN M.W., Assessing the performance of prediction models : a framework for some traditional and novel measures. *Epidemiology (Cambridge, Mass)*, 2010, vol. 21, p. 128
168. SUBTIL F., RABILLOUD M., Robust non-linear mixed modelling of longitudinal PSA levels after prostate cancer treatment. *Statistics in Medicine*, 2010, vol. 29, p. 573–587
169. SWEETING M.J., THOMPSON S.G., Joint modelling of longitudinal and time-to-event data with application to predicting abdominal aortic aneurysm growth and rupture. *Biometrical Journal*, 2011, vol. 53, p. 750–763
170. TANNOCK I.F., DE WIT R., BERRY W.R., HORTI J., PLUZANSKA A., CHI K.N., OUDARD S., THÉODORE C., JAMES N.D., TURESSON I., ROSENTHAL M.A., EISENBERGER M.A., Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. *New England Journal of Medicine*, 2004, vol. 351, p. 1502–1512
171. TANNOCK I.F., FIZAZI K., IVANOV S., KARLSSON C.T., FLÉCHON A., SKONECZNA I., ORLANDI F., GRAVIS G., MATVEEV V., BAVBEK S., GIL T., VIANA L., ARÉN O., KARYAKIN O., ELLIOTT T., BIRTLE A., MAGHERINI E., HATTEVILLE L., PETRYLAK D., TOMBAL B., ROSENTHAL M., Afibercept versus placebo in combination with docetaxel and prednisone for treatment of men with metastatic castration-resistant prostate cancer (venice) : a phase 3, double-blind randomised trial. *The lancet Oncology*, 2013, vol. 14, p. 760–768
172. TAYLOR J.M., PARK Y., ANKERST D.P., PROUST-LIMA C., WILLIAMS S., KESTIN L., BAE K., PICKLES T., SANDLER H., Real-time individual predictions of prostate cancer recurrence using joint models. *Biometrics*, 2013, vol. 69, p. 206–213
173. TAYLOR J.M., SHEN J., KENNEDY E.H., WANG L., SCHAUBEL D.E., Comparison of methods for estimating the effect of salvage therapy in prostate cancer when treatment is given by indication. *Statistics in Medicine*, 2014, vol. 33, p. 257–274
174. TAYLOR J.M., WANG Y., Surrogate markers and joint models for longitudinal and survival data. *Controlled Clinical Trials*, 2002, vol. 23, p. 626–634
175. TAYLOR J.M., YU M., SANDLER H.M., Individualized predictions of disease progression following radiation therapy for prostate cancer. *Journal of Clinical Oncology*, 2005, vol. 23, p. 816–825
176. THERASSE P., ARBUCK S.G., EISENHauer E.A., WANDERS J., KAPLAN R.S., RUBINSTEIN L., VERWEIJ J., VAN GLABBEKE M., VAN OOSTEROM A.T., CHRISTIAN M.C., GWYTHYER S.G., New guidelines to evaluate the response to treatment in solid tumors. *Journal of the National Cancer Institute*, 2000, vol. 92, p. 205–216
177. TSIATIS A., DEGRUTTOLA V., WULFSOHN M., 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*, 1995, vol. 90, p. 27–37

178. TSIATIS A.A., DAVIDIAN M., Joint modeling of longitudinal and time-to-event data : an overview. *Statistica Sinica*, 2004, vol. 14, p. 809–834
179. TU H., JACOBS S.C., BORKOWSKI A., KYPRIANOU N., Incidence of apoptosis and cell proliferation in prostate cancer : Relationship with TGF- β 1 and bcl-2 expression. *International Journal of Cancer*, 1996, vol. 69, p. 357–363
180. UECKERT S., RIVIÈRE-JOURDAN M.K., MENTRÉ F., Evaluation of the Fisher information matrix in nonlinear mixed effect models without linearization. *PODE : Population Optimum Design of Experiments*, 2015, http://www.maths.qmul.ac.uk/~bb/PODE/PODE2015Riviere_Ueckert.pdf
181. UNO H., CAI T., PENCINA M.J., D’AGOSTINO R.B., WEI L., On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Statistics in Medicine*, 2011, vol. 30, p. 1105–1117
182. VAN HOUWELINGEN H., PUTTER H., *Dynamic prediction in clinical survival analysis*. CRC Press, 2011
183. VIGAN M., *Modélisation de données longitudinales et de données d’événements répétés par des modèles non linéaires à effets mixtes : application à la maladie de Gaucher*. Ph.D. thesis, Paris 7, 2014
184. VIGAN M., STIRNEMANN J., MENTRÉ F., Evaluation of estimation methods and power of tests of discrete covariates in repeated time-to-event parametric models : application to gaucher patients treated by imiglucerase. *The AAPS Journal*, 2014, vol. 16, p. 415–423
185. VONESH E.F., A note on the use of Laplace’s approximation for nonlinear mixed-effects models. *Biometrika*, 1996, vol. 83, p. 447–452
186. WANG Y., TAYLOR J.M.G., Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association*, 2001, vol. 96, p. 895–905
187. WILBAUX M., TOD M., DE BONO J., LORENTE D., MATEO J., FREYER G., YOU B., HÉNIN E., A joint model for the kinetics of CTC count and PSA concentration during treatment in metastatic castration-resistant prostate cancer. *CPT : Pharmacometrics & Systems Pharmacology*, 2015, vol. 4, p. 277–285
188. WOLFINGER R., Laplace’s approximation for nonlinear mixed models. *Biometrika*, 1993, vol. 80, p. 791–795
189. WU L., LIU W., YI G.Y., HUANG Y., Analysis of longitudinal and survival data : joint modeling, inference methods, and issues. *Journal of Probability and Statistics*, 2011, vol. 2012
190. WULFSOHN M.S., TSIATIS A.A., A joint model for survival and longitudinal data measured with error. *Biometrics*, 1997, p. 330–339
191. YE W., LIN X., TAYLOR J.M., Semiparametric modeling of longitudinal measurements and time-to-event data—A two-stage regression calibration approach. *Biometrics*, 2008, vol. 64, p. 1238–1246

192. YOU B., PERRIN P., FREYER G., RUFFION A., TRANCHAND B., HÉNIN E., PAPAREL P., RIBBA B., DEVONEC M., FALANDRY C., FOURNAL C., TOD M., GIRARD P., Advantages of prostate-specific antigen (PSA) clearance model over simple PSA half-life computation to describe PSA decrease after prostate adenomectomy. *Clinical Biochemistry*, 2008, vol. 41, p. 785–795
193. YU M., LAW N.J., TAYLOR J.M., SANDLER H.M., Joint longitudinal-survival-cure models and their application to prostate cancer. *Statistica Sinica*, 2004, p. 835–862
194. YU M., TAYLOR J.M.G., SANDLER H.M., Individual prediction in prostate cancer studies using a joint longitudinal survival–cure model. *Journal of the American Statistical Association*, 2008, vol. 103, p. 178–187
195. ZAGARS G.K., POLLACK A., Kinetics of serum prostate-specific antigen after external beam radiation for clinically localized prostate cancer. *Radiotherapy and Oncology*, 1997, vol. 44, p. 213–221
196. ZHANG L., BEAL S.L., SHEINER L.B., Simultaneous vs. sequential analysis for population PK/PD data I : best-case performance. *Journal of Pharmacokinetics and Pharmacodynamics*, 2003, vol. 30, p. 387–404

Résumé

L'évaluation de traitements pour le cancer de la prostate métastatique hormono-résistant (CPmHR) repose sur le temps de décès. L'antigène spécifique de la prostate (PSA), supposé lié à la survie, est régulièrement mesuré. La modélisation conjointe qui consiste en l'analyse simultanée de l'évolution du biomarqueur et de la survie est alors particulièrement adaptée, mais souvent limitée à un processus longitudinal linéaire. L'objectif principal de cette thèse est d'étudier la modélisation conjointe quand la cinétique du biomarqueur est décrite par un modèle non-linéaire à effets mixtes (MNLEM). Tout d'abord, nous avons montré par simulations que l'algorithme SAEM de Monolix estimait sans biais les paramètres d'un modèle conjoint non-linéaire, avec un risque de première espèce et une puissance à détecter un lien entre les deux processus satisfaisants. Puis, nous avons développé un modèle conjoint mécanistique pour caractériser le lien entre la cinétique du PSA et la survie chez des patients ayant un CPmHR et traités par docetaxel. Le modèle structural du MNLEM a été défini par un système d'équations différentielles (ODEs) décrivant le mécanisme de production du PSA par des cellules sensibles au docetaxel et des cellules résistantes. La sélection et l'évaluation du modèle ont été détaillées. Le modèle conjoint final souligne le rôle prépondérant sur la survie des cellules résistantes, non-observées. Enfin, nous avons étendu des outils développés dans le contexte linéaire pour faire de la prédiction individuelle dynamique en utilisant un modèle conjoint non-linéaire. Une méthode bayésienne a été mise en place pour fournir la distribution des paramètres individuels. Les performances prédictives du modèle ont pu être évaluées à l'aide de métriques de discrimination et de calibration dépendant du temps. Ces travaux ouvrent la voie au développement de modèles conjoints mécanistiques, qui permettent de tenir compte de l'influence de plusieurs biomarqueurs sur la survie, au moyen d'ODEs, afin d'améliorer l'évaluation thérapeutique et la prédiction.

Mots-clés : *Modélisation conjointe, Modèle non-linéaire à effets mixtes, Modèle mécanistique, Analyse de survie, Cancer de la prostate métastatique, PSA, Prédiction individuelle dynamique, Algorithme SAEM*

Abstract

Treatment evaluation for metastatic Castration-Resistant Prostate Cancer (mCRPC) relies on time-to-death. Prostate-specific antigen (PSA), assumed to be linked to survival, is frequently measured. Joint modelling which consists in the simultaneous analyse of biomarker's evolution and survival is particularly adapted, but often limited to a linear longitudinal process. The main objective of this PhD is to study joint modelling when biomarker kinetics is described by a nonlinear mixed-effects model (NLMEM). First, we established by simulations that the SAEM algorithm of Monolix provided unbiased parameter estimations of a nonlinear joint model, with satisfying type 1 error and power to detect a link between the two processes. Then, we developed a mechanistic joint model to characterize the relationship between PSA kinetics and survival in mCRPC patients treated by docetaxel. The structural model of the NLMEM was defined by a system of differential equations (ODEs) describing the mechanism of PSA production by docetaxel-sensitive and -resistant cells. Model selection and evaluation were detailed. The final joint model showed the predominant role of the non-observed resistant cells on survival. Lastly, we expanded tools developed in a linear context for individual dynamic prediction using nonlinear joint model. A Bayesian method provided the distribution of individual parameters. Predictive performances of the model were assessed using time-dependent discrimination and calibration metrics. These works open the way for the development of mechanistic joint models, which enable to account for the impact of several biomarkers on survival through ODEs, in order to improve therapeutic evaluation and prediction.

Key-words : *Joint modelling, Nonlinear mixed-effects model, Mechanistic model, Survival analysis, Metastatic prostate cancer, PSA, Individual dynamic predictions, SAEM algorithm*