



**HAL**  
open science

# Méthodes de comparaisons de deux ou plusieurs groupes de données censurées par intervalle. Avec application en immunologie clinique.

Sarah Flora Jonas

## ► To cite this version:

Sarah Flora Jonas. Méthodes de comparaisons de deux ou plusieurs groupes de données censurées par intervalle. Avec application en immunologie clinique.. Méthodes et statistiques. Université Paris Saclay (COMUE), 2018. Français. NNT: 2018SACLS317 . tel-01902815

**HAL Id: tel-01902815**

**<https://theses.hal.science/tel-01902815>**

Submitted on 23 Oct 2018

**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.

# Méthodes de comparaisons de deux ou plusieurs groupes de données censurées par intervalle

Avec application en immunologie clinique

Thèse de doctorat de l'Université Paris-Saclay  
préparée à l'Université Paris-Sud

École Doctorale de Santé Publique  
Spécialité de doctorat: Biostatistique

Thèse présentée et soutenue à l'Université Paris-Sud, le 3 octobre 2018, par

**Sarah Flora Jonas**

Composition du jury:

Jean-Christophe Thalabard <i>PU-PH Université Paris-Descartes</i>	Rapporteur Président
Paul Landais <i>PU-PH Université de Montpellier</i>	Rapporteur
Jacques Benichou <i>PU-PH Université de Rouen</i>	Examineur
Pascale Tubert-Bitter <i>DR Université Paris-Saclay</i>	Examineur
Yann De Rycke <i>Université Pierre et Marie Curie</i>	Examineur
Philippe Broët <i>PU-PH Université Paris-Sud</i>	Directeur de thèse



# Remerciements

Toute ma gratitude va au Pr Philippe Broët, mon directeur de thèse qui m'a guidée tout au long de ces années avec une grande sagacité et une infinie patience. J'ai eu la chance d'apprendre auprès de lui qui m'a fait bénéficier de sa vaste érudition et m'a inspirée par sa profonde humanité. Il reste pour moi un modèle.

Un grand merci à Signè Hassler pour ses précieux conseils.

Je voudrais remercier les membres de mon jury de thèse pour le temps consacré à revoir mon travail. Le Pr Jean-Christophe Thalabard et le Pr Paul Landais pour leur relecture attentive. Ainsi que le Pr Jacques Benichou, le Dr Pascale Tubert-Bitter et M. Yann De Rycke d'avoir accepté de participer au jury de thèse.



# Introduction générale

L'analyse des données de survie consiste en l'étude du délai avant qu'un événement défini ne se produise. Bien que l'événement d'intérêt puisse être l'apparition d'anticorps, la guérison ou autre, on parle de données de survie car historiquement, l'événement d'intérêt considéré était le décès. Les études biomédicales et les études de fiabilité sont les deux domaines majeurs d'application de ce type d'analyse.

Dans les études de survie, les problématiques abordées sont principalement l'estimation des fonctions de survie et les analyses univariées et multivariées des relations entre temps de survie et variable explicatives.

La spécificité des études de survie en milieu médical est l'incomplétude de l'information. Si à la fin de l'étude, l'événement d'intérêt n'a pas eu lieu, la donnée est dite censurée. Pour autant, un patient censuré apporte une information : l'événement considéré ne s'est pas produit avant un certain temps, la date de dernières nouvelles. Ignorer les données censurées conduirait à perdre de l'information et à biaiser l'estimation de la fonction de survie. Le cas le plus classique de censure est celui de la censure aléatoire droite.

Dans de très nombreuses études, les patients ne sont pas suivis en continu. Ils effectuent une ou plusieurs visites au cours de l'étude. À chacune de ces visites, il est possible de déterminer si l'événement d'intérêt s'est ou non réalisé. Si l'événement d'intérêt est constaté lors de l'une de ces visites, il est cependant impossible de connaître avec précision la date de survenue de cet événement. L'information disponible est un intervalle de temps, l'événement s'étant produit entre les deux visites. Ce type de donnée est dit *censuré par intervalle*. Alors que la censure à droite a été largement étudiée, la censure par intervalle a fait l'objet de moins d'attention et d'utilisation pratique.

Lorsque les patients n'effectuent qu'une seule visite après le début de l'étude, les données sont dites censurées par intervalle de type 1 (ou *current status*). Lors de cette visite unique, est constatée l'absence ou la réalisation de l'événement d'intérêt. Ainsi, à la date de visite : si l'événement n'a pas eu lieu, le patient est *censuré à droite*; si l'événement a eu lieu, le patient est *censuré à gauche* (c'est à dire, que le patient a fait l'événement avant la date d'évaluation). Une base de données *current status* n'est constituée que de patients censurées, à droite ou à gauche. Certaines méthodes ont exclusivement été développées pour ce type de données.

Des approches ont été réalisées pour la censure par intervalle mais elles sont complexes, peu flexibles et dans certains cas très difficiles à mettre en œuvre. De plus, les méthodes développées pour la censure à droite ne s'adaptent pas aisément à la censure par intervalle. Une des raisons principales de cette difficulté d'adaptation réside dans l'impossibilité, avec la censure par intervalle, d'utiliser la vraisemblance partielle. Présentée par Sir David Cox pour la censure à droite et dans le cas d'un modèle multiplicatif à risques proportionnels, la vraisemblance partielle permet de s'affranchir de l'estimation de la fonction de risque instantané (paramètre de nuisance). Dans le cas de la censure par intervalle, l'utilisation de la vraisemblance partielle est la plupart du temps impossible.

Les tests de comparaisons des fonctions de survie de plusieurs groupes peuvent être construits à partir d'hypothèses paramétriques ou semi-paramétriques des fonctions de risques instantanés. Dans ce cas, le modèle le plus utilisé est le modèle multiplicatif à risques proportionnels ou

modèle de Cox. Pour ce modèle, entre deux groupes, le rapport des risques est proportionnel et constant dans le temps.

Cependant, il existe des alternatives au modèle de Cox lorsque l'hypothèse de proportionnalité des risques n'est pas valide. La non-proportionnalité des risques marginaux pour une variable explicative donnée peut, entre autre, être liée à l'existence d'interactions avec des variables latentes ou à la présence d'une population hétérogène comprenant des individus ayant un risque nul pour l'événement d'intérêt. En présence de risques non-proportionnels, les statistiques de test basées sur le modèle de Cox ne sont pas optimales.

Dans ce travail de thèse nous nous intéressons à deux situations de risques non-proportionnels en présence de données censurées par intervalle et proposons deux statistiques de test de comparaisons de deux (ou plusieurs groupes) de l'hypothèse nulle d'égalité des distributions de survie. Ces statistiques de test sont ensuite appliqués à des données réelles recueillies dans le cadre du consortium ABIRISK dont l'objectif principal était d'étudier l'immunogénicité de biothérapies données à des patients atteints de maladies auto-immunes.

Dans le premier chapitre, nous présentons un état de l'art et détaillons le cadre applicatif d'immunogénétique. Dans le second chapitre, nous présentons un nouveau test de comparaison de groupes de données censurées par intervalle et risques non-proportionnels avec une application sur des données du consortium ABIRISK portant sur les maladies inflammatoires chroniques de l'intestin. Dans le troisième chapitre, nous présentons un nouveau test de comparaison de groupes de données censurées par intervalle avec population à risques hétérogènes avec une application sur des données du consortium ABIRISK portant sur la sclérose en plaques. Nous terminons ce mémoire de thèse par une discussion générale.

# Table des matières

<b>Remerciements</b>	<b>i</b>
<b>Introduction</b>	<b>iii</b>
<b>Table des matières</b>	<b>v</b>
<b>Liste des figures</b>	<b>vii</b>
<b>Liste des tableaux</b>	<b>ix</b>
<b>PARTIE I ÉTAT DE L'ART ET CADRE APPLICATIF</b>	<b>1</b>
<b>1 Analyse de survie</b>	<b>3</b>
I. Généralités . . . . .	4
II. Types de censure . . . . .	10
III. Vraisemblance et censure par intervalle . . . . .	14
IV. Estimateur de la fonction de survie avec censure par intervalle . . . . .	22
V. Tests non-paramétriques de comparaison de groupes de données censurées . . . . .	24
VI. Références . . . . .	27
<b>2 Tests de comparaison de groupes de données censurées par intervalle</b>	<b>31</b>
I. Trois formes générales de tests pour données censurées par intervalle . . . . .	32
II. Les tests implémentés existants . . . . .	38
III. Références . . . . .	47
<b>3 Cadre applicatif : maladies auto-immunes et anticorps</b>	<b>51</b>
I. Système immunitaire . . . . .	52
II. Maladies auto-immunes . . . . .	60
III. Le projet ABIRISK . . . . .	63
IV. Références . . . . .	67
<b>PARTIE II COMPARAISON DE GROUPES DE DONNÉES CENSURÉES PAR INTERVALLE ET RISQUES NON-PROPORTIONNELS</b>	<b>69</b>
<b>4 Test de comparaison de deux ou plusieurs groupes de données censurées par intervalle avec risques non-proportionnels</b>	<b>71</b>
I. Introduction . . . . .	72
II. Proposition de test . . . . .	72
III. Comparaisons des tests et simulations . . . . .	79
IV. Discussion . . . . .	90
V. Références . . . . .	92

<b>5 Application sur données d'immunologie clinique censurées par intervalle</b>	<b>95</b>
I. Les maladies inflammatoires chroniques de l'intestin . . . . .	96
II. La cohorte MICI du projet ABIRISK . . . . .	102
III. Base de données et méthodes . . . . .	106
IV. Résultats . . . . .	109
V. Discussion . . . . .	114
VI. Références . . . . .	116
<b>PARTIE III COMPARAISON DE GROUPES DE DONNÉES CENSURÉES PAR INTERVALLE AVEC POPULATION À RISQUES HÉTÉROGÈNES</b>	<b>119</b>
<b>6 Test de comparaison de deux ou plusieurs groupes de données censurées par intervalle avec population à risques hétérogènes</b>	<b>121</b>
I. Introduction . . . . .	122
II. Proposition de test . . . . .	123
III. Étude de simulations sur données censurées par intervalle de type I . . . . .	128
IV. Discussion . . . . .	134
V. Références . . . . .	137
<b>7 Application sur données d'immunologie clinique censurées par intervalle de type I</b>	<b>141</b>
I. Introduction . . . . .	142
II. La sclérose en plaques . . . . .	142
III. Cohorte Munich . . . . .	143
IV. Cohorte Düsseldorf . . . . .	145
V. Discussion . . . . .	146
VI. Références . . . . .	147
<b>Conclusion</b>	<b>149</b>
<b>ANNEXES</b>	<b>III</b>
<b>A Vraisemblance marginale des rangs pour censure droite</b>	<b>III</b>
<b>B Code R du test de comparaison de groupes de données <i>current status</i> avec plateau</b>	<b>V</b>
<b>C Simulations supplémentaires sous modèle de mélange</b>	<b>IX</b>
<b>D Publications et valorisation de la thèse</b>	<b>XIII</b>

# Liste des figures

2.1	Fonctions de poids $W$ de la famille $G^{p,\gamma}$ . . . . .	43
3.1	Gènes du CMH de classe I et II sur le chromosome 6. . . . .	52
3.2	Structure de l'anticorps . . . . .	57
3.3	Origine et maturation des lymphocytes B et T . . . . .	58
3.4	Réaction du système immunitaire face à l'antigène . . . . .	59
4.1	Graphiques correspondant aux simulations sous modèle à risques croisés et censure exponentielle . . . . .	88
4.2	Graphiques correspondant aux simulations sous modèle à risques croisés et censure uniforme . . . . .	89
5.1	Schéma du tube digestif . . . . .	97
5.2	Facteurs environnementaux contribuant à la pathogenèse des MICI . . . . .	99
5.3	MICI en tant que trouble multifactoriel. . . . .	100
5.4	Schéma des visites dans la cohorte prospective MICI du projet ABIRISK . . . . .	103
5.5	Inclusion des patients dans la cohorte prospective MICI du projet ABIRISK . . . . .	104
5.6	NPMLE de la fonction de survie en fonction du traitement, cohorte MICI . . . . .	109
5.7	$p$ -value et $q$ -value des variables testées en univarié, cohorte MICI. . . . .	110
5.8	NPMLE de la fonction de survie en fonction de diverses variables sélectionnées par le test vgRC, cohorte MICI . . . . .	111
6.1	Données simulées <i>current status</i> avec plateau . . . . .	134
7.1	NPMLE de la fonction de survie en fonction du sexe et de l'âge, cohorte Munich . . . . .	144
7.2	NPMLE de la fonction de survie en fonction du sexe, cohorte Düsseldorf . . . . .	145



# Liste des tableaux

2.1	Implémentations du test <b>vgRP</b> . . . . .	39
2.2	Implémentations du test <b>vgOP</b> . . . . .	40
2.3	Implémentations du test <b>LR adapté</b> . . . . .	40
2.4	Implémentation du test <b>LR amélioré</b> . . . . .	41
2.5	Implémentation du test <b>LRP SZZ</b> . . . . .	42
2.6	Implémentations du test <b>LRP OG</b> . . . . .	43
2.7	Implémentation du test <b>LRPHyb</b> . . . . .	44
3.1	Molécules du CMH . . . . .	53
3.2	Exemples de cytokines et leurs actions . . . . .	56
3.3	Propriétés des différentes classes d'anticorps . . . . .	57
3.4	Exemples de maladies auto-immunes et leurs symptômes. . . . .	61
3.5	Quelques chiffres sur les pathologies auto-immunes en France. . . . .	61
3.6	Produits et voie d'injection pour la sclérose en plaques. . . . .	63
3.7	Biothérapies et prévalence des ADA de pathologies sélectionnées (ABIRISK) . . . . .	64
4.1	Tests du score pour censure par intervalle avec matrice d'information de Fisher . . . . .	79
4.2	Tests du score pour censure par intervalle avec variance permutacionnelle . . . . .	80
4.3	Tests du log-ranks pour censure par intervalle . . . . .	80
4.4	Tests comparés dans les simulations. . . . .	82
4.5	Modèle à risques croisés, censure exponentielle, $n = 300, \pi = 0.5$ . . . . .	83
4.6	Idem, censure uniforme, $n = 300, \pi = 0.5$ . . . . .	84
4.7	Idem, petits échantillons, $n = 100, \pi = 0.5, p = 30\%$ . . . . .	85
4.8	Modèle odds-proportionnel, risques accélérés, $n = 300, \pi = 0.5, p = 30\%$ . . . . .	86
4.9	Modèle à risques croisés, censure dépendante (ATD), $n = 300, \pi = 0.5, p = 30\%$ . . . . .	87
5.1	Les chiffres des MICI en France . . . . .	101
5.2	Traitements et indications administrés en cas de MICI . . . . .	102
5.3	Les chiffres des MICI en France . . . . .	103
5.4	Caractéristiques démographiques des patients MICI, stratifiées par biothérapie. . . . .	104
5.5	Caractéristiques cliniques des patients MICI, stratifiées par biothérapie . . . . .	105
5.6	Caractéristiques cliniques de la MC et de la RCH des patients de la cohorte MICI, stratifiées par biothérapie. . . . .	105
5.7	Développement des ADA dans la cohorte MICI, stratifié par biothérapie. . . . .	106
5.8	Variabes constituant la base de données de la cohorte MICI, Abirisk . . . . .	107
5.9	$q$ -value ( $< 0,25$ ) des variables de la cohorte MICI testées pour leur effet sur les ADA. . . . .	110
5.10	$p$ -value du test du score pour risques croisés et du pseudo log-rank pour la variable allaitement . . . . .	112
5.11	Comparaison de variables chez les patients exclus et inclus dans l'analyse multivariée . . . . .	113
5.12	Modèle multi-varié sur les variables sélectionnées dans la cohorte MICI . . . . .	113
5.13	Tableau récapitulatif des variables sélectionnées dans le modèle multivarié . . . . .	113
6.1	Modèle à risque cumulatif borné, cens. expo., $\tau_0 = 30\%, p = 20\%, n = 400, \xi = 0.5$ . . . . .	130

6.2	Idem, censure exponentielle, $\tau_0 = 50\%$ , $p = 20\%$ , $n = 400$ , $\xi = 0.5$ .	130
6.3	Idem, censure exponentielle, $\tau_0 = 70\%$ , $p = 20\%$ , $n = 400$ , $\xi = 0.5$ .	130
6.4	Idem, censure exponentielle, $\tau_0 = 30\%$ , $p = 40\%$ , $n = 400$ , $\xi = 0.5$ .	131
6.5	Idem, censure exponentielle, $\tau_0 = 50\%$ , $p = 40\%$ , $n = 400$ , $\xi = 0.5$ .	131
6.6	Idem, censure exponentielle, $\tau_0 = 70\%$ , $p = 40\%$ , $n = 400$ , $\xi = 0.5$ .	131
6.7	Idem, censure exponentielle, $\tau_0 = 30\%$ , $p = 20\%$ , $n = 400$ , $\xi = 0.3$ .	132
6.8	Idem, censure exponentielle, $\tau_0 = 30\%$ , $p = 20\%$ , $n = 400$ , $\xi = 0.7$ .	132
6.9	Idem, censure uniforme, suivi insuffisant, $\tau_0 = 30\%$ , $n = 400$ , $\xi = 0.5$ .	132
7.1	Traitements pour patients traités pour scléroses en plaques, cohorte Munich, ABIRISK	143
C.1	Modèle de mélange, censure exponentielle, $\tau_0 = 30\%$ , $p = 20\%$ , $n = 400$ , $\xi = 0.5$ .	IX
C.2	Idem, censure exponentielle, $\tau_0 = 30\%$ , $p = 20\%$ , $n = 400$ , $\xi = 0.5$ .	IX
C.3	Idem, censure exponentielle, $\tau_0 = 70\%$ , $p = 20\%$ , $n = 400$ , $\xi = 0.5$ .	X
C.4	Idem, censure exponentielle, $\tau_0 = 30\%$ , $p = 40\%$ , $n = 400$ , $\xi = 0.5$ .	X
C.5	Idem, censure exponentielle, $\tau_0 = 50\%$ , $p = 40\%$ , $n = 400$ , $\xi = 0.5$ .	X
C.6	Idem, censure exponentielle, $\tau_0 = 70\%$ , $p = 40\%$ , $n = 400$ , $\xi = 0.5$ .	XI
C.7	Idem, censure exponentielle, $\tau_0 = 30\%$ , $p = 20\%$ , $n = 400$ , $\xi = 0.3$ .	XI
C.8	Idem, censure exponentielle, $\tau_0 = 30\%$ , $p = 20\%$ , $n = 400$ , $\xi = 0.7$ .	XI

## **Première partie**

# **État de l'art et cadre applicatif**



# Chapitre 1

## Analyse de survie

### Sommaire

---

<b>I. Généralités</b> . . . . .	<b>4</b>
A. Fonctions liées à la survie . . . . .	4
B. Modèles paramétriques de la fonction de risque . . . . .	6
C. Modèles semi-paramétriques de la fonction de risque . . . . .	7
<b>II. Types de censure</b> . . . . .	<b>10</b>
A. Notations . . . . .	10
B. Censure droite et gauche . . . . .	10
C. Censure par intervalle . . . . .	11
D. Censure non-informative et indépendante . . . . .	13
<b>III. Vraisemblance et censure par intervalle</b> . . . . .	<b>14</b>
A. Le cas particulier de la censure à droite . . . . .	15
B. Le cas de la censure par intervalle . . . . .	17
C. Tests basés sur vraisemblance . . . . .	20
<b>IV. Estimateur de la fonction de survie avec censure par intervalle</b> . . . . .	<b>22</b>
A. Censure par intervalle de type 1 . . . . .	22
B. Censure par intervalle de type 2 . . . . .	23
C. Validité asymptotique de l'estimateur . . . . .	23
<b>V. Tests non-paramétriques de comparaison de groupes de données censurées</b> . . . . .	<b>24</b>
A. Introduction . . . . .	24
B. Principes généraux des tests linéaires de rangs pour deux échantillons . . . . .	24
C. Les tests logranks . . . . .	26
<b>VI. Références</b> . . . . .	<b>27</b>

---

## I. Généralités

Les études de survie traitent de l'analyse et de la modélisation de délais avant apparition d'un événement précis. L'événement d'intérêt peut être le décès, l'apparition d'anticorps, la guérison, la survenue d'une pathologie, la panne d'une machine, le déménagement, etc.

Les champs d'application majeurs de ce type d'études sont les études médicales et les études de fiabilité, bien qu'elles soient aussi utilisées en démographie, en assurance et en psychologie. Historiquement, l'événement d'intérêt étant le décès, on parle d'études de survie.

En pratique, les études de survie ont pour objectif :

- d'estimer l'effet d'une variable explicative sur les temps de survie;
- d'estimer la distribution des temps d'événements;
- de comparer les distributions de survie de plusieurs groupes.

Les données de survie ont une forme particulière; ce sont des variables aléatoires, positives et de distribution souvent dissymétrique. La problématique majeure des études de survie est l'existence de données incomplètes ou "censurées". La censure est due à un accès restreint à l'information complète.

Une première raison, structurelle, est cause de censure : les patients ne pouvant être suivis indéfiniment, il se peut qu'à la fin de l'étude, certains patients n'aient pas encore fait l'événement d'intérêt. Une seconde raison, non-structurelle, est l'abandon de l'étude par des patients au cours du suivi; ces patients sont appelés les perdus de vue.

Ainsi, en présence de censure, il est impossible d'utiliser des tests de comparaison classiques comme le test de comparaison des moyennes de Student, sous peine d'obtenir des résultats biaisés. Il apparaît indispensable d'avoir recours à une méthodologie spécifique aux données censurées.

### A. Fonctions liées à la survie

Nous notons  $T$ , la durée de survie; une variable positive ou nulle et absolument continue. Elle représente la durée écoulée entre la date d'origine (point de départ du délai étudié : date d'entrée dans l'étude, début d'exposition à un facteur de risque, date de naissance...) et la date de survenue de l'événement d'intérêt chez un individu.

La loi de probabilité de  $T$  peut être définie par cinq fonctions équivalentes :

#### Fonction de survie $S$

Pour  $t$  fixé, la fonction de survie est la probabilité de ne pas avoir fait l'événement jusqu'à l'instant  $t$ , soit la probabilité que  $T$  soit supérieur à une valeur  $t$  :

$$S(t) = \mathbb{P}(T > t), \quad 0 < t < +\infty.$$

$S(t)$  est une fonction décroissante avec  $S(0) = 1$  et  $\lim_{t \rightarrow +\infty} S(t) = 0$ .

Si  $\lim_{t \rightarrow +\infty} S(t) > 0$ , la fonction de survie correspondante est appelée *fonction de survie impropre*. Ce cas sera abordé dans les sections suivantes.

#### Fonction de répartition $F$

Pour  $t$  fixé, la fonction de répartition représente la probabilité d'avoir fait l'événement avant l'instant  $t$ , soit :

$$F(t) = \mathbb{P}(T \leq t) = 1 - S(t).$$

### Densité de probabilité $f$

La fonction correspondante à la densité de probabilité,  $f(t)$  est :

$$f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt}.$$

Lorsque  $F$  admet une dérivée au point  $t$ , la densité de probabilité s'exprime comme :

$$f(t) = \lim_{h \rightarrow 0} \frac{\mathbb{P}(t \leq T < t+h)}{h} = \underbrace{F'(t)}_{\text{probabilité de faire l'événement avant } t}$$

On note également que  $f(t) \geq 0$ ,  $\int_0^{+\infty} f(t)dt = 1$  et :

$$F(t) = \int_0^t f(u) du.$$

### Risque instantané $\lambda$

La fonction de risque instantané est définie comme :

$$\lambda(t) = \lim_{h \rightarrow 0^+} \frac{\mathbb{P}(t \leq T < t+h | T \geq t)}{h} = \frac{f(t)}{S(t)}.$$

Elle exprime le taux instantané d'apparition de l'événement pour les sujets encore présents et sans événement au temps  $t$ .

On peut l'associer à la limite d'une probabilité conditionnelle telle que :

$$\lambda(t)dt \approx \mathbb{P}(t \leq T < t+h | T \geq t).$$

### Risque cumulé $\Lambda$

Le taux de risque cumulé,  $\Lambda$  est l'intégrale du risque instantané  $\lambda$  :

$$\Lambda(t) = \int_0^t \lambda(u)du$$

C'est une fonction positive et non-décroissante.

### Liens entre les fonctions

$$S(t) = 1 - F(t)$$

$$S(t) = \exp(-\Lambda(t)) = \exp\left(-\int_0^t \lambda(u)du\right)$$

$$f(t) = F'(t) = -S'(t)$$

$$f(t) = \lambda(t) S(t) = \lambda(t) \exp(-\Lambda(t))$$

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = -\ln'(S(t))$$

$$\Lambda(t) = \int_0^t \lambda(u)du$$

$$\Lambda(t) = -\ln(S(t))$$

Bien que l'on puisse définir la distribution de  $T$  à partir de l'une des cinq fonctions équivalentes décrites précédemment; la fonction de risque instantané  $\lambda(t)$  en fournit la description la plus concrète. Sa forme : croissante, décroissante, en forme de U, constante... donne une information essentielle sur l'évolution du risque de survenue de l'événement dans la population.

Nous verrons dans la section suivante que la fonction  $\lambda(t)$  peut être entièrement non-spécifiée ou dépendante d'un ou plusieurs paramètres.

## B. Modèles paramétriques de la fonction de risque

La fonction de survie ou de manière équivalente, la fonction de risque, peut être modélisée à l'aide de modèles paramétriques.

Dans un modèle paramétrique, la distribution de la survie est celle d'une loi de distribution connue (exponentielle, Weibull, etc.). L'expérimentateur estime le(s) meilleur(s) paramètre(s) associé(s) à la loi choisie pour son cas.

Cependant, il est fréquent que la fonction de survie (ou de risque) dépende de variables explicatives. Dès lors, il est d'intérêt de considérer les généralisations des modèles paramétriques ou d'autres modèles prenant en compte les informations supplémentaires des individus constituant l'échantillon.

Dans un modèle semi-paramétrique, la forme de la fonction de survie (ou de risque) n'est pas spécifiée, mais il prend en compte l'effet d'une covariable. Il se distingue du modèle non-paramétrique par l'imposition d'une contrainte paramétrique.

Dans le cas paramétrique, il existe plusieurs lois de distribution permettant de modéliser le temps d'apparition d'un événement. Nous en présentons ci-dessous trois exemples classiques.

### 1) Distribution exponentielle

Dans la distribution exponentielle à un paramètre, la fonction de risque est constante, soit,  $\lambda(t) = \lambda > 0$ . La fonction de risque ne dépendant pas de  $t$ , on parle de loi sans mémoire; le risque de faire l'événement étant identique pour toute la période considérée. Nous avons :

$$F(t) = e^{-\lambda t} \quad \text{et} \quad f(t) = \lambda e^{-\lambda t}.$$

### 2) Distribution de Weibull

Ceci correspond à une généralisation à deux paramètres de la distribution exponentielle, la fonction de risque pouvant dépendre du temps. Nous avons :

$$\lambda(t) = \lambda\gamma(\lambda t)^{\gamma-1} \quad \text{et} \quad F(t) = \exp[-(\lambda t)^\gamma], \lambda, \gamma > 0$$

Si  $\gamma = 1$ , nous obtenons la distribution exponentielle. Le risque est monotone décroissant pour  $\gamma < 1$  et croissant pour  $\gamma > 1$ .

### 3) Distribution log-logistique

Comme la loi de Weibull, le modèle log-logistique dépend également de deux paramètres,  $\lambda$  et  $\gamma$ . Selon les valeurs de ses paramètres, cette loi correspond des situations où le risque de faire l'événement n'est pas monotone : il croît de 0 jusqu'à un maximum en  $t = (\gamma - 1)^{1/\gamma} / \lambda$  puis décroît ensuite. Cette loi a l'avantage de posséder une forme simple pour sa fonction de répartition. Nous avons :

$$\lambda(t) = \frac{\lambda\gamma(\lambda t)^{\gamma-1}}{1 + (\lambda t)^{\gamma-1}}, \lambda, \gamma > 0$$

$$F(t) = \frac{1}{1 + (\lambda t)^{\gamma-1}}$$

La fonction de risque est identique à celle de Weibull, si ce n'est son dénominateur  $1 + (\lambda t)^{\gamma-1}$ . Si  $\gamma \leq 1$ , la fonction de risque est monotone décroissante. Si  $\gamma > 1$ , le risque est croissant puis décroissant.

### C. Modèles semi-paramétriques de la fonction de risque

Nous avons présenté dans la section précédente, plusieurs distributions permettant de modéliser la survie d'une population homogène. Cependant, cette situation est réductrice. Il y a souvent des variables explicatives desquelles peut dépendre la survie.

L'approche semi-paramétrique est l'approche la plus utilisée depuis son introduction par Sir David COX [1972]. Elle ne pose pas d'hypothèse particulière sur la forme des distributions de survie et donc du risque instantané. Parallèlement, elle conduit à définir des modèles de régression en fonction des covariables. Le principe de base des modèles semi-paramétriques est la modélisation et la détermination de l'effet d'une covariable sur la fonction de survie (ou de risque), sans forme a priori.

Pour cela, il faut introduire  $Z = (Z_1, \dots, Z_n)$ , un vecteur de covariables dont la valeur est connue pour chaque individu; la variable explicative — quantitative ou qualitative — peut être un indicateur de traitement, d'âge, de sexe, de différentes grandeurs physiques ou biologiques, etc. Le cas le plus simple consiste en une variable indicative binaire.

Considérons également  $\beta$  qui est un vecteur de paramètres de la même dimension que  $Z_i$ . Si  $Z_i$  est un vecteur de taille  $1 \times k$ ,  $\beta$  est un vecteur de taille  $k \times 1$ . Il traduit l'effet de la covariable sur la fonction de risque (ou de survie).

Le **modèle de régression** ci-dessous ne spécifie pas la manière dont les covariables affectent le délai avant survenue de l'événement d'intérêt. Il permet d'exprimer le risque instantané comme suit :

$$\lambda(t; Z) = \phi [\lambda_0(t); Y(Z, \beta)] \quad (1.1)$$

- $\phi[\cdot]$  : fonction de lien entre  $\lambda_0(t)$  et la fonction de risque  $\lambda(t; Z)$ ;
- $\lambda_0(t)$  : fonction de risque instantané de base, non-paramétrée, générale et commune à tous les individus;
- $Y(\beta, Z)$  : fonction positive dépendante du vecteur de paramètre de régression,  $\beta$  et du vecteur de covariable,  $Z$ .

Dans le modèle (1.1), la fonction de risque est une fonction dépendante de la covariable  $Z$ . La fonction de lien  $\phi[\cdot]$  peut avoir différentes définitions comme nous allons le voir par la suite.

Il serait intéressant d'introduire des covariables dans les distributions des risques, présentées dans la section précédente. La généralisation de ces distributions permettent de suggérer plusieurs modèles de régression. La généralisation de la distribution exponentielle, par exemple, conduit à un modèle où  $\lambda_0(t)$  agit multiplicativement sur la fonction de risque (modèle de Cox). Mais d'autres formes d'action sont possibles, comme lorsque  $\lambda_0(t)$  agit additivement sur la fonction de risque (modèle de Aalen). Une généralisation des distributions de Weibull conduit à des modèles de régression où il y a une relation log-linéaire entre le temps et la covariable (modèles à risques accélérés).

#### 1) Modèle de Cox ou à risques proportionnels

Considéré comme une référence, le *modèle à risques proportionnels* ou *modèle de Cox* a été présenté par David Cox dans son article en 1972 (COX [1972]) et a servi de modèle de base pour tous les autres modèles de régression semi-paramétrique. C'est le modèle le plus répandu en analyse de survie notamment par son lien direct avec la vraisemblance partielle.

L'utilisation de la vraisemblance partielle — qui sera détaillée dans les sections suivantes — a été introduite pour la première fois dans cet article princeps. La vraisemblance partielle ne dépend que de la covariable  $\beta$  et il n'est pas nécessaire dans ce cas d'estimer la fonction de risque de base  $\lambda_0(t)$ .

La fonction de risque de T est définie par :

$$\lambda(t; Z) = \lambda_0(t) \exp(Z\beta), \beta > 0 \quad (1.2)$$

Dans ce modèle, la covariable a un effet multiplicatif sur la fonction de risque. Ainsi, dans la situation où il y a deux groupes à comparer ( $Z = 0$  ou  $1$ ), nous avons  $\lambda(t; Z=1)/\lambda(t; Z=0) = \exp(\beta)$ .

Plus généralement, nous avons :

$$S(t; Z) = S_0(t)^{\exp(Z\beta)}$$

$$\Lambda(t; Z) = \Lambda_0(t) \exp(Z\beta)$$

Si  $\lambda_0(t) = \lambda$ , le modèle (1.2) se réduit au modèle de régression exponentielle (généralisation de la distribution exponentielle) avec  $\lambda(t; Z) = \lambda \exp(Z\beta)$ .

Si  $\lambda_0(t) = \lambda \gamma (\lambda t)^{\gamma-1}$ , le modèle (1.2) se réduit au modèle de régression de Weibull (généralisation de la distribution de Weibull) avec  $\lambda(t; Z) = \lambda \gamma (\lambda t)^{\gamma-1} \exp(Z\beta)$ .

## 2) Autres modèles de régression

Différentes alternatives semi-paramétriques au modèle de Cox ont été proposées. Nous en présentons ci-dessous quelques-unes qui seront utiles par la suite.

### a) Risques additifs

Sous ce modèle additif, présenté par AALEN [1980], la fonction de risque a la forme additive suivante :

$$\lambda(t; Z) = \lambda_0(t) + Z\beta.$$

$\lambda_0(t)$  est la fonction de risque de base et  $\beta$  représente l'accroissement du risque. Le modèle est contraint par  $\lambda(t; Z) \geq 0$ . La covariable a un effet additif sur la fonction de risque. Dans la situation où il y a deux groupes à comparer, nous avons  $\lambda(t; Z = 1) = \lambda(t; Z = 0) + \beta$ .

### b) Survie accélérée (AFT)

L'expression générale du modèle est :

$$S(t; Z) = S_0(t e^{Z\beta}).$$

Nous en déduisons l'expression de la fonction de risque :

$$\lambda(t; Z) = \lambda_0(t e^{Z\beta}) e^{Z\beta}. \quad (1.3)$$

Dans un échantillon constitué de deux groupes, lorsque  $Z = 1$ , la fonction de survie associée  $S_1(t)$  au temps  $t$  est la même que celle de  $Z = 0$  au temps  $t e^{Z\beta}$ . On parle donc de *survie accélérée*.

Nous voyons de l'expression (1.3) que la covariable a un effet multiplicatif sur  $t$  plutôt que sur la fonction de risque, comme c'est le cas dans le modèle de Cox.

Le modèle AFT (pour *accelerated failure time*) peut également s'exprimer sous la forme d'une régression linéaire telle que :

$$\ln T = Z\beta + \epsilon \quad (1.4)$$

où  $\epsilon$  est une variable d'erreur d'espérance nulle dont la fonction de survie est telle que  $S_0[\ln(T)]$ . Si la distribution de T est log-normal, la distribution de la variable aléatoire  $\epsilon$  est normale. Le modèle (1.4) fait apparaître une relation linéaire entre  $\ln T$  et Z.

**c) Risques accélérés**

Ce modèle a été formalisé en 2001 par CHEN [2001]. Il permet d'étudier la relation entre les temps d'événement et la covariable à l'aide d'un changement d'échelle entre les fonctions de risque tel que :

$$\lambda(t;Z) = \lambda_0(te^{Z\beta}) \quad (1.5)$$

où  $\lambda(t;Z)$  est la fonction de risque pour la covariable  $Z$ , et  $\lambda_0(\cdot)$ , la fonction de risque de base non-spécifiée.

Il apparait du modèle (1.5) que lorsque  $Z = 1$ , sa fonction de risque instantané  $\lambda_1(t)$  au temps  $t$  a la même valeur que celle de  $Z = 0$  au temps  $te^{Z\beta}$ . On parle donc de *risque accéléré*.

Dans le modèle AFT, la covariable avait un effet (accélération, décélération) sur les fonctions de survie. Une comparaison détaillée du modèle de régression à risques accélérés avec les modèles de régression précédents a été faite par CHEN [2001].

**d) Risques odds-proportionnel**

La fonction de survie a la forme suivante :

$$\frac{S(t;Z)}{1-S(t;Z)} = \exp(Z\beta) \frac{S_0(t)}{1-S_0(t)}.$$

où  $S_0(t)$  est la fonction de survie de base (fonction de survie pour les sujets avec  $Z = 0$ ). On parle de modèle *odds-proportionnels* car la covariable a un effet multiplicatif sur les *odds* de la fonction de survie.

Pour obtenir une expression plus simple, on peut également considérer le logit de  $S$  telle que  $\text{logit}(S) = \ln(S/(1-S))$  ce qui nous donne :  $\text{logit}[S(t;Z)] = Z\beta + \text{logit}[S_0(t)]$ .

L'expression de la fonction de risque correspondante est :

$$\lambda(t;Z) = \lambda_0(t) \frac{1}{S_0(t)(e^{Z\beta} - 1) + 1}.$$

L'expression précédente nous permet de voir que les fonctions de risque  $\lambda(t;Z)$  et  $\lambda_0(t)$  se rapprochent tandis que  $t$  augmente puisque  $\lim_{t \rightarrow +\infty} S_0(t) = 0$ .

Le modèle odds-proportionnel, comme les autres modèles de régression, peut également être écrit sous la forme d'un *modèle de transformation linéaire*. Au lieu de  $\ln(T)$  comme en section b), nous utilisons une fonction inconnue, croissante  $h(T)$  telle que :

$$h(T) = Z\beta + \epsilon \quad (1.6)$$

avec  $\epsilon$  un terme d'erreur auquel est associée  $F_\epsilon$ , sa fonction de densité, complètement connue et  $h$  non-spécifié.

Si la distribution de l'erreur suit une loi logistique standard, l'équation (1.6) est celle d'un modèle odds-proportionnels. À titre de comparaison, si la distribution de l'erreur est celle de l'extremum généralisé, l'équation (1.6) est celle d'un modèle à risques proportionnels.

## II. Types de censure

On parle de données *censurées* lorsque l'information disponible sur : la date d'origine, la date d'événement ou la présence de l'événement est structurellement incomplète.

### A. Notations

Pour l'individu  $i$ , on note :

- $T_i$ , la variable aléatoire positive représentant la durée de vie avant événement, soit le temps écoulé entre la date d'origine et la date d'événement, avec :
  - fonction de survie  $S_T$
  - fonction de densité  $f_T$
  - fonction de risque instantané  $\lambda_T$
- $C_i$ , le temps de censure, c'est-à-dire la durée d'observation de l'individu entre le début de l'étude et la date de visite, avec :
  - fonction de survie  $S_C$
  - fonction de densité  $f_C$
  - fonction de risque instantané  $\lambda_C$

La variable aléatoire  $T$  n'est pas toujours connue avec précision. Ainsi, l'information sur  $T$  peut être exprimée par l'intervalle suivant :

$$T \in (L, R], L \leq R.$$

avec  $L$  et  $R$  des variables aléatoires, indépendantes de  $T$ .

Lorsque la date de survenue de l'événement est connue et exacte, alors  $L = R$ . Sinon, l'information est censurée; la censure peut être de plusieurs types : droite, gauche ou par intervalle.

### B. Censure droite et gauche

#### 1) Censure à droite

Les données sont censurées à droite en cas d'absence d'information sur le statut de l'individu après sa dernière observation. Deux cas sont possibles : l'individu n'a pas réalisé l'événement à la fin de l'étude (exclu-vivant) ou il a quitté l'étude en cours de route (perdu de vue ou exclu). Ainsi,  $T_i$  n'est pas connue, on sait juste qu'elle est supérieure à une certaine valeur. Une information censurée à droite est représentée par  $T_i \in (L_i, \infty]$ .

En présence de censure à droite, l'information disponible pour chaque individu est  $\{X_i; \delta_i\}$  avec :

- $X_i$ , la durée réellement observée :  $X_i = T_i \wedge C_i = \min(T_i, C_i)$ .
- $\delta_i = \mathbb{1}_{\{T_i \leq C_i\}}$ , un indicateur qui vaut 1 si l'événement est observé et 0 si l'individu est censuré.

#### 2) Censure à gauche

Une censure gauche survient si l'individu a déjà réalisé l'événement à la date d'observation. On a  $T_i \in (0, R_i]$ . Il y a donc une absence d'information sur la date exacte de survenue de l'événement; l'événement ayant déjà eu lieu à la date d'examen.

Par exemple, dans une étude sur l'âge d'apparition du trouble bipolaire de MONTLAHUC et al. [2017], si l'individu ne se souvient plus de sa première crise, on peut juste déduire que l'âge d'apparition est inférieur à l'âge actuel de l'individu. L'exemple classique de censure à gauche est l'âge de première consommation de cannabis présenté par TURNBULL et WEISS [1978].

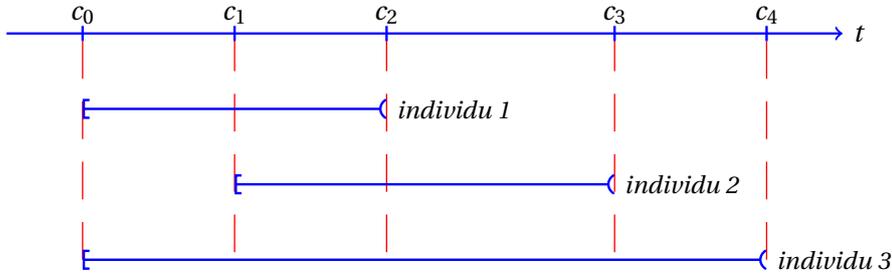
En présence de censure à gauche, l'information disponible pour chaque individu est  $\{X_i; \delta_i\}$  avec :

- $X_i$ , la durée réellement observée :  $X_i = T_i \vee C_i = \max(T_i, C_i)$ .
- $\delta_i = \mathbb{1}_{\{T_i \geq C_i\}}$ , un indicateur qui vaut 1 si l'événement est observé et 0 si l'individu est censuré.

### C. Censure par intervalle

Dans ce cas, l'événement n'est pas observé, mais il survient entre deux dates d'observation. La seule information disponible est qu'il a eu lieu entre deux dates connues. On a donc  $T_i \in (L_i, R_i]$ . Il est fréquent dans les études longitudinales que les patients ne soient pas suivis en continu; la date exacte de survenue de l'événement, s'il a lieu, ne sera jamais connue. La censure à droite et la censure à gauche sont des cas particuliers de censure par intervalle.

L'exemple — qui sera présenté par la suite — est constitué d'études de suivi de patients prenant des biothérapies. L'apparition d'anticorps anti-biothérapies (ADA) chez un patient ne peut être constaté que lors d'une visite. Sur le graphique ci-dessous, l'individu 2 a effectué deux visites : la première en  $c_1$  et la seconde, au temps  $c_3$ . Si lors de la seconde visite, il est déclaré ADA-positif, la seule information sur la date de positivité du patient est un intervalle de temps entre les deux visites, soit  $T_2 \in [c_1; c_3]$ .



#### 1) Censure par intervalle de type 1

La censure par intervalle de type 1 — *current status data* ou *cross-sectional data* — concerne les études où l'individu n'est observé qu'une fois durant l'étude (GROENEBOOM et WELLNER [1992]; JEWELL et VAN DER LAAN [2003]). Ce type de donnée est très courant en démographie, épidémiologie et en études de fiabilité. Encore une fois, la date de survenue de l'événement d'intérêt n'est pas connue. Seul, le *statut courant* du sujet est connu à la date de visite unique (avec événement ou sans).

Tous les intervalles contiennent soit zéro soit l'infini, c'est-à-dire que les observations sont soit censurées à gauche ( $L = 0$ ), soit censurées à droite ( $R = +\infty$ ).

Pour chaque patient, deux informations sont disponibles : le temps de suivi  $C$ , le statut du patient à cette date  $\delta$ . Ainsi, il est fréquent de représenter les censures par intervalle de type 1 par :

$$\{C, \delta = \mathbb{1}_{\{T \geq C\}}\}.$$

#### 2) Censure par intervalle de type 2

La censure par intervalle de type 2, souvent simplement appelée "censure par intervalle", inclut le cas d'une date d'événement appartenant à un intervalle, tel que  $T \in (L, R]$  (BOGAERTS et al. [2017]; GROENEBOOM et WELLNER [1992]; HUANG et WELLNER [1997]; LINDSEY et RYAN [1998]; SUN [2006]). Classiquement, la censure de type 2 inclut uniquement les études dans lesquels chaque patient n'effectue que deux visites.

L'information censurée par intervalle peut être exprimée par :

$$\{L, R, \delta_1 = \mathbb{1}_{\{T \leq L\}}, \delta_2 = \mathbb{1}_{\{L \leq T \leq R\}}, \delta_3 = 1 - \delta_1 - \delta_2\}.$$

Les  $\delta_1$  représentent les indicateurs de censures gauche, les  $\delta_2$  représentent les indicateurs de censure par intervalle et les  $\delta_3$ , les indicateurs de censures à droite.

Une extension naturelle de la censure de type 2, est la censure de type  $k$  où  $k$  est un nombre fixe de visites. Cependant, sachant qu'au plus deux intervalles d'observations parmi les  $(k + 1)$  seront nécessaires à l'encadrement de la variable  $T$ , il est possible de ne pas prendre en compte les autres intervalles qui n'apportent pas plus d'information sur  $T$  et de se ramener à un cas de censure de type 2. C'est le type de censure par intervalle le plus fréquent.

### 3) Remarques

#### Notations

Les variables  $L$  et  $R$  sont notées en majuscules ou en minuscules selon les auteurs. Par exemple, elles sont en majuscule chez LINDSEY et RYAN [1998], SUN [2006] et LI et MA [2013] et en minuscule chez BOGAERTS et al. [2017]. Généralement, les variables en minuscule correspondant aux observations (réalisations des variables aléatoires) tandis que les majuscules représentent les variables aléatoires sous-jacentes.

Or, avec la censure, cette convention peut présenter quelques difficultés de notations, car certains temps de censure peuvent être fixes et d'autres aléatoires. En effet, les variables  $L$  et  $R$  peuvent être fixes, selon le protocole de l'étude. Si  $L$  et  $R$  étaient fixées, cela ne modifierait pas la validité des vraisemblances, tests, etc. établis pour  $L$  et  $R$  aléatoires, tant que les choix de  $l$  et  $r$  ne sont pas liés à  $T$ .

Notons que l'on peut également rencontrer les notations  $L_i, R_i, l_i, r_i$ , qui correspondent respectivement aux variables aléatoires représentant les bornes de l'individu  $i$  et aux réalisations de ces variables.

Pour la suite, nous utiliserons, selon les situations, la forme majuscule ou minuscule des différentes variables ( $L, R$  et  $T$ ).

Il existe de même des différences de notations pour la covariable :  $Z, z, Z_i, z_i$ . Rigoureusement, du fait du conditionnement, il faudrait utiliser la réalisation de la variable  $Z$  dans les formules de vraisemblance, de score, etc.

#### Forme des intervalles

Il existe plusieurs manières d'écrire les intervalles :  $[L, R], [L, R), (L, R)$ . En pratique, l'utilisation d'intervalles fermés ou semi-ouverts dépend de la façon dont les observations ont été recueillies.

- L'utilisation d'intervalles fermés — utilisés depuis les articles pionniers de PETO [1973] et TURNBULL [1976] — permet d'inclure les informations exactes où  $L = R$  et des données groupées lorsque les temps d'inspection sont mesurés sur une échelle plutôt grossière (par exemple en années).
- L'utilisation d'intervalles semi-fermés — utilisés par GROENEBOOM et WELLNER [1992] — est la pratique la plus répandue et correspond au cas où les individus sont suivis par intermitence. Cette forme inclut les censures par intervalle de type 1 et 2.
- Enfin, YU et al. [2000] proposent un mélange de formes d'intervalles (ouverts, fermés, semi-ouverts) pour des données mixtes (censures droite, gauche et par intervalle, informations exactes).

On considère généralement qu'il n'y a pas de différence entre ces différentes formes d'intervalles tant que  $T$  est continu. Notons que NG [2002] montre que différentes interprétations des formes d'intervalles conduisent à des fonctions de vraisemblance différentes.

## D. Censure non-informative et indépendante

Les termes "censure indépendante" et "non-informative" sont souvent utilisées de manière similaire; nous définissons tout d'abord ces concepts pour la censure à droite puis pour la censure par intervalle.

### 1) Censure indépendante et non-informative droite

Le mécanisme de censure  $C$  est dit **indépendant** de la variable aléatoire  $T$  si le risque de faire l'événement s'appliquant à la population censurée est à chaque temps le même que le risque s'appliquant s'il n'y avait pas de censure. L'indépendance entre  $C$  et  $T$  implique :

$$\mathbb{P}(T \in [t, t + dt] | C \geq t, T \geq t) = \mathbb{P}(T \in [t, t + dt] | T \geq t), \forall t > 0.$$

Une forme moins restrictive portant sur le risque instantané est parfois également considéré (voir [FLEMING et HARRINGTON, 1991]) Cette hypothèse d'indépendance sous-tend en pratique que la visite ne dépende pas de l'état du patient. Considérons le cas d'une étude où l'événement d'intérêt est la rechute; si les patients les plus fragiles sont vus plus fréquemment, alors la condition d'indépendance entre censure et événement n'est évidemment pas vérifiée.

Le mécanisme de censure est dit **non-informatif** si la variable aléatoire de censure  $C$  ne dépend d'aucun paramètre utilisé pour modéliser la distribution de  $T$ .

L'hypothèse de censure non-informative est très importante tant du point de vue pratique que pour les méthodes d'estimation, elle permet notamment de simplifier nombre de modèles classiques d'analyse de survie. À l'inverse, si cette condition n'est pas vérifiée, l'estimation de la survie sera biaisée.

### 2) Censure indépendante et non-informative par intervalle

Considérons à présent le cas de la censure par intervalle, où les intervalles sont constitués de l'ensemble  $(L, R, T)$ . L'intervalle observé  $[l, r]$  implique que  $L = l$ ,  $R = r$  et  $T \in [l, R]$

L'hypothèse d'indépendance a été formalisée par SELF et GROSSMAN [1986] :

$$\mathbb{P}[T_i \leq t | L_i = l \leq T_i \leq r = R_i] = \mathbb{P}[T_i \leq t | l \leq T_i < r].$$

Le fait de savoir qu'une observation est censurée dans un intervalle donné ne fournit pas plus d'information sur la distribution de la variable d'intérêt,  $T$ , si ce n'est qu'elle appartient à un intervalle aux bornes définies. On pourrait également dire que la loi jointe de  $L$  et  $R$  ne contient aucune information relative à la distribution de  $T$ .

De manière équivalente, OLLER et al. [2004] définissent la censure non-informative par l'hypothèse que les observés  $(l, r)$  ne sont pas affectés par le fait que  $T$  appartienne à  $(l, r)$ . L'importance de la condition d'indépendance entre censure et événement apparaîtra lors de l'introduction de la vraisemblance.

### 3) Condition à somme constante

Il existe une condition plus faible que celle de non-informativité de la censure décrite précédemment, la condition à somme constante.

WILLIAMS et LAGAKOS [1977] ont originellement proposé des modèles censurés à droite dans lesquels les mécanismes de censure et de survie sont liés par ce qu'on appelle une condition à somme constante. Cette condition porte sur la loi jointe de  $T$  et  $C$  telle que  $X = \min(T, C)$  et a été redéfinie par KALBFLEISCH et MACKAY [1979] telle que :

$$\mathbb{P}(t \leq T \leq t + dt | T \geq t) = \mathbb{P}(t \leq T \leq t + dt | X \geq t).$$

BETENSKY [2000] présente les conditions de non-informativité et de somme constante pour les données censurées par intervalle de type 1. Puis, OLLER et al. [2004] ont étendu la définition de ces modèles aux données censurées par intervalle.

*Définition*

Un modèle censuré par intervalle est à somme constante si pour tout  $t$  tel que  $dF_T(t) \neq 0$  :

$$\int_{\{(l,r):t \in (l,r)\}} \frac{\mathbb{P}(L \in dl, R \in dr, T \in (l, r))}{\mathbb{P}(T \in (l, r))} = \int_{\{(l,r):t \in (l,r)\}} \mathbb{P}(L \in dl, R \in dr, T \in (l, r) | T \in (l, r)) = 1$$

où " $L \in dl$ " signifie que  $L$  est contenu dans un intervalle infinitésimal autour de  $l$ .

Ainsi, la condition de somme constante implique que le processus de modélisation peut omettre le caractère aléatoire des intervalles tandis que la condition de non-informativité implique que le mécanisme de censure ne peut affecter la distribution de la survie de la variable d'intérêt. Dans un modèle à somme constante, on considère que les intervalles observés sont fixés à l'avance et leur aspect aléatoire est ignoré. OLLER et al. [2004] montrent, dans le cas de la censure par intervalle, que la condition de somme constante suffit pour appliquer la vraisemblance simplifiée qui sera définie en section suivante.

À l'inverse, la présence de censure informative dans l'échantillon doit être prise en compte. Différents auteurs discutent de la présence de censure informative dans le cadre d'estimations ou de régressions sur des données censurées par intervalle (*par ex.* : BETENSKY et FINKELSTEIN [2002]; FARRINGTON et GAY [1999]; FINKELSTEIN et al. [2002]; WANG et al. [2010, 2016]; ZHANG et al. [2007]) et sur des données censurées par intervalle de type 1 (*par ex.* : CHEN et al. [2012]; ZHANG et al. [2005]; ZHAO et al. [2015]).

Sauf mention contraire, nous considérons dans la suite du document que les distributions des variables aléatoires  $L$  et  $R$  ne dépendent pas des paramètres utilisés pour modéliser la distribution de  $T$ ; ce qui garantit que la condition de censure indépendante soit satisfaite.

### III. Vraisemblance et censure par intervalle

Les méthodes statistiques d'estimation et de tests sont basées sur la vraisemblance introduite par FISHER [1922] qui s'établit à partir des données observées dont la distribution dépend d'un paramètre inconnu, ici noté  $\beta$ .

En présence de données censurées, la construction de la vraisemblance se fait en considérant l'information que peut fournir chaque observation; ce qui peut entraîner une forme complexe de vraisemblance. En 1972, COX [1972] propose une vraisemblance partielle pour données censurées à droite. Cette forme de "vraisemblance" simplifiée contribue à sa généralisation et à son emploi dans les tests basés sur la vraisemblance : les tests de Wald, du score et du rapport de vraisemblance.

Cependant, comme nous le verrons ultérieurement, la vraisemblance partielle s'applique dans un cas de censure à droite. Dans la majorité des autres cas, elle est impossible à spécifier et nécessite de considérer d'autres vraisemblances.

#### Généralités

Soit  $T_i, (i = 1, \dots, n)$ , un ensemble d'événements indépendants auxquels on associe une probabilité  $\mathbb{P}(i, \beta)$  avec  $\beta$  un vecteur de paramètre inconnu. La vraisemblance  $\mathcal{L}(\beta)$  d'un échantillon aléatoire constitué d'observations indépendantes est le produit de la contribution à la vraisemblance de chaque individu.

Une observation discrète a pour vraisemblance la probabilité  $\mathbb{P}(T = t)$ . Un temps d'événement exact fournit comme contribution à la vraisemblance la quantité  $f_T(t)$  où  $f_T(t)$  est la fonction de

densité à ce temps. Pour une observation censurée à droite, l'information disponible est que le temps d'événement est plus grand que le temps de surveillance; ce qui correspond à la fonction de survie évaluée en ce temps,  $S_T(t)$ . De même pour une observation censurée à gauche, l'information disponible est que l'événement a déjà eu lieu au temps  $t$ ; c'est donc la fonction de répartition,  $1 - S_T(t)$ . Et enfin pour finir, la censure par intervalle, correspond à un cas où l'événement a lieu dans un intervalle; l'information correspondante est la probabilité que  $T$  soit contenu dans cet intervalle, soit  $S_T(L) - S_T(R)$ . Les vraisemblances pour les divers types de censures peuvent être écrites en utilisant l'une des composantes suivantes dans lesquelles on a fait apparaître  $\beta$  :

Observation discrète exacte	—	$\mathbb{P}(T = t; \beta)$
Observation continue exacte	—	$f_T(t; \beta)$
Censure droite	—	$S_T(t; \beta)$
Censure gauche	—	$1 - S_T(t; \beta)$
Censure par intervalle	—	$S_T(L; \beta) - S_T(R; \beta)$

Nous introduisons tout d'abord la vraisemblance dans le cadre de la censure droite — un cas particulier de censure par intervalle — qui est la vraisemblance la plus fréquemment utilisée avant de s'intéresser à la vraisemblance dans le cadre de la censure par intervalle.

## A. Le cas particulier de la censure à droite

### 1) Vraisemblance complète

Les données disponibles sont  $(X_i, \delta_i)$  avec  $X_i = \min(T_i, C_i)$ , le temps d'événement ou de censure et  $\delta_i$ , un indicateur de présence ou d'absence d'événement chez l'individu  $i$  ( $i = 1, \dots, n$ ) (voir section 1) page 10). La censure est présente dans la fonction de vraisemblance.

La contribution à la vraisemblance pour un individu  $i$  est :

$$\mathcal{L}_i(\beta) = [f_T(T_i; \beta)S_C(T_i)]^{\delta_i} \times [f_C(T_i)S_T(T_i; \beta)]^{1-\delta_i}$$

Dans la vraisemblance ci-dessous, même les individus censurés apportent une contribution jusqu'à leur temps de censure.

### 2) Vraisemblance simplifiée

Si la censure est considérée comme *non-informative*, c'est-à-dire que le paramètre d'intérêt  $\beta$  n'apparaît pas dans la loi de la censure ( $S_C, f_C$ ), alors la vraisemblance pour une observation se réduit à :

$$\mathcal{L}_i(\beta) = f_T(T_i; \beta)^{\delta_i} S_T(T_i; \beta)^{1-\delta_i}. \quad (1.7)$$

Soit, pour l'ensemble des données :

$$\mathcal{L}(\beta) = \prod_{i=1}^n f_T(T_i; \beta)^{\delta_i} S_T(T_i; \beta)^{1-\delta_i}. \quad (1.8)$$

L'expression de la vraisemblance, dans laquelle le mécanisme de censure est ignoré, est appelée *vraisemblance simplifiée* par OLLER et al. [2004] dans le cadre de la censure par intervalle.

L'expression (1.8) peut également comporter des covariables via l'introduction d'un vecteur d'indicateur (soit,  $f_T(T_i; Z_i, \beta)$ ).

Si la fonction de vraisemblance dépend de plusieurs paramètres, mais que seule l'estimation d'un ou d'une partie des paramètres nous intéresse; les autres sont considérés comme des **paramètres de nuisance**.

### 3) Vraisemblance partielle

Dans l'objectif d'éliminer les paramètres de nuisance liés à la fonction de risque instantané de base, COX [1972] introduit dans son article princeps de 1972 la vraisemblance partielle dans le cadre du modèle de régression à risques proportionnels (voir section 1) page 7).

Cox se base sur le fait qu'aucune information sur le paramètre d'intérêt n'est apportée entre les temps d'événements. Au lieu de travailler sur l'ensemble des temps, on se place alors dans l'ensemble des temps où un événement a été constaté. Dans ce cas d'illustration, on suppose qu'un seul événement a lieu à chaque temps  $T_i (i = 1, \dots, d)$ .

De manière simplifiée, la probabilité qu'un événement ait lieu dans un petit intervalle de temps  $[T_i, T_i + dt[$  peut être approximée par :

$$\sum_{j \in R(T_i)} \lambda_0(T_i) \exp(Z_j \beta) \Delta t$$

avec  $R(T_i)$ , tous les individus à risque juste avant  $T_i$ .

La probabilité que ce soit l'individu  $i$  qui ait fait un événement sachant qu'un événement a eu lieu en  $T_i$  peut être approximée par :

$$\frac{\lambda_0(T_i) \exp(Z_i \beta) \Delta t}{\sum_{j \in R(T_i)} \lambda_0(T_i) \exp(Z_j \beta) \Delta t}$$

On peut supprimer le paramètre inconnu  $\lambda_0(T_i)$ . La vraisemblance ne dépend désormais que du paramètre  $\beta$  et plus de  $\lambda_0(T)$  dont l'estimation n'est pas nécessaire à chaque temps d'événement.

La vraisemblance partielle de Cox est alors exprimée comme le produit de ces probabilités conditionnelles :

$$\mathcal{L}_{VP}(\beta) = \prod_{i=1}^n \frac{\lambda_0(T_i) \exp(Z_i \beta)}{\sum_{j \in R(T_i)} \lambda_0(T_i) \exp(Z_j \beta)} = \prod_{i=1}^n \frac{\exp(Z_i \beta)}{\sum_{j \in R(T_i)} \exp(Z_j \beta)}$$

La vraisemblance partielle permet d'estimer et de tester le paramètre d'intérêt en ignorant la fonction de risque de base par maximisation de la vraisemblance. La justification théorique rigoureuse de la vraisemblance partielle a été apportée par les travaux sur les martingales et les processus de comptage initiés en France par BRÉMAUD et JACOD [1977] puis par les travaux d'application d'Andersen, Borgan, Gill et Keiding via le théorème Central Limite pour martingale (ANDERSEN P.K. et N. [1993]). Bien que la vraisemblance partielle ne soit pas une vraisemblance telle que définie par Fisher, elle en a ses propriétés asymptotiques et elle peut être utilisée pour estimer et tester les coefficients de régression avec de la censure à droite.

Cependant, la vraisemblance partielle n'est pas applicable dans le cas de la censure par intervalle. En effet, en ne connaissant pas la date d'événement exacte, il est impossible de connaître les temps d'événements et par conséquent le nombre de personnes à risques à chaque temps. Il existe toutefois une exception, LIN et al. [1998] ont développé une vraisemblance partielle sous le modèle additif pour des données censurées par intervalle de type 1 comme nous allons le présenter par la suite.

### 4) Vraisemblance marginale des rangs

Le développement de la vraisemblance marginale des rangs a été motivé par l'élimination des paramètres de nuisance de la vraisemblance dont le nombre augmente avec la taille de l'échantillon. Elle a tout d'abord été définie par KALBFLEISCH et SPROTT [1970] puis par KALBFLEISCH et PRENTICE [1973] pour la censure à droite sous un modèle de régression à risques proportionnels; le rang des individus est considéré plutôt que leur temps d'événement soit  $\mathbb{P}(r = [(1), \dots, (n)]; \beta)$ .

En présence de censure, la vraisemblance marginale de  $\beta$  est calculée comme la probabilité, au vu du sous-échantillon non censuré, d'observer l'échantillon total (avec censure). Plus l'échantillon est grand, plus le nombre de possibilités augmente. Considérons le sous-échantillon  $0 < t_1 < \dots < t_m < +\infty$ , des temps ordonnés où un événement a été constaté (non-censurés) et  $t_i \leq t_{i1}, \dots, t_{ik_i} < t_{i+1}$ , les temps de censure des individus censurés entre  $t_i$  et  $t_{i+1}$ .

*Nous rappelons que  $S(t) = \exp(-\int_0^t \lambda(u)du)$ , que dans un modèle de Cox, le risque associé à T la variable d'événement est  $\lambda(T, Z) = \lambda_0(T) \exp(Z\beta)$  et que dans la vraisemblance, une observation exacte est représentée par sa densité et une observation censurée est représentée par sa fonction de survie.*

On associe à l'événement  $t_i \leq t_{i1}, \dots, t_{ik_i}$  la probabilité conditionnelle :

$$\exp \left[ - \sum_{j=1}^{k_i} \exp(Z_{ij}\beta) \int_0^{t_i} \lambda_0(u) du \right],$$

afin d'obtenir la vraisemblance marginale de l'échantillon :

$$\begin{aligned} \mathcal{L}_{MR}(\beta) &\propto \int_0^\infty \int_{t_1}^\infty \dots \int_{t_{m-1}}^\infty \prod_{i=1}^m \lambda_0(t_i) \exp(Z_i\beta) \exp \left[ - \exp(Z_i\beta) \int_0^{t_i} \lambda_0(u) du \right] \exp \left[ - \sum_{j=1}^{k_i} \exp(Z_{ij}\beta) \int_0^{t_i} \lambda_0(u) du \right] dt_m \dots dt_1 \\ &\int_0^\infty \int_{t_1}^\infty \dots \int_{t_{m-1}}^\infty \prod_{i=1}^m \lambda_0(t_i) e^{Z_i\beta} e^{-\left( e^{Z_i\beta} + \sum_{j=1}^{k_i} e^{Z_{ij}\beta} \right) \int_0^{t_i} \lambda_0(u) du} dt_m \dots dt_1 \\ &= \prod_{i=1}^m \frac{\exp(Z_i\beta)}{\sum_{l \in R(t_i)} \exp(Z_l\beta)} \end{aligned} \tag{1.9}$$

où  $R(t_i)$  est le nombre de personnes à risque juste avant  $t_i$ . La démonstration de la simplification de l'expression (1.9) se trouve en annexe A page III. L'expression (1.9) est en parfait accord avec la vraisemblance partielle présentée plus haut.

## B. Le cas de la censure par intervalle

Nous allons à présent présenter l'écriture d'une vraisemblance adaptée au cas de la censure par intervalle.

### 1) Vraisemblance complète

La modélisation d'un échantillon constitué de  $n$  individus indépendants censurés par intervalle est réalisée à partir de la distribution jointe des variables aléatoires T et (L, R).

Puisque  $T \in (L, R]$  alors  $(T, L, R)$  est un sous-ensemble de  $\{(t, l, r) : 0 \leq l \leq t \leq r \leq +\infty\}$ . La contribution à la vraisemblance d'un individu censuré par intervalle sur  $(l, r]$  est donnée par la quantité suivante :

$$\mathcal{L}_i(\beta) = \mathbb{P}(L_i \in dl_i, R_i \in dr_i, T_i \in (l_i, r_i]; \beta) \tag{1.10}$$

où " $L_i \in dl_i$ " signifie que  $L_i$  appartient à un intervalle infinitésimal autour de  $l_i$ .

Tout comme pour la censure à droite, la vraisemblance complète inclut la censure. Dans l'expression simplifiée de la vraisemblance, la censure n'apparaît plus.

### 2) Vraisemblance simplifiée

Si la condition de non-informativité de la censure est satisfaite, la contribution (1.10) est proportionnelle à  $\mathbb{P}(T \in (l, r])$ , soit la probabilité que T appartienne à  $(l, r]$  en ignorant le mécanisme de censure. Cette probabilité est appelée *vraisemblance simplifiée* par OLLER et al. [2004].

La vraisemblance simplifiée d'une réalisation est :

$$\mathcal{L}_i(\beta) = \mathbb{P}(T_i \in (l_i, r_i]; \beta).$$

D'où la vraisemblance simplifiée des  $n$  réalisations indépendantes associées aux observés  $(l_1, r_1), \dots, (l_n, r_n)$  :

$$\mathcal{L}(\beta) = \prod_{i=1}^n \mathbb{P}(T_i \in (l_i, r_i]; \beta).$$

où la contribution à la vraisemblance des aléatoires,  $L_i, R_i$  est ignorée. Si la condition à somme-constante est satisfaite, la vraisemblance simplifiée peut également être utilisée.

### a) Censure par intervalle de type 1

Nous donnons à présent l'écriture d'une vraisemblance pour censure par intervalle de type 1 dont la population est constituée de  $n$  sujets indépendants avec  $T$ , le délai (inconnu) avant survenue de l'événement d'intérêt à laquelle on associe sa fonction de survie,  $S_T(T; Z, \beta)$  où  $Z$  est un vecteur de covariable. Et  $C$ , une variable aléatoire positive et continue, indépendante de  $T$ , représentant la date de visite. Pour un individu  $i$ , les informations disponibles sont de la forme :

$$\{(C_i, \delta_i, Z_i); i = 1, \dots, n\}, \quad (1.11)$$

avec  $\delta_i = \mathbb{1}_{(T_i \leq C_i)}$ , un indicateur d'événement.

La fonction de vraisemblance associée aux données (1.11) est :

$$\mathcal{L}(S(\cdot); \beta) = \prod_{i=1}^n [S(C_i|Z_i, \beta)]^{1-\delta_i} [1 - S(C_i|Z_i, \beta)]^{\delta_i}$$

Et la log-vraisemblance associée est :

$$\mathcal{L}_v(S(\cdot); \beta) = \sum_{i=1}^n (1 - \delta_i) \ln [S(C_i|Z_i, \beta)] + \delta_i \ln [1 - S(C_i|Z_i, \beta)].$$

### b) Censure par intervalle de type 2

Dans un échantillon constitué de  $n$  individus censurés par intervalle de type 2, les informations disponibles sont de la forme :

$$\{(L_i, R_i], Z_i; i = 1, \dots, n\}, \quad (1.12)$$

avec  $(L_i, R_i]$ , les intervalles observés auxquels  $T_i$  appartient.

La fonction de vraisemblance associée aux données (1.12) est :

$$\mathcal{L}(S(\cdot); \beta) = \prod_{i=1}^n \mathbb{P}(T_i \in (L_i, R_i]; Z_i, \beta) = \prod_{i=1}^n [S(L_i; Z_i, \beta) - S(R_i; Z_i, \beta)]$$

Et la log-vraisemblance associée est :

$$\mathcal{L}_v(S(\cdot); \beta) = \sum_{i=1}^n \ln [S(L_i; Z_i, \beta) - S(R_i; Z_i, \beta)]$$

avec  $L_i < R_i$  pour tout  $i = 1, \dots, n$ .

Tout comme pour la censure à droite, il existe plusieurs sortes de vraisemblances aux propriétés particulières. Nous en présentons quelques-unes ci-dessous.

### 3) Vraisemblance marginale des rangs

La vraisemblance marginale pour la censure à droite a été présentée en section 4) page 16. Avec la censure par intervalle, le calcul de la vraisemblance marginale est beaucoup plus compliqué; l'incertitude concernant la date d'événement augmentant le nombre de rangs possibles pour chaque individu.

La **vraisemblance marginale des rangs**, présentée dans SELF et GROSSMAN [1986] et SATTEN [1996] pour un modèle à risques proportionnels, est définie comme la somme des probabilités des rangs des temps d'événements (inconnus),  $T_i$ , sur l'ensemble de tous les rangs possibles des  $T_i$  consistant avec les données observées censurées par intervalle.

L'ensemble des rangs possibles des observations est noté  $\mathcal{R}$  ou  $R = (r_1, \dots, r_n)$  de  $\mathcal{R}$  est l'une des permutations possibles. Par exemple, si aucun intervalle ne se chevauche, il n'y aura dans  $\mathcal{R}$  qu'une permutation possible, donc un seul élément. La vraisemblance marginale est donnée par :

$$\mathcal{L}_{MR} = \sum_{R \in \mathcal{R}} \mathbb{P}(R; Z_i, \beta)$$

où  $\mathbb{P}(R; \beta, Z_i)$  est la probabilité du rang  $R$  dans le modèle à risques proportionnels classique en tenant en compte du paramètre d'intérêt  $\beta$  et de la variable explicative  $Z_i$ .

Cette procédure est une généralisation de la procédure équivalente disponible pour la censure à droite et a l'avantage de ne pas nécessiter une estimation de la fonction de survie de base,  $S_0(t)$ . Mais cette vraisemblance n'a pas une forme simple et est très longue et complexe à calculer.

#### Exemple

Si l'on considère un échantillon constitué de 3 individus  $(0, 5]; (3, 4]; (2, \infty]$ , l'ensemble  $\mathcal{R}$  contient 6 éléments vu qu'il y a 6 possibilités de rangs consistants avec les intervalles observés. Pour chacune de ces possibilités, il faut calculer la vraisemblance associée puis les sommer.

### 4) Vraisemblance groupée

La **vraisemblance groupée** est basée sur un groupement préalable des données.

En utilisant un modèle de transformation linéaire tel que  $g(T_i) = Z_i\beta + \epsilon_i$  où  $\epsilon_i$  est un terme d'erreur avec  $\mathbb{E}(\epsilon_i) = 0$  de fonction de répartition  $F$ , et  $g$  une fonction inconnue, croissante, la vraisemblance groupée s'écrit :

$$\begin{aligned} \mathcal{L}_{group}(\beta) &= \prod_{i=1}^n \int_{g(L_i)}^{g(R_i)} f(u_i - Z_i\beta) du_i \\ &= \prod_{i=1}^n [F\{g(R_i) - Z_i\beta\} - F\{g(L_i) - Z_i\beta\}] \end{aligned}$$

où  $f$  est la fonction de densité associée à la fonction de répartition  $F$ .

Soient  $\{q_j\}_{j=0}^m$ , les éléments uniques ordonnés de  $\{L_i, R_i; i = 1, \dots, n\}$ . Les seules contributions de  $g$  à la vraisemblance sont aux points  $g(q_j) = \gamma_{q_j}$ . Le vecteur  $\gamma = (\gamma_{q_1}, \dots, \gamma_{q_{m-1}})$  est un vecteur de paramètres de nuisance. Dans un modèle groupé,  $\gamma_{q_0} = -\infty$  et  $\gamma_{q_m} = +\infty$  pour que  $F(g(q_0) - Z_i\beta) = 0$  et  $F(g(q_m) - Z_i\beta) = 1$ . La vraisemblance devient :

$$\mathcal{L}_{group}(\beta, \gamma) = \prod_{i=1}^n [F(\gamma_{R_i} - Z_i\beta) - F(\gamma_{L_i} - Z_i\beta)]$$

La vraisemblance groupée permet de limiter le nombre d'estimations à effectuer puisqu'elle utilise un groupement préalable des données. Le nombre de paramètres à estimer est dès lors réduit par rapport à la vraisemblance "classique".

L'inconvénient de ce procédé, est qu'il n'existe pas de méthode unique de groupement de données. Grouper les données de manière trop fine nous ramène au problème du nombre trop important de paramètres à estimer tandis que les grouper de manière trop large entraîne une perte d'information et diminue la puissance du test (FAY [1996]). Ainsi, le choix des intervalles se fait en considérant ces deux problématiques.

Si la distribution du terme d'erreur est une distribution de type extreme value, on obtient un modèle à risques proportionnels comme proposé par FINKELSTEIN et WOLFE [1985].

### 5) Autres vraisemblances

HUANG et WELLNER [1995] utilisent une approche par **vraisemblance profilée** où l'estimation de la vraisemblance est réalisée par étapes successives. Une maximisation sur  $\Lambda$  (ou  $S$ ) est effectuée pour toutes les valeurs fixées de  $\beta$  pour obtenir  $\hat{\Lambda}(\cdot, \beta)$  puis une maximisation de la vraisemblance profilée  $\mathcal{L}_{prof}(\beta, \hat{\Lambda}(\cdot, \beta))$  est effectuée sur  $\beta$  pour identifier  $\hat{\beta}$  puis  $\hat{\Lambda}(\cdot, \hat{\beta})$ .

Enfin, comme nous l'avons mentionné précédemment, il n'est pas possible d'utiliser la **vraisemblance partielle** avec la censure par intervalle. Cependant, LIN et al. [1998] proposent une vraisemblance partielle pour données censurées par intervalle de type 1 sous un modèle de régression à risques additifs. L'approche proposée est basée sur le comptage des personnes qui n'ont pas fait leur événement et qui sont censurées. La date de censure est alors connue précisément permettant l'utilisation de la vraisemblance partielle. La vraisemblance partielle avec censure par intervalle et modèle additif est présentée de manière plus détaillée en section II. page 38.

Avant de clôturer cette section sur la vraisemblance, nous rappelons ici les expressions du score et de l'information de Fisher ainsi que des tests basés sur la vraisemblance qui seront largement utilisés par la suite.

## C. Tests basés sur vraisemblance

Nous supposons pour la suite que la vraisemblance du modèle considéré,  $\mathcal{L}_X(\beta)$ , est deux fois dérivable par rapport au paramètre  $\beta$ . Notons,  $\beta^*$ , la "vraie valeur" de  $\beta$  sous  $H_0$ , l'hypothèse nulle.

### 1) Score et information de Fisher

#### a) Score

Pour une observation  $i$  dont la log-vraisemblance est notée  $\mathcal{L}v_i(\beta)$ , on appelle score en  $\beta$ , la dérivée première de la log-vraisemblance par rapport au paramètre  $\beta$  :  $\frac{\partial \mathcal{L}v_i(\beta)}{\partial \beta}$ . Le score d'un échantillon de taille  $n$  est la somme des scores en  $\beta$  de chacune des  $n$  observations.

En la "vraie valeur" du paramètre  $\beta = \beta^*$ , l'espérance du score est nulle :

$$\mathbb{E} \left( \frac{\partial \mathcal{L}v_i(\beta)}{\partial \beta} \Big|_{\beta=\beta^*} \right) = 0$$

#### b) Information de Fisher

L'information de Fisher en  $\beta = \beta^*$  du modèle précédent est :

$$I(\beta) = \mathbb{E} \left[ - \frac{\partial^2 \mathcal{L}v_i(\beta)}{\partial \beta^2} \Big|_{\beta=\beta^*} \right] = \mathbb{E} \left[ \left( \frac{\partial \mathcal{L}v_i(\beta)}{\partial \beta} \Big|_{\beta=\beta^*} \right)^2 \right]$$

L'information de Fisher notée  $I(\beta^*)$  est le moment d'ordre deux de la fonction de score, soit la dérivée du logarithme de la vraisemblance par rapport au paramètre  $\beta$ . La fonction de score ayant une espérance nulle, l'information de Fisher correspond à **la variance de la fonction de score**.

La notion d'information de Fisher se généralise au cas où la distribution de probabilité de  $T$  dépend de plusieurs paramètres. Le score n'est plus un scalaire, mais un vecteur et l'information de Fisher, multidimensionnelle, est définie comme une matrice de covariance contenant les dérivées vis-à-vis des différentes composantes. Prenons l'exemple d'un score dépendant de deux paramètres,  $\beta$  et  $\gamma$ . En remplaçant les paramètres par leur valeur testée,  $\beta^*$  et  $\gamma^*$ , le score d'un individu  $i$  est :

$$U_i(\beta, \gamma) = \begin{pmatrix} \frac{\partial \mathcal{L}v_i(\beta, \gamma)}{\partial \beta} \Big|_{\substack{\beta=\beta^* \\ \gamma=\gamma^*}} \\ \frac{\partial \mathcal{L}v_i(\beta, \gamma)}{\partial \gamma} \Big|_{\substack{\beta=\beta^* \\ \gamma=\gamma^*}} \end{pmatrix}$$

La matrice d'information de Fisher correspondante est :

$$I_i(\beta, \gamma) = E \begin{bmatrix} \frac{\partial^2 \mathcal{L}v_i(\beta, \gamma)}{\partial \beta^2} \Big|_{\substack{\beta=\beta^* \\ \gamma=\gamma^*}} & \frac{\partial^2 \mathcal{L}v_i(\beta, \gamma)}{\partial \beta \partial \gamma} \Big|_{\substack{\beta=\beta^* \\ \gamma=\gamma^*}} \\ -\frac{\partial^2 \mathcal{L}v_i(\beta, \gamma)}{\partial \beta \partial \gamma} \Big|_{\substack{\beta=\beta^* \\ \gamma=\gamma^*}} & \frac{\partial^2 \mathcal{L}v_i(\beta, \gamma)}{\partial \gamma^2} \Big|_{\substack{\beta=\beta^* \\ \gamma=\gamma^*}} \end{bmatrix}.$$

## 2) Tests de $\beta = \beta^*$ basés sur la vraisemblance

Il existe trois tests asymptotiquement équivalents basés sur la vraisemblance permettant de tester l'hypothèse nulle  $\beta = \beta^*$ .

### a) Test du score

Pour tester l'hypothèse nulle  $H_0 : \beta = \beta^*$ , on se place sous  $H_0$  en remplaçant  $\beta$  par la valeur testée,  $\beta^*$ . La statistique du score est, dans le cas unidimensionnel, telle que  $W = U(\beta)^2 / I_n(\beta)$  où  $U(\beta)$  représente le score d'un  $n$ -échantillon dont l'expression apparait dans la section précédente et  $I_n(\beta)$  est la matrice d'information de Fisher correspondante. Dans le cas multidimensionnel (noté  $\beta_d$ ), le test du score est  $U(\beta_d)^T I_n(\beta_d)^{-1} U(\beta_d)$  avec  $I_n(\beta_d)^{-1}$ , l'inverse (ou pseudo-inverse) de la matrice d'information  $I_n(\beta_d)$ .

En  $\beta^*$ , la vraie valeur du paramètre, le score de l'observation  $i$ ,  $U_i(\beta^*)$  est d'espérance nulle et sa variance correspond à l'information de Fisher. En appliquant le théorème central limite au score du  $n$ -échantillon,  $U = \sum U_i$ , nous avons  $U(\beta^*) \sim \mathcal{N}(0; I_n(\beta^*))$ . Ainsi, si  $\beta = \beta^*$ , la statistique du score  $W = U(\beta^*)^2 / I_n(\beta^*)$  ou son équivalent pour le cas multidimensionnel, suit asymptotiquement un  $\chi^2$  à  $k-1$  degrés de liberté sous l'hypothèse nulle,  $k$  étant le nombre de groupes dans l'échantillon.

### b) Tests de Wald et du rapport de vraisemblance

Dans le test de Wald, sous l'hypothèse que la différence entre  $\hat{\beta}$  et  $\beta^*$  est normalement distribuée, le rapport  $(\hat{\beta} - \beta^*)^2 / V(\hat{\beta})$  suit une loi du  $\chi^2_{k-1}$  ddl.

Dans le test du rapport de vraisemblance, on compare deux vraisemblances : l'une, calculée sous l'hypothèse nulle  $\beta = \beta^*$  et l'autre, sous l'hypothèse alternative,  $\beta = \hat{\beta}$ . Sous  $H_0$ , on calcule la vraisemblance avec  $\beta = \beta^*$ . Sous  $H_1$ , on maximise la vraisemblance sans contrainte. Nous avons, asymptotiquement,  $2(\mathcal{L}v(\hat{\beta}) - \mathcal{L}v(\beta^*)) \sim \chi^2_{k-1}$  ddl.

Par la suite, nous présenterons un test développé sur la statistique du score. L'avantage du test du score par rapport aux deux autres tests est qu'il ne nécessite pas d'estimation du paramètre d'intérêt.

## IV. Estimateur de la fonction de survie avec censure par intervalle

L'estimation de la fonction de survie est un élément important dans l'analyse des données de survie.

Lorsque un modèle paramétrique peut raisonnablement être supposé pour la fonction de survie, son estimation est relativement simple (SUN [2006], Section 2.3). Lorsqu'on n'a aucun a priori, on cherche un estimateur non-paramétrique de la fonction de survie.

Dans le cas de la censure à droite, l'estimateur non-paramétrique du maximum de vraisemblance (*nonparametric maximum likelihood estimator* ou **NPMLE**) est celui de E. L. KAPLAN et MEIER [1958] avec sa variance estimée par la formule de GREENWOOD [1926]. Cependant, cet estimateur n'est pas applicable dans le cas de la censure par intervalle parce qu'on ne connaît pas la variable  $T_i$  ni le nombre de personnes à risque dans un intervalle proche de  $T_i$ ; des informations indispensables pour utiliser l'estimateur de Kaplan-Meier.

Pour des données censurées par intervalle, l'estimateur de la fonction de survie est plus compliqué à calculer. Il ne se fait qu'à l'aide d'algorithmes itératifs et ne possède généralement pas de solution unique, car sa détermination varie en fonction du choix des intervalles utilisés.

On notera qu'avec la censure droite, l'estimateur de Kaplan-Meier est considéré *indéfini* après la dernière observation si cette dernière est censurée à droite. En effet, le NPMLE n'est pas unique dans cette zone puisque des changements intervenant dans la fonction de survie après la dernière observation censurée n'affectent pas la vraisemblance des données observées. Un problème similaire et bien plus fréquent de non-unicité intervient avec la censure par intervalle.

Prenons le cas simple d'une étude où les patients ont été vus aux jours 5, 10, 15 et 20 et qu'à chaque date de visite un événement est constaté. Tout d'abord, nous n'avons aucune information sur la fonction de survie entre le jour 0 et le jour 5, entre le jour 5 et le jour 10, etc. Ensuite, si deux fonctions de survie correspondent aux jours 0, 5, etc. elles donneront la même vraisemblance. Pour l'exemple précédent, le NPMLE est ainsi représenté par des probabilités pondérées associées aux intervalles  $(0, 5], \dots, (15, 20]$ . Cela ne correspond pas à une fonction de survie unique, mais à n'importe quelle fonction qui maximisera la vraisemblance dans le cadre des intervalles définis plus haut.

### A. Censure par intervalle de type 1

Le cas des données censurées par intervalle de type 1 est plus simple puisqu'il existe une forme unique de l'estimateur du maximum de vraisemblance de la fonction de survie.

La vraisemblance simplifiée d'un échantillon de taille  $n$  est :

$$\mathcal{L}(S(.)) = \prod_{i=1}^n [S(C_i)]^{1-\delta_i} [1 - S(C_i)]^{\delta_i}$$

Il faut trouver l'estimateur du maximum de vraisemblance  $\hat{S} = 1 - \hat{F}$  qui maximise la vraisemblance. Notons que  $L(S(.))$  ne dépend de  $S$  qu'aux valeurs  $C_i$ .

Les  $\{q_j\}_{j=0}^m$  sont les éléments uniques ordonnés de  $\{0, C_i; i = 1, \dots, n\}$ . À chacun des temps de visite unique,  $q_j$ , nous disposons des informations suivantes :  $\mathbf{d}_j$ , le nombre d'événements ayant eu lieu à la date  $q_j$  (ou de manière équivalente, le nombre d'événements ayant eu lieu entre les dates de visite  $q_{j-1}$  et  $q_j$ ), et  $\mathbf{n}_j$ , le nombre de personnes encore en observation à la date  $q_j$ .

Une fonction isotonique respecte l'ordre des données comme la fonction de répartition  $F$  puisque  $q_1 \leq \dots \leq q_m$  et  $F(q_1) \leq \dots \leq F(q_m)$ . Pour obtenir une estimation  $\hat{F}$  qui maximisera la fonction de vraisemblance, nous allons donc minimiser l'expression suivante sous la contrainte  $F(q_1) \leq \dots \leq F(q_m)$  :

$$\sum_{j=1}^m n_j \left[ \frac{d_j}{n_j} - F(q_j) \right]^2 \tag{1.13}$$

L'expression (1.13) est une régression isotonique; c'est-à-dire que l'on cherche à trouver une fonction isotonique qui ajuste au mieux les observations. Les  $n_j$  (nombre de personnes encore à risque en  $q_j$ ) situés avant les crochets sont les poids de la régression isotonique. Afin d'estimer les  $F_{(q_j)}$ , nous utilisons l'algorithme du plus grand convexe-minorant (méthode graphique) ou de manière équivalente, la formule du max-min pour les régressions isotoniques.

## B. Censure par intervalle de type 2

Les données censurées par intervalle disponibles sont de la forme  $\{(L_i, R_i); i = 1, \dots, n\}$ . Comme précédemment, les  $\{q_j\}_{j=0}^m$  sont les éléments uniques et ordonnés de  $\{0, L_i, R_i; i = 1, \dots, n\}$  avec  $q_{m+1} = +\infty$ .

La vraisemblance simplifiée d'un échantillon de taille  $n$  est :

$$\mathcal{L}(p) = \prod_{i=1}^n [S(L_i)] - S(R_i) = \prod_{i=1}^n \sum_{j=1}^m \alpha_{ij} p_j$$

où  $\alpha_{ij} = \mathbb{1}_{\{q_j \in (L_i, R_i)\}}$  et  $p_j = S(q_{j-1}) - S(q_j)$ .

Comme pour la censure par intervalle de type 1, la fonction de vraisemblance ne dépend de la fonction de survie qu'à travers ses valeurs  $\{S(q_j)\}_{j=1}^m$  et la manière dont évolue la fonction de survie entre les  $q_j$  n'a pas d'effet sur l'estimation. La détermination du NPMLE de  $S$  est équivalent à la maximisation de  $\mathcal{L}(p)$  sachant que  $\mathbf{p}$  est contraint par  $\sum_{j=1}^m p_j = 1, p_j \geq 0$  (TURNBULL [1976]). Le NPMLE de  $S$  est ici représenté par un vecteur constitué de masses de probabilité et un ensemble d'intervalles sur lesquels les masses sont distribuées.

Il existe plusieurs approches pour déterminer le NPMLE de la fonction de survie. Celle de TURNBULL [1976] implique de tout d'abord déterminer quels sont les intervalles possibles puis de maximiser  $\mathcal{L}(p)$ . GENTLEMAN et GEYER [1994] énoncent les conditions de Kuhn-Tucker dont les deux premières sont  $\sum_{j=1}^m p_j = 1$  et  $p_j \geq 0$  pour vérifier que  $\hat{p}$  est le NPMLE.

Plusieurs algorithmes différents peuvent être utilisés de manière équivalentes pour déterminer le NPMLE de  $S$ ,  $F$  ou  $p$ . Parmi eux, l'algorithme auto-consistant — une application de l'algorithme EM — développé par TURNBULL [1976], l'algorithme itératif convexe-minorant ou ICM introduit par GROENEBOOM et WELLNER [1992], et l'algorithme EM-ICM — une combinaison de l'algorithme auto-consistant et de l'algorithme ICM — proposé par WELLNER et ZHAN [1997].

Enfin, il existe des situations plus complexes comme la situation de risques compétitifs où il y a plusieurs causes ou types d'événements d'intérêt (par exemple, dans les essais cliniques sur le cancer où l'on étudie la survenue de décès, de métastases, d'événements indésirables, etc.) avec éventuellement une covariable dépendante du temps, etc. Dans ce cas, la vraisemblance n'est pas la même que dans le cas simple et il faut avoir recours à des algorithmes adaptés comme celui de MAO et al. [2017].

## C. Validité asymptotique de l'estimateur

Pour la censure à droite, EFRON [1967] propose un estimateur "auto-consistant" (*self-consistent*) de  $F(t)$  tel que :

$$F^*(t) = \frac{1}{n} \sum_{i=1}^n \frac{\mathbb{P}(t_i \leq t | x, F^*)}{n} \quad (1.14)$$

où  $\mathbb{P}(t_i \leq t | x, F^*)$  est la probabilité conditionnelle que  $t_i \leq t$  en connaissant les données observées  $X$  et  $F^*$ . Pour identifier le meilleur  $F^*$ , l'équation (1.14) suggère l'utilisation d'un algorithme itératif.

Avec la censure à droite, il existe une relation d'équivalence entre les propriétés de l'estimation du maximum de vraisemblance et de l'estimation *self-consistent* (EFRON [1967]; LAIRD [2004]).

TURNBULL [1976] présente une extension de l'idée de Efron; il propose un *self-consistency algorithm* ("algorithme auto-consistant") pour données censurées par intervalle. L'estimation obtenue à l'aide de cet algorithme est équivalente à l'estimation du maximum de vraisemblance de F.

Une particularité des données censurées par intervalle est le très lent taux de convergence du NPMLE de F qui est de  $n^{1/3}$ , à comparer au taux de  $n^{1/2}$  des données censurées à droite (GROENEBOOM et WELLNER [1992]).

## V. Les tests non-paramétriques de comparaison de $k$ échantillons de données censurées

### A. Introduction

La comparaison de deux échantillons ou plus est un objectif majeur de nombre d'études cliniques. Dans le cas de la censure à droite, la majorité des tests non-paramétriques sont basés sur des **statistiques linéaires de rang** étendues au cas censuré; parmi lesquels, le test du log-rank et ses variantes sont les plus répandus. Il existe également des tests moins usités de comparaison de fonctions de survie tels ceux de Kaplan-Meier pondérés (PEPE et FLEMING [1989]).

Les tests de rangs pour données censurées à droite sont des généralisations de tests de rangs initialement proposés pour des données non-censurées. Ainsi, MANTEL [1966] propose en 1966 une extension de la statistique de Mantel-Haenszel pour données non-censurées. Ce test adapté est équivalent à la généralisation du test de Savage formalisé par RAO et al. [1960] et au test du log-rank proposé par PETO et PETO [1972].

De même, des généralisations des tests de Wilcoxon et de Kruskal-Wallis (test de Wilcoxon pour  $p + 1$  échantillons) ont été proposés par GEHAN [1965], BRESLOW [1970], PETO et PETO [1972] et PRENTICE [1978]. Par la suite, des adaptations de ces tests ont été développées pour des données censurées par intervalle comme nous le verrons par la suite.

Un aspect intéressant des tests de rangs est qu'il est possible avec une approche différente d'obtenir des statistiques équivalentes. Par exemple, le test de Wilcoxon peut être vu comme un test de log-rank pondéré ou considéré à l'aide d'une approche heuristique en admettant sous l'hypothèse nulle que les valeurs des groupes s'alternent de manière régulière dans l'échantillon total. Un autre exemple est l'équivalence entre la généralisation de la statistique de test de Savage aux données censurées, la statistique du score calculée sous la vraisemblance partielle avec risques proportionnels et le test du log-rank.

Dans les sections suivantes, nous présentons les principes généraux des tests de rangs pour données non-censurées et pour données censurées à droite. Dans le chapitre suivant, nous aborderons les tests de comparaison de deux groupes dans le cadre de la censure par intervalle.

### B. Principes généraux des tests linéaires de rangs pour deux échantillons

L'hypothèse nulle dans les tests de rangs est l'égalité des fonctions de survie dans les deux échantillons. Les deux séries d'observation,  $X_1$  et  $X_2$  de tailles respectives  $n_1$  et  $n_2$  ( $n = n_1 + n_2$ ) sont classées par ordre croissant :

$$x_{(1)} < \dots < x_{(n)}$$

À chaque observation  $x_{(i)}$  de rang ( $i$ ) est attribuée un score :  $c(i)$ . On peut construire la statistique linéaire de rang suivante :

$$R = \sum_{i=1}^n c(i)Z_{(i)} = \sum_{i=1}^n \omega_i Z_{(i)} \tag{1.15}$$

où  $Z_{(i)} = \mathbb{1}_{i_k}$  est une variable indicatrice qui vaut 1 si l'observation est issue du groupe  $k = 1$  et 0 sinon.

$c(\cdot)$  est la fonction de score associée à  $r_i$ , les rangs. Par la suite, elle est notée  $\omega_i$ <sup>1</sup>. Nous fixons  $\sum \omega_i = 0$ .

Par exemple :

- test de Wilcoxon pour données non-censurées :  $\omega_i = (i)^*$  ;
- test de Savage :  $\omega_i = 1 - \sum_{j=1}^i \frac{1}{n-j+1}$ .

\* :  $(i)$  dénote le rang des  $i$  observations.

### 1) Données non censurées

Sous l'hypothèse nulle, les temps d'événements ne sont pas associés à la covariable; ils sont donc indépendants et identiquement distribués. L'espérance  $\mathbb{E}(R)$  et la variance  $V(R)$  de  $R$  (expression 1.15) sont alors obtenues en considérant la permutation de la distribution des étiquettes des rangs  $(1), \dots, (n)$  (KALBFLEISCH et PRENTICE [2002]) :

$$\mathbb{E}(R) = \sum_{i=1}^n \omega_i \mathbb{E}(Z_{(i)}) = \sum_{i=1}^n \omega_i \bar{Z} = 0$$

$$V(R) = \mathbb{E}(RR') = \frac{1}{n-1} \sum_{i=1}^n (\omega_i^2 - \bar{\omega})^2 (Z_i - \bar{Z})' (Z_i - \bar{Z})$$

avec :

$$\bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i$$

On construit alors la statistique de test,  $\mathcal{T}$ , associée à la statistique de rang linéaire telle que :

$$\mathcal{T} = \frac{R - \mathbb{E}(R)}{\sqrt{V(R)}} \stackrel{H_0}{\sim} \mathcal{N}(0, 1)$$

$\mathcal{T}$  a une distribution asymptotique normale sous l'hypothèse nulle (RANDLES et WOLFE [1991]).

### 2) Données censurées

PRENTICE [1978] propose une généralisation des tests de rangs aux données censurées à droite. Considérons la suite ordonnée des temps d'événement notés  $t_1 < \dots < t_m$ . Avec  $t_{i1}, \dots, t_{ik}$  les valeurs censurées dans  $[t_i, t_{i+1}]$ ,  $i = 0, \dots, m$ . La statistique de rang linéaire associée est de la forme :

$$R = \sum_{i=1}^m \left( \omega_i Z_i + \sum_{j=1}^{k_j} \Omega_i Z_{ij} \right),$$

avec  $\omega_i$  et  $\Omega_i$  les scores pour des données non-censurées et censurées respectivement. Tous les individus dans  $[t_i, t_{i+1}]$  ont le même score  $\Omega_i$ .

Par exemple, pour le test de score exponentiel avec un modèle à risques proportionnels, les expressions des  $\omega_i$  et  $\Omega_i$  sont obtenues à partir de la fonction de score sous l'hypothèse nulle, soient  $\omega_i = \sum_{j=1}^i n_j^{-1} - 1$  et  $\Omega_i = \sum_{j=1}^i n_j^{-1}$  où  $n_j$  est le nombre de personnes à risque juste avant  $t_j$ .

La variance n'est pas calculée de manière permutationnelle comme précédemment, le score n'ayant plus de forme linéaire. La matrice d'information de Fisher observée fournit une variance généralement appropriée (KALBFLEISCH et PRENTICE [1980]).

1. Généralement, les scores dans une statistique linéaire de rangs sont notés  $c_i$ . Pour éviter une confusion avec la variable  $C$  représentant les temps de censure, nous les notons  $\omega_i$ .

Cependant, la variance peut être permutacionnelle si la statistique de test a une forme linéaire. Par exemple, la variance correspondante au test du score exponentiel avec un modèle à risques proportionnels ayant pour expression  $\sum_{i=1}^m (n_i^{-1} \sum_{l \in R(t_i)} Z_l - \bar{Z}_i)$  est :

$$V = \sum_{i=1}^m n_i^{-1} \sum_{l \in R(t_i)} (Z_l - \bar{Z}_i)(Z_l - \bar{Z}_i)'$$

où  $\bar{Z}_i = n_i^{-1} \sum_{l \in R(t_i)} Z_l$  est la moyenne des covariables des individus à risque juste avant  $t_i$ ,  $R(t_i)$  (PRENTICE [1978]).

### C. Les tests logranks

Parmi les tests de rangs, le test du log-rank et ses variantes, les log-rank pondérés occupent une place prépondérante, ils peuvent être envisagés selon plusieurs approches en tant que statistiques du score basés sur la vraisemblance partielle de Cox, pour une alternative à risques proportionnels ou en tant que statistiques des rangs. L'objectif du test est de tester l'hypothèse nulle  $H_0 : \lambda_0 = \lambda_1$ .

La première approche repose sur l'écriture de la fonction de score en supposant un modèle à risques proportionnels pour l'hypothèse alternative et à l'emploi de la vraisemblance partielle. La statistique obtenue est identique à celle obtenue avec le test du log-rank décrit ci-dessous.

Le test du logrank pour données censurées à droite peut être également présenté de manière heuristique, comme cela fut proposé par MANTEL [1966]. Le test consiste à comparer, à chaque temps d'événement,  $O$ , le nombre de décès observé, à  $E$ , le nombre de décès théorique que l'on aurait obtenu sous hypothèse nulle. Pour des données censurées à droite, la suite ordonnée des individus ayant fait l'événement dans l'échantillon commun est notée  $t_1 < \dots < t_m$ . À chaque  $t_j$  est associé  $d_{jl}$ , le nombre de décès et  $n_{jl}$ , l'effectif à risque dans le groupe  $l$ . On obtient alors les statistiques de test de la forme :

$$U = \sum_{j=1}^m w_j \left( d_{jl} - d_j \frac{n_{jl}}{n_j} \right) = \sum_{j=1}^m w_j U_j$$

où  $w_j$  est la pondération retenue.

#### Estimation de la variance

Il existe plusieurs manières de calculer la variance. Si on suit l'approche de Mantel avec indépendance des tableaux à chaque temps, on suppose sous l'hypothèse nulle d'égalité des distributions de survie dans les deux groupes, qu'à chaque temps,  $j$ , nous avons égalité des proportions de décès dans les deux groupes ( $\frac{d_{j1}}{n_{j1}} = \frac{d_{j2}}{n_{j2}}$ ) et  $d_{jl}$  suit une loi hypergéométrique  $H(n_j, d_j, \frac{n_{jl}}{n_j})$ .

On peut alors déduire l'espérance et la variance de  $d_{jl}$  et construire le test avec

$$v_0 = \sum_{j=1}^m w_j^2 d_j \frac{n_j - d_j}{n_j - 1} \frac{n_{j1} n_{j2}}{n_j^2}.$$

Ainsi, tous les tests non paramétriques de comparaison de groupes basés sur des statistiques de rang peuvent être mis sous cette forme. Par exemple, dans le test du log-rank,  $w_j = 1$ , dans celui de Wilcoxon-Gehan,  $w_j = n_j$ . La pondération choisie permet d'adapter le test à la forme des données testées.

Dans le chapitre suivant, nous présenterons les tests de comparaisons de plusieurs groupes destinés à la censure par intervalle. En particulier, le test de log-rank adapté à la censure par intervalle ainsi que des tests de pseudo-rangs linéaires seront exposés en détails.

## VI. Références

- AALLEN, O. 1980, «A model for nonparametric regression analysis of counting processes», dans *Mathematical Statistics and Probability Theory*, Springer New York, p. 1–25, doi :10.1007/978-1-4615-7397-5\_1. URL [https://doi.org/10.1007/978-1-4615-7397-5\\_1](https://doi.org/10.1007/978-1-4615-7397-5_1). 8
- ANDERSEN P.K., G. R., BORGAN O. et K. N. 1993, «Statistical models based on counting processes», Corr. 2. print éd., Springer series in statistics, Springer, New York, NY, ISBN 978-1-4612-4348-9. OCLC : 845452790. 16
- BETENSKY, R. 2000, «On nonidentifiability and noninformative censoring for current status data», *Biometrika*, vol. 87, n° 1, doi :10.1093/biomet/87.1.218, p. 218–221. URL <http://dx.doi.org/10.1093/biomet/87.1.218>. 14
- BETENSKY, R. A. et D. M. FINKELSTEIN. 2002, «Testing for dependence between failure time and visit compliance with interval-censored data», *Biometrics*, vol. 58, n° 1, p. 58–63, ISSN 0006-341X. 14
- BOGAERTS, K., A. KOMÁREK et E. LESAFFRE. 2017, *Survival analysis with interval-censored data : a practical approach with R, SAS and BUGS*, Chapman and Hall/CRC, ISBN 978-1420077476. 11, 12
- BRÉMAUD, P. et J. JACOD. 1977, «Processus ponctuels et martingales : résultats récents sur la modélisation et le filtrage», *Advances in Applied Probability*, vol. 9, n° 02, doi :10.2307/1426391, p. 362–416. URL <https://doi.org/10.2307/1426391>. 16
- BRESLOW, N. 1970, «A generalized kruskal-wallis test for comparing k samples subject to unequal patterns of censorship», *Biometrika*, vol. 57, n° 3, p. 579–594, ISSN 00063444. URL <http://www.jstor.org/stable/2334776>. 24
- CHEN, C.-M., T.-F. C. LU, M.-H. CHEN et C.-M. HSU. 2012, «Semiparametric transformation models for current status data with informative censoring», *Biometrical Journal. Biometrische Zeitschrift*, vol. 54, n° 5, doi :10.1002/bimj.201100131, p. 641–656, ISSN 1521-4036. 14
- CHEN, Y. Q. 2001, «Accelerated hazards regression model and its adequacy for censored survival data», *Biometrics*, vol. 57, n° 3, p. 853–860, ISSN 0006341X, 15410420. URL <http://www.jstor.org/stable/3068425>. 9
- COX, D. R. 1972, «Regression models and life-tables», *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 34, n° 2, p. 187–220, ISSN 00359246. URL <http://www.jstor.org/stable/2985181>. 7, 14, 16
- E. L. KAPLAN, E. et P. MEIER. 1958, «Nonparametric estimation from incomplete observations», *Journal of American Statistical Association*, vol. 53, p. 457–481. 22
- EFRON, B. 1967, *The two sample problem with censored data*, University of California Press, Berkeley, Calif., p. 831–853. URL <https://projecteuclid.org/euclid.bsmsp/1200513831>. 23
- FARRINGTON, C. P. et N. J. GAY. 1999, «Interval-censored survival data with informative examination times : parametric models and approximate inference», *Statistics in Medicine*, vol. 18, n° 10, p. 1235–1248, ISSN 0277-6715. 14
- FAY, M. P. 1996, «Rank invariant tests for interval censored data under the grouped continuous model», *Biometrics*, vol. 52, n° 3, p. 811–822, ISSN 0006-341X. 20
- FINKELSTEIN, D. M., W. B. GOGGINS et D. A. SCHOENFELD. 2002, «Analysis of failure time data with dependent interval censoring», *Biometrics*, vol. 58, n° 2, p. 298–304, ISSN 0006-341X. 14

- FINKELSTEIN, D. M. et R. A. WOLFE. 1985, «A Semiparametric Model for Regression Analysis of Interval-Censored Failure Time Data», *Biometrics*, vol. 41, n° 4, doi :10.2307/2530965, p. 933, ISSN 0006341X. URL <http://www.jstor.org/stable/2530965?origin=crossref>. 20
- FISHER, R. A. 1922, «On the mathematical foundations of theoretical statistics», *Philosophical Transactions of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, vol. 222, n° 594-604, doi :10.1098/rsta.1922.0009, p. 309–368, ISSN 0264-3952. URL <http://rsta.royalsocietypublishing.org/content/222/594-604/309>. 14
- FLEMING, T. R. et D. P. HARRINGTON. 1991, *Counting processes and survival analysis*, 2<sup>e</sup> éd., Wiley series in probability and mathematical statistics, Wiley, New York, ISBN 978-0-471-52218-8. 13
- GEHAN, E. A. 1965, «A generalized wilcoxon test for comparing arbitrarily singly-censored samples», *Biometrika*, vol. 52, n° 1/2, p. 203–223, ISSN 00063444. URL <http://www.jstor.org/stable/2333825>. 24
- GENTLEMAN, R. et C. GEYER. 1994, «Maximum likelihood for interval censored data : Consistency and computation», *Biometrika*, vol. 81, n° 3, doi :10.1093/biomet/81.3.618, p. 618–623, ISSN 0006-3444. 23
- GREENWOOD, M. 1926, *A Report on the Natural Duration of Cancer.*, 33, London : H.M.S.O. 22
- GROENEBOOM, P. et J. A. WELLNER. 1992, *Information bounds and nonparametric maximum likelihood estimation*, n° 19 dans DMV-Seminar, Birkhäuser, Basel Berlin, ISBN 978-3-7643-2794-1 978-0-8176-2794-2. OCLC : 246594378. 11, 12, 23, 24
- HUANG, J. et J. A. WELLNER. 1995, «Efficient estimation for the proportional hazards model with "case 2" interval censoring», . 20
- HUANG, J. et J. A. WELLNER. 1997, «Interval Censored Survival Data : A Review of Recent Progress», dans *Proceedings of the First Seattle Symposium in Biostatistics*, vol. 123, édité par P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg, I. Olkin, N. Wermuth, S. Zeger, D. Y. Lin et T. R. Fleming, Springer US, New York, NY, ISBN 978-0-387-94992-5 978-1-4684-6316-3, p. 123–169. URL [http://link.springer.com/10.1007/978-1-4684-6316-3\\_8](http://link.springer.com/10.1007/978-1-4684-6316-3_8). 11
- JEWELL, N. P. et M. VAN DER LAAN. 2003, «Current Status Data : Review, Recent Developments and Open Problems», dans *Handbook of Statistics*, vol. 23, Elsevier, ISBN 978-0-444-50079-3, p. 625–642. URL <http://linkinghub.elsevier.com/retrieve/pii/S0169716103230352>, DOI : 10.1016/S0169-7161(03)23035-2. 11
- KALBFLEISCH, J. D. et R. J. MACKEY. 1979, «On constant-sum models for censored survival data», *Biometrika*, vol. 66, n° 1, p. 87–90, ISSN 00063444. URL <http://www.jstor.org/stable/2335246>. 13
- KALBFLEISCH, J. D. et R. L. PRENTICE. 1973, «Marginal likelihoods based on cox's regression and life model», *Biometrika*, vol. 60, n° 2, doi :10.1093/biomet/60.2.267, p. 267–278. URL <http://dx.doi.org/10.1093/biomet/60.2.267>. 16
- KALBFLEISCH, J. D. et R. L. PRENTICE. 1980, *The statistical analysis of failure time data*, 15<sup>e</sup> éd., Wiley series in probability and mathematical statistics, Wiley, New York, NY, ISBN 978-0-471-05519-8. 25
- KALBFLEISCH, J. D. et R. L. PRENTICE. 2002, *The statistical analysis of failure time data*, 2<sup>e</sup> éd., Wiley series in probability and statistics, Wiley, Hoboken, NJ, ISBN 978-0-471-36357-6. 25
- KALBFLEISCH, J. D. et D. A. SPROTT. 1970, «Application of likelihood methods to models involving large numbers of parameters», *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 32, n° 2, p. 175–208, ISSN 00359246. URL <http://www.jstor.org/stable/2984524>. 16

- LAIRD, N. M. 2004, *Self-Consistency—I*, John Wiley & Sons, Inc., ISBN 9780471667193, p. 7538–7542, doi :10.1002/0471667196.ess2379.pub2. URL <http://dx.doi.org/10.1002/0471667196.ess2379.pub2>. 23
- LI, J. et S. MA. 2013, *Survival analysis in medicine and genetics*, Chapman & Hall/CRC biostatistics series, CRC Press, Boca Raton, Fla., ISBN 978-1-4398-9311-1. 12
- LIN, D. Y., D. OAKES et Z. YING. 1998, «Additive hazards regression with current status data», *Biometrika*, vol. 85, n° 2, doi :10.1093/biomet/85.2.289, p. 289–298. URL <http://biomet.oxfordjournals.org/content/85/2/289.abstract>. 16, 20
- LINDSEY, J. C. et L. M. RYAN. 1998, «Tutorial in biostatistics methods for interval-censored data», *Stat Med*, vol. 17, n° 2, p. 219–238. 11, 12
- MANTEL, N. 1966, «Evaluation of survival data and two new rank order statistics arising in its consideration», *Cancer Chemotherapy Reports. Part 1*, vol. 50, n° 3, p. 163–170, ISSN 0069-0112. 24, 26
- MAO, L., D. Y. LIN et D. ZENG. 2017, «Semiparametric regression analysis of interval-censored competing risks data», *Biometrics*, vol. 73, n° 3, p. 857–865. 23
- MONTLAHUC, C., E. CURIS, S. F. JONAS, F. BELLIVIER et S. CHEVRET. 2017, «Age-at-onset subsets of bipolar i disorders : A critical insight into admixture analyses», *International Journal of Methods in Psychiatric Research*, vol. 26, n° 3, doi :10.1002/mpr.1536, p. e1536–n/a, ISSN 1557-0657. URL <http://dx.doi.org/10.1002/mpr.1536>, e1536 IJMPR-May-2016-0054.R1. 10
- NG, M. P. 2002, «A modification of Peto’s nonparametric estimation of survival curves for interval-censored data», *Biometrics*, vol. 58, n° 2, p. 439–442, ISSN 0006-341X. 12
- OLLER, R., G. GOMEZ et M. CALLE. 2004, «Interval censoring : Model characterizations for the validity of the simplified likelihood», *Canadian Journal of Statistics*, vol. 32, n° 3, doi :10.2307/3315932, p. 315–326, ISSN 1708-945X. URL <http://dx.doi.org/10.2307/3315932>. 13, 14, 15, 17
- PEPE, M. S. et T. R. FLEMING. 1989, «Weighted Kaplan-Meier Statistics : A Class of Distance Tests for Censored Survival Data», *Biometrics*, vol. 45, n° 2, doi :10.2307/2531492, p. 497–507, ISSN 0006-341X. URL <http://www.jstor.org/stable/2531492>. 24
- PETO, R. 1973, «Experimental Survival Curves for Interval-Censored Data», *Applied Statistics*, vol. 22, n° 1, doi :10.2307/2346307, p. 86, ISSN 00359254. URL <http://www.jstor.org/stable/10.2307/2346307?origin=crossref>. 12
- PETO, R. et J. PETO. 1972, «Asymptotically Efficient Rank Invariant Test Procedures», *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, n° 2, doi :10.2307/2344317, p. 185, ISSN 00359238. URL <http://www.jstor.org/stable/10.2307/2344317?origin=crossref>. 24
- PRENTICE, R. L. 1978, «Linear rank tests with right censored data», *Biometrika*, vol. 65, n° 1, p. 167–179, ISSN 00063444. URL <http://www.jstor.org/stable/2335292>. 24, 25, 26
- RANDLES, R. H. et D. A. WOLFE. 1991, *Introduction to the theory of nonparametric statistics*, repr. éd., Krieger Publ, Malabar, Fla, ISBN 978-0-89464-543-3. 25
- RAO, U. V. R., I. R. SAVAGE et M. SOBEL. 1960, «Contributions to the theory of rank order statistics : The two-sample censored case», *Ann. Math. Statist.*, vol. 31, n° 2, doi :10.1214/aoms/1177705904, p. 415–426. URL <http://dx.doi.org/10.1214/aoms/1177705904>. 24
- SATTEN, G. A. 1996, «Rank-based inference in the proportional hazards model for interval censored data», *Biometrika*, vol. 83, n° 2, doi :10.1093/biomet/83.2.355, p. 355–370. URL <http://dx.doi.org/10.1093/biomet/83.2.355>. 19

- SELF, S. G. et E. A. GROSSMAN. 1986, «Linear Rank Tests for Interval-Censored Data with Application to PCB Levels in Adipose Tissue of Transformer Repair Workers», *Biometrics*, vol. 42, n° 3, doi :10.2307/2531202, p. 521, ISSN 0006341X. URL <http://www.jstor.org/stable/2531202?origin=crossref>. 13, 19
- SUN, J. 2006, *The Statistical Analysis of Interval-censored Failure Time Data*, Statistics for Biology and Health, Springer New York, New York, NY, ISBN 978-0-387-32905-5. URL <http://link.springer.com/10.1007/0-387-37119-2>. 11, 12, 22
- TURNBULL, B. W. 1976, «The empirical distribution function with arbitrarily grouped, censored and truncated data», *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 38, n° 3, p. 290–295, ISSN 00359246. URL <http://www.jstor.org/stable/2984980>. 12, 23, 24
- TURNBULL, B. W. et L. WEISS. 1978, «A likelihood ratio statistic for testing goodness of fit with randomly censored data», *Biometrics*, vol. 34, n° 3, p. 367–375, ISSN 0006341X, 15410420. URL <http://www.jstor.org/stable/2530599>. 10
- WANG, L., J. SUN et X. TONG. 2010, «Regression analysis of case 2 interval-censored failure time data with the additive hazards model», *Statistica Sinica*, vol. 20, n° 4, p. 1709–1723, ISSN 1017-0405. 14
- WANG, P., H. ZHAO et J. SUN. 2016, «Regression analysis of case  $K$  interval-censored failure time data in the presence of informative censoring», *Biometrics*, vol. 72, n° 4, doi :10.1111/biom.12527, p. 1103–1112, ISSN 0006341X. URL <http://doi.wiley.com/10.1111/biom.12527>. 14
- WELLNER, J. A. et Y. ZHAN. 1997, «A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data», *Journal of the American Statistical Association*, vol. 92, n° 439, doi :10.1080/01621459.1997.10474049, p. 945–959. URL <http://dx.doi.org/10.1080/01621459.1997.10474049>. 23
- WILLIAMS, J. S. et S. W. LAGAKOS. 1977, «Models for censored survival analysis : Constant-sum and variable-sum models», *Biometrika*, vol. 64, n° 2, doi :10.1093/biomet/64.2.215, p. 215–224. URL <http://dx.doi.org/10.1093/biomet/64.2.215>. 13
- YU, Q., L. LI et G. Y. C. WONG. 2000, «On consistency of the self-consistent estimator of survival functions with interval-censored data», *Scandinavian Journal of Statistics*, vol. 27, n° 1, doi :10.1111/1467-9469.00176, p. 35–44, ISSN 1467-9469. URL <http://dx.doi.org/10.1111/1467-9469.00176>. 12
- ZHANG, Z., J. SUN et L. SUN. 2005, «Statistical analysis of current status data with informative observation times», *Statistics in Medicine*, vol. 24, n° 9, doi :10.1002/sim.2001, p. 1399–1407, ISSN 02776715, 10970258. URL <http://doi.wiley.com/10.1002/sim.2001>. 14
- ZHANG, Z., L. SUN, J. SUN et D. M. FINKELSTEIN. 2007, «Regression analysis of failure time data with informative interval censoring», *Statistics in Medicine*, vol. 26, n° 12, doi :10.1002/sim.2721, p. 2533–2546, ISSN 0277-6715. 14
- ZHAO, S., T. HU, L. MA, P. WANG et J. SUN. 2015, «Regression analysis of informative current status data with the additive hazards model», *Lifetime Data Analysis*, vol. 21, n° 2, doi : 10.1007/s10985-014-9303-y, p. 241–258, ISSN 1572-9249. 14

## Chapitre 2

# Tests de comparaison de groupes de données censurées par intervalle

### Sommaire

---

<b>I. Trois formes générales de tests pour données censurées par intervalle</b> . . . . .	<b>32</b>
A. Test du score général pour censure par intervalle . . . . .	33
B. Test de permutation pour censure par intervalle . . . . .	34
C. Test du (pseudo) log-rank pour censure par intervalle . . . . .	35
D. Comparaisons des trois approches . . . . .	37
<b>II. Les tests implémentés existants</b> . . . . .	<b>38</b>
A. Les tests du score avec vraisemblance groupée . . . . .	38
B. Les (pseudo) log-ranks pour censure par intervalle . . . . .	39
C. Les (pseudo) log-ranks pondérés pour censure par intervalle . . . . .	41
D. Les tests basés sur la vraisemblance partielle . . . . .	44
<b>III. Références</b> . . . . .	<b>47</b>

---

## Introduction

Avec les données censurées à droite, la comparaison de deux groupes ou plus peut être réalisée à l'aide d'approches semi-paramétriques, basées sur la vraisemblance partielle, ou d'approches non-paramétriques qui coïncident comme dans le cas du log-rank. Cependant, cela n'est pas le cas pour les données censurées par intervalle.

En pratique les données censurées par intervalle sont souvent traitées de manière simplifiée. L'approche dite naïve, la plus fréquemment utilisée, est l'imputation simple qui consiste à choisir comme pseudo-date d'événement le milieu, voire la fin, de l'intervalle. Cette approche permet alors l'utilisation des tests de log-ranks classiques avec ou sans pondération ainsi que de tous les autres tests destinés à la censure à droite<sup>1</sup>.

Toutefois, la méthode d'imputation du milieu de l'intervalle est particulièrement sensible à la différence du mécanisme de censure pouvant exister entre les groupes. Cette méthode peut également être moins puissante que des tests du log-ranks adaptés à la censure par intervalle comme nous le montrerons à travers des simulations dans les sections suivantes. La perte de puissance est d'autant plus importante que les intervalles considérés sont grands. LAW et BROOKMEYER [1992] ont montré à travers des simulations que dans ce cas, le risque de première espèce augmentait et qu'il pouvait atteindre la valeur de 0,19. Enfin, il a été montré par PANAGEAS et al. [2007] que l'utilisation de la méthode d'imputation simple dans des études de cancérologie biaisait les estimations du temps de survie médian et de la fonction de survie.

Une approche alternative à l'imputation simple est l'imputation multiple. Les données censurées sont traitées comme des données manquantes auxquelles on essaie d'attribuer la date d'événement la plus plausible pour les transformer en données censurées à droite. PAN [2000] a proposé une méthode d'imputation multiple bayésienne avec bootstrap permettant d'estimer la date de survenue exacte de l'événement dans un modèle de Cox. Ces méthodes d'imputations multiples sont cependant lourdes à mettre en pratique.

Il apparait que des tests spécifiquement conçus pour les données censurées par intervalle sont nécessaires. Divers auteurs ont proposé des adaptations des tests du score, tests du log-ranks ou d'autres tests de rangs. De même que pour la censure à droite, il existe plusieurs approches possibles pour établir des tests de comparaison de deux groupes ou plus; des approches différentes pouvant donner lieu à des expressions équivalentes. Le choix de la variance se fait en fonction de l'approche choisie.

Par la suite, nous allons nous intéresser spécifiquement aux tests semi-paramétriques et plus particulièrement à celui basé sur la statistique du score qui a pour avantage de ne pas nécessiter d'estimation du paramètre d'intérêt. Des tests non-paramétriques, ici, des log-ranks adaptés à la censure par intervalle, seront également détaillés.

Nous présentons dans ce qui suit les différentes approches permettant de tester l'hypothèse nulle d'égalité des fonctions de survie. Nous introduirons dans la partie suivante le test que nous avons proposé spécifiquement conçu pour les situations de risques croisés.

## I. Trois formes générales de tests pour données censurées par intervalle

Les tests de comparaisons de deux ou plusieurs groupes pour données censurées par intervalle peuvent être rattachés à l'une des trois catégories suivantes : test du score "classique" avec matrice d'information de Fisher, test de permutation et test du (pseudo) log-rank. Nous détaillons par la suite ces trois catégories.

---

1. Le test du log-rank pour censure à droite est basé sur la vraisemblance partielle, inapplicable dans le cas de la censure par intervalle

### A. Test du score général pour censure par intervalle

FAY [1999] propose une formulation générale du test de score pour la censure par intervalle. Cette formulation permet d'inclure les tests du score sous un modèle à risques proportionnels, à odds-proportionnels et sous le modèle logistique.

Comme précédemment,  $T_i > 0$  est la variable représentant la durée écoulée avant l'événement de l'individu  $i$  et l'espace des temps est partitionné en  $(m + 1)$  intervalles avec  $Q_j = (q_{j-1}, q_j]$  et  $0 = q_0 < q_1 < \dots < q_m < q_{m+1} = +\infty$ . Comme vu en section B, page 23, à la place de  $T_i$ , nous observons que  $T_i \in (L_i, R_i]$  et  $L_i, R_i \in Q_j$ . Nous avons,  $\alpha_{ij} = \mathbb{1}_{\{Q_j \in (L_i, R_i]\}}$ , une fonction indicatrice qui prend la valeur 1 dans l'intervalle  $(L_i, R_i]$ .

Selon le test, les  $Q_j$  peuvent représenter un ensemble de points discrets ou une échelle discrète de temps continu (voir par exemple SUN [1996]) ou des intervalles de temps continu (comme dans FINKELSTEIN et WOLFE [1985], FAY [1996] ou OLLER et GÓMEZ [2012]). La manière dont est défini  $T$  ne modifie pas le calcul des tests.

Le test du score permet de tester l'effet de la covariable,  $Z$ , sur le délai avant survenue de l'événement,  $T$ . Pour un sujet  $i$ ,  $Z_i$  est un vecteur de taille  $1 \times k$  de covariable;  $k + 1$  étant le nombre de groupes dans l'échantillon. Il est constitué de 0 excepté dans la  $l^{\text{ième}}$  place (correspondant au groupe d'appartenance du patient  $i$ ) où il y a un 1.  $\beta$  est un vecteur de paramètre traduisant l'effet de la covariable de taille  $k \times 1$ .

La survie d'un individu  $i$  est modélisée par

$$\mathbb{P}(T > q_j; Z_i) = S(q_j; Z_i \beta, \gamma)$$

où  $\gamma$  est un vecteur de paramètres de nuisance de taille  $m \times 1$  associé à la fonction de survie de base.

Sous  $H_0 : \beta = 0$ , l'estimation non-paramétrique du maximum de vraisemblance, le NPMLE, est utilisé pour estimer  $\hat{\gamma}$ . La fonction de survie générale est une fonction dépendante de l'ensemble  $Z_i \beta$ . La manière dont agit la covariable sur la survie, c'est-à-dire la forme de  $S(q_j; Z_i \beta, \gamma)$  n'est pas précisée ici; elle dépend du modèle de régression choisi (risques proportionnels, odds proportionnels, additifs...). Par hypothèse, nous avons  $S(q_0; Z_i \beta, \gamma) = 1$  et  $S(q_{m+1}; Z_i \beta, \gamma) = 0$ . Le cas où  $S(q_{m+1}; Z_i \beta, \gamma) \neq 0$  sera présenté un peu plus tard lors de l'introduction des modèles à plateau.  $S$  est une fonction décroissante telle que  $1 > S(q_1) > \dots > S(q_m) > 0$ .

La vraisemblance d'un échantillon de données censurées par intervalle est :

$$\begin{aligned} \mathcal{L}(\beta) &= \prod_{i=1}^n \mathbb{P}(T_i \in (L_i, R_i] | \beta) \\ &= \prod_{i=1}^n [S(L_i; Z_i \beta, \gamma) - S(R_i; Z_i \beta, \gamma)] \\ &= \prod_{i=1}^n \sum_{j=1}^{m+1} \alpha_{ij} [S(q_{j-1}; Z_i \beta, \gamma) - S(q_j; Z_i \beta, \gamma)] \end{aligned}$$

Sous  $H_0 : \beta = 0$ , le score déduit de cette vraisemblance est un vecteur  $U_S$  de taille  $m + 1 \times 1$  où  $m$  est la dimension de  $\gamma$ . Le vecteur de score est défini ainsi :

$$U_S = \begin{pmatrix} \frac{\partial \mathcal{L}(\beta, \gamma)}{\partial \beta} \Big|_{\substack{\beta=0 \\ \gamma=\hat{\gamma}}} \\ \frac{\partial \mathcal{L}(\beta, \gamma)}{\partial \gamma_1} \Big|_{\substack{\beta=0 \\ \gamma=\hat{\gamma}}} \\ \dots \end{pmatrix} = \begin{pmatrix} \frac{\partial \mathcal{L}(\beta, \gamma)}{\partial \beta} \Big|_{\substack{\beta=0 \\ \gamma=\hat{\gamma}}} \\ 0 \\ \dots \end{pmatrix}$$

Le premier terme du vecteur de score est constitué de la dérivée première en  $\beta$  de  $\ln(\mathcal{L}(\beta, \gamma))$  évaluée en  $\beta = 0$  et  $\gamma = \hat{\gamma}$ . Les autres termes du vecteur de score sont nuls si les composantes du paramètre de nuisance sont substituées par leur estimation du maximum de vraisemblance,  $\hat{\gamma}$  (COX et HINKLEY [2000]).

L'estimateur du maximum de vraisemblance sous l'hypothèse nulle ( $\beta = 0$ ) est le *self-consistent estimator* (voir section C. page 23) de la survie,  $\hat{S}(q_j)$  avec  $\hat{S}(q_0) = 1$  et  $\hat{S}(q_{m+1}) = 0$ . Nous supposons qu'aucune estimation n'est dans les limites de l'espace de définition (c'est-à-dire, l'estimation  $\hat{S}(j-1)$  n'est pas trop proche de la suivante  $\hat{S}(j)$ ,  $\forall j \in [1, m]$ , soit  $1 > \hat{S}(q_1) > \dots > \hat{S}(q_m) > 0$ ) et que la censure ne dépend pas de l'événement d'intérêt. Les problèmes liés à l'estimation de la fonction de survie dans le cas d'un test du score avec censure par intervalle sont discutés un peu plus loin, en section D. page 37.

Au score,  $U_S$ , est associée sa variance,  $V$  définie par la matrice d'information de Fisher introduite en section 1) page 20. En reprenant les notations de FAY [1996], la matrice d'information de Fisher a la forme suivante :

$$A(\beta, \gamma) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

$$\text{où } A_{11} = \frac{-\partial^2 \mathcal{L}(\beta, \gamma)}{\partial \gamma^2}, A_{12} = A_{21}^T = \frac{-\partial^2 \mathcal{L}(\beta, \gamma)}{\partial \gamma \partial \beta}, A_{22} = \frac{-\partial^2 \mathcal{L}(\beta, \gamma)}{\partial \beta^2}.$$

À partir du score  $U_S$ , nous construisons le test du score tel que

$$W = U_S^T V_{0, \hat{\gamma}}^{-1} U_S$$

où  $V^{-1} = A_{22} - A_{21} A_{11}^{-1} A_{12}$  (voir Section 9.3 (iii) de COX et HINKLEY [2000]) et  $V_{0, \hat{\gamma}}^{-1}$  est  $V^{-1}$  évaluée en  $\beta = 0$  et  $\gamma = \hat{\gamma}$  (voir FINKELSTEIN et WOLFE [1985] ainsi que les annexes de SUN [1996] et FAY [1996, 1999]).

Sous l'hypothèse nulle et avec les conditions de régularité respectées,  $W$  suit asymptotiquement un  $\chi^2$  à  $k$  degrés de liberté,  $k + 1$  étant le nombre de groupe à comparer (FAY [1996, 1999]; FINKELSTEIN et WOLFE [1985]; SUN [1996]).

## B. Test de permutation pour censure par intervalle

Nous avons vu à la section précédente le test du score avec la matrice d'information de Fisher pour données censurées par intervalle. Le score associé à  $\beta$  peut également être écrit sous une forme permettant l'utilisation de la variance par permutation. Soit, le score de  $\beta$  tel que :

$$\begin{aligned} U &= \left. \frac{\partial \mathcal{L}(\beta, \gamma)}{\partial \beta} \right|_{\beta=0, \gamma=\hat{\gamma}} \\ &= \sum_{i=1}^n z_i \left( \frac{\sum_{j=1}^{m+1} \alpha_{ij} [\hat{S}'(q_{j-1}; \beta) - \hat{S}'(q_j; \beta)]}{\sum_{u=1}^{m+1} \alpha_{iu} [\hat{S}(q_{u-1}; \beta) - \hat{S}(q_u; \beta)]} \right) \\ &= \sum_{i=1}^n Z_i \left( \frac{\hat{S}'(L_i; \beta) - \hat{S}'(R_i; \beta)}{\hat{S}(L_i; \beta) - \hat{S}(R_i; \beta)} \right) \\ &= \sum_{i=1}^n Z_i \omega_i \end{aligned}$$

avec  $\hat{S}'$ , la dérivée première de la fonction de survie telle que :

$$\hat{S}'(q_j; \beta) = \left. \frac{\partial S(q_j | z_i \beta, \gamma)}{\partial \eta} \right|_{\beta=0, \gamma=\hat{\gamma}}$$

et  $\hat{S}'(q_0; \beta) = \hat{S}'(q_{m+1}; \beta) = 0$ .

La forme linéaire du vecteur de score obtenu suggère de considérer le test de l'hypothèse nulle à partir d'un test de permutation. Il est à noter que l'approche par test de permutation n'est valide que lorsque le mécanisme de censure est identique dans tous les groupes.

L'idée de l'approche permutationnelle est que si l'hypothèse nulle est vraie, les  $z_i$  associés aux scores  $\omega_i$  sont échangeables. La distribution permutée de  $U$  est alors obtenue en réalisant une permutation de tous les labels (les  $z_i$ ) et en calculant pour chaque permutation possible, la valeur du score.

Ainsi, lorsque l'échantillon est petit, il est possible de calculer exactement la distribution de la permutation. Lorsque  $n$  est grand, une version asymptotique est disponible avec le théorème central limite pour des variables aléatoires échangeables (SEN [2004]). Les  $\omega_i$  sont en effet non-indépendants, mais échangeables sous l'hypothèse nulle.

L'utilisation de la variance permutationnelle nécessite la condition supplémentaire que le mécanisme de censure ne soit pas lié à la covariable testée. Cela signifie que le mécanisme de censure doit être identique dans les différents groupes.

La variance permutationnelle asymptotique de  $U$  sous l'hypothèse nulle est :

$$V_p = (n-1)^{-1} \left\{ \sum_{i=1}^n (\omega_i - \bar{\omega})^2 \right\} \left\{ \sum_{i=1}^n (Z_i - \bar{Z})^T (Z_i - \bar{Z}) \right\}$$

(voir section 7.2 dans KALBFLEISCH et PRENTICE [2002] et OLLER et GÓMEZ [2012]).

L'obtention de l'expression de la variance permutationnelle sera discutée plus en détails lors de la présentation du test de score adapté aux risques croisés.

Lorsque  $n$  est grand, la  $p$ -value est estimée à l'aide de :

$$W_p = UV_p^{-1}U^T$$

sachant que  $W_p$  suit asymptotiquement un  $\chi^2$  à  $k$  degrés de liberté (FAY [1996]; OLLER et GÓMEZ [2012]).

### C. Test du (pseudo) log-rank pour censure par intervalle

Le test du log-rank pour censure à droite est basé sur la différence entre un nombre de personnes observées à risque ( $d$ ) et un nombre attendus de personnes à risque ( $e$ ). Dans le cas de la censure par intervalle, du fait de l'absence d'information sur la date exacte de survenue de l'événement, on ne connaît pas le nombre de personnes à risque aux différents temps. L'idée est donc de calculer des pseudo-nombres de personnes à risque comme dans SUN [1996, 2006], FAY [1999] ou ZHAO et SUN [2004].

Les log-rank pondérés sont de la forme :

$$\begin{aligned} U_l &= \sum_{j=1}^m w_j \left[ d_{jl}^* - e_{jl} \right] \\ &= \sum_{j=1}^m w_j \left[ d_{jl}^* - \frac{n_{jl}^* d_j^*}{n_j^*} \right] \end{aligned}$$

Avec  $d_{jl}^*$ , la valeur estimée sous l'hypothèse nulle du nombre d'événements au temps  $q_j$  pour le  $l^{\text{ième}}$  groupe,  $d_j^*$ , la valeur estimée sous l'hypothèse nulle du nombre total d'événements au temps  $q_j$ . De manière similaire, les  $n_{jl}^*$  et  $n_j^*$ , sont les valeurs estimées du nombre de personnes à risque au temps  $q_j$  dans le  $l^{\text{ième}}$  groupe et dans l'échantillon total, respectivement.

Comme précédemment, les  $\{q_j\}_{j=0}^m$  sont les éléments uniques et ordonnés de  $\{0, L_i, R_i; i = 1, \dots, n\}$  avec  $q_{m+1}$  le temps au-delà duquel  $\hat{S} = 0$ . Enfin, la pondération est notée  $w_j$ .

*Calcul du nombre de personnes à risques*

Différentes manières de calculer les pseudo-nombres de personnes à risque ont été développées. Par exemple, SUN [2006] propose :

$$d_j^* = \sum_{i=1}^n \delta_i \frac{\alpha_{ij} [\widehat{S}(q_{j-}) - \widehat{S}(q_j)]}{\sum_{u=1}^{m+1} \alpha_{iu} [\widehat{S}(q_{u-}) - \widehat{S}(q_u)]}$$

avec  $\delta_i$ , un indicateur qui vaut 1 sur le sujet est censuré à droite ( $\mathbb{1}_{R_i \leq q_m}$ ) et 0 sinon. Comme précédemment,  $\alpha_{ij} = \mathbb{1}_{\{q_j \in (L_i, R_i)\}}$ .

ZHAO et SUN [2004] en présentent une version améliorée, prenant en compte les données censurées et non censurées.

*Choix de la pondération*

La pondération permet de donner plus de poids aux décès précoces, tardifs, intermédiaires, etc. Comme pour la censure à droite, elle sert à adapter le test à la forme des données. FAY [1999] propose pour la pondération :

$$w_j = \frac{\widehat{S}(q_j) \widehat{S}'(q_{j-1}) - \widehat{S}'(q_j) \widehat{S}(q_{j-1})}{\widehat{S}(q_j) [\widehat{S}(q_{j-1}) - \widehat{S}(q_j)]}$$

où  $\widehat{S}'$  est la dérivée de  $S$  comme définie comme en section B. page 34.

Dans le test de SUN [1996], le poids  $w_j$  vaut 1.

*Calcul de la variance*

Le calcul de la variance n'est pas trivial dans le cas des log-ranks pour censure par intervalle. La différence entre la censure à droite et la censure par intervalle est qu'en cas d'événement, la date d'événement est connue exactement dans le premier cas tandis qu'on ne peut définir qu'un intervalle d'appartenance dans le second cas. Ainsi, avec des données censurées par intervalle, on peut considérer la date d'événement comme "manquante" et chercher alors par remplacer l'information manquante par la date la plus plausible. L'imputation est une approche générique pour faire face au problème de données manquantes. Le principe pour le calcul de la variance est d'imputer des données censurées à droite à partir des données censurées par intervalle et ensuite d'ajuster la variance.

La procédure est la suivante : notons  $N$  le nombre d'imputations réalisées. À chaque imputation  $I$ , une valeur "raisonnable" de date d'événement pour les patients censurés par intervalle est choisie. On construit alors le log-rank classique avec les données converties en données censurées à droite et on estime sa variance  $V^{(I)}$ . À partir des  $N$  variances disponibles, on estime la variance du log-rank pour données censurées.

Cette procédure, détaillée dans ZHAO et SUN [2004], est basée sur les imputations multiples de Rubin. La forme de l'estimateur de la matrice de covariance est une somme de deux parties, qui peuvent être considérées comme des estimations de la variabilité intra-imputation et de la variabilité entre imputations. Une procédure légèrement différente de celle de Zhao et Sun a été proposée par HUANG et al. [2008]; dans ce cas, la seconde partie de la matrice de covariance est soustraite et non sommée comme précédemment.

De manière assez similaire à la méthode d'imputations multiples, SUN [2001] propose une méthode de ré-échantillonnage pour calculer un ensemble de variances.

Une alternative à la méthode d'imputation consiste à utiliser un échantillonnage bootstrap sur les données observées. Bien que simple, cette approche ne se réduit pas à l'estimation de la covariance de la statistique de test du log-rank dans le cas de données censurées à droite.

Une autre alternative est de développer une méthode utilisant des algorithmes MCMC (méthodes de Monte-Carlo par chaînes de Markov). Cependant, en plus de ne pas se réduire à l'estimation de covariance de la statistique de test de log-rank dans le cas de données censurées à droite comme la méthode bootstrap, la méthode MCMC peut être longue et conceptuellement plus compliquée que la méthode d'imputations multiples.

## D. Comparaisons des trois approches

### Le test du score

Le test du score étant basé sur la vraisemblance, il est dépendant du type de vraisemblance considéré. HUANG et WELLNER [1995] proposent l'emploi d'une vraisemblance profilée avec pour variance, sa matrice d'information observée; mais sa validité asymptotique n'est pas prouvée (HUANG et WELLNER [1997]). Un test du score basé sur la vraisemblance marginale pour censure par intervalle a été proposé par SELF et GROSSMAN [1986]. Les difficultés computationnelles liées à cette forme de vraisemblance ont mené au développement du test du score basé sur la vraisemblance groupée dont le nombre de paramètres de nuisance réduits est plus approprié (FAY [1996, 1999]; FINKELSTEIN et WOLFE [1985]; SUN [1996]).

L'emploi de la vraisemblance groupée ne permet cependant pas de supprimer toutes les difficultés. Tout d'abord, le nombre de paramètres à estimer (qui est  $(m - 1)$ ,  $m$  étant la dimension du paramètre de nuisance) peut excéder  $n$ , le nombre d'individus dans l'échantillon. De plus, la vraisemblance groupée implique une contrainte sur le vecteur des paramètres de nuisance,  $\gamma$ , puisque ses composantes doivent être ordonnées. À cause de ces deux particularités, certaines estimations des paramètres de nuisance peuvent être trop proches, voire identiques. Le modèle est "sur-paramétré" (COX et HINKLEY [2000]; FAY [1996]).

Notons que la vitesse de convergence des estimateurs du maximum de vraisemblance de la fonction de répartition est lent (HUANG et WELLNER [1995]) et cela peut affecter les propriétés asymptotiques du test du score. Cependant, ce problème de convergence est commun à tous les tests nécessitant une estimation de la survie.

Pour éviter les problèmes évoqués plus haut, et bien que le test du score basé sur la vraisemblance a pour avantage de pas nécessiter d'indépendance entre censure et variable, il est fréquent d'avoir recours à l'approche permutationnelle pour le calcul de la variance.

### Le test permutationnel

L'avantage principal de cette méthode est sa simplicité d'écriture et d'application. Sa variance a également une forme simple. Cependant, cette méthode nécessite la condition supplémentaire d'indépendance entre censure et variable.

Des simulations effectuées sur des données censurées par intervalle de type 1 ne montrent pas d'inflation majeure du risque de première espèce avec ce type de méthode en cas de censure dépendante de  $Z$  (voir section C. page 83).

Ces résultats sont en accord avec FAY et SHIH [2012] qui ont conclu que les différentes méthodes peuvent être utilisées en présence de censure dépendante de la covariable à l'exception de celle basée sur l'imputation de la moitié de l'intervalle.

### Le test du (pseudo) log-rank

Les tests de log-rank pour censure par intervalle sont basés sur le même principe que les tests de log-rank destinés à la censure à droite. À chaque temps d'événement, le nombre de décès constaté est comparé à celui, théorique, qui aurait été constaté avec une absence de différence entre les deux groupes.

Plusieurs méthodes existent pour calculer les pseudos-nombres de personnes à risque dans le cas de la censure par intervalle. L'avantage de cette méthode est qu'elle possède une interprétation heuristique facilement compréhensible. De plus, la forme de ce test est similaire à celle disponible pour la censure à droite. Enfin, il est possible de choisir la pondération du test, ce qui permet de le faire correspondre à de nombreuses situations.

Pendant, le calcul de la variance dans les log-ranks pour données censurées par intervalle n'est pas trivial. Il nécessite le recours à des algorithmes d'imputations ou de ré-échantillonnage. Il n'existe pas de méthode unique et unanime pour le calcul de la variance. Les auteurs ont proposé différentes méthodes ainsi que nous l'avons vu plus haut.

De plus, certains log-ranks pour données censurées par intervalle ne se réduisent pas au test du log-rank classique en cas de censure droite (ZHAO et SUN [2004]). SUN [1996] préfère d'ailleurs retenir pour son test, originellement présenté comme un test de log-rank, une forme correspondante à un score pour pouvoir utiliser l'information de Fisher.

### Liens entre les différents tests

Un test écrit sous une certaine forme peut, dans certains cas, être réécrit sous une autre forme. Par exemple, SUN [1996] écrit son test sous la forme d'un log-rank puis sous celle d'un score. De manière similaire, FAY [1996] présente un test sous la forme d'un test de rang puis sous la forme d'un score.

Dans sa première annexe, FAY [1999] propose une procédure plus générale explicitant la transformation entre un test de permutation et un test de log-rank pondéré.

FAY et SHAW [2010] présentent pour les modèles de FINKELSTEIN et WOLFE [1985], SUN [1996] et FAY [1996], les formes linéaires et les formes de log-ranks pondérés des tests.

Enfin, OLLER et GÓMEZ [2012] en section 3, montrent l'équivalence entre l'expression du score sous la forme d'un log-rank pondéré et sa forme linéaire permettant l'utilisation d'une variance permutationnelle.

Ainsi, il existe des liens entre les différents tests et il peut être possible de passer de la forme d'un score à une autre. Le choix d'une statistique de test dépendra de la variance que l'on veut utiliser.

## II. Les tests implémentés existants

Suite à la nécessité de développer des tests spécifiques destinés à la censure par intervalle, plusieurs propositions ont été faites et certaines d'entre elles ont été implémentées et sont actuellement disponibles sur le logiciel *R* (R CORE TEAM [2015]).

Trois packages en particulier, proposent des statistiques de tests destinés à la censure par intervalle :

- *interval* par FAY et SHAW [2010],
- *glrt* par ZHAO et SUN [2015],
- *FHtest* par OLLER et LANGOHR [2015].

Nous allons effectuer un récapitulatif des divers tests et préciser lorsque c'est le cas, s'ils sont implémentés sur *R*.

### A. Les tests du score avec vraisemblance groupée

#### 1) Risques proportionnels

Cette statistique est basée sur un score avec une vraisemblance groupée sous un modèle de régression à risques proportionnels. Nous la dénommons **vgRP** dans la suite. Elle a été originellement proposée par FINKELSTEIN et WOLFE [1985] avec une variance basée sur la matrice d'information de Fisher.

Ainsi que l'ont montré FAY et SHAW [2010], cette statistique peut également être écrite sous une forme linéaire permettant l'utilisation d'une variance permutationnelle ou d'une forme de log-rank avec une variance basée sur les imputations multiples.

Ce test est implémenté sur le logiciel *R*; dans le tableau 2.1 sont indiqués les différentes possibilités.

Package	Fonction	Précision
Variance permutationnelle		
interval	ictest	scores = 'logrank2'
FHtest	FHtesticp	—
Variance information de Fisher		
interval	ictest	scores = 'logrank2', method = 'scoretest'
glrt	ScoreTest	—
FHtest	FHtestics	—
Variance par imputations multiples		
interval	ictest	scores = 'logrank2' , method = 'wsr.pclt'

TABLEAU 2.1 – Implémentations du test **vgRP**

La méthode de variance par défaut de la fonction `ictest` du package *interval* est la variance permutationnelle. Il n'est pas utile de préciser `method = 'pclt'` lorsqu'on veut l'utiliser.

## 2) Risques odds-proportionnels

Cette statistique — présentée par FAY [1996] — est assez similaire à la statistique précédente. Elle se base également sur la vraisemblance groupée, mais utilise un modèle de régression odds-proportionnel. Nous la dénommons **vgOP** dans la suite. Ce test a été proposé par PETO et PETO [1972] pour les données censurées à droite.

De même que la statistique précédente, cette statistique peut s'écrire sous la forme d'un score, d'un log-rank pondéré et d'une statistique linéaire.

Lorsque le score est écrit sous la forme d'une statistique linéaire, le test **vgOP** admet pour fonctions de score les  $\omega_i = \hat{S}(l_i) + \hat{S}(r_i) - 1$ . Lorsqu'il est sous la forme d'un log-rank pondéré, les poids du **vgOP** sont  $\hat{S}(q_{j-1})$ . Il apparaît que ce test est particulièrement adapté pour détecter des différences précoces dans la distribution, les premiers événements ayant plus de poids que ceux survenus plus tardivement.

Ce test est implémenté sur le logiciel *R*; dans le tableau 2.2 sont indiqués les différentes possibilités.

## B. Les (pseudo) log-ranks pour censure par intervalle

Sun propose en 1996 une adaptation aux données censurées par intervalle du log-rank classique pour données censurées à droite. D'autres versions de ce test ont par la suite été proposées.

### 1) Log-rank adapté à la censure par intervalle

Cette statistique est assez similaire au score avec vraisemblance groupée et risques proportionnels. Elle a été initialement proposée par SUN [1996] comme une généralisation du test de log-rank classique pour censure par intervalle.

Package	Fonction	Précision
Variance permutationnelle		
interval	ictest	scores = 'wmw'
FHtest	FHtesticp	rho = 1
Variance information de Fisher		
interval	ictest	scores = 'wmw', method = 'scoretest'
FHtest	FHtestics	rho = 1
Variance par imputations multiples		
interval	ictest	scores = 'wmw', method = 'wsr.HLY'

TABLEAU 2.2 – Implémentations du test **vgOP**

Ce test peut également être vu comme un test du score sous le modèle logistique avec des risques discrets. Originellement, il a été présenté pour des données censurées à droite discrètes (ou groupées) par Cox (voir section 6 dans COX [1972]).

Les *odds* des risques discrets des différents groupes à chaque temps  $q_j$  sont proportionnels à  $\exp(\beta Z)$  tels que :

$$\frac{\lambda(q_j|Z)}{1 - \lambda(q_j|Z)} = \exp(\beta Z) \frac{\lambda_0(q_j)}{1 - \lambda_0(q_j)}.$$

où  $\lambda_0(q_j)$  est la fonction de risque de base telle que  $\lambda_0(q_j) = S(q_{j-1}) - S(q_j) / S(q_{j-1})$ . Ce modèle n'est donc pas à risques proportionnels, mais dès lors que les données deviennent continues, les deux modèles sont équivalents.

De même que les tests précédents, cette statistique peut être écrite sous une forme linéaire permettant l'utilisation d'une variance permutationnelle ou d'un score avec l'information de Fisher (c'était d'ailleurs ainsi que SUN [1996] l'avait finalement présentée dans son article) et enfin sous la forme d'un log-rank avec une variance basée sur les imputations multiples.

Ce test, dénommé **LR adapté** par la suite, est implémenté sur le logiciel *R*; dans le tableau 2.3 sont indiqués les différentes possibilités.

Package	Fonction	Précision
Variance permutationnelle		
interval	ictest	scores = 'logrank1'
Variance information de Fisher		
interval	ictest	scores = 'logrank1', method = 'scoretest'
Variance par imputations multiples		
interval	ictest	scores = 'logrank1', method = 'wsrHLY',

TABLEAU 2.3 – Implémentations du test **LR adapté**

## 2) Log-rank adapté à la censure par intervalle, version 2

Cette statistique est un log-rank pour censure par intervalle. Proposée par ZHAO et SUN [2004], ce test ressemble beaucoup au test du log-rank de Sun, le **LR adapté**. Il est dénommé **LR amélioré** par la suite.

Les différences entre le LR adapté et le LR amélioré résident dans la manière de calculer les personnes à risque et donc également dans le calcul de la variance. Le LR amélioré peut être employé

en présence de temps de suivi continus et discrets et il se réduit au test du log-rank classique dans le cas de données censurées à droite contrairement au LR adapté. S’il n’y a pas de données censurées à droite (tous les patients considérés ont fait l’événement d’intérêt), les deux expressions de log-rank sont identiques.

Les auteurs proposent pour le LR amélioré une variance basée sur des imputations multiples de RUBIN [2004]. Cette méthode diffère de la méthode d’imputation proposée par HUANG et al. [2008], la différence entre les deux méthodes est exposée en détails dans l’article HUANG et al. [2008] et plus succinctement en section C, page 35.

Le package *interval* propose — lorsque c’est applicable — la variance basée sur l’imputation multiple telle que présentée par Huang, Lee et Yu, mais le package n’a pas de fonction correspondante aux tests de log-rank améliorés comme proposés par Zhao et Sun ou par Huang, Lee et Yu.

Le test du LR amélioré — version [ZHAO et SUN, 2004] — est implémenté dans le package *glrt*, voir tableau 2.4 :

Package	Fonction	Précision
glrt	gLRT1	—

TABLEAU 2.4 – Implémentation du test **LR amélioré**

## C. Les (pseudo) log-ranks pondérés pour censure par intervalle

### 1) Log-rank pondéré de Sun, Zhao et Zhao

Ce groupe de statistiques — proposé par Sun, Zhao et Zhao (SUN et al. [2005]) — correspond à des log-rank pondérés adaptés à la censure par intervalle. La fonction de lien est inspirée par la famille  $G^{\rho, \gamma}$  introduite par Fleming et Harrington pour des données censurées à droite (voir Définition 7.2.1 dans FLEMING et HARRINGTON [1991]).

Ce test se présente de la manière suivante :

$$U = \sum_{i=1}^n Z_i \frac{\xi\{\widehat{S}(L_i)\} - \xi\{\widehat{S}(R_i)\}}{\widehat{S}(L_i) - \widehat{S}(R_i)},$$

où  $\xi(x)$  est la fonction de lien connue sur  $(0, 1)$  telle que  $\lim_{x \rightarrow 0} \xi(x) = \lim_{x \rightarrow 1} \xi(x)$  soit une constante (e.g.  $c_0$ ).

Le choix de la fonction de lien détermine la statistique. Par exemple, avec  $\xi(x) = x \ln(x)$ , nous obtenons la statistique de test proposée par FINKELSTEIN et WOLFE [1985], le vgRP et la statistique du log-rank de PETO et PETO [1972] s’il n’y a que deux groupes à comparer.

Les auteurs proposent pour la fonction de lien :

$$\xi(x) = (x \ln x) x^{\rho} (1 - x)^{\gamma}.$$

D’après la forme de la fonction de lien, il apparait que  $\rho = 1$  donnera plus de poids aux événements précoces, tandis que  $\gamma = 1$  donnera plus de poids aux événements tardifs. Les valeurs  $\rho = 1, \gamma = 1$  correspondent aux événements intermédiaires. Il faut choisir les valeurs de ces paramètres en fonction de la forme des données à analyser.

#### Variance

Sun, Zhao et Zhao dérivent la matrice de covariance sous l’hypothèse nulle, après avoir identifié la distribution asymptotique normale centrée en 0 de la statistique de score,  $U / \sqrt{n}$ .

Les auteurs utilisent pour cela une décomposition de la statistique de test en censures à gauche, par intervalle et à droite telles que :

$$U_{\xi} = \sum_{i=1}^n Z_i \left[ \delta_i \frac{\xi\{\widehat{F}(L_i)\} - c_0}{\widehat{F}(L_i)} + \gamma_i \frac{\xi\{\widehat{F}(R_i)\} - \xi\{\widehat{F}(L_i)\}}{\widehat{F}(R_i) - \widehat{F}(L_i)} + (1 - \delta_i - \gamma_i) \frac{c_0 - \xi\{\widehat{F}(L_i)\}}{1 - \widehat{F}(R_i)} \right] \quad (2.1)$$

avec  $\delta$  et  $\gamma$ , les indicateurs d'une censure à gauche et par intervalle, respectivement et  $\widehat{F}$ , l'estimateur du maximum de vraisemblance de la fonction de répartition telle que  $F = 1 - S$ .

Les détails sur la distribution asymptotique de  $U/\sqrt{n}$  peuvent être trouvés dans l'annexe de l'article SUN et al. [2005]. Une fois la normalité du score établie, ainsi que sa moyenne nulle, l'expression de sa variance n'est pas compliquée et pourrait être très grossièrement résumée par une matrice de taille  $k \times k$  contenant les  $E \left( \sum \left[ \frac{\xi\{\widehat{S}(L_i)\} - \xi\{\widehat{S}(R_i)\}}{\widehat{S}(L_i) - \widehat{S}(R_i)} \right]^2 \right)$ ,  $k$  étant le nombre de groupes (voir SUN et al. [2005]).

Notons que la forme linéaire du score pourrait également encourager l'utilisation d'une variance permutationnelle; cette possibilité n'a pas été envisagée par les auteurs dans l'article.

Ce log-rank pondéré, dénommé **LRP SZZ** par la suite, est implémenté sur le logiciel *R*, voir tableau 2.5 :

Package	Fonction	Précision
glrt	gLRT2	—

TABLEAU 2.5 – Implémentation du test **LRP SZZ**

## 2) Log-rank pondéré de Oller & Gomez

Ce groupe de statistiques correspond également à des log-ranks pondérés pour données censurées par intervalle. Proposé par OLLER et GÓMEZ [2012] en 2012, ce test correspond à une adaptation des log-rank pondérés basés sur la famille  $G^{\rho, \gamma}$  de FLEMING et HARRINGTON [1991] initialement développés pour la censure droite.

Fleming et Harrington ont proposé la pondération suivante :

$$W(s) = \{\widehat{S}(q-)\}^{\rho} \{1 - \widehat{S}(q-)\}^{\gamma} \quad (2.2)$$

avec  $\rho, \gamma \geq 0$ .

Le test de Oller et Gomez est de la forme suivante :

$$U_l = \int_0^{\infty} w(t; \lambda, \rho) \left[ d_{jl}^* - \frac{n_{jl}^* d_j^*}{n_j^*} \right]$$

et

$$w(t; \lambda, \rho) = \widehat{S}(R_i) \frac{B_{(\lambda+1, \rho)}(1 - \widehat{S}(R_i)) - B_{(\lambda+1, \rho)}(1 - \widehat{S}(L_i))}{\widehat{S}(L_i) - \widehat{S}(R_i)}$$

avec  $\lambda, \rho \geq 0$  et  $B(t)_{a,b} = \int_0^t x^{a-1} (1-x)^{b-1} dx$ , une fonction beta incomplète. Asymptotiquement, lorsque  $d\widehat{S}(t) \rightarrow 0$  et  $\widehat{S}$  est presque continue au point  $t$ , la fonction de poids,  $w$  ressemble aux poids de la famille Fleming et Harrington comme vu dans l'expression (2.2).

Autant pour la censure à droite que pour la censure par intervalle, lorsque  $\rho = \gamma = 0$ , cela correspond au log-rank classique (**vgRP** en censure par intervalle). Lorsque  $\rho = 1, \gamma = 0$ , cela correspond au test de Wilcoxon (**vgOP** en censure par intervalle).

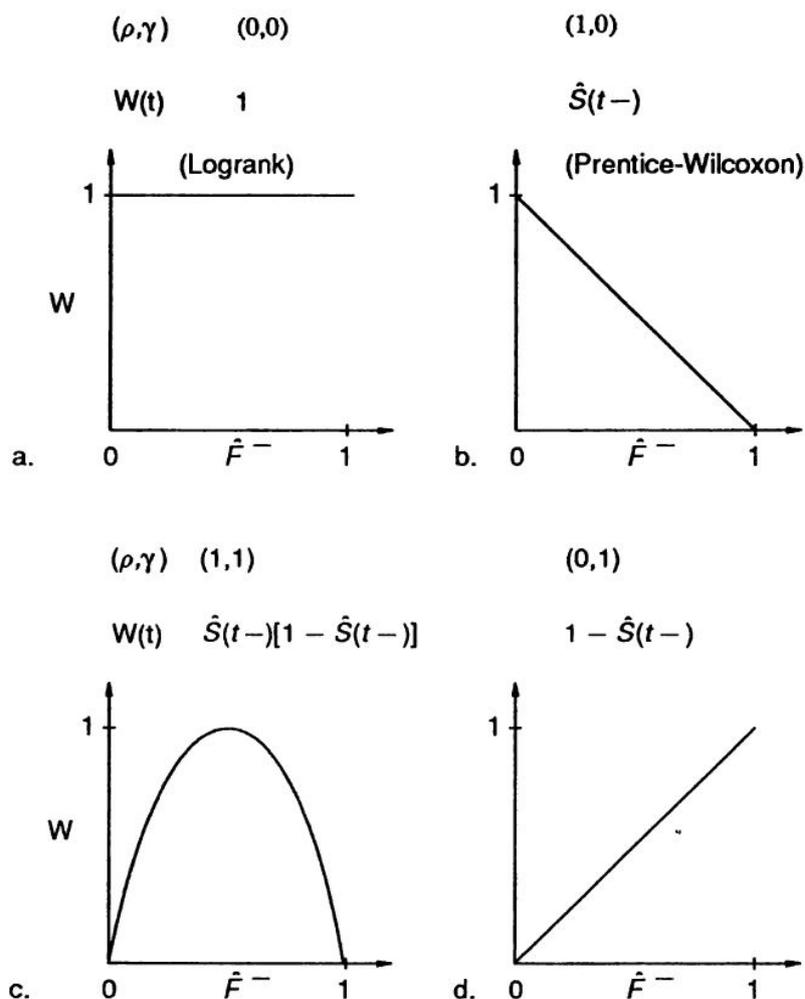


FIGURE 2.1 – Fonctions de poids  $W$  de la famille  $G^{\rho, \gamma}$ .

$W$  correspond à des poids uniformément distribués en (a), des poids accentuant les différences précoces en (b), des poids accentuant les événements intermédiaires en (c) et tardifs en (d)

Crédits : Counting Processes and Survival Analysis, FLEMING et HARRINGTON [1991].

Ces log-ranks pondérés peuvent être également écrits sous une forme linéaire permettant l'emploi d'une variance permutationnelle ou sous la forme d'un score avec l'information de Fisher (l'occurrence  $\lambda > 0$  est problématique dans ce cas).

Ce groupe de log-rank pondérés, dénommé **LRP OG** par la suite, est implémenté sur le logiciel *R*. Le tableau (2.6) en présente les détails :

Package	Fonction	Précision
Variance permutationnelle		
FHtest	FHtesticp	—
Variance information de Fisher		
FHtest	FHtestics	$\triangle$ $\lambda = 0$ obligatoirement

TABEAU 2.6 – Implémentations du test **LRP OG**

### 3) Log-rank pondéré hybride

Ce groupe de statistiques a été proposé en 2013 par ZHAO et al. [2013]. Il correspond à une modification du test du log-rank pondéré de SUN et al. [2005] (test **LRP SZZ**). Il est dénommé **LRPHyb** par la suite.

Le test **LRPHyb** est innovant par rapport aux tests précédents dans le sens où sont établies les distributions asymptotiques du score, à la fois sous l'hypothèse nulle et sous l'hypothèse alternative. Ainsi, il est possible de calculer la puissance du test employé.

Les auteurs reprennent la formulation du log-rank pondéré de Sun, Zhao et Zhao dans laquelle sont séparées les censures à gauche, à droite et par intervalle. Ils introduisent dans l'expression (2.1) de Sun, Zhao et Zhao, les fonctions de répartition du groupe 1 et du groupe 2,  $F_1$  et  $F_2$ . Un premier score consiste en  $U_{F_1, F_2}$  :  $F_1$  étant au dénominateur,  $F_2$  au numérateur. Puis l'inverse, pour le score  $U_{F_2, F_1}$ ,  $F_1$  est au numérateur et  $F_2$  au dénominateur. La statistique de test est finalement un vecteur constitué de ces deux éléments.

Les auteurs démontrent dans ZHAO et al. [2013] que les deux scores  $\bar{U}_{F_1, F_2}$  et  $\bar{U}_{F_2, F_1}$ , convergent vers deux variables normale, indépendantes et centrées en 0,  $U_{F_1, F_2} \sim \mathcal{N}(0, \sigma_1^2)$  et  $\bar{U}_{F_2, F_1} \sim \mathcal{N}(0, \sigma_2^2)$ . Ils construisent alors le test :

$$T = \frac{\bar{U}_{F_1, F_2}^2 / \sigma_1^2}{\bar{U}_{F_2, F_1}^2 / \sigma_2^2},$$

D'après ce qui a été exposé précédemment,  $T$  suit asymptotiquement une loi de Fisher (1, 1) sous l'hypothèse nulle.

L'utilisation de la statistique de test,  $T$ , requiert l'estimation de la fonction de survie dans les deux groupes comparés. Il est possible d'utiliser l'algorithme classique de TURNBULL [1976] à cet effet. De la même manière que dans SUN et al. [2005], il faut également choisir la fonction de lien,  $\xi(x)$ , qui sera utilisée.

Le log-rank pondéré tel que présenté par Zhao et al. est implémenté sur le logiciel *R*, voir tableau 2.7 :

Package	Fonction	Précision
glrt	gLRT4	—

TABLEAU 2.7 – Implémentation du test **LRPHyb**

## D. Les tests basés sur la vraisemblance partielle

### 1) Vraisemblance partielle et modèle additif

*Ce test est uniquement destiné aux données censurées par intervalle de type 1.*

Ce test, dénommé **vpMA** par la suite, a été proposé par Lin, Oakes et Ying dans LIN et al. [1998]. Le développement de cette approche est largement basé sur les processus de comptage et sur les propriétés des martingales.

Le **vpMA** est un test du score basé sur la vraisemblance partielle sous modèle additif pour *current status data*. L'astuce des auteurs est de compter les patients n'ayant pas eu d'événement constaté à la date de visite au lieu, comme classiquement, de prendre en considération ceux ayant fait l'événement d'intérêt.

On se place dans un modèle de régression à risques additifs présenté en section a) page 8 avec  $\Lambda_0(t)$  la fonction de risque cumulée de base associée à la variable  $T$ , le délai avant événement. La

variable associée au temps de visite est  $C$ . La covariable  $Z(t)$  peut dépendre du temps.<sup>2</sup>

Les auteurs utilisent un processus de comptage,  $N_i(t) = \mathbb{1}_{(C_i \leq t)}$  qui vaut 1 au temps de visite  $C_i$  si on ne constate pas d'événement chez le patient. Le saut du processus de comptage ( $dN_i(t) = 1$ ) résulte de deux événements simultanés : le patient fait sa visite au temps  $t$  (1) et le patient n'a pas fait d'événement à cette date (2) soit :

- (1) visite au temps  $t$  :  $C_i = t$  avec un risque associé  $d\Lambda_c(t)$ ;
- (2) pas d'événement au temps  $t$  :  $\mathbb{P}(T_i \geq t | Z_i(s), s \leq t) = e^{-\Lambda_0(t) - \beta Z_i(t)}$ .

Le produits de ces deux événements nous donne le risque de faire un saut :

$$\begin{aligned} dH_i(t) &= e^{-\beta Z_i(t) - \Lambda_0(t)} d\Lambda_c(t) \\ &= e^{-\beta Z_i(t)} dH_0(t) \text{ où } dH_0(t) = e^{-\Lambda_0(t)} d\Lambda_c(t). \end{aligned} \quad (2.3)$$

On reconnaît dans l'expression (2.3) le modèle de régression à risques proportionnels de Cox présenté en section 1) page 7 et avec lequel la vraisemblance partielle est applicable. La spécificité du modèle est que le risque est ici relatif au processus de comptage particulier décrit ci-dessus.

L'expression du score et sa matrice d'information de Fisher par rapport à  $\beta$  sont :

$$\begin{aligned} U_\beta &= \sum_{i=1}^n \int_0^\infty \left\{ Z_i(t) - \frac{S^{(1)}(t; \beta)}{S^{(0)}(t; \beta)} \right\} dN_i(t) \\ I_\beta &= \sum_{i=1}^n \int_0^\infty \left\{ \frac{S^{(2)}(t; \beta)}{S^{(0)}(t; \beta)} - \frac{S^{(1)}(t; \beta)^{\otimes 2}}{S^{(0)}(t; \beta)^2} \right\} dN_i(t) \end{aligned}$$

avec

$$\begin{aligned} S^{(k)}(t; \beta) &= \sum Y_j(t) e^{-\beta Z_j(t)} Z_j(t)^{\otimes k}, \quad (k = 0, 1, 2) \\ a^{\otimes 0} &= 1, \quad a^{\otimes 1} = a, \quad a^{\otimes 2} = aa'. \end{aligned}$$

En se basant sur les propriétés de  $U_\beta$  et  $I_\beta$  liées aux processus de comptage et à la théorie des martingales, Lin, Oakes et Young établissent que  $U_\beta' I_\beta U_\beta$  suit sous l'hypothèse nulle une loi du  $\chi^2$  dont le degré de liberté est celui du vecteur  $\beta$ .

Dans l'article LIN et al. [1998], les auteurs présentent également une version du test adaptée à la censure informative, sachant que le score décrit plus haut y est sensible; des simulations dont les résultats sont présentés dans les sections suivantes confirment la sensibilité du test à la censure informative.

Pour utiliser le test **vpMA**, on peut recourir à des packages destinés à la censure droite, tels que le package *survival*.

## 2) Imputation du milieu de l'intervalle

Une approche relativement simpliste pour traiter la censure par intervalle est d'imputer pour chaque patient ayant fait l'événement une pseudo-date d'événement en choisissant pour cela la moitié ou la fin de l'intervalle considéré.

Les données censurées par intervalle sont ainsi transformées en données censurées à droite classiques avec lesquelles on peut utiliser le test du score avec la vraisemblance partielle, dérivé sous un modèle à risques proportionnels comme proposé par COX [1972]; ou n'importe quel autre test de log-rank destiné à la censure à droite. Par la suite, **vpIMI** représentera le test du score pour données censurées à droite utilisé sur des données censurées par intervalle avec imputation du milieu de l'intervalle.

2. Le modèle additif est particulièrement adapté aux covariables dépendantes du temps.

L'utilisation de cette méthode — qui peut paraître plus facile à l'expérimentateur — n'est pas recommandée. Nous avons mentionné dans l'introduction du chapitre 3 que le test **vpIMI** était sensible au mécanisme de censure non-similaire dans les groupes. Le test peut également se révéler moins puissant selon les situations.

Pour utiliser le test **vpIMI**, on peut recourir à des packages destinés à la censure droite que le package *survival*.

### III. Références

- COX, D. R. 1972, «Regression models and life-tables», *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 34, n° 2, p. 187–220, ISSN 00359246. URL <http://www.jstor.org/stable/2985181>. 40, 45
- COX, D. R. et D. V. HINKLEY. 2000, *Theoretical statistics*, reprint éd., Chapman & Hall/CRC, Boca Raton, ISBN 978-0-412-16160-5 978-0-412-42860-9. OCLC : 247857300. 34, 37
- FAY, M. P. 1996, «Rank invariant tests for interval censored data under the grouped continuous model», *Biometrics*, vol. 52, n° 3, p. 811–822, ISSN 0006-341X. 33, 34, 35, 37, 38, 39
- FAY, M. P. 1999, «Comparing several score tests for interval censored data», *Statistics in Medicine*, vol. 18, n° 3, p. 273–285, ISSN 0277-6715. 33, 34, 35, 36, 37, 38
- FAY, M. P. et P. A. SHAW. 2010, «Exact and Asymptotic Weighted Logrank Tests for Interval Censored Data : The interval R Package», *Journal of Statistical Software*, vol. 36, n° 2, doi :10.18637/jss.v036.i02, ISSN 1548-7660. URL <http://www.jstatsoft.org/v36/i02/>. 38, 39
- FAY, M. P. et J. H. SHIH. 2012, «Weighted logrank tests for interval censored data when assessment times depend on treatment», *Statistics in Medicine*, vol. 31, n° 28, doi :10.1002/sim.5447, p. 3760–3772, ISSN 1097-0258. 37
- FINKELSTEIN, D. M. et R. A. WOLFE. 1985, «A Semiparametric Model for Regression Analysis of Interval-Censored Failure Time Data», *Biometrics*, vol. 41, n° 4, doi :10.2307/2530965, p. 933, ISSN 0006341X. URL <http://www.jstor.org/stable/2530965?origin=crossref>. 33, 34, 37, 38, 41
- FLEMING, T. R. et D. P. HARRINGTON. 1991, *Counting processes and survival analysis*, 2<sup>e</sup> éd., Wiley series in probability and mathematical statistics, Wiley, New York, ISBN 978-0-471-52218-8. 41, 42, 43
- HUANG, J., C. LEE et Q. YU. 2008, «A generalized log-rank test for interval-censored failure time data via multiple imputation», *Statistics in Medicine*, vol. 27, n° 17, doi :10.1002/sim.3211, p. 3217–3226, ISSN 02776715, 10970258. URL <http://doi.wiley.com/10.1002/sim.3211>. 36, 41
- HUANG, J. et J. A. WELLNER. 1995, «Efficient estimation for the proportional hazards model with "case 2" interval censoring», . 37
- HUANG, J. et J. A. WELLNER. 1997, «Interval Censored Survival Data : A Review of Recent Progress», dans *Proceedings of the First Seattle Symposium in Biostatistics*, vol. 123, édité par P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg, I. Olkin, N. Wermuth, S. Zeger, D. Y. Lin et T. R. Fleming, Springer US, New York, NY, ISBN 978-0-387-94992-5 978-1-4684-6316-3, p. 123–169. URL [http://link.springer.com/10.1007/978-1-4684-6316-3\\_8](http://link.springer.com/10.1007/978-1-4684-6316-3_8). 37
- KALBFLEISCH, J. D. et R. L. PRENTICE. 2002, *The statistical analysis of failure time data*, 2<sup>e</sup> éd., Wiley series in probability and statistics, Wiley, Hoboken, NJ, ISBN 978-0-471-36357-6. 35
- LAW, C. G. et R. BROOKMEYER. 1992, «Effects of mid-point imputation on the analysis of doubly censored data», *Statistics in Medicine*, vol. 11, n° 12, p. 1569–1578, ISSN 0277-6715. 32
- LIN, D. Y., D. OAKES et Z. YING. 1998, «Additive hazards regression with current status data», *Biometrika*, vol. 85, n° 2, doi :10.1093/biomet/85.2.289, p. 289–298. URL <http://biomet.oxfordjournals.org/content/85/2/289.abstract>. 44, 45

- OLLER, R. et G. GÓMEZ. 2012, «A generalized Fleming and Harrington's class of tests for interval-censored data», *Canadian Journal of Statistics*, vol. 40, n° 3, doi :10.1002/cjs.11139, p. 501–516, ISSN 03195724. URL <http://doi.wiley.com/10.1002/cjs.11139>. 33, 35, 38, 42
- OLLER, R. et K. LANGOHR. 2015, *FHtest : Tests for Right and Interval-Censored Survival Data Based on the Fleming-Harrington Class*. URL <https://CRAN.R-project.org/package=FHtest>, r package version 1.3. 38
- PAN, W. 2000, «A multiple imputation approach to Cox regression with interval-censored data», *Biometrics*, vol. 56, n° 1, p. 199–203, ISSN 0006-341X. 32
- PANAGEAS, K. S., L. BEN-PORAT, M. N. DICKLER, P. B. CHAPMAN et D. SCHRAG. 2007, «When you look matters : The effect of assessment schedule on progression-free survival», *JNCI : Journal of the National Cancer Institute*, vol. 99, n° 6, doi :10.1093/jnci/djk091, p. 428–432. URL [+http://dx.doi.org/10.1093/jnci/djk091](http://dx.doi.org/10.1093/jnci/djk091). 32
- PETO, R. et J. PETO. 1972, «Asymptotically Efficient Rank Invariant Test Procedures», *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, n° 2, doi :10.2307/2344317, p. 185, ISSN 00359238. URL <http://www.jstor.org/stable/10.2307/2344317?origin=crossref>. 39, 41
- R CORE TEAM. 2015, *R : A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. 38
- RUBIN, D. B. 2004, *Multiple imputation for nonresponse in surveys*, wiley classics library ed éd., Wiley classics library, Wiley-Interscience, Hoboken, NJ, ISBN 978-0-471-65574-9. OCLC : 834610580. 41
- SELF, S. G. et E. A. GROSSMAN. 1986, «Linear Rank Tests for Interval-Censored Data with Application to PCB Levels in Adipose Tissue of Transformer Repair Workers», *Biometrics*, vol. 42, n° 3, doi :10.2307/2531202, p. 521, ISSN 0006341X. URL <http://www.jstor.org/stable/2531202?origin=crossref>. 37
- SEN, P. K. 2004, «Permutational Central Limit Theorems», dans *Encyclopedia of Statistical Sciences*, édité par S. Kotz, C. B. Read, N. Balakrishnan, B. Vidakovic et N. L. Johnson, John Wiley & Sons, Inc., Hoboken, NJ, USA, ISBN 978-0-471-66719-3. URL <http://doi.wiley.com/10.1002/0471667196.ess1951>. 35
- SUN, J. 1996, «A non-parametric test for interval-censored failure time data with application to AIDS studies», *Statistics in Medicine*, vol. 15, n° 13, doi :10.1002/(SICI)1097-0258(19960715)15:13<1387::AID-SIM268>3.0.CO;2-R, p. 1387–1395, ISSN 0277-6715. 33, 34, 35, 36, 37, 38, 39, 40
- SUN, J. 2001, «Nonparametric test for doubly interval-censored failure time data», *Lifetime Data Analysis*, vol. 7, n° 4, p. 363–375, ISSN 1380-7870. 36
- SUN, J. 2006, *The Statistical Analysis of Interval-censored Failure Time Data*, Statistics for Biology and Health, Springer New York, New York, NY, ISBN 978-0-387-32905-5. URL <http://link.springer.com/10.1007/0-387-37119-2>. 35, 36
- SUN, J., Q. ZHAO et X. ZHAO. 2005, «Generalized log-rank tests for interval-censored failure time data», *Scandinavian Journal of Statistics*, vol. 32, n° 1, doi :10.1111/j.1467-9469.2005.00414.x, p. 49–57, ISSN 1467-9469. URL <http://dx.doi.org/10.1111/j.1467-9469.2005.00414.x>. 41, 42, 44
- TURNBULL, B. W. 1976, «The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data», *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 38, n° 3, p. 290–295. URL <http://www.jstor.org/stable/2984980>. 44

ZHAO, Q. et J. SUN. 2004, «Generalized log-rank test for mixed interval-censored failure time data», *Statistics in Medicine*, vol. 23, n° 10, doi :10.1002/sim.1746, p. 1621–1629, ISSN 0277-6715. 35, 36, 38, 40, 41

ZHAO, Q. et J. SUN. 2015, *glrt : Generalized Logrank Tests for Interval-censored Failure Time Data*. URL <https://CRAN.R-project.org/package=glrt>, r package version 2.0. 38

ZHAO, X., R. DUAN, Q. ZHAO et J. SUN. 2013, «A new class of generalized log rank tests for interval-censored failure time data», *Computational Statistics & Data Analysis*, vol. 60, doi : 10.1016/j.csda.2012.11.002, p. 123–131, ISSN 01679473. URL <http://linkinghub.elsevier.com/retrieve/pii/S0167947312003933>. 44



# Chapitre 3

## Cadre applicatif : maladies auto-immunes et anticorps

### Sommaire

---

<b>I. Système immunitaire</b> . . . . .	<b>52</b>
A. Reconnaissance du soi . . . . .	52
B. Les cellules du système immunitaire . . . . .	53
C. Les cytokines . . . . .	55
D. Les anticorps . . . . .	56
E. Production et maturation des lymphocytes . . . . .	57
F. Résumé général . . . . .	59
<b>II. Maladies auto-immunes</b> . . . . .	<b>60</b>
A. Les mécanismes pathogènes . . . . .	60
B. Facteurs de risque . . . . .	61
C. Thérapeutique des maladies auto-immunes . . . . .	62
D. Les anticorps anti-biothérapies . . . . .	63
<b>III. Le projet ABIRISK</b> . . . . .	<b>63</b>
A. Description générale . . . . .	63
B. Données utilisées . . . . .	65
<b>IV. Références</b> . . . . .	<b>67</b>

---

## I. Système immunitaire

Le système immunitaire est constitué d'un ensemble complexe de mécanismes impliquant protéines, cellules et réactions biochimiques, destiné à défendre l'organisme de toute agression d'agents pathogènes comme des infections et des cellules tumorales.

L'immunité comprend donc l'ensemble des réactions visant à éliminer des agents étrangers appartenant au *non-soi* tout en reconnaissant et en tolérant les composants de l'organisme appartenant au *soi*.

- Il existe une réponse innée ou **non spécifique** du système immunitaire. Lors de cette réaction, il y a une reconnaissance des structures communes à plusieurs agents pathogènes. En plus, des barrières physiques comme la peau et les muqueuses, les cellules en charge de l'immunité innée telles que les macrophages détruisent les agents étrangers de manière non spécifiques par phagocytose. La phagocytose consiste en l'ingestion de particules étrangères telles que des bactéries par des cellules spécialisées, les phagocytes (par ex. : macrophages).
- Il existe également une réponse acquise ou **spécifique** qui dépend de la reconnaissance particulière de la substance étrangère. La rencontre avec l'agent pathogène est gardée en mémoire par les lymphocytes, des cellules spécialisées issues des globules blancs. Parmi eux, les lymphocytes B et T y jouent un rôle primordial.

### A. Reconnaissance du soi

Les protéines membranaires, situées sur la membrane cytoplasmique des cellules, participent à la communication intercellulaire. Elles interviennent dans la reconnaissance intercellulaire et dans d'autres fonctions comme la délivrance de signaux à la cellule ou le transport de molécules à travers la membrane.

Parmi les protéines membranaires, se trouvent les antigènes du soi. Un antigène est une macromolécule capable de déclencher une réponse du système immunitaire lorsqu'elle est reconnue spécifiquement par des lymphocytes T ou B. Toutes les cellules nucléées possèdent sur leur membrane des protéines du complexe majeur d'histocompatibilité ou **CMH** qui expriment le soi de manière unique comme ce sera détaillé par la suite.

Deux groupes de gènes codent pour les molécules du CMH : les gènes de classe I et les gènes de classe II. Chez l'homme se retrouve trois gènes de classe I (A,B,C) et trois gènes de classe II (DR, DP, DQ) comme représenté sur la figure 3.1 . Les chaînes  $\alpha$  des molécules de classe I et les chaînes  $\alpha$  et  $\beta$  des molécules de classe II sont situées sur le chromosome 6. La chaîne  $\beta$ , ou  $\beta$ -2 microglobuline, des molécules de classe I est localisée sur le chromosome 15.

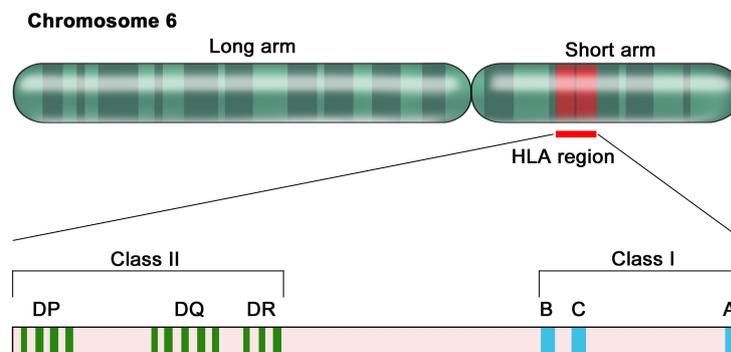


FIGURE 3.1 – Gènes du CMH de classe I et II sur le chromosome 6.

Crédits : Terese Winslow.

Chaque gène codant pour le CMH possède un grand nombre d'allèles. Il est le locus le plus polymorphique connu. Il s'ensuit que le nombre de combinaisons possibles est immense; c'est ce qui rend les molécules du CMH uniques à chaque individu. Le rôle des molécules du CMH dans le système immunitaire est fondamental. Les lymphocytes T (décrits dans la section suivante) ne peuvent reconnaître un antigène qu'après que ce dernier ait été fragmenté et présenté par une molécule CMH.

Les cellules présentatrices d'antigène (CPA) mentionnées dans le tableau 3.1 sont des cellules du système immunitaire (par ex. : lymphocyte B), exprimant des molécules CMH de classe II, qui présentent des antigènes, peptides aux lymphocytes T.

Groupe	Gènes	Présence	Antigène présenté à
Classe I	A, B, C	Toutes les cellules nucléées	lymphocytes T CD8
Classe II	DP, DQ, DR	Cellules présentatrices d'antigène (CPA) professionnelles	lymphocytes T CD4

TABLEAU 3.1 – Molécules du CMH

Toutes les cellules nucléées possèdent des molécules du CMH de classe I. Elles présentent l'antigène aux lymphocytes T CD8 mais ne déclenchent pas son activation. Seules les cellules présentatrices d'antigène (voir section 2) page 54), expriment des molécules CMH de classe II permettant d'activer les lymphocytes T CD4 et des molécules CMH de classe I permettant d'activer les lymphocytes T CD8.

## B. Les cellules du système immunitaire

Le rôle du système immunitaire est de défendre l'organisme contre des agents extérieurs ou pathogènes. Le système lymphoïde est constitué par l'ensemble des organes où résident les lymphocytes, les macrophages et les cellules spécialisées dans la présentation des antigènes Il existe de multiples cellules impliquées dans le système immunitaire : des lymphocytes, des monocytes et des granulocytes.

### 1) Les lymphocytes

Les lymphocytes, qui représentent un tiers des globules blancs du sang, sont divisés en deux types : les lymphocytes B et T. Chacun d'entre eux possède des protéines membranaires différentes. Elles participent à l'immunité acquise. L'immunocompétence du lymphocyte dépend de la synthèse d'un récepteur membranaire capable de reconnaître un antigène spécifique. Le répertoire immunologique d'un organisme est ainsi constitué de l'ensemble des récepteurs différents portés par les lymphocytes.

#### a) Lymphocyte B

Le nom de ces lymphocytes fait référence à l'organe où se fait leur maturation : la bourse de Fabricius des oiseaux, ce qui correspond chez les humains à la moelle osseuse.

Lors de la détection d'un agent étranger, les lymphocytes B sont stimulés. Suite à leur activation par un lymphocyte T, elles se multiplient (elles forment des clones) et produisent des anticorps qui serviront à la reconnaissance spécifique et à la destruction de l'agent pathogène<sup>1</sup>. Un anticorps est une protéine qui reconnaît spécifiquement un antigène. Chaque lymphocyte B ne peut reconnaître qu'un seul motif d'antigène et ne synthétise qu'une seule variété d'anticorps. Les

1. Il existe des lymphocytes B, T-indépendants, moins courants. Les anticorps étudiés dans la thèse sont exclusivement produits par des cellules B T-dépendantes.

lymphocytes B ont également un rôle de cellules présentatrices d'antigène, car elles possèdent des molécules CMH de classe II.

De manière générale, avant la rencontre initiale du lymphocyte B avec un antigène, le lymphocyte B possède à sa surface un récepteur constitué principalement d'un anticorps membranaire (aussi appelé BCR, *B-cell receptor*). Lorsqu'un antigène se lie à son site de réception, le lymphocyte B est activé. Le complexe antigène-récepteur est internalisé et l'antigène découpé en peptides. Le lymphocyte B exprime alors à sa surface des molécules CMH de classe II qui seront présentées à un lymphocyte T CD4. Ce dernier, activé, déclenchera par divers moyens (interactions membranaires ou sécrétion de cytokines qui sont des hormones du système immunitaire) la production d'anticorps par le lymphocyte B.

Suite à son activation par sa rencontre avec l'antigène et son signal de prolifération par le lymphocyte T, la cellule B se divise, se multiplie et se spécialise en deux types cellulaires :

- les plasmocytes, qui sécrètent des anticorps destinés à détruire les antigènes afin qu'ils deviennent des proies plus évidentes pour les phagocytes. Les phagocytes sont des cellules pouvant ingérer et détruire des particules étrangères telles que des microbes ou des cellules altérées;
- les cellules B à mémoire, formées de manière spécifique aux antigènes rencontrés. Leur longue durée de vie permet à ces cellules d'apporter une réponse rapide en cas de deuxième exposition à cet antigène spécifique.

## b) Lymphocyte T

Comme l'indique leur nom, les lymphocytes T sont produits dans le thymus. Il existe deux populations principales de lymphocytes T, possédant des protéines membranaires spécifiques.

- **CD8** : les cellules T cytotoxiques peuvent éliminer des cellules cibles transportant des antigènes telles que des virus ou des cellules cancéreuses en libérant une protéine appelée perforine.  
Elles reconnaissent l'antigène présenté par une molécule CMH de classe I. Bien que les CD8 reconnaissent les antigènes présentés par les molécules CMH de classe I, les CD8 ne sont activés que si cette présentation est effectuée par une cellule présentatrice d'antigène professionnelle (voir 2) page suivante). Les CD8 ne peuvent exercer leur activité cytotoxique sur des cellules cibles qu'une fois activées et différenciées.
- **CD4** : les cellules T facilitatrices signalent aux lymphocytes B de produire plus d'anticorps, aux macrophages et aux CD8 d'être actives pour phagocyter et tuer (soit, induire une apoptose). Elles reconnaissent l'antigène présenté par une molécule CMH de classe II sur une cellule présentatrice d'antigène.

Ainsi, il existe deux sortes d'immunité : l'immunité humorale comprenant les anticorps produits par les lymphocytes B et l'immunité cellulaire où les lymphocytes T agissent contre des cellules cibles directement en tuant les cellules infectées ou indirectement, en activant d'autres lymphocytes par exemple.

## 2) Le rôle des cellules présentatrices d'antigène

Les cellules présentatrices d'antigène (CPA) peuvent internaliser les antigènes afin d'en présenter des fragments aux lymphocytes T avec le complexe majeur d'histocompatibilité. Il existe deux catégories de CPA : les professionnelles et les non-professionnelles.

### a) CPA professionnelles

Elles internalisent les antigènes soit par phagocytose soit par endocytose<sup>2</sup> par récepteurs interposés, digèrent l'antigène en peptides puis présentent une partie de l'antigène, lié aux molécules CMH de classe II, sur leur membrane. Les lymphocytes T CD4 reconnaissent et réagissent avec le complexe moléculaire CMH de classe II présentant l'antigène. Les CPA professionnelles comme toutes les cellules nucléées possèdent également des molécules CMH de classe I; en conséquence, elles peuvent activer les lymphocytes T CD8. Il existe également des signaux stimulateurs supplémentaires produits par la CPA. Ce sont ces signaux qui définissent le caractère professionnel de ces CPA. Il existe trois types de CPA professionnelles : les cellules dendritiques qui possèdent la présentation d'antigène la plus large et qui sont de fait, les plus importantes CPA professionnelles, les macrophages et les lymphocytes B capables d'internaliser l'antigène.

### b) CPA non-professionnelles

Elles comprennent les fibroblastes (peau), les cellules épithéliales thymiques, les cellules épithéliales thyroïdiennes, les cellules gliales (cerveau), les cellules bêta pancréatiques, les lymphocytes T activés et les cellules endothéliales vasculaires.

Elles n'expriment pas de molécules CMH de classe II, nécessaires pour toutes interactions avec des cellules T CD4 naïves, tant qu'elles n'ont pas été stimulées par certaines cytokines comme les interféron- $\gamma$ . Les cytokines sont des protéines qui peuvent activer les cellules immunitaires et affecter des cellules hors du système immunitaire provoquant ainsi inflammations et lésions. N'étant pas accompagnées de signaux co-stimulateurs, les CPA non-professionnelles ne déclenchent pas d'activation des lymphocytes T CD4.

## C. Les cytokines

Les cytokines interviennent dans le dialogue entre lymphocytes, macrophages et d'autres cellules intervenant dans la réponse immunitaire. Ce sont les hormones du système immunitaire. Elles peuvent agir sur les cellules qui les ont produites, sur d'autres cellules ou même agir à distance sur des tissus ou organes.

Les cytokines sont majoritairement produites par les lymphocytes T CD4 mais peuvent également être produites par d'autres cellules comme les macrophages ou les CPA. Il existe une centaine de cytokines pouvant être regroupées par classe. Les principales cytokines connues sont les facteurs de nécrose des tumeurs (TNF- $\alpha$  et  $\beta$ ), les interférons (IFN- $\alpha$ ,  $\beta$ ,  $\gamma$ ) et les interleukines (IL-1 à IL-34).

Parmi les cytokines, les chimiokines sont de petites molécules qui attirent différentes cellules immunitaires et peuvent contribuer à l'invasion et à l'inflammation d'un organe cible. Par exemple, la surproduction de chimiokines dans les articulations de personnes atteintes de polyarthrite rhumatoïde peut entraîner une invasion de l'espace articulaire par des cellules destructrices du système immunitaire telles que les macrophages et les lymphocytes T.

- **TNF** : molécules naturellement présentes dans l'organisme provoquant l'apoptose des cellules anormales de type cancéreuse. La TNF- $\alpha$  est la plus importante des cytokines pro-inflammatoires, elle agit lors d'une infection déclenchant fièvre et inflammation. Elle est également impliquée dans la réponse immunitaire en cas de choc septique. Un choc septique est une insuffisance circulatoire aiguë secondaire à une infection bactérienne; c'est la forme la plus grave de réponse inflammatoire. Cependant, dans le cas de maladies inflammatoires chroniques (maladies inflammatoire chronique de l'intestin, polyarthrite rhumatoïde...), une synthèse trop abondante de cette molécule cause une inflammation des tissus (articulations, intestin...) voire leur altération.

---

2. endocytose : mécanisme de transport permettant à de grosses molécules de pénétrer dans la cellule.

- **IFN** : les interférons sont des cytokines protéiques ayant des fonctions antivirales. L'IFN- $\gamma$  active les macrophages et stimule la reconnaissance des antigènes par les lymphocytes T CD8 en augmentant l'expression des molécules du complexe majeur d'histocompatibilité.
- **Interleukines** : ce sont des glycoprotéines impliquées dans la signalisation de nombreux types de fonctions du système immunitaire. Il existe actuellement 36 familles différentes d'interleukines. Les plus importantes comprennent les médiateurs inflammatoires tels que l'IL-1, l'IL-4 et l'IL-6, l'anti-inflammatoire puissant, l'IL-10, et d'autres interleukines impliquées dans la présentation de l'antigène aux cellules T et B. L'IL-17 est un puissant pro-inflammatoire.

Le tableau 3.2 présente quelques exemples de cytokines en fonction de leur rôle.

Prolifération, différenciation des lymphocytes T	Activation, prolifération et différenciation des cellules B	Activation des macrophages	Activités cytotoxiques
IL-1, IL-2, IL-4, IL-6, IL-7	IL-1, IL-2, IL-4, IL-6, IL-7, IL-13, IFN $\gamma$	IFN $\gamma$ , IL-3, IL-8	TNF $\alpha$ , TNF $\beta$ , IFN $\gamma$ , IL-12

TABLEAU 3.2 – Exemples de cytokines et leurs actions

#### D. Les anticorps

Les anticorps sont des immunoglobulines (Ig), c'est-à-dire des molécules biologiques ayant un rôle fondamental dans la défense de l'organisme contre les agressions. Ce sont des protéines sécrétées par les lymphocytes B, capables de se lier spécifiquement avec des antigènes détectés par les lymphocytes B.

Selon leur activité biologique, ils peuvent être classifiés en l'une des cinq classes d'immunoglobulines. Une immunoglobuline est un anticorps de spécificité inconnue tandis qu'un anticorps est une immunoglobuline de spécificité connue<sup>3</sup>.

Les anticorps représentés schématiquement dans la figure 3.2 ont une structure en 'Y'. Les deux branches du Y contiennent à leur extrémité les sites de fixation à l'antigène. Le pied du Y — le fragment Fc (cristallisable) — porte la spécificité de classe de l'anticorps. Le tableau 3.3 renseigne les cinq classes d'anticorps; chacune d'elle ayant une fonction effectrice (qui répond à une stimulation) spécifique.

Les anticorps ont pour rôle de se fixer à l'antigène, d'activer le système du complément et de recruter des cellules immunocompétentes. La partie variable des anticorps leur permet de reconnaître et de se lier de façon spécifique à un antigène. Cette fixation des antigènes les empêche de se lier aux récepteurs cellulaires. Ainsi, les anticorps bloquent l'action des virus et des bactéries en se liant à leurs molécules de fixation.

L'activation du complément — ensemble de protéines de l'immunité innée — par les IgM principalement et les IgG mène à la destruction des bactéries par perforation, facilite la phagocytose et permet la libération de molécules chimiotactiques.

Enfin l'activation de cellules immunocompétentes se fait grâce à la partie fixe de l'anticorps. Cette action peut déclencher une phagocytose si l'anticorps s'est fixé sur un macrophage grâce à la partie constante de l'anticorps reconnue par le macrophage.

3. Immunoglobulines et anticorps indiquent la même classe de protéines, mais on utilise l'appellation *anticorps* lorsqu'on veut faire référence à l'antigène reconnu spécifiquement par l'anticorps, par exemple : un anticorps anti-biothérapie (qu'on évoquera plus tard); donc on se restreint aux immunoglobulines qui ont cette spécificité. On parle d'*immunoglobulines* lorsqu'on fait référence à la classe de protéines dont les anticorps font partie, qui ont certaines caractéristiques structurales, sans s'intéresser à l'antigène reconnu.

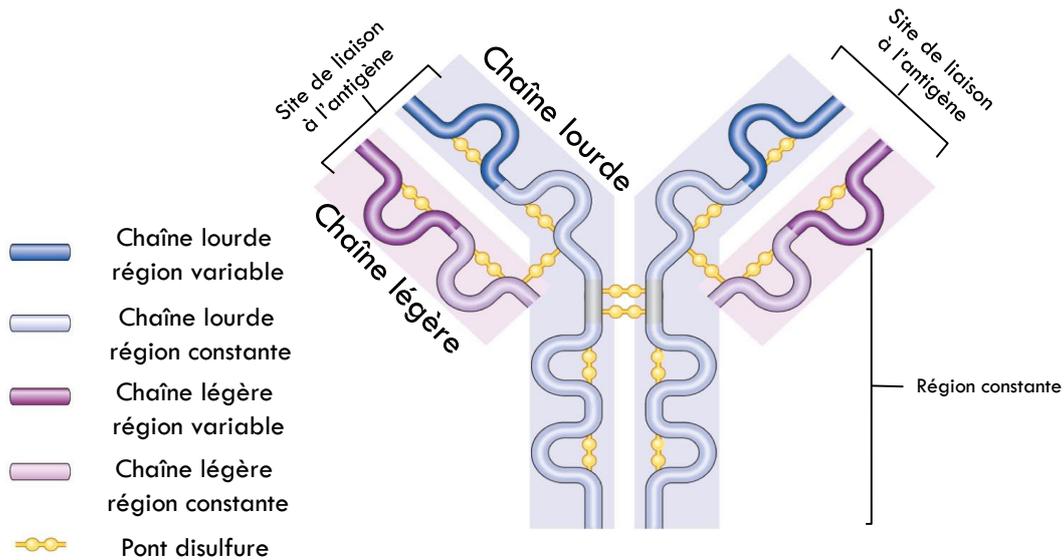


FIGURE 3.2 – Structure de l'anticorps

Classe	Localisation	Proportion	Rôles	Structure
IgG	Sang	70-75%	Neutralisation des toxines, bactéries et virus	
IgA	Muqueuses, sécrétions	15-20%	Agglutination, neutralisation des bactéries, virus	
IgM	Lymphocyte B, sang	10%	Agglutination, voie classique du complément	
IgE	Basophiles, mastocytes	moins de 1%	Allergies, infections parasitaires	
IgD	Lymphocyte B	moins de 1%	Activation du lymphocyte B	

TABLEAU 3.3 – Propriétés des différentes classes d'anticorps

Ainsi la formation complexe antigène-anticorps ne mène pas à la destruction de l'antigène, mais il prépare ce dernier à une destruction par d'autres agents du système immunitaire.

### E. Production et maturation des lymphocytes

La figure 3.3 résume l'origine des lymphocytes et leur évolution.

Il existe ainsi une différence entre le soi d'un individu et le non-soi. La majorité des cellules de l'organisme possèdent, sur leur membrane, des récepteurs du complexe majeur d'histocompatibilité qui est un système de reconnaissance du soi. En cas d'absence de ces récepteurs et si

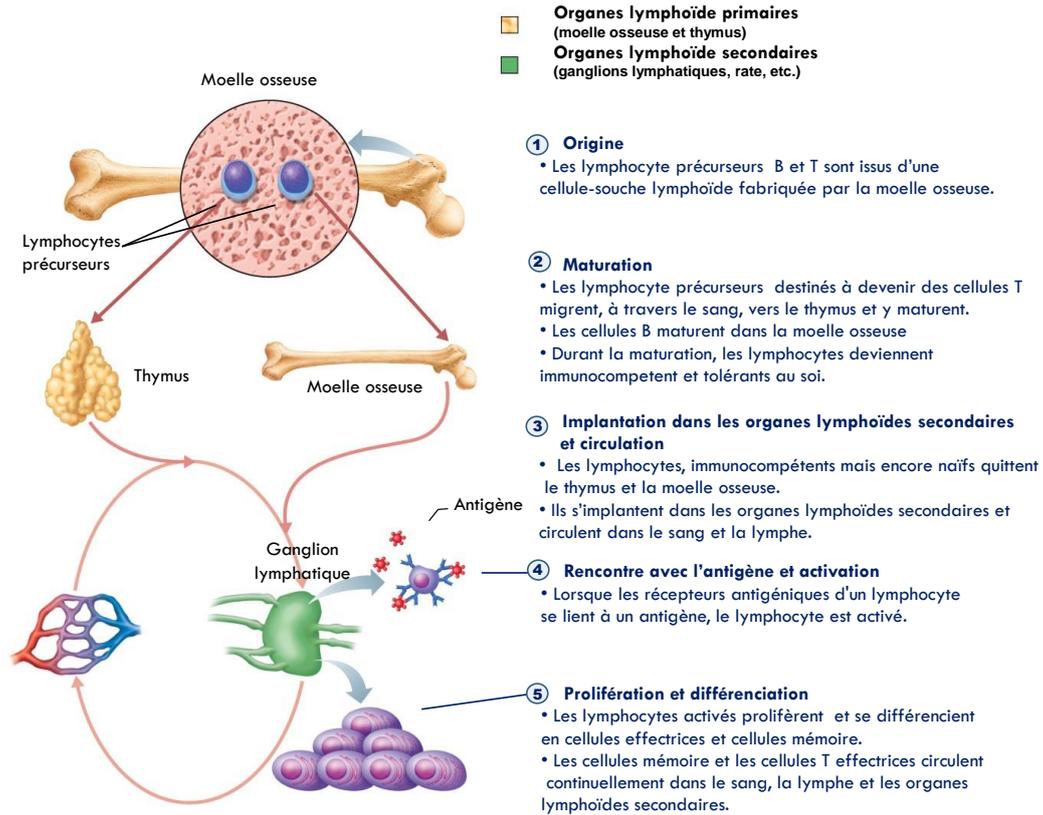


FIGURE 3.3 – Origine et maturation des lymphocytes B et T

la molécule est inconnue, les cellules présentatrices d'antigènes détectent l'élément comme du non-soi et alertent les cellules en charge de la réponse immunitaire.

Chez un sujet normal, des auto-anticorps, non pathogènes, sont présents à taux faible. En effet, lors de la maturation du système immunitaire, les cellules immunitaires réagissant aux éléments du soi sont éliminées, ceci, afin d'aboutir à un système immunitaire tolérant au soi. Le dérèglement de ce mécanisme, dû à une combinaison de prédisposition génétique et de facteurs environnementaux, contribue au développement de maladies auto-immunes.

L'*auto-immunité* est définie comme la réactivité du système immunitaire sous la forme d'auto-anticorps et d'actions de cellules lymphocytaires T à des éléments de soi.

Une *maladie auto-immune* traduit le fonctionnement anormal du système immunitaire résultant en la fabrication d'auto-anticorps attaquant leur propre organisme.

L'*immuno-tolérance* est une absence de réaction du système immunitaire à un antigène particulier, qui en d'autres circonstances déclencherait une réponse immunitaire.

F. Résumé général

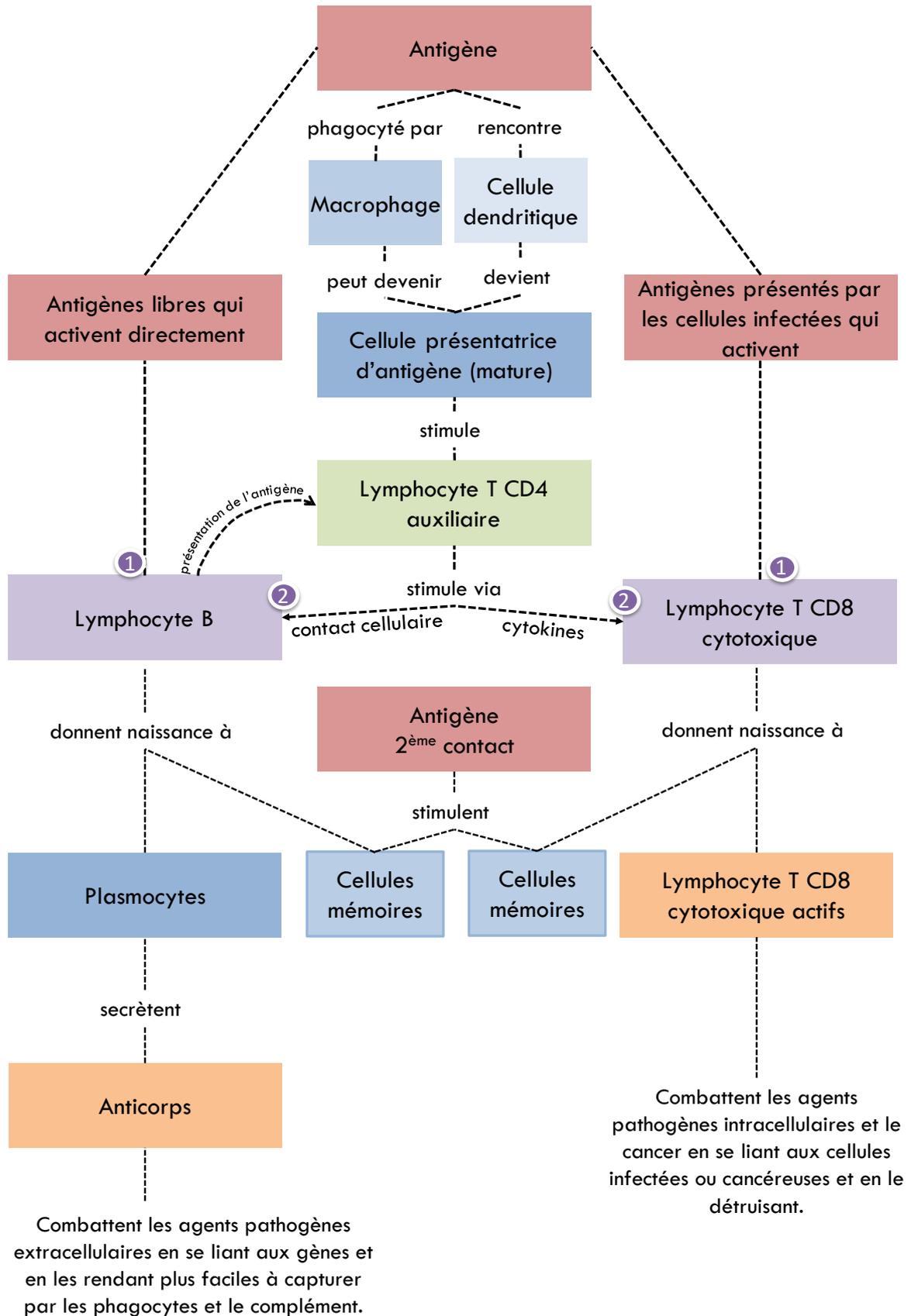


FIGURE 3.4 – Réaction du système immunitaire face à l'antigène

## II. Maladies auto-immunes

Les pathologies auto-immunes — qui surviennent lors de l'intolérance au soi du système immunitaire — peuvent concerner un organe spécifique ou être systémiques.

Lorsqu'un organe spécifique est concerné, le système immunitaire attaque les auto-antigènes de ce dernier comme la myéline (système nerveux central) dans le cas de la sclérose en plaques. Lorsque l'affection est systémique, l'attaque des auto-antigènes a lieu dans toutes les cellules de l'organisme comme dans la polyarthrite rhumatoïde où l'inflammation touche plusieurs articulations.

Lors de la maturation des lymphocytes dans le thymus et la moelle osseuse, les lymphocytes fortement auto-réactifs sont éliminés. Cependant, cette élimination n'est pas totale. Des mécanismes variés de tolérance périphérique assurent le contrôle des lymphocytes auto-réactifs n'ayant pas été éliminés préalablement. Ils comprennent la production de cytokines anti-inflammatoires, l'activation de lymphocytes B et T régulateurs, l'indifférence des lymphocytes auto-réactifs aux auto-antigènes (anergie). Il y a également des régions auxquelles les lymphocytes n'ont pas accès (par ex. : cristallin de l'œil) contenant des auto-antigènes. La défaillance de l'un ou l'autre de ces mécanismes mène à la prolifération des lymphocytes auto-réactifs et à l'apparition d'une maladie auto-immune.

### A. Les mécanismes pathogènes

Les lésions causées par le déclenchement d'une maladie auto-immune peuvent résulter d'un ou plusieurs mécanismes pathologiques.

#### *Mécanismes liés aux lymphocytes T auto-réactifs*

L'action des lymphocytes T auto-réactifs peut causer des lésions cellulaires par différents mécanismes de toxicité (libération de molécules cytotoxiques, déclenchement de l'apoptose de la cellule cibles, etc.). Pour mieux comprendre le mécanisme associé à la pathologie, il faut identifier l'auto-antigène ciblé par le système immunitaire.

#### *Mécanismes liés aux auto-anticorps*

Les auto-anticorps produits par les lymphocytes B peuvent avoir une action pathogène : selon la pathologie, il y a une cyto-toxicité dirigée contre les globules rouges (anémies hémolytiques auto-immunes), une destruction de tissu par accumulation de complexes auto-anticorps-auto-antigène (néphropathies glomérulaires du lupus) ou des interférences des auto-anticorps avec des récepteurs ou des structures cellulaires (myasthénie, hyperthyroïdie). L'identification de l'auto-anticorps est utile pour le diagnostic de la maladie. Par exemple, les anticorps anti-IgG sont des facteurs rhumatoïdes spécifiques à la polyarthrite rhumatoïde.

#### *Mécanismes liés aux cytokines*

La production anormale de cytokines favorise l'inflammation et un dérèglement de l'immunité. Par exemple, la sécrétion de cytokines pro-inflammatoires (interleukine-1 ou facteur TNF- $\alpha$ ) dans le cas de la polyarthrite rhumatoïde.

Le tableau 3.4 présente trois exemples de maladies auto-immunes avec leurs symptômes : les maladies inflammatoires chroniques de l'intestin, la polyarthrite rhumatoïde et la sclérose en plaques.

De nombreuses maladies auto-immunes commencent à un âge relativement jeune pour, la majorité d'entre elles, évoluer de manière chronique tout au long de la vie, avec des phases de poussées et de rémissions. Ces maladies nécessitent des soins longs. Elles concernent près de 10% de la population mondiale. Le tableau 3.5 rassemble quelques données sur les maladies auto-immunes présentées précédemment. Notons que la prévalence des maladies auto-immunes est nettement plus élevée chez les femmes.

Pathologies	Symptômes
<b>Maladies inflammatoire chronique de l'intestin :</b> inflammation chronique de l'intestin dû à une hyperactivité du système immunitaire digestif. Les deux principales formes en sont la maladie de Crohn et la recto-colite hémorragique.	<ul style="list-style-type: none"> <li>— Douleurs abdominales, diarrhées;</li> <li>— Fatigue, fièvre;</li> <li>— Saignements rectaux, amaigrissement (symptôme mineur).</li> </ul>
<b>Polyarthrite rhumatoïde :</b> attaque par le système immunitaire de plusieurs articulations à travers le corps; les mains et les pieds étant le plus souvent touchés.	<ul style="list-style-type: none"> <li>— Articulations douloureuses, raides, gonflées et déformées;</li> <li>— Fonctionnalité et mouvements réduits;</li> <li>— Fièvre, fatigue, perte de poids, inflammation oculaire.</li> </ul>
<b>Sclérose en plaques :</b> pathologie dans laquelle le système immunitaire attaque la myéline, la gaine protectrice de l'axone dans le système nerveux central;	<ul style="list-style-type: none"> <li>— Faiblesse et trouble de l'équilibre, du langage et de la motricité;</li> <li>— Paralysie;</li> <li>— Tremblements;</li> <li>— Engourdissement et sensation de picotements dans les extrémités.</li> </ul>

TABLEAU 3.4 – Exemples de maladies auto-immunes et leurs symptômes.

Pathologies	Prévalence	Incidence (/an)
Maladies inflammatoires chroniques de l'intestin [LOFTUS, 2004]	250 000	6 000
Polyarthrite rhumatoïde [GUILLEMIN et al., 2005; MINICHIELLO et al., 2016]	200 000 à 250 000	6 000 à 8 000
Sclérose en plaques [ARSEP, 2017]	100 000 à 200 000	4 000 à 6 000

TABLEAU 3.5 – Quelques chiffres sur les pathologies auto-immunes en France.

## B. Facteurs de risque

Les maladies auto-immunes sont nombreuses, variées et difficiles à diagnostiquer. De plus, elles sont souvent multifactorielles. Leur développement dépend d'une combinaison de facteurs génétiques et environnementaux.

### *Facteurs de risque génétiques*

La majorité d'entre elles sont polygéniques. Il existe ainsi une multitude de gènes qui lorsqu'ils s'expriment ensemble vont produire un phénotype anormal. Certains de ces gènes sont fortement liés au développement de maladies auto-immunes.

Parmi ceux-ci, il existe une classe particulière de gènes appelés HLA ou antigène des lymphocytes humains, responsables de l'identité biologique de l'organisme, qui sont le meilleur prédicteur de développement de maladies auto-immunes. L'haplotype HLA est un ensemble de gènes codant pour le complexe majeur d'histocompatibilité.

### *Facteurs de risque environnementaux*

Les facteurs environnementaux ont un rôle essentiel dans le développement des maladies auto-immunes. Ils contribuent non seulement à leur survenue, mais également au type de pathologie, à la nature des symptômes initiaux ainsi qu'au moment de survenue.

(1) *Hormones* : la majorité des maladies auto-immunes touchent plus de femmes que d'hommes. Les estimations indiquent que 80% des malades sont des femmes. Les hormones peuvent à la fois être produites par l'organisme (production de stéroïdes) et provenir de sources externes comme des médicaments (pilule contraceptive), des régimes alimentaires ou des produits cosmétiques. Les hormones sexuelles interagissent directement avec les cellules du système immunitaire via des récepteurs positionnés à leur surface. Les hormones stéroïdiennes sont quant à elles, connues pour leur influence sur la production d'anticorps et de cellules du système immunitaire.

(2) *Autres* : il existe d'autres facteurs de risque comme les infections, le tabac, l'exposition aux métaux toxiques ou des épisodes de stress majeurs [DOOLEY et HOGAN, 2003].

## **C. Thérapeutique des maladies auto-immunes**

L'objectif des traitements actuels pour les maladies auto-immunes est de minimiser les mécanismes inflammatoires spécifiques liés à la maladie. Il n'existe pas de traitement permettant de combattre la cause de la maladie, souvent inconnue. Ainsi, les traitements pour les pathologies auto-immunes doivent être pris sur le long terme et ne permettent pas de guérison totale.

Il existe des médicaments immunosuppresseurs agissant sur l'ensemble du système immunitaire. En limitant l'activation du système immunitaire, ces traitements augmentent le risque d'infection et nécessitent un suivi régulier. Parmi les immunosuppresseurs les plus courants, il y a des hormones comme les corticoïdes associées à d'autres molécules (cyclophosphamide, méthotrexate, azathioprine).

Des biomédicaments — produits à partir d'une cellule ou d'un organisme vivant et non par synthèse chimique — peuvent aussi être utilisés dans certaines maladies auto-immunes. Ce sont des anticorps monoclonaux qui ciblent des substances impliquées dans le processus pathologique, par exemple le TNF- $\alpha$  dans la polyarthrite rhumatoïde. Un anticorps monoclonal est un anticorps produit de façon industrielle soit par des bactéries transformées, soit par des lignées cellulaires des mammifères.

Les biomédicaments font actuellement l'objet de nombreuses recherches. Plusieurs dizaines d'entre eux sont en cours de développement à travers le monde. Il y a des molécules dirigées contre des cytokines pro-inflammatoires (par ex. : l'interleukine 17 active dans la polyarthrite rhumatoïde) telle que l'anti-IL-17 (secukinumab, brodalumab) ou les anti-TNF utilisés dans les maladies inflammatoires du colon ou dans la polyarthrite rhumatoïde (infliximab, adalimumab). D'autres agissent contre les lymphocytes B (rituximab). Certains ciblent les facteurs stimulants nécessaires à l'activation des lymphocytes T auto-réactifs (abatacept) ou bloquent la migration des lymphocytes vers les organes ciblés par la maladie (natalizumab).

Le traitement historique de la sclérose en plaques, et qui est toujours utilisé en traitement de première intention, consiste en la prise de cytokines : les interférons- $\beta$  (INF $\beta$ ). Ce sont des immunomodulateurs injectés par voie intramusculaire ou sous cutanée, selon le produit choisi. Le tableau 3.6 récapitule les interférons et leur mode d'injection, donnés dans le cadre de la sclérose en plaques.

Les interférons sont présents naturellement dans l'organisme, ils ont une activité antivirale et anti-inflammatoire. L'avantage des INF $\beta$  est qu'ils provoquent peu d'effets secondaires. Les autres traitements comme le natalizumab sont utilisés en seconde intention en cas d'échec du traitement par interférons.

Produits	Type d'interféron	Voie d'injection
Avonex	IFN $\beta$ -1a	Intra-musculaire
Rebif	IFN $\beta$ -1a	Sous-cutanée
Betaferon	IFN $\beta$ -1b	Sous-cutanée
Extavia	IFN $\beta$ -1b	Sous-cutanée

TABEAU 3.6 – Produits et voie d'injection pour la sclérose en plaques.  
 INF $\beta$  : interférons- $\beta$ .

Concernant les maladies inflammatoires chroniques de l'intestin, plusieurs traitements sont utilisés selon le stade de la pathologie. Les 5-aminosalicylés n'étant pas efficaces dans les deux maladies inflammatoires chroniques de l'intestin et les corticostéroïdes entraînant des effets secondaires, les patients dont la maladie est évolutive sont majoritairement traités par immunomodulateur. Les plus utilisés sont les biothérapies : des anti-TNF- $\alpha$ , dirigés contre le TNF- $\alpha$ , une cytokine pro-inflammatoire (exemple : infliximab, adalimumab).

#### D. Les anticorps anti-biothérapies

Les biothérapies peuvent avoir une structure complexe avec une variabilité structurale découlant de leur mode de production utilisant une technologie génétique recombinante dans les lignées cellulaires. De fait, ces traitements peuvent être reconnus comme étrangers par le système immunitaire hôte et induire une immunogénicité. Cette réaction se traduit par la production d'anticorps anti-médicament ou *anti-drug antibodies* (ADA) pouvant mener à l'inactivation des effets thérapeutiques du traitement et, dans de rares cas, causer des effets indésirables. L'immunogénicité des produits biopharmaceutiques est complexe et influencée à la fois par :

- des facteurs structurels et pharmacologiques : immunogénicité plus élevée avec des anticorps chimère humains-murins et plus faible avec les protéines de fusion...;
- le traitement : voie d'administration, dose, programme de traitement;
- des conditions liées au patient telles que la maladie traitée (par ex. : ADA plus élevé dans le polyarthrite rhumatoïde par rapport au psoriasis), les traitements antérieurs et concomitants et la réactivité immunitaire individuelle (par exemple le polymorphisme d'IL-10).

La prédiction du potentiel immunogène des nouvelles protéines thérapeutiques est donc un défi dans le développement de biothérapies.

### III. Le projet ABIRISK

#### A. Description générale

Dans le cadre des développements accrus des biothérapies, le projet européen ABIRISK a pour objectif d'analyser les mécanismes et les conséquences de l'immunogénicité induite par ces nouveaux traitements. Ce consortium public-privé rassemble des laboratoires (Gsk, Novartis, Pfizer, Sanofi, etc.), des instituts de recherche publics (Inserm, CEA, Institut de recherche en biomédecine Suisse, etc.), des universités (Medizinische Universität Innsbruck, Queen Mary and Westfield, University of London, etc.) et des hôpitaux (Assistance Publique-Hôpitaux de Paris, etc.). Le sigle ABIRISK signifie "Anti-Biopharmaceutical Immunization" soit la prédiction et l'analyse d'éléments cliniques permettant de réduire le risque (**RISK**) d'apparition d'une immunogénicité. Le consortium cible en particulier les anticorps anti-biothérapies.

Pour minimiser le risque d'apparition des ADA et améliorer l'efficacité de la biothérapie, il faut identifier les mécanismes par lesquels les biothérapies stimulent l'activation des cellules immunitaires. Cela comprend la compréhension des structures moléculaires des anticorps induisant

l'immunogénicité de la biothérapie, le rôle de l'inflammation lors de la réponse aux ADA et l'importance des caractéristiques du patient associées au développement des ADA.

Pour cela, des données ont été collectées rétrospectivement dans de multiples centres européens, puis de manière prospective durant les 5 ans du programme ABIRISK initié en 2012. Une banque de données dédiée, basée sur la plateforme tranSMART, rassemble les collectes des différents pays participants. Ce système permet le stockage, l'harmonisation, la comparaison et l'analyse des données d'ADA issues de différents pays. Trois types de pathologies sont étudiés en particulier : la sclérose en plaques, les maladies inflammatoires rhumatoïdes telle que la polyarthrite rhumatoïde et les maladies inflammatoires de l'intestin.

Le tableau 3.7 résume les traitements utilisés dans les cohortes ABIRISK et la prévalence des ADA, selon le traitement et la pathologie. Notons que la positivité aux ADA (donc la prévalence) est grandement dépendante de la méthode de dosage utilisé (KIM et al. [2015]; SONG et al. [2016]).

Pathologies	Traitements utilisés cohortes ABIRISK	Prévalence des ADA
Maladies inflammatoires chroniques de l'intestin	adalimumab, infliximab	— Infliximab : 6-60% UNGAR et al. [2014]
Polyarthrite rhumatoïde	adalimumab, infliximab, rituximab, tocilizumab, etanercept	— Infliximab 10-50% KEISERMAN et al. [2014] — Adalimumab : 1-87% KEISERMAN et al. [2014] — Etanercept : 0-6% KEISERMAN et al. [2014] — Autres : peu documenté KEISERMAN et al. [2014]
Sclérose en plaques	interferon, natalizumab	— Interferon — sous-cutanée : • INF $\beta$ -1a : 28% DEISENHAMMER [2009] • INF $\beta$ -1b : 47% DEISENHAMMER [2009] — intra-musculaire : 6% LUNDKVIST RYNER et al. [2014] — Natalizumab : 9-12% DEISENHAMMER [2009]

TABLEAU 3.7 – Biothérapies et prévalence des ADA associées aux pathologies étudiées dans le cadre d'ABIRISK

## B. Données utilisées

Les données utilisées dans le cadre de la thèse proviennent de données prospectives et rétrospectives collectées sur des patients atteints de polyarthrite rhumatoïde, de sclérose en plaques et de maladies inflammatoires de l'intestin. Nous décrivons ci-dessous les caractéristiques principales de cette base de données.

### *Variable d'intérêt*

La variable d'intérêt étudiée était le délai avant l'apparition d'anticorps anti-biothérapies chez le patient. Pour cela, il a fallu :

- déterminer le seuil de positivité aux ADA à partir duquel un patient était considéré *ADA positif* (ADA+);
- préciser la nature des ADA concernés. Il existe des anticorps pouvant bloquer la fonction de la biothérapie; ce sont les *neutralizing antibodies*, NAbs. Le second type d'anticorps, les *binding antibodies*, BAbs, entraîne une diminution de la demi-vie de la biothérapie. Les deux types d'anticorps peuvent potentiellement affecter l'efficacité du traitement;
- déterminer la période d'observation durant laquelle l'apparition des ADA était directement imputable à la prise des biothérapies.

Un des objectifs du consortium ABIRISK était entre autre d'uniformiser et de standardiser les éléments permettant de déclarer un patient comme ADA positif (RUP et al. [2015]).

### *Recueil des données*

Il est intéressant de noter que les schémas d'étude et de définition de l'événement d'intérêt pouvaient varier, selon le pays et le centre dans le cas des études rétrospectives. Dans le cas des études prospectives, ces éléments avaient été standardisés.

### **Dépistage des ADA (études rétrospectives)**

La pratique clinique consistant à tester les ADA varie selon le pays. Au Danemark par exemple, le dépistage des ADA est obligatoire depuis 1996 pour tous les patients atteints de sclérose en plaques et traités par biothérapies. Ce dépistage est organisé par les laboratoires pharmaceutiques. Tandis que dans d'autres pays, les patients ne sont testés qu'en cas d'échec du traitement. En Suède, les tests pour ADA étaient grandement recommandés, fournis gratuitement et très largement réalisés entre 2003 et 2006, avant d'être rendus obligatoires par les neurologues comme c'est le cas en Autriche et en Allemagne. En Suisse et en Espagne, les tests aux ADA n'ont été réalisés que dans le cadre de projet de recherche.

### **Nombre de visites (études rétrospectives)**

Dans les laboratoires de Munich et de Düsseldorf en Allemagne, entre 2003 et 2008, les patients ont effectué une seule visite afin d'être testés pour la présence d'ADA. Partout ailleurs, les patients effectuaient plusieurs visites; une dizaine en moyenne. À chacune d'entre elles, les patients étaient testés pour la présence d'ADA.

### **NAbs/BAbs (études rétrospectives)**

Par ailleurs, chaque laboratoire pouvait utiliser différentes méthodes d'analyse des ADA. Certains d'entre eux recherchaient les *neutralizing antibodies* et les *binding antibodies* comme par exemple l'Autriche, l'Allemagne ou le Danemark tandis que d'autres laboratoires comme la Suède ne testaient que les *neutralizing antibodies*. Étant donné que les NAbs sont plus fréquemment persistants et cliniquement pertinents que les BAbs, toutes les fois où le terme ADA sera mentionné par la suite, il fera référence aux NAbs.

### Méthode de dosage

La méthodologie utilisée pour déterminer la concentration en NAbs est la suivante : les échantillons positifs aux ADA (en *Laboratory Units* (LU)/ml) sont ensuite titrés suivant la méthode de KAWADE [1980] puis réanalysés afin d'obtenir la valeur titrée (en *Ten-fold Reduction Units* (TRU)/ml). Sachant que l'anticorps agit pour neutraliser une quantité fixe de molécules d'interféron (INF) biologiquement actives, la méthode Kawade permet d'exprimer le pouvoir neutralisant des ADA en mesurant la quantité d'ADA nécessaire à la réduction de la concentration en interféron de 10 LU/ml à 1 LU/ml.

Le seuil de positivité aux ADA est finalement exprimé en TRU/ml et varie selon le laboratoire dans les études rétrospectives :  $\geq 200$  TRU/ml en Suède,  $\geq 100$  TRU/ml en Autriche,  $\geq 20$  TRU/ml au Danemark,  $\geq 100$  TRU/ml en Allemagne-Düsseldorf, etc. Chaque laboratoire ayant défini son propre seuil de positivité, c'est ce dernier qui a été gardé dans toutes les analyses ultérieures.

Concernant les mesures des BAbs, le dosage était réalisé à l'aide de l'une des deux méthodes suivantes :

- la méthode immuno-enzymatique **ELISA** (*Enzyme Linked Immunosorbent Assay*). Dans les échantillons de sérum prélevés, les ADA forment des complexes avec l'IFN- $\beta$ ; les anticorps libres sont détectés par une enzyme (HRP) conjuguée à une protéine (Streptavidine). L'enzyme libérant un composant coloré, la concentration en ADA est évaluée par spectroscopie et exprimée en  $\mu\text{g/mL}$ .
- la méthode **MSD** (*Meso Scale Discovery*), plus récente. Cette méthode utilise des anticorps de détection (par exemple, une molécule de médicament) qui vont se lier aux ADA puis être détectés par électrochimiluminescence. La réaction entre l'anticorps et sa molécule vont permettre d'exciter un luminophore comme le  $[\text{Ru}(\text{bpy})_3]^{2+}$ , une substance qui émet de la lumière après avoir subi une excitation. La lumière émise est ensuite mesurée et permet de doser la quantité d'ADA.

Dans les études prospectives, les analyses de positivité aux ADA ont été centralisées dans un laboratoire unique.

## IV. Références

- ARSEP. 2017, «Fondation pour l'aide à la recherche sur la sclérose en plaque», <https://www.arsep.org>. 61
- DEISENHAMMER, F. 2009, «Neutralizing antibodies to interferon- $\beta$  and other immunological treatments for multiple sclerosis», *CNS Drugs*, vol. 23, n° 5, doi :10.2165/00023210-200923050-00003, p. 379–396, ISSN 1179-1934. URL <http://dx.doi.org/10.2165/00023210-200923050-00003>. 64
- DOOLEY, M. A. et S. L. HOGAN. 2003, «Environmental epidemiology and risk factors for autoimmune disease», *Current Opinion in Rheumatology*, vol. 15, n° 2, p. 99–103, ISSN 1040-8711. 62
- GUILLEMIN, F., A. SARAUX, P. GUGGENBUHL, C. H. ROUX, P. FARDELLONE, E. LE BIHAN, A. CANTAGREL, I. CHARY-VALCKENAERE, L. EULLER-ZIEGLER, R.-M. FLIPO, R. JUVIN, J.-M. BEHIER, B. FAUTREL, C. MASSON et J. COSTE. 2005, «Prevalence of rheumatoid arthritis in france : 2001», *Annals of the Rheumatic Diseases*, vol. 64, n° 10, doi :10.1136/ard.2004.029199, p. 1427–1430, ISSN 0003-4967. URL <http://ard.bmj.com/content/64/10/1427>. 61
- KAWADE, Y. 1980, «An analysis of neutralization reaction of interferon by antibody : a proposal on the expression of neutralization titer», *Journal of Interferon Research*, vol. 1, n° 1, doi :10.1089/jir.1980.1.61, p. 61–70, ISSN 0197-8357. 66
- KEISERMAN, M., C. CODREANU, R. HANDA, D. XIBILLÉ-FRIEDMANN, E. MYSLER, F. BRICEÑO et S. AKAR. 2014, «The effect of antidrug antibodies on the sustainable efficacy of biologic therapies in rheumatoid arthritis : practical consequences», *Expert Review of Clinical Immunology*, vol. 10, n° 8, doi :10.1586/1744666X.2014.926219, p. 1049–1057, ISSN 1744-8409. 64
- KIM, J. S., S. H. KIM, B. KWON et S. HONG. 2015, «Comparison of immunogenicity test methods used in clinical studies of infliximab and its biosimilar (CT-P13)», *Expert Rev Clin Immunol*, vol. 11 Suppl 1, p. 33–41. 64
- LOFTUS, E. V. 2004, «Clinical epidemiology of inflammatory bowel disease : incidence, prevalence, and environmental influences», *Gastroenterology*, vol. 126, n° 6, doi :10.1053/j.gastro.2004.01.063, p. 1504–1517, ISSN 00165085. URL <http://linkinghub.elsevier.com/retrieve/pii/S0016508504004627>. 61
- LUNDKVIST RYNER, M., R. A. FARRELL et A. FOGDELL-HAHN. 2014, «The case for measuring anti-drug antibodies in people with multiple sclerosis», *Expert Review of Clinical Immunology*, vol. 10, n° 6, doi :10.1586/1744666X.2014.914852, p. 697–699, ISSN 1744-8409. 64
- MINICHELLO, E., L. SEMERANO et M.-C. BOISSIER. 2016, «Time trends in the incidence, prevalence, and severity of rheumatoid arthritis : A systematic literature review», *Joint Bone Spine*, vol. 83, n° 6, doi :http://doi.org/10.1016/j.jbspin.2016.07.007, p. 625 – 630, ISSN 1297-319X. URL <http://www.sciencedirect.com/science/article/pii/S1297319X16301294>. 61
- RUP, B., M. PALLARDY, D. SIKKEMA, T. ALBERT, M. ALLEZ, P. BROET, C. CARINI, P. CREEKE, J. DAVIDSON, N. DE VRIES, D. FINCO, A. FOGDELL-HAHN, E. HAVRDOVA, A. HINCELIN-MERY, M. C HOLLAND, P. E. H JENSEN, E. C. JURY, H. KIRBY, D. KRAMER, S. LACROIX-DESMAZES, J. LEGRAND, E. MAGGI, B. MAILLERE, X. MARIETTE, C. MAURI, V. MIKOL, D. MULLEMAN, J. OLDENBURG, G. PAINTAUD, C. R PEDERSEN, N. RUPERTO, R. SEITZ, S. SPINDELDREHER et F. DEISENHAMMER. 2015, «Standardizing terms, definitions and concepts for describing and interpreting unwanted immunogenicity of biopharmaceuticals : recommendations of the Innovative Medicines Initiative ABIRISK consortium», *Clin. Exp. Immunol.*, vol. 181, n° 3, p. 385–400. 65

- SONG, S., L. YANG, W. L. TREPICCHIO et T. WYANT. 2016, «Understanding the supersensitive anti-drug antibody assay : Unexpected high anti-drug antibody incidence and its clinical relevance», *Journal of Immunology Research*, vol. 2016, doi :10.1155/2016/3072586, p. 1–8. URL <https://doi.org/10.1155/2016/3072586>. 64
- UNGAR, B., Y. CHOWERS, M. YAVZORI, O. PICARD, E. FUDIM, O. HAR-NOY, U. KOPYLOV, R. ELIAKIM, S. BEN-HORIN et ABIRISK CONSORTIUM. 2014, «The temporal evolution of antidrug antibodies in patients with inflammatory bowel disease treated with infliximab», *Gut*, vol. 63, n° 8, doi :10.1136/gutjnl-2013-305259, p. 1258–1264, ISSN 1468-3288. 64

## **Deuxième partie**

# **Comparaison de groupes de données censurées par intervalle et risques non-proportionnels**



## Chapitre 4

# Test de comparaison de deux ou plusieurs groupes de données censurées par intervalle avec risques non-proportionnels

### Sommaire

---

<b>I. Introduction</b> . . . . .	<b>72</b>
<b>II. Proposition de test</b> . . . . .	<b>72</b>
A. Notations . . . . .	72
B. Comparaison de deux groupes ( $k = 2$ ) . . . . .	73
C. Extension au cas de $k > 2$ groupes . . . . .	78
D. Implémentation sur R . . . . .	78
<b>III. Comparaisons des tests et simulations</b> . . . . .	<b>79</b>
A. Récapitulatif des tests pour données censurées par intervalle . . . . .	79
B. Simulations avec <i>current status</i> . . . . .	80
C. Commentaires des simulations . . . . .	83
<b>IV. Discussion</b> . . . . .	<b>90</b>
<b>V. Références</b> . . . . .	<b>92</b>

---

## I. Introduction

Jusqu’alors nous avons décrit différents tests, parmi lesquels, certains implémentés, basés sur la vraisemblance groupée avec des risques proportionnels ou odds-proportionnels ainsi que des log-rank pondérés. Cependant, aucun de ces tests ne correspond à une situation à risques croisés.

MANTEL et STABLEIN [1988] ont discuté dès 1988 le cas de l’analyse de fonctions de survie dont les courbes de risques se croisent. Ce type de comportement du rapport des fonctions de risque entre deux groupes peut se rencontrer dans de nombreuses situations. De manière non-exhaustive, nous pouvons citer le cas d’un effet dit à risque accéléré (KALBFLEISCH et PRENTICE [2002]), l’effet d’une interaction latente (ROUAM et al. [2011a]) ou la présence d’un groupe d’individus non-à risque. Dans ces deux derniers cas, si l’effet du groupe peut être conditionnellement à risque proportionnel, l’effet marginal ne l’est pas.

En présence de risques croisés avec des données censurées à droite, le log-rank classique perd en puissance; ce test étant optimal en présence d’une hypothèse alternative à risques proportionnels (STABLEIN et KOUTROUVELIS [1985]). Étant attendu qu’une problématique similaire existait pour les données censurées par intervalle, le développement d’un test statistique adapté à la censure par intervalle en présence de risques croisés était nécessaire.

Par la suite, nous commençons par présenter la version générale du test de comparaison de groupes pour données censurées par intervalle et risques croisés. Nous exposons ensuite les résultats des simulations effectuées sur des données censurées par intervalle de type 1 — les données ayant motivé le développement de ce test étaient censurées par intervalle de type 1 — en comparant le test développé avec les tests de comparaison de groupes déjà implémentés.

La présentation du test et les résultats de l’application du test sur données réelles censurées par intervalle de type 1 ont fait l’objet de l’article JONAS et al. [2017]. Une application du test proposé sur données censurées par intervalle classiques est présentée dans le chapitre suivant.

## II. Proposition de test

### A. Notations

*Nous utilisons les notations introduites en section A. page 10.*

Soient les variables aléatoires continues,  $T$  et  $C$ , représentant, respectivement, la durée écoulée jusqu’à la survenue de l’événement (inconnue) et la date de visite.

On note  $f(t)$ <sup>1</sup>, la fonction de densité de  $T$ . Ainsi que  $S(t) = \bar{F}(t)$  et  $F(t)$  les fonctions de survie et de répartition de  $T$ , respectivement. La fonction de risque de  $T$  est notée  $\lambda(t)$ . Et son taux de risque cumulé,  $\Lambda(t) = \int_0^t \lambda(u) du$ .  $Z$ , représente une variable catégorielle aléatoire, représentant le groupe d’appartenance.

Nous avons  $T \in (L, R]$ , avec  $L$  et  $R$ , les bornes aléatoires de l’intervalle;  $L$  pouvant être égal à 0 (censure gauche), et  $R$  pouvant être égal à  $+\infty$  (censure droite). Dans le cas de la censure par intervalle de type 1, tous les intervalles contiennent 0 ou  $+\infty$ .

Nous faisons l’hypothèse de l’indépendance entre censure et temps d’événement; ce qui nous permet d’ignorer le mécanisme de censure et d’utiliser la vraisemblance simplifiée (voir section 2) page 17).

Nous faisons l’hypothèse supplémentaire de l’indépendance entre temps de censure  $C$  et co-variable  $Z$ ; ce qui signifie que les temps de censures sont indépendants et issus d’une variable aléatoire identiquement distribuée pour tous les sujets.

1. Les seules fonctions de survie étudiées ici étant celles associées à la variable  $T$ ; au lieu de  $(\cdot)_T$  (comme en section A. page 10), nous utilisons plus simplement les notations sans indice pour décrire les fonctions associées à  $T$ .

Notons que sans information supplémentaire sur le processus de censure, il n'est généralement pas possible de tester l'informativité ou l'indépendance entre censure et temps d'événement. Dans le cas où l'informativité de la censure est avérée, des approches adaptées comme une vraisemblance complète présentée par FINKELSTEIN et al. [2002] sont nécessaires. Voir SUN [2006], 10.5.

Le non-respect de l'hypothèse d'indépendance entre censure et covariable, indispensable aux tests comportant une variance permutationnelle, est discuté par FAY et SHIH [2012] et a fait l'objet de travaux de simulation dans ce chapitre.

## B. Comparaison de deux groupes ( $k = 2$ )

Le modèle semi-paramétrique utilisé est un modèle à risques accélérés comme présenté en section c) page 9.

La fonction de survie considérée est :

$$S(t|Z = z) = S_z(t) = \exp \left[ -\Lambda_0(t) \exp(z\beta) \right] \quad (4.1)$$

où  $\Lambda_0(t)$  est une fonction positive croissante de 0 à l'infini.  $Z$  est une variable binaire avec ( $z = 0, 1$ ) et  $\beta$  est un paramètre d'intérêt unidimensionnel.

Il apparait de l'expression (4.1) que c'est une généralisation semi-paramétrique de deux distributions de Weibull dont le paramètre de forme varierait (voir section 2) page 6). Cette forme de fonctions de survie a été introduite par MOREAU et al. [1992]. Elle apparait également chez HSIEH [2001] où elle est décrite comme un modèle de régression à risques hétéroscédastiques (par opposition avec les modèles de Cox dont les risques sont homoscdastiques).

La distribution de Weibull généralisée pour un modèle de régression est une pratique courante. Il s'agit de généraliser une distribution de type Weibull en faisant dépendre le paramètre d'échelle, le paramètre de forme, voire les deux paramètres de la covariable (voir Chapitre 6 de LAWLESS [2002] et section 12.3 de PRABHAKAR MURTHY et al. [2003]). Lorsque c'est le paramètre d'échelle qui dépend de la covariable, nous sommes en présence d'un modèle AFT.

### 1) Fonctions de risque

*Forme des fonctions des risque*

En termes de fonctions de risque, le modèle (4.1) s'écrit :

$$\lambda(t|Z = z) = \lambda_z(t) = \lambda_0(t) \exp(z\beta) \Lambda_0(t)^{[\exp(z\beta)-1]},$$

avec  $\lambda_0(t)$  la fonction de risque instantané de  $T$ .

Soit,

$$h_0(t e^{z\beta}) = \lambda_z(t) = \lambda_0(t) \exp(z\beta) \Lambda_0(t)^{[\exp(z\beta)-1]}.$$

En posant,  $T = t e^{z\beta}$ , nous avons  $t = T e^{-z\beta}$  et nous obtenons :

$$h_0(T) = \lambda_z(t) = \lambda_0(T e^{-z\beta}) \exp(z\beta) \Lambda_0(T e^{-z\beta})^{[\exp(z\beta)-1]}. \quad (4.2)$$

Il apparait de l'expression (4.2) qu'elle correspond à un modèle à risques accélérés tel que défini en section c) page 9 où la covariable a un effet de changement d'échelle entre les fonctions de risque.

### *Croisement des fonctions de risque*

ZHANG et PENG [2009] décrivent différentes fonctions de risques; certaines formes mènent à un croisement des fonctions de risque des deux groupes. Ils se basent sur trois modèles : à risques proportionnels, à risques accélérés et AFT. Ils définissent les conditions sous lesquelles le modèle est un modèle à risques croisés.

Un modèle à risque accéléré a ses fonctions de risques qui se croisent et ne possède qu'un point de croisement si sa fonction de risque de base,  $h_0(t)$ , a une forme en U (voir Théorème 2, ZHANG et PENG [2009]).

Pour identifier le point de croisement des deux courbes de risque, nous étudions l'évolution du rapport de risque, ou plus simplement le logarithme du rapport de risque :

$$\frac{\lambda_1(t)}{\lambda_0(t)} = \frac{\lambda_0(t)e^{\beta} \Lambda_0(t)^{e^{\beta}-1}}{\lambda_0(t)},$$

$$\ln \left[ \frac{\lambda_1(t)}{\lambda_0(t)} \right] = \beta + (e^{\beta} - 1) \ln(\Lambda_0(t)).$$

Il apparait de l'expression précédente qu'avec  $\beta > 0$  et  $\Lambda_0(t)$ , une fonction inversible, monotone et non-décroissante, il existe un point de croisement,  $\tau$ , tel que le rapport  $\frac{\lambda_1(t)}{\lambda_0(t)} > 1$  lorsque  $t < \tau$  et  $\frac{\lambda_1(t)}{\lambda_0(t)} < 1$  lorsque  $t > \tau$ . L'inverse est vrai avec  $\beta < 0$ .

La valeur du point de croisement est alors :

$$\tau = \Lambda^{-1} \left\{ \exp \left[ \frac{-\beta}{e^{\beta} - 1} \right] \right\},$$

où  $\Lambda^{-1}$  est la fonction inverse de  $\Lambda_0$ .

## **2) Score pour censure par intervalle**

### *Hypothèse nulle*

Sous l'hypothèse nulle,  $H_0$ , il n'y a pas de différence entre les deux groupes, ce qui signifie que l'effet de la covariable est nulle :  $\beta = 0$ . Ceci est équivalent à  $\lambda_0(t) = \lambda_1(t) = \lambda(t) \Rightarrow S_0(t) = S_1(t) = S(t)$ .

L'hypothèse alternative  $H_1$  est  $\beta \neq 0$ , avec risques croisés.

### *Expression du score pour la censure par intervalle*

Ainsi que nous l'avons vu en section 2) page 17, la vraisemblance simplifiée sous le modèle (4.1) est :

$$\mathcal{L}(\beta, S_z(\cdot)) = \prod_{i=1}^n \mathbb{P}(T_i \in (l_i, r_i] | Z_i = z_i) = \prod_{i=1}^n [S(l_i | z_i) - S(r_i | z_i)].$$

Et la log-vraisemblance correspondante est :

$$\mathcal{L}_v(\beta, S_z(\cdot)) = \sum_{i=1}^n \ln [S(l_i | z_i) - S(r_i | z_i)].$$

Sous  $H_0$ , la statistique du score est la dérivée première de la log-vraisemblance par rapport au paramètre  $\beta$  évalué en  $\beta = 0$ .

Des équations précédentes, nous obtenons l'expression de la statistique du score,  $U_0$  telle que :

$$U_{0,(\bar{\beta})} = \frac{\partial \mathcal{L}_v(\beta, S_z(\cdot))}{\partial \beta} \Big|_{\beta=0} = \sum_{i=1}^n \frac{S'(l_i | z_i) - S'(r_i | z_i)}{S(l_i | z_i) - S(r_i | z_i)},$$

où  $S'(\cdot|z_i) = \left. \frac{\partial S(\cdot|z_i)}{\partial \beta} \right|_{\beta=0}$ .

Sous le modèle (4.1), nous avons :

$$\begin{aligned} U_{0,(\bar{F})} &= \sum_{i=1}^n z_i \left\{ \frac{S(l_i) \ln(S(l_i)) \ln[-\ln(S(l_i))] - S(r_i) \ln(S(r_i)) \ln[-\ln(S(r_i))]}{S(l_i) - S(r_i)} \right\} \\ &= \sum_{i=1}^n z_i a_{i,\bar{F}}. \end{aligned}$$

Ainsi,  $U_{0,(\bar{F})}$  peut être considéré comme un score de pseudo-rangs dépendant de  $\bar{F} = S$ .

### 3) Liens avec d'autres statistiques

En suivant l'article de OLLER et GÓMEZ [2012],  $U_{0,(\bar{F})}$  peut être ré-écrit sous la forme suivante :

$$U_{0,(\bar{F})} = \sum_{i=1}^n z_i \left\{ \frac{S(l_i) \eta(S(l_i)) - S(r_i) \eta(S(r_i))}{S(l_i) - S(r_i)} \right\}$$

où  $\eta(y) = \ln(y) \ln[-\ln(y)]$  est une fonction dérivable de  $[0 - 1]$  sur  $\mathbb{R}$ , croissante sur l'intervalle  $\left[0 - e^{-\frac{1}{e}}\right]$  et décroissante sur l'intervalle  $\left[e^{-\frac{1}{e}} - 1\right]$ . Notons que  $y \eta(y)$  est remplacé par sa limite en  $y = 0$  et  $y = 1$ .

#### a) Lien avec la censure à droite

D'après la proposition 1 de OLLER et GÓMEZ [2012], la statistique  $U_{0,(\bar{F})}$  pourrait être écrite sous la forme d'un log-rank pondéré dont les poids seraient tels que :

$$w(t) = \hat{S}(t-) \frac{\eta(\hat{S}(t-)) - \eta(\hat{S}(t))}{\hat{S}(t-) - \hat{S}(t)} = \hat{S}(t-) \frac{d\eta(\hat{S}(t))}{d\hat{S}(t)}.$$

Avec  $t- = \lim_{x \rightarrow t} x$  et

$$\frac{d\eta(\hat{S}(t))}{d\hat{S}(t)} = \frac{\ln[-\ln(\hat{S}(t))] + 1}{\hat{S}(t)}.$$

Nous obtenons alors lorsque  $\hat{S}(t-)/\hat{S}(t) \approx 1$  :

$$w(t) \approx \ln[-\ln(\hat{S}(t))] + 1. \quad (4.3)$$

L'expression  $\ln[-\ln(\hat{S}(t))] + 1$  correspond aux poids du log-rank pondéré de MOREAU et al. [1992] destiné à la censure à droite, adaptés à une survie de type Weibull généralisée.

Une remarque supplémentaire sur les poids (4.3) : ces valeurs sont positives lorsque  $\hat{S}(t) < e^{-\frac{1}{e}}$  et négatives autrement. De fait, la statistique  $U_{0,(\bar{F})}$  peut être également vue comme une différence pondérée de risques dont le signe change au point  $t = \exp(-1/e)$ . Le point de changement de signe est lié au modèle choisi, ici, le modèle (4.1).

#### b) Lien avec le log-rank pondéré

FAY [1999] propose une formulation générale de réécriture du score en log-rank pondéré, où le score correspondant au groupe  $l$  ( $l = 1, \dots, k$ ) est :

$$U_l = \sum_{j=1}^m w_j \left[ d_{jl}^* - \frac{n_{jl}^* d_j^*}{n_j^*} \right],$$

où

$$w_j = \frac{\widehat{S}(q_j)\widehat{S}'(q_{j-1}) - \widehat{S}(q_{j-1})\widehat{S}'(q_j)}{\widehat{S}(q_j)[\widehat{S}(q_{j-1}) - \widehat{S}(q_j)]}.$$

Nous avons  $S'(\cdot|z_i) = \left. \frac{\partial S(\cdot|z_i)}{\partial \beta} \right|_{\beta=0} = S(\cdot) \ln(S(\cdot)) \ln[-\ln(S(\cdot))]$ .

Donc, le poids correspondant au score  $U_{0,(\bar{F})}$  réécrit sous la forme d'un log-rank pondéré est :

$$w_j = \widehat{S}(q_{j-1}) \frac{\ln(\widehat{S}(q_{j-1})) \ln[-\ln(\widehat{S}(q_{j-1}))] - \ln(\widehat{S}(q_j)) \ln[-\ln(\widehat{S}(q_j))]}{\widehat{S}(q_{j-1}) - \widehat{S}(q_j)}.$$

#### 4) Score pour censure par intervalle de type 1

Nous présentons à présent la version du test pour risques croisés destinée à la censure par intervalle de type 1.

Les censures à gauche sont telles que  $S(l_i) = 1$  et  $S(r_i) = S(c_i)$  et les censures à droite sont telles que  $S(l_i) = S(c_i)$  et  $S(r_i) = 0$ .

La vraisemblance simplifiée est :

$$\mathcal{L}(\beta, S_z(\cdot)) = \sum_{i=1}^n (1 - \delta_i) \ln[S(c_i | z_i)] + \delta_i \ln[1 - S(c_i | z_i)]. \quad (4.4)$$

La statistique du score pour current status, notée  $U_{0,(\bar{F})}^{cs}$ , est alors telle que :

$$\begin{aligned} U_{0,(\bar{F})}^{cs} &= \sum_{i=1}^n z_i \left\{ (1 - \delta_i) [\eta(S(c_i))] + \delta_i \left[ \frac{-\eta(S(c_i)) S(c_i)}{1 - S(c_i)} \right] \right\} \\ &= \sum_{i=1}^n z_i a_{i,\bar{F}}^{cs} \end{aligned}$$

avec  $a_{i,\bar{F}}^{cs} = \frac{\ln[-\ln(S(c_i))] \ln(S(c_i))}{1 - S(c_i)} [1 - S(c_i) - \delta_i]$ .

#### 5) Distribution asymptotique du test

Nous pouvons relier les statistiques du score  $U_{0,(\bar{F})}$  ou  $U_{0,(\bar{F})}^{cs}$  aux statistiques de rangs linéaires pour censure par intervalle présentées dans la section 4.9.4 par Hájek, Sidák, and Sen dans HÁJEK et al. [1999] en faisant dépendre la fonction de score  $a_{i,\bar{F}}$  d'une fonction  $\phi(\cdot)$  telle que

$a_{i,\bar{F}} = a(i, \phi(\cdot, \bar{F}))$  avec  $\phi(u)$  définie sur  $u \in [0, 1]$ .

La primitive de  $\phi(u)$  est  $\Phi(u) = u \ln(u) \ln(-\ln(u))$ . La fonction  $\phi$  est de carré intégrable sur  $[0, 1]$  avec  $\int_0^1 |\phi(x)|^2 dx \approx 1,82 < \infty$ . Nous avons  $a_{i,\bar{F}} = \phi(\bar{F})$ .

Sous l'hypothèse que les  $C_i$  sont *i.i.d* pour tous les sujets et non liés à  $Z_i$ , l'espérance du score est  $E(U_{0,(\bar{F})}) = n\bar{a}_{\bar{F}}\bar{z}$  (voir également la section 8.5 dans PURI et SEN [1985] qui s'applique aux données groupées<sup>2</sup>).

Les scores étant centrés par définition, nous avons  $\bar{a}_{\bar{F}} = 0$ , ce qui implique que l'espérance du score est nulle également.

La forme linéaire du score  $U_{0,(\bar{F})}$  permet l'emploi de la variance permutationnelle asymptotique sous  $H_0$  qui est telle que (FAY [1996, 1999]; HÁJEK et al. [1999]; KALBFLEISCH et PRENTICE [2002]) :

2. Les données groupées sont un cas particulier de données censurées par intervalle. Dans le cas des données groupées, tous les intervalles se chevauchent complètement les uns sur les autres ou n'ont aucun chevauchement

$$\begin{aligned} V_{U_{0,(\bar{F})}} &= \int_0^1 (\phi(u) - \bar{\phi})^2 du \sum_{i=1}^n (z_i - \bar{z})^2 \\ &= (n-1)^{-1} \sum_{i=1}^n (a_{i,\bar{F}} - \bar{a}_{\bar{F}})^2 \sum_{i=1}^n (z_i - \bar{z})^2 \end{aligned}$$

avec  $\bar{a}_{\bar{F}} = \frac{1}{n} \sum_{i=1}^n a_{i,\bar{F}} = 0$  et  $\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$ .

Finalement, nous considérons la statistique de test suivante :

$$W = \frac{(U_{0,(\bar{F})})^2}{V_{U_{0,(\bar{F})}}}. \quad (4.5)$$

D'après la théorie des tests de rangs de HÁJEK et al. [1999], si nous considérons la statistique :

$$U_{0,(\bar{F})} = \sum_{i=1}^n z_i a_{i,\bar{F}}$$

avec  $a_{i,\bar{F}}$ , des fonctions de score telles que :

$$a_{i,\bar{F}} = n \int_{(i-1)/n}^{i/n} \phi(u) du, \quad i = 1, \dots, n. \quad (4.6)$$

où  $\phi(u)$  est de carré intégrable sur  $[0, 1]$  (lemme 2, section 6.1.6 de HÁJEK et al. [1999]) alors :

sous  $H_0$  et la fonction de score vérifiant la propriété 4.6, la statistique  $U_{0,(\bar{F})}$  est asymptotiquement normale  $\mathcal{N}(\mu_U, \sigma_U^2)$  avec  $\mu_U = E(U_{0,(\bar{F})}) = 0$  et  $\sigma_U^2 = V_{U_{0,(\bar{F})}}$ <sup>3</sup>.

La statistique de forme quadratique  $W$  (expression 4.5) est donc asymptotiquement distribuée sous l'hypothèse nulle, selon une loi du  $\chi^2$  à un degré de liberté. La distribution asymptotique de  $W$  peut être justifiée par le théorème central limite pour variables aléatoires interchangeables (FAY [1996]; PETO et PETO [1972]; SEN [2004]). Dans notre cas, sous l'hypothèse nulle d'invariance permutative, les fonctions de rangs sont en effet des variables aléatoires non indépendantes mais interchangeables.

En pratique,  $\bar{F}$ , inconnu, est remplacé par  $1 - \hat{F}(t)$  où  $\hat{F}(t)$  est l'estimateur du maximum de vraisemblance (NPMLE) de  $F(t)$  obtenu sur l'échantillon total (TURNBULL [1976]). Pour les données censurées par intervalle de type 1,  $\hat{F}(t)$  peut être estimé à l'aide de l'algorithme pool adjacent violators (ROBERTSON et al. [1995]).

La statistique de test dans laquelle la fonction de répartition est remplacée par son estimation, est une "statistique de rangs adaptative" ou *adaptive rank tests* comme décrit en section 8.5 de HÁJEK et al. [1999]. La statistique de test :

$$\tilde{W} = \frac{(U_{0,(\bar{F})})^2}{V_{U_{0,(\bar{F})}}}$$

est asymptotiquement distribuée sous l'hypothèse nulle comme un  $\chi^2$  à un degré de liberté.

Notons que le lent taux de convergence du NPMLE de  $F$  pourrait ralentir les résultats asymptotiques de la statistique de test  $\tilde{W}$  comme stipulé par les auteurs dans la section 8.5.2 de HÁJEK et al. [1999]. Son espérance est désormais d'approximativement 0 sous  $H_0$ <sup>4</sup>.

3. Référence : théorème 1, lemme 2, section 6.1.6 de HÁJEK et al. [1999]

4. la convergence asymptotique étant atteinte lentement. Une solution possible serait de centrer le score.

### C. Extension au cas de $k > 2$ groupes

Le test (4.5) peut être étendu au cas où il y a plus de deux groupes à comparer; on teste l'égalité de  $k$  courbes de survie comme dans FINKELSTEIN et WOLFE [1985] et FAY [1996].

Ici,  $\beta^T$  est un vecteur en ligne de taille  $(k - 1)$  de paramètres et  $Z$  est un vecteur colonne de taille  $(k - 1)$  d'indicateurs pour les  $k$  groupes, l'un des groupes étant choisi comme le groupe de référence.

Ainsi, pour le groupe  $j$  ( $j \in \{1, \dots, k - 1\}$ ) sous le modèle (4.1), le taux de risque instantané pour le groupe  $j$  est :

$$\lambda_j(t|Z_{i,j} = 1) = \lambda_0(t) \exp(\beta_j) \Lambda_0(t)^{[\exp(\beta_j)-1]}$$

où  $\beta_j$  est le  $j^{\text{ième}}$  élément de  $\beta$  associé au groupe  $j$ .

L'hypothèse nulle est  $H_0 : \beta_j = 0 \forall j$  et l'hypothèse alternative est  $H_1 : \exists j, \beta_j \neq 0$ , au moins l'une des fonctions de risque diffère de la fonction de risque de référence.

Alors, en utilisant l'expression de la log-vraisemblance donnée par (4.4) avec un vecteur en ligne de taille  $(k - 1)$  de  $z_i$  au lieu d'un scalaire, le vecteur du score est tel que :

$$U_{0,(\bar{F})} = \sum_{i=1}^n z_i a_{i,\bar{F}}.$$

Le cas où  $\beta$  et  $z_i$  sont des vecteurs et non des scalaires dans le score correspond à l'expression (8.5.38) du livre de PURI et SEN [1985] pour données groupées.

On peut associer au score  $U_{0,(\bar{F})}$  une variance permutationnelle telle que :

$$V_{U_{0,(\bar{F})}} = \frac{1}{n-1} \sum_{i=1}^n \left( a_{i,\bar{F}} - \bar{a}_{\bar{F}} \right)^2 \sum_{i=1}^n (z_i - \bar{z})^T (z_i - \bar{z})$$

avec  $\bar{a}_{\bar{F}} = \frac{1}{n} \sum_{i=1}^n a_{i,\bar{F}} = 0$  et  $\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$ .

Le test omnibus est finalement :

$$W = (U_{0,(\bar{F})})(V_{U_{0,(\bar{F})}})^{-1}(U_{0,(\bar{F})})^T.$$

La statistique de forme quadratique  $W$  est asymptotiquement distribuée comme un  $\chi^2$  à  $(k - 1)$  degrés de liberté (voir la section 4.10.3 dans HÁJEK et al. [1999] portant sur les tests de rangs multivariés).

De même que pour le cas où  $k = 2$  présenté précédemment,  $\bar{F}(t)$  est remplacé par  $1 - \hat{F}(t)$  et la statistique de test correspondante est asymptotiquement distribuée comme un  $\chi^2$  à  $(k - 1)$  degrés de liberté.

### D. Implémentation sur R

Le test présenté précédemment peut être utilisé à l'aide du package *interval*.

Il faut, pour cela, utiliser la fonction 'icest' avec les indications suivantes :

```
"scores='general', dqfunc=fonction(x) ifelse(x <=0,0,ifelse(x==1,0,-x*log(x)*log(-log(x))))".
```

Nous n'avons présenté précédemment que la variance permutationnelle. Mais il est possible d'associer au test développé une variance basée sur la matrice d'information de Fisher ou une variance basée sur des imputations lorsque le test est mis sous la forme d'un log-rank.

Le test présenté est dénommé **vgRC** pour la suite.

### III. Comparaisons des tests et simulations

Dans cette partie, nous présentons les résultats des simulations effectuées avec l'ensemble des tests sur des données censurées par intervalle de type 1. Nous présenterons dans le chapitre suivant les résultats obtenus par application du test développé sur des données réelles.

#### A. Récapitulatif des tests pour données censurées par intervalle

En se basant sur les tests présentés dans la section II. page 38, nous proposons un tableau récapitulatif de comparaisons des différents tests implémentés sur R pour données censurées par intervalle.

##### 1) Les tests du score

En nous inspirant de l'article de FAY et SHAW [2010], voici, dans le tableau (4.1), différents tests du score dont la variance correspond à une matrice d'information de Fisher.

La vraisemblance groupée est (voir section A. page 33) :

$$\prod_{i=1}^n [S(L_i|Z_i\beta, \gamma)] - S(R_i|Z_i\beta, \gamma).$$

Avec S remplacé par  $\hat{S}$ , le NPMLE de S. Dans la vraisemblance partielle, il n'y a pas d'estimation de la fonction de survie.

Test	Vraisemblance	Modèle de régression	Fonction de survie <sup>5</sup>
vgRC	groupée	risques croisés	$\exp(-\Lambda_0(q_j)^{\exp(Z_i\beta)})$
vgRP	groupée	risques proportionnels	$S_0(q_j)^{\exp(Z_i\beta)}$
vgOP	groupée	odds proportionnels	$\left\{1 + \left(\frac{1 - S_0(q_j)}{S_0(q_j)}\right) \exp(Z_i\beta)\right\}^{-1}$
LR adapté	groupée	logistiques, risques discrets	$\prod_{y=1}^j \left\{1 + \left(\frac{S_0(q_{y-1}) - S_0(q_y)}{S_0(q_y)}\right) \exp(Z_i\beta)\right\}^{-1}$
LRP OG	groupée	$\rho = 0$ : risques proportionnels, $\rho = 1$ : odds proportionnels*	
vpMA	partielle	risques proportionnels	$S_0(t)^{\exp(Z_i\beta)}$
vpIMI	partielle	risques additifs	$\exp\left(-\int_0^t \lambda_0(u) + \exp(Z_i\beta) du\right)$

TABLEAU 4.1 – Tests du score pour données censurées par intervalle avec variance basée sur la matrice d'information de Fisher

\* **LRP OG** : dans un modèle de transformation linéaire de la forme  $h(T) = -Z_i\beta + \epsilon_i$ ,  $\epsilon_i$  a pour fonction de survie  $S_{\epsilon}(t) = [1 + \rho \exp(t)]^{-1/\rho}$ .

##### 2) Les tests de permutation

Nous présentons ci-dessous les différents tests du score ayant une variance permutationnelle; c'est-à-dire que le score, U, peut être écrit sous la forme linéaire  $U = \sum Z_i \omega_i$ .

Les tests peuvent être résumés par les scores,  $\omega_i$ , ayant la forme suivante :

$$\omega_i = \frac{\hat{S}(l_i)\xi(\hat{S}(l_i)) - \hat{S}(r_i)\xi(\hat{S}(r_i))}{\hat{S}(l_i) - \hat{S}(r_i)}.$$

La fonction de survie est remplacée par  $\widehat{S}$ , son NPMLE sauf pour le **LR adapté** ou  $\widehat{S}$  est remplacé par  $\widetilde{S}$ , l'estimateur de Nelson-Aalen (signalé par un \* dans le tableau).

Le tableau 4.2 ci-dessous explicite les fonctions de lien  $\xi(x)$ .

Test	Fonctions de lien, $\xi$
vgRC	$\ln(x) \ln[-\ln(x)]$
vgRP	$\ln(x)$
vgOP	$x - 1$
LR adapté	$\ln(\widetilde{x})^*$
LRP SZZ	$\ln(x)x^\rho(1-x)^\gamma$
LRP OG	$-\int_0^{1-x}(1-u)^{\rho-1}u^\gamma du$

TABLEAU 4.2 – Tests du score pour données censurées par intervalle avec variance permutative

### 3) Les tests de (pseudo) log-rank

Nous présentons à présent un récapitulatif des tests pour données censurées par intervalle pouvant être mis sous la forme d'un log-rank tel que  $U = \sum w_j \left[ d_{jl}^* - \frac{n_{jl}^* d_j^*}{d_j^*} \right]$ .

La manière de calculer les personnes à risque peut varier d'un test à l'autre comme indiqué en section C. page 35. Ici, nous comparons les poids  $w_j$  des différents tests de log-ranks. Le tableau 4.3 résume les fonctions de poids  $w_j$  des tests présentés précédemment.

Test	Fonctions de poids, $w_j$
vgRC	$\widehat{S}(q_{j-1}) \frac{\ln(\widehat{S}(q_{j-1})) \ln[-\ln(\widehat{S}(q_{j-1}))] - \ln(\widehat{S}(q_j)) \ln[-\ln(\widehat{S}(q_j))]}{\widehat{S}(q_{j-1}) - \widehat{S}(q_j)}$
vgRP	$\widehat{S}(q_{j-1}) \frac{\ln(\widehat{S}(q_{j-1})) - \ln(\widehat{S}(q_j))}{\widehat{S}(q_{j-1}) - \widehat{S}(q_j)}$
vgOP	$\widehat{S}(q_{j-1})$
LR adapté LR améliore	1
LRP OG	$\widehat{S}(q_{j-1}) \frac{\int_0^{1-\widehat{S}(q_j)} x^\gamma (1-x)^{\rho-1} dx - \int_0^{1-\widehat{S}(q_{j-1})} x^\gamma (1-x)^{\rho-1} dx}{\widehat{S}(q_{j-1}) - \widehat{S}(q_j)}$

TABLEAU 4.3 – Tests du log-ranks pour données censurées par intervalle

## B. Simulations avec *current status*

Nous présentons dans cette section des simulations de Monte-Carlo incluant les différents tests présentés dans les sections précédentes ainsi que le test que nous avons développé.

Les simulations ont été effectuées sur des données *current status*; les données ayant motivé le développement du test étant de ce type.

### 1) Design des simulations

Pour évaluer les performances du test proposé (**vgRC**) sur des échantillons de tailles finies, tout en le comparant aux tests classiques décrits ci-dessus, nous avons effectué des simulations de Monte Carlo sur des échantillons de deux groupes incluant plusieurs scénarios avec différents types et pourcentages de censure.

Nous avons également effectué des simulations supplémentaires afin d'évaluer ses performances sous un modèle de survie aux risques croisés différent de celui sous lequel le test **vgRC** a été développé. Et enfin, des simulations ont été réalisées pour estimer la robustesse du test en présence de censure inégale.

Pour les trois scénarios suivants, la variable  $Z$  a été générée, selon une distribution de Bernoulli de paramètre  $\pi = 0,5$ , le paramètre  $\beta$  variait de 0 à 3 et, sauf mention contraire, la taille totale de l'échantillon pour chaque simulation était de  $n = 300$ .

Excepté pour le scénario 3, les temps de censure (visite)  $C$  ont été générés indépendamment soit à partir d'une distribution exponentielle avec un taux de risque  $\lambda_C$ , soit à partir d'une distribution uniforme s'étendant sur  $]0, c_{max}]$ . Les deux paramètres,  $\lambda_C$  et  $c_{max}$  variaient, selon  $p$ , le pourcentage désiré d'observations censurées à droite.

Le taux de censure,  $p$ , inclut la fraction de patients qui n'ont pas fait l'événement d'intérêt à leur temps de visite.

#### a) Scénario 1

Pour la première série de simulations, les temps d'évènement  $T$  ont été générés à partir d'une distribution de Weibull avec un taux de risque  $v\left(\frac{1}{\zeta}\right)^v t^{v-1}$  (paramètre de forme :  $v = \exp(\beta Z)$  et paramètre d'échelle :  $\zeta = 1$ ) afin de simuler un modèle à risques croisés. Deux niveaux de censures (à droite) ont été considérés ( $p = 30\%, 60\%$ ).

Nous avons effectué des simulations supplémentaires avec une petite taille d'échantillon ( $n = 100$ ) et  $p = 30\%$ .

Les résultats sous la censure exponentielle sont reportés dans la table 4.5. Les résultats sous la censure uniforme sont donnés dans la table 4.6. La table 4.7 présente les résultats des simulations avec un petit échantillon ( $n = 100$ ). Les graphiques correspondants aux tables 4.5 et 4.6 sont également présentés (respectivement, figures 4.1 et 4.2).

#### b) Scénario 2

Afin d'évaluer le comportement du test **vgRC** sous un modèle de survie aux risques croisés différent du précédent, nous avons effectué des simulations supplémentaires basées sur un modèle à risques accélérés (**AH** pour *accelerated hazard*) présenté en section c) page 9 et dans ZHANG et PENG [2009].

En pratique, les temps d'évènement ont été générés à partir d'une distribution log-logistique, avec une fonction de risque telle que

$$\lambda_T(t|Z) = \frac{\alpha (te^{\beta Z})^{\alpha-1} \lambda}{1 + (te^{\beta Z})^\alpha \lambda}.$$

Ici, nous avons considéré que  $\alpha = 2$ ,  $\lambda = 1$ .

Comme décrit dans ZHANG et PENG [2009], les fonctions de risque du modèle AH ne se croisent pas lorsque  $0 < \alpha < 1$  et ont un croisement unique lorsque  $\alpha > 1$ .

Ainsi, ici, les fonctions de risque se croisent une fois au point

$$\tau = \left( \frac{e^{-\theta(\alpha-1)} - 1}{\lambda(1 - e^\theta)} \right)^{\frac{1}{\alpha}}.$$

Les résultats de ces simulations, avec censures exponentielle et uniforme, sont présentées dans le tableau 4.8.

### c) Scénario 3

Afin d'évaluer la robustesse du test à l'hypothèse non-respectée d'un même mécanisme de censure pour tous les sujets, nous avons considéré une situation de censure inégale où la distribution des temps de censure est différente entre les groupes à comparer. Elle est désignée par FAY et SHIH [2012] comme *assessment-treatment dependence (ATD)*, une dépendance de la censure au traitement.

En pratique, deux distributions ont été considérées pour générer les temps de visite, avec  $v = 0, 25, 0, 75$  :

- une distribution exponentielle avec un taux de risque ( $e^{vZ}\lambda_C$ );
- une distribution uniforme variant sur  $]0, c_{max} + e^{vZ}]$ .

Les temps d'événement ont été générés de la même manière que dans le scénario 1. Les résultats sont donnés dans le tableau 4.9.

Pour les deux derniers scénarios, le pourcentage d'observations censurées à droite était  $p = 30\%$ .

Pour chaque configuration, nous avons généré 5 000 échantillons. Nous avons considéré un test bilatéral avec un niveau de signification de 0,05. Pour le test développé **vgRC**, le NPMLE de la fonction de survie a été obtenu en utilisant l'algorithme pool-adjacent violators disponible sur le package 'isotone' (LEEUW et al. [2009]).

Les résultats, en pourcentages, sont présentés dans la section suivante.

Les tests comparés dans les simulations sont présentés dans le tableau 4.4 :

Test	Variance	Remarque
vgRC	Variance permutacionnelle	—
vgOP		= LRP OG $_{\rho=0, \gamma=0}$
vgRP		= LRP OG $_{\rho=1, \gamma=0}$
LR adapté		—
LR amélioré	Variance par imputations (log-rank)	—
vgRP <sub>Fisher</sub>	Matrice observée de Fisher	= LRP OG $_{\rho=0, \gamma=0}$ + variance Fisher
LRP SZZ	Variance asymptotique normale	Paramètres : { $\rho = 0, \gamma = 0$ }; { $\rho = 0, \gamma = 1$ } { $\rho = 1, \gamma = 0$ }; { $\rho = 1, \gamma = 1$ }
LRP OG	Variance permutacionnelle	
LRPHyb	Variance asymptotique normale	
vpMA	Matrice observée de Fisher	—
vpIMI		—

TABEAU 4.4 – Tests comparés dans les simulations.

## 2) Résultats des simulations

Les résultats des simulations sont présentés dans les tableaux 4.5 à 4.9 et dans les graphiques 4.1 et 4.2.

	$\beta = 0$	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2	3	
<b>vgRC</b>	5.4	12.4	36.2	63.7	83.3	94.9	98.5	99.6	99.8	100.0	100.0	100.0	Censure droite <b>p = 30%</b>
vgOP	5.4	5.2	5.5	5.6	5.2	6.2	7.1	8.0	8.6	10.9	13.5	20.1	
vgRP	5.2	5.6	8.4	9.9	11.2	12.8	12.6	11.7	11.8	11.8	11.6	8.6	
vgRP <sub>Fisher</sub>	6.0	7.0	9.9	11.8	13.2	15.4	15.4	14.5	14.9	14.4	14.4	11.2	
LR adapté	5.2	5.4	7.6	8.5	8.7	9.9	9.4	8.2	7.8	7.9	7.5	5.7	
LR amélioré	0.1	0.1	0.1	0.1	0.2	0.0	0.1	0.0	0.0	0.0	0.1	0.0	
LRP SZZ $_{\rho=0,\gamma=0}$	5.2	5.7	8.5	10.0	11.3	12.9	12.7	11.7	11.9	11.9	11.7	8.6	
$\rho=0,\gamma=1$	4.7	7.7	15.6	24.2	31.9	43.2	50.8	56.1	62.2	64.4	66.7	71.9	
$\rho=1,\gamma=0$	5.4	6.3	8.5	12.1	19.1	29.0	38.7	48.3	58.1	65.1	71.5	87.9	
$\rho=1,\gamma=1$	5.2	5.4	6.1	6.3	6.8	7.9	9.0	9.4	9.6	12.2	13.6	18.7	
LRP OG $_{\rho=0,\gamma=1}$	4.8	7.7	17.1	26.8	35.9	48.5	56.5	62.6	68.8	70.8	73.3	78.4	
$\rho=1,\gamma=1$	5.0	6.3	10.0	14.2	17.4	22.4	25.9	27.5	31.0	32.6	33.2	36.6	
LRPHyb $_{\rho=0,\gamma=0}$	7.2	7.5	10.7	13.9	17.1	18.7	20.0	20.9	21.2	21.2	22.2	23.9	
$\rho=0,\gamma=1$	8.8	8.9	13.0	17.7	21.6	26.1	30.0	34.4	36.3	39.7	41.6	53.7	
$\rho=1,\gamma=0$	5.7	5.6	6.7	7.3	7.4	8.1	8.8	10.4	11.9	13.0	14.4	20.7	
$\rho=1,\gamma=1$	6.2	6.7	7.0	8.2	9.6	11.2	11.5	12.9	13.2	13.6	15.3	19.1	
vpMA	5.3	5.5	5.6	6.0	4.9	5.3	5.4	6.1	6.9	6.8	8.1	9.5	
vpIMI	6.1	5.7	6.6	6.6	6.3	6.5	6.2	5.4	5.8	5.5	6.1	6.1	
<b>vgRC</b>	5.3	14.3	39.9	71.4	92.2	98.5	99.8	100.0	100.0	100.0	100.0	100.0	
vgOP	5.9	8.0	17.0	26.1	40.0	52.6	64.0	74.5	80.1	84.3	88.8	95.0	
vgRP	5.8	6.7	10.1	12.7	16.1	19.0	22.3	24.8	26.0	27.1	30.8	35.5	
vgRP <sub>Fisher</sub>	6.6	8.0	11.3	13.9	18.2	21.1	24.9	28.0	29.2	30.3	34.2	39.6	
LR adapté	5.8	7.1	11.6	14.7	20.9	26.1	31.9	37.7	40.9	44.4	48.9	61.3	
LR amélioré	2.4	2.4	4.2	5.1	6.7	8.6	10.5	13.0	13.3	14.4	18.0	25.0	
LRP SZZ $_{\rho=0,\gamma=0}$	5.8	6.8	10.2	12.7	16.1	19.1	22.4	25.0	26.0	27.2	30.9	35.7	
$\rho=0,\gamma=1$	5.7	5.2	5.6	6.2	6.3	7.7	9.3	11.1	14.3	17.3	18.5	26.8	
$\rho=1,\gamma=0$	5.9	10.2	25.4	43.2	66.1	81.9	91.7	96.6	98.6	99.3	99.6	100.0	
$\rho=1,\gamma=1$	5.7	7.0	11.1	14.5	20.7	26.5	33.5	40.8	44.2	48.2	53.9	65.3	
LRP OG $_{\rho=0,\gamma=1}$	5.5	5.1	5.7	6.7	7.4	10.2	12.5	15.1	19.6	24.6	26.1	38.0	
$\rho=1,\gamma=1$	5.5	5.7	6.8	6.5	6.1	7.0	7.0	7.2	7.0	6.8	7.3	8.1	
LRPHyb $_{\rho=0,\gamma=0}$	5.4	5.0	6.5	7.2	9.1	10.1	10.9	12.5	13.4	15.5	17.2	29.9	
$\rho=0,\gamma=1$	6.0	5.7	7.0	7.2	8.6	9.8	10.6	11.4	12.8	14.6	14.3	24.0	
$\rho=1,\gamma=0$	6.9	6.5	8.3	10.8	14.2	18.4	22.7	28.0	32.7	36.6	39.2	54.8	
$\rho=1,\gamma=1$	5.5	5.7	6.0	6.2	8.5	9.3	11.0	12.3	15.7	18.0	19.9	37.0	
vpMA	5.3	6.6	8.6	10.5	13.8	17.2	19.6	23.7	24.7	26.8	29.8	34.4	
vpIMI	5.9	6.1	9.6	10.9	13.4	16.9	19.5	22.7	23.8	25.0	29.4	37.6	

TABEAU 4.5 – Simulations sous modèle à risques croisés avec censure exponentielle,  $n = 300, \pi = 0.5$ . Les résultats, en %, correspondent aux risques de première espèce et puissances des tests comparés (en ligne) en fonction des valeurs du  $\beta$  (en colonne) comprises entre 0 et 3. Le tableau du haut correspond à une censure droite de 30% et le tableau du bas à une censure droite de 60%.

### C. Commentaires des simulations

#### 1) Scénario 1

##### *Risque de première espèce et puissance*

À partir des simulations sous le modèle de Weibull (tableaux 4.5, 4.6 et 4.7), il apparaît que le risque de première espèce estimé du score **vgRC** appartient à l'intervalle de fluctuation pour un risque nominal de 5% construit à partir de la loi théorique sous  $H_0$  qui est un  $\chi^2$  a un degré de

CHAPITRE 4. TEST DE COMPARAISON DE DEUX OU PLUSIEURS GROUPES DE DONNÉES  
CENSURÉES PAR INTERVALLE AVEC RISQUES NON-PROPORTIONNELS

	$\beta = 0$	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2	3
<b>vgRC</b>	5.5	15.6	45.3	78.5	94.6	98.7	99.8	100.0	100.0	100.0	100.0	100.0
vgOP	5.7	6.3	7.8	9.8	9.8	10.1	9.0	8.6	7.3	6.8	6.6	5.9
vgRP	5.6	8.4	16.6	27.1	33.9	41.3	46.2	49.0	48.4	49.8	48.9	46.8
vgRP <sub>Fisher</sub>	6.6	9.5	18.6	30.0	38.0	45.7	51.0	53.4	53.7	55.5	54.0	52.7
LR adapté	5.5	8.0	15.2	23.8	28.5	33.9	37.7	39.0	36.3	37.2	34.6	28.3
LR amélioré	0.6	1.2	1.7	3.3	3.8	4.7	4.8	4.6	4.1	3.7	4.1	2.3
LRP SZZ $_{\rho=0,\gamma=0}$	5.7	8.5	16.6	27.1	34.3	41.5	46.4	49.2	48.5	49.9	49.0	47.0
$\rho=0,\gamma=1$	5.5	10.7	26.4	47.8	66.0	78.7	86.0	90.5	93.2	94.2	95.1	97.0
$\rho=1,\gamma=0$	5.5	6.2	6.2	8.0	11.5	15.4	22.5	30.3	38.4	46.8	54.4	75.6
$\rho=1,\gamma=1$	5.5	6.3	7.1	8.9	9.2	11.0	11.4	12.5	12.5	12.7	13.7	15.4
LRP OG $_{\rho=0,\gamma=1}$	5.6	10.9	28.8	51.6	71.4	83.4	90.0	93.8	95.8	96.5	97.0	98.5
$\rho=1,\gamma=1$	5.6	8.8	18.1	30.8	40.9	51.6	59.9	64.4	68.1	70.3	72.7	76.5
LRPHyb $_{\rho=0,\gamma=0}$	6.2	8.4	11.4	18.7	27.3	32.0	36.1	38.0	38.5	39.0	39.1	45.9
$\rho=0,\gamma=1$	7.3	8.2	13.4	22.7	33.3	39.4	45.5	49.9	52.4	57.6	60.0	74.8
$\rho=1,\gamma=0$	5.8	5.7	6.0	7.4	9.1	9.7	10.2	10.2	10.1	10.2	10.6	17.4
$\rho=1,\gamma=1$	5.3	6.4	8.0	10.8	14.7	16.8	20.7	21.4	21.3	21.6	22.3	26.5
vpMA	5.2	6.0	6.5	7.1	7.2	7.0	6.3	5.8	6.3	5.6	5.7	5.3
vpIMI	6.5	8.3	14.1	23.2	30.9	38.0	42.4	42.5	41.2	40.1	39.2	33.6
<b>vgRC</b>	5.4	11.1	31.1	57.7	81.8	95.0	99.0	99.9	100.0	100.0	100.0	100.0
vgOP	5.6	9.4	17.9	30.4	42.8	54.7	64.4	69.9	74.9	79.2	83.3	92.8
vgRP	5.3	8.6	13.5	20.4	27.1	31.8	34.4	32.0	30.5	27.2	26.5	24.8
vgRP <sub>Fisher</sub>	5.9	9.3	14.7	22.2	28.9	33.6	36.5	34.6	33.5	29.7	29.6	28.2
LR adapté	5.4	8.6	14.1	21.7	29.2	34.7	39.4	38.7	39.0	37.6	39.1	46.4
LR amélioré	3.8	6.5	10.1	16.1	21.5	25.6	28.1	26.1	25.4	22.6	22.8	26.0
LRP SZZ $_{\rho=0,\gamma=0}$	5.3	8.6	13.5	20.5	27.3	31.9	34.5	32.1	30.7	27.2	26.6	24.8
$\rho=0,\gamma=1$	5.3	7.0	7.7	8.4	7.5	6.7	6.2	5.8	9.5	14.4	21.9	45.8
$\rho=1,\gamma=0$	5.7	10.3	23.0	41.5	59.9	76.4	87.5	93.9	96.4	98.8	99.5	100.0
$\rho=1,\gamma=1$	5.3	8.2	11.8	17.1	23.4	28.9	34.1	36.0	39.2	40.5	44.2	57.0
LRP OG $_{\rho=0,\gamma=1}$	5.3	6.9	7.3	7.6	6.6	5.4	6.3	7.1	13.0	20.8	30.5	60.5
$\rho=1,\gamma=1$	5.2	7.5	9.2	11.4	12.3	12.4	11.1	8.2	8.3	6.5	7.3	9.0
LRPHyb $_{\rho=0,\gamma=0}$	4.4	4.5	5.0	5.7	7.8	10.2	13.3	14.1	14.8	14.2	15.0	21.4
$\rho=0,\gamma=1$	3.9	4.1	3.9	4.7	4.8	5.9	6.1	7.3	7.6	9.0	9.3	16.7
$\rho=1,\gamma=0$	6.5	6.2	7.5	10.0	12.8	15.8	20.0	23.2	24.6	28.0	30.2	41.1
$\rho=1,\gamma=1$	4.4	5.1	5.2	6.6	7.9	8.7	11.2	12.7	13.8	14.5	15.8	26.8
vpMA	5.8	6.6	9.3	13.4	18.4	22.7	25.2	27.7	31.0	31.6	34.0	40.3
vpIMI	5.8	9.2	15.8	26.2	36.7	47.0	56.2	60.8	66.3	69.6	73.8	86.0

Censure droite  $p = 30\%$

Censure droite  $p = 60\%$

TABEAU 4.6 – **Simulations sous modèle à risques croisés avec censure uniforme**,  $n = 300, \pi = 0.5$ . Les résultats, en %, correspondent aux risques de première espèce et puissances des tests comparés (en ligne) en fonction des valeurs du  $\beta$  (en colonne) comprises entre 0 et 3. Le tableau du haut correspond à une censure droite de 30% et le tableau du bas à une censure droite de 60%.

liberté qui est  $[0, 025; 0, 075]$  pour un échantillon de 300 individus .

Sous les hypothèses alternatives, les gains de puissance sont conséquents, quels que soient les types et les pourcentages de censure. Comme les autres tests, le score **vgRC** est néanmoins sensible au scénario de censure avec une puissance plus élevée sous une censure uniforme.

Concernant les tests comparés, nous notons que le test **LR amélioré** est extrêmement conservateur sous divers schémas de censure; avec une erreur de type I aussi faible que 0,10% pour la censure exponentielle. Ce risque de première espèce très faible est probablement lié au type de censure par intervalle (ici, censure par intervalle de type I); en effet, le risque de première espèce du **LR amélioré** est correct pour des simulations réalisées sur des données simulées censurées par

	$\beta = 0$	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2	3
<b>vgRC</b>	4.8	6.4	11.1	18.0	28.5	37.6	48.0	56.7	63.1	69.2	74.4	83.1
vgOP	5.8	5.7	5.2	5.9	6.3	6.1	6.5	6.4	7.6	8.3	8.4	10.8
vgRP	5.7	5.6	5.8	6.4	6.1	6.1	6.1	6.2	6.8	6.6	6.7	6.8
vgRP <sub>Fisher</sub>	7.7	8.2	8.1	8.9	8.7	9.3	8.9	9.1	9.6	9.5	9.4	8.9
LR adapté	5.6	5.5	5.6	6.2	6.0	6.1	6.0	5.8	6.7	6.7	6.8	7.5
LR amélioré	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LRP SZZ $_{\rho=0,\gamma=0}$	5.8	5.8	5.9	6.5	6.2	6.3	6.2	6.3	7.0	6.8	6.8	7.0
$\rho=0,\gamma=1$	5.6	5.6	6.3	8.5	9.0	10.4	11.8	12.8	13.4	13.9	15.6	15.6
$\rho=1,\gamma=0$	5.9	5.9	6.2	7.7	9.3	11.1	13.4	15.9	18.3	21.4	24.5	33.3
$\rho=1,\gamma=1$	6.0	5.1	5.7	6.8	6.6	7.3	7.4	7.7	9.0	9.8	9.9	13.1
LRP OG $_{\rho=0,\gamma=1}$	5.2	5.3	5.9	8.5	9.0	10.6	12.0	13.2	14.0	14.5	16.2	16.4
$\rho=1,\gamma=1$	5.8	5.7	6.2	7.2	7.6	8.7	9.0	9.6	10.0	10.5	11.0	10.6
LRPHyb $_{\rho=0,\gamma=0}$	6.8	7.6	8.4	9.8	10.8	12.3	12.7	14.0	15.9	16.3	18.3	27.9
$\rho=0,\gamma=1$	8.4	9.2	11.1	12.8	15.9	17.9	20.9	23.1	24.9	26.5	28.2	32.8
$\rho=1,\gamma=0$	6.1	6.0	6.9	7.3	8.8	9.3	10.7	12.5	13.6	14.2	16.3	27.8
$\rho=1,\gamma=1$	6.5	6.7	7.6	8.8	9.3	11.1	12.5	13.7	15.0	16.7	18.0	28.2
vpMA	5.4	5.5	5.2	5.5	5.6	6.4	5.7	6.2	6.5	6.0	6.2	7.1
vpIMI	6.2	6.2	6.2	5.8	5.5	6.1	6.6	6.3	6.1	6.0	5.9	5.9
<b>vgRC</b>	4.6	7.8	13.8	26.0	38.9	50.4	61.6	70.2	76.8	80.7	84.2	91.5
vgOP	6.0	5.8	6.8	6.7	6.9	7.3	6.6	7.2	7.1	6.2	6.9	6.2
vgRP	5.8	6.4	8.1	10.1	11.1	13.3	13.6	14.7	14.9	13.6	14.0	11.9
vgRP <sub>Fisher</sub>	7.6	8.6	10.9	13.6	14.9	17.6	18.4	19.4	18.8	18.1	18.5	16.0
LR adapté	6.0	6.4	8.0	9.5	10.3	12.0	11.7	12.5	11.9	10.6	10.9	8.8
LR amélioré	0.6	0.4	0.4	0.5	0.6	0.7	0.7	0.7	0.5	0.5	0.4	0.3
LRP SZZ $_{\rho=0,\gamma=0}$	6.0	6.6	8.4	10.5	11.2	13.5	13.8	15.0	15.1	13.7	14.3	12.0
$\rho=0,\gamma=1$	5.5	6.9	10.5	15.3	19.5	24.0	27.4	30.4	32.3	33.8	35.4	36.2
$\rho=1,\gamma=0$	5.9	5.8	7.2	6.1	6.4	8.0	9.3	10.4	13.1	15.1	16.7	23.3
$\rho=1,\gamma=1$	6.1	6.1	6.7	7.1	7.8	9.2	8.9	9.6	10.3	9.3	11.0	11.8
LRP OG $_{\rho=0,\gamma=1}$	5.2	6.7	10.6	15.3	20.2	24.5	28.7	31.6	33.5	35.1	36.8	37.9
$\rho=1,\gamma=1$	5.7	6.6	8.8	11.7	14.2	17.2	19.3	21.4	22.2	21.8	23.7	23.7
LRPHyb $_{\rho=0,\gamma=0}$	7.7	8.4	9.4	11.1	14.6	16.6	16.9	19.5	20.3	21.1	23.5	27.5
$\rho=0,\gamma=1$	9.1	10.3	12.0	15.0	20.5	24.0	26.8	31.0	34.0	36.5	38.5	42.3
$\rho=1,\gamma=0$	5.5	5.8	7.0	7.5	8.8	9.6	10.2	12.1	13.7	13.5	16.4	23.4
$\rho=1,\gamma=1$	7.3	7.1	7.9	8.7	11.7	13.3	14.1	16.3	17.0	17.4	20.2	25.3
vpMA	5.4	5.9	6.3	5.7	6.4	6.0	5.8	5.7	5.7	5.4	5.8	5.5
vpIMI	6.2	6.9	8.1	11.1	13.0	14.5	14.8	15.1	15.6	14.6	13.9	12.6

Censure exponentielle

Censure uniforme

TABLEAU 4.7 – Simulations sous modèle à risques croisés avec petits échantillons,  $n = 100, \pi = 0.5, p = 30\%$ . Les résultats, en %, correspondent aux risques de première espèce et puissances des tests comparés (en ligne) en fonction des valeurs du  $\beta$  (en colonne) comprises entre 0 et 3. Le tableau du haut correspond à une censure exponentielle et le tableau du bas à une censure uniforme.

intervalle de type II (résultats non-montrés ici).

Le taux d'erreur de type I du score **vgRP<sub>Fisher</sub>** est assez élevé même s'il se situe dans l'intervalle de fluctuation du risque nominal. En effet, OLLER et GÓMEZ [2012] ne recommandent pas l'utilisation de la matrice d'information de Fisher observée; le nombre de paramètres de nuisance pouvant affecter le comportement du test.

Le risque de première espèce du **LRPHyb** dépasse le niveau nominal de 5%.

En général, les autres tests ont une puissance plus faible que score **vgRC** dans les situations considérées ici. Ces différences de puissance sont encore plus prononcées en cas de censure exponentielle modérée ( $p = 30\%$ ). Par exemple, dans ce schéma, le modèle le plus puissant après

CHAPITRE 4. TEST DE COMPARAISON DE DEUX OU PLUSIEURS GROUPES DE DONNÉES  
CENSURÉES PAR INTERVALLE AVEC RISQUES NON-PROPORTIONNELS

	$\beta = 0$	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2	3	
<b>vgRC</b>	5.1	6.6	12.0	25.6	43.5	67.6	84.9	95.5	99.2	99.7	100.0	99.3	Censure exponentielle
vgOP	5.2	5.1	5.8	6.6	10.9	21.4	44.7	73.6	93.4	99.4	100.0	100.0	
vgRP	5.0	5.6	6.8	10.3	19.6	40.1	68.2	90.3	98.9	100.0	100.0	100.0	
vgRP <sub>Fisher</sub>	5.8	6.5	7.8	11.6	22.0	43.2	70.9	91.6	99.1	100.0	100.0	100.0	
LR adapté	5.0	5.4	6.4	9.6	18.3	37.5	65.9	89.1	98.7	100.0	100.0	100.0	
LR amélioré	0.1	0.1	0.1	0.7	2.9	10.7	32.2	65.6	91.3	99.2	100.0	100.0	
LRP SZZ <sub><math>\rho=0, \gamma=0</math></sub>	5.0	5.6	6.8	10.3	19.8	40.3	68.3	90.3	98.9	100.0	100.0	100.0	
$\rho=0, \gamma=1$	4.9	5.8	8.0	14.8	29.0	54.7	81.4	96.1	99.8	100.0	100.0	100.0	
$\rho=1, \gamma=0$	5.0	5.9	6.0	5.7	5.6	6.5	13.5	31.1	59.5	86.8	98.5	100.0	
$\rho=1, \gamma=1$	5.0	5.6	5.4	6.3	9.2	18.2	40.2	69.6	91.6	99.0	100.0	100.0	
LRP OG <sub><math>\rho=0, \gamma=1</math></sub>	4.9	5.7	8.3	15.7	30.6	57.1	83.1	96.6	99.8	100.0	100.0	100.0	
$\rho=1, \gamma=1$	4.9	5.8	6.3	11.0	20.9	42.0	71.0	91.3	99.0	100.0	100.0	100.0	
LRPHyb <sub><math>\rho=0, \gamma=0</math></sub>	7.4	6.7	7.1	7.9	11.2	12.2	16.3	20.8	25.8	31.2	38.0	73.1	
$\rho=0, \gamma=1$	7.7	7.4	8.3	9.3	12.7	16.0	19.1	23.5	28.2	31.2	37.3	56.0	
$\rho=1, \gamma=0$	5.1	5.5	5.3	5.8	6.1	6.2	8.7	11.3	16.4	21.9	29.8	85.6	
$\rho=1, \gamma=1$	6.4	5.6	6.1	6.9	7.8	9.3	12.2	18.5	24.0	31.2	39.0	81.6	
vpMA	5.1	6.0	5.2	6.3	7.5	11.2	21.9	41.4	65.8	87.1	97.5	100.0	
vpIMI	5.5	5.3	5.9	7.4	10.7	20.2	37.9	59.1	80.8	94.4	98.9	100.0	
<b>vgRC</b>	5.9	6.6	11.5	22.9	40.6	62.6	84.1	94.0	98.9	99.7	100.0	99.3	Censure uniforme
vgOP	5.9	5.9	5.9	7.5	13.0	27.1	53.3	81.2	96.6	99.9	100.0	100.0	
vgRP	5.8	5.9	6.4	10.5	21.0	42.1	71.6	91.5	99.3	100.0	100.0	100.0	
vgRP <sub>Fisher</sub>	6.5	6.9	7.5	11.8	22.9	44.6	73.9	92.6	99.4	100.0	100.0	100.0	
LR adapté	5.7	5.9	6.2	10.1	20.2	40.9	70.5	91.1	99.2	100.0	100.0	100.0	
LR amélioré	0.6	0.6	1.0	1.9	6.3	18.9	44.7	76.7	95.9	99.8	100.0	100.0	
LRP SZZ <sub><math>\rho=0, \gamma=0</math></sub>	5.8	5.9	6.4	10.6	21.0	42.2	71.7	91.5	99.3	100.0	100.0	100.0	
$\rho=0, \gamma=1$	5.8	6.1	8.0	14.2	28.9	54.2	81.4	95.3	99.7	100.0	100.0	100.0	
$\rho=1, \gamma=0$	5.9	6.1	6.3	5.9	5.7	9.8	21.7	46.4	76.0	94.9	99.7	100.0	
$\rho=1, \gamma=1$	6.0	5.8	6.2	7.3	12.1	25.5	50.7	79.6	95.8	99.8	100.0	100.0	
LRP OG <sub><math>\rho=0, \gamma=1</math></sub>	5.6	6.1	8.3	14.9	30.4	56.0	82.3	95.6	99.7	100.0	100.0	100.0	
$\rho=1, \gamma=1$	5.7	5.8	7.0	11.3	22.1	44.7	74.3	92.8	99.4	100.0	100.0	100.0	
LRPHyb <sub><math>\rho=0, \gamma=0</math></sub>	6.2	5.8	6.0	6.5	7.0	6.9	9.5	11.5	14.4	18.1	23.7	73.7	
$\rho=0, \gamma=1$	5.8	6.0	6.8	6.8	8.8	8.5	10.3	13.3	15.1	19.1	22.4	56.7	
$\rho=1, \gamma=0$	5.9	5.3	5.1	5.8	5.6	5.9	6.6	7.7	12.5	16.0	23.6	86.8	
$\rho=1, \gamma=1$	5.3	5.6	5.3	5.0	6.2	6.9	7.7	10.8	15.0	21.5	28.5	82.3	
vpMA	5.4	5.6	5.7	5.8	8.4	15.7	29.1	51.2	76.4	93.9	99.1	100.0	
vpIMI	6.6	6.3	6.8	9.9	17.5	33.3	59.2	82.5	95.9	99.7	100.0	100.0	

TABLEAU 4.8 – **Modèle odds-proportionnel, risques accélérés**,  $n = 300, \pi = 0.5, p = 30\%$ . Les résultats, en %, correspondent aux risques de première espèce et puissances des tests comparés (en ligne) en fonction des valeurs du  $\beta$  (en colonne) comprises entre 0 et 3. Le tableau du haut correspond à une censure exponentielle et le tableau du bas à une censure uniforme.

**vgRC** — le **LRP OG** <sub>$\rho=0, \gamma=1$</sub>  — a une puissance de 48,5% avec  $\beta = 1$  alors que le score **vgRC** a une puissance de 94,9%.

*Différences du niveau de censure*

Dans les situations de censure importante ( $p = 60\%$ ), le **LRP SZZ** <sub>$\rho=1, \gamma=0$</sub>  est performant avec une puissance supérieure à 90% pour  $\beta > 1,4$ . Cependant, ces bonnes performances ne sont pas maintenues lorsque le taux de censure diminue. Cela peut s'expliquer par le fait que cette statistique est conçue pour mettre plus de poids sur les différences précoces entre les groupes qui ne seront pas annulées par des différences tardives du signe opposé en présence d'une censure

	Censure exponentielle		Censure uniforme	
	$\omega = 0.25$	0.75	$\omega = 0.25$	0.75
<b>vgRC</b>	5.3	4.9	5.1	4.2
vgOP	5.3	4.5	4.9	4.8
vgRP	5.7	5.0	4.8	4.8
vgRP <sub>Fisher</sub>	6.8	6.0	5.8	6.5
LR adapté	5.7	5.0	5.0	4.8
LR amélioré	0.1	0.1	0.8	0.6
LRP SZZ $_{\rho=0,\gamma=0}$	5.7	5.0	4.9	4.8
$\rho=0,\gamma=1$	5.5	4.9	4.7	4.8
$\rho=1,\gamma=0$	5.0	4.1	5.5	4.7
$\rho=1,\gamma=1$	5.3	4.8	5.4	5.0
LRP OG $_{\rho=0,\gamma=1}$	5.2	4.8	4.8	4.8
$\rho=1,\gamma=1$	5.7	4.9	4.8	5.0
LRPHyb $_{\rho=0,\gamma=0}$	7.5	5.9	6.3	6.8
$\rho=0,\gamma=1$	7.8	7.3	7.3	8.2
$\rho=1,\gamma=0$	5.8	8.2	5.5	5.4
$\rho=1,\gamma=1$	5.9	5.8	5.9	5.6
vpMA	23.3	94.9	14.2	76.9
vpIMI	35.9	99.7	20.6	85.8

TABLEAU 4.9 – Simulations sous modèle à risques croisés avec censure dépendante (ATD),  $n = 300, \pi = 0.5, p = 30\%$ . Les risques de première espèce affichés sont en %.

élevée. La même tendance peut être observée avec **vgOP**, mais avec un effet moindre.

Inversement, la puissance des logranks pondérés avec ( $\rho = 0, \gamma = 1$ ) — qui mettent plus de poids sur les événements tardifs — augmente tandis que le taux de censure diminue. Toutes ces tendances sont plus prononcées en choisissant des valeurs plus élevées pour les paramètres tels que ( $\rho > 2, \gamma = 0$ ) et ( $\rho = 0, \gamma > 2$ ) (résultats non montrés ici).

#### *Différence du type de censure*

En comparant les censures exponentielles et uniformes, nous notons que les tests ont une puissance plus élevée sous une censure uniforme sauf pour le **LRP SZZ** $_{\rho=1,\gamma=0}$ . Ces différences observées sont atténuées lorsque le taux de censure augmente.

En particulier, le test d'imputation du milieu de l'intervalle (**vpIMI**) est très sensible au type de censure; sa puissance est multipliée par près de 8 avec  $\beta = 1,4$  et  $p = 30\%$  lorsqu'on considère la censure uniforme plutôt qu'exponentielle. Le **vpIMI** où une distribution de censure uniforme est attendue montre des performances médiocres lorsque les temps de censure sont générés à partir d'un modèle différent tel qu'un modèle exponentiel. Une tendance similaire est observée pour les tests **vgRP**, **LR adapté** et **LRP SZZ** $_{\rho=0,\gamma=0}$ .

#### *Petite taille d'échantillon*

À partir de simulations effectuées avec une petite taille d'échantillon, nous notons que — sachant que la statistique des **LRHyb** nécessite l'estimation de la fonction de survie séparément dans les deux groupes — elle peut échouer dans de nombreux cas.

Il est à noter que la plupart de ces statistiques peuvent être considérées comme des statistiques linéaires de rang qui résument les contributions des observations censurées par intervalles, calculées sur les intervalles de TURNBULL [1976]. Par conséquent, pour certaines configurations (petite

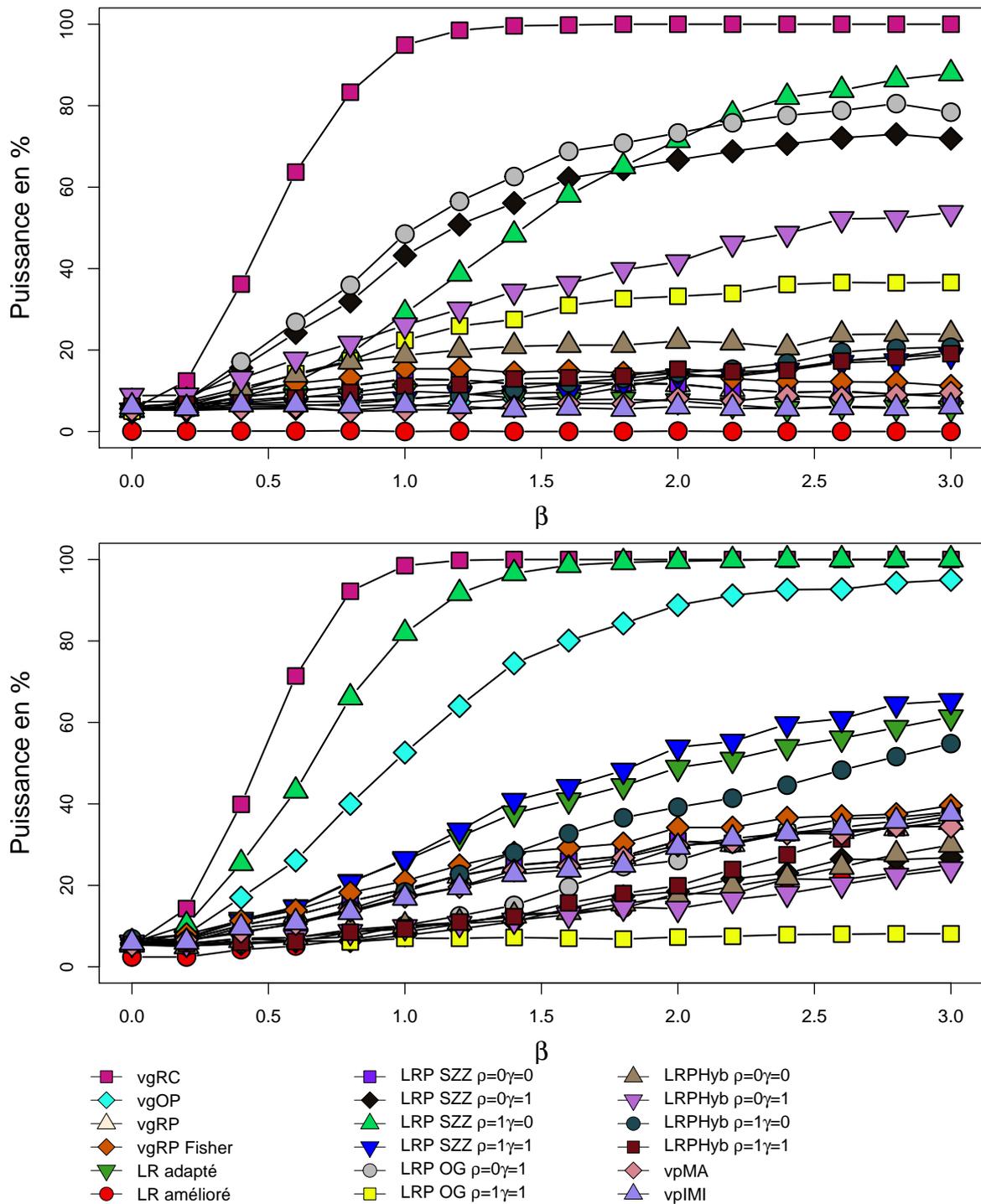


FIGURE 4.1 – Graphiques correspondant aux simulations sous modèle à risques croisés et censure exponentielle (tableau 4.5). Haut :  $p = 30\%$ . Bas :  $p = 60\%$

taille d'échantillon et faible censure), nous pouvons avoir un très petit nombre d'intervalles, ce qui mènera à des statistiques de test inefficaces.

Notons que, bien que le **vpMA** ne nécessite aucune estimation de la fonction survie puisqu'il repose sur le principe de la vraisemblance partielle, il montre une puissance statistique faible pour tous les scénarios.

Comme attendu, le **LRP SZZ** $_{\rho=0, \gamma=0}$  et le **vgRP** conduisent à des résultats très similaires quel que soit le scénario. En effet, ces deux tests ne diffèrent que par la façon dont la variance est calculée.

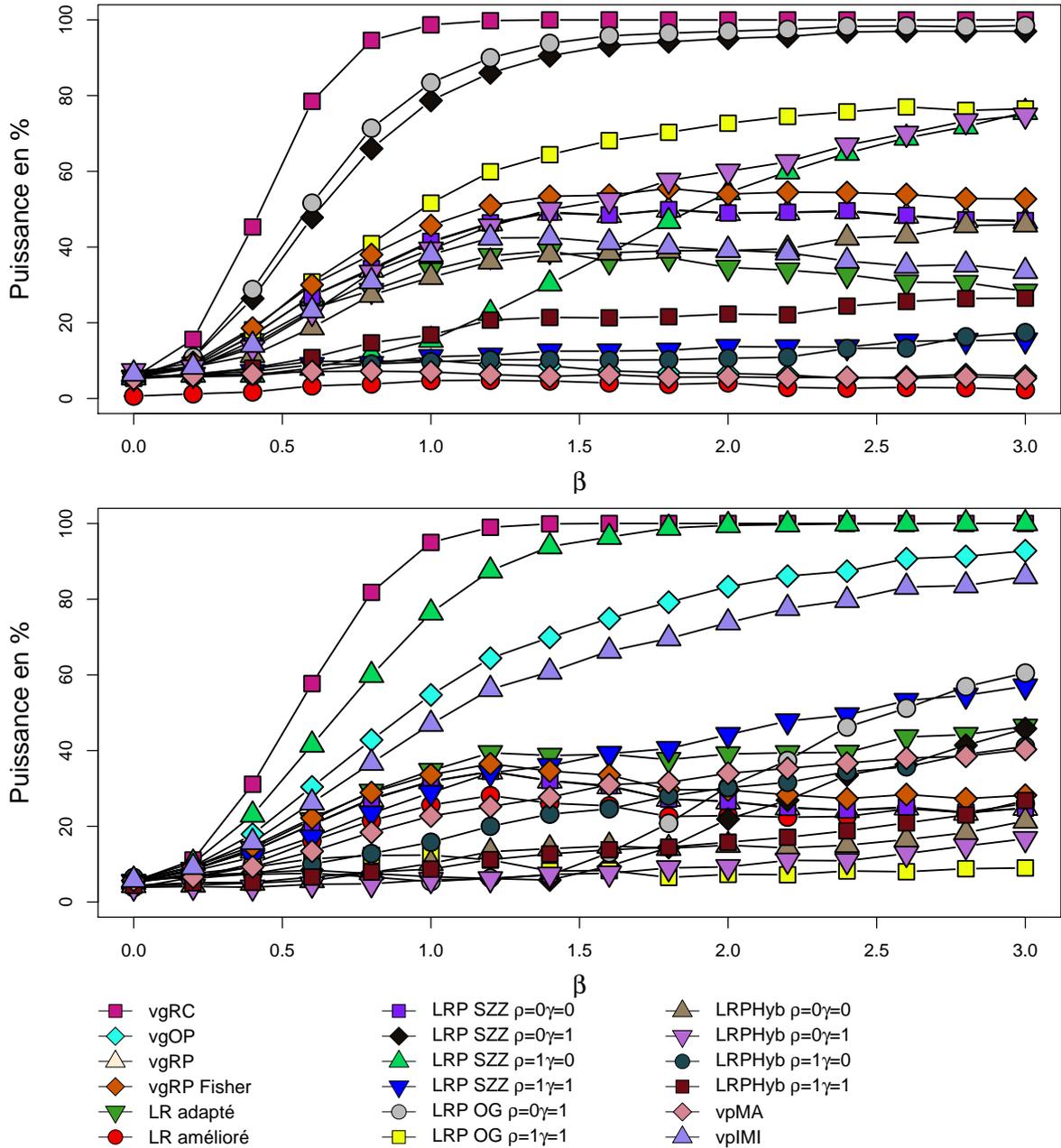


FIGURE 4.2 – Graphiques correspondant aux simulations sous modèle à risques croisés et censure uniforme (tableau 4.6). Haut :  $p = 30\%$ . Bas :  $p = 60\%$

## 2) Scénario 2

À partir des simulations sous le modèle log-logistique à risques accélérés (Tableau 4.8), nous notons que les risques de première espèce du score **vgRC** se situent dans l'intervalle de fluctuation du risque nominal (5%) d'une loi de  $\chi^2$  à un degré de liberté pour les censures exponentielle et uniforme. En ce qui concerne la puissance, le score **vgRC** présente encore de bonnes performances par rapport aux autres tests.

Nous avons des différences notables de puissance entre les simulations effectuées sous le 'bon' modèle (Weibull) et celles effectuées sous le modèle à risques accélérés. Certains tests ont de meilleures performances sous le modèle à risques accélérés (par ex. : **vgRP**, **Logrank adapté**, **LRP SZZ**  $\rho = 0, \gamma = 0$  et **vgRP<sub>Fisher</sub>** avec des gains de puissance allant jusqu'à 25% pour une même valeur de  $\beta$  avec une censure exponentielle) alors que d'autres affichent de meilleures performances sous le modèle de Weibull correspondant au scénario 1 (par ex. : les tests de logrank pondérés avec les

paramètres  $\rho = 0, \gamma = 1$  et une censure uniforme).

Notons que le **LRP OG** $_{\rho=0, \gamma=1}$  a des performances remarquables sous le modèle à risques accélérés autant avec une censure uniforme qu'exponentielle.

### 3) Scénario 3

À partir des simulations réalisées avec un mécanisme de censure dépendante (ATD) dans le Tableau 4.9, et même si la validité asymptotique du test de permutation est incorrecte sous cette dépendance, nous voyons que notre test a un risque de première espèce correct avec les censures exponentielle et uniforme.

Concernant les tests utilisés dans les autres schémas de simulation, à l'exception des **vpIMI** et **vpMA**, les tests ont une erreur de type I correcte. Comme on l'a vu précédemment sans ATD, les tests **LRPHyb** et **LR adapté** n'ont pas un taux d'erreur de type I correct.

Ces résultats concernant les tests de logrank pondéré pour censure par intervalle sont en accord avec ceux présentés par FAY et SHIH [2012]. Il est bien connu que le test d'imputation du milieu de l'intervalle est très sensible à la censure dépendante (LAW et BROOKMEYER [1992]) et le **vpMA** doit être modifié pour prendre en compte la dépendance de censure comme indiqué par LIN et al. [1998].

## IV. Discussion

Comme le montrent nos études de simulations, le test proposé **vgRC** a une erreur de type I contrôlée au seuil choisi pour toutes les configurations étudiées. En ce qui concerne les divers scénarios, le test proposé montre des gains de puissance notables par rapport aux tests classiques. Cela s'explique par le fait que pour les statistiques non pondérées du type «logrank», lorsque les risques se croisent, la perte de puissance est liée au fait que les différences précoces et tardives tendent à s'annuler mutuellement. La situation est plus complexe pour les logrank pondérés. En effet, en fonction du pourcentage de censurés et de la distribution des temps de censure, les poids conduisent à des performances très variables. En résumé, le test proposé est à recommander pour les situations où l'on suspecte un croisement des risques.

Le test proposé est une procédure de pseudo-rangs adaptative dont les propriétés asymptotiques reposent sur l'hypothèse d'une censure égale. Nous avons également évalué la robustesse de la procédure sous une censure dépendante (ATD). Les simulations montrent que même avec une censure inégale, notre procédure conserve un risque de première espèce correct. Cette conclusion est conforme aux résultats obtenus par FAY et SHIH [2012] sur des simulations effectuées sous censure dépendante pour les tests de type logrank pondéré.

De plus, et même si notre test proposé repose sur un modèle de Weibull semi-paramétrique, les simulations montrent qu'il a de bonnes performances pour d'autres modèles de risques croisés. Néanmoins, notre test peut ne pas être efficace lorsque le modèle sous-jacent est incorrect. Il est à noter que dans ce travail, nous avons considéré des résultats asymptotiques pour calculer les  $p$ -values, mais ces dernières peuvent être obtenues soit par énumération complète soit par échantillonnage de Monte Carlo.

Nous avons également réalisé des simulations avec des données censurées par intervalle de type II que nous n'avons pas présentées dans ce document. Dans ce cas, le test développé, **vgRC**, avait un risque de première espèce appartenant à l'intervalle de fluctuation attendu pour une loi du  $\chi^2$  à un degré de liberté.

Des simulations effectuées sous un modèle à risque proportionnel et données censurées par intervalle de type I (non-présentées ici) montraient la faible puissance du test utilisant la moitié de l'intervalle comme pseudo-date d'événement. Par exemple, pour un  $\beta = 0,6$ , une censure de 30%, sous un modèle à risque proportionnel et une base de données *current status* avec  $n =$

300, le test du milieu de l'intervalle avait une puissance de 48,12% pour 5000 simulations tandis que le log-rank pour censure par intervalle de Finkelstein (test vgRP) avait une puissance égale à 89,38%. Nous recommandons l'usage de tests adaptés à la censure par intervalle pour comparer des groupes dans un échantillon, quelle que soit la forme des courbes de risque.

La non-proportionnalité du risque peut avoir plusieurs explications. Il peut s'agir de l'effet d'une interaction entre la variable étudiée et un facteur latent. ROUAM et al. [2011b] en détaille le principe général : soient deux variables,  $Y_1$  et  $Y_2$  qui ont un effet global sur le risque. La première variable (inobservée),  $Y_1$ , n'a d'effet que si la seconde variable (étudiée),  $Y_2$  a un effet. Tandis que la seconde variable,  $Y_2$ , étudiée seule, a un effet proportionnel sur le risque, l'effet marginal des deux variables sur le risque n'est pas proportionnel et résulte en un croisement des courbes de risque.

Un exemple de ce mécanisme serait d'étudier l'effet du sexe (variable observée) sur la survenue d'un événement, tandis que cette variable serait liée à la production d'une certaine hormone (variable inobservée).

La non-proportionnalité des fonctions de risque peut également traduire un effet de l'hétérogénéité de la population qui comprend une proportion de sujets non-susceptibles. Ce cas sera discuté dans la partie suivante.

En conclusion, le test proposé est applicable à une situation de plus de deux groupes. Il est simple à mettre en œuvre à partir des logiciels statistiques existants. De plus, il peut être facilement étendu pour prendre en compte une variable de stratification.

## V. Références

- FAY, M. P. 1996, «Rank invariant tests for interval censored data under the grouped continuous model», *Biometrics*, vol. 52, n° 3, p. 811–822, ISSN 0006-341X. [76](#), [77](#), [78](#)
- FAY, M. P. 1999, «Comparing several score tests for interval censored data», *Statistics in Medicine*, vol. 18, n° 3, p. 273–285, ISSN 0277-6715. [75](#), [76](#)
- FAY, M. P. et P. A. SHAW. 2010, «Exact and Asymptotic Weighted Logrank Tests for Interval Censored Data : The interval *R* Package», *Journal of Statistical Software*, vol. 36, n° 2, doi :10.18637/jss.v036.i02, ISSN 1548-7660. URL <http://www.jstatsoft.org/v36/i02/>. [79](#)
- FAY, M. P. et J. H. SHIH. 2012, «Weighted logrank tests for interval censored data when assessment times depend on treatment», *Statistics in Medicine*, vol. 31, n° 28, doi :10.1002/sim.5447, p. 3760–3772, ISSN 1097-0258. [73](#), [82](#), [90](#)
- FINKELSTEIN, D. M., W. B. GOGGINS et D. A. SCHOENFELD. 2002, «Analysis of failure time data with dependent interval censoring», *Biometrics*, vol. 58, n° 2, p. 298–304. [73](#)
- FINKELSTEIN, D. M. et R. A. WOLFE. 1985, «A Semiparametric Model for Regression Analysis of Interval-Censored Failure Time Data», *Biometrics*, vol. 41, n° 4, doi :10.2307/2530965, p. 933, ISSN 0006341X. URL <http://www.jstor.org/stable/2530965?origin=crossref>. [78](#)
- HÁJEK, J., Z. SIDÁK et P. K. SEN. 1999, *Theory of rank tests*, 2<sup>e</sup> éd., Probability and mathematical statistics, Academic Press, San Diego, Calif., ISBN 978-0-12-642350-1. OCLC : 246275665. [76](#), [77](#), [78](#)
- HSIEH, F. 2001, «On heteroscedastic hazards regression models : theory and application», *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, vol. 63, n° 1, doi : 10.1111/1467-9868.00276, p. 63–79, ISSN 1369-7412, 1467-9868. URL <http://doi.wiley.com/10.1111/1467-9868.00276>. [73](#)
- JONAS, S. F., C. MBOGNING et P. BROËT. 2017, «A test for comparing current status survival data with crossing hazard functions and its application to immunogenicity of biotherapeutics», *Statistics in Medicine*, vol. 36, n° 27, doi :10.1002/sim.7439, p. 4364–4377, ISSN 1097-0258. URL <http://dx.doi.org/10.1002/sim.7439>. [72](#)
- KALBFLEISCH, J. D. et R. L. PRENTICE. 2002, *The statistical analysis of failure time data*, 2<sup>e</sup> éd., Wiley series in probability and statistics, Wiley, Hoboken, NJ, ISBN 978-0-471-36357-6. [72](#), [76](#)
- LAW, C. G. et R. BROOKMEYER. 1992, «Effects of mid-point imputation on the analysis of doubly censored data», *Statistics in Medicine*, vol. 11, n° 12, p. 1569–1578, ISSN 0277-6715. [90](#)
- LAWLESS, J. F. 2002, *Statistical Models and Methods for Lifetime Data : Lawless/Statistical*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ, USA, ISBN 978-1-118-03300-5 978-0-471-37215-8. URL <http://doi.wiley.com/10.1002/9781118033005>, DOI : 10.1002/9781118033005. [73](#)
- LEEUW, J. D., K. HORNIK et P. MAIR. 2009, «Isotone Optimization in *R* : Pool-Adjacent-Violators Algorithm (PAVA) and Active Set Methods», *Journal of Statistical Software*, vol. 32, n° 5, doi : 10.18637/jss.v032.i05, ISSN 1548-7660. URL <http://www.jstatsoft.org/v32/i05/>. [82](#)
- LIN, D. Y., D. OAKES et Z. YING. 1998, «Additive hazards regression with current status data», *Biometrika*, vol. 85, n° 2, doi :10.1093/biomet/85.2.289, p. 289–298. URL <http://biomet.oxfordjournals.org/content/85/2/289.abstract>. [90](#)

- MANTEL, N. et D. M. STABLEIN. 1988, «The Crossing Hazard Function Problem», *The Statistician*, vol. 37, n° 1, doi :10.2307/2348379, p. 59, ISSN 00390526. URL <http://www.jstor.org/stable/2348379?origin=crossref>. 72
- MOREAU, T., J. MACCARIO, J. LELLOUCH et C. HUBER. 1992, «Weighted log rank statistics for comparing two distributions», *Biometrika*, vol. 79, n° 1, doi :10.1093/biomet/79.1.195, p. 195–198, ISSN 0006-3444, 1464-3510. URL <http://biomet.oxfordjournals.org/cgi/doi/10.1093/biomet/79.1.195>. 73, 75
- OLLER, R. et G. GÓMEZ. 2012, «A generalized Fleming and Harrington's class of tests for interval-censored data», *Canadian Journal of Statistics*, vol. 40, n° 3, doi :10.1002/cjs.11139, p. 501–516, ISSN 03195724. URL <http://doi.wiley.com/10.1002/cjs.11139>. 75, 85
- PETO, R. et J. PETO. 1972, «Asymptotically Efficient Rank Invariant Test Procedures», *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, n° 2, doi :10.2307/2344317, p. 185, ISSN 00359238. URL <http://www.jstor.org/stable/10.2307/2344317?origin=crossref>. 77
- PRABHAKAR MURTHY, D. N., M. XIE et R. JIANG. 2003, dans *Weibull Models*, édité par W. A. Shewhart et S. S. Wilks, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ, USA, ISBN 978-0-471-36092-6 978-0-471-47326-8. URL <http://doi.wiley.com/10.1002/047147326X>, DOI : 10.1002/047147326X. 73
- PURI, M. et P. SEN. 1985, *Nonparametric methods in general linear models*, Wiley series in probability and mathematical statistics : Probability and mathematical statistics, Wiley, ISBN 9780471702276. URL <https://books.google.fr/books?id=0xSoAAAAIAAJ>. 76, 78
- ROBERTSON, T., F. T. WRIGHT et R. L. DYKSTRA. 1995, *Order restricted statistical inference*, repr éd., Wiley series in probability and mathematical statistics, Wiley, Chichester, ISBN 978-0-471-91787-8. OCLC : 258556159. 77
- ROUAM, S., T. MOREAU et P. BROËT. 2011a, «A pseudo-R2 measure for selecting genomic markers with crossing hazards functions», *BMC Medical Research Methodology*, vol. 11, n° 1, doi :10.1186/1471-2288-11-28, p. 28, ISSN 1471-2288. URL <http://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-11-28>. 72
- ROUAM, S., T. MOREAU et P. BROËT. 2011b, «A pseudo-r2measure for selecting genomic markers with crossing hazards functions», *BMC Medical Research Methodology*, vol. 11, n° 1, doi :10.1186/1471-2288-11-28. URL <https://doi.org/10.1186/1471-2288-11-28>. 91
- SEN, P. K. 2004, «Permutational Central Limit Theorems», dans *Encyclopedia of Statistical Sciences*, édité par S. Kotz, C. B. Read, N. Balakrishnan, B. Vidakovic et N. L. Johnson, John Wiley & Sons, Inc., Hoboken, NJ, USA, ISBN 978-0-471-66719-3. URL <http://doi.wiley.com/10.1002/0471667196.ess1951>. 77
- STABLEIN, D. M. et I. A. KOUTROUVELIS. 1985, «A Two-Sample Test Sensitive to Crossing Hazards in Uncensored and Singly Censored Data», *Biometrics*, vol. 41, n° 3, doi :10.2307/2531284, p. 643, ISSN 0006341X. URL <http://www.jstor.org/stable/2531284?origin=crossref>. 72
- SUN, J. 2006, *The Statistical Analysis of Interval-censored Failure Time Data*, Statistics for Biology and Health, Springer New York, New York, NY, ISBN 978-0-387-32905-5. URL <http://link.springer.com/10.1007/0-387-37119-2>. 73
- TURNBULL, B. W. 1976, «The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data», *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 38, n° 3, p. 290–295. URL <http://www.jstor.org/stable/2984980>. 77, 87
- ZHANG, J. et Y. PENG. 2009, «Crossing hazard functions in common survival models», *Statistics & Probability Letters*, vol. 79, n° 20, doi :10.1016/j.spl.2009.07.002, p. 2124–2130, ISSN 01677152. URL <http://linkinghub.elsevier.com/retrieve/pii/S0167715209002491>. 74, 81



# Chapitre 5

## Application sur données d'immunologie clinique censurées par intervalle

### Sommaire

---

<b>I. Les maladies inflammatoires chroniques de l'intestin</b> . . . . .	<b>96</b>
A. Localisation . . . . .	96
B. Symptômes . . . . .	96
C. Étiologie . . . . .	98
D. Incidence et prévalence . . . . .	99
E. Traitements . . . . .	101
<b>II. La cohorte MICI du projet ABIRISK</b> . . . . .	<b>102</b>
A. Mise en place de la cohorte . . . . .	102
B. Caractéristiques démographiques . . . . .	104
C. Caractéristiques cliniques . . . . .	105
<b>III. Base de données et méthodes</b> . . . . .	<b>106</b>
A. Description de la base de données . . . . .	106
B. Méthodes . . . . .	108
<b>IV. Résultats</b> . . . . .	<b>109</b>
A. Tests univariés avec le vgRC . . . . .	109
B. Modèle de régression multivarié . . . . .	112
<b>V. Discussion</b> . . . . .	<b>114</b>
<b>VI. Références</b> . . . . .	<b>116</b>

---

## Introduction

Nous avons présenté dans les chapitres précédents les tests déjà existants destinés à comparer des groupes de données censurées par intervalle ainsi qu'un nouveau test adapté à une situation de risques croisés. Nous allons à présent appliquer le test développé sur les données issues du consortium ABIRISK.

Ce projet européen avait pour objectif d'étudier l'immunogénicité de biothérapies (survenue d'anticorps anti-biothérapies, **ADA**) prescrites dans le cadre de certaines pathologies auto-immunes comme la sclérose en plaques ou les maladies inflammatoires de l'intestin. Le projet ABIRISK a été présenté plus en détails dans la première partie, dans la section III, page 63.

Parmi les pathologies incluses dans le projet ABIRISK, nous allons nous intéresser plus spécifiquement à la cohorte des patients traités pour les maladies inflammatoires de l'intestin. Plus homogène, cette cohorte a été mise en place dans un nombre de centre réduits et avec un nombre restreint de biothérapies. Cela nous permet ainsi de diminuer les facteurs de confusion associés à la survenue des ADA et d'identifier plus facilement les covariables ayant un effet significatif sur l'événement d'intérêt.

Après avoir exposé quelques éléments sur les maladies inflammatoires chroniques de l'intestin, puis présenté la cohorte, nous analyserons les données censurées par intervalle à l'aide du test développé dans le chapitre précédent pour identifier les covariables ayant un effet significatif sur la survenue des ADA

## I. Les maladies inflammatoires chroniques de l'intestin

Les maladies inflammatoires chroniques de l'intestin (**MICI**) regroupent deux pathologies : la maladie de Crohn et la rectocolite hémorragique, qui sont toutes deux caractérisées par une inflammation (partielle) de la paroi du tube digestif. Ces deux pathologies partagent des caractéristiques communes ; cependant, la localisation de la maladie, la morphologie et la physiopathologie sont clairement différentes entre elles.

### A. Localisation

La **maladie de Crohn** est une inflammation d'une partie du tube digestif schématisé en figure 5.1. Les lésions — segmentaires, asymétriques, souvent profondes — peuvent survenir à n'importe quel niveau du tube digestif, de la bouche à l'anus ; les parties les plus fréquemment atteintes étant situées sur l'intestin (iléon, côlon, zone ano-périnéale). Ces lésions peuvent parfois entraîner fissures et fistules.

La **rectocolite hémorragique** consiste en des inflammations situées au niveau du rectum et pouvant s'étendre dans tout le côlon. Dans ce cas, les lésions, continues, sont le plus fréquemment superficielles.

### B. Symptômes

Les maladies inflammatoires de l'intestin sont des maladies chroniques évoluant par poussées ; les phases de poussées — de durée et de fréquence très variables — alternant avec des phases de rémission. Il peut arriver que les premières poussées de la pathologie ne permettent pas de distinguer entre la maladie de Crohn et la rectocolite hémorragique. Cependant, chacune de ces pathologies a une évolution et des signes cliniques qui lui sont propres.

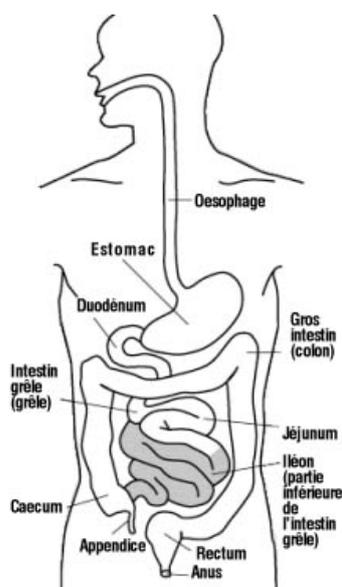


FIGURE 5.1 – Schéma du tube digestif

Crédits : [maladiedecrohn.eu](http://maladiedecrohn.eu)

### 1) Maladie de Crohn

La maladie de Crohn est associée aux symptômes suivants :

- douleurs abdominales dont la localisation la plus fréquente est le coin inférieur droit, accentuées après le repas ;
- diarrhée chronique, résultant des gonflements provoqués par l'inflammation ;
- fatigue, fièvre et perte de poids dû à la malabsorption ;
- saignements rectaux, anorexie (rarement).

Notons que les symptômes de la maladie de Crohn varient, selon la localisation. La localisation haute est par exemple caractérisée par des difficultés à avaler (dysphagie), des nausées, des aphtes, etc. De plus, le transit peut être normal si la pathologie n'atteint pas le côlon.

Hors du système digestif, la maladie de Crohn peut également causer des atteintes rénales, articulaires, cutanées, oculaires et hépatiques.

#### *Indice d'activité*

L'activité de la maladie de Crohn peut être évaluée à l'aide de deux scores. L'indice de BEST (ou CDAI pour *Crohn's Disease Activity Index*), un score essentiellement clinique et l'indice de Harvey-Bradshaw (HBI), corrélé au score CDAI. Ce second score est considéré comme plus aisément utilisable que le premier en pratique quotidienne, car les données sont recueillies lors des consultations.

Les deux scores évaluent la sévérité de la maladie suivant quatre stades : pathologie inactive (HBI < 4), poussées minimales ( $4 \leq \text{HBI} < 8$ ), poussées modérées ( $8 \leq \text{HBI} < 12$ ) et poussées sévères (HBI  $\geq 12$ ).

La maladie de Crohn est une entité hétérogène. La maladie peut également être caractérisée comme étant pénétrante (*penetrating*), sténosante<sup>1</sup> (*stricturing*) ou non-sténosante et non-pénétrante (c'est-à-dire inflammatoire) (*non-penetrating, non-stricturing*). C'est un phénotype cliniquement important, car il est associé à des complications et à un besoin éventuel de chirurgie. La maladie de Crohn évolue dans ces différents stades au cours de la maladie.

---

1. ici : rétrécissement de l'espace digestif.

Parmi les multiples sous-phénotypes proposés pour la maladie de Crohn, l'âge au diagnostic, la localisation et les caractéristiques pathologiques (y compris le développement de lésions pénétrantes ou sténosantes) ont été considérés comme pertinents et ont été inclus dans la classification de Vienne (1998) et dans celle de Montréal (2005) de la maladie de Crohn.

## 2) Rectocolite hémorragique

La rectocolite hémorragique est caractérisée par les symptômes suivants (certains, assez similaires à la maladie de Crohn) :

- selles sanglantes (rectorragies) ;
- diarrhée, douleurs abdominales ;
- lors des poussées sévères : hémorragies douloureuses, parfois accompagnées de glaires, amaigrissement, anémie, fatigue et fièvre ;
- brûlure rectales, coliques expulsives.

Tout comme la maladie de Crohn, la rectocolite hémorragique peut présenter des manifestations extra-intestinales et porter atteinte aux articulations, à la peau, au foie ou aux yeux.

### *Indice d'activité*

Le score de Mayo permet de classer la gravité de la poussée à l'aide d'appréciations cliniques et endoscopiques en quatre stades allant d'inactive ( $2 \leq$  score Mayo) à poussée sévère (score Mayo  $\geq 11$ ).

## C. Étiologie

À ce jour, les causes de la maladie de Crohn et de la rectocolite hémorragique sont encore partiellement inconnues. Ces pathologies auto-immunes — caractérisées par une inflammation chronique du tube digestif — pourraient s'expliquer comme une réponse anormale du système immunitaire aux composants de la flore intestinale. Les MICI sont depuis longtemps associées à des facteurs génétiques, mais elles impliquent probablement une réponse du système immunitaire à un ou plusieurs agents environnementaux ; elles sont de fait des pathologies multifactorielles.

### *Facteurs génétiques*

Plusieurs indices ont permis de mettre en évidence la susceptibilité génétique chez les malades. Tout d'abord, on retrouve un lien familial chez 5-10% des individus atteints de MICI. De plus, les études de jumeaux monozygotes ont permis de mettre en évidence un facteur génétique dans la maladie de Crohn avec 50-75% de concordance phénotypique dans les paires de jumeaux et dans la rectocolite hémorragique avec 10-20% de concordance ([KASER et al., 2010]). Le premier gène de susceptibilité aux MICI a été identifié en 2001 (NOD2/CARD15 sur le chromosome 16) ; depuis, les analyses du génome de personnes souffrantes de MICI ont montré le caractère polygénique de ces pathologies en identifiant plus de 150 gènes comme associés aux MICI avec, cependant, un impact modéré sur la survenue de ces pathologies non-héréditaires.

### *Facteurs environnementaux*

La discordance entre les MICI chez les jumeaux monozygotes, le développement des MICI chez les immigrants arrivant dans les pays à forte prévalence et l'augmentation de l'incidence dans les pays en cours d'occidentalisation met en exergue l'importance des facteurs environnementaux dans la pathogenèse de ces maladies. Le schéma 5.2 illustre l'influence de l'environnement sur le développement des MICI.

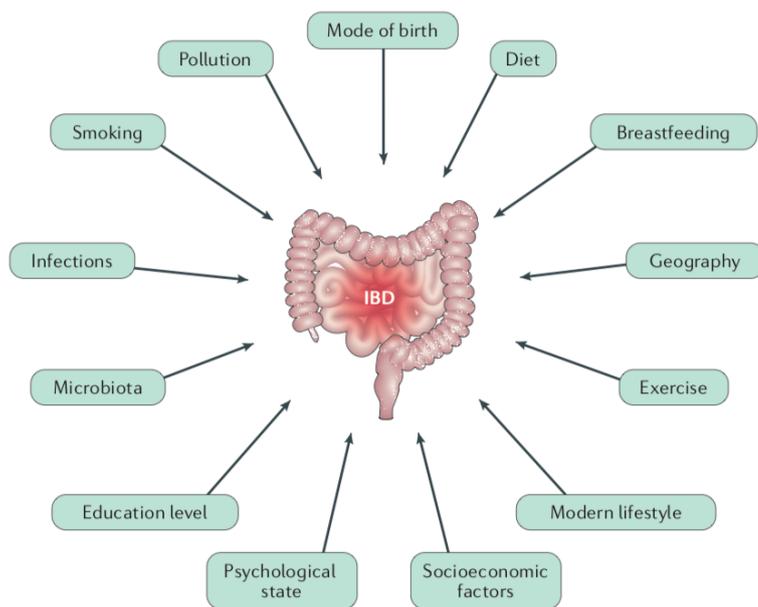


FIGURE 5.2 – **Facteurs environnementaux contribuant à la pathogenèse des MICI.** Les facteurs environnementaux tels que la conception, les influences maternelles sur le fœtus, le mode de naissance, la petite enfance, l'alimentation et les expositions diverses à l'âge adulte modifient le risque de développement de la maladie de Crohn et de la rectocolite hémorragique.  
ANANTHAKRISHNAN et al. [2017]

Les maladies inflammatoires de l'intestin sont ainsi la résultante de multiples facteurs tant génétiques qu'environnementaux; ces facteurs entraîneraient un dérèglement du microbiote intestinal causant une dysrégulation du système immunitaire.

La réponse immunitaire innée induit tout d'abord un ensemble de réactions conduisant à une perméabilité accrue de l'intestin vis-à-vis des bactéries se trouvant dans le tube digestif. Ces bactéries vont causer une inflammation, attirant de fait des cellules immunitaires spécialisées dans la paroi de l'intestin. Il y aura alors une réponse spécifique contre des composants du soi : la flore intestinale du patient, maintenant l'inflammation de l'intestin grêle et du colon. Le schéma 5.3 issu de l'article de KASER et al. [2010] met en évidence les nombreux éléments tant génétiques qu'environnementaux pouvant intervenir dans le développement des MICI.

Le développement des MICI est ainsi lié aux interactions (génétiques) entre le microbiote intestinal et le système immunitaire muqueux et la manière dont les facteurs environnementaux modifient ces relations. Il apparait important d'étudier les relations dynamiques de chacune de ces composantes et de les considérer dans leur totalité pour comprendre la pathogenèse de ces troubles.

#### D. Incidence et prévalence

Le patient type auquel est diagnostiqué une MICI est un jeune adulte, âgé de 20 à 30 ans. Cependant, les MICI peuvent être diagnostiqués à tout âge avec 15% des cas qui concernent des enfants et un pic d'incidence entre 60 et 70 ans.

La fréquence des MICI varie grandement d'un pays à l'autre. Les prévalences les plus importantes sont constatées dans les pays industrialisés (Europe du Nord et Ouest, États-Unis). Tandis que la prévalence reste stable dans ces zones, elle augmente très rapidement dans les pays en voie d'industrialisation comme les pays du Maghreb, en Asie, etc.

Le tableau 5.1 présente les taux d'incidence et de prévalence de la maladie de Crohn et de la rectocolite hémorragique en France. Ces chiffres sont issus de la cohorte EPIMAD, un registre français des maladies inflammatoires chroniques de l'intestin débuté en 1988 et décrit dans GOWER-ROUSSEAU et al. [2013], ainsi que de l'*afa*<sup>2</sup>.

2. *afa* (association François Aupetit) : organisation destinée au soutien aux malades atteints de maladies inflammatoires chroniques intestinales et à la recherche sur les MICI, [www.afa.asso.fr](http://www.afa.asso.fr)

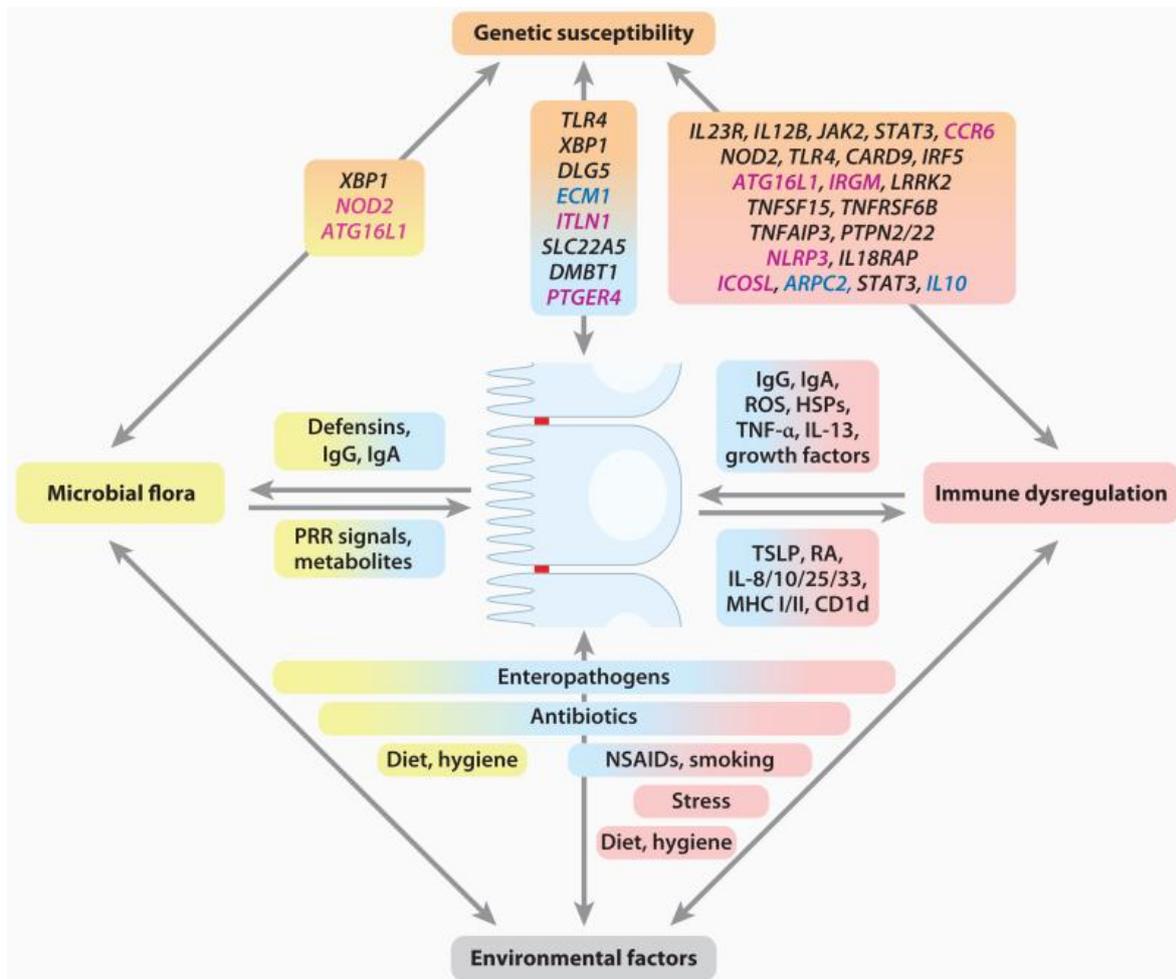


FIGURE 5.3 – **Maladies inflammatoires du colon et de l'intestin en tant que trouble multifactoriel.** Le développement et l'évolution des MICI sont influencés par plusieurs facteurs, y compris la susceptibilité génétique de l'hôte, le microbiote intestinal, d'autres facteurs environnementaux et le système immunitaire de l'hôte. De plus, ces facteurs se régulent l'un, l'autre de multiples façons, comme indiqué. Les gènes associés aux MICI sont organisés par voies moléculaires; des gènes appartenant à la même voie sont disposés les uns à côté des autres dans une ligne. Les polymorphismes des gènes spécifiques à la maladie de Crohn sont indiqués en magenta, alors que ceux spécifiques à la rectocolite hémorragique sont en bleu. Les associations génétiques partagées entre les deux maladies sont en noir. KASER et al. [2010]

Abréviations : HSPs, protéines de choc thermique; MHC, complexe majeur d'histocompatibilité; NSAIDs, médicaments anti-inflammatoires non stéroïdiens; PRR, récepteur de reconnaissance de motifs moléculaires; RA, acide rétinoïque; ROS, dérivés réactifs de l'oxygène; TSLP, lymphopoïétine stromale thymique.

Pathologie	Prévalence	Incidence (/an)
Maladie de Crohn	75 403	7,6/100000
Rectocolite hémorragique	61 175	4,4/100000

TABEAU 5.1 – Les chiffres des maladies inflammatoires du côlon et de l'intestin en France

Notons que le taux d'incidence et la prévalence varient grandement d'une catégorie d'âge à l'autre, d'une localisation à l'autre et d'une population à l'autre. Par exemple, la prévalence de la maladie de Crohn chez les juifs ashkénaze est 2 à 4 fois plus importante que dans la population générale (KENNY et al. [2012]). Les fumeurs ont près de deux fois plus de risque de développer la maladie de Crohn (MATSUOKA et al. [2018]).

En France, les MICI touchent un peu plus les femmes (56%) que les hommes (44%). La maladie de Crohn affecte 59 % de femmes et 41 % d'hommes tandis que la rectocolite hémorragique affecte les hommes et les femmes de manière égales (respectivement, 51% et 49%).

### E. Traitements

Il n'existe actuellement pas de traitement permettant de guérir définitivement de la maladie de Crohn ou de la rectocolite hémorragique. Les médicaments actuels visent un contrôle durable de la maladie et à améliorer la qualité de vie des patients pendant et en dehors des poussées. La table-  
bleau 5.2, basé sur une méta-analyse conduite par MATSUOKA et al. [2018], résume les traitements donnés pour les MICI et les indications de ces derniers.

Les patients peuvent également être amenés à suivre une thérapie nutritionnelle pour la maladie de Crohn uniquement. La nutrition entérale pourrait induire une rémission chez les patients atteints de MC active. Et le régime élémentaire, dans lequel les aliments sont donnés sous une forme liquide, est efficace pour maintenir la rémission dans la MC.

Enfin, la chirurgie peut améliorer le pronostic vital des patients atteints des formes les plus sévères de la maladie ou d'un cancer coexistant. En outre, elle peut également fournir une meilleure qualité de vie pour les patients qui souffrent d'effets secondaires de médicaments ou de manifestations extra-intestinales. Cependant, les interventions chirurgicales peuvent entraîner des complications postopératoires telles qu'une fistule anastomotique<sup>3</sup> et une occlusion intestinale chez les patients atteints de RCH et une insuffisance intestinale chez les patients atteints de la MC.

En conclusion, il existe de nombreux traitements pour les MICI, tous symptomatiques. Le traitement prescrit sera adapté à la forme de la maladie (légèrement à modérément active, extrêmement active, en rémission, MC avec complications intestinales), à l'historique médical du patient et à la manière dont le patient réagit au traitement.

3. Une fistule anastomotique résulte de la mauvaise cicatrisation de l'anastomose, soit la suture entre les deux parties du côlon.

Traitement	Indications
5-aminosalicylés (5-ASA)	<ul style="list-style-type: none"> <li>— Induction et maintien de la rémission (RCH);</li> <li>— Réduction de l'activité d'une MC active.</li> </ul>
Corticostéroïdes	<ul style="list-style-type: none"> <li>— Propriétés anti-inflammatoires puissantes permettant l'induction de la rémission dans la RCH et la MC;</li> <li>— Effets secondaires à moyen et long-terme.</li> </ul>
Immunosuppresseurs Azathioprine / 6-mercaptopurine / Methotrexate	<ul style="list-style-type: none"> <li>— Maintien de la rémission (RCH et MC);</li> <li>— Peut être combiné avec de l'infliximab (bio-thérapie);</li> <li>— Effets secondaires : symptômes gastro-intestinaux (nausée, pancréatites, etc.);</li> <li>— Autres immunomodulateurs : Tacrolimus (efficacité moins avérée), Cyclosporine (pour les formes les plus sévères) pour induction de rémission (RCH).</li> </ul>
Anti-TNF- $\alpha$ Infliximab / Adalimumab	<ul style="list-style-type: none"> <li>— Induction de la rémission pour RCH modérée à sévère réfractaire aux stéroïdes;</li> <li>— Induction et maintien de la rémission pour MC avec inflammation active.</li> </ul>
Antibiotiques et Probiotiques	<ul style="list-style-type: none"> <li>— (Possiblement efficaces) Induction de la rémission et réduction de l'écoulement de la fistule anale dans la MC;</li> <li>— (Possiblement efficaces) Induction de la rémission de la RCH sans que le type et la durée des antibiotiques ne fassent l'objet d'un consensus.</li> </ul>

TABLEAU 5.2 – Traitements et indications administrés en cas de MICI  
RCH : rectocolite hémorragique, MC : maladie de Crohn, TNF : *tumor necrosing factor*

## II. La cohorte MICI du projet ABIRISK

L'analyse statistique portera sur une base de données, prospective, constituée uniquement de patients atteints de maladies inflammatoires de l'intestin. Cette section présente les caractéristiques de cette cohorte.

### A. Mise en place de la cohorte

Dans cette cohorte, les patients "bio-thérapeutiquement naïfs" (c'est-à-dire, n'ayant pas été traités précédemment par la thérapie prescrite dans le cadre de l'étude) étaient traités par anti-TNF- $\alpha$ , infliximab et adalimumab. Le choix de l'une ou l'autre de ces biothérapies était laissé au praticien.

Ces biothérapies peuvent provoquer une réaction immunitaire du patient, qui se met à développer des anticorps-anti-biothérapies pouvant causer un échec thérapeutique. L'un des objectifs du projet était d'identifier des biomarqueurs précoces capables de prédire l'immunisation contre la biothérapie, au cours de la première année de traitement.

Le graphique 5.4 résume le schéma des visites; le nombre de visite variait, selon le type de biothérapies, infliximab ou adalimumab, prise par le patient.

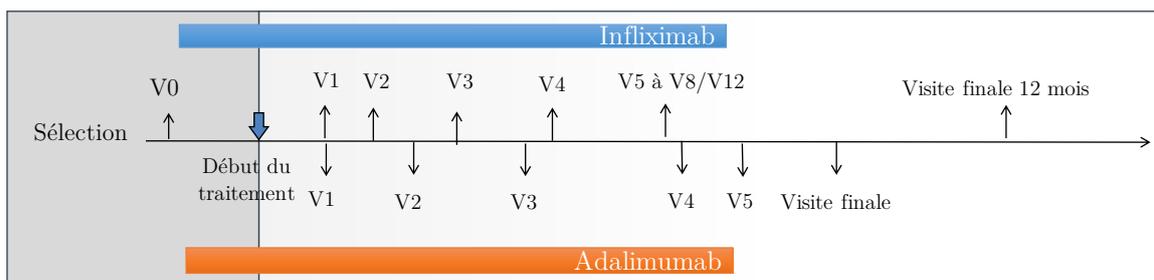


FIGURE 5.4 – Schéma des visites dans la cohorte prospective MICI du projet ABIRISK

Des données démographiques, épidémiologiques et cliniques ont été recueillies par des infirmières et des médecins au cours des visites de l'étude et consignées dans un formulaire électronique. Le contrôle qualité des données a été réalisé localement par les responsables de la collecte des données et supervisé par GETAID<sup>4</sup>. Une partie du travail a été coordonné par Sanofi.

À la fin du contrôle qualité, suivant les procédures standardisées dans le consortium et selon une terminologie internationale (CDISC), les données de l'étude ont été rassemblées dans une base de données unique (TranSMART) hébergée par le consortium eTRIKs et complétées avec les résultats des tests ADA fournis par les laboratoires centraux ABIRISK.

Le laboratoire ayant testé la présence d'ADA était l'UMR<sup>5</sup> 996 (INSERM) de la Faculté de Pharmacie, Université Paris-Saclay, France. La méthode de dosage utilisée pour détecter la positivité aux ADA était la méthode MSD (Meso Scale Discovery). Cette méthode utilise des anticorps de détection (par exemple, une molécule de médicament) qui vont se lier aux ADA puis être détectés par électrochimiluminescence. Cette technique consiste en l'emploi d'un luminophore comme le  $[Ru(bpy)_3]^{2+}$ , une substance qui émet de la lumière après avoir subi une excitation. La réaction entre l'anticorps et sa molécule vont produire des électrons qui exciteront le luminophore. La lumière émise est ensuite mesurée et permet de doser la quantité d'ADA.

Le tableau 5.3 résume les différents éléments de la cohorte :

Pays participants :	Belgique, France, Israël
Nombre de sites recruteurs :	14
Nombre de sujets recrutés :	204
Nombre de visites :	11 à 14 visites (Infliximab), 7 visites (Adalimumab)
Durée du suivi (prélèvements et test ADA) :	12 mois
Dates de l'étude :	2013 à 2017
Biothérapies :	Infliximab (ou similaires), Adalimumab

TABLEAU 5.3 – Cohorte MICI dans le projet ABIRISK

Les variables collectées incluaient : âge, durée de la maladie, IMC, sexe, statut tabagique ainsi que durée et intensité, antécédents médicaux et chirurgicaux, antécédents familiaux de maladie, vaccinations au cours de l'année précédant et pendant l'étude, traitements concomitants, traitements antérieurs, événements indésirables au cours de l'étude. Des variables spécifiques aux MICI telles que : score Mayo, HBI, localisation de la pathologie, etc. ont également été collectées.

4. GETAID (Groupe d'Étude Thérapeutique des Affections Inflammatoires du Tube Digestif) : association regroupant des centres universitaires français, belges et suisses, spécialisés dans la prise en charge de la maladie de Crohn et de la rectocolite hémorragique.

5. UMR : unité mixte de recherche

## B. Caractéristiques démographiques

La figure 5.5 récapitule le schéma d'inclusion des patients. Les données de 184 patients ont été utilisées dans l'analyse : 98 patients étaient traités par adalimumab et 86 patients, par infliximab.

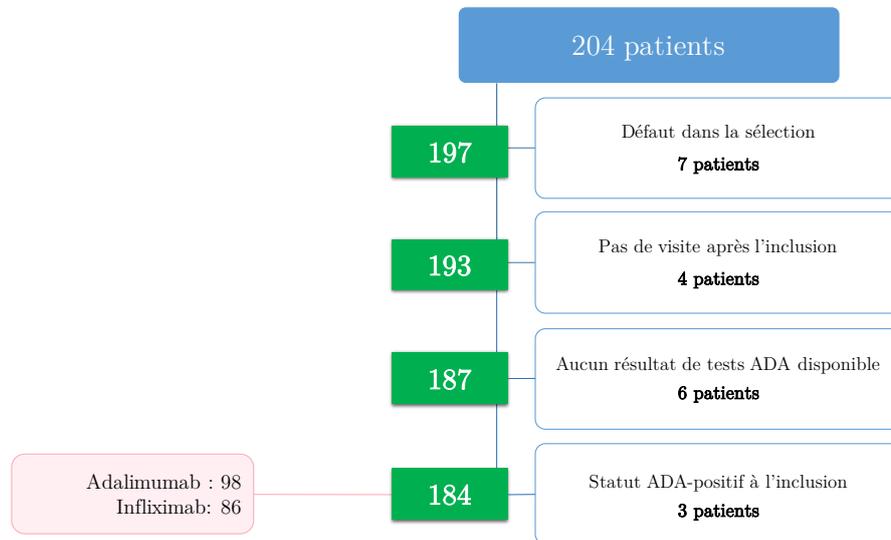


FIGURE 5.5 – Inclusion des patients dans la cohorte prospective MICI du projet ABIRISK

Le tableau 5.4 résume les caractéristiques démographiques des patients inclus dans la cohorte MICI, stratifiés par type de biothérapie. Le type de biothérapie est l'un des facteurs principaux liés à l'apparition des ADA. Dans les analyses suivantes, ce facteur sera pris en compte.

Caractéristiques	Adalimumab (n=98)	Infliximab (n=86)	Total (n=184)	p-value
Âge, moyenne (ET), ans	38,09 (14,08)	35,57 (13,25)	36,91 (13,72)	0,22
Femme, n (%)	47 (48,0)	42 (48,8)	89 (48,4)	1
Tabac, n (%)				0,14
Jamais fumé	37 (37,8)	42 (49,4)	79 (43,2)	
Ex-fumeur	30 (30,6)	16 (18,8)	46 (25,1)	
Fumeur	31 (31,6)	27 (31,8)	58 (31,7)	
IMC, moyenne (ET), kg/m <sup>2</sup>	24,00 (4,77)	22,17 (3,43)	23,14 (4,29)	0,004
Pays, n (%)				0,53
Belgique	4 (4,1)	2 (2,3)	6 (3,3)	
France	77 (78,6)	73 (84,9)	150 (81,5)	
Israël	17 (17,3)	11 (12,8)	28 (15,2)	
Suivi, médiane (EI), jours	357,5 (148,00)	360,0 (156,75)	357,5 (158,25)	0,45

TABEAU 5.4 – Caractéristiques démographiques des patients MICI, stratifiées par biothérapie. ET : écart-type, IMC : indice masse corporelle, EI : écart inter-quartile

Hormis l'IMC à l'entrée de l'étude, il n'y a pas de différence significative pour les caractéristiques démographiques entre les deux biothérapies.

Bien qu'aucun consensus n'ait été établi, les praticiens prescrivent de l'infliximab plutôt que de l'adalimumab pour les formes les plus sévères de la maladie. Cela explique la *p-value* significative associée à l'IMC entre les deux traitements. Les patients traités par infliximab ont un IMC plus faible, traduisant une forme plus sévère de la maladie.

### C. Caractéristiques cliniques

Le tableau 5.5 reprend les caractéristiques cliniques de la cohorte.

<b>Caractéristiques</b>	<b>Adalimumab</b> (n =98)	<b>Infliximab</b> (n =86)	<b>Total</b> (n=184)	<b>p-value</b>
Durée maladie, médiane (EI), ans	3,21 (8,85)	3,82 (12,01)	3,69 (10,37)	0,69
Type de maladie, n (%)				0,17
Crohn	83 (84,7)	65 (75,6)	148 (80,4)	
Rectocolite hémorragique	15 (15,3)	21 (24,4)	36 (19,6)	

TABEAU 5.5 – Caractéristiques cliniques des patients MICI, stratifiées par biothérapie. EI : écart inter-quartile

Le tableau 5.6 présente les variables spécifiques associées à chaque maladie : la maladie de Crohn ou la rectocolite hémorragique, stratifiées par traitement.

<b>Caractéristiques</b>	<b>Adalimumab</b> (n =98)	<b>Infliximab</b> (n =86)	<b>Total</b> (n=184)	<b>p-value</b>
<b>Rectocolite hémorragique</b>				
Mayo score, médiane (EI)	5,00 (5,50)	6,00 (4,00)	6,00 (5,00)	0,58
<b>Maladie de Crohn</b>				
HBI Index, médiane (EI)	5,00 (4,00)	5,00 (4,25)	6,00 (5,00)	0,43
Localisation, n (%)				0,15
Colique	11 (13,3)	16 (25,0)	27 (18,4)	
Iléale	42 (50,6)	24 (37,5)	66 (44,9)	
Iléocolique	30 (36,1)	23 (35,9)	53 (36,1)	
Localisation supérieure isolée	0 (0,0)	1 (1,6)	1 (0,6)	
Caractéristique de la maladie, n (%)				0,04
Non-sténosante, non-pénétrante	45 (54,2)	23 (35,9)	68 (46,3)	
Strénosante	27 (32,5)	24 (37,5)	51 (34,7)	
Pénétrante	11 (13,3)	17 (26,6)	28 (19,0)	

TABEAU 5.6 – Caractéristiques cliniques de la maladie de Crohn et de la rectocolite hémorragique des patients de la cohorte MICI, stratifiées par biothérapie. EI : écart inter-quartile

Excepté pour les caractéristiques de la maladie de Crohn, il n'y a pas de différence significative pour les caractéristiques cliniques entre les deux biothérapies.

La *p-value* significative associée aux caractéristiques de la maladie de Crohn confirme le fait que les patients atteints d'une forme plus sévère de la pathologie sont plutôt traités par infliximab : les patients prenant de l'adalimumab étant majoritairement atteints par la forme non-sténosante et non-pénétrante et les patients prenant de l'infliximab étant majoritairement atteints par la forme sténosante ou pénétrante de la maladie de Crohn.

#### *Précisions sur les traitements*

Les 98 patients étant traités par adalimumab recevaient une dose de 160 mg administrée en sous-cutanée toutes les deux semaines. Parmi les 86 patients traités par infliximab, 46 (54,1%) recevaient une dose comprise entre 200 mg et 320 mg, 13 (15,3%) recevaient une dose comprise entre 320 mg et 375 mg et 26 (30,6%) recevaient une dose comprise entre 375 mg et 600 mg. L'infliximab était administré en intra-veineuse.

#### *Description des ADA*

Le tableau 5.7 présente les caractéristiques de la cohorte en terme de développement des anticorps anti-biothérapies. 27,7% des patients de la cohorte MICI ont développé des ADA.

	Adalimumab (n =98)	Infliximab (n =86)	Total (n=184)
<b>ADA positif, n (%)</b>			
Cohorte entière	38 (38,8)	13 (15,1)	51 (27,7)
<i>Par maladie :</i>			
Crohn	32 (38,6)	11 (16,9)	43 (29,1)
Rectocolite hémorragique	11 (16,2)	2 (9,5)	8 (22,1)

TABLEAU 5.7 – Développement des ADA dans la cohorte MICI, stratifié par biothérapie.

### III. Base de données et méthodes

#### A. Description de la base de données

La base de données était constituée de 120 variables catégorielles et 21 variables quantitatives. Les variables quantitatives ont été catégorisées pour pouvoir être analysées avec un test univarié de comparaison de groupes. Les variables ont été collectées afin d'identifier celles ayant un effet sur les apparitions d'anticorps anti-biothérapies.

La base de données comprenait 12 variables non-spécifiques (âge, IMC, sexe, statut fumeur, pays, latitude, longitude) et 5 variables spécifiques portant sur la pathologie et le traitement (type de MICI, durée de la pathologie, manifestations extra-intestinales, type de biothérapie). Parmi ces variables, les variables quantitatives au nombre de 9, ont été catégorisées en une ou deux versions permettant la création de 11 variables qualitatives.

En plus de ces variables, la base de données comprenait 80 variables portant sur les traitements pris concomitamment ou récemment avant le début de la thérapie et des variables portant sur les pathologies concomitantes ou constatées après le début de la thérapie, ainsi que des informations sur la vaccination récente. Plus de détails sur ces variables seront donnés ultérieurement.

La base de données incluait 5 variables biologiques non-spécifiques portant sur les niveaux à l'entrée, relevés chez les patients (albumine, hémoglobine, leucocytes, lymphocytes, neutrophiles). Chacune de ces variables existait aussi en version catégorielle. Le niveau de la biothérapie a également été relevé à la sixième semaine après le début de l'étude pour les deux traitements; cette variable quantitative a été catégorisée.

Il y avait 4 variables portant sur la grossesse et l'allaitement et le temps écoulé depuis l'un ou l'autre de ces éléments. Les 2 variables quantitatives ont été catégorisées.

Onze variables ne concernaient que l'une ou l'autre des MICI (score Mayo, HBI, localisation et caractéristiques de la maladie). Les deux variables scores (Mayo et HBI) existaient en version catégorielle.

Le tableau 5.8 résume les variables constituant la base de données.

Type de variables	Quantitatives	Qualitatives
<i>Variables générales</i>		
Classiques : sexe, âge, etc.	8	14
Pathologie et traitement	1	5
Traitements et vaccins concomitants	1	39
Traitements et vaccins avant le début de l'étude	1	20
Pathologies concomitantes et effets indésirables	0	21
Mesures biologiques à la base	5	5
Niveau de la biothérapie (semaine 6)	1	1
<i>Variables spécifiques</i>		
Grossesse et allaitement	2	4
Maladie de Crohn	1	5
Rectocolite hémorragique	1	6
<b>TOTAL</b>	<b>21</b>	<b>120</b>

TABLEAU 5.8 – Variables constituant la base de données de la cohorte MICI, Abirisk

La base de données comprenait 184 patients. Pour chacun des patients, les informations suivantes étaient disponibles : son statut ADA (positif, négatif), un intervalle de dates de visite (s'il était ADA-positif) ou la date de sa dernière visite (s'il était ADA-négatif).

#### *Traitements concomitants et pris avant le début de l'étude*

Les médicaments antérieurs et concomitants, avec la date de début et de fin, la voie d'administration, la dose et l'indication ont été enregistrés à chaque visite d'étude à partir de l'inclusion. Les médicaments pris dans le mois précédent le début de la thérapie (ou dans les 3 mois précédents pour les immuno-suppresseurs) ont été inclus comme traitements antérieurs.

Les médicaments des différents pays ont été codés par Sanofi en utilisant le dictionnaire de classification internationale des médicaments de l'OMS (Organisation Mondiale de la Santé).

En raison de contraintes de temps et de ressources, il n'a pas été possible dans le projet ABI-RISK d'exploiter les données d'indication et d'administration pour attribuer chaque médicament à une catégorie unique. La classification ATC (anatomique, thérapeutique et chimique) de l'OMS a donc été utilisée. Elle est basée sur cinq niveaux de classification qui correspondent aux organes cibles, aux indications thérapeutiques et aux propriétés pharmacologiques et chimiques de la molécule. Par exemple, dans le premier niveau (ATC 1), il y a 14 groupes d'organes cibles

tels que "système digestif et métabolisme", "système nerveux", etc. Une molécule disposant de plusieurs voies d'administration ou ayant plusieurs indications principales appartient à plusieurs catégories. Ainsi, ici où les médicaments ont été classifiés, selon les deux premiers niveaux du dictionnaire de classification (ATC 1 et 2), la plupart des médicaments sont attribués à plusieurs catégories et il existe une très grande redondance entre les variables.

Enfin, une information sur les vaccins et sur celui de la grippe en particulier, effectués dans l'année précédente ou pendant l'étude, était disponible.

#### *Pathologies concomitantes et événements indésirables*

Les maladies concomitantes et antérieures ont été enregistrées. Une information supplémentaire sur l'historique familial était disponible permettant d'indiquer s'il existait une pathologie identique, s'il y avait du psoriasis ou des rhumatismes dans la famille.<sup>6</sup> Les effets indésirables très précoces (au premier mois de la thérapie) ont été enregistrés.

Les antécédents médicaux et les événements indésirables ont été codés par une CRO en utilisant le MedDRA (Medical Dictionary for Regulatory Activities), une terminologie médicale internationale. Ce dictionnaire comprend 5 niveaux dans la hiérarchie structurelle de la terminologie. Ici, le premier et le quatrième niveaux ont été utilisés.

Le premier niveau (SOC pour *system organ class*) comprend 27 classifications par discipline médicale. Une pathologie peut appartenir à plusieurs groupes comme les accidents vasculaires cérébraux qui relèvent à la fois des affections cardiaques et du système nerveux. Le quatrième niveau (PT pour *preferred term*) est beaucoup plus spécifique puisqu'il donne à chaque pathologie un descripteur distinct ; soit un concept médical unique pour un symptôme, un signe, une maladie, un diagnostic, une indication thérapeutique, une intervention chirurgicale ou médicale, ou un antécédent médical, social ou familial.

## **B. Méthodes**

### **1) Tests avec le vgRC**

Les 120 variables catégorielles des 184 patients ont été testées à l'aide du test que nous avons présenté au chapitre précédent et dénoté **vgRC**. Cette analyse avait pour but d'identifier les variables ayant un effet sur l'événement d'intérêt, soit le délai avant apparition des ADA.

#### *Stratification sur le type de traitement*

Le type de traitement (adalimumab ou infliximab) est très fortement lié au développement des ADA. La *p-value* du test du log-rank est  $1,19 \times 10^{-4}$  (valeur du score : 14,81). Le graphique 5.6 présente les courbes de survie de chaque traitement.

Or, il est admis que certains traitements sont plus immunogènes que d'autres. Dans cette cohorte, les patients traités par adalimumab ont trois fois plus de risque de développer des ADA rapport à ceux traités par infliximab (rapport de risque : 3,09 [1,65 ; 5,82]). Par la suite, les tests univariés ont été réalisés sur l'échantillon stratifié par traitement.

Pour prendre en compte la multiplicité des tests, nous avons considéré le FDR (*False Discovery Rate*) comme correction pour tests multiples avec une espérance du taux de faux positifs de moins de 25% (MCDONALD [2014]). L'estimation des *q-values* a été obtenue par la méthode LBE (*Location Based Estimator*) présentée par DALMASSO et al. [2004].

---

6. En plus de la cohorte MICI, le projet ABIRISK incluait une cohorte prospective de patients atteints de polyarthrite rhumatoïde ou de sclérose en plaques. Certaines variables, bien que renseignées pour les MICI, étaient plus pertinentes pour les deux autres cohortes.

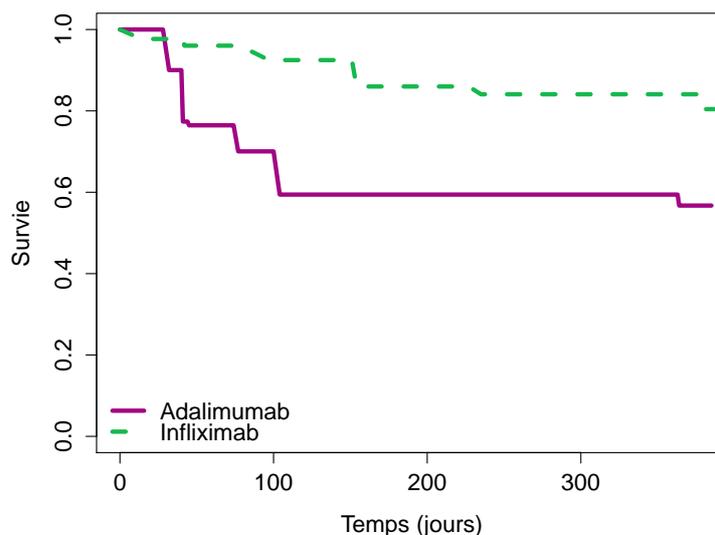


FIGURE 5.6 – Estimateur non-paramétrique du maximum de vraisemblance de la fonction de survie en fonction du traitement de la cohorte prospective MICI, ABIRISK.

## 2) Tests multivariés

Les analyses supplémentaires avaient pour but de mettre en place un modèle multivarié dans lequel les variables les plus pertinentes de l'étape précédente étaient incluses.

Notre travail n'ayant pas porté sur une extension multivariée, nous avons considéré ici une analyse multivariée avec un modèle de Cox restreint aux 7 variables à la fois détectées par le test vgRC et le test du log-rank pour données censurées par intervalle (FINKELSTEIN et WOLFE [1985]) et en prenant l'intervalle moyen. Une analyse complète et adéquate devrait envisager un nouveau type de modélisation adapté à la situation des risques croisés.<sup>7</sup>

Le schéma régulier des visites avec espacement faible, nous permet de raisonnablement supposer que la perte de puissance ne sera pas trop importante.

Les variables retenues ont été incluses dans un modèle de sélection *stepwise backward*.

## IV. Résultats

### A. Tests univariés avec le vgRC

#### 1) Variables associées significativement aux ADA

119 variables catégorielles ont été testées à l'aide du test vgRC de comparaison de groupes pour censure par intervalle, stratifié sur le traitement. L'événement d'intérêt était le délai avant apparition des ADA. Les *p-value* obtenues avec le vgRC ont ensuite été corrigées pour tests multiples. Nous avons préalablement retiré de l'analyse 15 variables dont l'effectif d'une catégorie représentait moins de 5% de l'échantillon total; dans ce cas, nous pouvions douter que les conditions asymptotiques soient remplies.

Le graphique 5.7 montre les *q-value* des 104 tests obtenues par la méthode LBE en fonction des *p-value*. Il apparaît qu'après correction, nombre de variables précédemment significatives ne le sont plus, selon le seuil choisi.

7. L'article de BOGAERTS et al. [2002] correspond au cas particulier où le modèle comprend une variable portant sur des données corrélées intra-sujet (symétrie de la croissance des dents chez l'enfant).

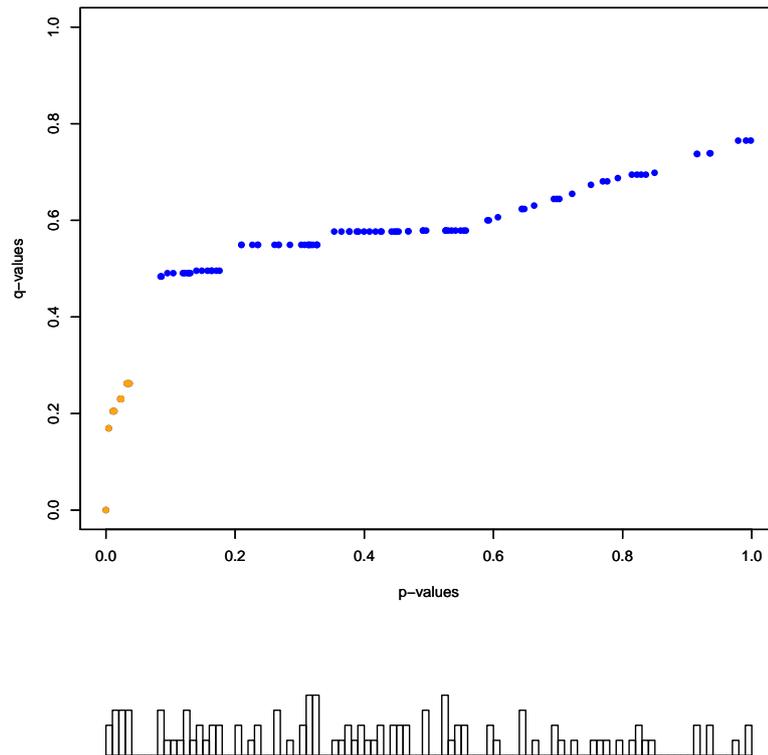


FIGURE 5.7 –  $p$ -value et  $q$ -value des variables testées pour leur association avec le développement des ADA, cohorte prospective MICI, ABIRISK.

En sélectionnant les tests avec une  $q$ -value  $< 0,25$ , 11 variables sont significativement associées au développement des ADA. Le tableau 5.9 présente les 11 variables sélectionnées.

Variabiles	$q$ -value
Niveau de biothérapie (6 semaines)	$1,5E^{-4}$
Immunomodulateur (pendant)	0,13
Allaitement*	0,18
Immunomodulateur 2 (pendant)	0,19
Traitement pour système nerveux (pendant)	0,19
Traitements pour système digestif et métabolique (avant)*	0,21
Années écoulées depuis l'allaitement*	0,21
Analgésiques (pendant)	0,21
Antidiarrhéiques, anti-inflammatoires & anti-infectieux intestinaux (avant) *	0,22
Immunosuppresseur (pendant)	0,22
Vaccins (dans l'année précédente)	0,22

TABLEAU 5.9 –  $q$ -value ( $<0,25$ ) des variables catégorielles de la cohorte MICI testées pour leur effet sur le développement des ADA. Les variables avec un astérisque sont uniquement détectées par le test adapté aux risques croisés.

## 2) Graphiques

Nous présentons ci-dessous, dans la Figure 5.8, les graphiques des estimateurs non-paramétriques du maximum de vraisemblance (NPMLE) de la fonction de survie du délai avant apparition des ADA en fonction des variables uniquement détectées par le test vgRC, soient : l'allaitement, les traitements pour le système digestif et métabolique pris avant le début de la biothérapie et les antidiarrhéiques, anti-inflammatoires et anti-infectieux intestinaux pris avant le début du traitement. La variable du délai écoulé depuis l'allaitement contient trop peu d'événements.

Nous présentons également le graphique des variables suivantes : niveau de biothérapie relevé 6 semaines après le début de la thérapie, vaccins effectués dans l'année précédant le début du traitement et immunosuppresseur pris pendant le traitement (Figure 5.8).

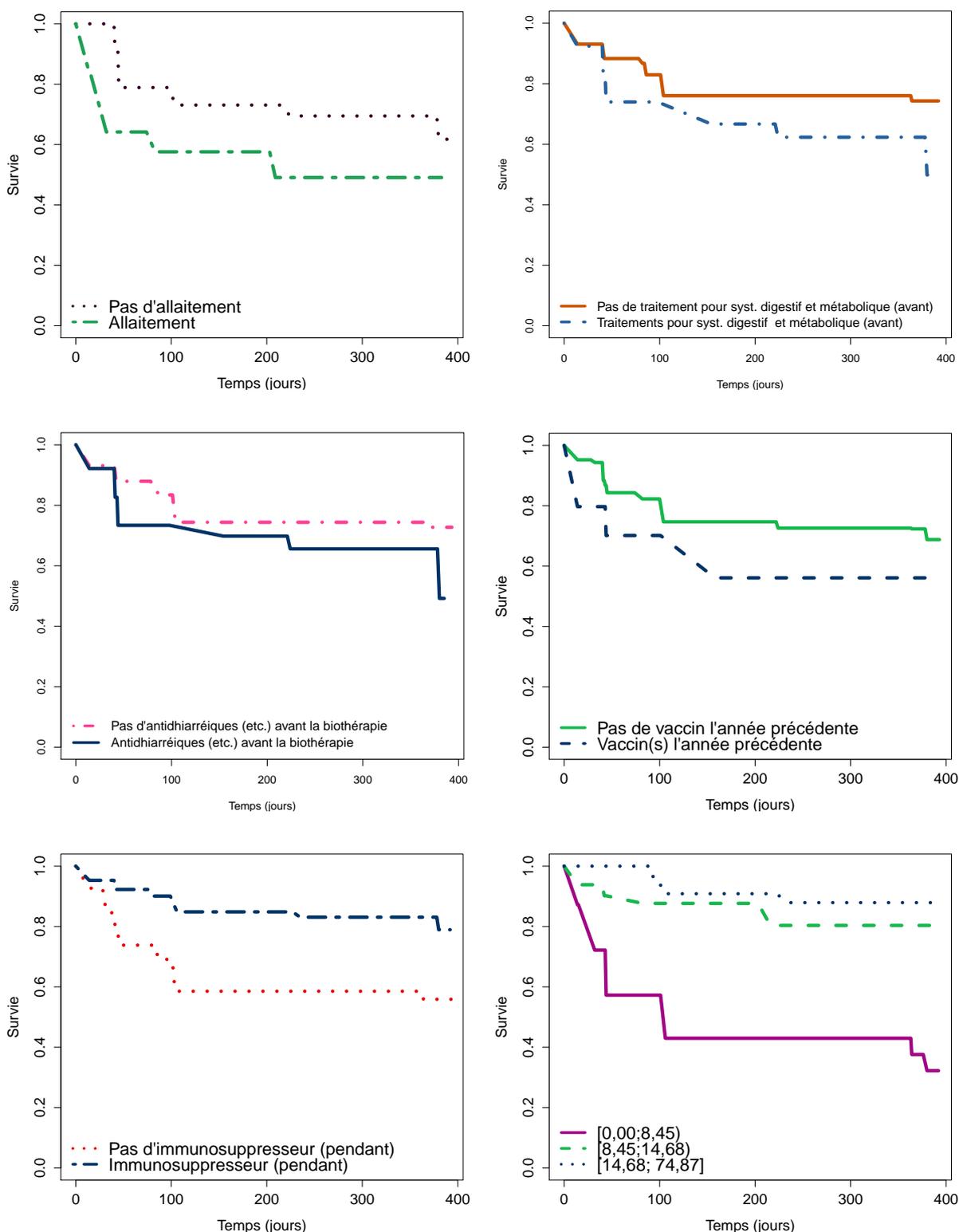


FIGURE 5.8 – Estimateur non-paramétrique du maximum de vraisemblance de la fonction de survie en fonction de l'allaitement (haut, gauche), de traitements pour système digestif et métabolique pris avant le début de la thérapie (haut, droite), d'antidiarrhéiques, anti-inflammatoires et anti-infectieux intestinaux pris avant le début de la thérapie (milieu, gauche), de vaccins réalisés l'année précédant le début du traitement (milieu, droite), d'immunosuppresseurs concomitants au traitement (bas, gauche) et du niveau de traitement en microgramme/mL (bas, droite). Cohorte prospective MICI, ABIRISK.

### 3) Commentaires

#### *La variable allaitement*

Dans cet échantillon, la proportion de femmes ayant allaité était de 26,1% ( $n = 23$ ). Le test destiné aux situations avec risques croisés a permis de détecter une association significative entre la variable allaitement et le développement des ADA. Les femmes ayant allaité développent plus d'ADA que les autres.

Ce résultat est confirmé avec la variable portant sur le temps écoulé depuis l'allaitement. Les femmes ayant récemment stoppé l'allaitement ( $n_{[0;3]} = 5, 21, 7\%$ ) développent plus d'ADA que les autres ( $n_{[3;39]} = 18, 78, 3\%$ ).

Ce résultat est d'autant plus intéressant que ces deux variables n'étaient pas détectées par le test du log-rank pour censure par intervalle. Les différentes *p-value* sont présentées dans le tableau 5.10.

Variabes	Tests	<i>p-value</i>
Allaitement	vgRC	0,01
	Log-rank	0,18
Années écoulées depuis l'allaitement	vgRC	0,02
	Log-rank	0,09

TABLEAU 5.10 – **p-value du test du score pour risques croisés et du pseudo log-rank pour la variable allaitement.** Le test du log-rank utilisé est le test du log-rank adapté à la censure par intervalle, présenté dans FINKELSTEIN et WOLFE [1985]. Le vgRC est le test du score, adapté à la situation de risques croisés. La variable "Années écoulées depuis l'allaitement" était divisée en deux catégories : [0;3) et [3;39].

#### *Autres variables*

Concernant les autres variables uniquement détectées par le test vgRC, nous voyons que les patients ayant pris des traitements pour le système digestif et métabolique avant le début de la biothérapie ( $n = 53, 29, 6\%$ ) développent plus d'ADA que les autres. Enfin, les patients ayant pris des antidiarrhéiques, anti-inflammatoires et anti-infectieux intestinaux avant le début du traitement ( $n = 47, 26, 3\%$ ) développent plus d'ADA que les autres.

Les trois variables suivantes seront commentées dans la section d'après.

### **B. Modèle de régression multivarié**

Parmi les variables sélectionnées précédemment, 7 variables à la fois détectées par le test que nous avons développé et par le log-rank ont été incluses dans un modèle de régression de Cox multivarié stratifié sur le traitement. L'analyse *stepwise* n'est faite que sur une base de données sans valeurs manquantes. Ici, elle est réalisée sur les données de 160 patients (dont 43, ADA-positifs) pour lesquels les valeurs des 7 variables sélectionnées sont disponibles.

Le tableau 5.11 présente la comparaison entre les 24 patients exclus et les 160 patients inclus dans l'analyse multivariée. Nous remarquons qu'aucune variable présentée n'est significativement différente entre les deux groupes.

Trois variables ont été sélectionnées en analyse pas-à-pas : vaccins (dans l'année précédant le début du traitement), immunosuppresseur (pendant) et niveau de biothérapie (6 semaines).

Le tableau 5.12 ci-dessous présente le modèle final incluant les variables sélectionnées.

Le tableau 5.13 résume la composition des variables dans l'échantillon.

Caractéristiques	Exclus (n =24)	Inclus (n =160)	p-value
ADA-positif, n (%)	8 (33,3)	51(27,7)	0,51
Type de maladie, n (%)			0,91
Crohn	20 (83,3)	128 (80,0)	
Rectocolite hémorragique	4 (16,7)	32 ( 20,0)	
Traitement, n (%)			0,90
Adalimumab	12 (50,0)	86 (53,7)	
Infliximab	12 (50,0)	74( 46,3)	
Sexe (femmes), n (%)	11 (45,8)	78(48,8)	0,96

TABLEAU 5.11 – **Comparaison de variables chez les patients exclus et inclus dans l'analyse multivariée.** La p-value du test ADA a été obtenue avec le test vgRP, pour censure par intervalle et risques croisés. Les autres p-value ont été obtenues avec un test du  $\chi^2$ .

<i>n</i> = 168, ADA+ = 46		RR	95% IC		p-value
Niveau de biothérapie (6 semaines)	[ 0,00, 8,45)	1	Référence		
	[ 8,45,14,68)	0,21	0,10	0,44	4,64e-05
	[14,68,74,87]	0,18	0,07	0,44	2,16e-04
Vaccin dans l'année précédente	Non	1	Référence		
	Oui	2,35	1,16	4,73	0,012
Immuno-suppresseurs (pendant)	Non	1	Référence		
	Oui	0,41	0,21	0,79	0,008

TABLEAU 5.12 – **Modèle multi-varié avec les variables sélectionnées dans la cohorte MICI, stratifié sur le type de biothérapie.** Le niveau de biothérapie est en  $\mu\text{g/mL}$ .  
RR : rapport de risque, IC : intervalle de confiance

<i>n</i> = 186, ADA+ = 51		<i>n</i>	%
Niveau de biothérapie (6 semaines)	[ 0,00, 8,45)	58	33,7
	[ 8,45,14,68)	57	33,15
	[14,68,74,87]	57	33,15
Vaccin dans l'année précédente	Non	156	84,8
	Oui	28	15,2
Immuno-suppresseurs (pendant)	Non	84	46,7
	Oui	96	53,3

TABLEAU 5.13 – **Tableau récapitulatif de composition dans l'échantillon, des variables sélectionnées dans le modèle multivarié, cohorte MICI.** Le niveau de biothérapie est en  $\mu\text{g/mL}$ .

Il apparaît que le niveau de biothérapie à 6 semaines est associé avec le développement des ADA : plus le niveau de biothérapie est élevé, plus le risque de développer des ADA est faible. De même, le fait de prendre des immunosuppresseurs concomitamment au traitement diminue la probabilité de développer des ADA. À l'inverse, le fait d'avoir été vacciné dans l'année précédant le traitement, augmente le risque de développer des ADA : les patients vaccinés ont deux fois plus de risque de développer des ADA que les autres.

## V. Discussion

Nous avons analysé, dans ce chapitre des données prospectives, issues du projet européen ABIRISK. La cohorte étudiée était constituée de patients atteints de maladies inflammatoires chroniques de l'intestin et traités par biothérapies. Le but de l'étude était d'identifier des facteurs cliniques associés au développement d'anticorps anti-biothérapies (ADA). La date d'apparition des ADA était censurée par intervalle.

Nous avons de fortes présomptions quant à la non-proportionnalité du rapport des risques liée à une interaction latente et/ou à la présence d'une fraction d'immuno-tolérants (non à risque). Nous avons donc analysé les variables explicatives avec le test que nous avons développé ; un test pour données censurées par intervalle et adapté à une situation de risques croisés. Une analyse multivariée a été réalisée par la suite et a permis de mettre en évidence trois variables significativement associées au développement des ADA.

Le niveau de biothérapie était significativement associé au développement des ADA ; un niveau plus élevé de biothérapie à la sixième semaine entraînant un risque plus faible de développer des ADA. Il serait intéressant de savoir si le niveau de biothérapie influence le développement des ADA, ou à l'inverse si le développement des ADA influence le niveau de biothérapie. L'absence d'anticorps anti-biothérapie pourrait en effet expliquer le niveau de biothérapie plus haut, à la sixième semaine. Des études supplémentaires seraient nécessaires pour préciser le sens du lien. Notons que la formation d'ADA est associée à une clairance accélérée de la biothérapie comme exposé par HINDRYCKX et al. [2017], qui étudient le développement des ADA chez des patients atteints de MICI et traités par biothérapies (infliximab, adalimumab, etc.). De plus, HINDRYCKX et al. [2017] montrent qu'une exposition à une concentration plus élevée de biothérapie a un effet protecteur sur le développement des ADA.

Une vaccination dans l'année précédant le début du traitement était associée à un risque plus élevé de développer des ADA. Nous pourrions supposer que la vaccination, qui déclenche une puissante réaction du système immunitaire, peut contribuer au développement des ADA. Les articles portant sur le sujet cherchent généralement à mesurer l'immunogénicité des vaccins et à son maintien dans le cadre de la prise d'immunomodulateurs pour le traitement des MICI comme chez MARÍN [2015] et chez LAUNAY et al. [2015]. Les mécanismes impliqués dans l'immunogénicité des vaccins chez les malades atteints de MICI restent peu clairs.

La prise d'immuno-suppresseurs pendant le traitement était associée à un risque plus faible de développer des ADA. L'effet protecteur des immuno-suppresseurs n'est pas surprenant. Dans HINDRYCKX et al. [2017], l'usage d'immunosuppresseurs chez des patients atteints de MICI et traités par biothérapies, est même recommandé pour prévenir la formation d'ADA. De nombreuses études ont confirmé l'effet protecteur des immunosuppresseurs pris concomitamment au traitement chez des patients atteints de MICI. Pour l'infliximab, le taux d'immunogénicité pouvant être divisé par 5 lorsqu'il était associé à des immunosuppresseurs. Pour l'adalimumab, le taux d'immunogénicité passait de 13,2% à 4% (HINDRYCKX et al. [2017]).

Une femme qui allaite produit des anticorps supplémentaires ; l'effet aggravant de l'allaitement sur le développement des ADA n'est donc pas étonnant. Le fait que cet effet s'estompe avec le temps — les femmes ayant cessé d'allaiter plus récemment ont plus d'ADA — est un argument supplémentaire en faveur de l'immunogénicité induite par l'allaitement. Ceci expliquerait aussi

pourquoi seul le test pour situation à risques croisés a détecté cet effet. Notons que de nombreux articles ont exploré le lien entre le fait d'avoir été allaité et les MICI (effet protecteur de l'allaitement, voir XU et al. [2017]), ainsi que de la nocivité potentielle de l'allaitement sur l'enfant par des femmes traitées par biothérapies, mais, à notre connaissance, il n'existe pas d'étude sur l'immunogénicité induite par l'allaitement dans les MICI.

Le test adapté à une situation de risques croisés a permis de mettre en évidence plusieurs variables significativement associées au développement des ADA. Parmi elles, l'effet de certaines variables comme l'allaitement n'était pas détecté par le log-rank. Des travaux supplémentaires seraient nécessaires pour obtenir une adaptation de tests multivariés aux données censurées par intervalle.

## VI. Références

- ABIRISK. «Anti-biopharmaceutical immunization : prediction and analysis of clinical relevance to minimize the risk», URL <http://www.abirisk.eu/>. 96
- ANANTHAKRISHNAN, A. N., C. N. BERNSTEIN, D. ILIOPOULOS, A. MACPHERSON, M. F. NEURATH, R. A. R. ALI, S. R. VAVRICKA et C. FIOCCHI. 2017, «Environmental triggers in IBD : a review of progress and evidence», *Nature Reviews Gastroenterology & Hepatology*, vol. 15, n° 1, doi : 10.1038/nrgastro.2017.136, p. 39–49. URL <https://doi.org/10.1038/nrgastro.2017.136>. 99
- BOGAERTS, K., R. LEROY, E. LESAFFRE et D. DECLERCK. 2002, «Modelling tooth emergence data based on multivariate interval-censored data», *Stat Med*, vol. 21, n° 24, p. 3775–3787. 109
- DALMASSO, C., P. B. T et T. MOREAU. 2004, «A simple procedure for estimating the false discovery rate», *Bioinformatics*, vol. 21, n° 5, doi :10.1093/bioinformatics/bti063, p. 660–668. URL <https://doi.org/10.1093/bioinformatics/bti063>. 108
- FINKELSTEIN, D. M. et R. A. WOLFE. 1985, «A Semiparametric Model for Regression Analysis of Interval-Censored Failure Time Data», *Biometrics*, vol. 41, n° 4, doi :10.2307/2530965, p. 933, ISSN 0006341X. URL <http://www.jstor.org/stable/2530965?origin=crossref>. 109, 112
- GOWER-ROUSSEAU, C., F. VASSEUR, M. FUMERY, G. SAVOYE, J. SALLERON, L. DAUCHET, D. TURCK, A. CORTOT, L. PEYRIN-BIROULET et J. F. COLOMBEL. 2013, «Epidemiology of inflammatory bowel diseases : New insights from a french population-based registry (EPIMAD)», *Digestive and Liver Disease*, vol. 45, n° 2, doi :10.1016/j.dld.2012.09.005, p. 89–94. URL <https://doi.org/10.1016/j.dld.2012.09.005>. 99
- HINDRYCKX, P., G. NOVAK, N. V. CASTEELE, R. KHANNA, D. LAUKENS, J. VIPUL et B. G. FEAGAN. 2017, «Incidence, prevention and management of anti-drug antibodies against therapeutic antibodies in inflammatory bowel disease : A practical overview», *Drugs*, vol. 77, n° 4, doi :10.1007/s40265-017-0693-5, p. 363–377. URL <https://doi.org/10.1007/s40265-017-0693-5>. 114
- KASER, A., S. ZEISSIG et R. S. BLUMBERG. 2010, «Inflammatory bowel disease», *Annual Review of Immunology*, vol. 28, n° 1, doi :10.1146/annurev-immunol-030409-101225, p. 573–621. URL <https://doi.org/10.1146/annurev-immunol-030409-101225>, pMID : 20192811. 98, 99, 100
- KENNY, E. E., I. PE'ER, A. KARBAN, L. OZELIUS, A. A. MITCHELL, S. M. NG, M. ERAZO, H. OSTRER, C. ABRAHAM, M. T. ABREU, G. ATZMON, N. BARZILAI, S. R. BRANT, S. BRESSMAN, E. R. BURNS, Y. CHOWERS, L. N. CLARK, A. DARVASI, D. DOHENY, R. H. DUERR, R. ELIAKIM, N. GILADI, P. K. GREGERSEN, H. HAKONARSON, M. R. JONES, K. MARDER, D. P. B. MCGOVERN, J. MULLE, A. ORR-URTREGER, D. D. PROCTOR, A. PULVER, J. I. ROTTER, M. S. SILVERBERG, T. ULLMAN, S. T. WARREN, M. WATERMAN, W. ZHANG, A. BERGMAN, L. MAYER, S. KATZ, R. J. DESNICK, J. H. CHO et I. PETER. 2012, «A genome-wide scan of ashkenazi jewish crohn's disease suggests novel susceptibility loci», *PLoS Genetics*, vol. 8, n° 3, doi :10.1371/journal.pgen.1002559, p. e1002559. URL <https://doi.org/10.1371/journal.pgen.1002559>. 101
- LAUNAY, O., V. ABITBOL, A. KRIVINE, L. B. SLAMA, A. BOURREILLE, J. L. DUPAS, X. HEBUTERNE, G. SAVOYE, D. DEPLANQUE, Y. BOUHNIC, A. L. PELLETIER, F. GALTIER, D. LAHARIE, M. NACHURY, F. ZERBIB, M. ALLEZ, G. BOMMELAER, B. DUCLOS, F. LUCHT, M. L. GOUGEON, E. TARTOUR, F. ROZENBERG, T. HANSLIK, L. BEAUGERIE, F. CARRAT, X. DUVAL, N. MOORE, J. F. COLOMBEL, S. SACHER-HUVELIN, M. VEYRAC, X. ROBLIN, C. DUBRAY, C. DUALE, A. F. CAILLEUX, G. H. BRUZZONI, J. M. CHALOPIN, C. BRANDT et P. BOUSQUET. 2015, «Immunogenicity and Safety of Influenza Vaccine in Inflammatory Bowel Disease Patients Treated or not with Immunomodulators and/or Biologics : A Two-year Prospective Study», *J Crohns Colitis*, vol. 9, n° 12, p. 1096–1107. 114

- MARÍN, A. C. 2015, «Immunogenicity and mechanisms impairing the response to vaccines in inflammatory bowel disease», *World Journal of Gastroenterology*, vol. 21, n° 40, doi :10.3748/wjg.v21.i40.11273, p. 11 273. URL <https://doi.org/10.3748/wjg.v21.i40.11273>. 114
- MATSUOKA, K., T. KOBAYASHI, F. UENO, T. MATSUI, F. HIRAI, N. INOUE, J. KATO, K. KOBAYASHI, K. KOBAYASHI, K. KOGANEI, R. KUNISAKI, S. MOTOYA, M. NAGAHORI, H. NAKASE, F. OMATA, M. SARUTA, T. WATANABE, T. TANAKA, T. KANAI, Y. NOGUCHI, K. ICHI TAKAHASHI, K. WATANABE, T. HIBI, Y. SUZUKI, M. WATANABE, K. SUGANO et T. SHIMOSEGAWA. 2018, «Evidence-based clinical practice guidelines for inflammatory bowel disease», *Journal of Gastroenterology*, vol. 53, n° 3, doi :10.1007/s00535-018-1439-1, p. 305–353. URL <https://doi.org/10.1007/s00535-018-1439-1>. 101
- MCDONALD, J. 2014, *Handbook of Biological Statistics*, 3<sup>e</sup> éd., Sparky House Publishing, Baltimore, Maryland. 108
- XU, L., P. LOCHHEAD, Y. KO, B. CLAGGETT, R. W. LEONG et A. N. ANANTHAKRISHNAN. 2017, «Systematic review with meta-analysis : breastfeeding and the risk of crohn's disease and ulcerative colitis», *Alimentary Pharmacology & Therapeutics*, vol. 46, n° 9, doi :10.1111/apt.14291, p. 780–789. URL <https://doi.org/10.1111/apt.14291>. 115



## **Troisième partie**

# **Comparaison de groupes de données censurées par intervalle avec population à risques hétérogènes**



## Chapitre 6

# Test de comparaison de deux ou plusieurs groupes de données censurées par intervalle avec population à risques hétérogènes

### Sommaire

---

<b>I. Introduction</b> . . . . .	<b>122</b>
A. Le modèle de mélange . . . . .	122
B. Le modèle à risque cumulatif borné . . . . .	123
<b>II. Proposition de test</b> . . . . .	<b>123</b>
A. Notations . . . . .	123
B. Fonction de survie . . . . .	124
C. Test du score pour données current status $H_0 : \alpha = \beta = 0$ . . . . .	125
D. Extension du score aux données censurées par intervalle . . . . .	126
<b>III. Étude de simulations sur données censurées par intervalle de type I</b> . . . . .	<b>128</b>
A. Les tests comparés . . . . .	128
B. Protocole de simulation . . . . .	128
C. Résultats de simulations . . . . .	130
D. Commentaires . . . . .	133
<b>IV. Discussion</b> . . . . .	<b>134</b>
<b>V. Références</b> . . . . .	<b>137</b>

---

## I. Introduction

Dans certaines études de survie, la population d'étude peut être hétérogène et comporter une fraction de sujets dits tolérants — les patients non-susceptibles — qui ne feront jamais l'événement d'intérêt.

Ce type de données peut se rencontrer dans différents domaines, dont :

1. l'oncologie : dans des études où la variable d'intérêt est le délai avant récurrence de tumeurs cancéreuses, il y a des évidences biologiques fortes portant sur certains patients — entièrement guéris — qui ne feront pas de récurrences. Ce sont les "survivants à long-terme" ou les *cured patients* qui ont donné leur nom aux modèles permettant de prendre en compte leur proportion : les *cure models* ou modèles à plateau. De très nombreuses analyses de données sur le sujet sont disponibles dans la littérature<sup>1</sup> ;
2. l'immunologie : une proportion de patients, les immuns sont tolérants vis-à-vis de l'anti-corps et ne développeront pas d'ADA.

L'une des hypothèses de base des analyses de survie est que tous les individus de l'échantillon feront l'événement d'intérêt. Cette condition n'est pas respectée lorsque l'échantillon contient des non-susceptibles. Ne pas prendre en compte l'existence de ces patients entraîne une perte de puissance des tests utilisés dû à la non-proportionnalité des risques.

Les modèles à plateau (*cure rate models*, *long-term survivors*) développés pour répondre à cette problématique, ont été beaucoup étudiés depuis la seconde moitié du vingtième siècle à partir des articles séminaux de BOAG [1949] et BERKSON et GAGE [1952].

Le modèle de mélange et le modèle à risque cumulatif borné (dénommé *promotion time cure model* par COONER et al. [2007]) sont les modèles à plateau les plus largement utilisés. Nous en présentons ci-après les caractéristiques générales.

### A. Le modèle de mélange

Le **modèle de mélange** à deux composantes a été proposé par BERKSON et GAGE [1952] tel que :

$$S_{\text{pop}}(t) = \phi + (1 - \phi) S^*(t) \quad (6.1)$$

avec  $\phi$ , la proportion d'individus non-susceptibles.  $S^*(t)$  représente la fonction de survie des individus susceptibles.  $S^*(t)$  est une fonction de survie propre, où  $\lim_{t \rightarrow +\infty} S^*(t) = 0$ .

$S_{\text{pop}}(t)$ , la fonction de survie marginale, est une fonction de survie impropre, car  $\lim_{t \rightarrow +\infty} S_{\text{pop}}(t) > 0$ .

La fonction  $\phi$  est souvent modélisée à l'aide d'une régression logistique lorsqu'on veut prendre en compte l'effet de covariables sur la fraction de non-susceptibles :

$$\phi(Z) = \frac{\exp(Z\gamma)}{1 + \exp(Z\gamma)},$$

où  $\gamma$  est un vecteur de coefficients pour la covariable  $Z$ .

Lorsque  $\phi = 0$ , le modèle (6.1) se réduit à un modèle de survie standard.

Le modèle de mélange à deux composantes a été largement étudié dans la littérature depuis son introduction<sup>2</sup>. Il est probablement le modèle le plus populaire du fait de son interprétation

1. BROET et al. [2004]; FAREWELL [1986]; KUK et CHEN [1992]; LAM et al. [2005] sur le cancer du sein, CHEN et al. [1999]; IBRAHIM et al. [2001a] sur les mélanomes.

2. Voir les livres de MALLER et ZHOU [1996] et PENG et TAYLOR [2013] ainsi que les articles de CHEN et al. [2017]; FAREWELL [1982, 1986]; GOLDMAN [1984]; GRAY et TSIATIS [1989]; GREENHOUSE et WOLFE [1984]; HALPERN et BROWN [1987]; KUK et CHEN [1992]; LASKA et MEISNER [1992]; MALLER et ZHOU [1995]; PENG et TAYLOR [2013]; SPOSTO et al. [1992]; SY et TAYLOR [2000]; TAYLOR [1995]

simple et de sa structure sous forme de combinaison de deux modèles aisée à utiliser et facilement généralisable à des situations plus complexes. Le modèle de mélange permet ainsi la prise en compte séparée de l'effet d'une covariable sur la proportion de non-susceptibles et l'effet d'une covariable sur la distribution du temps de survie pour les susceptibles. L'inconvénient de ce modèle est qu'il n'a pas de structure à risques proportionnels en présence des covariables, si  $\phi$  est modélisé à l'aide d'une régression binomiale classique (IBRAHIM et al. [2001b]).

### B. Le modèle à risque cumulatif borné

Par la suite, nous nous intéresserons plus spécifiquement au second type de modèle à plateau, le **modèle à risque cumulatif borné**, quelquefois appelé *nonmixture model*, par opposition au modèle précédent. Ce modèle a été introduit par YAKOVLEV et al. [1993] et YAKOVLEV et TSODIKOV [1996]. Il a été initialement considéré pour des applications en oncologie pour des stades précoces de cancers où une proportion d'individus peut être considérée comme guérie. Ce type de modèle peut également être utilisé dans d'autres situations où une proportion de la population n'est pas sujette à l'événement d'intérêt comme en immunologie.

Ce modèle peut s'écrire comme suit :

$$S(t) = \exp[-\theta(1 - e^{-H(t)})] \quad (6.2)$$

où  $H(t)$  est une fonction positive, croissante et non-bornée. Et  $e^{-H(t)}$  est une pseudo-fonction de survie propre<sup>3</sup>.

Nous avons  $S(t = 0) = 1$  et  $S(t = \infty) = \exp(-\theta)$ . Avec  $\exp(-\theta)$  représentant le plateau ou la fraction de sujets non-susceptibles.

De plus,

$$\Lambda(t) = \theta(1 - e^{-H(t)})$$

est la fonction de risque cumulé borné telle que  $S(t) = \exp(-\Lambda(t))$ .

## II. Proposition de test

Nous présentons ci-dessous un nouveau test pour données *current status* qui est une extension du test proposé par BROËT et al. [2001] dans le cas de la censure aléatoire droite. Il permet la comparaison de deux ou plusieurs groupes dans une population incluant une proportion de sujets non à risque qui ne feront pas l'événement d'intérêt. Le test a fait l'objet d'une publication dans JONAS et al. [2017]. Nous présentons également l'adaptation du test aux données censurées par intervalle classiques (type II).

### A. Notations

Nous reprenons les notations introduites en section A. page 10. Les variables aléatoires continues  $T$  et  $C$  représentent le délai avant détection des ADA (inobservé) et le temps écoulé entre le début de l'étude la visite, respectivement. Soient  $f(t)$ , la fonction de densité et  $S(t)$  (*resp.*  $\bar{S}(t)$ ), la fonction de survie (*resp.* fonction de répartition) de  $T$ . La fonction de risque de  $T$  est  $\lambda(t)$  avec  $\lambda(t) = f(t)/S(t)$ .

Pour les données censurées par intervalle de type I, nous savons uniquement si l'événement d'intérêt s'est produit avant un temps de surveillance aléatoire unique. Nous introduisons le test pour une situation où deux groupes sont comparés. Ainsi, pour chaque patient  $i$  ( $i = 1, \dots, n$ ),  $Z_i$  est une variable binaire qui indique l'appartenance au groupe ( $Z_i = 0$  ou  $1$ , avec  $0$  le groupe

3.  $S(t)$  peut également être interprété comme un modèle de fragilité avec une fragilité de type Poisson, d'espérance  $\theta$ .

de référence). Ainsi,  $\{C_i, \delta_i = \mathbb{1}_{(C_i \geq T_i)}, Z_i\}$  sont les informations disponibles. Si un événement s'est produit, nous savons que  $T_i$  appartient à  $[0, C_i]$  (censure gauche); sinon  $T_i$  appartient à  $[C_i, +\infty[$  (censure droite).

Pour les données censurées par intervalle de type II,  $T \in (L, R]$ , avec  $L$  et  $R$ , les bornes aléatoires gauches et droites de l'intervalle;  $L$  pouvant être égal à 0 (censure gauche), et  $R$  pouvant être égal à  $+\infty$  (censure droite).

Nous faisons l'hypothèse que les temps de censure et d'événement sont indépendants. Nous supposons également que les temps de censure sont des variables aléatoires indépendantes et identiquement distribuées pour tous les sujets.

## B. Fonction de survie

Dans ce développement, nous considérons la fonction de survie semi-paramétrique impropre telle que pour chaque patient  $i$ , nous avons :

$$S(t | Z_i = z) = S_z(t) = e^{-\theta e^{\alpha z} [1 - \exp(-H(t)e^{\beta z})]} \quad (6.3)$$

où  $H(t)$  est une fonction non-spécifiée croissante et positive de zéro à l'infini et  $\alpha$  et  $\beta$  sont des paramètres d'intérêt unidimensionnels.

Les fonctions de risque,  $\lambda_{Z=0}(t)$  et  $\lambda_{Z=1}(t)$  sont telles que :

$$\lambda_{Z=1}(t) = \lambda_{Z=0}(t) e^{\alpha + \beta} e^{-H(t)(e^{\beta} - 1)}.$$

La fonction de survie  $S_z(t)$  est impropre puisque  $\lim_{t \rightarrow +\infty} S_z(t) > 0$ . Sa valeur limite est appelée le plateau (quelquefois définie comme *tail defect*) et est ici égale à  $e^{-\theta e^{\alpha z}}$ .

Dans notre contexte, elle représente la probabilité d'être non à risque de réaliser l'événement d'intérêt. Les changements dans la fraction des non-susceptibles et dans le délai avant événement sont modélisés à travers les paramètres d'intérêt  $\alpha$  et  $\beta$ , respectivement. Ainsi, lorsque  $\alpha = 0$ , les deux groupes ont le même plateau (proportion de patients non-susceptibles). Et lorsque  $\beta = 0$ , le modèle (6.3) est un modèle à risques proportionnels avec un rapport de risque constant au cours du temps et une valeur de plateau différente selon le groupe.

### Log-vraisemblance pour données censurées par intervalle de type I

Sous le modèle (6.3), la log-vraisemblance simplifiée pour les  $n$  données *current status* est :

$$\begin{aligned} \mathcal{L}_n^{CS}(\alpha, \beta, \theta, H(\cdot)) = & \sum_{i=1}^n \delta_i \left\{ \ln \left( 1 - e^{-\theta e^{\alpha z_i} [1 - \exp(-H(c_i) e^{\beta z_i})]} \right) \right\} \\ & - (1 - \delta_i) \left\{ \theta e^{\alpha z_i} \left[ 1 - \exp \left( -H(c_i) e^{\beta z_i} \right) \right] \right\}. \end{aligned}$$

### Log-vraisemblance pour données censurées par intervalle de type II

Sous le modèle (6.3), la log-vraisemblance simplifiée pour les  $n$  données censurées par intervalle est :

$$\mathcal{L}_n(\alpha, \beta, \theta, H(\cdot)) = \sum_{i=1}^n \ln [S(l_i | z_i) - S(r_i | z_i)].$$

**C. Test du score pour données current status  $H_0 : \alpha = \beta = 0$**

L'hypothèse nulle ( $H_0 : \alpha = \beta = 0$ ) à tester, est l'égalité de deux distributions de survie impropres, ce qui équivaut à dire que l'appartenance à un groupe n'a aucun effet sur la distribution du délai avant événement ou sur la fraction non à risque. Sous l'hypothèse nulle  $H_0$ , le vecteur de score est constitué des éléments  $U^{cs} = (U_\alpha^{cs}, U_\beta^{cs})$  où :

$$U_\alpha^{cs} = \frac{\partial \mathcal{L}_n^{cs}}{\partial \alpha} \Big|_{\alpha=\beta=0} = \sum_{i=1}^n \delta_i \left\{ \frac{e^{-\theta[1-\exp(-H(c_i))]} \theta [1 - \exp(-H(c_i))]}{1 - e^{-\theta[1-\exp(-H(c_i))]}} z_i \right\} - (1 - \delta_i) \{ \theta [1 - \exp(-H(c_i))] z_i \}.$$

D'où,

$$\begin{aligned} U_\alpha^{cs} &= \sum_{i=1}^n z_i \left[ \frac{\theta [1 - \exp(-H(c_i))]}{1 - e^{-\theta[1-\exp(-H(c_i))]}} \right] \{ \delta_i - 1 + e^{-\theta[1-\exp(-H(c_i))]} \}, \\ &= \sum_{i=1}^n z_i w_{\alpha,i}^{cs}. \end{aligned}$$

Et,

$$U_\beta^{cs} = \frac{\partial \mathcal{L}_n^{cs}}{\partial \beta} \Big|_{\alpha=\beta=0} = \sum_{i=1}^n \delta_i \left\{ \frac{e^{-\theta[1-\exp(-H(c_i))]} \theta H(c_i) e^{-H(c_i)}}{1 - e^{-\theta[1-\exp(-H(c_i))]}} z_i \right\} - (1 - \delta_i) \{ \theta H(c_i) e^{-H(c_i)} z_i \}.$$

Donc,

$$\begin{aligned} U_\beta^{cs} &= \sum_{i=1}^n z_i \left[ \frac{\theta H(c_i) e^{-H(c_i)}}{1 - e^{-\theta[1-\exp(-H(c_i))]}} \right] \{ \delta_i - 1 + e^{-\theta[1-\exp(-H(c_i))]} \}, \\ &= \sum_{i=1}^n z_i w_{\beta,i}^{cs}. \end{aligned}$$

Comme le montrent les deux formules précédentes, ces statistiques de score peuvent être réécrites en tant que statistiques linéaires de rang avec les fonctions de pseudo-rangs  $w_{\alpha,i}^{cs}$  et  $w_{\beta,i}^{cs}$  pour le sujet  $i$ , qui dépendent de la fonction de répartition sous l'hypothèse nulle.

$$\begin{aligned} w_{\alpha,i}^{cs}(c_i) &= \left[ \frac{-\ln(1 - \bar{S}_{H_0}(c_i))}{\bar{S}_{H_0}(c_i)} \right] \{ \delta_i - \bar{S}_{H_0}(c_i) \}, \\ w_{\beta,i}^{cs}(c_i) &= \left[ \frac{-\theta \ln \left( 1 + \frac{\ln(1 - \bar{S}_{H_0}(c_i))}{\theta} \right) \left[ 1 + \frac{\ln(1 - \bar{S}_{H_0}(c_i))}{\theta} \right]}{\bar{S}_{H_0}(c_i)} \right] \{ \delta_i - \bar{S}_{H_0}(c_i) \}, \end{aligned}$$

où  $\bar{S}_{H_0}(\cdot) = F_{H_0}(\cdot)$  est la fonction de répartition sous l'hypothèse nulle  $H_0$ .

En pratique, nous remplaçons  $F_{H_0}(t)$  et  $\theta$  par, respectivement,  $\hat{F}_{H_0}(t)$  et  $\hat{\theta}$  en appliquant l'approche suivante : pour les données censurées par intervalles, un estimateur non paramétrique du maximum de vraisemblance (NPMLE) pour la fonction de répartition sous l'hypothèse nulle  $\hat{F}_{H_0}(t)$ , peut être obtenu en utilisant l'algorithme *pooled-adjacent-violators* sur l'ensemble complet de données (ROBERTSON et al. [1988]). Cet estimateur utilise les sauts estimés (masse de probabilité) se produisant sur les intervalles choisis comme ceux de TURNBULL [1976]. Dans notre

contexte, avec une distribution de survie incorrecte, nous considérons arbitrairement que pour le dernier intervalle (dont la limite supérieure est infinie), le saut est mis à zéro. Ainsi, sous l'hypothèse nulle, un estimateur de  $\theta$  est obtenu par  $\hat{\theta} = -\ln(\hat{F}_{H_0}(\infty))$  où  $\hat{F}_{H_0}(\infty)$  est la valeur estimée du plateau et sa valeur qui est calculée comme la différence entre la dernière estimation de  $\hat{F}_{H_0}(t)$  et 1. De plus, lorsque  $\bar{S}_{H_0}(\cdot) = 0$ , par convention  $\bar{S}_{H_0}(\cdot) = \min_{\bar{S}_{H_0} > 0}(\bar{S}_{H_0}(\cdot))$ .

Cette approche repose sur l'hypothèse que le mécanisme de censure vérifie la condition de suivi suffisant, c'est-à-dire que tous les sujets susceptibles de faire l'événement font l'événement (s'ils ne sont pas censurés) pendant la période de suivi (MALLER et ZHOU [1995]). En pratique, cela implique que nous devons nous assurer d'une période d'observation suffisamment longue pour détecter la présence d'individus non à risque dans la population étudiée (c'est-à-dire que le dernier intervalle devrait être sans événement).

Étant donné que les statistiques  $U_\alpha^{cs}$  et  $U_\beta^{cs}$  peuvent être exprimées comme des statistiques linéaires de rangs, nous pouvons obtenir leur variance permutationnelle ainsi que la matrice de covariance telles que (voir FAY [1996]; KALBFLEISCH et PRENTICE [1980]) :

$$\begin{aligned} V_{U_\alpha^{cs}} &= \frac{1}{n-1} \left[ \sum_{i=1}^n (w_{\alpha i}^{cs} - \bar{w}_\alpha^{cs})^2 \sum_{i=1}^n (z_i - \bar{z})^2 \right], \\ V_{U_\beta^{cs}} &= \frac{1}{n-1} \left[ \sum_{i=1}^n (w_{\beta i}^{cs} - \bar{w}_\beta^{cs})^2 \sum_{i=1}^n (z_i - \bar{z})^2 \right], \\ \text{cov}(U_\alpha^{cs}, U_\beta^{cs}) &= \frac{1}{n-1} \left[ \sum_{i=1}^n (w_{\alpha i}^{cs} - \bar{w}_\alpha^{cs}) (w_{\beta i}^{cs} - \bar{w}_\beta^{cs}) \sum_{i=1}^n (z_i - \bar{z})^2 \right], \end{aligned}$$

avec  $\bar{w}_\alpha = \frac{1}{n} \sum_{i=1}^n w_{\alpha i}$  et  $\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$ .

Finalement, la statistique de score proposée sous  $H_0$  est donnée par :

$$T_{H_0}^{cs} = \left( U_\alpha^{cs}, U_\beta^{cs} \right) V^{-1} \begin{pmatrix} U_\alpha^{cs} \\ U_\beta^{cs} \end{pmatrix}$$

où  $V$  est la matrice de variance-covariance obtenue avec les formules précédentes.

Sous l'hypothèse nulle, la statistique  $T_{H_0}^{cs}$  est asymptotiquement distribuée comme un  $\chi^2$  à deux degrés de liberté.

#### D. Extension du score aux données censurées par intervalle

Le test du score présenté dans la section précédente peut être étendu aux données censurées par intervalle classiques.

Le vecteur de score est constitué des scores par rapport aux paramètres d'intérêt,  $\alpha$  et  $\beta$ , soient les dérivées première de la log-vraisemblance par rapport aux paramètres considérés.

$$U_\alpha = \frac{\partial \mathcal{L}_n}{\partial \alpha} \Big|_{\alpha=\beta=0} = \sum_{i=1}^n \frac{S'(l_i|z_i) - S'(r_i|z_i)}{S(l_i|z_i) - S(r_i|z_i)},$$

où  $S'(\cdot|z_i) = \frac{\partial S(\cdot|z_i)}{\partial \alpha} \Big|_{\alpha=\beta=0}$ .

Sous le modèle (6.3), nous avons :

$$\begin{aligned} U_\alpha &= \sum_{i=1}^n z_i (-\theta) \left\{ \frac{e^{-\theta[1-\exp(-H(l_i))]} [1 - \exp(-H(l_i))] - e^{-\theta[1-\exp(-H(r_i))]} [1 - \exp(-H(r_i))]}{e^{-\theta[1-\exp(-H(l_i))]} - e^{-\theta[1-\exp(-H(r_i))]} \right\} \\ &= \sum_{i=1}^n z_i w_{\alpha,i}. \end{aligned}$$

Et,

$$U_\beta = \frac{\partial \mathcal{L}_n}{\partial \beta} \Big|_{\alpha=\beta=0} = \sum_{i=1}^n \frac{S^*(l_i|z_i) - S^*(r_i|z_i)}{S(l_i|z_i) - S(r_i|z_i)},$$

$$\text{où } S^*(\cdot|z_i) = \frac{\partial S(\cdot|z_i)}{\partial \beta} \Big|_{\alpha=\beta=0}.$$

Sous le modèle (6.3), nous avons :

$$\begin{aligned} U_\beta &= \sum_{i=1}^n z_i (-\theta) \left\{ \frac{e^{-\theta[1-\exp(-H(l_i))]} e^{-H(l_i)} H(l_i) - e^{-\theta[1-\exp(-H(r_i))]} e^{-H(r_i)} H(r_i)}{e^{-\theta[1-\exp(-H(l_i))]} - e^{-\theta[1-\exp(-H(r_i))]} } \right\} \\ &= \sum_{i=1}^n z_i w_{\beta,i}. \end{aligned}$$

Tout comme pour les données *current status*, ces statistiques de score peuvent être réécrites en tant que statistiques linéaires de rang avec les fonctions de pseudo-rangs  $w_{\alpha,i}$  et  $w_{\beta,i}$  pour le sujet  $i$ , qui dépendent de la fonction de répartition sous l'hypothèse nulle. Nous avons :

$$\begin{aligned} w_{\alpha,i} &= \frac{S_{H_0}(l_i) \ln[S_{H_0}(l_i)] - S_{H_0}(r_i) \ln[S_{H_0}(r_i)]}{S_{H_0}(l_i) - S_{H_0}(r_i)}, \\ w_{\beta,i} &= \frac{S_{H_0}(l_i) (\theta + \ln[S_{H_0}(l_i)]) \ln\left(1 + \frac{\ln[S_{H_0}(l_i)]}{\theta}\right) - S_{H_0}(r_i) (\theta + \ln[S_{H_0}(r_i)]) \ln\left(1 + \frac{\ln[S_{H_0}(r_i)]}{\theta}\right)}{S_{H_0}(l_i) - S_{H_0}(r_i)}, \end{aligned}$$

avec  $S_{H_0}(\cdot) = 1 - F_{H_0}(\cdot)$  où  $F_{H_0}$  est la fonction de répartition sous l'hypothèse nulle  $H_0$ .

Les statistiques  $U_\alpha$  et  $U_\beta$  ayant la forme de statistiques linéaires de rang, nous pouvons obtenir leur variance permutationnelle. Les expressions de variances et de covariance sont similaires à celles présentées dans le cadre des données censurées par intervalle de type 1.

La statistique de score pour données censurées par intervalle sous  $H_0$  est alors donnée par :

$$T_{H_0} = (U_\alpha, U_\beta) V^{-1} \begin{pmatrix} U_\alpha \\ U_\beta \end{pmatrix}$$

où  $V = \begin{bmatrix} V_{U_\alpha} & \text{cov}(U_\alpha, U_\beta) \\ \text{cov}(U_\alpha, U_\beta) & V_{U_\beta} \end{bmatrix}$  est la matrice de variance-covariance.

Sous l'hypothèse nulle, la statistique  $T_{H_0}$  est asymptotiquement distribuée comme un  $\chi^2$  à deux degrés de liberté (FAY [1996]; KALBFLEISCH et PRENTICE [1980]).

### III. Étude de simulations sur données censurées par intervalle de type I

Pour examiner les propriétés du test proposé dans la section précédente, nous avons effectué des simulations de Monte-Carlo avec des données censurées par intervalle de type I générées soit par un modèle à risque cumulatif borné, soit par un modèle de mélange à deux composantes.

#### A. Les tests comparés

Nous avons comparé le test de score que nous avons développé au test du log-rank pour données censurées par intervalle présenté en section C. page 35, qui est la statistique de test semi-paramétrique la plus fréquemment utilisée pour des données censurées par intervalle. Le test de pseudo log-rank est un test du score calculé sous une vraisemblance groupée avec une hypothèse alternative de risques proportionnels. Sachant que la vraisemblance simplifiée pour données *current status* est  $\mathcal{L}_n(\beta, S_z(\cdot)) = \sum_{i=1}^n (1 - \delta_i) \ln[S(c_i | z_i)] + \delta_i \ln[1 - S(c_i | z_i)]$ , nous rappelons ci-dessous l'expression du log-rank pour des données *current status* :

$$U_0^{cs} = \sum_{i=1}^n z_i \frac{\log \hat{S}(c_i)}{1 - \hat{S}(c_i)} [1 - \hat{S}(c_i) - \delta_i].$$

Les deux tests testent la même hypothèse nulle d'égalité des deux distributions de survie.

#### B. Protocole de simulation

Dans cette étude de simulation, nous avons étudié les effets de divers taux de censure, de déséquilibres dans la covariable ainsi que de suivis insuffisants.

##### *Temps de survie, T*

Les temps de survie ont été générés, selon les deux modèles décrits ci-dessous.

Le premier était un modèle cumulatif borné tel que :  $S(t | Z = z) = e^{-\theta e^{\alpha z} [1 - \exp(-te^{\beta z})]}$ . Le second était un modèle de mélange à deux composantes tel que :  $S(t | Z = z) = e^{-\theta e^{\alpha z}} + (1 - e^{-\theta e^{\alpha z}}) e^{-te^{\beta z}}$ .

##### *Covariable, Z*

La variable Z a été générée, selon une distribution de Bernoulli de paramètre  $\xi$ . Sauf indication contraire,  $\xi = 0,5$ .

##### *Temps de censure, C*

Les temps de censure (visite), C, ont été générés indépendamment à partir d'une distribution exponentielle de paramètre  $\lambda_C$ ; sa valeur a été choisie en fonction du pourcentage désiré d'observations censurées.

##### *Paramètres*

Le nombre total de sujets  $n_{\text{tot}}$  a été fixé à 400. Les configurations suivantes ont été considérées :  $\beta$  variait de  $-3,2$  à  $3,2$  avec un pas de  $0,4$ , et  $\alpha$  variait de  $-0,5$  à  $0,5$  avec un pas de  $0,25$ .

Les valeurs du plateau pour le groupe de référence,  $\tau_0 = S_{Z=0}(\infty) = e^{-\theta}$ , ont été fixées à  $0,3$ ,  $0,5$  et  $0,7$ . Nous avons indiqué pour chaque valeur de  $\alpha$  la valeur du plateau pour le groupe  $Z = 1$  tel que  $\tau_1 = \tau_0^{\exp(\alpha)}$ .

Les taux de censure à droite ( $p$ ) étaient de  $20\%$  et  $40\%$ . Ici,  $p$  représente le pourcentage d'observations censurées à droite parmi les sujets susceptibles uniquement. Ainsi, le pourcentage total de sujets censurés pour le groupe de référence est égal à  $\tau_0 + p(1 - \tau_0)$ . Les résultats des simulations effectuées sous le modèle à risque cumulatif borné sont présentés dans les tableaux 6.1 -

6.6. Les résultats des simulations effectuées sous le modèle de mélange à deux composantes sont disponibles en Annexe, section C page IX dans les tableaux C.1 - C.6.

#### *Scénarios supplémentaires*

Pour les scénarios dans lesquels la covariable était déséquilibrée,  $\xi$  a été fixé à 0,3 et 0,7. Les résultats des simulations effectuées sous le modèle à risque cumulatif borné sont présentés dans les tableaux 6.7 - 6.8. Les résultats des simulations effectuées sous le modèle de mélange à deux composantes sont disponibles en Annexe, section C page IX dans les tableaux C.7 - C.8.

Nous avons également évalué l'effet d'un suivi insuffisant sur les propriétés du test en générant des temps de censure à partir d'une distribution uniforme sur un petit intervalle  $(0, U_{max})$ , avec  $U_{max}$  variant selon la valeur du  $\beta$  et fixé pour s'assurer que 20% des patients susceptibles n'aient pas un suivi suffisant. Les résultats sont reportés dans le tableau 6.9.

Pour les deux derniers scénarios, la valeur du plateau était de  $\tau_0 = 30\%$  et celle de la censure à droite était de  $p = 20\%$ .

#### *Implémentation*

Comme indiqué ci-dessus, nous avons utilisé un estimateur de vraisemblance non paramétrique pour la fonction de répartition sous l'hypothèse nulle ( $\hat{F}_{H_0}(t)$ ) pour la statistique proposée. Cette estimation était obtenue à l'aide de l'algorithme pooled-adjacent-violators disponible dans le package 'isotone' de LEEUW et al. [2009].

Pour le test de pseudo-log-rank, nous avons choisi la fonction 'gLRT2' du package 'glrt' sur le logiciel R présenté en section 1) page 41 pour des raisons liées à la vitesse de calcul. D'autres procédures (méthodes de permutation exactes ou asymptotiques, test de score avec informations de Fisher observée et imputations multiples) sont disponibles dans différents packages R tels que 'interval' et 'FHtest' comme détaillé dans le chapitre 2.

La fonction R permettant de calculer le test du score proposé pour données censurées par intervalle de type 1 est disponible en Annexe en section B page V.

#### *Guide de lecture des tableaux de résultats*

Pour chaque scénario, nous avons généré 5 000 ensembles de données. Nous avons considéré un test bilatéral avec un niveau de signification à 0,05. Les résultats, en pourcentages, sont présentés dans les tableaux ci-dessous. Nous y rapportons les probabilités de rejeter  $H_0$  pour le test de score proposé ainsi qu'entre parenthèses, en dessous de chaque résultat, les différences observées en terme de puissance statistique entre le test du score proposé et le test de pseudo log-rank; les valeurs positives dans les différences indiquant une meilleure puissance du test de score proposé. Les valeurs estimées sous l'hypothèse nulle sont également reportées, surlignées en gris dans les tableaux.

Les lignes correspondent aux différentes valeurs de plateaux ( $\alpha$ ) testées. La valeur du plateau pour le second groupe ( $\tau_1$ ) correspondante à chaque valeur de  $\alpha$  est indiquée.

Les colonnes correspondent aux différentes valeurs du paramètre  $\beta$  testées, variant de -3,2 à 3,2.

**C. Résultats de simulations**

TABLEAU 6.1 – Modèle à risque cumulatif borné, censure exponentielle,  $\tau_0 = 30\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.48$	100.0 (0.0)	100.0 (0.1)	99.8 (-0.1)	99.8 (-0.1)	99.6 (-0.2)	98.9 (-0.4)	97.8 (-0.9)	94.7 (-2.5)	89.8 (-4.5)	84.4 (-6.2)	78.8 (-5.7)	74.3 (-3.3)	71.1 (0.7)	68.2 (4.2)	69.9 (9.5)	69.1 (10.6)	68.0 (13.5)
-0.25 $\tau_1 = 0.39$	99.9 (0.1)	99.9 (0.1)	99.8 (0.3)	99.4 (0.4)	98.1 (0.8)	94.0 (0.7)	81.5 (-1.3)	58.3 (-7.5)	34.0 (-9.6)	19.2 (-3.1)	19.0 (8.9)	27.6 (22.4)	39.8 (33.9)	54.8 (44.8)	65.0 (50.8)	72.6 (52.1)	77.6 (53.1)
0	99.9 (0.4)	99.8 (0.7)	99.7 (1.4)	98.6 (3.5)	93.6 (7.7)	77.5 (11.6)	43.2 (6.1)	13.9 (1.0)	5.1	14.1 (0.9)	45.0 (7.9)	78.1 (11.9)	93.2 (8.4)	98.7 (4.3)	99.6 (1.7)	99.9 (0.6)	99.9 (0.3)
0.25 $\tau_1 = 0.21$	100.0 (1.2)	99.9 (2.6)	99.0 (6.8)	95.7 (15.3)	83.0 (28.4)	51.1 (28.3)	19.6 (14.2)	11.9 (1.1)	34.0 (-9.1)	73.6 (-6.4)	96.2 (-0.9)	99.9 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.14$	99.9 (2.7)	99.6 (7.7)	97.7 (19.3)	89.0 (40.1)	66.6 (47.4)	37.3 (32.1)	29.3 (11.0)	53.6 (-8.2)	88.6 (-5.1)	99.4 (-0.3)	100.0 (0.0)						

$\tau_1 = \tau_{Z=1} = \tau_0^{\text{exp}(\alpha)} / \tau_0 = \tau_{Z=0} = 0.3$

TABLEAU 6.2 – Modèle à risque cumulatif borné, censure exponentielle,  $\tau_0 = 50\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.66$	96.5 (-0.9)	95.7 (-1.3)	95.3 (-1.6)	94.3 (-2.2)	92.9 (-2.0)	90.8 (-3.1)	88.1 (-4.8)	84.5 (-6.1)	79.4 (-7.5)	75.2 (-8.7)	71.5 (-8.8)	68.3 (-8.1)	66.5 (-7.0)	65.2 (-7.4)	64.3 (-6.4)	64.1 (-5.8)	62.9 (-5.6)
-0.25 $\tau_1 = 0.58$	97.4 (1.4)	96.0 (1.5)	94.0 (1.5)	91.2 (1.6)	83.5 (0.1)	73.6 (-2.5)	59.2 (-5.2)	41.9 (-8.0)	26.4 (-7.8)	17.2 (-4.3)	14.8 (2.1)	16.7 (8.6)	20.4 (15.0)	25.5 (20.3)	29.8 (25.2)	34.7 (28.5)	37.4 (30.8)
0	97.2 (4.5)	95.3 (6.2)	91.8 (8.5)	84.4 (9.4)	70.2 (11.2)	49.8 (9.9)	25.9 (4.2)	10.0 (0.5)	4.6	10.2 (0.7)	26.6 (2.7)	49.8 (8.2)	71.8 (11.5)	84.6 (10.7)	91.8 (7.6)	95.1 (6.4)	96.9 (4.4)
0.25 $\tau_1 = 0.41$	97.0 (9.2)	93.7 (14.0)	88.4 (20.1)	77.1 (25.6)	55.9 (25.6)	30.0 (17.0)	13.1 (8.1)	11.4 (0.5)	26.7 (-8.3)	57.6 (-8.0)	87.0 (-2.6)	97.8 (0.1)	99.8 (0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.32$	96.0 (16.7)	92.4 (25.5)	84.0 (35.6)	68.5 (41.8)	45.3 (36.5)	27.3 (22.0)	24.4 (6.0)	44.5 (-8.4)	79.0 (-8.1)	96.2 (-2.0)	99.9 (-0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\text{exp}(\alpha)} / \tau_0 = \tau_{Z=0} = 0.5$

TABLEAU 6.3 – Modèle à risque cumulatif borné, censure exponentielle,  $\tau_0 = 70\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.81$	69.9 (-7.7)	70.7 (-8.5)	68.2 (-8.7)	67.6 (-8.6)	64.4 (-9.5)	64.3 (-9.2)	61.6 (-10.5)	59.5 (-10.4)	56.3 (-11.0)	55.2 (-10.3)	52.9 (-10.5)	50.9 (-11.2)	49.7 (-10.3)	49.8 (-10.2)	49.2 (-9.8)	49.1 (-10.5)	48.9 (-9.7)
-0.25 $\tau_1 = 0.76$	73.2 (-0.3)	69.9 (-1.5)	64.1 (-3.1)	58.6 (-3.2)	52.9 (-3.7)	42.8 (-5.3)	33.1 (-5.7)	23.8 (-7.9)	17.2 (-6.1)	12.8 (-3.2)	10.9 (-0.5)	10.4 (1.5)	11.2 (4.0)	12.8 (6.7)	13.5 (7.8)	13.8 (8.6)	15.2 (10.1)
0	75.0 (7.4)	70.1 (9.4)	62.1 (8.0)	51.4 (6.9)	38.8 (4.6)	26.3 (2.6)	13.8 (0.7)	7.4 (-0.1)	4.3	6.1 (-1.0)	13.5 (0.9)	25.3 (2.3)	38.8 (5.0)	52.2 (6.4)	64.7 (7.9)	70.1 (7.5)	74.1 (8.0)
0.25 $\tau_1 = 0.63$	75.4 (18.1)	67.4 (16.8)	58.4 (18.6)	43.9 (16.8)	28.8 (13.0)	16.5 (9.1)	9.2 (4.3)	8.7 (0.0)	17.3 (-5.8)	36.3 (-7.8)	61.8 (-7.1)	81.7 (-2.9)	94.1 (-0.8)	98.3 (0.0)	99.4 (-0.1)	99.6 (0.0)	99.9 (0.0)
0.5 $\tau_1 = 0.56$	73.5 (26.1)	64.6 (27.2)	53.4 (28.1)	37.9 (24.1)	24.3 (18.1)	16.1 (9.9)	17.7 (2.2)	29.5 (-6.9)	55.7 (-9.6)	81.6 (-6.5)	95.9 (-1.7)	99.6 (-0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\text{exp}(\alpha)} / \tau_0 = \tau_{Z=0} = 0.7$

CHAPITRE 6. TEST DE DONNÉES CENSURÉES PAR INTERVALLE, POPULATION À RISQUES HÉTÉROGÈNES

TABLEAU 6.4 – Modèle à risque cumulatif borné, censure exponentielle,  $\tau_0 = 30\%$ ,  $p = 40\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.48$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.9 (-0.1)	97.2 (-1.3)	80.7 (-7.3)	48.5 (-8.5)	27.7 (6.3)	27.9 (21.9)	46.5 (38.0)	68.7 (45.3)	83.8 (39.6)	92.6 (28.4)	96.2 (20.0)
-0.25 $\tau_1 = 0.39$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.8 (-0.1)	96.5 (-1.1)	70.8 (-7.1)	26.7 (-8.1)	8.6 (1.6)	20.6 (8.9)	58.5 (12.3)	88.9 (8.7)	98.4 (2.2)	99.9 (0.4)	100.0 (0.1)	100.0 (0.0)
0	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.1)	97.3 (0.2)	72.7 (-2.9)	24.0 (-2.3)	4.7	22.9 (-3.9)	72.2 (-3.1)	97.5 (0.2)	99.9 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.25 $\tau_1 = 0.21$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	98.8 (0.3)	80.7 (3.8)	30.6 (4.1)	6.8 (2.0)	27.1 (-7.4)	79.4 (-6.3)	98.9 (-0.4)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.14$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.5 (0.5)	89.2 (4.6)	43.8 (10.0)	11.5 (6.3)	26.1 (-6.1)	79.5 (-7.2)	98.9 (-0.6)	100.0 (0.0)						

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.3$

TABLEAU 6.5 – Modèle à risque cumulatif borné, censure exponentielle,  $\tau_0 = 50\%$ ,  $p = 40\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.66$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.7 (-0.1)	97.1 (-1.1)	87.6 (-5.1)	66.0 (-10.0)	41.8 (-8.4)	26.8 (1.2)	23.6 (11.1)	27.5 (21.9)	36.3 (30.3)	47.4 (38.1)	56.5 (40.6)	62.1 (41.1)
-0.25 $\tau_1 = 0.58$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.6 (-0.2)	97.0 (-0.7)	83.6 (-3.6)	51.3 (-8.0)	20.9 (-6.5)	8.4 (0.7)	12.5 (6.2)	33.0 (11.5)	60.2 (15.0)	82.1 (11.1)	93.4 (7.3)	97.4 (3.0)	98.9 (1.5)
0	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.7 (-0.1)	97.4 (0.3)	84.4 (0.5)	48.6 (-3.4)	15.8 (-2.2)	4.3	14.8 (-2.8)	49.1 (-2.5)	84.1 (-0.4)	97.7 (0.3)	99.9 (0.2)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.25 $\tau_1 = 0.41$	100.0 (0.0)	100.0 (0.0)	99.9 (0.0)	98.4 (0.5)	88.5 (1.9)	57.6 (4.8)	19.2 (3.0)	6.7 (1.0)	20.6 (-7.1)	62.6 (-7.3)	93.2 (-2.7)	99.4 (-0.3)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.32$	100.0 (0.0)	99.9 (0.1)	99.2 (0.4)	93.1 (4.4)	66.5 (9.1)	26.9 (10.0)	10.6 (5.7)	22.6 (-6.3)	66.2 (-10.2)	94.8 (-2.3)	99.9 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.5$

TABLEAU 6.6 – Modèle à risque cumulatif borné, censure exponentielle,  $\tau_0 = 70\%$ ,  $p = 40\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.81$	99.9 (0.0)	99.9 (-0.1)	99.8 (0.1)	99.1 (-0.3)	97.5 (-0.8)	91.7 (-2.6)	80.5 (-5.7)	63.1 (-9.8)	43.9 (-10.9)	28.8 (-7.8)	18.5 (-3.0)	16.0 (3.8)	15.1 (7.8)	18.4 (12.7)	21.4 (16.8)	23.4 (18.4)	27.5 (20.8)
-0.25 $\tau_1 = 0.76$	100.0 (0.0)	99.9 (0.0)	99.0 (-0.2)	97.1 (-0.4)	91.1 (-1.6)	76.6 (-5.6)	54.5 (-6.2)	29.1 (-8.1)	13.9 (-4.6)	6.7 (-0.6)	7.6 (2.8)	16.1 (5.5)	29.9 (8.4)	46.2 (9.2)	60.3 (10.1)	72.4 (8.9)	79.6 (7.3)
0	99.9 (0.1)	99.5 (0.3)	97.6 (-0.2)	92.8 (-0.7)	78.4 (-0.6)	54.5 (-2.1)	26.8 (-3.5)	10.1 (-1.7)	4.6	10.8 (-0.9)	27.0 (-2.9)	53.6 (-3.8)	78.2 (-1.4)	92.0 (0.2)	97.9 (0.0)	99.5 (0.0)	99.9 (0.0)
0.25 $\tau_1 = 0.63$	99.5 (0.1)	98.0 (0.3)	93.8 (0.1)	82.0 (0.7)	60.1 (1.6)	30.0 (1.8)	11.9 (2.0)	6.2 (0.4)	14.1 (-4.5)	37.0 (-9.0)	71.0 (-6.8)	92.6 (-2.2)	98.8 (-0.4)	99.9 (-0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.56$	98.8 (0.3)	95.4 (0.5)	86.7 (3.3)	65.7 (4.5)	37.6 (6.7)	14.8 (5.2)	8.3 (2.9)	17.1 (-4.2)	45.8 (-9.4)	76.4 (-7.9)	95.3 (-2.3)	99.5 (-0.3)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.7$

TABLEAU 6.7 – Modèle à risque cumulatif borné, censure exponentielle, covariable déséquilibrée,  $\tau_0 = 30\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.3$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.48$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.9 (-0.1)	99.6 (-0.1)	97.9 (-0.9)	92.0 (-3.2)	82.0 (-6.9)	69.9 (-8.5)	63.6 (-5.0)	59.3 (0.5)	58.0 (8.1)	58.4 (13.4)	59.9 (18.9)	61.4 (24.0)	62.3 (25.6)
-0.25 $\tau_1 = 0.39$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.9 (0.0)	99.6 (0.1)	96.1 (0.2)	82.8 (-2.4)	54.5 (-7.9)	30.0 (-8.0)	17.4 (-2.1)	17.4 (8.7)	24.5 (18.5)	35.1 (29.5)	45.3 (38.5)	53.4 (45.1)	59.0 (49.0)	62.0 (50.8)
0	100.0 (0.0)	100.0 (0.0)	100.0 (0.1)	99.5 (0.5)	95.9 (2.9)	78.0 (6.0)	42.5 (4.1)	13.6 (1.0)	4.6	12.5 (1.4)	34.2 (6.9)	60.4 (15.8)	79.3 (17.8)	88.5 (15.9)	93.8 (14.9)	96.0 (12.4)	97.0 (11.6)
0.25 $\tau_1 = 0.21$	100.0 (0.0)	99.9 (0.1)	99.7 (0.8)	97.3 (6.2)	82.4 (14.6)	46.5 (19.6)	15.7 (9.6)	10.4 (0.4)	28.1 (-7.4)	61.9 (-7.3)	89.1 (-0.9)	97.8 (0.7)	99.5 (0.3)	99.9 (0.1)	99.9 (0.0)	100.0 (0.1)	100.0 (0.0)
0.5 $\tau_1 = 0.14$	100.0 (0.0)	99.9 (0.6)	98.2 (5.8)	89.2 (21.7)	58.9 (34.2)	29.1 (24.5)	22.2 (5.8)	46.4 (-9.5)	83.2 (-6.8)	97.8 (-1.0)	99.9 (-0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.3$

TABLEAU 6.8 – Modèle à risque cumulatif borné, censure exponentielle, covariable déséquilibrée,  $\tau_0 = 30\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.7$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.48$	92.2 (-3.1)	91.9 (-3.0)	92.2 (-2.9)	91.9 (-3.7)	90.8 (-4.1)	90.4 (-4.0)	88.4 (-5.3)	88.6 (-4.6)	85.9 (-6.3)	84.9 (-6.0)	83.0 (-6.3)	79.1 (-7.2)	76.6 (-7.9)	76.8 (-6.7)	75.2 (-6.9)	73.2 (-6.7)	72.9 (-6.9)
-0.25 $\tau_1 = 0.39$	95.4 (4.3)	94.3 (3.9)	93.3 (3.4)	90.3 (2.7)	84.9 (1.9)	76.5 (-0.5)	63.1 (-4.1)	46.9 (-7.6)	29.6 (-8.0)	16.9 (-4.1)	13.8 (4.8)	20.5 (15.7)	33.3 (26.8)	49.8 (35.6)	64.5 (38.8)	76.2 (37.9)	82.5 (32.7)
0	96.6 (10.9)	95.3 (12.1)	93.2 (14.3)	88.9 (16.7)	79.7 (17.4)	60.2 (14.5)	34.0 (7.4)	12.0 (1.3)	4.8	13.1 (0.3)	42.0 (3.4)	77.1 (5.5)	96.1 (3.3)	99.6 (0.7)	99.9 (0.0)	100.0 (0.0)	100.0 (0.0)
0.25 $\tau_1 = 0.21$	97.4 (18.7)	95.8 (22.1)	92.3 (29.3)	84.4 (33.2)	68.6 (35.3)	43.1 (26.7)	20.3 (13.6)	13.0 (2.1)	29.3 (-7.3)	69.2 (-6.2)	94.8 (-1.7)	99.6 (-0.2)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.14$	97.6 (28.2)	94.7 (35.0)	90.6 (44.5)	79.8 (50.6)	60.9 (47.2)	36.4 (30.7)	27.4 (10.1)	46.7 (-6.0)	82.1 (-6.4)	98.2 (-0.7)	100.0 (0.0)						

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.3$

TABLEAU 6.9 – Modèle à risque cumulatif borné, censure uniforme et suivi insuffisant,  $\tau_0 = 30\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.48$	100.0 (0.0)	98.3 (-0.9)	75.4 (-8.0)	26.5 (-6.3)	7.5 (1.6)	19.2 (4.5)	51.7 (5.1)	79.6 (5.8)	90.7 (4.6)	96.6 (3.3)	98.4 (2.1)						
-0.25 $\tau_1 = 0.39$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	98.5 (-0.8)	78.8 (-6.2)	24.5 (-8.4)	5.3 (0.3)	23.4 (-2.2)	68.7 (-3.1)	93.8 (-1.0)	99.1 (0.0)	99.9 (0.1)	100.0 (0.0)	100.0 (0.0)
0	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	98.2 (-0.7)	78.7 (-6.1)	26.8 (-7.3)	4.8	25.7 (-6.9)	79.6 (-5.8)	98.4 (-0.5)	99.9 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.25 $\tau_1 = 0.21$	100.0 (0.0)	100.0 (0.0)	100.0 (0.1)	99.7 (0.1)	96.9 (-0.3)	76.4 (-2.2)	28.4 (-1.1)	5.8 (0.8)	28.6 (-8.7)	85.1 (-5.2)	99.4 (-0.4)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.14$	99.9 (0.6)	99.7 (1.0)	98.5 (2.3)	93.5 (6.1)	70.9 (9.0)	28.7 (7.8)	8.9 (3.3)	33.6 (-7.8)	87.8 (-5.1)	99.7 (-0.2)	100.0 (0.0)						

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.3$

## D. Commentaires

Dans toutes les situations considérées ici, sous le modèle cumulatif borné (Tableaux 6.1 - 6.6), les valeurs estimées du risque de première espèce pour le test de score proposé, sous l'hypothèse nulle ( $\alpha = \beta = 0$ ), appartiennent à l'intervalle de fluctuation du risque nominal basé sur la distribution du khi-deux avec deux degrés de liberté qui est  $[2, 86; 7, 14]$  pour un échantillon de 400 individus.

Comme on le voit dans les tableaux 6.1 - 6.3 (20 % censure, covariable équilibré), le test de score proposé est plus puissant que le test de pseudo log-rank dans la plupart des scénarios. Cependant, comme prévu, lorsque  $\beta = 0$ , le modèle à plateau correspond à un modèle à risque proportionnel<sup>4</sup> et le test du logrank est dans ce cas toujours plus puissant. Ce dernier suit en effet sous l'hypothèse nulle, la loi d'un khi-deux à un degré de liberté, tandis que le test de score proposé, avec ses deux paramètres, suit sous l'hypothèse nulle une loi du khi-deux à deux degrés de liberté.

Le tableau 6.1 montre que la puissance du test de score proposé peut être jusqu'à 53,1% plus élevée que celle du test du logrank, tandis qu'elle n'est jamais inférieure de plus de 9,6 % à cette dernière. Dans les tableaux 6.2 et 6.3, la puissance du test de score proposé est la plus élevée, respectivement, de 41,8% et de 28,1% et pas inférieure de plus de 8,8% et de 11,2% à la puissance du test du log-rank. Dans toutes les configurations étudiées, la puissance du test diminue à mesure que la valeur du plateau augmente. Cette puissance réduite n'est pas surprenante, car moins d'événements sont attendus avec une valeur de plateau plus élevée (comme dans les configurations présentées dans les tableaux 6.2 et 6.3). De plus, pour les configurations avec des valeurs faibles de  $\beta$ , la puissance du score proposé est plus faible que celle du test du log-rank, comme rapporté ci-dessus avec  $\beta = 0$ .

Il convient de noter que pour toute combinaison donnée de  $\alpha$  et  $\beta$ , la puissance du test de score proposé est plus élevée que celle du test du log-rank lorsque  $\alpha \beta > 0$ ; par ailleurs, les différences de puissance entre les deux tests sont les plus élevées en faveur du test de score proposé lorsque  $\alpha \beta < 0$ .

Ces tendances de puissance en faveur du test de score proposé peuvent être expliquées par le modèle à partir duquel le test est dérivé : il inclut deux paramètres qui ne sont liés qu'à une seule covariable. Cela signifie que, dans une certaine mesure, les deux paramètres sont en concurrence l'un avec l'autre. De plus, le développement limité de premier ordre de  $\beta$  autour de zéro nous donne :  $\lambda_{Z=1}(t) \approx \lambda_{Z=0}(t)e^{\alpha+\beta(1-H(t))}$ . Nous avons alors un rapport de risque dépendant du temps. Globalement, cela signifie qu'à des temps d'événement précoces (lorsque  $H(t) < 1$ ),  $\alpha$  et  $\beta$  se compensent l'un l'autre lorsqu'ils ont des signes opposés.

Les différences de puissance élevées observées lorsque  $\alpha \beta < 0$  peuvent aussi s'expliquer par le fait que ces configurations ne sont clairement pas favorables au test du log-rank, car les courbes de survie se croisent. Par exemple, lorsque  $\alpha$  est négatif et  $\beta$  est positif, le groupe de référence a une valeur de plateau inférieure et un risque de faire l'événement plus faible que le second groupe comme cela est illustré sur la figure 6.1 ci-dessous :

À partir de tableaux 6.4 - 6.6, où la censure à droite est de 40 %, nous observons les mêmes tendances en comparant les puissances des deux tests. Notons que la puissance du test de score proposé par rapport à celle du test du log-rank est légèrement inférieure à ce qu'elle était dans le même scénario avec 20% de censure à droite. Le gain de puissance est, cependant, toujours intéressant puisqu'il peut être jusqu'à 45,3 % supérieur et pas plus de 8,5 % inférieur ( $\tau_0 = 30\%$ ).

En général, pour une configuration donnée, la puissance du test de score proposé est plus élevée avec un taux de censure à droite de 40% comparativement à celle observée avec une censure de 20%. Cette augmentation de puissance observée lorsque le taux de censure augmente peut s'expliquer par la forme du rapport de risque, lorsque nous avons un mélange de contributions positives et négatives associées au  $\beta$  et dépendantes du schéma de censure.

4. Nous pouvons d'ailleurs remarquer que l'expression du score correspondant au paramètre  $\alpha$  est identique à celle du log-rank.

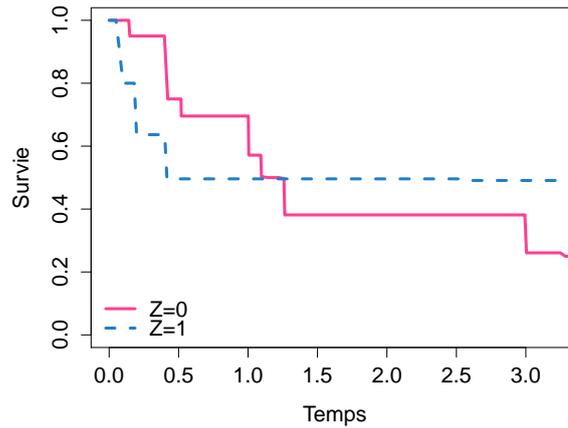


FIGURE 6.1 – Données simulées censurées par intervalle de type 1 comportant une fraction de non-susceptibles, plateau du groupe de référence  $\tau = 0,4$ ,  $n = 400$ , avec  $\beta = 2$ ,  $\alpha = -0,25$

En analysant les scénarios avec covariable déséquilibrée ( $\xi = 0.3$  et  $\xi = 0.7$ ) dans les tableaux 6.7 - 6.8, nous notons que les différences de puissance sont assez semblables à celles observées dans les scénarios avec covariable équilibrée. Le test de score proposé a de meilleures performances avec  $\xi = 0,3$  que dans les situations avec  $\xi = 0,7$ , selon le signe de  $\alpha \beta$ .

Pour tous les scénarios simulés sous un modèle de mélange (Annexe, section C page IX, tableaux C.1 - C.8), les valeurs estimées du risque de première espèce du test de score proposé, sous l'hypothèse nulle ( $\alpha = \beta = 0$ ), appartiennent également à l'intervalle de fluctuation du risque nominal basé sur la distribution du khi-deux avec deux degrés de liberté. En observant les performances du test proposé (sa puissance) et les différences de puissance par rapport au test du log-rank, nous voyons que les résultats sont légèrement meilleurs que ceux obtenus sous le modèle à risque cumulatif borné. À titre d'exemple, dans le tableau C.1 (Annexe) avec  $p = 20\%$ ,  $\tau_0 = 30\%$ ,  $\xi = 0,5$ , la puissance du test de score proposé, comparée à celle du test de pseudo log-rank, peut être jusqu'à 58,1 % plus élevée et pas plus de 8,7 % inférieure.

Dans le cas où la condition de suivi suffisant n'a pas été respectée (tableau 6.9), le taux d'erreur de type I est proche du niveau de signification nominal (4,8%) et il appartient à son intervalle de fluctuation. Les performances du test de score proposé sont légèrement inférieures à celles obtenues avec un suivi suffisant et des taux de censure comparables. De plus, les gains de puissance du test de score proposé par rapport au test du log-rank sont également plus faibles dans ce cas.

## IV. Discussion

L'analyse de données de populations comportant une fraction non à risque nécessite une extension des statistiques de test développées pour le cas de la censure aléatoire droite, pour les données de type *current status* et censurées par intervalle; les statistiques classiques n'étant pas conçues pour faire face à des populations mixtes.

Les modèles à plateau avec censure par intervalle ont été peu étudiés. Nous pouvons citer comme exemples LAM et al. [2013]; LIU et SHEN [2009] qui établissent une régression semi-paramétrique pour censure par intervalle dans le cadre d'un modèle à risque cumulatif borné et XIANG et al. [2011] dans le cadre d'un modèle de mélange.

Concernant les données censurées par intervalle de type 1, les études sur les modèles à plateau sont encore plus limitées. À titre d'exemple, nous pouvons citer MA [2008], dans lequel la modélisation du temps avant événement dans la population susceptible est celle d'un modèle à risque additif et MA [2009], dans lequel la survie est modélisée à l'aide d'un modèle de Cox. Dans ces deux articles, l'auteur a utilisé un modèle de mélange.

Les approches pour modèles à plateau adaptées aux données censurées par intervalle sont généralement orientées vers l'estimation et ne proposent pas de solutions simples pour tester l'hypothèse d'absence d'effet sur le plateau et sur le délai avant survenue de l'événement chez les susceptibles. Ces procédures sont complexes, parfois inefficaces ou ad-hoc pour l'estimation de la variance. À notre connaissance, aucune de ces approches, avec une option clairement établie de test d'hypothèse, n'a fait l'objet d'une implémentation

Dans les sections précédentes, nous avons proposé un nouveau test du score permettant de comparer deux groupes basé sur un modèle de survie impropre; le test est adapté pour détecter les écarts par rapport à l'égalité des distributions de survie dans les populations de patients mixtes. Ce test est une extension du test de BROËT et al. [2001], destiné à la censure aléatoire droite. De plus, le choix d'utiliser un modèle à risque cumulatif borné a été motivé par son interprétation mécanistique intéressante dans le cadre de l'immunogénicité des anticorps anti-biothérapies.

Le test du score proposé est conçu pour détecter des différences entre les deux groupes soit dans la proportion de patients non-susceptibles de faire l'événement (c'est-à-dire, la valeur du plateau), soit dans la distribution des temps de survie chez les patients susceptibles. Comme le montrent nos simulations, sous l'hypothèse nulle, le test proposé a une erreur de type I correcte dans toutes les configurations étudiées. De plus, en analysant les résultats de simulation sous les hypothèses alternatives, nous voyons que la puissance du test proposé est meilleure que celle du test de type log-rank dans la plupart des situations. À titre d'exemple, avec un plateau de 30 % et un taux de censure de 20 %, la puissance du test proposé peut être jusqu'à 53 % plus importante, tout en n'étant jamais inférieure à plus de 10 % à celle du test du log-rank.

Des simulations préliminaires, non présentées ici, réalisées sur des données censurées par intervalle de type II montraient pour le test développé un risque de première espèce estimé appartenant à l'intervalle de fluctuation d'un risque nominal de 5% obtenu à partir de la loi d'un  $\chi^2$  à deux degrés de liberté. Cependant, le test pour données censurées par intervalle nécessite des évaluations supplémentaires pour ses performances en termes de puissance.

Notre test est dérivé d'un modèle qui prend en charge deux paramètres pour une seule covariable. Ainsi, dans une certaine mesure, les deux paramètres sont en concurrence l'un avec l'autre; ce qui signifie que l'effet d'un paramètre peut annuler l'effet du second paramètre. Ceci explique pourquoi la puissance du test de score proposé est plus faible dans les situations où il y a un effet directionnel opposé de la covariable testée sur la distribution du temps avec événement et sur la fraction des non-susceptibles. Cela se traduit ainsi : le fait d'appartenir au groupe 1, comparé au groupe de référence, entraîne un risque plus faible d'être susceptible, mais un risque plus élevé de faire l'événement pour ceux qui sont susceptibles. Cette situation est peu réalisée en pratique puisqu'elle suppose qu'une variable favoriserait le fait d'être non à risque, tout en augmentant le risque de faire l'événement.

Les résultats de la simulation montrent également que lorsque la censure est élevée chez les individus susceptibles, les gains de puissance par rapport au test de type log-rank sont préservés. Lorsque la condition de suivi suffisante n'est pas vérifiée — c'est-à-dire lorsque la durée du suivi est plus courte que la fenêtre de temps pendant laquelle un événement peut survenir — la puissance du test est réduite. Dans ce cas, son taux d'erreur de type I estimé est proche du niveau nominal.

Une limitation du test proposé est qu'il ne peut détecter qu'une différence globale entre les groupes; il ne peut pas distinguer de changements dans la fraction de non-susceptibles ou dans la distribution de survie parmi les individus susceptibles. Ainsi, d'autres travaux supplémentaires sont nécessaires pour dériver des solutions simples pour résoudre ce problème. Néanmoins, compte tenu de l'amélioration de la puissance obtenue lorsque la distribution de survie est modifiée parmi les individus susceptibles, l'utilisation du test de score proposé peut être recommandée lorsqu'une fraction d'individus non-susceptibles est attendue.

D'un point de vue pratique, il convient de garder à l'esprit que ce test nécessite une durée de suivi suffisante, c'est-à-dire avec une fenêtre d'observation suffisamment longue pour que le dernier événement potentiel se produise. En d'autres termes, il faut fournir un suivi adéquat pour détecter la présence d'individus non-susceptibles dans la population étudiée.

Comme pour le test de type log-rank classique, le test du score proposé repose sur l'hypothèse que la distribution des temps de visite est la même pour les différents groupes à comparer. Il convient également de noter que le test proposé peut être étendu pour prendre en compte d'autres facteurs, en développant une version stratifiée avec des strates définies par les niveaux des facteurs.

En conclusion, nous pensons que le test présenté dans les sections précédentes est à la fois facile à mettre en œuvre (le code *R* pour le test adapté aux données censurées par intervalle de type I est disponible en Annexe, section B page V) et intéressant pour les études transversales de survie avec des populations mixtes.

## V. Références

- BERKSON, J. et R. P. GAGE. 1952, «Survival curve for cancer patients following treatment», *Journal of the American Statistical Association*, vol. 47, n° 259, doi :10.1080/01621459.1952.10501187, p. 501–515. URL <http://www.tandfonline.com/doi/abs/10.1080/01621459.1952.10501187>. 122
- BOAG, J. W. 1949, «Maximum likelihood estimates of the proportion of patients cured by cancer therapy», *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 11, n° 1, p. 15–53, ISSN 00359246. URL <http://www.jstor.org/stable/2983694>. 122
- BROET, P., A. TSODIKOV, Y. DE RYCKE et T. MOREAU. 2004, «Two-sample statistics for testing the equality of survival functions against improper semi-parametric accelerated failure time alternatives : an application to the analysis of a breast cancer clinical trial», *Lifetime Data Anal.*, vol. 10, n° 2, p. 103–120. 122
- BROËT, P., Y. RYCKE, P. TUBERT-BITTER, J. LELLOUCH, B. ASSELAIN et T. MOREAU. 2001, «A semi-parametric approach for the two-sample comparison of survival times with long-term survivors», *Biometrics*, vol. 57, n° 3, doi :10.1111/j.0006-341x.2001.00844.x, p. 844–852. URL <https://doi.org/10.1111/j.0006-341x.2001.00844.x>. 123, 135
- CHEN, C.-M., P.-S. SHEN, J. C.-C. WEI et L. LIN. 2017, «A semiparametric mixture cure survival model for left-truncated and right-censored data», *Biometrical Journal. Biometrische Zeitschrift*, vol. 59, n° 2, doi :10.1002/bimj.201500267, p. 270–290, ISSN 1521-4036. 122
- CHEN, M.-H., J. G. IBRAHIM et D. SINHA. 1999, «A new bayesian model for survival data with a surviving fraction», *Journal of the American Statistical Association*, vol. 94, n° 447, p. 909–919, ISSN 01621459. URL <http://www.jstor.org/stable/2670006>. 122
- COONER, F., S. BANERJEE, B. P. CARLIN et D. SINHA. 2007, «Flexible Cure Rate Modeling Under Latent Activation Schemes», *Journal of the American Statistical Association*, vol. 102, n° 478, doi :10.1198/016214507000000112, p. 560–572, ISSN 0162-1459 1537-274X. 122
- FAREWELL, V. T. 1982, «The use of mixture models for the analysis of survival data with long-term survivors», *Biometrics*, vol. 38, n° 4, p. 1041–1046, ISSN 0006341X, 15410420. URL <http://www.jstor.org/stable/2529885>. 122
- FAREWELL, V. T. 1986, «Mixture models in survival analysis : Are they worth the risk?», *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, vol. 14, n° 3, p. 257–262, ISSN 03195724. URL <http://www.jstor.org/stable/3314804>. 122
- FAY, M. P. 1996, «Rank invariant tests for interval censored data under the grouped continuous model», *Biometrics*, vol. 52, n° 3, p. 811–822, ISSN 0006-341X. 126, 127
- GOLDMAN, A. I. 1984, «Survivorship analysis when cure is a possibility : a Monte Carlo study», *Statistics in Medicine*, vol. 3, n° 2, p. 153–163, ISSN 0277-6715. 122
- GRAY, R. J. et A. A. TSIATIS. 1989, «A linear rank test for use when the main interest is in differences in cure rates», *Biometrics*, vol. 45, n° 3, p. 899–904, ISSN 0006-341X. 122
- GREENHOUSE, J. B. et R. A. WOLFE. 1984, «A competing risks derivation of a mixture model for the analysis of survival data», *Communications in Statistics - Theory and Methods*, vol. 13, n° 25, doi :10.1080/03610928408828883, p. 3133–3154, ISSN 0361-0926, 1532-415X. URL <http://www.tandfonline.com/doi/abs/10.1080/03610928408828883>. 122
- HALPERN, J. et B. W. BROWN. 1987, «Cure rate models : power of the logrank and generalized Wilcoxon tests», *Statistics in Medicine*, vol. 6, n° 4, p. 483–489, ISSN 0277-6715. 122

- IBRAHIM, J. G., M.-H. CHEN et D. SINHA. 2001a, «Bayesian semiparametric models for survival data with a cure fraction», *Biometrics*, vol. 57, n° 2, p. 383–388, ISSN 0006341X, 15410420. URL <http://www.jstor.org/stable/3068343>. 122
- IBRAHIM, J. G., M.-H. CHEN et D. SINHA. 2001b, «Cure Rate Models», dans *Bayesian Survival Analysis*, Springer Series in Statistics, Springer New York, New York, NY, ISBN 978-1-4419-2933-4 978-1-4757-3447-8, p. 155–207. URL <http://link.springer.com/10.1007/978-1-4757-3447-8>, DOI : 10.1007/978-1-4757-3447-8. 123
- JONAS, S. F., C. MBOGNING, S. HÄSSLER et P. BROËT. 2017, «A score test for comparing cross-sectional survival data with a fraction of non-susceptible patients and its application in clinical immunology», *PLOS ONE*, vol. 12, n° 6, doi :10.1371/journal.pone.0179896, p. e0179896. URL <https://doi.org/10.1371/journal.pone.0179896>. 123
- KALBFLEISCH, J. D. et R. L. PRENTICE. 1980, *The statistical analysis of failure time data*, 15<sup>e</sup> éd., Wiley series in probability and mathematical statistics, Wiley, New York, NY, ISBN 978-0-471-05519-8. 126, 127
- KUK, A. Y. C. et C.-H. CHEN. 1992, «A mixture model combining logistic regression with proportional hazards regression», *Biometrika*, vol. 79, n° 3, p. 531–541, ISSN 00063444. URL <http://www.jstor.org/stable/2336784>. 122
- LAM, K. F., D. Y. T. FONG et O. Y. TANG. 2005, «Estimating the proportion of cured patients in a censored sample», *Statistics in Medicine*, vol. 24, n° 12, doi :10.1002/sim.2137, p. 1865–1879, ISSN 1097-0258. URL <http://dx.doi.org/10.1002/sim.2137>. 122
- LAM, K. F., K. Y. WONG et F. ZHOU. 2013, «A semiparametric cure model for interval-censored data», *Biometrical Journal*, vol. 55, n° 5, doi :10.1002/bimj.201300004, p. 771–788, ISSN 1521-4036. URL <http://dx.doi.org/10.1002/bimj.201300004>. 134
- LASKA, E. M. et M. J. MEISNER. 1992, «Nonparametric Estimation and Testing in a Cure Model», *Biometrics*, vol. 48, n° 4, doi :10.2307/2532714, p. 1223, ISSN 0006341X. URL <http://www.jstor.org/stable/2532714?origin=crossref>. 122
- LEEuw, J. D., K. HORNIK et P. MAIR. 2009, «Isotone Optimization in R : Pool-Adjacent-Violators Algorithm (PAVA) and Active Set Methods», *Journal of Statistical Software*, vol. 32, n° 5, doi : 10.18637/jss.v032.i05, ISSN 1548-7660. URL <http://www.jstatsoft.org/v32/i05/>. 129
- LIU, H. et Y. SHEN. 2009, «A semiparametric regression cure model for interval-censored data», *Journal of the American Statistical Association*, vol. 104, n° 487, p. 1168–1178, ISSN 01621459. URL <http://www.jstor.org/stable/40592284>. 134
- MA, S. 2008, «Additive risk model for current status data with a cured subgroup», *Annals of the Institute of Statistical Mathematics*, vol. 63, n° 1, doi :10.1007/s10463-008-0212-2, p. 117–134. URL <https://doi.org/10.1007/s10463-008-0212-2>. 134
- MA, S. 2009, «Cure model with current status data», *Statistica Sinica*, vol. 19, n° 1, p. 233–249, ISSN 10170405, 19968507. URL <http://www.jstor.org/stable/24308717>. 134
- MALLER, R. A. et S. ZHOU. 1995, «Testing for the presence of immune or cured individuals in censored survival data», *Biometrics*, vol. 51, n° 4, doi :10.2307/2533253, p. 1197, ISSN 0006341X. URL <http://www.jstor.org/stable/2533253?origin=crossref>. 122, 126
- MALLER, R. A. et X. ZHOU. 1996, *Survival analysis with long-term survivors*, Wiley series in probability and statistics, Wiley, Chichester, ISBN 978-0-471-96201-4. OCLC : 833038479. 122

- PENG, Y. et J. G. TAYLOR. 2013, «Cure Models», dans *Handbook of Survival Analysis*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods, Chapman and Hall/CRC, ISBN 978-1-4665-5566-2, p. 113–134. URL <http://dx.doi.org/10.1201/b16248-8>, DOI : 10.1201/b16248-8 DOI : 10.1201/b16248-8. 122
- ROBERTSON, T., F. T. WRIGHT et R. L. DYKSTRA. 1988, *Order restricted statistical inference*, repr éd., Wiley series in probability and mathematical statistics, Wiley, Chichester, ISBN 978-0-471-91787-8. OCLC : 258556159. 125
- SPOSTO, R., H. N. SATHER et S. A. BAKER. 1992, «A comparison of tests of the difference in the proportion of patients who are cured», *Biometrics*, vol. 48, n° 1, p. 87–99, ISSN 0006-341X. 122
- SY, J. P. et J. M. TAYLOR. 2000, «Estimation in a Cox proportional hazards cure model», *Biometrics*, vol. 56, n° 1, p. 227–236, ISSN 0006-341X. 122
- TAYLOR, J. M. G. 1995, «Semi-parametric estimation in failure time mixture models», *Biometrics*, vol. 51, n° 3, p. 899–907, ISSN 0006341X, 15410420. URL <http://www.jstor.org/stable/2532991>. 122
- TURNBULL, B. W. 1976, «The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data», *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 38, n° 3, p. 290–295. URL <http://www.jstor.org/stable/2984980>. 125
- XIANG, L., X. MA et K. K. W. YAU. 2011, «Mixture cure model with random effects for clustered interval-censored survival data», *Statistics in Medicine*, vol. 30, n° 9, doi :10.1002/sim.4170, p. 995–1006, ISSN 1097-0258. URL <http://dx.doi.org/10.1002/sim.4170>. 134
- YAKOVLEV, A. Y., B. ASSELAIN, V.-J. BARDOU, A. FOURQUET, T. HOANG, A. ROCHEFODIÈRE et A. D. TSODIKOV. 1993, «A simple stochastic model of tumor recurrence and its application to data on premenopausal breast cancer», , n° 12, p. 66–82. 123
- YAKOVLEV, A. Y. et A. D. TSODIKOV. 1996, *Stochastic models of tumor latency and their biostatistical applications*, n° 1 dans Series in mathematical biology and medicine, World Scientific, Singapore, ISBN 978-981-02-1831-7. OCLC : 832737186. 123



## Chapitre 7

# Application sur données d'immunologie clinique censurées par intervalle de type I

### Sommaire

---

<b>I. Introduction</b> . . . . .	<b>142</b>
<b>II. La sclérose en plaques</b> . . . . .	<b>142</b>
<b>III. Cohorte Munich</b> . . . . .	<b>143</b>
A. Description de la cohorte . . . . .	143
B. Test de l'influence du sexe et de l'âge sur le délai avant apparition des ADA . . . . .	144
<b>IV. Cohorte Düsseldorf</b> . . . . .	<b>145</b>
A. Description de la cohorte . . . . .	145
B. Test de l'influence du sexe sur le délai avant apparition des ADA . . . . .	145
<b>V. Discussion</b> . . . . .	<b>146</b>
<b>VI. Références</b> . . . . .	<b>147</b>

---

## I. Introduction

Le test présenté dans le chapitre précédent convient à des données *current status*, c'est-à-dire censurées par intervalle de type 1, dans lesquelles une partie de la population étudiée ne fera pas l'événement d'intérêt. Nous avons appliqué le test développé sur une des cohortes du projet ABIRISK.

Les données analysées correspondent à deux cohortes rétrospectives relatives à la sclérose en plaques mises en place lors du projet ABIRISK. Le mode de recueil des données était de type *current status* à savoir qu'un seul dosage par patient était disponible. Le dosage avait été réalisé lors d'un envoi du prélèvement au laboratoire de la ville dans laquelle s'était déroulée la visite dans le cadre d'une étude pilote sur l'immunogénicité des biothérapies.

La première étude a été réalisée dans un groupe d'étude de la ville de Munich et la seconde étude a été réalisée par le laboratoire de référence en immunologie de la ville de Düsseldorf.

L'introduction des biothérapies a été une étape cruciale dans le traitement de la sclérose en plaques (SEP), avec des produits tels que l'interféron bêta (IFN- $\beta$ ) utilisés depuis les années 1990 et le natalizumab disponible depuis 2006 dans l'Union Européenne. L'échec de la réponse au traitement peut être dû à une réponse immunitaire indésirable contre la biothérapie. Le mécanisme immunogène des biothérapies n'est pas entièrement compris et la fraction de patients développant des anticorps anti-biothérapie (ADA) dans le cadre de la SEP varie de 2 à 47%, selon les méthodes de dosage et les seuils utilisés pour définir la positivité des anticorps (DEISENHAMMER [2009]).

Dans ces cohortes, nous nous attendions à une proportion non négligeable de patients immuno-tolérants; c'est-à-dire ne développant jamais d'ADA. De plus, un seul échantillon de test pour détecter la positivité aux ADA par patient a été fourni; ce qui signifie que l'événement d'intérêt était censuré par intervalle de type 1. Le test que nous avons développé correspondait donc à l'analyse des variables collectées dans ces cohortes.

Étant dans une étude rétrospective pilote, seules quelques variables démographiques avaient été recueillies. Nous présentons par la suite les résultats de l'analyse des variables sexe et âge dans la cohorte de Munich et de la variable sexe dans celle de Düsseldorf.

## II. La sclérose en plaques

La sclérose en plaques est une maladie auto-immune qui affecte le système nerveux central. Elle est caractérisée par la démyélinisation — l'attaque de la gaine protectrice — des cellules nerveuses du cerveau et de la moelle épinière. Ces dommages perturbent la capacité de communication du système nerveux, entraînant une série de troubles physiques, mentaux et parfois psychiatriques. À long terme, ces troubles peuvent entraîner un handicap irréversible.

Les symptômes peuvent également inclure des troubles visuels (double vision, cécité dans un œil), sensitifs, de l'équilibre. Ainsi que des troubles moteurs (faiblesse musculaire, problèmes de coordination, troubles urinaires et sexuels).

La sclérose en plaques prend deux formes principales : récurrente-rémittente (85% des cas) et progressive (15% des cas). La première forme est constituée d'une alternance de phases de poussées et de rémissions. Entre les attaques, les symptômes peuvent disparaître complètement; cependant, certaines troubles neurologiques demeurent, surtout à mesure que la maladie progresse. La seconde forme de la maladie implique une absence de poussée et une évolution graduelle de la gravité des symptômes.

Bien que la cause ne soit pas clairement identifiée, on pense que le mécanisme sous-jacent est la destruction, par le système immunitaire, de la myéline ou l'inefficacité des cellules productrices de myéline. Parmi les causes de la sclérose en plaques, il y a la susceptibilité génétique et les

facteurs environnementaux tels que le déclenchement d'une infection virale; c'est une maladie multifactorielle.

Il n'y a pas de remède connu pour la sclérose en plaques. Les traitements visent à améliorer la qualité de vie des malades après une attaque et à empêcher l'apparition de nouvelles attaques. Ils sont prescrits dans la forme récurrente-rémittente, car ils sont généralement inefficaces pour la forme progressive de la maladie. Le traitement de première intention est constitué d'immunomodulateurs comme les interféron- $\beta$ . En seconde intention, le patient peut aussi être traité par immunosuppresseurs comme le natalizumab.

La sclérose en plaques touche principalement le jeune adulte. L'âge moyen de début des symptômes est 30 ans. Les femmes représentent près de 70% des malades. Environ 80 000 personnes sont touchées en France par la SEP.

### III. Cohorte Munich

#### A. Description de la cohorte

L'étude de cette cohorte a été approuvée par le comité d'éthique de l'école de médecine de l'université technique de Munich. Des tests ADA ont été réalisés lors d'une visite unique pour 2254 patients entre 2005 à 2008. La positivité aux ADA était déterminée à l'aide de la méthode de dosage ELISA. C'est une méthode immuno-enzymatique. Dans les échantillons de sérum prélevés, les ADA forment des complexes avec la molécule de médicament. Les ADA libres sont détectés par une enzyme conjuguée à une protéine. L'enzyme libérant un composant coloré, la concentration en ADA est évaluée par spectroscopie.

Les patients inclus étaient "bio-thérapeutiquement naïfs" (c'est-à-dire, n'ayant pas été préalablement traités par la biothérapie prise dans le cadre de l'étude) pour la biothérapie prescrite (ils pouvaient avoir pris d'autres biothérapies précédemment), âgés de plus de 18 ans. Nous avons retenu pour l'analyse, les patients dont l'échantillon avait été prélevé dans les 22 premiers mois, ceci afin de nous assurer d'être dans la fenêtre d'immunogénicité des ADA. Le schéma d'inclusion des patients de cette cohorte est décrit dans BACHELET et al. [2016].

Finalement, 408 patients ont été inclus dans l'étude. Ils étaient traités par immunomodulateurs comme détaillé dans le tableau 7.1. 89,3% des patients étaient atteint d'une SEP récurrente-rémittente.

Interferon- $\beta$ 1a		Interferon- $\beta$ 1b
AVONEX intra-musculaire	REBIF sous-cutanée	BETAFERON sous-cutanée
117	137	154
28,7%	33,6%	37,8%

TABLEAU 7.1 – Traitements pour patients traités pour scléroses en plaques, cohorte Munich, ABIRISK. La deuxième ligne indique le nom commercial du médicament ainsi que sa voie d'administration.

## B. Test de l'influence du sexe et de l'âge sur le délai avant apparition des ADA

Parmi les 408 patients analysés, 10,5% avaient développé des ADA ( $n_{\text{ADA}} = 43$ ). L'échantillon contenait 73,5% de femmes ( $n_{\text{femmes}} = 300$ ). 82,1% des patients étaient âgés de moins de 45 ans à la date de prélèvement ( $n_{<45} = 335$ ). Les variables sexe et âge (<45) ont été testées.

Le test du score avec variance permutationnelle suppose une distribution similaire des temps de visite dans les groupes comparés. La condition de censure similaire dans les groupes comparés est vérifiée ( $p\text{-value}_{\text{sexe}} = 0,1$ ,  $p\text{-value}_{\text{âge}} = 0,4$ ).

La figure 7.1 montre l'estimateur du maximum de vraisemblance non-paramétrique de la distribution de survie du délai avant apparition des ADA pour chaque groupe en fonction du sexe et de l'âge (<45).

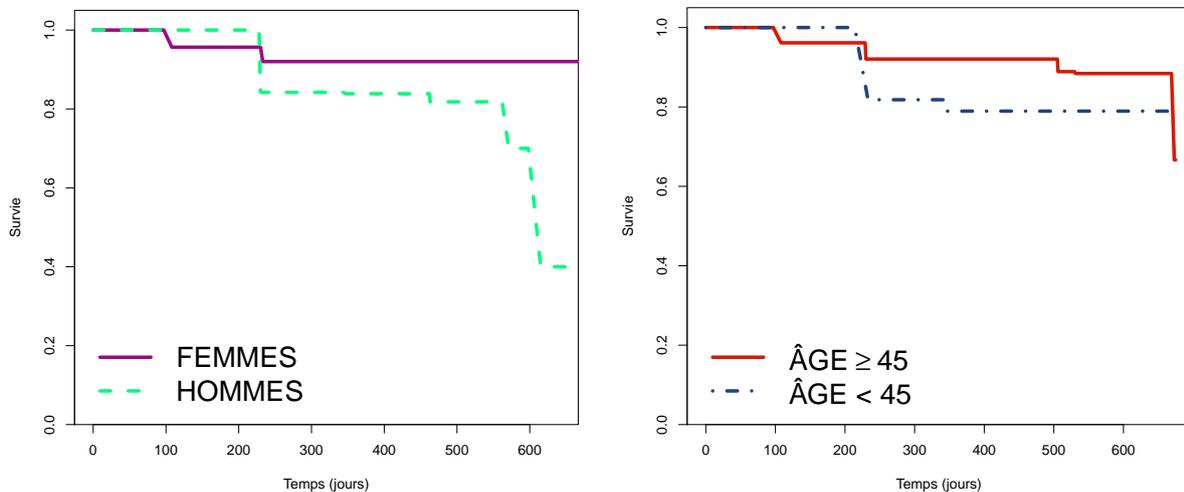


FIGURE 7.1 – Estimateur non-paramétrique du maximum de vraisemblance de la fonction de survie en fonction du sexe (gauche) et en fonction de l'âge < 45 (droite). Cohorte rétrospective Munich, ABIRISK.

### Variable sexe

Le test du score proposé montre une différence significative entre les deux groupes au niveau de signification 5% avec une valeur de 12,85 ( $p\text{-value} = 0,0016$ ). Sur la figure 7.1, il apparaît un risque plus faible pour les femmes de développer des ADA, comparé aux hommes. La valeur du plateau à 22 mois était différente entre les deux groupes : chez les femmes, elle était de 0,92 et chez les hommes de 0,40.

### Variable âge

Le test du score proposé montre une différence significative entre les deux groupes au niveau de signification 5% avec une valeur de 6,96 ( $p\text{-value} = 0,03$ ). Sur la figure 7.1, il apparaît un risque plus faible pour les personnes âgées de moins de 45 ans de développer des ADA, comparé aux personnes âgées de plus de 45 ans. La valeur du plateau à 22 mois était différente entre les deux groupes : chez les patients âgés de moins de 45 ans, elle était de 0,66, et chez les patients âgés de plus de 45 ans, de 0,79.

## IV. Cohorte Düsseldorf

### A. Description de la cohorte

Cet ensemble de données provient d'une étude transversale réalisée entre 2008 et 2014 par le Laboratoire de référence en immunologie de Düsseldorf; il est constitué d'échantillons ayant été prélevés sur des patients durant les deux ans suivant le début de leur prise de traitement. La collecte de données a été approuvée par le comité d'éthique de la faculté de médecine de l'université de Düsseldorf, dans le cadre de la banque de données créée par le consortium européen ABIRISK. L'étude concernait 5500 patients.

Ici, nous avons considéré des patients adultes, "biothérapeutiquement naïfs", dont les échantillons biologiques ont été prélevés au cours des 22 premiers mois de traitement; ceci afin de nous assurer d'être dans la fenêtre d'immunogénicité des biothérapies étudiées. Un échantillon par patient a été fourni. Le schéma d'inclusion des patients de cette cohorte est décrit dans BACHELET et al. [2016].

Dans cet ensemble de données, les patients étaient traités par interféron- $\beta$  : 63,3% prenaient de l'interféron- $\beta$ -1a et 36,7% de l'interféron- $\beta$ -1b. L'interféron- $\beta$ -1b n'est administré que par voie sous-cutanée alors que l'interféron- $\beta$ -1a peut l'être par voies sous-cutanée (54,7%) ou intramusculaire (45,3%).

### B. Test de l'influence du sexe sur le délai avant apparition des ADA

Au total, 984 patients ont été analysés, dont 9,9% avaient développé des ADA ( $n_{\text{ADA}} = 97$ ). L'échantillon était constitué de 72,6% de femmes ( $n_{\text{femmes}} = 714$ ). La variable sexe a été testée.

Comme précédemment, la condition de censure similaire dans les deux groupes est vérifiée ( $p\text{-value}_{\text{sexe}} = 0,8$ ).

La figure 7.2 montre l'estimateur du maximum de vraisemblance non-paramétrique de la distribution de survie du délai avant apparition des ADA pour chaque groupe.

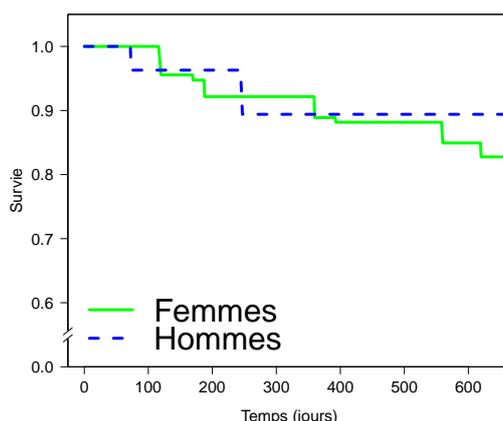


FIGURE 7.2 – Estimateur non-paramétrique du maximum de vraisemblance de la fonction de survie en fonction du sexe, cohorte Düsseldorf, ABIRISK

Le test du score proposé montre une différence significative entre les deux groupes au niveau de signification 5% avec une valeur de 7,42 ( $p\text{-value} = 0,02$ ). Notons, qu'ici le test de type log-rank n'était pas significatif avec une valeur de 3,29 ( $p\text{-value} = 0,07$ ). Comme le montre la figure 7.2, les hommes ont un risque plus élevé de développer des ADA que les femmes et un risque similaire de susceptibilité. La valeur du plateau, à 22 mois, chez les femmes était de 0,80 et chez les hommes de 0,84 ( $p\text{-value}$  du test du  $\chi^2$  à 0,14).

## V. Discussion

Nous avons analysé les données rétrospectives de deux cohortes allemandes du projet ABI-RISK à l'aide du test que nous avons développé et présenté dans le chapitre précédent. Le test est destiné à détecter l'effet de covariable dans une population hétérogène, où l'événement d'intérêt est censuré par intervalle. Dans ces cohortes, un seul dosage de test de positivité aux ADA a été réalisé après le début du traitement; l'étude étant une étude pilote portant sur l'immunogénicité des biothérapies.

Dans notre étude sur les données de Düsseldorf et de Munich, dont on peut considérer qu'elles comportent un mélange de sujets susceptibles et non-susceptibles, nous avons identifié une association significative entre le sexe et l'apparition des ADA dans les deux cohortes analysées. De plus, nous avons identifié un effet de l'âge dans la cohorte de Munich.

Nous avons déterminé que le sexe est lié à l'immunogénicité des ADA dans les deux cohortes analysés; un effet non détecté par le test de type log-rank classique dans la cohorte de Düsseldorf. Ce résultat est rapporté dans d'autres cohortes traitées par interféron- $\beta$  pour la sclérose en plaques comme chez BACHELET et al. [2016].

Nous avons également identifié un effet de l'âge dans la cohorte de Munich; les patients plus âgés développant plus d'ADA. Cet effet n'a pas été retrouvé dans la seconde cohorte analysée, mais a été identifié dans d'autres cohortes du projet ABIRISK décrites dans BACHELET et al. [2016].

Notons que la cohorte de Munich traitée par interféron- $\beta$ , comportait 10,7% de patients atteints des formes progressives de la sclérose en plaques. Or, l'interféron- $\beta$  est actuellement prescrit pour les formes rémittentes, car il est probablement inefficace dans les formes progressives de la maladie. Cependant, pendant longtemps, en l'absence d'autres médicaments approuvés pour les formes progressives, l'interféron- $\beta$  était malgré tout prescrit dans ce cas. La cohorte de Munich, étant une étude pilote relativement ancienne, elle inclut des cas progressifs traités par interféron- $\beta$ .

Dans cette cohorte, la forme de la maladie n'est pas liée au type d'interféron- $\beta$  prescrit (1a intra-musculaire, 1a sous-cutanée, 1b), ni au développement d'ADA, ni au sexe. Comme attendu, la forme de la maladie est liée à l'âge; les formes progressives de la maladie survenant chez des patients plus âgés. Afin de prendre en compte cet effet, nous avons vérifié les résultats des tests précédents en stratifiant sur la forme de la maladie (progressive, récurrente-rémittente).

Comme pour le test de type log-rank classique, le test du score utilisé repose sur l'hypothèse que la distribution des temps de visite est la même pour les différents groupes à comparer. Dans notre étude, nous avons vérifié qu'il n'y a pas de différence dans le schéma des temps de visite des hommes et des femmes, et des patients plus ou moins âgés. De plus, nous supposons le respect de la condition de suivi suffisant du test du score; les patients ayant été prélevés aléatoirement sur 22 mois.

Le test développé a permis de mettre en évidence deux variables liées au développement des ADA. L'utilisation du test peut être recommandée pour des données *current status* où une partie de l'échantillon ne fait pas l'événement d'intérêt.

## VI. Références

- ABIRISK. «Anti-biopharmaceutical immunization : prediction and analysis of clinical relevance to minimize the risk», URL <http://www.abirisk.eu/>. 142, 145
- BACHELET, D., S. HÄSSLER, C. MBOGNING, J. LINK, M. RYNER, R. RAMANUJAM, M. AUER, P. E. HYLDGAARD JENSEN, N. KOCH-HENRIKSEN, C. WARNKE, K. INGENHOVEN, D. BUCK, V. GRUMMEL, A. LAWTON, N. DONNELLAN, A. HINCELIN-MERY, D. SIKKEMA, M. PALLARDY, B. KIESEIER, B. HEMMER, H. P. HARTUNG, P. SOELBERG SORENSEN, F. DEISENHAMMER, P. DONNES, J. DAVIDSON, A. FOGDELL-HAHN, P. BROËT et ON BEHALF OF THE ABIRISK CONSORTIUM. 2016, «Occurrence of Anti-Drug Antibodies against Interferon-Beta and Natalizumab in Multiple Sclerosis : A Collaborative Cohort Analysis», *PLOS ONE*, vol. 11, n° 11, doi :10.1371/journal.pone.0162752, p. e0162 752, ISSN 1932-6203. URL <http://dx.plos.org/10.1371/journal.pone.0162752>. 143, 145, 146
- DEISENHAMMER, F. 2009, «Neutralizing antibodies to interferon- $\beta$  and other immunological treatments for multiple sclerosis», *CNS Drugs*, vol. 23, n° 5, doi :10.2165/00023210-200923050-00003, p. 379–396. URL <https://doi.org/10.2165/00023210-200923050-00003>. 142



# Conclusion générale

Ce travail de thèse était motivé par la nécessité de développer des statistiques de tests pour la comparaison de plusieurs groupes, adaptées au cas de la censure par intervalle simple (*current status*) ou générale, pour des alternatives à risques non proportionnels. Les données recueillies dans le cadre du projet ABIRISK sur l'immunogénicité des biothérapies nous apportaient des exemples pratiques nécessitant ce type de développement.

Dans la première partie de ce travail, nous avons présenté le développement d'un test adapté aux données censurées par intervalle pour des alternatives à risque non proportionnels avec croisement des fonctions de risques instantanés. Ce test a fait l'objet d'une étude de simulations extensive, dans lesquelles ses performances étaient comparées avec celles des autres tests pour données censurées par intervalle présentés dans la partie introductive.

Le test avait un risque nominal appartenant à l'intervalle de fluctuation attendu sous l'hypothèse nulle, quel que soit le schéma testé. De plus, il avait de très bonnes performances en termes de puissance. Lorsque les données étaient simulées sous un autre modèle de croisement de risque (odds-proportionnel et non-Weibull), notre test avait également de bons résultats de puissance, comparativement aux autres tests implémentés.

Le test développé a pu être évalué sur des données réelles : la cohorte des maladies inflammatoires de l'intestin, une des cohortes prospectives du projet ABIRISK. En plus d'avoir de nombreuses variables significatives détectées, communément avec le log-rank, notre test nous a permis d'identifier des variables dont l'effet n'était pas détecté par le log-rank. La variable allaitement en particulier est intéressante, puisque l'effet d'une variable latente inobservée (production d'hormones) est fortement suspecté. Les femmes ayant allaité (récemment), développant plus d'anticorps que les autres femmes.

La problématique liée à la présence de patients non-susceptibles a été largement étudiée dans le cas d'une censure aléatoire droite. Le modèle de mélange et le modèle à risque cumulatif borné sont les deux principaux modèles considérés. Cependant, à notre connaissance, lorsque les données sont censurées par intervalle, il n'existe, actuellement, pas de méthode de comparaisons de groupes qui soit simple d'utilisation et dont les performances aient été étudiées.

Le test développé a pu être évalué sur des données réelles du consortium ABIRISK issues de deux cohortes rétrospectives allemandes dans lesquelles un seul dosage des anticorps anti-biothérapies par patient avait été réalisé lors d'une visite unique. Nous avons utilisé ces données pour appliquer le test développé pour population hétérogène et données de type *current status*. Ces données provenaient de deux cohortes de patients traités pour sclérose en plaques et avaient été analysés dans les laboratoires de Munich et de Düsseldorf. L'événement d'intérêt était comme précédemment, le développement d'anticorps anti-biothérapie (anti-interféron). La présence de patients immuno-tolérants dans l'échantillon était attendue.

Le test développé a permis de mettre en évidence un effet significatif du sexe sur l'apparition des ADA; un effet non-détecté par le test de log-rank pour censure par intervalle dans les deux cohortes, mais qui avait été signalé dans d'autres cohortes plus importantes. Le test a également permis de mettre en évidence un effet de l'âge sur le développement des ADA.

Comme nous l'avons précédemment mentionné, le croisement des courbes de risque peut, entre autres, être dû à l'effet d'une interaction latente ou à la présence d'une fraction de patients

non à risque de faire l'événement d'intérêt. Nous avons proposé deux tests permettant de faire face à cette problématique, en présence de données censurées par intervalle.

Ces deux tests peuvent être facilement implémentés en  $R$  et les codes  $R$  correspondants sont inclus dans ce document. En outre, ces deux tests peuvent être aisément étendus, pour obtenir une version stratifiée. En pratique, le test pour risque croisés a été utilisé dans sa version stratifiée pour tester les données de la cohorte des maladies inflammatoires de l'intestin.

Il est utile de noter que le test pour population hétérogène, développé pour le cas de données de type *current status* peut être étendu aux données censurées par intervalle (cas général). Une étude de simulation extensive est cependant nécessaire avant de pouvoir préciser son cadre applicatif. Cette étude de simulation est l'une des extensions prévues de ce travail. Le cadre de la comparaison de deux groupes pourrait être étendu et évalué pour la comparaison de  $k > 2$  groupes. Une autre extension concerne le développement d'un test partiel permettant de tester l'effet d'une covariable sur la proportion de sujets non susceptibles quelque soit l'effet sur le délai d'apparition de l'événement parmi les susceptibles.

Au cours de la thèse, nous avons été amené à nous intéresser aux méthodes de comparaisons de plusieurs groupes proposés pour la censure par intervalle. Nous avons pu dresser un panorama que nous espérons le plus complet possible des différents tests de comparaisons de groupe existants et utilisables facilement. Les deux tests proposés dans ce travail enrichissent l'arsenal méthodologique qui peut être utilisé en recherche clinique.

À la fin de ce travail, nous pensons que le développement de nouvelles stratégies uni et multivariées, dédiées aux données censurées par intervalle, représentent un champ méthodologique très prometteur, notamment dans le cadre d'applications en immunologie clinique.

# **Annexes**



## Annexe A

# Vraisemblance marginale des rangs pour censure droite

Pour obtenir l'expression simplifiée de la vraisemblance marginale des rangs, il faut effectuer la simplification suivante :

$$\begin{aligned}\mathcal{L}_{\text{MR}}(\beta) &\propto \int_0^\infty \int_{t_1}^\infty \dots \int_{t_{m-1}}^\infty \prod_{i=1}^m \lambda_0(t_i) e^{Z_i \beta} e^{-\left(e^{Z_i \beta} + \sum_{j=1}^{k_i} e^{Z_j i \beta}\right) \int_0^{t_i} \lambda_0(u) du} dt_m \dots dt_1 \\ &= \int_0^\infty \int_{t_1}^\infty \dots \int_{t_{m-2}}^\infty \prod_{i=1}^{m-1} \left\{ e^{-\left(e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta}\right) \Lambda_0(t_{m-1})} \right\} \\ &\quad \left[ \int_{t_{m-1}}^\infty \lambda_0(t_m) e^{Z_m \beta} e^{-\left(e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta}\right) \Lambda_0(t_m)} dt_m \right] dt_{m-1} \dots dt_1\end{aligned}$$

Le terme entre accolades :

$$\left\{ e^{-\Lambda_0(t_{m-1}) \left( e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta} \right)} \right\}$$

représente la probabilité qu'un individu dont le risque instantané est  $\lambda_0(u) \left( e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta} \right)$  n'ait pas fait l'événement jusqu'à  $t_{m-1}$ , soit, sa fonction de survie. Pour faciliter la lecture de la démonstration, nous le notons  $\left\{ \right\}$ .

$$\begin{aligned}\mathcal{L}_{\text{MR}}(\beta) &\propto \int_0^\infty \int_{t_1}^\infty \dots \int_{t_{m-2}}^\infty \prod_{i=1}^{m-1} \left\{ \right\} \left[ e^{-\left( e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta} \right) \Lambda_0(t_{m-1})} \right] dt_{m-1} \dots dt_1 \\ &= \int_0^\infty \int_{t_1}^\infty \dots \int_{t_{m-3}}^\infty \prod_{i=1}^{m-2} \left\{ \right\} \left[ \int_{t_{m-2}}^\infty \lambda_0(t_{m-1}) e^{Z_{m-1} \beta} e^{-\left( e^{Z_{m-1} \beta} + \sum_{j=1}^{k_{m-1}} e^{Z_j m-1 \beta} \right) \Lambda_0(t_{m-1})} \right. \\ &\quad \left. e^{-\left( e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta} \right) \Lambda_0(t_{m-1})} dt_{m-1} \right] dt_{m-2} \dots dt_1 \\ &= \frac{e^{Z_{m-1} \beta}}{e^{Z_{m-1} \beta} + \sum_{j=1}^{k_{m-1}} e^{Z_j m-1 \beta} + e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta}} \int_0^\infty \int_{t_1}^\infty \dots \int_{t_{m-3}}^\infty \prod_{i=1}^{m-2} \left\{ \right\} \left[ \right] dt_{m-2} \dots dt_1\end{aligned}$$

avec :

$$\begin{aligned}\left[ \right] &= \int_{t_{m-2}}^\infty \left( e^{Z_{m-1} \beta} + \sum_{j=1}^{k_{m-1}} e^{Z_j m-1 \beta} + e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta} \right) \lambda_0(t_{m-1}) e^{-\Lambda_0(t_{m-1}) \left[ \left( e^{Z_{m-1} \beta} + \sum_{j=1}^{k_{m-1}} e^{Z_j m-1 \beta} + e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta} \right) \right]} dt_{m-1} \\ &= e^{-\Lambda_0(t_{m-2}) \left[ \left( e^{Z_{m-1} \beta} + \sum_{j=1}^{k_{m-1}} e^{Z_j m-1 \beta} + e^{Z_m \beta} + \sum_{j=1}^{k_m} e^{Z_j m \beta} \right) \right]}\end{aligned}$$

D'où :

$$\mathcal{L}_{MR}(\beta) = \frac{e^{Z_{m-1}\beta}}{e^{Z_{m-1}\beta} + \sum_{j=1}^{k_{m-1}} e^{Z_{jm-1}\beta} + e^{Z_m\beta} + \sum_{j=1}^{k_m} e^{Z_{jm}\beta}} \int_0^\infty \int_{t_1}^\infty \cdots \int_{t_{m-3}}^\infty \prod_{i=1}^{m-2} \left\{ \right. \\ \left. e^{-\Lambda_0(t_{m-2})} \left[ e^{Z_{m-1}\beta} + \sum_{j=1}^{k_{m-1}} e^{Z_{jm-1}\beta} + e^{Z_m\beta} + \sum_{j=1}^{k_m} e^{Z_{jm}\beta} \right] \right\} dt_{m-2} \dots dt_1$$

En poursuivant ainsi, nous voyons que :

$$\mathcal{L}_{MR}(\beta) = \dots = \prod_{i=1}^{m-1} \frac{e^{Z_i\beta}}{\sum_{l=i}^m e^{Z_l\beta} + \sum_{j=1}^{k_l} e^{Z_{jl}\beta}} = \prod_{i=1}^m \frac{\exp(Z_i\beta)}{\sum_{l \in R(t_i)} \exp(Z_l\beta)}$$

où  $R(t_i)$  est le nombre de personnes à risque juste avant  $t_i$ .

## Annexe B

# Code R du test de comparaison de deux groupes de données censurées par intervalle de type 1 incluant une fraction de non-susceptibles

Nous détaillons dans cette section le code permettant le calcul du score destiné à la comparaison de deux groupes de données censurées par intervalle de type 1 comportant une fraction de non-susceptibles. Pour rappel, pour chaque individu  $i$ , les informations disponibles sont les suivantes :

$$\{C_i, \delta_i = \mathbb{1}_{(C_i \geq T_i)}, Z_i\}.$$

- La variable *datvisite* est un vecteur constitué des dates de visites (uniques) des patients, les  $C_i$ .
- La variable *cov* est un vecteur de covariable constitué de 0 et de 1 selon le groupe d'appartenance du patient.
- La variable *evenement* est un vecteur constitué de 0 (sans événement) et de 1 (événement constaté à la visite).

Voici ci-dessous le code de la fonction de score, la fonction calcule le score de  $\beta$  et celui de  $\alpha$ , leurs variances permutationnelles respectives et leur covariance. Une fois la valeur du test du score obtenue, on compare cette valeur à une loi du  $\chi^2$  à deux degrés de liberté pour obtenir la  $p$ -value :

```
library(isotone) # pour l'estimat° de la fonct° de répartition
library(MASS)    # pour la fonct° d'inversion de matrice

score_beta_alpha <- fonction ( datvisite, cov, evenement )

{

  # Estimation de la fonction de repartition marginale

  pav      = list( gpava ( datvisite, evenement ) )
  f.repart = pav[[1]][[1]]
```

```

f.repart[ f.repart==1 ] = max( f.repart[ f.repart!=1 ] )
f.repart[ f.repart==0 ] = min( f.repart[ f.repart!=0 ] )

# Estimation du theta (plateau) en prenant la valeur
# maximum atteinte par la fonct° de repartit° marginale

theta      = - log ( min ( 1- f.repart ) )

# Estimation de la fonction de survie conditionnelle

f.cond     = - log ( 1 - f.repart ) / theta
f.cond     = ifelse (f.cond > 1, 1, f.cond)

Lambda = ifelse( f.cond==1, 0, -log( 1 - f.cond ) )

poids.beta = theta * Lambda * ( 1-f.cond ) / f.repart*
            ( evenement - f.repart )
poids.beta = ifelse(is.na(poids.beta)==T, 0, poids.beta)

score.beta = sum( cov*( poids.beta - mean( poids.beta ) ) )
variance.beta = sum( ( cov - mean( cov ) ) ^ 2 ) / ( length( cov ) - 1 ) *
                sum( ( poids.beta - mean( poids.beta ) ) ^ 2 )

poids.alpha = theta*f.cond/f.repart*(evenement - f.repart)
poids.alpha = ifelse(is.na(poids.alpha)==T, 0, poids.alpha)

score.alpha = sum ( cov*(poids.alpha - mean( poids.alpha ) ) )
variance.alpha = sum ( (cov - mean(cov))^2 ) / ( length(cov) - 1 ) *
                    sum ( (poids.alpha - mean(poids.alpha))^2 )

var.beta.alpha = sum( ( cov - mean( cov ) ) ^ 2 ) / ( length( cov ) - 1 ) *
                  sum( ( poids.beta - mean( poids.beta ) ) *
                      ( poids.alpha - mean( poids.alpha ) ) )

# Calcul du vecteur de score et de sa variance

score = matrix ( c ( score.beta, score.alpha ), ncol = 2 )

variance_score = matrix( c( variance.beta, var.beta.alpha,
                           var.beta.alpha, variance.alpha ), ncol=2)

valeur.test = score %*% ginv( variance_score ) %*% t( score )

p.value      = 1 - pchisq ( valeur.test, 2 )

res          = c ( valeur.test , p.value )
names(res)   = c ( "Valeur du test", "p-value" )

return ( res ) }

```

La fonction peut être testée avec une base de données censurée par intervalle de type 1 simulée telle que :

(NB : dans cette base de données très simplifiée, il n'y a pas de plateau ni d'effet de la covariable sur le délai avant événement. La survie et la date de visite suivent une distribution uniforme.)

```
sim_data <- function ( n ){  
  
  Tevenement = runif ( n, 0, 2 )  
  Tvisite    = runif ( n, 0.1, 1.5 )  
  evenement  = ifelse ( Tevenement > Tvisite, 0, 1 )  
  covariable = rbinom ( n, 1, 0.5 )  
  
  don = as.data.frame(cbind (Tvisite, evenement, covariable))  
  colnames(don) = c( "datvisite", "evenement", "covariable" )  
  
  return ( don )  
  
}
```

Exemple :

```
n = 300  
AA = sim_data ( n )  
score_beta_alpha(AA$datvisite, AA$covariable, AA$evenement)
```

Valeur du test	p-value
1.552731	0.460075



## Annexe C

# Simulations supplémentaires sous modèle de mélange

Ce chapitre présente les simulations supplémentaires au Chapitre 5. Elles correspondent aux simulations effectuées sous un modèle de mélange, en censure par intervalle de type 1.

TABLEAU C.1 – Modèle de mélange, censure exponentielle,  $\tau_0 = 30\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.48$	100.0 (0.0)	100.0 (0.1)	100.0 (0.1)	99.9 (0.0)	99.6 (0.0)	98.7 (-0.6)	96.9 (-1.3)	94.0 (-2.5)	88.8 (-4.8)	83.0 (-6.1)	77.7 (-4.8)	75.7 (-0.2)	73.6 (3.1)	72.4 (8.2)	72.2 (11.5)	73.2 (14.7)	73.0 (17.1)
-0.25 $\tau_1 = 0.39$	100.0 (0.1)	100.0 (0.3)	99.9 (0.2)	99.5 (0.9)	98.4 (1.4)	93.8 (0.9)	79.1 (-2.4)	53.8 (-8.7)	30.7 (-8.3)	20.4 (0.1)	22.7 (14.0)	32.4 (27.2)	48.9 (41.6)	62.1 (52.0)	71.4 (57.1)	77.1 (58.1)	82.0 (58.1)
0	100.0 (0.4)	100.0 (0.8)	99.8 (1.4)	98.8 (4.2)	95.0 (9.7)	82.0 (14.3)	49.3 (10.0)	15.6 (2.3)	4.8	15.5 (2.2)	48.4 (10.6)	80.2 (13.6)	95.2 (9.1)	99.1 (3.9)	99.9 (1.5)	100.0 (1.0)	100.0 (0.3)
0.25 $\tau_1 = 0.21$	100.0 (0.8)	99.9 (2.0)	99.7 (4.6)	98.4 (12.8)	92.9 (28.7)	69.8 (38.1)	34.1 (25.5)	14.0 (6.8)	24.7 (-7.6)	66.0 (-8.0)	94.7 (-0.9)	99.8 (0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.14$	100.0 (1.4)	99.9 (3.8)	99.7 (10.4)	98.2 (25.1)	90.4 (49.2)	68.7 (57.0)	41.9 (35.8)	38.9 (8.1)	71.8 (-7.5)	96.4 (-1.9)	99.9 (-0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.3$

TABLEAU C.2 – Modèle de mélange, censure exponentielle,  $\tau_0 = 50\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.66$	96.2 (-1.3)	96.3 (-0.8)	95.5 (-1.1)	94.0 (-1.9)	92.8 (-2.4)	90.5 (-3.3)	86.1 (-5.6)	82.3 (-7.0)	78.8 (-7.9)	73.6 (-8.3)	70.8 (-8.0)	68.9 (-7.9)	66.7 (-7.6)	65.4 (-6.3)	65.9 (-4.7)	63.7 (-5.4)	64.4 (-4.7)
-0.25 $\tau_1 = 0.58$	97.8 (1.9)	95.7 (1.4)	94.1 (1.3)	90.4 (2.0)	83.7 (0.7)	72.6 (-1.8)	56.9 (-5.1)	39.0 (-8.8)	24.3 (-7.7)	17.6 (-3.2)	15.8 (3.1)	18.1 (10.4)	24.0 (18.2)	27.8 (22.5)	33.8 (28.2)	37.8 (31.9)	39.9 (33.6)
0	97.6 (5.4)	96.6 (6.3)	93.4 (9.7)	86.0 (11.9)	73.9 (12.7)	52.4 (10.4)	27.7 (4.8)	10.0 (0.6)	4.6	9.5 (0.5)	27.7 (5.5)	51.9 (11.0)	73.5 (12.2)	86.8 (10.9)	93.0 (9.2)	95.7 (6.5)	97.3 (4.6)
0.25 $\tau_1 = 0.41$	97.6 (8.7)	96.6 (13.6)	92.3 (19.9)	82.9 (26.9)	66.5 (31.2)	39.5 (25.1)	17.2 (11.7)	11.4 (1.9)	24.0 (-6.2)	53.5 (-8.5)	84.4 (-2.2)	97.2 (0.1)	99.6 (-0.1)	100.0 (0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.32$	97.8 (14.7)	96.0 (21.9)	90.5 (33.7)	79.8 (43.6)	62.2 (46.5)	38.9 (33.8)	27.6 (16.0)	36.4 (-3.3)	67.5 (-9.8)	93.6 (-2.8)	99.6 (-0.2)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.5$

TABLEAU C.3 – Modèle de mélange, censure exponentielle,  $\tau_0 = 70\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.81$	70.4 (-8.0)	69.5 (-8.6)	69.2 (-8.1)	67.7 (-8.9)	64.9 (-8.3)	63.6 (-9.2)	61.3 (-10.2)	59.1 (-10.9)	55.8 (-11.3)	53.9 (-9.2)	51.2 (-10.9)	49.9 (-10.5)	50.8 (-10.0)	48.7 (-10.2)	48.1 (-9.7)	47.9 (-10.0)	47.4 (-10.2)
-0.25 $\tau_1 = 0.76$	72.6 (-0.2)	70.1 (-0.6)	66.4 (-0.8)	58.7 (-2.9)	52.8 (-3.7)	41.6 (-5.7)	30.9 (-7.4)	23.7 (-6.3)	16.4 (-5.6)	13.4 (-2.8)	10.8 (-0.5)	10.4 (1.6)	12.0 (5.2)	13.4 (7.5)	13.8 (9.0)	16.0 (10.5)	15.6 (10.8)
0	75.5 (9.4)	70.7 (8.8)	62.9 (7.2)	52.8 (8.3)	39.4 (5.9)	25.7 (3.6)	14.2 (0.2)	7.4 (-0.4)	4.7	7.3 (0.2)	14.5 (1.0)	27.4 (3.5)	40.0 (5.7)	53.7 (7.8)	63.3 (8.8)	71.5 (10.6)	75.2 (8.3)
0.25 $\tau_1 = 0.63$	77.2 (17.9)	69.3 (18.2)	60.4 (19.2)	48.8 (18.7)	33.3 (16.3)	18.8 (9.5)	9.9 (5.2)	8.4 (-0.7)	15.3 (-4.7)	35.1 (-8.4)	60.4 (-5.8)	81.4 (-2.6)	93.5 (-0.7)	98.3 (0.2)	99.5 (0.0)	99.9 (0.1)	100.0 (0.1)
0.5 $\tau_1 = 0.56$	78.6 (27.0)	69.7 (27.9)	58.4 (30.0)	45.7 (29.6)	30.6 (23.4)	19.5 (14.1)	17.1 (5.0)	27.3 (-5.7)	51.1 (-10.6)	77.9 (-7.7)	94.5 (-2.5)	99.4 (-0.3)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.7$

TABLEAU C.4 – Modèle de mélange, censure exponentielle,  $\tau_0 = 30\%$ ,  $p = 40\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.48$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.6 (-0.2)	94.2 (-2.6)	69.1 (-9.5)	38.9 (-2.6)	30.1 (17.6)	42.1 (37.1)	62.6 (49.1)	81.1 (47.5)	91.6 (36.5)	96.5 (26.1)	97.9 (17.4)
-0.25 $\tau_1 = 0.39$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.8 (0.0)	95.0 (-1.4)	65.0 (-7.8)	20.0 (-6.9)	9.5 (4.3)	34.2 (15.2)	75.2 (15.6)	96.0 (7.4)	99.6 (1.6)	100.0 (0.1)	100.0 (0.0)	100.0 (0.0)
0	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.1)	98.5 (0.4)	77.8 (-0.7)	25.2 (-3.2)	4.5	26.5 (-2.0)	78.6 (0.0)	98.6 (-0.1)	99.9 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.25 $\tau_1 = 0.21$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.9 (0.3)	93.5 (3.0)	55.2 (6.6)	10.5 (4.4)	16.8 (-4.2)	69.8 (-8.1)	98.1 (-0.6)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.14$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.5 (1.0)	87.0 (9.7)	38.4 (16.1)	13.0 (6.4)	46.1 (-9.3)	93.1 (-3.3)	99.9 (-0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.3$

TABLEAU C.5 – Modèle de mélange, censure exponentielle,  $\tau_0 = 50\%$ ,  $p = 40\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.66$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.9 (-0.1)	99.6 (-0.1)	96.3 (-1.7)	85.4 (-5.5)	61.9 (-10.2)	37.8 (-7.0)	25.5 (4.5)	26.1 (17.4)	33.1 (27.7)	42.9 (36.0)	52.5 (40.8)	61.8 (44.7)	68.8 (45.7)
-0.25 $\tau_1 = 0.58$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.1)	99.6 (0.0)	96.5 (-0.7)	82.2 (-4.3)	48.5 (-8.9)	17.9 (-5.3)	8.1 (1.8)	15.5 (8.4)	39.8 (15.5)	69.7 (17.1)	87.0 (12.0)	95.0 (6.3)	98.7 (2.8)	99.5 (1.4)
0	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.9 (0.0)	98.0 (0.0)	87.1 (1.1)	53.4 (-1.6)	17.6 (-1.5)	4.6	16.4 (-2.3)	52.5 (-1.4)	87.2 (0.1)	98.1 (0.4)	99.8 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.25 $\tau_1 = 0.41$	100.0 (0.0)	100.0 (0.0)	100.0 (0.1)	99.3 (0.3)	93.8 (3.0)	68.3 (4.0)	26.8 (4.7)	7.7 (2.3)	16.5 (-4.3)	55.8 (-8.8)	91.7 (-2.9)	99.6 (-0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.32$	100.0 (0.0)	100.0 (0.0)	99.9 (0.4)	98.6 (2.5)	85.3 (8.8)	50.9 (14.6)	17.6 (10.6)	13.9 (-0.5)	49.0 (-10.6)	88.8 (-4.5)	99.5 (-0.4)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\exp(\alpha)} / \tau_0 = \tau_{Z=0} = 0.5$

TABLEAU C.6 – Modèle de mélange, censure exponentielle,  $\tau_0 = 70\%$ ,  $p = 40\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.5$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.81$	100.0 (0.0)	100.0 (0.0)	99.7 (-0.1)	98.9 (-0.4)	96.9 (-0.8)	91.3 (-2.6)	79.8 (-6.0)	61.9 (-9.9)	41.2 (-11.3)	27.4 (-6.6)	19.1 (-1.3)	15.4 (4.0)	16.5 (9.8)	19.4 (14.2)	24.0 (18.4)	25.5 (20.3)	28.6 (22.0)
-0.25 $\tau_1 = 0.76$	99.9 (0.0)	99.9 (0.0)	99.3 (0.0)	97.4 (-0.3)	91.1 (-1.5)	76.7 (-4.6)	54.5 (-7.2)	29.8 (-7.3)	12.4 (-4.4)	6.8 (0.2)	8.3 (3.1)	18.2 (7.1)	32.0 (10.2)	48.6 (10.8)	64.0 (10.6)	72.9 (8.5)	80.8 (7.5)
0	99.9 (0.0)	99.5 (-0.1)	98.4 (0.1)	93.0 (0.1)	81.4 (-0.2)	55.1 (-3.0)	28.0 (-2.3)	10.2 (-1.2)	4.9	10.4 (-1.2)	28.5 (-2.3)	56.2 (-2.4)	80.6 (-0.3)	92.7 (0.0)	98.3 (0.3)	99.6 (0.1)	99.8 (-0.1)
0.25 $\tau_1 = 0.63$	99.8 (0.0)	98.9 (0.2)	95.9 (0.5)	85.4 (1.4)	65.2 (3.2)	35.5 (3.0)	13.3 (2.2)	6.3 (0.6)	12.2 (-4.1)	36.7 (-8.9)	70.0 (-7.0)	91.6 (-2.6)	98.6 (-0.4)	99.9 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.56$	99.2 (0.0)	97.8 (0.5)	91.5 (2.8)	75.2 (6.0)	48.6 (8.5)	21.4 (7.1)	8.7 (3.5)	13.3 (-1.8)	37.2 (-10.5)	71.7 (-8.6)	94.3 (-2.7)	99.5 (-0.2)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\text{exp}(\alpha)} / \tau_0 = \tau_{Z=0} = 0.7$

TABLEAU C.7 – Modèle de mélange, censure exponentielle, covariable déséquilibrée,  $\tau_0 = 30\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.3$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.48$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.9 (0.0)	99.5 (-0.1)	96.6 (-1.6)	90.5 (-4.2)	79.1 (-7.4)	69.0 (-7.1)	63.3 (-3.0)	62.3 (5.1)	62.3 (12.8)	63.9 (19.1)	64.3 (23.9)	65.0 (27.6)	66.3 (28.7)
-0.25 $\tau_1 = 0.39$	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.9 (0.0)	99.5 (0.4)	95.3 (0.3)	79.1 (-3.3)	50.2 (-7.5)	26.9 (-6.5)	17.3 (1.3)	22.1 (13.1)	32.2 (27.2)	43.4 (38.0)	53.3 (46.0)	60.5 (52.7)	65.1 (55.0)	69.1 (57.2)
0	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	99.8 (0.4)	97.0 (3.1)	83.0 (8.0)	46.1 (6.5)	14.2 (1.2)	5.4	13.5 (2.1)	37.8 (11.2)	65.8 (19.3)	83.0 (21.9)	91.3 (19.2)	94.7 (16.1)	96.1 (14.5)	97.3 (11.9)
0.25 $\tau_1 = 0.21$	100.0 (0.0)	100.0 (0.0)	100.0 (0.3)	99.4 (2.5)	93.6 (14.4)	67.7 (26.6)	28.8 (18.9)	10.6 (4.7)	20.4 (-6.8)	56.1 (-7.5)	86.5 (-1.1)	97.4 (0.9)	99.5 (0.7)	100.0 (0.4)	100.0 (0.1)	100.0 (0.1)	100.0 (0.0)
0.5 $\tau_1 = 0.14$	100.0 (0.0)	100.0 (0.0)	99.9 (0.8)	98.9 (7.3)	89.6 (31.2)	60.8 (45.4)	31.4 (26.9)	32.0 (4.1)	64.5 (-11.0)	92.5 (-3.6)	99.5 (-0.2)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\text{exp}(\alpha)} / \tau_0 = \tau_{Z=0} = 0.3$

TABLEAU C.8 – Modèle de mélange, censure exponentielle, covariable déséquilibrée,  $\tau_0 = 30\%$ ,  $p = 20\%$ ,  $n_{\text{tot}} = 400$ ,  $\xi = 0.7$ .

$\alpha \backslash \beta$	-3.2	-2.8	-2.4	-2.0	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2
-0.5 $\tau_1 = 0.48$	92.7 (-2.7)	91.5 (-3.4)	91.8 (-2.8)	91.8 (-3.1)	91.6 (-3.3)	90.2 (-4.0)	88.7 (-5.1)	87.5 (-5.0)	85.6 (-5.9)	83.2 (-7.1)	81.0 (-7.7)	78.6 (-7.3)	77.4 (-6.3)	74.9 (-7.4)	74.2 (-6.6)	73.7 (-5.7)	73.7 (-5.4)
-0.25 $\tau_1 = 0.39$	95.3 (4.2)	95.2 (5.8)	93.2 (3.7)	90.8 (4.1)	86.4 (3.2)	77.8 (0.6)	63.2 (-2.5)	42.5 (-9.9)	24.5 (-8.3)	16.9 (-2.7)	17.1 (9.0)	26.3 (21.3)	41.7 (34.0)	59.9 (42.8)	72.8 (44.1)	81.6 (40.6)	87.1 (36.8)
0	97.5 (11.4)	95.9 (12.6)	94.8 (16.7)	90.5 (18.4)	81.6 (18.2)	63.6 (17.5)	38.5 (10.6)	14.0 (1.9)	4.6	13.6 (0.4)	45.5 (6.0)	82.7 (8.5)	97.4 (3.2)	99.8 (0.5)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.25 $\tau_1 = 0.21$	98.4 (19.8)	97.4 (23.8)	96.3 (28.7)	91.7 (36.3)	81.9 (42.7)	60.3 (38.9)	31.7 (23.7)	13.9 (6.5)	21.8 (-7.1)	60.5 (-7.4)	93.7 (-0.9)	99.9 (0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)
0.5 $\tau_1 = 0.14$	98.9 (26.8)	98.3 (34.7)	96.3 (41.6)	91.7 (51.9)	83.0 (61.0)	61.7 (53.4)	40.1 (32.5)	35.4 (6.8)	62.0 (-7.9)	93.1 (-3.2)	99.8 (-0.1)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)	100.0 (0.0)

$\tau_1 = \tau_{Z=1} = \tau_0^{\text{exp}(\alpha)} / \tau_0 = \tau_{Z=0} = 0.3$



## Annexe D

# Publications et valorisation de la thèse

Voici, ci-dessous, la liste des publications et des communications issues du travail de thèse :

### Publications

JONAS, S. F., C. MBOGNING et P. BROËT. 2017a, «A test for comparing current status survival data with crossing hazard functions and its application to immunogenicity of biotherapeutics», *Statistics in Medicine*, vol. 36, n° 27, doi :10.1002/sim.7439, p. 4364–4377, ISSN 1097-0258. URL <http://dx.doi.org/10.1002/sim.7439>.

JONAS, S. F., C. MBOGNING, S. HÄSSLER et P. BROËT. 2017b, «A score test for comparing cross-sectional survival data with a fraction of non-susceptible patients and its application in clinical immunology», *PLOS ONE*, vol. 12, n° 6, doi :10.1371/journal.pone.0179896, p. e0179896. URL <https://doi.org/10.1371/journal.pone.0179896>.

### Communications

Présentation orale, Immunogenicity of Biopharmaceuticals, Final ABIRISK Open meeting, Lisbonne (Portugal), novembre 2017.





**Titre:** Méthodes de comparaisons de deux ou plusieurs groupes de données censurées par intervalle. Avec application en immunologie clinique.

**Mots clés:** Censure par intervalle, tests de comparaison de groupes, données current status, risques croisés, modèle à plateau, immunologie clinique.

**Résumé:** Dans le cadre des analyses des données de survie, la comparaison de plusieurs groupes d'individus, où l'événement d'intérêt est censuré par intervalle, représente un défi méthodologique. Lorsque le suivi des patients au cours de l'étude n'est pas continu, l'événement d'intérêt pourra survenir entre deux dates d'observation; il est dit censuré par intervalle. Des tests de comparaisons des distributions des temps de survie pour plusieurs groupes, adaptés à la censure par intervalle, ont été développés (tests du score, tests de pseudo log-rank pondérés, tests des rangs). C'est dans ce contexte que nous avons proposé deux nouveaux tests de comparaisons de groupes adaptés à des situations particulières de censure par intervalle. Le premier test concerne une situation où l'hypothèse alternative considère que les fonctions de risque instantané se croisent. Le second test concerne une situation où la population étudiée comporte une fraction non à risque pour l'événement d'intérêt. Ces deux tests ont fait l'objet d'une application sur des données réelles d'immunologie clinique.

**Title:** Methods of comparisons of two or more groups of interval censored data. With application in clinical immunology.

**Keywords:** Interval censoring, group comparison tests, current status data, crossing hazard, cure model, clinical immunology.

**Abstract:** In the context of analysis of survival data, the comparison of several groups of individuals, where the event of interest is interval censored, represents a methodological challenge. When the monitoring of patients during the study is not continuous, the event of interest may occur between two observation dates; it is said "interval censored". Tests of comparisons of survival time distributions for several groups, adapted for interval censoring, have been developed (score tests, weighted pseudo log-rank tests, rank tests). In this context, we have developed two new group comparison tests adapted to the particular situations of interval censoring. The first test apply to a situation where the alternative hypothesis considers that the hazard functions cross. The second test concerns a situation where the study population has a fraction not at risk for the event of interest. Both of these tests have been applied to real clinical immunology dataset.

